Dr hab. inż. Damian Grzechca, prof. Pol. Śl.
Politechnika Śląska
Wydział Automatyki, Elektroniki i Informatyki
Katedra Elektroniki, Elektrotechniki i Mikroelektroniki

Gliwice, 10 stycznia 2024 r.

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy: Zastosowanie metod uczenia maszynowego i pomiarów hiperspektralnych w diagnostyce uprawy Solanum lycopersicum

Autor rozprawy: mgr inż. Krzysztof Smykała

Promotor rozprawy: dr hab. inż. Michał Tomaszewski, prof. uczelni

Promotor pomocniczy: dr inż. Bogdan Ruszczak

Niniejsza recenzja została przygotowana na zlecenie Rady Naukowej Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Opolskiej, na podstawie decyzji Komisji Awansu z dnia 8.11.2023 i pisma z dnia 15.11.2023r. Przewodniczącego Rady Naukowej Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Pana Profesora doktora habilitowanego inżyniera Andrzeja Waindoka.

1. Informacje wstępne

Rozprawa doktorska mgr inż. Krzysztofa Smykały dotyczy metod automatycznej diagnostyki chorób z wykorzystaniem metod przetwarzania danych hiperspektralnych pozyskanych za pomocą spektroradiometru w zakresie 350-2500nm. Obiektem badań diagnostycznych była roślina Solanum lycopersicum, czyli pomidor zwyczajny w różnych odmianach. Rozprawa ma związek z pracą Doktoranta w przedsiębiorstwie produkującym żywność i jest jednym z elementów projektu finansowanego przez Narodowe Centrum Badań i Rozwoju.

Praca została podzielona na 11 rozdziałów, które obejmują ogólne informacje dotyczące uprawy, opis konwencjonalnych metod diagnostycznych, przegląd literatury, opis koncepcji proponowanej metody, opis eksperymentu pomiarowego, wybór algorytmów uczenia maszynowego, propozycję przetwarzania danych hiperspektralnych na potrzeby diagnostyki chorób w uprawie roślin oraz opis i analizę wyników zrealizowanych badań naukowych. Praca napisana jest w języku polskim i uważam, że łacińska nazwa występująca w tytule powinna być tłumaczona na nazwę polską, czyli pomidor zwyczajny lub wprost odniesienie się do chorób, np. “Zastosowanie metod uczenia maszynowego i pomiarów hiperspektralnych w diagnostyce wybranych chorób roślin na przykładzie uprawy pomidora zwyczajnego”.

Zauważyć również należy, że Autor rozprawy jest współautorem 5 publikacji, z czego 3 pozycje są cytowane w rozprawie (nr 31, 32 i 97). Całość dorobku Doktoranta w ujęciu liczba punktów MNiSzW wynosi 250 przy łącznym współczynniku wpływu IF=6,18. Mgr inż. Krzysztof Smykała jest pierwszym autorem w jednej publikacji w czasopiśmie Ambient Intelligence and Smart Environments (40 pkt.).

2. Zagadnienia naukowe rozpatrywane w rozprawie

(Jakie zagadnienia naukowe/badawcze jest rozpatrywane w pracy (cel i teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora?)

Praca doktorska ma charakter interdyscyplinarny i dotyczy automatycznej diagnostyki chorobowej roślin za pomocą metod uczenia maszynowego w konkretnym zastosowaniu, tj. uprawie pomidora zwyczajnego. Obiektem badań jest więc pomidor zwyczajny w dwóch odmianach Bonito i Polfast (choć

później Autor wprowadza kilka innych odmian), dla którego podjęto wysiłek opracowania automatycznej metody klasyfikacji wybranych chorób za pomocą pomiarów hiperspektralnych. Motywacją do podjęcia działań w tym zakresie jest zwiększenie produkcji żywności, bez znaczącej szkody dla środowiska naturalnego. Doktorant zajmuje się więc zwiększeniem produktywności w rolnictwie, które może odbywać się w wielu obszarach, od tych typowo chemiczno-rolniczych, jak środki ochrony roślin, nawozy, odżywki, poprzez aspekty mechaniczno-elektryczne, jak ciągniki oraz maszyny rolnicze, aż po inteligentne systemu diagnostyki i monitoringu tak maszyn, jak i upraw. I ten właśnie aspekt dotyczący automatycznej inteligentnej diagnostyki oraz monitoringu stanowią główny wątek naukowo-badawczy zawarty w rozprawie.

Tezę główną pracy przedstawiono w rozdziale 2.2:

Istnieje możliwość wykorzystania pomiarów hiperspektralnych reflektancji z zakresu fal o długości 350-2500nm i algorytmów uczenia maszynowego w procesie automatycznej diagnostyki Solanum lycopersicum dla potrzeb klasyfikacji wybranych jednostek chorobowych. Tezę główną uzupełniono o trzy tezy pomocnicze:

1. Możliwe jest określenie algorytmów klasyfikacyjnych, które z większą efektywnością pozwalają na rozpoznanie następujących chorób pomidora zwyczajnego: alternarioza, antraknoza, bakteryjna cętkowatość pomidora, septorioza, zaraza ziemniaka w pomidorze z zastosowaniem ograniczonego zbioru danych uczących.
2. Możliwe jest określenie specyficznych długości zakresu widma odpowiadających określonym jednostkom chorobowym Solanum lycopersicum.
3. Hiperspektralne pomiary reflektancji pozwalają na klasyfikację jednostek chorobowych Solanum lycopersicum w czasie szybszym niż metoda oceny wizualnej stosowana w praktyce.

O ile druga teza pomocnicza nie budzi większych zastrzeżeń, to teza pierwsza i trzecia są w mojej ocenie zbyt ogólne, a ich prawdziwość w kontekście metod uczenia maszynowego jest znana. Teza pierwsza zawierająca sformułowanie "z większą efektywnością", które można rozumieć, jako np. szybsza lub bardziej dokładna, co w ogólnym stwierdzeniu jest łatwe do udowodnienie. Brakuje tutaj również odniesienia, zapewne do metody wizualnej – tradycyjnej. Natomiast trzecia teza pomocnicza jest modyfikacją głównego celu pracy i również jest zawsze spełniona, ponieważ nie budzi wątpliwości fakt szybszej klasyfikacji za pomocą metod uczenia maszynowego lub ogólnie rozumianą sztuczną inteligencją w stosunku do metody oceny wizualnej, klasycznej, stosowanej w praktyce.

Reasumując, stwierdzić można, że rozprawa dotyczy opracowania metodologii do automatycznej klasyfikacji i identyfikacji wybranych chorób roślin na podstawie obrazów hiperspektralnych na przykładzie pomidora zwyczajnego w różnych odmianach. Doktorant postawił sobie za cel opracowanie modelu klasyfikatora, który w konkretnym zastosowaniu rozpozna stan chorobowy pomidora zwyczajnego. Praca ma silne przełożenie na produktywność i powinna znaleźć zastosowanie praktyczne w przedsiębiorstwie, w którym pracuje Doktorant.

3. Analiza źródeł, stan wiedzy i zastosowań w przemyśle

(Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle?)

Lista publikacji zawiera 103 pozycje naukowe, które dobrze komponują się z zakresem prac realizowanym przez Doktoranta. W rozdziale 3 opisane są informacje na temat obiektu badań, tj. uprawy wykorzystanej w procesie pozyskiwania zbioru danych hiperspektralnych, wskazano wybrane, typowe jednostki chorowe, jednocześnie uwypuklając skutki braku prawidłowej diagnostyki i reagowania na zagrożenia chorobowe. Następnie, w rozdziale 4 przedstawione zostały różne typy danych spektralnych wykorzystywanych w nieinwazyjnych metodach diagnostycznych, przeanalizowano aktualny stan

wiedzy w zakresie metod łączących zalety danych hiperspektralnych oraz uczenia maszynowego w celu prawidłowej diagnostyki chorób roślin.

Przegląd aktualnego stanu wiedzy związany z diagnostyką chorób roślin z wykorzystaniem nieinwazyjnych metod teledetekcyjnych, takich jak obrazy w przestrzeni barwnej RGB, obrazowanie wielospektralne, obrazowanie hiperspektrlane, w tym pomiary hiperspektralne został przedstawiony w rozdziale 4. Doktorant systematyzuje choroby roślin ze względu na ich pochodzenie, tj. choroby pochodzenia wirusowego, grzybowego oraz bakteryjnego. Wskazuje, że roślinom zagrażają również niedobory i nadmiar składników odżywczych, szkodniki oraz rośliny pasożytnicze. Analiza wykonana jest wielopłaszczyznowo ze szczególnym naciskiem na metodologię pozyskiwania danych, ich częstość, zakres spektralny, parametr DPI (dni po inokulacji). Doktorant słusznie stwierdza, że określone uprawy, jednostki chorobowe oraz położenie geograficzne wymagają dedykowanego podejścia diagnostycznego i dlatego należy zbudować model diagnostyczny pod wskazany w pracy obiekt, składający się z szeregu elementów, procedury, której wejściem są dane hiperskalrane, a wyjściem stan chorobowy obiektu, w tym wypadku pomidora zwyczajnego.

Stwierdzić więc należy, że do tej pory nie udało się wypracować narzędzia pozwalającego na diagnozowanie chorób roślin z dużą dokładnością za pomocą danych satelitarnych, co wynika między innymi z ograniczonej rozdzielczości spektralnej dostępnych instrumentów obrazujących (np. Sentinel-2). Dużą szansę na rozwój w tej dziedzinie oferuje technologia obrazowania hiperspektralnego, która pozwala na badanie obiektów z wyższą, o rząd wielkości, rozdzielczością spektralną w stosunku do obrazowania multispektralnego. Dzięki postępowi technologicznemu i nieustannej miniaturyzacji oraz intensywnej popularyzacji dronów w ostatnich latach możliwe jest wykonywanie lotniczych zdjęć hiperspektralnych w rozdzielczości do paru milimetrów na piksel, co otwiera nowe możliwości analityczne dla zjawisk rozważanych w rozprawie. Zanim jednak odpowiednie metody będą upowszechnione na szeroką skalę wymagane jest opracowanie metod i modeli diagnostycznych w laboratorium. Dlatego celem pracy jest stworzenie metody pozwalającej na wstępną, automatyczną diagnostykę obiektu za pomocą danych hiperspektralnych oraz algorytmów uczenia maszynowego na przykładzie diagnostyki wybranych chorób w uprawach roślin z wykorzystaniem procedur kontrolowanych w laboratorium. W związku z tym, wpływ wyników prac badawczych na implementację i wdrożenie jest jednym z celów pracy, co powinno pozwolić na zautomatyzowanie procesu diagnostycznego i zwiększyć produktywność wskazanej uprawy pomidora zwyczajnego.

Podsumowując, przedstawiona analiza aktualnego stanu wiedzy jest przeprowadzona właściwie i dobrze odzwierciedla aktualne trendy w obszarze stosowania metod uczenia maszynowego do automatycznej diagnostyki chorobowej roślin.

4. Metody i założenia użytych metod

(Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?)

Założeniem podstawowym pracy, wynikającym z tytułu i postawionego problemu jest obiekt badań (pomidor zwyczajny) w kilku odmianach Benito, Polfast, Dyno, Guadivia, który oceniano z pomocą algorytmów uczenia maszynowego pod kątem zakażenia pięcioma różnymi chorobami, z czego 4 to choroby grzybowe: Alternarioza lub sucha plamistość liści, Antraknoza, Septorioza, a jedna to choroba bakteryjna: bakteryjna cętkowatość. W celu przygotowania danych do dalszej analizy przeprowadzono, starannie zaprojektowany eksperyment laboratoryjny, który zawiera wartościowe dane wejściowe dla modeli uczenia maszynowego. Zauważyć należy, że zbiór obiektów został podzielony na dane kontrolne i chorobowe, rośliny przetrzymywano w komorach wegetacyjnych, w których utrzymywano stałą

temperaturę 18-20°C (z dokładnością do 0,5°C), wilgotnością 60-80% (z dokładnością do 5%) oraz doświetleniem (minimum 200 $\mu mol \cdot m^{-2} \cdot s^{-1}$) w 16-godzinnym cyklu naświetlenia. Pomiary wykonywano w ciągu 33 kolejnych dni w okresie od DPI = 0 do DPI = 63. Bardzo klarownie przedstawiono proces pozyskiwania danych na diagramie Gantta (rys. 13).

Zbiór danych wykorzystywanych w badaniu obejmował pomiary hiperspektralne roślin dwóch odmian uprawy Solanum lycopersicum: Benito oraz Polfast (Bejo Zaden B.V.). Dla każdej z odmian przygotowano 36 roślin doświadczalnych, które zarażono pięcioma różnymi patogenami w grupach po 6 sztuk. Pozostałe 6 roślin było obiektami kontrolnymi, niezakażonymi żadnym patogenem (tzw. grupa kontrolna). Każdy z patogenów, zarówno pochodzenia bakteryjnego, jak i grzybiczego, został wyizolowany i wyhodowany na 6 niezależnych płytkach. Warunki hodowli i przechowywania patogenów były stałe, tj. temperatura powietrza 20±1°C, a wilgotność wynosiła co najmniej 70% utrzymywane były na tym samym poziomie. Rośliny zakażone zostały poprzez zanurzenie ich w inokulum rozcieńczonym do wskazanego stężenia 105 do 106 zarodników na 1 ml. Obecność patogenów była potwierdzana na liściach i/lub owocach za pomocą analizy mikroskopowej, izolacji patogenu i określeniu gatunków. Fragmenty roślin z objawami choroby wycinano z liści lub owoców, po czym je dezynfekowano powierzchniowo przez 3 min w 2% podchlorynie sodu. Fragmenty suszono, krojono na 4 mniejsze fragmenty i wykładano na pożywkę PDA. Po 7-10 dniach oceniano wyrosłe kultury i przypisywano do gatunków sprawców.

Do wykonywania pomiarów hiperspektralnych wykorzystany został spektroradiometr ASD FieldSpec 4 Hi-Res. Urządzenie wykonuje pomiar reflektancji, transmisji lub energii promieniowania elektromagnetycznego, tj. radiancji (luminacja energetyczna) lub irradiancji (natężenie promieniowania), badanej próbki w zakresie spektralnym 350-2500nm z rozdzielczością spektralną 3nm oraz 8nm odpowiednio dla długości 700nm oraz 1400nm i 2100nm. Dane analizowane przez Doktoranta są pomiarami reflektancji, reprezentującymi ilość światła odbitego w skali od 0 do 1 w poszczególnych kanałach spektralnych badanego zakresu. Obiekt oświetlano lampą halogenową (dwoma lampami) Eurostar Reflektor MR16 firmy Ushio, a pomiary wykonano dla wysokości równej 5 cm, 30 cm i 60 cm, co bezpośrednio przekłada się na średnicę skanowanego obszaru odpowiednio 2,2 cm, 13,3 cm, oraz 26,6 cm, co znacząco przekracza przeciętne wymiary liści uprawy pomidora zwyczajnego. Z tego powodu wykorzystano wyłącznie pomiar dla wysokości 5cm.

W celach automatycznej diagnostyki danych hiperspektlarnych mgr inż. Krzysztof Smykała proponuje 6 różnych klasyfikatorów: liniowy klasyfikator wektorów nośnych, regresję w kilku odmianach, np. regresja liniowa z regularyzacją Tichonowa, regresję logistyczną (klasyfikator binarny), oraz metody agregacji klasyfikatorów tj. las losowy (random forest) oparty na drzewach decyzyjnych w odmianach zaimplementowanych za pomocą pakietów LightGBM oraz XGBoost. Algorytmy te przystosowuje do swoich celów i danych pomiarowych, które w konsekwencji stanowią kompletną procedurę diagnostyczną, w której najważniejszą rolę odgrywa budowa modelu klasyfikatora opartego na uczeniu maszynowym.

Stwierdzić zatem można, że eksperyment został przygotowany z należytą starannością z dbałością o możliwość wieloaspektowej analizy otrzymanych wyników, a wykorzystane narzędzie, tj. klasyfikatory oparte na uczeniu maszynowym są właściwe.

5. Oryginalność rozprawy, samodzielny i oryginalny dorobek Autora
(Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy i poziomu techniki reprezentowanych przez literaturę światową?)

Rozprawa doktorska mgr inż. Krzysztofa Smykały dotyczy automatycznej diagnostyki chorobowej uprawy pomidora zwyczajnego na podstawie danych widmowych z wykorzystaniem metod uczenia maszynowego. Przeprowadzone badania nie tylko potwierdzają przydatność danych hiperspektralnych w procesie diagnostycznym i wykrywania niektórych chorób uprawy pomidora, ale także w procesie wczesnego wykrywania chorób tej uprawy. Zaproponowana w pracy metoda przetwarzania danych hiperspektralnych została podzielona na etapy, co w konsekwencji pozwoliła na uzyskanie poprawy wyników klasyfikacji, również poprzez odpowiednie "dostrojenie" procesu diagnostycznego, a w zasadzie dostrojenie poszczególnych klasyfikatorów zaproponowanych w rozprawie. W związku z tym, do najważniejszych osiągnięć Autora zaliczyć można:

- opracowanie metody automatycznej diagnostycznej wykorzystującej elementy uczenia maszynowego do wykrywania chorób na podstawie danych hiperspektralnych, w skład której wchodzą:
    - przetwarzanie danych hiperspektralnych głównie poprzez redukcję wymiarowości danych wejściowych,
    - wyznaczenie, w tym identyfikacja i weryfikacja, zakresów spektralnych w przestrzeni bliskiej podczerwieni i podczerwieni (VNIR) oraz krótkofalowej podczerwieni (SWIR) istotnych w klasyfikacji chorób obiektu badań,
    - koncepcję tworzenia danych hiperspektralnych z uwzględnieniem zmian dynamicznych w czasie, wymaganych dla dalszego procesu budowy modelu klasyfikatora,
    - opracowanie modeli diagnostycznych poprzez wybór klasyfikatorów prostych i złożonych opartych na metodach uczenia maszynowego dedykowanych dla poprawnej diagnostyki chorobowej pomidora zwyczajnego.

Szczególnie należy podkreślić staranność przeprowadzenia eksperymentu laboratoryjnego, który umożliwił przygotowania danych wejściowych pod tworzone modele, redukcję wymiarowości oraz analizę zakresów spektralnych, które umożliwiają stworzenie modeli klasyfikatorów metodami uczenia maszynowego. Na tej podstawie stwierdzić można, że przedstawione przez Doktoranta rozwiązanie stanowi kompletne, dedykowane rozwiązanie naukowe, które zaliczyć można do dyscypliny automatyka, elektronika, elektrotechnika i technologie kosmiczne.

6. Umiejętność przedstawienia uzyskanych wyników
(Czy autor wykazał umiejętność poprawnego i przekonywującego przedstawienia uzyskanych przez siebie wyników - zwięzłość, jasność, poprawność redakcyjna rozprawy.)

Redakcja samej rozprawy nie budzi zastrzeżeń, jest napisana starannie i nie wymaga poprawek, choć Doktorant nie ustrzegł się literówek i drobnych błędów stylistycznych, czy języka żargonowego, np.: „wydobycie cech sygnałów" zamiast np. „wyznaczenie ….". Doktorant bardzo dobrze opisuje sposób przeprowadzenia eksperymentu, ciąg przyczynowo skutkowy, jednak w mojej ocenie opis ten powinien mieć raczej charakter metodologiczny, czyli propozycji rozwiązania problemu braku zautomatyzowanych modeli diagnostycznych dla roślin uprawnych z wykorzystaniem analizy hiperspektralnej oraz metod uczenia maszynowego, a ponadto opis proponowanej metody (rozwiązania) powinien być na wyższym poziomie ogólności. Oczekiwałbym raczej od tego typu pracy, pokazania pewnej metodyki, z której mogą korzystać inni potencjalni odbiorcy, a sam opis swojego eksperymentu powinien służyć jako ilustracja proponowanej metody. W mojej ocenie w pracy brakuje kompleksowego schematu blokowego lub opisu proceduralnego proponowanej metody, co nie tylko umożliwiłoby płynne przejście przez poszczególne rozdziały pracy, ale również pokazałoby generalną koncepcję rozwiązania problemu. Diagram z rys. 41 jest tylko częściową informacją związaną z proponowaną koncepcją i stanowi raczej opis eksperymentu (doktorant pisze: "schemat postępowania w poniższym

eksperymencie”). Opis ten zaliczam raczej do pewnej niezręczności językowej, a schemat na rys. 41 traktuję jako rekomendowaną procedura przetwarzania danych.

Wyniki eksperymentów przedstawione są prawidłowo i w oparciu o właściwie przyjęte założenia, choć w mojej ocenie brakuje dyskusji wyników w kontekście wieloaspektowym, wykorzystania metod uczenia maszynowego, o czym piszę więcej w punkcie uwagi krytyczne.

7. Przydatność rozprawy w naukach inżynieryjno-technicznych

Wyniki prac zawartych w rozprawie z zakresu analizy sygnałów hiperspektralnych i diagnostyki obiektów organicznych z wykorzystaniem metod uczenia maszynowego częściowo powstały w ramach realizacji przez QZ Solutions Sp. z o.o. projektu dofinansowanego z funduszy Narodowego Centrum Badań i Rozwoju (POIR.01.01.01-00- 1317/17) przy współpracy z Wydziałem Ogrodnictwa i Warzywnictwa Uniwersytetu Przyrodniczego w Poznaniu oraz Centrum Badań Kosmicznych Polskiej Akademii Nauk. Zaznaczyć należy, że wykorzystanie metod uczenia maszynowego do klasyfikacji nie jest zagadnieniem nowym, metody te są niejako dedykowane do tego typu problemów. Największą trudnością jest jednak przygotowanie odpowiednich danych wejściowych tak, aby modele dostarczały pożądanej informacji wyjściowej. W aspekcie przygotowania danych hiperspektralnych, praca posiada istotną przydatność w diagnostyce chorób pomidora zwyczajnego, którą Doktorant potwierdził badaniami. Stwierdzić można, że rozprawa doktorska mgr. inż. Krzysztofa Smykały jest interdyscyplinarna i zawiera wątki nauk rolniczych oraz inżynieryjno-technicznych. Uważam, że specyfika pracy jest elementem nie tylko ważnym praktycznie, ale przedstawia logiczny sposób łączenia, przetwarzania (w tym redukcję) danych hiperspektralnych pod określone algorytmy uczenia maszynowego w kontekście specyficznego zastosowania w klasyfikacji chorobowej roślin uprawnych. W ten sposób powstaje dedykowana procedura, która prowadzi do zautomatyzowanej diagnostyki chorobowej obiektu badań.

8. Uwagi szczegółowe i dyskusyjne

Podczas czytania rozprawy nasunęły się uwagi, głównie o charakterze polemicznym i dyskusyjnym, które przedstawione są poniżej:

Ocena jakości poszczególnych metod opracowanych przez Autora oparta została na wskaźniku F1, który uwzględnia precyzję i czułość, a zdefiniowany jest wzorem (7.23). Autor postanowił tak dobrać parametr, aby czułość miała większy wpływ na wskaźnik F1, a następnie arbitralnie przyjął zadawalającą wartość współczynnika F1 na poziomie 0.8. Zauważyć należy, że operowanie pojedynczym współczynnikiem jest wygodne, jednak znacząco ogranicza dyskusję uzyskanych wyników. W mojej ocenie wartościowe byłoby również określenie procentowej dokładności wykonanej diagnozy, czyli operowanie wprost na wzorach 7.19, 7.20, 7.21. Prowadzi to moim zdaniem do uproszczonej analizy otrzymanych wyników, ponieważ otrzymane modele, ich skuteczność działania jest odnoszona do wskaźnika F1 ustalonego arbitralnie i odnoszący się do badań laboratoryjnych. Operowanie na wzorach 7.19, 7.20 i 7.21 pozwoliłoby dostroić algorytmy pod oczekiwanie rezultaty, np. wczesnej diagnozy.

Innym elementem dyskusyjnym jest sposób pomiaru lub odniesienia procentowej zawartości zarażonej powierzchni do wielkości obszaru badanego, który ma przecież wpływ na wartości w poszczególnych pasmach spektralnych. Wartości te są zależne od analizowanej powierzchni, która nie jest jednorodna, w związku z tym kluczowe wydaje się pytanie o odległość od analizowanego obszaru w aspekcie wdrożeniowym np. pułap drona. W jaki sposób oszacować wpływu obszaru

objętego badaniem na jakość predykcji oraz klasyfikacji wskazanych chorób? Wydaje się, że Autor zauważa ten problem redukując odległości do jednej najbliższej 5cm wyjaśniając, że występuje tutaj najmniejsze zaszumienie obrazem tła. Zakładając praktyczną implementację i wdrożenie, czy zdaniem Autora można określić procentowy udział obszaru chorobowego do całości obszaru, dla którego metoda prawidłowo zdiagnozuje chorobę? W jakim okresie od inokulacji jest to możliwe (chodzi o zależność udziału obszaru chorobowego do całości obszaru, dla którego wykonano pomiar widmowy)? Czy spodziewa się, że również w tym aspekcie klasyfikator regresji logistycznej będzie najlepszym modelem diagnostycznym?

Czytając rozprawę można odczuć pewien niedosyt związany z pobieżną analizą i dyskusją wyników dla poszczególnych algorytmów klasyfikacyjnych. W rozdziale 7.1.3 przedstawiono zakresy rozważanych hiperparametrów, niestety nie znalazłem w pracy końcowych modeli, wyników dostrajania, jest informacja o tym, że po dostrojeniu uzyskano poprawę wyników działania, co jest oczywiste, ponieważ proces ten jest wykonywany w tym właśnie celu. Cenna byłaby informacja, dostarczona czytelnikowi o wpływie wybranych parametrów na końcowy wynik klasyfikacji. Innymi słowy, w rozważanym problemie, który z parametrów należy w pierwszej kolejności dostrajać i dlaczego? Czy w procesie strojenia Doktorant zauważył prawidłowości związane polepszeniem docelowej klasyfikacji?

Zastanawiającym jest dla mnie proces filtracji, który zgodnie z informacją wstępną powinien być przedstawiony w rozdziale 8, ale za wyjątkiem ogólnych stwierdzeń oraz wystąpienia tego słowa w podsumowaniu, na próżno szukać sposobu filtracji danych wejściowych w rozumieniu przetwarzania sygnałów, czy obrazów. Naciągane wydaje się włączenie procesu "czyszczenia danych" do filtracji w wyżej wskazanym rozumieniu. W jaki sposób w takim razie należy rozumieć słowo filtracja? Czy Doktorant faktycznie wykorzystał filtrację, czy to należy to rozumieć, jako czyszczenie danych?

Uwagi szczegółowe zauważone podczas czytania rozprawy:

a) Jeśli coś jest widoczne, to nie ma potrzeby dodawania "gołym okiem" - str. 24, 60, 62, 63
b) DPI - tłumaczone jest raz w jęz. Polskim innym razem w jęz. Angielskim.
c) Str. 44: Do prawidłowego doświetlenia badanych obiektów wykorzystano sztuczne źródło światła w postaci dwóch identycznych lamp halogenowych, widocznych na rys. 20. (błędny odnośnik).
d) Część tabel zawiera nazwy anglojęzyczne, które powinny być tłumaczone na jęz. polski. Doktorant dodatkowo umieścił tłumaczenia nazw anglojęzycznych - brakuje tutaj konsekwencji przez co całość pracy jest niespójna.
e) Część rysunków jest zbędna, np. Rys. 28, rys. 29, rys. 31 i inne podobne wykresy słupkowe. Informacja o liczbie wykonanych pomiarów w postaci tabel byłaby wystarczająca, a sama wizualizacja zwiększa objętość pracy, nie wnosząc dodatkowych informacji, czy rozszerzając dyskusję ogólną Doktoranta.
f) Autor w pracy dość swobodnie operuje terminologią wymiar, wielkość, a nawet cecha, które czasem występują wymiennie, np. W rozdziale 9.3.1 kontekst odnosi się do czegoś innego niż zawiera ostatnie zdanie. Odnieść można wrażenie, że wymiar i wielkość są dla Doktoranta synonimami.
g) Niespójne są też informacje w różnych miejscach na temat odległości od badanego obiektu, ponieważ w jednym miejscu pracy Autor pisze o redukcji przypadków testowych do odległości wynoszącej 5cm, a w tabelach np. w rozdziale 8, pozostałe wysokości również są wpisane.
h) Str. 91: "ogrywa" zamiast "odgrywa"
i) W rozdziale 9.5 dotyczącym klasyfikacji wieloklasowej nie przedstawiono informacji liczbowych dotyczących zbiorów, klas, a tabele od 22 do 24 zapewne w domyśle operują na

wskaźniku F1. Zastanawiający jest brak szerszej dyskusji wyników i w jaki sposób interpretować wskaźnik F1 dla klasyfikacji wieloklasowej.

9. Wniosek końcowy

W świetle powyższych uwag stwierdzam, że opiniowana rozprawa doktorska, pomimo uwag o charakterze dyskusyjnym i polemicznym, spełnia warunki określone w art. 13 *Ustawy o stopniach naukowych i tytule naukowym* z dnia 14 marca 2003 r. z późn. zm. (Dz.U. 2017, poz. 1789) oraz ustawy *Prawo o szkolnictwie wyższym i nauce* z dnia 20 lipca 2018r. i wnioskuje o dopuszczenie mgr. inż. Krzysztofa Smykały do dalszych etapów przewodu doktorskiego.