dr hab. inż. Maciej Jaworski, prof. PK
Katedra Informatyki
Wydział Informatyki i Telekomunikacji
Politechnika Krakowska
Warszawska 24, 31-155 Kraków

Kraków, 03.11.2026

Recenzja rozprawy doktorskiej mgr inż. Michała Wieczorka pt. „Semi-Supervised Training with Fuzzy Logic for Transformer-Based Semantic Segmentation”

Podstawa sporządzenia recenzji

Recenzję rozprawy doktorskiej mgr inż. Michała Wieczorka pt. „Semi-Supervised Training with Fuzzy Logic for Transformer-Based Semantic Segmentation”, przygotowanej pod kierunkiem prof. dr hab. inż. Marcina Woźniaka, sporządziłem na podstawie umowy nr RN-UC-63/26 z dnia 30.07.2026, zawartej z Politechniką Częstochowską.

Recenzja dotyczy egzemplarza papierowego rozprawy. Numery stron w recenzji były podawane według tego egzemplarza.

1. Dane o kandydacie

Według materiałów przekazanych recenzentowi Kandydat uzyskał tytuł zawodowy inżyniera w 2022 roku, a tytuł magistra inżyniera w 2024 roku. Z materiałów tych nie wynika, aby ubiegał się wcześniej o nadanie stopnia doktora.

Od 2024 roku jest asystentem w Katedrze Modelowania Sztucznej Inteligencji na Wydziale Matematyki Stosowanej Politechniki Śląskiej, a równolegle pracuje w przemyśle jako Tech Lead przy symulacjach fizycznych w grafice 3D. Od 2025 roku

kieruje projektem PN/02/0084/2023 „Semantyczno-segmentacyjne sieci transformatorowe w opisie zrównoleglonego uczenia adaptacyjnego”. W roku 2024 otrzymał zespołową nagrodę Rektora II stopnia za osiągnięcia naukowe.

Jak na ten etap kariery Doktorant ma wyjątkowo bogaty dorobek: według dokumentacji postępowania indeks Hirscha wynosi 12 w Web of Science (830 cytowań) oraz 13 w Scopus (1123 cytowania) i Google Scholar (1517 cytowań), sumaryczny IF 74,98. Mgr inż. Michał Wieczorek publikował prace m.in. w IEEE Transactions on Industrial Informatics, Neural Computing and Applications, Future Generation Computer Systems, Chaos, Solitons & Fractals i International Journal of Applied Mathematics and Computer Science. Współpracuje z Uniwersytetem Przyrodniczym w Lublinie przy analizie obrazów weterynaryjnych.

2. Ocena rozprawy

Na wstępie chcę podkreślić najmocniejsze strony pracy. Doktorant podjął ważny i praktycznie istotny problem segmentacji rozmazów krwi zwierzęcej przy bardzo małej liczbie adnotacji eksperckich, zbudował kompletny, działający system i zastosował go przy zaledwie 80 maskach eksperckich. Wspólnie ze środowiskiem weterynaryjnym przygotował nowe zbiory danych, część badań opublikował, a praca pokazuje szeroki warsztat: od sieci konwolucyjnych, przez transformery i syntetyzowanie danych, po logikę rozmytą typu 2. Uwagi krytyczne, w części poważne, dotyczą przede wszystkim udokumentowania i ewaluacji wyników rozdziału 9.

2.1 Tytuł i układ rozprawy

Rozprawa, napisana po angielsku, liczy 129 stron tekstu zasadniczego i 142 pozycje bibliografii. Tytuł trafnie oddaje treść rozdziału 9, który jest jej rdzeniem. Po trzech rozdziałach wprowadzających (wstęp, słowniczek, podstawy teoretyczne; s. 1–18) Autor przedstawia pięć studiów przygotowawczych: segmentację rozmazów z malarią (rozdz. 4), transformer TSSN do segmentacji krętków (rozdz. 5), segmentację broni na danych syntetycznych (rozdz. 6), rozmytą reputację klientów w uczeniu federacyjnym (rozdz. 7) i klasyfikację tkanek przewodu pokarmowego zwierząt (rozdz. 8). Rozdział 9 (s. 90–

121) przedstawia półnadzorowany, transformerowy system segmentacji rozmazów krwi z rozmytymi funkcjami straty, a rozdziały 10–11 zawierają dyskusję i weryfikację tez.

Taki układ pracy dobrze dokumentuje drogę badawczą, jaką przeszedł Doktorant dochodząc do finalnej postaci proponowanego systemu. Rozdziały 4, 5 i 8 tworzą spójną linię prowadzącą do rozdziału 9, natomiast rozdziały 6 i 7 mają z tematem związek luźny i z korzyścią mogłyby je zastąpić bliższe tematowi prace Doktoranta (zob. uwagi dodatkowe przed konkluzją).

2.2 Piśmiennictwo

Bibliografia jest obszerna i stanowi szeroki przegląd podejmowanej tematyki. Obejmuje kluczowe architektury segmentacyjne i transformerowe (m.in. U-Net, nnU-Net, TransUNet, SegFormer, Swin-Unet), metody półnadzorowane (Mean Teacher, FixMatch, CPS, U2PL) oraz rozmyte metody segmentacji, w tym pracę Chena i in. [127], z której pochodzą funkcje straty użyte jako bazowe w doktoracie. Brakuje jednak prac najbliższych zaproponowanemu mechanizmowi uwagi (Pyramid Vision Transformer v1 i v2) i filtracji w dziedzinie Fouriera (np. GFNet), co utrudnia ocenę nowości bloku HybridGlobalAttention. Zbiory z rozdziałów 4 i 5 (s. 23 i 38) nie mają pozycji w bibliografii, chociaż w treści pracy podano ich nazwy. Dla zbioru z rozdziału 9 także nie podano źródła. Nie zacytowano również wcześniejszej pracy Doktoranta na zbiorze z rozdziału 5 (IVUS 2022). Rozdział 3, będący wstępem teoretycznym pracy, nie zawiera żadnych odwołań. W części tabel porównawczych (np. tab. 5 i 24) zestawiono wyniki uzyskane na innych zbiorach danych i wyciągnięto z nich wnioski o przewadze proponowanych rozwiązań (np. s. 44).

2.3 Cel pracy i tezy

Celem rozprawy jest opracowanie metod analizy obrazów medycznych i biologicznych, które są dokładne, oszczędne w danych, wykonalne obliczeniowo i odporne na niepewność (s. 125). Jest on istotny z praktycznego punktu widzenia i dobrze umotywowany.

Według trzech tez (s. iii) mechanizmy uwagi z kompresją w przestrzeni ukrytej mogą poprawić uogólnianie transformerowej segmentacji próbek krwi zwierząt, rozmyte uczenie półnadzorowane może poprawić klasyfikację pikseli w warunkach niejednoznacznych danych, a logika rozmyta może zmniejszyć liczbę błędów na danych niejednoznacznych i na granicach obiektów. Tezy dają się zweryfikować eksperymentalnie, jednak mają formę możliwościową („can improve”) i nie wskazują, względem czego ma nastąpić poprawa. Teza 1 dotyczy wprost analizy próbek krwi zwierząt, a tezy 2 i 3 są sformułowane ogólniej, jednak jedynym rozwiązaniem w rozprawie realizującym ich przedmiot jest system z rozdziału 9. Autor sam wskazuje go jako źródło bezpośredniej weryfikacji tez 1 i 2 (s. 125–126), a rozdziały 4–8 mogą ją co najwyżej pośrednio wspierać.

2.4 Metody badawcze

System z rozdziału 9 łączy normalizację barwienia (Alg. 8), hierarchiczny enkoder z blokiem HybridGlobalAttention, w którym klucze i wartości są kompresowane do 65, 37 i 17 tokenów i filtrowane w dziedzinie częstotliwości (s. 98–100, Alg. 10), dekoder z modułem MedicalContextRefiner oraz uczenie na mieszanych partiach z nadzorowaną stratą rozmytą typu FCM i Dice oraz nienadzorowaną stratą RFCM+ (Alg. 11–13). Doceniam przemyślaną integrację tych elementów, niewielki rozmiar modelu (6,9 mln parametrów, s. 102), pseudokod i schemat dla każdego składnika oraz staranny opis procedury uczenia, w tym hiperparametrów i ustawień sprzętowych (s. 105, tab. 23). Wartościowe są też generator danych syntetycznych w środowisku Blender (rozdz. 6) i procedura budowy zbioru histologicznego (rozdz. 8, s. 76–77).

Mimo to odtworzenie samego modelu byłoby trudne, bo jego kluczowe elementy są niedookreślone: filtrację i decymację FFT opisano tylko słownie (Alg. 10), nie zdefiniowano parametru τ w równ. (16) ani modyfikacji straty z pracy [127] (s. 102), a podpisy rys. 12–14 są sprzeczne z algorytmami (szczegóły w sekcji o nieprawidłowościach). Rozprawa nie informuje też o dostępności kodu.

Szkielet sieci odpowiada architekturze SegFormer-B0, redukcja kluczy i wartości jest bliska mechanizmowi z Pyramid Vision Transformer v2, a funkcje straty pochodzą z

pracy [127]. Autor wspomina wprawdzie o koncepcji architektur typu SegFormer (s. 93), ale nie pokazuje, jak bliskie są te odpowiedniki. Wkład metodyczny ma więc przede wszystkim charakter integracyjny i aplikacyjny, co należało powiedzieć wprost.

2.5 Wyniki badań i ich omówienie

Wyniki studiów przygotowawczych przedstawiono w większości poprawnie: z wydzielonym zbiorem testowym (rozdz. 4 i 5, w tym IoU 0,842 dla rzadkiej klasy krętków), z walidacją krzyżową i ablacją (rozdz. 8, 98,12% ± 0,23%) oraz z rozbudowaną ablacją danych syntetycznych (rozdz. 6, tab. 8). W rozdziale 9 Autor raportuje przy 80 maskach 99,73% dokładności pikselowej i micro-Dice oraz 99,46% micro-IoU (s. 106), przykłady na rys. 17–21 wyglądają dobrze, a wstępne zastosowanie do wykrywania Babesia pokazuje praktyczny potencjał metody. Doceniam też, że Autor sam wskazuje ograniczenia dokładności pikselowej (s. 106), odróżnia porównanie eksperymentalne od zestawień z literaturą (s. 108) i wymienia potrzebną dalszą walidację (s. 129).

Omówienie głównego eksperymentu ma jednak istotne luki. Na s. 107–108 Autor zapowiada analizę przy stopniowo zmniejszanej liczbie etykiet oraz porównanie z wariantem w pełni nadzorowanym i z architekturami takimi jak U-Net, nnU-Net, TransUNet i UNETR. W rozprawie nie ma jednak liczbowych wyników żadnego z tych porównań na tych samych danych. Wariant w pełni nadzorowany, uczony wyłącznie na oznaczonym podzbiorze, porównano jedynie jakościowo, na przykładowych obrazach (rys. 21), a Autor stwierdza na tej podstawie, że daje on „much worse results” (s. 109). Pozostałe architektury omówiono tylko słownie (s. 109–112) i zestawiono z wynikami z literatury uzyskanymi na innych danych (tab. 24). Nie pozwala to uzasadnić deklarowanej na s. 112 przewagi zaproponowanego rozwiązania. Sekcja 10.6 „Statistical results and ablation studies” (s. 124) również opisuje wyniki ablacji wyłącznie słownie, bez żadnej liczby, a jedyny opisany efekt usunięcia etapu FFT to wydłużenie czasu uczenia, bez zmiany jakości segmentacji (s. 124, 125–126). Taki wynik uzasadnia FFT względami obliczeniowymi, ale nie wspiera tezy 1 o poprawie uogólniania.

Zbiór walidacyjny służył do sterowania harmonogramem uczenia, wyboru punktu kontrolnego (s. 104 oraz Alg. 13) i doboru konfiguracji (s. 123), a następnie na nim raportowano wynik końcowy (s. 106, 113 i 125). Zbioru testowego nie wydzielono, a podziału danych nie opisano. Metryki mikro są zdominowane przez tło, a mIoU i IoU dla poszczególnych klas, zapowiedziane na s. 106 i deklarowane jako raportowane w tab. 24 (s. 111), nie zostały nigdzie podane. Wyniku 99,73% nie można zatem traktować jako miarodajnej oceny jakości modelu. O segmentacji samych komórek mówi on niewiele. Przywołane w rozdz. 11.1 dodatkowo wyniki z rozdziałów 6 i 7 (s. 126) mają ograniczony związek z przedmiotem tez 2 i 3, a argumenty odwołujące się do rozdziału 9 są wyłącznie słowne. Zysków na granicach obiektów (s. 126) nie zmierzono żadną metryką brzegową. Jednocześnie na s. 108 Autor przyznaje, że pozostałe błędy koncentrują się właśnie na brzegach obiektów, i określa je jako mniej istotne w praktyce, co trudno pogodzić z tezą 3. Wyniki pochodzą z pojedynczych uruchomień, a w rozdziale 5 metodę Autora dostrajano optymalizacją bayesowską (s. 41), modeli bazowych zaś nie.

W rezultacie rozprawa pokazuje, że system działa przy skrajnym niedoborze etykiet, ale przewagę nad alternatywami i wpływ składników wykazano tylko jakościowo. Tezy uznaję więc za wykazane częściowo.

2.6 Praktyczne zastosowanie wyników

Wyniki mają wyraźnie aplikacyjny charakter: system segmentacji rozmazów krwi uczony na kilkudziesięciu maskach, jego zastosowanie do wykrywania Babesia, istotnej klinicznie u psów w Polsce (s. 107), oraz klasyfikator pochodzenia tkanek z rozdziału 8. Szczególnie doceniam to, że Doktorant we współpracy z profesjonalistami zbudował własne zbiory danych: zbiór Babesia (ponad 1000 obrazów, ponad 100 oznaczonych), powstały z pomocą naukowców z Uniwersytetu Przyrodniczego w Lublinie i lekarzy weterynarii (s. 106–107) i określony przez Autora jako drugi zbiór przygotowany z weterynarzami, oraz wielogatunkowy zbiór histologiczny (3680 obrazów, 25 klas; s. 76–77). To praca żmudna, a takich danych w weterynarii bardzo brakuje. Do zastosowań

klinicznych potrzebna jest jeszcze, walidacja zewnętrzna i badania z udziałem ekspertów, a także analiza błędów istotnych klinicznie.

2.7 Nieprawidłowości

W rozprawie występują nieprawidłowości, które wymagają wyjaśnienia przez Doktoranta. Najpoważniejsza dotyczy danych z rozdziału 9: podano jedynie liczbę obrazów i masek, rozdzielczość i liczbę klas (s. 90 i 95). Z kontekstu (s. 107) można wnioskować, że to zbiór własny, ale nie opisano jego źródła, gatunków, barwienia, nazw i udziału klas, protokołu adnotacji ani podziału. Gatunki i barwienie pojawiają się tylko mimochodem (feline, canine, s. 91, 122 i 124). W rozdziałach 5, 6 i 8 występują ponadto niezgodności liczbowe. Najistotniejsza dotyczy rozdziału 8: liczba przykładów uczących „3680 × 201 ≈ 740,000” (s. 78) i liczba iteracji na epokę (ok. 11 562 przy partii 64, s. 82) odpowiadają augmentacji wszystkich 3680 obrazów, choć 15% z nich miało stanowić nienaruszony zbiór walidacyjny (s. 82–83). Jeśli augmentowane kopie obrazów walidacyjnych trafiły do zbioru uczącego, wyniki tego rozdziału mogą być zawyżone; wymaga to wyjaśnienia.

Poniżej, rozdział po rozdziale, wymieniam pozostałe zauważone błędy oraz niesłuszne lub niedokładne sformułowania.

Rozdziały 1–3. Na s. 3 Autor zapowiada, że przynależności rozmyte posłużą do udoskonalania pseudoetykiet i poprawy kalibracji, tymczasem rozdział 9 nie korzysta z pseudoetykiet (s. 94), a kalibracji nigdzie nie mierzono. Rozdział 3 (s. 16–18), który jest wstępem teoretycznym do pracy, nie zawiera żadnych odwołań do literatury.

Rozdział 4. Maski wygenerowane heurystycznie z ramek nazwano „high-quality, pixel-level ground truth” (s. 25) i „critical and novel component” (s. 24), bez odniesienia do znanych w literaturze metod segmentacji z ramek (np. BoxSup, BoxInst). Miara 99,68% (s. 29) opiera się na porównaniu liczby wykrytych komórek z liczbą komórek w adnotacjach, ale nie podano, jak z tego porównania wyznaczono dokładność. Zgodność liczebności nie wyklucza przy tym wzajemnie znoszących się błędów.

Rozdział 5. Enkoder z 296 blokami transformera (s. 40) wymaga podania liczby parametrów i uzasadnienia przy zbiorze liczącym 366 obrazów. Wiersz „Full TSSN” w tab. 4, opisanej jako ablacje na zbiorze walidacyjnym (s. 42), jest co do trzeciego miejsca po przecinku identyczny z wynikami na zbiorze testowym z tab. 3. Wymaga wyjaśnienia, na jakim zbiorze faktycznie przeprowadzono ablacje.

Rozdział 6. Recall wynosi 100% w tab. 7 i 99,54% w tab. 9 (s. 58–59) przy tej samej precyzji 98,66% i tym samym F1 = 0,9932, choć przy recall 99,54% F1 wynosiłby ok. 0,991. Wiersz G w tab. 8, przedstawiony jako „the most compelling evidence” przewagi danych syntetycznych (s. 59, także s. 126), dotyczy modelu uczonego na danych rzeczywistych, ale ocenianego na syntetycznym zbiorze walidacyjnym, a więc w innej domenie; nie dowodzi on przewagi danych syntetycznych na obrazach rzeczywistych. Wniosek o „impressive generalization” na obrazach z internetu (s. 60) nie jest poparty żadnym przykładem ani informacją o liczbie i doborze obrazów.

Rozdział 7. Druga nierówność w dowodzie twierdzenia 1 (s. 70) zachodzi tylko dla klientów, których liczba danych nie przekracza średniej, podczas gdy w eksperymentach (s. 70) liczności klientów mają rozkład potęgowy, a więc największy klient ma ich wyraźnie więcej. Ponadto uzyskane ograniczenie nie jest skończone, gdy r_min może być równe zeru. Tab. 11 (s. 70) zapowiada trzy zbiory danych, a wyniki (s. 72) dotyczą tylko CIFAR-10.

Rozdział 8. Warto sprawdzić podaną na s. 81 liczbę 1 162 809 parametrów, gdyż z architektury opisanej w tab. 18 wynika raczej liczba ok. 4-krotnie mniejsza. Sformułowanie „SGD with the NAdam optimiser” jest wewnętrznie sprzeczne, a współczynnik kappa Cohena, przyjmujący wartości od −1 do 1, podano w zapisie procentowym („98.30”, tab. 19 i s. 127). Wynik główny i ablacja pochodzą z pojedynczego uruchomienia.

Rozdział 9. W równ. (16) na s. 99 nie zdefiniowano τ (nie wiadomo m.in., czy czynnik $\exp(\tau)$ zastępuje standardowe skalowanie $1/\sqrt{d}$), a skompresowane klucze i wartości oznaczono tam K' i V', podczas gdy w Alg. 10 (s. 100, linie 8–9) oznaczono je K i V. Podpisy trzech rysunków opisują działanie modułów inaczej niż odpowiadające im algorytmy. Według podpisu rys. 12 (s. 101) zapytanie jest kompresowane („projecting

the novel query into a smaller, compressed latent space”) i dzielone na gałąź lokalną i globalną, podczas gdy w Alg. 10 (s. 100) zapytanie powstaje bez kompresji (linia 2), kompresowane są tylko klucze i wartości (linie 7–8), a gałąź lokalna jest liczona z X_pos, nie z zapytania (linia 10). Podpis rys. 13 (s. 103) mówi o „fuzzy type-2 logic operators, which are learned during training”, tymczasem w Alg. 11 (s. 102) przynależność to μ ← P^m (linia 7), bez logiki typu 2 i bez parametrów uczonych. Podpis rys. 14 (s. 114) określa przynależności jako „trainable parameters” i mówi o zwracanym „loss vector”, podczas gdy w Alg. 12 (s. 104) przynależności są wyliczane z wyjścia sieci, μ_k ← P_k^m (linia 6), a zwracana jest pojedyncza wartość straty (linia 14). Strata L_FCM-sup w równ. (18) nie ma definicji, a sposoby obliczania składników J_1 i J_2 (s. 103) pojawiają się dopiero na słabo czytelnym rys. 16 (s. 116). Na stronie 108 Autor uznaje pozostałe błędy, skupione na brzegach obiektów, za mało istotne w praktyce. Warto to zestawić z s. 93, gdzie struktura granic jest wskazywana jako podstawa decyzji klinicznych, oraz z tezą 3, która dotyczy właśnie redukcji błędów na granicach. Powołanie się na „comparative discussion and ablation analysis” (s. 113) nie ma oparcia w przedstawionych wynikach.

Rozdziały 10 i 11. Zadanie “Removing the FFT stage restored the quadratic cost” (s. 124) budzi wątpliwości, gdyż zgodnie z Alg. 10 FFT działa na pamięci już zredukowanej do 65/37/17 tokenów, więc złożoność kwadratową względem liczby pikseli usuwa sama ta redukcja (równ. (17)). Sformułowanie „sufficient to treat all three stated theses as proved” (s. 127) jest zbyt mocne wobec przedstawionych dowodów.

2.8 Oryginalność rozwiązania problemu naukowego

Uważam, że rozprawa zawiera oryginalne rozwiązanie problemu naukowego, o charakterze przede wszystkim integracyjnym i aplikacyjnym. Prezentuje połączenie efektywnej uwagi globalnej, rozmytych funkcji straty i uczenia na mieszanych partiach w systemie do segmentacji rozmazów krwi zwierzęcej przy skrajnym niedoborze etykiet, wraz z zastosowaniem do wykrywania Babesia i nowymi zbiorami danych. Składniki mają bliskie odpowiedniki w literaturze (SegFormer, PVT v2, praca [127]), a przewagi ich połączenia nie wykazano liczbowo, dlatego stwierdzenie o „superiority … at the architectural level” (s. 112) uważam za słabo uzasadnione.

2.9 Wiedza teoretyczna i umiejętność samodzielnego prowadzenia pracy naukowej

Rozprawa świadczy o szerokiej znajomości głębokiego uczenia, metod półnadzorowanych i rozmytych oraz o umiejętności ich stosowania w różnych dziedzinach. Słabiej natomiast wypada skrótowy rozdział teoretyczny, bez cytowań z literatury. O samodzielności najlepiej świadczy dorobek (zob. punkt 1), imponujący jak na badacza, który tytuł magistra uzyskał w 2024 roku. Znajdują się w nim prace, w których mgr inż. Michał Wieczorek jest pierwszym autorem, m.in. w IEEE Transactions on Industrial Informatics, Neural Computing and Applications i International Journal of Applied Mathematics and Computer Science. Na podkreślenie zasługuje też fakt kierowania własnym projektem oraz nawiązanie współpracy ze środowiskiem weterynaryjnym. Oceniam, że rozprawa wraz z dorobkiem potwierdza ogólną wiedzę teoretyczną kandydata w dyscyplinie i umiejętność samodzielnego prowadzenia pracy naukowej.

3. Pytania do Doktoranta i uwagi dodatkowe

Proszę Doktoranta o odniesienie się podczas obrony do następujących kwestii:

1. Skąd pochodzą dane z rozdziału 9? Jakich gatunków dotyczą, jak barwiono i rejestrowano preparaty, jak nazywają się klasy i jaki mają udział pikselowy? Kto i według jakiego protokołu wykonał maski?

2. Jak podzielono dane na część uczącą i walidacyjną i na jakim poziomie (zwierzęcia, rozmazu, obrazu)? Jakie wyniki osiąga model na danych, które nie brały udziału w wyborze punktu kontrolnego i hiperparametrów, z IoU i Dice dla poszczególnych klas? Czy w rozdziale 8 augmentację wykonano przed podziałem danych, czy po nim?

3. Jakie są liczbowe wyniki porównań na tych samych danych z modelami zapowiedzianymi na s. 108 (U-Net, nnU-Net, TransUNet, UNETR), z wariantem w pełni nadzorowanym oraz z nierozmytą metodą półnadzorowaną (np. Mean Teacher lub CPS)?

4. Jaki jest liczbowy wpływ poszczególnych składników (straty RFCM+, bloku HybridGlobalAttention, liczby etykiet) i na czym matematycznie polegają modyfikacje straty z pracy [127]?

5. Na czym dokładnie polega filtracja i decymacja w dziedzinie FFT w bloku HybridGlobalAttention i jak zdefiniowano parametr τ w równ. (16)?

6. Dlaczego modele uczono od zera, zamiast wykorzystać dostępne modele wstępnie wytrenowane (np. SegFormer-B0 z wagami z ImageNet)? Czy porównywano oba podejścia?

7. Na jakiej podstawie stwierdzono, że największe zyski z zastosowania logiki rozmytej wystąpiły na granicach obiektów (s. 126), skoro pozostałe błędy koncentrują się właśnie na brzegach (s. 108)? Czy mierzono metryki brzegowe, np. boundary IoU lub odległość Hausdorffa (HD95)?

8. Czy kod źródłowy i zbiory danych z rozdziału 9 są lub będą publicznie dostępne?

4. Konkluzja recenzji

Rozprawa podejmuje ważny i istotny z praktycznego punktu widzenia problem segmentacji obrazów mikroskopowych przy niedoborze adnotacji eksperckich, w diagnostyce weterynaryjnej, gdzie narzędzi automatycznych wciąż brakuje. Doktorant zbudował spójny system łączący efektywną uwagę globalną z rozmytym uczeniem półnadzorowanym, który przy 80 maskach daje obiecujące wyniki. Przygotował we współpracy z weterynarzami nowe zbiory danych i opublikował część badań. Na wyraźne podkreślenie zasługuje jego dorobek, imponujący jak na ten etap kariery.

Recenzja zawiera istotne zastrzeżenia: niepełny opis danych użytych w rozdziale 9, raportowanie wyników na zbiorze walidacyjnym bez wydzielonego zbioru testowego oraz brak liczbowych wyników porównań z innymi metodami na tych samych danych i liczbowych wyników ablacji. W tej postaci rozdział 9 nie dostarcza wystarczających dowodów ilościowych dla tez, dlatego uznaję je za wykazane częściowo. Zastrzeżenia te dotyczą jednak przede wszystkim udokumentowania i ewaluacji wyników, a nie samej

koncepcji, i nie podważają oryginalności rozwiązania. O pozytywnej ocenie przesądzają oryginalność koncepcji, wkład w budowę zbiorów danych oraz dorobek Kandydata.

Mimo tych zastrzeżeń, oceniając pracę jako całość, wraz z dorobkiem Doktoranta, stwierdzam, że rozprawa doktorska mgr inż. Michała Wieczorka pt. „Semi-Supervised Training with Fuzzy Logic for Transformer-Based Semantic Segmentation” stanowi oryginalne rozwiązanie problemu naukowego oraz wykazuje ogólną wiedzę teoretyczną Kandydata w dyscyplinie informatyka techniczna i telekomunikacja, a także umiejętność samodzielnego prowadzenia pracy naukowej. Spełnia tym samym wymagania określone w art. 187 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce. Wnoszę o dopuszczenie rozprawy do publicznej obrony.