

4 września 2023 r.

dr hab. inż. Artur Klepaczko, prof. ucz.  
Instytut Elektroniki  
Politechniki Łódzkiej

RECENZJA ROZPRAWY DOKTORSKIEJ  
DLA RADY DYSCYPLINY INŻYNIERIA BIOMEDYCZNA  
W POLITECHNICE ŚLĄSKIEJ

*Tytuł rozprawy:* Deep learning applications in biomedical engineering

*Autor rozprawy:* mgr inż. Konrad Duraj

*Promotor:* dr hab. inż. Paweł Kostka, prof. PŚ

**I. Ocena układu rozprawy, w tym informacja o jej poszczególnych częściach składowych. Czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie oraz umiejętność samodzielnego prowadzenia pracy naukowej?**

Przedłożona do recenzji rozprawa jest napisana w języku angielskim i obejmuje 5 rozdziałów oraz bibliografię. Zasadniczy trzon rozprawy stanowi cykl 4 współautorskich publikacji Doktoranta, na który składają się 2 referaty opublikowane w materiałach konferencyjnych oraz 2 artykuły opublikowane w czasopiśmie międzynarodowych. Rozdział 1., zatytułowany *Introduction*, zawiera krótki opis rozwoju metod sztucznej inteligencji, następnie prezentuje podstawy systemów ekspertowych, w tym systemów logiki rozmytej, algorytmów uczenia maszynowego oraz, w odrębnym podrozdziale, uczenia głębokiego. W rozdziale tym brakuje jednak rzeczywistego wprowadzenia czytelnika do faktycznej tematyki rozprawy, czy choćby nawiązania do jej tytułu. Rozdział 2, zatytułowany *Aim of the doctoral thesis*, prezentuje cel pracy, zarysowany jednak bardzo szeroko i jednocześnie zdawkowo – Autor ograniczył się zasadniczo do powtórzenia informacji zawartych w tytule rozprawy. Rozdział 3. zawiera zestawienie wspomnianych publikacji należących do cyklu. Każda publikacja opatrzona została wstępem, w którym syntetycznie podsumowano oryginalność danego rozwiązania oraz merytoryczny wkład Autora. Nie podano procentowego udziału Doktoranta w danej publikacji. Rozdział 4. stanowi uzupełnienie cyklu publikacji o projekt systemu łączącego elementy uczenia głębokiego i wnioskowania rozmytego. W rozdziale 5. zamieszczono podsumowanie badań wraz ze wskazaniem możliwych kierunków dalszych prac. Chociaż rozdział ten zatytułowano *Conclusions*, to nie zawiera on jednak oryginalnych obserwacji wykraczających poza istniejącą powszechną wiedzę o potencjalnie szerokim zakresie zastosowań uczenia głębokiego w inżynierii biomedycznej i problemach z tym związanych.

Podsumowując, układ rozdziałów jest zasadniczo poprawny. Jednak w pracy brakuje dobrego wprowadzenia do kontekstu rozprawy, uwzględniającego wyraźne i logiczne uzasadnienie do podjęcia takich, a nie innych prac, uszczegółowienia celu pracy, spójnego podsumowania osiągnięć zaprezentowanych w poszczególnych pozycjach cyklu publikacji oraz krytycznej analizy uzyskanych rezultatów. Z drugiej strony, z całą pewnością zawartość

merytoryczna rozprawy świadczy o osiągnięciu przez Autora umiejętności do samodzielnego prowadzenia prac naukowych oraz pozyskaniu odpowiedniej wiedzy w dyscyplinie, w której realizowany jest doktorat.

**II. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Tematyka podjęta w recenzowanej rozprawie doktorskiej obejmuje szereg różnych kierunków badawczych dotyczących zastosowania algorytmów uczenia głębokiego w tzw. rozpoznawaniu wzorców (ang. *pattern recognition*) w danych, w szczególności w sygnałach, biomedycznych. Dynamiczny rozwój wspomnianych algorytmów, jaki obserwujemy w ostatnich latach, w tym pojawienie się różnorodnych architektur splotowych oraz rekurencyjnych sieci neuronowych, a także metod ich uczenia, w sposób naturalny predestynuje je do wykorzystania do rozwiązywania złożonych, nieliniowo uwarunkowanych problemów klasyfikacji lub segmentacji danych wymagających jednocześnie wytworzenia ich użytecznej reprezentacji numerycznej. W literaturze przedmiotu odnaleźć można obecnie liczne przykłady takich zastosowań. Tym niemniej, rozprawa prezentuje opis eksperymentów badawczych przeprowadzonych z udziałem Autora, które mimo wszystko zawierają elementy nowatorskie na tle aktualnego stanu wiedzy.

Cel pracy został zdefiniowany jako próba oceny użyteczności oraz ograniczeń uczenia głębokiego w różnych dziedzinach inżynierii biomedycznej. Cel ten został sformułowany w rozdziale 2. stosunkowo jasno, moim zdaniem jednak zbyt ogólnie. Na podstawie lakonicznego tekstu tego rozdziału trudno jest ocenić, jaki jest szczególny wkład Autora do dyscypliny i w czym, w ramach bardzo pojemnej dziedziny uczenia maszynowego, Doktorant osiągnął ekspercki poziom wiedzy.

Tezy badawcze w pracy nie zostały postawione, jednak w kolejnych załączonych do tekstu rozprawy publikacjach, których Doktorant jest współautorem, określono następujące zagadnienia naukowe:

1. W publikacji [43] podjęto próbę wytworzenia modelu obliczeniowego do oceny toksyczności substancji chemicznych na podstawie cech struktury molekularnej.
2. Publikacja [44] prezentuje model głębokiej sieci neuronowej do rozpoznawania rodzaju aktywności kierowcy podczas jazdy samochodem na podstawie sygnału elektrookulograficznego (EOG).
3. W ramach publikacji [45] wykonano ocenę skuteczności detekcji akcji bicia serca w sygnale sejsmokardiograficznym (SCG) z wykorzystaniem jednowymiarowej sieci splotowej do semantycznej segmentacji szeregów czasowych.
4. W publikacji [46] zaprezentowano model sieci do semantycznej segmentacji sygnału elektrokardiograficznego (ECG), tzn. podziału zapisu ECG na tło oraz sekcje odpowiadające załamkom P, T i zespołowi QRS.

Ponadto, w rozdziale 4. zatytułowanym *A new Paradigm*, przedstawiono oryginalną koncepcję systemu łączącego elementy logiki rozmytej oraz uczenia głębokiego. Zaproponowany system ma w założeniu być odpowiedzią na ograniczenia sieci splotowych, jak również systemów wnioskowania rozmytego.

Rozprawa ma więc charakter eksperymentalno-teoretyczny. Zarówno „nowy paradygmat” przedstawiony w rozdziale 4., jak i publikacje [43]-[46] zawierają oryginalne projekty sieci neuronowych, których skuteczność potwierdzono w eksperymentach wykonanych dla danych pozyskanych z otwartych repozytoriów lub wytworzonych w ramach prowadzonych badań.

### III. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy, świadczący o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Przegląd doniesień literaturowych dotyczących zagadnień rozpatrywanych w poszczególnych pozycjach cyklu artykułów odpowiada wymogom formatu danej publikacji. Stąd w przypadku publikacji konferencyjnych [43] i [45] przegląd ten obejmuje odpowiednio jedynie 2 i 6 pozycji odnoszących się bezpośrednio do nowych rozwiązań proponowanych przez Autora, natomiast w przypadku artykułów [44] i [46] liczba cytowanych konkurencyjnych podejść wzrasta odpowiednio do 15 i 17. Na podstawie przeglądu, Doktorant wraz z pozostałymi współautorami publikacji wyciągają wnioski odnośnie do oryginalności własnych metod, w tym:

- Ad 43. **Hybrydowej architektury głębokiej sieci neuronowej zawierającej warstwy splotowe i rekurencyjne.** O ile nowatorskość tego rozwiązania nie budzi wątpliwości, to jednak zamieszczony tu przegląd literatury jest bardzo skromny. Kwerenda *molecule toxicity prediction deep learning* tylko w czasopiśmie *BMC Bioinformatics* wydawnictwa Springer zwraca kilkanaście wyników powiązanych z tematem publikacji.
- Ad 44. **Modelu klasyfikatora do rozpoznawania aktywności kierowcy na podstawie sygnału EOG.** Zgodnie z informacjami zawartymi w p. 3.2.3 przegląd literatury w tej publikacji nie należał do zadań Autora rozprawy, jednakże ewaluacja zaproponowanego modelu już tak. W tym zakresie, Autor dokonał rzetelnego porównania zaprojektowanego modelu z innymi opublikowanymi rozwiązaniami.
- Ad 45. **Modelu sieci neuronowej do detekcji uderzeń serca w sygnale SCG.** Przegląd literatury wykonany przez Autora jest tu stosunkowo szeroki jak na referat konferencyjny i obejmuje pozycje dotyczące różnych zastosowań sygnału SCG, jak również metod jego przetwarzania. Jednakże konkurencyjne rozwiązania wykorzystujące metody uczenia maszynowego zostały jedynie wskazane w cytowaniach [23]-[27] i nie przeprowadzono ich dogłębnej analizy. Szczegółowe różnice wykazano jedynie między proponowanym modelem a podejściem opisanym w pozycji [28] i obejmowały one m.in. algorytm przetwarzania wstępnego sygnału, rodzaj normalizacji, czy też szczegóły architektury i parametry sieci neuronowej.
- Ad 46. **Modelu sieci do semantycznej segmentacji sygnału ECG.** W tej publikacji przegląd literatury wygląda na najbardziej rzetelny i spina kłamrą wprowadzenie do artykułu oraz podsumowanie wyników.

Ze względu na znaczną dywersyfikację kierunków badań Autora trudno jednoznacznie ocenić prawidłowość wykonanej analizy źródłowej. Jednakże można zauważyć pozytywny trend w sposobie prezentacji stanu wiedzy światowej w kolejnych publikacjach, jak również w

rozdziale 4. rozprawy, co świadczy ostatecznie o zdobyciu przez Autora dostatecznej wiedzy z zakresu tematycznego doktoratu.

#### IV. Czy autor rozwiązał poprawnie postawione zagadnienia, czy użył właściwej do tego metody?

W rozdziale 1 rozprawy Autor opisał systemy eksperckie, w tym algorytmy logiki rozmytej oraz metody uczenia głębokiego wykorzystane później w przeprowadzonych badaniach. Rozdział ten świadczy o tym, iż Autor posiada kompetencje w zakresie eksploracji danych oraz opanował w stopniu zaawansowanym niezbędny warsztat narzędziowy. Wybór metod klasyfikacji danych jest siłą rzeczy arbitralny, chociaż zawiera grupę algorytmów należących do kanonu uczenia maszynowego. Z punktu widzenia celu rozprawy najistotniejsze techniki inteligencji obliczeniowej zostały opisane w podrozdziale 1.4 (*Deep learning*), w którym przedstawiono podstawowe typy warstw sieci neuronowych, metody uczenia, regularyzacji, czy też normalizacji danych. Techniki te zostały wykorzystane do realizacji następujących projektów:

1. Opracowanie sieci neuronowej do predykcji toksyczności substancji chemicznej na podstawie jej cech strukturalnych i zapisu symbolicznego [43].

Materiał badawczy w tym zadaniu pracy stanowił otwarty zbiór danych o nazwie „SMILES Toxicity”. Zgodnie z podanymi informacjami zbiór ten charakteryzuje się znaczną nierównowagą między instancjami należącymi do dwóch klas – cząstki toksyczne (964 przykłady) oraz nietoksyczne (6698 przykładów). Problem ten oznacza konieczność przetworzenia bazy danych do postaci zrównoważonej lub zastosowania odpowiedniej ważonej funkcji straty. Autor zaproponował tu dwa rozwiązania – modele oznaczone jako I i II zawierające dwukierunkowe jednostki rekurencyjne (właściwe do przetwarzania sekwencyjnych danych będących przedmiotem badania) lub filtry splotowe. Oba modele należy w ogólności uznać za prawidłowe podejścia do rozwiązania zagadnienia, jednakże przyjęta metoda badawcza zawiera pewne mankamenty:

- Model II zawiera istotne różnice w architekturze w stosunku do modelu I, m.in. warstwy normalizacyjne (*BatchNormalization*), warstwy splotowe oraz warstwę redukującą (*MaxPooling*). Jednocześnie zastosowano inną technikę uczenia, obejmującą ważoną funkcję straty, tzw. funkcję ogniskową (ang. *focal loss*), czy też cykliczne zmniejszanie współczynnika uczenia. Mamy tu więc nałożenie na siebie dwóch osi zmian – architektonicznych i procesowych. Tymczasem ocena modelu II i porównanie z modelem I zostały wykonane jednocześnie. Nie można więc ocenić, który czynnik miał wpływ na poprawę działania modelu II, a teza, że zadziałały oba te czynniki jest nieuprawniona.
- Model I wydaje się nadmiernie uproszczony. Nie uwzględniono w nim choćby warstwy normalizacyjnej, co sugeruje, że został on intencjonalnie osłabiony, aby wykazać dobroczynne skutki modyfikacji zastosowanych w modelu II.
- Autorzy zwracają uwagę na problem nadmiernego dopasowania (ang. *overfitting*). Z jednej strony tym bardziej niezrozumiałą jest brak warstwy normalizacyjnej w modelu I (warstwa ta uczy się bowiem pewnego przesunięcia i rozproszenia w przestrzeni cech utajonych dla losowych pakietów danych uczących, uodparniając sieć na zależność od konkretnego rozkładu danych). Z drugiej zaś strony, zarówno w artykule [43], jak i we

wprowadzeniu rozprawy (sekcja 1.4.2) nie znajdujemy wyjaśnienia dla zastosowania warstwy typu *Dropout*.

- Samo zjawisko nadmiernego dopasowania zostało ujęte w kontekście nierównowagi klas w zbiorze uczącym. Z definicji jednak, problem ten oznacza, że przetrenowany model bardzo dobrze rozdziela klasy w zbiorze treningowym, ale nie generalizuje się do danych testowych. Jednakże w przypadku niezbalansowanych danych uczących podstawowym problemem nie jest zbyt dokładne dopasowanie się modelu do rozkładu klas. Przeciwnie nawet – model uczy się ignorować klasę mniejszościową ponieważ błąd na niej popełniany „kosztuje” stosunkowo mało. W tym sensie, ani idealne zrównoważenie klas, ani ważona funkcja straty nie są technikami przeciwdziałającymi nadmiernemu dopasowaniu i problem ten pozostanie nierozwiązany w przypadku np. rozbudowanych architektur sieci neuronowych o zbyt dużej liczbie warstw i przez to dużej pojemności pamięci.
- Co więcej, naiwne nadpróbkowanie zbioru danych w celu sztucznego powiększenia liczności klasy mniejszościowej – a takie działanie podjęto w przypadku uczenia modelu I – zwiększa ryzyko nadmiernego dopasowania ponieważ hiperpłaszczyzna decyzyjna modelu znajduje numerycznie silne podparcie w de facto niedostatecznie reprezentowanej kategorii danych. Do pewnego stopnia wolne od tego problemu są techniki augmentacji danych, np. SMOTE (ang. *synthetic minority oversampling technique*). Jak Autor widzi możliwość wykorzystania tego rodzaju algorytmów w zastosowaniu do danych będących przedmiotem badań w publikacji [43]?

## 2. Opracowanie sieci neuronowej do klasyfikacji rodzaju aktywności kierowcy na podstawie sygnału EOG [44].

Celem tego zadania było rozwinięcie modelu klasyfikatora, który rozpoznawałby jedną z czterech predefiniowanych aktywności kierowcy (parkowanie, jazda w mieście, przejazd przez skrzyżowanie oraz przez rondo) na podstawie sygnału EOG mierzonego za pomocą inteligentnych okularów. Na potrzeby projektu utworzono własną bazę danych sygnałów pozyskanych w badaniach z udziałem 10 doświadczonych kierowców i takiej samej liczby słuchaczy szkoły jazdy. Zgodnie z wykazem zamieszczonym w pkt. 3.2.3, pozyskanie tej bazy danych nie należy jednak do osiągnięć Autor rozprawy. Ponownie, zadaniem Doktoranta było opracowanie i ewaluacja samego modelu klasyfikatora. Autor zaproponował stosunkowo prostą architekturę sieci złożoną z dwóch szeregowo połączonych bloków zbudowanych z jednowymiarowej warstwy splotowej z funkcją aktywacji typu *ReLU* oraz warstwy redukcji danych. Bloki zostały dodatkowo rozdzielone warstwą *Dropout*. Ta prosta architektura okazała się bardzo skuteczna i jednocześnie wydaje się odpowiednia do relatywnie małego zbioru danych zawierającego w sumie 520 pomiarów. Wątpliwości dotyczą:

- przetwarzania wstępnego bazy danych – zgodnie z zaproponowaną metodą, długości poszczególnych sygnałów wyrównywano do najdłuższego pomiaru poprzez ich  $n$ -krotną replikację, gdzie  $n$  było liczone jako stosunek liczby próbek najdłuższego pomiaru (32,357) do liczby próbek danego fragmentu. Czy powstałe w ten sposób ciągi danych nie zawierają zatem sztucznych wzorców, nie odzwierciedlających faktycznego zachowania kierowcy? Jaki był rozkład długości pomiarów w poszczególnych klasach?

Jeśli był to rozkład nierównomierny, to istnieje ryzyko, że hiperpłaszczyzna decyzyjna klasyfikatora opiera się na cechach związanych z replikacją ciągów próbek sygnału, a nie ich charakterystycznym przebiegiem.

- użycia w tym samym badaniu zarówno danych pochodzących od kierowców doświadczonych, jak i uczących się jazdy – zostało to umotywowane obserwacją, że pomiędzy obiema grupami kierowców nie występują zasadnicze różnice w zdolnościach poznawczych i motorycznych. Należy się z tym stwierdzeniem zgodzić, jednakże reakcje kierowców doświadczonych na sytuację drogową są inne, często wyuczone i może zachodzić w ich przypadku inny mechanizm przetwarzania bodźców. Skoro autorzy posiadali informację o podziale kierowców ze względu na doświadczenie, interesujące byłoby sprawdzenie efektywności modelu osobno dla każdej z grup.

3. Opracowanie sieci neuronowej do detekcji uderzeń serca w sygnale SCG [45].

W pracy wykorzystano publicznie dostępny zbiór danych CEBS, zawierający m.in. sygnały ECG oraz odpowiadające im pomiary SCG. Autor przeprowadził wstępne przetwarzanie danych, w efekcie którego oznaczył w sygnałach sejsmokardiograficznych akcje uderzeń serca. Oznaczeń tych dokonał na podstawie detekcji załamek R w sygnale EKG. W architekturze opracowanej sieci wykorzystano jednostki *Squeeze-Excitation* (SE). Nie wyjaśniono jednak w jakim celu skorzystano z tego rozwiązania i czy podstawowa architektura sieci do semantycznej segmentacji typu U-Net okazała się niewystarczająca.

4. Opracowanie sieci neuronowej do semantycznej segmentacji sygnału ECG [46].

Podobnie do projektu opisanego w publikacji [45] wykorzystano tu istniejącą otwartą bazę sygnałów ECG oraz zastosowano architekturę sieci U-Net. Artykuł ten wyróżnia się najbardziej wszechstronnym podejściem do oceny wytrenowanego modelu. Oprócz metryk typowych dla segmentacji (*IoU* oraz współczynnik  $F_1$ ) wyliczono m.in. dokładność, czułość, swoistość i precyzję. W odróżnieniu od publikacji [45] porównano także skuteczność segmentacji dla sieci zawierającej jednostki SE, jak również w wariancie podstawowym.

5. Opracowanie głębokiej sieci neuronowej do generacji hipotez na potrzeby wnioskowania opartego o logikę rozmytą. Projekt ten stanowi integralną część recenzowanej rozprawy doktorskiej i ma stanowić odpowiedź na ograniczenia metod uczenia głębokiego w kontekście zastosowań w inżynierii biomedycznej. Autor rzetelnie przedstawił rodzinę systemów neuronowo-rozmytych pozwalających na automatyczne tworzenie reguł na podstawie przykładów ze zbioru uczącego, przechodząc od systemów kooperatywnych, poprzez modele współbieżne, aż po systemy hybrydowe. Spośród tych ostatnich Autor opisał m.in. systemy typu Mamdaniego, Takagi-Sugeno oraz model adaptacyjny (ANIFS). Następnie przedstawił założenia własnego rozwiązania, które w odróżnieniu od istniejących pozwala na zastosowanie dowolnie głębokiej sieci neuronowej i pozwala na wytworzenie zrozumiałej, interpretowalnej hipotezy wyjaśniającej w miejsce ściśle określonej predykcji. Działanie systemu zostało zilustrowane na przykładzie tzw. problemu bramki XOR, który stanowi klasyczne nieliniowe zagadnienie klasyfikacji danych. Jednakże takie zastosowanie odbiega od tytułu rozprawy i jako takie nie wyjaśnia

użyteczności zaproponowanego podejścia w zadaniach analizy i rozpoznawania danych biomedycznych.

**V. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?**

Literatura naukowa w zakresie przedmiotu rozprawy doktorskiej jest niezwykle bogata i prezentuje liczne przykłady zastosowań metod uczenia głębokiego do rozwiązywania różnorodnych zagadnień przetwarzania danych medycznych. Tym bardziej wykazanie oryginalności osiągnięć wymaga od Autora bardzo precyzyjnego i rzetelnego przedstawienia stanu wiedzy i konfrontacji własnych wyników względem niego. Niestety tekst rozprawy takiego podsumowania nie zawiera. Tym niemniej jako oryginalne można wskazać modele sieci neuronowych głębokiego uczenia wytrenowane do poszczególnych zadań opisanych w sekcji IV recenzji. Same modele nie zawierają nowatorskich elementów architektonicznych, ale stanowią przykłady skutecznego zastosowania głębokiego uczenia w obszarach, które jak dotąd nie zostały intensywnie w tym kontekście przebadane. Najbardziej oryginalnym rozwiązaniem wydaje się natomiast system neuronowo-rozmyty zaproponowany w rozdziale 4. System ten dokonuje syntezy hipotezy, np. o przynależności tensora wejściowego do określonej klasy, zamiast dokonywać jednoznacznej predykcji. Takie ujęcie problemu jest nieoczywiste i charakteryzuje się dużym potencjałem publikacyjnym. Nie wykazano jednak doświadczalnie związku między tym rozwiązaniem a domeną inżynierii biomedycznej, co powinno być zwornikiem prac prezentowanych w rozprawie.

**VI. Czy autor wykazał umiejętność poprawnego przedstawienia uzyskanych wyników i ich interpretacji (zwięzłość, jasność)?**

Opis osiągnięć badawczych Autora rozprawy został przedstawiony w postaci 4 współautorskich publikacji oraz jednego rozdziału zawierającego nieopublikowany wcześniej opis koncepcji systemu neuronowo-rozmytego. Oceniając te dwa składniki pracy jako całość należy uznać, iż rozprawa jest napisana dobrym językiem technicznym o odpowiednim poziomie szczegółowości, chociaż zawiera pewne niedociągnięcia w zakresie poprawnej prezentacji wyników wskazane w sekcji VII niniejszej recenzji.

**VII. Jakie są słabe strony rozprawy i ewentualnie jej główne wady?**

W mojej ocenie najsłabszą stroną rozprawy jest jej format. Trzonem rozprawy jest cykl 4 publikacji, w których Autor opisał najważniejsze osiągnięcia badawcze. Dwie z tych publikacji stanowią referaty konferencyjne. Siłą rzeczy muszą być one zwięzłe ze względu na ograniczenia redakcyjne tego rodzaju publikacji. Dlatego Autor nie mógł zaprezentować szerzej wyników swoich prac, ani też przeprowadzić głębszej analizy i porównania ze stanem wiedzy. Tego rodzaju informacje mogłyby się znaleźć lub zostać rozwinięte w autoreferacie, który stanowiłby przewodnik po publikacjach. Ponadto taki autoreferat pozwoliłby na doprecyzowanie celu rozprawy, który jak wskazałem w sekcji II został sformułowany bardzo ogólnie. Faktem jest przecież, że rozprawa nie wyczerpuje tak szeroko zakrojonego tematu jak zastosowania w biomedycynie metod uczenia głębokiego. W rozprawie brakuje również łącznika pomiędzy poszczególnymi publikacjami oraz logicznej syntezy technicznych osiągnięć Autora. Wreszcie,

decydując się na sporządzenie rozprawy częściowo w postaci cyklu współautorskich publikacji, oczekiwałbym określenia procentowego udziału Autora w każdej z nich.

W odniesieniu do poszczególnych części rozprawy można zauważyć natomiast pewne szczegółowe niedociągnięcia edycyjne lub merytoryczne wymienione poniżej.

- W publikacji [43]:
  - Brakuje wyjaśnienia kolorów przebiegów (treningowy, walidacyjny) oraz opisów osi na rys. 3 i 4. Ponadto, pomimo, iż funkcja straty maleje w całym zakresie epok, to jednak dokładność stabilizuje się od epoki ok. 30. Może to być oznaką wchodzenia modelu w stan przetrenowania.
  - Nie podano, w jaki sposób liczona była dokładność klasyfikacji, tzn. czy do jej wyznaczenia wykorzystano tzw. metrykę zrównoważoną (ang. *balanced accuracy*), co ma znaczenie w przypadku klas o bardzo różnej liczności.
- W publikacji [44]:
  - W pkt. 3.2.3 Autor używa terminu *methodology* do określenia swojego wkładu do publikacji polegającego na opracowaniu modelu sieci neuronowej i przeprowadzeniu całego procesu jego treningu i ewaluacji. Z definicji jednak metodologia jest nauką o metodach lub badaniem metod. Wkładem Autora nie jest więc metodologia, ale sama metoda lub ich zestaw.
  - Motywacją badań ma być tu poprawa bezpieczeństwa w ruchu na drodze. Nie wykazano, w jaki sposób opracowany model mógłby przyczynić się do osiągnięcia tego celu. Rozpoznawane przez klasyfikator aktywności kierowców są wolnozmiennie. Tymczasem sytuacje zagrożenia bezpieczeństwa charakteryzują się nagłymi i krótkotrwałymi reakcjami. Czy opracowany system jest zdolny do wykrywania tego rodzaju zachowań?
    - Na rysunku 11 nie podano nazw klas, a w tekście nie ma powiązanie pomiędzy numerami a poszczególnymi aktywnościami kierowców.
- W publikacji [46]:
  - W dyskusji wyników podniesiono kwestię ograniczeń zaproponowanego rozwiązania. Wśród nich wymieniono zależność detekcji uderzenia serca w sygnale SCG od obecności załamka R w sygnale ECG. Nie wyjaśniono, na czym ta zależność polega poza faktem zaetykietowania sygnału SCG na podstawie analizy sygnału ECG.
- W publikacji [46]:
  - Na rys. 8 widać wyraźnie moment (12-13 epoka), po którym następuje wzrost funkcji straty dla zbioru walidacyjnego przy jednocześnie malejącej wartości tej funkcji dla zbioru treningowego. Oznacza to, iż model został przetrenowany. Nie zostało to jednak w żaden sposób skomentowane w artykule. Czy wyniki zamieszczone w tabelach 1 i 2 oraz na rys. 9 zostały obliczone dla modelu z najmniejszą funkcją straty dla zbioru walidacyjnego, czy dla modelu odpowiadającemu ostatniej epoce?
  - Macierze pomyłek pokazane na rys. 9 zostały błędnie znormalizowane względem całkowitej liczby wektorów danych. Tymczasem normalizacja powinna odbyć się w poszczególnych kategoriach, tak aby wartości w rzędach sumowały się do 1.



## **VIII. Co wniosła rozprawa do nauki i/lub techniki oraz ocena rozprawy**

Autor rozprawy wykazał, że zastosowanie metod inteligencji obliczeniowej umożliwia skuteczne rozpoznawanie sygnałów i danych biomedycznych. Opracowane metody zostały poddane weryfikacji dla danych zebranych podczas badań lub pozyskanych z ogólnodostępnych repozytoriów danych. Za oryginalny wkład Autora do dyscypliny inżynieria biomedyczna uważam opracowanie zestawu algorytmów do klasyfikacji toksyczności molekuł, rozpoznawania aktywności kierowców na podstawie sygnału EOG, detekcji uderzeń serca w sygnale SCG oraz semantycznej segmentacji sygnału ECG. Ponadto Autor zaproponował oryginalną koncepcję systemu neuronowo-rozmytego. Prace te inicjują nowy kierunek badań nad algorytmami uczenia głębokiego, które zapewniają większy niż obecnie dostępny poziom wyjaśnialności i transparentności procesu decyzyjnego.

### **Wniosek końcowy**

Stwierdzam, że przedłożona do recenzji rozprawa doktorska spełnia wymagania Ustawy Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 roku, Dziennik Ustaw Nr 2018, poz. 1668 z późn. zm. Wnioskuje o dopuszczenie rozprawy doktorskiej mgra inż. Konrada Duraja do publicznej obrony.

Podpisał Artur Klepaczko (podpis odręczny)