

Recenzja rozprawy doktorskiej mgr inż. Marka Sochy p.t. "Computer-Assisted Detection and Quantification of Low-Attenuation Lung Patterns in CT Studies of COPD Patients"

1. Tematyka rozprawy

Przedmiotem rozprawy doktorskiej mgr inż. Marka Sochy są metody przetwarzania i segmentacji obrazów dla obrazowania medycznego, w szczególności w zastosowaniu do tomografii komputerowej (TK, CT) w diagnostyce wybranych schorzeń płuc. Tematyka ta plasuje ją w obszarze przetwarzania obrazów i sygnałów, widzenia komputerowego, rozpoznawania obrazów, oraz w pewnym stopniu odkrywania wiedzy i uczenia maszynowego, co pozwala mi stwierdzić że praca jednoznacznie wpisuje się w obszar dziedziny Informatyka Techniczna i Telekomunikacja.

2. Ocena treści rozprawy i wkładu oryginalnego

2.1 Treść rozprawy

Rozprawa mgr Sochy składa się z 9 rozdziałów. W rozdziale 1 Autor wprowadza tematykę rozprawy i formułuje hipotezy badawcze. Rozdział 2 charakteryzuje TK jako narzędzie obrazowania medycznego. W rozdziale 3 Autor opisuje anatomię płuc oraz charakterystykę schorzenia Chronic Obstructive Pulmonary Disease (COPD), tj. przewlekłej obturacyjnej choroby płuc, stanowiącej przedmiot badań rozprawy. Rozdział 4 krótko charakteryzuje zbiory danych obrazowych wykorzystane w dalszych częściach pracy. Kolejne cztery rozdziały stanowią właściwe przyczynki pracy i dotyczą poszczególnych etapów analizy obrazu, tj. segmentacji płuc i ich płatów (rozdział 5), wyodrębnienia dróg oskrzelowych (rozdział 6), właściwej identyfikacji znamion (biomarkerów) COPD charakteryzujących się niskim tłumieniem sygnału (rozdział 7), oraz analizą ilościową (rozdział 8). Rozprawę zamyka podsumowujący rozdział 9.

2.2 Ocena wkładu oryginalnego i prezentacji pracy

Dr Socha sformułował trzy hipotezy (s. 3), pierwszą o charakterze ogólnym, pozostałe dwie bardziej konkretne i o znacznej w mojej ocenie wartości badawczej. Wszystkie one zostały w mojej ocenie potwierdzone w rozprawie. Autor wykazuje się bardzo dobrą znajomością tematu, w tym od strony medycznej, odnośnie anatomii płuc, klinicznych aspektów obrazowania KT, oraz sposobów w jaki COPD obrazuje się w TK – co zaznaczmy nie jest normą w pracach w dyscyplinie informatyka, gdzie wielu autorów koncentruje się bardziej na danych i ich przetwarzaniu, przywiązując mniejszą wagę do ich źródła i procesu ich pozyskiwania. Znajomość ta znajduje odzwierciedlenie także w licznych odwołaniach do wcześniejszych prac, co przełożyło się na znaczną liczbę 233 prac w bibliografii rozprawy. Dr Socha nie ograniczył się zresztą jedynie do aspektów obrazowych, prezentując także nie-obrazowe metody diagnozowania COPD i rozedmy (emphysema) w rozdziale 3.

Głównymi przyczynkami pracy są w mojej ocenie umotywowanie, zaprojektowanie, implementacja i rzetelne przeanalizowanie wieloetapowej metody segmentacji i analizy ilościowej wzorców o niskim tłumieniu w obrazach płuc w obrazowaniu TK, sprofilowanej na zastosowanie w diagnostyce COPD (a właściwie kilku metod, z wariantami). W szczególności, mgr Socha zdecydował się zaproponować metodę która angażuje algorytmy uczenia maszynowego jedynie w niewielkim stopniu (m.in. etap segmentacji płuc w rozdziale 5 czy małe sieci neuronowe w na potrzeby wizualizacji w rozdziale 8). Szczególną zaletą proponowanej metody jest jej bardzo silne oparcie na wiedzy dziedziny oraz wynikające z jej konstrukcji znaczne możliwości wyjaśniania sugerowanych decyzji. Doceniam też próby wykorzystania metody jako narzędzia diagnostyki różnicowej oraz celem scharakteryzowania różnych typów i stopni zaawansowania COPD oraz wyjaśniania istotności wybranych cech obrazu w rozdziale 8. Jedną z istotnych motywacji dla przyjętego toku postępowania w projektowaniu metody jest brak publicznie dostępnych etykietowanych badań z COPD – wobec takiego stanu rzeczy, przyjęta przez Doktoranta strategia wydaje się adekwatna.

Praca zredagowana jest starannie i ma dobrze przemyślaną strukturę. Podczas jej lektury zidentyfikowałem relatywnie mało niedopatrzeń. W rozdziale 1 skrót COPD wprowadzany jest i przynajmniej trzykrotnie, co wydaje się nadmiarowe jak 4-stronicowy rozdział. Na str. 40 i 45 napotkałem wadliwe 'wiszące' referencje ('??'). Poziom języka angielskiego jest dobry i nie budzi zastrzeżeń, poza nielicznymi wyjątkami; np. fraza 'on images' na s. 16; 'when COPD is suspected' na s. 19; czy 'The input was compromised from the 4 canonical CT views'. Zdarzają się drobne uchybienia terminologiczne, np. termin 'train set' na s. 30 zamiast 'training set'. Przydałaby się też większa konsekwencja w użyciu terminologii; np. etapy 'slices' i 'grouping' w rysunkach 7.1 i 7.2 nazywane są w tekście odpowiednio 'cross-sections' i 'clustering'. Podobnie wskazane byłoby częstsze używanie precyzyjnych referencji skrótnych do innych rozdziałów i sekcji, bo sformułowania typu 'As described in the "Anatomy" chapter and the "State-of-the-art" sub-section' (s. 57) czy 'based on previously developed algorithms' (s. 59), a także w innych częściach pracy (np. s. 72) nie ułatwiają sprawnego nawigowania po tekście rozprawy. Niemniej te drobne braki nie stanowią znacznej przeszkody w lekturze pracy.

Uważam że rozprawa mogłaby zyskać na przystępności i atrakcyjności dzięki większej liczbie diagramów, np. dla wyjaśnienia charakterystyki filtrów splotowych w sekcji 2.8; w początkowych sekcjach rozdziału 3, gdzie tłumacząc anatomię płuc Autor przytacza wiele terminów medycznych; w prezentacji bardziej wyrafinowanych metod (np. metoda segmentacji dróg oskrzelowych w sekcji 6.3, zwłaszcza w okolicach sekcji 6.3.5-7, gdzie bardzo przydałby się diagram ilustrujący jak częściowe wyniki agregowane są w finalne).

Rekomendowałbym też bardziej klarowne wskazywanie które części rozprawy stanowią oryginalne przyczynki Doktoranta – w obecnym manuskrypcie jest to miejscami niejasne, co po części bierze się z polegania na gramatycznym trybie biernym, oraz być może na milczącym założeniu że sekcje zatytułowane 'Methods' zawierają takie przyczynki. Na przykład w pierwszym akapicie sekcji 7.3 pada zdanie "The second method tested was Exact Threshold Estimation (ETE) method" – nie pada tu jednoznaczne stwierdzenie że jest to metoda autorska.

Pewne (niemniej nieliczne) wątki w rozprawie wydają się być potraktowane zbyt zdawkowo. W rozdziale 4 Autor charakteryzuje krótko jedynie dwa zbiory danych TK COPD, podczas gdy nawet pobieżny przegląd internetowy wydaje się sugerować że takich kolekcji jest więcej (np. zbiór związany z publikacją [204]). Uważam że należałoby wymienić choćby krótko te pozostałe, albo przynajmniej przyznać się do wiedzy o ich istnieniu, a co ważniejsze uzasadnić dlaczego wybór Doktoranta padł właśnie na COPDGene i MOLTEST-BIS. Brakowało mi także komentarza dotyczącego zaskakującej właściwości pierwszej z tych kolekcji, tj. znikomej obecności w nim badań o wartości skali GOLD równej 1 (rys. 4.2) – ta nieregularność bardzo rzuca się w oczy i nasuwa się pytanie czy odzwierciedla naturalny stan rzeczy w populacji generalnej, czy też jest to artefakt tej konkretnej bazy danych. Zastanawiam się także jaki ma ona wpływ na konkluzje rozprawy.

W rozdziałach 5 i 6, Autor umiejętnie motywuje projekt metody i algorytmów, opierając się na wiedzy anatomicznej i charakterystyce schorzenia COPD i jego objawów. Następnie ilustruje przykładami obrazowymi skuteczność wybranych rozwiązań. W prezentacji tej niestety nie stawia jednak przysłowiowej kropki nad i, pomijając ocenę ilościową poszczególnych etapów obliczeń (przez co rozumiem tu w szczególności próbę skonfrontowania metody z anotacjami płatów i dróg oskrzelowych, które spodziewam się że są dostępne w innych kolekcjach obrazów niż te wymienione w rozdziale 4). Brakuje jej zupełnie w rozdziale 5; w rozdziale 6 pojawia się tabela 6.1 prezentująca wyniki w pod kątem dokładności (accuracy), ale jest ona w tekście wprowadzona jednym krótkim zdaniem, a interpretacja metryki nie jest jasna (i nie jest jasne na podstawie którego ze zbiorów danych zaprezentowanych w rozdziale 4 zostały one obliczone).

Z drugiej strony pewne (relatywnie niewielkie) części rozprawy uważam za nadmiernie detaliczne. Uważam m.in. że nie ma konieczności prezentacji wyników badań wstępnych które są raczej oczywiste i spodziewane – na przykład nie jest niczym zaskakującym że segmentacja przez progowanie obrazu (sekcja 5.3) w nietrywialnych zastosowaniach nie ma prawa osiągnąć zadowalających wyników. Procedura wstępnego przetwarzania danych w sekcji 8.3.3 prezentowana jest na zbyt podstawowym i detalicznym poziomie – standaryzacja zmiennych, macierz kowariancji czy własności i wektory własne to elementarne pojęcia które nie wymagają prezentacji wzorów.

Podobna uwaga dotyczy prezentacji metody UMAP w sekcji 8.3.4 – jest ona co prawda dużo bardziej wyrafinowana, ale popularna, dobrze znana w środowisku i traktowana jest gotowe narzędzie z półki (cytowana dziesiątki tysięcy razy). Detale dot. takich elementów warsztatu powinny znaleźć się raczej w załącznikach do rozprawy.

W rozprawie napotkałem też kilka stwierdzeń które wzbudziły moje wątpliwości.

1. Na s. 3 pada stwierdzenie “This research concentrates on classical methods, which have shown better generalisation abilities than neural networks”. Być może intencją Autora było zapowiedzenie w ten sposób wybranych konkluzji rozprawy, niemniej w kontekście wprowadzenia zdanie to brzmi jak ogólny osąd, który nie znajduje potwierdzenia w faktach: choć sam jestem zwolennikiem stosowania prostych i skutecznych metod tam gdzie się one sprawdzają, sieci neuronowe osiągają często dużo lepsze wyniki.
2. Sekcja 2.3 definiuje DICOM jako typ/format plików, podczas gdy jest to dużo szerszy standard, obejmujący także protokoły komunikacyjne.
3. Akapit na dole s. 44 sugeruje że poddanie obrazu rastrowego filtrowaniu filtrem Gaussa adresuje brak ciągłości i dwukrotnej różniczkowalności. Oczywiście tak nie jest – raster otrzymany z tego filtrowania pozostaje rastrem, nieciągłym i nieróżniczkowalnym.
4. W sekcji 5.4 Autor wydaje się przeczyć samemu sobie: w drugim akapicie argumentuje za metodami klasycznymi (nie angażującymi uczenia maszynowego), ale już na początku kolejnego akapitu wskazuje na metodę Hofmanningera i współautorów, tj. architekturę neuronową, jako komponent proponowanego potoku przetwarzania (do segmentacji płuc).
5. Kluczowy etap potoku obliczeniowego proponowanego przez Autora, tj. wyodrębnianie wzorców o niskim tłumieniu (rozdział 7), bazuje na progowaniu obrazu. Dr Socha poddał wyniki tej metody (w dwóch wariantach, ATE i ETE, różniących się sposobem agregowania informacji, tj. per tomogram i globalnie) solidnej analizie statystycznej, w tym pod kątem zgodności ze skalą GOLD (rozdział 7.4), i wyniki te wskazują na pomocniczą zdolność metody. Ale zauważmy że tego typu analiza ilościowa obecności poszczególnych przedziałów intensywności sygnału (tłumienia), bazująca na statystykach pikseli, jest nadal relatywnie mało wyrafinowana, ponieważ ignoruje charakterystyki/cechy przestrzenne obrazu (piksele traktowane są po prostu jako zbiór), na przykład charakterystyki tekstury. W tym kontekście twierdzenie (pojawiające się także w tytule pracy) że mamy tu do czynienia z wyodrębnianiem wzorców o niskim tłumieniu jest nieco na wyrost.
6. Doktorant używa algorytmu k-średnich (K-means) do grupowania obserwacji w metodach opisywanych w rozdziałach 7 i 8. Jest to jeden z najbardziej podstawowych algorytmów grupowania, i nasuwa się pytanie czy nie warto by zdać się na bardziej wyrafinowane i nowsze metody (np. DBSCAN, który wydaje się dobrze sprawdzać w wielu zastosowaniach).

Powyższe wątpliwości nie mają jednak krytycznego charakteru i nie wpływają na moją ogólnie pozytywną ocenę rozprawy.

3. Konkluzja końcowa

Rozprawa doktorska mgr inż. Marka Sochy zawiera interesujące i wartościowe przyczynki do algorytmów analizy obrazowania medycznego, charakteryzujące się dobrym ugruntowaniem w stanie wiedzy o schorzeniu COPD i jego obrazowaniu w TK, i znacznym potencjale aplikacyjnym, potwierdzonym wstępnymi badaniami empirycznymi. Uważam zatem że **spełnia ona warunki stawiane przez ustawę o tytule naukowym i stopniach naukowych w odniesieniu do rozpraw doktorskich, a zatem powinna być dopuszczona do publicznej obrony, o co wnoszę do Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej.**