

## Recenzja

pracy doktorskiej mgr. inż. Marka Wodzińskiego

pt. „Medical image registration methods focused on the problem of missing data”.

Promotor: dr hab. inż. Andrzej Skalski, prof. AGH

Niniejsze recenzja została przygotowana w odpowiedzi na pismo prof. dr. hab. inż. Piotra Augustyniaka, Przewodniczącego Rady Dyscypliny Inżynieria, Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie wraz z pismem przewodnim L.Dz. RDIB\0012\8\2021 z dnia 19 kwietnia 2021. W przesłanej dokumentacji załączono załączniki do Uchwały nr 12/RDIB/2021 w sprawie ustalenia zasad wyróżniania rozpraw doktorskich.

### 1. Tematyka, cele i tezy rozprawy

Tematyka pracy jest aktualna, łączy bowiem aspekty dynamicznie rozwijanych metod uczenia maszynowego stosującego głębokie sieci neuronowe z praktycznym i ważnym w inżynierii biomedycznej problemem dopasowania obrazów. Poszukiwanie algorytmów, które prowadzą do dokładnego dopasowania obrazów, w szczególności w przypadku istotnych zmian w dopasowywanych obrazach (np. z powodu interwencji chirurgicznej) ma duże znaczenie praktyczne w diagnostyce medycznej.

Pan mgr inż. Marek Wodziński, przedstawiając motywację do podjęcia prac, wskazał trzy istotne problemy medyczne, jakimi zajął się w ramach prowadzonych badań:

- dopasowanie obrazów histologicznych uzyskanych z zastosowaniem różnych czynników barwiących,
- lokalizacja w obrazach CT łoży nowotworowej po resekcji nowotworu piersi,
- dopasowanie obrazów USG uzyskiwanych w trakcie resekcji glejaka mózgu.

Cele rozprawy nie zostały zapisane w sposób jawny, niemniej przedstawiony opis w rozdziale pierwszym wyraźnie wskazuje kierunki prowadzonych prac oraz tezę rozprawy.

Autor postawił hipotezę, że „metody dedykowane dla problemu brakujących danych mogą poprawić dokładność estymacji transformacji geometrycznej w procesie dopasowania obrazów medycznych.” (str. 6).

Sformułował również następującą tezę: „Medical image registration algorithms dedicated to the missing data problem may significantly improve the alignment quality” (str. 9). Teza jest interesująca, uważam jednak, że została zdefiniowana zbyt ogólnie i wymagałaby doprecyzowania w celu podkreślenia oryginalnych propozycji przedstawionych przez mgr. inż. Marka Wodzińskiego w dalszych częściach rozprawy.

## 2. Zawartość rozprawy

Rozprawa napisana jest w języku angielskim i składa się z dwóch części.

Pierwsza część (rozdziały 2, 3 i 4) stanowi wprowadzenie do tematyki rozprawy wraz z przedstawieniem stanu wiedzy oraz omawia problemy związane z metodami dopasowania obrazów medycznych, w szczególności w przypadku brakujących danych.

W rozdziale 2 autor omówił podstawowe zagadnienia związane z klasycznymi metodami dopasowania obrazów, m.in. stosowane przekształcenia geometryczne, miary podobieństwa, metody optymalizacji, itp.

Rozdział 3 wprowadza czytelnika w tematykę modeli głębokich sieci neuronowych w odniesieniu do problemu dopasowania obrazów.

W rozdziale 4 doktorant przedstawia ważne w kontekście pracy definicje dotyczące typów brakujących danych. Wyróżnił 3 typy brakujących danych w obrazach medycznych w zakresie rozpatrywanej tematyki pracy: brak struktury w obrazie np. ze względu na ograniczenia metody obrazowej - typ I, brak struktury w obrazie w wyniku interwencji (np. chirurgicznej, radioterapii, itp. - typ II, różnice w odpowiadających regionach obrazów (typ III). Dla typu danych brakujących I i III autor zasadniczo zakłada ignorowanie brakujących struktur (nie nakładających się regionów), ponieważ nie wnoszą istotnej informacji. Niemniej istnieją i dlatego muszą być uwzględnione w procesie dopasowania. Uwzględnienie brakujących danych stosowane jest z wykorzystaniem czynników regularyzacji (ang. regularization/penalty terms) na czym skupia się doktorant w drugiej części rozprawy.

Druga część pracy (rozdziały 5, 6, 7 i 8) przedstawia główne osiągnięcia autora, w szczególności w trzech wspomnianych wcześniej obszarach aplikacyjnych.

Rozdział 5 poświęcony jest dopasowaniu obrazów histologicznych uzyskanych z zastosowaniem różnych czynników barwiących. Przedstawia dwa algorytmy stanowiące osiągnięcia pracy, zbiór danych wykorzystywany w procesie uczenia i walidacji oraz wyniki, porównując je do osiągnięć innych zespołów.

W rozdziale 6 mgr inż. Marek Wodziński proponuje metodę lokalizacji łoża nowotworowej po resekcji nowotworu piersi w obrazach TK. Ponownie przedstawia proponowany algorytm, czynnik regularyzacji, zbiór danych do walidacji oraz uzyskane wyniki wraz z prostą oceną wzrokową uzyskanych wyników.

Podobną formę zastosował doktorant do przedstawienia ostatniej proponowanej metody (rozdział 7) dopasowania obrazów USG uzyskiwanych w trakcie resekcji glejaka mózgu.

Ostatni rozdział pracy stanowi podsumowanie rozprawy wraz ze wskazaniem oryginalnych osiągnięć oraz planami dalszych prac badawczych.

Autor cytuje 192. pozycje literatury stanowiące aktualne i powiązane z tematyką pracy artykuły i inne zasoby literaturowe. W załączniku A do rozprawy autor przedstawił osiągnięcia publikacyjne i inne związane z tematyką rozprawy oraz pozostałe efekty prac badawczych. Rozprawa liczy 166 stron.

Rozprawa jest właściwie uporządkowana strukturalnie, a prezentowane treści są przedstawione w sposób czytelny i łatwy do lektury.

### 3. Oryginalne osiągnięcia i ocena merytoryczna rozprawy

Oryginalne osiągnięcia pracy omówione zostały w części II rozprawy. Ujęto je w formie trzech rozdziałów, każdy skupiony na innym problemie aplikacyjnym.

Do oryginalnych osiągnięć zaliczam:

- O1: Opracowanie złożonego algorytmu iteracyjnego, wykorzystującego kolejne przybliżenie dopasowania obrazów z zastosowaniem 1) detekcji cech, 2) dedykowanej transformacji afinicznej z uwzględnieniem brakujących danych i 3) własnego rozwinięcia algorytmu MIND Demons do problemu brakujących danych wraz z walidacją algorytmu dla obrazów preparatów histologicznych.
- O2: Opracowanie i walidacja złożonego modelu dopasowania obrazów poprzez propozycję zbioru modeli głębokich (splotowych) realizujących segmentację obrazów, dopasowanie na bazie przekształcenia afinicznego oraz dopasowanie obrazów realizującego przekształcenie „nie-szytywne” wraz z walidacją algorytmu dla obrazów preparatów histologicznych.
- O3: Opracowanie algorytmu z częściowym nadzorem w procesie dopasowania obrazów wraz z propozycją i weryfikacją czynnika regularyzacji (ang. volume penalty term) w funkcji celu, zdefiniowanego jako stosunek objętości guza po i przed dopasowaniem obrazu (po zabiegu nie ma guza zatem objętość ta powinna dążyć do zera) wraz z walidacją dla obrazów TK piersi.
- O4: Opracowanie metody rejestracji obrazów z wykorzystaniem podejścia uczenia przeciwstawnego, niewymagającego stosowania predefiniowanej miary podobieństwa i walidację dla śródoperacyjnego dopasowania obrazów 3D w technice USG z korektą danych dla usuwanego guza oraz głowicy USG.

Zaproponowane algorytmy są poprawnie przedstawione, zastosowana metodologia właściwa, a weryfikacja eksperymentalna wskazuje wysokie wartości miar jakości uzyskiwanych wyników.

Niemniej, oceniając pracę należy zwrócić uwagę na kilka aspektów:

- 1.) W omawianiu każdego z algorytmów autor wskazuje istotny aspekt oceny wzrokowej uzyskanych wyników (5.5.1, 6.4.1, 7.4.1). Szkoda, że doktorant nie przedstawił

wyników oceny przez doświadczonych lekarzy specjalistów, którzy w praktyce mogą być beneficjentami efektów prac mgr. inż. Marka Wodzińskiego.

- 2.) Brak przedstawienia w rozprawie porównania wyników prac dla O3 i O4 względem innych nowszych metod referencyjnych. Przykładowo dla problemu dopasowania obrazów USG (O4) prace zespołu Canalini L. et al., ([Canalini, L., Klein, J., Miller, D. et al. Enhanced registration of ultrasound volumes by segmentation of resection cavity in neurosurgical procedures. Int J CARS 15, 1963–1974 (2020). <https://doi.org/10.1007/s11548-020-02273-1>; publikacja 7.10.2020]) demonstrują redukcję średniego błędu TRE do poziomu 1,21mm.
- 3.) Zbiory danych wykorzystywane w O3 i O4 są mało liczne. Uzyskane modele nauczone zostały i zweryfikowane na mało licznych zbiorach, co wymagało precyzyjnego doboru hiperparametrów oraz znacznego (sztucznego) poszerzenia zbioru danych. Uzyskane zatem modele mogą okazać się (ale nie muszą) niezdolne do uogólnień przy analizie innych zestawów danych (np. uzyskiwanych z innych urządzeń diagnostycznych).

Opracowane algorytmy uważam za oryginalne osiągnięcie doktoranta i cenne w rozwoju metod automatycznego dopasowania obrazów medycznych. Autor zaprezentował posiadanie odpowiedniej wiedzy i umiejętności z zakresu uczenia maszynowego, przetwarzania obrazów i inżynierii biomedycznej. Liczne publikacje wyników prac związanych z osiągnięciami świadczą o docenieniu opracowanych rozwiązań w środowisku międzynarodowym.

#### 4. Uwagi szczegółowe i pytania

Praca jest napisana poprawnie i nie zawiera istotnych błędów. Niemniej chciałbym zwrócić uwagę na szereg drobnych elementów dotyczących zarówno metodologii jak i samej rozprawy.

W szczególności:

- Opis symboli przedstawiany pod wzorem warto prezentować według kolejności ich pojawiania się (np. wzór 2.1.).
- Stosowanie kategoriycznych stwierdzeń np. „The B-Splines Free Form Deformations (FFD) model is the most frequently used (...)” (str. 21) wymagałoby przedstawienia uzasadnienia poprzez np. cytowanie odpowiedniego źródła.
- Zapis  $p(\cdot)$  określający funkcje gęstości (masy) prawdopodobieństwa we wzorach 2.20 i 2.21 wydaje się nadmiarowy, gdyż z definicji entropii  $H(X)$  wynika, że takie funkcje są stosowane względem  $x$  (np. patrz Strudholme et. al., 1999).
- Warto stosować dodatkowe indeksy w stosowanych symbolach w celu rozróżnienia sytuacji występującej np., we wzorze 2.29:  $y(x)=y(x)*G(x, \sigma)$ .
- Na stronie 39 autor stwierdza: „The target registration error (TRE) is definitely the most common quantitative evaluation metric in IR [Sotiras et al., 2013 (...)]”. Ponownie doktorant posługuje się kategoriycznym stwierdzeniem, a cytowana praca przeglądowa z [IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 32, NO. 7, pp. 1153-1190, JULY 2013] w ogóle nie przedstawia w sposób jawny TRE, cytuje tylko powiązane prace

wskazując „Lastly, fiducial markers are also used to guide image registration. Some recent studies regarding the errors in the process are given in [284]–[286].”. Uważam zatem, że zbyt kateryczne stwierdzenia nie powinny być stosowane w pracy, chyba że są właściwie udokumentowane.

- W pracy stanowiącej spójne dzieło warto, aby każdy rozdział kończył się krótkim podsumowaniem, stanowiącym pomost pomiędzy kończącym się tematem a zagadnieniami przedstawianymi w kolejnej części.
- W części 3.2 poświęconej sieciom splotowym zabrakło mi omówienia warstw typu pooling i znaczenia pola odbiorczego neuronów (ang. receptive field), szczególnie w odniesieniu do zakresu deformacji w dopasowywanych obrazach.
- „Sztuczne” poszerzanie zbioru uczącego (ang. data augmentation) służy również, a może przede wszystkim, utworzeniu większego zróżnicowania w zbiorze uczącym w celu uzyskania większej zdolności do uogólnień uczonego modelu. Zabrakło w mojej ocenie podkreślenia tego aspektu w opisie na stronie 50.
- Podrozdział 3.3.2 wprowadza aspekty architektury sztucznych sieci neuronowych związane z zestawieniem kodera i dekodera. Autor stwierdza, że dekodery bazuje na „so-called transposed convolutions” co jest oczywiście prawdą dla sieci splotowych, ale ogólnie modele tego typu nie muszą bazować na warstwach splotowych.
- Stosowanie uniwersalnej legendy „P: padding/pooling” nie oznacza jednoznacznej interpretacji oznaczeń na ilustracjach (np. P=1 oznacza padding o rozmiarze 1, czy stosowanie operacji typu pooling? Jakiej?). Oczywiście można się tego domyślić z kontekstu i znajomości modeli bazowych, ale lepiej stosować jednoznaczne oznaczenia.
- Na stronie 114 wskazana jest referencja do tablicy 6.1, która znajduje się dopiero 5 stron dalej. Trochę utrudnia to lekturę, tym bardziej, że kolejne strony zawierają również opis tekstowy kolejnych zagadnień.

#### **Pytania do autora rozprawy:**

- P1: Proszę o wyjaśnienie jak przedstawiona definicja TRE we wzorze 2.32 ma się do stosowanej wcześniej miary Fiducial Registration Error (FRE) i dlaczego TRE jest obecnie częściej stosowaną miarą.

- P2: Proszę uszczegółwić procedurę normalizacji wartości w odniesieniu do przetwarzania wstępnego opisanego w 6.3.1. Czy normalizacji podlegały oryginalne dane pomiarowe (np. z DICOM – Hounsfield Units)? Jakże były dobrane okna obserwacji (ang. window level, window width)?

- P3: W punkcie 7.3.2 na stronie 125 autor podaje „The training set needs to be strongly augmented (...)”. Bardzo proszę o przedstawienie charakterystyki operacji augmentacji oraz rozmiaru zbioru z uwzględnieniem określonych rodzajów metod poszerzania danych.

- P4: Dla modelu nazwanego w rozprawie „affine registration network” proszę uszczegółwić proces wykorzystania mechanizmu uwagi.

## 5. Konkluzja recenzji

Podsumowując, stwierdzam, że przedstawione w rozprawie cele badawcze zostały osiągnięte, metody badawcze właściwie dobrane i zastosowane, a eksperymenty należycie przeprowadzone. Teza została potwierdzona.

Opracowane przez autora nowe metody i algorytmy dopasowania obrazów (medycznych) stanowią oryginalne osiągnięcie, a ich weryfikacja doprowadziła do wykazania, że zaproponowane metody charakteryzują się nowością oraz parametrami wyróżniającymi je spośród znanych rozwiązań.

Biorąc pod uwagę przedstawione wyżej wnioski z recenzji, uważam, że rozprawa mgr. inż. Marka Wodzińskiego spełnia wymagania sformułowane w Ustawie z dnia 20 lipca 2018 roku – Prawo o szkolnictwie wyższym i nauce, Dz. U. 2018, poz. 1668, z późniejszymi zmianami. Wnioskuje do Rady Dyscypliny Inżynieria Biomedyczna Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie o dopuszczenie mgr. inż. Marka Wodzińskiego do dalszych etapów postępowania w sprawie nadania stopnia naukowego doktora.

Jednocześnie, ze względu na istotny wpływ osiągnięcia doktoranta na rozwój nauki wyrażający się opublikowanymi pracami w wysoko punktowanych czasopismach (4) oraz w materiałach konferencyjnych (9) jak również kierowanie dwoma projektami NCN (w ramach konkursów Etiuda 7 i Preludium 15) wnioskuje o wyróżnienie rozprawy.



dr hab. inż. Jacek Rumiński, prof. PG  
Katedra Inżynierii Biomedycznej, ETI, PG