

**dr hab. inż. Robert Salat, prof. PK**  
Katedra Inżynierii Elektrycznej  
Wydział Inżynierii Elektrycznej i Komputerowej  
Politechnika Krakowska im. Tadeusza Kościuszki  
ul. Warszawska 24, 31-155 Kraków  
Tel. 12 628 26 14, e-mail: [robert.salat@pk.edu.pl](mailto:robert.salat@pk.edu.pl)

Kraków, 16.09.2024 r.

## **RECENZJA ROZPRAWY DOKTORSKIEJ**

Pana kpt. mgr. inż. **Fabiana Gila**, pt.:

*„Zespoły sieci głębokich w rozpoznawaniu wybranych klas obrazów medycznych”*,

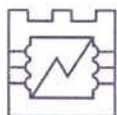
wykonanej na Wydziale Elektroniki Wojskowej Akademii Technicznej  
im. Jarosława Dąbrowskiego

Promotor: prof. dr hab. inż. Stanisław Osowski

Recenzję wykonano na podstawie uchwały nr 15/RDN AEEiTK/2024 podjętej przez Radę  
Dyscypliny Naukowej Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne

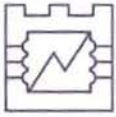
Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego

z dnia 19 czerwca 2024 r



## 1. Uzasadnienie podjęcia tematu

Rozpoznawanie obrazów medycznych nie jest nową metodą stosowaną w różnych dziedzinach nauki jednak jak ważne jest to zagadnienie, świadczyć może liczba artykułów na ten temat, która do roku 2023 w różnych bazach przekracza kilkadziesiąt tysięcy pozycji. Niezmiennym kierunkiem od samego początku jest poszukiwanie coraz skuteczniejszych metod analizy obrazów medycznych oraz rozwiązywania problemów z nimi związanych. Dlatego zasadne jest rozwijanie nowoczesnych metod stosowanych w analizie obrazów, które będą wspomagać lekarzy w stawianiu diagnozy, co przyczyni się do optymalizacji leczenia pacjenta. Różne podejścia do powyższych zagadnień można znaleźć w literaturze - niektóre oparte są na analizie tekstur, modelowaniu statystycznym, jak i coraz częściej wykorzystujące algorytmy sztucznej inteligencji. Mamy tu do czynienia z podejściem klasycznym wykorzystującym tzw. płytkie sieci typu MLP (ang. *Multilayer Perceptron*), SVM (ang. *Support Vector Machines*), RBF (ang. *Radial Basis Function*), drzewa decyzyjne (ang. *Random Forest*), jak i nowoczesne metody, w szczególności głębokie sieci neuronowe, np. CNN (ang. *Convolutional Neural Network*). Wady używania płytkich sieci neuronowych w analizie obrazów wynikają głównie z ich ograniczeń dotyczących zdolności do wykrywania złożonych wzorców, problemów z przetwarzaniem dużych zbiorów danych, niższej dokładności, niższej tolerancja na zakłócenia, braku hierarchicznego uczenia się cech oraz ograniczenia w transferze wiedzy. Ogólnie mówiąc, płytkie sieci są mniej efektywne w analizie obrazów, zwłaszcza tych o wysokiej złożoności, ze względu na brak zaawansowanych mechanizmów uczenia się złożonych cech oraz ograniczone możliwości generalizacji. Ich następcy, czyli głębokie sieci, choć bardziej zasobożerne, znacznie lepiej radzą sobie z analizą skomplikowanych obrazów medycznych i innych typów danych wizualnych. Obecnie popularność zyskało, szczególnie w medycynie, budowanie zespołów klasyfikatorów dzięki postępom w sztucznej inteligencji i uczeniu maszynowym. Łączenie wyników wielu modeli, w tym głębokich sieci neuronowych np. CNN, pozwala uzyskać bardziej niezawodne wyniki w diagnostyce obrazowej. Zastosowanie zespołów klasyfikatorów w analizie obrazów przynosi wiele korzyści, szczególnie w medycynie, gdzie dokładność i niezawodność klasyfikacji są kluczowe. Do głównych zalet możemy zaliczyć poprawę dokładności, redukcję nadmiernego dopasowania, zwiększenie odporności na szumy, łączenie zalet różnych modeli, lepszego uogólnienia wyników, zwiększenia stabilności wyników oraz skalowalność. Stosowanie zespołów klasyfikatorów w analizie obrazów ma wiele zalet, ale są również pewne wady, o których warto wspomnieć. Głównie są to: wysoka złożoność obliczeniowa, zwiększone wymagania pamięci, trudność interpretacji wyników, ryzyko



nadmiernego dopasowania, problemy z wagami poszczególnych modeli, problemy z generalizacją, wysokie koszty utrzymania i wdrożenia.

Podsumowując, uważam, że podjęty temat rozprawy dotyczący zastosowania sieci neuronowych z uczeniem głębokim w systemach rozpoznawania obrazów medycznych jest bardzo ważnym i aktualnym problemem. Bardzo obszerny zakres tematyczny rozprawy oraz motywacja do podjęcia tak trudnego tematu zasługują na duże uznanie.

## **2. Ogólna charakterystyka rozprawy doktorskiej**

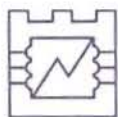
Recenzowana rozprawa doktorska napisana w postaci monografii obejmuje 109 stron i składa się ze streszczenia w języku polskim i angielskim, spisu treści, oraz 7 rozdziałów. Poszczególne rozdziały zawierają wprowadzenie, wyniki badań, podsumowanie, wnioski końcowe, wykaz literatury zawierający 72 pozycje, w tym 4 pozycje współautorskie Autora rozprawy. Na końcu rozprawy zamieszczono spis rysunków i tabel. Przedstawione w rozprawie wyodrębnione i ponumerowane rozdziały to:

### **Rozdział 1. Wstęp**

Rozdział 1 zawiera wprowadzenie w tematykę rozprawy oraz przegląd aktualnego stanu wiedzy w zakresie automatycznej analizy obrazów medycznych. Doktorant na początku rozdziału przedstawił podstawowe sieci neuronowe, które są obecnie zaliczane do klasycznych metod rozpoznawania obrazu a następnie omówił ich zalety i wady. W dalszej części rozdziału zaprezentował ich następców w postaci głębokich sieci neuronowych oraz dokonał szerokiego przeglądu literaturowego odnośnie tematyki rozprawy. Szczególnie skupił się na metodach budowania zespołów klasyfikatorów oraz przedstawił i omówił wyniki uzyskane przez autorów poszczególnych publikacji wykazanych w wykazie literatury zbliżonej tematycznie to ocenianej rozprawy. Rozdział kończy się przedstawieniem celu pracy i zwięzłym omówieniem zawartości wszystkich kolejnych rozdziałów.

### **Rozdział 2. Baza danych medycznych**

W rozdziale 2 przedstawiono analizę obrazów zawartych w trzech bazach danych medycznych wykorzystanych w badaniach Doktoranta. Doktorant skupił się na omówieniu zbiorów danych i przedstawił problemy jakie pojawiają się w analizie danych z obrazów. Przedstawił



szczegółowo porównanie wartości statystycznych analizowanych obrazów, tj. średniej, odchylenia standardowego, energii, kurtozy i skośności. Wykazał w Tabelach od 2.1 do 2.5, że obrazy zaliczane do różnych klas są podobne statystycznie według wyżej wymienionych parametrów statystycznych. Potraktowanie opisanych parametrów jako zbioru cech wejściowych dla klasyfikatorów spowodują trudności rozpoznania klas. Analizy poddane ocenie dotyczyły zarówno nowotworowych zmian skórnych, jak i raka piersi.

### **Rozdział 3. Indywidualne systemy klasyfikacyjne**

W rozdziale 3 przedstawiono i scharakteryzowano wszystkie klasyfikatory użyte w badaniach. Dokonany został podział na klasyfikatory, które były bezpośrednio użyte do wskazania klas oraz takie, gdzie głębokie sieci neuronowe typu konwolucyjnego użyte zostały jedynie do generacji cech opisujących obraz wejściowy. Opisano dokładnie głębokie sieci neuronowe oraz ich metody uczenia skupiając się na architekturze, procesie generacji cech a kończąc na klasyfikatorze neuronowym typu softmax. Przedstawiono również klasyczne metody klasyfikacji oparte na sieciach SVM oraz drzewa decyzyjne. Rozdział kończy się omówieniem deterministycznych metod selekcji cech diagnostycznych.

### **Rozdział 4. Zespoły klasyfikatorów**

Rozdział 4 zawiera omówienie metod budowania zespołów klasyfikatorów. Doktorant skupił się tu na przedstawieniu, na czym polega fuzja klasyfikatorów a następnie omówił dostępne w Matlabie struktury sieciowe. W dalszej części rozdziału rozwinął dwa podejścia zastosowane w pracy do tworzenia zespołu klasyfikatorów. Pierwsze - oparte na połączeniu sieci SVM i RF zasilanych przez przedstawione metody selekcji cech opisanych w poprzednim rozdziale i połączone z sieciami CNN, a drugie bazujące wyłącznie na strukturach głębokich CNN. Rozdział kończy się omówieniem metod oceny jakości klasyfikacji zastosowanej w badaniach.

### **Rozdział 5 Wyniki badań eksperymentalnych w rozpoznawaniu czerniaka**

Rozdział piąty jest jednym z dwóch kluczowych rozdziałów rozprawy. Autor dysertacji w tym rozdziale przedstawił wyniki badań eksperymentalnych dotyczących skuteczności różnych systemów klasyfikacyjnych odnośnie rozpoznawania czerniaka. Badania odnoszą się do dwóch niezależnych baz danych, które są osobno analizowane. W pierwszej części rozdziału, Doktorant przedstawił uczenie 19 sieci CNN dostępnych w Matlabie i ocenił ich jakość klasyfikacji. Następnie zbadał wpływ zamrożenia poszczególnych warstw na wynik oraz ich czas uczenia. W dalszej części



rozdziału rozszerzył badania o dodanie do CNN sieci SVM i RF i wyłonienie ze zbioru deskryptorów mniejszej liczby cech podawanych na klasyfikatory klasyczne. Badania porównał z zespołami złożonymi wyłącznie z sieci CNN. W ostatniej części rozdziału skupił się na klasyfikacji czerniaka w ogólnodostępnej bazie ISIC w oparciu tylko o klasyfikatory CNN. Zbudował 11 zespołów a ich wyniki ostateczne przedstawił w Tabeli 5.11.

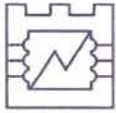
## **Rozdział 6. Wyniki badań eksperymentalnych w rozpoznawaniu raka piersi**

Rozdział szósty jest drugim głównym rozdziałem rozprawy, w którym Doktorant skupił się na przedstawieniu metod zastosowanych do wykrywania zmian nowotworowych piersi na podstawie obrazów mammograficznych, które są odmienne niż obrazy zmian skórnych. Badania przeprowadził z wykorzystaniem bazy danych DDSM; zostały one podzielone na trzy zadania. Dwa dotyczyły dwuklasowej klasyfikacji a jedno trójklasowej. Następnie dla tak sformułowanych zadań wykonał uczenie dla poszczególnych architektur sieci w celu wybrania najlepszych zespołów klasyfikatorów. Rozdział kończy się wynikami, które są przedstawione w Tabelach 6.4, 6.6, 6.8, 6.10.

## **Rozdział 7. Podsumowanie i wnioski końcowe**

W rozdziale 7 Doktorant przedstawił podsumowanie, wnioski końcowe oraz omówił najważniejsze osiągnięcia przedstawione w niniejszej rozprawie. Uzyskane wyniki potwierdzają słuszność zastosowanych zespołów klasyfikatorów. Podsumowanie kończy się przedstawieniem kierunku dalszych badań oraz możliwości ich osiągnięcia.

Podsumowując, na podstawie przedstawionego opisu zawartości poszczególnych rozdziałów w recenzowanej rozprawie, można stwierdzić, że praca ma charakter teoretyczno-praktyczny i przedstawia obszerne wyniki badań, które w stopniu wystarczającym pozwalają zrealizować postawione główne cele pracy i udowodnić tezę. Praca napisana jest poprawnie pod względem językowym, układ i struktura recenzowanej rozprawy są przedstawione logicznie i we właściwy sposób.



### 3. Najważniejsze osiągnięcia rozprawy

Zaplanowane w ramach pracy doktorskiej badania miały nowatorski charakter. Cele jakie Doktorant założył na początku rozprawy konsekwentnie realizował w poszczególnych etapach rozprawy doktorskiej. Do najważniejszych osiągnięć Doktoranta, które również sam podkreślił, można zaliczyć:

1. Opracowanie i przebadanie skuteczności działania zespołu klasyfikatorów opartych na głębokich sieciach CNN.
2. Opracowanie sposobu optymalnego doboru członków zespołów bazującego na analizie wartości trzech głównych miar jakości.
3. Przeprowadzenie ogromnej liczby eksperymentów numerycznych przy testowaniu zaproponowanych rozwiązań zespołowych.

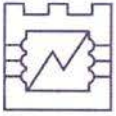
Ostateczne wyniki, które zostały uzyskane, zostały porównane z najlepszymi rezultatami dla baz, dla których istniały wyniki innych zespołów badawczych. Dla bazy ISIC2017 były aż o 6% lepsze dla wskaźnika ACC, o ponad 12% dla TPR i 5% dla TNR. Dla bazy DDSM, wyniki, które mogły być obiektywnie porównane, były lepsze o ponad 9% dla ACC, o ponad 4% dla TPR i 18% dla TNR.

Podsumowując nie wszystkie dane z baz danych dało się porównać, ponieważ Doktorant nie dysponował wynikami otrzymanymi przez inne zespoły badawcze. Jednak wyniki, które Doktorant dla wszystkich badanych przypadków uzyskał na poziomie 96% i wyższym, można uznać, że są na światowym poziomie.

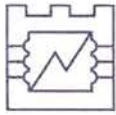
### 4. Uwagi do rozprawy

**Uwagi o charakterze ogólnym, które są zarówno pytaniami do Doktoranta, komentarzami, jak również uwagami, które wymagają wyjaśnienia. Kolejność uwag nie jest tożsama z ich pojawianiem się w rozprawie.**

1. Na stronie 95 pojawia się pojęcie hipotezy. W tekście nie ma jednak informacji odnośnie postawionych hipotez. Czy chodziło o tezę?
2. W części teoretycznej rozprawy wprowadzono miary F1 oraz F1(+) i F1(-), które nie były użyte w badaniach. Jakie były przesłanki do ich wprowadzenia?



3. Uczenie sieci CNN w Matlabie używa algorytmów gradientowych, które mają tendencję do utykania w minimach lokalnych. Jednym z najprostszych sposobów ich unikania jest kilkukrotne uruchomienie procesu uczenia i zamrożenia wag dla najlepszych wyników. Czy były wykonywane próby ponownego uczenia tej samej sieci, aby sprawdzić, czy można otrzymać jeszcze lepsze wyniki?
4. W niektórych tabelach w rozprawie podawane są czasy uczenia, ale brak podania dokładnych danych sprzętu komputerowego, na którym je wykonywano. Uważam to za dość istotne niedociągnięcie. Istotna jest również informacja, czy obliczenia były wykonywane na procesorze, czy karcie graficznej, których wybór Matlab umożliwia. Takie podejście może znacznie przyspieszyć obliczenia.
5. W posumowaniu Doktorant napisał, że współpraca SVM, RF z CNN pokazała stosunkowo niewielką poprawę wyników w stosunku do najlepszego indywidualnego rozwiązania sieci CNN. Tabele 5.2, 5.5. i 5.6 pokazują, że jest odwrotnie. Proszę o komentarz do tej uwagi.
6. Brakuje informacji na temat, jaka była funkcja strat, współczynnik uczenia, rodzaj sieci SVM i jego parametry, itd. W samym Matlabie można użyć kilkunastu różnych sieci SVM, które mogą dać bardzo różniące się od siebie wyniki.
7. Proszę o wyjaśnienie, dlaczego SVM i RF nie użyto do rozpoznawania raka piersi. Mamy tu do czynienia z innym zbiorem, innym typem zdjęć. Jaka jest pewność, że nic się nie zyska na poprawie wyników?
8. Rozmiar obrazów ma duży wpływ na działanie sieci CNN, szczególnie jeśli sieć była trenowana na obrazach o innym rozmiarze. W Tabeli 4.1 przedstawiono wymiary obrazów wejściowych wstępnie przetrenowanych sieci. Czy był badany wpływ rozmiarów zdjęcia z użytych baz danych na wyniki wstępnie przetrenowanych sieci?
9. Czy były wykonane próby zastosowania wszystkich deskryptorów generowanych przez sieci CNN na wejście sieci SVM? Dla sieci *Resnet18* liczba 512 nie jest bardzo duża. Z doбором (redukcją) wektora cech należy być ostrożnym, gdyż dla problemów o dużym wymiarze wektora mogą pojawić się przypadki, gdzie wszystkie cechy są równie ważne i są tego samego typu, a redukcja ich liczby wiąże się ze znacznym pogorszeniem wyników klasyfikacji. Istnieją również przypadki, gdzie cechy, które indywidualnie uważane są za nieistotne, w połączeniu z innymi cechami, znacząco poprawiają wyniki klasyfikacji. Niekiedy nawet połączenie razem cech mało istotnych indywidualnie powoduje, że stają się one łącznie istotne w rozwiązaniu problemu.



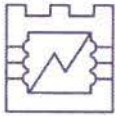
10. Brakuje mi w pracy chociaż jeden struktury sieci, która zostałaby zaproponowana przez Doktoranta. Łatwość budowania struktur oraz uczenia w Matlabie pod konkretne zadanie nie powinny stanowić dużego wyzwania. Proszę o komentarz do tej uwagi.

## 5. Uwagi redakcyjne

**Pragnę zaznaczyć, że poniższe uwagi redakcyjne nie mają istotnego wpływu na ocenę merytoryczną pracy i nie komplikują jej odbioru.**

1. Praca napisana jest bardzo poprawnym językiem polskim, jedynie w kilku miejscach pojawiają się pojedyncze błędy literowe lub źle dobrane słowa (np. strona 42, pierwsze zdanie).
2. W niektórych miejscach w pracy opis do tabel i rysunków znajduje się w dalszej części za tabelą/rysunkiem, co zdecydowanie utrudnia zrozumienie tego fragmentu pracy.
3. Literatura zawiera dużo pozycji arXiv. (11 pozycji). Czy nie dało się dołączyć już opublikowanej pozycji tych artykułów wiedząc że te artykuły są już kilkuletnie. Wiadomo, że artykuły tam zamieszczone nie zawsze przeszły pełną recenzję naukową (peer review), co jest standardowym procesem w renomowanych czasopismach naukowych.
4. Brak konsekwencji w ułożeniu literatury, czasami data wydania jest ostatnia, a niekiedy przed numerem wydania, itd.
5. Brak rozwinięcia skrótów po angielsku pojawiających się w pierwszym wystąpieniu w pracy np. CNN, SVM itd,
6. Brak konsekwencji w przedstawianiu wykresów, raz są łączone raz nie.
7. Brak konsekwencji w oznaczeniach raz ACC jest z % raz bez (strona 96).





#### 4. Wniosek końcowy

Przedstawiona do recenzji rozprawa doktorska kpt. mgr. inż. **Fabiana Gila** dotyczy bardzo aktualnego zagadania naukowego jakim jest rozpoznawanie obrazów medycznych za pomocą sztucznych sieci neuronowych.

Cele założone w pracy były konsekwentnie realizowane i w mojej ocenie **zostały osiągnięte**.

Zamieszczone uwagi w niniejszej recenzji nie mają istotnego wpływu na moją **pozytywną ocenę**. Zatem stwierdzam, że rozprawa doktorska Pana kpt. mgr. inż. **Fabiana Gila** zatytułowana „**Zespoły sieci głębokich w rozpoznawaniu wybranych klas obrazów medycznych**”, stanowi oryginalne rozwiązanie zagadnienia naukowego, a spójnie i kompleksowo przedstawiony sposób rozwiązania problemów dowodzi umiejętności prowadzenia prac naukowych.

Uważam, że opiniowana praca spełnia warunki i wymagania stawiane rozprawom doktorskim, określone w art. 187 ustawy z dnia 20 lipca 2018r. - Prawo o szkolnictwie wyższym i nauce. (Dz. U. 2023 poz. 742 - tekst ujednoczony z dnia 08.03.2023r.).

W związku z tym, przedkładam Radzie Naukowej dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Wydziału Elektroniki Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego wniosek o przyjęcie i dopuszczenie recenzowanej rozprawy doktorskiej do publicznej obrony.

*SauA Robert*

