



Warszawa, dn. 23 sierpnia 2025

Dr hab. n. med. i n. o zdr, inż. Piotr Regulski
Kierownik Pracowni Obrazowania Cyfrowego i VR
Zakład Radiologii Stomatologicznej i Szczękowo-Twarzowej
Warszawski Uniwersytet Medyczny
Ul. Binieckiego 6, Warszawa

Recenzja rozprawy doktorskiej mgr Juliena Issy pt. „*Evaluation of Artificial Intelligence for Detecting Anatomical Landmarks and their Dimensions on Cone-beam Computed Tomography Scans*”

Rozprawa została przygotowana w Katedrze Praktycznej Stomatologii Klinicznej, Zakładzie Diagnostyki Obrazowej, Uniwersytetu Medycznego w Poznaniu pod opieką prof. dr hab. n. med. Marty Dyszkiewicz-Konwińskiej oraz dr. hab. n. med. Raphaëla Olszewskiego. Praca ma formę cyklu publikacji, obejmującego cztery artykuły: przegląd systematyczny piśmiennictwa oraz trzy prace oryginalne. Całość uzupełnia część syntetyczna zawierająca wprowadzenie, cele, ogólną dyskusję i wnioski, a także kopie publikacji wraz z oświadczeniami współautorów oraz oświadczenie Komisji Bioetycznej. Artykuły wchodzące w skład rozprawy posiadają łączny współczynnik oddziaływania (Impact Factor) na poziomie 10,3 i łączną punktację Ministra Nauki i Szkolnictwa Wyższego na poziomie 350 punktów.

Publikacje w cyklu są powiązane tematycznie, tworzą spójną całość i spełniają kryteria cyklu prac, będącego podstawą do nadania stopnia doktora nauk medycznych w macierzystej jednostce. Doktorant jest pierwszym autorem we wszystkich pracach wchodzących w skład cyklu, co potwierdza to Jego wiodący udział w przygotowaniu materiału i opracowaniu wyników badań. Ponadto Doktorant, zgodnie z oświadczeniami współautorów, odpowiadał za koncepcję, metodykę, analizę formalną/statystyczną zbiorów danych, opracowanie danych oraz przygotowanie manuskryptów.

W skład cyklu wchodzi artykuły opublikowane w recenzowanych czasopismach naukowych:

1. Issa J, Olszewski R, Dyszkiewicz-Konwińska M. The Effectiveness of SemiAutomated and Fully Automatic Segmentation for Inferior Alveolar Canal Localization on CBCT Scans: A Systematic Review. Int J Environ Res Public Health. 2022;19(1):560. doi:10.3390/ijerph19010560

IF = 3,3

Punkty MNiSW = 140



2. Issa J, Kulczyk T, Rychlik M, Czajka-Jakubowska A, Olszewski R, Dyszkiewicz-Konwińska M. Artificial intelligence versus semi-automatic segmentation of the inferior alveolar canal on cone-beam computed tomography scans: A pilot study. *Dent Med Probl.* 2024;61(6):893-899. doi:10.17219/dmp/175968

IF = 2,7

Punkty MNiSW = 70

3. Issa J, Riad A, Olszewski R, Dyszkiewicz-Konwińska M. The Influence of Slice Thickness, Sharpness, and Contrast Adjustments on Inferior Alveolar Canal Segmentation on Cone-Beam Computed Tomography Scans: A Retrospective Study. *J Pers Med.* 2023;13(10):1518. doi:10.3390/jpm13101518

IF = 3,4

Punkty MNiSW = 70

4. Issa J, Dyszkiewicz-Konwińska M, Kazimierczak N, Olszewski R. Assessing the accuracy of artificial intelligence in mandibular canal segmentation compared to semiautomatic segmentation on cone-beam computed tomography images. *Pol J Radiol.* 2025;90. doi:10.5114/pjr/202477

IF = 0,9

Punkty MNiSW = 70

W części wprowadzającej Doktorant klarownie prezentuje temat w szerszym kontekście radiologii stomatologicznej: przedstawia miejsce CBCT w praktyce klinicznej, rolę segmentacji kanału żuchwy dla bezpieczeństwa zabiegów chirurgicznych oraz potencjał metod sztucznej inteligencji (AI) w automatyzacji segmentacji kanału żuchwy. Autor wskazuje również na ograniczenia dotychczasowych badań, w tym różnorodność metod porównawczych i brak zunifikowanych metryk oceny. Zwraca uwagę, że wyniki segmentacji zależą od wielu czynników, takich jak warianty anatomiczne, jakość obrazu CBCT czy zmienność zależną od obserwatora. Cele rozprawy zostały przedstawione w sposób jednoznaczny i obejmują: przegląd systematyczny piśmiennictwa dotyczącego dokładności AI w segmentacji kanału żuchwy, porównanie AI z metodą półautomatyczną na małym zbiorze danych, ocenę wpływu parametrów obrazowania CBCT i ustalenie standardu segmentacji półautomatycznej, a także walidację AI na większej próbie klinicznej.

Pierwszy artykuł, stanowiący przegląd systematyczny przeprowadzony zgodnie z wytycznymi PRISMA, objął kwerendę pięciu baz literaturowych na dzień 22 sierpnia 2021 r. (uzyskano 990 rekordów, z których 7 spełniło kryteria włączenia). Autorzy stwierdzają, iż w pracach pierwotnych wykazywana była wysoka dokładność detekcji i segmentacji kanału żuchwy (czułość ok. 90%, swoistość ok. 95%), przy jednoczesnej heterogeniczności stosowanych metod walidacji, standardów



referencyjnych oraz raportowanych metryk, co znacząco utrudniało porównywanie algorytmów. Do oceny jakości badań wykorzystano narzędzie QUADAS-2, wskazując jednak na jego ograniczoną przydatność w odniesieniu do badań nad sztuczną inteligencją, ponieważ nie obejmuje ono specyficznych źródeł biasu algorytmicznego. Artykuł ten nie tylko syntetyzuje aktualny stan wiedzy, lecz przede wszystkim uzasadnia kolejne etapy cyklu badawczego: potrzebę standaryzacji metod referencyjnych oraz konieczność wyjścia poza wewnętrzne walidacje algorytmów.

Drugi artykuł przedstawia badanie pilotażowe obejmujące 15 badań CBCT (30 kanałów), w którym porównano segmentacje półautomatyczne wykonane w programie *Romexis* z segmentacjami AI uzyskanymi przy użyciu systemu *Diagnocat* (architektura sieci neuronowej typu „U-Net-like”). Do analizy zastosowano nakładanie modeli 3D oraz pomiar średniej odległości powierzchnia–powierzchnia (Geomagic), a istotność różnic oceniano testem Shapiro–Wilka oraz testem U Manna–Whitneya w odniesieniu do czasu segmentacji. Wyniki wskazują, że średnie odchylenie segmentacji AI względem metody referencyjnej wyniosło $0,275 \pm 0,475$ mm, natomiast czasy segmentacji były porównywalne (AI: $175,00 \pm 68,08$ s vs. półautomatyczna: $175,67 \pm 49,08$ s; $p = 0,389$). Autor rzetelnie odnotowuje przypadki niepełnej segmentacji kanału, podkreślając potrzebę badań na większej próbie i w szerszym spektrum przypadków. Trafnie zwraca również uwagę, że nawet niewielkie odchylenia w przebiegu kanału żuchwy mogą mieć znaczenie kliniczne, co stanowi istotne i właściwe spostrzeżenie.

Trzeci artykuł ma charakter metodologiczny i koncentruje się na fundamentalnym problemie zidentyfikowanym w przeglądzie systematycznym – braku jednolitej referencji dla segmentacji kanału żuchwy. Na podstawie 25 badań CBCT (50 kanałów) autor przeanalizował wpływ parametrów obrazowania w programie *Romexis (Planmeca)* – grubości warstwy, poziomu ostrości i kontrastu – na jakość i powtarzalność segmentacji półautomatycznej. Wyniki analizy superimpozycji potwierdziły wysoką zgodność ze średnią różnicą powierzchnia–powierzchnia $0,295$ mm, a zgodność ocen między pomiarami była wysoka. Choć problematyka może wydawać się techniczna, znaczenie uzyskanych wyników jest zasadnicze: praca ta definiuje i standaryzuje protokół przygotowania segmentacji referencyjnej w segmentacji półautomatycznej, który staje się niezbędnym punktem odniesienia dla dalszych porównań w artykule czwartym.

Czwarty artykuł stanowi ostatni etap cyklu badawczego i obejmuje walidację na materiale 150 badań CBCT (300 kanałów). Autor porównał segmentację opartą na sztucznej inteligencji (*Diagnocat*) z półautomatyczną segmentacją referencyjną, przygotowaną zgodnie ze standardem wypracowanym w artykule trzecim. Do oceny wykorzystano metryki odległości powierzchnia–powierzchnia oraz testy nieparametryczne (Kruskal–Wallis, U Manna–Whitneya), a wiarygodność segmentacji referencyjnej potwierdzono współczynnikami zgodności ICC (intra-rater 95,5%, inter-rater 84,5%). Wyniki wskazują na wysoką zgodność AI z referencją: mediana różnicy powierzchniowej wyniosła



0,29 mm, a 88% przypadków mieściło się w granicach klinicznej akceptowalności ($\pm 0,5$ mm). Analiza dodatkowa ujawniła jednak wrażliwość algorytmu na złożoność anatomiczną – obecność trzecich trzonowców, zwłaszcza zatrzymanych, prowadziła do wzrostu błędów segmentacji, co autor interpretuje jako efekt artefaktów typu beam hardening, nakładania się struktur oraz ograniczeń zbioru treningowego. Wniosek praktyczny jest jasny i wyważony: segmentacja AI osiąga poziom zgodności bliski referencji w typowych przypadkach i może stanowić realne wsparcie w planowaniu przedzabiegowym, jednak w sytuacjach złożonych wymaga nadzoru eksperckiego. Trafnie podkreślono również potrzebę walidacji wielośrodkowej oraz poszerzenia różnorodności zbiorów danych, co warunkuje dalszą wiarygodną implementację AI w praktyce klinicznej.

Znaczenie naukowe cyklu rozprawy należy ocenić wysoko. Po pierwsze, Doktorant bardzo świadomie przesuwa akcent z nieprecyzyjnego „AI działa/nie działa” na kwantyfikowalne pytanie „kiedy, w jakich warunkach i w jakiej relacji do metody referencyjnej AI osiąga akceptowalną zgodność”, a odpowiedzi dawkowane są w kolejnych artykułach w sposób konsekwentny. Po drugie, projekt badania czwartego – obejmujący dużą próbę, solidny standard odniesienia i analizę czynników anatomicznych – odpowiada na zarzuty podniesione w przeglądzie, w tym na brak spójnych protokołów walidacyjnych; wnioski mają zasięg praktyczny (kliniczne ustawienia progu akceptowalności) i metodologiczny (rola parametryzacji rekonstrukcji oraz kontroli jakości segmentacji). Po trzecie, Doktorant nie ucieka od tematu ograniczeń: wskazuje na wrażliwość modeli na przypadki o zwiększonej złożoności i na potrzebę trenowania na bardziej zróżnicowanych danych oraz prowadzenia badań wielośrodkowych – co czyni narrację dojrzałą, wolną od nieuzasadnionego entuzjazmu technologicznego.

Warto podkreślić, że dobór narzędzi i metryk analitycznych jest adekwatny do postawionych pytań. Zastosowanie oprogramowania Romexis do segmentacji półautomatycznej i CloudCompare do rejestracji oraz porównań 3D odpowiada współczesnym standardom analiz wolumetrycznych. Użyte testy statystyczne (Shapiro–Wilk, U Manna–Whitneya, Kruskala–Wallisa) oraz wskaźniki rzetelności (ICC, Kappa) dobrano racjonalnie i spójnie z charakterem danych. Taki wachlarz metod potwierdza kompetencje Doktoranta i jego dbałość o wiarygodność wyników. Należy jednak zaznaczyć, że w prezentacji wyników cennym uzupełnieniem mogłyby być dodatkowe metryki objętościowo-geometryczne stosowane powszechnie w ocenie segmentacji (np. współczynnik Dice’a, odległość Hausdorffa), które ułatwiłyby porównywalność z literaturą międzynarodową i wykraczałyby poza samą analizę odległości powierzchniowych.

Na koniec warto zauważyć, że znaczenie cyklu wykracza poza samą walidację AI. Autor aktywnie uczestniczy w międzynarodowej dyskusji nad standaryzacją raportowania badań w dziedzinie stomatologicznego computer vision, czego dowodem jest współautorstwo wytycznych Core Outcomes Measures in Dental Computer Vision Studies (DentalCOMS) (Journal of Dentistry, 2024).



Choć praca ta nie wchodzi bezpośrednio w skład cyklu, wzmacnia wiarygodność całego dorobku. Podkreślić należy również całościowy dorobek Doktoranta. Doktorant jest współautorem, łącznie 24 publikacji naukowych o łącznym współczynniku oddziaływania 79.4 i łączną punktacją Ministra Nauki i Szkolnictwa Wyższego na poziomie 1980 punktów.

Reasumując, rozprawa mgr Juliana Issy spełnia wymogi stawiane doktoratom w formie cyklu publikacji. Zestaw publikacji ukazał się w czasopismach recenzowanych o adekwatnej randze i łącznej punktacji, a część syntetyczna skutecznie łączy osobne artykuły w jedną narrację badawczą. W mojej ocenie przedstawiony cykl stanowi oryginalne osiągnięcie naukowe w obszarze zastosowań AI w radiologii stomatologicznej.

Na podstawie przedstawionej do recenzji rozprawy stwierdzam, że praca mgr Julien Issa „Evaluation of Artificial Intelligence for Detecting Anatomical Landmarks and their Dimensions on Cone-beam Computed Tomography Scans” spełnia warunki określone w art. 187 ustawy z dnia 20.07.2018 roku Prawo o Szkolnictwie Wyższym i Nauce (t.j. Dz.U. z 2018 r. poz 1668). W związku z tym przedstawiam Wysokiej Radzie Dyscypliny Nauk Medycznych Kolegium Nauk Medycznych Uniwersytetu Medycznego im. Karola Marcinkowskiego w Poznaniu wnioski o dopuszczenie mgr. Julien Issa do dalszych etapów przewodu doktorskiego. Ze względu na innowacyjny i nowatorski charakter pracy, wnioskuję o wyróżnienie.