



**Politechnika
Śląska**

POLITECHNIKA ŚLĄSKA

WYDZIAŁ AUTOMATYKI, ELEKTRONIKI I INFORMATYKI

**DYSCYPLINA: INFORMATYKA TECHNICZNA
I TELEKOMUNIKACJA**

Rozprawa doktorska

Hybrydowa wizyjno-radiowa metoda wspomagania rozpoznawania
i pozycjonowania komponentów inteligentnego środowiska domowego
oparta na uczeniu maszynowym

Autor: mgr inż. Sebastian Pokuciński

Promotor: prof. dr hab. inż. Dariusz Mrozek

Opiekun pomocniczy: dr inż. Marek Koźlak

Gliwice, wrzesień 2025

*Rozprawę dedykuję żonie Magdalenie,
która pozwoliła zamienić mieszkanie w poligon doświadczalny,
oraz córce Katarzynie, która zaczęła z narodzinami,
aż tata skończył pisać doktorat*

Podziękowania

Powstanie niniejszej rozprawy nie mogłoby się udać, gdyby nie wsparcie i pomoc, którą otrzymałem od niezliczonej rzeszy osób. Jestem przekonany, że wymienienie ich wszystkich wykracza poza zdolności mojej pamięci. Jeżeli więc nie odnajdziesz swojego imienia poniżej, a przyczyniłeś się w choć minimalnym stopniu do końcowego efektu (nawet samą dobrą myślą!), wiedz, że serdecznie Ci dziękuję.

W pierwszej kolejności chciałbym podziękować mojemu promotorowi naukowemu, profesorowi Dariuszowi Mrozkowi, oraz opiekunowi ze strony firmy Simlab, Markowi Koźlakowi. Bez ich wiary we mnie i stałego zaangażowania na przestrzeni kolejnych lat, nie udałoby mi się doprowadzić tak dużego projektu do końca, o jego jakości naukowej i wartości przemysłowej nie wspominając.

Serdeczne podziękowania należą się całej najbliższej rodzinie, z żoną Magdaleną na czele, rodzicom, teściom, ciociom i wujkom (nawet tym przyszywanym!). Wszą pomoc na przestrzeni właściwie pięciu lat nie ma ceny, bo i obejmowała rzeczy bezcenne: motywację, wyrobienie poczucia wartości dla edukacji i nauki, bezkresną wiarę w sukces (nawet wtedy, gdy ja ją traciłem!), a także zwykłą obecność i dobre słowo (obiady i kawę z dostawą do komputera pomijam celowo).

W podziękowaniach nie mogę zapomnieć o moich współpracownikach, którzy towarzyszyli mi właściwie każdego dnia. Za współpracę akademicką chciałbym szczególnie podziękować Krzysztofowi Hanzłowi oraz Katarzynie Filus. Wielokrotnie tworzyliśmy coś razem, często w zdecydowanie nieprzyzwoitych godzinach pracy. Za biurową codzienność dziękuję natomiast Aleksandrowi, Michałowi, Markom, Jakubowi, Rafałowi, Przemysławom, a także całej załodze, która nawet przy kawie dopytywała “czy praca już oddana?”.

Końcowe podziękowania kieruję do profesora Jacka Pieprzycy, którego bluesowe podejście stanowiło dla mnie oazę spokoju w momentach paniki i zwątpienia.

Spis treści

Streszczenie	1
1 Wstęp	3
1.1 Zarządzanie nieruchomościami	4
1.1.1 Digitalizacja budynków	6
1.1.2 Cyfrowy bliźniak w kontekście inteligentnych budynków . . .	8
1.2 Motywacja i cele przemysłowe	10
1.3 Problematyka naukowa	11
1.3.1 Hipotezy pracy	14
1.4 Układ i metodyka pracy	14
2 Analiza zagadnienia	17
2.1 Analiza literaturowa	19
2.1.1 Metodologia przeglądu literatury	19
2.1.2 Taksonomia typów danych wejściowych	25
2.1.3 Taksonomia kategorii przetwarzania	27
2.1.4 Taksonomia realizowanych zadań	30
2.1.5 Dostrzegane wyzwania	39
2.1.6 Podsumowanie	40
2.2 Badania rynku i oczekiwań użytkowników	42
2.2.1 Analiza produktu	42
2.2.2 Analiza SWOT	44
2.2.3 Analiza konkurencji rynkowej	46
2.2.4 Badania sondażowe	53
2.2.5 Podsumowanie	60

3	Systemy radiowe	61
3.1	Osadzenie pracy w kontekście	63
3.1.1	Radiowe wyznaczanie dystansu	63
3.1.2	Pozycjonowanie przestrzenne	69
3.1.3	Technologie i protokoły radiowe	77
3.2	Elementy autorskie w zakresie pozycjonowania radiowego	80
3.2.1	Digitalizacja stanowiska badawczego	80
3.2.2	Skonstruowane urządzenia	82
3.3	Prototyp systemu lokalizacji radiowej	91
3.3.1	Fizyczna budowa systemu	92
3.3.2	Środowisko testowe i warunki eksperymentalne	94
3.3.3	Obsługa protokołu Matter	101
3.3.4	Konfiguracja eksperymentów	102
3.3.5	Wyniki	104
3.3.6	Wnioski	112
4	Systemy wizyjne	115
4.1	Osadzenie pracy w kontekście	117
4.1.1	Sztuczne sieci neuronowe i architektura YOLO	117
4.1.2	Miary skuteczności detektorów obiektów	123
4.1.3	Obrazy sferyczne i ich reprezentacja	126
4.1.4	Publiczne zbiory danych	129
4.1.5	Etykietowanie danych własnych	131
4.2	Elementy autorskie	132
4.2.1	Badanie skuteczności dostępnych rozwiązań	132
4.2.2	Algorytm doboru punktu rozcięcia sfery	141
4.2.3	Oryginalne zbiory danych	150
4.3	Prototyp systemu lokalizacji wizyjnej	158
4.3.1	Detektor głębi	159
4.3.2	Środowisko testowe	161
4.3.3	Eksperymenty	163
4.3.4	Wyniki i wnioski	166

5	System fuzyjny	169
5.1	Hybrydowa metoda pozycjonowania obiektów środowiska domowego	170
5.2	Skonstruowane urządzenia	173
5.3	Środowisko badawcze	179
5.4	Pozycjonowanie ogólne	184
5.5	Pozycjonowanie precyzyjne	188
5.6	Pozycjonowanie wizyjne	192
5.7	Przeprowadzone eksperymenty końcowe	198
5.7.1	Konfiguracja sprzętu i środowiska	198
5.7.2	Przebieg eksperymentów	201
5.7.3	Wyniki	206
6	Podsumowanie	227
6.1	Weryfikacja hipotez pracy	232
6.1.1	Hipoteza główna	232
6.1.2	Hipoteza dodatkowa 1	233
6.1.3	Hipoteza dodatkowa 2	234
6.2	Aspekt wdrożeniowy projektu	234
6.3	Dalszy rozwój projektu	236
6.4	Dorobek publikacyjny autora	238
6.4.1	Publikacje naukowe	238
6.4.2	Publikacje popularnonaukowe	240
6.5	Inna aktywność naukowo-badawcza autora	243
	Spis pojęć i skrótów	247
	Spis rysunków	255
	Spis tabel	265
	Bibliografia	269

Streszczenie

Digitalizacja otoczenia człowieka postępuje, a wraz z nią powszechniejsze tworzenie jego cyfrowych reprezentacji. Modele trójwymiarowe budynków są coraz szerzej wykorzystywane, a tworzy się je z użyciem nie tylko profesjonalnych skanerów przestrzeni, ale również urządzeń codziennego użytku, jak kamery sportowe lub smartfony. Niezależnie od urządzenia akwizycji danych, w opracowanych modelach brakuje jednak najważniejszego: semantycznego znaczenia ich komponentów. Są to pewne cyfrowe odwzorowania rzeczywistości, np. chmury punktów, pozbawione jednak użytecznych informacji o rodzajach i funkcjach swoich elementów.

Aby tę lukę wypełnić, w ramach rozprawy opracowano hybrydową, wizyjno-radiową metodę wspomagania rozpoznawania i pozycjonowania komponentów inteligentnego środowiska domowego opartą na uczeniu maszynowym. Pozwala ona odzyskać tę utraconą semantykę i nadać surowym danym znaczenie oraz określić ich funkcje. Wykorzystane zostały w tym celu elementy dwóch gałęzi rozwiązań: pozycjonowania radiowego, a także wizji komputerowej wspomaganej sztuczną inteligencją, które scalono do postaci metody hybrydowej i poddano badaniom.

W rozprawie omówiony został cały proces rozwoju stworzonego rozwiązania: od analizy literatury i stanu wiedzy w dziedzinie, przez opracowanie prototypów i badania ich skuteczności, aż po sformułowanie autorskiej metody postępowania i jej testy. Prace prowadzono w sposób usystematyzowany, każdorazowo definiując pewien problem, analizując proponowane rozwiązania i weryfikując ich użyteczność, szczególnie pod kątem przemysłowym. Pracę rozpoczęto od przeglądu literatury, a także badań rynkowych oraz oczekiwań użytkowników. Wyciągnięte z nich wnioski posłużyły do opracowania pierwszych prototypów systemów radiowych i wizyjnych. W zakresie systemów radiowych rozpoczęto od badania procesu estymacji dystansu, który dostarczył wiedzy o skuteczności i zasadach działa-

nia technologii radiowych Bluetooth Low Energy (BLE), Ultra-Wideband (UWB) oraz protokołu Matter. Na tej podstawie opracowano dwa algorytmy: pozycjonowania ogólnego, służący do wskazywania pomieszczeń prowadzenia badań i oparty na BLE, a także pozycjonowania precyzyjnego, wykorzystujący UWB oraz metodę trilateracji do określania współrzędnych śledzonych obiektów. Dla systemów wizyjnych rozpoczęto od badań skuteczności różnych generacji detektorów obiektów z rodziny YOLO (z ang. *You Only Look Once*) na publicznie dostępnych zbiorach zdjęć, a następnie przeanalizowano proces opracowywania zbiorów własnych. Zaproponowano autorski algorytm do wskazywania najlepszego południka do rozcięcia sfery w procesie rzutowania jej na reprezentację panoramiczną, a także zbadano zastosowanie estymatora głębi do oceny odległości do wykrytych na obrazach panoramicznych zdjęć. Wyniki zestawiono z wiedzą o sferycznym przekształcaniu współrzędnych i doprowadzono do postaci metody pozycjonowania wizyjnego, która określała położenie wykrytych obiektów w układzie współrzędnych urządzenia digitalizującego przestrzeń. Finalnie dokonano syntezy wyników opisanych badań, wypracowanych metod postępowania, a także skonstruowanych urządzeń.

Rozbudowane testy eksperymentalne potwierdziły wysoki stopień uzyskiwanej automatyzacji procesu odzyskiwania semantycznego znaczenia w surowych modelach cyfrowych nadzorowanych nieruchomości. Wykorzystanie proponowanej metody hybrydowej w skonstruowanym środowisku badawczym pozwoliło na automatyczne wykrycie 62% wszystkich jego komponentów. Co ważne, 95% tak wykrytych obiektów zostało jednocześnie zlokalizowanych w przestrzeni trójwymiarowej ze średnią precyzją na poziomie 71,8 cm. Pozwala to użytkownikowi końcowemu odnaleźć nawet niewielkich rozmiarów przedmiot w promieniu poniżej jednego metra od wskazanego punktu i to niezależnie, czy jest to urządzenie elektroniczne z własnymi komponentami radiowymi, czy obiekt pasywny zupełnie pozbawiony zasilania. Opracowana metoda może być zatem stosowana przez szerokie grono odbiorców, tak użytkowników prywatnych, jak i przemysłowych. Z jej wykorzystaniem realizować można różnorodne zadania wymagające lokalizowania zasobów, takie jak identyfikacja urządzeń biurowych wymagających serwisu (np. jednostek klimatyzacji) lub inwentaryzacja zasobów (jak gaśnice).

Słowa kluczowe: wizja komputerowa, pozycjonowanie radiowe, uczenie maszynowe, internet rzeczy, cyfrowy bliźniak, inteligentne budynki

Rozdział 1

Wstęp

Zagadnienie rozumienia otoczenia oraz najbliższego człowiekowi środowiska stale zyskuje na znaczeniu. Dotyczy to tak aspektu naukowego, jak i praktycznego. Obszarami, w których jest to szczególnie widoczne, są koncepcje tzw. inteligentnych domów (z ang. *Smart Home*) oraz całych miast, określanых mianem inteligentnych (z ang. *Smart City*). Terminy te odnoszą się do przestrzeni działających w pewnym stopniu w sposób zautomatyzowany, wyposażonych w sieci czujników i aktorów wykonawczych, które poprzez systemy komunikacji i analizy danych pozwalają dynamicznie reagować na zmieniające się warunki otoczenia i dostosowywać się do oczekiwań mieszkańców. Upowszechnienie zastosowania szeroko rozumianej Sztucznej Inteligencji (SI, z ang. *Artificial Intelligence*, AI) w niemal wszystkich dziedzinach nauki sprawiło, że i tutaj rozwiązania oparte na Sztucznych Sieciach Neuronowych (SSN, z ang. *Artificial Neural Network*, ANN) stały się właściwie standardowym wyborem. Wspomagać możemy w ten sposób takie procesy, jak identyfikacja obecności osób niepowołanych na nadzorowanych obszarach, monitorowanie rozmieszczenia cennych zasobów czy dynamiczne zarządzanie urządzeniami w oparciu o analizę danych pochodzących z otoczenia. Zakres zastosowań tego typu systemów może obejmować zarówno monitoring pojedynczego pomieszczenia, jak i całych budynków, a nawet rozproszonych zespołów obiektów, takich jak osiedla czy квартаły miejskie.

Jednocześnie rozwiązania te nie muszą już ograniczać się wyłącznie do analizy pojedynczego rodzaju danych, takiego jak obraz z kamer czy detekcja obecności.

Współcześnie mogą wykorzystywać różnorodne źródła informacji, takie jak modele trójwymiarowe, pomiary radiowe, skany laserowe, historia zużycia energii, a nawet analiza dźwięków otoczenia. Niniejsza rozprawa jest zwięźczeniem prac badawczych realizowanych na ścieżce doktoratu wdrożeniowego. Szczególny nacisk położono tu więc na biznesowe aspekty opracowywanych rozwiązań, co znajduje odzwierciedlenie w kolejnych rozdziałach, poprzez konsekwentne odniesienia i nawiązania do ponoszonych kosztów oraz dążenie do minimalizacji wymaganych nakładów pracy własnej, zwłaszcza w porównaniu z już istniejącymi i dostępnymi na rynku rozwiązaniami. Analizowano takie zagadnienia, jak dostępność rynkowa (czy dane rozwiązanie można nabyć), koszty wdrożenia (jaka jest cena zakupu), zgodność z wymaganiami użytkowników, stopień automatyzacji procesu implementacji (ile czasu i zasobów wymaga konfiguracja), a także możliwości dalszego utrzymania i rozwoju (na ile rozwiązanie jest skalowalne i elastyczne).

Dla przeciętnego użytkownika domowego dostępne na rynku rozwiązania są często nieosiągalne z uwagi na wysokie koszty lub nieprzystające do rzeczywistych potrzeb, przez co nie spełniają one oczekiwań odbiorców. Dotyczy to również środowisk o charakterze półprywatnym lub niskoskalowym, takich jak pojedyncze biura, gabinety lekarskie, salony sprzedaży i usługowe czy domy opieki, a także wszystkich przestrzeni, w których codzienne funkcjonowanie i otoczenie przypominają realia znane z gospodarstw domowych. Dzieje się tak mimo stale rosnącej liczby urządzeń Internetu Rzeczy (z ang. *Internet of Things*, IoT) obecnych w typowym, nowoczesnym środowisku oraz coraz wyższym stopniu jego digitalizacji. Odnalezienie intuicyjnego systemu wiążącego różne źródła danych i przenoszącego wyciągnięte wnioski do faktycznie funkcjonującej przestrzeni, a który nie wymaga opracowania kosztownego rozwiązania dedykowanego, pozostaje wyzwaniem. Pótwierdza to zasadność dalszych badań w tym zakresie i poszukiwań skutecznych rozwiązań służących integracji inteligentnych technologii z codzienną rzeczywistością.

1.1 Zarządzanie nieruchomościami

Nawet najbardziej zaawansowane i wysoce autonomiczne rozwiązania z zakresu rozpoznawania, monitorowania, pozycjonowania czy zarządzania inteligentnymi

przestrzeniami wymagają osadzenia w określonym kontekście, aby mogły funkcjonować skutecznie i efektywnie. Potrzebne są nie tylko zaawansowane algorytmy czy złożone systemy kontroli, lecz także wiarygodne punkty odniesienia, które pokazać można w przystępny sposób człowiekowi, będącemu użytkownikiem końcowym danego systemu. Kluczowe staje się zatem efektywne pozyskiwanie, integracja i utrzymywanie danych, obejmujących zarówno dynamiczne zmiany zachodzące w otoczeniu, jak i stałe cechy nadzorowanych budynków. Niezależnie od tego, czy mowa o systemie monitoringu przeciwwkradzieżowego, planowaniu rozmieszczenia ekspozycji towaru czy śledzeniu lokalizacji pacjentów, to właśnie jakość i dostępność danych kontekstowych przesądza o skuteczności wdrożenia nowoczesnych systemów zarządzania przestrzenią.

Dominującą praktyką w środowisku architektów oraz szerzej w branży AEC (z ang. *Architecture, Engineering, Construction*), czyli sektorze zajmującym się projektowaniem, realizacją oraz zarządzaniem budynkami i infrastrukturą, jest tworzenie tzw. modelu BIM (z ang. *Building Information Modeling*). Model BIM to koncepcja integrująca informacje o budynku, obejmujące dane geometryczne, przestrzenne, funkcjonalne i techniczne, w jedną, spójną strukturę cyfrową. Powstaje on na etapie projektowania i realizacji inwestycji przy udziale architektów, projektantów oraz inżynierów. Jego głównym założeniem jest zapewnienie kompleksowego obrazu obiektu z perspektywy zarówno technicznej, jak i eksploatacyjnej. W praktyce to właśnie na bazie modelu BIM bardzo często opiera się późniejsze wdrażanie systemów zarządzania budynkiem, które wykorzystują importowane schematy, plany oraz dodatkowe informacje niezbędne do organizacji pracy.

Choć współczesne modele BIM odgrywają istotną rolę w zarządzaniu inwestycjami budowlanymi, to w codziennej praktyce ujawniają się liczne ograniczenia tej koncepcji. Już samo korzystanie z informacji zawartych w modelu, niezależnie od kwestii ich regularnej aktualizacji, wymaga specjalistycznej wiedzy oraz dostępu do często kosztownego oprogramowania. Przeciętny użytkownik budynku, zarządca lub najemca może mieć trudności zarówno z dostępem do takiej dokumentacji, jak i z jej interpretacją. Dodatkowo, wiele elementów istotnych dla praktycznego funkcjonowania już użytkowanego obiektu, takich jak rozmieszczenie mebli, sprzętów RTV, AGD czy urządzeń elektronicznych, nie jest uwzględnionych w modelach BIM lub uwzględniane są w bardzo ograniczonym zakresie. W efekcie

nawet najbardziej zaawansowana dokumentacja projektowa nie zapewnia kompleksowego wsparcia podczas codziennego zarządzania przestrzenią.

Po zakończeniu etapu realizacji inwestycji aktualizacja modeli BIM zazwyczaj ustaje. Opracowane modele, nawet jeśli pierwotnie były bardzo szczegółowe, z biegiem czasu tracą aktualność w wyniku zmian zachodzących w obiekcie. Ręczne nanoszenie korekt dotyczących rodzaju i rozmieszczenia wyposażenia czy wprowadzanych zmian układu pomieszczeń pojawia się sporadycznie i jedynie przy dużych nakładach czasu i środków. W rezultacie narasta potrzeba ponownej weryfikacji zgodności planu ze stanem faktycznym. Tego typu działania są kosztowne i realizowane rzadko, co skutkuje sytuacją, w której dokumentacja przestaje odwzorowywać rzeczywisty stan obiektu, a jej przydatność operacyjna systematycznie maleje. Jest to zjawisko powszechne i właściwie oczekiwane, szczególnie w obiektach komercyjnych, których przestrzeń z upływem lat przechodzi wiele przebudów i reorganizacji, wynikających z decyzji i potrzeb zmieniających się użytkowników lub właścicieli.

O ile problem starzejącej się dokumentacji nie jest unikalny dla branży budowlanej, to w kontekście modeli BIM brakuje szeroko dostępnych narzędzi i praktyk umożliwiających jej automatyczne odświeżanie, najlepiej na podstawie samej nieruchomości, bez udziału specjalistów. Brakuje również bezpłatnych rozwiązań, które pozwalałyby użytkownikom końcowym na uaktualnianie dokumentacji własnym nakładem pracy. Wszystko to prowadzi do wniosku, że jednym z głównych wyzwań współczesnego zarządzania nieruchomościami jest pełna integracja dokumentacji projektowej z dynamicznie zmieniającym się stanem faktycznym obiektu. Dodatkowym wyzwaniem jest fakt, że w wielu starszych budynkach nie istnieje zupełnie żadna, nawet najprostsza dokumentacja cyfrowa, nie mówiąc już o zaawansowanych modelach BIM.

1.1.1 Digitalizacja budynków

Nie jest możliwym, aby pojedyncze nieruchomości lub całe miasta dysponowały rozbudowaną dokumentacją cyfrową, ponieważ są one znacznie starsze niż sama technologia komputerowa. Często najpełniejszą “dokumentacją” takiego obiektu bywa sam budynek. Dopóki on fizycznie istnieje, jego struktura pozostaje źród-

dłem wiedzy o układzie, funkcji i zmianach wprowadzanych przez kolejne dekady użytkowania.

W przypadku braku cyfrowych danych historycznych, podstawową metodą inwentaryzacji pozostaje ręczne tworzenie dokumentacji. Proces ten obejmuje wizyty na miejscu, dokładne pomiary, opracowywanie szkiców, fotograficzną rejestrację przestrzeni i szczegółowe opisywanie rozkładu pomieszczeń oraz wyposażenia. Niestety, metoda ta jest bardzo pracochłonna, kosztowna i podatna na błędy. Aktualizowanie powstałej dokumentacji wymaga regularnych inspekcji i stałego pilnowania spójności zapisu z faktycznym stanem obiektu. W praktyce prowadzi to do licznych opóźnień i rozbieżności, szczególnie w przypadku obiektów o dużej powierzchni lub często podlegających zmianom funkcjonalnym.

Współczesne technologie pozwalają jednak nawet bardzo stare budynki odwzorować w formie cyfrowej bez konieczności żmudnej, ręcznej pracy. Typowym podejściem jest wykonywanie skanów trójwymiarowych struktury i wyposażenia obiektu z wykorzystaniem nowoczesnych skanerów przestrzennych lub realizowanie tzw. interaktywnych spacerów wirtualnych (z ang. *Virtual Tour*, VT), które umożliwiają odtworzenie układu i wyglądu wnętrza na podstawie zdjęć wykonanych podczas przemieszczania się po budynku. Umożliwia to względnie szybkie i stosunkowo tanie odtworzenie geometrii wnętrza oraz ich wyposażenia, a uzyskany model może posłużyć jako punkt wyjścia do dalszych prac projektowych lub inwentaryzacyjnych.

Jeszcze większy postęp przyniosła upowszechniająca się technologia fotogrametrii. Pozwala ona na stworzenie szczegółowego modelu trójwymiarowego pomieszczeń i budynków przy użyciu sekwencji zdjęć lub rejestracji wideo wykonywanych za pomocą zaawansowanych kamer. Nową jakość i prostotę przyniosło w niej powszechne zastosowanie dronów do realizacji akwizycji obrazu z zewnątrz budynków, a nawet w dużych kompleksach miejskich. Fotogrametria znakomicie sprawdziła się chociażby w przypadku rekonstrukcji katedry Notre-Dame w Paryżu, gdzie dzięki cyfrowemu modelowi, wykonanemu przed pożarem, możliwe było szczegółowe odtworzenie zniszczonych fragmentów i wsparcie procesu renowacji. Obecnie podobne rozwiązania wykorzystywane są zarówno w architekturze zabytkowej, jak i w inwentaryzacji współczesnych przestrzeni użyteczności publicznej oraz przemysłowej.

Nawet najlepiej wykonane modele trójwymiarowe, powstałe dzięki skanowaniu

laserowemu czy fotogrametrii, mają jednak istotne ograniczenie: są jedynie “suchą” reprezentacją geometrii, pozbawioną warstwy semantycznej. Rezultaty uzyskane tą metodą mają charakter wyłącznie wizualny (lub przestrzenny) i nie zawierają informacji o prezentowanych obiektach. Otrzymujemy pewne cyfrowe odwzorowanie (siatkę wielokątów, chmurę punktów, ciąg zdjęć sferycznych, wizualizację zdjęciową, itd.), które nie informuje jednak automatycznie, które fragmenty przedstawiają ściany, drzwi czy wyposażenie, ani jakie funkcje pełnią poszczególne przestrzenie. Wszelkie interpretacje czy decyzje wymagają dodatkowo ręcznego nadania im takiego znaczenia i przejścia właściwie w pełni manualnego procesu klasyfikacji i etykietowania fragmentów modelu.

Fotogrametria i skany przestrzenne nie stanowią zatem kompletnego rozwiązania problemu dynamicznej cyfryzacji istniejących obiektów. Pokonują część trudności znanych z modeli BIM, tj. nieaktualność zgromadzonych danych oraz manualne ich odtwarzanie, wprowadzają natomiast nowe zagadnienie: sposoby odzyskania ich semantycznego znaczenia. Automatyzacja tego procesu, na przykład za pomocą rozwiązań uczenia maszynowego, jest obecnie jednym z kluczowych kierunków rozwoju, a który stanowi również motyw przewodni niniejszej rozprawy.

1.1.2 Cyfrowy bliźniak w kontekście inteligentnych budynków

Dotychczas omówione technologie, takie jak modele BIM, trójwymiarowe skany czy fotogrametria, umożliwiają szczegółowe odwzorowanie nieruchomości w postaci elektronicznej. Bez wsparcia mechanizmów rozpoznawania, interpretacji oraz automatycznej aktualizacji danych, modele te pełnią jedynie funkcję statycznych wizualizacji. Nie są dynamicznymi, interaktywnymi narzędziami zarządzania przestrzenią i nie potrafią samodzielnie dbać o potrzeby mieszkańców. To ograniczenie jest szczególnie istotne w kontekście utrzymywania aktualności informacji i adaptacji do sterowania automatyką budynkową.

Idea Cyfrowego Bliźniaka (z ang. *Digital Twin*, DT) stanowi rozwinięcie tych pojęć. Cyfrowy bliźniak to model cyfrowy, który nie tylko odtwarza geometryczną i funkcjonalną strukturę obiektu, ale także stanowi żywe, nieustannie aktualizowane odwzorowanie jego rzeczywistego stanu. Jest to zintegrowany system, w któ-

rym łączą się dane pochodzące z dokumentacji architektonicznej, technicznej, inwentaryzacji oraz z czujników i urządzeń działających w czasie rzeczywistym. Dzięki temu model można wykorzystywać nie tylko do bieżącego monitorowania stanu budynku, ale także do analiz predykcyjnych, przeciwdziałania zagrożeniom, zarządzania zasobami czy planowania remontów i modernizacji.

W naturalny sposób rozwój cyfrowych bliźniaków łączy się z koncepcją urządzeń Internetu Rzeczy oraz inteligentnych budynków. W ostatnich latach jest to właściwie standardem w nowoczesnych budynkach mieszkalnych i użytkowych. Opiera się na rozbudowanej sieci czujników i aktorów wykonawczych (urządzeń IoT), które monitorują i autonomicznie sterują parametrami środowiska, dopasowując je do potrzeb użytkowników. Najczęściej obserwacji podlegają takie parametry, jak temperatura, poziom hałasu, czy ruch i obecność osób. Systemy te poprawiają komfort i bezpieczeństwo użytkowników, a także umożliwiają optymalizację zużycia zasobów. Powiązanie analizy struktury budynku, obecności i aktywności mieszkańców, prognozy pogody, nasłonecznienia i temperatury pozwala np. na inteligentne zarządzanie ogrzewaniem, wentylacją, oświetleniem czy klimatyzacją.

Współczesne systemy automatyki domowej sięgają dalej niż tradycyjne pomiary. Oprócz podstawowych danych środowiskowych integrują również transmisję obrazów z kamer monitoringu czy uproszczonych skanów 3D przestrzeni, np. pochodzących ze skanerów zamontowanych w robotach odkurzających. Integracja tych danych w ujednoliconym systemie pozwala automatyce budynkowej “ożywić” cyfrowe modele nieruchomości. Realny obiekt zostaje połączony z jego wirtualnym odpowiednikiem, który jest stale uzupełniany o bieżące informacje sensoryczne. To pozwala na szybkie wykrywanie nieprawidłowości, takich jak awarie instalacji, nieuprawniony dostęp czy nieefektywne zużycie zasobów.

Zgromadzone w dobrze funkcjonującym cyfrowym bliźniaku informacje mogą być więc wystarczające do zautomatyzowania procesu digitalizacji zasobów nieruchomości oraz ich aktualizacji przy minimalnym zaangażowaniu użytkownika. Temu też procesowi, tj. automatyzacji rozpoznawania i pozycjonowania zasobów nieruchomości, poświęcone są badania przedstawione w niniejszej rozprawie. Stanowi ono jedno z kluczowych wyzwań partnera biznesowego doktoratu (firmy Simlab), motywuje realizację całego projektu i definiuje jego poszczególne cele.

1.2 Motywacja i cele przemysłowe

Przedmiotem niniejszej rozprawy są badania realizowane w ramach doktoratu wdrożeniowego, którego celem jest opracowanie naukowo-technicznego rozwiązania konkretnego problemu zgłoszonego przez partnera przemysłowego. Praca powstaje we ścisłej współpracy z firmą Simlab Sp. z o.o. z Gliwic, która rozwija platformę SIM-ON, tj. kompleksowe narzędzie do przestrzennego zarządzania nieruchomościami. Platforma ta nie tylko umożliwia wizualizację szczegółowego modelu obiektu, ale także oferuje funkcje nadzoru i analiz w czasie rzeczywistym, wykorzystując szeroko dostępne technologie automatyki domowej oraz ideę cyfrowego bliźniaka. W ramach systemu integrowane są skany trójwymiarowe, interaktywne spacery, dane wizyjne oraz odczyty z sensorów IoT, a finalny model udostępniany jest na jednej, dostępnej z dowolnego miejsca na świecie platformie chmurowej.

Pomimo zaawansowania i wysokiej jakości modeli trójwymiarowych nieruchomości, z których korzysta oferowany system, firma napotyka istotne wyzwanie. Proces pełnej konfiguracji cyfrowego bliźniaka nadzorowanej przestrzeni wymaga od użytkowników żmudnego, powtarzalnego przypisywania semantycznego znaczenia poszczególnym, interesującym ich zasobom. Brak automatyzacji oraz inteligentnego wsparcia na tym etapie często prowadzi do porzucenia konfiguracji, co skutkuje powstaniem cyfrowego bliźniaka pozbawionego pełnej wartości użytkowej i analitycznej. W efekcie produkt ogranicza się do samej wizualizacji, która nie wyróżnia się na tle konkurencji i nie spełnia oczekiwań użytkowników.

W powyższym kontekście nadrzędnym celem pracy jest zwiększenie stopnia automatyzacji procesu rozpoznawania i pozycjonowania obiektów oraz struktur inteligentnego środowiska domowego. Szczególną uwagę poświęcono urządzeniom IoT, które stanowią komponenty natywnie wyposażone w układy radiowe, potencjalnie umożliwiające nie tylko komunikację, ale również wzajemne lokalizowanie przestrzenne. Wykorzystanie tego potencjału, wraz z integracją analiz wizualnych opartych na uczeniu maszynowym, stanowi podstawę proponowanej metody hybrydowej i całego systemu fuzyjnego, które mają na celu ograniczenie manualnej pracy użytkownika oraz podniesienie jakości i kompletności modelu cyfrowego. Jednocześnie kluczowe jest precyzyjne określenie zasobów ważnych z perspektywy użytkownika, gdyż nawet w pełni zautomatyzowane rozwiązanie nie spełni swojego

zadania, jeśli nie będzie odpowiadać na rzeczywiste potrzeby odbiorców.

W ramach celów przemysłowych pracy istotne jest przeprowadzenie szerokiego rozeznania rynkowego i analizy otoczenia, które wykracza poza sam rozwój wdrożeniowego rozwiązania technicznego. Takie podejście pozwala nie tylko na lepsze zrozumienie rzeczywistych potrzeb użytkowników, ale również na ocenę aktualnego stanu dziedziny, w której funkcjonuje opracowywany produkt. Platforma chmurowa, dzięki swojej elastyczności, daje możliwość integracji różnorodnych źródeł danych i pozyskiwania nowych informacji. Jednak z biznesowego punktu widzenia nie wszystkie dane są warte uwagi. Ważne jest sprecyzowanie, które kategorie informacji i typy danych mają największe znaczenie dla funkcjonalności systemu oraz oczekiwań rynku. Dlatego tak istotna staje się systematyczna klasyfikacja danych wejściowych, metod ich przetwarzania oraz realizowanych zadań, które pozwolą ukierunkować rozwój produktu. Równolegle konieczna jest gruntowna analiza rynku, obejmująca badanie aktualnej oferty konkurencji, ocenę własnego produktu oraz przeprowadzenie badań sondażowych wśród potencjalnych użytkowników. Takie kompleksowe badania zapewnią zbudowanie fuzyjnego systemu automatyzacji rozpoznawania przestrzeni, który opiera się na rzetelnej i usystematyzowanej wiedzy o źródłach danych oraz realnych oczekiwaniach klientów, zwiększając tym samym szanse na sukces wdrożeniowy i rynkowy.

1.3 Problematyka naukowa

Dotychczasowe rozdziały jasno ukazały podstawy biznesowe i praktyczne motywacje stojące za realizacją niniejszej rozprawy, przedstawiając aktualne wyzwania rynkowe oraz potrzeby użytkowników końcowych. Jednak aby przejść do systematycznego badania problemu i precyzyjnego definiowania zakresu prowadzonych prac, konieczne jest przekształcenie tych motywacji na problematykę naukową.

Zagadnieniem głównym i naukowo najważniejszym z elementów jest opracowanie tytułowej hybrydowej wizyjno-radiowej metody wsparcia rozpoznawania i pozycjonowania komponentów inteligentnego środowiska domowego, w której wykorzystane zostaną algorytmy uczenia maszynowego. Opis ten rozszerzyć możemy o uwzględnienie, że docelowy system powinien być rozwiązaniem fuzyjnym, łączącym w sobie różne źródła danych i holistycznie wyciągającym z nich wnioski.

W niniejszej pracy hybrydyzacji podlegać będą dwie główne gałęzie tematyczne. Są to systemy wizji komputerowej oraz rozwiązania pozycjonowania radiowego. W gałęzi wizyjnej najistotniejszym aspektem badawczym jest zastosowanie opartych na uczeniu maszynowym algorytmów do rozpoznawania obiektów oraz estymacji odległości do nich, a także przetwarzania wstępnego oraz interpretacji obrazów. W kontekście systemów radiowych zagadnienia pracy dotyczą głównie analizy zdadności i łatwości użycia komponentów elektronicznych różnych technologii oraz protokołów komunikacyjnych do realizacji procesu estymacji odległości pomiędzy urządzeniami, a docelowo określania ich wzajemnego położenia. Dla obu zagadnień istotnym aspektem będzie wykorzystanie powszechnie dostępnych metod, danych, narzędzi, platform rozwojowych czy prototypów udostępnianych zgodnie z ideą wolnego oprogramowania (w tym o bezpłatnym użytku komercyjnym).

W procesie projektowania badań oraz wyznaczania kierunków eksperymentów istotne jest posiłkowanie się pytaniami badawczymi, które umożliwiają precyzyjne określenie obszaru analizy oraz wyznaczają kontekst dla formułowania hipotez pracy. Pytania te pełnią jednocześnie rolę przewodnika, utrzymując spójność tematyczną wyводу oraz monitorując zgodność prowadzonych prac z rzeczywistymi potrzebami użytkowników i rynku. W trakcie opracowywania rozwiązania zidentyfikowano szereg dodatkowych zagadnień i wyzwań, które dotyczą zarówno aspektów technologicznych, jak i użytkowych.

Do najistotniejszych pytań badawczych należą:

1. **W jaki sposób można zwiększyć odsetek użytkowników kończących proces konfiguracji systemu oraz kontynuujących jego późniejsze użytkowanie?** Pytanie to odnosi się do motywacji przewodniej całej rozprawy. Proponowanym rozwiązaniem jest opracowanie i wdrożenie możliwie rozbudowanej automatyzacji. Warto mieć jednak na uwadze, że prowadzone badania mogą zasugerować również inne rozwiązania, potencjalnie warte uwagi.
2. **Czy i w jakim stopniu możliwe jest odciążenie użytkownika z konieczności ręcznego aktualizowania informacji o posiadanym wyposażeniu oraz jego rozmieszczeniu?** Zakładając, że faktycznie rozwiąza-

niem najbardziej wspierającym użytkowników jest automatyzacja, to pytanie kieruje uwagę na stopień jej osiągalności. Potencjalnie istnieć może metoda pół-automatyczna, która połączy zadowolenie użytkownika z prostotą wdrożenia i niskimi kosztami.

3. **Jakie cechy systemu najbardziej zachęcają użytkowników końcowych do jego wdrożenia i codziennego stosowania?** Podczas analizy obieranych rozwiązań warto pamiętać, że istnieć może wiele różnych czynników decyzyjnych. Może to być cena, nowoczesność, trwałość, rozbudowany wachlarz możliwości, a może wręcz przeciwnie, prostota obsługi.
4. **Czy uzyskane wyniki można z powodzeniem odwzorować w interaktywnym modelu trójwymiarowym nadzorowanej nieruchomości?** Pytanie podkreśla wagę przeniesienia wyników badań naukowych na pole przemysłowe i uzyskanie faktycznego wdrożenia. Jeżeli opracowane rozwiązania będą wymagały ściśle kontrolowanych warunków laboratoryjnych do poprawnego działania, ich wartość zauważalnie spadnie.
5. **Czy proponowane rozwiązanie może być skonstruowane i wdrożone przy użyciu powszechnie dostępnych komponentów sprzętowych, a zatem i niższych kosztach niż dostępne obecnie produkty dedykowane?** To pytanie odnosi się do ekonomicznej efektywności projektu, a także jego potencjalnej konkurencyjności rynkowej. Aby wyróżnić swoje rozwiązanie od konkurencji musi ono być albo zauważalnie lepsze jakościowo albo przystępne cenowo.
6. **Czy opracowanie syntetycznego zbioru danych panoramicznych może zastąpić proces manualnego pozyskiwania i etykietowania zdjęć rzeczywistego środowiska domowego?** Obserwowany rozwój generatywnej sztucznej inteligencji, szczególnie w kontekście wzrostu obaw o prywatność danych w internecie, prowadzi do pytań o zastępowalność prywatnych zdjęć obrazami syntetycznymi. Nie jest jednak znana ich jakość.

1.3.1 Hipotezy pracy

Przedstawione zagadnienia badawcze, dostrzegane wyzwania, a także cele naukowe oraz motywacja wdrożeniowa umożliwiają sformułowanie hipotez naukowych rozprawy, które stanowią podstawę dalszych prac. Ich weryfikacja uzyskana zostanie przy użyciu opracowanych prototypów oraz końcowych eksperymentów.

- **Hipoteza główna:** Zastosowanie hybrydowego systemu wizyjno-radiowego do analizy cyfrowego modelu inteligentnego środowiska domowego pozwala na usprawnienie procesu odzyskania semantycznego znaczenia jego komponentów poprzez automatyzację ich rozpoznawania i pozycjonowania.
- **Hipoteza dodatkowa 1:** Zastosowanie cyfrowego modelu środowiska badawczego pozwala na radiową estymację dystansu pomiędzy punktami w przestrzeni bryły budynku o równie wysokiej jakości, co jej wyznaczenie bazując na pomiarach manualnych.
- **Hipoteza dodatkowa 2:** Przetwarzanie wstępne zbiorów zdjęć, polegające na doborze południka rozcięcia sfery (punktu początkowego w procesie rzutowania obrazów sferycznych na reprezentację panoramiczną), pozwala na zwiększenie efektywności wyszkolonego modelu detektora obiektów i jego działania w środowisku domowym.

1.4 Układ i metodyka pracy

Całość rozprawy podzielona została na sześć rozdziałów. W niniejszym omówiono już kontekst, motywacje, cele, hipotezy i pytania badawcze oraz propozycje rozwiązań.

W rozdziale drugim (Roz. 2) przedstawiona zostaje analiza literatury oraz badania rynku i oczekiwań klientów. W analizie literatury, przy pomocy metodyki przeglądu systematycznego, nakreślono obecny stan wiedzy w przedmiocie. W części badań rynkowych zastosowano analizę strategiczną i omówiono oferowany produkt partnera przemysłowego doktoratu, jego charakterystykę i dostrzeganą konkurencję. Do oceny potrzeb użytkowników zastosowano metodę sondażu diagnostycznego i technikę ankiety, której rezultaty skonfrontowano ze stanem literatury.

W rozdziale trzecim (Roz. 3) przedstawiono zagadnienie systemów radiowych. Jako pierwszy nakreślony został kontekst, szczególnie w aspekcie radiowego wyznaczania dystansu oraz zastosowania metody trilateracji w zadaniu pozycjonowania przestrzennego. Następnie omówione zostały elementy autorskie: przygotowane środowisko badawcze, przeprowadzone badania porównawcze systemów radiowej estymacji odległości z użyciem różnych technologii i protokołów komunikacyjnych, a także końcowy skonstruowany prototyp systemu radiowego. Skupiono się na określeniu silnych i słabych stron technologii BLE oraz UWB, które warto było wziąć pod uwagę przy opracowywaniu końcowego systemu fuzyjnego, a także na ocenie przydatności protokołu Matter do koordynacji wymiany wiadomości w warunkach heterogenicznej infrastruktury IoT nieruchomości. Od tego rozdziału wiodącą metodą badawczą była metoda eksperymentalna, a prowadzone prace skupiały się na projektowaniu, przeprowadzaniu i wnioskowaniu z eksperymentów.

Rozdział czwarty (Roz. 4) stanowi analogię rozdziału trzeciego, ale dedykowaną systemom wizyjnego rozpoznawania i lokalizowania obiektów. Ponownie, jako pierwszy opisany został kontekst prac: sztuczne sieci neuronowe, miary skuteczności detektorów obiektów, sposób reprezentacji obrazów sferycznych, dostępne publiczne zbiory zdjęć sferycznych oraz proces tworzenia zbiorów własnych. Następnie przedstawiono elementy autorskie. Rozpoczęto od badań skuteczności detektorów obiektów z rodziny YOLO (jako powszechnie dostępnego rozwiązania) w przetwarzaniu zdjęć panoramicznych. Później omówiono opracowany algorytm doboru najlepszego południka sfery do rozpoczęcia procesu rzutowania na obraz panoramiczny oraz badania jego wpływu na wyniki detekcji obiektów. Następnie przedstawiono zgromadzone autorskie zbiory danych zdjęciowych i badania próby ich zastąpienia obrazami generowanymi syntetycznie. Rozdział zamyka przedstawienie stworzonego prototypu systemu wizyjnego, w ramach którego skupiono się na analizie i konfiguracji parametrów procesu estymacji głębi na zdjęciach.

Rozdział piąty (Roz. 5) poświęcony jest już całkowicie omówieniu pracy własnej. Przedstawiona zostaje w nim synteza prototypów obu podsystemów w jeden większy system fuzyjny, a przede wszystkim stanowiąca esencję całej rozprawy hybrydowa metoda rozpoznawania i pozycjonowania komponentów środowiska domowego. Na jej opis składają się kolejne fragmenty, dotyczące takich zagadnień, jak: skonstruowane urządzenia lokalizacyjne, przygotowany cyfrowy model śro-

dowiska eksperymentalnego (typowych warunków domowych, tj. dwupokojowego mieszkania z kuchnią i łazienką), a przede wszystkim poszczególne etapy realizacji. Opis zamykają rozbudowane badania eksperymentalne proponowanej metody hybrydowej i analiza wyników jakości jej działania.

Rozdział szósty (Roz. 6) stanowi zwieńczenie rozprawy i podsumowanie jej wyników. Umieszczono w nim weryfikację hipotez pracy, krótko przedstawiono aspekt wdrożeniowy całego projektu, a także omówiono jego potencjalny dalszy rozwój i skrótowo nakreślono dotychczasowy dorobek oraz aktywność naukową autora.

W pracy stosowane jest możliwie jednolite nazewnictwo i skróty dla wprowadzanych pojęć. Zwroty techniczne, tam gdzie było to możliwe, zostały przetłumaczone na język polski. Zagadnienia, dla których nie odnaleziono polskich odpowiedników lub dla których naturalniejszym i powszechnie przyjętym określeniem jest słownictwo anglojęzyczne, używane są w oryginalnej postaci. Zestawienie wykorzystanych w rozprawie pojęć, ich skrótów i nazw angielskich umieszczono na końcu pracy (Tab. 6.3).

Rozdział 2

Analiza zagadnienia

Digitalizacja środowiska wewnętrznego (z ang. *Indoor Environment*, IE) jest odpowiedzią na różnorodne potrzeby cyfryzacji i modelowania świata rzeczywistego. Efektywne wdrożenie opracowanych rozwiązań wymaga ich dobrego zrozumienia i usystematyzowania. Przemysł oczekuje optymalizacji procesów produkcyjnych, ułatwienia zarządzania przestrzenią, a także monitorowania posiadanych zasobów, w tym nieruchomości wielkopowierzchniowych i magazynowych. Potrzeby użytkowników indywidualnych obejmują poprawę wygody życia codziennego i proste, małoskalowe usprawnienia codziennych obowiązków. Różnorodność wymagań przekłada się na szeroki wachlarz dostępnych technik gromadzenia danych, reprezentacji zasobów i procesów, a także sposobów rozumienia przestrzeni i rozpoznawania jej funkcji. W ramach niniejszej rozprawy wspólny pozostaje natomiast cel, tj. rozpoznanie semantycznie wartościowych typów i ustalenie położenia posiadanych zasobów w surowym modelu inteligentnego środowiska domowego: jego cyfrowym bliźniaku (z ang. *Digital Twin*, DT).

Aby ten cel osiągnąć, w niniejszym rozdziale przedstawiono analizę zagadnienia z perspektywy naukowej i przemysłowej. Rozdział otwarto systematycznym przeglądem literatury (Roz. 2.1), zrealizowanym według określonego protokołu postępowania (Roz. 2.1.1). Jego wyniki opracowano w postaci trzech taksonomii: typów danych wejściowych (Roz. 2.1.2), kategorii zastosowanej metody (Roz. 2.1.3) oraz realizowanego zadania (Roz. 2.1.4). Przegląd domknęto analizą dostrzeganych wyzwań (Roz. 2.1.5) i podsumowaniem (Roz. 2.1.6). W drugiej części ujęto

badania rynku i oczekiwań klientów (Roz. 2.2). Poprzedzono je rozeznaniem produktu partnera przemysłowego doktoratu (Roz. 2.2.1), analizą jego potencjału (Roz. 2.2.2) oraz oceną konkurencji (Roz. 2.2.3). Następnie przedstawiono badania sondażowe potrzeb i oczekiwań klientów końcowych (Roz. 2.2.4), które zwieńczono podsumowaniem (Roz. 2.2.5).

Na potrzeby niniejszej rozprawy, w zakresie analizy literatury przedmiotu, przyjęto podejście dwustopniowe. Prezentowany dalej przegląd systematyczny miał na celu objęcie możliwie szerokiego wachlarza zagadnień na wysokim poziomie ogólności. Jego osią przewodnią są klasyfikacja, segmentacja oraz analiza większych komponentów przestrzeni, w szczególności całych pomieszczeń. Rozpoznanie typu i układu pomieszczenia było istotne z punktu widzenia pozycjonowania wstępnego oraz uzyskania globalnego kontekstu organizacji przestrzeni. Uzyskane w ten sposób rozeznanie okazało się jednak na tyle rozbudowane, że obejmowało również potrzeby precyzyjnego pozycjonowania obiektów obecnych w środowisku na poziomie poszczególnych obszarów, a także wykorzystywane w tym celu rozwiązania i technologie. Drugi stopień analizy, tj. rozeznanie pogłębione, prezentowane jest więc jedynie zgodnie z potrzebami, odrębnie dla rozwiązań wybranych do implementacji i w ramach opisu ich osadzenia w kontekście: dla systemów radiowych w Roz. 3.1 oraz dla systemów wizyjnych w Roz. 4.1.

Podsumowując, wkład rozdziału w przeprowadzone prace naukowe i wdrożeniowe to:

- autorski protokół systematycznego przeglądu literatury, zapewniający powtarzalność doboru źródeł i transparentność ich filtracji;
- trzy taksonomie organizujące uzyskane wyniki i wskazujące rodzaje rozwiązań do zastosowania na kolejnych etapach prac;
- analiza przemysłowa produktu, jego potencjału i konkurencji, ukierunkowująca decyzje rozwojowe firmy na realne wymagania rynku;
- lista klas obiektów i rodzajów pomieszczeń, oczekiwanych przez użytkowników końcowych do rozpoznawania w sposób automatyczny (autorski zbiór etykiet klas wykorzystany w dalszych badaniach);
- matryca typowości obiektów w poszczególnych pomieszczeniach, pomagająca ustalać priorytety i kierunki dalszego rozwoju proponowanych rozwiązań.

2.1 Analiza literaturowa

Na potrzeby przeglądu bibliograficznego postanowiono wykorzystać ideę Systematycznego Przeglądu Literatury (z ang. *Systematic Literature Review*, SLR)[1–3]. Metodyka ta opisuje kolejne kroki postępowania i planowania badań tak, aby zminimalizować wpływ subiektywnej opinii badacza na końcowy dobór analizowanych materiałów. Zakłada ona zdefiniowanie konkretnych pytań badawczych (na które szukamy odpowiedzi), źródeł wiedzy (z których będziemy korzystać), warunków filtracji publikacji (pozwalających na odrzucanie niedopasowanych opracowań) oraz całego protokołu postępowania (opisującego wykonanie prac krok po kroku)[4]. W kolejnych podrozdziałach elementy te zostają przedstawione i omówione, włączając w to analizę wyników. Ich autorskie opracowanie do postaci taksonomii stanowi istotny wkład w literaturę przedmiotu, w którym odnaleziono tylko pojedyncze przekrojowe analizy dostępnych rozwiązań [5–7].

2.1.1 Metodologia przeglądu literatury

Do sprecyzowania poszukiwanych zagadnień wykorzystane zostały trzy proste frazy, które celnie opisują esencję prezentowanego przeglądu literatury: **co?**, **jak?** i **po co?** Frazy te rozwinięto, a na ich podstawie sformułowane zostały trzy właściwe pytania badawcze (z ang. *Research Question*, RQ):

- **RQ1: Jakie struktury danych wejściowych są wykorzystywane?**
- **RQ2: Jakie ogólne kategorie przetwarzania są stosowane?**
- **RQ3: Jakie konkretne zadania realizują analizowane systemy?**

Aby na te pytania odpowiedzieć, skonstruowano zapytanie wyszukiwania (z ang. *Search Query*, SQ), wykorzystane następnie w repozytoriach literatury. W celu zrównoważenia istotności odnalezionych artykułów oraz ich liczby, zaproponowano dwie składnie konstrukcji zapytania (prezentowane w oryginalnej postaci w języku angielskim): *room [activity]* oraz *floor plan [activity]*. W pierwszej grupie użyto następujących słów określających aktywności: *segmentation*, *classification*, *clustering*, *structure*, *detection*, *recognition*, oraz *labeling*. W drugiej grupie, skupionej na planach pomieszczeń, ograniczono się do: *segmentation*, *clustering* i *analysis*.

Celem takiej konstrukcji było zebranie publikacji opisujących różnorodne typy danych wejściowych (grupa “*room*”) z jednoczesnym naciskiem na przetwarzanie płaskich reprezentacji nieruchomości (grupa “*floor plan*”).

Tak opracowane zapytanie skierowano do sześciu renomowanych baz danych publikacji naukowych:

- ACM Digital Library,
- IEEE Xplore,
- DBLP (Digital Bibliography & Library Project),
- Scopus,
- Elsevier ScienceDirect,
- SpringerLink.

W procesie filtrowania odnalezionych publikacji zastosowano określone kryteria inkluzywne (z ang. *Inclusion Criteria*, IC) oraz kryteria wykluczające (z ang. *Exclusion Criteria*, EC). Kryteria IC określały warunki, które musiały być spełnione (np. zastosowanie w środowisku wewnętrznym), natomiast kryteria EC definiowały cechy, które dyskwalifikowały publikację (np. język inny niż angielski, brak recenzji). Szczegóły wszystkich kryteriów przedstawiono w Tab. 2.1.

Wszystkie wymagania, źródła danych oraz kryteria filtrowania połączono w jeden spójny protokół przeglądu literatury. Proces został podzielony na dwa etapy: główny oraz uzupełniający. Etap główny obejmował analizę publikacji pozyskanych bezpośrednio z baz danych. Etap uzupełniający skupiał się na przetwarzaniu obrazów 2D (jako subiektywnie łatwych do pozyskania i stanowiących podstawę działania wielu algorytmów detekcji obiektów) oraz dodawał publikacje odnalezione jako cytowane w bazowym zestawie artykułów.

Tabela 2.1: Zestawienie zastosowanych w procesie przeglądu literatury kryteriów inkluzywnych (IC) i wykluczających (EC) wraz z ich opisami.

Nr	Tytuł	Opis
EC1	Publikacja sprzed 2012	Uwzględniano jedynie publikacje z ostatniej dekady, aby zachować aktualność przeglądu.
EC2	Zduplikowany artykuł	Artykuły znalezione w wielu bazach mogły się powtarzać. Każdy tekst analizowano tylko raz.
EC3	Język inny niż angielski	Akceptowano wyłącznie język angielski.
EC4	Brak powiązania z segmentacją lub klasyfikacją pomieszczeń	Publikacje musiały dotyczyć zagadnień związanych z analizą planów, danymi przestrzennymi lub problemami mogącymi prowadzić do segmentacji przestrzeni.
EC5	Brak pełnego tekstu	Artykuły bez pełnej treści były odrzucane.
EC6	Brak opisu metodyki	Publikacje zawierające wyłącznie wyniki, bez opisu zastosowanej metody, nie były uwzględniane.
EC7	Artykuły koncepcyjne lub dyskusyjne	Uwzględniano tylko teksty zawierające implementację, testy i wyniki.
IC1	Zastosowanie w środowisku wewnętrznym	Rozwiązania musiały dotyczyć przestrzeni zamkniętych, a nie analiz terenów otwartych.
IC2	Automatyczne przetwarzanie danych	Uwzględniano tylko rozwiązania wykorzystujące automatyczne lub półautomatyczne przetwarzanie.
IC3	Co najmniej 10 cytowanych prac	Aby zapewnić wiarygodność, publikacja musiała odwoływać się do minimum dziesięciu innych źródeł.
IC4	Analiza na poziomie pomieszczeń lub wyżej	Wymagano, by system potrafił analizować przynajmniej jedną jednostkę przestrzenną odpowiadającą pomieszczeniu.
IC5	Opis wyników i zbioru danych	Artykuły musiały zawierać informację o użytym zbiorze danych oraz metodzie oceny skuteczności.

Cała procedura składała się z dwunastu kroków: dziesięciu w głównym przepływie oraz dwóch w rozszerzeniu przeglądu, wykonywanych sekwencyjnie.

Krok 1: Wyszukiwanie w bazach danych. Zapytanie wyszukiwania skierowano do wszystkich sześciu baz danych. Bez wprowadzenia ograniczenia na rok publikacji otrzymano ponad 3000 wyników.

Krok 2: Filtrowanie po atrybutach. Wprowadzono pierwsze kryterium wykluczające (EC1), dotyczące daty publikacji. Liczba wyników została zredukowana o ok. 35%, do analizy pozostawało ponad 2000 artykułów.

Krok 3: Scalanie wyników i usunięcie duplikatów. Wszystkie wyniki połączono w jeden zestaw, z którego usunięto duplikaty (EC2). Pozostawiono do analizy 1854 unikalnych rekordów.

Krok 4: Selekcja na podstawie tytułów. Sprawdzono język (EC3) oraz zgodność tematyczną (EC4, IC1). Usunięto ponad dwie trzecie rekordów. Pozostawiono do analizy 634 publikacje.

Krok 5: Selekcja na podstawie streszczeń. Streszczenia oceniono pod względem języka i tematu (EC3, EC4, IC1) oraz obecności automatyzacji (IC2). Pozostawiono do analizy 402 prace.

Krok 6 i 7: Dostępność oraz analiza pełnych tekstów. Pozyskano całość publikacji (EC5), a następnie oceniono zgodnie z wcześniejszymi kryteriami. Pozostawiono do analizy 172 artykuły.

Krok 8: Grupowanie według typu danych wejściowych. Uzyskano odpowiedź na pytanie RQ1. Artykuły podzielono według rodzaju danych wejściowych: obrazy 2D, dane przestrzenne, struktury grafowe, artykuły przeglądowe oraz zbiory cech. Największą grupę stanowiły obrazy płaskie, tj. 87 publikacji.

Krok 9: Grupowanie według kategorii rozwiązania. Uzyskano odpowiedź na pytanie RQ2. Artykuły przypisano do czterech kategorii według ogólnego typu zastosowanej metody: samej segmentacji, segmentacji z klasyfikacją uproszczoną, segmentacji z klasyfikacją precyzyjną albo samej klasyfikacji precyzyjnej.

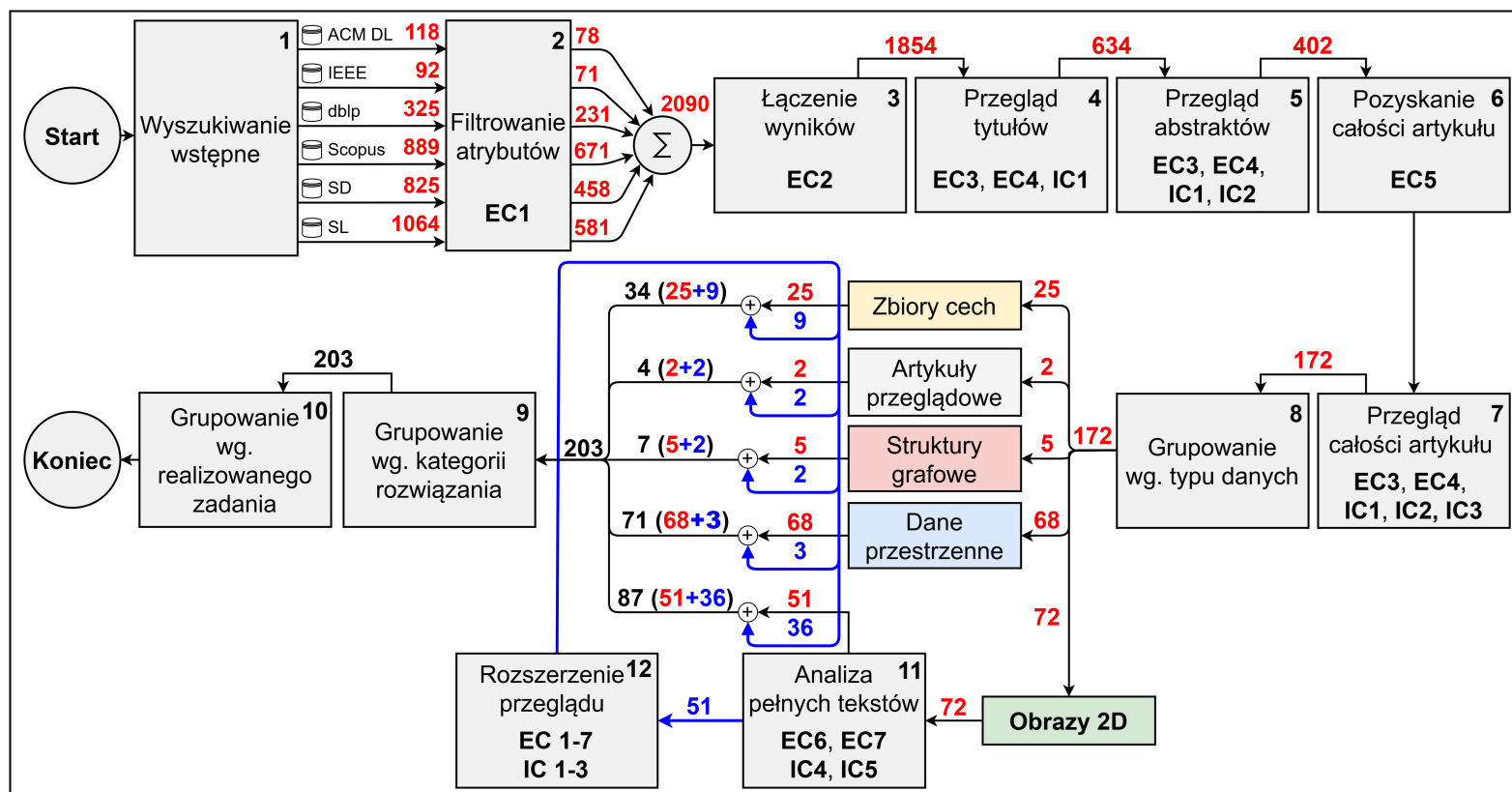
Krok 10: Grupowanie według realizowanego zadania. Uzyskano odpowiedź na pytanie RQ3. Publikacje przypisano do konkretnych zastosowań, takich jak rekonstrukcja modelu trójwymiarowego czy klasyfikacja pomieszczeń.

Krok 11: Szczegółowa analiza pełnego tekstu. Rozszerzenie przeglądu. Artykuły przetwarzające obrazy 2D poddano głębszej analizie pod kątem zgodno-

ści z kryteriami EC6 i EC7 oraz IC4 i IC5. Spośród 72 prac, 51 spełniło wymagania.

Krok 12: Rozszerzenie przeglądu. Ostatni krok protokołu przeglądu miał na celu uzupełnienie zbioru publikacji o artykuły istotne, powiązane z tematem, a przy tym cytowane w analizowanych pracach jako źródła bazowe lub kluczowe odniesienia. Ten krok wykonano tylko raz, aby uniknąć rekurencji.

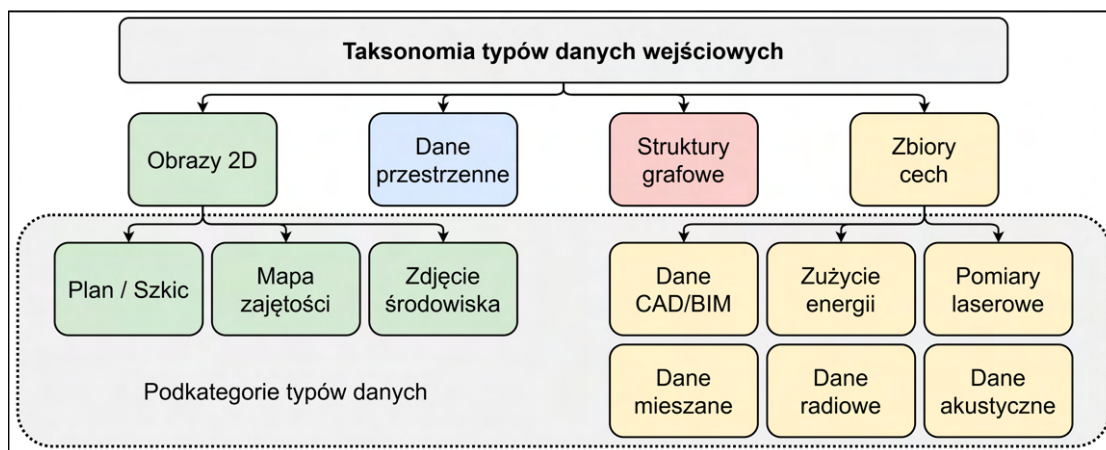
Całość protokołu została przedstawiona w formie schematu na Rys. 2.1. Poszczególne bloki odpowiadają kolejnym krokom, zawierają ich numer, nazwę oraz zastosowane kryteria filtrowania (jeśli wystąpiły). Kroki główne oznaczono kolorem czerwonym, a rozszerzenie kolorem ciemnoniebieskim. Na skutek wykonania przedstawionego protokołu analizy literatury opracowane zostały trzy taksonomie, opisane w kolejnych rozdziałach.



Rysunek 2.1: Schemat protokołu przeglądu literatury z informacją o liczbie publikacji po każdym kolejnym kroku. Czerwone liczby oznaczają główny przebieg protokołu, a niebieskie jego rozszerzenie. Wprowadzono kolorystyczne oznaczenia typów danych wejściowych: obrazy 2D (zielony), dane przestrzenne (jasnoniebieski), struktury grafowe (czerwony), artykuły przeglądowe (szary) oraz zbiory cech (żółty).

2.1.2 Taksonomia typów danych wejściowych

Pierwsza zaproponowana taksonomia porządkuje publikacje według rodzaju przetwarzanych danych wejściowych. Przedstawiono ją na Rys. 2.2. Jako że struktura danych wpływa zarówno na dobór sprzętu pomiarowego, jak i metod analizy, klasyfikacja ta pozwala lepiej zrozumieć kontekst i ograniczenia stosowanych rozwiązań.



Rysunek 2.2: Diagram pierwszej opracowanej taksonomii: grupowanie publikacji na podstawie typów przetwarzanych danych.

Z analizy wykluczono prace przeglądowe, ponieważ nie odnoszą się one bezpośrednio do żadnej konkretnej struktury danych. Zidentyfikowane typy danych wejściowych podzielono na cztery główne grupy: obrazy 2D, dane przestrzenne, struktury grafowe, zbiory cech. W ramach obrazów 2D sprecyzowano trzy, a dla zbiorów cech sześć podtypów danych, co łącznie dało jedenaście wyróżnionych struktur. Szczegółowe zestawienie publikacji przypisanych do odpowiednich kategorii przedstawiono w Tab. 2.2.

Obrazy 2D obejmują wszystkie dane wejściowe zapisane w formie obrazów, skanów, zdjęć lub szkiców. Grupa ta zawiera zarówno symboliczne reprezentacje przestrzeni (np. plany pięter), jak i realistyczne fotografie środowiska wewnętrznego. Kluczową cechą jest możliwość przetwarzania obrazu jako macierzy pikseli z przypisanymi wartościami. Do tej kategorii zaliczono: obrazy czarno-białe, kolorowe (RGB), mapy zajętości oraz obrazy z dodatkowym kanałem głębi (RGBD). W ramach tej grupy wyróżniono trzy podtypy:

Tabela 2.2: Zestawienie artykułów według rodzaju danych wejściowych. Liczby w nawiasach oznaczają liczbę artykułów uwzględnionych w głównym przeglądzie (pierwsza liczba) oraz w jego rozszerzeniu (druga liczba).

Typ danych	Podtyp danych	Publikacje
Dane przestrzenne (68 + 3)	-	[8–78]
Struktury grafowe (5 + 2)	-	[79–85]
Obrazy 2D (51 + 36)	Plan / Szkic (26 + 15)	[86–126]
	Mapa zajętości (17 + 6)	[5, 127–148]
	Zdjęcie środowiska (8 + 15)	[149–171]
Zbiory cech (25 + 9)	Dane CAD/BIM (1 + 0)	[172]
	Zużycie energii (2 + 0)	[173, 174]
	Pomiary laserowe (5 + 0)	[175–179]
	Dane mieszane (10 + 9)	[180–198]
	Dane radiowe (2 + 0)	[199, 200]
	Dane akustyczne (5 + 0)	[201–205]

- Plany i szkice: skany planów architektonicznych, rysunki odręczne;
- Mapy zajętości: dane z czujników laserowych, np. robotów sprzątających;
- Zdjęcia środowiska: fotografie wnętrz, panoramy, klatki z nagrań wideo.

Dane przestrzenne 3D obejmują wszystkie struktury, które odnoszą się do położenia w przestrzeni. Dominującym przypadkiem są chmury punktów (np. ze skanerów typu LiDAR), ale do grupy tej zaliczono także modele siatkowe 3D, trajektorie urządzeń (np. z czujników inercyjnych) oraz obrazy RGBD wykorzystywane nie jako matryce pikseli, lecz jako źródło do budowy punktów w przestrzeni.

Struktury grafowe obejmują dane, w których elementy przestrzeni (np. pomieszczenia, korytarze) są przedstawione jako wierzchołki grafu, a relacje między nimi (np. sąsiedztwo, połączenia) jako krawędzie. Warunkiem przypisania rozwiązania do tej grupy było jawne wykorzystanie teorii grafów jako metody przetwarzania danych.

Zbiory cech to najbardziej zróżnicowana grupa, obejmująca dane nieobrazowe i niemające formy przestrzennej. Wyróżniono sześć podtypów:

- Dane CAD/BIM: modele eksportowane z oprogramowania projektowego,
- Zużycie energii: dane z liczników energii w budynku,
- Pomiary laserowe: dane o dystansie do przeszkód (inne niż mapy zajętości),

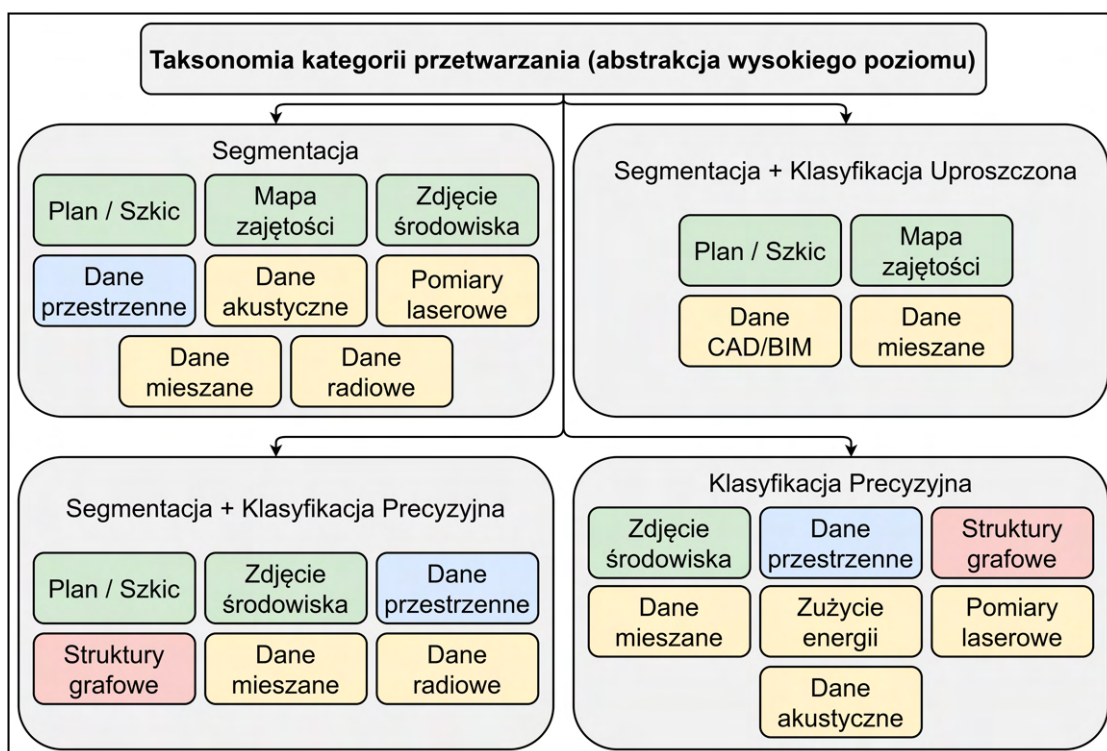
- Dane mieszane: zestaw różnych parametrów, np. zdjęcie fasady i rozmiar pomieszczenia,
- Dane radiowe: np. siła sygnału z nadajników w odniesieniu do lokalizacji,
- Dane akustyczne: pomiary natężenia dźwięku, echa, dedykowanych sygnałów dźwiękowych.

Stworzona taksonomia pokazuje, że zakres wykorzystywanych danych i ich złożoność jest właściwie dowolna. Rozciąga się ona od prostych szeregów liczbowych (jednowymiarowych) i macierzy (dwuwymiarowych), przez dane przestrzenne (trzy wymiary), aż po złożone kombinacje różnych informacji zawierających dowolną liczbę cech. Tak duża różnorodność umożliwia dobór struktury danych odpowiedniej do możliwości sprzętowych danego systemu. Najprostsze urządzenia rejestrują tylko podstawowe dane, np. zużycie energii lub siłę sygnału radiowego. Wraz ze wzrostem mocy obliczeniowej czujnika rośnie też złożoność rejestrowanych danych. Możliwe staje się przetwarzanie obrazów, strumieni wideo czy danych ze skanerów przestrzeni. Warto zwrócić uwagę na ewolucję podejścia: od rejestrowania danych offline i ich późniejszego przetwarzania po analizę na urządzeniu i przesyłanie do chmury jedynie gotowych wniosków.

2.1.3 Taksonomia kategorii przetwarzania

Druga zaproponowana taksonomia porządkuje analizowane publikacje według ogólnej kategorii sposobu przetwarzania danych. Taki podział pozwala zrozumieć, jakiego rodzaju operacje były realizowane i jaki poziom semantycznego zrozumienia danych uzyskano. Artykuły korzystające z tego samego typu danych mogły zostać przypisane do różnych kategorii – w zależności od celu badania. Zachowano ciągłość nazewnictwa oraz identyfikacji kolorystycznej, jak w poprzedniej taksonomii. Diagram ilustrujący opracowany podział zaprezentowano na Rys. 2.3.

Wyróżnione zostały cztery kategorie rozwiązań: segmentacja (bez rozpoznania typu pomieszczeń), segmentacja z uproszczoną klasyfikacją (podstawowe rozróżnienie typu przestrzeni, np. pokój / korytarz), segmentacja z klasyfikacją precyzyjną (rozpoznanie typu pomieszczenia, np. kuchnia, łazienka), klasyfikacja precyzyjna (bez segmentacji przestrzennej). Granica między kategoriami wynika z poziomu semantycznego przypisanego do analizowanych danych: od czystej struktury geo-



Rysunek 2.3: Diagram drugiej opracowanej taksonomii: grupowanie publikacji na podstawie kategorii sposobu przetwarzania danych.

metrycznej po konkretne znaczenie funkcjonalne przestrzeni. Zestawienie z przypisaniem publikacji do typu danych oraz kategorii przetwarzania przedstawiono w Tab. 2.3.

Segmentacja obejmuje metody, które identyfikują określone struktury przestrzenne (np. pokoje) w danych źródłowych, ale bez przypisywania im konkretnego znaczenia. Segmentowane obiekty traktowane są jako jednostki równoważne, przykładowo: *pokój1*, *pokój2*, *pokój3*. Często wykorzystywaną metodą w tej grupie jest klasteryzacja.

Segmentacja z klasyfikacją uproszczoną rozszerza powyższe podejście o możliwość rozróżnienia pomieszczeń na podstawowe typy. Choć nie dochodzi jeszcze do pełnego przypisania znaczenia, możliwe jest rozdzielenie przestrzeni na klasy takie jak: *pomieszczenie*, *korytarz*, *przestrzeń zewnętrzną*.

Segmentacja z precyzyjną klasyfikacją obejmuje rozwiązania, które nie tylko segmentują przestrzeń na poszczególne pomieszczenia, ale także potrafią przy-

Tabela 2.3: Zestawienie artykułów według rodzaju danych wejściowych i kategorii przetwarzania, kolejno: segmentacja (S), segmentacja z klasyfikacją uproszczoną (S + KU), segmentacja z klasyfikacją precyzyjną (S + KP) oraz klasyfikacja precyzyjna (KP).

Typ danych	Kategoria przetwarzania			
	S	S + KU	S + KP	KP
Plan/Szkic	[86–88, 90–97, 103, 108, 113, 116, 118, 120–122, 124, 125]	[102, 106, 110, 126]	[89, 98–101, 104, 105, 107, 109, 111, 112, 114, 115, 117, 119, 123]	-
Mapa zajętości	[5, 129–132, 134, 135, 137–139, 142, 145, 146, 148]	[127, 128, 133, 136, 140, 141, 143, 144, 147]	-	-
Zdjęcie środowiska	[151, 153, 154, 170]	-	[156]	[149, 150, 152, 155, 157–169, 171]
Dane przestrzenne	[8–14, 16–21, 23, 25–32, 34–49, 51–55, 57–68, 70, 71, 74–77]	-	[22, 24, 33, 50, 56, 69, 78]	[15, 72, 73]
Pomiary laserowe	[176, 177]	-	-	[175, 178, 179]
Dane mieszane	[181–185]	[189]	[180, 187, 190–195, 197, 198]	[186, 188, 196]
Dane radiowe	[200]	-	[199]	-
Dane akustyczne	[204]	-	-	[201–203, 205]
Dane CAD/BIM	-	[172]	-	-
Struktury grafowe	-	-	[79, 80, 82, 83]	[81, 84, 85]
Zużycie energii	-	-	-	[173, 174]

pisać im konkretne funkcje. W tym przypadku każdy wykryty fragment przestrzeni otrzymuje znaczenie semantyczne, przykładowo: *kuchnia*, *salon*, *sypialnia*, *łazienka*. Zakres dostępnych klas różnił się w zależności od badanego rozwiązania.

Klasyfikacja precyzyjna grupuje metody, które realizują klasyfikację wieloklasową na podstawie całego zestawu danych wejściowych, bez wcześniejszego etapu segmentacji. W przeciwieństwie do metod omawianych wcześniej, klasyfikacja dotyczy całości danych i nie wydziela się w niej fragmentów przestrzeni. Przykładem może być rozpoznanie typu pomieszczenia na podstawie pojedynczego zdjęcia środowiska. Wynikiem jest przypisanie klasy do całego obrazu, przykładowo: *“to jest zdjęcie sypialni”*.

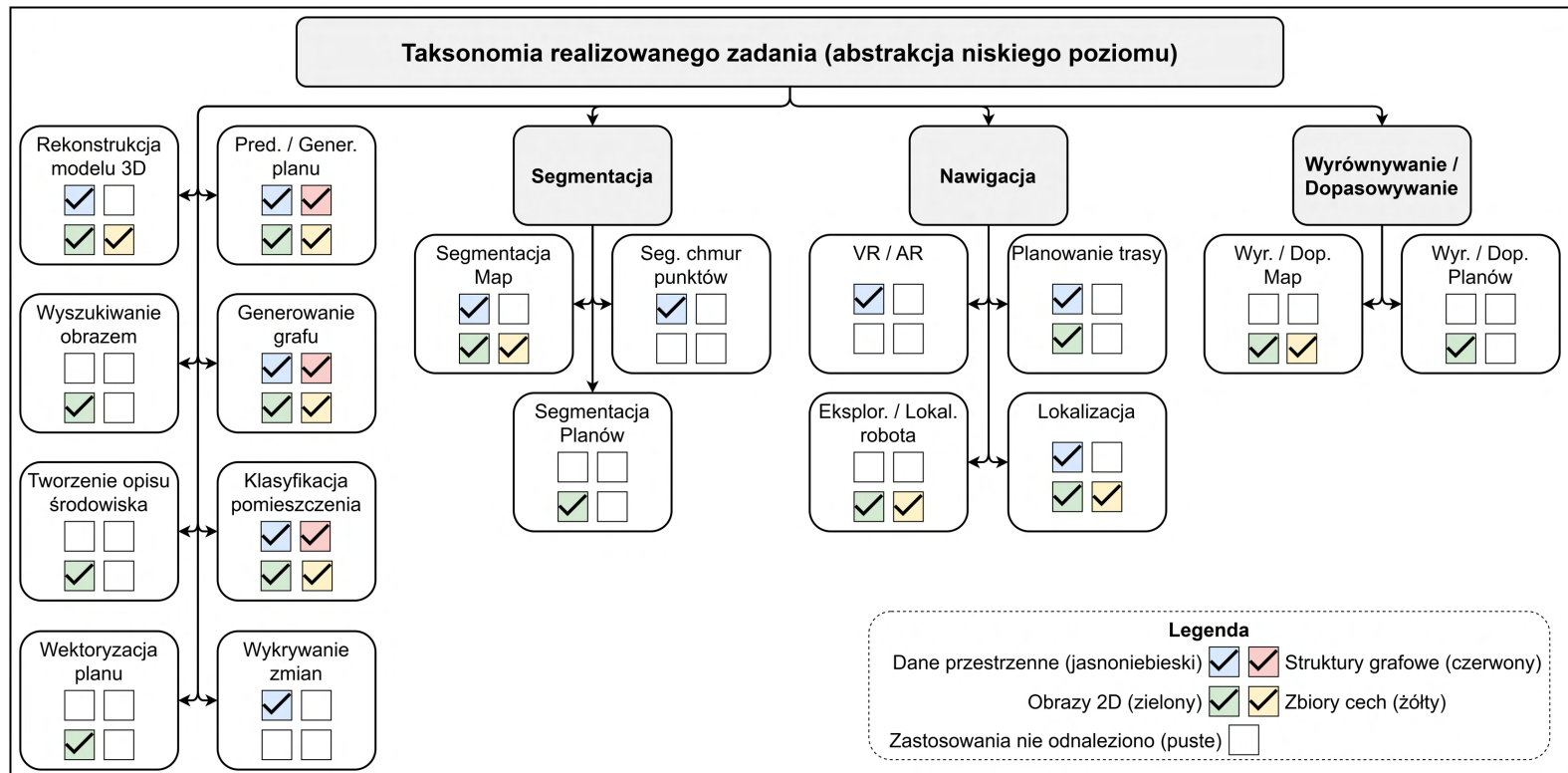
Analiza metod przetwarzania ujawniła wyraźne zależności pomiędzy charakterem zadania, typem danych wejściowych, stosowanymi technikami oraz wymaganiami sprzętowymi. W przypadku prostych zadań segmentacji, zwłaszcza opartych na danych abstrakcyjnych (np. mapach zajętości), dominowały klasyczne algorytmy, takie jak klasteryzacja czy metody oparte na grafach. Jednak wraz ze wzrostem złożoności danych oraz potrzebą uzyskania wyników semantycznych, rosło znaczenie rozwiązań opartych na uczeniu maszynowym, które potrafią wykrywać ukryte wzorce i lepiej radzą sobie z rzeczywistymi danymi, takimi jak zdjęcia czy chmury punktów. Coraz większa popularność metod AI, widoczna szczególnie w nowszych pracach, wynika nie tylko z ich wyższej odporności na zakłócenia, ale też z możliwości generowania stabilnych wyników przy niższej jakości danych. Wybór podejścia jest również silnie powiązany z dostępem do zasobów obliczeniowych: proste metody mogą być uruchamiane na urządzeniach o ograniczonej mocy, natomiast przetwarzanie złożonych zbiorów danych często wymaga dedykowanego, wydajnego sprzętu.

2.1.4 Taksonomia realizowanych zadań

Trzecia opracowana taksonomia porządkuje analizowane prace według konkretnego zastosowania, w którym wykorzystano dany typ danych wejściowych. Dzięki temu zestawieniu można lepiej zrozumieć zakres praktycznych zastosowań oraz problemy, które były faktycznie rozwiązywane. Diagram taksonomii przedstawiono na Rys. 2.4. Zidentyfikowano jedenaście głównych zagadnień, z których trzy

zostały dodatkowo podzielone na mniejsze zadania. W sumie wyróżniono 17 typów realizowanych zadań. W Tab. 2.4 przedstawiono przypisania analizowanych publikacji do typu przetwarzanych danych oraz realizowanego zadania. Pozwala ona na szybkie odnalezienie referencji do opisu konkretnego zastosowania lub sprawdzenie, jakie zadania są możliwe do realizacji w oparciu o dany typ danych.

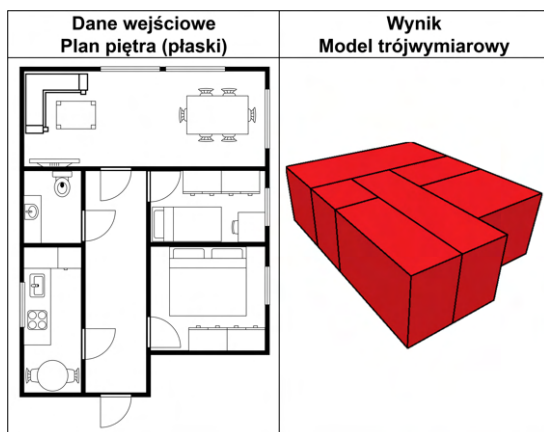
Dalsza część niniejszego rozdziału omawia proponowaną taksonomię na graficznych przykładach. Aby ułatwić zrozumienie różnic w realizowanych zadaniach, układ referencyjny jest możliwie jednakowy: przedstawia rzut mieszkania z salonem, dwoma sypialniami, kuchnią, łazienką i korytarzem. Taki układ danych został wybrany z trzech powodów: (1) głównym typem danych w tej pracy są obrazy 2D, (2) przykład ten może być zastosowany do każdego z opisywanych zadań, (3) jest subiektywnie najprostszy w zrozumieniu.



Rysunek 2.4: Diagram trzeciej opracowanej taksonomii: grupowanie publikacji na podstawie realizowanego zadania.

Tabela 2.4: Zestawienie artykułów według zadania i typu danych wejściowych.

Zadanie	Typ przetwarzanych danych			
	Dane przestrzenne	Obrazy 2D	Struktury grafowe	Zbiory cech
Rekonstrukcja modelu 3D	[9–12, 17, 18, 20, 21, 23, 26–30, 35, 36, 42, 45–49, 51, 52, 54, 55, 57, 60, 61, 63, 64, 74–76]	[94, 110, 111, 116, 154]	-	[195]
Wyszukiwanie obrazem	-	[97, 98, 114]	-	-
Generacja opisu środow.	-	[90, 95, 96, 99, 103, 107, 115, 119]	-	-
Wektoryzacja planu	-	[104, 117, 120, 122]	-	-
Gener./Pred. planu	[8, 25, 31, 33, 37, 44, 53, 58, 62, 68, 71, 78]	[89, 151, 153, 170]	[79, 83]	[77, 172, 176, 177, 183–185, 189, 204]
Gener. grafu	[13, 14, 19, 67]	[138, 139]	[80, 82]	[180, 190, 191, 194]
Klasyfikacja pokoiów	[15, 73]	[143, 149, 150, 152, 155, 157–169]	[81, 84, 85]	[173–175, 178, 179, 186, 188, 201–203, 205]
Detekcja zmian	[43]	-	-	-
Segmentacja mapy	[50]	[5, 82, 127, 128, 135, 136, 140–142, 144–147]	-	[192, 193, 197, 198, 206]
Segmentacja planu	-	[86–88, 91, 100–102, 105, 106, 108, 109, 112, 113, 118, 121, 123–126]	-	-
Segmentacja chmury pkt.	[16, 32, 34, 38, 39, 41, 59, 65, 66, 69, 207, 208]	-	-	-
VR/AR	[40]	-	-	-
Eksplo./Lokal. robota	-	[129, 131, 132, 134, 156, 171]	-	[182]
Planowanie trasy	[70]	[133]	-	-
Lokalizacja	[22, 56]	[93]	-	[24, 181, 196, 199, 200, 209]
Wyrówn./Dopas. map	-	[187]	-	[130, 148]
Wyrówn./Dopas. planów	-	[92]	-	-



Rysunek 2.5: Wizualizacja zadania rekonstrukcji modelu 3D. Na podstawie wejściowego obrazu 2D (planu pomieszczeń, **po lewej**) generowany jest uproszczony model trójwymiarowy (**po prawej**).



Rysunek 2.6: Wizualizacja zadania wyszukiwania obrazem. Obraz zapytania (**u góry**) służy do przeszukania bazy danych w celu znalezienia 3 najbardziej podobnych planów pięter (**na dole**).

Rekonstrukcja modelu 3D

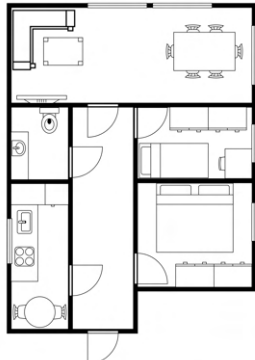
Zadanie to obejmowało generowanie przestrzennych modeli budynków z chmur punktów, obrazów 2D lub zestawów cech. Wykorzystano różne typy modeli: siatkowe [17], szczelne [61], BIM [27], czy wokselowe [64]. Wśród źródeł danych znalazły się skanery statyczne [30], mobilne [45], plecakowe [57], a także urządzenia AR, np. HoloLens [75]. Inne podejścia bazowały na obrazach planów pięter [110, 116] lub obrazach panoramicznych [154]. Zadanie przedstawiono graficznie na Rys. 2.5.

Wyszukiwanie obrazem

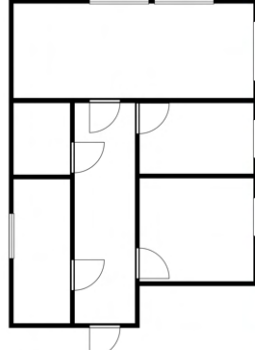
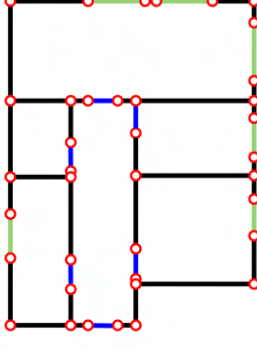
Mimo rosnącej popularności metod analizy obrazów, zagadnienie wyszukiwania podobnych rzutów kondygnacji na podstawie przykładowej grafiki było rzadko podejmowane. Wszystkie znalezione rozwiązania wykorzystywały analizę obrazu wejściowego, tworzyły jego reprezentację grafową i porównywały ją z bazą innych planów [97, 98]. W jednej z prac umożliwiono także użycie odrębnego szkicu jako zapytania [114]. Zauważa się potencjał rozwoju tej grupy metod, szczególnie w kontekście systemów rekomendacyjnych. Zadanie przedstawiono graficznie na Rys. 2.6.

Tworzenie opisu środowiska. Zadanie to obejmowało generowanie zrozumiałych opisów tekstowych analizowanych budynków. Rozwiązania te wspierały osoby niewidome w eksploracji otoczenia, planowaniu ścieżek lub poznawaniu układu budynku przed wizytą. Goncu i in. [90] stworzyli dotykowy interfejs z opisami planu piętra. Madugalla i in. [95, 103] rozbudowali go o opisy całych pięter i narzędzie korekty. Paladugu i in. [115] zaproponowali system przeglądarkowy. Goyal i in. [96, 99, 107, 119] rozwijali serię narzędzi do opisu i nawigacji, opartych m.in. o repozytorium planów BRIDGE. Zadanie przedstawiono graficznie na Rys. 2.7.

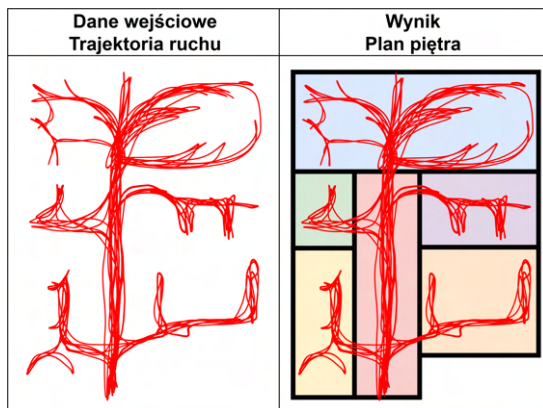
Wektoryzacja planów. Zadanie to polegało na konwersji rastrowych obrazów planów pięter na format wektorowy, ułatwiający skalowanie oraz dalsze przetwarzanie danych. Liu i in. [117] zaproponowali metodę "Raster-to-Vector", opartą na sieciach konwolucyjnych i programowaniu całkowitoliczbowym. Jang i in. [120] użyli m.in. sieci PSPNet, GCN i Deeplab v3. Surikov i in. [122] połączyli sieć UNet, Faster-RCNN i klasyczne algorytmy geometryczne. Dong i in. [104] opracowali sieć EdgeGAN, osiągając lepszą jakość i szybkość. Zadanie przedstawiono graficznie na Rys. 2.8.

Dane wejściowe Plan piętra	Wynik Opis tekstowy
	"Mieszkanie składa się z sześciu pomieszczeń. Pośrodku znajduje się długi korytarz. W prawym dolnym rogu mieści się sypialnia główna. Nad sypialnią główną, w prawym górnym rogu, znajduje się mniejsza sypialnia. Na końcu korytarza znajduje się duży salon. W lewym dolnym rogu znajduje się kuchnia, a nad nią, w lewym górnym rogu, mieści się toaleta."

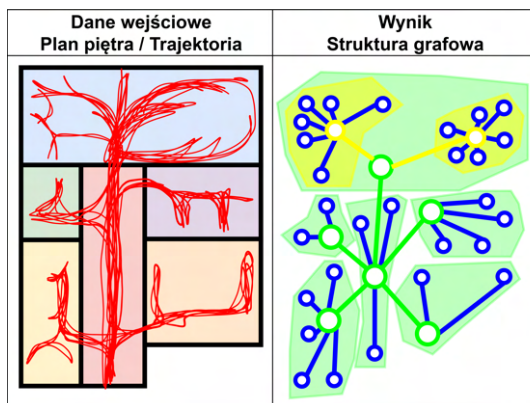
Rysunek 2.7: Wizualizacja zadania tworzenia opisu środowiska. Plan piętra wejściowego (**po lewej**) jest analizowany i przekształcany w zestaw zdań opisujących jego elementy oraz ich położenie (**po prawej**).

Dane wejściowe Obraz rastrowy	Wynik Obraz wektorowy
	

Rysunek 2.8: Wizualizacja zadania wektoryzacji planu pomieszczeń. Wejściowy obraz rastrowy (**po lewej**) zostaje przekształcony na reprezentację wektorową (**po prawej**). Przekształcone ściany oznaczono czarnymi liniami, drzwi niebieskimi, okna zielonymi, a połączenia czerwonymi okręgami.



Rysunek 2.9: Wizualizacja zadania generacji planu pomieszczeń. Dane wejściowe (zarejestrowana trajektoria użytkownika, **po lewej**) są wykorzystywane do oszacowania możliwego planu piętra (na podstawie gęstości ścieżek i ich klasteryzacji **po prawej**).



Rysunek 2.10: Wizualizacja zadania tworzenia grafu. Dane wejściowe (zarejestrowana trajektoria użytkownika wraz z pewnym zrozumieniem układu budynku, **po lewej**) są wykorzystywane do utworzenia struktury grafu (**po prawej**), w której pomieszczenia i obszary reprezentowane są jako węzły, a ich sąsiedztwo jako krawędzie grafu.

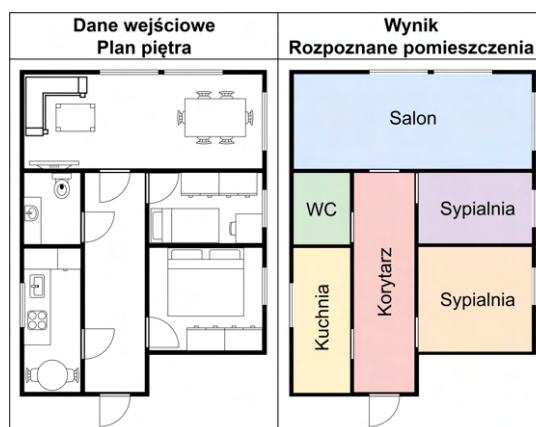
Predykcja/generowanie planu.

Zadanie polegało na tworzeniu planów pomieszczeń na podstawie różnych danych: chmur punktów, obrazów 360°, szkiców, trajektorii czy danych sensorycznych z urządzeń mobilnych. Niektóre prace wykorzystywały podejście proceduralne [89], inne analizowały dane przestrzenne w czasie rzeczywistym [77]. Coraz większą rolę odgrywały algorytmy uczenia maszynowego [58], przetwarzające dane z kamer, LiDAR czy czujników inercyjnych. Celem było uproszczenie tworzenia planów budynków i ograniczenie udziału człowieka w procesie. Zadanie przedstawiono graficznie na Rys. 2.9.

Generowanie grafu. W tym zadaniu środowisko wewnętrzne modelowano jako graf, gdzie wierzchołki reprezentowały pomieszczenia lub obserwacje, a krawędzie relacje przestrzenne. Grafy tworzone z trajektorii robotów [190], skanów 3D [13], obrazów RGBD [19], a nawet siatek Voronoi [139]. Na podstawie takiej struktury przewidywano podział przestrzeni lub typ pomieszczeń [82]. Nowsze podejścia łączyły przetwarzanie grafowe z sieciami neuronowymi [67], co pozwalało na semantyczne zrozumienie układu budynku. Zadanie przedstawiono graficznie na Rys. 2.10.

Klasyfikacja pomieszczeń. Zadanie polegało na określeniu funkcji pomieszczenia na podstawie różnych danych wejściowych. Najczęściej analizowano obrazy RGB i głębi, przetwarzane za pomocą sieci neuronowych [152, 155], klasycznych klasyfikatorów [149] lub metod hierarchicznych [178]. Rozwiązania uwzględniały przetwarzanie całych obrazów [169] lub ich fragmentów [150], a także dane topologiczne [84], akustyczne [202], sygnały radiowe i rozkłady zużycia energii [173]. Istotnym krokiem było tworzenie reprezentacji sceny jako zbioru obiektów, cech geometrycznych lub parametrów przestrzennych. Wzrost skuteczności klasyfikacji wspierany był również przez rozwój zbiorów danych (np. sieć Places-CNN [168]), oraz podejścia hybrydowe łączące różne źródła danych. Zadanie przedstawiono graficznie na Rys. 2.11.

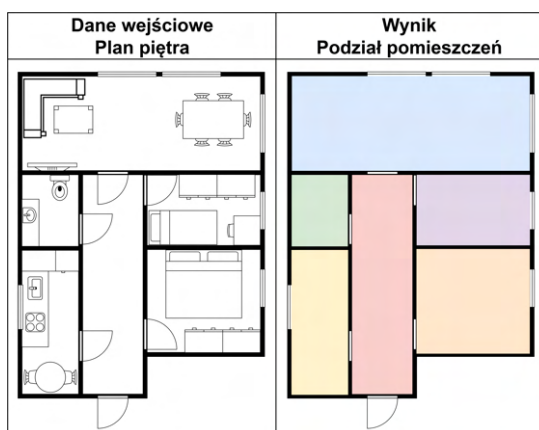
Wykrywanie zmian. Zadanie to miało na celu wykrycie zmian w strukturze budynku poprzez porównanie danych zebranych w różnych momentach. Analizowano różnice geometryczne i topologiczne, wspierające np. dokumentację remontów. W literaturze znaleziono jedno eksperymentalne podejście: system Koeva i in. [43]. Zadanie przedstawiono graficznie na Rys. 2.12.



Rysunek 2.11: Wizualizacja zadania klasyfikacji pomieszczeń. Przykładowy plan piętra wejściowego (**po lewej**) jest analizowany w celu wykrycia obszarów pomieszczeń. Każdy obszar zostaje sklasyfikowany i oznaczony etykietą typu pomieszczenia (**po prawej**).

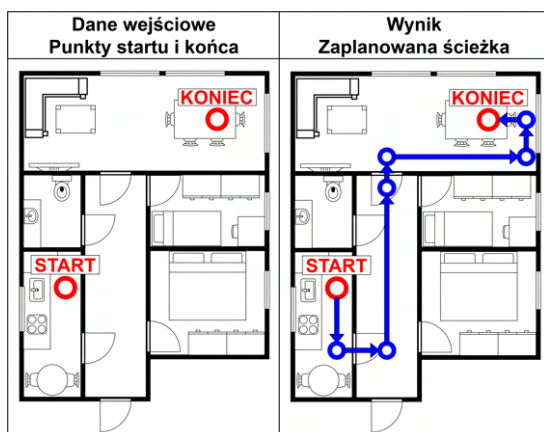


Rysunek 2.12: Wizualizacja zadania wykrywania zmian. Obrazy początkowy i zaktualizowany (**u góry**) są analizowane w celu wykrycia różnic między nimi. Wykryte niezgodne elementy zostały zaznaczone na czerwono (**na dole**).



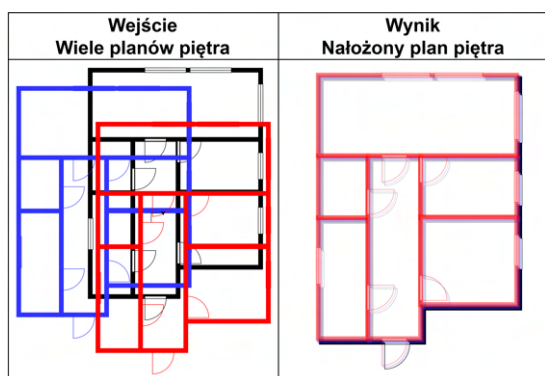
Rysunek 2.13: Wizualizacja zadania segmentacji. Obraz wejściowy (**po lewej**) jest analizowany w celu wykrycia oddzielnych obszarów reprezentujących poszczególne pomieszczenia. Każde pomieszczenie zostało oznaczone wielokątem w innym kolorze (**po prawej**).

Segmentacja. Zadanie to polegało na podziale przestrzeni na odrębne pomieszczenia na podstawie planów, map siatkowych lub chmur punktów. W przypadku danych 2D wykorzystywano klasyczne metody geometryczne i sieci neuronowe [100, 109], często wspomagane analizą grafową [126]. Segmentacja chmur punktów opierała się na klasteryzacji przestrzennej i analizie trajektorii skanowania [16, 59]. Prace różniły się podejściem, ale ich celem było wydzielenie czytelnych, semantycznie spójnych obszarów przestrzeni. Zadanie przedstawiono graficznie na Rys. 2.13.



Rysunek 2.14: Wizualizacja zadania nawigacji. Dane wejściowe obejmują wybrane punkty początkowy i końcowy (**po lewej**). W wyniku przetwarzania generowana jest proponowana ścieżka, przedstawiona jako zestaw instrukcji (**po prawej**).

Nawigacja wewnętrzna. Zadanie to obejmowało lokalizację użytkownika, planowanie ścieżek, eksplorację robotów i rozszerzoną wizualizację przestrzeni. Wykorzystywano dane radiowe [199], obrazy RGBD [196], chmury punktów [56], sensory smartfonów [24], czy trajektorie z czujników inercyjnych [181]. Łączono różne źródła informacji, by zwiększyć dokładność. Typowe podejścia obejmowały klasyfikację sceny, tworzenie map semantycznych, analizę ścieżek i filtrowanie pomiarów. Coraz częściej korzystano z metod uczenia maszynowego i fuzji danych. Zadanie przedstawiono graficznie na Rys. 2.14.



Rysunek 2.15: Wizualizacja zadania dopasowania mapy. Zbiór nieskalibrowanych planów wejściowych (**po lewej**) zostaje połączony w jeden wspólnie wyrównany rezultat (**po prawej**).

Dopasowywanie/wyrównywanie.

Zadanie polegało na precyzyjnym wyrównaniu dwóch reprezentacji tej samej przestrzeni: mapy siatkowej lub planu piętra. W przypadku map stosowano segmentację, analizę grafów oraz wyznaczanie macierzy transformacji [130, 148]. Rozważano również dopasowanie planu do zdeformowanej mapy sensorycznej [187]. Dopasowanie planów dotyczyło ich porównywania i wyszukiwania podobnych układów, z użyciem segmentacji i grafów topologicznych [92]. W obu przypadkach celem było umożliwienie automatycznej synchronizacji różnych źródeł danych. Zadanie przedstawiono graficznie na Rys. 2.15.

2.1.5 Dostrzegane wyzwania

Analiza literatury pozwoliła zidentyfikować szereg wspólnych problemów, napotykanym przez badaczy zajmujących się segmentacją przestrzeni i przypisywaniem jej znaczenia semantycznego w środowiskach wewnętrznych. Choć szczegóły różniły się w zależności od rozwiązania, pewne wyzwania pojawiały się regularnie, niezależnie od konkretnego zadania czy struktury danych.

Różnorodność danych wejściowych. Największym wyzwaniem jest niejednorodność formatu i struktury danych. Dotyczy to m.in. planów budynków, które mogą mieć zupełnie różne postacie: uproszczone lub techniczne, czarno-białe lub kolorowe, puste lub pełne ikon mebli. Podobna różnorodność dotyczy modeli przestrzennych (siatki, chmury punktów). Brakuje jednolitych standardów reprezentacji, co utrudnia tworzenie uniwersalnych rozwiązań.

Skalowalność i ograniczenia sprzętowe. Współczesne systemy muszą obsługiwać ogromne ilości danych: miliony próbek i terabajty pomiarów. Choć w przy-

padku złożonych systemów (np. rekonstrukcji cyfrowych bliźniaków) możliwe jest zastosowanie potężnego sprzętu, nie zawsze jest to wykonalne. Rosnąca liczba małych, energooszczędnych i niedrogich czujników (np. w systemach IoT) powoduje, że dane są zbierane w ogromnej ilości, ale sprzęt analizujący ma ograniczoną moc obliczeniową. Potrzeba rozwiązań opartych na przetwarzaniu brzegowym (z ang. *Edge Computing*) jest coraz bardziej paląca.

Udział człowieka. Wielu autorów nadal zakłada ręczne zbieranie danych: skanowanie pomieszczeń, robienie zdjęć, nanoszenie znaczników. Tymczasem najbardziej pożądane są systemy autonomiczne, które minimalizują udział operatora. Rozwiązania bazujące na robotach, dronach lub inteligentnych sensorach mogą w przyszłości znacząco zautomatyzować ten proces.

Utrata informacji. Niektóre dane mogą zostać bezpowrotnie utracone, jeśli nie zostaną uwzględnione we wczesnym etapie przetwarzania. Przykładem może być wysokość kondygnacji, pomijana w planach 2D. Jej późniejsze odtworzenie jest trudne lub niemożliwe bez dodatkowych danych. Oznacza to, że już na etapie akwizycji należy świadomie zaplanować, jakie informacje muszą zostać zachowane.

Wnioski. Zidentyfikowane wyzwania pokazują wyraźny kierunek dalszych badań: rozwój rozwiązań hybrydowych, integrujących różne typy danych oraz wykorzystujących potencjał urządzeń IoT. Konieczne jest projektowanie systemów, które:

- potrafią łączyć różnorodne źródła danych (np. obrazy 2D, pomiary przestrzenne, dane radiowe),
- są odporne na zmienność formatu i struktury danych wejściowych,
- potrafią działać efektywnie na ograniczonych platformach sprzętowych,
- umożliwiają automatyczne i niskokosztowe pozyskiwanie danych.

2.1.6 Podsumowanie

Analiza literatury wskazuje na duży potencjał badawczy poruszanych zagadnień i ich praktyczne znaczenie dla rozwoju systemów funkcjonujących w środowiskach inteligentnych (z ang. *Smart Environment*).

Przeprowadzona analiza pozwoliła wyróżnić trzy główne perspektywy badawcze: rodzaj wykorzystywanych danych, charakter prowadzonych analiz (ich kate-

gorie wysokopoziomowe) oraz typ realizowanego zadania (implementacje niskopoziomowe).

To właśnie z tych powodów tematyka niniejszej pracy doktorskiej skupia się na systemach lokalizacji i rozpoznawania przestrzeni, które bazują na danych pochodzących z wielu źródeł, są w stanie działać w środowisku IoT oraz potrafią dynamicznie dostosowywać się do dostępnych informacji. Przeprowadzone badania literatury jednoznacznie wskazują na dynamiczny rozwój obszaru. Współczesne rozwiązania coraz częściej opierają się na złożonych strukturach danych, reprezentacjach wielowymiarowych oraz metodach wykorzystujących sztuczną inteligencję. Co ważne, coraz więcej systemów opiera swoje działanie na więcej niż jednym źródle i formacie danych.

Zaproponowano klasyfikację istniejących rozwiązań, obejmującą zarówno techniki rozpoznawania planów pomieszczeń, rekonstrukcję modeli 3D, rozpoznawanie scen, jak i nawigację wewnątrzbudynkową i inne, niskopoziomowe realizowane zadania. Jednym z kluczowych wniosków jest to, że warto świadomie dobierać reprezentację danych wejściowych i nie bać się jej zmieniać. Zarówno wybór urządzenia pomiarowego, jak i konwersja formatu danych mogą znacząco uprościć przetwarzanie. Reprezentacja danych nie musi być stała – można ją dopasować do potrzeb zadania. Ciekawym zjawiskiem jest także możliwość przenoszenia rozwiązań między wymiarami danych – wiele algorytmów dobrze radzi sobie z przetwarzaniem danych o jeden wymiar wyższych niż te, dla których były pierwotnie zaprojektowane. Segmentacja i klasyfikacja stają się nie tyle osobnymi zadaniami, co elementami większych systemów.

Przedstawione w niniejszym rozdziale omówienie analizy literatury stanowi syntezę i skrót większego opracowania bibliograficznego i zrealizowanych prac. Dogłębne omówienia poszczególnych referowanych artykułów, a także rozbudowana dyskusja, obserwowane dla każdej z kolejnych taksonomii wyzwania i wnioski oraz statystyki bibliometryczne omówione zostały w publikacji naukowej *“Methods and applications of space understanding in indoor environment - a decade survey”* [210] do czasopisma Applied Sciences (100 pkt. MNiSW).

2.2 Badania rynku i oczekiwań użytkowników

Opracowywanie rozwiązań i produktów, które skutecznie spełniają oczekiwania biznesowe, wymaga dobrej znajomości potrzeb użytkowników końcowych. Ponieważ prowadzone badania mają na celu uzyskanie konkretnego wdrożenia przemysłowego opracowanych rozwiązań, konieczne staje się zdefiniowanie oczekiwań i dopasowanie do nich kierunku realizowanych analiz. Żeby taki zestaw oczekiwań uzyskać, przeprowadzone zostały badania rynkowe, w ramach których przeanalizowano otoczenie gospodarcze firmy, jej produktów oraz potencjalnych klientów. Wykonano badania sondażowe, a także określono potencjalne dalsze kierunki analiz. Przedstawieniu tego zagadnienia poświęcona jest dalsza część tego rozdziału.

Szczególne znaczenie przypisano w nim problemowi wysokiego nakładu czasu pracy, który wymagany jest na uzyskanie bogatego semantycznie modelu cyfrowego bliźniaka. To właśnie ten aspekt, wskazywany przez dotychczasowych użytkowników jako uciążliwy, stanowi główny temat przewodni prowadzonego rozeznania rynku.

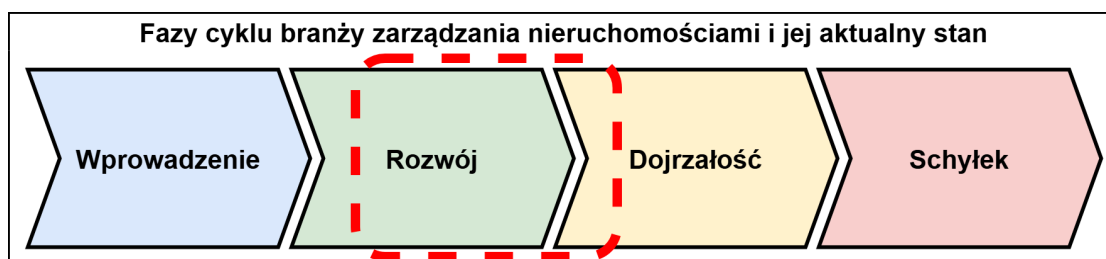
2.2.1 Analiza produktu

Przedsiębiorstwo partnerskie realizowanego doktoratu wdrożeniowego to spółka tworząca, udoskonalająca, a także wdrażająca na rynek oprogramowanie z dziedziny Internetu Rzeczy (z ang. *Internet of Things*, IoT) oraz Architektury, Inżynierii i Budownictwa (z ang. *Architecture Engineering Construction*, AEC). Wraz ze swoimi partnerami technologicznymi dąży do popularyzacji idei Cyfrowego Bliźniaka, tak do zastosowań domowych, jak i przemysłowych. Z użyciem zaawansowanego modelowania trójwymiarowego wprowadza do użytku cyfrowe modele obiektów infrastruktury domów, firm, biur, a także przestrzeni wielkopowierzchniowej, jak hotele, szpitale oraz fabryki.

Główny przedmiot działalności spółki to *Informacja i komunikacja* (Sekcja J w ramach Polskiej Klasyfikacji Działalności, PKD), ze szczególnym naciskiem na oprogramowanie, doradztwo w zakresie informatyki i inne powiązane (kody PKD 62.01.Z, 62.02.Z). Oferowane produkty swoją tematyką wpisują się w realizację zadań związanych z zarządzaniem nieruchomościami (z ang. *Facility Ma-*

nagement, FM). Przedsiębiorstwo funkcjonuje głównie na rynku zagranicznym, co utrudnia analizę rynku (dane Głównego Urzędu Statystycznego, GUS, dotyczą jedynie obrotu krajowego). Do rozeznania konieczne było więc wykorzystanie zagranicznych materiałów i opracowań o zasięgu międzynarodowym [211–213].

Ich wykorzystanie pozwala zaobserwować trend wzrostowy całej branży. Należy spodziewać się obecności dużych graczy konkurencji rynkowej, tj. przedsiębiorstw o ugruntowanej renomie, które nadal bardzo intensywnie rozwijają swoje nowe produkty. Wysokie zainteresowanie zagadnieniami sztucznej inteligencji sprawia, że mimo pewnego stopnia dojrzałości systemów FM nadal na rynek wchodzi nowe firmy, oferujące przełomowe rozwiązania, a popyt na inteligentnych asystentów rośnie. W klasyfikacji branżowej na podstawie faz cyklu, systemy zarządzania nieruchomościami cechuje głównie rozwój, jednakże dostrzec można także elementy dojrzałości. Przedstawiono to na Rys. 2.16. Raporty prognozują wzrost branży i nie przewidują nadejścia fazy schyłkowej w najbliższych dekadach.



Rysunek 2.16: Diagram faz cyklu rynkowego z zaznaczeniem (czerwona ramka) aktualnego umiejscowienia branży zarządzania nieruchomościami.

Na portfel produktowy przedsiębiorstwa partnera składają się dwa systemy oprogramowania: SIM-ON oraz Stages. Platforma SIM-ON pozwala na zarządzanie już istniejącymi nieruchomościami, zawiera interfejs do sterowania urządzeniami IoT, planowania konserwacji, komunikacji między użytkownikami, itp. System Stages pozwala natomiast monitorować proces konstrukcji budynku na przestrzeni czasu. Docelowym produktem, do którego wdrażane są opracowane systemy lokalizacji i rozpoznawania obiektów, jest platforma SIM-ON. Tym samym, w dalszych analizach rynkowych badana jest Strategiczna Jednostka Biznesu (SJB) w postaci działu sprzedaży produktu SIM-ON. Jest to produkt cyfrowy, pakiet oprogramowania sprzedawany w modelu subskrypcyjnym, służący do zarządzania nierucho-

mościami w oparciu o ideę cyfrowego bliźniaka. Według klasyfikacji z użyciem macierzy BCG (metoda analizy portfelowej opracowanej przez firmę Boston Consulting Group) SIM-ON to produkt kategorii “znaku zapytania”, tj. o niskim zysku i bardzo wysokim popycie inwestycyjnym.

Podsumowując, analizie podlega dział sprzedaży produktu SIM-ON. Jest to nowinka technologiczna, której potencjał i tempo zwrotu z inwestycji są ciężkie do oszacowania. Na rynku spodziewać się należy innych produktów o ugruntowanej pozycji oraz potencjalnie wielokrotnie większym budżecie na innowacje. Proponuje się zastosowanie strategii sprzedaży poprzez walkę cenową (strategia kosztowa). Konkurencyjne produkty są kosztowne, bardzo profesjonalne i prezentują wysoki stopień komplikacji w użytkowaniu. Rynek pozostaje dynamicznie rozwijającym się, co pozwala znaleźć niszę sprzedażową dla produktu łatwego w użyciu, o intuicyjnym interfejsie, a także niskiej cenie.

2.2.2 Analiza SWOT

Analiza SWOT (od ang. *Strengths*, *Siły*; *Weaknesses*, *Słabości*; *Opportunities*, *Szanse*; *Threats*, *Zagrożenia*) jest jednym ze sposobów na usystematyzowaną ocenę strategiczną tworzonego produktu. Pozwala ona w skróty sposób zidentyfikować czynniki wewnętrzne: sukcesu (*Siły*) i ograniczenia (*Słabości*) oraz zewnętrzne uwarunkowania rozwoju: pozytywne (*Szanse*) i negatywne (*Zagrożenia*). W pierwszym kroku analizy wytypowano pięć najistotniejszych elementów w każdej z tych czterech kategorii. Przedstawiono je w Tab. 2.5).

W następnym kroku określono wzajemne relacje i zależności między czynnikami wewnętrznymi i zewnętrznymi. W tym celu zastosowano macierz ocen, w której każda para (np. Siła I i Szansa III) została opisana wartością od 0 do 3, gdzie 0 to brak interakcji i zależności, a 3 to zależność silna o wysokiej istotności. Suma punktów w poszczególnych wierszach pokazuje względną istotność par czynników pomiędzy sobą. Przypisaną punktację przedstawiono w Tab. 2.6. Pięć najistotniejszych aspektów (wierszy, które uzyskały największą liczbę punktów) wyróżniono pogrubieniem. Uzyskane wyniki umożliwiają jakościowy opis pozycji rynkowej produktu SIM-ON, a także określenie i priorytetyzację przyszłych działań rozwojowych, które powinny zminimalizować podejmowane ryzyko ryn-

Tabela 2.5: Macierz analizy SWOT dla produktu SIM-ON

Siły	Słabości
Innowacyjność produktu Połączenie technologii IoT, cyfrowych bliźniaków i automatyki budynkowej stanowi unikalne rozwiązanie w skali świata. Intuicyjny interfejs 3D Przestrzenność ułatwia zarządzanie zasobami i podnosi atrakcyjność produktu w oczach klientów. Szeroka kompatybilność Obsługa urządzeń IoT różnych ekosystemów: KNX, SmartThings, Schneider Electric i innych. Potencjał globalny Dynamiczny wzrost rynku zarządzania nieruchomościami sprzyja ekspansji. Wsparcie techniczne i doradztwo Szkolenia, dodatkowe integracje i obsługa na życzenie podnoszą rzetelność w oczach klientów.	Wysokie koszty wdrożenia Pozyskanie skanów 3D i cyfrowego bliźniaka to wysoka bariera wejścia. Brak rozpoznawalności globalnej marki SIM-ON, jako nowinka technologiczna, dopiero buduje pozycję międzynarodową. Odmienność rozwiązania Konieczność nauki nowego narzędzia przez użytkowników, co może napotkać opór i spowodować zniechęcenie. Ograniczona oferta subskrypcji Brak dedykowanych, szytych na miarę, produktów dla dużych przedsiębiorstw (tzw. <i>white label</i>). Zależność od partnerów technologicznych Integracje wymagają współpracy i zgody dostawców zewnętrznych.
Szanse	Zagrożenia
Rozwój rynku zarządzania nieruchomościami Raporty prognozują wzrost rynku globalnego na poziomie kilku-kilkunastu procent rocznie. Rosnąca popularność IoT Coraz szersze wykorzystanie w zarządzaniu nieruchomościami, upowszechnienie technologii i dostępności do Internetu. Regulacje efektywności energetycznej SIM-ON może być narzędziem zapewniającym zgodność z wymogami formalnymi i oczekiwaniami prawnymi. Ekspansja międzynarodowa Szansa wejścia na nowe, rozwijające się rynki, np. Ameryki Południowej i Azji. Partnerstwa strategiczne Kooperacja z dużymi graczami branży IoT oraz automatyki budynkowej zapewnia dostęp do puli klientów.	Silna konkurencja Globalne platformy z podobną funkcjonalnością, a większym dostępnym budżetem na badania i rozwój. Wysoka dynamika zmian technologicznych Konieczna ciągła aktualizacja produktu i stałe nakłady na możliwie szybkie wdrażanie nowych rozwiązań. Zagrożenia cyberbezpieczeństwa Rozwiązania IoT są silnie narażone na ataki hakerskie, wzrost popularności produktu może je spotęgować. Silna zależność od globalnej sytuacji gospodarczej W kryzysie finansowym optymalizacja nie jest priorytetem. Spowolnienie ograniczy inwestycje klientów. Zmiany prawne Nowe przepisy mogą wymusić zmianę sposobu działania systemu, np. w związku z ochroną danych osobowych.

Tabela 2.6: Macierz oceny zależności pomiędzy czynnikami analizy SWOT.

		Szanse						Zagrożenia					
		I	II	III	IV	V	Σ	I	II	III	IV	V	Σ
Siły	I	2	3	1	2	3	11	3	2	2	3	0	10
	II	2	3	0	0	1	6	3	3	0	0	0	6
	III	3	3	1	2	3	12	1	0	1	0	1	3
	IV	2	2	2	3	3	12	1	1	0	1	0	3
	V	1	1	1	0	0	3	0	0	1	1	3	5
Słabości	I	1	1	1	1	0	4	1	2	0	2	3	8
	II	1	2	0	3	3	9	3	1	1	3	0	8
	III	0	0	1	2	0	3	0	0	1	1	1	3
	IV	1	0	0	3	3	7	0	0	1	2	1	4
	V	0	0	0	3	3	6	3	0	0	2	0	5

kowe.

Przeprowadzona analiza wskazuje, że kluczowymi aspektami produktu SIM-ON są jego silne strony oraz szanse rozwoju. Trzy z pięciu par o największej liczbie przypisanych punktów pochodziły właśnie z tego kwadranta macierzy. Rozwiązanie wyróżnia się innowacyjnością oraz wysokim potencjałem wdrożeniowym na globalnym rynku, szczególnie w kontekście dynamicznego rozwoju branży. Jednocześnie te same obszary stanowią źródło największych wyzwań, związanych z koniecznością budowania rozpoznawalności i skutecznego dotarcia do odbiorców końcowych. Wskazane jest zatem nawiązanie współpracy z wybranymi partnerami strategicznymi o globalnym zasięgu, tak technologicznymi, jak i inwestycyjnymi. Takie podejście pozwoli zarówno na efektywne wykorzystanie mocnych stron systemu, jak i na ograniczenie barier związanych z wejściem na rynek oraz rozpoczęciem procesu komercjalizacji.

2.2.3 Analiza konkurencji rynkowej

Platforma SIM-ON oferowana jest tak użytkownikom indywidualnym, jak i komercyjnym. Poza marketingiem celowanym, który w sposób aktywny poszukuje sprzedaży, istotne jest dogłębne rozeznanie ofert konkurencji, na które natrafić może potencjalny klient. Podobnie jak w analizie literatury, przeprowadzone

zostało również rozeznanie rynkowe. Pod uwagę wzięto systemy zarządzania budynkami i ich wyposażeniem, ze szczególnym uwzględnieniem funkcji związanych z analizą obrazów, przetwarzaniem modeli 2D i 3D oraz śledzeniem położenia obiektów w przestrzeni. W zestawieniu pominięto rozwiązania oferujące **tylko** lokalizację zasobów przy użyciu danych radiowych, które uznano za odrębną kategorię szerzej omawianą w Roz. 3.

Analizowane systemy oceniono pod kątem dwudziestu kryteriów funkcjonalnych, m.in. możliwości importu planów i modeli budynków, automatycznej ekstrakcji wyposażenia, ręcznego i automatycznego pozycjonowania obiektów w przestrzeni oraz obecności komponentów przetwarzania obrazu. Lista kryteriów z pełnym opisem przedstawiona została w Tab. 2.7. Celem rozeznania było nie tylko wskazanie najbliższych odpowiedników oferowanego oprogramowania, lecz także identyfikacja potencjalnej niszy rynkowej wynikającej z niedoborów funkcjonalnych w istniejących systemach.

Pierwszą kategorię analiz stanowiły systemy dedykowane zarządzaniu obiektami z dwóch głównych sektorów: hotelarskiego oraz operacyjno-eksploatacyjnego. Systemy hotelarskie dostarczają narzędzi do obsługi procesów rezerwacji, zarządzania pokojami, płatnościami oraz komunikacją z gośćmi. Przykłady takich systemów to Cloudbeds, Mews, HotelTime PMS czy Oracle OPERA. Analiza wykazała, że choć niektóre z nich oferują podstawową funkcjonalność związaną z planami pomieszczeń i całych pięter (np. adnotacje na planach), to ich rola ogranicza się głównie do prezentowania dostępności pokoi. Brakuje zaawansowanego modelowania przestrzeni, integracji z danymi przestrzennymi, a także jakichkolwiek funkcji związanych z analizą obrazu. Pomieszczenia i ich wyposażenie traktowane są jako byty organizacyjne, a nie elementy posiadające konkretne położenie w przestrzeni. Systemy hotelarskie, mimo znacznego stopnia rozbudowania, nie oferują funkcji istotnych z punktu widzenia systemów wizualnego nadzoru przestrzeni mieszkalnych. Systemy sektora operacyjno-eksploatacyjnego oferują natomiast znacznie szersze wsparcie dla modeli typu BIM. Wśród analizowanych rozwiązań znalazły się m.in. produkty firm Autodesk (Plangrid, BIM 360 Ops), Procore, a także platformy, takie jak Disperse czy BuildingEngines. Ich głównym założeniem jest jednak praca na już istniejących modelach, w których projektant wcześniej umieścił istotne obiekty i przypisał im właściwe pozycje w przestrzeni.

Tabela 2.7: Zestawienie analizowanych w ramach rozeznania konkurencji rynkowej funkcji systemów zarządzania nieruchomościami.

Funkcja	Opis
F1	Wprowadzanie dwuwymiarowych planów architektonicznych nadzorowanych budynków do systemu
F2	Automatyczne generowanie listy wyposażenia na podstawie planu pomieszczeń / pięter
F3	Ręczne wprowadzanie listy wyposażenia do planu pomieszczeń / pięter
F4	Ręczne wydzielanie pokoiów na planach pomieszczeń / pięter zarządzanych budynków
F5	Automatyczne wydzielanie pokoiów na planach pomieszczeń / pięter zarządzanych budynków
F6	Ręczne wprowadzanie precyzyjnego położenia konkretnych obiektów na planach pomieszczeń / pięter zarządzanych budynków
F7	Automatyczne wprowadzanie precyzyjnego położenia konkretnych obiektów na planach pomieszczeń / pięter zarządzanych budynków
F8	Wprowadzanie modeli trójwymiarowych zarządzanych budynków do systemu
F9	Automatyczne generowanie listy wyposażenia na podstawie modelu trójwymiarowego
F10	Ręczne wprowadzanie listy wyposażenia do modelu trójwymiarowego
F11	Ręczne wydzielanie pomieszczeń w modelu trójwymiarowym
F12	Automatyczne wydzielanie pomieszczeń w modelu trójwymiarowym
F13	Ręczne wprowadzanie precyzyjnego położenia nadzorowanego obiektu w przestrzeni trójwymiarowej modelu budynku
F14	Automatyczne wprowadzanie precyzyjnego położenia nadzorowanego obiektu w przestrzeni trójwymiarowej modelu budynku
F15	Rozpoznawanie wyposażenia pomieszczeń na podstawie zdjęć
F16	Rozpoznawanie wyposażenia pomieszczeń na podstawie obrazu “na żywo” z kamery
F17	Rozpoznawanie wyposażenia pomieszczeń na podstawie skanu trójwymiarowego, tzw. spaceru interaktywnego
F18	Nadzorowanie położenia śledzonego obiektu w czasie rzeczywistym na podstawie obrazu z kamer
F19	Nadzorowanie położenia śledzonego obiektu w czasie rzeczywistym na podstawie zamontowanych czujników radiowych
F20	Ostrzeganie o opuszczeniu przez śledzony obiekt określonej przestrzeni

Zwykle nie występuje w nich jednak ani analiza obrazu, ani aktualizacja modelu na podstawie danych wizualnych. Systemy te wykazują wysoką funkcjonalność w zakresie przeglądania, dokumentowania i wprowadzania do systemu nowych informacji użytkowych, jednak pozostają biernymi odbiorcami danych w kontekście zarządzania modelem nieruchomości. Brakuje w nich także możliwości śledzenia zmian wyposażenia wnętrz na podstawie kamer lub zdjęć. Zestawienie i podsumowanie oferowanych przez analizowane systemy funkcji przedstawiono w Tab. 2.8.

Druga kategoria analiz ukierunkowana została na systemy przetwarzające dane napływające strumieniowo. Przykładem źródeł takich danych są systemy monitoringu z użyciem kamer lub dedykowane systemy wizyjne do nadzoru przestrzeni. Analizie poddano rozwiązania firm Situm, Navigine, Intermodalics, DragonflyCV, Indoors i Eliko. Systemy te jako jedne z pierwszych korzystają z danych optycznych w procesie lokalizacji zasobów, jednak ich przeznaczenie jest często ograniczone do obiektów przemysłowych lub przestrzeni publicznych. Zamiast pełnej rekonstrukcji wnętrz, skupiają się one na określaniu pozycji obiektów ruchomych w czasie rzeczywistym, najczęściej za pomocą kamer zainstalowanych na tych obiektach. Przykładowo, Intermodalics oraz DragonflyCV oferują lokalizację opartą na analizie obrazu w czasie rzeczywistym, jednak ich zastosowanie obejmuje głównie śledzenie dużych obiektów, jak wózki widłowe czy pojazdy typu AGV (z ang. *Automated Guided Vehicle*, zautomatyzowane pojazdy bezzałogowe). Choć kierunek rozwoju jest zbliżony z zakładanym celem opracowywanego systemu, brak im finezji wymaganej do nadzoru typowo mieszkaniowego oraz precyzyjnego rozpoznawania wyposażenia. Podobnym zagadnieniem są systemy jednoczesnego lokalizowania i mapowania (z ang. *Simultaneous Localization and Mapping*, SLAM), stosowane m.in. w robotyce i nawigacji wewnętrznej. Wśród analizowanych produktów znalazły się m.in. Kudan, 3dVisionLabs, SevenSense oraz SteerPath. Rozwiązania te charakteryzują się zaawansowaną analizą przestrzeni na podstawie danych z kamer stereoskopowych, jednak ich zastosowanie koncentruje się głównie na potrzebach autonomicznych robotów lub systemów mobilnych. Wyjątek stanowi platforma Hemi Stereo DK1 firmy 3dVisionLabs, będąca systemem kamer sferycznych ogólnego zastosowania, umożliwiającą przetwarzanie danych od razu na urządzeniu, bez konieczności transferu danych na zewnętrzne serwery. Choć sprzęt ten jest interesujący z punktu widzenia badań, analizowane produkty nie oferują funk-

Tabela 2.8: Zestawienie funkcji obsługiwanych przez analizowane systemy do zarządzania nieruchomościami. Fakt wsparcia konkretnej funkcji oznaczony został symbolem +.

Platforma	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20
Cloudbeds																				
PMS																				
Mews																				
Clock PMS+	+			+																
HotelTime																				
PMS																				
innRoad																				
RMS Cloud	+			+																
PMS																				
RoomRaccoon																				
PMS																				
WebRezPro	+			+																
OPERA PMS	+		+	+		+														
StayNTouc																				
Aero Guest	+			+																
Protel PMS																				
Little	+																			
Hotelier																				
Slope																				
Agilysys																				
Aquicore																				
Verdigris																				
Eden	+		+	+																
Disperse								+		+	+		+							
SMS Assist																				
Budderfly																				
Building	+		+	+		+														
Engines																				
Breezeway																				
ResMan																				
Buildium																				
AD Takeoff	+		+	+		+		+		+	+		+							
AD Plangrid	+		+	+		+														
Build																				
StructionSite	+		+	+		+		+		+	+		+							
Procore BIM	+		+	+		+		+		+	+		+							
AD BIM 360																				
Ops	+		+	+		+		+		+	+		+							

cji tworzenia i aktualizacji planów pomieszczeń ani identyfikacji typowych obiektów domowego użytku. Są to technologie niskopoziomowe, które wymagają dalszej integracji, aby mogły stanowić bazę dla kompleksowego systemu nadzoru. Uzupełnieniem analizy były systemy dedykowane, takie jak: Sortly, IBM Sterling, VisualOne, Ambient.ai czy Meru Software. Rozwiązania te często korzystają z komponentów sztucznej inteligencji i oferują zaawansowaną analizę scen, jednak w wąskich, wyspecjalizowanych zastosowaniach, np. monitoringu regałów sklepowych, nadzoru magazynów czy wykrywania nietypowych zachowań. Przykładowo, VisualOne umożliwia użytkownikowi trening własnej sieci neuronowej do detekcji sytuacji, takich jak np. wejście psa na kanapę. Z kolei Ambient.ai analizuje obraz z kamer pod kątem obecności osób i potencjalnie niepożądanych zdarzeń. Pomimo wysokiego poziomu technologicznego, systemy te nie oferują wsparcia dla modeli przestrzennych ani integracji z planami budynków. Wciąż brakuje im funkcji, które umożliwiałyby pełną reprezentację stanu przestrzeni mieszkalnej. Zestawienie i podsumowanie oferowanych przez analizowane systemy funkcji przedstawiono w Tab. 2.9.

Podsumowując, żaden z przeanalizowanych systemów nie spełniał w pełni zestawu założonych kryteriów funkcjonalnych. Systemy hotelarskie i przemysłowe operują na danych opisowych lub wcześniej przygotowanych modelach. Nie integrują się one z dynamicznie pozyskiwanymi danymi wizualnymi. Rozwiązania lokalizacji i optyczne systemy śledzenia oferują zaawansowane technologie przetwarzania obrazu, lecz ich zastosowanie ogranicza się do specjalistycznych urządzeń mobilnych, a nie stacjonarnych scenariuszy domowych. Część z badanych systemów dostarcza komponentów opartych na uczeniu maszynowym do analizy scen i detekcji obiektów, ale brakuje w nich spójnej architektury umożliwiającej reprezentację stanu przestrzeni w czasie rzeczywistym. Taka luka rynkowa stanowi uzasadnienie dla opracowania autorskiego rozwiązania łączącego analizę obrazu, modelowanie przestrzeni i bieżący nadzór nad położeniem obiektów. System tego typu wymaga nie tylko rozwiązań sprzętowych (np. kamer szerokokątnych), ale przede wszystkim platformy integracyjnej, która umożliwi tworzenie cyfrowego odwzorowania wnętrza z wykorzystaniem danych wizyjnych.

2.2.4 Badania sondażowe

Jednym z kluczowych aspektów i celem prowadzonych prac jest podniesienie zgodności oferowanych produktów z oczekiwaniami klientów końcowych. Aby określić, jakie te oczekiwania są, postanowiono o wykonaniu krótkich badań sondażowych wśród ekspertów branżowych. Rozpoczęto od wywiadu z kilkoma reprezentantami partnerów biznesowych przedsiębiorstwa i wytypowano główny problem, z którym borykają się użytkownicy. Okazała się nim być monotonna, zajmująca wiele godzin pracy konfiguracja wstępna modelu cyfrowego bliźniaka nadzorowanej nieruchomości. O ile akwizycja modelu trójwymiarowego i przetworzenie danych do postaci interaktywnej prezentacji jest subiektywnie proste i krótkie (w kontekście nakładów czasu pracy), o tyle nadanie surowym danym przestrzennym ich semantycznego znaczenia przekracza wymaganym nakładem czasowym próg akceptowalny dla typowych klientów. Porzucają oni wtedy proces konfiguracji i nie odwzorowują wszystkich posiadanych zasobów w modelu cyfrowym, co z założenia ogranicza użyteczność oferowanego oprogramowania.

Aby temu zaradzić, w niniejszej rozprawie zaproponowano opracowanie systemu automatyzującego taką konfigurację, a dalsze badania sondażowe ukierunkowano na uzyskanie odpowiedzi na pytania dotyczące tego, jakie elementy takiej konfiguracji są dla użytkowników najistotniejsze. Zaprojektowano badania z użyciem metody sondażu diagnostycznego i techniki ankiety, opracowano zestaw pytań, a następnie zgromadzono pulę użytkowników systemu chętnych na wzięcie udziału w badaniach. Uzyskano odpowiedzi od piętnaściorga respondentów, którzy odpowiadali najpierw na dwa pytania wstępne, a po pewnym czasie na trzecie pytanie podsumowujące. Konstrukcji kwestionariusza ankiety, omówieniu jego pytań oraz uzyskanych odpowiedzi i wyników poświęcona jest pozostała część tego rozdziału.

Oczekiwane rodzaje obiektów

Pierwszym zagadnieniem, na które poszukiwano odpowiedzi, była lista klas obiektów, które rozpoznawane powinny być automatycznie. W kwestionariuszu ankiety zadano zatem pytanie: *“Jakie obiekty inteligentny asystent powinien rozpoznawać w domu?”*. Pytanie było typu otwartego, respondenci odpowiadali bez ogra-

niczeń dotyczących minimalnej lub maksymalnej liczby wymienionych obiektów, a także w ich języku naturalnym. Uzyskane odpowiedzi zostały następnie przetłumaczone na język angielski, a także ujednolicone pod względem synonimów i określeń bliskoznacznych. Scalono takie frazy, jak np. *Lampa*, *Lampka*, *Lampka nocna* i *Lampka biurkowa*. Na podstawie uzyskanych odpowiedzi, poprzez uszeregowanie po najczęściej powtarzających się rodzajach, opracowany został zestaw 55 oczekiwanych do rozpoznawania obiektów w środowisku domowym. Nadano mu nazwę **CODE55**, od angielskiego *Common Objects in Domestic Environment*, tj. 55 obiektów powszechnie występujących w środowisku domowym. Zestawienie proponowanych do rozpoznawania obiektów przedstawiono w Tab. 2.10. Najpopularniejszymi wskazywanymi do rozpoznania obiektami były: *Lampa*, *Krzesło*, *Stół*, *TV* oraz *Pralka*. Uzyskano łącznie 477 odpowiedzi, które przełożyły się na 173 niepowtarzalne propozycje typów obiektów do rozpoznawania. W odpowiedziach przeważały sugestie wyposażenia z kategorii mebli oraz sprzętu gospodarstwa domowego.

Tabela 2.10: Spis klas obiektów w proponowanym zbiorze CODE55.

L.p.	Klasa	L.p.	Klasa	L.p.	Klasa
1	Krzesło	19	Lodówka	37	Akwarium
2	Stół	20	Mikrofalówka	38	Obraz
3	TV	21	Zmywarka	39	Oczyszczacz powietrza
4	Pralka	22	Kuchenka	40	Włącznik
5	Szafka	23	Czajnik	41	Bojler
6	Konsola do gier	24	Ekspres do kawy	42	Dywan
7	Kanapa	25	Toster	43	Tablica
8	Głośnik	26	Piekarnik	44	Wazon
9	Kominek	27	Lampa	45	Kran
10	Łóżko	28	Klimatyzacja	46	Zasłona
11	Szafa	29	Komputer	47	Roleta
12	Poduszka	30	Roślina	48	Półka
13	Stolik nocny	31	Okno	49	Gaśnica
14	Wanna	32	Biurko	50	Wentylator
15	Toaleta	33	Drzwi	51	Grzejnik
16	Prysznic	34	Lustro	52	Samochód
17	Suszarka na pranie	35	Gniazdko	53	Telefon
18	Suszarka do włosów	36	Zlew	54	Zegar
↪	↪	↪	↪	55	Czujnik alarmowy

Oczekiwane rodzaje pomieszczeń

Drugie z pytań miało na celu określenie rodzajów pomieszczeń, które warto wydzielać w surowych danych modeli nieruchomości. Sformułowano je jako: “Jakie typy pokoi znajdują się w typowym domu?”. Podobnie jak poprzednio, pytanie miało postać otwartą, bez ograniczeń i wymagań dotyczących udzielanych odpowiedzi. Zgromadzone wyniki przetłumaczono na język angielski, ujednolicono synonimy i przedstawiono w postaci listy najczęściej występujących rodzajów pomieszczeń. Łącznie uzyskano 228 odpowiedzi, które przełożyły się na 60 niepowtarzalnych rodzajów pomieszczeń. W syntezie przedstawiono wyniki po dodatkowej filtracji, tj. takie wyniki, które pojawiły się wśród odpowiedzi **więcej niż raz**. Ograniczyło to liczbę oczekiwanych do rozpoznania rodzajów pomieszczeń do 30 rodzajów. Zestawienie opracowanego zbioru typów pomieszczeń prezentowane jest w Tab. 2.11.

Tabela 2.11: Spis klas pomieszczeń w proponowanym zbiorze CODE55.

L.p.	Typ	L.p.	Typ	L.p.	Typ
1	Salon	11	Piwnica	21	Sauna
2	Kuchnia	12	Jadalnia	22	Schody
3	Łazienka	13	Spizarnia	23	Schowek
4	Sypialnia	14	Poddasze	24	Taras
5	Hol	15	Balkon	25	Audytorium
6	Garaż	16	Korytarz	26	Pokój gościnny
7	Pokój dziecięcy	17	Studio	27	Kino domowe
8	Biuro	18	Pralnia	28	Salon z kuchnią
9	Garderoba	19	Pom. techniczne	29	Pokój telewizyjny
10	Toaleta	20	Kotłownia	30	Biblioteka
↔		↔			

Matryca obiektów w pomieszczeniach

Ostatnim, najbardziej złożonym pytaniem formularza ankiety była matryca łącząca odpowiedzi z obu poprzednich pytań. Odpowiedzi zebrano wśród tej samej grupy ankietowanych, jednakże po pewnym odstępie czasowym, który konieczny był na opracowanie wyników wstępnych. Ankietowanych poproszono tym razem o określenie: “Czy dany rodzaj przedmiotu typowo występuje w danym rodzaju po-

mieszkania?”. Utworzono w ten sposób matrycę, składającą się z 550 pól: 55 klas obiektów * 10 klas pomieszczeń.

Jej analiza wierszami, kolumnami, a także przez wszystkie uzyskane odpowiedzi pozwala wnioskować o popularności i tendencjach, jakie zaobserwować możemy w środowisku domowym, jego podziale i sposobie wyposażania (meblowania). W pierwszej kolejności zsumowano i posortowano liczby wystąpień poszczególnych rodzajów obiektów w odpowiedziach ankietowanych. Najczęściej wskazywanymi, jako typowe, były takie klasy jak: *Gniazdko*, *Drzwi*, *Włącznik*, *Półka* oraz *Lampa*. Dla każdego z tych przedmiotów średnia liczba pomieszczeń, w których respondenci spodziewają się jego umieszczenia, wynosi niewiele poniżej 10. Przy właśnie dziesięciu rodzajach pokoi, o które pytano, pozwala to określić ten zestaw obiektów jako “uniwersalny” lub “powszechny”. Najrzadziej występowały natomiast: *Samochód*, *Piekarnik*, *Zmywarka*, *Wanna* oraz *Kuchenska*. Tutaj średnia liczba pomieszczeń, dla których obiekty te zostały uznane za typowe, wynosi niewiele ponad 1. Pozwala to wyciągnąć wnioski, że są to obiekty silnie specyficzne. Obserwowany rozkład odpowiedzi sugeruje możliwość wykorzystania pewnych klas obiektów do wnioskowania o funkcji pomieszczenia, co może być istotnym mechanizmem wspomagającym automatyczną klasyfikację danych przestrzennych. Obserwacje te są zgodne z ludzką intuicją, a wyniki ilościowe z przeprowadzonych badań je potwierdzają. Dziesięć pierwszych (najbardziej popularnych) oraz dziesięć ostatnich (najmniej popularnych) obiektów wraz z sumą, średnią i odchyleniem standardowym ich wystąpień, prezentowany jest w Tab. 2.12.

W drugiej kolejności przeanalizowane zostało, ile obiektów znajduje się w których pomieszczeniach. Uzyskane wyniki pokazują różnice przekraczające dwukrotność liczby spodziewanych obiektów. Pomieszczeniem o największej oczekiwanej liczbie typowych przedmiotów jest kuchnia, a najmniej garderoba. Zestawienie przedstawiające sumę odpowiedzi ankietowanych, średnią oraz odchylenie standardowe, prezentowane jest w Tab. 2.13.

Jako trzecie sprawdzone zostało, jak zmienia się oczekiwana przez użytkowników pula obiektów w zależności od pomieszczenia, o które pytamy. W uzyskanych wynikach dostrzec można kontynuację wcześniejszej obserwacji dotyczącej podziału obiektów na grupę uniwersalną oraz wysoce specyficzną. Obiekty powszechne wytypowane poprzednio, również tym razem zajmują wysokie pozycje w zesta-

Tabela 2.12: Najbardziej i najmniej popularne obiekty według liczby wystąpień w odpowiedziach uczestników.

Najbardziej popularne obiekty					Najmniej popularne obiekty				
Poz.	Obiekt	Suma	Śr.	Std.	Poz.	Obiekt	Suma	Śr.	Std.
1	Gniazdko	148	9,87	0,52	46	Czajnik	28	1,87	1,46
2	Drzwi	142	9,47	1,55	47	Lodówka	24	1,60	0,51
3	Włącznik	135	9,00	2,80	48	Mikrofal.	19	1,27	0,46
4	Półka	135	9,00	1,73	49	Prysznic	18	1,20	0,41
5	Lampa	130	8,67	2,06	50	Toster	18	1,20	0,41
6	Okno	123	8,20	2,18	51	Kuchenska	17	1,13	0,35
7	Grzejnik	121	8,07	2,12	52	Wanna	16	1,07	0,26
8	Roleta	120	8,00	2,00	53	Zmywarka	16	1,07	0,26
9	Roślina	108	7,20	2,60	54	Piekarnik	16	1,07	0,26
10	Cz. alarmowy	104	6,93	2,81	55	Samochód	15	1,00	0,00

Tabela 2.13: Łączna liczba odpowiedzi o obiektach w podziale na pomieszczenia.

Poz.	Pomieszczenie	Suma	Śr.	Std.
1	Kuchnia	481	32,07	4,50
2	Salon	469	31,27	3,77
3	Biuro	446	29,73	4,68
4	Sypialnia	440	29,33	5,25
5	Pokój dziecięcy	437	29,13	2,83
6	Łazienka	351	23,40	4,21
7	Hol	283	18,87	6,46
8	Garaż	278	18,53	8,26
9	Toaleta	249	16,60	4,75
10	Garderoba	191	12,73	4,48

wieniu. Obiekty wysoce specyficzne natomiast pojawiają się jedynie w określonych rodzajach pomieszczeń, zwykle w pierwszej trójce najpopularniejszych obiektów. Zestawienie wszystkich pokoi oraz pierwszych dziesięciu najpopularniejszych w nich obiektów przedstawiono w Tab. 2.15.

Na koniec zrealizowano przeliczenie odwrotne do popularności obiektów w pomieszczeniach, tj. określono, jakie pomieszczenia są najpopularniejsze dla danych obiektów. Łącznie uzyskano 55 różnych kombinacji, w Tab. 2.14 przedstawiono po dwa przykłady obiektów silnie- i słabo-powiązanych z konkretnym pomieszczeniem. Po raz kolejny wyraźnie dostrzegalne są dwie grupy obiektów, tj. takie o występowaniu powszechnym (rozkład ich obecności jest rozproszony po rodza-

jach pomieszczeń), a także wysoce specyficzne (gdzie występowanie obserwujemy w pojedynczych pomieszczeniach).

Tabela 2.14: Rozkład liczby wskazań wybranych obiektów w poszczególnych pomieszczeniach.

Obiekty silnie związane z pomieszczeniem				Obiekty słabo związane z pomieszczeniem			
Pom.	Obiekt: Wanna	Pom.	Obiekt: Piekarnik	Pom.	Obiekt: Roślina	Pom.	Obiekt: Lampa
Łazienka	15	Kuchnia	15	Salon	14	Salon	15
Toaleta	1	Salon	1	Kuchnia	14	Sypialnia	15
Salon	0	Łazienka	0	Sypialnia	14	Pok. dziecięcy	15
Kuchnia	0	Sypialnia	0	Hol	13	Biuro	15
Sypialnia	0	Hol	0	Pok. dziecięcy	13	Kuchnia	13
Hol	0	Garaż	0	Biuro	13	Hol	12
Garaż	0	Pok. dziecięcy	0	Łazienka	10	Garaż	12
Pok. dziecięcy	0	Biuro	0	Toaleta	9	Garderoba	12
Biuro	0	Garderoba	0	Garaż	4	Łazienka	11
Garderoba	0	Toaleta	0	Garderoba	4	Toaleta	10

Porównanie z literaturą

Aby podkreślić istotność wykonanego badania sondażowego w kontekście odpowiedzi na faktyczne oczekiwania klientów, przeprowadzono porównanie pokrycia uzyskanych odpowiedzi z innymi, publicznie dostępnymi zbiorami danych. W tym celu wykorzystano dwa zbiory: zaczerpnięty z projektu “What’s in my Room? Object Recognition on Indoor Panoramic Images” [214] (nazywany dalej skrótowo *WiMR*), a także “PANDORA” zaproponowany przez Xu i in. w publikacji “Panoramic Detection Dataset for Object with Orientation” [215]. Porównano, jaki procent etykiet klas dostępnych w tych zbiorach pokrywa się z wynikami zestawienia ankiety. Analizowano pięć kolejnych progów, tj. najczęściej oczekiwanych 10, 20, 30, 40 i 50 rodzajów obiektów. Zestawienie przedstawiono w Tab. 2.16. Zbiór *WiMR* wykazał niskie pokrycie oferowanych etykiet klas nawet z najbardziej podstawowym zbiorem oczekiwań. Zbiór *PANDORA* stanowić może natomiast dobry próg wyjścia do dalszych badań i rozwoju, w przypadku 20 najczęściej występujących obiektów odnaleźć w nim możemy 85% z nich. Niestety, powyżej progu 20 klas etykiet pokrycie zauważalnie spada, a zbiór wymaga dalszego rozwoju.

Tabela 2.15: Dziesięć najczęściej wskazywanych obiektów w poszczególnych pomieszczeniach.

Pomieszczenie (Część I)					
Poz.	Salon	Kuchnia	Łazienka	Sypialnia	Hol
1	Krzesło	Lodówka	Wanna	Łóżko	Drzwi
2	Stół	Mikrofalówka	Toaleta	Szafa	Lustro
3	TV	Zmywarka	Prysznic	Poduszka	Gniazdko
4	Konsola	Kuchenka	Suszarka do wł.	Stolik nocny	Włącznik
5	Kanapa	Ekspres do kawy	Lustro	Lampa	Cz. alarmowy
6	Głośnik	Toster	Gniazdko	Okno	Roślina
7	Poduszka	Piekarnik	Kran	Drzwi	Lampa
8	Lampa	Okno	Grzejnik	Gniazdko	Obraz
9	Okno	Gniazdko	Drzwi	Półka	Zegar
10	Drzwi	Zlew	Zlew	Klimatyzacja	Wazon

Pomieszczenie (Część II)					
Poz.	Garaż	Pokój dziecięcy	Biuro	Garderoba	Toaleta
1	Gniazdko	Krzesło	Krzesło	Szafa	Drzwi
2	Samochód	Konsola	Lampa	Drzwi	Gniazdko
3	Włącznik	Łóżko	Komputer	Gniazdko	Kran
4	Cz. alarmowy	Szafa	Biurko	Półka	Toaleta
5	Drzwi	Poduszka	Gniazdko	Włącznik	Lustro
6	Półka	Lampa	Oczyszczacz pow.	Suszarka na pr.	Zlew
7	Gaśnica	Komputer	Zegar	Lampa	Włącznik
8	Lampa	Gniazdko	Głośnik	Lustro	Półka
9	Okno	Dywan	Klimatyzacja	Szafka	Grzejnik
10	Bojler	Roleta	Okno	Roleta	Okno

Tabela 2.16: Porównanie pokrycia oczekiwanych przez użytkowników etykiet obiektów z etykietami dostępnymi w wybranych zbiorach danych.

	TOP	WiMR	PANDORA
10		50%	100%
20		45%	85%
30		37%	63%
40		28%	48%
50		22%	42%

2.2.5 Podsumowanie

Przeprowadzone badania rynku i oczekiwań użytkowników wykazały silną potrzebę integracji systemów typu cyfrowy bliźniak z rzeczywistym wyposażeniem przestrzeni mieszkalnych i użytkowych. Analiza branży wskazuje na dynamiczny rozwój rozwiązań z zakresu zarządzania nieruchomościami oraz wysoką aktywność konkurencji, szczególnie w obszarach powiązanych z technologiami BIM, IoT i sztuczną inteligencją. Produkt SIM-ON, jako nowy gracz na rynku, wymaga sprecyzowanej strategii ukierunkowanej na prostotę, integrację oraz niską barierę wejścia. Analiza SWOT i porównanie z konkurencją potwierdziły istnienie luki funkcjonalnej: brak systemu łączącego bieżącą analizę wizualną z dynamiczną aktualizacją modeli przestrzennych.

Wyniki badań sondażowych dobitnie wskazują, że oczekiwania użytkowników wykraczają poza typowe obiekty smart czy urządzenia IoT. Wśród najczęściej wskazywanych elementów znalazły się klasyczne, analogowe komponenty wyposażenia: krzesła, stoły, lampy czy szafy. Sugeruje to konieczność budowy systemów rozpoznających nie tylko nowoczesne urządzenia elektroniczne, ale również codzienne meble, stanowiące istotny element semantyki przestrzeni. Takie podejście nie tylko zwiększa użyteczność platformy SIM-ON, ale również podnosi jej wartość z punktu widzenia końcowego odbiorcy.

Rozdział 3

Systemy pozycjonowania radiowego

Technologie lokalizacyjne odgrywają kluczową rolę w wielu współczesnych zastosowaniach, od systemów *Przemysłu 4.0*, takich jak pojazdy autonomiczne [216], poprzez monitoring środowiskowy czy prognozowanie pogody, aż po zastosowania militarne [217]. W zależności od wymagań, można wyróżnić techniki pozycjonowania w przestrzeniach zewnętrznych (z ang. *Outdoor Positioning*) oraz wewnętrznych (z ang. *Indoor Positioning*) [218]. Choć globalne metody lokalizacji na zewnątrz, takie jak system GPS, stały się powszechnym elementem codziennego życia i są wykorzystywane m.in. w topografii, geodezji, hydrografii czy transporcie [219], ich skuteczność drastycznie spada w środowiskach zabudowanych i wewnątrz budynków [220]. Z tego powodu rozwijane są alternatywne rozwiązania dedykowane przestrzeniom wewnętrznym, bazujące na takich technologiach, jak dźwięk [221], obrazowanie optyczne [222], fale elektromagnetyczne [223] czy pole magnetyczne [224]. Choć różnią się one zasadami działania, wszystkie koncentrują się na precyzyjnym określaniu położenia obiektów w mniejszych skalach dystansów, tj. w najbliższym otoczeniu człowieka. Coraz częściej, podobnie jak w przypadku systemów wizji komputerowej, elementem tych rozwiązań staje się integracja metod sztucznej inteligencji oraz ich powszechna dostępność, np. w smartfonach. Pomimo tych postępów, zapewnienie wysokiej precyzji, niskich kosztów i łatwości wdrożenia w zróżnicowanych warunkach wciąż pozostaje aktywnym obszarem ba-

dań.

Celem niniejszego rozdziału jest prezentacja kompletnego procesu badawczego, od analizy teoretycznej, poprzez konstrukcję urządzeń, aż po walidację eksperymentalną w warunkach zbliżonych do rzeczywistych. Dokonano w nim przeglądu, implementacji i oceny technologii radiowego wyznaczania odległości oraz pozycjonowania obiektów. Rozdział otwarto osadzeniem pracy w kontekście literaturowym (Roz. 3.1), omawiając wybrane metody wyznaczania dystansu (Roz. 3.1.1), a także geometryczne i algebraiczne podejścia do pozycjonowania przestrzennego (Roz. 3.1.2). W tej części przedstawiono również przegląd kluczowych technologii i protokołów radiowych (Roz. 3.1.3). Następnie, w Roz. 3.2 skupiono się na prezentacji elementów autorskich. Obejmowały one digitalizację stanowiska badawczego (Roz. 3.2.1) oraz konstrukcję dedykowanych urządzeń do radiowego wyznaczania dystansu (Roz. 3.2.2). Rozdział zamyka opis budowy i walidacji pełnego prototypu systemu lokalizacji (Roz. 3.3), przedstawiając jego architekturę fizyczną (Roz. 3.3.1 i 3.3.2), konfigurację i realizację eksperymentów (Roz. 3.3.3 i 3.3.4) oraz szczegółową analizę wyników (Roz. 3.3.5), w tym porównanie zastosowanych technologii, ocenę zmienności między urządzeniami oraz analizę pomiarów w funkcji odległości, która prowadzi do końcowych wniosków (Roz. 3.3.6).

Podsumowując, wkład rozdziału w przeprowadzone prace naukowe i wdrożeniowe to:

- dogłębny przegląd i usystematyzowanie metod pozycjonowania radiowego, stanowiące podstawę merytoryczną dla podjętych decyzji projektowych;
- projekt i fizyczna realizacja prototypu systemu pozycjonowania radiowego;
- opracowanie i konstrukcja autorskich urządzeń pomiarowych (kotwic i lokalizatorów) dostosowanych do specyfiki eksperymentów;
- implementacja rozwiązania opartego na standardzie Matter, umożliwiającego zarządzanie heterogeniczną siecią urządzeń IoT;
- przeprowadzenie kompleksowej analizy porównawczej wybranych technologii radiowych w kontekście pomiaru odległości wewnątrz pomieszczeń;
- walidacja eksperymentalna dokładności i stabilności opracowanego systemu, której wyniki pozwoliły sformułować konkretne rekomendacje do budowy docelowego systemu fuzyjnego.

3.1 Osadzenie pracy w kontekście

Zakres zagadnień, metod oraz kierunków rozwoju w obszarze pozycjonowania radiowego jest na tyle rozległy, że w niniejszym opracowaniu przedstawione zostały wyłącznie te elementy, które pozostają bezpośrednio powiązane z rzeczywiście przeprowadzonymi badaniami oraz z rozwiązaniami zaimplementowanymi w końcowym systemie fuzyjnym. Dla uproszczenia są to: radiowe wyznaczanie dystansu (omówione w Roz. 3.1.1), korzystające z jego wyników pozycjonowanie w przestrzeni (Roz. 3.1.2), a także stosowane do ich realizacji technologie (Roz. 3.1.3).

3.1.1 Radiowe wyznaczanie dystansu

Proponowany prototyp systemu pozycjonowania radiowego opiera swoje działanie na ustaleniu odległości śledzonego obiektu, zaopatrzonego w lokalizator, od pewnego zestawu znanych punktów referencyjnych (tzw. kotwic). W praktyce istota problemu sprowadza się do wiarygodnej obserwacji i powiązania odbieranego sygnału radiowego z konkretną wartością fizyczną. Zależnie od zaimplementowanej metody, może to być moc sygnału, czas przelotu sygnału radiowego, różnica czasu pomiędzy kolejnymi połączeniami, a nawet sam kąt nadejścia fali radiowej.

W ramach prowadzonych prac wykorzystano dwa rozwiązania, zależnie od technologii radiowej (szczegóły w Roz. 3.1.3). Pierwszym była realizacja oparta o analizę czasu, w jakim nadany sygnał pokonuje drogę od nadajnika do odbiornika (z ang. *Time of Flight*, ToF). Na jego podstawie, przy założeniu synchronizacji zegarów urządzeń radiowych lub poprzez odpowiedni algorytm wymiany wiadomości, wyznaczyć można również konkretny czas odbioru sygnału (z ang. *Time of Arrival*, ToA). Drugim rozwiązaniem była natomiast analiza wskaźnika RSSI, tj. mocy odbieranego sygnału radiowego (z ang. *Received Signal Strength Indicator*). Realizacja przebiegała w zróżnicowanych warunkach środowiska, zarówno w sytuacji bezpośredniej widoczności komunikujących się urządzeń, jak również przy obecności przeszkód fizycznych (mebli, ścian, itp.).

Time of Flight

Zasada wyznaczania odległości w metodzie *Time of Flight* opiera się na powszechnie znanym równaniu: dystans jest iloczynem prędkości, z którą coś się porusza, oraz czasu, przez który to robi. W przypadku fal elektromagnetycznych prędkość ta jest równa prędkości światła c , będącej stałą fizyczną. Kluczowym wyzwaniem pozostaje więc możliwie najdokładniejsze wyznaczenie czasu propagacji sygnału, ponieważ nawet niewielkie błędy, po przemnożeniu przez c , przekładają się na znaczące rozbieżności w oszacowanej odległości.

Popularną metodą wyznaczania tego czasu jest dwukierunkowy pomiar odległości (z ang. *Two-Way Ranging*, TWR). Realizuje się go zwykle na jeden z dwóch sposobów: poprzez komunikację jednostronną (z ang. *Single-Sided TWR*, SS-TWR) albo symetryczną dwustronną (z ang. *Double-Sided TWR*, DS-TWR). Różnią się one głównie sposobem kompensacji błędów wynikających z braku synchronizacji pomiędzy zegarami urządzenia lokalizatora i kotwicy. Niemniej, wartym nakreślenia jest, że w literaturze odnaleźć można wiele metod pokrewnych oraz modyfikacji wyżej wymienionych, np. wersje: dwustronną asynchroniczną [225], z wielokrotnym potwierdzeniem [226], z nierównym czasem odpowiedzi [227], czy też równoległą [228].

Single-Sided TWR W metodzie *Single-Sided Two-Way Ranging* lokalizator wysyła sygnał wywołania do kotwicy, która po odebraniu wiadomości odczekuje określony czas przetwarzania T_{proc} i odsyła odpowiedź [229]. Zakłada się, że czas przelotu sygnału w obu kierunkach jest równy, co pozwala obliczyć czas propagacji jako połowę całkowitego czasu pomiędzy wysłaniem wywołania a odebraniem odpowiedzi. Wzór na czas przelotu przyjmuje wtedy postać:

$$T_{\text{tof}} = \frac{T_{\text{całk}} - T_{\text{proc}}}{2}, \quad (3.1)$$

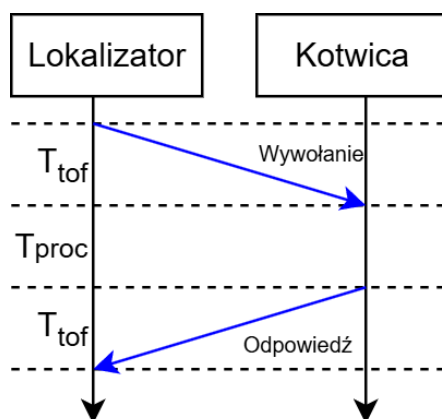
gdzie $T_{\text{całk}}$ oznacza czas całkowity pomiędzy nadaniem wywołania a odebraniem odpowiedzi, natomiast T_{proc} to czas przetwarzania w kotwicy.

Dystans wyznacza się następnie ze wspomnianej zależności prędkości i czasu:

$$Dystans = c \cdot T_{\text{tof}}, \quad (3.2)$$

gdzie c jest wspomnianą prędkością światła.

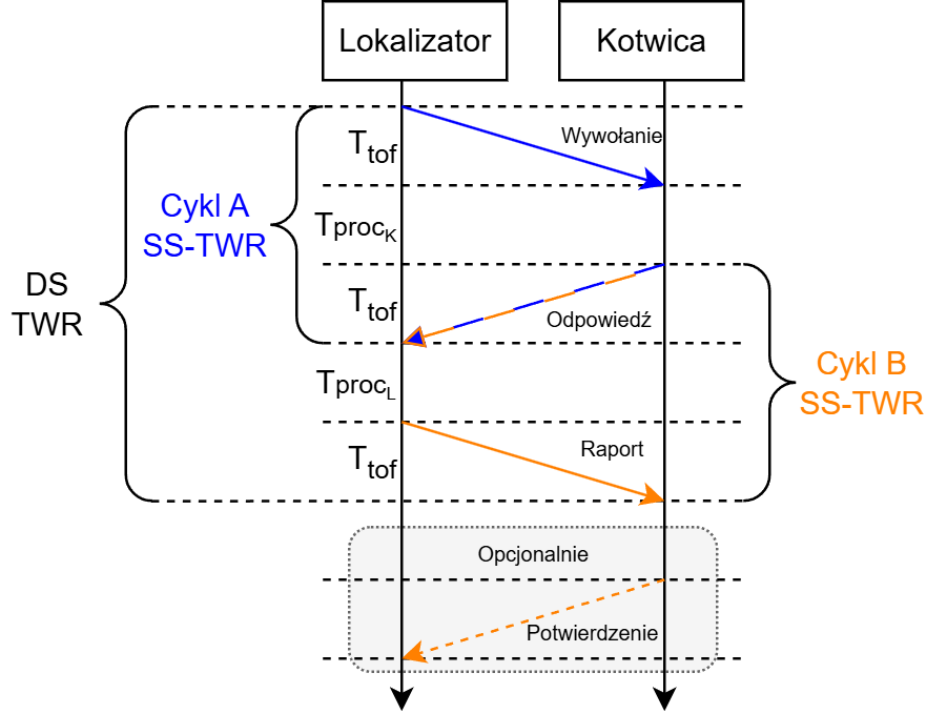
Metoda SS-TWR jest prosta w implementacji, a jej główną zaletą jest brak konieczności synchronizacji zegarów pomiędzy wykorzystanymi urządzeniami. Do cech negatywnych tego rozwiązania należy jednak jego podatność na błędy. Wynikać one mogą np. z niedokładnego oszacowania czasu przetwarzania T_{proc} albo asymetrii w czasie propagacji sygnału w obu kierunkach. Schemat wymiany wiadomości w metodzie SS-TWR prezentowany jest na Rys. 3.1.



Rysunek 3.1: Schemat komunikacji w procesie dwukierunkowego, jednostronnego pomiaru odległości (Single-Sided Two-Way Ranging, SS-TWR). Czas przelotu oznaczono jako T_{tof} , a czas przetwarzania jako T_{proc} .

Double-Sided TWR Metoda *Double-Sided Two-Way Ranging* rozszerza wcześniejsze rozwiązanie o dodatkowe przesłanie wiadomości powrotnej do urządzenia kotwicy (raport). Dla ułatwienia opisu postępowania przedstawiona zostaje ona na Rys. 3.2. Pierwszy etap przetwarzania, tj. wywołanie oraz odpowiedź, jest taki sam jak dla SS-TWR. Określić to możemy mianem pierwszego cyklu komunikacji (*Cykl A*, oznaczony kolorem niebieskim). Druga wiadomość cyklu A (odpowiedź) stanowi jednak tym razem jednocześnie pierwszą wiadomość drugiego cyklu komunikacji (*Cykl B*, oznaczony kolorem pomarańczowym). Dodanie do pro-

cesu dodatkowej, trzeciej wiadomości (raportu) pozwala dokończyć drugi cykl komunikacji i uzyskać efekt dwustronnego wyznaczania czasu propagacji sygnału. Na schemacie uwzględniono również, że całość procedury może kończyć się jeszcze jedną wiadomością zwrotną, potwierdzającą wyznaczone wartości (szary obszar rysunku).



Rysunek 3.2: Schemat komunikacji w metodzie DS-TWR. Czas przelotu oznaczono jako T_{tof} . Czas przetwarzania po stronie urządzenia kotwicy jako T_{proc_K} , a po stronie lokalizatora jako T_{proc_L} .

Na podstawie tak zdefiniowanych operacji wyprowadzić możemy następujące równania, opisujące zależności czasowe [225]:

$$T_{\text{cykl}_A} = 2T_{\text{tof}} + T_{\text{proc}_K}, \quad (3.3)$$

$$T_{\text{cykl}_B} = 2T_{\text{tof}} + T_{\text{proc}_L}, \quad (3.4)$$

gdzie T_{cykl_A} oznacza całkowity czas trwania cyklu A, liczony od momentu nadania

wiadomości wywołania przez lokalizator do momentu odebrania wiadomości odpowiedzi, a T_{procK} jest czasem przetwarzania wiadomości w kotwicy. Analogicznie, T_{cyklB} to całkowity czas trwania cyklu B, liczony od chwili nadania wiadomości odpowiedzi przez kotwicę do momentu odebrania przez nią wiadomości raportu, a T_{procL} jest czasem przetwarzania wiadomości w lokalizatorze.

Na podstawie tak określonych czasów trwania obu cykli komunikacji i zachowując ciągłość przyjętych oznaczeń, wyznaczyć możemy czas pojedynczej propagacji sygnału jako:

$$T_{\text{tof}} = \frac{T_{\text{cyklA}} - T_{\text{procK}} + T_{\text{cyklB}} - T_{\text{procL}}}{4} \quad (3.5)$$

Uzyskany wynik podstawić możemy do wzoru 3.2 i wyznaczyć szukaną odległość kotwicy od lokalizatora. W porównaniu do SS-TWR, metoda DS-TWR cechuje się większą odpornością na błędy komunikacji oraz niedokładnego pomiaru czasu przetwarzania. Wadą jest jednak większa liczba wymian ramek, co zwiększa złożoność i czas realizacji pojedynczego pomiaru.

Received Signal Strength

Rozwiązaniem alternatywnym do analizy czasu przelotu sygnału od nadajnika do odbiornika jest obserwacja wskaźnika mocy odbieranego sygnału (z ang. *Received Signal Strength Indicator*, RSSI). Ideologicznie jest to rozwiązanie zauważalnie prostsze i opiera się na założeniu, że im bliżej do nadajnika, tym mocniejszy jest odbierany sygnał. Wyznaczenie odległości realizowane może być na podstawie logarytmicznego modelu strat. Konieczne jest wtedy jednak określenie parametrów wejściowych do pozycjonowania, tj. referencyjnej siły sygnału przy określonej odległości (typowo jednego metra), a także eksperymentalnie dobranej wartości współczynnika środowiska. Wykorzystywany w badaniach wzór zaczerpnięto z [230], gdzie zdefiniowany jest on jako:

Definicja 1. Niech **RSSI** będzie dynamicznie odczytywaną wartością siły sygnału w danej chwili czasu, **RSSI_{ref}** wartością odniesienia RSSI mierzoną w znanej odległości (1 m), a **n_{env}** współczynnikiem środowiskowym, opisującym tłumienie sygnału w danych warunkach otoczenia. Wtedy:

$$Dystans = 10^{\left(\frac{RSSI_{ref} - RSSI}{10 * n_{env}}\right)}. \quad (3.6)$$

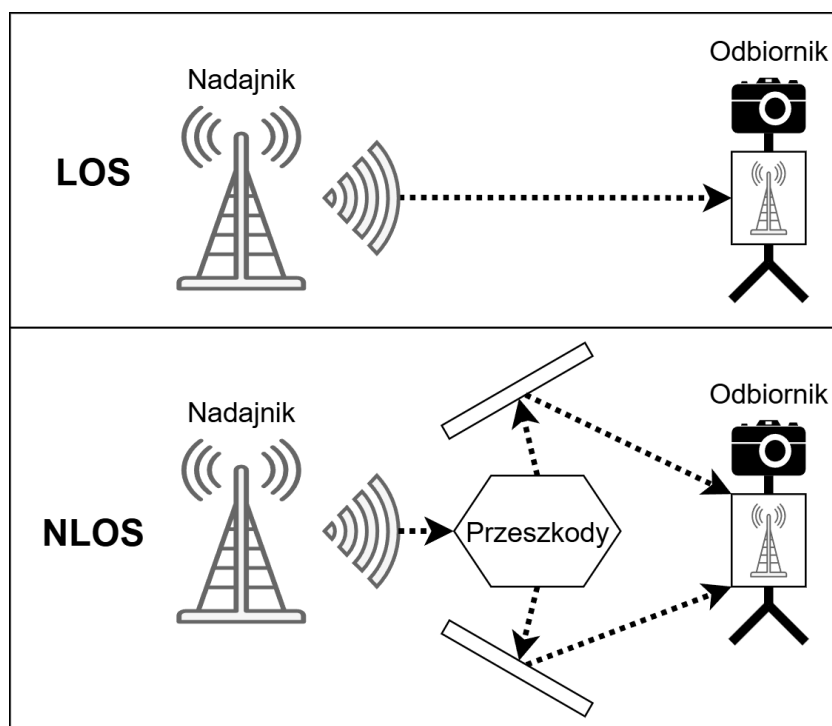
Wykorzystując RSSI do wyznaczania odległości zgodnie z tym wzorem, należy zdawać sobie sprawę z dużych różnic parametrów fizycznych pomiędzy kolejnymi sztukami modułów elektronicznych (nawet dla tego samego modelu produktu), a także pamiętać o konieczności pozostawienia użytkownikowi końcowemu możliwości manipulacji wartością współczynnika środowiska. Wdrożenie rozwiązania zwykle wymagać będzie kalibracji poszczególnych urządzeń celem określenia ich faktycznych parametrów fizycznych.

Do pozytywnych cech tego rozwiązania należy uniwersalność, a także potencjalna redukcja intensywności komunikacji radiowej w scenariuszach wielkoskalowych. Nadajnik sygnału może propagować go niezależnie od liczby nasłuchujących na komunikację urządzeń lokalizacyjnych. Nie jest konieczne zachowanie procedury wymiany informacji, a tym samym wraz ze wzrostem liczby urządzeń końcowych nie rośnie narzut obliczeniowy po stronie kotwic.

Warunki środowiska

Istotnym czynnikiem działania systemów pozycjonowania jest środowisko i warunki, w których jest ono realizowane. Głównego podziału sytuacji dokonać możemy na dwa scenariusze. W warunkach określanych mianem LOS (z ang. *Line-of-Sight*) sygnał od nadajnika do odbiornika przebywa drogę bezpośrednią, a na jego ścieżce nie znajdują się żadne przeszkody fizyczne. W warunkach NLOS (z ang. *Non-Line-of-Sight*) dominują odbicia sygnału od napotkanych obiektów oraz jego silne rozpraszanie. Schemat ideowy obu scenariuszy prezentowany jest na Rys. 3.3.

W zagadnieniu systemów stosowanych wewnątrz pomieszczeń, szczególnie domowych, trudno jest uzyskać warunki LOS. Ściany, meble, a także przemieszczający się ludzie skutecznie rozpraszają sygnał i generują jego odbicia. Efektywnie prowadzi to do możliwości odebrania przez lokalizator więcej niż jednego sygnału dotyczącego tej samej komunikacji, co skutkuje z kolei błędami wyznaczania czasów propagacji sygnału. Ze względu na specyfikę są to jednak zwykle błędy dodatnie, tj. wyniki są przeszacowane.



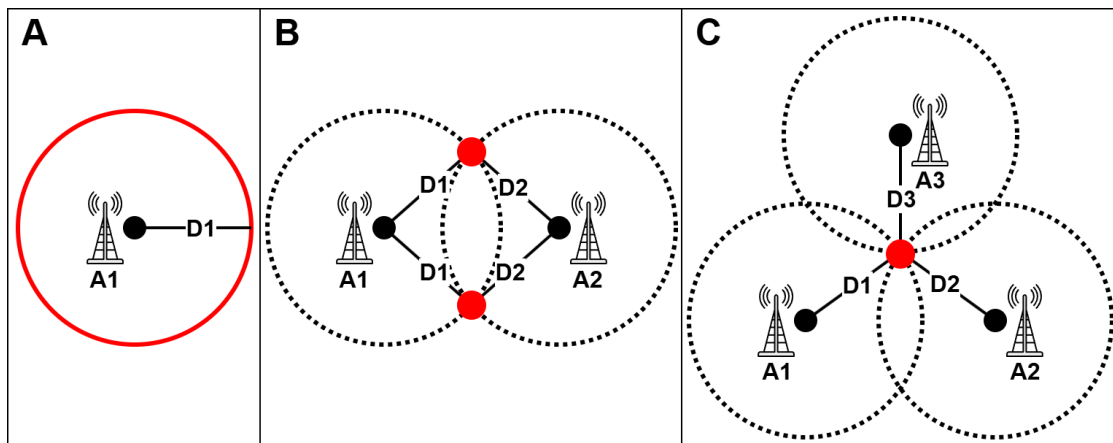
Rysunek 3.3: Schemat komunikacji w warunkach bez przeszkód (Line-of-Sight, LOS), a także z odbiciami od obiektów (Non-Line-of-Sight, NLOS).

3.1.2 Pozycjonowanie przestrzenne

Znając pewien zbiór odległości, jakie dzieli urządzenie lokalizatora od kolejnych punktów referencyjnych, a także precyzyjne współrzędne tychże punktów, przejść można do określenia pozycji samego lokalizatora. W niniejszej pracy zastosowaną metodą lokalizacji w przestrzeni jest trilateracja. Metoda ta posiada wiele różnych realizacji pod względem obranego podejścia obliczeniowego. Niemniej, zasady jej działania można łatwo wyjaśnić, dla pozycjonowania na płaszczyźnie również graficznie. W zależności od liczby wymiarów, w których ma nastąpić pozycjonowanie, zmienia się liczba niezależnych pomiarów, niezbędnych do uzyskania wyników. W minimalnym scenariuszu niezbędnych jest tyle pomiarów, ile wymiarowa jest przestrzeń lokalizacji. Uzyskujemy wtedy jednak wiele potencjalnych wyznaczonych pozycji, a końcowy wybór musi oprzeć się na dodatkowej znajomości realiów zastosowania (np. odrzucając wyniki pod podłogą pomieszczenia). Aby uzyskać konkretny wynik, pomiarów musi być o co najmniej jeden więcej.

Rozwiązanie geometryczne

Prezentację zasady działania metody trilateracji, w scenariuszu pozycjonowania na płaszczyźnie, przedstawiono na Rys. 3.4. Punktem wyjścia do rozważań jest, że znajomość odległości od lokalizatora do kotwicy pozwala na wykreślenie sfery, na której znajduje się lokalizator. Sprowadzone na płaszczyznę, prezentowane jest to w części A. Kolejne kroki postępowania dokładać będą kolejne pomiary, a co za tym idzie kolejne sfery, których punktów wspólnych będziemy szukać. Dysponując dwoma zestawami pomiarów, na płaszczyźnie wskazać już możemy konkretne położenie lokalizatora. Wyniki są jednak dwa (część B). W zależności od dodatkowej wiedzy o realiach zastosowania jeden z wyników musimy albo ręcznie odrzucić, albo zwiększyć liczbę pomiarów. Korzystając z trzech pomiarów, wyniki zacieśnić możemy do pojedynczego wskazanego punktu (część C). Przy założeniu znajomości położenia wszystkich kotwic, wyznaczamy współrzędne lokalizatora. Przedstawione postępowanie stanowi rozwiązanie analityczne metody trilateracji i zastosowane może być również do przestrzeni trójwymiarowej. Skuteczna wizualizacja tego scenariusza jest jednak kłopotliwa.



Rysunek 3.4: Wizualizacja operacji trilateracji na płaszczyźnie w scenariuszach z oparciem pozycjonowania na pomiarach z pojedynczej kotwicy (scenariusz A), z dwóch kotwic (scenariusz B) oraz z trzech kotwic (scenariusz C). Wyznaczone położenie lokalizatora oznaczono kolorem czerwonym, kolejne kotwice jako A1, A2 i A3, a wyznaczone dystanse do nich jako D1, D2 i D3.

Rozwiązanie algebraiczne

Wyprowadzenie rozwiązania algebraicznego należy rozpocząć od podstawowego elementu, tj. równania i -tej sfery, które dla sfery o środku w punkcie $P_i = (x_i, y_i, z_i)$ oraz promieniu r_i ma postać:

$$(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 = r_i^2, \quad (3.7)$$

gdzie (x, y, z) oznacza współrzędne dowolnego punktu leżącego na powierzchni sfery. W prowadzonych rozważaniach opisywać będziemy scenariusz wyznaczenia położenia lokalizatora w trójwymiarze i w oparciu o trzy pomiary odległości. Zakładamy zatem istnienie trzech punktów odniesienia (środków trzech sfer): $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, $P_3 = (x_3, y_3, z_3)$, oraz odpowiadających im promieni r_1, r_2, r_3 . Szukana pozycja lokalizatora (x, y, z) spełnia wtedy układ równań kwadratowych:

$$\begin{cases} (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = r_1^2, \\ (x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = r_2^2, \\ (x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = r_3^2, \end{cases} \quad (3.8)$$

który po rozpisaniu i uporządkowaniu zapisu przyjmuje postać:

$$\begin{cases} x^2 + y^2 + z^2 - 2x_1x - 2y_1y - 2z_1z = r_1^2 - (x_1^2 + y_1^2 + z_1^2), \\ x^2 + y^2 + z^2 - 2x_2x - 2y_2y - 2z_2z = r_2^2 - (x_2^2 + y_2^2 + z_2^2), \\ x^2 + y^2 + z^2 - 2x_3x - 2y_3y - 2z_3z = r_3^2 - (x_3^2 + y_3^2 + z_3^2). \end{cases} \quad (3.9)$$

Chcąc sprowadzić problem do układu równań liniowych, które będzie się dało następnie rozwiązać metodą eliminacji Gaussa, wprowadzamy do układu nową zmienną:

$$x_0 = x^2 + y^2 + z^2, \quad (3.10)$$

która po podstawieniu modyfikuje układ do układu o trzech równaniach oraz czte-

rech niewiadomych i przyjmuje postać:

$$\begin{cases} x_0 - 2x_1x - 2y_1y - 2z_1z = r_1^2 - (x_1^2 + y_1^2 + z_1^2), \\ x_0 - 2x_2x - 2y_2y - 2z_2z = r_2^2 - (x_2^2 + y_2^2 + z_2^2), \\ x_0 - 2x_3x - 2y_3y - 2z_3z = r_3^2 - (x_3^2 + y_3^2 + z_3^2). \end{cases} \quad (3.11)$$

Dla metody eliminacji Gaussa charakterystycznym zapisem jest macierzowy, po sprowadzeniu do którego układ przyjmuje postać:

$$\underbrace{\begin{bmatrix} 1 & -2x_1 & -2y_1 & -2z_1 \\ 1 & -2x_2 & -2y_2 & -2z_2 \\ 1 & -2x_3 & -2y_3 & -2z_3 \end{bmatrix}}_A \cdot \underbrace{\begin{bmatrix} x_0 \\ x \\ y \\ z \end{bmatrix}}_{\mathbf{x}_{\text{vec}}} = \underbrace{\begin{bmatrix} r_1^2 - (x_1^2 + y_1^2 + z_1^2) \\ r_2^2 - (x_2^2 + y_2^2 + z_2^2) \\ r_3^2 - (x_3^2 + y_3^2 + z_3^2) \end{bmatrix}}_{\mathbf{b}}, \quad (3.12)$$

którą znamy wprost z algebry, jako opis układu równań liniowych:

$$A\mathbf{x}_{\text{vec}} = \mathbf{b}, \quad (3.13)$$

gdzie A to macierz współczynników, x_{vec} to wektor niewiadomych, a b to wektor wyrazów wolnych.

Przed przystąpieniem do dalszego omawiania konieczne jest przypomnienie kilku pojęć z algebry liniowej oraz wskazanie dodatkowych założeń przyjętych w analizie.

Rangą macierzy nazywamy liczbę liniowo niezależnych wierszy (lub kolumn) tej macierzy, co w praktyce oznacza liczbę niezależnych równań w opisywanym układzie (dalej: $\text{ranga}(A)$).

Jądro macierzy A (dalej: $\text{jądro}(A)$) to zbiór wszystkich wektorów $\mathbf{x}_{\text{vec}}$, spełniających równanie jednorodne $A\mathbf{x}_{\text{vec}} = \mathbf{0}$.

Wymiar jądra macierzy A określa natomiast, ile parametrów swobodnych posiada rozwiązanie takiego układu jednorodnego (dalej: $\text{wymiar}(A)$).

Na potrzeby rozprawy zakłada się uproszczone scenariusze trilateracji, ponieważ sam proces jej realizacji nie jest tematem przewodnim prowadzonych badań. Przyjmuje się, że rozstawione kotwice nie leżą wszystkie na jednej prostej. W takiej

sytuacji każda z kotwic dostarcza niezależnych informacji pomiarowych, a macierz układu równań ma rangę równą liczbie kotwic.

Pomiędzy rangą macierzy a wymiarem jej jądra zachodzi zależność wynikająca z twierdzenia o wymiarach:

$$\text{ranga}(A) + \text{wymiar}(\text{jądro}(A)) = \text{liczba niewiadomych}. \quad (3.14)$$

W naszym przypadku:

$$\text{ranga}(A) = 3, \quad \text{liczba niewiadomych} = 4, \quad (3.15)$$

stąd po podstawieniu do 3.14:

$$\text{wymiar}(\text{jądro}(A)) = 4 - 3 = 1. \quad (3.16)$$

Uzyskanie takiego wymiaru jądra oznacza, że zbiór rozwiązań liniowych ma postać prostej, wyrażonej w $\mathbb{R}^4$, co z kolei pozwala sprowadzić dalsze obliczenia do jednego parametru i w efekcie wyznaczyć maksymalnie dwa punkty spełniające warunki trilateracji w przestrzeni trójwymiarowej. Korzystając z wyprowadzonej obserwacji, można zastosować ogólny wzór na równanie prostej w postaci parametrycznej:

$$\mathbf{x}_{vec} = \mathbf{x}_p + t \mathbf{x}_h, \quad (3.17)$$

gdzie:

- $\mathbf{x}_p$ to rozwiązanie szczególne układu $A\mathbf{x}_{vec} = \mathbf{b}$ (z ang. *Particular Solution*),
- $\mathbf{x}_h$ to rozwiązanie jednorodne układu $A\mathbf{x}_{vec} = \mathbf{0}$ (z ang. *Homogeneous Solution*),
- $t \in \mathbb{R}$ to parametr rzeczywisty określający położenie na prostej.

Dla tak zdefiniowanego równania należy następnie wyznaczyć wartości $\mathbf{x}_p$ oraz $\mathbf{x}_h$. Krok ten zrealizować można na wiele sposobów, w ramach prowadzonych badań zastosowano jednak klasyczną metodę eliminacji Gaussa. Dla zachowania przejrzystości wywodu zakłada się, że czytelnik zaznajomiony jest z tym procesem, a wynikami obliczeń są wyznaczone wektory $\mathbf{x}_p$ i $\mathbf{x}_h$. Wraz ze zdefiniowanym wcześniej

wektorem szukanych współrzędnych (3.12), stanowią one trójkę czterowymiarowych wektorów:

$$\mathbf{x}_{vec} = \begin{bmatrix} x_0 \\ x \\ y \\ z \end{bmatrix}, \quad \mathbf{x}_p = \begin{bmatrix} x_{p0} \\ x_{p1} \\ x_{p2} \\ x_{p3} \end{bmatrix}, \quad \mathbf{x}_h = \begin{bmatrix} x_{h0} \\ x_{h1} \\ x_{h2} \\ x_{h3} \end{bmatrix}. \quad (3.18)$$

Tak przygotowany zestaw parametrów podstawiamy do równania ogólnego (3.17) i uzyskujemy układ równań, który przedstawić możemy macierzowo:

$$\begin{bmatrix} x_0 \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{p0} \\ x_{p1} \\ x_{p2} \\ x_{p3} \end{bmatrix} + t \cdot \begin{bmatrix} x_{h0} \\ x_{h1} \\ x_{h2} \\ x_{h3} \end{bmatrix}, \quad (3.19)$$

lub w postaci układu czterech równań:

$$\begin{cases} x_0 = x_{p0} + t x_{h0}, \\ x = x_{p1} + t x_{h1}, \\ y = x_{p2} + t x_{h2}, \\ z = x_{p3} + t x_{h3}. \end{cases} \quad (3.20)$$

Pamiętając o wprowadzonej wcześniej zmiennej pomocniczej (3.10):

$$x_0 = x^2 + y^2 + z^2, \quad (3.21)$$

i podstawiając do niej wartości zgodnie z układem równań (3.20), otrzymujemy:

$$x_{p0} + t x_{h0} = (x_{p1} + t x_{h1})^2 + (x_{p2} + t x_{h2})^2 + (x_{p3} + t x_{h3})^2, \quad (3.22)$$

co po rozwinięciu kwadratów oraz grupując wyrazy równania względem potęg parametru t , daje klasyczne równanie kwadratowe:

$$\underbrace{(x_{h1}^2 + x_{h2}^2 + x_{h3}^2)}_{a_w} t^2 + \underbrace{(2x_{p1}x_{h1} + 2x_{p2}x_{h2} + 2x_{p3}x_{h3} - x_{h0})}_{b_w} t + \underbrace{(x_{p1}^2 + x_{p2}^2 + x_{p3}^2 - x_{p0})}_{c_w} = 0. \quad (3.23)$$

W ten sposób docieramy do ostatniego kroku obliczeń, tj. rozwiązania równania kwadratowego, które spodziewamy się, że zwróci dwie końcowe wartości parametru t . Rozwiązujemy je na podstawie klasycznego wzoru opartego na wyróżniku równania kwadratowego:

$$t_{1,2} = \frac{-b_w \pm \sqrt{b_w^2 - 4a_w c_w}}{2a_w}, \quad (3.24)$$

co po ponownym podstawieniu do równania ogólnego (3.17) wyznaczonych metodą eliminacji Gaussa wektorów $\mathbf{x}_p$ i $\mathbf{x}_h$ (3.18) oraz powyższych wartości $\mathbf{t}_1$ i $\mathbf{t}_2$, daje końcowe wyznaczone punkty położenia lokalizatora:

$$\mathbf{x}_{vec}(t_1) = \mathbf{x}_p + t_1 \mathbf{x}_h, \quad (3.25)$$

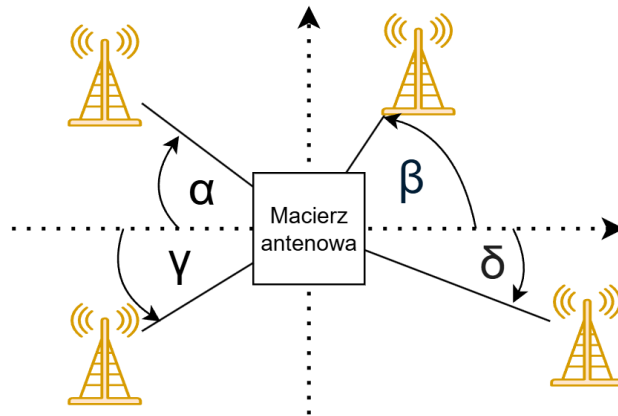
$$\mathbf{x}_{vec}(t_2) = \mathbf{x}_p + t_2 \mathbf{x}_h. \quad (3.26)$$

Jeżeli równanie nie może być rozwiązane w zbiorze liczb rzeczywistych, a sfery nie mają punktów wspólnych, np. na skutek zbyt krótkich wyznaczonych dystansów (niedoszacowania wyznaczonej odległości), to wykorzystanie tylko rzeczywistej części uzyskanych wyników może posłużyć za pewne przybliżenie wyznaczonej pozycji. Szczegółowa analiza takich przypadków, a także wyprowadzenia algebraiczne dla scenariuszy z większą liczbą pomiarów, wykraczają poza zakres niniejszego omówienia. Ich pełną treść odnaleźć można w publikacji, która stanowiła źródło opracowania matematycznego zrealizowanego systemu [231].

Alternatywne metody lokalizacji

Zakres sposobów określania położenia w przestrzeni jest bardzo rozległy. Przedstawione zostają skrótowo dwa alternatywne rozwiązania, nie wyczerpuje to jednak całego zagadnienia, które samodzielnie stanowi odrębną gałąź nauki.

Kąt nadejścia sygnału W metodzie analizy kąta nadejścia sygnału radiowego (z ang. *Angle-of-Arrival*, AoA) stosuje się układy elektroniczne integrujące więcej niż jedną antenę radiową [232]. Docierająca do urządzenia fala radiowa jest analizowana w poszukiwaniu różnic w fazie lub czasie jej odbioru pomiędzy kolejnymi antenami tego samego lokalizatora. Znając precyzyjnie budowę urządzenia i charakterystykę jego anten, wyznaczyć można położenie źródła sygnału. Schematycznie przedstawiono to na Rys. 3.5. Metoda ta jest silnie podatna na zakłócenia przestrzenne i zwykle łączona jest w systemy hybrydowe wraz z innymi rozwiązaniami [233, 234].



Rysunek 3.5: Schemat ideowy procesu lokalizacji z użyciem analizy kąta nadejścia sygnału radiowego.

Charakterystyka siły sygnału Omówiony wcześniej indeks siły sygnału (RSSI) wykorzystać można inaczej, niż do samego wyznaczania dystansu. Posiadając odpowiednio liczną pulę pomiarów referencyjnych (odczytów w stosunku do wielu kotwic) i dla wielu kolejnych punktów na badanym obszarze, stworzyć możemy mapę spodziewanych wartości (z ang. tzw. *Heatmap*). Odczyty wartości RSSI traktujemy wtedy jako swoistą charakterystykę i analizujemy podobieństwo uzyskiwanych odczytów do wartości spodziewanych [235]. Takie podejście stanowi podstawę metody tzw. “*fingerprintingu radiowego*”, w której każdemu punktowi w przestrzeni przypisywany jest unikalny wzorec siły sygnałów z różnych źródeł. Ich późniejsze porównywanie z zapisanymi wzorcami umożliwia określenie położenia.

3.1.3 Technologie i protokoły radiowe

Na potrzeby rozprawy skoncentrowano się na analizie trzech technologii komunikacji bezprzewodowej: Ultra-Wideband (UWB), Bluetooth Low Energy (BLE) oraz Matter. W niniejszym rozdziale protokoły UWB i BLE, jako standardy o ugruntowanej pozycji na rynku, zostają przedstawione jedynie ogólnikowo. Szczegóły sposobu ich wykorzystania omówiono w kontekście prototypu systemu lokalizacji radiowej (Roz. 3.3.1). Protokół Matter omówiony zostaje dogłębnie, od jego wprowadzenia, przez model danych, aż po prototypowanie.

UWB

UWB to technologia komunikacji bezprzewodowej, a także konkretna warstwa fizyczna (radio), wykorzystująca bardzo szerokie pasmo częstotliwości (zwykle od 3,1 do 10,6 GHz). Znajduje ona szerokie zastosowanie w komunikacji, radarach, wojsku oraz systemach bezpieczeństwa [236]. Jej użycie do precyzyjnego pomiaru odległości i estymacji pozycji zostało wielokrotnie opisane w literaturze. Przykładowo, Margiani et al. [237] badają pomiary kąta nadejścia fali radiowej z wykorzystaniem komercyjnych modułów oferujących funkcję estymacji różnicy faz sygnałów (z ang. *Phase Difference of Arrival*, PDoA). Cossette et al. [238] osiągają estymację relatywnej pozycji robotów autonomicznych z błędem poniżej jednego metra. Z kolei Liu et al. [239] osiągają wysoką precyzję lokalizacji dzięki hybrydowej kompensacji błędów temperaturowych i dystansowych. Inni autorzy proponują podejścia łączące trilaterację i sygnaturę radiową, aby zapewnić dokładność w złożonych środowiskach wewnętrznych [240]. Co więcej, rozwijane są metody oparte na uczeniu maszynowym, zdolne do korekty błędów wynikających z ograniczonej widoczności bezpośredniej sygnалу [241], czy błędów orientacji urządzeń [242]. Porównania komercyjnych systemów UWB (m.in. Ubisense, BeSpoon, DecaWave) przedstawiono natomiast w [243].

BLE

Alternatywą dla UWB są rozwiązania oparte na protokole Bluetooth. Jako odmiana głównego standardu, BLE (z ang. *Bluetooth Low Energy*) to technologia komunikacji radiowej o niskim zużyciu energii, opisana m.in. w [244]. Lokalizacja

za pomocą BLE opiera się zwykle na analizie poziomu wskaźnika mocy odebranego sygnału. Choć metoda ta cechuje się mniejszą precyzją niż UWB, może osiągać dokładność rzędu metra przy odpowiednim przetwarzaniu danych [245–247]. W ostatnich latach rozwijano wiele podejść, które pozwalają uzyskać dokładność poniżej 1 metra w warunkach rzeczywistych [248]. Inne badania pokazują, że BLE może być przydatne w mniej precyzyjnych, ale nadal użytecznych aplikacjach, jak systemy ewakuacji z dokładnością rzędu 10 metrów [249].

Matter

W odpowiedzi na rosnące zapotrzebowanie na interoperacyjność w środowiskach IoT, protokół Matter został zaprojektowany jako ujednolicona warstwa komunikacji i modelowania dla urządzeń pochodzących od różnych producentów. Stanowi on inicjatywę organizacji *Connectivity Standards Alliance* (CSA), wspieraną m.in. przez firmy *Apple*, *Google*, *Amazon* i *Samsung*. Celem protokołu jest zapewnienie spójnego modelu współpracy urządzeń w złożonych środowiskach IoT (np. inteligentnych budynkach), przy jednoczesnym zachowaniu prostoty integracji i wysokiego bezpieczeństwa [250–252].

Jednym z filarów architektury tego protokołu jest jego model danych (z ang. *Data Model*), który hierarchicznie organizuje strukturę urządzeń. Z perspektywy tego modelu, każde urządzenie (z ang. *Device*) składa się z jednego lub więcej węzłów (z ang. *Nodes*). Każdy węzeł dzieli się na zestaw punktów dostępu (z ang. *Endpoints*), z których każdy odpowiada konkretnej funkcji urządzenia, przykładowo sterowaniu ogrzewaniem. Punkty końcowe zawierają klastry (z ang. *Clusters*), które grupują określone funkcjonalności, takie jak ustawienie temperatury lub włączanie/wyłączanie urządzenia. Każdy klaster zawiera atrybuty (z ang. *Attributes*) opisujące bieżący jego stan, np. zadaną temperaturę. Taka modułarna struktura pozwala na izolację poszczególnych komponentów, a tym samym na lepszą organizację kodu źródłowego, ułatwiając zarówno rozwój, jak i utrzymanie oprogramowania. Jest to szczególnie ważne, gdyż problem heterogeniczności struktury urządzeń IoT był już przedmiotem wielu badań [253–257].

Poza statycznym odwzorowaniem budowy urządzeń, w protokole Matter zdefiniowany jest również model interakcji (z ang. *Interaction Model*), odpowiadający

za komunikację między węzłami. Kluczowe elementy tego modelu to: wspomniane już atrybuty, zdarzenia (z ang. *Events*) oraz komendy (z ang. *Commands*). Komendy umożliwiają wywołanie określonych funkcji urządzenia, a ich wykonanie może generować odpowiedzi zwrotne. Atrybuty odwzorowują aktualny stan klastra, natomiast zdarzenia przechowują historię zmian, co pozwala na analizę przebiegu działań i ich skutków w czasie. Dzięki takiemu modelowi interakcji, węzły mogą m.in. odczytywać i zapisywać atrybuty, subskrybować zdarzenia oraz inicjować komendy. Efektywnie pozwala to na realizację właściwie dowolnej logiki funkcjonalnej urządzenia.

Aby zapewnić bezpieczne i spójne dodawanie nowych urządzeń do istniejącej sieci, Matter wspiera proces tzw. komisjonowania (z ang. *Commissioning*), polegający na przypisaniu poświadczeń bezpieczeństwa i włączeniu urządzenia do wspólnej domeny komunikacyjnej. Grupa urządzeń, które przeszły proces komisjonowania i dzielą ten sam kontekst bezpieczeństwa, tworzy tzw. *Fabric*. Każda taka sieć posiada unikalny, 64-bitowy identyfikator, który umożliwia bezpieczną i wzajemnie autoryzowaną komunikację pomiędzy węzłami, a także wspólne źródło zaufania (z ang. *Root of Trust*). Protokół wspiera pełne szyfrowanie i identyfikację urządzeń, tworząc tym samym bezpieczną i odporną na ataki strukturę sieciową. Niemniej, kwestie bezpieczeństwa protokołu Matter poruszono już m.in. w pracach dotyczących uwierzytelniania [258] oraz potencjalnych ataków zagłuszających [259].

Rozwiązania zgodne z Matter mogą być rozwijane na wielu popularnych platformach sprzętowych, takich jak *ESP32* (firmy Espressif), *Ameba D Series* (firmy Realtek), czy *EFR32* (firmy Silicon Labs). Co więcej, zgodność z wiodącymi asystentami głosowymi i rozwiązaniami sterowania automatyką domową, takimi jak *Google Home*, *Apple HomeKit* czy *Amazon Alexa*, czyni z Matter atrakcyjne rozwiązanie dla twórców systemów wbudowanych i aplikacji IoT. W literaturze odnaleziono już przykład integracji protokołu Matter z innym ekosystemem, *ECHO-NET Lite* [260, 261].

Pomimo dynamicznego rozwoju, protokół Matter i jego wykorzystanie w systemach lokalizacji wewnętrznej pozostają słabo opisane w literaturze. Jego elastyczna architektura danych, wspieranie interakcji urządzeń oraz zorientowanie na bezpieczeństwo czynią go obiecującą technologią również i w tym zastosowaniu, a w kontekście prowadzonych badań Matter może stanowić dobry łącznik

między różnymi technologiami w ramach zunifikowanej architektury komunikacyjnej. Jego zastosowanie, stanowiące nowatorskie podejście do zagadnienia, zasługuje na dalszą analizę, jako że może przyczynić się do rozwoju powszechnie dostępnych, tanich i uniwersalnych systemów lokalizacji radiowej.

3.2 Elementy autorskie w zakresie pozycjonowania radiowego

Niniejszy rozdział poświęcony jest omówieniu przeprowadzonych badań własnych oraz autorskich fragmentów prototypu systemu radiowego, stanowiącego pierwszą część końcowego systemu fuzyjnego. Na początku przedstawione zostaje podejście do stworzenia cyfrowego modelu środowiska testowego, sposób jego digitalizacji oraz uzyskany dla pierwszego prototypu model trójwymiarowy. Następnie skrótowo przedstawiony zostaje opracowany koordynator wymiany danych oparty na protokole Matter.

3.2.1 Digitalizacja stanowiska badawczego

Jednym z zagadnień pojawiających się w różnych kontekstach niniejszej pracy jest sposób pozyskiwania danych referencyjnych (z ang. *Ground Truth*, GT). Dotyczy to tak oznaczania zbiorów zdjęć, jak i danych niezbędnych do oceny jakości działania systemów lokalizacji. Każdy eksperyment, którego celem jest porównanie wartości estymowanych z rzeczywistymi, wymaga precyzyjnego zdefiniowania wartości oczekiwanej. W przypadku estymacji odległości przy użyciu technologii radiowych, a także w docelowych systemach pozycjonowania, często wystarczające są pomiary wykonane za pomocą dalmierza laserowego lub taśmy mierniczej. Urządzenia te są powszechnie uważane za satysfakcjonująco precyzyjne, a ewentualne błędy wynikają głównie z czynnika ludzkiego.

Wraz ze wzrostem złożoności środowiska pomiarowego, rośnie jednak także stopień trudności w uzyskaniu takich danych. W realiach otoczenia domowego, nawet specjalnie posprzątany i przygotowany obszar testowy pozostaje w pełni umeblowany, co niejednokrotnie uniemożliwia wykonanie bezpośredniego pomiaru odległości między punktem odniesienia a urządzeniem śledzącym. W tradycyjnych

metodach, w takich przypadkach konieczne jest wykonanie dodatkowych pomiarów pośrednich, co nie tylko wydłuża czas eksperymentu, ale także zwiększa ryzyko popełnienia błędu. Zaproponowano zatem wykonanie i wykorzystanie cyfrowego modelu środowiska badawczego. Jego wizualizacja, po wczytaniu do programu Blender w wersji 4.2, prezentowana jest na Rys. 3.6.

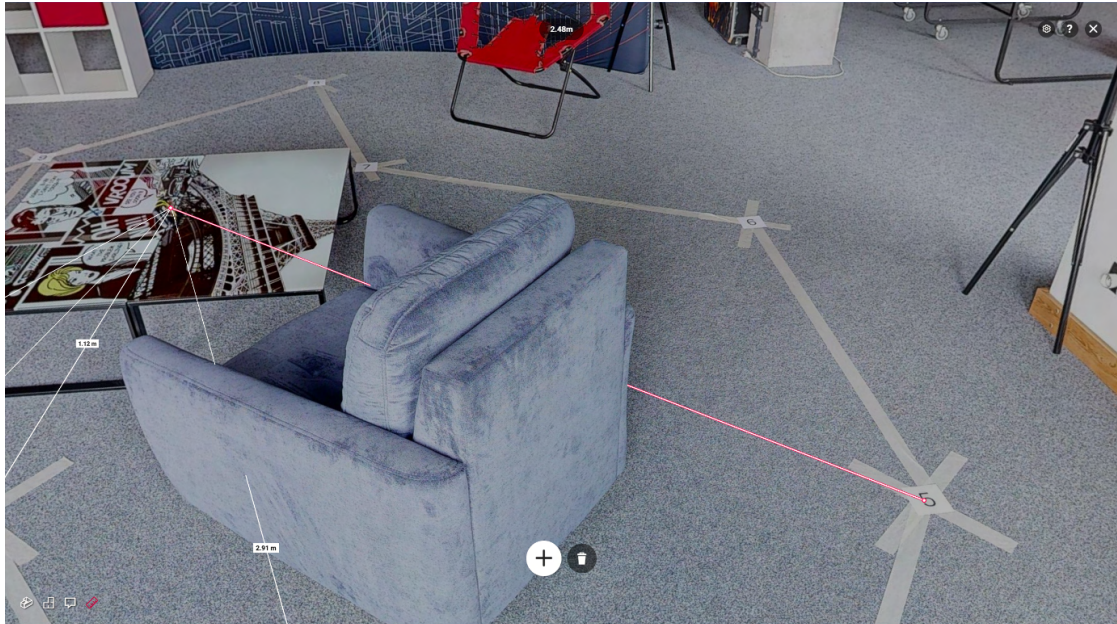


Rysunek 3.6: Zrzut ekranu z narzędzia Blender 4.2 z wizualizacją uzyskanego modelu cyfrowego.

Efektywnie, dzięki zastosowaniu technologii przedsiębiorstwa partnera doktora, uzyskano pełen model trójwymiarowy stanowiska pomiarowego. Pozwoliło to na całkowite wyeliminowanie wspomnianych problemów i wyznaczanie odległości pomiędzy dowolnie wybranymi punktami przestrzeni. Przykład takiej sytuacji, w której tradycyjny pomiar byłby utrudniony ze względu na przeszkodę (fotel zasłaniający bezpośrednią linię pomiaru), przedstawiono na Rys. 3.7.

Dalsze szczegóły opracowanego modelu, a także realia jego faktycznego wykorzystania w przeprowadzonych badaniach opisane zostają w Roz. 3.3.2. W analogiczny sposób opracowane zostały cyfrowe modele nieruchomości, tak dla prototypu systemu radiowego (dla środowiska o skali jednego pomieszczenia), jak i w do-

celowym systemie fuzyjnym (jako model większej przestrzeni, tj. kilkupokojowego mieszkania).



Rysunek 3.7: Zrzut ekranu z narzędzia spaceru interaktywnego firmy Matterport z wizualizacją wykorzystania funkcji pomiaru odległości.

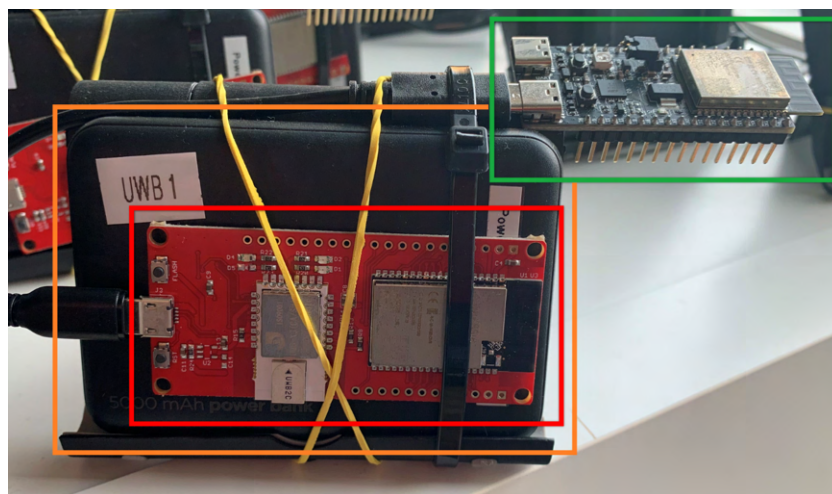
3.2.2 Skonstruowane urządzenia

Druga z dodatkowych hipotez niniejszej pracy dotyczy opracowania systemu lokalizacji, który oparty może zostać o powszechnie dostępne moduły elektroniczne. W ramach prowadzonych prac zaprojektowano więc, wykonano i przetestowano pierwsze prototypy urządzeń dedykowanych dwóm technologiom radiowym: UWB i BLE. Wykorzystano je następnie w procesie estymacji dystansu pomiędzy punktami o znanym położeniu (statycznymi punktami referencyjnymi zwanymi kotwicami, z ang. *Anchors*), a mobilnymi lokalizatorami (punktami o zmiennym położeniu, z ang. *Trackers*). Dla oceny zdolności koordynacji wymiany danych pomiędzy heterogenicznymi systemami IoT, zaproponowane zostało również urządzenie pośredniczące, implementujące w swoim działaniu protokół Matter. Wszystkie te elementy, jako pewnie niezależne elementy systemu, opisane zostają dalej w tym rozdziale. Ich połączenie i zastosowanie w badaniach omówione zostaje w Roz. 3.3.

Podsystem UWB

Wykorzystany podsystem oparty na technologii Ultra-Wideband (UWB) stanowi rozszerzenie istniejącego rozwiązania lokalizacyjnego, wcześniej opracowanego i przetestowanego w ramach pracy inżynierskiej, studiów w modelu *Project Based Learning*, a także dalszych prowadzonych badań [262, 263]. W obecnej pracy urządzenie zostało rozbudowane, a jego działanie dostosowano do współpracy z pozostałymi komponentami systemu oraz protokołem Matter.

Oba rodzaje urządzeń, tak kotwice, jak i lokalizatory, zbudowano na bazie gotowej płytki rozwojowej *Makerfabs ESP32 UWB* [264]. Platforma ta integruje mikrokontroler *ESP32* z układem komunikacyjnym UWB (*Decawave DWM3000*) i zapewnia dodatkowo obsługę Wi-Fi oraz BLE. Zasilanie zapewniające było z banku energii przez standardowy port *Micro-USB*. W zależności od przypisanej roli, urządzenia odpowiednio programowano do realizacji swoich zadań. W przypadku kotwic moduły UWB montowano fizycznie razem z modułami BLE (dla oszczędności miejsca). W przypadku lokalizatorów łączono je natomiast z dedykowaną płytką koordynatora Matter. Całość, jako jedynie prototyp, przymocowano na plastikowym uchwycie na telefon. Przykładowy zmontowany lokalizator podsystemu UWB zaprezentowano na Rys. 3.8.



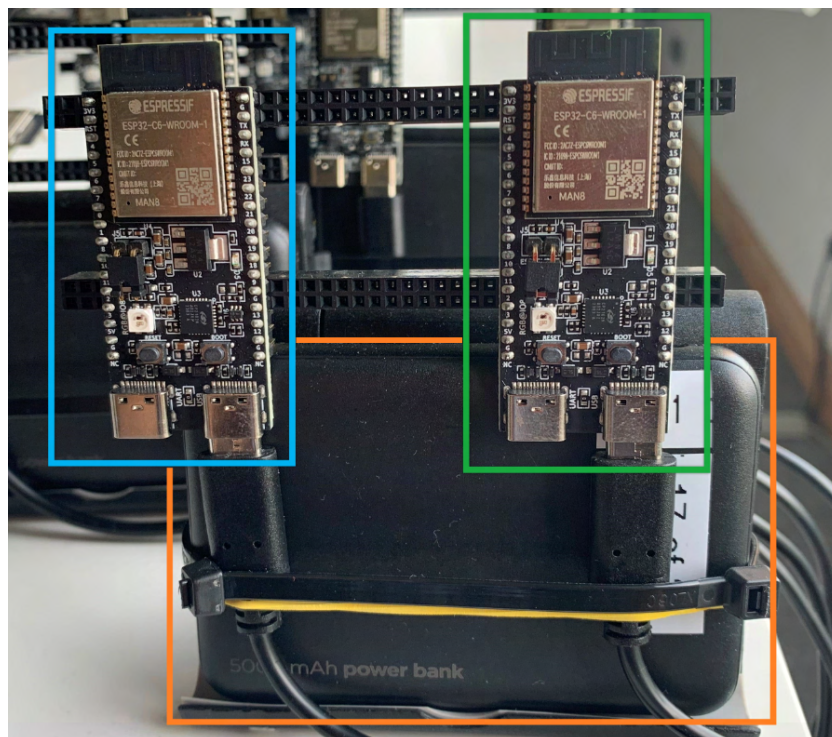
Rysunek 3.8: Zbudowane urządzenie lokalizacyjne w technologii UWB: płytka UWB (ramka czerwona), koordynator Matter (ramka zielona), zasilanie (ramka pomarańczowa).

W opracowanym systemie zastosowano opisaną już w Roz. 3.1 metodę radiowego wyznaczania dystansu, tj. Double-Sided Two-Way Ranging. Zaimplementowano podejście, w którym to kotwica inicjuje komunikację (wysyła pierwsze zapytanie pomiarowe), a następnie analizuje otrzymywane odpowiedzi. Lokalizatory pełnią natomiast rolę odbiorników, nasłuchujących i odpowiadających na wiadomości. W ten sposób zwiększa się poziom prywatności, jako że końcowa informacja o dystansie obliczana jest po stronie urządzenia śledzonego, a nie punktów referencyjnych. Aby przekazać tę informację z powrotem do kotwicy, potrzebna jest jawna, dodatkowa wymiana wiadomości. W celu zachowania poufności wyznaczonych odległości wystarczy z niej zrezygnować, a system referencyjny będzie posiadał pomiar z jedynie połowy całego procesu (tylko z pierwszego cyklu komunikacji). Początkowa transmisja odbywa się w trybie rozgłoszeniowym (z ang. *Broadcast*), a każda kotwica może odpowiedzieć tylko raz w ramach pojedynczej sekwencji wymiany wiadomości. Runda pomiarów kończy się dopiero po zebraniu odpowiedzi od wszystkich dostępnych w pomieszczeniu kotwic lub po upływie określonego czasu oczekiwania (z ang. *Time-out*). Jeśli odpowiedź pochodzi od kotwicy, która została już uwzględniona w bieżącej rundzie pomiarowej, system resetuje bieżącą sekwencję i rozpoczyna kolejną. Dodatkowo zastosowano mechanizm czasowej blokady odpowiedzi ze strony kotwic po każdym udanym pomiarze. Zapobiega to przeciążeniom komunikacyjnym i zwiększa stabilność całego systemu. Zebrane pomiary mogą być dalej przetwarzane: filtrowane, uśredniane, czy też mogą posłużyć do obliczenia pozycji w trójwymiarze na podstawie algorytmu trilateracji. Pomiary odległości mogą pełnić rolę nośnika dodatkowych informacji, takich jak identyfikacja urządzeń czy binarna detekcja obecności użytkownika w zasięgu wybranego fragmentu systemu.

Podsystem BLE

Analogicznie do urządzeń UWB, lokalizatory i kotwice systemu BLE zbudowano z użyciem gotowych płytek deweloperskich. W szczególności wykorzystano płytkę *ESP32-C6-DevKitC-1* firmy *Espressif* [265]. Obsługa wielu protokołów komunikacyjnych, w tym BLE, pozwoliła na szybką i efektywną implementację algorytmu estymacji odległości. Tak jak w przypadku podsystemu UWB, przygoto-

wano dwa warianty oprogramowania: pierwszy przekształcający płytkę w urządzenie referencyjne (kotwicę), a drugi konfiguruje ją jako lokalizator. W przypadku lokalizatorów, w jedno urządzenie fizyczne łączony był moduł BLE oraz płytka koordynatora protokołu Matter. Dla kotwic moduł BLE łączony był z modulem kotwicy UWB, aby zaoszczędzić miejsce w przestrzeni. Przykładowe zbudowane urządzenie lokalizatora w technologii BLE zaprezentowano na Rys. 3.9.



Rysunek 3.9: Zbudowane urządzenie lokalizacyjne w technologii BLE: płytka BLE (ramka niebieska), koordynator Matter (ramka zielona), zasilanie (ramka pomarańczowa).

Zastosowana metoda lokalizacji oparta została na analizie wskaźnika siły sygnału, RSSI, zgodnie z podejściem opisanym w Roz. 3.1.1. Kalibrację prowadzono indywidualnie dla każdego zmontowanego urządzenia. W pierwszym kroku lokalizator umieszczano w odległości jednego metra od kotwicy, zbierając po 25 odczytów dla trzech różnych kotwic. Uśredniona wartość ze wszystkich 75 pomiarów stanowiła końcową wartość $RSSI_{ref}$. Drugi etap kalibracji polegał na wyznaczeniu wartości n_{env} , poprzez rozwiązanie wzoru (3.6) dla odległości trzech metrów

i wcześniej wyznaczonego RSSI_{ref} . Również w tym przypadku zbierano 75 wartości, a następnie je uśredniano. Wyniki kalibracji przedstawiono w Tab. 3.1.

Tabela 3.1: Wartości kalibracji i odchylenie standardowe dla zmontowanych lokalizatorów technologii BLE.

Lokalizator	Skalibrowane wartości			
	RSSI_{ref} [dBm]	Std.	n_{env}	Std.
1	-40	6,67	3,41	0,98
2	-37	7,57	4,21	2,11
3	-45	5,27	1,16	0,87
4	-51	6,15	1,45	0,68

Po zakończeniu kalibracji i skonfigurowaniu urządzeń, kotwice uruchomiono jako nadajniki rozgłaszające (z ang. *Advertisers*), które wysyłały sygnały z ustaloną częstotliwością 20 milisekund. Algorytm ich działania przedstawiono w Alg. 1. Z uwagi na przejrzystość, pominięto w nim szczegóły implementacji procedur połączeniowych, raportowania statusu czy prób przywrócenia połączenia. Po uruchomieniu, celem lokalnego podglądu komunikacji, ułatwienia usuwania błędów oraz nadzoru nad działaniem urządzenia, mikrokontroler kotwicy inicjował diodę LED oraz komunikację szeregową (linie 10-11). Odpowiednie kolory diody informowały o wykonywanej operacji, a na porcie szeregowym nadawano wiadomości diagnostyczne. Następnie aktywacji podlegał moduł BLE, odpowiedzialny za właściwy proces rozgłaszania wiadomości i biorący udział w estymacji odległości (linia 12). W jego ramach diodę przestawiano na kolor niebieski, a także konfigurowano i uruchamiano nadawanie (linie 2-7). Konfigurację kończyło połączenie z lokalną siecią Wi-Fi oraz brokerem protokołu wymiany wiadomości MQTT (z ang. *Message Queuing Telemetry Transport*), do którego przesyłano adres MAC (z ang. *Media Access Control*) urządzenia BLE, pozwalający na przypisanie go do określonego pomieszczenia (linie 13-15). W ramach głównej pętli programu podtrzymywano działanie urządzenia (a zatem i jego funkcje komunikacji poprzez BLE) oraz połączenie z brokerem MQTT (linie 18-21). Dla celów diagnostycznych kolor diody ustawiano na zielony, a także cyklicznie publikowano wiadomość o poprawnym działaniu urządzenia (sygnał “alive”, linie 22-25).

Urządzenia lokalizatorów odpowiedzialne były za właściwą estymację odległo-

ści. Algorytm ich działania przedstawiono w Alg. 2. Urządzenia te pracowały jako aktywne skanery, które dla możliwie szybkiego wykrywania otaczających ich kotwic skonfigurowano do ciągłego nasłuchu wiadomości rozgłoszeniowych BLE z interwałem skanowania 110 milisekund i oknem 100 milisekund. Wstępna konfiguracja urządzenia lokalizatora była analogiczna do tej, przedstawionej dla kotwicy: inicjalizowano diodę i port szeregowy (linie 7-8), uruchamiano i konfigurowano moduł BLE (linia 9 oraz 1-5), a także łączono się z WiFi oraz brokerem MQTT (linie 10-11). Jediną różnicą była dodatkowa inicjalizacja tablicy znanych do komunikacji kotwic i zgłoszenie gotowości urządzenia (linie 12-13). W głównej pętli programu wyszczególnić możemy trzy bloki kodu. Pierwszy odpowiedzialny był za podtrzymanie komunikacji poprzez MQTT (linie 16-19). W drugim, pod warunkiem upłynięcia określonego interwału (linia 21), realizowano właściwą estymację dystansu. Ustawiano diodę na kolor zielony (linia 22), przez jedną sekundę skanowano w poszukiwaniu dostępnych kotwic (linia 23 oraz 33-42), a następnie wyznaczano dystanse do nich (linia 24 oraz 43-49). Blok kończyło zgłoszenie odczytów do serwera MQTT (przekazującego je dalej do urządzenia koordynatora protokołu Matter) oraz aktualizacja ostatniego znanego czasu przeprowadzonego skanowania (linie 25-26). Trzeci blok kodu odpowiadał natomiast za diagnostyczne monitorowanie działania urządzenia i cykliczne zgłaszanie jego statusu poprzez MQTT (linie 28-31).

Algorytm 1 Program urządzenia kotwicy w technologii BLE: inicjalizacja, konfiguracja i główna pętla programu.

Dane wejściowe: Dane dostępne do Wi-Fi, adres brokera MQTT, interwały nadawania BLE

Dane wyjściowe: Okresowe raportowanie statusu przez MQTT

```
1: procedure INITIALIZEBLE
2:   Ustaw kolor diody LED na niebieski
3:   Uruchom usługę BLE z identyfikatorem urządzenia
4:   Ustaw BLE w tryb wykrywalności
5:   Skonfiguruj dane nadawcze BLE i interwały
6:   Włącz odpowiedź na skanowanie
7:   Rozpocznij nadawanie
8: end procedure

9: procedure SETUP
10:  Zainicjalizuj wbudowaną diodę LED
11:  Zainicjalizuj komunikację przez port szeregowy
12:  Uruchom usługę BLE i ustaw parametry nadawania
13:  Połącz z siecią Wi-Fi
14:  Połącz z brokerem MQTT
15:  Wyślij adres MAC BLE do serwera
16: end procedure

17: procedure MAINLOOP
18:   if klient MQTT jest rozłączony then
19:     Ponownie połącz z brokerem MQTT
20:   end if
21:   Utrzymuj połączenie MQTT
22:   if upłynął interwał alive then
23:     Zmień kolor LED na zielony
24:     Wyślij wiadomość alive przez MQTT
25:   end if
26: end procedure
```

Algorytm 2 Program urządzenia lokalizatora w technologii BLE: inicjalizacja, konfiguracja i główna pętla programu.

Dane wejściowe: Dane dostępne do Wi-Fi, adres brokera MQTT, adresy MAC znanych kotwic BLE, moc nadawcza, wykładnik strat propagacyjnych

Dane wyjściowe: Okresowe odczyty kotwic i raporty statusu przez MQTT

```

1: procedure INITIALIZEBLE
2:   Ustaw kolor diody LED na niebieski
3:   Uruchom usługę BLE z identyfikatorem lokalizatora
4:   Skonfiguruj parametry skanowania BLE (okno i interwał)
5: end procedure

6: procedure SETUP
7:   Zainicjalizuj wbudowaną diodę LED
8:   Zainicjalizuj komunikację przez port szeregowy
9:   Rozpocznij skanowanie BLE z zadanymi parametrami
10:  Połącz z siecią Wi-Fi
11:  Połącz z brokerem MQTT
12:  Zainicjalizuj odczyty znanych kotwic
13:  Zgłoś zakończenie konfiguracji lokalizatora
14: end procedure

15: procedure MAINLOOP
16:   if klient MQTT jest rozłączony then
17:     Ponownie połącz z brokerem MQTT
18:   end if
19:   Utrzymuj połączenie MQTT
20:   Bieżący czas  $\leftarrow$  aktualny czas systemowy w milisekundach
21:   if Bieżący czas – czas ostatniego skanu  $\geq$  interwał skanu then
22:     Zmień kolor LED na zielony
23:     Przeprowadź skan BLE w poszukiwaniu kotwic
24:     Zaktualizuj wartości RSSI kotwic i oblicz odległości
25:     Zgłoś odczyty kotwic przez MQTT
26:     Zaktualizuj czas ostatniego skanu
27:   end if
28:   if Bieżący czas – czas ostatniego statusu  $\geq$  interwał alive then
29:     Zgłoś status alive przez MQTT
30:     Zaktualizuj czas ostatniego statusu
31:   end if
32: end procedure

33: procedure SCANFORANCHORS
34:   Rozpocznij skan BLE przez zdefiniowany czas trwania
35:   for all odkryte urządzenie do
36:     if urządzenie jest znaną kotwicą then
37:       Zapisz wartość RSSI danej kotwicy
38:       Oblicz odległość za pomocą modelu strat propagacyjnych
39:       Zaktualizuj odczyty danej kotwicy
40:     end if
41:   end for
42: end procedure

43: procedure CALCULATEDISTANCE( $rssi$ )
44:   Wzór: Odległość  $\leftarrow 10^{\left(\frac{RSSI_{ref}-RSSI}{10 \cdot n_{env}}\right)}$ 
45:   if RSSI = 0 lub równe mocy nadawczej then
46:     Zwróć odległość jako 0
47:   end if
48:   Zwróć obliczoną odległość
49: end procedure

```

Koordinator Matter

Wychodząc z założenia przemysłowych realiów wdrożenia opracowanych prototypów, postanowiono, że istniejąca platforma lokalizacyjna powinna zostać rozszerzona, a nie zastąpiona całkowicie. Dlatego do już komunikującego się z użyciem protokołu MQTT systemu dodano nowy, dedykowany protokołowi Matter element, tzw. koordinator Matter. Urządzenie to stanowi ogniwo pośredniczące, umożliwiające integrację dwóch sposobów komunikacji: istniejącego wcześniej i nadającego wiadomości poprzez MQTT oraz nowo wprowadzanego, udostępniającego pomiary jako węzeł protokołu Matter. W formie platformy sprzętowej wykorzystano ponownie płytkę *ESP32-C6-DevKitC-1*, czyli identyczną jak w przypadku urządzeń BLE, jednak z zupełnie odmiennym oprogramowaniem. Fizycznie, płytki układów koordynatora umieszczone zostały razem z układami lokalizatorów pozostałych protokołów. Prezentowano to już na Rys. 3.8 i Rys. 3.9, każdorazowo oznaczając płytki koordynatora Matter zieloną ramką.

Napisane oprogramowanie bazowało na oficjalnym repozytorium kodu źródłowego protokołu Matter [266]. Projekt skompilowano od podstaw wraz ze wszystkimi zależnościami, systemem budowania rozwiązań i implementacjami przykładowych urządzeń IoT. W oparciu o oficjalny model danych, zdefiniowano strukturę własnego urządzenia, której celem było możliwie jak najwierniejsze odzwierciedlenie realnych przypadków zastosowań. Zaproponowane urządzenie protokołu Matter rozwijało udostępniony przykład systemu klimatyzacji. Uznano je za dobrze dopasowane do kontekstu, jako popularnego urządzenia domowego, najczęściej statycznie umiejscowionego w przestrzeni.

Z perspektywy wspomnianego modelu danych, przechodząc do szczegółów implementacyjnych, stworzono jedno urządzenie protokołu Matter, złożone z pojedynczego węzła i zawierającego jeden punkt dostępu. Konfiguracja węzła pozostała domyślna, bez dodatkowych rozszerzeń. Zdefiniowanemu punktowi dostępu nadano typ `room_air_conditioner`, z jednym ustawionym ręcznie parametrem, tj. wartością domyślnej mocy urządzenia. Kluczową częścią konfiguracji było wykorzystanie klastra `esp_matter::cluster::diagnostic_logs`, który umożliwia zapis i odczyt otrzymywanych komunikatów, pierwotnie wykorzystywanych do analizy działania urządzeń.

Chociaż standard Matter oferuje wiele gotowych do użycia klastrów i atrybutów, nie znaleziono żadnego natywnie obsługującego funkcjonalność związaną z pozycjonowaniem lub odległością. Rozważano ich modelowanie jako temperatury, jednak ostatecznie uznano, że takie użycie mogłoby prowadzić do nieporozumień oraz niepoprawnej semantyki. Zamiast tego zaimplementowano własny mechanizm obsługi rejestru zdarzeń, definiując własny oczekiwany rozmiar bufora komunikatów oraz metody: zapisu, odczytu i czyszczenia wiadomości. Umożliwiło to przechowywanie danych odebranych przez protokół MQTT w postaci wpisów rejestru, udostępnianych następnie przez protokół Matter.

W praktyce, koordynator nasłuchiwał na komunikację po protokole MQTT, którym dane o oszacowanych odległościach przesyłane były przez urządzenia śledzące (lokalizatory). Każdy odebrany komunikat zapisywano w buforze cyklicznym o stałej wielkości, mieszczącym dokładnie pięć najnowszych odczytów. Rozmiar ten dobrano eksperymentalnie, zgodnie z ograniczeniami sprzętowymi oraz z maksymalną dopuszczalną długością wiadomości odczytywanych przez narzędzie CHIP-TOOL (szczegóły w Roz. 3.3.3).

Dzięki opisanej architekturze, zbudowany koordynator spełniał wszystkie zakładane wymagania:

- zapewniał zgodność z istniejącym systemem wykorzystującym MQTT (bez konieczności jego modyfikacji),
- umożliwiał udostępnienie tych samych danych za pośrednictwem drugiego, bardziej nowoczesnego protokołu,
- stanowił faktyczne urządzenie Matter, w pełni zgodne ze specyfikacją i gotowe do integracji z systemami domowymi nowej generacji.

3.3 Prototyp systemu lokalizacji radiowej

Wykorzystując wszystkie opisane elementy autorskie, skonstruowano pierwszy prototyp radiowego systemu estymacji odległości. Jego celem było opracowanie i przetestowanie kompletnego środowiska lokalizacyjnego w warunkach mieszkalnych, z użyciem tanich i ogólnodostępnych komponentów IoT. Aby to osiągnąć:

- zrealizowano system estymacji odległości, oparty na technologiach BLE i UWB,

- połączony za pomocą warstwy komunikacji wykorzystującej protokół Matter,
- stworzono cyfrowy model rzeczywistego środowiska testowego,
 - porównano jakość estymacji odległości przy różnych źródłach danych referencyjnych (pomiar manualny, a model cyfrowy),
 - opracowano zbiór danych zawierający 264 serie pomiarowe z 17 zestawów eksperymentów.

W niniejszym rozdziale przedstawiono szczegóły jego realizacji i wykonania, przeprowadzone eksperymenty oraz uzyskane wyniki. Na podstawie tak opracowanych danych możliwe stało się dalsze projektowanie systemu pozycjonowania trójwymiarowego, które znalazło zastosowanie w końcowym systemie fuzyjnym. Prezentowany prototyp łączył obszary integracji komunikacji z użyciem wielu protokołów, pozwalał na ich porównanie, a przede wszystkim operował w ramach stworzonego cyfrowego modelu całego stanowiska badawczego. Zastosowanie modelu cyfrowego wyeliminowało natomiast potrzebę ręcznego pozyskiwania pomiarów odniesienia, pozwoliło na ponowną analizę danych nawet po zakończeniu eksperymentu oraz umożliwiło zwiększenie powtarzalności testów.

3.3.1 Fizyczna budowa systemu

Zaprojektowany system lokalizacji składał się łącznie z 24 fizycznych modułów komunikacyjnych, zorganizowanych w cztery grupy funkcjonalne:

- 4 mobilne lokalizatory UWB,
- 4 mobilne lokalizatory BLE,
- 8 statycznych kotwic UWB,
- 8 statycznych kotwic BLE.

Lokalizatory odwzorowywały urządzenia śledzące, których położenie chcieliśmy wyznaczyć, a kotwice pełniły rolę punktów referencyjnych o znanym umiejscowieniu. Wszystkie urządzenia zostały samodzielnie zmontowane i skonfigurowane, a następnie rozlokowane na stanowisku badawczym, które stanowiła umeblowana przestrzeń mieszkalna. Zdjęcia przedstawiające rzeczywisty wygląd urządzeń wykorzystanych w eksperymentach zaprezentowano na Rys. 3.10.



(a) Zbudowane urządzenia lokalizacyjne obu technologii radiowych: BLE (po lewej) i UWB (po prawej).

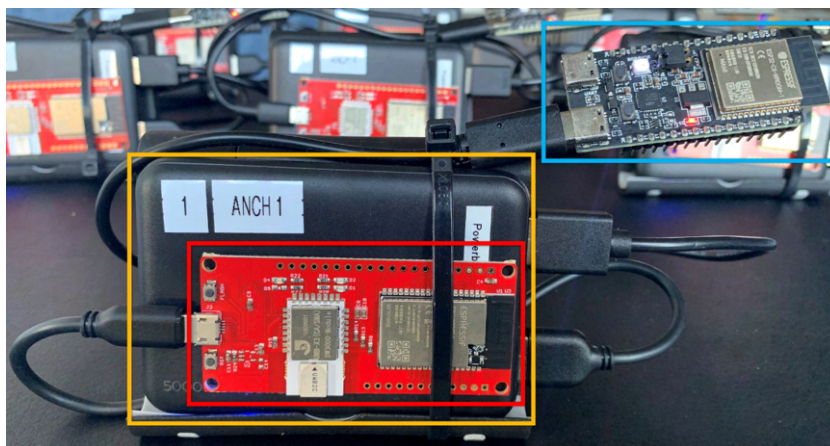


(b) Zbudowane urządzenia kotwic, łączące komponenty obu technologii: UWB (płytki czerwone) i BLE (płytki czarne).

Rysunek 3.10: Zdjęcie wszystkich skonstruowanych prototypowo urządzeń systemu pozycjonowania.

Na Rys. 3.10a przedstawiono lokalizatory. Po lewej stronie znajdują się urządzenia BLE, a po prawej UWB. Każdy z tych znaczników został zaprogramowany tak, aby realizować zadania związane z pomiarem odległości oraz transmisją danych do odpowiednich odbiorników, w zależności od zastosowanej technologii radiowej. Kotwice zostały przedstawione na Rys. 3.10b. Płytki koloru czerwonego to moduły, które pracowały w technologii UWB, natomiast czarne odpowiadały za komunikację BLE. Równomierne rozmieszczenie kotwic w całym środowisku testowym umożliwiło uzyskanie stabilnych i porównywalnych pomiarów niezależnie od lokalizacji znacznika. Jak opisano w Roz. 3.2.2, aby zoptymalizować wykorzystanie przestrzeni oraz zasobów sprzętowych, kotwice systemu zostały fizycznie zmontowane jako urządzenia hybrydowe, łączące na jednym uchwycie moduły komunikacyjne obu technologii. Zamiast budować szesnaście osobnych urządzeń: po osiem dla UWB i BLE, skonstruowano osiem samodzielnych, fizycznych kotwic, z których każda zawierała jednocześnie moduł BLE i moduł UWB. Budowa pojedynczego takiego urządzenia prezentowana jest na Rys. 3.11.

Rozwiązanie to pozwoliło nie tylko ograniczyć liczbę elementów rozmieszczonych w przestrzeni eksperymentalnej, ale również uprościć proces instalacji i konfiguracji systemu, zachowując pełną funkcjonalność obu podsystemów. Takie podejście umożliwiło również równoległe prowadzenie eksperymentów z wykorzystaniem



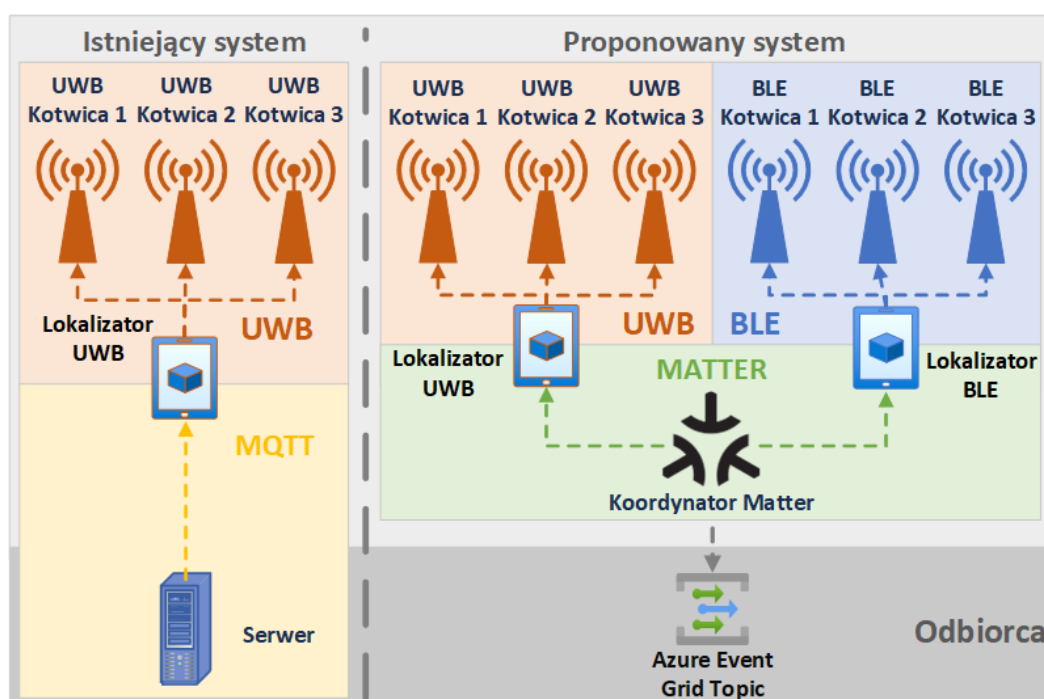
Rysunek 3.11: Zbliżenie na skonstruowane urządzenie kotwicy (punktu referencyjnego). Uchwyt mocujący i bank energii oznaczone pomarańczową ramką, kotwica systemu UWB czerwoną ramką, a kotwica systemu BLE niebieską ramką.

różnych technologii radiowych w tej samej lokalizacji, przy zachowaniu spójności geometrycznej rozmieszczenia punktów referencyjnych.

Aby dopasować skonstruowany prototyp do potencjalnej ścieżki wdrożenia w produktach partnera doktoratu, rozbudowano jego funkcje o raportowanie wyznaczonych odległości nie tylko do lokalnego serwera, ale również na serwery chmurowe. Jako platformę docelową wybrano rozwiązanie firmy Microsoft w postaci usługi obsługi zdarzeń *Azure Event Grid*. Przygotowano konkretny interfejs komunikacji, dostępny swobodnie z Internetu, do którego przesyłano odczyty, efektywnie uzyskując efekt agregacji pomiarów. Umożliwiono w ten sposób dalsze przetwarzanie pomiarów oraz przekierowywanie ich do innych usług. Rozszerzoną architekturę systemu oraz jej porównanie do już istniejącej zaprezentowano na Rys. 3.12.

3.3.2 Środowisko testowe i warunki eksperymentalne

W celu jak najdokładniejszego odwzorowania warunków domowych, przestrzeń testowa została wydzielona w budynku biurowym z cegły i zaadaptowana tak, by przypominała wnętrze mieszkalne. Umieszczono w niej m.in. sofę, trzy rodzaje krzeseł, regał, stolik kawowy, szafkę RTV oraz ściankę ekspozycyjną. Ponadto, jako że część przestrzeni przylegała do pozostałej części biura, obecne były również elementy typowe dla mieszkań, takie jak drzwi, okna, lampy czy klimatyzacja.



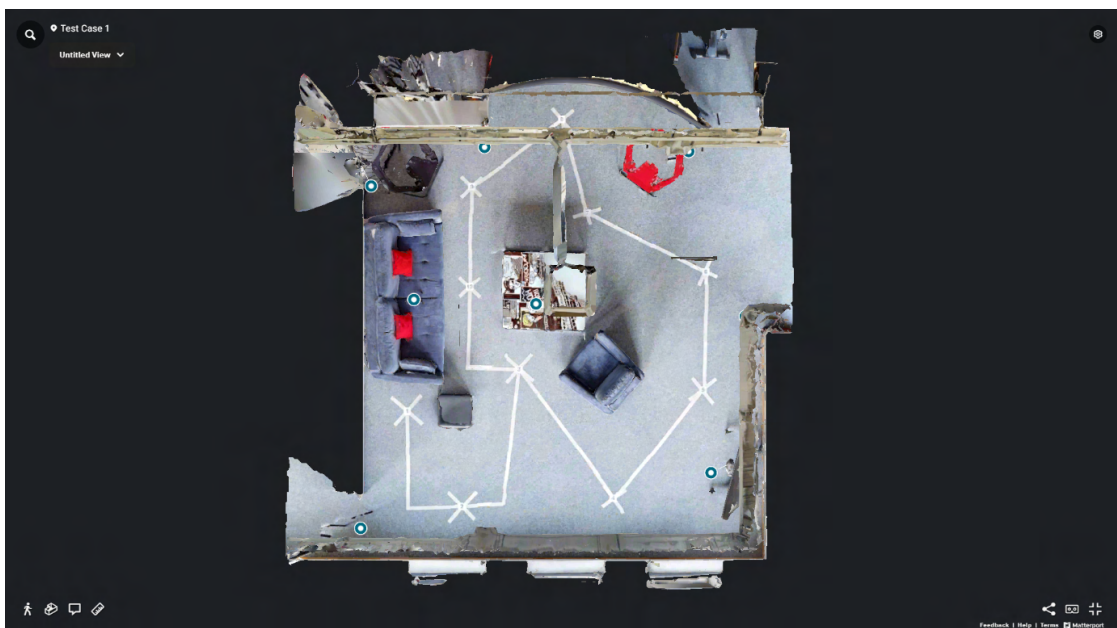
Rysunek 3.12: Diagram przedstawiający architekturę proponowanego systemu lokalizacyjnego. Rozwiązanie źródłowe prezentowane jest po lewej, a rozszerzona postać po prawej. Podsystem UWB oznaczono w odcieniach czerwonego, BLE w odcieniach niebieskiego, a koordynator protokołu Matter oznaczono odcieniami zielonego. Oryginalny odbiornik danych (serwer) zastąpiono usługą chmurową Azure Event Grid (szary obszar).

W celu precyzyjnego odwzorowania tej przestrzeni wykonano jej cyfrową reprezentację z wykorzystaniem technologii digitalizacji przestrzeni firmy *Matterport*. Zastosowanie skanera *Matterport Pro3*, wykorzystującego zarówno zdjęcia panoramiczne, jak i pomiary laserowe, pozwoliło na uzyskanie wysokiej jakości modelu cyfrowego. Zrzut ekranu z widoku spaceru interaktywnego przedstawiono na Rys. 3.13.

Dzięki przygotowanemu modelowi możliwe było wyodrębnienie widoku z góry (Rys. 3.14). W ten sposób określono dokładny plan pomieszczenia oraz rozmieszczenie dziesięciu punktów pomiarowych, oznaczonych na podłodze, ponumerowanych i połączonych ustaloną ścieżką, przewidzianą do dalszych badań z zakresu lokalizacji dynamicznej i śledzenia ruchu w przestrzeni trójwymiarowej.



Rysunek 3.13: Fizyczne środowisko, w którym przeprowadzono eksperyment.



Rysunek 3.14: Wizualizacja rzutu pomieszczenia wraz z punktami pomiarowymi, wygenerowana przy użyciu wirtualnej wycieczki Matterport.

Każdy z punktów pomiarowych był wykorzystywany w dwóch konfiguracjach: jako stacja przyziemna oraz jako stacja umieszczona na wysokości człowieka. W przypadku pierwszym, urządzenie lokalizacyjne umieszczano bezpośrednio na podłodze. W drugim, lokalizator instalowano na statywie, co podnosiło jego wysokość do 150 cm nad poziomem podłogi. Obie konfiguracje przedstawiono na Rys. 3.15.



Rysunek 3.15: Urządzenie śledzące umieszczone w punkcie pomiarowym P1 w konfiguracji na podłodze (po lewej) oraz na wysokości 150 cm (po prawej).

W celu uzyskania zróżnicowanych wysokości kotwic referencyjnych zastosowano dwa typy statywów: niski, zwiększający wysokość o kilkanaście centymetrów, oraz regulowany, pozwalający na montaż kotwic tuż pod sufitem. Oba typy przedstawiono na Rys. 3.16.

Choć samo oszacowanie odległości nie wymagało znajomości współrzędnych 3D wszystkich kotwic, to dzięki zastosowaniu modelu cyfrowego możliwe było ich łatwe określenie. W modelu 3D każdej kotwicy przypisano trójwymiarowe współrzędne (X_k, Y_k, Z_k) odpowiadające środkowi sfery o promieniu 15 cm, ręcznie umieszczonej tak, aby w pełni obejmowała urządzenie śledzące. Stopniowe nakładanie sfery na urządzenie przedstawiono na Rys. 3.17, a uzyskane współrzędne kotwic zestawiono w Tab. 3.2.



Rysunek 3.16: Urządzenie kotwicy zamontowane na statywie niskim (po lewej) oraz regulowanym wysokim (po prawej).



Rysunek 3.17: Wizualizacja obudowania lokalizatora sferą (z lewej: 100% nieprzezroczystości, środek: 50%, z prawej: 0%).

W celu analizy wpływu metody pozyskiwania danych referencyjnych na jakość wyników estymacji, w badaniach uwzględniono trzy niezależne ich źródła: ręczny pomiar dalmierzem laserowym, interaktywny pomiar z poziomu wirtualnego spaceru (tzw. *Matterport 3D Walk*) oraz pełny eksport utworzonego modelu przestrzeni do formatu *Wavefront OBJ*.

Pomiary manualne przeprowadzono z użyciem dalmierza laserowego *Uni-T LM50A*. Jest to urządzenie klasy 2 o zasięgu do 50 metrów i deklarowanej do-

Tabela 3.2: Pozycja kotwic w przestrzeni trójwymiarowej stanowiska pomiarowego.

Kotwica	X_k (cm)	Y_k (cm)	Z_k (cm)
1	-139,20	60,76	255,00
2	330,87	67,77	255,00
3	110,04	-469,94	218,00
4	372,73	-392,84	192,00
5	387,97	-100,83	150,00
6	-70,14	-427,68	88,00
7	154,26	-15,51	55,00
8	160,96	-179,49	35,00

kładności $\pm (2, \text{mm} + 5 \times 10^{-5} \times \text{Dystans})$. Tam, gdzie możliwe było wykonanie pomiaru bezpośredniego, rejestrowany był pojedynczy odczyt. W pozostałych przypadkach, w których przeszkody zasłaniały wiązkę lasera, obliczano wartość oczekiwaną na podstawie dwóch pomiarów pomocniczych z wykorzystaniem twierdzenia Pitagorasa. Ze względu na trudność w precyzyjnym ustawieniu dalmierza (dokładne umieszczenie promienia w pozycji odpowiadającej antenie urządzenia) oraz rzeczywiste wymiary anteny ($6 \text{ mm} \times 9 \text{ mm}$), zdecydowano się zrezygnować z dokładności milimetrowej i przyjęto za wystarczające zaokrąglenie do pełnych centymetrów.

Pomiary z użyciem wirtualnego spaceru były realizowane z poziomu przeglądarki internetowej. Platforma podglądu i obsługi utworzonych modeli firmy *Matterport*, a w szczególności samo narzędzie wizualizacji spacerów, pozwala na wyznaczanie pomiarów odległości w modelu trójwymiarowym. Pomimo swojej intuicyjności, narzędzie to ma pewne ograniczenia technologiczne, dotyczące precyzji wskazywania punktów pomiarowych. Dostępna w nim reprezentacja modelu posiada skończoną liczbę punktów umieszczonych na siatce przestrzennej, co ogranicza możliwość wskazania początku i końca pomiaru.

Dlatego też trzecie podejście polegało na eksporcie pełnego modelu trójwymiarowego przestrzeni do formatu *.obj* oraz dalszej jego analizie w programie *Blender*. W tym przypadku możliwe było już w pełni swobodne definiowanie i modelowanie punktów odniesienia, niezależnie od ich dostępności z poziomu wirtualnego spaceru. Pozwoliło to na dokładniejsze odwzorowanie położenia elementów systemu

względem siebie, a tym samym na uzyskanie najbardziej wiarygodnych danych odniesienia.

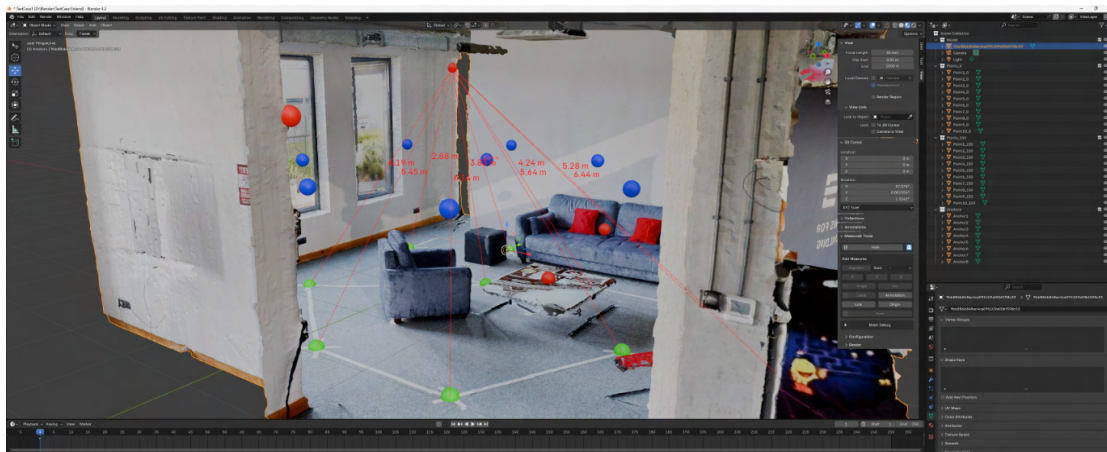
Wyniki eksperymentów, przedstawiane w kolejnych rozdziałach, uwzględniają odniesienie do każdego z tych trzech źródeł danych referencyjnych. Dzięki temu możliwa stała się wzajemna weryfikacja poprawności uzyskanych wartości i wykrycie błędów ludzkich, uproszczenie pomiarów w scenariuszach z przeszkodami, a także powtarzalna analiza środowiska testowego i planowanie dalszych badań z wyprzedzeniem.

Jako przykład wstępnej oceny spójności danych uzyskanych różnymi metodami, w Tab. 3.3 przedstawiono przykładowe porównanie odczytów wysokości zamocowania kotwic. Wartości wyrażone zostały w centymetrach i odnoszą się do wysokości nad poziomem podłogi. Różnice między metodami były minimalne, zazwyczaj nie większe niż 1 cm. Największa rozbieżność dotyczyła przypadku, w którym dostęp do punktu pomiarowego był zasłonięty przez mebel (kotwica zamontowana na telewizorze), co uniemożliwiało wykonanie dokładnego pomiaru metodą interaktywnego spaceru. W tym przypadku jedynie pomiary ręczne oraz model trójwymiarowy dawały rzetelnie porównywalne rezultaty.

Tabela 3.3: Wysokości kotwic ponad poziomem podłogi (w centymetrach), wyznaczone przy użyciu metody manualnego pomiaru dalmierzem (Laser), interaktywnego spaceru trójwymiarowego (Spacer) oraz odczytane wprost z modelu (Model).

Kotwica	Wysokość [cm]		
	Laser	Spacer	Model
1	255	257	255
2	255	256	255
3	218	217	218
4	192	193	192
5	150	151	150
6	88	92	88
7	55	-	55
8	35	34	35

Do wizualizacji końcowego rozmieszczenia wszystkich punktów pomiarowych i kotwic wykorzystano program *Blender*. Dzięki niemu możliwe było jednoczesne przedstawienie punktów przyziemnych, lokalizatorów umieszczonych na statywie oraz kotwic. Pełna wizualizacja tej konfiguracji prezentowana jest na Rys. 3.18.



Rysunek 3.18: Wizualizacja eksperymentalnego środowiska w programie Blender: lokalizatory na podłodze oznaczone na zielono, na statywie na niebiesko, a kotwice na czerwono.

3.3.3 Obsługa protokołu Matter

Eksperymentalna obsługa i komunikacja z urządzeniami protokołu Matter, ze względu na wbudowane mechanizmy bezpieczeństwa i bardziej złożony proces uwierzytelniania, wymagały użycia dedykowanego narzędzia. Wykorzystano w tym celu *CHIP-TOOL*. Jest to interfejs wiersza poleceń (z ang. *Command-line Interface*, CLI) wspierający konfigurowanie, komunikację i sterowanie urządzeniami IoT poprzez protokół Matter. Narzędzie to pozwalało m.in. na przeprowadzenie procesu włączania urządzenia do lokalnej sieci, wysyłanie zapytań do wybranych klastrów urządzenia (w tym diagnostycznych) oraz odbiór i zapis odpowiedzi tekstowych.

CHIP-TOOL był używany we wszystkich fazach eksperymentów, tak do konfiguracji wstępnej, jak i podczas właściwego zbierania danych. W trybie interaktywnym pozwalał na ręczne dodanie urządzenia do sieci oraz na jego konfigurację. W trybie jednego polecenia (z ang. *Single-command Mode*) umożliwiał odczyt danych w sposób w pełni zautomatyzowany. W momencie prowadzenia badań narzędzie to było w fazie rozwojowej oraz oznaczone jako eksperymentalne. Posiadało wiele ograniczeń, np. jednoczesne wsparcie tylko jednego urządzenia lub brak wsparcia dla wielowątkowości. Z użyciem dedykowanych skryptów języka *Python* możliwe było ominięcie tych trudności. W ramach każdego eksperymentu uruchamiano wielowątkowo osiem niezależnych instancji *CHIP-TOOL*, z których każda

była przypisana do konkretnego lokalizatora. Klient łączył się z urządzeniem, odczytywał dane z klastra `diagnostic_logs`, zapisywał wynik w wybranym katalogu i kończył działanie.

Mimo że narzędzie to nie było jeszcze przystosowane ani do użytku komercyjnego, ani masowego, umożliwiło skuteczne wdrożenie i przetestowanie mechanizmu transmisji danych za pomocą protokołu Matter. Zebrano w ten sposób pełen zestaw pomiarów odległości za pomocą dwóch niezależnych protokołów (zarówno przez MQTT, jak i przez Matter) oraz z wykorzystaniem jednego spójnego zestawu urządzeń.

3.3.4 Konfiguracja eksperymentów

Celem eksperymentów była ocena dokładności i funkcjonalności systemów lokalizacji w przestrzeni wewnętrznej. W tym celu zaplanowano szereg scenariuszy pomiarowych, różniących się pod względem użytej technologii radiowej (UWB, BLE), protokołu komunikacyjnego (MQTT, Matter), wysokości pomiarów, źródła danych referencyjnych oraz dynamiki ruchu. Środowisko eksperymentalne zostało przygotowane w sposób ułatwiający rozmieszczenie punktów odniesienia (kotwic), przy jednoczesnym zachowaniu cech zbliżonych do typowego układu mieszkania. Realizacja testów w warunkach odwzorowujących rzeczywiste środowisko domowe była istotna z punktu widzenia późniejszej ścieżki wdrożenia przemysłowego opracowanego rozwiązania.

W ramach badań przeprowadzono siedemnaście różnych eksperymentów, które podzielono na dwie główne grupy: statyczne oraz dynamiczne. Eksperymenty statyczne (numery: 1–3, 6–10, 12–16) polegały na analizie sygnałów z urządzenia lokalizacyjnego znajdującego się w ustalonym miejscu. Eksperymenty dynamiczne (numery: 4, 5, 11, 17) odwzorowywały scenariusze z przemieszczającym się obiektem. Każdy scenariusz został powtórzony trzykrotnie w celu uśrednienia wyników i ograniczenia wpływu przypadkowych zakłóceń. Zbiorcze zestawienie konfiguracji przedstawiono w Tab. 3.4.

W przypadku eksperymentów statycznych urządzenie śledzące umieszczano w jednym z wcześniej zdefiniowanych punktów pomiarowych. Przedstawiono i omówiono to już w Roz. 3.3.2. Pojedynczy pomiar trwał około jednej minuty. W ekspe-

rymentach dynamicznych urządzenie poruszało się po wcześniej określonej ścieżce, a czas potrzebny na rejestrację danych był znacząco dłuższy, tj. około pięciu minut. Dla zachowania zbliżonej wysokości względem scenariuszy statycznych urządzenie również umieszczano na statywie, tym razem postawionym na mobilnej platformie, efektywnie uzyskując wysokość około 140 cm. W zależności od scenariusza eksperymenty w obu technologiach radiowych były wykonywane równocześnie lub osobno. Przedstawiona w rozprawie analiza omawia jedynie wyniki eksperymentów statycznych, głównie ze względu na możliwość ich prostego i dokładnego porównania z wartościami oczekiwanymi.

Tabela 3.4: Konfiguracje i opis przeprowadzonych eksperymentów. Wykorzystana technologia radiowa (w kolumnie Tech.): UWB, BLE albo U+B (obie). Typ eksperymentu (w kolumnie Typ): statyczny (S) albo dynamiczny (D). Wysokość umieszczenia lokalizatora podana w centymetrach ponad poziomem podłogi.

#	Tech.	Typ	Wys.	Opis
E1	U+B	S	0	Lokalizatory 1–4 w punktach P1–P4: UWB1 i BLE1 w P1, UWB2 i BLE2 w P2, itd.
E2	U+B	S	0	Lokalizatory 1–4 w punktach P5–P8: UWB1 i BLE1 w P5, UWB2 i BLE2 w P6, itd.
E3	U+B	S	0	Lokalizatory 1–4 w punktach P9, P10, P1, P2: UWB1 i BLE1 w P9, UWB2 i BLE2 w P10, itd.
E4	UWB	D	15	Lokalizatory UWB 1–4 na platformie mobilnej. Pętla: od P1 do P10 - P3 - P2 - P1.
E5	BLE	D	15	Lokalizatory BLE 1–4 na platformie mobilnej. Pętla: od P1 do P10 - P3 - P2 - P1.
E6	UWB	S	150	Lokalizator UWB1 na trójkącie w punkcie P1.
E7	UWB	S	150	Lokalizator UWB1 na trójkącie w punkcie P3.
E8	UWB	S	150	Lokalizator UWB1 na trójkącie w punkcie P5.
E9	UWB	S	150	Lokalizator UWB1 na trójkącie w punkcie P7.
E10	UWB	S	150	Lokalizator UWB1 na trójkącie w P10.
E11	UWB	D	140	Lokalizator UWB1 na trójkącie na platformie mobilnej. Pętla: od P1 do P10 - P3 - P2 - P1.
E12	BLE	S	150	Lokalizator BLE1 na trójkącie w punkcie P1.
E13	BLE	S	150	Lokalizator BLE1 na trójkącie w punkcie P3.
E14	BLE	S	150	Lokalizator BLE1 na trójkącie w punkcie P5.
E15	BLE	S	150	Lokalizator BLE1 na trójkącie w punkcie P7.
E16	BLE	S	150	Lokalizator BLE1 na trójkącie w P10.
E17	BLE	D	140	Lokalizator BLE1 na trójkącie na platformie mobilnej. Pętla: od P1 do P10 - P3 - P2 - P1.

3.3.5 Wyniki

Zgromadzony materiał pomiarowy umożliwił ocenę skuteczności oraz precyzji działania zaproponowanego systemu lokalizacji wewnętrznej w różnych warunkach środowiskowych. W szczególności przeanalizowano wpływ typu technologii radiowej (UWB i BLE), zastosowanego protokołu komunikacyjnego (MQTT i Matter), wysokości umieszczenia urządzenia śledzącego, jego unikalnych właściwości kalibracyjnych, a także odległości pomiędzy punktem odniesienia a lokalizatorem. Wyniki zostały zebrane w formie surowych odczytów z każdego eksperymentu, a następnie poddane analizie statystycznej, obejmującej m.in. obliczenie wartości średnich, odchyłeń standardowych oraz błędów bezwzględnych w stosunku do różnych źródeł danych referencyjnych. Szczególną uwagę poświęcono ocenie spójności działania systemów opartych na różnych technologiach w obrębie tej samej architektury pomiarowej oraz ich potencjalnej wymienności.

Dla każdego z eksperymentów porównano odczyty uzyskane poprzez dwie niezależne ścieżki komunikacyjne: wcześniejsze podejście oparte na MQTT oraz zaproponowane, rozszerzone rozwiązanie z wykorzystaniem protokołu Matter. Dzięki temu możliwe było nie tylko sprawdzenie poprawności implementacji, ale również ocenienie przydatności protokołu Matter jako warstwy komunikacyjnej dla systemów lokalizacji w środowiskach domowych. Analizie poddano zarówno dokładność pojedynczych pomiarów, jak i stabilność oraz powtarzalność wyników w ramach każdego zestawu urządzeń i warunków pomiarowych. Dodatkowo, na potrzeby porównań podzielono wyniki na grupy względem technologii oraz dystansu dzielącego punkty pomiarowe. W trakcie przeprowadzonych eksperymentów zebrano łącznie 264 serie pomiarowe. Każdy eksperyment był trzykrotnie powtarzany i wykorzystywał od jednego do ośmiu lokalizatorów jednocześnie. W zależności od scenariusza liczba odczytów w jednej serii wahała się od kilkudziesięciu (najprostsze scenariusze statyczne) do niemal tysiąca (rozbudowane przypadki dynamiczne), a każdy pojedynczy pomiar zawierał estymaty odległości względem wszystkich ośmiu rozmieszczonych w przestrzeni punktów referencyjnych.

Porównanie technologii

Pierwszym przeanalizowanym aspektem był wpływ wybranej technologii na jakość uzyskanych estymacji odległości. Zestawienie wyników pokazano w Tab. 3.5, gdzie porównaniu podlegają cztery kolejne aspekty:

- technologia radiowa: BLE / UWB,
- protokół komunikacyjny: MQTT / Matter,
- wysokość lokalizatora: na podłodze (0 cm) / na statywie (150 cm),
- źródło danych referencyjnych: pomiar laserowy / spacer interaktywny / model trójwymiarowy.

Tabela 3.5: Porównanie błędów estymacji dystansu dla kombinacji źródła danych referencyjnych, wysokości umieszczenia lokalizatora, wykorzystanego protokołu komunikacyjnego oraz technologii radiowej.

Ref.	Stat. błędu	Podłoga				Trójnóg			
		MQTT		Matter		MQTT		Matter	
		UWB	BLE	UWB	BLE	UWB	BLE	UWB	BLE
Laser	Średnio [m]	0.14	5.18	0.14	2.38	0.14	2.44	0.14	2.44
	Odch. Std. [m]	0.11	7.57	0.11	1.56	0.10	1.35	0.10	1.35
	Średnio [%]	4.30	336.56	4.30	338.00	5.99	554.11	6.02	555.42
	Odch. Std. [%]	3.33	352.51	3.30	350.44	5.64	593.24	5.64	594.40
Spacer	Średnio [m]	0.13	5.18	0.13	2.38	0.09	2.36	0.09	2.36
	Odch. Std. [m]	0.12	7.57	0.12	1.56	0.09	1.36	0.09	1.36
	Średnio [%]	4.11	339.20	4.10	340.92	3.66	535.49	3.68	536.83
	Odch. Std. [%]	3.44	356.47	3.43	354.43	3.66	581.27	3.65	582.54
Model	Średnio [m]	0.14	2.47	0.14	2.40	0.11	2.38	0.11	2.39
	Odch. Std. [m]	0.11	1.55	0.11	1.57	0.10	1.35	0.10	1.36
	Średnio [%]	4.23	348.91	4.20	342.45	4.49	540.85	4.52	542.18
	Odch. Std. [%]	3.23	359.83	3.24	355.43	4.36	584.65	4.33	585.90

Wyniki z poziomego podłogi pochodzą z eksperymentów E1–E3, natomiast odczyty ze statywu z eksperymentów E6–E10 (UWB) i E12–E16 (BLE). Dla każdej konfiguracji obliczono błędy względem wszystkich trzech źródeł danych referencyjnych. Jak wynika z zestawienia, system oparty na technologii UWB cechował się wyraźnie wyższą dokładnością estymacji w porównaniu z rozwiązaniami BLE,

zarówno w wartościach bezwzględnych, jak i procentowych. Różnice były szczególnie widoczne przy większych odległościach oraz w scenariuszach z lokalizatorem umieszczonym na trójnogu.

Technologia radiowa. Wybór technologii radiowej okazał się najistotniejszym czynnikiem wpływającym na skuteczność estymacji odległości. Zgromadzone dane jednoznacznie potwierdzają wyższość technologii Ultra-Wideband nad Bluetooth Low Energy. Dotyczy to zarówno dokładności estymaty dystansu, jak i stabilności uzyskiwanych wyników. Dla UWB średni błąd bezwzględny mieścił się w zakresie od 9 cm do 14 cm, przy odchyleniu standardowym od 9 cm do 12 cm. W wartościach względnych błędy średnie nie przekraczały 7%, a ich zakres wahał się od 3,66% do 6,02% (przy odchyleniu standardowym w zakresie od 3,23% do 5,64%). W przypadku technologii BLE wartości te były zdecydowanie gorsze. Średnie błędy bezwzględne osiągały od 236 cm do 518 cm, a odchylenia standardowe sięgały aż 757 cm. Dla błędów względnych średnie mieściły się między 336,56% a 555,42%, przy odchyleniach standardowych w zakresie od 350,44% do 594,40%.

Protokół komunikacyjny. W większości testowanych konfiguracji nie zaobserwowano istotnych różnic pomiędzy wykorzystaniem protokołów MQTT i Matter. W przypadku systemów UWB, niezależnie od wysokości lokalizatora ani źródła danych referencyjnych, wyniki błędów bezwzględnych były identyczne, a błędy względne różniły się o mniej niż 0,1%. W przypadku systemów BLE zauważono niewielką poprawę precyzji dla pomiarów na poziomie podłogi, gdy dane zbierano protokołem Matter i porównywano z dalmierzem laserowym lub spacerem interaktywnym. Błąd bezwzględny spadł z 518 cm do 238 cm, lecz błędy względne pozostały bez zmian. Sugeruje to, że poprawa może wynikać z mniejszej liczby odebranych próbek, co, przy dużej niestabilności danych BLE, wpłynęło na chwilową stabilizację wyników.

Wysokość lokalizatora. Analiza wpływu wysokości zamontowania urządzenia śledzącego wykazała w przypadku UWB bardzo zbliżoną precyzję, niezależnie od wysokości. Różnice w błędach bezwzględnych nie przekraczały 4 cm, natomiast w błędach względnych zaobserwowano niewielki wzrost (z 3–4% przy pomiarach z poziomu podłogi do 5–6% na statywie). Potencjalnym wytłumaczeniem tego zjawiska jest skrócenie rzeczywistego dystansu między lokalizatorami a punktami referencyjnymi, które były rozmieszczone wyłącznie powyżej poziomu gruntu.

Przy stałej wartości błędu bezwzględnego zwiększa to jego wartość procentową. W systemie BLE wyniki były bardziej nieprzewidywalne. Dla protokołu Matter nie zaobserwowano wyraźnej różnicy w błędach bezwzględnych, choć podobnie jak w przypadku UWB nastąpił wzrost błędu względnego. W przypadku komunikacji MQTT zauważalna była poprawa. Średni błąd bezwzględny spadł z 518 cm do 244 cm, a odchylenie standardowe z 757 cm do 135 cm. Dla źródła danych w postaci pełnego modelu trójwymiarowego wyniki pozostały jednak bez zmian, co potwierdza wysoką zmienność i nieprzewidywalność pomiarów w systemie BLE.

Źródło danych referencyjnych. W analizowanych konfiguracjach nie stwierdzono istotnego wpływu metody pozyskiwania danych referencyjnych na końcową jakość estymacji. Jedynym wyjątkiem była konfiguracja pomiarów z poziomu podłogi w systemie BLE z komunikacją MQTT. Jak omówiono wcześniej, efekt ten przypisać należy jednak niestabilności pomiarów, a nie wykorzystanemu źródłu danych referencyjnych. W przypadku systemów UWB różnice między źródłami danych referencyjnych wynosiły maksymalnie 5 cm (dla błędów bezwzględnych, z odchyleniem 1 cm) i 3% (dla błędów względnych, z odchyleniem poniżej 2%). Dla BLE różnice były wyraźnie większe, szczególnie w błędach względnych: wahały się od 2 cm do 8 cm (dla błędów bezwzględnych, z odchyleniem 1 cm) oraz od 3% do 19% (dla błędów względnych, z odchyleniem od 2% do 12%).

Zmienność pomiędzy urządzeniami

Kolejnym analizowanym aspektem była zmienność wyników pomiędzy poszczególnymi urządzeniami tej samej technologii. Celem tej części badań było ustalenie, czy przed wdrożeniem systemu lokalizacji konieczne może być wcześniejsze skalibrowanie lub wyselekcjonowanie najlepiej działających modułów. Wyniki przedstawiono w Tab. 3.6.

Tabela 3.6: Porównanie błędów estymaty dystansu pomiędzy wszystkimi czterema lokalizatorami UWB (górna część) i BLE (dolna część) w zależności od źródła danych referencyjnych oraz wykorzystanego protokołu komunikacyjnego.

Ref.	Stat. błędu	UWB							
		MQTT				Matter			
		UWB	UWB	UWB	UWB	UWB	UWB	UWB	UWB
		1	2	3	4	1	2	3	4
Laser	Średnio [m]	0.17	0.16	0.11	0.13	0.18	0.16	0.11	0.12
	Odch. Std. [m]	0.12	0.13	0.10	0.11	0.12	0.13	0.09	0.10
	Średnio [%]	4.97	5.24	3.78	3.23	5.00	5.20	3.79	3.22
	Odch. Std. [%]	3.09	4.27	3.30	2.66	3.13	4.27	3.25	2.56
Spacer	Średnio [m]	0.16	0.15	0.12	0.10	0.16	0.15	0.12	0.10
	Odch. Std. [m]	0.11	0.12	0.15	0.10	0.11	0.12	0.15	0.10
	Średnio [%]	4.61	5.40	3.90	2.53	4.65	5.38	3.92	2.46
	Odch. Std. [%]	2.76	4.80	3.92	2.30	2.77	4.78	3.89	2.28
Model	Średnio [m]	0.17	0.17	0.11	0.11	0.17	0.17	0.11	0.11
	Odch. Std. [m]	0.12	0.12	0.09	0.11	0.12	0.12	0.09	0.11
	Średnio [%]	4.86	5.62	3.68	2.75	4.90	5.63	3.64	2.62
	Odch. Std. [%]	3.04	4.40	2.92	2.54	3.09	4.38	2.93	2.54
Ref.	Stat. błędu	BLE							
		MQTT				Matter			
		BLE	BLE	BLE	BLE	BLE	BLE	BLE	BLE
		1	2	3	4	1	2	3	4
Laser	Średnio [m]	2.39	13.57	2.17	2.59	2.15	2.63	2.15	2.58
	Odch. Std. [m]	2.28	25.32	1.40	1.29	2.33	1.24	1.38	1.31
	Średnio [%]	167.97	520.84	208.83	448.59	154.57	545.94	205.01	446.47
	Odch. Std. [%]	161.74	617.90	172.96	457.45	161.95	614.16	170.24	455.42
Spacer	Średnio [m]	2.40	13.58	2.16	2.60	2.16	2.64	2.13	2.59
	Odch. Std. [m]	2.28	25.31	1.41	1.28	2.33	1.22	1.39	1.30
	Średnio [%]	169.37	525.64	209.28	452.52	155.73	551.95	205.57	450.42
	Odch. Std. [%]	162.47	622.82	177.36	463.24	162.52	619.12	174.72	461.37
Model	Średnio [m]	2.41	2.65	2.20	2.61	2.17	2.64	2.18	2.60
	Odch. Std. [m]	2.29	1.23	1.41	1.29	2.34	1.23	1.39	1.30
	Średnio [%]	171.15	559.32	211.22	453.96	157.37	553.09	207.41	451.92
	Odch. Std. [%]	164.70	636.17	173.68	464.76	164.77	622.73	171.04	463.19

W przypadku urządzeń UWB całkowity średni błąd bezwzględny mieścił się w zakresie od 10 cm do 18 cm, z odchyleniem standardowym wynoszącym od 9 cm do 15 cm. Błąd względny wahał się od 2,62% do 5,63%, a jego odchylenie wynosiło od 2,28% do 4,80%. Dla pojedynczego urządzenia nie zaobserwowano wyraźnego wpływu ani źródła danych referencyjnych, ani protokołu komunikacyjnego na uzyskane wyniki. Wartości błędu bezwzględnego różniły się między protokołami komunikacyjnymi o maksymalnie 2 cm, a najczęściej były identyczne. Zaobserwowane różnice między kolejnymi urządzeniami wskazują jednak, że wcześniejsze wyselekcjonowanie najlepiej działających jednostek mogłoby poprawić końcową dokładność systemu. Choć różnice bezwzględne nie były duże, to ich względna wielkość w kontekście średnich wartości błędu jest zauważalna. Przykładowo, różnica między lokalizatorami UWB4 i UWB1 wyniosła 6 cm przy błędzie średnim równym 11 cm.

W przypadku urządzeń BLE zmienność pomiędzy modułami była znacznie większa niż w systemie UWB. Całkowity średni błąd bezwzględny wahał się od 215 cm do 265 cm, z odchyleniem standardowym w przedziale od 122 cm do 234 cm. Błąd względny mieścił się natomiast w zakresie od 154% do 559%, z odchyleniem od 161% do 636%. Podobnie jak w przypadku UWB, nie zaobserwowano wpływu ani protokołu komunikacyjnego, ani źródła danych referencyjnych na poprawę wyników pojedynczych urządzeń. Pomimo wcześniejszej kalibracji wszystkich modułów BLE, uzyskane wyniki analizy pomiędzy kolejnymi urządzeniami sugerują konieczność zastosowania dodatkowej procedury selekcji sprzętu. Różnica pomiędzy średnim błędem najlepiej (BLE3) a najgorzej (BLE2) działającego lokalizatora sięgnęła 45 cm. Choć w wartościach bezwzględnych jest to ponad siedmiokrotnie więcej niż w przypadku UWB, to różnica względna dla technologii BLE (nawet w najgorszym scenariuszu) nie przekroczyła jednak 20%.

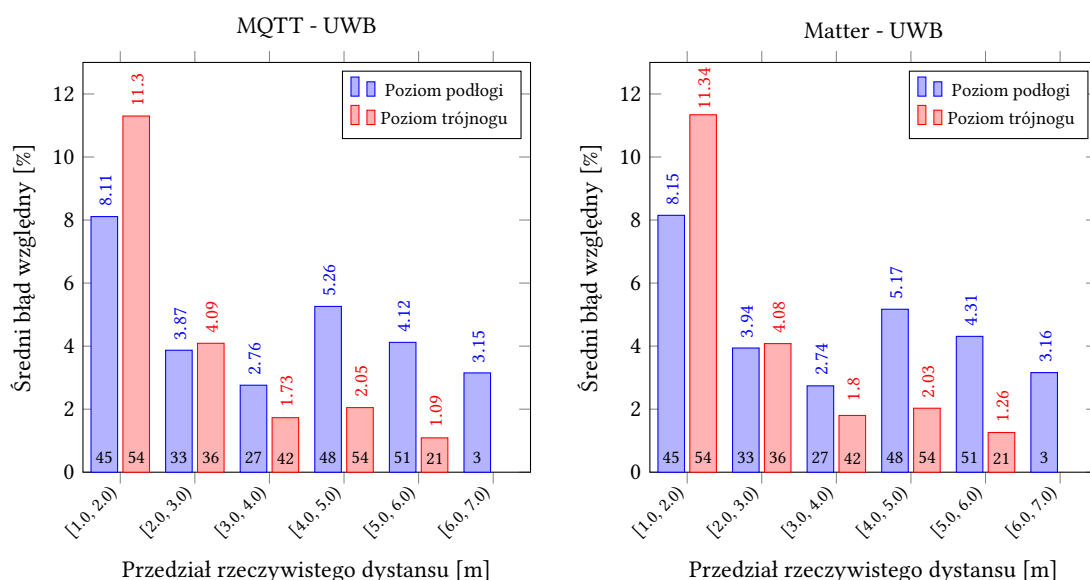
Podział pomiarów według odległości

Ostatnią przeprowadzoną analizą było zbadanie wpływu rzeczywistego dystansu pomiędzy lokalizatorem a punktem referencyjnym na dokładność estymacji. Wszystkie pojedyncze pomiary zostały przyporządkowane do jednego z sześciu przedziałów odległości (co 1 metr), a następnie obliczono dla nich statystyki błędów

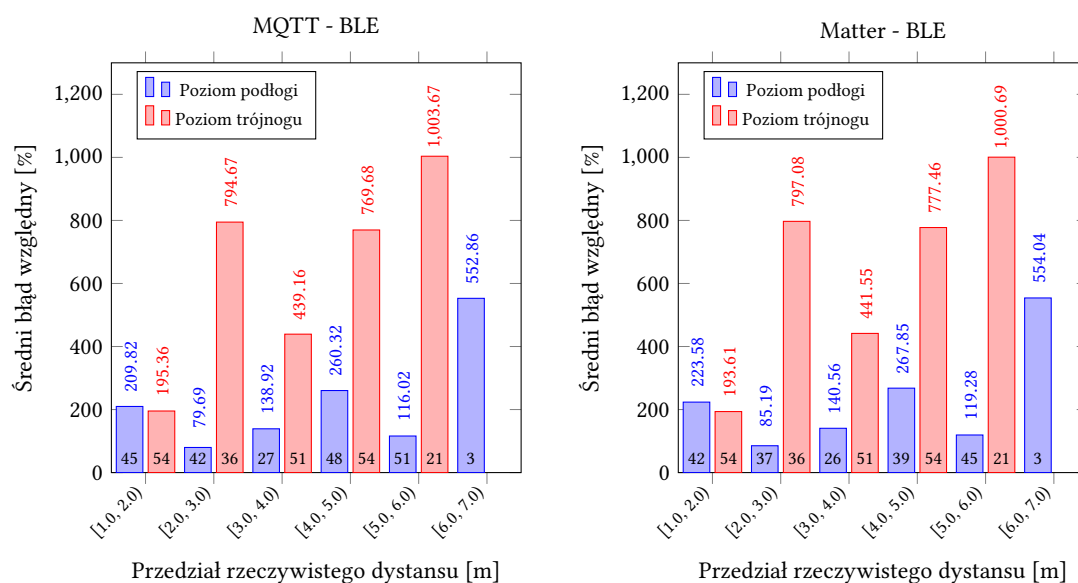
względnych. Ze względu na różną charakterystykę danych, wyniki przedstawiono osobno dla urządzeń UWB (Rys. 3.19) oraz BLE (Rys. 3.20). Każdorazowo przygotowano po dwa wykresy: osobno dla protokołu MQTT i Matter. Dla każdego przedziału odległości pokazano dwa słupki: lewy odpowiada pomiarom z poziomu podłogi, a prawy pomiarom na statywie. Nad słupkiem umieszczono wartość błędu, a pod nim liczbę próbek, z których obliczono wynik. Analizę źródła danych referencyjnych pominięto, jako że wcześniejsze analizy wykazały pomijalny jego wpływ na wyniki końcowe.

Dla lokalizatorów UWB zaobserwowano efekt stabilizacji wyników wraz ze wzrostem odległości. Największy błąd względny wystąpił w przedziale o najmniejszej rzeczywistej odległości i wynosił około 8% dla pomiarów z podłogi oraz 11% dla pomiarów ze statywu. Dla większych odległości wartości te ustabilizowały się odpowiednio na poziomie około 4% i 2%. Różnice między pomiarami z podłogi a ze statywu mogą wynikać z wpływu zakłóceń wywoływanych przez betonowe podłoże, lecz temat ten wymaga dalszych badań. W analizie nie zaobserwowano istotnego wpływu protokołu komunikacyjnego na wyniki.

Dla lokalizatorów BLE wyniki były znacznie mniej stabilne. Błędy względne mieściły się w zakresie od około 80% dla najlepszego scenariusza pomiaru z podłogi, aż do ponad 1000% dla najgorszego wyniku pomiaru na statywie. Mimo dużej zmienności można dostrzec wyraźny trend: wraz ze wzrostem odległości rośnie także średni błąd estymacji. Najmniejsze błędy wystąpiły w drugim najkrótszym przedziale odległości, a największe w najdłuższym. Wpływ wysokości lokalizatora był zauważalny, a pomiary z poziomu podłogi dawały lepsze wyniki. Znaczących różnic między protokołami komunikacyjnymi nie zaobserwowano.



Rysunek 3.19: Średni błąd względny dla kolejnych przedziałów rzeczywistej odległości lokalizatora UWB od kotwicy dla dwóch poziomów jego wysokości: podłogi oraz trójkątu. Na dole słupka przedstawiona liczność pomiarów w danym przedziale dystansu.



Rysunek 3.20: Średni błąd względny dla kolejnych przedziałów rzeczywistej odległości lokalizatora BLE od kotwicy dla dwóch poziomów jego wysokości: podłogi oraz trójkątu. Na dole słupka przedstawiona liczność pomiarów w danym przedziale dystansu.

3.3.6 Wnioski

Skonstruowane rozwiązanie prototypowe, zmontowany system lokalizacji, przeprowadzone eksperymenty oraz zebrane dane pomiarowe pozwoliły na sformułowanie szeregu wniosków dotyczących funkcjonowania zaproponowanego systemu radiowego szacowania odległości w warunkach środowiska domowego. Ich uwzględnienie w końcowym systemie fuzyjnym zauważalnie podnosi jego rzetelność i wartość wdrożeniową.

Wyniki wskazują, że protokół Matter może z powodzeniem zostać wykorzystany jako warstwa integracyjna w systemach lokalizacji wewnętrznej, a jego użycie nie ma negatywnego wpływu na wyniki estymacji odległości. W większości testowanych przypadków nie zaobserwowano różnicy pomiędzy wynikami uzyskanymi za pomocą protokołów MQTT i Matter. W niektórych konfiguracjach protokół Matter nawet nieznacznie stabilizował wyniki, co jednak potencjalnie może wynikać z liczby przetworzonych próbek oraz konkretnej wykonanej implementacji, a nie ze specyfiki samego protokołu. Należy jednak podkreślić, że to rozwój komponentów opartych o konkurencyjny protokół MQTT był szybszy i mniej problematyczny, głównie ze względu na jego powszechność, dostępność materiałów, a także generalnie niższy poziom skomplikowania. Wyniki potwierdzają, że protokół Matter może być z powodzeniem stosowany w analogicznych scenariuszach. Należy jednak mieć świadomość narzutów z tym związanych. Jego stosunkowo wczesny etap wdrożeniowy, rzadkość na rynku oraz ograniczona liczba przykładów i materiałów dla programistów stanowią zauważalne wyzwanie. Etap prototypowania rozwiązań opartych na protokole Matter wymaga więcej czasu i nakładów pracy, a docelowe wdrożenie na rynek odpowiednich certyfikacji. **W związku z powyższym, do docelowego systemu fuzyjnego jako protokół komunikacyjny sugeruje się wykorzystanie MQTT.**

W scenariuszach obejmujących środowisko domowe rozwiązanie bazujące na technologii UWB osiągnęło satysfakcjonującą precyzję, oscylującą wokół kilkunastu centymetrów. Potwierdza to jego przydatność jako technologii bazowej dla systemów pozycjonowania wysokiej dokładności, a także tych operujących w przestrzeni trójwymiarowej, np. przy odnajdowaniu przedmiotów codziennego użytku w mieszkaniu. Poprawna obsługa tej technologii jest jednak bardziej wymagająca

implementacyjnie, a komponenty są mniej popularne na rynku niż BLE. Powoduje to wyższy narzut czasu i kosztów na opracowanie takich urządzeń. Niemniej, rozwiązanie oparte na technologii BLE charakteryzowało się wysoką niestabilnością wyników i błędami estymacji dystansu o ponad rząd wielkości gorszymi niż w przypadku systemu UWB. Nie może on być zatem rekomendowany jako podstawa do wdrożeń o niskim marginesie błędu. Uwzględniając jednak jego niskie koszty, popularność oraz łatwość zakupu komponentów rozwojowych, podsystem BLE może być użyteczny w prostszych scenariuszach. Dostrzegane jego zastosowania obejmować mogą wykrycie najbliższego punktu odniesienia, identyfikację pomieszczenia, w którym znajduje się lokalizator, lub jako etap wstępnego przetwarzania w bardziej złożonych systemach pozycjonowania. Istotniejsze staje się tutaj binarne określenie faktu nawiązania połączenia, aniżeli oszacowanie rzeczywistej odległości do kotwicy. **Do docelowego systemu fuzyjnego rekomenduje się zatem opracowanie systemu integrującego obie technologie, np. dla określenia obecności w danym pomieszczeniu poprzez BLE, a następnie określenie precyzyjnej pozycji w jego obrębie z użyciem UWB.**

Wpływ wysokości lokalizatora na jakość estymacji wymaga dalszych badań. Obecnie zebrane pomiary nie pozwalają na jednoznaczną interpretację wpływu tej zmiennej. W przypadku urządzeń UWB nie zaobserwowano wyraźnej zależności, a w przypadku BLE wyniki wydają się być bardziej związane z charakterystyką technologii niż z samą wysokością. W dalszych badaniach warto uwzględnić bardziej zróżnicowane środowisko pomiarowe oraz różne konfiguracje przesłonięcia punktów referencyjnych. **Takie zalecenie, tj. zróżnicowanie umiejscowienia punktów referencyjnych radiowego systemu pozycjonującego, należy sformułować dla rozwiązania docelowego.**

W ramach skonstruowanego prototypu systemu radiowego potwierdzono pierwszą hipotezę dodatkową rozprawy. Zastosowanie cyfrowego modelu środowiska badawczego pozwoliło na opracowanie danych referencyjnych o równie wysokiej jakości, co ich określenie poprzez pomiary manualne. Efektywnie, ich późniejsze użycie nie miało wpływu na końcowe wyniki estymacji odległości. Otrzymane estymaty były porównywalne z tymi uzyskanymi przy pomiarach wykonywanych manualnie za pomocą dalmierza laserowego, przy czym proces ich pozyskiwania okazał się znacznie szybszy i wy-

godniejszy. W kontekście planowania, monitorowania i analizy eksperymentów potwierdzono wysoką użyteczność modelu cyfrowego, przy jednoczesnym zachowaniu jakości wyników. Potwierdza to, że tego typu model może z powodzeniem wspierać realizację badań naukowych, zarówno na etapie przygotowania i prowadzenia pomiarów, jak i w późniejszej analizie danych, bez konieczności powrotu do fizycznej przestrzeni testowej. Zachowanie cyfrowej wersji środowiska badawczego umożliwia jego analizę w dowolnym czasie, także po jego demontażu lub zmianie aranżacji. Zastosowanie cyfrowego modelu nie tylko nie pogarsza jakości uzyskanych danych, ale także znacząco usprawnia pracę i ułatwia prowadzenie nawet bardziej złożonych pomiarów. **Tym samym, docelowy system lokalizacji i rozpoznawania obiektów również zaleca się oprzeć na cyfrowym modelu nadzorowanego środowiska.**

Rozdział 4

Systemy lokalizacji oparte na technikach wizji komputerowej

Wizja komputerowa (z ang. *Computer Vision*, CV), jako nauka i technika poświęcona automatycznej analizie obrazów, to dziedzina o wieloletniej i bardzo bogatej historii. Początki sięgają lat pięćdziesiątych i sześćdziesiątych ubiegłego stulecia, nieprzerwanie pozostając od tego czasu w centrum zainteresowania badaczy [267]. Dziś stanowi nieodzowny element zarówno nauki, jak i przemysłu, prowadząc do optymalizacji i automatyzacji procesów produkcyjnych, wsparcia diagnostyki medycznej, czy tak prozaicznych czynności, jak automatyczne otwieranie bram parkingowych dla pojazdów o znanych numerach rejestracyjnych. Na przestrzeni dekad zmieniała się metodyka postępowania i wykorzystywane techniki. Początkowo, standardowa realizacja analizy i przetwarzania obrazu opierała się na podejściu algorytmicznym, wykorzystywała przekształcenia geometryczne i punktowe, filtry logiczne, morfologię obrazu czy transformację Fouriera [268]. Znaczącą uwagę kierowano wtedy na sam proces akwizycji danych oraz wykorzystywany do tego sprzęt. Oferowana przez ówczesne kamery jakość obrazu była na tyle niska, że stanowiło to istotny czynnik wpływający na końcowe rezultaty. Współczesna wizja komputerowa zdominowana jest już jednak przez rozwiązania oparte na uczeniu maszynowym, zwłaszcza sztucznych sieciach neuronowych, a poszczególne zastosowania stanowią samodzielnie rozwijające się kierunki badań (np. detekcja obiektów [269], klasyfikacja obiektów [270], segmentacja [271]).

Niniejszy rozdział, w celu zachowania spójności narracyjnej z Roz. 3, ma analogiczną strukturę, która przedstawia kompletny proces badawczy: od analizy teoretycznej i rozwoju autorskich metod, aż po walidację prototypowego systemu lokalizacji wizyjnej działającego w środowisku rzeczywistym. Wychodząc od przeglądu stanu wiedzy w dziedzinie detekcji obiektów i przetwarzania obrazów sferycznych, zaproponowano i zwalidowano autorskie metody usprawniające ten proces. Kluczowym założeniem było zbadanie skuteczności gotowych, przemysłowych rozwiązań detekcyjnych, a także omówienie procesu tworzenia eksperymentalnych zbiorów danych. Szczególny nacisk położono na weryfikację wpływu przetwarzania wstępnego na jakość detekcji oraz na ocenę możliwości wykorzystania najnowszych modeli estymacji głębi do realizacji prototypu pełnego rozwiązania wizyjnego do szacowania odległości w środowisku domowym.

Rozdział otwarto osadzeniem pracy w kontekście literaturowym (Roz. 4.1), omawiając sztuczne sieci neuronowe i ich architekturę (Roz. 4.1.1), miary skuteczności detektorów obiektów (Roz. 4.1.2), zagadnienie obrazowania sferycznego (Roz. 4.1.3), a także problematykę dostępności zbiorów danych (Roz. 4.1.4) i tworzenia własnych zbiorów zdjęć (Roz. 4.1.5). Następnie, w Roz. 4.2, przedstawiono wkład autorski: badania skuteczności już istniejących rozwiązań (Roz. 4.2.1, algorytm doboru najlepszego punktu rozcinań sfery w procesie rzutowania na obraz perspektywiczny (Roz. 4.2.2), a także opracowane zbiory rzeczywistych danych eksperymentalnych oraz badania zbiorów syntetycznych (Roz. 4.2.3). Rozdział zamyka opis prototypu systemu lokalizacji wizyjnej (Roz. 4.3), wykorzystany w jego ramach detektor głębi (Roz. 4.3.1), środowisko testowe (Roz. 4.3.2), przeprowadzone eksperymenty (Roz. 4.3.3) oraz szczegółowa analiza wyników (Roz. 4.3.4).

Podsumowując, wkład rozdziału w przeprowadzone prace naukowe i wdrożeniowe to:

- dogłębny przegląd i usystematyzowanie metod wizji komputerowej oraz technik detekcji obiektów w kontekście obrazów sferycznych;
- opracowanie i walidacja autorskiej metody doboru punktu rozcięcia panoram równokątnych, minimalizującej zniekształcenia obiektów na krawędziach;
- pozyskanie i przygotowanie dedykowanych, obszernych zbiorów danych: zarówno rzeczywistych (*PanoSilesia*, *Office*, *EquiB&B*), jak i syntetycznych

(*SPHEREA*) do trenowania i testowania modeli;

- przeprowadzenie kompleksowej analizy porównawczej architektur detektorów obiektów (YOLOv5, YOLOv8, Faster-RCNN) oraz wpływu hiperparametrów na ich skuteczność w pracy z obrazami panoramicznymi;
- projekt i weryfikacja eksperymentalna prototypu systemu lokalizacji wizyjnej, zdolnego do estymacji odległości do wykrytych obiektów, wraz z rekomendacjami konfiguracyjnymi dla docelowego systemu fuzyjnego.

4.1 Osadzenie pracy w kontekście

Analogicznie do zagadnienia pozycjonowania radiowego, tematyka wizji komputerowej i analizy obrazów jest bardzo rozległa. W niniejszym rozdziale przedstawione zostają tylko te aspekty, które faktycznie zastosowano w rozwiązaniu prototypowym lub końcowym systemie fuzyjnym, a dalsze kierunki rozeznania wskazywane są poprzez referencje do innych publikacji.

Kluczowymi cechami, których oczekuje się od rozważanych systemów analizy wizyjnej, są: wysoka wydajność pod względem czasowym, mierzalność uzyskiwanych wyników, dostępność publicznych zbiorów danych do trenowania modeli, a także prostota opracowania własnych zbiorów treningowych. Wynikają one głównie z realiów przemysłowych docelowego rozwiązania, ale również z wykorzystywanych w przedsiębiorstwie partnerskim technologii oraz posiadanych kompetencji.

4.1.1 Sztuczne sieci neuronowe i architektura YOLO

Sztuczne sieci neuronowe stanowią fundamentalny model uczenia maszynowego, inspirowany działaniem układu nerwowego człowieka. Model ten składa się z prostych jednostek obliczeniowych, tytułowych neuronów, które przekazują sygnały za pośrednictwem połączeń o określonych wagach. Stanowią one uniwersalne narzędzie do szacowania wartości praktycznie dowolnej funkcji, dzięki czemu znajdują zastosowanie w tak różnorodnych zadaniach, jak klasyfikacja obrazów, analiza tekstu czy sterowanie robotami. Kluczową cechą sieci jest zdolność do uczenia się reprezentacji danych oraz automatycznego wydobywania złożonych cech. Popularność tych rozwiązań wynika z ich wysokiej skuteczności i elastyczności, umożliwia-

jących efektywną pracę nawet na dużych i zróżnicowanych zbiorach danych.

Od strony formalnej, działanie pojedynczego neuronu można opisać modelem matematycznym. Dla wektora $\mathbf{x}_{\text{neur}} = [x_{\text{neur}_1}, x_{\text{neur}_2}, \dots, x_{\text{neur}_n}]^T$ danych wejściowych oraz odpowiadającego mu wektora wag $\mathbf{w}_{\text{neur}} = [w_{\text{neur}_1}, w_{\text{neur}_2}, \dots, w_{\text{neur}_n}]^T$, sygnał wyjściowy y_{neur} jest zdefiniowany jako [272]:

$$y_{\text{neur}} = f_{\text{akt}} \left(\sum_{i_{\text{neur}}=1}^{n_{\text{neur}}} w_{\text{neur}_i} x_{\text{neur}_i} + b_{\text{neur}} \right) = f_{\text{akt}}(\mathbf{w}_{\text{neur}}^T \mathbf{x}_{\text{neur}} + b_{\text{neur}}), \quad (4.1)$$

gdzie:

- x_{neur_i} to i_{neur} -ta wartość wejściowa,
- w_{neur_i} to waga przypisana do i_{neur} -tego połączenia wejściowego,
- b_{neur} to tzw. obciążenie (z ang. *bias*), parametr pozwalający na przesunięcie funkcji aktywacji,
- $f_{\text{akt}}()$ to nieliniowa funkcja aktywacji.

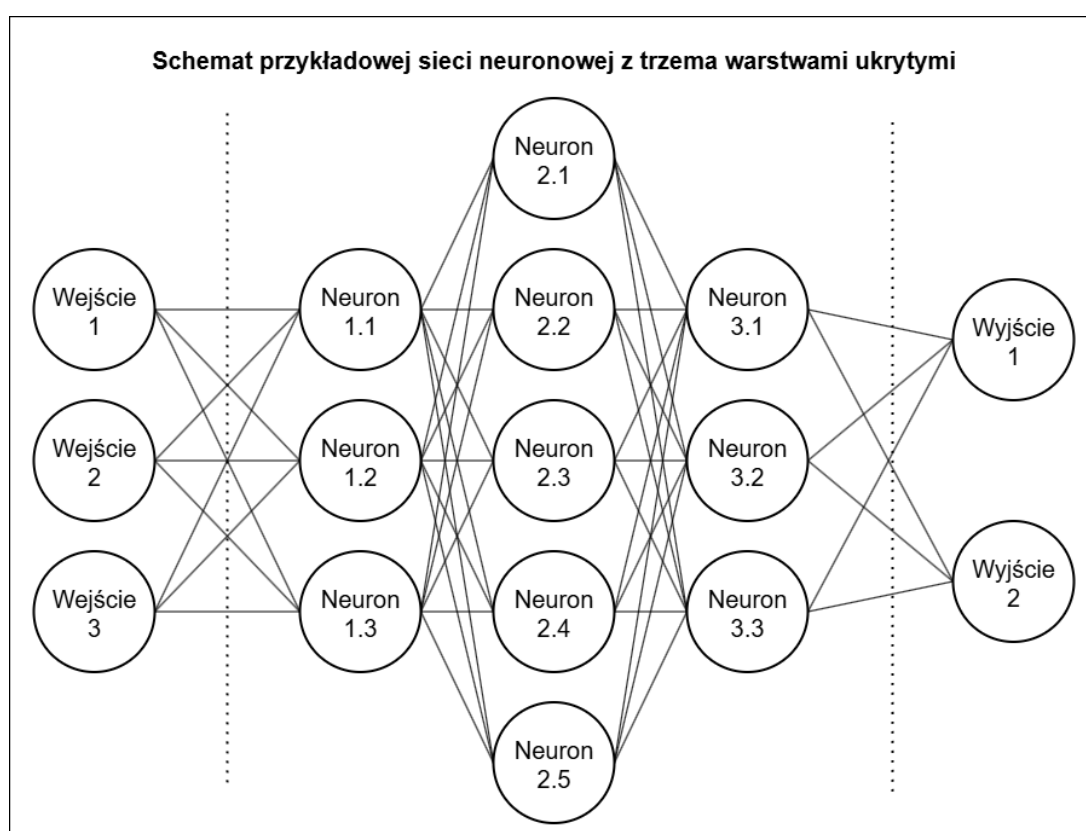
Wybór funkcji aktywacji ma kluczowe znaczenie dla zdolności sieci do modelowania złożonych zależności. Wśród najczęściej stosowanych funkcji aktywacji w głębokich sieciach neuronowych, szczególnie w kontekście wizji komputerowej, należy wymienić ReLU i jej warianty. Funkcja ReLU (z ang. *Rectified Linear Unit*), która została spopularyzowana m.in. w pracach Glorota i in. [273], zdefiniowana jest wzorem:

$$\text{ReLU}(z_{\text{neur}}) = \max(0, z_{\text{neur}}). \quad (4.2)$$

Jej popularność wynika z prostoty obliczeniowej oraz dużej skuteczności w procesie uczenia. W praktyce, funkcja ta wprowadza do modelu nieliniowość w sposób niezwykle efektywny: wszystkie wartości ujemne na jej wejściu są zerowane, natomiast wartości dodatnie przepuszczane są bez zmian. Takie działanie nie tylko znacząco upraszcza obliczenia, ale również przeciwdziała problemowi zanikającego gradientu, co jest kluczowe przy trenowaniu głębokich architektur sieci neuronowych.

Podstawowy sztuczny neuron wykonuje trzy główne operacje: przyjmuje sygnały wejściowe, mnoży je przez odpowiadające im wagi, a następnie scala wyniki

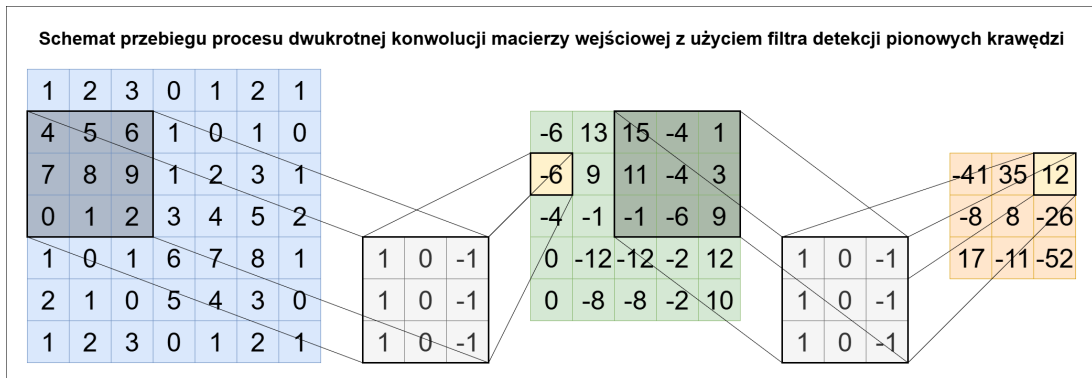
poprzez funkcję aktywacji, generując wyjście. W praktyce neurony organizowane są w rozbudowane struktury. Klasyczny model warstwowy sztucznej sieci neuronowej obejmuje trzy typy warstw: warstwę wejściową, zestaw połączonych warstw ukrytych oraz pojedynczą warstwę wyjściową. Na wejściu dostarczane są surowe wartości, np. pomiary z czujników lub piksele obrazu. Warstwy ukryte odpowiadają za nieliniową transformację sygnałów, a ich liczba i rozmiar decydują o głębokości i złożoności sieci. Warstwa wyjściowa generuje końcowy wynik modelu, np. przypisaną do obrazu klasę rozpoznanego obiektu. Schemat ideowy tego typu sieci przedstawiono na Rys. 4.1.



Rysunek 4.1: Schemat ideowy struktury warstwowej sieci neuronowej o trzech warstwach neuronów ukrytych.

W zadaniu rozpoznawania obiektów oraz generalnie w procesie analizy obrazu, popularnym rozwiązaniem są konwolucyjne sieci neuronowe (nazywane również sieciami splotowymi, z ang. *Convolutional Neural Network*, CNN). Jest to typ głą-

bokich sztucznych sieci neuronowych, które wyspecjalizowane zostały do analizy danych o strukturze matrycowej, w szczególności obrazów i wideo. Główną innowacją tej architektury, w stosunku do omówionego klasycznego modelu warstwowego, jest zastosowanie filtrów konwolucyjnych (zwanych też jądrami konwolucyjnymi, z ang. *Kernels*) przesuwanych po matrycy danych wejściowych. W ramach każdej warstwy splotowej wagi takiego filtra mnożone są z elementami obrazu, a następnie łączone w jeden wynik dla całego analizowanego fragmentu. Zastosowanie wielu różnych filtrów pozwala na wyodrębnianie poszczególnych cech, np. krawędzi czy wzorców kształtów, w sposób automatyczny, tj. bez konieczności ich ręcznego określania jeszcze przed szkoleniem modelu. Uproszczoną wizualizację tego procesu przedstawiono na Rys. 4.2. Idea splotu jest koncepcyjnie prosta, jej przedstawienie graficzne przysparza jednak problemów, a skuteczne wyjaśnienie najłatwiej uzyskać na nagraniach lub grafikach interaktywnych. Przebieg procesu konwolucji, sposoby jej właściwej realizacji, a także dobór wykorzystywanych filtrów to odrębne zagadnienia, którym poświęca się dużo uwagi w literaturze przedmiotu [274–278].



Rysunek 4.2: Schemat ideowy procesu konwolucji obrazu.

Z matematycznego punktu widzenia, operacja dyskretnego splotu dwuwymiarowego, stosowana w warstwach konwolucyjnych, jest operacją na dwóch macierzach: macierzy wejściowej I_k (reprezentującej obraz lub mapę cech z poprzedniej warstwy) oraz macierzy filtra (jądra) K_k o wymiarach $m_k \times n_k$. Wynikiem jest nowa macierz, tzw. mapa cech (z ang. *Feature Map*) S_k , której wartości obliczane są zgodnie z następującą zależnością (zaczerpniętą z [272, 274]):

$$S_k(i_k, j_k) = (I_k * K_k)(i_k, j_k) = \sum_{u_k=1}^{m_k} \sum_{v_k=1}^{n_k} I_k(i_k - u_k, j_k - v_k) K_k(u_k, v_k). \quad (4.3)$$

W praktyce, w kontekście sieci neuronowych, operację tą realizuje się jako korelację krzyżową, co eliminuje konieczność odwracania jądra K_k . Wówczas wzór przyjmuje postać:

$$S_k(i_k, j_k) = \sum_{u_k=1}^{m_k} \sum_{v_k=1}^{n_k} I_k(i_k + u_k, j_k + v_k) K_k(u_k, v_k). \quad (4.4)$$

Elementy jądra K_k są w rzeczywistości wagami, których wartości są optymalizowane w procesie uczenia sieci. Umożliwia to automatyczne dostosowanie filtrów do wykrywania specyficznych, istotnych cech w danych wejściowych, np. krawędzi, tekstur czy bardziej złożonych wzorów. Parametry operacji splotu, takie jak krok (z ang. *Stride*) oraz dopełnienie (z ang. *Padding*), determinują natomiast wymiary wyjściowej mapy cech i pole recepcyjne neuronów.

W konwolucyjnych sieciach neuronowych strukturę modelu budują różne typy warstw, z których każda odpowiada za inny etap przetwarzania danych. Kluczowe znaczenie mają warstwy splotowe, które automatycznie wyodrębniają istotne cechy obrazu poprzez filtrowanie i analizę lokalnych wzorców. Następnie warstwy aktywacji wprowadzają do modelu nieliniowość, umożliwiając mu modelowanie złożonych zależności. Dzięki warstwom łączenia (z ang. *Pooling Layer*) sieć zyskuje zdolność redukcji wymiarów oraz agregowania informacji. Na końcu stosuje się natomiast warstwy w pełni połączone, które przetwarzają zgromadzone cechy i dostarczają finalnej klasyfikacji, określając ostateczny wynik działania sieci.

Sieci konwolucyjne odnalazły swoje zastosowanie w procesie detekcji obiektów, czyli kluczowym zadaniu z zakresu wizji komputerowej, którego celem nie jest przypisanie całego zdjęcia do pewnej pojedynczej kategorii, ale odnalezienie i klasyfikacja wielu obiektów w ramach tego samego zdjęcia. Wymagane jest oznaczenie każdego z wykrytych obiektów przy użyciu etykiety, zwykle w postaci tzw. ramki ograniczającej (z ang. *Bounding Box*, BB). Ramka taka otacza możliwie przyległe wskazany obiekt, a towarzysząca jej klasa opisuje przynależność do wyuczonej kategorii. Operacja detekcji jest zauważalnie bardziej złożonym problemem,

niż sama klasyfikacja. Do głównych wyzwań w tym procesie zaliczamy: skalowanie i rozmiar obiektów, zmienność ich położenia, a także złożoność obliczeniową.

W odpowiedzi na te wyzwania powstało kilka kolejnych sposobów projektowania architektury modeli splotowych. Definiują one ogólne zasady ich działania i stosowane są zależnie od potrzeb oraz dostępnych zasobów. Są to tzw. detektory jednostopniowe (z ang. *Single Shot Detector*, SSD) oraz detektory dwustopniowe (z ang. *Two-Stage Detectors*). Główną różnicą pomiędzy nimi jest sposób realizacji procesu wskazania i klasyfikacji regionu na obrazie. Ideowo są to dwie operacje: osobno wskazać należy obszar, w którym spodziewamy się odnaleźć jakiś obiekt, a następnie określić, jakiego rodzaju jest to obiekt. W ten właśnie sposób funkcjonują detektory dwustopniowe. Najpierw generowana jest w nich tzw. propozycja obszaru (z ang. *Region Proposal*), a następnie klasyfikuje się ją i dokonuje regresji granic wykrytych obiektów. Rozwiązania takie zapewniają wysoką precyzję wykrywania, jednak charakteryzują się relatywnie dużym zapotrzebowaniem na zasoby obliczeniowe oraz czas przetwarzania. Detektory jednostopniowe wykonują obie te operacje w pojedynczym kroku, tj. bez osobnej fazy generowania regionów. Sieć bezpośrednio przewiduje położenie oraz klasy wszystkich obiektów na podstawie pełnego obrazu wejściowego. Największą zaletą takiego podejścia jest bardzo duża szybkość działania przy zachowaniu akceptowalnej jakości wyników, co czyni je szczególnie atrakcyjnym wyborem w systemach wymagających przetwarzania w czasie rzeczywistym. W literaturze odnaleźć można również pojęcie detektorów bezstopniowych (z ang. *Zero-shot Detector*, ZSD), ich zastosowanie jest jednak wysoce specyficzne i dotyczy rozpoznawania obiektów o klasach nieobecnych w zbiorach treningowych [279–281].

Typowym scenariuszem w procesie tworzenia cyfrowych modeli nieruchomości jest przemieszczanie się człowieka z urządzeniem akwizycji obrazu po budynku. Kluczowe jest więc możliwie szybkie przetwarzanie danych, nawet kosztem jakości detekcji. Nierozpoznany z dużej odległości obiekt najpewniej i tak w kilku następnych chwilach zostanie zeskanowany z mniejszego dystansu, a łączny efekt płynności przeważa nad skutecznością wnioskowania z pojedynczego kadru obrazu. Na tej podstawie to właśnie detektory jednostopniowe wytypowane zostały do dalszych badań omówionych w rozprawie.

Wśród detektorów jednostopniowych dużą popularność zyskała architektura

YOLO (z ang. *You Only Look Once*). Model ten stanowił swoisty przełom w projektowaniu rozwiązań detekcji obiektów. Łączy on charakterystyczną dla tych modeli prędkość działania z niespotykaną zdolnością do precyzyjnej lokalizacji i klasyfikacji obiektów o różnorodnych rozmiarach w ramach tego samego obrazu. Zasadniczą ideą działania architektury YOLO jest podział całego przekazanego obrazu na siatkę komórek, w której dla każdej z nich przewiduje się współrzędne ramki ograniczającej, prawdopodobieństwo obecności obiektu oraz jego klasę. Proces ten traktowany jest jak operacja regresji, co zasadniczo kontrastuje z metodyką znaną z innych detektorów, szczególnie dwustopniowych.

Od pierwotnej wersji opublikowanej w 2016 roku [282], architektura ta była wielokrotnie rozwijana [283, 284]. W ciągu zaledwie kilku lat realizacji badań i gromadzenia materiału do niniejszej rozprawy przybrała już siedem kolejnych form (generacji). Badania wstępne oparte zostały o generację piątą (YOLOv5 [285]), rozszerzenie badań obejmowało wersję ósmą (YOLOv8 [286]), a końcowe wdrożenie wykorzystuje już generację jedenastą (YOLOv11 [287]). Architektura YOLO jest z powodzeniem wykorzystywana w realizacji różnorodnych zadań, od jazdy autonomicznej [288], przez produkcję układów scalonych [289], aż po diagnostykę nowotworów [290]. Jako pewna rozległa rodzina modeli jest rozwijana przez wiele różnych ośrodków, osób i instytucji, a poszczególne jej generacje rozpowszechniane są dla przemysłu w oparciu o różne licencje, tak darmowe, jak i płatne.

4.1.2 Miary skuteczności detektorów obiektów

Ocena skuteczności detektorów obiektów wymaga precyzyjnych, dobrze zdefiniowanych miar ilościowych. Jedną z najbardziej rozpowszechnionych metryk, jakie odnajdziemy w literaturze oraz praktyce badawczej, jest uśredniona precyzja średnia (z ang. *Mean Average Precision*, mAP). W uproszczeniu nazywana również po prostu precyzją średnią. Podstawą do obliczenia wskaźnika mAP jest ocena każdej pojedynczej detekcji modelu w odniesieniu do wspomnianego już stanu faktycznego (z ang. *Ground Truth*, GT). W kontekście detekcji obiektów zbiór GT stanowi zestaw wzorcowych ramek ograniczających wraz z przypisanymi do nich etykietami klas, które zostały wcześniej ręcznie oznaczone przez człowieka. W tym celu definiuje się wskaźnik przecięcia nad sumą (z ang. *Intersection over Union*, IoU),

który mierzy stopień pokrycia między ramką ograniczającą przewidzianą przez model (BB_{pred}) a ramką wzorcową (BB_{gt}). Wartość IoU obliczana jest jako:

$$\text{IoU}(BB_{pred}, BB_{gt}) = \frac{\text{obszar}(BB_{pred} \cap BB_{gt})}{\text{obszar}(BB_{pred} \cup BB_{gt})}, \quad (4.5)$$

gdzie licznik reprezentuje pole części wspólnej obu ramek, a mianownik pole ich sumy. Wartość IoU mieści się w zakresie od 0 (brak pokrycia) do 1 (idealne dopasowanie).

Na podstawie wartości IoU, każdą detekcję ocenia się w odniesieniu do ustalonego progu zgodności τ (np. $\tau = 0,5$), jako:

- **Prawdziwie pozytywna** (z ang. *True Positive*, TP): poprawna detekcja, dla której istnieje obiekt GT tej samej klasy, a ich IoU $\geq \tau$;
- **Fałszywie pozytywna** (z ang. *False Positive*, FP): błędna detekcja, która nie spełnia warunków TP. Dzieje się tak, gdy detekcja wskazuje na tło (brak obiektu GT) lub gdy jej IoU z pasującym obiektem GT jest niższe od progu τ ;
- **Fałszywie negatywna** (z ang. *False Negative*, FN): pominięty obiekt GT, dla którego model nie wygenerował żadnej poprawnej detekcji (tj. żadnej z wynikiem ocenionym jako TP).

Wartości TP, FP oraz FN zliczane są dla każdej z klas osobno. Należy przy tym zaznaczyć, że w procesie oceny każda niedokładna detekcja (niespełniająca progu IoU) skutkuje jednocześnie zliczeniem jednej fałszywej detekcji (FP) oraz jednego pominiętego obiektu (FN).

Proces obliczania wartości mAP rozpoczyna się od wyznaczenia dla każdej klasy obiektów dwóch miar fundamentalnych: precyzji (z ang. *Precision*) oraz czułości (z ang. *Recall*). Precyzja określa, jaki odsetek detekcji wskazanych przez model był poprawny, natomiast czułość informuje, jaki odsetek wszystkich istniejących obiektów został prawidłowo wykryty. Dysponując zliczeniami TP, FP i FN dla danej klasy i ustalonego progu IoU, miary te definiuje się wzorami:

$$\text{Precyzja} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (4.6)$$

$$\text{Czułość} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (4.7)$$

Aby wykreślić krzywą precyzja-czułość (z ang. *Precision-Recall Curve*, PRC), wszystkie detekcje dla danej klasy sortuje się malejąco według ich wyniku pewności (z ang. *Confidence Score*, CS), przypisanego przez model. Następnie, poprzez iteracyjne obniżanie progu dla tego wyniku, oblicza się wartości precyzji i czułości, uzyskując kolejne punkty na wykresie. Pole pod tak wyznaczoną krzywą definiuje metrykę średniej precyzji (z ang. *Average Precision*, AP) dla pojedynczej klasy:

$$AP = \int_0^1 p_{det}(r_{det}) dr_{det}, \quad (4.8)$$

gdzie $p_{det}(r_{det})$ jest funkcją precyzji w zależności od czułości r_{det} . Ostatecznie, wskaźnik mAP stanowi średnią arytmetyczną wartości AP obliczonych dla wszystkich N_{det} klas w zbiorze danych:

$$mAP = \frac{1}{N_{det}} \sum_{i_{det}=1}^{N_{det}} AP_{i_{det}}. \quad (4.9)$$

W zapisie tekstowym obrany próg zgodności IoU przedstawia się wraz ze skrótem samego wskaźnika, np. mAP@50, co oznacza uśrednioną precyzję średnią przy progu zgodności $\tau = 0,5$.

Stosowanie miary mAP pozwala na porównywanie skuteczności detektorów obiektów niezależnie od liczby czy rozkładu klas, a także umożliwia analizę działania modeli przy różnych poziomach tolerancji na nakładanie się ramek ograniczających, określanych przez próg IoU. W praktyce najczęściej stosowane są dwa podejścia: próg IoU równy 0,5, zgodnie z protokołem zbioru danych PASCAL VOC [291], lub zakres progów od 0,5 do 0,95 ze skokiem 0,05, zalecany przy pracy ze zbiorem Microsoft COCO [292].

Warto zaznaczyć, że oprócz średniej precyzji stosowane są również inne miary oceny skuteczności detekcji, takie jak miara F1 (z ang. *F1-Score* [293, 294]), średnia czułość (z ang. *Average Recall*, AR [291, 292]), logarytmiczna średnia błędnych detekcji (z ang. *Log-Average Miss Rate*, LAMR [295, 296]), czy krzywa odpowiedzi swobodnej (z ang. *Free-Response Receiver Operating Characteristic*, FROC [297]). W niniejszej rozprawie głównym wskaźnikiem pozostaje jednak *Mean Average Precision*, który przedstawiany jest dla dwóch progów: 0,5 (mAP@50 lub mAP50) oraz zakresu od 0,5 do 0,95 ze skokiem 0,05 (mAP@50-95 lub mAP50-95).

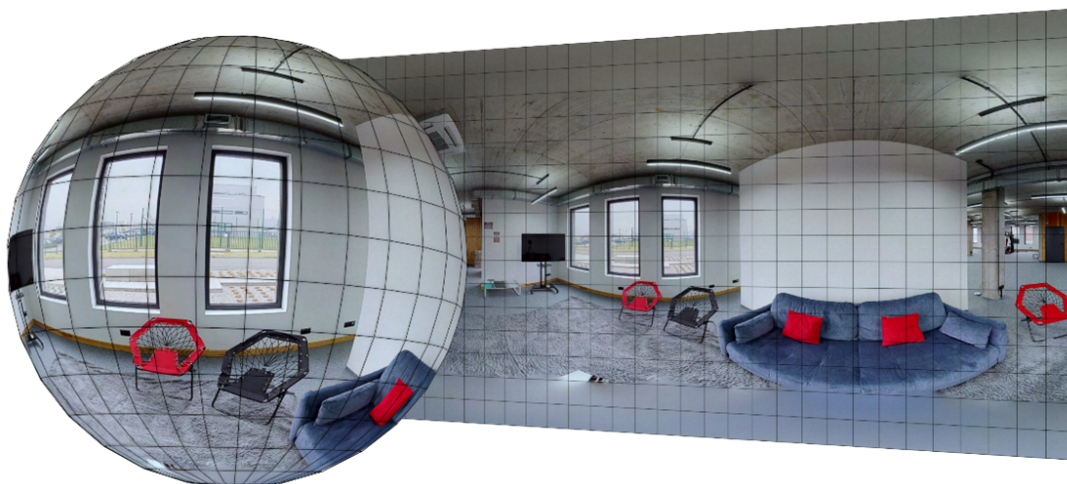
4.1.3 Obrazy sferyczne i ich reprezentacja

Tradycyjnie rozumiane zdjęcia to płaskie obrazy, które oddają wybrany przez autora wycinek rzeczywistości, widziany z jednego konkretnego punktu i pod określonym kątem. Najczęściej przedstawione w sposób celowo zbliżony do tego, jak otoczenie w danym momencie postrzega będący na miejscu człowiek. Na potrzeby niniejszej pracy nazywać je będziemy zdjęciami perspektywicznymi i traktujemy jako domyślny format gromadzenia danych wizyjnych. Matematycznie reprezentuje je macierz pikseli (lub kilka macierzy, dla zdjęć kolorowych), gdzie każdy piksel odpowiada konkretnemu punktowi na matrycy aparatu.

Taka reprezentacja nie spełnia jednak potrzeb omówionych we wcześniejszych rozdziałach, szczególnie dotyczących realistycznego odwzorowania całego otoczenia. Do stworzenia modeli trójwymiarowych oraz spacerów interaktywnych niezbędna jest akwizycja nie tylko konkretnego punktu widzenia kamery, ale całego otaczającego nas świata, rośnie więc znaczenie akwizycji danych o odpowiednich formatach. Odpowiedzią na tę potrzebę wydają się być szeroko rozumiane zdjęcia sferyczne. Idea takiej reprezentacji jest dość intuicyjna, jednakże jej zapisanie stanowi pewne wyzwanie. Zdjęcie sferyczne to reprezentacja całego otoczenia wokół punktu, w którym znajduje się kamera, obejmując pełne pole widzenia w zakresie 360° w poziomie i 180° w pionie. W praktyce oznacza to, że niezależnie od tego, w którym kierunku spojrzymy, czy to przed siebie, za siebie, w górę czy w dół, obraz, który zobaczymy, zostaje utrwalony i możliwe będzie jego odтворzenie. Matematycznie jest to odwzorowanie wszystkich możliwych kierunków obserwacji na powierzchni sfery o promieniu jednostkowym, której środek pokrywa się z pozycją kamery.

Wspomniane wyzwanie w zapisie reprezentacji sferycznych związane jest z nieprzygotowaniem istniejących systemów do pracy z inną strukturą danych niż macierz pikseli. Konieczne staje się więc sprowadzenie wszelkich pozyskanych danych do ujednoliconych formatów, które składać się będą z pewnych wariacji tychże macierzy. Jednym z najpopularniejszych takich formatów jest panorama równokątna (z ang. *Equirectangular Panorama*, ERP). Technicznie, ERP to projekcja sfery na płaszczyznę, a zatem rzutowanie bardzo wysokiej rozdzielczości materiału wizyjnego na prostokątne zdjęcie perspektywiczne o proporcjach 2:1 (sze-

rokość:wysokość). Dla łatwiejszego zobrazowania zagadnienia możemy wskazać wysokie podobieństwo takiego rzutowania do projekcji Merkatora, znanej z map świata. Tam również pewna (w przybliżeniu) sfera odwzorowywana zostaje na płaszczyźnie.

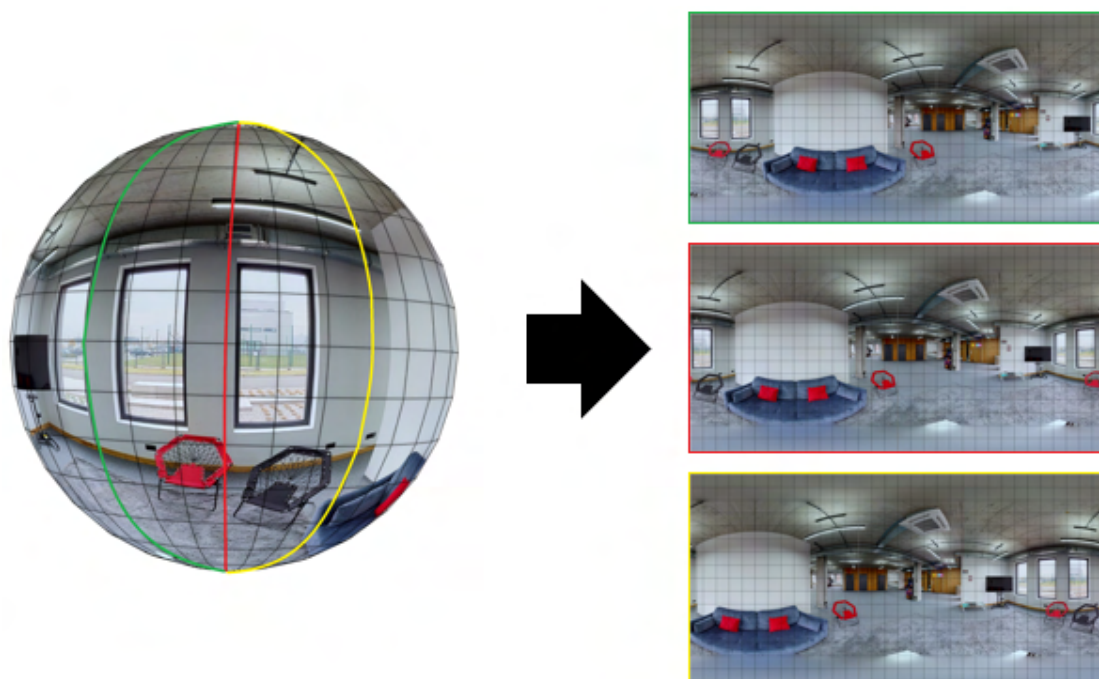


Rysunek 4.3: Zobrazowanie zdjęcia sferycznego i jego reprezentacja w postaci panoramy równokątnej.

Jako format docelowy, ERP jest niezależny od źródła danych. Może pochodzić z szerokiego zakresu urządzeń: od przemysłowych skanerów przestrzeni, przez półprofesjonalne kamery sportowe, aż po sprzęt zupełnie amatorski, tj. smartfon z odpowiednim oprogramowaniem. Wsparcie dla wyświetlania i pracy z takim zapisem zobaczyć możemy w mediach społecznościowych, systemach monitoringu, a przede wszystkim jest on popularnym elementem procesu tworzenia cyfrowych bliźniaków i spacerów interaktywnych. Co więcej, ERP stanowią bardzo istotny sposób reprezentacji świata, jako że gromadzą bardzo dużą ilość informacji w każdym zdjęciu. Dzięki pełnemu polu widzenia szczególnie widoczne jest to w kontekście liczby obiektów znajdujących się na każdym kadrze.

Istotną cechą takiej reprezentacji obrazu sferycznego jest jego zdolność do zawijania na krawędziach ekranu. Ułożenie wielu kopii tej samej panoramy, jedna obok drugiej, odwzorowuje efekt nieograniczonego obrotu w osi poziomej. Niemniej, trzeba mieć również na uwadze, że proces rzutowania ze sfery na panoramę rozpoczyna się na określonym południku, co wpływa na końcowe wyniki. Zagadnie-

nie to, ze szczegółami, omówione zostaje w Roz. 4.2.2. Jako wprowadzenie, przedstawia się natomiast wizualizację tego zjawiska. Na Rys. 4.4, przy pomocy trzech kolorów (kolejno: zielonego, czerwonego i żółtego), przedstawiono trzy kolejne południki źródłowej sfery. Skutek wyboru każdego z nich, jako punktu rozpoczęcia rzutowania ze sfery na płaszczyznę, prezentowany jest w postaci wynikowej panoramy obramowanej analogicznym kolorem. Zaobserwować możemy, że przesuwanie południka rozcięcia sfery sprawia, że wynikowe panoramy mają inne początki i końce. Istotnym wnioskiem tej obserwacji jest, że może prowadzić to do efektu rozcięcia obiektu umieszczonego na skraju i rozrzucenia go po przeciwległych krańcach obrazu.



Rysunek 4.4: Wizualizacja wpływu wybranego miejsca rozcięcia sfery obrazu sferycznego przy rzutowaniu do reprezentacji dwuwymiarowej.

Choć formaty panoramiczne oferują szerokie pokrycie sceny, wprowadzają również znaczące wyzwania. Dla ERP są to głównie zniekształcenia geometryczne wynikające z algorytmu projekcji sferycznej. Panoramy te są efektywnie silnie zaburzone, szczególnie w swoich obszarach polarnych (górna i dolna krawędź obrazu), a także pokrywają różnorodne warunki oświetleniowe jednocześnie. Rzutowanie za-

burza kształt i pozorne odwzorowanie położenia obiektów, np. prosta lampa sufitowa zostaje zakrzywiona, a ściany pomieszczenia wydają się wypukłe. Ciągłość obrazu sprawia natomiast, że jednocześnie silnie może fluktuować natężenie światła. Dodatkowym wyzwaniem staje się przesył strumieniowy gromadzonego na żywo sferycznego materiału wideo, które ze względu na rozdzielczość obrazu wymaga bardzo wydajnego łącza. Efektywnie jest to zauważalne ograniczenie zastosowania takiego rozwiązania przy niedoborach mocy obliczeniowej urządzeń brzegowych lub przepustowości łącza internetowego.

Dla zagadnienia obrazów sferycznych i ich przetwarzania, w literaturze wskazuje się wiele innych rozwiązań, tak dla samego ich odwzorowania, jak i szerokiego wachlarza potencjalnych zastosowań. Dwa główne formaty alternatywne dla danych sferycznych to tzw. siatki trójkątów (z ang. *Spherical Polyhedron*, SP) oraz projekcje sześciennie (czasem zwane kostkowymi, z ang. *Cubemap*, CP). Rozbudowaną dalszą analizę, omówienie kamer, formatów danych, szczegółów algorytmów rzutowania, a także zbiory danych i przykładowe zastosowania omówiono w pracy [298].

Na potrzeby niniejszej rozprawy analizowane są przede wszystkim zdjęcia środowiska domowego w formacie panoram równokątnych, nazywane dalej IERP (z ang. *Indoor Equirectangular Panorama*). Uwarunkowane jest to faktycznie stosowaną przez firmę partnera projektu technologią odtwarzania przestrzeni, powszechnością zastosowania tego formatu w przemyśle, a także dostępnością alternatywnych, publicznych zbiorów danych i urządzeń akwizycji obrazu do opracowania własnych.

4.1.4 Publiczne zbiory danych

Obrazowanie sferyczne, w kontekście pozyskiwania, oznaczania i udostępniania zbiorów danych, jest tematem stosunkowo nowym, jednocześnie coraz popularniejszym, a także coraz bardziej restrykcyjnie monitorowanym. Na przestrzeni lat tworzenie zbiorów zdjęć przestrzennych, szczególnie w ich odwzorowaniu panoramicznym, uwarunkowane było urządzeniami akwizycji obrazu. Zaporowe ceny kamer o 360-stopniowym polu widzenia mocno ograniczały ich popularność. Spadek cen sprzętu sprawił, że dziedzina ta zaczęła się szybko rozwijać i równie szybko do-

tarła do problemu natury ludzkiej: prywatności użytkowników. Szczególnie w kontekście prowadzonych badań środowiska domowego zagadnienie poufności informacji zawartych w skanach przestrzeni i na zdjęciach sferycznych nabiera kluczowego znaczenia. Dostęp do wcześniej publicznych zbiorów zdjęć jest obecnie obwarowany wymaganiami licencyjnymi, deklaracjami uzasadnionego dostępu akademickiego, a po wielokroć również zakazem wykorzystania do celów komercyjnych.

Niemniej, w literaturze odnaleźć można nadal dostępne zbiory, na których bazować można swoje rozwiązania. Jednym z pierwszych odnalezionych podczas analizy zbiorów sferycznych był “*SUN360*” [299], rozszerzony później o adnotacje dedykowane środowisku wewnątrzbudynkowemu i wykorzystany w badaniach rekonstrukcji sceny trójwymiarowej [300, 301]. Chronologicznie późniejszy zbiór, “*F-360iSOD*” [302], dedykowano detekcji obiektów charakterystycznych, jednakże liczność jego próbek (nieco ponad 100) uniemożliwiała skuteczne wykorzystanie do szkolenia dużych sieci neuronowych. Bliższy tematycznie środowisku domowemu zbiór “*PANDORA*” [215], o zauważalnie większym rozmiarze (3000 zdjęć), rozszerzył oryginalnych 37 klas obiektów ze zbioru “*360-Indoor*” [303] o 10 kolejnych, jednocześnie jednak zmieniając wykorzystywany rodzaj adnotacji obiektów na obrócone pole widzenia (z ang. *Rotated Bounding-Field-of-View*, RBFoV), co skutecznie utrudnia jego zastosowanie w klasycznych detektorach obiektów i zwiększa oczekiwane nakłady pracy na pierwsze prototypy.

Wśród przybliżonych zbiorów ciężko się jednak doszukać jednego przyjętego zbioru standardowego, który wykorzystywany byłby powszechnie do porównywania rozwiązań i ich jakości między sobą. Na swój sposób popularnym odniesieniem w literaturze jest zbiór “*Stanford 2D-3D*” [304]. Sam jest on jednak opracowany na podstawie wcześniejszej wersji, tj. zbiorze “*Stanford Large-Scale 3D Indoor Spaces*” [20], a także posiada swoich następców, np. “*Pano3D*” [305]. Każde z tych opracowań różni się zawartością i zakresem rodzajów scen, które możemy w nich odnaleźć. Podkreśla to konieczność dobrego dopasowania wybranego zbioru danych do konkretnego wdrożenia docelowego, ponieważ brak tu typowych zbiorów wskaźnikowych, jak wcześniej wspomniane *PASCAL VOC* [291] albo *Microsoft COCO* [292] dla zdjęć perspektywicznych.

4.1.5 Etykietowanie danych własnych

Przy okazji omawiania sieci neuronowych przytoczono już domyślny sposób przygotowywania etykiet wykrywanych obiektów z użyciem ramek ograniczających. W kontekście wizji sferycznej rozwiązanie to staje się jednak słabo dopasowane do potrzeb. Zniekształcenia wynikające z projekcji sfery na płaszczyznę utrudniają ich wykorzystanie i obniżają jakość oznaczeń. W literaturze zaproponowano alternatywne sposoby etykietowania obiektów, np. z użyciem ograniczającego pola widzenia (z ang. *Bounding Field of View*, BFoV), które polega na oznaczaniu obiektów z uwzględnieniem kąтового pola widzenia w przestrzeni sferycznej. Niemniej, BFoV nie stanowi standardu w powszechnie stosowanych narzędziach do etykietowania danych. Popularne platformy i narzędzia do zarządzania zbiorami zdjęć, np. “CVAT” [306], “LabelMe” [307], czy “Roboflow” [308], nie wspierają oznaczeń o tego typu geometrii. Stąd też dominująca większość odnalezionych rozwiązań przemysłowych ogranicza się do klasycznego podejścia opartego na dwuwymiarowych prostokątnych ramkach, nawet w przypadku danych panoramicznych.

Prowadząc analizę sposobu realizacji oznaczeń dla wizji sferycznej, zidentyfikowane zostały trzy proponowane narzędzia do ich edycji: “PanoAnnotator” [309], “Label360” [310] oraz “360RAT” [311], kolejno z lat: 2018, 2020 i 2022. Oferują one funkcje nanoszenia etykiet, które swoim kształtem dopasowują się do zniekształceń panoramy, pozostają jednak projektami badawczymi o charakterze eksperymentalnym, które wymagają od użytkownika istotnych kompetencji technicznych oraz manualnej konfiguracji środowiska. Brak wsparcia dla oznaczeń sferycznych w przemysłowo wdrażanych systemach etykietujących stanowi istotne wyzwanie w procesie przygotowywania zbiorów danych do szkolenia sieci neuronowych.

W literaturze odnaleźć można badania sposobów redukcji zniekształcenia rzutowanych obrazów [312–314], łagodzenia wpływu takich zaburzeń na pracę detektorów obiektów albo opracowywania dedykowanych rozwiązań specjalnie przygotowanych do pracy z obrazami sferycznymi. Niemniej, wszystkie one prowadzą do dodatkowych nakładów pracy: od przetwarzania wstępnego surowych danych, przez dedykowane architektury sieci neuronowych, aż po specjalne narzędzia wizualizujące wyniki.

Z powyższych względów, na potrzeby niniejszej pracy przyjęto za akcepto-

walne rozwiązanie oparte na klasycznych prostokątnych ramach ograniczających, przy jednoczesnym zachowaniu świadomości ograniczeń tej reprezentacji w przypadku użycia odwzorowania typu ERP.

4.2 Elementy autorskie

W niniejszym rozdziale przedstawiono badania własne, obejmujące przekrój zagadnień związanych z opracowaniem autorskiego prototypu systemu wizji komputerowej. Prace rozpoczęto od analizy dostępnych na rynku rozwiązań opartych na sieciach neuronowych architektury YOLO. Następnie omówiono opracowaną procedurę usprawnienia procesu rzutowania danych sferycznych do formatu ERP, a także opisano proces pozyskiwania, przygotowania i generacji autorskich zbiorów zdjęć panoramicznych. Całość zamyka charakterystyka i omówienie wizyjnej części finalnego systemu fuzyjnego.

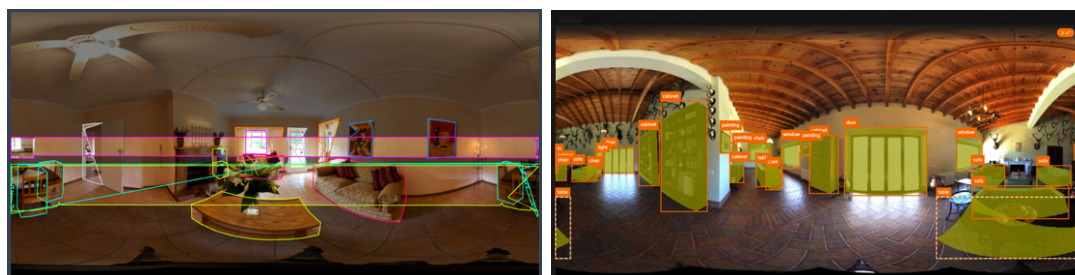
4.2.1 Badanie skuteczności dostępnych rozwiązań

Jednym z oczekiwań realizacji projektów wdrażających rozwiązania uczenia maszynowego do przemysłu jest ich osadzenie w już dostępnych na rynku, komercyjnych produktach lub systemach o otwartym kodzie źródłowym. Unika się w ten sposób potencjalnie niepotrzebnych nakładów na opracowywanie rozwiązań autorskich. W tym celu wytypowano bezpłatny, publiczny zbiór danych o tematyce środowiska domowego, a następnie wykorzystano go w badaniach skuteczności detekcji obiektów przez modele wybranych architektur sztucznych sieci neuronowych.

Zbiór danych

W realizacji eksperymentów postanowiono wykorzystać omówiony wcześniej zbiór obrazów panoramicznych *WiMR* (Roz. 2.2.4) oraz konwolucyjne sieci neuronowe z rodziny *YOLO* (Roz. 4.1.1). Wczytanie danych do chmurowego systemu zarządzania zbiorami *Roboflow* [308] zwróciło niespodziewane rezultaty, tj. błędne odwzorowanie oznaczonych obiektów, które umieszczone były na skrajach obrazu. Przyczyną błędów była reprezentacja oznaczeń w przestrzeni ciągłej (zawiniętej

panoramie), które prowadziły do wyjścia poza zakres zdjęcia w jego reprezentacji perspektywicznej. Zbiór danych przetworzono, oznaczenia umieszczone na krawędziach rozdzielono, a poprawiony zbiór danych wgrano do systemu ponownie. Wizualizację obu sytuacji przedstawiono na Rys. 4.5.

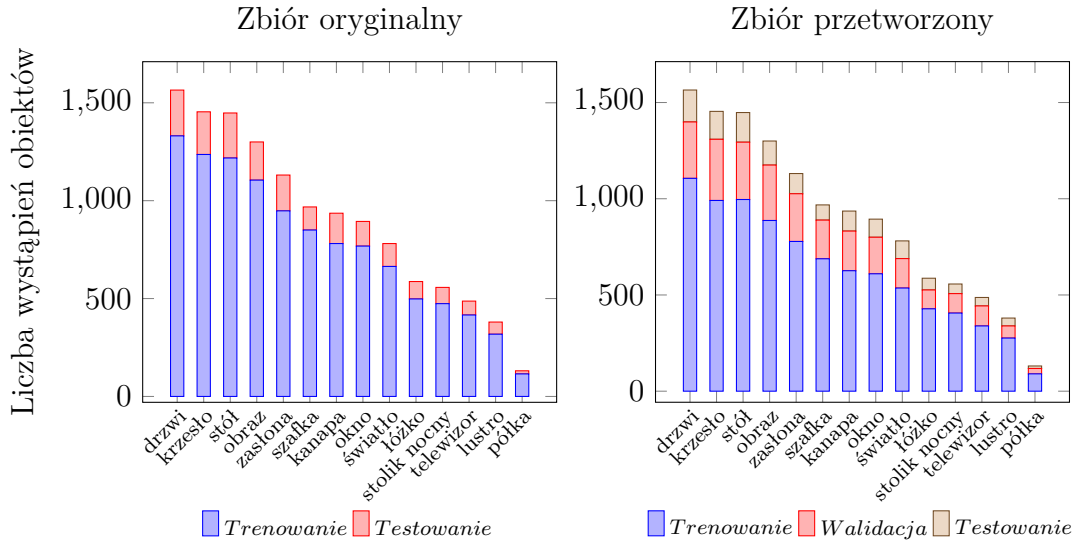


Rysunek 4.5: Wizualizacja prezentacji oznaczonego sferycznie zbioru danych przed (lewo) i po (prawo) wprowadzeniu poprawek.

W procesie budowy modeli uczenia maszynowego dane zwyczajowo dzieli się na trzy podzbiory: treningowy, walidacyjny i testowy. Odpowiednio służące do: trenowania, strojenia hiperparametrów i oceny działania. Oryginalny zbiór *WiMR* zawierał podział tylko na dane treningowe i testowe. W ramach zrealizowanych prac połączono je i dokonano nowego, losowego podziału na trzy części. Zestawienie liczebności adnotacji według klas obiektów przed i po przetworzeniu pokazano na Rys. 4.6. Zbiór *WiMR* obejmuje 666 obrazów ERP, głównie wnętrz różnych budynków biurowych, hotelowych, muzealnych oraz sakralnych. Oryginalnie dane podzielono w proporcji 85%/15% (566/100), a nowy podział wynosi 70%/20%/10% (468/132/66).

Przeprowadzone eksperymenty

Przygotowane etykiety oraz zbiór zdjęć wykorzystano w eksperymentach z użyciem detektorów obiektów dwóch generacji sieci neuronowych architektury YOLO, tj. generacji piątej i ósmej. Do oceny skuteczności detekcji obiektów dla wszystkich klas zastosowano miarę uśrednionej precyzji średniej (mAP). Obie wykorzystane generacje sieci oferują po pięć wariantów modeli różniących się rozmiarem (liczbą parametrów): *Nano* (n), *Small* (s), *Medium* (m), *Large* (l) oraz *Extra-Large* (x). W ramach analizy przetestowano wszystkie dostępne warianty, a badania rozsze-



Rysunek 4.6: Porównanie liczności obiektów i sposobu ich podziału w zbiorach źródłowym oraz przetworzonym.

rzo o eksperymenty dotyczące wpływu wybranych hiperparametrów na jakość detekcji. Analizowano rozmiar partii danych (z ang. *Batch Size*, BS), rozdzielczość obrazów oraz długość trenowania modeli (liczbę epok szkolenia).

Badania rozpoczęto od analizy wielkości partii danych i rozdzielczości obrazów. W analizie uwzględniono dwa progi detekcji: 0,50 oraz zakres od 0,50 do 0,95 z krokiem 0,05, a także dwa rozmiary zdjęć: 512 pikseli i 1024 pikseli. Modele szkolono przez 50 epok i z użyciem domyślnych wartości pozostałych hiperparametrów. Eksperymenty prowadzono tak długo, jak pozwalały na to ograniczenia sprzętowe. Rozpoczynając od pojedynczej próbki, wielkość partii była podwajana aż do momentu, gdy pamięć karty graficznej nie była w stanie obsłużyć dalszego wzrostu, co prowadziło do błędów typu brak pamięci (z ang. *Out of Memory*, OOM). Podejście to zaowocowało macierzą testową obejmującą 200 możliwych konfiguracji pomiarowych ($10 \text{ modeli} \times 2 \text{ rozdzielczości obrazów} \times 10 \text{ rozmiarów paczki wsadu}$). Dla 122 konfiguracji udało się uzyskać wyniki, a 78 skutkowało błędami.

Wyniki uzyskane dla modeli z rodziny YOLOv5 wskazują na wzrost wartości mAP wraz z początkowym zwiększaniem rozmiaru wsadu, przy czym optymalny zakres mieści się między wartościami 16 a 32. Dalsze zwiększanie wielkości partii

prowarziło do pogorszenia wyników. Przy rozmiarze wsadu równym 512 wszystkie modele powodowały przeciążenie karty graficznej i kończyły się błędem pamięci. Wpływ rozdzielczości obrazów na osiągnięte wyniki był wyraźnie widoczny w każdej konfiguracji testowej. Zaobserwowano również, że zależność ta jest silniejsza w mniejszych modelach (*Nano*, *Small*) niż w większych (*Large*, *Extra-Large*). Dla rozmiaru paczki równemu jednej próbce, przyrost wartości mAP@50 przy przejściu z obrazu rozmiaru 512 pikseli do 1024 pikseli wynosił odpowiednio: 40% (*Nano*), 24% (*Small*), 8% (*Medium*), 7% (*Large*) i 6% (*Extra-Large*). Podobna zależność występowała również dla mAP@50-95. Różnice sięgały odpowiednio: 50%, 33%, 16%, 13% i 12%. Nie zaobserwowano zależności między rozdzielczością obrazu a optymalnym rozmiarem wsadu. Szczegółowe wyniki dla mAP@50 przedstawiono w Tab. 4.1. Wnioski są spójne niezależnie od przyjętego progu detekcji, dlatego też tabela wyników dla mAP@50-95 zostaje pominięta.

Tabela 4.1: Wpływ rozmiaru wsadu i rozmiaru obrazu na uzyskaną wartość mAP@50 dla rodziny modeli YOLOv5. Najlepsze wyniki wyróżniono pogrubieniem, błędy braku pamięci oznaczono kursywą.

Model	Rozmiar obrazu	YOLOv5 mAP@50								
		Rozmiar wsadu								
		1	2	4	8	16	32	64	128	256
YOLOv5n	512	0.355	0.359	0.352	0.356	0.367	0.372	0.359	0.235	0.137
	1024	0.496	0.510	0.496	0.483	0.503	0.490	0.486	<i>OOM</i>	<i>OOM</i>
YOLOv5s	512	0.477	0.490	0.494	0.479	0.521	0.519	0.506	0.322	<i>OOM</i>
	1024	0.591	0.601	0.592	0.585	0.612	0.613	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv5m	512	0.587	0.573	0.575	0.574	0.598	0.590	0.594	<i>OOM</i>	<i>OOM</i>
	1024	0.635	0.640	0.637	0.636	0.661	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv5l	512	0.604	0.608	0.607	0.611	0.621	0.626	0.627	<i>OOM</i>	<i>OOM</i>
	1024	0.646	0.641	0.644	0.635	0.667	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv5x	512	0.611	0.613	0.616	0.621	0.643	0.646	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
	1024	0.649	0.661	0.647	0.651	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>

Wyniki uzyskane dla modeli z rodziny YOLOv8 są zbliżone do wyników modeli YOLOv5. Zaobserwowano podobną zależność między wielkością wsadu, a osiąganą wartością mAP, tj. początkowy wzrost wydajności, a następnie osiągnięcie maksimum, które w przypadku YOLOv8 przesuwają się w stronę wsadów o większej liczbie próbek. Najbardziej wyraźnym przypadkiem jest partia licząca 64 próbki, która okazała się osiągać najlepsze wyniki mAP@50-95 w pięciu z dziesięciu konfiguracji. Modele z rodziny YOLOv8 zużywały więcej pamięci niż ich odpowiedniki z rodziny YOLOv5. Błędy pamięci występowały o jeden krok wcześniej, przy wsadzie rozmiaru 256 próbek, niezależnie od odmiany modelu. Wnioski

te pozostają spójne dla obu analizowanych progów detekcji. Dla modeli z rodziny YOLOv8 wpływ rozdzielczości obrazu na wyniki detekcji również był wyraźnie widoczny, przy czym okazał się bardziej przewidywalny i mniej podatny na wahania niż w przypadku rodziny YOLOv5. Zwiększenie rozdzielczości obrazu poprawiało skuteczność detekcji we wszystkich przeprowadzonych eksperymentach. Dla wsadu w postaci pojedynczej próbki zaobserwowano wzrost wartości mAP@50 odpowiednio o: 27% (*Nano*), 18% (*Small*), 20% (*Medium*), 20% (*Large*) oraz 23% (*Extra-Large*). W przypadku mAP@50-95 różnice były jeszcze większe i wyniosły: 39%, 32%, 30%, 29% oraz 34% dla kolejnych wariantów modeli. Można zatem podsumować, że dla rodziny YOLOv8 wpływ rozdzielczości obrazu na osiąganą skuteczność detekcji jest bardziej przewidywalny i ustabilizowany, a zyski jakości detekcji wynoszą średnio ponad 20% dla mAP@50 i ponad 30% dla mAP@50-95. Szczegółowe wyniki dla mAP@50 przedstawiono w Tab. 4.2.

Tabela 4.2: Wpływ rozmiaru wsadu i rozmiaru obrazu na uzyskaną wartość mAP@50 dla rodziny modeli YOLOv8. Najlepsze wyniki wyróżniono pogrubieniem, błędy braku pamięci oznaczono kursywą.

Model	Rozmiar obrazu	YOLOv8 mAP@50								
		Rozmiar wsadu								
		1	2	4	8	16	32	64	128	256
YOLOv8n	512	0.314	0.344	0.331	0.347	0.348	0.338	0.359	0.210	<i>OOM</i>
	1024	0.399	0.412	0.416	0.417	0.425	0.429	0.486	<i>OOM</i>	<i>OOM</i>
YOLOv8s	512	0.381	0.406	0.408	0.416	0.429	0.419	0.506	<i>OOM</i>	<i>OOM</i>
	1024	0.451	0.478	0.480	0.503	0.488	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv8m	512	0.410	0.434	0.450	0.453	0.458	0.450	0.594	<i>OOM</i>	<i>OOM</i>
	1024	0.492	0.528	0.530	0.550	0.555	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv8l	512	0.421	0.454	0.442	0.453	0.454	0.440	0.627	<i>OOM</i>	<i>OOM</i>
	1024	0.504	0.536	0.555	0.542	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
YOLOv8x	512	0.425	0.452	0.460	0.464	0.472	0.472	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>
	1024	0.523	0.550	0.554	0.554	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>	<i>OOM</i>

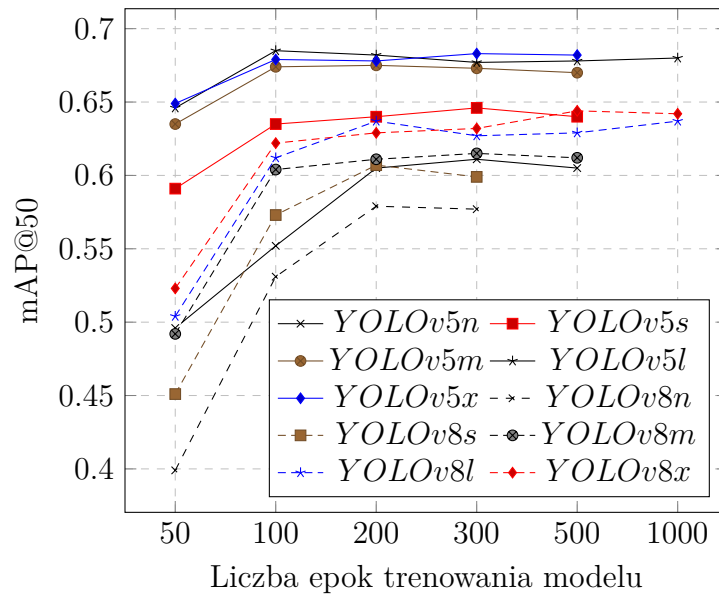
Tempo uczenia się sieci zostało przeanalizowane w sposób zbliżony do badania wielkości rozmiaru wsadu. Trening każdorazowo kontynuowano aż do momentu jego automatycznego przerwania przez mechanizm wczesnego zatrzymywania, aktywowany po wykryciu braku poprawy przez określoną liczbę epok (domyślnie 100 dla YOLOv5 i 50 dla YOLOv8). Uzyskiwaną skuteczność detekcji mierzono dla kolejnych liczb epok szkolenia: 50, 100, 200, 300, 500 oraz 1000. Na podstawie wcześniejszych eksperymentów z rozmiarem wsadu, w których większy rozmiar skutkowało lepszymi rezultatami, wszystkie testy w tej części przeprowadzono z wykorzystaniem obrazów w rozmiarze 1024 pikseli. Aby skrócić czas trenowa-

nia, dla każdego modelu zastosowano największą obsługiwaną wielkość rozmiaru wsadu.

W przypadku modeli YOLOv5 najlepsze wyniki osiągnięto po 300 epokach treningu, choć różnice po przekroczeniu 200 epok były niewielkie. Mniejsze modele wykazywały wyraźnie gorszą skuteczność niż większe, przy czym nie zaobserwowano istotnych różnic pomiędzy dużym (YOLOv5l) a bardzo dużym modelem (YOLOv5x). Co więcej, model *Large* osiągnął nieco lepsze rezultaty niż model *Extra-Large*, co może sugerować, że osiągnięto granice możliwości zbioru danych.

Modele YOLOv8 badano analogicznie. Podobnie jak w analizie rozmiaru wsadu, również tutaj zauważono, że nowsze modele szybciej osiągały granice poprawy wyników: zazwyczaj o jeden krok wcześniej niż ich odpowiedniki z rodziny v5. Tylko dla największych modeli konieczne było przeprowadzenie pełnego, 1000-epokowego treningu, aby aktywowany został mechanizm wczesnego zatrzymania. Wszystkie modele wykazywały ten sam ogólny trend: szybka poprawa skuteczności detekcji w pierwszych 200 epokach i stabilizacja wyników w kolejnych. Choć czysto liczbowo dłuższe treningi przynosiły nieco lepsze rezultaty, różnice były minimalne, a dla mAP@50-95 powyżej 500-ej epoki szkolenia zaczęto zauważać pierwsze objawy przetrenowania modeli.

Zestawienie wyników obu rodzin modeli dla mAP@50 przedstawiono na Rys. 4.7. Dla mAP@50-95 uzyskane obserwacje są analogiczne, a wykres zostaje pominięty. Dla wszystkich badanych konfiguracji widoczny jest niedostatek treningu przy użyciu 50 epok, jednakże już przy 100 epokach rezultaty ulegają znaczącej poprawie, a po osiągnięciu 200 epok stabilizują się. Dalsze wydłużanie treningu, nawet pięciokrotnie, nie przekłada się na lepsze wyniki, a w niektórych przypadkach prowadzi do przetrenowania sieci. W sumie uzyskano 173 zestawy pomiarowe: 122 w ramach analizy wpływu wielkości rozmiaru wsadu oraz 51 dotyczących długości treningu. Najlepszym do przetwarzania wybranego zbioru zdjęć panoramicznych okazał się model YOLOv5x, trenowany przez 300 epok, z obrazami wejściowymi o rozdzielczości 1024 pikseli. Najlepsze wyniki pozostałych modeli zostały w Tab. 4.3.



Rysunek 4.7: Porównanie uzyskiwanych wartości mAP@50 dla modeli YOLOv5 i YOLOv8 w zależności od liczby epok trenowania.

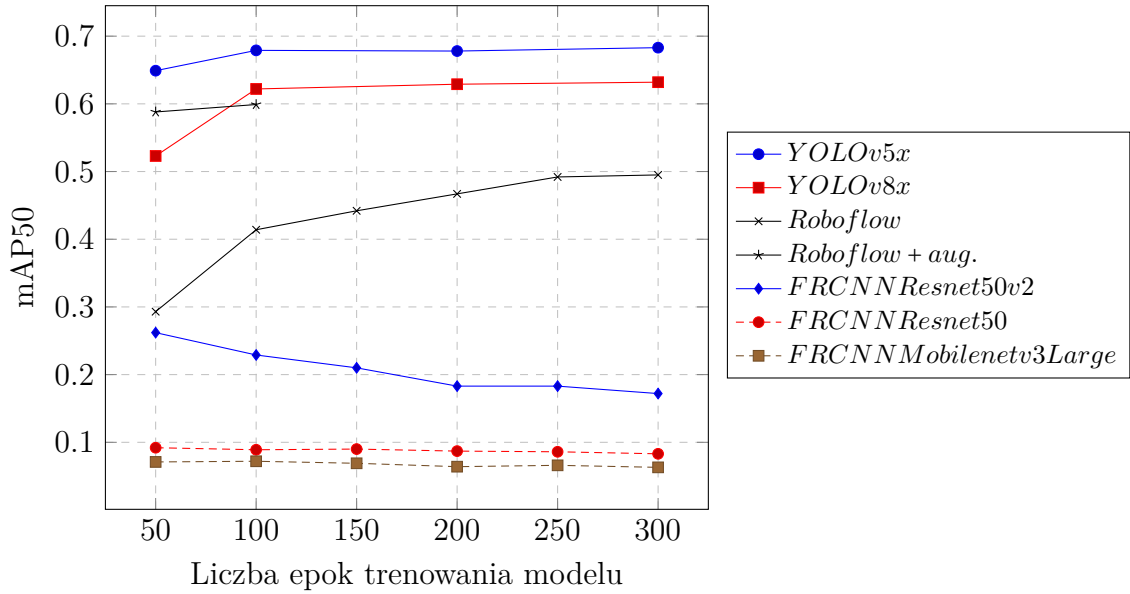
Tabela 4.3: Porównanie skuteczności sieci neuronowych rodzin YOLOv5 i YOLOv8 w zadaniu detekcji obiektów na zdjęciach panoramicznych w formacie ERP.

Model	mAP@50
YOLOv8n	0,579
YOLOv8s	0,607
YOLOv5n	0,611
YOLOv8m	0,615
YOLOv8l	0,637
YOLOv8x	0,644
YOLOv5s	0,646
YOLOv5m	0,675
YOLOv5l	0,685
YOLOv5x	0,683

Aby osadzić wyniki detektorów rodziny YOLO w szerszym kontekście oraz uzyskać wartości bazowe do porównań, przeprowadzono dodatkowe testy z użyciem dwóch innych rozwiązań. Pierwszym z nich było wykorzystanie modelu dostępnego na platformie *Roboflow*. Jako kompleksowe rozwiązanie do prowadzenia projektów, *Roboflow* oferuje zarówno narzędzia do zarządzania zbiorem danych, jak i możliwość trenowania własnych modeli. Po przygotowaniu zbioru danych skorzystano z opcji trenowania z użyciem modelu “*Roboflow 3.0 Object Detection (FAST)*”. Sprawdzono dwie konfiguracje: model trenowany od podstaw, bez użycia techniki rozszerzania zbioru danych oraz model z wykorzystaniem wag wstępnie wytrenowanych na zbiorze *Microsoft COCO* oraz z zastosowaniem pięciu typów manipulacji obrazem: zmianą odcienia szarości, korektą nasycenia (w zakresie $\pm 25\%$), modyfikacją jasności ($\pm 25\%$), wprowadzeniem rozmycia (do 2 pikseli) oraz dodaniem szumu (do 3% pikseli). Dzięki zabiegom tym, liczba przykładów w zbiorze została potrojona. Drugim z rozwiązań do uzyskania odniesienia było użycie innego, gotowego modelu detekcyjnego, tj. sieci neuronowej architektury *Faster-RCNN* [315], opartej na trzech różnych bazowych strukturach sieci wewnętrznych (z ang. *Backbone*): *Resnet50* [316], *Resnet50 v2* [317] oraz *MobileNet v3 Large* [318]. Dla porównania przyjęto stałe warunki testowe, tj. długość treningu wynosząca 300 epok oraz metryka mAP@50. Wyniki przedstawiono na Rys. 4.8.

Model platformy *Roboflow*, przy pracy bez rozszerzenia zbioru danych, osiągnął rezultaty wyraźnie słabsze od rozwiązań opartych na architekturze YOLO i osiągnął mAP@50 równy 0,50. Wprowadzenie zmienności do zbioru zdjęć znacząco poprawiło wyniki, a mAP wzrósł do poziomu 0,62. Pozwoliło to zbliżyć się do najlepszych modeli architektury YOLO, choć wciąż pozostawało poniżej wyników modelu YOLOv5x (0,68) oraz YOLOv8x (0,64). Sieci *Faster-RCNN* osiągnęły znacznie niższe wyniki: mAP@50 na poziomie 0,28 dla szkieletu w postaci sieci *Resnet50v2*, 0,11 dla *Resnet50* oraz 0,08 dla *MobileNetv3Large*.

Jako, że celem prowadzonych badań było oszacowanie sensowności implementacji rozwiązań specjalnie opracowywanych dla obrazów sferycznych, uzyskane wyniki odniesiono również do rozwiązania dedykowanego. W oryginalnej pracy wprowadzającej zbiór *WiMR* [214], najlepszy wynik osiągnięty został przy progu detekcji mAP@30 i wynosił 0,78. Najlepszy wyznaczony model, YOLOv5x, skonfigurowano ponownie i oceniono z użyciem metryki mAP@30.



Rysunek 4.8: Porównanie skuteczności najlepszych modeli architektury YOLO generacji piątej i ósmej oraz innych wybranych modeli detekcji obiektów w zadaniu przetwarzania obrazów typu ERP.

Uzyskany wynik wyniósł 0,69. Różnica między modelem dedykowanym, przygotowanym specjalnie do tego zastosowania, a rozwiązaniem gotowym do użycia wyniosła 9 punktów procentowych. Kosztem dodatkowych narzutów prac, względna poprawa wyniku wyniosła zatem jedynie około 13%. W kontekście prowadzonych prac wdrożeniowych i ich przemysłowego aspektu, uzyskane rezultaty udowadniają, że detektory obiektów stworzone do pracy z obrazami perspektywicznymi dostarczają akceptowalnych wyników również w realiach pracy z obrazami typu ERP. Opracowywanie rozwiązań szytych na miarę nie znajduje zatem uzasadnienia.

Badania opisane w niniejszym rozdziale upublicznione zostały najpierw w ramach artykułu naukowego “*Object Detection with YOLOv5 in Indoor Equirectangular Panoramas*” [319] (70 pkt. MNiSW) i zaprezentowane podczas “*27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)*”, a następnie rozwinięte i opisane w artykule “*YOLO-based object detection in panoramic images of smart buildings*” [320] (140 pkt. MNiSW) zaprezentowanym podczas “*2023 IEEE 10th International Conference on Data Science & Advanced Analytics (DSAA)*”.

4.2.2 Algorytm doboru punktu rozcięcia sfery

Obrazy sferyczne, w swojej oryginalnej reprezentacji, stanowią pewną zapętloną formę, tj. pozbawione są pojęcia krawędzi, nie posiadają początku ani końca. Taka postać ma swoje zalety, przede wszystkim w wizualizacjach przestrzeni, nie jest jednak zgodna z klasycznymi architekturami sieci neuronowych i prowadzi do trudności z jej wykorzystaniem w powszechnie dostępnych rozwiązaniach. Aby przekształcić sferę do panoramy, niezbędne jest wybranie półokręgu jej rozcięcia. W niniejszej pracy określa się go jako “południk początkowy”, a wstępnie zaprezentowany został w Roz. 4.1.3. Jego dobór w sposób przypadkowy doprowadzić może do niespodziewanych skutków ubocznych i trudności, jak miało to miejsce w opisanych badaniach (Roz. 4.2.1), gdzie rzutowanie na obraz perspektywiczny doprowadziło do przecięcia obiektów umieszczonych na krawędziach obrazu. Przemyślany dobór jest o tyle istotny, że oddziaływać może na końcową skuteczność detektorów obiektów. Rozcięte i rozrzucone fragmenty obiektów są trudne w rozpoznaniu nawet dla człowieka. Aby temu problemowi zaradzić i przeanalizować jego faktyczny wpływ na uzyskiwane wyniki, postanowiono o sformalizowaniu algorytmu doboru najlepszego punktu na rozcięcie sfery i przeprowadzeniu badań jego wpływu na skuteczność detekcji obiektów.

Proponowana metoda

Proponowana *Metoda Doboru Najlepszego Punktu Rozcięcia* (z ang. *Best Cutting Point Selection Method*, BCPSM) to autorska propozycja algorytmu postępowania. Ma ona na celu takie wyznaczenie punktu rozcięcia panoramy (południka początkowego), które ograniczy efekt przecinania obiektów umieszczonych na krawędzi uzyskanego obrazu. Spodziewanym rezultatem jego zastosowania jest poprawa skuteczności detektorów obiektów, a także stabilizacja ich wyników. Metoda opiera się na analizie liczby obiektów (a dokładniej ramek je ograniczających), które zostają przecięte pionową linią w danym punkcie poziomej osi obrazu, a także wagi informacji w niej zawartej. Zakładamy, że najlepszą kolumną do rozcięcia sfery jest taka kolumna, która ma najmniejszą sumaryczną wagę przeciętych etykiet obiektów, określaną dla poszczególnych pikseli z użyciem wagi opartej na funkcji cosinus. Kroki postępowania opisane zostają dalej w tym rozdziale.

Krok 0: Dane wejściowe. W ramach proponowanej metody przetwarzaniu podlega pojedynczy obraz $\mathbf{o}_{erp} \in \mathbf{O}_{erp}$ (przedstawiony na Rys. 4.9a) oraz towarzyszące mu adnotacje obiektów $\mathbf{A}_{erp}$, które można na nim wykryć.

Definicja 2. Niech $\mathbf{O}_{erp}$ będzie zbiorem kolorowych obrazów panoramicznych (paleta barw RGB) takich, że:

$$\forall \mathbf{o}_{erp} \in \mathbf{O}_{erp}, \quad \mathbf{o}_{erp} \in \{0, \dots, 255\}^{H_{erp} \times W_{erp} \times C_{erp}},$$

gdzie H_{erp} i W_{erp} to odpowiednio wysokość i szerokość obrazu panoramicznego, wyrażone w pikselach, a C_{erp} to liczba kanałów kolorów.

Definicja 3. Na podstawie [321], niech $\mathbf{A}_{erp} = \{a_{erp1}, a_{erp2}, \dots, a_{erpN_{erp}}\}$ będzie zbiorem o liczebności N_{erp} , adnotacji odpowiadających obiektom obecnym na obrazie, gdzie każda $\mathbf{i}_{erp}$ -ta adnotacja ma postać:

$$a_{erp_i} = (x_{erp_i}, y_{erp_i}, w_{erp_i}, h_{erp_i}) \in [0, 1]^4.$$

Współrzędne (x_{erp_i}, y_{erp_i}) oznaczają środek obiektu w układzie współrzędnych normalizowanym względem szerokości i wysokości obrazu, zaś w_{erp_i} i h_{erp_i} odpowiadają odpowiednio znormalizowanej szerokości i wysokości ramki ograniczającej obiekt.

Krok 1: Mapa adnotacji. Celem pierwszego kroku jest utworzenie maski, która dla każdego piksela przetwarzanego obrazu zlicza liczbę adnotacji obejmujących dany punkt. Na podstawie Def. 2 i 3, dla każdej adnotacji $\mathbf{a}_{erp_i} \in \mathbf{A}_{erp}$, znajdującej się na obrazie $\mathbf{o}_{erp} \in \mathbf{O}_{erp}$, niech $\mathbf{R}_{erp_i}$ oznacza odpowiadający jej prostokąt, o współrzędnych takich że:

$$\begin{aligned} \mathbf{R}_{erp_i} = & \left(\left\lfloor \left(x_{erp_i} - \frac{w_{erp_i}}{2} \right) \cdot W_{erp} \right\rfloor, \left\lfloor \left(x_{erp_i} + \frac{w_{erp_i}}{2} \right) \cdot W_{erp} \right\rfloor \right) \times \\ & \times \left(\left\lfloor \left(y_{erp_i} - \frac{h_{erp_i}}{2} \right) \cdot H_{erp} \right\rfloor, \left\lfloor \left(y_{erp_i} + \frac{h_{erp_i}}{2} \right) \cdot H_{erp} \right\rfloor \right), \end{aligned} \quad (4.10)$$

opisany funkcją charakterystyczną:

$$F_{\text{adn}}^{(i_{erp})}(x_{erp}, y_{erp}) = \begin{cases} 1 & \text{jeśli } (x_{erp}, y_{erp}) \in \mathbf{R}_{erp_i}, \\ 0 & \text{w przeciwnym razie.} \end{cases} \quad (4.11)$$

Maskę adnotacji $\mathbf{M}_{erp_a} : \{0, \dots, H_{erp}-1\} \times \{0, \dots, W_{erp}-1\} \rightarrow \mathbb{N}_0$ definiujemy wówczas jako sumę wartości funkcji $F_{\text{adn}}^{(i_{erp})}$, liczonej po wszystkich adnotacjach:

$$\mathbf{M}_{erp_a}(y_{erp}, x_{erp}) = \sum_{i_{erp}=1}^{N_{erp}} F_{\text{adn}}^{(i_{erp})}(x_{erp}, y_{erp}). \quad (4.12)$$

Przykładową wizualizację maski adnotacji przedstawiono na Rys. 4.9b.

Krok 2: Mapa wag. Dla każdej wartości y_{erp} , tj. każdego wiersza obrazu, obliczana jest wartość wagowa na podstawie funkcji cosinus, która większe wagi przypisuje wierszom pikseli położonym bliżej środka obrazu, co odzwierciedla założenie, że zawierają one mniej zniekształconych informacji wizualnych:

$$w_{erp}(y_{erp}) = \cos\left(-\frac{y_{erp} - H_{erp}/2}{H_{erp}} \cdot \pi\right). \quad (4.13)$$

Następnie, dla każdego piksela w tym wierszu przypisujemy obliczoną wartość, tworząc macierz wag $\mathbf{W}_{erp} \in \mathbb{N}^{H_{erp} \times W_{erp}}$ taką, że:

$$\mathbf{W}_{erp}[y_{erp}, x_{erp}] = w_{erp}(y_{erp}) \quad \forall x_{erp} \in \{0, \dots, W_{erp}-1\},$$

Przykładową wizualizację wyznaczonej maski wag przedstawiono na Rys. 4.9c.

Krok 3: Ważona mapa adnotacji. Uzyskaną wcześniej maskę adnotacji $\mathbf{M}_{erp_a}$ łączymy z mapą wag $\mathbf{W}_{erp}$ poprzez elementowy iloczyn Hadamarda:

$$\mathbf{M}_{erp_w} = \mathbf{M}_{erp_a} \odot \mathbf{W}_{erp}. \quad (4.14)$$

Macierz $\mathbf{M}_{erp_w} \in \mathbb{R}^{H_{erp} \times W_{erp}}$ zawiera wartości odpowiadające ważonej liczbie adnotacji przypisanych do każdego piksela obrazu. Wizualizacje utworzonej maski przedstawiono w postaci znormalizowanej kolorystycznie na Rys. 4.9d oraz z podkreśleniem zmiany wag wierszy obrazu na Rys. 4.9e.

Krok 4: Sumaryczna waga kolumn. Sumujemy wartości w każdej kolumnie zważonej maski $\mathbf{M}_{erp_w}$, uzyskując jednowymiarowy wektor wag kolumnowych $\mathbf{c}_{erp} \in \mathbb{R}^{W_{erp}}$:

$$\mathbf{c}_{erp}[x_{erp}] = \sum_{y_{erp}=0}^{H_{erp}-1} \mathbf{M}_{erp_w}[y_{erp}, x_{erp}], \quad \forall x_{erp} \in \{0, \dots, W_{erp}-1\}. \quad (4.15)$$

Krok 5: Grupowanie kolumn o tej samej wadze. Dzielimy zbiór kolumn na nieprzecinające się przedziały o stałej wartości wagowej:

$$\mathcal{R} = \left\{ R_{erp_k} = [x_{erp_k}^{(1)}, x_{erp_k}^{(2)}] \mid \forall x_{erp} \in R_{erp_k}, \mathbf{c}_{erp}[x_{erp}] = w_{erp_k} \right\}.$$

Spośród tych przedziałów wybieramy taki, który charakteryzuje się najmniejszą wagą w_{erp_k} , a w przypadku remisów wybieramy ten o największej długości. Formalnie:

$$R_{erp}^* = \arg \min_{R_{erp_k} \in \mathcal{R}} (w_{erp_k}, -|R_{erp_k}|). \quad (4.16)$$

gdzie $|R_{erp_k}| = x_{erp_k}^{(2)} - x_{erp_k}^{(1)}$ oznacza długość przedziału R_{erp_k} .

Wizualizacja uzyskanych wartości kolumn na przykładowej panoramie prezentowana jest na Rys. 4.9f.

Krok 6: Obliczenie południka rozcięcia. Na podstawie wybranego przedziału $R_{erp}^* = [x_{erp_{start}}^*, x_{erp_{end}}^*]$ obliczana jest (jako jego środek) ostateczna kolumna rozcięcia:

$$x_{erp_{best}} = \left\lfloor \frac{x_{erp_{start}}^* + x_{erp_{end}}^*}{2} \right\rfloor. \quad (4.17)$$

Kolumna x_{best} określa południk początkowy, względem którego wykonane zostanie przekształcenie panoramiczne.

Wynik. Algorytm zwraca indeks x_{best} jako najlepszy sugerowany południk rozpoczęcia rzutowania ze sfery na panoramę. Wariant alternatywny tej metody pozwala na wyznaczenie tzw. najgorszego punktu rozcięcia. Celem jest wtedy prze-

cięcie obrazu w miejscu, które generuje największe zniekształcenia, czyli przecina jak najwięcej obiektów o dużym znaczeniu detekcyjnym. Spodziewanym zastosowaniem takiego podejścia jest analiza odporności detektorów obiektów na warunki skrajne, typowe dla przetwarzania danych panoramicznych. Obraz wczytany do stworzonego narzędzia wizualizacji, wraz z naniesioną sugestią najlepszego punktu rozcięcia oraz aktualnie wybranym południkiem, prezentowany jest na Rys. 4.9g, a efekt przesunięcia rozcięcia i adnotacji obiektów na Rys. 4.9h.

Zastosowane strategie cięcia panoram i wykonane eksperymenty

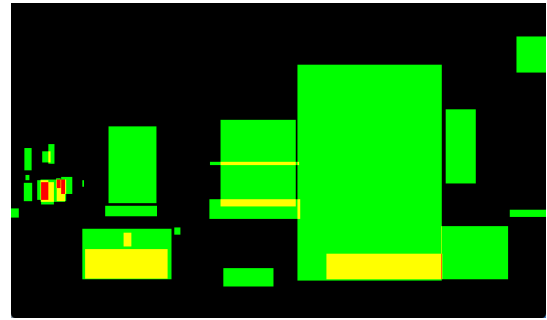
Zbiór danych *EquiBEB*, opracowany autorsko i wykorzystany na potrzeby niniejszych badań, omówiony zostaje w rozdziale 4.2.3. Na jego podstawie przygotowano sześć zestawów danych, każdy z zastosowaniem innej metody doboru punktu rozcięcia sfery. Były to:

- automatyczny najlepszy punkt cięcia (**AutoMin**),
- ręcznie wybrany najlepszy punkt cięcia (**ManualMin**),
- losowy punkt cięcia (**Random**),
- domyślny punkt cięcia (**Default**),
- ręcznie wybrany najgorszy punkt cięcia (**ManualMax**),
- automatyczny najgorszy punkt cięcia (**AutoMax**).

Metody automatyczne (AutoMin, AutoMax) bazują na proponowanej metodzie *BCPSM*. Metody ręczne (ManualMin, ManualMax) wynikają z subiektywnej oceny eksperta. W pierwszym przypadku (ManualMin) użytkownik wybiera punkt cięcia, który jego zdaniem najlepiej zachowuje spójność obiektów i sprzyja nauce modelu. W drugim przypadku (ManualMax) celem jest wybór punktu, który celowo utrudnia analizę, np. przecina wiele małych obiektów. Takie podejście, choć czasochłonne, pozwala na konfrontację wyników uzyskanych intuicyjnie z efektami podejścia algorytmicznego. W przypadku metody losowej (Random), punkt cięcia wybierany jest losowo z rozkładu jednostajnego. Proces ten powtarzany jest pięć razy, a wyniki walidacji i testów prezentowane są jako średnie oraz odchylenia standardowe. Strategia domyślna (Default) oznacza brak jakiegokolwiek przetwarzania, tj. wykorzystano w niej oryginalne panoramy ze zbioru danych. Na Rys. 4.10



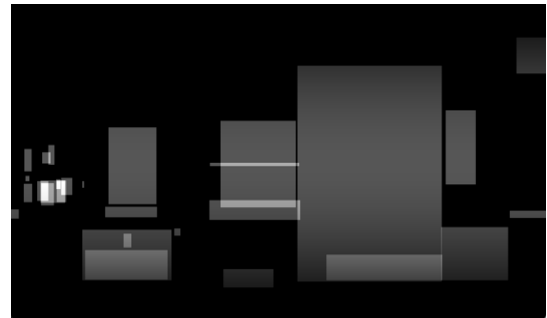
(a) Zdjęcie źródłowe



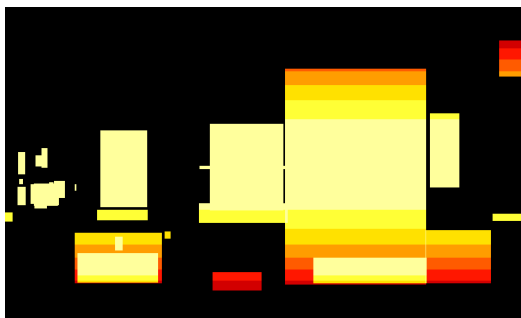
(b) Maska adnotacji



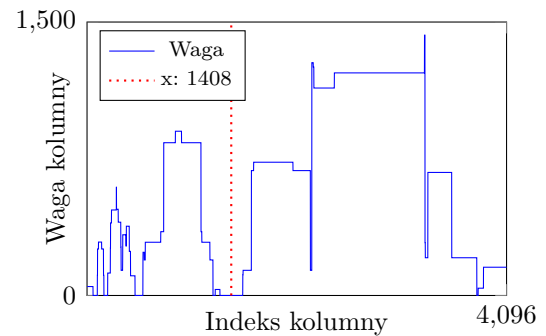
(c) Maska wagi pikseli



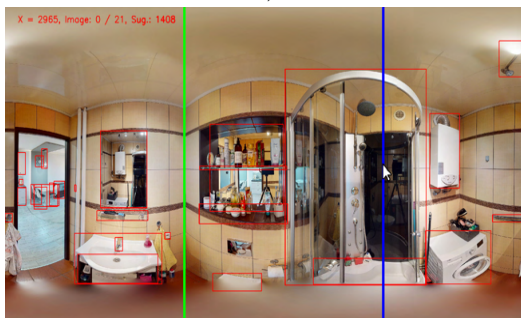
(d) Znormalizowana maska ważonych adnotacji



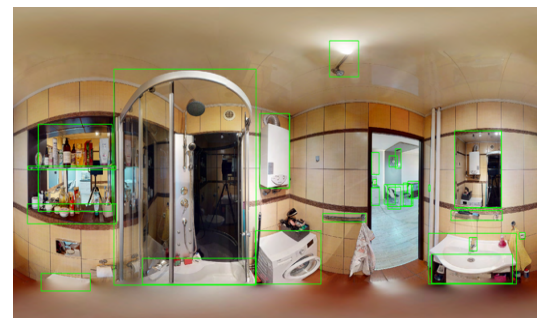
(e) Kolorowa maska ważonych adnotacji (do celów wizualizacji)



(f) Wykres wag kolumn



(g) Sugerowany najlepszy punkt rozcięcia (zielona linia) i aktualnego położenia kursora myszki (niebieska linia)



(h) Przetworzona panorama

Rysunek 4.9: Przykład procesu przetwarzania obrazu panoramicznego z proponowanego zbioru *EquiB&B*.

przedstawiono przykłady panoram przetworzonych różnymi metodami. Zaobserwować można zmianę położenia obiektów, a dla przypadku metody najgorszego cięcia istotne obiekty, takie jak stół, szafka czy lampa, są przecinane i umieszczane na krawędziach obrazu.

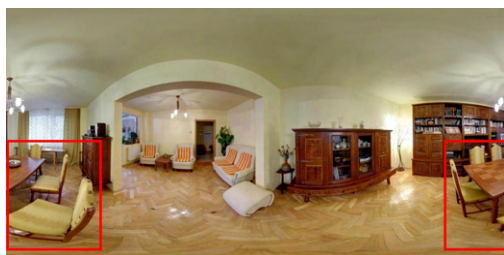
Do detekcji obiektów wykorzystano sieci neuronowe architektury YOLO. Na podstawie wyników badań wstępnych (Roz. 4.2.1) do użycia wytypowano trzy modele generacji piątej: *Nano*, *Medium* i *Extra-Large*. Scenariusze testowe opracowano jako wyuczenie każdego z modeli na każdym z przygotowanych zbiorów, a następnie testy krzyżowe każdej konfiguracji.



(a) Automatyczny najlepszy punkt cięcia



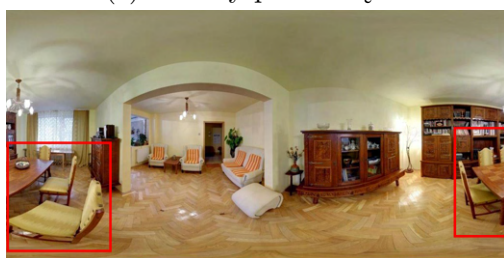
(b) Ręcznie wybrany najlepszy punkt cięcia



(c) Losowy punkt cięcia



(d) Domyślny punkt cięcia



(e) Ręcznie wybrany najgorszy punkt cięcia



(f) Automatyczny najgorszy punkt cięcia

Rysunek 4.10: Przykładowe obrazy (ze zbioru WiMR [322]) przetworzone przy użyciu kolejnych metod doboru południka rzutowania.

Wyniki eksperymentów i wnioski

W Tab. 4.4 i Tab. 4.5 zaprezentowano wyniki metryki mAP uzyskane dla dwóch zbiorów danych, kolejno: *WiMR* oraz *EquiB&B*. Przedstawiono wyniki fazy walidacji (na zbiorach uczących), a także testowania (na zbiorach testowych) dla wszystkich konfiguracji cięcia panoramy. Dla metody losowej podano średnie wartości z pięciu powtórzeń wraz z odchyleniem standardowym.

Tabela 4.4: Wyniki miary mAP50 dla zbioru danych *WiMR* dla kolejnych konfiguracji eksperymentalnych. Najlepsze wyniki w pionie pogrubiono, a najlepsze w poziomie podkreślono.

Model	Metoda	Walidacja	Testowanie					
	cięcia		AutoMin	ManualMin	Default	Random	ManualMax	AutoMax
Extra-large	AutoMin	67.2	65.3	64.4	64.2	<u>65.2±0.8</u>	64.5	63.0
	ManualMin	66.6	63.6	63.0	63.8	64.7±0.5	63.7	61.4
	Default	64.3	62.3	62.3	62.7	<u>63.4±0.7</u>	63.2	60.2
	Random	64.4±0.5	60.5±0.5	60.1±0.2	<u>61.9±0.7</u>	61.3±0.5	61.7±0.4	59.5±0.7
	ManualMax	61.5	55.6	55.1	56.6	58.2±0.8	<u>59.9</u>	57.0
	AutoMax	58.9	53.5	54.4	55.7	56.8±0.7	<u>57.7</u>	55.5
Medium	AutoMin	64.7	63.4	61.3	61.4	60.8±0.9	60.7	59.4
	ManualMin	63.8	61.1	59.3	<u>61.4</u>	60.6±0.9	59.7	60.1
	Default	62.8	<u>59.6</u>	59.2	58.2	58.7±0.7	57.5	58.9
	Random	64.4±0.5	60.5±0.5	60.1±0.2	<u>61.9±0.7</u>	61.3±0.5	61.7±0.4	59.5±0.7
	ManualMax	58.4	51.6	53.0	<u>55.0</u>	54.1±0.6	54.3	54.2
	AutoMax	56.3	51.5	51.2	53.7	52.9±0.8	<u>54.0</u>	52.6
Nano	AutoMin	51.3	<u>49.6</u>	48.4	47.7	47.0±1.3	45.2	44.5
	ManualMin	48.6	<u>47.5</u>	47.0	46.9	47.2±1.1	45.4	45.5
	Default	47.8	45.1	<u>45.2</u>	45.0	45.1±0.9	44.4	43.4
	Random	46.3±0.6	43.4±0.6	42.8±0.4	<u>43.8±1.2</u>	43.4±0.5	42.8±0.9	42.6±0.4
	ManualMax	42.7	37.4	37.8	38.9	<u>39.4±1.8</u>	39.0	37.9
	AutoMax	41.6	38.1	38.2	38.7	39.5±1.2	38.9	<u>39.7</u>

Analizę wyników rozważać można w dwóch wymiarach, tj. pionowo i poziomo. Analiza pionowa przedstawia zmianę wyników na skutek trenowania modeli na kolejnych zbiorach danych. Uzyskujemy w ten sposób wiedzę o procesie uczenia się sieci. W większości scenariuszy najwyższe wartości mAP uzyskano z wykorzystaniem automatycznej metody doboru najlepszego południka rozcięcia. Potwierdza to sensowność przetwarzania wstępnego używanych zbiorów danych i skuteczność proponowanej metody *BCPSM*. Analiza pozioma pokazuje, jak dana konfiguracja treningowa (model i zbiór) radzi sobie z różnymi scenariuszami testowymi. Pokazują one, że w zdecydowanej większości przypadków metoda automatycznego najgorszego rozcięcia (**AutoMax**) pozwalała wygenerować najtrudniejsze zestawy

Tabela 4.5: Wyniki miary mAP50 dla zbioru danych *EquiBEB* dla kolejnych konfiguracji eksperymentalnych. Najlepsze wyniki w pionie pogrubiono, a najlepsze w poziomie podkreślono.

Model	Metoda		Testowanie					
	cięcia	Walidacja	AutoMin	ManualMin	Default	Random	ManualMax	AutoMax
Extra-large	AutoMin	82.1	87.0	86.3	83.9	85.0±0.9	85.8	79.6
	ManualMin	81.0	82.7	<u>85.3</u>	83.9	85.1±1.1	82.1	83.6
	Default	81.7	83.6	<u>85.5</u>	82.9	84.5±1.5	82.5	83.2
	Random	82.6 ±0.6	<u>84.3±0.6</u>	83.9±1.3	83.4±1.8	83.9 ±1.2	82.5±1.5	82.2±1.7
	ManualMax	74.7	<u>78.1</u>	75.6	75.0	77.3±1.5	76.2	79.7
	AutoMax	76.3	81.4	80.3	79.7	80.5±1.1	79.3	<u>84.4</u>
Medium	AutoMin	76.2	81.1	78.8	78.5	82.5±2.2	78.6	74.2
	ManualMin	73.0	82.0	78.7	78.3	80.3±1.4	76.5	75.2
	Default	77.1	<u>81.2</u>	78.3	78.2	80.3±1.6	75.9	74.8
	Random	77.5±2.2	79.4±2.2	77.3±2.7	79.9±2.1	76.7±2.0	78.2±2.0	73.9±0.9
	ManualMax	69.9	71.6	<u>76.6</u>	74.5	75.4±0.8	70.3	72.5
	AutoMax	70.6	75.2	76.6	71.9	<u>77.1±3.3</u>	74.3	72.1
Nano	AutoMin	54.5	56.9	<u>60.9</u>	49.5	56.5±4.0	52.9	47.7
	ManualMin	51.0	57.0	58.2	55.4	<u>58.4±2.8</u>	53.6	48.0
	Default	49.0	58.0	61.0	53.9	58.9±2.9	51.1	48.1
	Random	51.3±1.4	60.7±1.4	58.7±1.1	55.9±3.6	56.9±0.4	49.5±2.7	45.6±1.1
	ManualMax	41.8	<u>57.0</u>	53.1	46.5	53.2±3.6	46.4	40.3
	AutoMax	35.0	<u>62.0</u>	53.9	48.5	54.4±3.2	47.6	39.3

testowe. Skuteczność detekcji była w nich najniższa. Warto dodać, że różnice w wynikach mAP pomiędzy **AutoMin** i **AutoMax** były tym większe, im mniejszy był użyty model, tj. mniejsze sieci były bardziej wrażliwe na jakość danych testowych.

Eksperymenty pokazały, że stosowanie strategii cięcia w najlepszym możliwym punkcie poprawia proces uczenia sieci i pozwala uzyskać wyższą skuteczność detekcji. Proponowana metoda automatyczna okazała się lepsza lub porównywalna z rozwiązaniem ręcznym, a jednocześnie wymagała mniej czasu i zaangażowania użytkownika. Pozwala to określić ją jako zdatną i wartościową do zastosowania, ponieważ nie wymaga istotnego zaangażowania użytkownika, a prowadzi do poprawy uzyskiwanych wyników detekcji obiektów. Opracowany system przewijania panoram stanowi również użyteczne narzędzie do ich wizualizacji oraz dalszej modyfikacji, np. celem uwspólnienia orientacji zdjęć.

Potwierdzono tym samym drugą hipotezę dodatkową rozprawy i pokazano, że przetwarzanie wstępne zbiorów zdjęć, polegające na doborze najlepszego południka rozcięcia sfery, pozwala na zwiększenie efektyw-

ności wyszkolonego modelu detektora obiektów i jego działania w środowisku domowym.

Badania opisane w niniejszym rozdziale upublicznione zostały w ramach artykułu naukowego “*Finding the perfect cut: selection of the best cutting point in equirectangular panoramas for object detection*” [323] (70 pkt. MNiSW) i zaprezentowane podczas “*28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)*”.

4.2.3 Oryginalne zbiory danych

Przygotowanie rozwiązania wizyjnego, które dobrze odpowiada na zidentyfikowane potrzeby użytkowników (Roz. 2.2) wymaga rozbudowanych zbiorów danych o odpowiedniej tematyce. Jak wykazano poprzez rozeznanie literatury (Roz. 2.1) oraz analizę samych zbiorów (Roz. 4.1.4), dostęp do satysfakcjonująco dużych, zgodnych w reprezentacji i rodzaju scen, a przede wszystkim o licencji do użytku komercyjnego, zbiorów obrazów jest ograniczony. Postanowiono zatem o pozyskaniu takich danych samodzielnie. Dalej w tym rozdziale przedstawione zostają zgromadzone zbiory zdjęć własnych, pozyskane różnymi urządzeniami, w różnych scenariuszach, zawsze etykietowane jednak taką samą pulą klas obiektów do analizy, tj. autorskim zestawem *CODE55* (Roz. 2.2.4). Pod taką też nazwą funkcjonować będzie końcowy zbiór danych, opracowany na potrzeby całej rozprawy i stworzonego systemu fuzyjnego. Dla zachowania prywatności osób udostępniających swoje nieruchomości do badań, ze zbiorów prezentowane są tylko przykładowe, subiektywnie wybrane obrazy o celowo obniżonej jakości.

Zmodyfikowany zbiór PANDORA

W pierwszej kolejności prace skupiono na stworzeniu dodatkowych etykiet obiektów dla dostępnego publicznie zbioru *PANDORA* [215], który wykorzystywany był na etapie badań wstępnych i rozeznania zagadnienia. Podczas prac przeprowadzono dodatkowe operacje porządkujące, tj. usunięto wszelkie wykryte duplikaty zdjęć, a także rozdzielono cały zbiór na dwa podzbiory: *REAL* i *SYNT*. Zbiór *REAL* liczy 1205 obrazów rzeczywistych, a zbiór *SYNT* to 255 obrazów wygenerowanych syntetycznie. Łącznie w zbiorze tym oznaczono ponad 53 000 obiek-

tów zgodnych z etykietami systemu *CODE55*. Przykładowe obrazy ze zbioru *REAL* przedstawiono na Rys. 4.11, a ze zbioru *SYNT* na Rys. 4.12.



Rysunek 4.11: Przykładowe obrazy ze zbioru *PANDORA REAL*.



Rysunek 4.12: Przykładowe obrazy ze zbioru *PANDORA SYNT*.

Zbiór zdjęciowy PanoSilesia

Pierwszym zestawem danych, który zgromadzono w całości autorsko, były zdjęcia sferyczne wykonane kamerą *Ricoh Theta SC2*. Materiał wizyjny zgromadzony został w trzech kolejnych nieruchomościach, każdorazowo w Polsce, na terenie Śląska. Były to: dom rekreacyjny w obszarze górskim (podzbiór *Resort*), piętrowy dom jednorodzinny (podzbiór *House*) oraz kilkupokojowe mieszkanie w zabudowie wielorodzinnej (podzbiór *Apartment*). Najliczniejszy podzbiór to *Resort*, stanowiący 386 obrazów, na których oznaczonych zostało ponad 11 000 obiektów (Rys. 4.13). Drugi co do wielkości jest podzbiór *House*, tj. 122 obrazy i około 5500 oznaczonych obiektów (Rys. 4.14). Najmniejszy jest podzbiór *Apartment*, w którym odnajdziemy 102 zdjęcia i około 5000 obiektów (Rys. 4.15). Łącznie na zbiór *PanoSilesia* składa się 610 obrazów IERP, na których oznaczono ponad 21 000 obiektów.

Zbiór skanów trójwymiarowych

Motywy przewodnim i jednym z zagadnień całej rozprawy są spacer trójwymiarowe po nieruchomościach. Zawierają one nie tylko pomiary laserowe, ale rów-



Rysunek 4.13: Przykładowe obrazy ze zbioru PanoSilesia RESORT.



Rysunek 4.14: Przykładowe obrazy ze zbioru PanoSilesia HOUSE.

niez obrazy sferyczne, które gromadzone są podczas digitalizacji przestrzeni. Wykorzystano dwa kolejne takie modele, z których wyeksportowano zbiory zdjęć i utworzono dwa dodatkowe zestawy danych: *Office* oraz *EquiB&B* (wspomniany w Roz. 4.2.2). Obrazy zostały zarejestrowane kamerą *Matterport Pro2* i wyeksportowane jako panoramy ERP. W podzbiorze *Office* znalazło się 36 zdjęć nieruchomości stanowiącej biuro kancelarii prawnej (Rys. 4.16). Oznaczono na nich ponad 1600 obiektów. Podzbiór *EquiB&B* prezentuje małe mieszkanie pod wynajmem krótkoterminowy (Rys. 4.17). Zawiera 22 obrazy oraz ponad 900 oznaczonych na nich obiektów.

Zbiór badawczy SPHEREA

Tematyka generatywnych sieci neuronowych, które pozwalają tworzyć nowe obrazy czy to na podstawie zapytań tekstowych, czy na bazie przekazanych zdjęć, zyskuje popularność. Postanowiono więc nie tylko stworzyć taki autorski zbiór syntetyczny, ale również przeprowadzić jego badania. Analizowano zdatność zbiorów sztucznych do szkolenia modeli detektorów obiektów. Proponowany zbiór, nazwany **SPHEREA** (od ang. **S**ynthetic **P**anoramas for **H**ome **E**nvironments and **R**esearch in **E**quirectangular **A**pplications), wygenerowano z użyciem architektury dużych modeli generatywnych, konkretnie DALL-E 3 [324]. Do badań wybrano pięć rodzajów pomieszczeń domowych, na których oparto zapytania: *Salon*,



Rysunek 4.15: Przykładowe obrazy ze zbioru PanoSilesia APARTMENT.



Rysunek 4.16: Przykładowe obrazy ze zbioru OFFICE.

Kuchnia, Łazienka, Sypialnia oraz *Biuro*. Połączono je z opracowanym zbiorem etykiet *CODE55* (Roz. 2.2) i przygotowano dwa schematy generowania zapytań, efektywnie dwa podzbiory obrazów pomieszczeń domowych:

- **ITEMS**, gdzie polecenie zostało wzbogacone o jawną listę przedmiotów, które powinny znaleźć się na wygenerowanym obrazie,
- **NO-ITEMS**, gdzie modelowi pozostawiono pełną swobodę doboru elementów wyposażenia typowego dla danego rodzaju pomieszczenia.

Na potrzeby porównań ilościowych w stosunku do zbiorów rzeczywistych, zaproponowano wykorzystanie fragmentu zbioru *REAL* z *PANDORA*, omówionego już w Roz. 4.2.3. Aby zapewnić rzetelność porównania, ze zbioru tego wybrano po 30 zdjęć o największej liczbie oznaczonych obiektów dla każdego z pięciu rodzajów pomieszczeń, tj. łącznie 150 zdjęć, taką też liczbę obrazów wygenerowano dla każdego ze schematów generacji.

Wstępne testy wygenerowanych obrazów ujawniły szereg niedoskonałości, szczególnie błędne przekształcenia geometryczne: nieodpowiednią perspektywę widoku (Rys. 4.18a), efekt typu “rybie oko” (Rys. 4.18b) lub brak pełnego pokrycia obrazu (Rys. 4.18c). Odpowiednie poprawki wprowadzono do treści zapytań generujących obrazy, a każdy przykład weryfikowano dodatkowo manualnie. Odrzucano obrazy, które przeczyły ludzkiej intuicji i logice, przykładowo zawierające nierealne obiekty (pralka wewnątrz lodówki wewnątrz drzwi, Rys. 4.18d), zbyt wiele nietypowych



Rysunek 4.17: Przykładowe obrazy ze zbioru EquiB&B.

obiektów (np. siedem pralek w salonie, Rys. 4.18e) lub o zaburzonych rozmiarach (np. pralka rozmiaru zabawki, Rys. 4.18f). Dla większej heterogeniczności zbioru każde zapytanie zawierało dodatkowe parametry sceny: porę dnia (dzień albo noc) oraz styl aranżacji wnętrza.



(a) Zła perspektywa.

(b) Efekt “rybiego oka”.

(c) Brak pełnego pokrycia.



(d) Nierealny obiekt (pralka wewnątrz lodówki wewnątrz drzwi).

(e) Zbyt wiele nietypowych obiektów (siedem pralek).

(f) Zbyt mały obiekt (pralka).

Rysunek 4.18: Przykłady błędnie wygenerowanych obrazów odrzuconych z procesu tworzenia zbioru *SPHEREA*.

Na tak przygotowanym zbiorze danych przystąpiono do jego analizy i badań. Rozpoczęto od porównania zgodności zawartości zbiorów syntetycznych w stosunku do zbioru rzeczywistego, głównie pod względem statystyki liczby obiektów w nich zawartych. Zbadano, czy wzbogacenie zapytań o wymagania użytkowników może poprawić jakość generowanych danych i zbliżyć je do rzeczywistych scen.

W analizie uwzględniono parametry, takie jak:

- całkowita liczba obiektów w zbiorze oraz średnia i odchylenie standardowe

- liczby obiektów na obraz (niezależnie od typu obiektu i pomieszczenia),
- liczba obiektów dla każdego typu pomieszczenia (bez rozróżniania obiektów),
- liczba konkretnych obiektów w zbiorze oraz ich średnia liczba i zmienność na obraz (niezależnie od typu pomieszczenia),
- liczba i zmienność konkretnych obiektów przypisanych do konkretnych pomieszczeń,
- rozkład przestrzenny obiektów w panoramach,
- rozkład rozmiarów obiektów w obrazach.

W Tab. 4.6 pokazano, że wzbogacenie zapytań o listę pożądaných obiektów skutkuje wygenerowaniem o około 7% (618) więcej obiektów (zbiór *ITEMS* w porównaniu do wariantu *NO-ITEMS*). Zbiór ten jest także subiektywnie bliższy danym rzeczywistym (REAL), zachowując przy tym mniejszą zmienność.

Tabela 4.6: Statystyki ogólne opracowanych zbiorów syntetycznych oraz zbioru referencyjnego.

Zbiór danych	Liczba obiektów	Średnia na obraz	Odchylenie std.
ITEMS	9296	61,97	20,28
NO-ITEMS	8678	57,85	22,53
REAL	9430	62,87	25,82

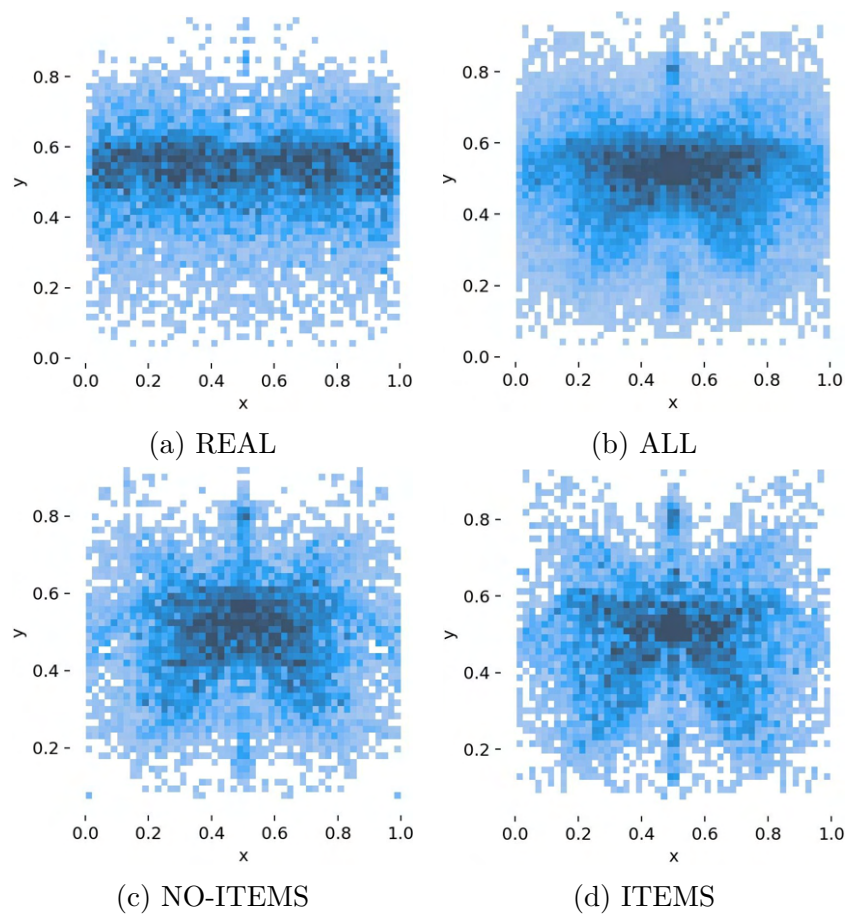
W ramach towarzyszącej badaniom publikacji naukowej przygotowane zostały tabelaryczne analizy liczby obiektów w konkretnych rodzajach pokoi, a także rozkład liczby obiektów dla poszczególnych klas. Dla uproszczenia, w rozprawie zostają one jedynie krótko omówione, bez prezentacji w całości.

Największe różnice w liczbie obiektów pomiędzy zbiorem sztucznym a prawdziwym zostały zaobserwowane dla pomieszczeń typu *Kuchnia* i *Biuro*. W przypadku kuchni syntetyczne zbiory zawierają więcej obiektów głównie ze względu na obecność dodatkowych elementów dekoracyjnych. Dla biur główną różnicą jest brak danych przedstawiających stricte domowe gabinety w zbiorze REAL. Zamiast tego dominują tam zdjęcia z sal konferencyjnych i przestrzeni korporacyjnych, które zawierają więcej obiektów niż typowe pokoje domowe. Ogólnie generatywny model DALL · E ma tendencję do sugerowania bardziej dekoracyjnych przestrzeni, a najczęściej występującymi na zdjęciach obiektami były: *Roślina*, *Wazon* i *Obraz*.

Szczegółowa analiza pokazuje, że wzbogacenie zapytań o konkretne listy obiektów poprawia reprezentatywność typowych przedmiotów w odpowiednich rodzajach pomieszczeń (np. *Ekspres do kawy* w *Kuchni*, *Kominek* w *Salonie*, *Prysznic* w *Łazience*), co czyni zbiór *ITEMS* bliższym rzeczywistym środowiskom domowym.

Na Rys. 4.19 przedstawiono graficznie analizę położenia oznaczonych na panoramach obiektów. Przedstawiono cztery dwuwymiarowe histogramy, po jednym dla każdego ze zbiorów: *ITEMS*, *NO-ITEMS*, *REAL* oraz dodatkowy wynik zbiorczy (*ALL*), dla zbioru utworzonego z połączenia wszystkich wymienionych. Rozkład w zbiorze *REAL* pokazuje charakterystyczny, poziomy układ obiektów, odzwierciedlający ciągłość obrazu panoramicznego oraz znikomą liczbę obiektów umieszczanych na sufitach i bezpośrednio pod nogami fotografa. Dane syntetyczne prezentują układ zauważalnie bardziej centralny i pokazują trudność modelu generatywnego w uzyskaniu wyraźnego efektu zniekształceń panoramicznych. Gwiazdzisty kształt rozkładu (najmocniej widoczny dla zbioru *ITEMS*) sugeruje, że obrazy syntetyczne przypominają bardziej rzut perspektywiczny niż obraz typu ERP, a także że zauważalnie różnią się od rzeczywistych zdjęć środowiska domowego.

W celu dalszej oceny zdolności zastosowania danych syntetycznych do szkolenia detektorów obiektów, przeszkolono i przetestowano po trzy rozmiary modeli rodziny YOLOv5 (*Small*, *Medium*, *Extra-Large*) dla wszystkich możliwych kombinacji zbiorów: *ITEMS*, *NO-ITEMS*, *REAL*, *MIXED* (*ITEMS* + *NO-ITEMS*) oraz *ALL* (*MIXED* + *REAL*). Wyniki dla mniejszych modeli, dla uproszczenia, zostają pominięte, a w Tab. 4.7 przedstawiono wartości mAP@50 uzyskane dla modelu YOLOv5 w wersji *Extra-Large*. Pierwszą istotną obserwacją jest fakt, że modele wyszkolone tylko na danych syntetycznych działają nieakceptowalnie słabo przy zastosowaniu do obrazów rzeczywistych. Osiągane wyniki są akceptowalne tylko przy zastosowaniu w obrębie tej samej grupy, tj. detekcja wśród obrazów syntetycznych po szkoleniu na danych syntetycznych. Charakterystyki danych zbiorów syntetycznych są na tyle wyraziste, że sieć potrafi się ich nauczyć, jednakże równocześnie są na tyle rozbieżne od danych rzeczywistych, że ich zastosowanie nie przynosi dobrych rezultatów. Pomimo wysokiego poziomu pozornego, wizualnego podobieństwa obrazów, przejście między domenami stanowi wyzwanie.



Rysunek 4.19: Wykresy rozkładów przestrzennych etykiet obiektów na zdjęciach ERP dla kolejnych analizowanych zbiorów danych.

Tabela 4.7: Porównanie skuteczności detekcji (mAP@50) w różnych konfiguracjach zbiorów treningowych i testowych dla modelu YOLOv5 *Extra-Large*.

Trenowanie \ Testowanie	NO-ITEMS	ITEMS	MIXED	REAL	ALL
NO-ITEMS	0.466	0.445	0.410	0.179	0.313
ITEMS	0.484	0.544	0.494	0.178	0.348
MIXED	0.509	0.554	0.515	0.196	0.368
REAL	0.360	0.297	0.299	0.363	0.317
ALL	0.549	0.568	0.529	0.344	0.441

Pozostając w domenie danych syntetycznych, trenowanie modeli na zbiorach *ITEMS* oraz *NO-ITEMS* daje bardzo zbliżone wyniki podczas testowania na danych rzeczywistych. Zbiór *ITEMS* stanowi istotnie trudniejszy przypadek testowy dla modeli trenowanych na danych rzeczywistych niż zbiór *NO-ITEMS* (różnica na poziomie ok. 5 punktów procentowych). Zapewne wynika to z obecności w zbiorze *NO-ITEMS* bardziej typowych i przewidywalnych przedmiotów, które model generatywny znalazł lepiej i był w stanie wykorzystać bez dodatkowego wymagania ze strony list oczekiwanych obiektów. Jest to potencjalny kierunek dalszych badań. Dla testowania systemów pod kątem zgodności z wymaganiami użytkowników, zaleca się zatem wykorzystywanie zbiorów syntetycznych tworzonych z uwzględnieniem tychże, ponieważ sieć generatywna bez precyzyjnych wymagań wprowadzi do obrazów jedynie najpopularniejsze i najprostsze do rozpoznania obiekty.

Podsumowując, eksperymenty pokazały, że wzbogacenie poleceń generacji obrazów syntetycznych pozwala na uzyskanie wyższej jakości danych. Dotyczy to zarówno zawartości obrazów, jak i ich zdolności do dalszych badań. Niemniej, detektory szkolone na zbiorach czysto syntetycznych osiągają zauważalnie gorsze wyniki detekcji obiektów w realiach analizy obrazów typu ERP. Ich zastosowanie sugerować można do celów rozszerzania i heterogenizacji posiadanych zbiorów rzeczywistych, ale nie jako ich zamienniki. **Uzyskano zatem odpowiedź na jedno z pytań badawczych rozprawy. Zastosowanie, nawet ukierunkowanej precyzyjnymi oczekiwaniami użytkowników, generatywnej sztucznej sieci neuronowej nie pozwala na opracowanie syntetycznego zbioru danych panoramicznych wystarczająco wysokiej jakości, aby w pełni zastąpić manualne pozyskiwanie zdjęć rzeczywistego środowiska domowego.**

4.3 Prototyp systemu lokalizacji wizyjnej

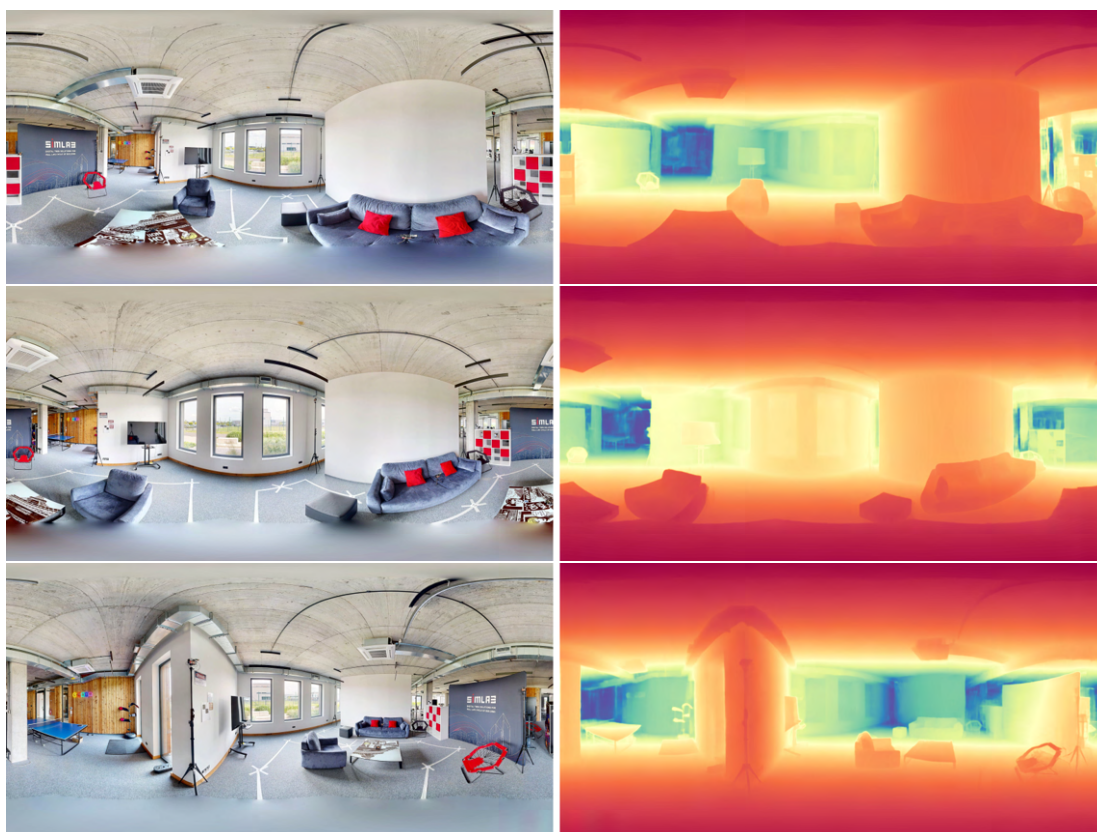
Podobnie jak w przypadku systemów radiowych, rozwiązanie pozycjonowania wizyjnego najpierw przetestowano w ramach opracowanego prototypu. Potwierdzić miało to zdolność poszczególnych proponowanych rozwiązań (detekcji i klasyfikacji obiektów oraz estymacji odległości do nich) do wdrożenia w końcowym systemie pozycjonowania hybrydowego. W ramach przygotowań oraz opracowywania zagadnień związanych z wizją komputerową, już na etapie badań wstępnych prze-

prowadzono wiele eksperymentów doświadczalnych dotyczących detekcji obiektów. Ich wyniki zweryfikowały wstępne założenia i dostarczyły konkretnych danych liczbowych (Roz. 4.2.1). Powtarzanie analogicznych badań w ramach kolejnego prototypu systemu wizyjnego uznano za nieefektywne tak merytorycznie, jak i wdrożeniowo. Z tego też względu dalsze kroki postępowania skierowano ku analizie wdrożenia do rozwiązania nowego zagadnienia, tj. estymacji głębi na obrazie. W oparciu o stanowisko pomiarowe z badań radiowych oraz jego model trójwymiarowy i interaktywny spacer, przeanalizowano skuteczność wybranego detektora głębi w pracy z obrazami formatu ERP, tj. w zadaniu szacowania odległości od obiektu skanera przestrzeni do wykrytego obiektu. W rezultacie zaproponowano autorski prototyp systemu, który stanowi odpowiednik funkcjonalny rozwiązania radiowego, jednak oparty na technikach wizji komputerowej. Opracowane rozwiązanie integruje już przebadane elementy (realizację cyfrowego bliźniaka, pracę z obrazami panoramicznymi, detekcję obiektów) rozszerzając je o zagadnienie dodatkowe, tj. wyznaczanie odległości do wykrytych obiektów. Omówione zostaje to w dalszej części niniejszego rozdziału.

4.3.1 Detektor głębi

Estymacja głębi polega na przypisaniu każdemu pikselowi obrazu wartości liczbowej, która określa odległość reprezentowanego przez niego punktu sceny od kamery. W typowych zdjęciach perspektywicznych pozwala to uzyskać efekt poczucia głębi obrazu, a w kontekście panoram ERP (w przypadku idealnym) umożliwia pełne odtworzenie sceny trójwymiarowej i jej struktury. Tak opracowane odległości przedstawić można np. przy pomocy gradientowej palety barw, która konkretnym odległościom przypisuje pewne odcienie koloru. Powstały w ten sposób obraz nazywamy mapą głębi (Rys. 4.20).

Wśród opisanych w literaturze przedmiotu opracowań odnaleziono dwa główne rodzaje wygenerowanych map głębi: relatywne i metryczne. W pomiarach relatywnych skalą wartości jest zwykle przedział od 0 do 1, a obiekty na obrazie oceniane są tylko względem siebie. Uzyskuje się dzięki temu poczucie, który obiekt umieszczony był bliżej, a który dalej. Nie daje to jednak konkretnych wartości liczbowych, a analiza wyników jest zasadniczo inna w scenariuszu zdjęć z poszczególnych pokoi



Rysunek 4.20: Przykładowe zdjęcia panoramiczne z badań prototypu wizyjnego z wyznaczonymi dla nich mapami głębi. Paleta barw określa odległość do obiektu, kolory ciepłe podkreślają krótki dystans, kolory zimne duży dystans.

małego mieszkania, a np. stadionu sportowego lub przestrzeni otwartej. Metryczne mapy głębi rozwiązują ten problem i podają odległość do obiektu w postaci konkretnej wartości liczbowej, wyrażonej zwykle w centymetrach lub metrach. Opracowanie, a także późniejsze użycie detektora metrycznego jest jednak zauważalnym wyzwaniem. Wielokrotnie do poprawnego działania metrycznej estymacji głębi konieczna jest znajomość parametrów soczewki obiektywu kamery lub innych jej cech charakterystycznych. W najnowszej literaturze przedmiotu odnaleziono jedynie kilka publicznie dostępnych i zdalnych do użycia przemysłowych projektów, które prezentowałyby wystarczającą elastyczność pod względem sprzętowym. Do najważniejszych rozwiązań, które odnaleziono, należały *Depth Pro* firmy Apple [325], *ZoeDepth* firmy Intel [326], a także *Depth Anything* tworzone pod przewodnictwem

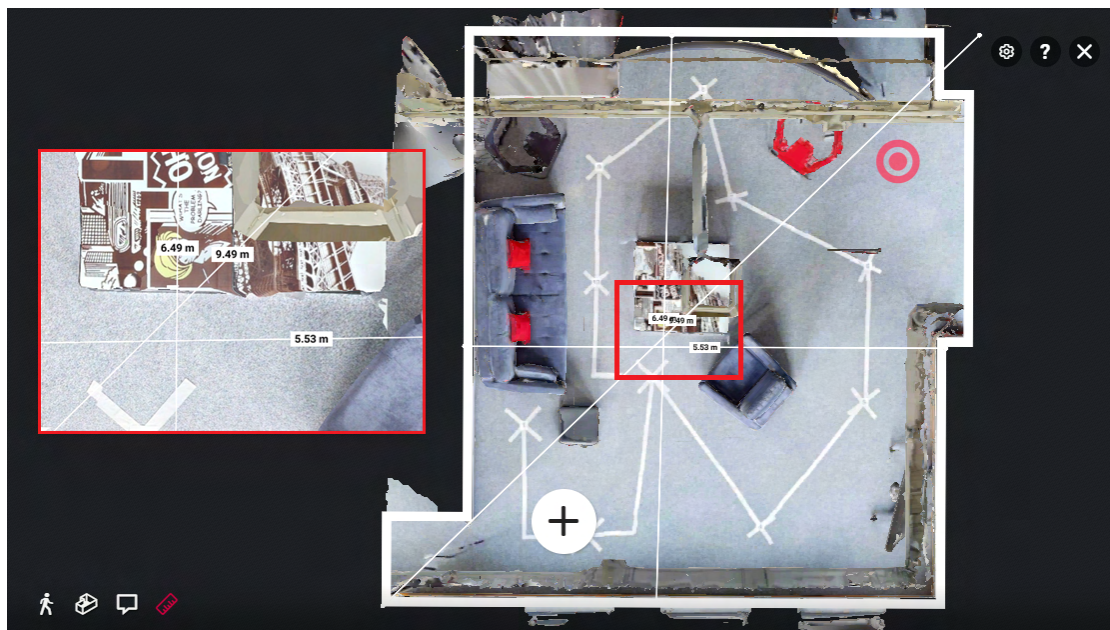
pracowników firmy TikTok [327]. Rozwiązanie *Depth Pro* odrzucono ze względu na specyficzną licencję udostępnionego oprogramowania i zasady jego użycia. Projekt *ZoeDepth* został porzucony przez autorów rozwiązania, a repozytorium kodu źródłowego zarchiwizowano. Jedynym projektem, nadal kontynuowanym oraz takim, który doczekał się następcy, jest *Depth Anything*. Tym samym do opracowanego prototypu systemu wizyjnego obrano detektor głębi *Depth Anything V2* [328], będący ulepszoną wersją pierwowzoru i prezentującą (na moment prowadzenia badań) rezultaty demonstrujące aktualny stan wiedzy w tej dziedzinie.

Do poprawnego działania detektor ten wymaga od użytkownika sprecyzowania oczekiwanej maksymalnej odległości do najbardziej oddalonego obiektu na scenie. Określona zostaje ona poprzez parametr `max_depth`. W realiach wdrożeniowych opracowywanego rozwiązania nie jest jednak jasne, w jaki sposób parametr ten należy wyznaczyć. Typowy scenariusz digitalizacji przestrzeni domowej prowadzi do dużych rozbieżności obserwowanej na obrazie głębi. Mowa tu o np. otwartych drzwiach pomiędzy kolejnymi pokojami albo cyfryzacji dużej przestrzeni biurowej. Sytuację tę wyraźnie wskazać możemy na przykładowych zdjęciach ze środowiska pomiarowego, gdzie widok przez otwarte drzwi sprawia, że oczekiwana estymacja dystansu wzrasta skokowo, z kilku do nawet kilkunastu metrów. Intuicyjnym parametrem wydaje się dobór parametru `max_depth` zgodnie z przeciętną szerokością pomieszczeń, w których się poruszamy. Zagadnienie to postanowiono jednak głębiej przebadać, co przedstawiono w dalszej części rozdziału.

4.3.2 Środowisko testowe

Jak wspomniano, za środowisko testowe obrano przestrzeń znaną z badań systemu radiowego (omówiono ją już ze szczegółami w Roz. 3.3.2). Dzięki stworzeniu jej cyfrowej reprezentacji, możliwy stał się powrót do badań nawet kilka miesięcy później i na długo po zmianie aranżacji pomieszczenia. Ze względu na specyfikę zastosowanego detektora głębi, konieczne stało się doprecyzowanie faktycznych wymiarów przestrzeni pod względem długości i szerokości. Pomieszczenie, w którym prowadzono badania, miało szerokość około 5,5 metra, a także 6,5 metra długości. Pomiary z użyciem planu pomieszczenia ze spaceru interaktywnego przedstawiono na Rys. 4.21. Tak określone wymiary środowiska wykorzystane zostały następnie

do zaproponowania wartości badawczych dla parametru maksymalnej odległości ustawionej w programie detektora głębi. Istotnym aspektem, wartym podkreślenia, jest fakt, że obszar badawczy to pewien wycinek całego skrzydła budynku, który funkcjonuje w systemie “open-space”. Sprawia to, że określone fragmenty zdjęcia mogą prezentować skokowy wzrost głębi do poziomu nawet 24 metrów (długości przekątnej całego piętra).



Rysunek 4.21: Zrzut ekranu narzędzia *Matterport Viewer* z naniesionymi wymiarami pomieszczenia oraz zbliżeniem na etykiety długości.

Dla każdego z dziesięciu umieszczonych w przestrzeni punktów pomiarowych wyeksportowane zostały odpowiadające im obrazy panoramiczne. Efektywnie uży skano w ten sposób nowy zbiór testowy, dedykowany badaniom detektora głębi. Przykładowe jego obrazy prezentowane są na Rys. 4.22. Przyglądając się wybranym fragmentom obrazu, dostrzec można wspomniane “przebicia” widoczności części budynku, wykraczającej poza przestrzeń badawczą.

Opracowując zestaw danych referencyjnych dla prototypu wizyjnego, konieczne było uzyskanie informacji o faktycznym umiejscowieniu obiektywu kamery (urządzenia skanującego). Jako że wyniki badań wstępnych zwróciły wysoki stopień precyzji pomiarów dokonywanych w oparciu o model cyfrowy, położenie kamery



Rysunek 4.22: Przykładowe panoramy ze zbioru punktów pomiarowych prototypu systemu wizyjnego.

w płaszczyźnie podłogi odczytywano każdorazowo ze spaceru trójwymiarowego. Zastosowanie oryginalnego skanera *Matterport Pro 3* oraz dedykowanego dla niego trójnogu sprawiło jednak, że konieczne stało się wyznaczenie wysokości w osi pionowej, na jakiej znajdowało się urządzenie. Rozstawiono je więc dodatkowo po zakończeniu badań i zmierzono wysokość taśmą mierniczą. Na podstawie dokumentacji technicznej urządzenia spodziewano się wartości 137 cm nad poziomem podłogi (ok. 4,5 stopy). Faktyczna wysokość umieszczenia obiektywu okazała się jednak większa o ok. 10 cm. Wynikało to z różnicy pomiędzy zakończeniem stelaża (faktycznie zmierzonego na 137 cm), a dodatkowo obudową samego skanera. Wizualizacja urządzenia *Pro 3* oraz jego rozstawienie z naniesionymi pomiarami wysokości przedstawiono na Rys. 4.23.

4.3.3 Eksperymenty

W ramach eksperymentów pomiarowych przeanalizowano proces doboru parametru maksymalnej głębi detektora z projektu *Depth Anything V2*[328]. Dla pokrycia przypadkami testowymi możliwie szerokiego wachlarza wartości, za maksymalną spodziewaną odległość do obiektu obrano kolejno: 50%, 100%, 150% oraz 200% zmierzonej średniej szerokości pomieszczenia badawczego. Na podstawie pomiarów środowiska, faktyczną szerokość pomieszczenia określono na 6 metrów, a zatem za wartości badawcze parametru `max_depth` wybrano kolejno: 3 m, 6 m, 9 m oraz 12 m. Dla każdego z punktów pomiarowych (dla każdej panoramy) wyznaczane były w ten sposób cztery mapy głębi, które wykorzystywano następnie w połączeniu z detekcją obiektów.

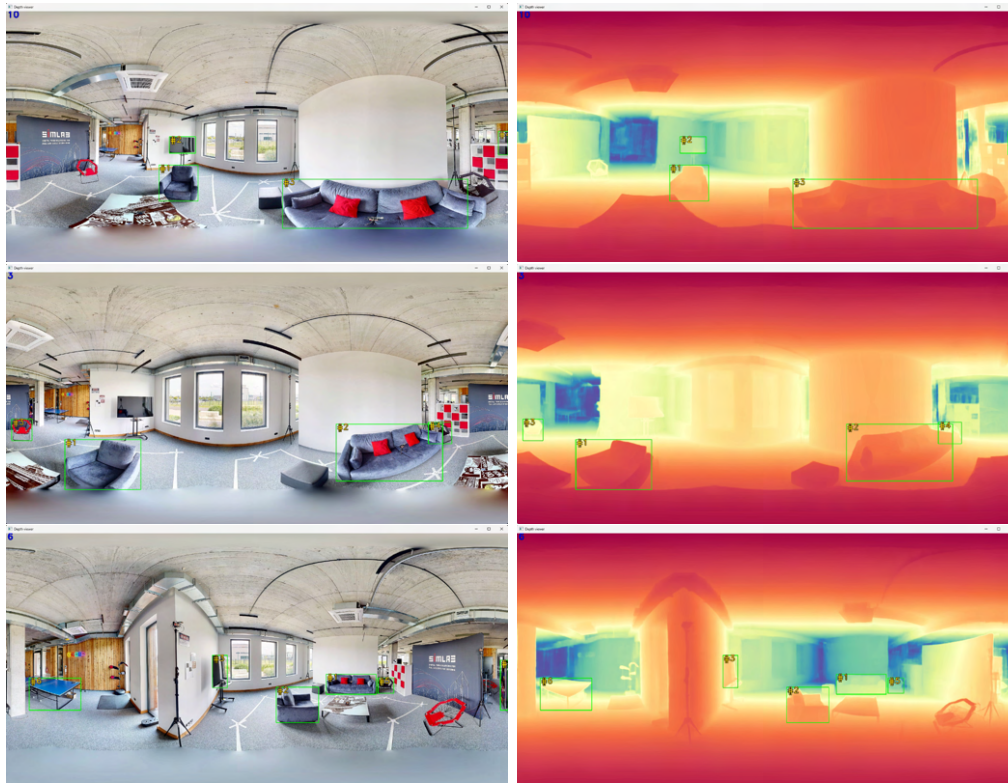
Proces rozpoczynano od wykorzystania sieci neuronowej detektora obiektów (dla zaktualizowania wykorzystywanych rozwiązań zastosowano sieć YOLOv11



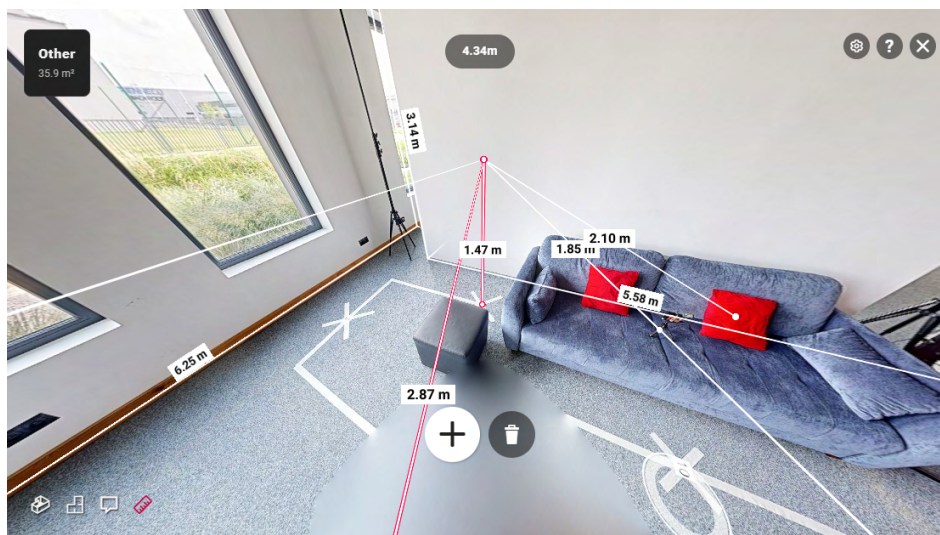
Rysunek 4.23: Zdjęcie skanera przestrzeni Matterport Pro 3 (po lewej) oraz jego umieszczenie na dedykowanym trójnogu z naniesionymi na zdjęcie pomiarami wysokości (po prawej).

w wersji *Extra-Large*) do wykrycia i wskazania położenia dostrzeganych na obrazie obiektów. Następnie, dla każdego z tak wykrytych obiektów odczytywano odległość. Za punkt pomiaru obierano środek ramki ograniczającej. W ten sposób pozyskano cztery odczyty głębokości, dla kolejnych map głębi, stworzonych dla kolejnych wartości parametru maksymalnej głębi. Wizualizacja detekcji obiektów oraz odczytu odpowiadających im odległości z map głębi przedstawiona jest na Rys. 4.24.

Aby porównać wartości odległości z detektora głębi do rzeczywistości, konieczne było opracowanie danych oczekiwanych. Zgromadzono je z użyciem spaceru interaktywnego. Odczytując położenie w płaszczyźnie podłogi na podstawie umiejscowienia punktu skanowania przestrzeni i zestawiając je z wyznaczoną na 147 cm nad podłogą wysokością obiektywu kamery, uzyskiwano punkt referencyjny. Dla każdej z panoram i dla każdego wykrytego obiektu odczytywano oczekiwaną wartość odległości. Wizualizację takiego procesu pokazano na Rys. 4.25.



Rysunek 4.24: Przykładowe wizualizacje przetworzonych z użyciem detektora obiektów zdjęć panoramicznych oraz odpowiadających im map głębi.

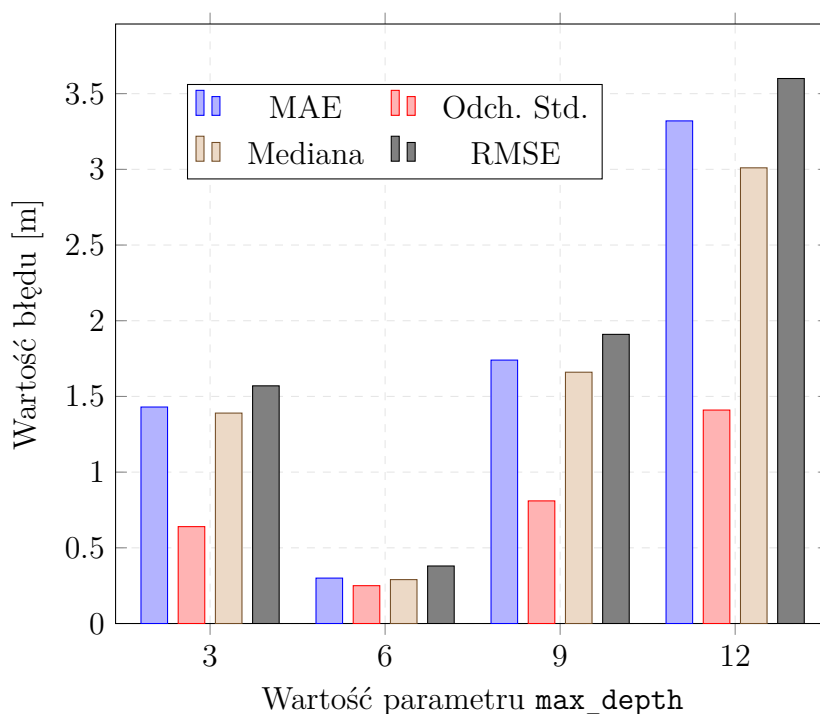


Rysunek 4.25: Zrzut ekranu z narzędzia *Matterport Viewer* przedstawiający proces pomiaru oczekiwanej odległości od skanera przestrzeni do wykrytego obiektu.

4.3.4 Wyniki i wnioski

Na zgromadzonym zbiorze dziesięciu zdjęć panoramicznych wykrytych zostało łącznie 53 obiekty. Dla 12 z nich pomiar odległości oczekiwanej nie był możliwy. Ich oznaczenie miało miejsce we fragmentach panoram, które odnosiły się do wspomnianych “przebić głębi”, tj. oznaczano obiekty spoza przestrzeni badawczej, a dostrzegalne na zdjęciu. W takiej sytuacji nie posiadano modelu trójwymiarowego niezbędnego fragmentu piętra, a co za tym idzie, brak było możliwości pomiaru odległości. Efektywnie badania oparto więc na 41 wykrytych obiektach. Nie analizowano detekcji pod względem jakościowym, a jedynie ilościowym, tj. jako pewne wytyczne odnośnie interesujących obiektów na zdjęciach.

Dla każdej z wartości parametru `max_depth` wyznaczono zbiorcze wyniki: średni błąd bezwzględny (z ang. *Mean Absolute Error*, MAE), jego odchylenie standardowe, a także medianę błędu bezwzględnego oraz błąd średniokwadratowy (z ang. *Root Mean Square Error*, RMSE). Przedstawiono je na Rys. 4.26.



Rysunek 4.26: Porównanie metryk błędu dla różnych wartości parametru `max_depth` detektora obiektów *Depth Anything V2*.

Wyniki jasno pokazują, że najlepszą skutecznością detekcji głębi okazał się wybór parametru `max_depth` zgodnie z przeciętną szerokością pomieszczenia, w którym prowadzimy badania. Skoki głębi należy zignorować, ponieważ próba dostosowania parametru do tzw. najgorszego scenariusza zauważalnie pogarsza wyniki końcowe. Jako całość, zaproponowany prototyp wizyjnej estymacji głębi dobrze spełnił swoje zadanie. Połączenie omówionych zagadnień cyfryzacji przestrzeni, detekcji obiektów, estymacji głębi oraz przetwarzania zdjęć panoramicznych pozwala na skuteczne wykorzystanie takiego podejścia w docelowym systemie fuzyjnym.

Na podstawie przeprowadzonych badań można wyciągnąć konkretną rekomendację dla docelowego systemu fuzyjnego: **detektor głębi dla części wizyjnej skonfigurować należy zgodnie z przeciętnym rozmiarem pomieszczenia badawczego.**

Rozdział 5

Fuzyjny, wizyjno-radiowy system pozycjonowania komponentów środowiska domowego

Indywidualne wykorzystanie systemów wizyjnych lub radiowych w środowisku domowym prowadzi do powstania istotnych ograniczeń: obrazy uzyskiwane za pomocą kamer sferycznych ogólnego zastosowania nie zawierają powiązania z globalnym układem współrzędnych całego środowiska, natomiast profesjonalne skanery, mimo wysokiej jakości pomiaru geometrii, nie dostarczają semantycznych informacji o obiektach w swoim otoczeniu. Połączenie obu podejść stwarza możliwość odzyskania brakującej semantyki oraz automatyzacji procesu lokalizacji. Na podstawie przeprowadzonych badań wstępnych i skonstruowanych prototypów opracowany został wynikowy system fuzyjny. Łączy on oba rodzaje rozwiązań w jednej, nadrzędnej metodzie wyznaczania pozycji i rodzaju obiektów, napotykanym w środowisku domowym. Zaproponowane rozwiązanie wykorzystuje wnioski ze skonstruowanych i przebadanych prototypów, a także prac badawczych zapoczątkowanych jeszcze przed rozpoczęciem realizacji doktoratu [329–331]. Końcowe wdrożenie dopasowane zostaje natomiast do oczekiwań użytkowników oraz realiów przemysłowych partnera doktoratu.

W następnych rozdziałach kompleksowo przedstawiono opracowaną metodę. Rozpoczęto od opisanie jej założeń i etapów realizacji, a także architektury całego

systemu pozycjonowania. Następnie przybliżono skonstruowany i wykorzystywany w jego ramach sprzęt oraz opracowany model cyfrowy środowiska, w którym przeprowadzono końcowe badania eksperymentalne. Kolejne trzy rozdziały poświęcono omówieniu poszczególnych etapów realizacji metody hybrydowej, a całość prac podsumowują rozbudowane badania eksperymentalne.

Podsumowując, niniejszy rozdział stanowi zwieńczenie prac badawczych całej rozprawy. Ich realizacja obejmowała:

- opracowanie nowej, autorskiej metody hybrydowego pozycjonowania obiektów środowiska domowego;
- skonstruowanie niezbędnych do jej działania urządzeń oraz przygotowanie środowiska badawczego;
- przeprowadzenie rozbudowanych eksperymentów, które dostarczyły wyników niezbędnych do weryfikacji stawianej w rozprawie hipotezy głównej.

5.1 Hybrydowa metoda pozycjonowania obiektów środowiska domowego

Podstawowym celem proponowanej metody hybrydowej było rozpoznanie i wyznaczenie położenia możliwie największej liczby obiektów, faktycznie obecnych w nadzorowanym środowisku. Dla ułatwienia opisu rozróżniano dwa rodzaje takich obiektów: pasywne i aktywne. Do grupy obiektów pasywnych należały takie obiekty, które nie posiadały komponentów elektronicznych, a w szczególności modułów komunikacji radiowej. Były to np. krzesła, stoły, kanapy i inne meble. Nie były one w stanie samodzielnie określić swojego położenia w globalnym układzie współrzędnych całego środowiska. Obiektami aktywnymi były natomiast urządzenia elektryczne i elektroniczne, szczególnie te, których położenie ustalić można było przy użyciu samego tylko radiowego systemu pozycjonowania (na potrzeby badań dzięki przypisaniu dedykowanego lokalizatora do każdego takiego obiektu). Dla ułatwienia klasyfikacji i dopasowania się do realiów systemów IoT, do grupy aktywnej zaliczano wszystkie obiekty współcześnie wykorzystujące pewne źródło zasilania, niezależnie czy bateryjne (akumulatorowe) czy sieciowe.

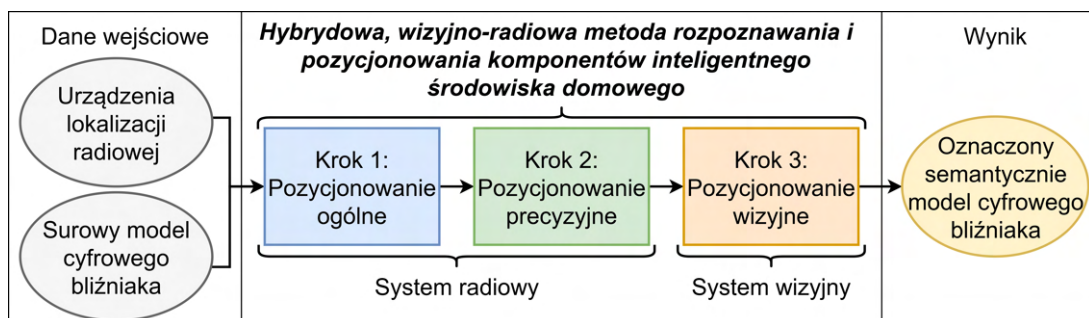
Zasadniczym elementem autorskim proponowanego systemu fuzyjnego było mobilne urządzenie pomiarowe, wprowadzone do środowiska w celu kompensacji ograniczeń obiektów pasywnych i umożliwienia ich pozycjonowania. Urządzeniem tym była kamera sferyczna wyposażona w moduły radiowe BLE oraz UWB, której zadaniem była detekcja obiektów z wykorzystaniem technik wizji komputerowej. Oznaczenie obiektów na zgromadzonych zdjęciach, dzięki zastosowaniu przeliczeń sferycznych, pozwoliło na wskazanie ich położenia w trójwymiarze. Pozycjonowanie wizyjne zapewniło jednak jedynie współrzędne względem lokalnego układu kamery, a nie względem globalnego układu całego środowiska. Z tego względu kamera zintegrowana została z lokalizatorami radiowymi i pełniła jednocześnie rolę obiektu aktywnego, który śledzono w przestrzeni za pomocą systemu pozycjonowania radiowego. Dzięki temu znane były jej współrzędne w układzie globalnym. Zestawienie obu układów współrzędnych pozwoliło na końcowe wyznaczenie globalnych współrzędnych wszystkich wykrytych wizyjnie obiektów, a system fuzyjny zrealizował syntezę detekcji i pozycjonowania zarówno obiektów aktywnych, jak i pasywnych.

Pod względem sposobu realizacji, proponowana metoda hybrydowa ma charakter wieloetapowy. Można ją modyfikować, zgodnie z dostępnymi technologiami radiowymi oraz sprzętem wizyjnym, a także dostosowywać do potrzeb wdrożeniowych. Poszczególne kroki zrealizować można na kilka różnych sposobów, dopasowując stopień uzyskanej automatyzacji do oczekiwanych, akceptowalnych nakładów pracy manualnej.

Poprawne działanie fuzyjnego systemu pozycjonowania obiektów wymagało spełnienia określonych warunków wejściowych, użycia odpowiedniego sprzętu oraz zachowania sekwencyjności operacji, w której każda opierała się na wynikach poprzedniej. Przedstawiona w rozprawie implementacja wiązała się więc z opracowaniem następujących elementów:

- zestawu urządzeń lokalizacji radiowej (Roz. 5.2),
- środowiska pracy w postaci modelu cyfrowego (Roz. 5.3),
- procesu radiowego pozycjonowania ogólnego (Roz. 5.4),
- procesu radiowego pozycjonowania precyzyjnego (Roz. 5.5),
- procesu pozycjonowania wizyjnego (Roz. 5.6).

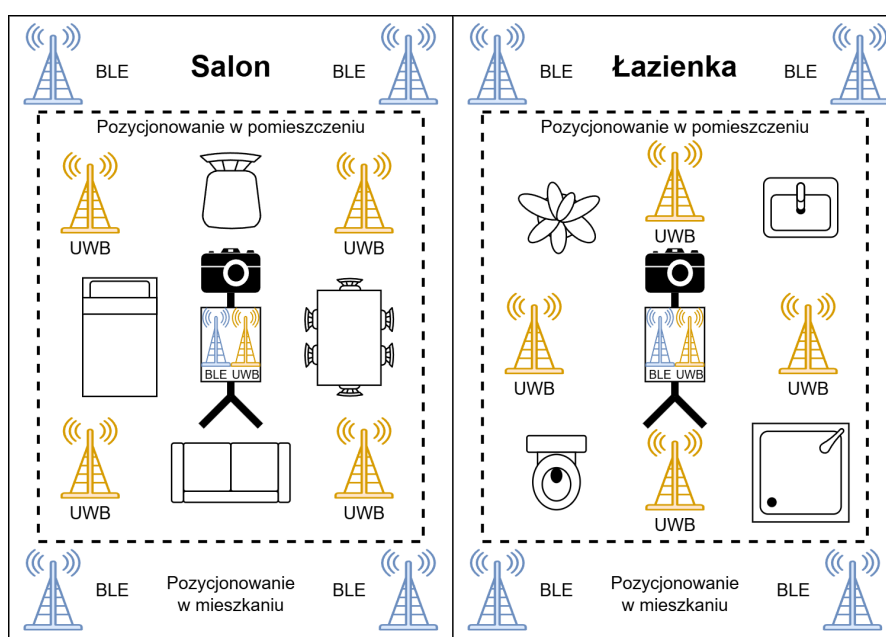
Schemat ogólny postępowania w ramach proponowanej metody hybrydowej przedstawiono na Rys. 5.1. Dwa pierwsze kroki miały na celu określenie współrzędnych kamery, a trzeci przedstawia jej faktyczne zastosowanie do wykrywania obiektów.



Rysunek 5.1: Schemat logiczny kroków realizacji proponowanej hybrydowej metody lokalizacji.

Słowem wstępu nakreślona zostaje ogólna architektura opracowanego systemu oraz zastosowane rozwiązania, a ich dokładne omówienie przedstawiono w kolejnych rozdziałach. Pod nazwą pozycjonowania ogólnego przedstawiono uzyskanie wstępnej informacji o pozycji śledzonego obiektu na poziomie całej przestrzeni. Efektywnie wskazuje to pokój lub pomieszczenie, gdzie należy kontynuować lokalizację. Zastosowano w tym celu podsystem radiowy oparty o technologię BLE. Pozycjonowanie precyzyjne uzupełnia informacje wstępne o konkretne współrzędne lokalizatora w odniesieniu do punktów referencyjnych umieszczonych w poszczególnych pomieszczeniach. W tej części zastosowane zostały komponenty radiowe technologii UWB. Całość podsumowuje proces pozycjonowania wizyjnego, który odpowiada za detekcję i pozycjonowanie obiektów, szczególnie pasywnych. Do tego celu zastosowano sportową kamerę sferyczną. Schemat architektury skonstruowanego systemu prezentowany jest na Rys. 5.2. Oznaczono na nim kolorem niebieskim komponenty systemu BLE (pozycjonowanie ogólne), kolorem pomarańczowym systemu UWB (pozycjonowanie precyzyjne), a mobilne urządzenie pomiarowe przedstawiono jako kamerę. Dla uproszczenia schematu pominięto na nim urządzenie odbierające pomiary i pełniące rolę miejscowego serwera. Do realizacji tej części wybrano minikomputer *Raspberry Pi 4B*. Obecny w sieci lokalnej mieszkania, stanowił serwer komunikacyjny protokołu MQTT, z użyciem którego raportowano,

zbierano i opcjonalnie przesyłano dalej zgromadzone pomiary. Takie rozwiązanie, tj. ujednolicony i niezależny od komponentów systemu punkt komunikacyjny, umożliwia swobodną zmianę implementacji wybranych funkcji i np. zastąpienie w pozycjonowaniu ogólnym technologii BLE dowolną inną. Istotne jest jedynie zachowanie analogicznego interfejsu.



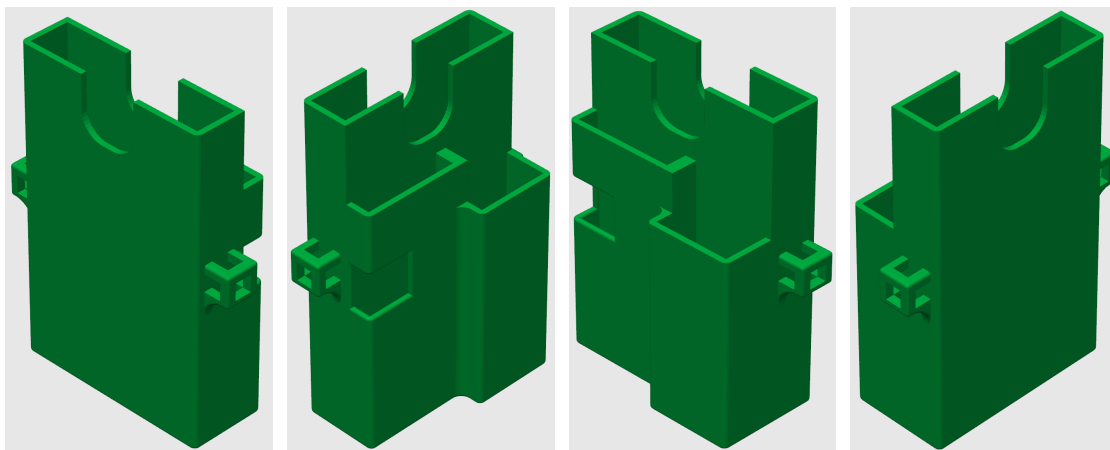
Rysunek 5.2: Schemat logiczny architektury proponowanego systemu fuzyjnego. Pozycjonowanie na poziomie mieszkania z użyciem technologii BLE, na poziomie pomieszczenia w oparciu o UWB oraz rozpoznawanie obiektów z użyciem kamery.

5.2 Skonstruowane urządzenia

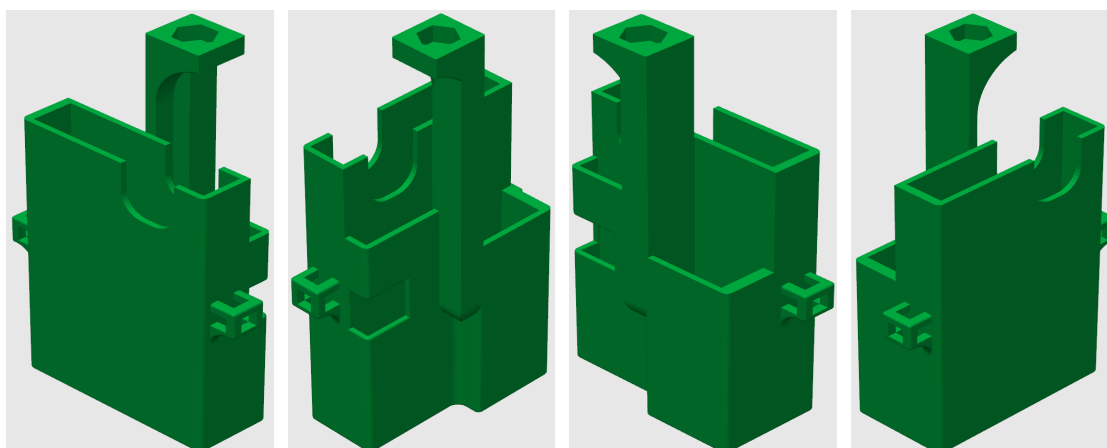
Realizacją doskonałą wdrożenia przemysłowego byłoby bezinwazyjne wpisanie się w obecny proces cyfryzacji już istniejącej nieruchomości. W scenariuszu idealnym zakłada się, że akwizycja danych realizowana będzie przy użyciu mobilnego urządzenia pomiarowego (zintegrowanego ze skanerem przestrzeni), a także otaczających go urządzeń IoT (już zamontowanych i używanych w budynku). Lokalizacja obiektów nie powodowałaby wtedy żadnych dodatkowych narzutów pracy człowieka. Systemy skanowania to jednak ściśle strzeżona tajemnica przemysłowa pro-

ducentów takich rozwiązań, a same urządzenia nie oferują indywidualnych modyfikacji. Podobnie wygląda sytuacja dostępnych na rynku urządzeń IoT oraz sprzętu gospodarstwa domowego. O ile technicznie wykorzystanie obecnych w nich komponentów jest możliwe i wystarczające, o tyle oficjalne zamknięcie samej platformy sprzętowej na autorskie modyfikacje stanowi przeszkodę natury prawnej. Po odnalezieniu producenta sprzętu chętnego do współpracy, proponowane rozwiązanie będzie można zintegrować z jego produktami. Nie ma jednak pewności, ile takie poszukiwania mogą zająć ani czy w ogóle zakończą się powodzeniem. Dlatego też opracowano własne zestawy urządzeń, które oparte są na powszechnie dostępnych i relatywnie tanich komponentach elektronicznych. Zaakceptowano dodatkowy nakład pracy i uznano dostępność sprzętu do lokalizacji radiowej (oraz samego modelu cyfrowego nieruchomości) za założenia wejściowe metody hybrydowej.

Uchwyty na moduły elektroniczne, wykorzystywane w trakcie przygotowywania prototypu radiowego systemu estymacji dystansu, były zbyt prowizoryczne, aby zastosować je w końcowym systemie hybrydowym, a także uznać za akceptowalne do prezentacji potencjalnym klientom. Z tego względu zaprojektowane zostały dedykowane obudowy, których modele prezentowane są na Rys. 5.3 i Rys. 5.4. Wykorzystano technikę druku trójwymiarowego oraz nakrętki i śruby w standardzie fotograficznym i przygotowano nowe uchwyty dla każdego ze zmontowanych urządzeń (tak lokalizatorów, jak i kotwic oraz urządzeń pomiarowych).



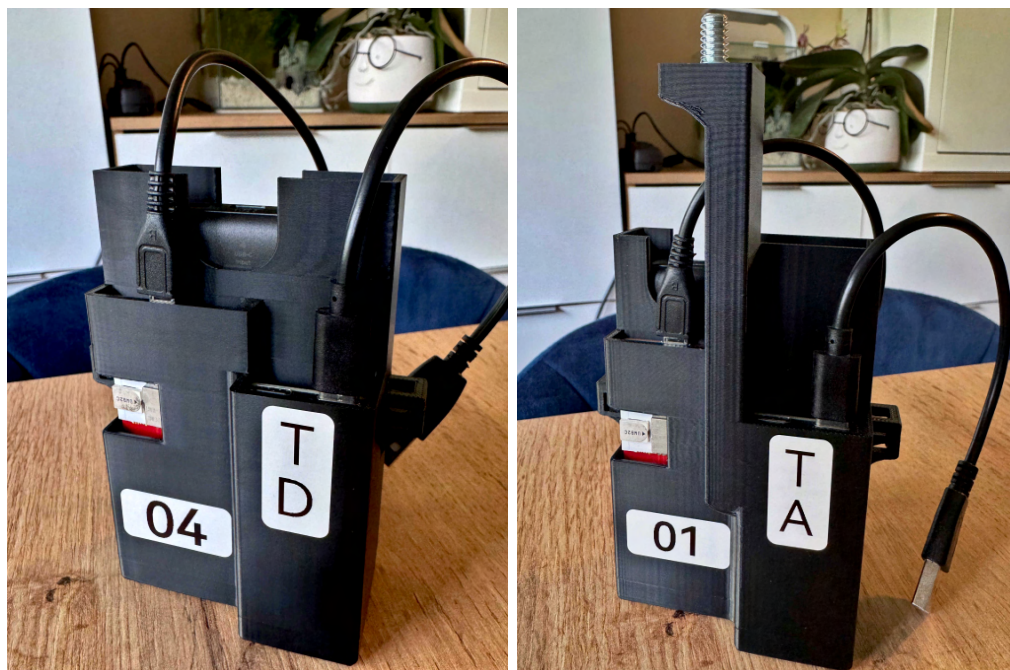
Rysunek 5.3: Zrzuty ekranu opracowanego modelu obudowy urządzenia kotwicy lub lokalizatora radiowej części systemu hybrydowego.



Rysunek 5.4: Zrzuty ekranu opracowanego modelu obudowy mobilnego urządzenia pomiarowego (ze słupkiem do montażu kamery sferycznej).

Dla urządzeń kotwicy oraz lokalizatora obudowa jest taka sama, wynika to z analogicznych wykorzystanych komponentów elektronicznych, które różnią się jedynie wgranym programem. Dla mobilnych urządzeń pomiarowych opracowano obudowę powiększoną o dodatkowy słupek do montażu kamery sferycznej. Przykłady gotowych urządzeń prezentowane są na Rys. 5.5. W podstawy uchwytów (czego nie widać na zrzutach ani zdjęciach) wbudowane zostały nakrętki fotograficzne, umożliwiające montaż uchwytów na uniwersalnych statywach fotograficznych.

Wspomniane urządzenia mobilne to zestawienie trzech komponentów, które wykorzystano w trakcie końcowych badań i eksperymentów. Każde z nich składało się z kamery sferycznej, radiowego lokalizatora i regulowanego trójnożu. W formie kamery sferycznej wykorzystano urządzenia Insta360 w wersji X3. Łączono je następnie z wydrukowaną obudową, a całość umieszczano na trójnożu o możliwie wąskiej podstawie. Skonstruowane urządzenia przedstawiono na Rys. 5.6.



Rysunek 5.5: Zdjęcia zrealizowanych w technologii druku 3D obudów urządzeń systemu radiowego: statycznych (po lewej) i mobilnych (po prawej).



Rysunek 5.6: Zdjęcia mobilnych urządzeń pomiarowych. Na lewym zdjęciu, w zbliżeniu, od góry: kamera sferyczna, lokalizator radiowy, trójnóg montażowy. Po prawej zdjęcie wszystkich trzech przygotowanych urządzeń pomiarowych.

Zastosowanie standardowych nakrętek i śrub fotograficznych w przygotowanych obudowach umożliwiło ich swobodne łączenie z dowolnym statywem lub trójnogiem. W większości przypadków do rozmieszczenia kotwic w środowisku pomiarowym wykorzystano obecne na miejscu meble i konstrukcję budynku (np. parapety okienne). Dla dywersyfikacji wysokości kotwic zastosowano jednak również dedykowane statywy dużej wysokości, które (podobnie jak podczas prac nad prototypem systemu radiowego) pozwoliły wynieść urządzenia właściwie pod sam sufit pomieszczenia. W badaniach wykorzystano cztery takie statywy, a montaż kotwic i przykład ich zamontowania przedstawiono na Rys. 5.7.



Rysunek 5.7: Po lewej: zdjęcie sposobu montażu punktów referencyjnych umieszczanych na dużych wysokościach, tj. z wykorzystaniem statywów, które pozwoliły na stabilne wyniesienie kotwic radiowych pod sufit pomieszczenia. Po prawej: zdjęcie faktycznego użycia takiego statywu.

Każde przygotowane urządzenie etykietowane było odpowiednimi oznaczeniami na dwa sposoby. Po pierwsze zgodnie z kolejnym numerem urządzenia. Dodatkowo, dla zapewnienia redundancji, alfabetycznie i zgodnie z wgranym na moduły oprogramowaniem: kotwice z prefiksem *A* (urządzenia *AA*, *AB*, *AC*, itd.), a lokalizatory z prefiksem *T* (urządzenia *TA*, *TB*, *TC*, itd.). W ten sposób przygotowano 35 urządzeń końcowych (o numerach od 01 do 35), podzielonych na dwie kategorie. Dla wygody etykiety umieszczono również na kamerach, oznaczając je kolejno: *C1*, *C2* oraz *C3*. Całokształt skonstruowanych urządzeń i wykorzystanego wyposażenia do radiowej części systemu hybrydowego przedstawiono na Rys. 5.8. Dodatkowo, w ujęciu ze zbliżeniem na Rys. 5.9, pokazano faktycznie uruchomione i aktywne urządzenia lokalizacyjne po podłączeniu modułów do zasilania.



Rysunek 5.8: Zdjęcie zbiorcze wszystkich 35 skonstruowanych urządzeń, budujących radiową część hybrydowego systemu pozycjonowania. Na zdjęciu, kolejno: trzy mobilne urządzenia pomiarowe (kamera, lokalizator, trójnóg) po lewej stronie, cztery zmontowane punkty referencyjne dużej wysokości (kotwice wyniesione pod sufit) w tle, 28 pozostałych urządzeń (20 kotwic i 8 lokalizatorów) z przodu.



Rysunek 5.9: Zdjęcie skonstruowanych urządzeń radiowych podczas pracy i z podłączonym zasilaniem.

5.3 Środowisko badawcze

Typowym środowiskiem pracy dla opracowanej metody była pewna zamknięta przestrzeń wewnątrz budynku, nosząca cechy wykorzystania mieszkaniowego, biurowego lub innego nieprzemysłowego. Spodziewanym przykładem może być pokój hotelowy, mieszkanie, dom rodzinny lub małe biuro. O ile zastosowanie proponowanego postępowania powinno dobrze sprawdzić się również w scenariuszach hal produkcyjnych, magazynów czy boisk sportowych, o tyle nie było ono w takich warunkach sprawdzane na żadnym etapie prac.

Do badań systemu hybrydowego w pełni wykorzystano ideę cyfrowego bliźniaka przestrzeni pomiarowej. Ze względu na jakość wyników uzyskiwanych w oparciu o model trójwymiarowy (co przedstawiono w Roz. 3) zaniechano manualnego wyznaczania pozycji, wymiarowania przestrzeni, czy określania położenia urządzeń. Wszelkie pomiary i analizę położenia prowadzono na podstawie punktów siatki modelu przestrzennego. Wizualizację opracowanej reprezentacji przedstawiono na Rys. 5.10.



Rysunek 5.10: Wizualizacja modelu trójwymiarowego całego środowiska badawczego z użyciem narzędzia *Matterport Viewer*.

Zdigitalizowano łącznie pięć pomieszczeń: cztery pokoje wyposażone w radiowy system pozycjonowania, a także łączący je korytarz (pozbawiony radiowych punktów referencyjnych). Plan pomieszczeń, w widoku “z góry”, prezentowany jest na Rys. 5.11. Pogrubiono granice pomieszczeń i oznaczono kolorami: salon na czerwono, łazienka na zielono, sypialnia na niebiesko, kuchnia na fioletowo, korytarz na żółto. Widoczne na planie są również naniesione punkty pomiarowe, w których umieszczano urządzenia i dokonywano akwizycji danych (zdjęć oraz estymacji odległości). Do realizacji cyfrowego bliźniaka wykorzystano skaner przestrzeni firmy *Matterport* (model *Pro3*), a także oprogramowanie platformy chmurowej, dostarczającej spacer interaktywny oraz właściwy opracowany model trójwymiarowy.

Na modelu wczytanym do programu *Blender* przystąpiono do oznaczania znajdujących się w mieszkaniu obiektów. Wyznaczenie pozycji każdego obiektu poprzez otoczenie go możliwie ściśle przylegającym prostopadłościanem (odpowiednikiem ramki ograniczającej, ale dla danych trójwymiarowych) było niezwykle mozolnym i czasochłonnym zadaniem. Łącznie wyszczególniono 208 obiektów obecnych w całym mieszkaniu. Tak utworzone adnotacje stanowiły dane referencyjne, umożliwia-



Rysunek 5.11: Rzut planu mieszkania (widok z góry) z narzędzia *Matterport Viewer*.

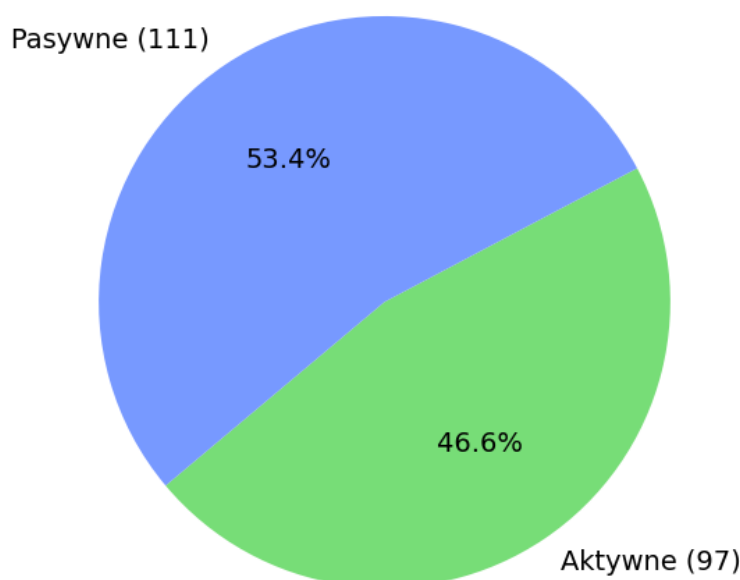
jące późniejszą ocenę jakości procesów automatycznego wykrywania i pozycjonowania obiektów w ramach systemu hybrydowego. Wizualizacja modelu z uruchomionym widokiem przedstawiającym naniesione prostopadłościanny ograniczające, prezentowana jest na Rys. 5.12.



Rysunek 5.12: Model mieszkania z naniesionymi oznaczeniami obiektów.

Oznaczając obiekty, dokonano również analizy ich rodzajów, rozróżniając obiekty pasywne i aktywne. Wyniki zaprezentowano na Rys. 5.13. Spośród 208 obecnych w mieszkaniu obiektów, 111 sklasyfikowano jako obiekty pasywne (53,4%), a 97 jako aktywne (46,6%). Podkreśla to istotność rozpoznawania obiektów, które z założenia nie będą w stanie samodzielnie określić swojego położenia na podstawie komunikacji radiowej. Obiekty pasywne, które wykryć trzeba bez interakcji z nimi, stanowią większość typowego wyposażenia domowego.

Podział komponentów środowiska badawczego na podstawie ich rodzaju



Rysunek 5.13: Wykres kołowy podziału oznaczonych komponentów środowiska domowego na obiekty pasywne i aktywne.

Najwięcej oznaczonych zostało lamp (30), szafek (14), czujników alarmowych (14), gniazdek elektrycznych (13) oraz włączników i wazonów (po 11). Jest to wysoce zgodne z ludzką intuicją i pokrywa się z oczekiwaniami dotyczącymi środowiska domowego. Dość zaskakujące może się wydawać liczne występowanie wazonów, wynika ono jednak z przyjętego systemu oznaczania obecnych w przestrzeni

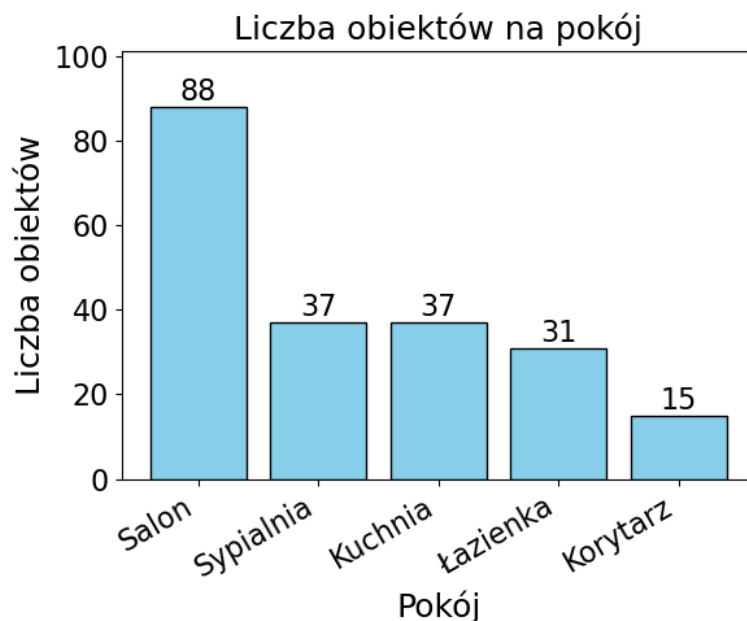
kwiatów: osobno samej rośliny, a osobno jej donicy (utożsamianej z wazonem oraz wazą). W zdigitalizowanym mieszkaniu niektóre obiekty ze zbioru etykiet *CODE55* nie występują również w ogóle. Są to np. kominek (palenisko), samochód, wiatrak, mikrofalówka czy też bojler elektryczny. Zestawienie liczności poszczególnych rodzajów obiektów, posortowane malejąco, prezentowane jest w Tab. 5.1.

Tabela 5.1: Zestawienie liczby obiektów określonego rodzaju (zgodnych z etykietami *CODE55*) w całym środowisku testowym.

Typ obiektu	Liczba wystąpień	Typ obiektu	Liczba wystąpień
Lampa	30	Piekarnik	1
Szafka	14	Toster	1
Sensor alarmowy	14	Łóżko	1
Gniazdko elektryczne	13	Wanna	1
Włącznik	11	Prysznic	1
Wazon	11	Toaleta	1
Roślina	10	Akwarium	1
Dywan	9	Biurko	1
Okno	7	Czajnik	1
Roleta	6	Oczyszczacz powietrza	1
Lustro	6	Ekspres do kawy	1
Drzwi	6	Kanapa	1
Krzesło	5	Obraz	1
Klimatyzacja	5	Zmywarka	1
TV	4	Lodówka	1
Grzejnik	4	Kuchenska	1
Stół	4	Telefon	1
Głośnik	4	Konsola do gier	1
Półka	4	Komputer	1
Kran	3	Tablica	1
Zegar	3	Samochód	0
Poduszka	3	Wiatrak	0
Gaśnica	2	Kominek	0
Pralka	2	Bojler elektryczny	0
Suszarka na pranie	2	Mikrofalówka	0
Zlew	2	Suszarka do włosów	0
Szafa	2	Zasłona	0
Szafka nocna	2		
↪	↪		

Analiza wartości agregowanych również przedstawia dużą różnorodność. Pomiedzy poszczególnymi pokojami występuje nawet ponad pięciokrotny rozrzut liczby znajdujących się w nich obiektów. Najmniej oznaczeń naniesiono w korytarzu (15), a najwięcej w salonie (88). Ponownie, obserwacje potwierdzają intuicję. Korytarz,

jako pomieszczenie małe oraz o wtórnym znaczeniu, nie posiada zbyt wielu przypisanych przedmiotów. Dokładnie odwrotna obserwacja dotyczy salonu, pomieszczenia dużego, o wszechstronnych funkcjach i takiego, gdzie mieszkańcy spędzają najwięcej czasu. Wykres słupkowy licznosci oznaczonych obiektów w poszczególnych pomieszczeniach przedstawiono na Rys. 5.14.



Rysunek 5.14: Wykres słupkowy liczby obiektów obecnych w konkretnych pokojach całego środowiska testowego.

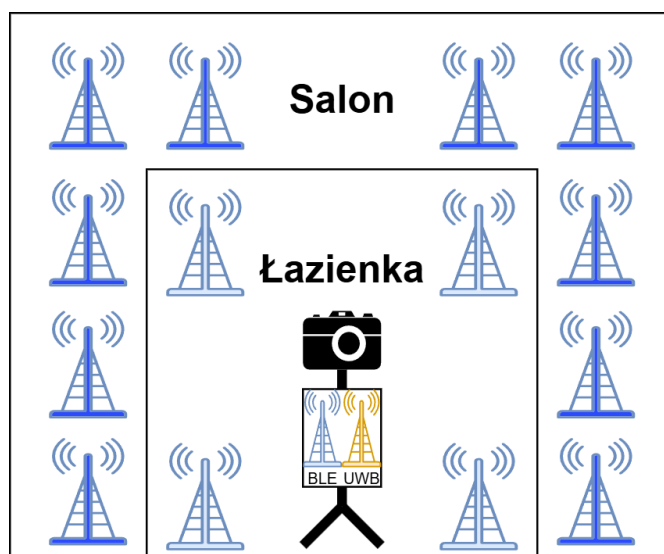
5.4 Pozycjonowanie ogólne

Pierwszy etap metody hybrydowej, określanany jako pozycjonowanie ogólne, pełni funkcję wstępnej filtracji przestrzeni i umożliwia realizację podejścia “od góry do dołu” (czy też: “od ogółu do szczegółu”). Jego celem jest określenie położenia lokalizatora z dokładnością do pojedynczego pomieszczenia. Do realizacji proponuje się wykorzystanie technologii BLE, ze względu na jej powszechną obecność w urządzeniach wyposażenia domowego, niskie koszty wdrożenia oraz potencjalną możliwość integracji z istniejącą infrastrukturą. Badania prototypu wykazały, że choć BLE nie nadaje się do realizacji precyzyjnego pozycjonowania,

to zapewnić może skuteczne monitorowanie otoczenia w poszukiwaniu dostępnych do komunikacji punktów referencyjnych. Na tej podstawie rozstrzygnąć możemy, czy lokalizator znajduje się na danym obszarze, czy też nie. Efektywnie uzyskujemy zatem pożądaną dokładność na poziomie poszczególnych pokoi. Opracowany algorytm gromadzi odczyty wskaźnika RSSI z sygnałów odbieranych przez lokalizator, a następnie agreguje je w celu nadania pojedynczej etykiety pokoju. Odbywa się to na podstawie analizy średnich wartości mocy sygnału z otaczających kotwic. Wynikiem etapu pozycjonowania ogólnego jest określenie identyfikatora pomieszczenia oraz przesłanie go w postaci wiadomości MQTT na ustalony temat komunikacji do lokalnego serwera.

W zaproponowanym rozwiązaniu algorytm postępowania zakłada komunikowanie jedynie zmiany pokoju (lub brak komunikacji, jeżeli lokalizator stale przebywa w tym samym pomieszczeniu), a także że znane są oczekiwane punkty referencyjne, które wstępnie przypisano do konkretnych pokoi. Zachowano jednosekundowe okno skanowania sygnałów rozgłoszeniowych kotwic, zaimplementowano jednak dodatkową metodę agregacji predykcji pokoju. W jej pierwotnej postaci zliczano ilościowo otrzymywane wiadomości, a następnie grupowano po pokojach. Najczęściej pojawiający się wynik uznawany był za aktualne położenie lokalizatora. Rozwiązanie to było jednak obciążone zauważalnymi problemami, wynikającymi z nierównomiernego rozłożenia kotwic po pokojach, a także zmienności aranżacji wnętrz. Szczególnie widoczne stało się to w pomieszczeniach blisko otoczonych innymi obszarami. Problem przedstawiono graficznie na Rys. 5.15. Urządzenie pomiarowe, mimo że blisko otoczone kotwicami jednego z pomieszczeń, co do liczby uzyskanych sygnałów może wskazywać inne jako swoje położenie.

Algorytm zmieniono i zaimplementowano rozwiązanie oparte na średnim odczycie wartości RSSI spośród kotwic danego pomieszczenia. Zakłada się, że ściany pomieszczeń wprowadzają na tyle silne tłumienie sygnału, że nawet znikoma liczba, ale silnych odczytów stanowi rzetelniejszą referencję. W celu dodatkowej stabilizacji wyników predykcje są analizowane w określonym przedziale czasowym, a komunikat o zmianie położenia publikowany jest dopiero po utrzymaniu się wskazań odpowiadających innemu pomieszczeniu niż pierwotnie. Wprowadza to pewną konfigurowalną bezwładność systemu i poprawia przewidywalność wyników. Dostrzega się potencjał dalszego rozwoju proponowanej metody i np. zastosowanie do wyzna-



Rysunek 5.15: Wizualizacja problemu ustalenia położenia na podstawie liczby odbieranych sygnałów radiowych w scenariuszu pomieszczenia-enklawy.

czania pomieszczenia klasyfikatorów opartych na uczeniu maszynowym. Algorytm postępowania metody wyznaczania i raportowania pokoju przedstawiono w Alg. 3. Procedura rozpoczyna się od nasłuchu na nadchodzące rozgłoszenia od okolicznych kotwic (dane wejściowe), a następnie grupowane są uzyskane wartości zgodnie z pokojami (linie 3-8), w których były one umieszczone. Tworzony jest w ten sposób słownik, w którym przechowane są listy odczytów wartości RSSI dla poszczególnych obszarów. Brak odczytów oznacza brak zmiany pokoju (linie 9-11). Zgrupowanie pomiarów umożliwia obliczenie średniej, a tym samym określenie, czy istnieje inny pokój (niż obecnie najlepszy) o mocniejszej średniej sile odbieranego sygnału (linie 12-18). Jeżeli wykryto lepszego kandydata na aktualną pozycję (linie 19-21), a sytuacja ta utrzymuje się przez ustalony czas (wybrano 10 cykli nasłuchu, po 1 sekundzie każdy, tj. 10 sekund łącznie) to publikowana jest na ustalonym temacie komunikacji serwera MQTT wiadomość o nowym aktualnym pomieszczeniu (linie 22-25). Dodatkowo informacja ta wypisywana jest również na port szeregowy, co pomogło w procesie uruchomienia systemu i wykrywania błędów (linia 26).

Algorytm 3 Algorytm wyznaczenia pokoju na podstawie odczytów RSSI.

Dane wejściowe: `anchorReadings` (lista odczytów: `address`, `rss`, `anchorId`, `distance`),
`anchorRoomMap` (słownik: `MAC` $\leftrightarrow$ `Room`), `lastEstimatedRoom`, `pendingRoom`,
`requiredConfirmations`, `mqttClient`, `mqttRoomTopic`

Dane wyjściowe: Ewentualna aktualizacja `lastEstimatedRoom` i publikacja komunikatu
`"RoomX"` przez MQTT

```

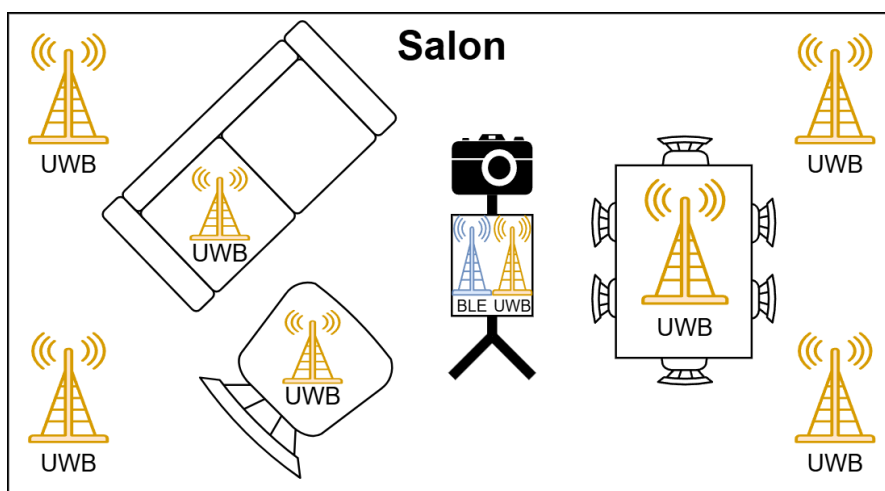
1: procedure ESTIMATEANDREPORTROOM
2:   roomRssiMap  $\leftarrow$  pusty słownik (Room  $\mapsto$  lista wartości RSSI)
3:   for all odczyt  $a$  w anchorReadings do                                 $\triangleright$  zbierz odczyty per pokój
4:     if  $a.rssi \neq 0$  then                                               $\triangleright$  ignoruj brak detekcji
5:        $r \leftarrow \text{anchorRoomMap}[a.address]$                          $\triangleright$  określ w którym pokoju była kotwica
6:       dołącz  $a.rssi$  do roomRssiMap[r]                                 $\triangleright$  wprowadź wpis do słownika
7:     end if
8:   end for
9:   if roomRssiMap jest puste then
10:    return                                                             $\triangleright$  brak nowych danych  $\Rightarrow$  brak zmiany pokoju
11:  end if
12:   $bestRoom \leftarrow lastEstimatedRoom$ ;  $bestMean \leftarrow -\infty$ 
13:  for all para  $(r, L)$  w roomRssiMap do                                 $\triangleright$  analizuj pokój po pokoju
14:     $\mu \leftarrow \frac{1}{|L|} \sum_{x \in L} x$                                  $\triangleright$  oblicz średni RSSI dla pokoju  $r$ 
15:    if  $\mu > bestMean$  then                                             $\triangleright$  sprawdź czy lepszy
16:       $bestMean \leftarrow \mu$ ;  $bestRoom \leftarrow r$                      $\triangleright$  zaktualizuj
17:    end if
18:  end for
19:  if  $bestRoom \neq lastEstimatedRoom$  then                                 $\triangleright$  jeżeli zmiana predykcji
20:    if  $bestRoom = pendingRoom$  then                                 $\triangleright$  potwierdź stabilność zmiany
21:      confirmCounter  $\leftarrow$  confirmCounter + 1
22:      if confirmCounter  $\geq$  requiredConfirmations then                 $\triangleright$  jeżeli stabilna
23:         $lastEstimatedRoom \leftarrow bestRoom$ ; confirmCounter  $\leftarrow$  0
24:         $message \leftarrow \text{"Room\{bestRoom+1\}"}$                      $\triangleright$  przygotuj komunikat
25:        mqttClient.publish(mqttRoomTopic, message)                 $\triangleright$  publikuj wiadomość
26:        Wypisz na port szeregowy: "Estimated Room: {message}"       $\triangleright$  wypisz log
27:      end if
28:    else                                                             $\triangleright$  jeżeli zmiana niestabilna
29:       $pendingRoom \leftarrow bestRoom$ ; confirmCounter  $\leftarrow$  0       $\triangleright$  zmiana kandydata
30:    end if
31:  end if
32: end procedure

```

5.5 Pozycjonowanie precyzyjne

Pozycjonowanie precyzyjne stanowi drugi etap metody hybrydowej. Odbywa się na wstępnie zawężonym obszarze, ograniczonym do jednego pomieszczenia. Jego bezpośrednim celem jest wskazanie konkretnych współrzędnych trójwymiarowych (w układzie odniesienia opracowanego modelu cyfrowego całej nieruchomości), w których znajduje się urządzenie pomiarowe. Do realizacji tego kroku konieczne jest zatem użycie technologii radiowej uzyskującej wysoką dokładność i precyzję pozycjonowania, nawet kosztem ograniczonego zasięgu. Ponownie, bazując na badaniach prototypów, do realizacji tej części wybrano technologię UWB. Wyższy koszt pojedynczych modułów elektronicznych rekompensowany jest możliwością zmniejszenia ich liczby i wprowadzenia tylko do tych pomieszczeń, w których faktycznie konieczna jest lokalizacja precyzyjna. Po wstępnej detekcji pomieszczenia w oparciu o system BLE, lokalizator może wybrać znane sobie konfiguracje dostępnych kotwic systemu UWB i prowadzić komunikację tylko z tymi modułami, od których spodziewa się uzyskania wartościowych informacji. Pozwala to na zmniejszenie obciążenia urządzenia, a także wstępną filtrację pomiarów odległości. Wiedząc, że pewna kotwica znajduje się poza interesującym nas pomieszczeniem, pominąć możemy komunikację z nią, a potencjalny wyznaczony odczyt odległości traktować jako z założenia obarczony dużym błędem.

W razie braku komponentu wizyjnego we wdrażanym systemie, pozycjonowanie precyzyjne może być końcowym punktem działania całego rozwiązania. Wymagane jest jednak wtedy wyposażenie każdego interesującego nas obiektu w dedykowany lokalizator radiowy oraz manualne nadanie mu identyfikatora rodzaju obiektu. Każdy tak oznaczony obiekt będzie pełnił rolę własnego lokalizatora w przestrzeni i sam będzie nawiązywał połączenia z otaczającymi go kotwicami. Nie spełni to ani realizacji zadania analizy całego modelu cyfrowego, ani nie będzie systemem realizującym metodę hybrydową. Przy tak dużym narzucie w liczbie urządzeń radiowych można jednak w zamian uzyskać efekt stałego monitorowania położenia zasobów (tzw. trybu “online”). System taki swoje zastosowanie odnaleźć może w scenariuszach nadzoru obiektów krytycznie istotnych lub o bardzo dużej wartości, gdzie wysoką wagę nadaje się częstotliwości odświeżania informacji. Symbolicznie przedstawiono to na Rys. 5.16.



Rysunek 5.16: Scenariusz stałego radiowego monitorowania położenia obiektów środowiska domowego w pełni wyposażonych w lokalizatory UWB.

Proponowane w metodzie hybrydowej postępowanie jest, co do zasady, analogiczne do zrealizowanego prototypu systemu pozycjonowania radiowego w technologii UWB (przedstawionego w Roz. 3.2.2). Pomiary odległości od nadajnika do odbiornika wykonano poprzez implementację metody DS-TWR (omówionej w Roz. 3.1.1) z zastrzeżeniem zmiany końcowego wykorzystanego wzoru na czas przelotu sygnału radiowego. Efektywnie, zrealizowane fizycznie rozwiązanie oparte zostało na metodzie *Alternative DS-TWR*, które przedstawione zostało w [332]. Jego wybór uzasadniony jest wykorzystanymi układami elektronicznymi oraz sugerowanym przez producenta rozwiązaniem. Powyższa metoda alternatywna zaproponowana została przez pracowników firmy Decawave, tj. producenta wykorzystanych układów UWB i stanowi domyślną implementację wykorzystywaną w dokumentacji i przykładach wyznaczania dystansu.

Na jej podstawie zmianie podlega wzór 3.5, który w domyślnej postaci przedstawiony był jako:

$$T_{\text{tof}} = \frac{T_{\text{cykl}_A} - T_{\text{proc}_K} + T_{\text{cykl}_B} - T_{\text{proc}_L}}{4}, \quad (5.1)$$

gdzie T_{cykl_A} i T_{cykl_B} oznaczały całkowite czasy trwania cykli A i B, a T_{proc_K} i T_{proc_L} oznaczały czasy przetwarzania wiadomości, kolejno: w kotwicy i w lokalizatorze.

Zastąpiony on zostaje różnicą iloczynów dzielonych przez sumę:

$$T_{\text{tof}} = \frac{T_{\text{cyklA}} \cdot T_{\text{cyklB}} - T_{\text{procK}} \cdot T_{\text{procL}}}{T_{\text{cyklA}} + T_{\text{procK}} + T_{\text{cyklB}} + T_{\text{procL}}}, \quad (5.2)$$

gdzie wykorzystane oznaczenia pozostają bez zmian, tak samo, jak pozostałe wykonywane operacje, ich kolejność, a także rezultaty.

Nie dostrzega się celowości przedstawiania zrealizowanej implementacji w postaci pseudokodów napisanych programów, jako że podążają one ściśle za przykładami dostępnymi w repozytorium kodu producenta [333]. Zostają one pominięte, a ze szczegółami przedstawia się wykorzystanie algorytmu trilateracji, który zintegrowany został z całą metodą hybrydową i posłużył do końcowego wyznaczenia współrzędnych trójwymiarowych.

Podczas prowadzenia eksperymentów oraz faktycznego użycia mobilnych urządzeń pomiarowych, dla każdego z interesujących nas punktów zbierane są całe serie wyznaczonych odległości do kotwic, a nie pojedynczy zestaw danych. Każdy cykl pomiarów trwa domyślnie jedną minutę. Przez ten czas wszystkie wyznaczone odległości do kotwic raportowane są przez lokalizator do lokalnego serwera MQTT z użyciem ustalonego tematu komunikacji oraz odpowiednio sformatowanej wiadomości. Sposób realizacji serwera nie jest istotny, a w badaniach wykorzystano w tym celu minikomputer Raspberry. Serwer agreguje gromadzone dla danej sesji pomiarowej wyniki, a realizowane pozycjonowanie opiera się na wartościach średnich, obliczonych na podstawie wszystkich zebranych pomiarów. Stanowi to bardzo uproszczoną realizację, w której dostrzega się daleko idący potencjał dalszych badań i rozwoju, głównie w kierunku filtrowania pomiarów. Dla tak zagregowanych odczytów odległości, końcowe współrzędne lokalizatora wyznacza się zgodnie z zasadami opisanymi w Roz. 3.1.2, tj. przy użyciu metody trilateracji. Jej opracowanie matematyczne wraz z wyprowadzeniem algebraicznym omówiono już wcześniej, dla podsumowania opisu przedstawia się więc jedynie pseudokod programu, którym realizowano faktyczne obliczenia (Alg. 4).

Algorytm 4 Algorytm trilateracji z $n \geq 3$ odległości, składnia języka Python.

Dane wejściowe: P (lista pozycji kotwic: $P_i = (x_i, y_i, z_i)$), S (zmierzone odległości r_i),
(opcj.) W (macierz wag, domyślnie $W = I$)

Dane wyjściowe: $(\mathbf{x}_{\text{vec}}^{(1)}, \mathbf{x}_{\text{vec}}^{(2)})$ $\triangleright$ wyznaczone położenie, $\mathbf{x}_{\text{vec}} = [x_0, x, y, z]^T$

```

1:  $(m_p, n_p) \leftarrow \text{shape}(P)$ ;  $n_s \leftarrow \text{length}(S)$   $\triangleright$  odczyt kształtu danych wejściowych
2: if  $m_p \neq 3$  lub  $n_s \neq n_p$  then  $\triangleright$  kontrola spójności wymiarów
3:   raise ValueError
4: end if

5:  $A \leftarrow \text{zeros}(n_p, 4)$ ;  $\mathbf{b} \leftarrow \text{zeros}(n_p)$   $\triangleright$  Konstrukcja układu liniowego  $A\mathbf{x}_{\text{vec}} = \mathbf{b}$ , zgodnie
  z (3.12)
6: for  $i \leftarrow 1$  to  $n_p$  do
7:    $(x_i, y_i, z_i) \leftarrow P[:, i]$ ;  $r_i \leftarrow S[i]$ 
8:    $A[i, :] \leftarrow [1, -2x_i, -2y_i, -2z_i]$ 
9:    $\mathbf{b}[i] \leftarrow r_i^2 - (x_i^2 + y_i^2 + z_i^2)$ 
10: end for

11: if  $n_p \leq 3$  then
12:   return  $(\text{null}, \text{null})$   $\triangleright$  trilateracja wymaga co najmniej trzech pomiarów

13: else if  $n_p = 3$  then  $\triangleright$  wyznaczanie z 3 pomiarów, zgodnie z (3.16)
14:    $\mathbf{x}_p \leftarrow \text{lstsq}(A, \mathbf{b})$   $\triangleright$  wyznaczenie rozwiązania szczególnego
15:    $Z \leftarrow \text{null\_space}(A)$   $\triangleright$  wyznaczenie zbioru rozwiązań jednorodnych
16:   if  $\text{rank}(A) < 3$  lub  $Z$  nie ma dokładnie jednej kolumny then
17:     return  $(\text{null}, \text{null})$   $\triangleright$  zdegenerowana geometria, np. kotwice współliniowe
18:   end if
19:    $\mathbf{x}_h \leftarrow Z[:, 1]$   $\triangleright$  przypisanie wektora rozwiązania jednorodnego

    $\triangleright$  parametryzacja prostej rozwiązań:  $\mathbf{x}(t) = \mathbf{x}_p + t\mathbf{x}_h$ , zgodnie z (3.17)
    $\triangleright$  podstawienie:  $x_0 = x^2 + y^2 + z^2$ , równanie kwadratowe zgodnie z (3.23)
20:    $a \leftarrow \|\mathbf{x}_h[2:4]\|_2^2$   $\triangleright x_h^2 + y_h^2 + z_h^2$ 
21:    $b \leftarrow 2\mathbf{x}_p[2:4]^T \mathbf{x}_h[2:4] - \mathbf{x}_h[1]$   $\triangleright 2(x_p x_h + y_p y_h + z_p z_h) - x_{h0}$ 
22:    $c \leftarrow \|\mathbf{x}_p[2:4]\|_2^2 - \mathbf{x}_p[1]$   $\triangleright x_p^2 + y_p^2 + z_p^2 - x_{p0}$ 

23:    $(t_1, t_2) \leftarrow \text{roots}(a t^2 + b t + c)$   $\triangleright$  wyznaczenie pierwiastków równania
24:    $\mathbf{x}^{(1)} \leftarrow \mathbf{x}_p + t_1 \mathbf{x}_h$ ;  $\mathbf{x}^{(2)} \leftarrow \mathbf{x}_p + t_2 \mathbf{x}_h$   $\triangleright$  określenie wyników
25:   return  $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ 

26: else if  $n_p > 3$  then  $\triangleright$  Więcej pomiarów niż wymagane minimum. Szczegóły w [231]
27:   if  $W$  podano i nie jest równa  $I$  then
28:      $C \leftarrow W^T W$   $\triangleright$  ogólna metryka wag
29:      $\mathbf{x}_{\text{WLS}} \leftarrow (A^T C A)^\dagger A^T C \mathbf{b}$   $\triangleright$  estymuj ważoną metodą najmniejszych kwadratów
30:   else
31:      $\mathbf{x}_{\text{WLS}} \leftarrow A^\dagger \mathbf{b}$   $\triangleright$  estymuj zwykłą metodą najmniejszych kwadratów (bez wag)
32:   end if
33:   return  $(\mathbf{x}_{\text{WLS}}, \mathbf{x}_{\text{WLS}})$   $\triangleright$  dla  $n_p > 3$  mamy jedno rozwiązanie, zwracamy dwa identyczne
34: end if

```

Podobnie jak dla pozycjonowania ogólnego, wynik kroku pozycjonowania precyzyjnego może być uzyskiwany różnymi metodami, co pozwala na elastyczność w realizacji tego zadania. W sytuacjach, gdy proponowany system UWB jest niedostępny, nie działa prawidłowo lub jego jakość jest niewystarczająca, możliwe jest zastosowanie scenariusza ręcznego określenia położenia obiektów. Choć rozwiązanie to przeczy idei pełnej automatyzacji, zapewnia kontynuację metody hybrydowej i umożliwia uzupełnienie braków przy zastosowaniu pracy manualnej. Użycie ręcznego pozycjonowania pozwala również na eliminację błędów obecnych w radiowej estymacji położenia oraz daje możliwość szybkiej weryfikacji i korekty wyników automatycznych.

5.6 Pozycjonowanie wizyjne

Finalnym krokiem analizy środowiska jest pozycjonowanie wizyjne. Wiedząc, gdzie znajduje się mobilne urządzenie pomiarowe (którego położenie wyznaczono poprzez pozycjonowanie precyzyjne), wykonać możemy zdjęcia sferyczne otaczającego nas środowiska, a następnie wykorzystać wizję komputerową do detekcji obiektów oraz rozmieszczenia ich w przestrzeni. Zagadnienie odwzorowywania obrazów sferycznych w innych formatach omówione zostało w Roz. 4.1.3. Pominęto tam jednak szczegóły matematyczne, które teraz są konieczne do wyjaśnienia przyjętego sposobu postępowania.

Opisując reprezentację sferyczną, posłużyć możemy się trzema parametrami, które definiują punkt na sferze w następujący sposób:

Definicja 4. Niech $\mathbf{P}_{\text{sfer}}(\theta, \phi, \rho)$ oznacza punkt opisany w sferycznym układzie współrzędnych, taki że:

- $\theta \in (0, 2\pi)$ to kąt azymutu, odpowiadający długości geograficznej (lub szerokości obrazu w rozumieniu perspektywicznym),
- $\phi \in (0, \pi)$ to kąt elewacji, odpowiadający szerokości geograficznej (lub wysokości obrazu w rozumieniu perspektywicznym),
- $\rho \in \mathbb{R}_{>0}$ to promień sfery, czyli odległość punktu od początku układu współrzędnych (charakterystyczna dla obrazowania sferycznego).

Jednocześnie ten sam punkt odwzorować możemy w klasycznej trójwymiarowej przestrzeni kartezjańskiej i oznaczyć jako $\mathbf{P}_{\text{kart}}(x_{\text{kart}}, y_{\text{kart}}, z_{\text{kart}})$. Koordynaty przeliczamy swobodnie pomiędzy obiema reprezentacjami, zgodnie z zależnościami zaczerpniętymi z [334]:

$$\begin{cases} \rho \\ \theta \\ \phi \end{cases} = \begin{cases} (x_{\text{kart}}^2 + y_{\text{kart}}^2 + z_{\text{kart}}^2)^{1/2} \\ \arctan(x_{\text{kart}}/z_{\text{kart}}) \\ \arccos(y_{\text{kart}}/\rho) \end{cases}, \quad \begin{cases} x_{\text{kart}} \\ y_{\text{kart}} \\ z_{\text{kart}} \end{cases} = \begin{cases} \rho \sin(\theta) \sin(\phi) \\ \rho \cos(\phi) \\ \rho \cos(\theta) \sin(\phi) \end{cases}. \quad (5.3)$$

Tak określony punkt na sferze posiada właściwie nieskończoną dokładność reprezentacji, jako że układ współrzędnych jest układem ciągłym. Do przetwarzania w systemach wizji komputerowej niezbędna jest jednak reprezentacja dyskretna. Stosowane w pracy rzutowanie panoramy równokątnej (ERP) wprowadza do procesu dwa nowe parametry (kąty jednostkowe) oraz systematyzuje nazewnictwo położenia pikseli na panoramie (zmienne $\mathbf{u}_{\text{pano}}$ i $\mathbf{v}_{\text{pano}}$). Opis na podstawie [298]:

Definicja 5. Niech $\mathbf{w}_{\text{pano}}$ oraz $\mathbf{h}_{\text{pano}}$ oznaczają odpowiednio szerokość i wysokość docelowego obrazu panoramicznego w odwzorowaniu równokątnym (ERP). Wtedy kąty jednostkowe wyznaczające rozdzielczość kątową w obu kierunkach opisujemy jako:

$$\vartheta_{\text{unit}} = \frac{2\pi}{w_{\text{pano}}}, \quad \varphi_{\text{unit}} = \frac{\pi}{h_{\text{pano}}}, \quad (5.4)$$

gdzie ϑ_{unit} jest kątem jednostkowym poziomym, a φ_{unit} kątem jednostkowym pionowym.

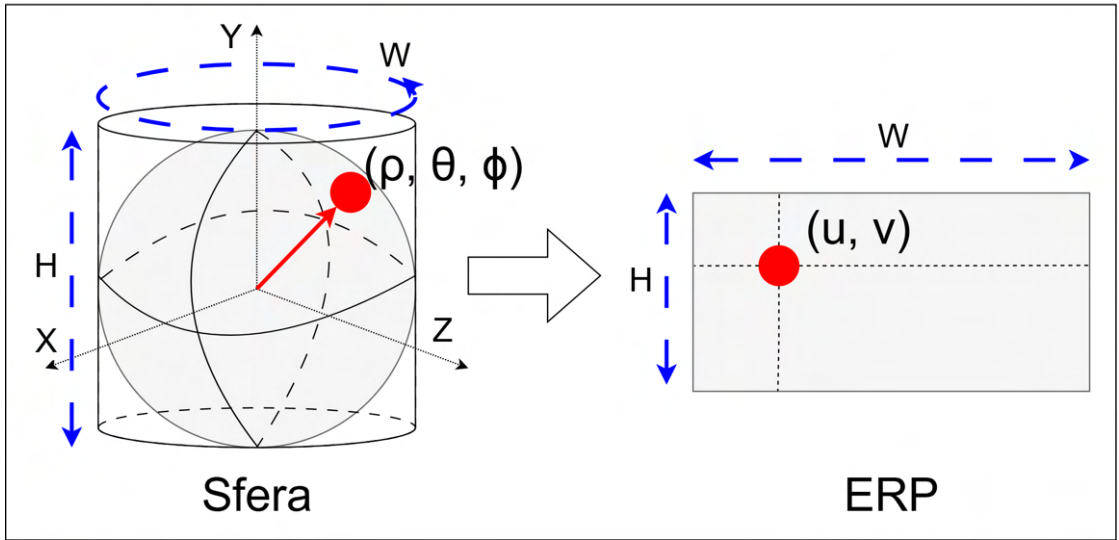
Definicja 6. Niech $\mathbf{u}_{\text{pano}} \in \{0, 1, \dots, w_{\text{pano}} - 1\}$ oraz $\mathbf{v}_{\text{pano}} \in \{0, 1, \dots, h_{\text{pano}} - 1\}$ oznaczają odpowiednio współrzędne pikseli w odwzorowaniu równokątnym (ERP), gdzie:

- $\mathbf{u}_{\text{pano}}$ określa pozycję w poziomie (kolumnę obrazu),
- $\mathbf{v}_{\text{pano}}$ określa pozycję w pionie (wiersz obrazu).

Zakładając rzutowanie na sferę o promieniu jednostkowym, każdemu pikselowi o współrzędnych (u_{pano}, v_{pano}) w obrazie ERP można wtedy przypisać odpowiadające współrzędne sferyczne:

$$(\theta, \phi) = (u_{pano} \cdot \vartheta_{unit}, v_{pano} \cdot \varphi_{unit}). \quad (5.5)$$

Uzyskujemy w ten sposób skończoną liczbę punktów do odwzorowania i późniejszego przetwarzania. Położenie punktu w obu reprezentacjach przedstawione zostało symbolicznie na Rys. 5.17.

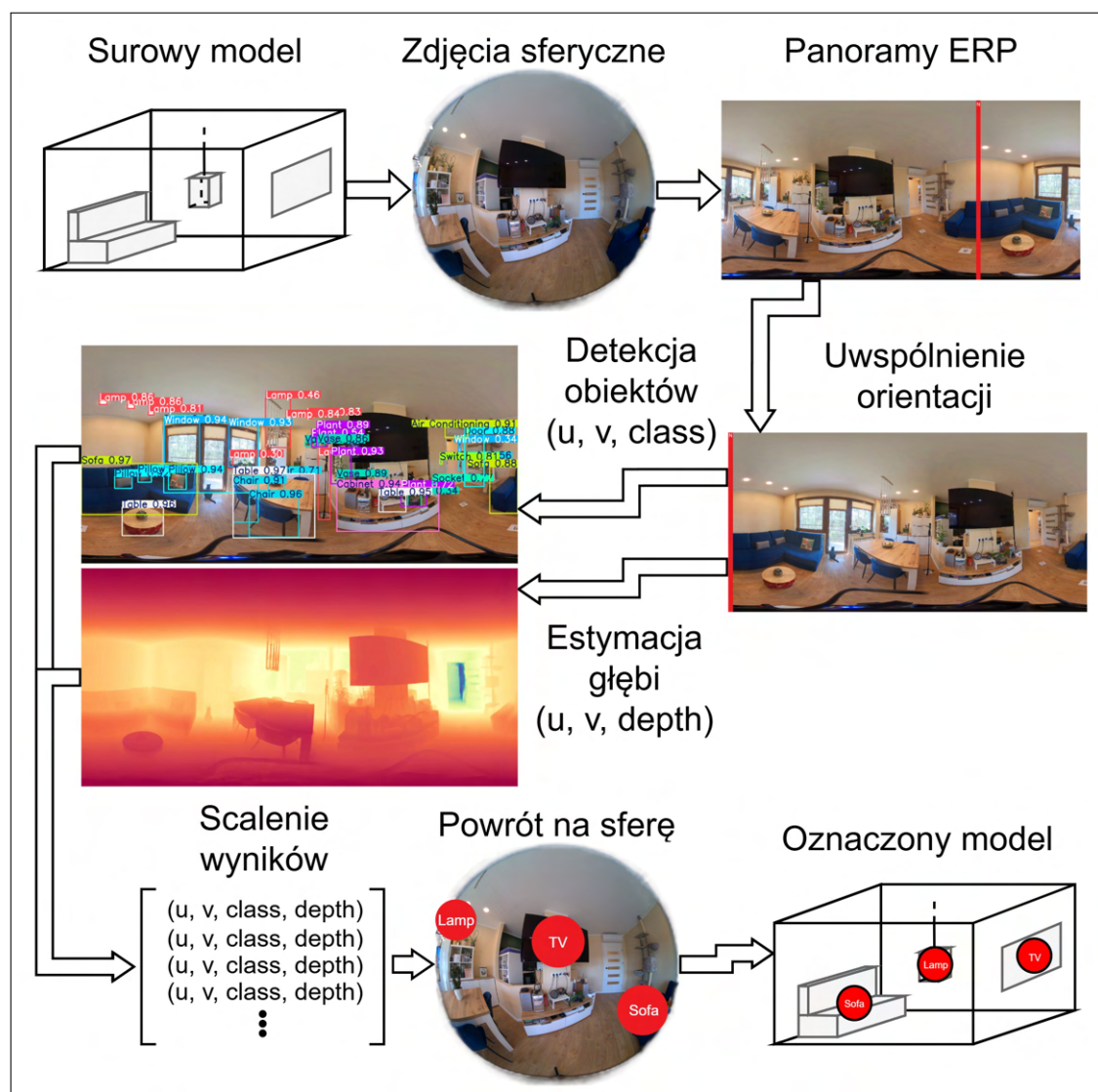


Rysunek 5.17: Wizualizacja układów współrzędnych i oznaczenia analogicznego punktu na sferze (po lewej) i na panoramie ERP (po prawej).

Mając tak zarysowany kontekst oznaczeń i opracowania matematycznego, przejść możemy do właściwego opisu postępowania. Uzyskana procedura wizyjna przedstawiona zostaje schematycznie na Rys. 5.18. Scala ona znakomitą część całej rozprawy i odwołuje się do wcześniej opisanych zagadnień, takich jak:

- modelu cyfrowego bliźniaka i jego reprezentacji,
- obrazowania sferycznego,
- przetwarzania panoram ERP,

- oczekiwania użytkowników,
- detekcji obiektów (sieci neuronowych, szkolenia, zbiorów danych),
- estymacji głębi.



Rysunek 5.18: Schemat realizacji wizyjnego rozpoznawania i lokalizowania komponentów modelu cyfrowego bliźniaka.

Przetwarzanie rozpoczyna się od założenia istnienia pewnego modelu trójwymiarowego cyfrowego bliźniaka nieruchomości, którego semantyczne znaczenie chcemy odzyskać (Roz. 1.1.2). W omawianych realiach przemysłowych będzie to model zdigitalizowany z użyciem technologii spaceru interaktywnego firmy Matterport.

Z tak przygotowanego modelu pozyskujemy następnie obrazy sferyczne. W zależności od faktycznego dostępnego ich formatu i technicznej realizacji, proces ten może być dwustopniowy (akwizycja zdjęć sferycznych, a następnie rzutowanie na panoramy ERP) albo jednostopniowy (jeżeli system źródłowy od razu udostępnia panoramy ERP). Panoramy te przedstawiają najbliższe otoczenie kamery, którą zostały zarejestrowane. W sytuacji wykorzystania dedykowanych skanerów przestrzeni są one automatycznie sprowadzane do wspólnej orientacji w przestrzeni. Na schemacie południk początkowy sfery oznaczono czerwonym paskiem. W przypadku zbierania zdjęć ręcznie i z użyciem powszechnie dostępnych kamer sferycznych, uzyskane zdjęcia trzeba będzie manualnie sprowadzić do wspólnej orientacji. W przeciwnym razie, końcowe wyniki pozycjonowania będą dla każdej panoramy odniesione tylko do jej własnego układu współrzędnych. Postępować można tu analogicznie do procesu przewijania panoram, omówionego w badaniach z Roz. 4.2.2.

Gdy zgromadzony zbiór obrazów typu ERP jest już gotowy, postępowanie można zrównoleglic. Jednocześnie uzyskać potrzebujemy dwa rodzaje informacji: detekcję obiektów oraz estymację głębi. Do realizacji zadania wykrywania i rozpoznawania obiektów wykorzystane zostaną wnioski z badań dotyczących modeli sieci neuronowych (Roz. 4.2.1), a także analizy i opracowania zbiorów danych szkoleniowych (Roz. 4.2.3). Na potrzeby realizacji tego kroku scalono wszystkie dostępne zbiory danych, przetworzono z użyciem procedury przewijania panoram, a następnie wykorzystano do wyszkolenia modelu sieci neuronowej architektury YOLOv11 w wersji *Extra-Large*. Tak przygotowanego detektora użyto do wskazania zestawów informacji: położenia oraz etykiety wykrytych obiektów. Do uzyskania głębi obrazu wykorzystano opisany w ramach prototypu podsystemu wizyjnego detektor *Depth Anything V2* (Roz. 4.3). Zwrócił on mapę głębi, która dla każdego z pikseli obrazu wejściowego określiła jego metryczną odległość od kamery. Na mapę tę nałożono następnie wyniki detekcji obiektów i efektywnie utworzono listę czwórek wartości, opisujących wykryte na obrazie obiekty: położenie, etykietę i odległość od kamery.

Tak opracowany zestaw wyników przeniesiono następnie do reprezentacji sferycznej. Zakładając sferę docelową jako o jednostkowym promieniu, wyznaczano wektor kierunkowy do zlokalizowanego obiektu o początku w początku układu współrzędnych sfery. Następnie przemieszczano się wzdłuż tego wektora na odległość określoną z estymacji głębi, efektywnie uzyskując efekt położenia w trójwymiarze.

Posiadając zestawy wykrytych dla każdej ze sfer obiektów, a także zestawiając ich układy współrzędnych z wynikami pozycjonowania precyzyjnego z poprzedniego kroku metody hybrydowej, wyznaczyć można końcowe położenie dla każdego wykrytego obiektu w układzie współrzędnych modelu cyfrowego bliźniaka. W ten sposób kończy się proces detekcji wizyjnej, a wejściowy model wzbogacony zostaje o semantyczne znaczenie jego elementów. Stanowi to esencję idei metody hybrydowej, która, bazując na odczytanych radiowo położeniach kamery, scala wyniki z przetwarzaniem wizyjnym i redukuje nakłady pracy manualnej użytkowników.

Wspomnieć należy, że ten ostatni element pozycjonowania wizyjnego, tj. zestawienie układów współrzędnych, może być oparty zarówno na wynikach radiowego pozycjonowania precyzyjnego (sposób automatyczny), jak i na ręcznie określonych współrzędnych (sposób manualny). Proponowana metoda do poprawnej realizacji wymaga jedynie znajomości położenia przetwarzanych sfer, nie jest jednak istotne, skąd te informacje pochodzą. Oznaczanie ręczne sfer jest operacją czasochłonną, lecz możliwą, znajdującą szczególne zastosowanie przy konieczności eliminacji błędów radiowej estymacji położenia.

Podsumowując, przy funkcjonowaniu w realiach systemu hybrydowego, końcową pozycję obiektu wyznaczyć możemy całkowicie automatycznie na dwa sposoby: poprzez pozycjonowanie radiowe (jeśli obiekt posiada dedykowany lokalizator radiowy) lub poprzez pozycjonowanie wizyjne (o ile pozycja kamery skanującej pomieszczenie została wcześniej określona metodą radiową lub wskazana ręcznie). W szczególności, gdy zarówno obiekt, jak i kamera są wyposażone w lokalizatory, możliwe jest zatem jednoczesne wyznaczenie pozycji obiektu obiema metodami, co umożliwia wzajemną weryfikację rezultatów oraz zwiększa wiarygodność wyników. Takie podejście pozwala elastycznie wybierać strategię pozycjonowania w zależności od dostępności urządzeń oraz akceptowalnych nakładów pracy manualnej.

5.7 Przeprowadzone eksperymenty końcowe

Celem przeprowadzonych badań końcowych było zweryfikowanie skuteczności opracowanej metody hybrydowej, ze szczególnym uwzględnieniem precyzji i stabilności lokalizacji obiektów w środowisku domowym. Istotną część badań stanowiła ocena stopnia automatyzacji procesu tworzenia cyfrowego modelu nieruchomości (tzw. Cyfrowego Bliźniaka), zwłaszcza w kontekście redukcji liczby manualnych operacji w konfiguracji systemu. Badania końcowe są zatem kontynuacją i integracją wszystkich wcześniej wykonanych eksperymentów oraz opracowanych prototypów, łącząc ich wyniki w celu całościowej oceny proponowanej metody hybrydowej. Analiza objęła porównanie wyników pozycjonowania uzyskiwanych z modułów radiowego pozycjonowania precyzyjnego oraz systemu wizyjnego, a także ocenę jakości detekcji i estymacji głębi przy wykorzystaniu opracowanych modeli i algorytmów.

5.7.1 Konfiguracja sprzętu i środowiska

Za przestrzeń badawczą wybrano w pełni umeblowane, dwupokojowe mieszkanie z osobną łazienką, kuchnią i łączącym je korytarzem. Opisane zostało ono ze szczegółami w Roz. 5.3. W tej przestrzeni wskazanych zostało łącznie 39 stanowisk pomiarowych (12 w salonie, 4 w łazience, 12 w sypialni, 7 w kuchni, 4 w korytarzu), w których dokonywano akwizycji danych z użyciem mobilnego urządzenia pomiarowego (składającego się z kamery sferycznej oraz lokalizatorów BLE i UWB, przedstawionego w Roz. 5.2). Nadano im kolejne numery porządkowe, np. P1, P2, P3, itd. Ich rozmieszczenie wykonano zgodnie z intuicyjną potrzebą pokrycia przestrzeni zdjęciami sferycznymi, efektywnie nierównomiernie, co do ich liczby per pomieszczenie. Na każde stanowisko składały się trzy charakterystyczne punkty pomiarowe (zestawy współrzędnych trójwymiarowych): położenie w płaszczyźnie podłogi, na wysokości lokalizatora radiowego oraz na faktycznej wysokości oka kamery. Różnice wynikają z konstrukcji urządzenia pomiarowego: trójnóg ustawiono na podłodze, lokalizator radiowy znajdował się na jego szczycie, tj. o 112,5 cm wyżej, a zamocowana na jego obudowie kamera dodatkowo 19 cm ponad nim. Efektywnie uzyskano w ten sposób trzy wysokości charaktery-

styczne, używane w późniejszym oznaczaniu punktów: 0 cm (np. P13_0), 112,5 cm (np. P13_1125) oraz 131,5 cm (np. P13_1315). W mieszkaniu rozmieszczone zostały również wszystkie 24 urządzenia kotwic radiowych oraz 8 statycznych lokalizatorów (wirtualnie reprezentujących zaopatrzone w lokalizatory radiowe aktywne obiekty inteligentnego środowiska domowego). Wybierając miejsce ich rozłożenia, a zatem i kolejne punkty charakterystyczne w środowisku, kierowano się przede wszystkim różnorodnością wysokości nad poziomem podłogi, widocznością na zdjęciach sferycznych, a dla statycznych lokalizatorów również zgodnością z konkretnym przedmiotem w przestrzeni, który miały reprezentować. W ten sposób w salonie umieszczono 8 kotwic i 3 lokalizatory, w łazience 5 kotwic i 2 lokalizatory, w sypialni 5 kotwic i jeden lokalizator, w kuchni 6 kotwic i dwa lokalizatory. W korytarzu, chcąc uzyskać dodatkowy scenariusz testowy, celowo nie umieszczono żadnych elementów radiowego systemu pomiarowego. Wszystkie tak opisane punkty charakterystyczne przestrzeni badawczej odnaleziono i oznaczono w opracowanym modelu trójwymiarowym cyfrowego bliźniaka. Zaprezentowano to na Rys. 5.19.



Rysunek 5.19: Wizualizacja modelu trójwymiarowego mieszkania w programie Blender, widok z naniesieniem punktów charakterystycznych. Sfery czerwone oznaczają kotwice systemu radiowego. Sfery niebieskie pokazują położenie punktu charakterystycznego w płaszczyźnie podłogi dla każdego ze stanowisk pomiarowych. Sfery zielone reprezentują położenie mobilnych lokalizatorów radiowych (punktów charakterystycznych na poziomie lokalizatorów umieszczonych na trójnogach). Sfery fioletowe to lokalizatory umiejscowione statycznie.

Dzięki wykorzystaniu programu Blender oraz dedykowanemu skryptowi eksportującemu dane, możliwe było odczytanie współrzędnych trójwymiarowych (w układzie współrzędnych modelu) wszystkich 149 tak oznaczonych punktów charakterystycznych i przedstawienie ich w postaci tabelarycznej. Przykład zgromadzonych informacji, dla pomieszczenia łazienki, pokazano w Tab. 5.2. Statyczne lokalizatory połączono z łatwo rozpoznawalnymi obiektami, silnie charakterystycznymi dla rodzaju pomieszczenia: toaletą oraz umywalką. Wartym zauważenia jest kolejny benefit zastosowania modelu cyfrowego, tj. łatwość wyznaczania położenia odstającego od normy. Jedno ze stanowisk pomiarowych wybrane zostało dosłownie w wannie, co efektywnie zmieniło jego wysokość nad podłogą o kilkadziesiąt centymetrów i zwiększyło spodziewane wysokości lokalizatora radiowego oraz kamery. Dla pozostałych pokoi punkty oznaczano analogicznie.

Tabela 5.2: Zestawienie współrzędnych punktów charakterystycznych dla przykładowego pomieszczenia łazienki.

Typ punktu	ID punktu	X [m]	Y [m]	Z [m]	Opis
Kotwica statyczna	AO	-5,30	-0,01	2,39	Statyw, poziom sufitu
	AP	-4,50	-2,26	0,81	Róg pod lustrem, obok wanny
	AQ	-4,50	-0,64	0,81	Róg przy umywalce
	AR	-6,05	-1,20	1,79	Pralka, górna powierzchnia
	AS	-5,68	-0,05	1,86	Grzejnik, góra
Lokalizator statyczny	TG	-4,68	-0,97	0,73	Umywalka
	TH	-5,98	-0,40	0,47	Toaleta
Podłoga	P13_0	-4,95	0,02	0,00	Stanowisko P13, poziom podłogi
	P14_0	-5,33	-0,67	0,00	Stanowisko P14, poziom podłogi
	P15_0	-5,34	-1,11	0,00	Stanowisko P15, poziom podłogi
	P16_0	-5,62	-1,93	0,20	Stanowisko P16, poziom podłogi
Pomiar radiowy	P13_1125	-4,95	0,02	1,13	Stanowisko P13, poziom lokalizatora
	P14_1125	-5,33	-0,67	1,13	Stanowisko P14, poziom lokalizatora
	P15_1125	-5,34	-1,11	1,13	Stanowisko P15, poziom lokalizatora
	P16_1125	-5,62	-1,93	1,33	Stanowisko P16, poziom lokalizatora
Obiektyw kamery	P13_1315	-4,95	0,02	1,32	Stanowisko P13, poziom kamery
	P14_1315	-5,33	-0,67	1,32	Stanowisko P14, poziom kamery
	P15_1315	-5,34	-1,11	1,32	Stanowisko P15, poziom kamery
	P16_1315	-5,62	-1,93	1,52	Stanowisko P16, poziom kamery

5.7.2 Przebieg eksperymentów

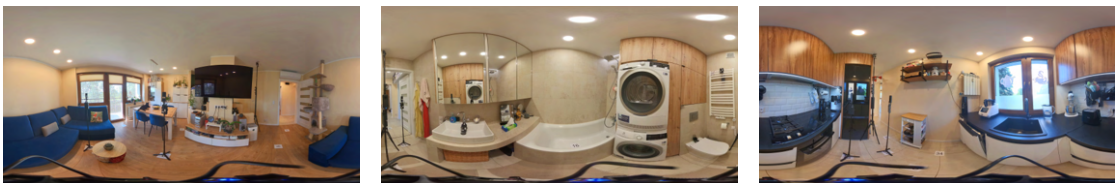
Procedura gromadzenia danych pomiarowych polegała na stopniowym przemieszczaniu się zestawu mobilnych urządzeń pomiarowych (kamery i lokalizatorów) od stanowiska do stanowiska, każdorazowo przeskakując o jedno stanowisko dalej, aż do całkowitego zapętlenia i powrotu do stanu wyjściowego. Wykorzystano cały dostępny potencjał sprzętowy, tj. trzy zestawy urządzeń pomiarowych, które umieszczane były jeden po drugim na kolejnych stanowiskach. W ten sposób przykładowym scenariuszem początkowym było umieszczenie kamery C1 na stanowisku P3, kamery C2 na P2, a kamery C3 na P1. Przedstawiono to na Rys. 5.20, tj. obrazie wykonanym kamerą C2 umieszczoną na stanowisku pomiarowym P2. W następnym kroku kamery przemieszczano na stanowiska o rosnących indeksach, tj. kamerę C1 na stanowisko P4, C2 na P3, C3 na P2. Uzyskano w ten sposób efekt krokowego przemieszczania się po stanowiskach pomiarowych, podobnego do tzw. “wężyka”.



Rysunek 5.20: Zdjęcie sferyczne w formacie ERP wykonane przez urządzenie pomiarowe C2 na stanowisku pomiarowym P2 z zaznaczonymi towarzyszącymi mu urządzeniami pomiarowymi C1 oraz C3.

Proces taki obrano ze względu na dostrzegany dalszy potencjał badawczy obranej metody i dodatkową jego automatyzację. W postępowaniu opisanym w ramach procedury pozycjonowania wizyjnego (Roz. 5.6), synchronizacja układu współrzędnych kolejnych zdjęć została przedstawiona jako osiągalna automatycznie (jeżeli system cyfryzacji przestrzeni robi to samodzielnie) lub wykonywana była ręcznie. Przyjęty sposób przemieszczania urządzeń umożliwia jednak dalszy rozwój projektu w kierunku automatyzacji synchronizacji układów współrzędnych nawet dla surowych, oderwanych od większego kontekstu zdjęć sferycznych. Dzięki znajomości współrzędnych stanowisk pomiarowych możliwe powinno być zautomatyzowanie wykrywania pozostałych kamer i dostosowanie przewinieć panoram do ich rozmieszczenia. Rozłożystość realizacji takiej operacji przekroczyła jednak zakres prowadzonych badań, głównie ze względu na niedobór zdjęć przedstawiających ten konkretny scenariusz i brak możliwości wyszkolenia satysfakcjonująco skutecznego detektora obiektów. W przeprowadzonych eksperymentach proces wykonano więc ręcznie. Niemniej, zebrany materiał badawczy może zostać wykorzystany w przyszłości.

Całość przebiegu akwizycji danych pomiarowych (dla każdego ze stanowisk i dla każdej z kamer) realizowana była w środowisku aktywnie działającego, dostępnego i raportującego położenie urządzeń mobilnych systemu pozycjonowania ogólnego BLE. Każde z urządzeń mobilnych, po umieszczeniu na stanowisku pomiarowym, uzyskiwało więc wiedzę o swoim położeniu na poziomie poszczególnych pokoi, a następnie rozpoczynało pozycjonowanie precyzyjne na podstawie znanych sobie kotwic podsystemu UWB. Akwizycji danych radiowych dokonywano każdorazowo przez okres jednej minuty. Takie okno czasowe pozwalało na zebranie satysfakcjonująco dużej liczby pomiarów, a podsumowywane było wykonaniem zdjęcia sferycznego. Przykłady wykonanych zdjęć prezentowane są na Rys. 5.21.



Rysunek 5.21: Przykładowe obrazy ze zbioru zdjęć wykonanych w ramach badań systemu hybrydowego.

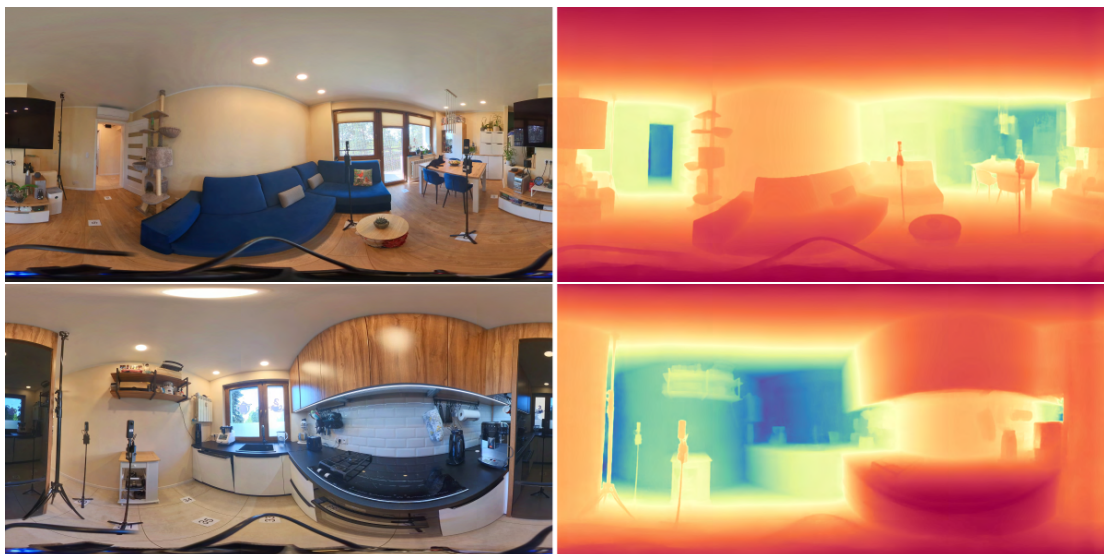
Po zakończeniu procedury dla jednego ze stanowisk przechodzono dalej, a całość powtarzano. Trwająca procedura pomiarowa zapisywała do plików pomiary odczytywane ze wszystkich aktywnych lokalizatorów systemu UWB, w tym lokalizatorów statycznie przypisanych do konkretnych przedmiotów. Efektywnie uzyskano zatem po 3 zestawy danych dla każdego z 39 stanowisk pomiarowych (łącznie 117 zestawów danych), a także po 39 zestawów pomiarów dla każdego z 8 lokalizatorów statycznych (kolejnych 312 zestawów danych). Łącznie na całość pozycjonowania radiowego w przeprowadzonych badaniach złożyło się 429 zestawów pomiarów.

Redundancja zdjęć sferycznych dla każdego ze stanowisk pomiarowych (każda z kamer wykonywała własne zdjęcie) sprawiła, że do końcowych badań wytypować można było subiektywnie najlepszej jakości zestaw. O ile użycie wszystkich pomiarów radiowych można było łatwo zintegrować, o tyle nadmiar zdjęć wprowadziłby niewspółmiernie duży narzut pracy dodatkowej, a sposób scalenia wyników ich analizy ciężko przewidzieć. Do końcowego zastosowania wytypowano zestaw zdjęć wykonanych przez kamerę C3. W porównaniu z pozostałymi kamerami, zdjęcia nią wykonane były najlepiej doświetlone, a także najbardziej ostre i wyraźne. Różnice, spodziewa się, że wynikają z faktu funkcjonowania wszystkich kamer w czasie badań w trybie automatycznego doboru parametrów, który w zależności od punktu początkowego mógł inaczej reagować, np. na rozbłyśki światła.

Po zakończeniu procesu pozycjonowania precyzyjnego i posiadając zgromadzony materiał zdjęciowy, przystąpiono do realizacji pozycjonowania wizyjnego. Dla każdego z wynikowych 37 zdjęć użyto dwóch rodzajów sieci neuronowych do dalszego przetwarzania: detektora obiektów oraz estymatora głębi. Przykładowe wyniki działania detektora obiektów prezentowane są na Rys.5.22, a estymatora głębi (w postaci wynikowych map głębi) na Rys. 5.23.



Rysunek 5.22: Przykładowe zdjęcia panoramiczne z badań systemu hybrydowego wraz z zaznaczonymi na nich detekcjami obiektów.



Rysunek 5.23: Przykładowe zdjęcia panoramiczne z badań systemu hybrydowego wraz z wyznaczonymi dla nich mapami głębi. Paleta barw określa odległość do obiektu, kolory ciepłe podkreślają krótki dystans, kolory zimne daleki dystans.

W końcowym kroku postępowania wykonano scalenie informacji oraz przeliczenie ich na końcowe położenia i etykiety obiektów. Przyjęto postępowanie zgodne z omówionymi w Roz. 5.6 wzorami, jednakże trudne w bieżącej kontroli poprawności. Aby zwizualizować otrzymywane wyniki, konieczne jest albo przenoszenie ich od razu do modelu trójwymiarowego, albo opracowanie innego narzędzia, które przedstawiłoby wyniki cząstkowe. Jako że wcześniej opracowano i wykorzystywano już pomocnicze programy wizualizacyjne, i tym razem dostosowano już istniejące rozwiązanie do potrzeb. Wykorzystano znane z Roz. 4.2.2 narzędzie wizualizacji detekcji na obrazie i wprowadzono do niego dodatkową funkcję, tj. wyświetlanie przeliczeń matematycznych głębi i kątów sferycznych dla aktualnego położenia kursora myszy. Wyświetlanie jednocześnie: detekcji obiektów oraz dotyczących ich odległości z mapy głębi, zobrazowano na Rys. 5.24.



Rysunek 5.24: Zrzut ekranu opracowanego narzędzia deweloperskiego do wizualizacji detekcji obiektów na panoramach ERP wraz z przeliczeniem położenia do przestrzeni trójwymiarowej.

5.7.3 Wyniki

Każda z faz hybrydowej metody pozycjonowania obiektów została przebadana oddzielnie, a następnie zestawiono poszczególne etapy do postaci końcowych wyników zbiorczych. Przedstawiają one rezultat całego procesu analizy środowiska w kontekście kompletnego modelu badanej nieruchomości. Jako pierwszy omówione zostało zastosowanie technologii BLE w procesie ogólnego wskazania pomieszczenia, w którym prowadzone były pomiary. Następnie przedstawiono wyniki procesu radiowego pozycjonowania precyzyjnego z wykorzystaniem technologii UWB. Badania poszczególnych etapów zamyka opis aspektu przetwarzania obrazów i detekcji głębi zrealizowany w ramach pozycjonowania wizyjnego. Cały rozdział podsumowuje natomiast przedstawienie najważniejszych (z punktu widzenia niniejszej rozprawy) wyników, tj. działania całego systemu fuzyjnego i realizacji metody hybrydowej, radiowo-wizyjnej.

Powyższe wyniki wykorzystane do weryfikacji hipotezy głównej całej pracy (określane dalej mianem hybrydowych lub fuzyjnych), uzyskano na podstawie fuzyji wszystkich opisanych składników i przejścia całej procedury pozycjonowania, w której:

- zastosowano technologię BLE do określenia pomieszczenia, w którym znajdowało się urządzenie pomiarowe;
- wskazano jego precyzyjne współrzędne w układzie odniesienia cyfrowego modelu nieruchomości z użyciem technologii UWB;
- wyeksportowano zbiór zdjęć panoramicznych, poddano je estymacji głębi, a następnie wykryto obecne na nich obiekty;
- wykonano proces pozycjonowania wizyjnego, pozyskując współrzędne obiektów w układzie odniesienia poszczególnych zdjęć sferycznych;
- zestawiono wyniki wizyjne z resztą pomiarów radiowych, efektywnie wyznaczając położenie końcowe wszystkich wykrytych komponentów środowiska.

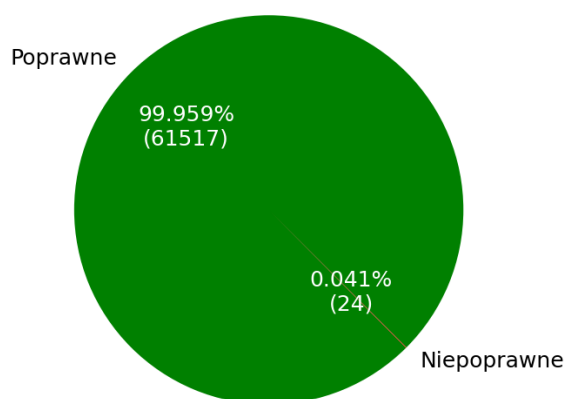
Wskazywanie pomieszczenia z wykorzystaniem systemu BLE

Badania końcowe w zakresie procesu pozycjonowania ogólnego polegały na określeniu poprawności przypisania próbkom pomiarowym właściwego pokoju, w którym znajdowało się mobilne urządzenie pomiarowe. Ponieważ podsystem oparty

na technologii BLE publikował wiadomości protokołu MQTT o nowej predykcji pomieszczenia wyłącznie w przypadku wykrycia zmiany, w analizie wykorzystano próbki pomiarowe z kolejnego etapu, tj. zapisy pochodzące z systemu UWB. Uwzględniono wszystkie zgromadzone próbki dotyczące wyznaczania odległości do kotwic i zweryfikowano, czy komunikacja była nawiązywana z prawidłowym zestawem urządzeń. Pozycjonowanie realizowane z wykorzystaniem kotwic umieszczonych w innym pomieszczeniu niż rzeczywiste położenie punktu pomiarowego klasyfikowano jako niepoprawne działanie systemu (tj. błędne wskazanie pomieszczenia). Jeżeli próbka pomiarowa zawierała komunikację z zestawem kotwic zgodnym z oczekiwanym dla danego pomieszczenia, była klasyfikowana jako poprawna.

Wykres kołowy wyników prezentowany jest na Rys. 5.25. Łącznie zgromadzonych zostało 61 541 próbek pomiarowych, z czego jedynie 24 próbki sklasyfikowane zostały jako niepoprawne. Dokładna analiza tych nielicznych przypadków wykazała, że sytuacjami powodującymi błędy była inercja całego rozwiązania, a pojawiały się one na samym początku lub na samym końcu serii pomiarowych. Wynikało to z czasu niezbędnego do ustabilizowania predykcji zmiany pomieszczenia. Jako że czas ten był konfigurowalny, a scenariusze pomiarowe przewidywały statyczne zbieranie sygnałów radiowych, spodziewanym jest, że wydłużenie tego czasu pozwoliłoby całkowicie wyeliminować nawet pojedyncze błędy.

Poprawność wskazania pomieszczenia dla próbki pomiarowej
w procesie pozycjonowania ogólnego (BLE)



Rysunek 5.25: Wykres kołowy poprawności przypisania pomieszczenia dla próbek pomiarowych systemu pozycjonowania ogólnego w technologii BLE.

Uzyskane rezultaty wyraźnie pokazują, że proponowany system pozycjonowania ogólnego, oparty na technologii BLE, spełnił swoje zadanie. Uzyskano praktycznie pełną poprawność określania pomieszczeń. Eksperymentalnie udowodniono wnioski wyciągnięte z opracowanych prototypów systemów radiowych:

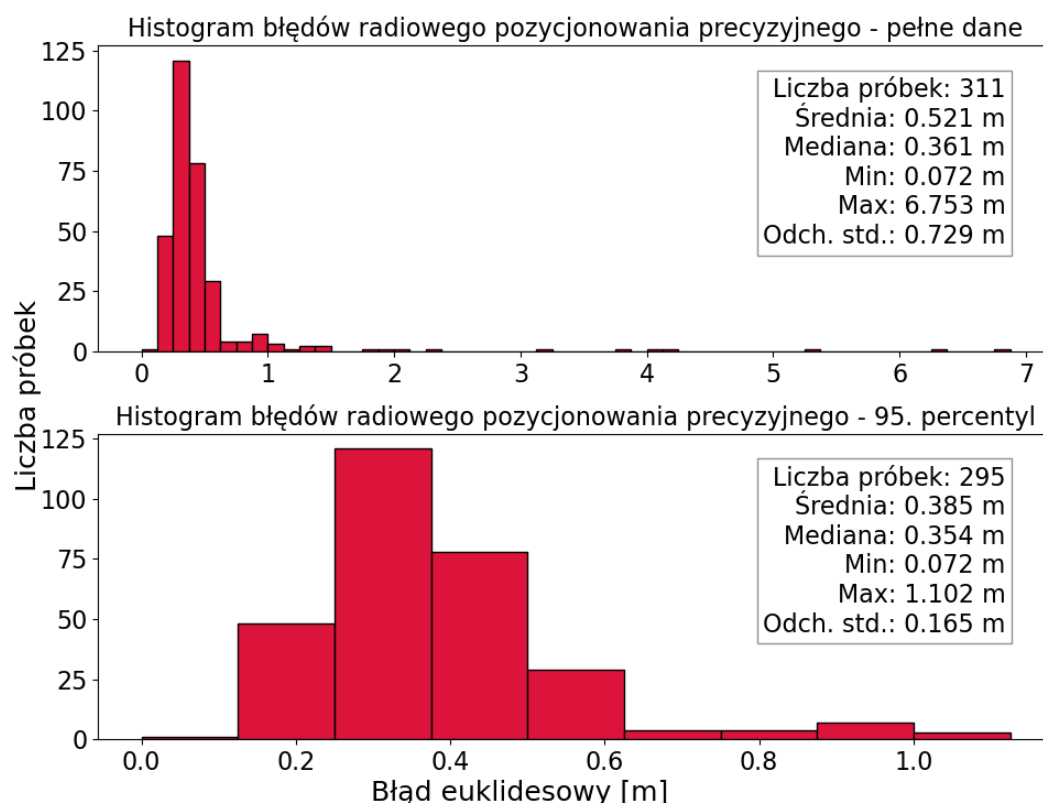
- dopasowanie realizowanego zadania do możliwości technologicznych pozwala na spełnienie wymagań funkcjonalnych przy jednoczesnym zastosowaniu powszechnie używanych komponentów elektronicznych,
- zastosowanie algorytmu opartego o średnią siłę odbieranego sygnału (zamiast liczby kotwic w zasięgu) właściwie eliminuje błędy spowodowane przystawianiem obszarów badawczych do siebie.

Zastosowanie technologii UWB do określania współrzędnych precyzyjnych

W ramach testów drugiego etapu metody hybrydowej zbadano proces radiowego pozycjonowania precyzyjnego z wykorzystaniem technologii UWB w rzeczywistym środowisku domowym. Najpierw przeanalizowano ogólną charakterystykę błędów wyznaczonego położenia, a następnie ich wielkość w zależności od pomieszczenia. Wyniki spodziewano się uzyskać na podstawie 312 próbek danych (39 serii pomiarowych * 8 urządzeń lokalizatorów = 312 próbek). Dla jednego z urządzeń, w jednej z serii, wystąpiły problemy z komunikacją do serwera MQTT, efektywnie redukując końcową pulę pomiarów do 311 próbek. Za miarę jakości wybrano odległość euklidesową w przestrzeni trójwymiarowej.

Histogramy rozkładu błędu pozycjonowania w skali całego mieszkania przedstawiono na Rys. 5.26. Uzyskane wyniki pokazują, że zastosowanie technologii UWB umożliwiło skuteczną realizację wyznaczania położenia na poziomie poszczególnych pokoi, nawet przy wysokich wymaganiach dotyczących jego precyzji. Mediana błędu euklidesowego wyliczona na podstawie pełnego zbioru wszystkich pomiarów wyniosła 36,1 cm, a błąd średni był niewiele większy i osiągnął 52,1 cm (górny histogram). Po odrzuceniu 5% pomiarów o największym błędzie, obliczając tzw. 95-ty percentyl, błąd średni spadł do 38,5 cm, zauważalnie zbliżając się do wynoszącej 35,4 cm mediany (dolny histogram). Wyniki pozwoliły zatem na dokładne określenie współrzędnych nawet niewielkich przedmiotów i w mocno

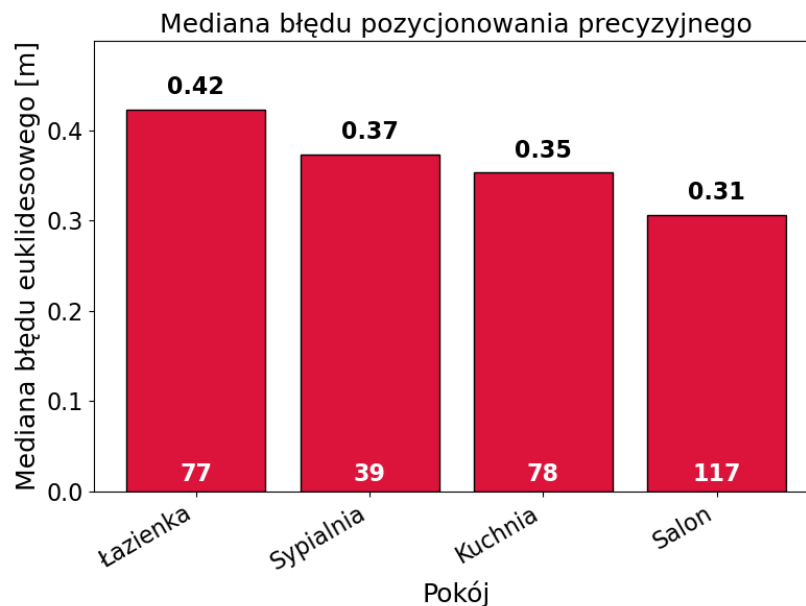
zacieśnionym obszarze dalszych poszukiwań, np. rośliny doniczkowej postawionej na stoliku czy poduszki leżącej na kanapie.



Rysunek 5.26: Histogramy błędów radiowego pozycjonowania precyzyjnego w ramach końcowych badań systemu fuzyjnego.

Zagregowanie wyników według pomieszczeń dostarczyło obserwacji analogicznych do tych, poczynionych na etapie pierwszych prototypów systemu radiowego (Roz. 3.3): zwiększenie dystansów w wykonywanych pomiarach radiowych stabilizowało odczyty i poprawiało precyzję pozycjonowania. Zjawisko to zilustrowano na diagramie słupkowym na Rys. 5.27. Najlepsze wyniki, czyli najniższą medianę błędu, osiągnięto w salonie. Było to fizycznie największe pomieszczenie całego środowiska badawczego, a także z największą liczbą umieszczonych w nim obiektów. Najgorsze wyniki, czyli najwyższą medianę błędu, odnotowano natomiast w najmniejszym z pomieszczeń i najsłabiej umeblowanym, tj. łazience. Pozostałe

dwa pokoje, fizycznie zbliżone rozmiarami do siebie, uzyskały podobne wyniki błędu pozycjonowania precyzyjnego. Z przedstawienia na wykresie wykluczone zostało pomieszczenie korytarza. Zgodnie ze spodziewaniami, brak rozmieszczenia dedykowanych urządzeń radiowych i pozycjonowanie na podstawie komunikacji z kotwicami pomieszczeń ściennych skutkowało uzyskaniem nierzetelnych pomiarów. Zgromadzone w ten sposób wartości mediany błędu były większe niż fizyczne rozmiary korytarza. Ich próba umieszczenia na wykresie zaburzałaby jego czytelność.



Rysunek 5.27: Wykres słupkowy mediany błędu pozycjonowania precyzyjnego dla każdego z pokoi z osobna.

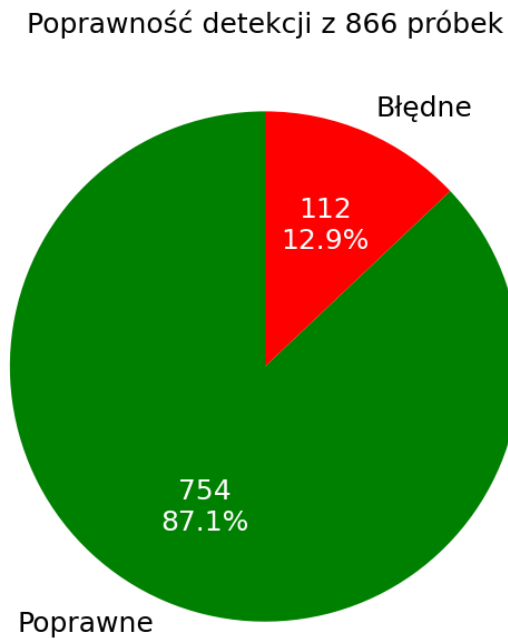
Uzyskane wyniki pozycjonowania precyzyjnego udowodniły, że:

- skutecznie zastosowano technologię radiową UWB do opracowania systemu monitorowania położenia obiektów w przestrzeni trójwymiarowej,
- uzyskane położenie śledzonego obiektu cechowało się precyzją pozwalającą na wykorzystanie tych informacji w kolejnych krokach metody hybrydowej,
- wdrożenie w warunkach małego mieszkania ma tendencję do poprawy wyników wraz ze wzrostem mierzonych odległości, co pozwala wnioskować o wysokiej przydatności systemu w pomieszczeniach o większej powierzchni.

Wizyjna detekcja obiektów i estymacja głębi

W tej części badań skoncentrowano się na kluczowych elementach budujących końcowy system pozycjonowania wizyjnego. Najpierw zweryfikowano działanie wytrenowanego modelu detektora obiektów, a następnie skupiono się na procesie estymacji głębi. Efektywność detektora oceniono najpierw globalnie, po czym przeprowadzono szczegółową analizę detekcji, grupując wyniki według rozpoznawanych klas obiektów oraz poszczególnych typów pomieszczeń. Rozdział kończą badania estymatora głębi, w ramach których sprawdzono zgodność rozkładu wyznaczonych odległości do obiektów z rzeczywistą strukturą środowiska pomiarowego na podstawie jego modelu trójwymiarowego.

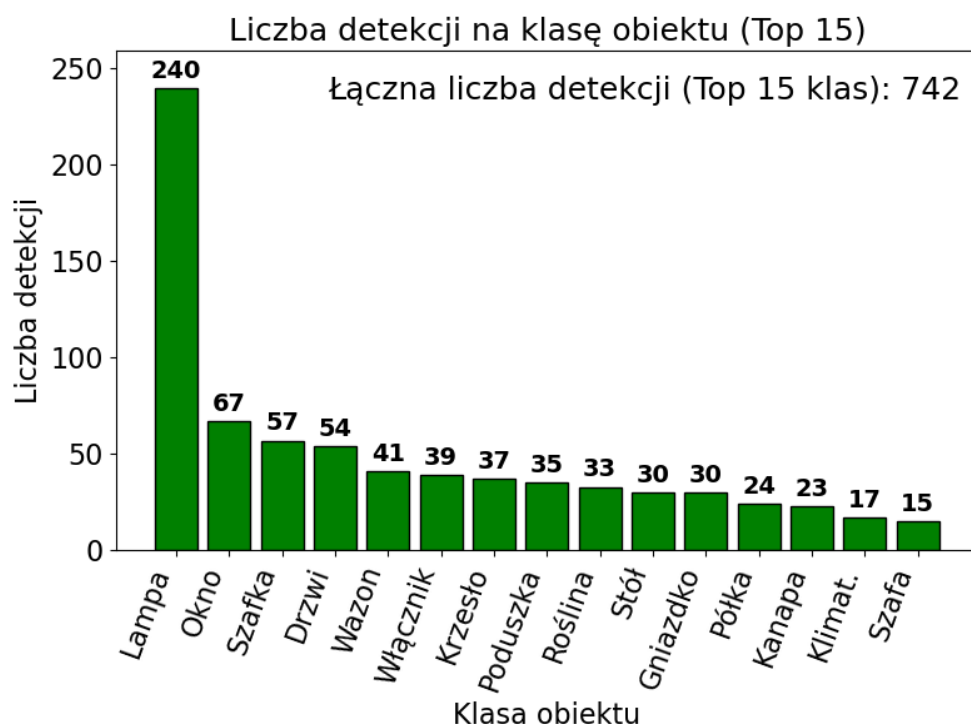
Badania końcowe dostarczyły zupełnie nowych obrazów panoramicznych środowiska domowego, które nie były etykietowane, w odróżnieniu od opisanych wcześniej zbiorów zdjęć z Roz. 4.2.3. Wyniki zawierały więc wykryte na dostarczonych zdjęciach obiekty, jednak brakowało danych referencyjnych, niezbędnych do wyliczenia standardowych miar jakości detekcji. Podejście postanowiono uprościć i ograniczyć do wyznaczenia miary precyzji, a analizę przeprowadzono manualnie. Wykorzystano do tego wcześniej omówione narzędzie wizualizacji (Roz. 5.7.2, Rys. 5.24). Detekcje obiektów klasyfikowano binarnie, jako poprawne albo błędne. Wyniki przedstawiono na wykresie kołowym na Rys. 5.28. Ogółem, spośród wszystkich 866 detekcji, 754 uznano za poprawne, a jedynie 112 za błędne. Potwierdzono w ten sposób wysoki stopień precyzji wyszkolonego detektora, który wyniósł 87,1%. Różnica w liczbie obiektów referencyjnych a ich detekcjami wynika z mnogości stanowisk pomiarowych umieszczonych w tych samych pomieszczeniach. Powodowało to, że pola widzenia kamery częściowo nakładały się na kolejnych zdjęciach, a ten sam fizyczny obiekt mógł być zarejestrowany wielokrotnie i wykryty z różnych perspektyw.



Rysunek 5.28: Diagram ogólnej poprawności detekcji obiektów.

W kolejnej fazie analiz dokonano oceny detekcji z podziałem na klasy obiektów. Dla zwiększenia przejrzystości wykresów uwzględniono wyłącznie 15 najczęściej wykrywanych rodzajów obiektów, wybranych na podstawie całkowitej liczby detekcji, niezależnie od ich poprawności. Tak ograniczony zbiór analiz obejmował łącznie 742 obiekty wykryte na obrazach. Liczbę detekcji dla każdej klasy przedstawiono na wykresie słupkowym na Rys. 5.29. Porównanie z rzeczywistym występowaniem obiektów w środowisku badawczym wskazało, że 10 z 15 najczęściej wykrywanych klas było wspólnych dla obu zestawień (na podstawie Tab. 5.1). Najczęściej pomijanymi przez detektor kategoriami były natomiast: sensory alarmowe, dywany, rolety, lustra oraz ekrany. Dostrzega się pewne cechy charakterystyczne pomijanych obiektów, które odpowiadać mogły za trudności z ich wykrywaniem. Sensory alarmowe były obiektami bardzo małymi w stosunku do rozmiaru zdjęć panoramicznych, pole ich ramek ograniczających stanowi znikomy odsetek powierzchni całego zdjęcia. Dywany zwykle umiejscawiane były w obszarach o najsilniejszych zaburzeniach wynikających z projekcji ze sfery na obraz perspektywiczny. Lustra oraz ekrany to natomiast obiekty silnie połyskliwe, które odbijały obraz otoczenia

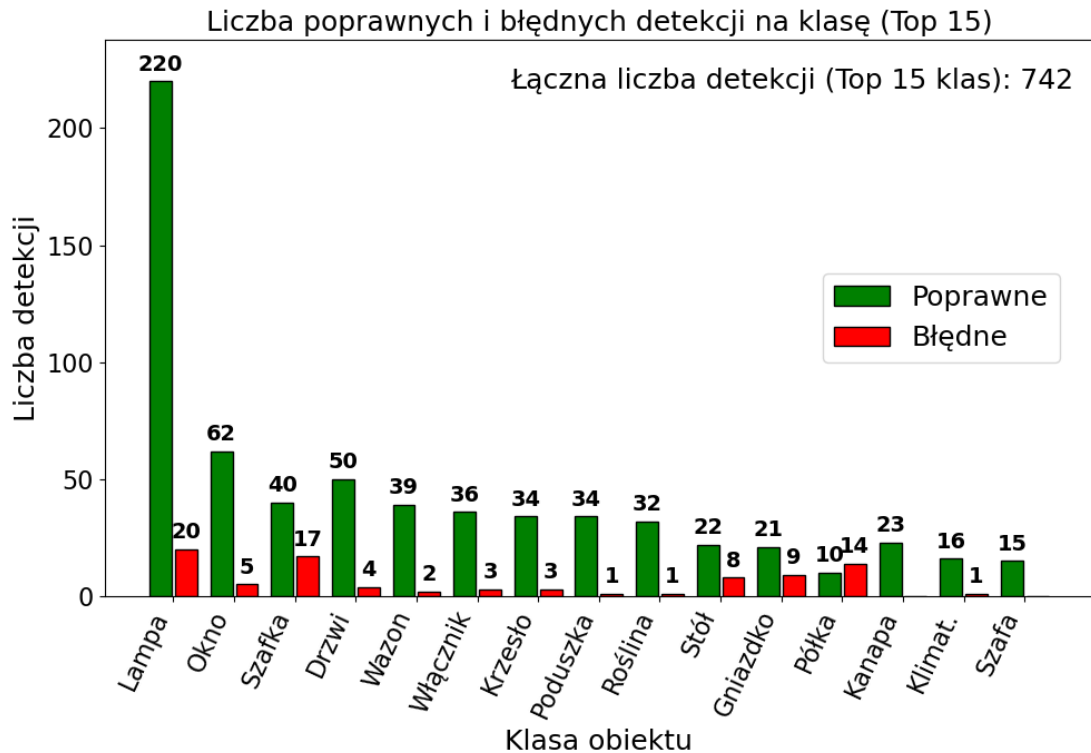
i prowadziły do powstawania złudzeń optycznych. Potwierdzenie zależności charakterystyki komponentów środowiska domowego i trudności z ich wykrywaniem wymaga jednak dodatkowych, dedykowanych badań.



Rysunek 5.29: Wykres słupkowy liczby detekcji obiektów w podziale na klasy obiektów dla 15 najpopularniejszych rodzajów obiektów.

Aby zakończyć analizę detekcji w podziale na klasy rozpoznawanych obiektów, każdą detekcję sklasyfikowano również pod kątem poprawności. Uzyskane rozkłady, ponownie tylko dla 15 najczęściej wykrytych rodzajów obiektów, przedstawiono na wykresie słupkowym na Rys. 5.30. Średnia częstość błędnej detekcji dla jednej klasy wyniosła 13,4% (przy odchyleniu standardowym równym 16% i medianie wynoszącej 7,5%). Istotną obserwacją była nadspodziewanie wysoka liczba błędów w rozpoznawaniu obiektu typu “Półka”. Po weryfikacji macierzy pomyłek detektora zaobserwowano, że występuje duża zależność błędów pomiędzy obiektami oznaczonymi w zbiorach treningowych jako “Półka” oraz “Szafka”. Dalsze rozeznanie wykazało, że wielokrotnie spotykanym w treningowym zbiorze obrazów rodzajem mebla była “Witryna”. Nie została ona jednak wydzielona

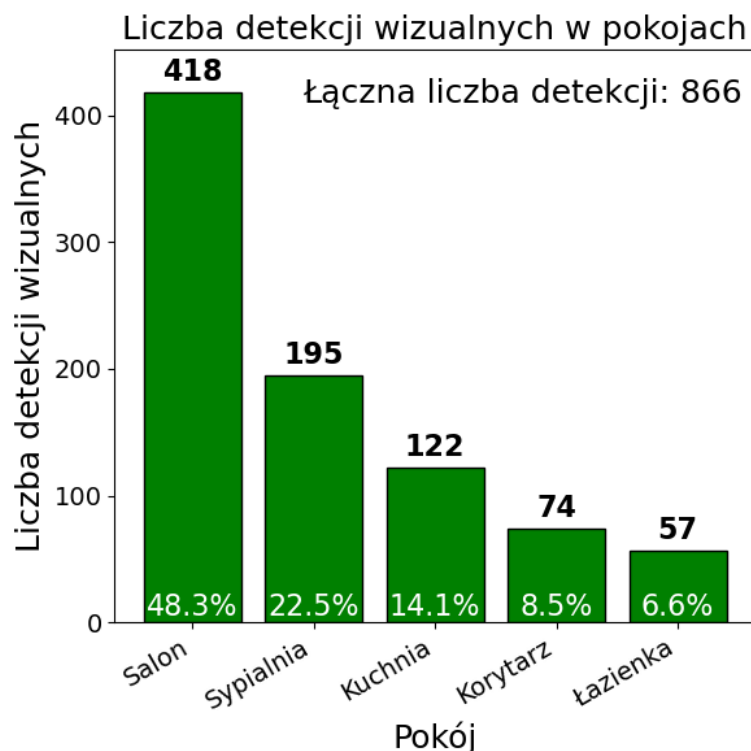
jako osobna klasa obiektów, a jako zestawienie całego mebla (“Szafki”) oraz jego fragmentu (“Półki”). Podobnych zależności nie zaobserwowano wśród pozostałych klas. Pozwala to podsumować, że wyszkolony detektor obiektów dobrze rozpoznawał oczekiwaną pulę klas przedmiotów, zgodną z opracowanym zbiorem etykiet.



Rysunek 5.30: Wykres słupkowy poprawnych i niepoprawnych detekcji obiektów w podziale na klasy obiektów dla 15 najpopularniejszych rodzajów obiektów.

Badania działania detektora obiektów podsumowano analizą w podziale na pokoje. Wykres słupkowy liczby obiektów znalezionych w poszczególnych pomieszczeniach przedstawiono na Rys. 5.31. Ponownie, wartości oczekiwane zostały już omówione wcześniej, w Roz. 5.3, i zaprezentowane na Rys. 5.14. Główne spostrzeżenie pozostaje analogiczne, tj. najwięcej obiektów ulokowanych było w salonie. Analiza pozostałych pomieszczeń nie jest już jednak tak jednoznaczna. Mimo podobnej liczby faktycznie umieszczonych w sypialni, kuchni i łazience obiektów, widzimy duże rozbieżności w liczbie ich detekcji. Na podstawie danych referencyjnych oczekiwany był podobny rozkład pomiędzy pozostałymi pokojami (po około 15-18%

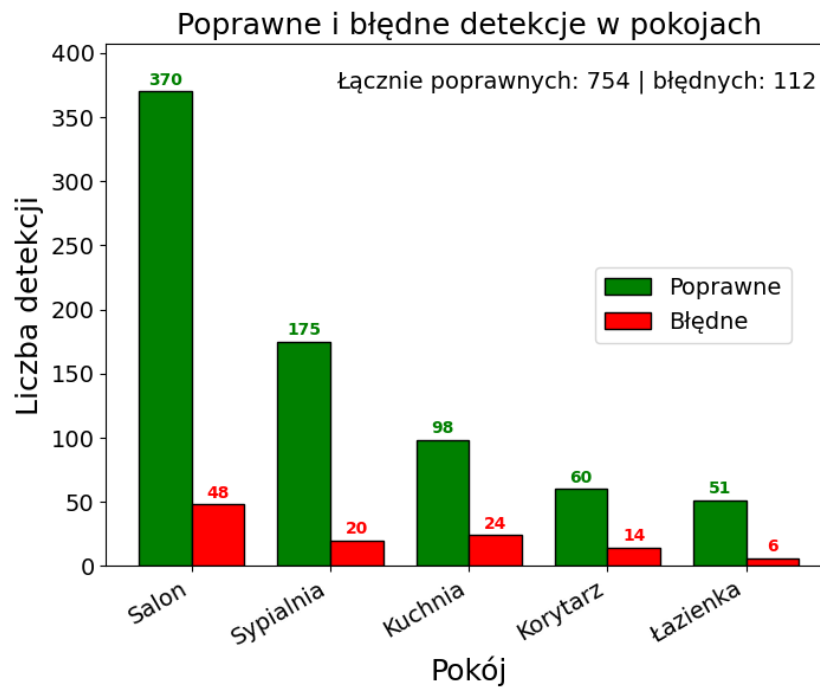
wszystkich obiektów), a jednak to w sypialni detekcji było zauważalnie więcej niż np. w łazience (22,5% całości w sypialni, a jedynie 6,6% w łazience). Pozwala to wnioskować, że rodzaje obiektów umieszczanych w łazienkach są zauważalnie trudniejsze do rozpoznania niż te z sypialni.



Rysunek 5.31: Wykres słupkowy liczby detekcji obiektów w podziale na pokoje (góra słupka) wraz z procentowym udziałem poszczególnych pokoi w łącznej liczbie detekcji (dół słupka).

Aby pogłębić analizę związaną z trudnościami w rozpoznawaniu obiektów w poszczególnych pomieszczeniach, wykonana została analiza uwzględniająca liczby detekcji uznanych za błędne. Przedstawiono ją na Rys. 5.32. Liczby błędnych detekcji wahały się między 10% a 20%. Najwyższą skuteczność osiągnięto dla pomieszczeń sypialni, salonu i łazienki (kolejno: 10% i po 11% błędów), a najniższą dla korytarza i kuchni (kolejno: 19% i 20% błędów). Dostrzec można zatem bardziej zbliżone wartości, jednakże potwierdzenie przypuszczeń o wpływie rodzaju pomieszczenia na detektor obiektów wymaga dalszych badań, głównie dedykowa-

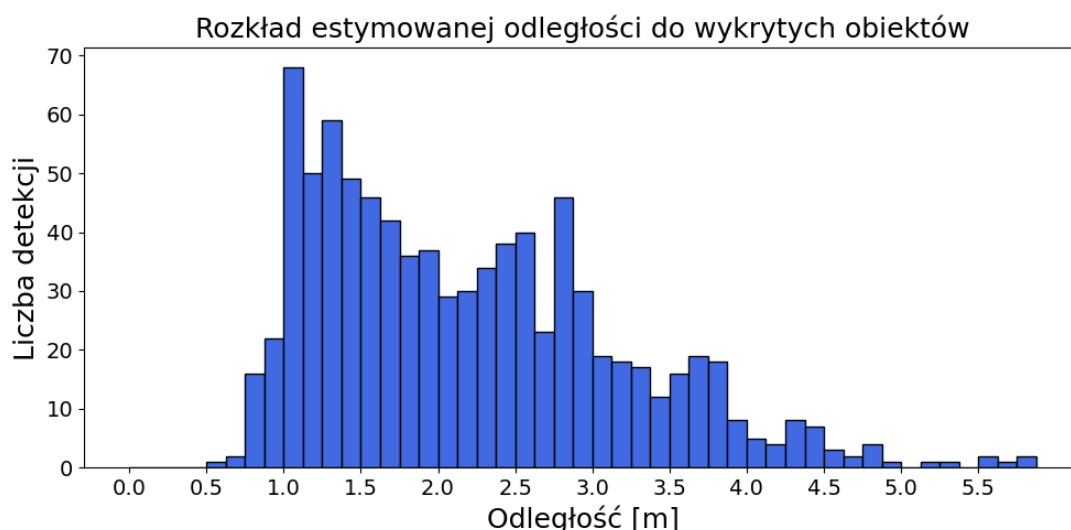
nych analizie przynależności określonych rodzajów obiektów (a co za tym idzie, trudności ich rozpoznania) do pewnych charakterystycznych pomieszczeń środowiska domowego. Niemniej, każde z pomieszczeń określić można było jako zdatne do prowadzenia badań, a w żadnym z nich odsetek błędnych detekcji nie sugerował jego wykluczenia z użycia w dalszych analizach i przebiegu metody hybrydowej.



Rysunek 5.32: Wykresy słupkowe poprawnych i niepoprawnych detekcji obiektów w podziale na pokoje.

Po zakończeniu przetwarzania wyników pracy detektora obiektów przystąpiono do analizy rezultatów działania estymatora głębi. Jego dokładność została już przebadana w Roz. 4.3, a wyniki badań eksperymentalnych w tym zakresie przedstawiono na Rys.4.26. Na ich podstawie wartość parametru maksymalnej głębi ustalono na 6 m, a wyniki pomiarów głębokości uzyskane podczas testów przedstawionej w niniejszym rozdziale metody hybrydowej przedstawiono na histogramie na Rys. 5.33.

Wyniki są silnie zgodne z ludzką intuicją i odpowiadają parametrom środowiska eksperymentalnego. Większość estymowanych głębokości przyjmuje wartość z przedziału od 1,0 m do 3,0 m, co dobrze wpisuje się w spodziewane wartości



Rysunek 5.33: Histogram rozkładu wyników estymacji głębokości z badań systemu hybrydowego.

od środka wolnej przestrzeni do najbliższych mebli w scenariuszu pomieszczeń o szerokości nie większej niż 6 metrów. Warto zwrócić jednak uwagę na wprost widoczną istotność doboru parametru maksymalnej głębi. Jego wartość wyznaczona z modelu cyfrowego bliźniaka na wspomniane 6 metrów sprawia, że wszystkie wyniki estymacji mieszczą się poniżej tego progu. Nie jest to jednak całkowicie zgodne z prawdą i manualnie obserwowanymi detekcjami obiektów. Niektóre obiekty wykrywane były na widoku przez otwarte drzwi i na przekroju dwóch pokoi oraz korytarza. Rzeczywista odległość była wtedy zauważalnie większa, a jednak nie pojawia się ona wśród uzyskanych wyników.

Podsumowując, komponenty systemu dedykowane zagadnieniom wizji komputerowej spełniły swoje zadanie poprawnie i wykazały gotowość do wykorzystania w ramach metody hybrydowej. Na podstawie końcowych badań eksperymentalnych pokazano, że:

- znakomita większość obiektów (87,1%), wykrytych z użyciem wyszkolonego detektora, rozpoznana została poprawnie, co potwierdziło wysoką precyzję opracowanego rozwiązania;
- 10 z 15 rodzajów najczęściej wykrywanych obiektów pokryło się z rodzajami oczekiwanymi na podstawie danych referencyjnych;

- mediana liczby błędnych detekcji obiektów względem każdej z klas wyniosła 7,5%, co pokazuje niskie wymagania dotyczące zaangażowania użytkownika końcowego w manualne korekcje błędów;
- wyznaczenie wartości parametru maksymalnej głębi, na podstawie szerokości pomieszczenia odczytanej z modelu trójwymiarowego nieruchomości, pozwala na skuteczne wyskalowanie wszystkich wyników estymatora głębi;
- potwierdzenie obserwacji o zależności skuteczności detekcji obiektów od rodzaju pokoju, którego ona dotyczy, wymaga dodatkowych badań.

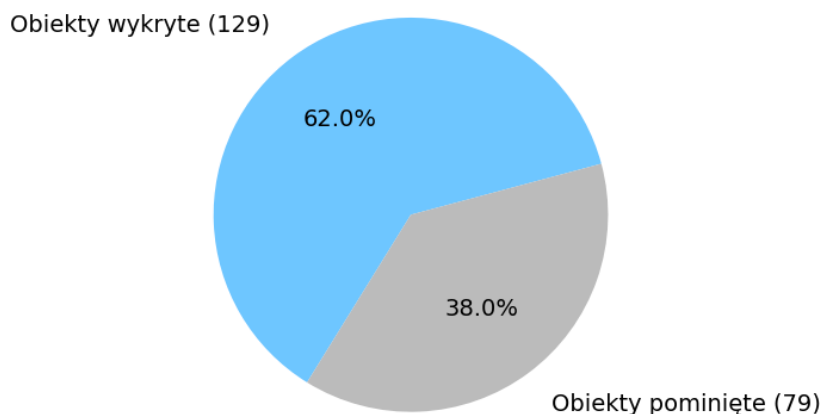
Metoda hybrydowa

Ostatni rozdział omówienia wyników poświęcony jest już całkowicie prezentacji sumarycznej, końcowej sprawności zastosowania proponowanej w rozprawie, hybrydowej metody rozpoznawania i pozycjonowania komponentów inteligentnego środowiska domowego. Rozdział rozpoczyna analiza stopnia automatyzacji i jakości całego procesu. Określono udział wykrytych obiektów względem stanu referencyjnego i całego wyposażenia środowiska badawczego, jego poprawność, a także rozkład liczebności detekcji. Następnie sprawdzona została precyzja uzyskanego położenia w trójwymiarze, tak w wymiarze globalnym całego mieszkania, jak również w podziale na poszczególne pomieszczenia. Na zakończenie przedstawiono wizualizację wyników w postaci ich odwzorowania w modelu trójwymiarowym.

W pierwszej kolejności omówione zostają agregacje: stopień automatyzacji procesu wykrywania obiektów w całym badanym środowisku, jego poprawność oraz liczebność detekcji na każdy obiekt. Rozpoczęto od zgrupowania zlokalizowanych hybrydowo obiektów (wykrytych wizyjnie i z wyznaczonym położeniem w trójwymiarze) po identyfikatory odpowiadających im obiektów referencyjnych. Porównano, ile faktycznie obecnych w mieszkaniu obiektów posiadało chociaż jedną poprawną detekcję. Wyniki pokazano na wykresie kołowym na Rys. 5.34.

Spośród wszystkich 208 oczekiwanych do wykrycia komponentów, 129 zlokalizowanych zostało automatycznie (62%), a 79 zostało pominiętych (38%). **Wyniki pozwalają zatem określić, że większość procesu oznaczania obiektów została zautomatyzowana.** Aby pogłębić analizę, w ramach każdej grupy detekcji oceniono je również pod kątem liczby poprawnych i błędnych klasyfi-

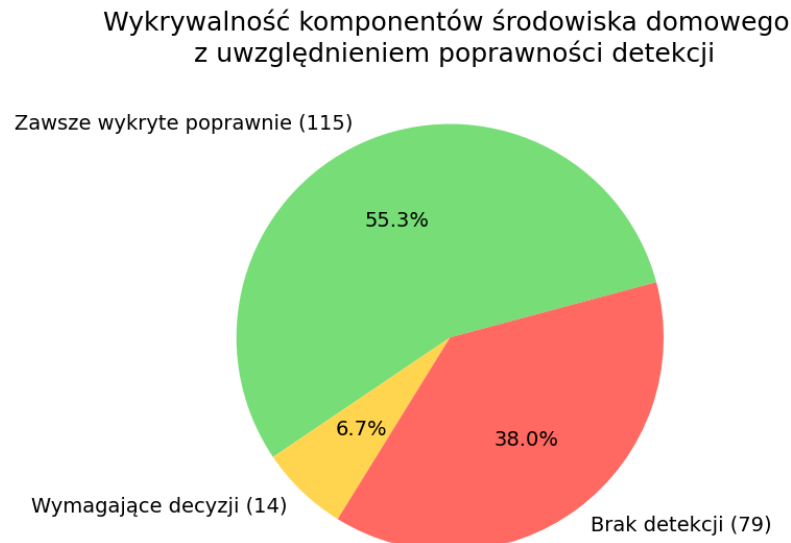
Wykrywalność komponentów środowiska domowego przy zastosowaniu metody hybrydowej



Rysunek 5.34: Wykres końcowej wykrywalności komponentów środowiska domowego przy zastosowaniu metody hybrydowej.

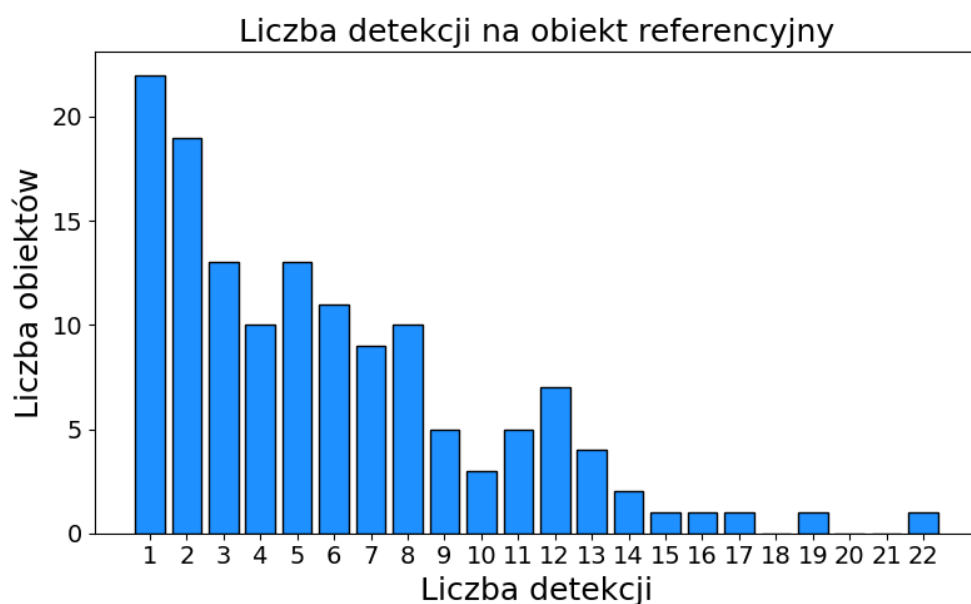
kacji. Miało to na celu uzyskanie wiedzy o stopniu pewności opracowanego systemu co do klasyfikacji rodzajów wykrytych obiektów. Wyniki zaprezentowano na Rys. 5.35.

Wśród wspomnianych 129 obiektów zlokalizowanych automatycznie, 115 zawsze klasyfikowano poprawnie co do ich rodzaju (55,3% wszystkich obiektów), a 14 wymagało dalszego przetwarzania (6,7% wszystkich obiektów). Dogłębnie sprawdzono tych 14 przypadków wymagających przetwarzania. W 13 z 14 przypadków stwierdzono, że wybranie najczęściej powtarzającej się klasyfikacji obiektu pozwoliłoby na uzyskanie poprawnego wyniku. Tylko w jednym z przypadków (obiekt "Table.002", tj. stół w sypialni) liczba detekcji poprawnych i błędnych była sobie równa (po jednej detekcji), a obiekt należałoby wykluczyć z procesu automatycznego i oznaczyć manualnie.



Rysunek 5.35: Wykres końcowej wykrywalności komponentów środowiska domowego z uwzględnieniem poprawności detekcji.

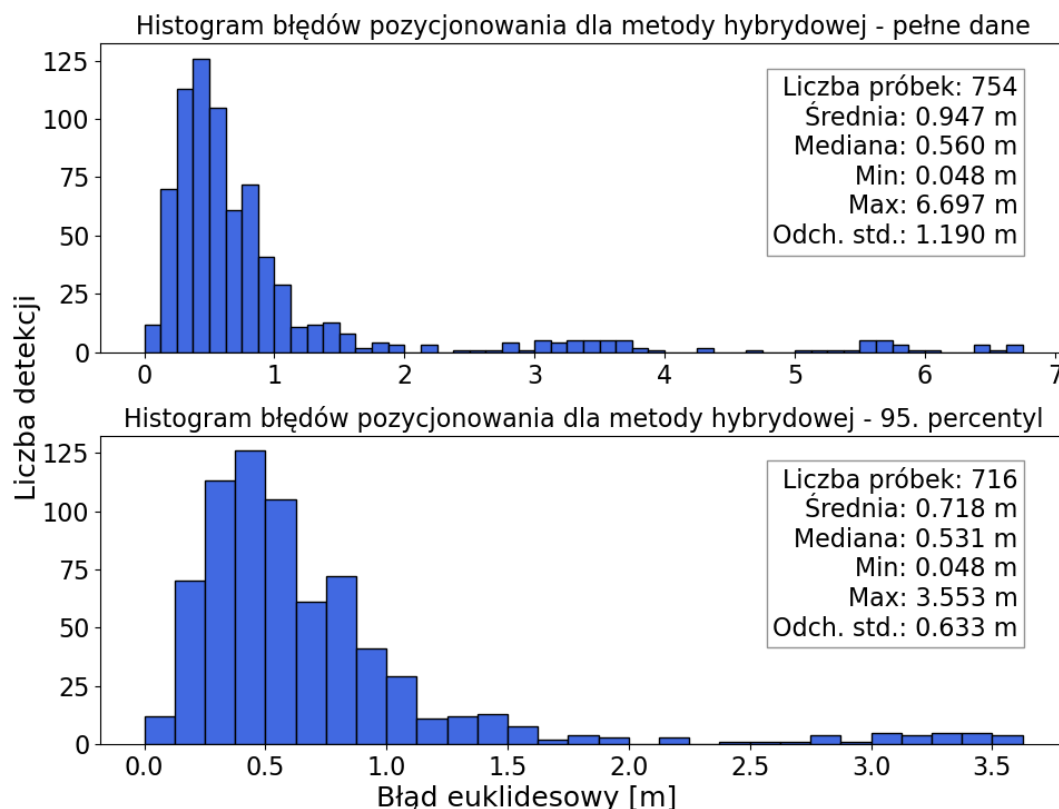
W kolejnym kroku sprawdzono, jak często wykrywano poszczególne obiekty, tj. na ilu zdjęciach możliwe było uzyskanie informacji o położeniu tego samego komponentu środowiska. Im więcej takich detekcji, tym z większej liczby próbek można wyciągać wnioski, a tym samym projektować bardziej zaawansowane rozwiązania dalszej poprawy precyzji. Przedstawiono to na wykresie słupkowym na Rys. 5.36. Dla czytelności wykresu, w obliczeniach i omówieniu pominięto 70 obiektów, których położenia nie wskazano ani razu. Najczęściej pojawiającym się wynikiem była pojedyncza detekcja, a mediana wyniosła pięć wykryć. Stwierdzić zatem można, że dla zgromadzenia większej próbki pomiarowej należałoby zagęścić pokrycie powierzchni mieszkania punktami pomiarowymi, a co za tym idzie, zdjęciami panoramicznymi. Dominanta detekcji wynosząca jedynie pojedyncze wykrycie wskazuje, że istnieje duże ryzyko pominięcia obiektu w razie utraty któregoś zestawu pomiarów.



Rysunek 5.36: Wykres licznosci detekcji na pojedynczy obiekt w metodzie hybrydowej.

Aby definitywnie potwierdzić hipotezę główną rozprawy, przeanalizowano również precyzję wyznaczonego położenia obiektów. Obliczone zostały miary charakterystyczne błędów, tj. średnia, mediana, minimum, maksimum oraz odchylenie standardowe, a także przygotowano histogramy ich rozkładu, co przedstawiono na Rys. 5.37. Podobnie jak wcześniej, za miarę jakości wybrano odległość euklidesową w przestrzeni trójwymiarowej, a obliczenia przeprowadzono na pełnym zbiorze pomiarów (górny histogram) oraz dla 95-ego percentyla (dolny histogram). Łącznie zgromadzono 754 predykcje położenia obiektów, wartość średnia błędu pozycjonowania wyniosła 94,7 cm, a jego mediana osiągnęła 56,0 cm. Po odrzuceniu 5% najbardziej odstających pomiarów, 95% obiektów udało się zlokalizować z zauważalnie niższym średnim błędem, wynoszącym 71,8 cm, a także przy medianie równej 53,1 cm. **Wyniki pozwalają zatem potwierdzić, że większość pomiarów wskazała poprawnie położenie nawet niewielkiego obiektu i z precyzją umożliwiającą jego skuteczne odnalezienie w przestrzeni przez człowieka.** W kształcie obu prezentowanych histogramów widać charakterystyczny tzw. “długi ogon” wykresu, co pozwala zaobserwować, że istotnym

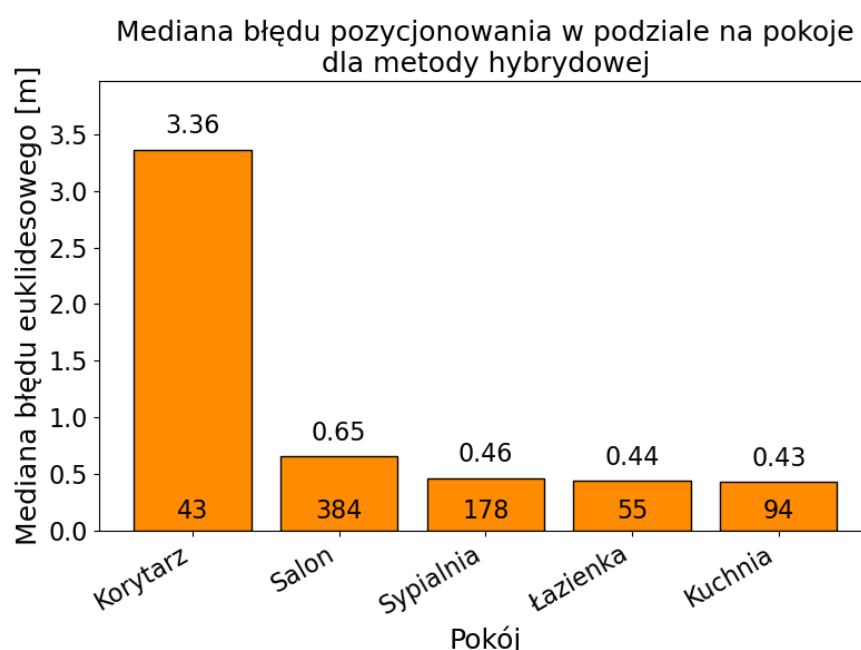
kierunkiem dalszych badań powinno być wprowadzenie do rozwiązania bardziej zaawansowanego systemu filtracji wartości odstających.



Rysunek 5.37: Histogramy błędów pozycjonowania wizyjnego uzyskanego w hybrydowej, wizyjno-radiowej metodzie pozycjonowania obiektów.

Analizy liczbowe zamyka omówienie błędów pozycjonowania w podziale na pokoje, w których miało ono miejsce. Wykres słupkowy zaprezentowano na Rys. 5.38. Pomijając na moment korytarz, najwyższą medianę błędu pozycjonowania obiektów odnotowano w największym z pomieszczeń, tj. salonie. Wyniosła ona 65 cm. Dla pozostałych pomieszczeń mediana waha się od 43 cm w kuchni, przez 44 cm w łazience, po 46 cm w sypialni. Różnice są więc niewielkie i wynoszą nie więcej niż 3 cm. Zwrócenia szczególnej uwagi wymaga wspomniane pomieszczenie korytarza. Już w opisie wyników dla radiowego pozycjonowania precyzyjnego (Rys. 5.27) pomieszczenie to zostało celowo pominięte, ponieważ uzyskane wyniki były na tyle

niepoprawne, że zaburzyłyby czytelność wykresu. Dla analizy błędu metody hybrydowej postanowiono je jednak uwzględnić. Miało to na celu choć szczątkową prezentację skali rozbieżności wyników w sytuacji, gdy końcowe pozycjonowanie hybrydowe miało miejsce na podstawie położenia kamery uzyskanego z systemu radiowego, którego urządzenia referencyjne (kotwice) umieszczono w pomieszczeniach ościennych, a nie właściwie badanym. Mediana błędu w takiej sytuacji była ponad pięciokrotnie wyższa niż w największym z analizowanych pokoiów.

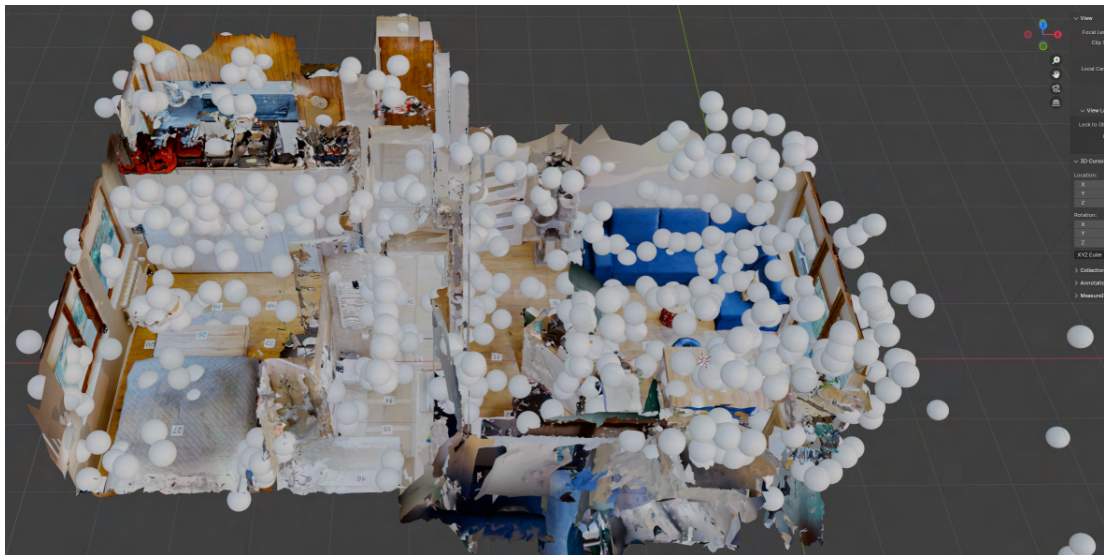


Rysunek 5.38: Wykres słupkowy mediany błędu pozycjonowania w podziale na pokoje dla metody hybrydowej. Na dole słupka przedstawiono liczbę próbek w danym pokoju.

Podkreśla to ważność doboru realizowanego wdrożenia do potrzeb. W sytuacji niedoboru kotwic systemu radiowego proponuje się więc łączenie rozwiązań. W przedstawionej sytuacji dla 4 z 5 pomieszczeń zastosować można metodę w pełni hybrydową. Dla pomieszczenia korytarza wynik działania pozycjonowania precyzyjnego, tj. współrzędne kamery, sugeruje się jednak uzupełnić ręcznie. W ten sposób osiągnąć można satysfakcjonującą jakość wyników przy ograniczeniu nakładów pracy manualnej. Niemniej, przy zapewnieniu oczekiwanych realiów wdrożeniowych, tj. przy pełnej dostępności wszystkich podsystemów rozwiązania hybrydo-

wego, wyniki pokazały, że realizację systemu lokalizowania przedmiotów jest możliwa niezależnie od rodzaju pomieszczenia, a nawet w największym z pomieszczeń poszukiwane obiekty były lokalizowane z medianą błędu pozycjonowania mniejszą niż jeden metr.

Na zakończenie omówienia, przedstawiona zostaje bardziej wizualna reprezentacja wyników. Współrzędne wszystkich wyznaczonych obiektów przetworzone zostały do postaci listy o ujednoliconych rekordach. Następnie napisano skrypt języka Python, który uruchomiono w ramach programu Blender. Każdy zestaw współrzędnych (praktycznie: każdy obiekt) odwzorowany został z powrotem wprost w modelu trójwymiarowym przy użyciu oznaczenia symbolicznego w postaci sfery o średnicy 30 cm. Przedstawiono to na Rys. 5.39.



Rysunek 5.39: Wizualizacja położenia wszystkich wykrytych w mieszkaniu obiektów (oznaczonych przy pomocy sfer). Współrzędne obiektów określone na podstawie systemu hybrydowego z punktem położenia kamery wyznaczonej automatycznie przez radiowy system pozycjonowania.

Przedstawiona analiza i uzyskane wyniki potwierdzają główną hipotezę rozprawy. Zastosowanie proponowanej metody hybrydowego, wizyjno-radiowego rozpoznawania i pozycjonowania obiektów w środowisku domowym pozwoliło na automatyczne odzyskanie semantycznego znaczenia większości komponentów opracowanego modelu cyfrowego bliźniaka

nieruchomości. Zachowana została wysoka precyzja określenia ich położenia, która w większości przypadków umożliwia użytkownikowi końcowemu skuteczne odnalezienie nadzorowanych przedmiotów w promieniu poniżej jednego metra od wskazanego położenia. Co ważne, opracowany system fuzyjny pozwala wygodnie i szybko odnaleźć poszukiwane komponenty środowiska niezależnie od ich rodzaju i faktu wyposażenia w dedykowane urządzenia lokalizacyjne, a także przy użyciu powszechnie dostępnych układów elektronicznych i komercyjnych kamer sferycznych. Dzięki temu możliwe staje się realizowanie różnorodnych zadań konserwatorskich w oparciu o cyfrowy model nieruchomości i proces wyszukiwania jego wyposażenia, np. celem odnalezienia i nawigacji do wszystkich jednostek klimatyzacji wymagających konserwacji (przykład urządzeń aktywnych) albo gaśnic zbliżających się do wygaśnięcia przeglądu technicznego (przykład urządzeń pasywnych).

Rozdział 6

Wyniki i podsumowanie

Poruszona w rozprawie tematyka była wprost powiązana z ideą digitalizacji i rozumienia otoczenia człowieka oraz tworzenia jego cyfrowej reprezentacji. Jednocześnie, bezpośrednią motywacją przeprowadzonych prac badawczych i wdrożeniowych była konkretna potrzeba przemysłowa firmy partnera doktoratu (Simlab Sp. z o.o.): nadanie surowemu modelowi cyfrowemu nadzorowanej nieruchomości semantycznego znaczenia jej komponentów i przekształcenie go do faktycznie użytecznego cyfrowego bliźniaka. Dotychczas dostępne na rynku komercyjnym rozwiązania, które mogły zrealizować tę potrzebę, były zwykle nieosiągalne dla przeciętnego użytkownika końcowego produktów firmy. Spowodowane było to ich wysokimi kosztami lub niedopasowaniem do realnych potrzeb. Wykrytą luką badawczo-przemysłową było więc opracowanie opartego na powszechnie dostępnych komponentach sprzętowych, prostego w użyciu systemu rozpoznawania i pozycjonowania obiektów w typowym środowisku domowym. Rozwiązanie takie miało usprawnić proces konfiguracji wstępnej i przejścia od surowych danych do wypełnionego informacjami cyfrowego bliźniaka. Dotychczas brak było jakiegokolwiek automatyzacji procesu identyfikacji kluczowych dla użytkownika obiektów, a stanowiących zarazem komponenty jego otoczenia, odwzorowane w modelu cyfrowym. Konieczne było wskazywanie ich rodzaju i położenia ręcznie, jeden po drugim.

Aby na potrzeby te odpowiedzieć, zasadnicza część pracy poświęcona została opracowaniu hybrydowego, wizyjno-radiowego systemu rozpoznawania i pozycjonowania komponentów inteligentnego środowiska domowego. Jego realizacja i testy

stanowią zwięźcenie prac, których wachlarz najistotniejszych elementów objął:

- definicję metody postępowania, jej kolejnych etapów oraz sposobów ich realizacji;
- konstrukcję zestawu dedykowanych urządzeń, realizujących proces pozycjonowania radiowego;
- digitalizację przestrzeni badawczej i wykorzystanie jej cyfrowego modelu w projektowaniu i ewaluacji eksperymentów;
- rozbudowane testy opracowanego rozwiązania, zrealizowane w warunkach możliwie zbliżonych do docelowych wdrożeń u klientów przemysłowych.

Oceniając stworzone rozwiązanie w kontekście pytań badawczych pracy z Roz. 1.3, podsumowanie rozpocząć możemy od śmiałego stwierdzenia, że dotychczas manualny proces nadawania komponentom pozyskanego cyfrowego modelu nieruchomości ich semantycznego znaczenia może zostać **w większości zautomatyzowany**. Dla przykładowego scenariusza stanowiącego dwupokojowe mieszkanie z osobną kuchnią i łazienką, 62% wszystkich obecnych w środowisku obiektów zostało choć raz przeanalizowanych automatycznie, a ponad 55% obiektów było identyfikowanych zawsze poprawnie. Wdrożenie proponowanego rozwiązania do użycia pozwala zatem **odciążyć użytkowników końcowych produktów firmy partnera z ponad połowy manualnych operacji niezbędnych w ramach konfiguracji wstępnej modelu cyfrowego bliźniaka**. Proponowany system nie zmusza użytkowników do pozyskania całego wykorzystywanego osprzętu od razu (obniża koszty początkowe), a poszczególne etapy metody hybrydowej mogą zostać zastąpione manualną pracą operatorów. **Można w ten sposób wdrożyć system półautomatyczny**, w którym np. detekcja i pozycjonowanie wizyjne realizowane jest automatycznie, a położenie kamer sferycznych określone jest manualnie. Dzięki wyszkoleniu modelu detektora obiektów w oparciu o zbiory danych etykietowane zgodnie z wynikami badań sondażowych pośród użytkowników, **zwiększony został stopień zgodności oferowanych automatyzacji z faktycznymi potrzebami rynku**, co zachęca do ich popularyzacji i codziennego stosowania. Doświadczalnie potwierdzona końcowa precyzja wyznaczonego położenia wykrytych obiektów pozwala na skuteczne ich odwzorowanie w modelu cyfrowym, a także praktyczne zastosowanie: **uzyskane wyniki pozwalają zlokalizować większość**

nawet niewielkich rozmiarów obiektów (jak plecak, krzesło lub wózek inwalidzki) w obszarze o promieniu poniżej jednego metra od wskazanego punktu. Wszystko to zrealizowane zostało natomiast w oparciu o powszechnie dostępne komponenty elektroniczne, gotowe płytki układów scalonych, a także popularne sportowe kamery sferyczne. Obniżono tym samym koszty realizacji proponowanego systemu w stosunku do konkurencyjnych rozwiązań przemysłowych i podniesiono jego efektywność ekonomiczną. Jest to o tyle ważne, że przeprowadzone badania nad wykorzystaniem w pełni syntetycznych zbiorów danych do trenowania detektorów obiektów wykazały, że nie są one jeszcze zdolne w pełni zastąpić ręcznie gromadzonych zbiorów zdjęć rzeczywistych. Wyniki uzyskane na zbiorach w pełni syntetycznych były właściwie o połowę gorsze (pod względem wskaźnika mAP) w porównaniu ze zbiorami rzeczywistymi. Prowadzi to do wniosku, że obecnie nie jest jeszcze możliwe pominięcie procesu manualnego gromadzenia i etykietowania zbiorów treningowych, z kosztami opracowania których należy się liczyć. Jest to szczególnie istotne w kontekście rosnących oczekiwań właścicieli nieruchomości w zakresie dbałości o prywatność sferycznych zdjęć ich środowiska domowego.

Realizacja systemu hybrydowego i jego końcowa skuteczność możliwe były dzięki rozbudowanym pracom wstępnym. Do ich najważniejszych aspektów i wartych wspomnienia obserwacji należą:

- rozłożystość dostępnych do przetwarzania źródeł i typów danych, a także wachlarz ich zastosowań;
- różnorodność oczekiwań użytkowników końcowych w odniesieniu do typowości obiektów i rodzajów pomieszczeń, które definiuje się w środowisku domowym;
- istotność dopasowania wykorzystywanej technologii radiowej i protokołu komunikacyjnego do realizowanego zadania i sposobu jego wdrożenia do użycia;
- czasochłonność gromadzenia i etykietowania sferycznych zbiorów zdjęć, a także znaczenie ich przetwarzania wstępnego i jakości;
- powszechna dostępność gotowych do zastosowania rozwiązań, które mimo swojego teoretycznie ogólnego zastosowania, dobrze sprawdzają się również w specyficznych scenariuszach wdrożeniowych;

Analiza literatury zwróciła tak obszerne wyniki, że podzielono je na trzy odrębne taksonomie tematyczne. W każdym kolejnym aspekcie: typie danych wejściowych, kategorii rozwiązania, a także końcowym zadaniu, możliwe było przedstawienie jego rozbudowanego kontekstu naukowego i wachlarza możliwości. Podobnie różnorodne okazały się być oczekiwania użytkowników końcowych i respondentów przeprowadzonej ankiety. Chcąc pokryć możliwościami detektora obiektów wszystkie oczekiwania, konieczne byłoby wyszczególnienie ponad 170 klas obiektów komponentów środowiska domowego, a klasyfikacja rodzaju pomieszczenia musiałaby objąć 60 ich rodzajów. Niemniej, dostrzec można było również elementy wspólne wszystkich potrzeb, a do końcowego zbioru etykiet obiektów wybrano 55 najpopularniejszych. Pewne obiekty charakterystyczne nosiły też cechy silnego przywiązania do konkretnych pomieszczeń, np. wanna do łazienki czy samochód do garażu. Wnioski z opracowanego prototypu radiowego systemu estymacji dystansu, wdrożone w końcowym systemie hybrydowym, pokazały natomiast, jak istotna jest dogłębna znajomość wykorzystywanych narzędzi i technologii. Oceniony początkowo za zupełnie niezdatny do pozycjonowania precyzyjnego, system oparty na technologii Bluetooth Low Energy uzyskał prawie 100% skuteczności w realizacji zadania pozycjonowania ogólnego i odpowiednio zastosowany właściwie bezbłędnie klasyfikował pomieszczenie, w którym prowadzono badania. Jednym z najważniejszych, a jednocześnie bardzo czasochłonnych i często niedostrzeganych aspektów prowadzonych prac było rzetelne opracowanie danych treningowych dla detektora obiektów. Wysoka specyfika domeny wdrożeniowej, tj. obrazy sferyczne w reprezentacji panoramicznej, powodowała niedobór publicznie dostępnych, wielkoskalowych zbiorów szkoleniowych. Zgromadzono i oznaczono je więc ręcznie, co pozwoliło pozyskać wiedzę praktyczną o systemach zarządzania zbiorami, platformach internetowych do ich przetwarzania, a także doświadczenie z zakresu ich dalszego przygotowania do użycia. Przeprowadzone badania wykazały, że odpowiednie przygotowanie zbiorów, np. poprzez dobór południka początkowego w procesie rzutowania ze sfery na obraz panoramiczny, poprawia końcową skuteczność wyszkolonych przy ich użyciu detektorów obiektów. Co istotne, pokazano również, że powszechnie używane modele sieci neuronowych, których licencje zezwalają na bezpłatny użytek akademicki lub nawet komercyjny, uzyskują tylko nieznacznie gorsze rezultaty niż rozwiązania dedykowane. Samo opracowanie wewnętrznej struktury

takiego uniwersalnego detektora obiektów (jego architektury) to wyzwanie, któremu uwagę poświęcają najwięksi światowi potentaci przemysłowi oraz instytucje naukowe. Dzięki ich publicznej dostępności możliwe było jednak szybkie opracowanie prototypów, weryfikacja ich skuteczności, a następnie doszkolenie modeli i wdrożenie do użycia.

Do autorskich oryginalnych wyników uzyskanych w rozprawie można zaliczyć:

- metodę automatyzacji procesu nadania semantycznego znaczenia surowemu modelowi cyfrowemu nieruchomości poprzez zastosowanie hybrydowej, wizyjno-radiowej metody rozpoznawania i pozycjonowania jego komponentów;
- algorytm radiowego pozycjonowania ogólnego, który w oparciu o technologię BLE pozwala na identyfikację pomieszczeń, w których prowadzone są badania;
- metodę radiowego pozycjonowania precyzyjnego, która dzięki zastosowaniu technologii UWB i algorytmowi trilateracji pozwala na wyznaczenie współrzędnych trójwymiarowych śledzonego obiektu w bezwzględnym układzie współrzędnych modelu cyfrowego;
- algorytm pozycjonowania wizyjnego, wykorzystujący syntezę działania detektora obiektów, estymatora głębi i przekształcenia sferyczne do wskazania położenia wykrytych komponentów środowiska we względnym układzie odniesienia kamery;
- wykazanie niezdatności sztucznych sieci generatywnych do opracowywania syntetycznych zbiorów zdjęć mających całkowicie zastąpić zbiory rzeczywiste;
- algorytm doboru najlepszego południka rozcięcia dla procesu rzutowania obrazu sferycznego na panoramę typu ERP;
- potwierdzenie bliskości wyników detekcji obiektów pomiędzy rozwiązaniami dedykowanymi, a uniwersalnego zastosowania;
- określenie charakterystyki skutecznego użycia technologii BLE i UWB do opracowania systemu hybrydowej lokalizacji w środowisku domowym;
- potwierdzenie zdatności protokołu Matter do realizacji wymiany danych w ramach systemów pozycjonowania opartych na domowych urządzeniach IoT;
- opracowany zbiór etykiet klas, pomieszczeń oraz matryca ich typowości, cha-

rakterystycznych dla środowiska domowego i oczekiwanych przez użytkowników;

- rozbudowany zestaw oznaczonych zbiorów zdjęć panoramicznych środowiska domowego dedykowanych procesowi szkolenia detektora obiektów;
- opracowane taksonomie systematyzujące wnioski z przeglądu literatury, opisujące trzy kolejne poziomy wiedzy: typy danych wejściowych, kategorie ich przetwarzania oraz realizowane zadania;
- analizy: konkurencji rynkowej partnera doktoratu oraz SWOT produktu, do którego wdrażane są proponowane w rozprawie ulepszenia;

Podsumowując, przedstawione w niniejszej pracy podejście pozwoliło osiągnąć zamierzone cele. Przeprowadzone badania i uzyskane wyniki dają podstawę do sformułowania wniosku o słuszności postawionej w pracy tezy głównej oraz tez dodatkowych. Ich dokładne omówienie przedstawiono poniżej (w Roz. 6.1). Otwiera to również ścieżkę do skutecznej implementacji wdrożeniowej opracowanego rozwiązania u szerokiej gamy klientów firmy partnera doktoratu: od właścicieli i nadzorców obiektów prywatnych, przez małe przedsiębiorstwa komercyjne, aż po wielkoskalowe obiekty hotelarskie.

6.1 Weryfikacja hipotez pracy

Wszystkie sformułowane w rozprawie hipotezy zostały w przebiegu prowadzonych badań oraz prac wdrożeniowych potwierdzone. Ich szczegółowe omówienie zaprezentowano dla każdej z osobna, rozpoczynając od przypomnienia jej treści, a następnie opisu sposobu i wyników ich weryfikacji.

6.1.1 Hipoteza główna

Zastosowanie hybrydowego systemu wizyjno-radiowego do analizy cyfrowego modelu inteligentnego środowiska domowego pozwala na usprawnienie procesu odzyskania semantycznego znaczenia jego komponentów poprzez automatyzację ich rozpoznawania i pozycjonowania.

Potwierdzenie hipotezy głównej całej pracy miało miejsce w ramach rozdziału o eksperymentalnych badaniach końcowych (Roz. 5.7), szczególnie w Roz. 5.7.3,

gdzie przedstawiono konkretne wyniki liczbowe. Spośród wszystkich 208, obecnych w środowisku pomiarowym obiektów, 62% zostało wykrytych na chociaż jednym zdjęciu panoramicznym, a ponad 55% wykrywanych obiektów było zawsze klasyfikowanych zgodnie z ich rzeczywistym rodzajem, tj. etykietą klasy (Rys. 5.34 i Rys. 5.35). Jednocześnie, znakomita większość obiektów była pozycjonowana z precyzją umożliwiającą ich zlokalizowanie w promieniu poniżej jednego metra od wskazanego punktu w przestrzeni trójwymiarowej modelu budynku (Rys. 5.37). Wdrożeniowo pozwala to na odnalezienie w przestrzeni różnorodnych komponentów środowiska, od okien i drzwi, przez meble, aż po pojedyncze przedmioty. Przedstawiono również odwzorowanie wyznaczonych współrzędnych z powrotem w modelu trójwymiarowym (Rys. 5.39). Uzyskane wyniki uznaje się więc za wystarczające do potwierdzenia i przyjęcia hipotezy głównej pracy, jako w całości prawdziwej.

6.1.2 Hipoteza dodatkowa 1

Zastosowanie cyfrowego modelu środowiska badawczego pozwala na radiową estymację dystansu pomiędzy punktami w przestrzeni bryły budynku o równie wysokiej jakości, co jej wyznaczenie bazując na pomiarach manualnych.

Potwierdzenie pierwszej hipotezy dodatkowej uzyskano na skutek opracowanego prototypu systemu radiowego, omówionego w Roz. 3.3, szczególnie dzięki analizie porównawczej uzyskanych wyników estymacji dystansu z Roz. 3.3.5. Różnice błędów estymacji odległości pomiędzy wybranymi punktami i dla różnych kombinacji źródeł danych referencyjnych, a także technologii radiowych i wykorzystywanych protokołów, przedstawione zostały w postaci rozbudowanych tabel porównawczych (Tab. 3.5 i Tab. 3.6). Za najważniejszy przykład uznać możemy jednakże estymację dystansu dla technologii UWB, dla której bezwzględna różnica wyznaczonych odległości, pomiędzy pomiarami manualnymi z użyciem dalmierza laserowego a odczytami z modelu trójwymiarowego, wyniosła maksymalnie 5 cm, tj. około 3% wartości. Wyniki takie uznaje się za potwierdzające hipotezę o zdolności wykorzystania cyfrowego modelu środowiska do procesu estymacji odległości i braku jego istotnego wpływu na końcowe wyniki.

6.1.3 Hipoteza dodatkowa 2

Przetwarzanie wstępne zbiorów zdjęć, polegające na doborze południka rozcięcia sfery (punktu początkowego w procesie rzutowania obrazów sferycznych na reprezentację panoramiczną), pozwala na zwiększenie efektywności wyszkolonego modelu detektora obiektów i jego działania w środowisku domowym.

Potwierdzenie drugiej hipotezy dodatkowej rozprawy wypracowano na skutek implementacji oraz badań opracowanego algorytmu doboru najlepszego południka rozcięcia sfery, który ze szczegółami omówiono w Roz. 4.2.2. Wyniki liczbowe przedstawiono natomiast w zestawieniach tabelarycznych (Tab. 4.4 i Tab. 4.5) w Roz. 4.2.2. Zbiór danych przetworzony zgodnie z opracowanym algorytmem (oznaczony jako *AutoMin*) w większości proponowanych konfiguracji testowych ogółem, a także we wszystkich konfiguracjach dotyczących największego z badanych modeli sieci neuronowej detektora obiektów (*Extra-Large*), dostarczył najwyższej skuteczności wykrywania obiektów. Pomiedzy scenariuszem najbardziej pozytywnym (zastosowaniem proponowanego algorytmu doboru najlepszego południka rozcięcia), a najbardziej negatywnym (algorytmie zmodyfikowanym na dobór najgorszej propozycji rozcięcia) różnica wartości wskaźnika mAP sięgnęła poziomu 8 p.p. Uzyskane wyniki uznaje się więc za potwierdzające drugą hipotezę dodatkową pracy i podkreślające istotność przetwarzania wstępnego posiadanych zbiorów danych.

6.2 Aspekt wdrożeniowy projektu

Spełnienie oczekiwań przemysłowych partnera doktoratu (Simlab Sp. z o.o.) było kluczowym elementem wszystkich przeprowadzonych badań i opracowanych rozwiązań. Wdrażanie ich wyników do codziennego funkcjonowania firmy było natomiast obecne na każdym etapie prac i systematycznie kierunkowało rozwój produktów. Najbardziej widoczne jest to w platformie chmurowej do zarządzania inteligentną przestrzenią domową SIM-ON. To właśnie w tym produkcie dostosowana wersja hybrydowej metody rozpoznawania i pozycjonowania komponentów cyfrowego bliźniaka została udostępniona pierwszym klientom. Opracowany został nowy

panel aplikacji, który pozwala na przejście przez procedurę automatycznej detekcji obiektów oraz oznaczenie ich położenia w przestrzeni przy użyciu dedykowanych struktur danych (tzw. „*Assetów*”). Rozwiązanie to oparto na interaktywnym spacerze oraz budującym go modelu trójwymiarowym, z którego eksportowane są obrazy sferyczne w postaci panoram ERP. Dzięki wykorzystaniu profesjonalnego skanera przestrzeni mają one precyzyjnie określone położenie kamery, w stosunku do którego przeprowadzany jest następnie proces pozycjonowania wizyjnego oraz ostateczne wyznaczenie położenia obiektów. Dzięki wprowadzeniu systemu wprost w środowisko cyfrowego bliźniaka możliwe stało się również zastąpienie detektora głębi procesem śledzenia promienia (z ang. „*Ray tracing*”), a za końcowe położenie obiektu wybierana jest pierwsza kolizja śledzonego promienia z otaczającym kamerę modelem budynku. Rozwiązanie to nie wymaga od użytkowników posiadania radiowego komponentu systemu hybrydowego i może zostać zaprezentowane na bazie samej tylko zdigitalizowanej nieruchomości. Niemniej, do oferty firmy wprowadzony został również opracowany system radiowy oraz urządzenia go budujące. Poszukiwani są obecnie pierwsi klienci, którzy chcieliby dostosować go do swoich potrzeb i wprowadzić do stałego użytku.

W sposób pośredni, dzięki kierunkowaniu prowadzonych prac, swoje wdrożenie do produktów firmy odnalazły również pozostałe elementy opisane w rozprawie. Wykonana analiza literatury pozwoliła na odniesienie opracowywanych funkcji do szerszego kontekstu obecnie znanego stanu nauki. Jej uzupełnienie o rozeznanie konkurencji rynkowej oraz pełną analizę SWOT produktów wywarło wymierny skutek na płaszczyźnie prowadzonych działań marketingowych, pozycjonowania produktów i oferty firmy na rynku. Pozwoliło to również na określenie priorytetów rozwoju firmy, które znalazły swoje ugruntowanie w rozbudowanej analizie potrzeb i oczekiwań klientów. Opracowana na tej podstawie pula klas obiektów oraz zgromadzony materiał wizyjny (zbiory zdjęć panoramicznych) stanowią dobry punkt wyjścia do dalszych prac i szkolenia kolejnych modeli uczenia maszynowego.

Dodatkowym rezultatem przeprowadzonych prac jest wykształcenie wewnętrznego schematu postępowania, który stanowi bezpośrednią wartość dla przedsiębiorstwa. Dotychczas firma nie posiadała rozwiązań bazujących na metodach sztucznej inteligencji ani doświadczenia w zakresie ich tworzenia. W ramach doktoratu wypracowano pewien konkretny proces postępowania, obejmujący cały cykl życia no-

wego elementu oprogramowania oraz poruszający bardzo rozległy wachlarz zagadnień: od identyfikacji potrzeb oraz wymagań funkcjonalnych klientów, przez analizę literatury i rozwiązań konkurencyjnych, gromadzenie oraz przetwarzanie materiału źródłowego, budowę zbiorów danych, aż po opracowywanie i testowanie prototypów oraz wytyczne utrzymania produktu. Tak zdefiniowany schemat ma charakter uniwersalny i może być stosowany w przyszłych projektach badawczo-rozwojowych. Co istotne, jego realizacja umożliwiła zbudowanie w przedsiębiorstwie podstaw kompetencyjnych w zakresie prowadzenia prac wykorzystujących narzędzia sztucznej inteligencji oraz systematyzację kolejnych etapów rozwoju produktu. W praktyce oznacza to, że kolejne podobne przedsięwzięcia będą mogły być realizowane w sposób bardziej efektywny, z wykorzystaniem doświadczeń oraz metodologii opracowanej w toku niniejszej rozprawy.

6.3 Dalszy rozwój projektu

Całokształt zrealizowanych prac wskazuje wiele kierunków dalszych badań oraz nowych funkcji, których realizację warto przewidzieć w produktach firmy partnera doktoratu. Najważniejsze z nich przedstawione zostają w tym rozdziale.

Omówiony w Roz. 5.4 sposób określania pomieszczenia, w którym znajduje się mobilne urządzenie pomiarowe, oparty jest obecnie na średniej arytmetycznej odczytanych wartości RSSI. Rozwiązanie to jest skuteczne, lecz pozostawia wiele możliwości dalszego udoskonalenia. Potencjalne usprawnienia mogą dotyczyć różnych aspektów: od filtracji gromadzonych pomiarów, przez dedykowane klasyfikatory pomieszczeń, aż po zupełną zmianę przyjętego postępowania. Wprowadzenie dodatkowej filtracji mogłoby posłużyć stabilizacji wyników. Specjalnie wyszkolony klasyfikator (np. oparty na sztucznych sieciach neuronowych) byłby w stanie uchwycić zależności niewidoczne dla człowieka. Proponowana metoda hybrydowa nie ogranicza również końcowego wdrożenia tylko i wyłącznie do pozycjonowania radiowego. Potencjalnie, selekcja pomieszczenia mogłaby zostać zrealizowana również wizyjnie, np. poprzez klasyfikację wykonanych zdjęć.

Analogiczne uwagi oraz możliwości dalszych usprawnień odnoszą się do etapu pozycjonowania precyzyjnego (opisanego w Roz. 5.5). Obecnie pozycja wyznaczana jest na podstawie średniej wartości uzyskanej w procesie trilateracji, jednak zasto-

sowanie dodatkowych mechanizmów, takich jak wykrywanie wartości odstających, filtracja pomiarów czy przewidywanie zmian odczytów (np. w scenariuszu monitorowania obiektów ruchomych), mogłoby znacząco poprawić jakość i stabilność wyników.

W aspekcie wdrożeniowym, interesującym krokiem w dalszych badaniach może być sposób doboru implementowanego rozwiązania do faktycznych potrzeb użytkowników. Jak pokazano przy omówieniu wyników systemu fuzyjnego (Roz. 5.7.3), niedobór kotwic referencyjnych stanowi zauważalne wyzwanie i przeszkodę dla skutecznego wykorzystania lokalizatorów radiowych. Spodziewać się jednak można, że całkowita automatyzacja każdego kroku analizy środowiska nie jest konieczna, a dla użytkowników końcowych pewien stopień pracy manualnej będzie akceptowalny. Otwartym pytaniem pozostaje sposób doboru elementów koniecznych wykonywanych automatycznie, a tych, które warto zrealizować z pomocą człowieka.

W ostatnim kroku metody hybrydowej, pozycjonowaniu wizyjnym z Roz. 5.6, dość naturalnym krokiem dalszego rozwoju projektu jest faktyczne wdrożenie wspomnianej automatyzacji procesu synchronizacji orientacji zdjęć sferycznych, nawet w sytuacji użycia amatorskiego sprzętu fotograficznego. Wykonana w obecnej wersji rozwiązania praca manualna zastąpiona może zostać detektorem obiektów, który dostarczy informacji o wzajemnej orientacji kolejnych zdjęć i umożliwi ich automatyczne przewinięcie do uśrednionego układu odniesienia. Konieczne będzie do tego jednak zgromadzenie wystarczająco dużego i różnorodnego zbioru zdjęć, a także wyszkolenie dedykowanego detektora obiektów.

W zagadnieniu detektorów obiektów, dalsze badania przeprowadzić można również w kierunku potwierdzenia hipotezy o wpływie rodzaju pomieszczenia na typowość umieszczanych w nim przedmiotów, a tym samym na końcową skuteczność detekcji. Dostrzeżona została różnica w jakości wyników pomiędzy kolejnymi rodzajami pomieszczeń, co pozwala wysnuć wniosek, że przedmioty typowo wykorzystywane np. w łazienkach są trudniejsze do rozpoznania niż te z sypialni czy salonu.

Za oczywiste uznaje się również konieczność dalszego aktualizowania wykorzystywanych modeli uczenia maszynowego do ich najnowszych architektur i wersji. Traktować to jednak należy raczej jako element utrzymania oprogramowania niż jego rzeczywistego rozwoju. Przewidując kontynuację badań, warto zwró-

cić uwagę na wyzwania związane przede wszystkim z bezpieczeństwem danych oraz prywatnością użytkowników. Już obecnie pozyskiwanie zbiorów danych odwzorowujących środowisko domowe stanowi poważne utrudnienie, a w przyszłości należy spodziewać się dalszych ograniczeń w tym zakresie. Jest to jednak zmiana uzasadniona troską o dobro użytkowników, a jednocześnie wskazująca kierunek rozwoju rozwiązań opartych na danych w pełni syntetycznych (podobnie do badań opisanych w Roz. 4.2.3).

6.4 Dorobek publikacyjny autora

Prowadzone prace badawcze oraz ich wyniki były regularnie upubliczniane możliwie szerokiemu gronu odbiorców pod dwoma postaciami: publikacji naukowych (posiadających określoną punktację Ministerstwa Nauki i Szkolnictwa Wyższego) i publikacji popularnonaukowych (kierowanych do odbiorców nieakademickich).

6.4.1 Publikacje naukowe

W czasie doktoratu zrealizowano cztery publikacje o łącznej wartości 380 punktów MNiSW. W spisie uwzględniono również dwie publikacje zrealizowane wcześniej (dodatkowe 110 punktów MNiSW), bezpośrednio jednak związane tematycznie z prowadzonymi badaniami i stanowiące fundament oraz punkt wyjścia dla opisanych w rozprawie prac. Na moment złożenia rozprawy całkowity dorobek naukowy autora uzyskał łącznie 490 punktów MNiSW i prezentowany jest w Tab. 6.1.

Tabela 6.1: Publikacje z punktacją ministerialną

Rok	Pełne cytowanie	Punkty
2024	Pokuciński Sebastian, Mrozek Dariusz: <i>Methods and applications of space understanding in indoor environment - a decade survey</i> , <i>Applied Sciences-Basel</i> , MDPI, vol. 14, no. 10, 2024, Article number: 3974, pp. 1–55, DOI: 10.3390/app14103974.	100

Rok	Pełne cytowanie	Punkty
2024	Pokuciński Sebastian, Filus Katarzyna, Mrozek Dariusz: <i>Finding the perfect cut: selection of the best cutting point in equirectangular panoramas for object detection</i> , <i>Procedia Computer Science</i> , Elsevier BV, vol. 246, 2024, pp. 519–528, DOI: 10.1016/j.procs.2024.09.432.	70
2023	Pokuciński Sebastian, Mrozek Dariusz: <i>YOLO-based object detection in panoramic images of smart buildings</i> , In: <i>2023 IEEE 10th International Conference on Data Science & Advanced Analytics (DSAA)</i> , October 9–12, 2023, Thessaloniki, Greece. Proceedings / Manolopoulos Yannis, Zhou Zhi-Hua (eds.), 2023, Institute of Electrical and Electronics Engineers, ISBN 979-8-3503-4503-2, pp. 1–9, DOI: 10.1109/DSAA60987.2023.10302568.	140
2023	Pokuciński Sebastian, Mrozek Dariusz: <i>Object detection with YOLOv5 in indoor equirectangular panoramas</i> , <i>Procedia Computer Science</i> , Elsevier BV, vol. 225, 2023, pp. 2420–2428, DOI: 10.1016/j.procs.2023.10.233.	70
2020	Grzechca Damian, Tokarz Krzysztof, Hanzel Krzysztof, Pokuciński Sebastian: <i>Fundamental issues when creating an autonomous platform for a UWB navigation system</i> , <i>Control Engineering and Applied Informatics</i> , vol. 22, no. 1, 2020, pp. 84–93.	40

Rok	Pełne cytowanie	Punkty
2020	Grzechca Damian, Ziębiński Adam, Komorowski Dariusz, Hanzel Krzysztof, Pokuciński Sebastian, Gorywoda Michał, Klonowski Kamil, Siwek Sławomir, Kristof Robert, Czerwiński Arkadiusz: <i>Monitoring the gait process during the rehabilitation of patients using computer vision techniques and UWB technology</i> , In: <i>Information Systems: 16th European, Mediterranean, and Middle Eastern Conference, EMCIS 2019, Dubai, UAE, December 9–10, 2019. Proceedings</i> / Themistocleous M., Papadaki M. (eds.), Lecture Notes in Business Information Processing, 2020, Springer, ISBN 978-3-030-44321-4, pp. 419–437, DOI: 10.1007/978-3-030-44322-1_31.	70

6.4.2 Publikacje popularnonaukowe

Uzupełnieniem publikacji naukowych był udział autora w wielu edycjach krajowych konferencji, gdzie wyniki badań przedstawiano zarówno w formie prezentacji, jak i posterów. Przedstawiane w ten sposób opracowania nosiły charakter popularnonaukowy i nie posiadały punktacji MNiSW. Przedstawiono je w Tab. 6.2. Na liście przedstawiono 14 pozycji, w tym również jedną publikację sprzed rozpoczęcia doktoratu, ponownie, ściśle związaną z tematyką rozprawy.

Tabela 6.2: Publikacje popularnonaukowe (bez punktacji ministerialnej).

Rok	Pełne cytowanie
2024	Pokuciński Sebastian: <i>Komponenty środowiska domowego: klasteryzacja, segmentacja i klasyfikacja pokojów</i> , In: <i>Analiza zagadnień i wyników. Materiały Konferencji Młodych Naukowców: Analiza zagadnienia, analiza wyników – Edycja VI</i> , 15–17.03.2024, Kraków. <i>Streszczenia wystąpień młodych naukowców 6. Materiały konferencyjne</i> / Kuczera Marcin (eds.), 2024, Creativetime, ISBN 978-83-66772-33-5, pp. 45–45.

Rok	Pełne cytowanie
2024	Pokuciński Sebastian: <i>Narzędzia i platformy rozwojowe sztucznej inteligencji</i> , In: <i>Analiza zagadnień i wyników. Materiały Konferencji Młodych Naukowców: Analiza zagadnienia, analiza wyników – Edycja VI</i> , 15–17.03.2024, Kraków. <i>Streszczenia wystąpień młodych naukowców 6. Materiały konferencyjne</i> / Kuczera Marcin (eds.), 2024, Creativetime, ISBN 978-83-66772-33-5, pp. 54–54.
2023	Pokuciński Sebastian: <i>Boom na AI – szansa czy klątwa</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIII. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-29-8, pp. 82–82.
2023	Pokuciński Sebastian: <i>Gdy wizja komputerowa spotyka lokalizację radiową – hybrydowe podejście do monitorowania pacjenta</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIV. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-31-1, pp. 98–98.
2023	Pokuciński Sebastian: <i>Gdzie jestem i skąd to wiem? – analiza lokalizacji radiowej</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIV. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-31-1, pp. 96–96.
2023	Pokuciński Sebastian: <i>Morze możliwości, a może morze problemów? Formaty danych dla Sztucznej Inteligencji</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIII. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-29-8, pp. 81–81.

Rok	Pełne cytowanie
2023	Pokuciński Sebastian: <i>Narzędzia akwizycji danych dla spacerów trójwymiarowych</i> , In: <i>Inżynieria – spojrzenie młodych naukowców. Edycja II. Streszczenia wystąpień. Materiały konferencyjne. Materiały Konferencji Młodych Naukowców: Inżynieria – spojrzenie młodych naukowców, 20–21.05.2023, Kraków</i> / Słowik Magdalena, Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-25-0, pp. 32–32.
2023	Pokuciński Sebastian: <i>Protokoły komunikacyjne urządzeń Internetu Rzeczy – standardy i trendy</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIV. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-31-1, pp. 97–97.
2023	Pokuciński Sebastian, Pietryga Filip: <i>Skakanie po ekranie – techniki wizji komputerowej i SI w analizie szkieletu człowieka</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIV. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-31-1, pp. 95–95.
2023	Pokuciński Sebastian: <i>Sztuczna inteligencja a prywatność</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIII. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-29-8, pp. 83–83.
2023	Pokuciński Sebastian: <i>Zamek, a zamek – homonimy i inne wyzwania dla klasyfikatorów</i> , In: <i>Nowe wyzwania dla polskiej nauki. Edycja XIII. Materiały konferencyjne. Streszczenia wystąpień</i> / Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-29-8, pp. 84–84.

Rok	Pełne cytowanie
2023	Pokuciński Sebastian: <i>Zarządzanie i oznaczanie zbiorów danych dla rozwiązań sztucznej inteligencji</i> , In: <i>Inżynieria – spojrzenie młodych naukowców. Edycja II. Streszczenia wystąpień. Materiały konferencyjne. Materiały Konferencji Młodych Naukowców: Inżynieria – spojrzenie młodych naukowców, 20–21.05.2023, Kraków / Słowik Magdalena, Kuczera Marcin (eds.), 2023, Creativetime, ISBN 978-83-66772-25-0, pp. 31–31.</i>
2022	Pokuciński Sebastian: <i>Segmentacja i klasyfikacja pokoiów w środowisku domowym</i> , In: <i>Nowe trendy w badaniach naukowych 5, Kraków, 21–23.10.2022: Streszczenia wystąpień młodych naukowców. Materiały konferencyjne / Słowik Magdalena, Kuczera Marcin (eds.), 2022, Creativetime, ISBN 978-83-66772-19-9, pp. 53–53.</i>
2019	Grzechca Damian, Ziębiński Adam, Komorowski Dariusz, Hanzel Krzysztof, Pokuciński Sebastian, Gorywoda Michał, Klonowski Kamil, Siwek Sławomir, Kristof Robert, Czerwiński Arkadiusz: <i>Stanowisko badawcze do oceny postępów rehabilitacyjnych kończyn dolnych pacjenta oparte na analizie wizyjnej, systemie pozycjonowania UWB oraz wkładce tensometrycznej</i> , <i>PM News</i> , no. 21, 2019, pp. 9.

6.5 Inna aktywność naukowo-badawcza autora

Jeszcze przed rozpoczęciem studiów doktoranckich autor brał czynny udział w realizacji grantu badawczego prowadzonego w przedsiębiorstwie partnerskim. W trakcie doktoratu angażował się także w proces przygotowywania nowych wniosków grantowych, pełniąc rolę wewnętrznego konsultanta. Podjęte aktywności zostały szczegółowo omówione w dalszej części rozdziału.

Grant badawczy Realizacja grantu finansowanego przez Narodowe Centrum Badań i Rozwoju (projekt nr POIR.04.01.01-00-0028), prowadzonego przez firmę

Simlab Sp. z o.o. w okresie od maja 2020 do stycznia 2022, dotyczyła opracowania systemu “Efektywnego zarządzania energią w inteligentnych budynkach z rozwiązaniami IoT, bazującego na modelu Digital Twin”. Celem przedsięwzięcia było stworzenie skalowalnej i otwartej platformy wspomagającej optymalizację zużycia energii w budynkach mieszkalnych i biurowych. Zakres prac obejmował m.in. opracowanie metodyki tworzenia parametrycznych modeli obiektów i urządzeń, metodyki pomiaru strat cieplnych odbiorników energii oraz algorytmów automatycznej integracji skanów 3D z modelami funkcjonalnymi budynków. Kluczowym rezultatem było przygotowanie systemu do zdalnego monitoringu zużycia energii w sieci ośmiu obiektów testowych oraz opracowanie wirtualnego modelu energetycznego obiektu, powiązanego z jego trójwymiarowym odwzorowaniem. Udział autora koncentrował się na integracji modeli cyfrowych z infrastrukturą IoT oraz wsparciu prac nad rozwojem algorytmów i środowiska testowego, co umożliwiło zdobycie praktycznego doświadczenia w zastosowaniach koncepcji Digital Twin do zagadnień efektywności energetycznej w inteligentnych budynkach.

Wnioski grantowe Po zakończeniu realizacji pierwszego projektu badawczego autor został włączony w prace zespołu zajmującego się przygotowywaniem kolejnych wniosków o finansowanie badań, zarówno na poziomie krajowym, jak i międzynarodowym. Uczestniczył w projektowaniu koncepcji, przygotowywaniu opisów technicznych, analizie aspektów badawczych oraz redakcji treści aplikacji. Do najważniejszych zgłoszeń, w których brał udział jako współtwórca, należały m.in.:

- wniosek do programu **Fundusze Europejskie dla Śląskiego 2021–2027** (projekt pt. *Nowe rozwiązania umożliwiające zarządzanie wielodzielinowym zbiorem danych dla sektora konstrukcyjno-budowlanego*, nr FESL.01.02-IP.01-08F6/24-001),
- wniosek do programu **Fundusze Europejskie dla Nowoczesnej Gospodarki – Ścieżka SMART** (projekt pt. *Środowisko Digital Twin dla dowolnej nieruchomości, wspierające komunikację interesariuszy w trakcie całego życia budynku, zarządzanie IoT oraz procesy FM*, nr FENG.01.01-IP.02-001/23),
- wcześniejszy wniosek do **NCBiR** w ramach Programu Operacyjnego Inteli-

gentny Rozwój (*STAGES – cyfrowa platforma do zarządzania wielodzielniowym zbiorem danych*, nr POIR.01.01.01-00-0348/21),

- aplikacja do programu **Horizon Europe – EIC Accelerator** (*SimOn – Digital Home Experience*, nr 190186420), która, mimo braku finansowania, uzyskała prestiżowe wyróżnienie Komisji Europejskiej w postaci odznaki *Seal of Excellence*.

Choć żaden z opisanych wniosków nie zakończył się przyznaniem środków, udział w ich opracowywaniu pozwolił autorowi na zdobycie cennego doświadczenia w zakresie konstruowania aplikacji grantowych, analizy kryteriów oceny projektów oraz pracy zespołowej w międzynarodowym środowisku.

Inna aktywność Równolegle autor pozostawał aktywnym członkiem Studenckiego Koła Naukowego Data Science. Zaangażowanie obejmowało m.in. doradzanie młodszym studentom w realizacji projektów, prowadzenie dyskusji tematycznych oraz współorganizację warsztatów i wydarzeń popularyzujących naukę. Efektem tej działalności było m.in. przygotowanie interaktywnego stanowiska prezentującego możliwości wizyjnej analizy szkieletu człowieka, które zostało zaprezentowane podczas *Nocy Naukowców* na Politechnice Śląskiej. Wcześniej, w ramach współpracy Koła, autor brał udział również w inicjatywach związanych z popularyzacją sztucznej inteligencji w codziennym życiu, czego przykładem były demonstratory systemów śledzących sylwetkę użytkownika i zachęcających do aktywności fizycznej.

Spis pojęć i skrótów

Tabela 6.3: Zestawienie skrótów i objaśnień wykorzystywanych w rozprawie.

Skrót	Objaśnienie
AEC	Architecture, Engineering, Construction Architektura, Inżynieria, Budownictwo Branża zajmująca się projektowaniem, realizacją i zarządzaniem budynkami oraz infrastrukturą.
AGV	Automated Guided Vehicle Zautomatyzowany pojazd bezzałogowy Samobieżny wózek transportowy, zdalnie sterowany pojazd.
AI	Artificial Intelligence Sztuczna inteligencja Opis patrz: SI.
ANN	Artificial Neural Network Sztuczna sieć neuronowa Opis patrz: SSN.
AoA	Angle of Arrival Kąt nadejścia fali radiowej Metoda lokalizacji na podstawie analizy kierunku, z którego nadchodzi fala radiowa.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
BB	Bounding Box Ramka ograniczająca Sposób oznaczenia obiektu na zdjęciu przy użyciu ramki okalającej obiekt.
BCPSM	Best Cutting Point Selection Method Metoda Doboru Najlepszego Punktu Rozcięcia Autorski algorytm doboru punktu rozpoczęcia rzutowania sfery na płaszczyznę.
BFoV	Bounding Field of View Ograniczające pole widzenia Rodzaj etykiety obiektu dla procesu detekcji na obrazach sferycznych.
BIM	Building Information Modeling Modelowanie informacji o budynku Cyfrowy model łączący dane geometryczne, techniczne i funkcjonalne na potrzeby całego cyklu życia obiektu.
BLE	Bluetooth Low Energy Protokół komunikacji bezprzewodowej zaprojektowany do niskiego zużycia energii.
BS	Batch Size Rozmiar wsadu Liczba próbek przetwarzanych w ramach jednej paczki danych.
CLI	Command-line Interface Interfejs wiersza poleceń Interfejs tekstowy do wprowadzania komend.
CNN	Convolutional Neural Network Konwolucyjna sieć neuronowa Klasa głębokich sieci neuronowych wykorzystujących proces konwolucji do uzyskania wyników.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
CS	Confidence Score Wynik pewności Miara przekonania modelu co do poprawności detekcji.
CSA	Connectivity Standards Alliance Organizacja standaryzacyjna Związek odpowiedzialny m.in. za rozwój protokołu Matter.
CV	Computer Vision Wizja komputerowa Dziedzina nauki poświęcona komputerowej analizie obrazów.
DS-TWR	Double-Sided TWR Dwustronny TWR Komunikacja TWR realizowana dwustronnie (z opcjonalnym potwierdzeniem).
DT	Digital Twin Cyfrowy bliźniak Elektroniczny model rzeczywistości.
EC	Exclusion Criterium Kryterium wykluczające Warunek logiczny, którego nie wolno spełnić.
ERP	Equirectangular Panorama Panorama równokątna Format zapisu obrazu sferycznego do macierzy pikseli o proporcjach 2:1 (szer:wys).
GT	Ground Truth Dane referencyjne Wartości oczekiwane wykorzystywane do oceny jakości detekcji.
IC	Inclusion Criterium Kryterium inkluzywne Warunek logiczny konieczny do spełnienia.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
IE	Indoor Environment Środowisko wewnętrzne Szeroko rozumiane wnętrza nieruchomości, przestrzenie wewnętrzzbudynkowe.
IERP	Indoor Equirectangular Panorama Panorama równokątna wewnętrzzbudynkowa Odmiana zdjęć typu ERP obrazująca środowiska wewnątrz budynków.
IoT	Internet of Things Internet rzeczy Idea tworzenia przedmiotów z dostępem do internetu.
IoU	Intersection over Union Iloraz przecięcia do sumy Wskaźnik pokrycia w procesie detekcji obiektów.
MAC	Media Access Control Identyfikator sprzętowy urządzenia.
MAE	Mean Absolute Error Średni błąd bezwzględny Miara agregacyjna błędu pomiaru.
mAP	Mean Average Precision Uśredniona precyzja średnia Miara precyzji detektora obiektów.
MQTT	Message Queuing Telemetry Transport Protokół wymiany wiadomości w systemach IoT.
NLoS	Non-Line-of-Sight Brak widoczności bezpośredniej Warunki otoczenia podczas pozycjonowania, ograniczające widoczność sygnału.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
OOM	Out of Memory Brak pamięci Typowy błąd określający brak pojemności pamięci operacyjnej do wykonania żądanej operacji.
PDoA	Phase Difference of Arrival Różnica fazy nadejścia fali radiowej Metoda estymacji położenia nadajnika z użyciem różnicy fazy sygnałów radiowych.
PRC	Precision-Recall Curve Krzywa precyzja-czułość Wykres zależności precyzji od czułości dla wyników detektora obiektów.
RBFoV	Rotated Bounding Field of View Obrócone ograniczające pole widzenia Rodzaj etykiety obiektu w procesie detekcji obiektów.
ReLU	Rectified Linear Unit Jedna z funkcji aktywacji sztucznego neuronu.
RMSE	Root Mean Square Error Błąd średniokwadratowy Miara agregacyjna błędu pomiaru.
RSSI	Received Signal Strength Indicator Wskaźnik mocy odebranego sygnału Wartość określająca moc sygnału odbieranego przez urządzenie radiowe.
RQ	Research Question Pytanie badawcze Formalnie sformułowane pytanie, na które w badaniach poszukiwana jest odpowiedź.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
SI	Sztuczna Inteligencja Dziedzina nauki i technologii zajmująca się tworzeniem systemów zdolnych do naśladowania ludzkiej inteligencji.
SLAM	Simultaneous Localization and Mapping Jednoczesna lokalizacja i mapowanie Technologia do autonomicznej nawigacji robotów i modelowania przestrzeni.
SLR	Systematic Literature Review Systematyczny przegląd literatury Metodyka ściśle określonego i sformalizowanego prowadzenia przeglądu literatury.
SP	Spherical Polyhedron Siatka trójkątów Rodzaj odwzorowania obrazu sferycznego na płaszczyźnie przy użyciu siatki trójkątów.
SQ	Search Query Zapytanie wyszukiwania Fraza podawana do silników wyszukiwania baz danych w celu odnalezienia rekordów o żądanych parametrach.
SSD	Single Shot Detector Detektor jednostopniowy Detektor obiektu dokonujący wskazania i klasyfikacji obszaru przy jednym przebiegu sieci.
SSN	Sztuczna Sieć Neuronowa Model obliczeniowy złożony z warstw sztucznych neuronów, stosowany w uczeniu maszynowym.
SS-TWR	Single-Sided TWR Jednostronny TWR Komunikacja TWR realizowana jednostronnie.

Kontynuacja na następnej stronie

Tabela 6.3 — kontynuacja z poprzedniej strony

Skrót	Objaśnienie
ToA	Time of Arrival Czas nadejścia sygnału Opis patrz: ToF.
ToF	Time of Flight Czas przelotu sygnału Czas potrzebny na dostarczenie wiadomości radiowej od nadajnika do odbiornika.
TWR	Two-Way Ranging Dwukierunkowy pomiar odległości Technika pomiaru odległości między dwoma urządzeniami radiowymi.
UWB	Ultra-Wideband Technologia komunikacji bezprzewodowej bardzo szerokiego pasma częstotliwości.
VT	Virtual Tour Interaktywny spacer wirtualny Interaktywny sposób eksploracji wirtualnej przestrzeni.
YOLO	You Only Look Once Rodzina architektur konwolucyjnych sieci neuronowych typu detektora jednostopniowego.
ZSD	Zero-shot Detector Detektor bezstopniowy Architektura sieci konwolucyjnej dedykowana rozpoznawaniu klas nieobecnych w zbiorach treningowych.

Spis rysunków

2.1	Schemat protokołu przeglądu literatury z informacją o liczbie publikacji po każdym kolejnym kroku. Czerwone liczby oznaczają główny przebieg protokołu, a niebieskie jego rozszerzenie. Wprowadzono kolorystyczne oznaczenia typów danych wejściowych: obrazy 2D (zielony), dane przestrzenne (jasnoniebieski), struktury grafowe (czerwony), artykuły przeglądowe (szary) oraz zbiory cech (żółty).	24
2.2	Diagram pierwszej opracowanej taksonomii: grupowanie publikacji na podstawie typów przetwarzanych danych.	25
2.3	Diagram drugiej opracowanej taksonomii: grupowanie publikacji na podstawie kategorii sposobu przetwarzania danych.	28
2.4	Diagram trzeciej opracowanej taksonomii: grupowanie publikacji na podstawie realizowanego zadania.	32
2.5	Wizualizacja zadania rekonstrukcji modelu 3D. Na podstawie wejściowego obrazu 2D (planu pomieszczeń, po lewej) generowany jest uproszczony model trójwymiarowy (po prawej).	34
2.6	Wizualizacja zadania wyszukiwania obrazem. Obraz zapytania (u góry) służy do przeszukania bazy danych w celu znalezienia 3 najbardziej podobnych planów pięter (na dole).	34
2.7	Wizualizacja zadania tworzenia opisu środowiska. Plan piętra wejściowego (po lewej) jest analizowany i przekształcany w zestaw zdań opisujących jego elementy oraz ich położenie (po prawej). . .	35

- 2.8 Wizualizacja zadania wektoryzacji planu pomieszczeń. Wejściowy obraz rastrowy (**po lewej**) zostaje przekształcony na reprezentację wektorową (**po prawej**). Przekształcone ściany oznaczono czarnymi liniami, drzwi niebieskimi, okna zielonymi, a połączenia czerwonymi okręgami. 35
- 2.9 Wizualizacja zadania generacji planu pomieszczeń. Dane wejściowe (zarejestrowana trajektoria użytkownika, **po lewej**) są wykorzystywane do oszacowania możliwego planu piętra (na podstawie gęstości ścieżek i ich klasteryzacji **po prawej**). 36
- 2.10 Wizualizacja zadania tworzenia grafu. Dane wejściowe (zarejestrowana trajektoria użytkownika wraz z pewnym zrozumieniem układu budynku, **po lewej**) są wykorzystywane do utworzenia struktury grafu (**po prawej**), w której pomieszczenia i obszary reprezentowane są jako węzły, a ich sąsiedztwo jako krawędzie grafu. 36
- 2.11 Wizualizacja zadania klasyfikacji pomieszczeń. Przykładowy plan piętra wejściowego (**po lewej**) jest analizowany w celu wykrycia obszarów pomieszczeń. Każdy obszar zostaje sklasyfikowany i oznaczony etykietą typu pomieszczenia (**po prawej**). 37
- 2.12 Wizualizacja zadania wykrywania zmian. Obrazy początkowy i zaktualizowany (**u góry**) są analizowane w celu wykrycia różnic między nimi. Wykryte niezgodne elementy zostały zaznaczone na czerwono (**na dole**). 37
- 2.13 Wizualizacja zadania segmentacji. Obraz wejściowy (**po lewej**) jest analizowany w celu wykrycia oddzielnych obszarów reprezentujących poszczególne pomieszczenia. Każde pomieszczenie zostało oznaczone wielokątem w innym kolorze (**po prawej**). 38
- 2.14 Wizualizacja zadania nawigacji. Dane wejściowe obejmują wybrane punkty początkowy i końcowy (**po lewej**). W wyniku przetwarzania generowana jest proponowana ścieżka, przedstawiona jako zestaw instrukcji (**po prawej**). 38
- 2.15 Wizualizacja zadania dopasowania mapy. Zbiór nieskalibrowanych planów wejściowych (**po lewej**) zostaje połączony w jeden wspólnie wyrównany rezultat (**po prawej**). 39

2.16	Diagram faz cyklu rynkowego z zaznaczeniem (czerwona ramka) aktualnego umiejscowienia branży zarządzania nieruchomościami. .	43
3.1	Schemat komunikacji w procesie dwukierunkowego, jednostronnego pomiaru odległości (Single-Sided Two-Way Ranging, SS-TWR). Czas przelotu oznaczono jako T_{tof} , a czas przetwarzania jako T_{proc}	65
3.2	Schemat komunikacji w metodzie DS-TWR. Czas przelotu oznaczono jako T_{tof} . Czas przetwarzania po stronie urządzenia kotwicy jako T_{proc_K} , a po stronie lokalizatora jako T_{proc_L}	66
3.3	Schemat komunikacji w warunkach bez przeszkód (Line-of-Sight, LOS), a także z odbiciami od obiektów (Non-Line-of-Sight, NLOS). .	69
3.4	Wizualizacja operacji trilateracji na płaszczyźnie w scenariuszach z oparciem pozycjonowania na pomiarach z pojedynczej kotwicy (scenariusz A), z dwóch kotwic (scenariusz B) oraz z trzech kotwic (scenariusz C). Wyznaczone położenie lokalizatora oznaczono kolorem czerwonym, kolejne kotwice jako A1, A2 i A3, a wyznaczone dystanse do nich jako D1, D2 i D3.	70
3.5	Schemat ideowy procesu lokalizacji z użyciem analizy kąta nadejścia sygnału radiowego.	76
3.6	Zrzut ekranu z narzędzia Blender 4.2 z wizualizacją uzyskanego modelu cyfrowego.	81
3.7	Zrzut ekranu z narzędzia spaceru interaktywnego firmy Matterport z wizualizacją wykorzystania funkcji pomiaru odległości.	82
3.8	Zbudowane urządzenie lokalizacyjne w technologii UWB: płytką UWB (ramka czerwona), koordynator Matter (ramka zielona), zasilanie (ramka pomarańczowa).	83
3.9	Zbudowane urządzenie lokalizacyjne w technologii BLE: płytką BLE (ramka niebieska), koordynator Matter (ramka zielona), zasilanie (ramka pomarańczowa).	85
3.10	Zdjęcie wszystkich skonstruowanych prototypowo urządzeń systemu pozycjonowania.	93

3.11 Zbliżenie na skonstruowane urządzenie kotwicy (punktu referencyjnego). Uchwyt mocujący i bank energii oznaczone pomarańczową ramką, kotwica systemu UWB czerwoną ramką, a kotwica systemu BLE niebieską ramką.	94
3.12 Diagram przedstawiający architekturę proponowanego systemu lokalizacyjnego. Rozwiązanie źródłowe prezentowane jest po lewej, a rozszerzona postać po prawej. Podsystem UWB oznaczono w odcieniach czerwonego, BLE w odcieniach niebieskiego, a koordynator protokołu Matter oznaczono odcieniami zielonego. Oryginalny odbiornik danych (serwer) zastąpiono usługą chmurową Azure Event Grid (szary obszar).	95
3.13 Fizyczne środowisko, w którym przeprowadzono eksperyment.	96
3.14 Wizualizacja rzutu pomieszczenia wraz z punktami pomiarowymi, wygenerowana przy użyciu wirtualnej wycieczki Matterport.	96
3.15 Urządzenie śledzące umieszczone w punkcie pomiarowym P1 w konfiguracji na podłodze (po lewej) oraz na wysokości 150 cm (po prawej).	97
3.16 Urządzenie kotwicy zamontowane na statywie niskim (po lewej) oraz regulowanym wysokim (po prawej).	98
3.17 Wizualizacja obudowania lokalizatora sferą (z lewej: 100% nieprzezroczystości, środek: 50%, z prawej: 0%).	98
3.18 Wizualizacja eksperymentalnego środowiska w programie Blender: lokalizatory na podłodze oznaczone na zielono, na statywie na niebiesko, a kotwice na czerwono.	101
3.19 Średni błąd względny dla kolejnych przedziałów rzeczywistej odległości lokalizatora UWB od kotwicy dla dwóch poziomów jego wysokości: podłogi oraz trójnogu. Na dole słupka przedstawiona liczność pomiarów w danym przedziale dystansu.	111
3.20 Średni błąd względny dla kolejnych przedziałów rzeczywistej odległości lokalizatora BLE od kotwicy dla dwóch poziomów jego wysokości: podłogi oraz trójnogu. Na dole słupka przedstawiona liczność pomiarów w danym przedziale dystansu.	111

4.1	Schemat ideowy struktury warstwowej sieci neuronowej o trzech warstwach neuronów ukrytych.	119
4.2	Schemat ideowy procesu konwolucji obrazu.	120
4.3	Zobrazowanie zdjęcia sferycznego i jego reprezentacja w postaci panoramy równokątnej.	127
4.4	Wizualizacja wpływu wybranego miejsca rozcięcia sfery obrazu sferycznego przy rzutowaniu do reprezentacji dwuwymiarowej.	128
4.5	Wizualizacja prezentacji oznaczonego sferycznie zbioru danych przed (lewo) i po (pravo) wprowadzeniu poprawek.	133
4.6	Porównanie liczności obiektów i sposobu ich podziału w zbiorach źródłowym oraz przetworzonym.	134
4.7	Porównanie uzyskiwanych wartości mAP@50 dla modeli YOLOv5 i YOLOv8 w zależności od liczby epok trenowania.	138
4.8	Porównanie skuteczności najlepszych modeli architektury YOLO generacji piątej i ósmej oraz innych wybranych modeli detekcji obiektów w zadaniu przetwarzania obrazów typu ERP.	140
4.9	Przykład procesu przetwarzania obrazu panoramicznego z proponowanego zbioru <i>EquiB&B</i>	146
4.10	Przykładowe obrazy (ze zbioru WiMR [322]) przetworzone przy użyciu kolejnych metod doboru południka rzutowania.	147
4.11	Przykładowe obrazy ze zbioru <i>PANDORA REAL</i>	151
4.12	Przykładowe obrazy ze zbioru <i>PANDORA SYNT</i>	151
4.13	Przykładowe obrazy ze zbioru PanoSilesia RESORT.	152
4.14	Przykładowe obrazy ze zbioru PanoSilesia HOUSE.	152
4.15	Przykładowe obrazy ze zbioru PanoSilesia APARTMENT.	153
4.16	Przykładowe obrazy ze zbioru OFFICE.	153
4.17	Przykładowe obrazy ze zbioru EquiB&B.	154
4.18	Przykłady błędnie wygenerowanych obrazów odrzuconych z procesu tworzenia zbioru <i>SPHEREA</i>	154
4.19	Wykresy rozkładów przestrzennych etykiet obiektów na zdjęciach ERP dla kolejnych analizowanych zbiorów danych.	157

4.20	Przykładowe zdjęcia panoramiczne z badań prototypu wizyjnego z wyznaczonymi dla nich mapami głębi. Paleta barw określa odległość do obiektu, kolory ciepłe podkreślają krótki dystans, kolory zimne duży dystans.	160
4.21	Zrzut ekranu narzędzia <i>Matterport Viewer</i> z naniesionymi wymiarami pomieszczenia oraz zbliżeniem na etykiety długości.	162
4.22	Przykładowe panoramy ze zbioru punktów pomiarowych prototypu systemu wizyjnego.	163
4.23	Zdjęcie skanera przestrzeni Matterport Pro 3 (po lewej) oraz jego umieszczenie na dedykowanym trójnogu z naniesionymi na zdjęcie pomiarami wysokości (po prawej).	164
4.24	Przykładowe wizualizacje przetworzonych z użyciem detektora obiektów zdjęć panoramicznych oraz odpowiadających im map głębi. . . .	165
4.25	Zrzut ekranu z narzędzia <i>Matterport Viewer</i> przedstawiający proces pomiaru oczekiwanej odległości od skanera przestrzeni do wykrytego obiektu.	165
4.26	Porównanie metryk błędu dla różnych wartości parametru <code>max_depth</code> detektora obiektów <i>Depth Anything V2</i>	166
5.1	Schemat logiczny kroków realizacji proponowanej hybrydowej metody lokalizacji.	172
5.2	Schemat logiczny architektury proponowanego systemu fuzyjnego. Pozycjonowanie na poziomie mieszkania z użyciem technologii BLE, na poziomie pomieszczenia w oparciu o UWB oraz rozpoznawanie obiektów z użyciem kamery.	173
5.3	Zrzuty ekranu opracowanego modelu obudowy urządzenia kotwicy lub lokalizatora radiowej części systemu hybrydowego.	174
5.4	Zrzuty ekranu opracowanego modelu obudowy mobilnego urządzenia pomiarowego (ze słupkiem do montażu kamery sferycznej). . . .	175
5.5	Zdjęcia zrealizowanych w technologii druku 3D obudów urządzeń systemu radiowego: statycznych (po lewej) i mobilnych (po prawej). .	176

- 5.6 Zdjęcia mobilnych urządzeń pomiarowych. Na lewym zdjęciu, w zbliżeniu, od góry: kamera sferyczna, lokalizator radiowy, trójnóg montażowy. Po prawej zdjęcie wszystkich trzech przygotowanych urządzeń pomiarowych. 176
- 5.7 Po lewej: zdjęcie sposobu montażu punktów referencyjnych umieszczanych na dużych wysokościach, tj. z wykorzystaniem statywów, które pozwoliły na stabilne wyniesienie kotwic radiowych pod sufit pomieszczenia. Po prawej: zdjęcie faktycznego użycia takiego statywu. 177
- 5.8 Zdjęcie zbiorcze wszystkich 35 skonstruowanych urządzeń, budujących radiową część hybrydowego systemu pozycjonowania. Na zdjęciu, kolejno: trzy mobilne urządzenia pomiarowe (kamera, lokalizator, trójnóg) po lewej stronie, cztery zmontowane punkty referencyjne dużej wysokości (kotwice wyniesione pod sufit) w tle, 28 pozostałych urządzeń (20 kotwic i 8 lokalizatorów) z przodu. 178
- 5.9 Zdjęcie skonstruowanych urządzeń radiowych podczas pracy i z podłączonym zasilaniem. 179
- 5.10 Wizualizacja modelu trójwymiarowego całego środowiska badawczego z użyciem narzędzia *Matterport Viewer*. 180
- 5.11 Rzut planu mieszkania (widok z góry) z narzędzia *Matterport Viewer*. 181
- 5.12 Model mieszkania z naniesionymi oznaczeniami obiektów. 181
- 5.13 Wykres kołowy podziału oznaczonych komponentów środowiska domowego na obiekty pasywne i aktywne. 182
- 5.14 Wykres słupkowy liczby obiektów obecnych w konkretnych pokojach całego środowiska testowego. 184
- 5.15 Wizualizacja problemu ustalenia położenia na podstawie liczby odbieranych sygnałów radiowych w scenariuszu pomieszczenia-enklawy. 186
- 5.16 Scenariusz stałego radiowego monitorowania położenia obiektów środowiska domowego w pełni wyposażonych w lokalizatory UWB. . . 189
- 5.17 Wizualizacja układów współrzędnych i oznaczenia analogicznego punktu na sferze (po lewej) i na panoramie ERP (po prawej). 194
- 5.18 Schemat realizacji wizyjnego rozpoznawania i lokalizowania komponentów modelu cyfrowego bliźniaka. 195

- 5.19 Wizualizacja modelu trójwymiarowego mieszkania w programie Blender, widok z naniesieniem punktów charakterystycznych. Sfery czerwone oznaczają kotwice systemu radiowego. Sfery niebieskie pokazują położenie punktu charakterystycznego w płaszczyźnie podłogi dla każdego ze stanowisk pomiarowych. Sfery zielone reprezentują położenie mobilnych lokalizatorów radiowych (punktów charakterystycznych na poziomie lokalizatorów umieszczonych na trójnogach). Sfery fioletowe to lokalizatory umiejscowione statycznie. 199
- 5.20 Zdjęcie sferyczne w formacie ERP wykonane przez urządzenie pomiarowe C2 na stanowisku pomiarowym P2 z zaznaczonymi towarzyszącymi mu urządzeniami pomiarowymi C1 oraz C3. 201
- 5.21 Przykładowe obrazy ze zbioru zdjęć wykonanych w ramach badań systemu hybrydowego. 202
- 5.22 Przykładowe zdjęcia panoramiczne z badań systemu hybrydowego wraz z zaznaczonymi na nich detekcjami obiektów. 204
- 5.23 Przykładowe zdjęcia panoramiczne z badań systemu hybrydowego wraz z wyznaczonymi dla nich mapami głębi. Paleta barw określa odległość do obiektu, kolory ciepłe podkreślają krótki dystans, kolory zimne daleki dystans. 204
- 5.24 Zrzut ekranu opracowanego narzędzia deweloperskiego do wizualizacji detekcji obiektów na panoramach ERP wraz z przeliczeniem położenia do przestrzeni trójwymiarowej. 205
- 5.25 Wykres kołowy poprawności przypisania pomieszczenia dla próbek pomiarowych systemu pozycjonowania ogólnego w technologii BLE. 207
- 5.26 Histogramy błędów radiowego pozycjonowania precyzyjnego w ramach końcowych badań systemu fuzyjnego. 209
- 5.27 Wykres słupkowy mediany błędu pozycjonowania precyzyjnego dla każdego z pokoi z osobna. 210
- 5.28 Diagram ogólnej poprawności detekcji obiektów. 212
- 5.29 Wykres słupkowy liczby detekcji obiektów w podziale na klasy obiektów dla 15 najpopularniejszych rodzajów obiektów. 213

5.30 Wykres słupkowy poprawnych i niepoprawnych detekcji obiektów w podziale na klasy obiektów dla 15 najpopularniejszych rodzajów obiektów.	214
5.31 Wykres słupkowy liczby detekcji obiektów w podziale na pokoje (góra słupka) wraz z procentowym udziałem poszczególnych pokoi w łącznej liczbie detekcji (dół słupka).	215
5.32 Wykresy słupkowe poprawnych i niepoprawnych detekcji obiektów w podziale na pokoje.	216
5.33 Histogram rozkładu wyników estymacji głębokości z badań systemu hybrydowego.	217
5.34 Wykres końcowej wykrywalności komponentów środowiska domowego przy zastosowaniu metody hybrydowej.	219
5.35 Wykres końcowej wykrywalności komponentów środowiska domowego z uwzględnieniem poprawności detekcji.	220
5.36 Wykres liczności detekcji na pojedynczy obiekt w metodzie hybrydowej.	221
5.37 Histogramy błędów pozycjonowania wizyjnego uzyskanego w hybrydowej, wizyjno-radiowej metodzie pozycjonowania obiektów.	222
5.38 Wykres słupkowy mediany błędu pozycjonowania w podziale na pokoje dla metody hybrydowej. Na dole słupka przedstawiono licznosc próbek w danym pokoju.	223
5.39 Wizualizacja położenia wszystkich wykrytych w mieszkaniu obiektów (oznaczonych przy pomocy sfer). Współrzędne obiektów określone na podstawie systemu hybrydowego z punktem położenia kamery wyznaczonej automatycznie przez radiowy system pozycjonowania.	224

Spis tabel

2.1	Zestawienie zastosowanych w procesie przeglądu literatury kryteriów inkluzywnych (IC) i wykluczających (EC) wraz z ich opisami. .	21
2.2	Zestawienie artykułów według rodzaju danych wejściowych. Liczby w nawiasach oznaczają liczbę artykułów uwzględnionych w głównym przeglądzie (pierwsza liczba) oraz w jego rozszerzeniu (druga liczba).	26
2.3	Zestawienie artykułów według rodzaju danych wejściowych i kategorii przetwarzania, kolejno: segmentacja (S), segmentacja z klasyfikacją uproszczoną (S + KU), segmentacja z klasyfikacją precyzyjną (S + KP) oraz klasyfikacja precyzyjna (KP).	29
2.4	Zestawienie artykułów według zadania i typu danych wejściowych. .	33
2.5	Macierz analizy SWOT dla produktu SIM-ON	45
2.6	Macierz oceny zależności pomiędzy czynnikami analizy SWOT. . . .	46
2.7	Zestawienie analizowanych w ramach rozeznania konkurencji rynkowej funkcji systemów zarządzania nieruchomościami.	48
2.8	Zestawienie funkcji obsługiwanych przez analizowane systemy do zarządzania nieruchomościami. Fakt wsparcia konkretnej funkcji oznaczony został symbolem +.	50
2.9	Zestawienie funkcji obsługiwanych przez analizowane systemy lokalizacji zasobów. Fakt wsparcia konkretnej funkcji oznaczony został symbolem +.	52
2.10	Spis klas obiektów w proponowanym zbiorze CODE55.	54
2.11	Spis klas pomieszczeń w proponowanym zbiorze CODE55.	55
2.12	Najbardziej i najmniej popularne obiekty według liczby wystąpień w odpowiedziach uczestników.	57

2.13	Łączna liczba odpowiedzi o obiektach w podziale na pomieszczenia.	57
2.14	Rozkład liczby wskazań wybranych obiektów w poszczególnych pomieszczeniach.	58
2.15	Dziesięć najczęściej wskazywanych obiektów w poszczególnych pomieszczeniach.	59
2.16	Porównanie pokrycia oczekiwanych przez użytkowników etykiet obiektów z etykietami dostępnymi w wybranych zbiorach danych.	59
3.1	Wartości kalibracji i odchylenie standardowe dla zmontowanych lokalizatorów technologii BLE.	86
3.2	Pozycja kotwic w przestrzeni trójwymiarowej stanowiska pomiarowego.	99
3.3	Wysokości kotwic ponad poziomem podłogi (w centymetrach), wyznaczone przy użyciu metody manualnego pomiaru dalmierzem (Laser), interaktywnego spaceru trójwymiarowego (Spacer) oraz odczytane wprost z modelu (Model).	100
3.4	Konfiguracje i opis przeprowadzonych eksperymentów. Wykorzystana technologia radiowa (w kolumnie Tech.): UWB, BLE albo U+B (obie). Typ eksperymentu (w kolumnie Typ): statyczny (S) albo dynamiczny (D). Wysokość umieszczenia lokalizatora podana w centymetrach ponad poziomem podłogi.	103
3.5	Porównanie błędów estymacji dystansu dla kombinacji źródła danych referencyjnych, wysokości umieszczenia lokalizatora, wykorzystanego protokołu komunikacyjnego oraz technologii radiowej.	105
3.6	Porównanie błędów estymaty dystansu pomiędzy wszystkimi czterema lokalizatorami UWB (górna część) i BLE (dolna część) w zależności od źródła danych referencyjnych oraz wykorzystanego protokołu komunikacyjnego.	108
4.1	Wpływ rozmiaru wsadu i rozmiaru obrazu na uzyskaną wartość mAP@50 dla rodziny modeli YOLOv5. Najlepsze wyniki wyróżniono pogrubieniem, błędy braku pamięci oznaczono kursywą.	135

4.2	Wpływ rozmiaru wsadu i rozmiaru obrazu na uzyskaną wartość mAP@50 dla rodziny modeli YOLOv8. Najlepsze wyniki wyróżniono pogrubieniem, błędy braku pamięci oznaczono kursywą.	136
4.3	Porównanie skuteczności sieci neuronowych rodzin YOLOv5 i YOLOv8 w zadaniu detekcji obiektów na zdjęciach panoramicznych w formie ERP.	138
4.4	Wyniki miary mAP50 dla zbioru danych <i>WiMR</i> dla kolejnych konfiguracji eksperymentalnych. Najlepsze wyniki w pionie pogrubiono, a najlepsze w poziomie podkreślono.	148
4.5	Wyniki miary mAP50 dla zbioru danych <i>EquiBEB</i> dla kolejnych konfiguracji eksperymentalnych. Najlepsze wyniki w pionie pogrubiono, a najlepsze w poziomie podkreślono.	149
4.6	Statystyki ogólne opracowanych zbiorów syntetycznych oraz zbioru referencyjnego.	155
4.7	Porównanie skuteczności detekcji (mAP@50) w różnych konfiguracjach zbiorów treningowych i testowych dla modelu YOLOv5 <i>Extra-Large</i>	157
5.1	Zestawienie liczby obiektów określonego rodzaju (zgodnych z etykietami CODE55) w całym środowisku testowym.	183
5.2	Zestawienie współrzędnych punktów charakterystycznych dla przykładowego pomieszczenia łazienki.	200
6.1	Publikacje z punktacją ministerialną	238
6.2	Publikacje popularnonaukowe (bez punktacji ministerialnej).	240
6.3	Zestawienie skrótów i objaśnień wykorzystywanych w rozprawie.	247

Bibliografia

- [1] H. Mackenzie, A. Dewey, A. Drahota, S. Kilburn, P. Kalra, C. Fogg, and D. Zachariah, “Systematic reviews: what they are, why they are important, and how to get involved,” *Journal of Clinical and Preventive Cardiology*, vol. 1, no. 10, 2012.
- [2] D. Gough, S. Oliver, and J. Thomas, *An introduction to systematic reviews*. SAGE, 2nd edition ed., 2017. OCLC: ocn963634161.
- [3] M. J. Grant and A. Booth, “A typology of reviews: an analysis of 14 review types and associated methodologies: A typology of reviews,” *Maria J. Grant & Andrew Booth*, *Health Information & Libraries Journal*, vol. 26, no. 2, pp. 91–108, 2009.
- [4] Y. Xiao and M. Watson, “Guidance on conducting a systematic literature review,” *Journal of Planning Education and Research*, vol. 39, no. 1, pp. 93–112, 2019.
- [5] R. Bormann, F. Jordan, W. Li, J. Hampp, and M. Hagele, “Room segmentation: Survey, implementation, and analysis,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1019–1026, IEEE, 2016.
- [6] L. Gimenez, J.-L. Hippolyte, S. Robert, F. Suard, and K. Zreik, “Review: reconstruction of 3d building information models from 2d scanned plans,” *Journal of Building Engineering*, vol. 2, pp. 24–35, 2015.
- [7] Z. Kang, J. Yang, Z. Yang, and S. Cheng, “A review of techniques for 3d re-

- construction of indoor environments,” *ISPRS International Journal of Geo-Information*, vol. 9, no. 5, 2020. Publisher: MDPI AG.
- [8] F. Santos, A. Moreira, and P. Costa, “Towards extraction of topological maps from 2d and 3d occupancy grids,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8154 LNAI, pp. 307–318, 2013. ISBN: 9783642406683.
- [9] C. Mura, O. Mattausch, A. J. Villanueva, E. Gobbetti, and R. Pajarola, “Robust reconstruction of interior building structures with multiple rooms under clutter and occlusions,” in *2013 International Conference on Computer-Aided Design and Computer Graphics*, pp. 52–59, 2013.
- [10] D. Borrmann, A. Nüchter, M. Đakulović, I. Maurović, I. Petrović, D. Osmanović, and J. Velagić, “A mobile robot based system for fully automated thermal 3d mapping,” *Advanced Engineering Informatics*, vol. 28, no. 4, pp. 425–440, 2014.
- [11] I. Maurović, M. Đakulović, and I. Petrović, “Autonomous exploration of large unknown indoor environments for dense 3d model building,” *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 10188–10193, 2014.
- [12] C. Mura, O. Mattausch, A. J. Villanueva, E. Gobbetti, and R. Pajarola, “Automatic room detection and reconstruction in cluttered indoor environments with complex room layouts,” *Comput. Graph.*, vol. 44, pp. 20–32, 2014.
- [13] S. Ochmann, R. Vock, R. Wessel, M. Tamke, and R. Klein, “Automatic generation of structural building descriptions from 3d point cloud scans,” in *GRAPP 2014 - Proceedings of the 9th International Conference on Computer Graphics Theory and Applications*, pp. 120–127, SciTePress, 2014.
- [14] S. Ochmann, R. Vock, R. Wessel, and R. Klein, “Towards the extraction of hierarchical building descriptions from 3d indoor scans,” in *Eurographics Workshop on 3D Object Retrieval, EG 3DOR*, Eurographics Association, 2014. ISSN: 19970463.

- [15] A. Swadzba and S. Wachsmuth, “A detailed analysis of a new 3d spatial feature vector for indoor scene classification,” *Robotics and Autonomous Systems*, vol. 62, no. 5, pp. 646–662, 2014. Publisher: Elsevier.
- [16] H. Macher, T. Landes, and P. Grussenmeyer, “Point clouds segmentation as base for as-built BIM creation,” in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (C. H.-M. Yen A.Y.-N., Weng K.-H., ed.), vol. 2, pp. 191–197, Copernicus GmbH, 2015. ISSN: 21949042 Issue: 5W3.
- [17] E. Turner, P. Cheng, and A. Zakhor, “Fast, automated, scalable generation of textured 3d models of indoor environments,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 9, no. 3, pp. 409–421, 2015. Publisher: Institute of Electrical and Electronics Engineers Inc.
- [18] E. Turner and A. Zakhor, “Automatic indoor 3d surface reconstruction with segmented building and object elements,” in *Proceedings - 2015 International Conference on 3D Vision, 3DV 2015* (T. C. Brown M., Kosecka J., ed.), pp. 362–370, Institute of Electrical and Electronics Engineers Inc., 2015.
- [19] S. Ikehata, H. Yang, and Y. Furukawa, “Structured indoor modeling,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1323–1331, 2015. ISSN: 2380-7504.
- [20] I. Armeni, O. Sener, A. Zamir, H. Jiang, I. Brilakis, M. Fischer, and S. Savarese, “3d semantic parsing of large-scale indoor spaces,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 1534–1543, IEEE Computer Society, 2016. ISSN: 10636919.
- [21] G. Manfredi, S. Devin, M. Devy, and D. Sidobre, “Autonomous apartment exploration, modelling and segmentation for service robotics,” *IFAC-PapersOnLine*, vol. 49, no. 15, pp. 120–125, 2016.
- [22] J. Martínez-Gómez, V. Morell, M. Cazorla, and I. García-Varea, “Semantic localization in the PCL library,” *Robotics and Autonomous Systems*, vol. 75, pp. 641–648, 2016.

- [23] S. Ochmann, R. Vock, R. Wessel, and R. Klein, "Automatic reconstruction of parametric building models from indoor point clouds," *Computers and Graphics (Pergamon)*, vol. 54, pp. 94–103, 2016. Publisher: Elsevier Ltd.
- [24] Q. Liu, R. Li, H. Hu, and D. Gu, "Using semantic maps for room recognition to aid visually impaired people," in *2016 22nd International Conference on Automation and Computing (ICAC)*, pp. 89–94, 2016.
- [25] J. Jung, C. Stachniss, and C. Kim, "Automatic room segmentation of 3d laser data using morphological processing," *ISPRS Int. J. Geo Inf.*, vol. 6, no. 7, p. 206, 2017.
- [26] H. Macher, T. Landes, and P. Grussenmeyer, "From point clouds to building information models: 3d semi-automatic reconstruction of indoors of existing buildings," *Applied Sciences (Switzerland)*, vol. 7, no. 10, 2017. Publisher: MDPI AG.
- [27] S. Murali, P. Speciale, M. Oswald, and M. Pollefeys, "Indoor scan2bim: Building information models of house interiors," in *IEEE International Conference on Intelligent Robots and Systems*, vol. 2017-September, pp. 6126–6133, Institute of Electrical and Electronics Engineers Inc., 2017. ISSN: 21530858.
- [28] S. Nikoohemat, M. Peter, S. Oude Elberink, and G. Vosselman, "EXPLOITING INDOOR MOBILE LASER SCANNER TRAJECTORIES for SEMANTIC INTERPRETATION of POINT CLOUDS," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 355–362, Copernicus GmbH, 2017. ISSN: 21949042 Issue: 2W4.
- [29] R. Wang, L. Xie, and D. Chen, "Modeling indoor spaces using decomposition and reconstruction of structural elements," *Photogrammetric Engineering and Remote Sensing*, vol. 83, no. 12, pp. 827–841, 2017. Publisher: American Society for Photogrammetry and Remote Sensing.
- [30] L. Xie and R. Wang, "Automatic indoor building reconstruction from mobile laser scanning data," in *International Archives of the Photogrammetry*,

- Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 42, pp. 417–422, International Society for Photogrammetry and Remote Sensing, 2017. ISSN: 16821750 Issue: 2W7.
- [31] R. Ambrus, S. Claici, and A. Wendt, “Automatic room segmentation from unstructured 3-d data of indoor environments,” *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 749–756, 2017.
- [32] D. Bobkov, M. Kiechle, S. Hilsenbeck, and E. Steinbach, “Room segmentation in 3d point clouds using anisotropic potential fields,” in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 727–732, 2017. ISSN: 1945-788X.
- [33] M. Brucker, M. Durner, R. Ambrus, Z. Marton, A. Wendt, P. Jensfelt, K. Arras, and R. Triebel, “Semantic labeling of indoor environments from 3d RGB maps,” in *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 1871–1878, Institute of Electrical and Electronics Engineers Inc., 2018. ISSN: 10504729.
- [34] A. Elseicy, S. Nikoohemat, M. Peter, and S. Elberink, “Space subdivision of indoor mobile laser scanning data based on the scanner trajectory,” *Remote Sensing*, vol. 10, no. 11, 2018. Publisher: MDPI AG.
- [35] J. Jung, C. Stachniss, S. Ju, and J. Heo, “Automated 3d volumetric reconstruction of multiple-room building interiors for as-built BIM,” *Advanced Engineering Informatics*, vol. 38, pp. 811–825, 2018. Publisher: Elsevier Ltd.
- [36] L. Li, F. Su, F. Yang, H. Zhu, D. Li, X. Zuo, F. Li, Y. Liu, and S. Ying, “Reconstruction of three-dimensional (3d) indoor interiors with multiple stories via comprehensive segmentation,” *Remote Sensing*, vol. 10, no. 8, 2018. Publisher: MDPI AG.
- [37] L. Magri and A. Fusiello, “Reconstruction of interior walls from point cloud data with min-hashed j-linkage,” in *Proceedings - 2018 International Conference on 3D Vision, 3DV 2018*, pp. 131–139, Institute of Electrical and Electronics Engineers Inc., 2018.

- [38] S. Nikoohemat, M. Peter, S. Elberink, and G. Vosselman, "Semantic interpretation of mobile laser scanner point clouds in indoor scenes using trajectories," *Remote Sensing*, vol. 10, no. 11, 2018. Publisher: MDPI AG.
- [39] Y. Zheng, M. Peter, R. Zhong, S. Elberink, and Q. Zhou, "Space subdivision in indoor mobile laser scanning point clouds based on scanline analysis," *Sensors (Switzerland)*, vol. 18, no. 6, 2018. Publisher: MDPI AG.
- [40] O. Sharma, J. Pandey, H. Akhtar, and G. Rathee, "Navigation in AR based on digital replicas," *The Visual Computer*, vol. 34, no. 6, pp. 925–936, 2018.
- [41] Y. Cui, Q. Li, and Z. Dong, "Structural 3d reconstruction of indoor space for 5g signal simulation with mobile laser scanning point clouds," *Remote Sensing*, vol. 11, no. 19, 2019. Publisher: MDPI AG.
- [42] Y. Cui, Q. Li, B. Yang, W. Xiao, C. Chen, and Z. Dong, "Automatic 3-d reconstruction of indoor environment with mobile laser scanning point clouds," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 3117–3130, 2019. Publisher: Institute of Electrical and Electronics Engineers.
- [43] M. Koeva, S. Nikoohemat, S. Elberink, J. Morales, C. Lemmen, and J. Zevenbergen, "Towards 3d indoor cadastre based on change detection from point clouds," *Remote Sensing*, vol. 11, no. 17, 2019. Publisher: MDPI AG.
- [44] E. Maset, L. Magri, and A. Fusiello, "Improving automatic reconstruction of interior walls from point cloud data," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* (Y. M. Vosselman G., Oude Elberink S.J., ed.), vol. 42, pp. 849–855, International Society for Photogrammetry and Remote Sensing, 2019. ISSN: 16821750 Issue: 2/W13.
- [45] S. Nikoohemat, A. Diakit , S. Zlatanova, and G. Vosselman, "INDOOR 3d MODELING and FLEXIBLE SPACE SUBDIVISION from POINT CLOUDS," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Y. M. Vosselman G., Oude Elberink S.J., ed.), vol. 4, pp. 285–292, Copernicus GmbH, 2019. ISSN: 21949042 Issue: 2/W5.

- [46] S. Ochmann, R. Vock, and R. Klein, “Automatic reconstruction of fully volumetric 3d building models from oriented point clouds,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 151, pp. 251–262, 2019. Publisher: Elsevier B.V.
- [47] M. Previtali, L. Barazzetti, and F. Roncoroni, “AUTOMATED DETECTION and LAYOUT REGULARIZATION of SIMILAR FEATURES in INDOOR POINT CLOUD,” in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (R. F. Fassi F., Cardaci A., ed.), vol. 42, pp. 631–638, Copernicus GmbH, 2019. ISSN: 21949042 Issue: 2/W9.
- [48] W. Shi, W. Ahmed, N. Li, W. Fan, H. Xiang, and M. Wang, “Semantic geometric modelling of unstructured indoor point cloud,” *ISPRS International Journal of Geo-Information*, vol. 8, no. 1, 2019. Publisher: MDPI AG.
- [49] S. Tang, Y. Zhang, Y. Li, Z. Yuan, Y. Wang, X. Zhang, X. Li, Y. Zhang, R. Guo, and W. Wang, “Fast and automatic reconstruction of semantically rich 3d indoor maps from low-quality RGB-d sequences,” *Sensors (Switzerland)*, vol. 19, no. 3, 2019. Publisher: MDPI AG.
- [50] P. Wang, J. Cheng, and W. Feng, “An approach for construct semantic map with scene classification and object semantic segmentation,” in *2018 IEEE International Conference on Real-Time Computing and Robotics, RCAR 2018*, pp. 270–275, Institute of Electrical and Electronics Engineers Inc., 2019.
- [51] F. Yang, L. Li, F. Su, D. Li, H. Zhu, S. Ying, X. Zuo, and L. Tang, “Semantic decomposition and recognition of indoor spaces with structural constraints for 3d indoor modelling,” *Automation in Construction*, vol. 106, 2019. Publisher: Elsevier B.V.
- [52] F. Yang, G. Zhou, F. Su, X. Zuo, L. Tang, Y. Liang, H. Zhu, and L. Li, “Automatic indoor reconstruction from point clouds in multi-room environments with curved walls,” *Sensors (Switzerland)*, vol. 19, no. 17, 2019. Publisher: MDPI AG.

- [53] Z. He, J. Hou, and S. Schwertfeger, “Furniture free mapping using 3d lidars,” in *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 583–589, 2019.
- [54] S. Nikoohemat, A. Diakité, S. Zlatanova, and G. Vosselman, “Indoor 3d reconstruction from point clouds for optimal routing in complex buildings to support disaster management,” *Automation in Construction*, vol. 113, 2020. Publisher: Elsevier B.V.
- [55] R. Otero, E. Frías, S. Lagüela, and P. Arias, “Automatic gbXML modeling from LiDAR data for energy studies,” *Remote Sensing*, vol. 12, no. 17, 2020. Publisher: MDPI AG.
- [56] I. Rusli, B. R. Trilaksono, and W. Adiprawita, “RoomSLAM: Simultaneous localization and mapping with objects and indoor layout structure,” *IEEE Access*, vol. 8, pp. 196992–197004, 2020.
- [57] M. Ryu, S. Oh, M. Kim, H. Cho, C. Son, and T. Kim, “Algorithm for generating 3d geometric representation based on indoor point cloud data,” *Applied Sciences (Switzerland)*, vol. 10, no. 22, pp. 1–13, 2020. Publisher: MDPI AG.
- [58] A. Phalak, V. Badrinarayanan, and A. Rabinovich, “Scan2plan: Efficient floorplan generation from 3d scans of indoor scenes,” *arXiv:2003.07356 [cs]*, 2020.
- [59] E. Frías, J. Balado, L. Díaz-Vilariño, and H. Lorenzo, “POINT CLOUD ROOM SEGMENTATION BASED ON INDOOR SPACES AND 3d MATHEMATICAL MORPHOLOGY,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIV-4/W1-2020, pp. 49–55, 2020.
- [60] M. Ai, Z. Li, and J. Shan, “Topologically consistent reconstruction for complex indoor structures from point clouds,” *Remote Sensing*, vol. 13, no. 19, 2021. Publisher: MDPI.

- [61] Y. Cai and L. Fan, “An efficient approach to automatic construction of 3d watertight geometry of buildings using point clouds,” *Remote Sensing*, vol. 13, no. 10, 2021. Publisher: MDPI AG.
- [62] H. Fang, F. Lafarge, C. Pan, and H. Huang, “Floorplan generation from 3d point clouds: A space partitioning approach,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, pp. 44–55, 2021. Publisher: Elsevier B.V.
- [63] Z. He, H. Sun, J. Hou, Y. Ha, and S. Schwertfeger, “Hierarchical topometric representation of 3d robotic maps,” *Autonomous Robots*, vol. 45, no. 5, pp. 755–771, 2021. Publisher: Springer.
- [64] P. Hübner, M. Weinmann, S. Wursthorn, and S. Hinz, “Automatic voxel-based 3d indoor reconstruction and room partitioning from triangle meshes,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 254–278, 2021.
- [65] Y. Wang, M. Ramezani, M. Mattamala, and M. Fallon, “Scalable and elastic LiDAR reconstruction in complex environments through spatial analysis,” in *2021 10th European Conference on Mobile Robots, ECMR 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2021.
- [66] F. Yang, M. Che, X. Zuo, L. Li, J. Zhang, and C. Zhang, “Volumetric representation and sphere packing of indoor space for three-dimensional room segmentation,” *ISPRS Int. J. Geo Inf.*, vol. 10, no. 11, p. 739, 2021.
- [67] J. Yang, Z. Kang, L. Zeng, P. Hope Akwensi, and M. Sester, “Semantics-guided reconstruction of indoor navigation elements from 3d colorized points,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173, pp. 238–261, 2021. Publisher: Elsevier B.V.
- [68] R. Cai, H. Li, J. Xie, and X. Jin, “Accurate floorplan reconstruction using geometric priors,” *Computers and Graphics (Pergamon)*, vol. 102, pp. 360–369, 2022. Publisher: Elsevier Ltd.

- [69] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Nießner, M. Savva, S. Song, A. Zeng, and Y. Zhang, “Matterport3d: Learning from RGB-d data in indoor environments,” in *Proceedings of the 2017 International Conference on 3D Vision (3DV)*, pp. 667–676, IEEE, 2017. type: article.
- [70] F. Blöchliger, M. Fehr, M. Dymczyk, T. Schneider, and R. Siegwart, “Topomap: Topological mapping and navigation based on visual SLAM maps,” *arXiv:1709.05533 [cs]*, 2018.
- [71] J. Chen, C. Liu, J. Wu, and Y. Furukawa, “Floor-SP: Inverse CAD for floorplans by sequential room-wise shortest path,” in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2661–2670, IEEE, 2019.
- [72] J. L. Carrera V., Z. Zhao, T. Braun, Z. Li, and A. Neto, “A real-time robust indoor tracking system in smartphones,” *Computer Communications*, vol. 117, pp. 104–115, 2018.
- [73] J. Matez-Bandera, J. Monroy, and J. Gonzalez-Jimenez, “Efficient semantic place categorization by a robot through active line-of-sight selection,” *Knowledge-Based Systems*, vol. 240, 2022. Publisher: Elsevier B.V.
- [74] N. Liu, B. Lin, L. Yuan, G. Lv, Z. Yu, and L. Zhou, “An interactive indoor 3d reconstruction method based on conformal geometry algebra,” *Advances in Applied Clifford Algebras*, vol. 28, no. 4, 2018. Publisher: Birkhauser Verlag AG.
- [75] M. Weinmann, S. Wursthorn, M. Weinmann, and P. Hübner, “Efficient 3d mapping and modelling of indoor scenes with the microsoft HoloLens: A survey,” *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 89, no. 4, pp. 319–333, 2021.
- [76] S. Franz, R. Irmeler, and U. Rüppel, “Real-time collaborative reconstruction of digital building models with mobile devices,” *Advanced Engineering Informatics*, vol. 38, pp. 569–580, 2018. Publisher: Elsevier Ltd.
- [77] R. Gao, M. Zhao, T. Ye, F. Ye, Y. Wang, K. Bian, T. Wang, and X. Li, “Jigsaw: Indoor floor plan reconstruction via mobile crowdsensing,” in *Pro-*

- ceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom '14, pp. 249–260, Association for Computing Machinery, 2014. event-place: Maui, Hawaii, USA.
- [78] C. Liu, J. Wu, and Y. Furukawa, “FloorNet: A unified framework for floor-plan reconstruction from 3d scans,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11210 LNCS, pp. 203–219, 2018. ISBN: 9783030012304 Publisher: Springer Verlag.
- [79] H. Liu, Y.-L. Yang, S. AlHalawani, and N. J. Mitra, “Constraint-aware interior layout exploration for pre-cast concrete-based buildings,” *The Visual Computer*, vol. 29, no. 6, pp. 663–673, 2013.
- [80] M. Luperto, L. D’Emilio, and F. Amigoni, “A generative spectral model for semantic mapping of buildings,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4451–4458, IEEE, 2015.
- [81] M. Luperto, A. Riva, and F. Amigoni, “Semantic classification by reasoning on the whole structure of buildings using statistical relational learning techniques,” in *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2562–2568, Institute of Electrical and Electronics Engineers Inc., 2017. ISSN: 10504729.
- [82] M. Luperto and F. Amigoni, “Predicting the global structure of indoor environments: A constructive machine learning approach,” *Autonomous Robots*, vol. 43, no. 4, pp. 813–835, 2019.
- [83] N. Nauata, K.-H. Chang, C.-Y. Cheng, G. Mori, and Y. Furukawa, “HouseGAN: Relational generative adversarial networks for graph-constrained house layout generation,” in *Computer Vision – ECCV 2020* (A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, eds.), vol. 12346, pp. 162–177, Springer International Publishing, 2020. Series Title: Lecture Notes in Computer Science.

- [84] A. Paudel, R. Dhakal, and S. Bhattarai, “Room classification on floor plan graphs using graph neural networks,” *CoRR*, vol. abs/2108.05947, 2021.
- [85] Z. Wang, R. Sacks, and T. Yeung, “Exploring graph neural networks for semantic enrichment: Room type classification,” *Automation in Construction*, vol. 134, 2022. Publisher: Elsevier B.V.
- [86] S. Ahmed, M. Liwicki, M. Weber, and A. Dengel, “Automatic room detection and room labeling from architectural floor plans,” in *2012 10th IAPR International Workshop on Document Analysis Systems*, pp. 339–343, IEEE, 2012.
- [87] L.-P. d. l. Heras, S. Ahmed, M. Liwicki, E. Valveny, and G. Sánchez, “Statistical segmentation and structural recognition for floor plan interpretation - notation invariant structural element recognition,” *Int. J. Document Anal. Recognit.*, vol. 17, no. 3, pp. 221–237, 2014.
- [88] Z. Liu and G. Von Wichert, “Extracting semantic indoor maps from occupancy grids,” *Robotics and Autonomous Systems*, vol. 62, no. 5, pp. 663–674, 2014. Publisher: Elsevier.
- [89] D. Camozzato, L. Dihl, I. Silveira, F. Marson, and S. Musse, “Procedural floor plan generation from building sketches,” *Visual Computer*, vol. 31, no. 6, pp. 753–763, 2015. Publisher: Springer Verlag.
- [90] C. Goncu, A. Madugalla, S. Marinai, and K. Marriott, “Accessible on-line floor plans,” in *Proceedings of the 24th International Conference on World Wide Web, WWW ’15*, pp. 388–398, International World Wide Web Conferences Steering Committee, 2015. event-place: Florence, Italy.
- [91] L.-P. de las Heras, O. R. Terrades, and J. Lladós, “Attributed graph grammar for floor plan analysis,” in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 726–730, 2015.
- [92] D. Sharma, C. Chattopadhyay, and G. Harit, “A unified framework for semantic matching of architectural floorplans,” in *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 2422–2427, 2016.

- [93] X. Zhang, A. K.-S. Wong, and C.-T. Lea, "Automatic floor plan analysis for adaptive indoor wi-fi positioning system," in *2016 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 869–874, 2016.
- [94] S. Dodge, J. Xu, and B. Stenger, "Parsing floor plan images," in *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, pp. 358–361, IEEE, 2017.
- [95] A. Madugalla, K. Marriott, and S. Marinai, "Partitioning open plan areas in floor plans," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 01, pp. 47–52, 2017. ISSN: 2379-2140.
- [96] S. Goyal, C. Chattopadhyay, and G. Bhatnagar, "Plan2text: A framework for describing building floor plan images from first person perspective," in *Proceedings - 2018 IEEE 14th International Colloquium on Signal Processing and its Application, CSPA 2018*, pp. 35–40, Institute of Electrical and Electronics Engineers Inc., 2018.
- [97] D. Sharma and C. Chattopadhyay, "High-level feature aggregation for fine-grained architectural floor plan retrieval," *IET Computer Vision*, vol. 12, no. 5, pp. 702–709, 2018. Publisher: Institution of Engineering and Technology.
- [98] T. Yamasaki, J. Zhang, and Y. Takada, "Apartment structure estimation using fully convolutional networks and graph model," in *Proceedings of the 2018 ACM Workshop on Multimedia for Real Estate Tech, RETech'18*, pp. 1–6, Association for Computing Machinery, 2018. event-place: Yokohama, Japan.
- [99] S. Goyal, S. Bhavsar, S. Patel, C. Chattopadhyay, and G. Bhatnagar, "Sugaman: Describing floor plans for visually impaired by annotation learning and proximity-based grammar," *IET Image Processing*, vol. 13, no. 13, pp. 2623–2635, 2019. Publisher: Institution of Engineering and Technology.

- [100] A. Kalervo, J. Ylioinas, M. Häikiö, A. Karhu, and J. Kannala, “CubiCasa5k: A dataset and an improved multi-task model for floorplan image analysis,” *arXiv:1904.01920 [cs]*, 2019.
- [101] Z. Zeng, X. Li, Y. K. Yu, and C.-W. Fu, “Deep floor plan recognition using a multi-task network with room-boundary-guided attention,” *arXiv:1908.11025 [cs]*, 2019.
- [102] H. K. Mewada, A. V. Patel, J. P. Chaudhari, K. K. Mahant, and A. Vala, “Automatic room information retrieval and classification from floor plan using linear regression model,” *Int. J. Document Anal. Recognit.*, vol. 23, no. 4, pp. 253–266, 2020.
- [103] A. Madugalla, K. Marriott, S. Marinai, S. Capobianco, and C. Goncu, “Creating accessible online floor plans for visually impaired readers,” *ACM Trans. Access. Comput.*, vol. 13, no. 4, 2020. Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [104] S. Dong, W. Wang, W. Li, and K. Zou, “Vectorization of floor plans based on EdgeGAN,” *Information (Switzerland)*, vol. 12, no. 5, 2021. Publisher: MDPI AG.
- [105] F. Foroughi, J. Wang, A. Nemati, Z. Chen, and H. Pei, “MapSegNet: A fully automated model based on the encoder-decoder architecture for indoor map segmentation,” *IEEE Access*, vol. 9, pp. 101530–101542, 2021. Publisher: Institute of Electrical and Electronics Engineers Inc.
- [106] Y. Gan, S.-Y. Wang, C.-E. Huang, Y.-C. Hsieh, H.-Y. Wang, W.-H. Lin, S.-N. Chong, and S.-T. Liong, “How many bedrooms do you need? a real-estate recommender system from architectural floor plan images,” *Scientific Programming*, vol. 2021, 2021. Publisher: Hindawi Limited.
- [107] S. Goyal, C. Chattopadhyay, and G. Bhatnagar, “Knowledge-driven description synthesis for floor plan interpretation,” *International Journal on Document Analysis and Recognition*, vol. 24, no. 1, pp. 19–32, 2021. Publisher: Springer Science and Business Media Deutschland GmbH.

- [108] Z. Lu, T. Wang, J. Guo, W. Meng, J. Xiao, W. Zhang, and X. Zhang, “Data-driven floor plan understanding in rural residential buildings via deep recognition,” *Information Sciences*, vol. 567, pp. 58–74, 2021. Publisher: Elsevier Inc.
- [109] G. Murugan, V. Moyal, P. Nandankar, O. Pandithurai, and E. S. J. Pimo, “A novel CNN method for the accurate spatial data recovery from digital images,” *Materials Today: Proceedings*, 2021.
- [110] S. Park and H. Kim, “3dplannet: Generating 3d models from 2d floor plan images using ensemble methods,” *Electronics (Switzerland)*, vol. 10, no. 22, 2021. Publisher: MDPI.
- [111] X. Lv, S. Zhao, X. Yu, and B. Zhao, “Residential floor plan recognition and reconstruction,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16712–16721, 2021. ISSN: 2575-7075.
- [112] S. Ahmed, M. Liwicki, M. Weber, and A. Dengel, “Improved automatic analysis of architectural floor plans,” in *2011 International Conference on Document Analysis and Recognition, ICDAR 2011, Beijing, China, September 18-21, 2011*, pp. 864–869, IEEE Computer Society, 2011.
- [113] Z. Liu, D. Chen, and G. von Wichert, “2d semantic mapping on occupancy grids,” in *ROBOTIK 2012; 7th German Conference on Robotics*, (Munich, Germany), pp. 1–6, VDE-Verlag, May 2012.
- [114] S. Ahmed, M. Weber, M. Liwicki, C. Langenhan, A. Dengel, and F. Petzold, “Automatic analysis and sketch-based retrieval of architectural floor plans,” *Pattern Recognition Letters*, vol. 35, pp. 91–100, 2014.
- [115] A. Paladugu, Q. Tian, H. Maguluri, and B. Li, “Towards building an automated system for describing indoor floor maps for individuals with visual impairment,” *Cyber-Physical Systems*, vol. 1, pp. 132–159, 2015.
- [116] L. Gimenez, S. Robert, F. Suard, and K. Zreik, “Automatic reconstruction of 3d building models from scanned 2d floor plans,” *Automation in Construction*, vol. 63, pp. 48–56, 2016.

- [117] C. Liu, J. Wu, P. Kohli, and Y. Furukawa, "Raster-to-vector: Revisiting floorplan transformation," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2017-October, pp. 2214–2222, Institute of Electrical and Electronics Engineers Inc., 2017. ISSN: 15505499.
- [118] F. Sandelin, "Semantic and instance segmentation of room features in floor plans using mask r-cnn," Master's thesis, Uppsala University, Department of Information Technology, Uppsala, Sweden, Sept. 2019. [Online]. Available: <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-393348>.
- [119] S. Goyal, V. Mistry, C. Chattopadhyay, and G. Bhatnagar, "BRIDGE: Building plan repository for image description generation, and evaluation," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1071–1076, IEEE, 2019.
- [120] H. Jang, K. Yu, and J. Yang, "Indoor reconstruction from floorplan images with a deep learning approach," *ISPRS International Journal of Geo-Information*, vol. 9, no. 2, 2020. Publisher: MDPI AG.
- [121] S. Kim, S. Park, H. Kim, and K. Yu, "Deep floor plan analysis for complicated drawings based on style transfer," *Journal of Computing in Civil Engineering*, vol. 35, no. 2, 2020. Publisher: American Society of Civil Engineers (ASCE).
- [122] I. Surikov, M. Nakhatovich, S. Belyaev, and D. Savchuk, "Floor plan recognition and vectorization using combination unet, faster-rcnn, statistical component analysis and rammer-douglas-peucker," *Communications in Computer and Information Science*, vol. 1235 CCIS, pp. 16–28, 2020. ISBN: 9789811566479 Publisher: Springer.
- [123] Y. Zhang, Y. He, S. Zhu, and X. Di, "The direction-aware, learnable, additive kernels and the adversarial network for deep floor plan recognition," 2020. type: article.
- [124] Z. Liu and G. von Wichert, "A generalizable knowledge framework for semantic indoor mapping based on markov logic networks and data driven MCMC," *arXiv:2002.08402 [cs]*, 2020.

- [125] H. Kim, S. Kim, and K. Yu, “Automatic extraction of indoor spatial information from floor plan image: A patch-based deep learning methodology application on large-scale complex buildings,” *ISPRS International Journal of Geo-Information*, vol. 10, no. 12, 2021. Publisher: MDPI.
- [126] J. Song and K. Yu, “Framework for indoor elements classification via inductive learning on floor plan graphs,” *ISPRS International Journal of Geo-Information*, vol. 10, no. 2, 2021. Publisher: MDPI AG.
- [127] S. Hellbach, M. Himstedt, F. Bahrmann, M. Riedel, T. Villmann, and H.-J. Böhme, “Some room for GLVQ: Semantic labeling of occupancy grid maps,” in *Advances in Self-Organizing Maps and Learning Vector Quantization - Proceedings of the 10th International Workshop, WSOM 2014, Mittweida, Germany, July, 2-4, 2014* (T. Villmann, F.-M. Schleif, M. Kaden, and M. Lange, eds.), vol. 295 of *Advances in Intelligent Systems and Computing*, pp. 133–143, Springer, 2014.
- [128] M. Luperto, A. Quattrini Li, and F. Amigoni, “A system for building semantic maps of indoor environments exploiting the concept of building typology,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8371 LNAI, pp. 504–515, 2014. ISBN: 9783662444672 Publisher: Springer Verlag.
- [129] L. Fermin-Leon, J. Neira, and J. Castellanos, “TIGRE: Topological graph based robotic exploration,” in *2017 European Conference on Mobile Robots, ECMR 2017*, Institute of Electrical and Electronics Engineers Inc., 2017.
- [130] D. Kakuma, S. Tsuichihara, G. Ricardez, J. Takamatsu, and T. Ogasawara, “Alignment of occupancy grid and floor maps using graph matching,” in *Proceedings - IEEE 11th International Conference on Semantic Computing, ICSC 2017*, pp. 57–60, Institute of Electrical and Electronics Engineers Inc., 2017.
- [131] A. Kleiner, R. Baravalle, A. Kolling, P. Pilotti, and M. Munich, “A solution to room-by-room coverage for autonomous cleaning robots,” in *IEEE International Conference on Intelligent Robots and Systems*, vol. 2017-September,

- pp. 5346–5352, Institute of Electrical and Electronics Engineers Inc., 2017. ISSN: 21530858.
- [132] L. Fermin-Leon, J. Neira, and J. A. Castellanos, “Incremental contour-based topological segmentation for robot exploration,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2554–2561, 2017.
- [133] M. Hang, M. Lin, S. Li, Z. Chen, and R. Ding, “A multi-strategy path planner based on space accessibility,” in *2017 IEEE International Conference on Robotics and Biomimetics, ROBIO 2017*, vol. 2018-January, pp. 2154–2161, Institute of Electrical and Electronics Engineers Inc., 2018.
- [134] B. Liu, L. Zuo, C.-H. Zhang, and Y. Liu, “An approach to graph-based grid map segmentation for robot global localization,” in *Proceedings of 2018 IEEE International Conference on Mechatronics and Automation, ICMA 2018*, pp. 1812–1817, Institute of Electrical and Electronics Engineers Inc., 2018.
- [135] M. Mielle, M. Magnusson, and A. Lilienthal, “A method to segment maps from different modalities using free space layout maoris: Map of ripples segmentation,” in *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4993–4999, Institute of Electrical and Electronics Engineers Inc., 2018. ISSN: 10504729.
- [136] M. Hiller, C. Qiu, F. Particke, C. Hofmann, and J. Thielecke, “Learning topometric semantic maps from occupancy grids,” in *IEEE International Conference on Intelligent Robots and Systems*, pp. 4190–4197, Institute of Electrical and Electronics Engineers Inc., 2019. ISSN: 21530858.
- [137] M. Luperto, V. Arcerito, and F. Amigoni, “Predicting the layout of partially observed rooms from grid maps,” in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2019-May, pp. 6898–6904, Institute of Electrical and Electronics Engineers Inc., 2019. ISSN: 10504729.
- [138] S. Schwertfeger and T. Yu, “Room detection for topological maps,” *CoRR*, vol. abs/1912.01279, 2019.

- [139] J. Hou, Y. Yuan, and S. Schwertfeger, “Area graph: Generation of topological maps using the voronoi diagram,” in *2019 19th International Conference on Advanced Robotics (ICAR)*, pp. 509–515, 2019.
- [140] M. Tien, Y. Park, K.-H. Jung, S.-Y. Kim, and J.-E. Kye, “Performance evaluation on the accuracy of the semantic map of an autonomous robot equipped with p2p communication module,” *Peer-to-Peer Networking and Applications*, vol. 13, no. 2, pp. 704–716, 2020. Publisher: Springer.
- [141] T. Zheng, Z. Duan, J. Wang, G. Lu, S. Li, and Z. Yu, “Research on distance transform and neural network lidar information sampling classification-based semantic segmentation of 2d indoor room maps,” *Sensors*, vol. 21, no. 4, p. 1365, 2021.
- [142] M. Luperto, T. P. Kucner, A. Tassi, M. Magnusson, and F. Amigoni, “Robust structure identification and room segmentation of cluttered indoor environments from occupancy grid maps,” *CoRR*, vol. abs/2203.03519, 2022.
- [143] L. Shi, S. Kodagoda, and G. Dissanayake, “Application of semi-supervised learning with voronoi graph for place classification,” in *IEEE International Conference on Intelligent Robots and Systems*, pp. 2991–2996, 2012. ISSN: 21530858.
- [144] K. Sjøo, “Semantic map segmentation using function-based energy maximization,” in *2012 IEEE International Conference on Robotics and Automation*, pp. 4066–4073, IEEE, 2012.
- [145] R. Capobianco, G. Gemignani, D. Bloisi, D. Nardi, and L. Iocchi, “Automatic extraction of structural representations of environments,” *Advances in Intelligent Systems and Computing*, vol. 302, pp. 721–733, 2015. ISBN: 9783319083377 Publisher: Springer Verlag QID: Q62042069.
- [146] M. Liu, F. Colas, L. Oth, and R. Siegwart, “Incremental topological segmentation for semi-structured environments using discretized GVG,” *Autonomous Robots*, vol. 38, no. 2, pp. 143–160, 2015. QID: Q61457771.

- [147] R. Goeddel and E. Olson, “Learning semantic place labels from occupancy grids using CNNs,” in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3999–4004, IEEE, 2016.
- [148] J. Hou, H. Kuang, and S. Schwertfeger, “Fast 2d map matching based on area graphs,” in *IEEE International Conference on Robotics and Biomimetics, ROBIO 2019*, pp. 1723–1729, Institute of Electrical and Electronics Engineers Inc., 2019.
- [149] F. Rubio, J. Martínez-Gómez, M. J. Flores, and J. M. Puerta, “Comparison between bayesian network classifiers and SVMs for semantic localization,” *Expert Systems with Applications*, vol. 64, pp. 434–443, 2016.
- [150] P. Ursic, R. Mandeljc, A. Leonardis, and M. Kristan, “Part-based room categorization for household service robots,” in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2016-June, pp. 2287–2294, Institute of Electrical and Electronics Engineers Inc., 2016. ISSN: 10504729.
- [151] D. Flier, “Human-like room segmentation for domestic cleaning robots,” *Robotics*, vol. 6, no. 4, p. 35, 2017.
- [152] J. Young, V. Basile, M. Suchi, L. Kunze, N. Hawes, M. Vincze, and B. Caputo, “Making sense of indoor spaces using semantic web mining and situated robot perception,” *The Semantic Web: ESWC 2017 Satellite Events*, vol. 10577, pp. 299–313, 2017.
- [153] G. Pintore, F. Ganovelli, R. Pintus, R. Scopigno, and E. Gobbetti, “3d floor plan recovery from overlapping spherical images,” *Computational Visual Media*, vol. 4, no. 4, pp. 367–383, 2018. Publisher: Tsinghua University Press.
- [154] G. Pintore, F. Ganovelli, R. Pintus, R. Scopigno, and E. Gobbetti, “Recovering 3d indoor floor plans by exploiting low-cost spherical photography,” in *Pacific Graphics 2018 Short Papers*, (Hong Kong), pp. 45–48, Eurographics Association, Oct. 2018. [Online]. Available: <https://diglib.eg.org/handle/10.2312/pg20181277>.

- [155] K. Othman and A. Rad, “An indoor room classification system for social robots via integration of CNN and ECOC,” *Applied Sciences (Switzerland)*, vol. 9, no. 3, 2019. Publisher: MDPI AG.
- [156] V. Balaska, L. Bampis, M. Boudourides, and A. Gasteratos, “Unsupervised semantic clustering and localization for mobile robotics tasks,” *Robotics and Autonomous Systems*, vol. 131, 2020. Publisher: Elsevier B.V.
- [157] F. Sadeghi and M. F. Tappen, “Latent pyramidal regions for recognizing scenes,” in *Computer Vision – ECCV 2012*, Lecture Notes in Computer Science, pp. 228–241, Springer, 2012.
- [158] O. Erkent and I. Bozma, “Place representation in topological maps based on bubble space,” in *2012 IEEE International Conference on Robotics and Automation*, pp. 3497–3502, 2012. ISSN: 1050-4729.
- [159] A. Ranganathan, “PLISS: labeling places using online changepoint detection,” *Autonomous Robots*, vol. 32, no. 4, pp. 351–368, 2012.
- [160] S. N. Parizi, J. G. Oberlin, and P. F. Felzenszwalb, “Reconfigurable models for scene recognition,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2775–2782, IEEE, 2012.
- [161] A. Sadvnik and T. Chen, “Hierarchical object groups for scene classification,” in *2012 19th IEEE International Conference on Image Processing*, pp. 1881–1884, IEEE, 2012.
- [162] O. Mozos, H. Mizutani, R. Kurazume, and T. Hasegawa, “Categorization of indoor places using the kinect sensor,” *Sensors (Basel, Switzerland)*, vol. 12, pp. 6695–711, 2012. QID: Q41892808.
- [163] M. Juneja, A. Vedaldi, C. Jawahar, and A. Zisserman, “Blocks that shout: Distinctive parts for scene classification,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 923–930, IEEE, 2013.
- [164] R. Margolin, L. Zelnik-Manor, and A. Tal, “OTC: A novel local descriptor for scene classification,” in *Computer Vision – ECCV 2014*, Lecture Notes in Computer Science, pp. 377–391, Springer International Publishing, 2014.

- [165] Z. Zuo, G. Wang, B. Shuai, L. Zhao, Q. Yang, and X. Jiang, “Learning discriminative and shareable features for scene classification,” in *Computer Vision – ECCV 2014*, Lecture Notes in Computer Science, pp. 552–568, Springer International Publishing, 2014.
- [166] Z. Jie and S. Yan, “Robust scene classification with cross-level LLC coding on CNN features,” in *Computer Vision – ACCV 2014*, Lecture Notes in Computer Science, pp. 376–390, Springer International Publishing, 2015.
- [167] G. Mesnil, S. Rifai, A. Bordes, X. Glorot, Y. Bengio, and P. Vincent, “Unsupervised learning of semantics of object detections for scene categorization,” in *Pattern Recognition Applications and Methods* (A. Fred and M. De Marsico, eds.), Advances in Intelligent Systems and Computing, pp. 209–224, Springer International Publishing, 2015.
- [168] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, “Learning deep features for scene recognition using places database,” *Advances in Neural Information Processing Systems*, vol. 1, 2015. QID: Q77651936.
- [169] M. Dixit, Si Chen, Dashan Gao, N. Rasiwasia, and N. Vasconcelos, “Scene classification with semantic fisher vectors,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2974–2983, IEEE, 2015.
- [170] G. Pintore, V. Garro, F. Ganovelli, E. Gobbetti, and M. Agus, “Omnidirectional image capture on mobile devices for fast automatic generation of 2.5d indoor maps,” in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–9, 2016.
- [171] E. Cruz, J. C. Rangel, F. Gomez-Donoso, Z. Bauer, M. Cazorla, and J. Garcia-Rodriguez, “Finding the place: How to train and use convolutional neural networks for a dynamically learning robot,” in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2018.
- [172] C. Simonsen, F. Thiesson, M. Philipsen, and T. Moeslund, “GENERALIZING FLOOR PLANS USING GRAPH NEURAL NETWORKS,” in *Pro-*

- ceedings - International Conference on Image Processing, ICIP*, vol. 2021-September, pp. 654–658, IEEE Computer Society, 2021. ISSN: 15224880.
- [173] Q. Wei, Q. Wei, Y. Liu, Q. Guan, and D. Liu, “Data-driven room classification for office buildings based on echo state network,” in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, pp. 2602–2607, 2015. ISSN: 1948-9447.
- [174] G. Shi, B. Zhao, C. Li, Q. Wei, and D. Liu, “An echo state network based approach to room classification of office buildings,” *Neurocomputing*, vol. 333, pp. 319–328, 2019.
- [175] P. Uršič, M. Kristan, D. Skočaj, and A. Leonardis, “Room classification using a hierarchical representation of space,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1371–1378, 2012. ISSN: 2153-0866.
- [176] E. Turner and A. Zakhor, “Floor plan generation and room labeling of indoor environments from laser range data,” in *2014 International Conference on Computer Graphics Theory and Applications (GRAPP)*, pp. 1–12, 2014.
- [177] E. Turner and A. Zakhor, “Multistory floor plan generation and room labeling of building interiors from laser range data,” *Communications in Computer and Information Science*, vol. 550, pp. 29–44, 2015. ISBN: 9783319251165 Publisher: Springer Verlag.
- [178] P. Ursic, A. Leonardis, D. Skocaj, and M. Kristan, “Hierarchical spatial model for 2d range data based room categorization,” in *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2016-June, pp. 4514–4521, Institute of Electrical and Electronics Engineers Inc., 2016. ISSN: 10504729.
- [179] X. He, H. Liu, and W. Huang, “Room categorization using local receptive fields-based extreme learning machine,” in *2017 2nd International Conference on Advanced Robotics and Mechatronics, ICARM 2017*, vol. 2018-January, pp. 620–625, Institute of Electrical and Electronics Engineers Inc., 2018.

- [180] H. Wu, G.-h. Tian, Y. Li, F.-y. Zhou, and P. Duan, "Spatial semantic hybrid map building and application of mobile service robot," *Robotics and Autonomous Systems*, vol. 62, no. 6, pp. 923–941, 2014.
- [181] M. Hardegger, D. Roggen, and G. Tröster, "3d ActionSLAM: Wearable person tracking in multi-floor environments," *Personal Ubiquitous Comput.*, vol. 19, no. 1, pp. 123–141, 2015. Place: Berlin, Heidelberg Publisher: Springer-Verlag.
- [182] D. Rojas Castro, A. Revel, and M. Ménard, "Document image analysis by a mobile robot for autonomous indoor navigation," in *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pp. 156–160, 2015.
- [183] S. Loch-Dehbi, Y. Dehbi, G. Gröger, and L. Plümer, "PREDICTION of BUILDING FLOORPLANS USING LOGICAL and STOCHASTIC REASONING BASED on SPARSE OBSERVATIONS," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 265–270, Copernicus GmbH, 2016. ISSN: 21949042 Issue: 2W1.
- [184] Y. Dehbi, S. Loch-Dehbi, and L. Plümer, "PARAMETER ESTIMATION and MODEL SELECTION for INDOOR ENVIRONMENTS BASED on SPARSE OBSERVATIONS," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 303–310, Copernicus GmbH, 2017. ISSN: 21949042 Issue: 2W4.
- [185] S. Loch-Dehbi, Y. Dehbi, and L. Plümer, "Estimation of 3d indoor models with constraint propagation and stochastic reasoning in the absence of indoor measurements," *ISPRS International Journal of Geo-Information*, vol. 6, no. 3, 2017. Publisher: MDPI AG.
- [186] Y. Dehbi, N. Gojayeva, A. Pickert, J.-H. Haunert, and L. Plümer, "Room shapes and functional uses predicted from sparse data," in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Z. S. Dragicevic S., Sithole G., ed.), vol. 4, pp. 33–40, Copernicus GmbH, 2018. ISSN: 21949042 Issue: 4.

- [187] S. G. Shahbandi, M. Magnusson, and K. Iagnemma, “Nonlinear optimization of multimodal two-dimensional map alignment with application to prior knowledge transfer,” *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2040–2047, 2018. Conference Name: IEEE Robotics and Automation Letters.
- [188] X. Hu, H. Fan, A. Noskov, A. Zipf, Z. Wang, and J. Shang, “Feasibility of using grammars to infer room semantics,” *Remote Sensing*, vol. 11, no. 13, 2019. Publisher: MDPI AG.
- [189] R. Zhou, X. Lu, H.-S. Zhao, Y. Fu, and M.-J. Tang, “Automatic construction of floor plan with smartphone sensorsb,” *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 13–25, 2019.
- [190] A. Pronobis and P. Jensfelt, “Large-scale semantic mapping and reasoning with heterogeneous modalities,” in *2012 IEEE International Conference on Robotics and Automation*, pp. 3515–3522, IEEE, 2012.
- [191] I. Kostavelis, K. Charalampous, and A. Gasteratos, “Online spatiotemporal-coherent semantic maps for advanced robot navigation,” in *5th Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV), in conjunction with IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2013)*, (Tokyo, Japan), Nov. 2013. Workshop paper.
- [192] S. Hemachandra, M. R. Walter, S. Tellex, and S. Teller, “Learning spatial-semantic representations from natural language descriptions and scene classifications,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2623–2630, IEEE, 2014.
- [193] N. Sünderhauf, F. Dayoub, S. McMahon, B. Talbot, R. Schulz, P. Corke, G. Wyeth, B. Upcroft, and M. Milford, “Place categorization and semantic mapping on a mobile robot,” *arXiv:1507.02428 [cs]*, 2015.
- [194] I. Kostavelis, K. Charalampous, A. Gasteratos, and J. K. Tsotsos, “Robot navigation via spatial and temporal coherent semantic maps,” *Engineering Applications of Artificial Intelligence*, vol. 48, pp. 173–187, 2016.

- [195] I. Kostavelis and A. Gasteratos, “Semantic maps from multiple visual cues,” *Expert Systems with Applications*, vol. 68, pp. 45–57, 2017.
- [196] M. Liu, R. Chen, D. Li, Y. Chen, G. Guo, Z. Cao, and Y. Pan, “Scene recognition for indoor localization using a multi-sensor fusion approach,” *Sensors*, vol. 17, 2017.
- [197] R. C. Luo and M. Chiou, “Hierarchical semantic mapping using convolutional neural networks for intelligent service robotics,” *IEEE Access*, vol. 6, pp. 61287–61294, 2018.
- [198] C. Jin, A. Elibol, P. Zhu, and N. Y. Chong, “Semantic mapping based on image feature fusion in indoor environments,” in *2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 693–698, IEEE, 2021.
- [199] J. Schäfer, “Practical concerns of implementing machine learning algorithms for w-LAN location fingerprinting,” in *2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 310–317, 2014. ISSN: 2157-023X.
- [200] M. Laska, J. Blankenbach, and R. Klamma, “Adaptive indoor area localization for perpetual crowdsourced data collection,” *Sensors (Switzerland)*, vol. 20, no. 5, 2020. Publisher: MDPI AG.
- [201] N. Peters, H. Lei, and G. Friedland, “Name that room: Room identification using acoustic features in a recording,” in *Proceedings of the 20th ACM International Conference on Multimedia, MM ’12*, pp. 841–844, Association for Computing Machinery, 2012. event-place: Nara, Japan.
- [202] Q. Song, C. Gu, and R. Tan, “Deep room recognition using inaudible echos,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, 2018. Place: New York, NY, USA Publisher: Association for Computing Machinery.
- [203] J. Au-Yeung, M. K. Banavar, and M. Vanitha, “Room classification using acoustic signals,” in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pp. 1–4, 2020.

- [204] N. Resuli, M. Skubic, and S. Kovaleski, “Learning room structure and activity patterns using RF sensing for in-home monitoring of older adults,” in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2054–2061, 2020.
- [205] D. Dziwis, S. Zimmermann, T. Lübeck, J. M. Arend, D. Bau, and C. Pörschmann, “Machine learning-based room classification for selecting binaural room impulse responses in augmented reality applications,” in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, pp. 1–8, 2021.
- [206] M. Walter, S. Hemachandra, B. Homberg, S. Tellex, and S. Teller, “Learning semantic maps from natural language descriptions,” in *Robotics: Science and Systems IX*, Robotics: Science and Systems Foundation, 2013.
- [207] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, IEEE, 2016.
- [208] Y. Wang, N. Funk, M. Ramezani, S. Papatheodorou, M. Popovic, M. Camurri, S. Leutenegger, and M. Fallon, “Elastic and efficient LiDAR reconstruction for large-scale exploration tasks,” in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5035–5041, IEEE, 2021.
- [209] J. L. Carrera, Z. Li, Z. Zhao, T. Braun, and A. Neto, “A real-time indoor tracking system in smartphones,” in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWiM ’16, pp. 292–301, Association for Computing Machinery, 2016. event-place: Malta, Malta.
- [210] S. Pokuciński and D. Mrozek, “Methods and applications of space understanding in indoor environment - a decade survey,” *Applied Sciences-Basel*, vol. 14, no. 10, pp. 1–55, 2024.
- [211] D. B. M. Research, “Facility management market size, share & trends report by 2031,” tech. rep., Data Bridge Market Research, 2024.

- [212] F. B. Insights, “Facility management market size, share, trends | report [2032],” Tech. Rep. FBI101658, Fortune Business Insights, 2023.
- [213] A. Mehra, “Facility management market by offering (solutions (IWMS, BIM, facility operations & security management) and services), vertical (BFSI, retail, construction & real estate, healthcare & life sciences) and region - global forecast to 2028,” Tech. Rep. TC 2470, Markets and Markets, 2023.
- [214] J. Guerrero-Viu, C. Fernandez-Labrador, C. Demonceaux, and J. J. Guerrero, “What’s in my room? object recognition on indoor panoramic images,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 567–573, 2020.
- [215] H. Xu, Q. Zhao, Y. Ma, X. Li, P. Yuan, B. Feng, C. Yan, and F. Dai, “PANDORA: A Panoramic Detection Dataset for Object with Orientation,” in *Computer Vision – ECCV 2022*, Lecture Notes in Computer Science, (Cham), pp. 237–252, Springer Nature Switzerland, 2022.
- [216] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, “A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, 2018.
- [217] T. J. Chowdhury, C. Elkin, V. Devabhaktuni, D. B. Rawat, and J. Oluoch, “Advances on localization techniques for wireless sensor networks: A survey,” *Computer Networks*, vol. 110, pp. 284–305, 2016.
- [218] H. Obeidat, W. Shuaieb, O. Obeidat, and R. Abd-Alhameed, “A review of indoor localization techniques and wireless technologies,” *Wireless Personal Communications*, 2021.
- [219] G. Mintsis, S. Basbas, P. Papaioannou, C. Taxiltaris, and I. N. Tziavos, “Applications of gps technology in the land transportation system,” *European journal of operational Research*, vol. 152, no. 2, pp. 399–409, 2004.

- [220] Y. Gu, A. Lo, and I. Niemegeers, “A survey of indoor positioning systems for wireless personal networks,” *IEEE Communications surveys & tutorials*, vol. 11, no. 1, pp. 13–32, 2009.
- [221] J. Moutinho, R. Araújo, and D. Freitas, “Indoor localization with audible sound — towards practical implementation,” *Pervasive and Mobile Computing*, vol. 29, pp. 1–16, 2016.
- [222] D. Khan, Z. Cheng, H. Uchiyama, S. Ali, M. Asshad, and K. Kiyokawa, “Recent advances in vision-based indoor navigation: A systematic literature review,” *Computers & Graphics*, vol. 104, pp. 24–45, 2022.
- [223] A. Sheinker, B. Ginzburg, N. Salomonski, L. Frumkis, B.-Z. Kaplan, and M. B. Moldwin, “A method for indoor navigation based on magnetic beacons using smartphones and tablets,” *Measurement*, vol. 81, pp. 197–209, 2016.
- [224] Z. Farid, R. Nordin, and M. Ismail, “Recent advances in wireless indoor localization techniques and system.,” *Journal of Computer Networks and Communications*, 2013.
- [225] C. Lian Sang, M. Adams, T. Hörmann, M. Hesse, M. Porrmann, and U. Rückert, “Numerical and experimental evaluation of error estimation for two-way ranging methods,” *Sensors*, vol. 19, no. 3, 2019.
- [226] H. Kim, “Double-sided two-way ranging algorithm to reduce ranging time,” *IEEE Communications Letters*, vol. 13, no. 7, pp. 486–488, 2009.
- [227] J. X. Lee, Z. Lin, and F. C. P. Shin, “Symmetric double side two way ranging with unequal reply time,” in *2007 IEEE 66th Vehicular Technology Conference*, pp. 1980–1983, 2007.
- [228] R. Dalce, A. van den Bossche, and T. Val, “A study of the ranging error for parallel double sided-two way ranging protocol,” in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, 2016.
- [229] M. Utter, “Indoor positioning using ultra-wideband technology,” 2015.

- [230] O. Şeker, B. Şahin, T. Akdoğan, and G. Dalkılıç, “A physical tracking of esp32 iot devices with rssi based indoor position calculation,” *Journal of Millimeterwave Communication, Optimization and Modelling*, vol. 4, pp. 13–16, February 2024.
- [231] A. Norrdine, “An algebraic solution to the multilateration problem,” in *3rd International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2012.
- [232] S. O. Al-Jazzar, A. Muchkaev, A. Al-Nimrat, and M. Smadi, “Low complexity and high accuracy angle of arrival estimation using eigenvalue decomposition with extension to 2d AOA and power estimation,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, p. 123, 2011.
- [233] C. Reddy, M. Tech, and S. B R, “Tdoa computation using multicarrier modulation for sensor networks,” *Int. J. Comput. Sci. Commun. Netw*, vol. 1, 01 2011.
- [234] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, “Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks,” *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.
- [235] N. S. Kodippili and D. Dias, “Integration of fingerprinting and trilateration techniques for improved indoor localization,” in *2010 7th International Conference on Wireless and Optical Communications Networks (WOCN)*, (Colombo, Sri Lanka), pp. 1–6, IEEE, Sept. 2010.
- [236] I. Oppermann, M. Hämäläinen, and J. Iinatti, *UWB: Theory and Applications*. John Wiley & Sons, October 2004. Google-Books-ID: WQbuVCl-NQgC.
- [237] T. Margiani, S. Cortesi, M. Keller, C. Vogt, T. Polonelli, and M. Magno, “Angle of arrival and centimeter distance estimation on a smart uwb sensor node,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–10, 2023.

- [238] C. C. Cossette, M. Shalaby, D. Saussié, J. R. Forbes, and J. Le Ny, “Relative position estimation between two uwb devices with imus,” *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4313–4320, 2021.
- [239] S. Liu, H. Yang, Z. Mei, X. Xu, and Q. He, “Ultra-wideband high accuracy distance measurement based on hybrid compensation of temperature and distance error,” *Measurement*, vol. 206, p. 112276, 2023.
- [240] S. Djosic, I. Stojanovic, M. Jovanovic, T. Nikolic, and G. L. Djordjevic, “Fingerprinting-assisted UWB-based localization technique for complex indoor environments,” *Expert Systems with Applications*, vol. 167, p. 114188, 2021.
- [241] B. Yang, J. Li, Z. Shao, and H. Zhang, “Robust uwb indoor localization for nlos scenes via learning spatial-temporal features,” *IEEE Sensors Journal*, vol. 22, no. 8, pp. 7990–8000, 2022.
- [242] P. Krapež, M. Vidmar, and M. Munih, “Distance Measurements in UWB-Radio Localization Systems Corrected with a Feedforward Neural Network Model,” *Sensors*, vol. 21, p. 2294, January 2021. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.
- [243] A. R. Jiménez Ruiz and F. Seco Granja, “Comparing ubisense, bespoon, and decawave uwb location systems: Indoor performance analysis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 2106–2117, 2017.
- [244] K. Sairam, N. Gunasekaran, and S. Redd, “Bluetooth in wireless communication,” *IEEE Communications Magazine*, vol. 40, no. 6, pp. 90–96, 2002.
- [245] Y.-G. Jan, H.-W. Tseng, Y.-H. Lee, C.-Y. Lo, and L.-Y. Yen, “Accurate Bluetooth Positioning Using Weighting and Large Number of Devices Measurements,” *Wireless Personal Communications*, vol. 79, pp. 1129–1143, November 2014.
- [246] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. Hamalainen, “Experiments on local positioning with bluetooth,” in *Proceedings ITCC 2003. In-*

- ternational Conference on Information Technology: Coding and Computing*, pp. 297–303, 2003.
- [247] S. Zhou and J. Pollard, “Position measurement using bluetooth,” *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 555–558, 2006.
- [248] K. Phutcharoen, M. Chamchoy, and P. Supanakoon, “Accuracy Study of Indoor Positioning with Bluetooth Low Energy Beacons,” in *2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, pp. 24–27, March 2020.
- [249] D. Janczak, W. Walendziuk, M. Sadowski, A. Zankiewicz, K. Konopko, and A. Idzkowski, “Accuracy Analysis of the Indoor Location System Based on Bluetooth Low-Energy RSSI Measurements,” *Energies*, vol. 15, p. 8832, January 2022. Number: 23 Publisher: Multidisciplinary Digital Publishing Institute.
- [250] I. Holguin and S. M. Errapotu, “Smart home iot communication protocols and advances in their security and interoperability,” in *2023 7th Cyber Security in Networking Conference (CSNet)*, pp. 208–211, 2023.
- [251] V. C. Pham, T. Xin, T. Nguyen-Mau, G. D. Le Trong, and Y. Tan, “Towards echonet lite web api as a restful api for matter devices,” in *2024 IEEE 13th Global Conference on Consumer Electronics (GCCE)*, pp. 1251–1252, 2024.
- [252] S. B. Prabhu, M. I., and S. Phatak, “Vision controlled matter smart home solution for dysdiadochokinesia patients,” in *2023 14th International Conference on Computing Communication and Networking Technologies (ICC-CNT)*, pp. 1–7, 2023.
- [253] E. Gambi, L. Montanini, L. Raffaeli, S. Spinsante, and L. Lambrinos, “Interoperability in iot infrastructures for enhanced living environments,” in *2016 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, pp. 1–5, 2016.

- [254] F. Montori, L. Bedogni, F. Morselli, and L. Bononi, “Achieving iot interoperability through a service oriented in-home appliance,” in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pp. 1–6, 2017.
- [255] E. S. Pramukantoro, F. Andri Bakhtiar, B. Aji, and R. Pratama, “Middleware for network interoperability in iot,” in *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 499–502, 2018.
- [256] I. P. Zarko, “Why interoperability matters to your or any iot solution,” in *2019 15th International Conference on Telecommunications (ConTEL)*, pp. 1–1, 2019.
- [257] H. Shahinzadeh, A.-s. Mirhedayati, M. Shaneh, H. Nafisi, G. B. Gharehpetian, and J. Moradi, “Role of joint 5g-iot framework for smart grid interoperability enhancement,” in *2020 15th International Conference on Protection and Automation of Power Systems (IPAPS)*, pp. 12–18, 2020.
- [258] K. Shashwat, F. Hahn, X. Ou, and A. Singhal, “Security Analysis of Trust on the Controller in the Matter Protocol Specification,” in *2023 IEEE Conference on Communications and Network Security (CNS)*, pp. 1–6, 2023.
- [259] F. Klement, E. Vorderwülbecke, and S. Katzenbeisser, “One Standard to Rule Them All? Assessing the Disruptive Potential of Jamming Attacks on Matter Networks,” in *2023 IEEE International Workshop on Information Forensics and Security (WIFS)*, pp. 1–6, 2023. ISSN: 2157-4774.
- [260] V. C. Pham, T. Nguyen-Mau, M. Sioutis, and Y. Tan, “Enabling ECHONET Lite Devices in the Matter Ecosystem: A Bridge Solution,” in *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE)*, pp. 867–868, 2023. ISSN: 2693-0854.
- [261] S. Nagashige and H. Suzuki, “Design of Protocol Converter to Interconnect Matter and ECHONET Lite,” in *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE)*, pp. 871–872, 2023. ISSN: 2693-0854.

- [262] K. Hanzel and D. Grzechca, “Increasing the security of smart cities of the future thanks to uwb technology,” in *Information Systems* (M. Themistocleous and M. Papadaki, eds.), pp. 585–596, Springer International Publishing, 2022.
- [263] K. Hanzel, K. Paszek, and D. Grzechca, “The influence of the data packet size on positioning parameters of uwb system for the purpose of tagging smart city infrastructure,” *Bulletin of the Polish Academy of Sciences Technical Sciences*, vol. 68, no. No. 4 (i.a. Special Section on Advances in Electrical Power Engineering), pp. 857–868, 2020.
- [264] Makerfabs, “ESP32 UWB – Ultra Wideband Module.” <https://www.makerfabs.com/esp32-uwband-ultra-wideband.html>, n.d. Accessed: 2025-07-27.
- [265] Espressif Systems, “Esp32-c6-devkitc-1 development board.” <https://docs.espressif.com/projects/esp-dev-kits/en/latest/esp32c6/esp32-c6-devkitc-1/index.html>, 2023. Accessed: 2025-07-27.
- [266] Project CHIP Contributors, “Matter: The Official Matter Protocol Implementation.” <https://github.com/project-chip/connectedhomeip>, 2024. Accessed: 2025-07-28.
- [267] J. E. DOBSON, *The Birth of Computer Vision*. University of Minnesota Press, 2023.
- [268] R. Tadeusiewicz and P. Korohoda, *Komputerowa analiza i przetwarzanie obrazów*. Społeczeństwo Globalnej Informacji, Wydawnictwo Fundacji Postępu Telekomunikacji, 1997.
- [269] L. Jiao, F. Zhang, F. Liu, S. Yang, L. Li, Z. Feng, and R. Qu, “A survey of deep learning-based object detection,” *CoRR*, vol. abs/1907.09408, 2019.
- [270] K. Stapor, *Metody klasyfikacji obiektów w wizji komputerowej*. Warszawa: Wydawnictwo Naukowe PWN, 2011.

- [271] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image segmentation using deep learning: A survey,” *CoRR*, vol. abs/2001.05566, 2020.
- [272] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [273] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, vol. 15 of *JMLR Workshop and Conference Proceedings*, pp. 315–323, 2011.
- [274] V. Dumoulin and F. Visin, “A guide to convolution arithmetic for deep learning,” 2018.
- [275] F. Yu and V. Koltun, “Multi-scale context aggregation by dilated convolutions,” 2016.
- [276] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” 2017.
- [277] F. Yu, V. Koltun, and T. Funkhouser, “Dilated residual networks,” 2017.
- [278] S. Schubert, P. Neubert, J. Pöschmann, and P. Protzel, “Circular convolutional neural networks for panoramic images and laser data,” in *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 653–660, 2019.
- [279] A. Bansal, K. Sikka, G. Sharma, R. Chellappa, and A. Divakaran, “Zero-shot object detection,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.
- [280] W. Cao, X. Yao, Z. Xu, Y. Liu, Y. Pan, and Z. Ming, “A survey of zero-shot object detection,” *Big Data Mining and Analytics*, vol. 8, no. 3, pp. 726–750, 2025.
- [281] L. Zhang, X. Wang, L. Yao, L. Wu, and F. Zheng, “Zero-shot object detection via learning an embedding from semantic space to visual space,”

- in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20* (C. Bessiere, ed.), pp. 906–912, International Joint Conferences on Artificial Intelligence Organization, 7 2020. Main track.
- [282] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.
- [283] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A Review of Yolo Algorithm Developments,” *Procedia Computer Science*, vol. 199, pp. 1066–1073, 2022.
- [284] M. Hussain, “Yolo-v1 to yolo-v8, the rise of yolo and its complementary nature toward digital manufacturing and industrial defect detection,” *Machines*, vol. 11, no. 7, 2023.
- [285] G. Jocher, “Yolov5.” <https://github.com/ultralytics/yolov5>, 2020.
- [286] G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics yolov8,” 2023.
- [287] G. Jocher and J. Qiu, “Ultralytics yolo11,” 2024.
- [288] B. Mahaur and K. Mishra, “Small-object detection based on YOLOv5 in autonomous driving systems,” *Pattern Recognition Letters*, vol. 168, pp. 115–122, Apr. 2023.
- [289] J. Chen, E. Bao, and J. Pan, “Classification and Positioning of Circuit Board Components Based on Improved YOLOv5,” *Procedia Computer Science*, vol. 208, pp. 613–626, 2022.
- [290] N. Aishwarya, K. Manoj Prabhakaran, F. T. Debebe, M. S. S. A. Reddy, and P. Pranavee, “Skin Cancer diagnosis with Yolo Deep Neural Network,” *Procedia Computer Science*, vol. 220, pp. 651–658, 2023.
- [291] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes (voc) challenge,” *International Journal of Computer Vision*, vol. 88, pp. 303–338, June 2010.

- [292] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft COCO: Common Objects in Context,” in *Proceedings of the European Conference on Computer Vision*, pp. 740–755, Springer, 2014.
- [293] D. Powers, “Evaluation: From precision, recall and f-measure to roc, informedness, markedness & correlation,” *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [294] P. Christen, D. J. Hand, and N. Kirielle, “A review of the f-measure: Its history, properties, criticism, and alternatives,” *ACM Comput. Surv.*, vol. 56, Oct. 2023.
- [295] M. Braun, S. Krebs, F. Flohr, and D. Gavrilu, “Eurocity persons: A novel benchmark for person detection in traffic scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, pp. 1–1, 02 2019.
- [296] Y. Hamzeh, A. Mohammadi, and S. A. Rawashdeh, “Improving the performance of automotive vision-based applications under rainy conditions,” *IET Image Processing*, vol. 16, no. 5, pp. 1457–1472, 2022.
- [297] A. I. Bandos, H. E. Rockette, T. Song, and D. Gur, “Area under the free-response roc curve (froc) and a related summary index,” *Biometrics*, vol. 65, pp. 247–256, 03 2009.
- [298] H. Ai, Z. Cao, J. Zhu, H. Bai, Y. Chen, and L. Wang, “Deep learning for omnidirectional vision: A survey and new perspectives.”
- [299] J. Xiao, K. A. Ehinger, A. Oliva, and A. Torralba, “Recognizing scene viewpoint using panoramic place representation,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2695–2702, June 2012. ISSN: 1063-6919.
- [300] Y. Zhang, S. Song, P. Tan, and J. Xiao, “PanoContext: A Whole-Room 3D Context Model for Panoramic Scene Understanding,” in *Computer Vision – ECCV 2014*, Lecture Notes in Computer Science, (Cham), pp. 668–686, Springer International Publishing, 2014.

- [301] C. Zhang, Z. Cui, C. Chen, S. Liu, B. Zeng, H. Bao, and Y. Zhang, “Deep-panocontext: Panoramic 3d scene understanding with holistic scene context graph and relation-based optimization,” in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12612–12621, 2021.
- [302] Y. Zhang, L. Zhang, W. Hamidouche, and O. Deforges, “A Fixation-Based 360° Benchmark Dataset For Salient Object Detection,” in *2020 IEEE International Conference on Image Processing (ICIP)*, pp. 3458–3462, Oct. 2020. ISSN: 2381-8549.
- [303] S.-H. Chou, C. Sun, W.-Y. Chang, W.-T. Hsu, M. Sun, and J. Fu, “360-Indoor: Towards Learning Real-World Objects in 360° Indoor Equirectangular Images,” in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 834–842, Mar. 2020. ISSN: 2642-9381.
- [304] S. D. S. of Sustainability, “Stanford 2d-3d-semantic dataset (2d-3d-s),” June 2024.
- [305] G. Albanis, N. Zioulis, P. Drakoulis, V. Gkitsas, V. Sterzentsenko, F. Alvarez, D. Zarpalas, and P. Daras, “Pano3d: A holistic benchmark and a solid baseline for 360° depth estimation,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 3722–3732, 2021.
- [306] B. Sekachev, N. Manovich, M. Zhiltsov, A. Zhavoronkov, D. Kalinin, B. Hoff, T. Osmanov, D. Kruchinin, A. Zankevich, D. DmitriySidnev, M. Markelov, Johannes222, M. Chenuet, a andre, telenachos, A. Melnikov, J. Kim, L. Ilouz, N. Glazov, Priya4607, R. Tehrani, S. Jeong, V. Skubriev, S. Yonekura, vugia truong, zliang7, lizhming, and T. Truong, “opencv/cvat: v1.1.0,” Aug. 2020.
- [307] A. Torralba, B. C. Russell, and J. Yuen, “Labelme: Online image annotation and applications,” *Proceedings of the IEEE*, vol. 98, no. 8, pp. 1467–1484, 2010.
- [308] B. Dwyer, J. Nelson, T. Hansen, *et al.*, “Roboflow (version 1.0) [software],” 2025. Computer vision.

- [309] S.-T. Yang, C.-H. Peng, P. Wonka, and H.-K. Chu, “PanoAnnotator: a semi-automatic tool for indoor panorama layout annotation,” in *SIGGRAPH Asia 2018 Posters*, (Tokyo Japan), pp. 1–2, ACM, Dec. 2018.
- [310] L.-H. Lin, H.-K. Wen, M.-H. Kao, E. Chen, T.-H. Lin, and M. Ouhyoung, “Label360: An Annotation Interface for Labeling Instance-Aware Semantic Labels on Panoramic Full Images,” in *SIGGRAPH Asia 2020 Posters*, (Virtual Event Republic of Korea), pp. 1–2, ACM, Dec. 2020.
- [311] L. Althoff, M. Prado, S. Alamgeer, A. R. E. Silva, R. Prakash, M. M. Carvalho, and M. C. Q. Farias, “360RAT: A Tool for Annotating Regions of Interest in 360-degree Videos,” in *Brazilian Symposium on Multimedia and Web*, (Curitiba Brazil), pp. 272–280, ACM, Nov. 2022.
- [312] J. Kopf, D. Lischinski, O. Deussen, D. Cohen-Or, and M. Cohen, “Locally Adapted Projections to Reduce Panorama Distortions,” *Computer Graphics Forum*, vol. 28, pp. 1083–1089, June 2009.
- [313] J. Gonçalves, D. Ribeiro, and F. Soares, “Perspective correction of panoramic images created by parallel motion stitching,” in *WSCG '2015: Short Communications Proceedings — The 23rd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2015)*, (Plzen, Czech Republic), pp. 125–132, Václav Skala – UNION Agency, June 2015. [Online]. Available: <http://hdl.handle.net/11025/29674>.
- [314] W. Yan and C. Hou, “Reducing perspective distortion for stereoscopic image stitching,” in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pp. 1–6, 2016.
- [315] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” 2016.
- [316] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015.
- [317] K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” 2016.

- [318] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, “Searching for mobilenetv3,” 2019.
- [319] S. Pokuciński and D. Mrozek, “Object detection with yolov5 in indoor equirectangular panoramas,” *Procedia Computer Science*, vol. 225, pp. 2420–2428, 2023.
- [320] S. Pokuciński and D. Mrozek, “Yolo-based object detection in panoramic images of smart buildings,” in *2023 IEEE 10th International Conference on Data Science & Advanced Analytics (DSAA), October 9-12, 2023, Thessaloniki, Greece. Proceedings* (Y. Manolopoulos and Z.-H. Zhou, eds.), pp. 1–9, 2023.
- [321] Ultralytics, “Object detection datasets overview,” 2023. Dostęp: 13 maja 2025.
- [322] J. Guerrero-Viu, C. Fernandez-Labrador, C. Demonceaux, and J. J. Guerrero, “What’s in my Room? Object Recognition on Indoor Panoramic Images,” in *Proceedings of the International Conference on Robotics and Automation*, pp. 567–573, IEEE, 2020.
- [323] S. Pokuciński, K. Filus, and D. Mrozek, “Finding the perfect cut: selection of the best cutting point in equirectangular panoramas for object detection,” *Procedia Computer Science*, vol. 246, pp. 519–528, 2024.
- [324] A. Ramesh, M. Pavlov, G. Goh, S. Gray, C. Voss, A. Radford, M. Chen, and I. Sutskever, “Zero-shot text-to-image generation,” 2021.
- [325] A. Bochkovskii, A. Delaunoy, H. Germain, M. Santos, Y. Zhou, S. R. Richter, and V. Koltun, “Depth pro: Sharp monocular metric depth in less than a second,” 2025.
- [326] S. F. Bhat, R. Birkel, D. Wofk, P. Wonka, and M. Müller, “Zoedepth: Zero-shot transfer by combining relative and metric depth,” 2023.
- [327] L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng, and H. Zhao, “Depth anything: Unleashing the power of large-scale unlabeled data,” in *CVPR*, 2024.

- [328] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, J. Feng, and H. Zhao, “Depth anything v2,” *arXiv:2406.09414*, 2024.
- [329] D. Grzechca, A. Ziebinski, D. Komorowski, K. Hanzel, S. Pokucinski, K. Klonowski, and S. Siwek, “Monitoring the gait process during the rehabilitation of patients using computer vision techniques and UWB technology,” in *Information Systems*, Lecture Notes in Business Information Processing, pp. 419–437, Springer International Publishing, 2020.
- [330] D. Grzechca, K. Tokarz, K. Hanzel, and S. Pokucinski, “On fundamental issues on creating autonomous platform for UWB navigation system,” *Journal of Control Engineering and Applied Informatics*, vol. 22, no. 1, pp. 84–93, 2020. Number: 1.
- [331] D. Grzechca, A. Ziębiński, D. Komorowski, S. Pokuciński, M. Gorywoda, K. Klonowski, S. Siwek, R. Kristof, and A. Czerwiński, “Stanowisko badawcze do oceny postępów rehabilitacyjnych kończyn dolnych pacjenta oparte na analizie wizyjnej, systemie pozycjonowania uwb oraz wkładce tensometrycznej,” *PM News*, no. 21, p. 9, 2019.
- [332] D. Neiryneck, E. Luk, and M. McLaughlin, “An alternative double-sided two-way ranging method,” in *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*, pp. 1–4, 2016.
- [333] Decawave, “Decawave uwb legacy software — github organization,” 2025. Archiwalne repozytoria UWB dla DW1000/DWM1001 (np. `uwb-core`, `uwb-apps`, `dwm1001-examples`); organizacja przejęta przez Qorvo.
- [334] N. Zioulis, A. Karakottas, D. Zarpalas, F. Alvarez, and P. Daras, “Spherical view synthesis for self-supervised 360° depth estimation,” in *International Conference on 3D Vision (3DV)*, September 2019.