

# Recenzja

**rozprawy doktorskiej mgra inż. Tomasza Krzywickiego**

pod tytułem:

**“Metody Rozpoznawania Obrazów W Badaniu Dna Oka Dla Celów Diagnostyki Medycznej”**

przygotowanej pod kierunkiem promotora – Profesora dr hab. **Lecha Polkowskiego**

## Podstawa faktyczna recenzji

Recenzja sporządzona została na podstawie oceny rozprawy doktorskiej przedłożonej przez Pana mgra inż Tomasza Krzywickiego. Praca liczy 225 stron, a w tym spis treści oraz bibliografia (ponad 200 pozycji).

## Teza rozprawy doktorskiej

Autor przedłożonej rozprawy wskazywał cztery główne cele:

1. Pokazanie, że nowoczesna technika oparta o sieci neuronowe i uczenie transferowe będzie lepsza niż tradycyjne techniki przetwarzania obrazów w lokalizacji cech anatomicznych na zdjęciach dna oka pod względem dokładności i pokrycia.
2. Pokazanie, że czas wykonania obliczeń i zużycie pamięci operacyjnej będzie niższe w przypadku lokalizacji cech anatomicznych na zdjęciach dna oka przy użyciu technik przetwarzania obrazów, niż w przypadku zastosowania sieci neuronowych.

3. Pokazanie, że zastosowanie przyspieszania sprzętowego zaproponowanych metod za pomocą zrównoleglania obliczeń na procesorach graficznych skróci czas obliczeń potrzebny na wyznaczenie lokalizacji cech anatomicznych na zdjęciach dna oka.
4. Pokazanie, że dokładność i pokrycie lokalizacji cech anatomicznych na zdjęciach dna oka przy użyciu zespołu sieci neuronowych będą wyższe, niż w przypadku pojedynczej sieci neuronowej.

## **Wybór tematu rozprawy doktorskiej**

Tematem przedłożonej rozprawy było opracowanie i zaproponowanie nowych metod lokalizujących dołeczek (punkt centralny plamki żółtej) oraz tarczę nerwu wzrokowego na zdjęciu dna oka. Do opracowania proponowanych metod wykorzystano dwa podejścia:

- wykorzystujące techniki przetwarzania obrazów oraz
- wykorzystujące sieci neuronowe.

Przy opracowywaniu proponowanych metod uwzględniono aspekty takie jak dokładność i pokrycie lokalizacji cech anatomicznych będących przedmiotem zainteresowania, efektywność obliczeniową w postaci zużytego czasu i pamięci i możliwości przyspieszania wykonania za pomocą rozwiązań sprzętowych. Realizacja badań miała na celu również dostarczenie wkładu w rozwój diagnostyki okulistycznej wspieranej za pomocą AI oraz poszerzenie bieżącego stanu wiedzy naukowej.

W mojej ocenie, wybór tematu pracy był bardzo trafny. Podjęta problematyka rozprawy jest aktualna i ważna dla zastosowania nowoczesnych technik sztucznej inteligencji w medycynie.

## **Zakres analizy w pracy**

Stwierdzam, że zakres analizy jest zgodny z tematem pracy. Pan mgr inż. Tomasz Krzywicki opracował dwie metody wykrywania lokalizacji współrzędnych środków dołeczka i tarczy nerwu wzrokowego na zdjęciach dna oka: metoda MSiRL wykorzystująca techniki przetwarzania obrazów ze specjalnie dobranymi parametrami, dobranymi przez eksperta, oraz metoda MSLW oparta o technikę detekcji obiektów na zdjęciach z użyciem głębokich sieci neuronowych (konkretnie sieci YOLOv8).

Autor rozprawy poświęcił dużą uwagę na efektywność pod względem obliczeniowym proponowanych rozwiązań. Przedstawił m.i. dwie implementacje: wersja sekwencyjna i

wersja równoległa dla algorytmu MSiRL. W obu przypadkach, obliczenia równoległe zostały realizowane za pomocą GPU na platformie NVIDIA Cuda.

Opracowane metody zostały przetestowane na dwóch zbiorach danych: baza Fundacji Okulistyka 21 (BFO21) oraz Baza zdjęć dna oka Indian Diabetic Retinopathy Image Dataset (IDRID). Każda opracowana metoda była testowana pod względem czasu obliczeń, dokładności i pokrycia lokalizacji cech anatomicznych.

## **Struktura rozprawy doktorskiej**

Główna część przedłożonej rozprawy zawiera siedem rozdziałów oraz cztery dodatki (A, B, C, D).

Pierwszy rozdział opisuje główny problem badawczy rozprawy, t.j. diagnostyka retinopatii cukrzycowej. Jest to jednym z głównych zastosowań sztucznej inteligencji w okulistyce. W tym rozdziale oprócz opisu celu, tezy i głównych wyników rozprawy, autor przedstawił również listę programów i bibliotek zewnętrznych w języku Python oraz zawartość pracy.

Rozdział drugi składa się z dwóch części: przeglądu tradycyjnych i współczesnych technik w sztucznej inteligencji, oraz omówienia jej zastosowań jako pomocy w diagnostyce okulistycznej. Omówiono schemat budowy anatomicznej i proces diagnostyczny, a także techniki obrazowania dna oka. Rozdział zamyka podkreślenie znaczenia plamki żółtej i tarczy wzrokowej w identyfikacji strony oka, stawianiu diagnozy klinicznej oraz ocenie standardu i jakości obrazu.

W trzecim rozdziale omówiono metody i ogólne algorytmy przetwarzania i rozpoznawania obrazów, które posłużyły do opracowania proponowanych rozwiązań lokalizacji środków dołka i tarczy wzrokowej. Autor omówił pojęcie złożoności obliczeniowej algorytmów oraz koncepcję obrazu cyfrowego wraz z modelami kolorów, takimi jak RGB, CIELab i YUV. Następnie, omówiono kilka kluczowych technik przetwarzania obrazu, w tym filtrowanie obrazu, binaryzację, wyrównywanie histogramu, określanie wypukłych powłok konturów wielokątów i analizę komponentów połączonych. Kolejnym zestawem algorytmów są algorytmy topologiczne oraz algorytmy inspirowane naturą, takie jak sieci neuronowe i morfologia matematyczna. Pod koniec omówiono współczesne podejścia do głębokiego uczenia się, identyfikacji obiektów, uczenia się transferowego i rozpoznawania obrazów, metody uczenia zespołowego oraz techniki przyspieszania algorytmów uczących się.

Zbiory danych zawierające obrazy dna oka są tematem czwartego rozdziału „Bazy danych wykorzystane w badaniach” i zostały wykorzystane do opracowania sugerowanych technik identyfikacji dołka i tarczy wzrokowej na zdjęciach dna oka. Oprócz zapewnienia unikalnej technologii umożliwiającej automatyczną ocenę porównywalnych zbiorów danych dostępnych online, w rozdziale tym szczegółowo opisano dwa zbiory danych.

Rozdział piąty przedstawia i szczegółowo omawia techniki identyfikacji środków tarczy i dołka wzrokowego. Rozdział rozpoczął się od przeglądu obecnego poziomu wiedzy naukowej oraz wyników uzyskanych przy zastosowaniu aktualnych rozwiązań. Podano uzasadnienie stworzenia każdego proponowanego podejścia, dokładne wyjaśnienie jego działania, diagramy przepływu danych i ilustracje przedstawiające, w jaki sposób poszczególne etapy metody są wdrażane na rzeczywistych obrazach i fotografiach.

Wraz z rozdziałem piątym, rozdział szósty przedstawia i omawia wnioski płynące z proponowanych metod identyfikacji ośrodków dołka i tarczy nerwu wzrokowego na zdjęciach dna oka. Wyniki proponowanych technik zostały zbadane i porównywane z wynikami aktualnych rozwiązań.

Rozdział siódmy zatytułowany „Konkluzje” podsumowuje rozprawę i przedstawia wnioski, które wynikają z celów pracy i założonych tez.

Rozprawa doktorska zawiera również 4 dodatki, które stanowią uzupełnienia głównych rozdziałów. Na przykład, załącznik A zawiera wyniki eksperymentów, na podstawie których wnioski są omawiane w rozdziale szóstym.

Specyfika wdrażania przedstawionych technik i narzędzi dodatkowych została omówiona w Załączniku B. Oprócz instrukcji dotyczących instalacji narzędzi startowych środowiska uruchomieniowego, dodatek zawiera również niezbędne procedury umożliwiające dostęp do kodów źródłowych.

Postęp matematyczny w metodologii stosowanej do generowania sugerowanych procedur oraz wnioski przedstawiono i omówiono w Załączniku C.

Zagadnienia, które nie zostały omówione w tej pracy, są wymienione w Załączniku D pod tytułem „Otwarte problemy”. W załączniku znajdują się odpowiednie wskazówki dotyczące rozwiązywania poszczególnych zadań.

## Zalety rozprawy doktorskiej

Retinopatia cukrzycowa (ang. Diabetic Retinopathy, DR) jest bardzo poważnym schorzeniem, szczególnie dla pacjentów z cukrzycą, mogącym prowadzić do trwałej utraty wzroku. Diagnostyka i wcześniejsze wykrycie retinopatii cukrzycowej pozwala na szybkie wdrożenie terapii i leczenia. W ostatnich latach, metody i narzędzia sztucznej inteligencji, zwłaszcza techniki głębokiego uczenia pokazują ogromny potencjał we wspomaganie analizie i przetwarzania zdjęć dna oka. Te techniki mogą okazać się bardzo skuteczne w wykrywaniu drobnych zmian w siatkówce, które mogą być związane z różnymi choroby układu krążenia lub układu oddechowego. Pod tym względem, podjęty temat badawczy jest szczególnie ważny dla nowoczesnego rozwoju medycyny, a każdy postęp w tych badaniach ma ogromne znaczenia dla ludzkiego zdrowia.

Najważniejszym wynikiem rozprawy są dwa opracowane algorytmy lokalizacji współrzędnych środków dołączka i tarczy nerwu wzrokowego na zdjęciach dna oka o nazwach: MSiLR (Metoda Szybkiej i Równoległej) oraz MSLW (Metoda Symultanicznej przy użyciu Zespołu Sieci Neuronowych). Algorytm MSiLR to sekwencja tradycyjnych technik przetwarzania obrazów takich jak konwersja koloru, redukcja szumu, wyrównanie histogramu, wygładzanie konturów, ... parametry wewnętrznych modułów zostały dobrane ręcznie lub na podstawie informacji z literatury.

W drugim algorytmie (MSLW), oprócz algorytmów wstępnego przetwarzania zdjęć, Pan mgr inż. Tomasz Krzywicki korzystał z algorytmu lokalizacji obiektów na zdjęciach o nazwie YOLOv8. Obecnie, to jest jedna z najskuteczniejszych metod detekcji obiektów opartych o głębokich sieci neuronowych i uczenia transferowego. Poza tym, MSLW został wzmocniony o moduł uczenia zespołowe, który agregował wyniki z 16 sieci neuronowych realizujących zadanie lokalizacji anatomicznych dna oka.

Bardzo cenne są eksperymenty przeprowadzone na prawdziwych zdjęciach dna oka. Są to zdjęcia z bazy Fundacji Okulistyka 21 oraz z bazy IDRID służącej do konkursu "ImageNet – Large Scale Visual Recognition Challenge". Skoro algorytmy zostały przetestowane na prawdziwych danych konkursowych, wnioski z tych eksperymentów są trafniejsze.

Eksperymenty na wyżej wymienionych danych pokazują, że algorytm MSiLR jest znacznie słabszy niż MSLW pod względem jakościowym. Za szczególnie cenne uważam osiągnięcie przez Autora rozprawy bardzo dobrej dokładności lokalizacji

współrzędnych środków dołączka za pomocą MSLW. Wynik był lepszy od wyników zwycięzców konkursu analizy danych IDRID 2023.

Za wartościowe należy uznać fakt, że oba algorytmy zostały zaimplementowane zarówno w wersji sekwencyjnej, jak i równoległej. Dzięki równoległej implementacji na procesorach graficznych, proces rozpoznawania spieszyl się średnio ok. 2 razy w przypadku MSiLR i prawie 8 razy w przypadku MSLW.

## **Uwagi polemiczne i krytyczne po lekturze rozprawy doktorskiej**

Uważam, że rozdział 3 o tytule “metody rozpoznawania obrazów” jest najsłabszą częścią rozprawy. Ten rozdział miał być omówieniem pojęć i algorytmów, które będą używane w następnych rozdziałach. Niestety zawartości i poziomy szczegółowości jej podrozdziałów nie były odpowiednio dobrane. Na przykład podrozdziały 3.2 zawiera zbyt szczegółowe omówienia teorii złożoności obliczeniowej włącznie z pojęciem maszyny Turinga i klasami złożoności DTIME, SPACE, NSPACE, P, NP podczas gdy te pojęcia nigdzie indziej w rozprawie nie były wspomniane. Co więcej, autor nie przeprowadził analizy złożoności obliczeniowej dla dwóch głównych algorytmów w rozprawie. W podrozdziale 3.5 p.t. “Algorytmy topologiczne i algorytmy inspirowane naturą”, autor poświęcił zbyt dużo uwagi i bardzo szczegółowo opisał teorię uczenia wraz z wymiarem Vapnik-Chervonenkis i PAC uczenia, podczas gdy, metoda uczenia transferowego – jedna z ważniejszych metod uczenia maszynowego obecnie – była bardzo pobieżnie omówiona. Brakuje również prezentacji struktur głębokich sieci neuronowych, w szczególności Doktorant powinien bardziej rozbudować część omówienia algorytmu YOLOv8, skoro algorytm ten odgrywa bardzo istotną rolę w rozprawie.

Podrozdział 3.6 o tytule “Nowoczesne metody uczenia maszynowego i rozpoznawania obrazu” zawiera 3 podpodrozdziały: 3.6.1 “Uczenie transferowe”, 3.6.2 “Uczenie zespołowe” oraz 3.6.3 “Detekcja i lokalizacja obiektów”. Jeśli dwa podpodrozdziały 3.6.1 i 3.6.3 są bardzo ważne dla rozprawy i doskonale pasują do 3.6, to 3.6.2 nie bardzo pasuje. Omówione techniki w 3.6.2 pochodzą z lat 90-tych, czyli ponad 25 lat temu, trudno więc traktować je jako nowoczesne techniki. Co gorsze, znów poziomy szczegółowości tych podpodrozdziałów nie zostały odpowiednio zadbane, 3.6.1 i 3.6.3 były opisane bardzo pobieżnie i mają po jednej stronie podczas, gdy 3.6.2 był rozpisany na 4 stronach.

Wśród dużej liczby odnośników rozprawy doktorskiej, znalazłem jeden artykuł wydany w czasopiśmie Journal of Clinical Medicine, w którym doktorantem był współautorem. Pan mgr inż Tomasz Krzywicki był samodzielnym autorem dwóch innych publikacji (jeden rozdział w książce i jednego monografa wydanego przez UWM).

Zaniepokojony tą skromną ilością publikacji, zacząłem szukać i zauważyłem, że Pan mgr inż. Tomasz Krzywicki jest autorem kilka innych publikacji. Wśród tych niewymienionych w rozprawie, są dwa artykuły wydane w roku 2023 w czasopiśmie Journal of Clinical Medicine oraz Ophthalmic Research. Sądząc po tytułach, uważam, że oba artykuły są ściśle związane z tematyką rozprawy doktorskiej, dlatego one powinny znaleźć się w spisie literatury.

## **Aspekty formalne i językowe rozprawy doktorskiej**

Praca pod względem formalnym przygotowana została poprawnie, choć wykazuje też drobne usterki.

Przed wszystkim w pracy brakuje wykazu skrótów, w którym powinny znaleźć się skróty powszechnie używane (takie jak: s. np.) oraz skróty metod, pojęć i algorytmów.

Język pracy jest klarowny, zrozumiały. Składnia i stylistyka są bardzo dobre. Praca jest wzorowa z perspektywy językowej, na docenienie zasługuje świetna korekta. W efekcie pracę bardzo dobrze się czyta, bo w zasadzie nie ma ona żadnych błędów utrudniających percepcję.

## **Bibliografia rozprawy doktorskiej**

Spis literatury liczy 22 strony i zawiera ponad 200 odnośników do artykułów naukowych i do źródeł zasobów dostępnych w internecie.

Dużą część bibliografii (ok. 70 pozycji) stanowią publikacje z okresu 2020-2023. Pod względem ilościowym uważam, że doktorant dość starannie śledzi rozwój i trendy w swojej dziedzinie badawczej.

Mam jednak krytyczne opinie na temat zawartości i sposobu prezentacji pozycji literatury. Według mnie autor rozprawy powinien rozdzielić artykuły naukowe od zasobów internetowych takich jak kody źródłowe, biblioteki oprogramowania, dokumentacje, instrukcje lub źródła danych.

## **Uwagi i wnioski końcowe**

W świetle przeprowadzonej analizy i oceny rozprawy doktorskiej p. mgr inż. Tomasza Krzywickiego, w kontekście przewidzianych prawem kryteriów wymagających spełnienia dla uzyskania stopnia naukowego doktora, należy stwierdzić, że kryteria te zostały spełnione.

Przedstawiona do recenzji rozprawa doktorska stanowi oryginalne rozwiązanie problemu lokalizacji cech anatomicznych na zdjęciach dna oka.

Po lekturze rozprawy można również stwierdzić, że Autor rozprawy wykazał się bardzo dobrą ogólną wiedzą w swojej dyscyplinie. Mimo, że w recenzji zostały wskazane pewne niedociągnięcia oraz zgłoszone pewne zastrzeżenia, praca stanowi jednak ciekawy, oryginalny przykład zastosowania uczenia maszynowego w praktyce oraz stanowić cenny materiał dla praktycznych zastosowań. Praca jest wartościowa, interesująca, na wysokim poziomie merytorycznym. Zdecydowanie zasługuje na publikację. Ponadto, analiza zawarta w pracy dowodzi pogłębionej ogólnej wiedzy Doktoranta w zakresie informatyki technicznej

Podsumowując, recenzowana praca spełnia wszystkie kryteria właściwe dla rozpraw doktorskich, co uzasadnia postawienie wniosku o przyjęcie rozprawy doktorskiej, dopuszczenie jej do publicznej obrony i kontynuowanie czynności w ramach przewodu doktorskiego p. mgra inż. Tomasza Krzywickiego.



Nguyen Hung Son