

dr hab. inż. Mirosław Szmaja, prof. uczelni
Politechnika Opolska
Wydział Elektrotechniki, Automatyki i Informatyki
Katedra Automatyki
ul. Prószkowska 76, 45-758 Opole

Opole, 31 maja 2025 r.

RECENZJA

rozprawy doktorskiej mgr inż. Kamili Białek

pt. „Strategie przygotowania obrazów pisma odręcznego na potrzeby rozpoznawania choroby Parkinsona z wykorzystaniem sieci konwolucyjnej”

Promotor: dr hab. inż. Jacek Jakubowski, prof. Uczelni

Dyscyplina naukowa: automatyka, elektronika, elektrotechnika i technologie kosmiczne

1. WSTĘP I PODSTAWA PRAWNA

Recenzja została zlecona przez Przewodniczącemu Rady Dyscypliny Naukowej Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Wojskowej Akademii Technicznej w piśmie z dnia 31.03.2025 r. w oparciu o dokumentację, w skład której wchodzi:

- rozprawa doktorska pt.: „Strategie przygotowania obrazów pisma odręcznego na potrzeby rozpoznawania choroby Parkinsona z wykorzystaniem sieci konwolucyjnej”, której promotorem jest dr hab. inż. Jacek Jakubowski.

2. OCENA FORMALNA PRACY

Praca doktorska pt.: „Strategie przygotowania obrazów pisma odręcznego na potrzeby rozpoznawania choroby Parkinsona z wykorzystaniem sieci konwolucyjnej” obejmuje 100 stron i składa się z 11 zasadniczych rozdziałów, bibliografii z wykazem 113 prac oraz dwóch załączników. Ponadto praca uzupełniona jest o streszczenie w języku polskim oraz angielskim. Praca jest napisana w języku polskim.

Rozdział 1 stanowi wstęp do pracy, uzasadniający potrzebę zajęcia się zagadnieniem, przedstawiający problem oraz metodykę badań. W rozdziale zdefiniowano cel, dwie tezy oraz strukturę pracy.

W rozdziale 2 przedstawiono przegląd literaturowy.

Rozdział 3 jest wprowadzeniem do stosowanych w dysertacji konwolucyjnych sieci neuronowych, techniki poolingu, augmentacji danych oraz transfer learningu.

W rozdziale 4 przedstawiono grupę pacjentów, których poddano badaniom przez zespół medyczny w Klinice i Katedrze Neurologii Warszawskiego Uniwersytetu Medycznego. Zaprezentowano również stanowisko badawcze.

W rozdziale 5 zaprezentowano metody przygotowania obrazów do dalszego zastosowania w trenowaniu i walidacji modeli sieci.

Rozdział 6 przedstawiono sposób zarządzania danymi oraz wprowadzenie do miar jakości klasyfikacji.

Rozdziały 7, 8, 9 reprezentują opis metodyki badań oraz osiągniętych rezultatów prowadzących do udowodnienia postawionych tez. Pierwsza z tez została udowodniona w rozdziale 8, natomiast druga w 9.

W rozdziale 10 wskazano na możliwości dalszej optymalizacji sieci AlexNet, uzyskując przy tym najlepsze wyniki klasyfikacji. Ponadto przeanalizowano wyniki uzyskane za pomocą trzech alternatywnych do AlexNet konwolucyjnych sieci neuronowych.

Rozdziały 4 do 10 stanowią zasadniczą część merytoryczną dysertacji.

Ostatecznie rozdział 11 jest podsumowaniem dysertacji, w którym podsumowano wątek naukowy oraz cel i tezę pracy, zestawiono wyniki zrealizowanych prac, uwypuklono autorskie podejście do rozwiązywania problemów badawczych oraz zarysowano tematykę dalszych prac naukowych.

Uwagi na temat poprawności językowej i redakcyjnej rozprawy.

- Struktura pracy jest logiczna i czytelna.
- Ogólnie praca jest napisana poprawnym językiem polskim bez większych błędów stylistycznych.
- Pomniejsze uwagi:
 - str. 18: zaleca się rozwinięcie skrótów NCV oraz NCA,
 - str. 24: w tabeli 2.3 podane są wartości parametru dokładności w procentach, natomiast w tabeli 2.2 w wartościach znormalizowanych w zakresie 0 do 1; dla czytelności rozprawy można by jednostki ujednolicić i podać, bądź w procentach, bądź w zakresie 0 do 1,
 - w opisie wzoru 3.6 warto wprost zdefiniować parametr Θ .

3. OCENA MERYTORYCZNA PRACY

3.1. Analiza istotność zagadnienia.

Problematyka badawcza niniejszej dysertacji dotyczy bardzo istotnego zagadnienia społeczno-medycznego, jakim jest detekcja choroby Parkinsona, jednej z najliczniej występującej choroby neurodegeneracyjnej na świecie. Wczesna diagnoza oraz implementacja stosownego leczenia pozwala zdecydowanie dłużej utrzymać sprawność ruchową i niezależność chorego. Typowo diagnostyka choroby Parkinsona obejmuje tzw. "triadę parkinsonowską", obejmującą drżenie spoczynkowe (najczęściej rąk), bradykinezję (spowolnienie ruchowe) oraz sztywność mięśniową. W celu wykluczenia innych chorób stosuje się również pomocniczo m.in. rezonans magnetyczny, badanie elektroencefalograficzne oraz elektromiografię.

Cały proces diagnostyczny jest zatem dosyć złożony i kosztowny, dlatego uzasadnionym jest dostarczenie personelowi medycznemu łatwego w obsłudze, szybkiego i taniego narzędzia do diagnostyki przesiewowej. Niniejsza dysertacja jest jednym z etapów związanych z odpowiedzią na postawione zadanie, wykorzystując zjawisko drżenia rąk, zarejestrowanego w postaci zniekształconego pisma chorego. Doktorantka rozpoznając stan badań w tym zagadnieniu, nawiązuje polemikę z innymi badaczami, i proponuje swoje rozwiązanie, zwiększające poziom wiarygodności diagnostycznej.

Rozwiązanie problemu jest trudne i nieoczywiste. Jest to związane z samą naturą badań medycznych, gdzie obiektem są pacjenci o różnych charakterystycznych dla siebie cechach fizjologicznych. Zatem detekcja choroby neurodegeneracyjnej z dużym prawdopodobieństwem jest sama w sobie osiągnięciem. Z pomocą przychodzi tu rozwój technologii, a w szczególności układów akwizycji danych, przetwarzania obrazów oraz zastosowań sztucznej inteligencji.

Ostatecznie postawione sobie przez Doktorantkę zagadnienia badawcze, stanowią próbę stworzenia środowiska diagnostycznego, umożliwiającego w sposób bezinwazyjny i szybki wykrycie wczesnego stadium rozwoju jednej z najczęstszych chorób neurodegeneracyjnych, jaką jest choroba Parkinsona. Tak postawione zadanie uważam za trudne oraz ambitne i jego rozwiązanie stanowi zagadnienie badawcze o sporym potencjale naukowym na poziomie dysertacji doktorskiej. Zagadnienie to wpisuje się w rozwój dyscypliny naukowej automatyka, elektronika, elektrotechnika i technologie kosmiczne.

3.2. Zawartość merytoryczna pracy

Rozdziały 1 do 3 stanowią wprowadzenie do zagadnień, jakie mają być poruszane w ramach pracy.

W szczególności rozdział 1 identyfikuje:

- główny cel rozprawy:

„zbadać zagadnienie wpływu procesu przygotowania obrazów pisma na dokładność systemu rozpoznawania choroby Parkinsona wykorzystującego jak klasyfikator konwolucyjną sieć neuronową”,

- zakres pracy:

- opracowanie elementów stanowiska badawczego z zastosowaniem tabletu graficznego,
- przeprowadzenie eksperymentów rejestracyjnych próbek pisma z udziałem pacjentów cierpiących na chorobę Parkinsona oraz grupy kontrolnej,
- zaproponowanie strategii przygotowania danych obrazowych dla sieci konwolucyjnej,
- zaproponowanie metodyki podejmowania decyzji przez sieć neuronową w przypadkach, gdy danymi wejściowymi są sekwencje obrazów zdań i sekwencje obrazów słów,
- zastosowanie walidacji krzyżowej w celu oceny procesu klasyfikacji,

- oraz tezy pracy:

- *„użycie sekwencji obrazów zdań daje lepsze wyniki klasyfikacji binarnej w diagnostyce choroby Parkinsona z wykorzystaniem konwolucyjnej sieci neuronowej niż użycie obrazu zdania pojedynczego”,*
- *„możliwa jest dalsza poprawa dokładności klasyfikacji binarnej w diagnostyce choroby Parkinsona z wykorzystaniem konwolucyjnej sieci neuronowej na drodze podziału obrazów sekwencji zdań na obrazy sekwencji słów”.*

W rozdziale 2 przedstawiono przegląd literaturowy. Zaprezentowano ewaluację tematyki – od roku 1972 po stan bieżący badań – 2024: od podejść statystycznych po techniki uczenia maszynowego i implementację algorytmów klasyfikacyjnych. Zwrócono szczególną uwagę na zastosowania elektronicznych rejestratorów pisma: tabletów graficznych oraz piór elektronicznych, umożliwiających, oprócz rejestracji samego kształtu liter, wyrazów oraz zdań,

badanie siły nacisku czy parametrów dynamicznych procesu pisania. Wymieniono również różne metody przetwarzania uzyskanych grafik oraz parametrów dynamicznych w celu ekstrakcji i selekcji stosownych cech, będących podstawą zastosowania klasyfikatorów. Ostatecznie przytoczono zastosowanie różnych typów sieci neuronowych w badaniu szeregów czasowych (badania z 2023r. i 2024r.), oraz w szczególności sieci konwolucyjnych, umożliwiających podejmowanie bezpośrednich decyzji klasyfikacyjnych. Te ostatnie badania opublikowano w roku 2018. Ponadto scharakteryzowano dwie grupy pacjentów, stosowane przez autorów publikacji. Zestawiono również zdania, wyrazy oraz sekwencje liter, stosowane podczas testów. Ostatecznie przedstawiono aktualne kierunki badań nad implementacją konwolucyjnych sieci neuronowych w diagnostyce choroby Parkinsona. Doktorantka powołuje się przy tym na aktualne wyniki badań, w których szczegółowo przytacza dwie publikacje z 2019 i 2021 r., będące dla niej punktem odniesienia do badań własnych.

Rozdział 3 stanowi wprowadzenie do konwolucyjnych sieci neuronowych wraz z wyjaśnieniem operacji funkcji splotu, funkcji aktywacji i poolingu, uczenia sieci konwolucyjnych, augmentacji danych oraz transfer learningu.

Podsumowując rozdziały 1-3 stwierdzam, iż:

- prawidłowo uargumentowano potrzebę zajęcia się zagadnieniem,
- jasno określono cel badawczy,
- tezy pracy zostały postawione w sposób odpowiedni,
- dokonano poprawnego przeglądu literatury pod kątem badania obiektu – choroba Parkinsona, oraz metod związanych z przetwarzaniem pisma stosowanych w jej diagnostyce,
- prawidłowo zidentyfikowano bazy danych pod kątem możliwości pozyskania przebiegów zapisów graficznych liter, słów oraz całych zdań osób cierpiących na chorobę Parkinsona oraz grupy kontrolnej.

Dalsze rozdziały 4-10 są już autorską częścią dysertacji, zawierającą opis stanowiska, stosowanej bazy danych, metodologii oraz wyników badań.

W rozdziale 4 Doktorantka przedstawia grupę pacjentów, których poddano badaniom przez zespół medyczny w Klinice i Katedrze Neurologii Warszawskiego Uniwersytetu Medycznego. Badania objęły grupę 48 osób, wśród których były 24 osoby zdrowe (11 kobiet i 13 mężczyzn) i tyle samo osób chorych na chorobę Parkinsona (16 kobiet i 8 mężczyzn). Osoby chore zostały farmakologicznie przygotowane tak, aby objawy choroby były niewielkie. Takie podejście warunkuje zbadanie czułości metody rozwijanej dalej przez Doktorantkę. Zaprezentowano następnie stanowisko badawcze, składające się z tabletu graficznego, mikrofonu, kamer (kamera RGB, termowizyjna oraz audio-video) oraz komputera do akwizycji i przetwarzania danych. Ponadto przygotowano oprogramowanie w środowisku Matlab do przeprowadzania obliczeń oraz stworzenia interfejsu użytkownika. **Jako zadanie, mające być wykonane przez pacjenta, Doktorantka zaproponowała 5-krotne napisanie jedno pod drugim zdanie o treści: „Dzisiaj jest ładna pogoda”.** Tu nasuwa się pytanie, dlaczego akurat takie zdanie, skoro w literaturze stosuje się wyrazy zawierające litery zbliżone do kształtu pętli, np. „hello”, „hell” lub zdania np.: „Ein helles grelles Licht”. Doktorantka również podczas wykrywania całych zdań stosuje założenie a’priori o szerokości zdania. Pytanie zatem, czy stosowany algorytm działa tylko w przypadku stosowania tego konkretnego zdania, czy też dla innych zdań o innej liczbie wyrazów, oraz co w przypadku, gdy stosuje się standardowe zdanie, ale pacjent pisze bardzo „rozwlekłe” - np. dziecko lub osoba w spektrum autyzmu?



Rozdział 5 prezentuje metodykę przygotowania danych będących wejściem dla konwolucyjnej sieci neuronowej. Metodyka składała się z trzech zasadniczych części: wyodrębnienia zdań oraz wyodrębnienia poszczególnych wyrazów. W tym celu wykorzystana została analiza ruchu rysika na powierzchni oraz ponad powierzchnią tabletu z wykorzystaniem monitorowania nacisku rysika.

W rozdziale 6 zaprezentowano sposób wykorzystania zarejestrowanych obrazów pisma po wstępnym przetwarzaniu. Analizowano:

- pojedyncze zdania,
- pojedyncze zdania pochodzące z zapisu ich sekwencji,
- pojedyncze słowa w zdaniu.

Tak przygotowane dane wejściowe pogrupowano w zbiory uczące i testujące z wykorzystaniem 8-krotnej walidacji krzyżowej. Takie podejście umożliwiło w każdej z iteracji do uczenia wykorzystać dane od 42 osób a do testowania dane od 6. W tym rozdziale wprowadzono również pojęcia szeroko stosowanych miar jakości klasyfikacji, krzywej ROC oraz metod optymalnego dobru jej parametrów.

W rozdziale 7 Doktorantka prezentuje metodologię i wyniki badań w przypadku użycia pojedynczego zdania w celu przeprowadzenia klasyfikacji binarnej. Celem badania było zestawienie wyników dokładności klasyfikacji uzyskanych z zastosowaniem zbioru uczącego pochodzącego z własnych rejestracji, przedstawionych w rozdziale 4 z wynikami literaturowymi implementującymi bazę PaHaW.

Zaprezentowano trzy metodologiczne podejścia wykorzystując sieć AlexNet, przetrenowaną na podstawowej bazie ImageNet:

- 1) wykorzystanie warstw: konwolucyjnej 5 (*conv5*) oraz warstw 6 i 7 pełnego połączenia (*fc6* i *fc7*) wraz z ekstrakcją cech i klasyfikatorami maszyny wektorów nośnych SVM (ang. Support Vector Machine),
- 2) jak wyżej, z douczeniem sieci autorską bazą KN_19,
- 3) wykorzystanie dostrajania za pomocą zastąpienia warstwy *fc8* i warstwy *softmax* nową warstwą „nowa *fc8*” i douczeniu sieci bazą KN_19.

Jako klasyfikator w pierwszych dwóch metodologiach zastosowano algorytm maszyny wektorów nośnych SVM. Rozumiem, że algorytm ten został wybrany, aby nawiązać do publikacji [79]. Czy jednak były rozważane ewentualnie inne algorytmy klasyfikacyjne?

W podrozdziale 7.2 przedstawiono wyniki eksperymentów dla podanych trzech metodologii mierząc główne parametry macierzy błędów: *dokładność*, *czułość*, *precyzję* i *specyficzność*.

W pierwszym etapie do badań przyjęto tylko pierwsze z pięciu napisanych przez pacjenta zdań (podrozdział 7.2.1). Badania wykazują najlepsze wartości parametru *dokładność* dla metody 2 przy wykorzystaniu warstwy *fc7* oraz w metodzie 3 (nieznacznie gorsze parametry *precyzji* i *czułości*). Aczkolwiek dostrzegalna jest spora wartość odchyłeń standardowych dla wszystkich mierzonych parametrów, co nakazuje ich wartości średnie traktować raczej ostrożnie. Największe wartości *dokładności* plasują się na poziomie 0,6 a *czułości* na poziomie 0,67. Nie są to niestety wartości zbyt wysokie.

W drugim etapie rozważano analizę wszystkich pięciu zdań (podrozdział 7.2.2). Widoczny jest wzrost parametru *dokładności* wraz ze wzrostem numeru zdania (za wyjątkiem zdania nr 2). Ostatecznie *dokładność* dla zdania nr 5 osiąga dla metodologii 3 (dostrajanie) wartość 0,69

z odchyleniem standardowym $\pm 0,19$, co już jest znacznie lepszym rezultatem, w stosunku do pierwszego etapu. Analizując pozostałe parametry dla tego przypadku, *precyzja* oraz *specyficzność* posiadają również bardziej obiecujące wartości (odpowiednio 0,85 i 0,83) natomiast zauważalna jest niska *czułość* klasyfikacji (0,54). **Czy rozważane były przez Doktorantkę przyczyny takiej sytuacji?**

Rozdział 8 modyfikuje metodykę badań nr 2, analizując wszystkie pięć zdań, uzupełniwszy o strategię decyzyjną (rys. 8.1 blok „GŁOSOWANIE”). Ponadto zaimplementowano technikę augmentacji w celu zwiększenia liczby danych wejściowych do nauczania sieci neuronowych, stosując operacje obrotu w zakresie -10° do 10° . Eksperyment z tak wstępnie wyuczoną siecią neuronową podzielono na dwie części: bez zastosowania „algorytmu integrowania decyzji” oraz z wykorzystaniem „algorytmu decyzyjnego”. Być może nazewnictwo jest tutaj niezbyt dobrze dobrane, ponieważ na pierwszy rzut oka, są problemy z interpretowaniem tego pojęcia. Zasugerowałbym tutaj raczej np.: „algorytm głosowania” (nawiązując do rys. 8.1 i podpisu rys. 8.3).

Dla algorytmu bez stosowania „integrowania decyzji”, na rys. 8.3 i w tabeli 8.2 zaprezentowano macierz błędów oraz parametry jakości klasyfikacji. Nie jest jednoznacznie opisane, czy parametry te dotyczą zdania nr 1, czy też są to uśrednione wyniki wszystkich zdań. **W jaki sposób je obliczano?** Generalnie widoczny jest wzrost parametru *dokładność* oraz *czułość* przy gorszej *precyzji* i *specyficzności*. Zatem w mojej opinii nie do końca się zgodzę z Doktorantką, iż „wyniki w tym przypadku są zdecydowanie lepsze”, ale istotnie jest niewielka pewna poprawa w stosunku do poprzednich badań.

W przypadku zastosowania „algorytmu integrowania decyzji” (wyniki w tab. 8.3) sprawa ma się już faktycznie dużo lepiej. Sam wymieniony algorytm wykorzystuje wartość progową N , która determinuje wartość, będącą „minimalną liczbą przypisanych do grupy chorych”. Ze względu na 5 napisanych przez pacjentów zdań wartość N może wahać się w zakresie 1 – 5. Dla takich wartości, zostały obliczone parametry jakości klasyfikacji. Widoczny jest wzrost *precyzji* i *specyficzności* w funkcji wzrostu wartości N , przy dość dobrych wartościach *dokładności* (za wyjątkiem $N = 5$) i malejących wartościach *czułości*. W celu optymalnego doboru N , Doktorantka posłużyła się analizą krzywej ROC wraz z wyznaczeniem parametrów: *indeksu Youdena*, *punktu będącego najbliższym punktu (0,1)* oraz *prawdopodobieństwa zgodności*. **Warto w kilku krótkich zdaniach, przytoczyć definicje wykorzystanych współczynników, zwłaszcza, że w przypadku prawdopodobieństwa zgodności prawdopodobnie chodzi po prostu o parametr *dokładność*, którym Doktorantka bardzo często się posługuje.** Ostatecznie, na podstawie analizy wymienionych współczynników, Doktorantka wybiera wartość $N=3$, jako wartość optymalną. Istotnie, wszystkie parametry jakościowe są lepsze w stosunku do badań z rozdziału 7.

Ponieważ w powyższym podejściu Doktorantka zastosowała, w stosunku do analiz z rozdziału 7, dodatkową augmentację oraz przetwarzanie sekwencji zdań wraz z „algorytmem integrowania decyzji”, nie można jednoznacznie określić, które z czynników mają decydujący wpływ: augmentacja, czy przetwarzanie sekwencji zdań. W tym celu w podrozdziale 8.2.2 zostało dokonane dodatkowe badanie, w którym zastosowano model dostrojony z augmentacją danych dla zdania nr 1. **Choć na podstawie lektury rozdziału 8.2.2 trudno było ocenić o który model i które zdanie z rozdziału 7 chodzi. Dopiero po analizie rys. 8.7 i wartości macierzy błędów można było jednoznacznie określić do czego Doktorantka się odnosi.** Po wyznaczeniu macierzy błędów oraz obliczenia parametrów jakości, ostatecznie stwierdzono, iż to zastosowanie sekwencji zdań, w odróżnieniu do augmentacji zdania nr 1, ma decydujące znaczenie na polepszenie parametrów jakości klasyfikacji.

Na podstawie dotychczasowych przeprowadzonych badań można stwierdzić, iż:

- zastosowanie sekwencji zdań pochodzących z zapisu w ramach jednej rejestracji daje lepsze wartości wskaźników jakości klasyfikacji niż w przypadku tylko jednego zdania,
- zastosowanie strategii decyzyjnej - poszukiwanie optymalnej wartości parametru N , zwiększa *dokładność* do wartości 81% z odchyleniem standardowym 0,17% przy *precyzji* i *specyficzności* powyżej 80% i *czułości* 79%.

Powyższe wnioski tym samym udowadniają pierwszą z tez, mówiącą iż:

„użycie sekwencji obrazów zdań daje lepsze wyniki klasyfikacji binarnej w diagnostyce choroby Parkinsona z wykorzystaniem konwolucyjnej sieci neuronowej niż użycie obrazu zdania pojedynczego”.

Na podstawie lektury rozdziałów 4 – 8 nasuwa się kilka zagadnień, które mogłyby być, lub być może już były rozważane przez Doktorantkę.

- **Czy rozważane było wytrenowanie zaproponowanych sieci z wykorzystaniem zarówno autorskiej bazy KN_19 jak i bazy PaHaW? Czy jest to możliwe i czy mogłoby podnieść jakość klasyfikacji?**
- **Na podstawie analizy rys. 7.7 – 7.9 oraz tabeli 7.5 Doktorantka zauważyła, iż parametry jakości klasyfikacji są gorsze dla zdania nr 2. Doktorantka tłumaczy że ma to związek z „doświadczenia u pacjentów sytuacji stresowej (...) i potrzeby wykonania zapisu w nowych warunkach”. Niewątpliwie sytuacja ta ma związek ze stresem i przyzwyczajaniem się do pisanja na nieznanym sobie wcześniej urządzeniu. Ponieważ dopiero od trzeciego zdania pacjent oswaja się z badaniem, to być może pierwsze zdanie, lub pierwsze dwa zdania powinny zostać potraktowane jako treningowe, a sam proces trenowania i diagnostyki powinien uwzględniać dopiero kolejne zdania. Czy takie podejście mogłoby polepszyć jakość klasyfikacji?**

W rozdziale 9 kontynuowano prace nad dalszą poprawą parametrów jakości klasyfikacji wyodrębniając tym razem poszczególne słowa z sekwencji zdań. Jest to podejście nietypowe, w kontekście diagnostyki choroby Parkinsona z zastosowaniem konwolucyjnych sieci neuronowych. Ponadto rozważano badanie wyrazów, powstałych z zastosowaniem interpolacji pomiędzy poszczególnymi punktami, powstałymi w wyniku 200 Hz próbkowania tabletu. Doktorantka zauważyła, iż gęstość powstałych w wyniku próbkowania punktów różni się w funkcji prędkości pisanja wyrazów przez badanego. Istotnie może to nieść ze sobą dodatkową informację diagnostyczną, ze względu na spowolnienie procesu pisanja w wyniku rozwijającej się choroby. Dlatego też porównano parametry jakości klasyfikacji, uzyskanych za pomocą dwóch sieci AlexNet z zastosowaną interpolacją (standardowe podejście) oraz bez interpolacji. Obydwie sieci uwzględniały dostrojenie modelu i wytrenowanie za pomocą poszczególnych wyrazów w zdaniu. Tym razem parametr N , mający być zoptymalizowany w celu uzyskania minimalnej liczby przypisań do grupy chorych, mógł przyjąć wartości od 1 do 20. Dla każdej wartości tego parametru, oraz dla każdej z metod (z interpolacją i bez interpolacji), wygenerowano zestaw parametrów jakości klasyfikacji oraz obliczono parametry optymalizacyjne krzywych ROC. Najlepszy rezultat został osiągnięty w metodzie wizualizacji pisma bez interpolacji dla wartości parametru N równego 10, osiągając *dokładność* 0,9, *czułość* 0,96, *specyficzność* 0,83 oraz *precyzję* 0,88. **Wniosek ten ostatecznie potwierdza drugą z tez, postawioną w podrozdziale 1.2, która mówi, iż:**

„możliwa jest dalsza poprawa dokładności klasyfikacji binarnej w diagnostyce choroby Parkinsona z wykorzystaniem konwolucyjnej sieci neuronowej na drodze podziału obrazów sekwencji zdań na obrazy sekwencji słów”.

Doktorantka zauważyła, iż powyższy wniosek powstał w oparciu o rezultaty uzyskane z zastosowania metody bez interpolacji, podczas gdy badania w poprzednich rozdziałach dotyczyły analizy zdań z zastosowaniem interpolacji. Aby sprawdzić wpływ interpolacji na wyniki, Doktorantka wykonała serię badań z wykorzystaniem pojedynczego zdania i sekwencji zdań bez interpolacji. Badania wykazują niewielką poprawę niektórych parametrów jakości klasyfikacji, ale uzyskane wyniki były gorsze niż te, które osiągnięto w przypadku potwierdzającym tezę nr 2.

W rozdziale 10 Doktorantka podejmuje próbę dokonania dalszych optymalizacji modeli sieci podążając w dwóch kierunkach:

- z dalszym zastosowaniem sieci AlexNet, ale przy zamrożeniu poszczególnych warstw sieci głębokiej,
- stosując inne modele sieci konwolucyjnych: VGG16, GoogLeNet, ResNet18.

Pierwsze z podejść polegało na przetrenowaniu sieci AlexNet z dostrajaniem z sekwencją słów, ale przy zamrożeniu poszczególnych warstw. Przeprowadzono w sumie 6 badań: dla zamrożonej tylko warstwy conv1, następnie conv1 i conv2 aż do zamrożenia wszystkich warstw konwolucyjnych oraz jednej warstwy pełnego połączenia (fc6). Analiza wyników wykazuje, iż najlepsze rezultaty osiąga się przy zamrożeniu warstw conv1 i conv2. Wszystkie miary klasyfikacji nieznacznie wzrosły, stając się tym samym najlepszymi osiągniętymi parametrami w rozprawie: **dokładność - 0,94, precyzja - 0,94, czułość - 0,96 oraz specyficzność - 0,92.**

Drugie podejście, a mianowicie zastosowanie alternatywnej do AlexNet sieci konwolucyjnej, jest odpowiedzią na nasuwające się pytanie o to, dlaczego zastosowano tylko sieć *AlexNet* w kontekście tytułu rozprawy. Sam tytuł sformułowany jest dosyć szeroko i nie ogranicza sieci konwolucyjnych tylko do sieci AlexNet. Dlatego też, rozdział ten jest istotny z punktu widzenia globalnego spojrzenia na sieci konwolucyjne. Ostatecznie, najlepsze rezultaty, spośród proponowanych alternatywnych sieci, osiągnięto stosując sieć *ResNet18*: **dokładność - 0,92, precyzja - 0,93, czułość - 0,92 oraz specyficzność - 0,92.** Wyniki te są nieznacznie gorsze od uzyskanych przy stosowaniu sieci *AlexNet*, co potwierdza zasadność szczególnego uwzględnienia sieci *AlexNet* w grupie sieci konwolucyjnych.

Rozdział 11 stanowi podsumowanie, wykaz głównych elementów naukowych oraz wnioski końcowe.

3.3. Uwagi dyskusyjne

Uwagi dyskusyjne prezentowane były bezpośrednio w analizie części merytorycznej pracy (punkt 3.2 niniejszej recenzji) i zostały wyróżnione pogrubioną czcionką.

Pozostałe uwagi dyskusyjne.

1. Rozdział 4. Na rys. 4.3 przedstawiono schemat blokowy opracowanego stanowiska do rejestracji danych. Spośród pokazanych elementów, w badaniach Doktorantka stosuje laptop oraz tablet graficzny. Czy podczas badań były wykorzystywane pozostałe elementy stanowiska, np. kamery oraz mikrofon? W dysertacji brak jest wyników uzyskanych za pomocą tych elementów.

2. Czy możliwe jest uogólnienie uzyskanych wyników na inne choroby neurodegeneracyjne np.: chorobę Alzheimera, stwardnienie rozsiane czy też stwardnienie zanikowe boczne?

3.4. Ogólna ocena merytoryczna pracy

Praca ma charakter wybitnie interdyscyplinarny z pogranicza nauk medycznych, inżynierii biomedycznej oraz zagadnień cyfrowego przetwarzania obrazów i zastosowań sztucznej inteligencji. Praca jest napisana w sposób skondensowany z niezbędnym wprowadzeniem do narzędzi i zagadnień naukowych, wykorzystywanych w dysertacji. Na uwagę zwraca logiczny ciąg działań, poczynawszy od przygotowania studium literaturowego, poprzez nawiązanie polemiki z wybranymi publikacjami stanowiącymi punkt inicjujący oraz wyznaczającymi kierunek badań Doktorantki, aż po wykonanie autorskiego zestawu badań wraz z ich podsumowaniem. Kolejne etapy badań potwierdzały trafność obranego kierunku badawczego, wykazując wzrost konkretnych parametrów jakości klasyfikacji, aż do udowodnienia tez pracy. Wyniki te okazały się konkurencyjne do rezultatów badań uzyskanych przez autorów, z którymi Doktorantka nawiązała polemikę.

Poszukiwania naukowe Doktorantki zaowocowały nowymi, autorskimi aspektami badawczymi, które stanowią niewątpliwą wkład w rozwój dyscypliny automatyka, elektronika, elektrotechnika i technologie kosmiczne, a do najważniejszych z nich należą:

- stworzenie stanowiska laboratoryjnego oraz przygotowanie autorskiej aplikacji do akwizycji i przetwarzania danych oraz implementacji modeli sztucznych sieci neuronowych,
- przygotowanie metodyki przetwarzania obrazów w celu dostosowania ich do procesu trenowania i walidacji wyników pracy modeli sieci neuronowych,
- opracowanie autorskiej koncepcji badań z zastosowaniem konwolucyjnych sieci neuronowych na podstawie analizy sekwencji zdań oraz sekwencji wyrazów wraz z dostrojeniem sieci,
- przetestowanie możliwości zastosowania obrazów wyrazów złożonych z punktów (metoda bez interpolacji) jako zbiorów wejściowych sieci neuronowej,
- wykonanie badań z wykorzystaniem bazy danych, powstałej w wyniku współpracy z Warszawskim Uniwersytetem Medycznym, umożliwiających udowodnienie postawionych tez.

Zaprezentowane badania mają charakter rozwojowy, zmierzając w kierunku opracowania pomocniczego narzędzia diagnostycznego do wykrywania wczesnych stadiów choroby Parkinsona.

Ostatecznie Doktorantka przeprowadziła kompletny proces badawczy, poczynawszy od zdefiniowania problemu badawczego, przez zaproponowanie metodologii jego rozwiązania, implementacji a skończywszy na weryfikacji badań.

Uważam, że Doktorantka dysponuje odpowiednim warsztatem badawczym oraz posiada umiejętności publikowania wyników badań, czego dowodzą parametry bibliometryczne:

- 11 publikacji indeksowanych w bazie WoS w tym:
 - 6 jako pierwszy autor,
 - 3 publikacje w Q2 oraz 4 publikacje w Q4,
 - 4 publikacje w konferencjach międzynarodowych,
- 15 cytowań (12 bez autocytowań),
- indeks Hirscha 2.

4. WNIOSEK KOŃCOWY

Po zapoznaniu się z pracą doktorską pt.: „Strategie przygotowania obrazów pisma odręcznego na potrzeby rozpoznawania choroby Parkinsona z wykorzystaniem sieci konwolucyjnej” stwierdzam, iż

przedstawiona rozprawa spełnia

wymagania określone w ustawie "Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 r." (Dz.U. 2024 poz. 1571 - Obwieszczenie Marszałka Sejmu Rzeczypospolitej Polskiej z dnia 11 września 2024 r. w sprawie ogłoszenia jednolitego tekstu ustawy - Prawo o szkolnictwie wyższym i nauce).

W związku z powyższym

wnioskuję o przyjęcie rozprawy doktorskiej oraz o dopuszczenie

mgr inż. Kamili Białek do publicznej obrony.

Mirosław Szmojła