Prof. dr hab. inż. Jerzy Stefanowski
Instytut Informatyki Politechniki Poznańskiej

11 sierpnia 2022

Recenzja osiągnięcia habilitacyjnego, dorobku naukowego, dydaktycznego i organizacyjnego dra Bartosza Zielińskiego w związku z postępowaniem o nadanie stopnia naukowego doktora habilitowanego w dyscyplinie informatyka techniczna i telekomunikacja

Niniejsza recenzja została przygotowana w odpowiedzi na pismo Przewodniczącego Komisji ds. stopni naukowych w dyscyplinie informatyka techniczna i telekomunikacja P.Wroc. prof. dr hab. inż. Michała Woźniaka, przesłane w związku z postępowaniem o nadanie stopnia naukowego doktora habilitowanego nauk inżynieryjno-technicznych Panu dr Bartoszowi Zielińskiemu oraz powołaniem mnie w skład komisji habilitacyjnej jako recenzenta.

Pan dr Bartosz Zieliński jest pracownikiem Instytutu Informatyki i Matematyki Komputerowej Uniwersytetu Jagielońskiego. Tamże ukończył studia magisterskie na kierunku informatyka w 2007 roku i od 2009 był zatrudniony jako asystent, a następnie jako adiunkt. Jego doktorat, prowadzony pod opieką prof. M.Skomorowskiego dotyczył techniki wykrywania wybranych objawów chorób reumatycznych na podstawie zdjęć rentgenowskich rąk (obrona 2012 w IPPT PAN). Analizując dostarczoną dokumentacje można zauważyć długoletnie zainteresowanie Habilitanta problematyką rozpoznawania obrazów oraz zastosowań medycznych lub biologicznych.

1. Informacje ogólne

Przedmiotem recenzji jest osiągniecie naukowe pt. „*Wyjaśniane i interpretowalne uczenie maszynowe w analizie obrazów medycznych*". W skład osiągnięcia Pan dr Zieliński włączył 9 prac opublikowanych zarówno w czasopismach naukowych (3 artykuły), jak i materiałach znanych konferencji naukowych. Lista tych prac jest następująca (w kolejności wskazanej w autoreferacie i dostarczonych materiałach oraz jego notacji):

[A1] Dominika Basaj, Witold Oleszkiewicz, Igor Sieradzki, Michał Górszczak, Barbara Rychalska, Tomasz Trzciński, Bartosz Zieliński. Explaining Self-Supervised Image Representations with Visual Probing.International Joint Conference on Artificial Intelligence (IJCAI), str. 592–598, 2021.

[A2] Bartosz Zieliński, Michał Górszczak. What Pushes Self-Supervised Image Representations Away? International Conference on Neural Information Processing (ICONIP), Communications in Computer and Information Science, str 514–521, 2021.

[A3] Bartosz Zieliński, Anna Plichta, Krzysztof Misztal, Przemysław Spurek, Monika Brzychczy-Włoch, Przemysław Spurek, Dorota Ochońska. Deep learning approach to bacterial colony classification. PloS one 12 (9), str. 1–14, 2017.

[A4] Bartosz Zieliński, Agnieszka Sroka-Oleksiak, Dawid Rymarczyk, Adam Piekarczyk, Monika Brzychczy-Włoch.Deep Learning approach to description and classification of fungi microscopic images. PloS one 15 (6), str 1–16, 2020.

[A5] Adriana Borowa, Dawid Rymarczyk, Dorota Ochońska, Monika Brzychczy-Włoch, Bartosz Zieliński. Deep learning classification of bacteria clones explained by persistence homology. International Joint International Joint Conference on Neural Networks (IJCNN), str. 1–8, 2021.

[A6] Dawid Rymarczyk, Łukasz Struski, Jacek Tabor, Bartosz Zieliński. ProtoPShare: Prototypical Parts Sharing for Similarity Discovery in Interpretable Image Classification. Proceedings of ACM SIGKDD Conference on Knowledge Discovery & Data Mining, str. 1420–1430, 2021.

[A7] Dawid Rymarczyk, Adriana Borowa, Jacek Tabor, Bartosz Zieliński.

WPŁYNĘŁO
24-08-2022
RDN ITiT/222/2022

Kernel Self-Attention for Weakly-supervised Image Classification using Deep Multiple Instance Learning.Winter Conference on Applications of Computer Vision (WACV), str. 1720–1729, 2021.
[A8] Bartosz Zieliński, Michał Lipiński, Mateusz Juda, Matthias Zeppelzauer, Paweł Dłotko. Persistence Bag-of-Words for Topological Data Analysis. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJ- CAI), str. 4489–4495, 2019.
[A9] Bartosz Zieliński, Michał Lipiński, Mateusz Juda, Matthias Zeppelzauer, Paweł Dłotko. Persistence codebooks for topological data analysis. Artificial Intelligence Review 54 (3), str. 1969–2009, 2021.

Wszystkie powyższe prace są współautorskie. W przekazanych materiałach znajdują się oświadczenia współautorów nt. ich udziału. Opis w autoreferacie wskazuje na udział tzw. wkładu Pana dra Zielińskiego w zakresie od 25% do 70%. W pięciu pracach ten udział przekracza 50%. Biorąc pod uwagę, że większość prac jest silnie wielo-autorska (czasami 7 osób) powyższe udziały są relatywnie wysokie. W opisach prac i swojego udziału, Habilitant często pisze „zdefiniowałem problem badawczy i kierowałem pracami (..)", albo „nadzorowałem postępy prac" czy „opracowałem metodę klasyfikacji (..) — można więc uznać na tej podstawie, że w większości prac pełnił rolę tzw. wiodącego, głównego autora. Patrząc na listę współautorów mogę stwierdzić, że większość pochodzi z macierzystego ośrodka UJ lub współpracowników medycznych także z UJ. Widzę że p. dr Zielińskiemu udało się nawiązać współpracę z prof. T. Trzcińskim z Politechniki Warszawskiej oraz prof. M. Zeppelzaurem z Austrii. W autoreferacie dostrzegłem także informacje o innej współpracy międzynarodowej oraz publikacje z wieloma innymi współautorami, co także oceniam bardzo dobrze.

2. Ocena osiągnięcia habilitacyjnego

Analizując zawartość przedstawionych prac oraz ich opis w autoreferacie uważam, że przedstawiają one różne rozwiązania z zakresu głębokiego uczenia sieci neuronowych przeznaczonego głównie do rozpoznawania i klasyfikacji obrazów (w wielu przypadkach biomedycznych, gdzie ich specyfika jest twórczo wykorzystana w operacjach ich przetwarzania). W zakresie wykorzystanych metod uczenia sieci, są to różne architektury głębokich sieci neuronowych obecnie często rozważanych w literaturze, łączonych w bardziej złożone systemy, gdzie integruje się je z innymi metodami przeznaczonymi do przetwarzania obrazów. Habilitant często korzysta z metod związanych z tzw. modelem „worka" wizualnych słów - cech w obrazie (ang. bag-of-visual-words model) popularnego w wizji komputerowej albo z mechanizmów uwagi poznawczej (ang. attention), realizowanych w głębokim uczeniu się. Ponadto występują koncepcyjnie inne rozwiązania, przede wszystkim w sieciach prototypowych dla wyjaśnialności predykcji oraz tzw. diagramach persystencji w metodach topologicznych analizy danych.
Ogólnie są to różne metody i mogę się zgodzić z Panem dr Bartoszem Zielińskim, który w tekście autoreferatu (rozdział 3.1 strona 8) podzielił je na cztery w sumie rozłączne metodycznie kategorie prezentowanych osiągnięć:

- Prace [A1,A2] dotyczące analizy reprezentacji obrazów otrzymywanych przez sieci z samo-nadzorujące się oraz adaptacje wizyjnych klasyfikatorów diagnostycznych.
- Prace [A3,A4, A5] wprowadzające złożone systemy neuronowe w zastosowaniach mikrobiologicznych do klasyfikacji wybranych bakterii lub grzybów (w trzech różnych wariantach)
- Prace [A6 i A7] bezpośrednio związane z wyjaśnialnością predykcji sieci neuronowych, tj. dwie w sumie mocno różne propozycje: uogólnienie sieci prototypowych z współdzieleniem prototypów i redukcją ich liczby oraz wykorzystania neuronowego mechanizmu skupiania uwagi do zadania klasyfikacji wielokrotnych instancji (ang.

MIL). Dla mnie są to jednak dwie zupełnie niezależne metody raczej nie nadające się do połączenia w jedną grupę metod.

- Prace [A8 i A9] w zakresie topologicznej analizy danych wysoce wielowymiarowych, gdzie wprowadza się nowe adaptacyjne reprezentacje diagramów persystencji.

W stosunku do wszystkich z tych prac Pan dr Bartosz Zieliński zaproponował w autoreferacie wspólne określenie tematyki opisywanych badań jako opracowanie wyjaśnianych metod uczenia maszynowego (patrz bardziej szczegółowo przedstawienie przez habilitanta w podrozdziale 3.1 wskazanie osiągnięcia habilitacyjnego w treści autoreferatu), co w mojej opinii jest jednak dyskusyjne analizując zawartość konkretnych prac [A1-A9], gdyż nie wszystkie one są na to ukierunkowane. Powrócę do tego wątku po krótkim omówieniu oryginalnych dokonań w tych pracach, co przedstawiam poniżej.

W pierwszej z tych kategorii najoryginalniejsza i ciekawsza jest praca [A1] gdzie autorzy wprowadzili własną adaptacje wizyjnych klasyfikatorów diagnostycznych (oryginalnie pochodzących z NLP, a tutaj realizujących operacje nazwaną po ang. visual probing tasks) do analizy reprezentacji obrazów, uzyskanej za pomocą metod samo-nadzorujących się (ang. self-supervised). W moim zrozumieniu zaproponowanego podejścia istotnymi elementami jest podział obrazów na obszary tzw. superpikseli oraz tworzenie ich reprezentacji, które są przypisywane do zadanych cech obrazu, tzw. słów wizualnych - co może być dalej analizowane przez zaadaptowane klasyfikatory diagnostyczne (gdzie autorzy pokazują związki z odpowiednikami wizualnymi słów, zdań). Praca zawiera także eksperymentalną ocenę proponowanego podejścia i porównanie z innymi. Praca [A2] jest kontynuacją pierwszej, gdzie celem było eksperymentalne zbadanie które z cech obrazu zawarte w ich cząstkowych reprezentacjach mają większe znaczenia. Autorzy zastosowali specyficzne podejście do relatywnie naturalnej modyfikacji obrazu i oceny eksperymentalnej wpływu tej modyfikacji. Pozwoliło im to na eksperymentalne sprawdzenie, które cechy-fragmenty obrazu / wizualne słowa mają większy wpływ dla rozważanych reprezentacji - sam eksperyment jest dość ciekawy i powiązany z dyskusją różnić pomiędzy rozważanymi metodami samo-nadzorującymi.

Druga grupa trzech prac [A3,A4 i A5] jest związane z prowadzonymi zastosowaniami rozpoznawania obrazów w diagnostyce mikrobiologicznej (przy współpracy z naukowcami z Kolegium Medycznego UJ). Mają one sensownie przedstawioną motywacje przyspieszenia procedury diagnostycznej w sytuacjach zakażenia pacjentów i poprawy rozpoznawania i skutecznej klasyfikacji różnych rodzajów bakterii lub później w pracy [A4] także rodzajów grzybów. Za najważniejszą metodycznie uważam pracę [A3], gdzie autorzy wprowadzali swoje podejścia wykorzystujące głębokie konwolucyjne sieci neuronowe CNN. Są one inspirowane wcześniejszymi pracami innych autorów i wykorzystują poszukiwanie małych obszarów (nazywanych łatami - ang. patches) w obrazie, tworzenia głębokich reprezentacji w końcowych warstwach CNN, które są wyjęte z sieci do dalszego przetwarzania - tutaj autorzy wykorzystali metodę wektorów Fisher'a i stworzenia nowej wektorowej reprezentacji (z elementami wizualnych słów, kodów), która ostatecznie jest przekazana na standardowe klasyfikatory uczenia maszynowego. Wydaje mi się, że jest to twórcze połączenie wielu znanych metod w złożony, specjalizowany system. Przedstawione wyniki oceny eksperymentalnej są dość wysokie - i w mojej opinii nacisk w artykule jest postawiony na stworzenie możliwie dokładnego systemu klasyfikacji wybranych gatunków bakterii, które mogą powodować infekcje. Sam artykuł jest napisany bardziej w stylu zastosowania bio-medycznego.
Praca [A4] to w zasadzie wykorzystanie tego samego pomysłu na budowę systemu do rozpoznawania gatunków grzybów. Są drobne zmiany innych pre-trenowanych architektur sieci neuronowej oraz połączenia modelu worka słów wizualnych w obrazie z metodą wektorów Fishera. Także głównym osiągnięciem jest wytrenowanie systemu o wysokiej

ocenie klasyfikacyjnej. W stosunku do pracy [A3] poświecono więcej miejsca dokładniejszej analizie wyników wraz z implementacją wizualnej macierzy pomyłek oraz podobieństwa elementów łat (czyli fragmentów obrazu) do środka skupiska podczas grupowania słów wizualnych. Jak autorzy dyskutują jeśli połączy się to z predefiniowana przez mikrobiologów wiedzą na temat znanych własności gatunków, to można dokonać eksperckiej analizy ich cech na obrazach.
Ostatnia prac [A5] z tej grupy zastosowań dotyczy klasyfikacji klonów tego samego gatunku bakterii. Jak piszą autorzy ze względu na inną charakterystykę zdjęć mikroskopowych, zastosowano inną metodę w której agregacja reprezentacji łat jest wykonana za pomocą neuronowej uwagi (wykorzystano w miarę nową propozycje innych autorów znaną jako Attention-based Multiple Instance Learnng AbMIL). W stosunku do wcześniejszych prac [A3,A4] zaproponowano ustalanie łat obrazu o największej wadze i dalsze przetwarzanie z wykorzystaniem modelu persystentnego worka słów wizualnych oraz bardziej typowej segmentacji obrazu, które wizualizując swoje wyniki mogą pomóc w analizie przestrzennego położenie bakterii w obrazie i parametrów ich wielkości oraz kształtu. Jeśli konsultować to z ekspertami z mikrobiologii, to może prowadzić do dokładniejszej analizy różnic pomiędzy badanymi bakteriami.

Praca [A6] jest najbardziej i bezpośrednio związana z lokalną wyjaśnialnością predykcji sieci neuronowych. Autorzy uogólniają znaną propozycje sieci prototypowej ProtoNet (oryginalnie wprowadzonej przez zespół p. C.Rudin) w celu redukcji liczby rozważanych prototypów poprzez łączenie (ang. merge) wcześniej odkrytych prototypów, które są najbardziej do siebie podobne (autorzy rozważyli metrykę podobieństwa ich aktywacji w danych uczących). Ocena eksperymentalna (na znanych obrazowych bazach) i porównanie do wcześniejszych podobnych metod wskazało na trafność zbliżoną do nich przy użyciu mniejszej liczby prototypów. Ponadto autorzy przeprowadzili interesującą i dogłębniejszą analizę łączonych prototypów (odpowiadających im obrazów) i ich wzajemnego podobieństwa (w fazie merge) powiązane z studium analizy opinii ludzkich. Dodatek do pracy zawiera także analizę teoretyczną warunków operacji łączenia. W mojej opinii, pomimo że jest to modyfikacja znanej propozycji sieci prototypowej ten artykuł jest napisany w sposób najlepiej realizujący postulaty wyjaśniania predykcji neuronowych modeli klasyfikacji obrazów.

Kolejny artykuł [A7] dotyczy zupełnienie innego zadania uczenia się klas z zbiorów wielokrotnej instancji (ang. Multiple Instance Learning, MIL). Autorzy zaproponowali rozszerzenie znanej już wcześniej metody AbMiL agregacji wykorzystującej neuronowy mechanizm uwag (ang. attention-based pooling) poprzez wcześniejszy mechanizm self-attention znany z sieci tzw. transformerów dla dokładniejszego odkrywania zależności pomiędzy instancjami. Ich połączenie w odpowiedni wektor jest przekazywane na typowy neuronowy klasyfikator. Przeprowadzono ocenę eksperymentalną na pięciu różnych obrazowych dostępnych publicznie bazach osiągając dobre wyniki predykcyjnej.

Ostatnia grupa artykułów [A8,A9] dotyczy topologicznej analizy struktury danych wielowymiarowych i metodologicznie jest to mocno inny obszar badawczy od powyżej omawianych. W zakresie homologicznych narzędzi w topologii danych autorzy zajmowali się metodami efektywnego generowania reprezentacji w postaci diagramów persystencji. Wskazując na ograniczenia istniejących metod wykorzystujących przekształceniach jądrowe oraz innych wektorowych zaproponowali alternatywne adaptacyjne reprezentacje diagramów persystencji, które mają tworzyć wektora lepiej pasujące do metod uczenia maszynowego. W pracy [A8] zaproponowano metodę PBoW będącą adaptacją metody worka słów do tego rodzaju diagramów punktów w danych. Praca ma charakter bardziej teoretyczny niż poprzednie artykuły i zawiera wiele analiz stabilności rozwiązania. Jak autorzy pokazują jest ona efektywniejsza obliczeniowo i wartości uzyskanego wektora

odpowiadają poszczególnym słowom persystentnym w modelu (ang. persistence codewords), które można wykorzystać do interpretacji różnić pomiędzy klasami punktów (w artykule zawarte wizualizacje różnić pomiędzy średnimi wektorów uzyskanych reprezentacji w metodzie PBoW). Praca [A9] jest rozszerzeniem zaproponowanej metody o inne sposoby kodowania tych diagramów. Autorzy wprowadzali tzw. persystentny VLAD (ang. Vectors of Locally Aggregated Descriptors) oraz persystentny model Fisher vector. Szczegółowo badali stabilność zaproponowanych metod względem metryki Wassersteina oraz inne własności teoretyczne.

Podsumowując ten opis, uważam że *przedstawione prace z osiągnięcia habilitacyjnego są ciekawe, choć mocno zróżnicowane metodycznie*.

Niektóre z nich wydają się być bardziej interesujące lub oryginalne (za najciekawsze dla mnie uznałbym przede wszystkim [A1], [A3], [A6] i [A8]) a pozostałe są w jakimś stopniu kontynuowaniem wątku z innej pracy lub jej innym zastosowaniem w stosunku do innych obrazów lub innego sposobu kodowania (zwłaszcza w wersjach modeli słów wizualnych) lub agregacji ich fragmentów w rozważanych modelach.
W przedstawionych pomysłach na złożone neuronowe systemy uczące się można też często dostrzec łączenie wielu wcześniejszych metod lub technik zaproponowanych przez innych badaczy, choć odniosłem wrażenia że część takich integracji ma charakter dość twórczy i dobrze przełożyła się na wyniki eksperymentalne. Nie czuję się być ekspertem od rozpoznawania obrazów, ale taki trend widzę w wielu obecnych pracach dotyczących zastosowania modyfikacji głębokich sieci neuronowych, chociażby w bliższym mnie przetwarzaniu tekstów.
Ponadto opisy propozycji metod, dla mnie czasami na zbyt ogólnym poziomie, co na podstawie mojego doświadczenia jest także typowe dla wielu innych prac w takich obszarach badawczych. Na tzw. plus mogę także zauważyć, że oprócz pokazywania określonych wyników eksperymentalnych, w niektórych z artykułów (zwłaszcza ostatnich z listy) widać też elementy większej analizy teoretycznej. Niektóre z przedstawionych dodatkowych wyników albo analiz rozszerzają wiedzę o działaniu elementów badanych metod albo o trudnościach badanych danych — co trzeba dobrze ocenić, że autorzy nie poprzestali na przedstawieniu bardziej suchych wyników tabelarycznych jakieś miary oceny predykcyjnej.

Ponadto stwierdzam, że prace wchodzące do tzw. osiągnięcia habilitacyjnego są często ulokowane w tzw. dobrych miejscach. W szczególności dotyczy to:

- Dwóch prac na bardzo prestiżowej konferencji IJCAI oraz jednej pracy na ACM SIGKDD - chcę tutaj skomentować, że dotychczas nieliczne prace polskich autorów były przyjmowane na te tzw. „najważniejsze" konferencje z zakresu sztucznej inteligencji - dlatego trzeba to mocno docenić.
- Prac w znanych czasopismach międzynarodowych (odpowiednio wysoko indeksowanych) - przede wszystkim Artificial Intelligence Review oraz dwukrotnie PLoS one (choć tutaj po tzw. linii zastosowania mikrobiologicznego - klasyfikacji obrazów bakterii oraz grzybów powodujących wybrane choroby).

Inne prace konferencyjne są też w miarę dobrze ulokowane (w zakresie sieci neuronowych albo wizji komputerowej).
Ponadto jak piszę w dalszej ocenie dodatkowego dorobku, zauważyłem że p. dr Zieliński i jego współautorzy kontynuują badania w niektórych wątkach wyjaśnialności - patrz np. najnowsze prace przyjęte na najlepsze konferencje jak ECML PKDD i ECCV w tym roku - co oceniam bardzo pozytywnie.

Zauważam także, że w części artykułów autorzy udostępniają linki do kodu implementacji swoich metod - co jest zgodne z postulatami otwartej nauki i rekomendacjami niektórych konferencji. W kilku pracach zamieszono też bardziej szczegółowe informacje o

architekturze sieci oraz parametrach ich uczenia - co z punktu widzenia możliwego odtwarzania rezultatów oceniam bardzo pozytywnie. Choć biorąc pod uwagę znane rekomendacje do powtarzalności i odtwarzalności obecnych modelu uczenia maszynowego (tzw. ang. reproducibility), w szczególności problemów z powtarzaniem rezultatów dla głębokich sieci neuronowych powyższe informacje są na ogół nie wystarczające i w dodatkowych załącznikach elektronicznych powinno się udostępniać więcej dokumentacji oraz informacji, patrz np. Joelle Pineau et al The Machine Learning Reproducibility Checklist wynikające z konferencji NeurIPS lub IJCAI Reproducibility Guidelines oraz inne specjalizowane wytyczne dla prac z głębokimi sieciami.

Niestety muszę też wyrazić także trochę *uwag krytycznych*.

Główna trudność, którą napotkałem czytając prace w ramach osiągnięcia habilitacyjnego i odnosząc je do treści autoreferatu dotyczy ujęcia tam wyjaśnień predykcji modeli. Trzeba tutaj przypomnieć, że Habilitant zatytułował swoje osiągnięcie „Wyjaśnialne i interpretowalne uczenie maszynowe z zastosowaniami w analizie obrazów medycznych" a jako cel swoich badań wymienił (cytat ze str 5 autoreferatu) „Celem opisywanych badań było opracowanie narzędzi uczenia maszynowego, które dostarczają wyjaśnień predykcji, ze szczególnym ukierunkowaniem na dane obrazowe i medyczne.". Ponadto w wielu miejscach autoreferatu Habilitant formułuje myśli tak, że wyjaśnialność ma być głównym elementem omawianej przez Niego pracy. Jako czytelnik przedstawionych prac oraz badacz zajmujący się także tzw. wyjaśnialną sztuczną inteligencją (ang. XAI) nie podzielam takiego ujęcia dokonania habilitacyjnego, gdyż tylko część z prac [A1-A9] mogę tak scharakteryzować.

Zdecydowanie w nurcie XAI mieści się głównie praca [A6]. Sieci prototypowe oryginalnie wymyślono w zespole prof. Rudin do w miarę naturalnych dla człowieka wyjaśnień lokalnych predykcji głębokich sieci neuronowych. Są one rozważane zarówno dla danych obrazowych, lecz także w mniejszym stopniu klasyfikacji dokumentów tekstowych (znane są mi przynajmniej trzy różne propozycje). Ch.Molnar w swojej książce Interpretable Machine Learning w rozdziale 7 oraz 8 podaje także wiele argumentów na korzyść wykorzystywania rzeczywistych przykładów oraz ich prototypów w różnych metodach wyjaśnień predykcji - lecz także podaje jako element krytyki trudności doboru liczby prototypów. Ponadto w pracy [A6] bardzo ciekawie rozszerzono interpretacje podobnych prototypów na wybranych obrazach oraz przedstawiono studium oceny jakościowej z ludźmi. Dlatego praca p. Dr Zielińskiego i współautorów pomimo, że skupia się głównie na pewnej modyfikacji (tzw. ang. merge) podobnych prototypów w uogólnieniu znanego modelu wręcz idealnie się wpisuje w przedstawiony temat XAI i cel badań. Niestety w mojej ocenie z pozostałymi pracami jest już w różnym stopniu gorzej. Prace [A8-A9] głównie dotyczą nowych sposobów konstrukcji bardzo specyficznych deskryptorów - diagramów persystencji. Pomimo, że cel tych prac jest napisany w duchu „this paper introduces peristatence bag-of-words" + wskazania, że będą dokładniejsze i efektywniejsze obliczeniowo niż poprzednie, to mieszczą się w analizie topologicznej danych, która jako całość metodyki jest bardzo przydatna w specjalnej wizualnej analizie struktury danych. Nie nazwałbym jednak tego jak Habilitant interpretowalnym uczeniem maszynowym - to ciagle jest dla mnie podejście do wizualnej analizy danych, co jest metodologicznie czymś innym. Lecz pokazane w pracach możliwości użycia słów kodowych w proponowanych histogramach mogą być przydatne do wizualnej analizy skupisk punktów w diagramach - więc poprzez specjalizowane wizualizacje mogą można zaliczać do narzędzi wspomagających objaśnienia.
Natomiast mam krytyczne spojrzenie na nadmierne „lansowanie" w autoreferacie aspektu wyjaśnialności jako celu propozycji dla zadań diagnostyce mikrobiologicznej. Zaproponowane w pracach [A3-A5] podejścia w moim zrozumieniu maja na celu przede

wszystkim jak najlepszą trafność klasyfikacji w wybranym problemie zastosowania (bakterie lub grzyby powodujące zakażenia) i czytając zarówno streszczenia jak i treści w tych artykułach wyraźnie czytamy „the purpose of this article is to propose a computer-aided recognition system for classifying (..)" i nie mamy jawnie zapewnienia wyjaśnialności predykcji. Tak też jest opisana główna część eksperymentalna, zwłaszcza w pracy [A3]. W pracy [A4] mamy wprawdzie dodatkowo podrozdział nt. implementacji wizualnej macierzy pomyłek - ale jest ona raczej dodatkiem do głównej części pracy i można bardziej na nią spojrzeć jako dość prostą i naturalną pogłębioną analizę wyników klasyfikacji. Jeśli do niej, jak pisze się Habilitant doda się elementy powiązania do dodatkowego opisu cech bakterii przez mikrobiologa może ona być ciekawsza do analizy przez ekspertów dziedzinowych, lecz nie jest to aż taki poziom wyjaśnień działania systemu, który się omawia w pracach z zakresu XAI. Trochę bardziej zaawansowane podejście pojawiło się w pracy [A5], gdzie mogę uznać że element wykorzystania PBoW we wizualizacji przestrzennego położenia bakterii względem sieci oraz druga metoda identyfikacji różnic w budowie rozpoznanych bakterii w ramach tzw. klonów jest mocniej wprowadzony do propozycji całości podejścia opisanego w artykule i ma większy potencjał wspomagania interakcji z ekspertem mikrobiologiem w analizie rozważanych klonów bakterii - dlatego w większym stopniu element potencjalnej wyjaśnialności jest częścią zaproponowanej metody w [A5].

Mam niestety też trudność w uznaniu, że wyjaśnialność predykcji jest częścią ciekawego podejścia w pracach [A1 i A2]. W tekstach prac nie jest to zasadnicza cześć opisanej propozycji dla samo nadzorujących podejść w rozpoznawaniu obrazów. Dodatkowe eksperymenty nt. wagi słów wizualnych dla reprezentacji to w moim zrozumieniu tekstów jest raczej narzędzie do pogłębionej analizy tych reprezentacji i proponowanych metod. Podobne zastrzeżenia mogę sformułować do analizy wag przypisanych tzw. łatkom obrazów w pracy [A7]. Mogę to ocenić jako przydatną dodatkową analizę różnic między wprowadzonym podejściem do neuronowej uwagi SA-AbMIL a podstawowym AbMIL. Natomiast w obu przypadkach nie rozpoznaje tego jako bardziej całościowego podejścia do wyjaśnień predykcji modelu neuronowego, w stosunku do którego oczekuje się więcej - patrz np. dyskusja metod wyjaśniania w takich pracach przeglądowych jak np. R.Guidotti et al.: A survey of methods for explaining black box models 2018. Zresztą sam pan dr Zieliński w wstępnych częściach autoreferatu dość dobrze charakteryzował ogólną taksonomie metod XAI (np. str 8). Ponadto także wskazywał na różnicę w rozumieniu interpretowanego uczenia maszynowego, które powinno z samej formy reprezentacji być dostatecznie czytelne dla odbiorcy - takie spojrzenie na wykorzystanie jawnego symbolicznego języka reprezentacji było obecne już dawno w uczeniu maszynowego i np. w poprzednim wieku wyraźnie widoczne w pracach prof. Ryszarda Michalskiego (w inny niebezpośredni sposób częściowo odniósł Habilitant w wprowadzaniu do autoreferatu 3.1 w tzw. przetargu trafność vs. intepretowalnośc). Pomimo tego w swoim autoreferacie część propozycji od [A6] nazwał interpretowanym uczeniem maszynowym - co moim zdaniem jest nieuzasadnione (jak już to są one bliższe wyjaśnialności). Przypomnę także, że w oczekiwaniach wobec wyjaśnialności modeli uczenia maszynowego podkreśla się często, że jest to cały proces i system udziela wyjaśnień dostatecznie zrozumiałych dla osoby w kategoriach jego wiedzy i słownictwa (patrz np. prace F.Giannotti et al, czy K.Mullera). Chciałbym dodać, że osobiście podzielam opinie innych badaczy że sam proces oceny wyjaśnień i współpracy z odbiorcą jest także bardziej złożony niż prosta wizualizacja lub analiza wyników klasyfikacji, patrz np. dyskusja w tzw. position paper „F.Doshi-Velez, B.Kim: Towards a rigorous science of interpretable machine learning 2017".

Podsumowując moje pewnie zbyt długie wypowiedzi, nie mogę niestety uznać, że takie rozumienie wyjaśnialności oraz przede wszystkiem interpretowanego uczenia

maszynowego są obecne w części w przedstawionych prac, które w mojej opinie ciekawe dokonanie mają w zupełnie innym zakresie. I przede wszystkim część zdań napisanych w autoreferacie, zwłaszcza w duchu frazy „Celem opisywanych badań było opracowanie narzędzi uczenia maszynowego, które dostarczają wyjaśnień predykcji" nie ma jawnego odzwierciedlenia w rzeczywistej treści wszystkich przedstawionych artykułów, a co najwyżej występuje w mniejszej lub dodatkowej części tych tekstów.

Mogę domniemywać, że Habilitant starał się w miarę spójnie połączyć pod wspólnym „parasolem tematycznym" wybrane prace w osiągnięciu habilitacyjnym, co nie było dla Niego takie łatwe, bo są one zróżnicowane metodycznie (co najmniej 4 zupełnie różne kategorie propozycji). Komentując, osobiście obserwuje także od lat, że wiele (także dużo słabszych dorobków habilitacyjnych) jest obecnie złożenie dość różnych prac i trudniej jest wychwycić ich wspólną podstawę proponowanej metodyki niż w przypadku pisania kiedyś książki - monografii. Pomimo tego, muszę „ponarzekać", że w treści autoreferatu Habilitant poszedł za daleko z wieloma sformułowaniami w zakresie wyjaśnialności oraz nadmiernie zinterpretował niektóre z własnych propozycji.

Sama dostarczona dokumentach jest poza tym kompletna i dobrze przygotowana - czytając ją mogłem zrozumieć i przeanalizować różne osiągnięcia pana dr Bartosza Zielińskiego. Przy czym oprócz polemiki nt. nadinterpretacji i rozumienia realizacji zasad wyjaśnialnego i interpretowalnego uczenia maszynowe dostrzegam także pewną przesadę stylu opisu wyników - np. niezwykle dobrej skuteczności rozpoznawania lub że jego propozycja pozwala na „bezsporny" wgląd w reprezentaci obrazów. Myślę że powinno używać innych bardziej bezstronnych sformułowań i ocenę pozostawić recenzentom. Także pewność sformułowania że jedna (zgadzam się, że ciekawych) praca [A3] miała wysokie cytowania „przyczyniając się do popularyzacji metod uczenia maszynowego w mikrobiologii", jeśli obecnie ma ona w Google Scholar tylko 23 cytowania jest dla mnie zdecydowaną za silną opinią i może za zasadą racjonalności uznawania przekonań Ajdukiewicza powinno pisać się jednak inaczej.
Pomimo, że w ogólności język polski jest poprawnie użyty i czytelny, to pojawiają się też dziwne „kalki" z języka angielskiego, np. kilkakrotnie Habilitant metody GMM i podobne określa jako miksturę Gaussów - poprawnie są to mieszaniny rozkładów Gaussowskich; klastry to skupiska itd.

Pomimo powyższych zastrzeżeń co do części z aspektów wyjaśnialności przedstawionych modeli oraz sposobu w jaki Habilitant opisał je w autoreferacie, uważam że same publikacje i ich ważność zarówno teoretycznie-poznawcza, jak i zastosowania (zwłaszcza biochemiczne) są zdecydowanie dobre, a część z propozycji jest dla mnie bardzo interesująca.

Podsumowując moją ocenę tej części czyli samego osiągnięcia habilitacyjnego Pana dra. Bartosza Zielińskiego, **jest ona bardzo pozytywna**. Stwierdzam, że zostały spełnione wymagania stawiane osiągnięciom habilitacyjnym i publikacje wnoszą tzw. znaczący wkład do obecnych badań w zakresie głębokich sieci neuronowych w odniesieniu do ich zastosowań do rozpoznawania i klasyfikacji obrazów.

3. Ocena pozostałego dorobku naukowego i innych aktywności organizacyjnych oraz dydaktycznych

W tej części oceny opieram się na dostarczonej dokumentacji obejmującej autoreferat, listę publikacji, wykaz innych osiągnięć oraz także na podstawie samodzielnego przeszukiwania zasobów WWW w zakresie publikacji i ich tzw. cytowalności. W ogólności *moja ocena jest bardzo pozytywna*. W autoreferacie oraz wykazie dorobku dr Bartosz

Zieliński przedstawił wiele informacji świadczących o dużej aktywności naukowej, organizacyjnych oraz dydaktycznej.

Po pierwsze jest on autorem lub współautorem wielu prac niewchodzących w skład wybranego osiągnięcia habilitacyjnego. W autoreferacie (pkt 2) wykazał on 10 prac przed doktoratem, w tym 3 artykuły w czasopismach (jedno bardzo dobre międzynarodowe Pattern Recognition, dwa w lokalnym (UJ) czasopiśmie Schedae Informaticae); prace konferencyjne w różnych miejscach, raczej różnego prestiżu, w tym część krajowych = na uwagę zasługuje jednak publikowanie kilku prac samodzielnie. *Po doktoracie nastąpił znaczący przyrost liczby publikacji.* Poza pracami wybranymi do osiągnięcia habilitacyjnego jest on współautorem 7 artykułów w czasopismach. Za najbardziej selektywne i prestiżowe mogę uznać Expert Systems with Applications i Pattern Recognition, oraz w mniejszym stopniu Computer vision and image understanding. Pozostałe czasopisma są bardziej typowe dla zastosowań biologicznych lub medycznych. Ponadto mamy wykaz 8 prac konferencyjnych (różnego prestiżu) oraz niepoliczone wystąpienia na tzw. workshopach (lecz klika przy dobrych konferencjach z zakresu uczenia maszynowego lub sieci neuronowych jak np. NeuroIPS). Habilitant zliczył łącznie 26 wystąpień na konferencjach (uwzględniając też okres przed doktoratem).

Patrząc na ostatnie dane w internecie ponadto zauważam, że w tym roku przyjęto dwie nowe prace na bardzo ważne i selektywne (z własnego doświadczenia PC) konferencje międzynarodowe: ECML PKDD 2022 ProtoMIL: Multiple Instance Learning with Prototypical Parts for Whole-Slide Image Classification (najprawdopodobniej kontynuowanie badań z osiągnięcia habilitacyjnego) oraz European Conference on Computer Vision (ECCV), 2022 (pt. Interpretable Image Classification with Differentiable Prototypes Assignment - także wygląda na dalsze badania w duchu osiągnięcia habilitacyjnego). W obu przypadkach jest to ważny sukces i świadczy o ciągłej aktywności naukowej pana dr Zielińskiego oraz dobrym wyborze miejsc zgłaszania ostatnich prac - co nie jest takie częste w polskim środowisku informatycznym.

Liczba publikacji po doktoracie (patrz też zestawienia dostępne na https://bartoszzielinski.github.io lub https://scholar.google.pl/citations?user=AJHaOpkAAAAJ&hl=pl) jest duża i przede wszystkim tematycznie wyraźnie różni się od prac z okresu przed doktoratem - co w mojej opinii jest bardzo ważne dla właściwej oceny postępu badań naukowych przy habilitacji. Ponadto bardzo dobrze oceniam zmiany w wyborze konferencji oraz czasopism. Wyraźnie Habilitant oraz jego współautorzy skierowali zainteresowanie na najlepsze konferencje w zakresie uczenia maszynowego, wizji komputerowej i sztucznej inteligencji. Podsumowując tą analizę publikacji i swoją powyższą wypowiedz - stwierdzam bardzo jednoznacznie, że z pełnym przekonaniem oceniam *bardzo pozytywnie jego aktywność naukową i publikacyjną.*

Informacja bibliometryczna przedstawiona w autoreferacie (pkt 4) jest dość dobra, choć jej prezentacja akurat nie do końca precyzyjna, przede wszystkim nie pokazano zróżnicowania na fazę doktorat i po doktoracie oraz rozbicie punktów na czasopisma i konferencje. Łączna liczba punktów Min. jest wysoka (ponad 2tys.), także wysoki łączny współczynnik wpływu źródeł tzw. impact factor (raczej przedstawiony w stylu nauk przyrodniczych). Troszkę gorzej jest z oceną tzw. widzialności prac poprzez cytowania — 427 GoogleScholar 174 WoS i 240 Scopus (z autocytowaniami). Indeksy Hirscha są trochę lepsze niż typowe dla ostatnich habilitacji, które recenzowałem - GoogleScholar 13, WoS 7 i podobnie Scopus.
Wartości te uległy poprawie w ostatnich miesiącach = np. w momencie pisania tej recenzji na początku sierpnia GoogleScholar podawał 534 cytowania i h index = 14. Najbardziej cytowane prace to M.Smieja et al .Processing of missing data by neural networks z bardzo znanej konferencji Advances in neural information processing systems NeurIPS

2018 (87 cytowań). Z prac zaliczonych do osiągnięcia habilitacyjnego największą cytowalność zarejestrowano dla artykułów z zastosowań biochemicznych z PLoS one (Zieliński, Bartosz, et al. "Deep learning approach to bacterial colony classification." PloS one 12.9 (2017) około 80, a drugi artykuł trochę ponad 20). Inne prace z osiągnięcia habilitacyjnego są niestety gorzej cytowane (max. kilka cytowań). Mogę wyrazić swoje zdanie - mogłoby być lepiej patrząc na miejsca lokalizacji publikacji i aktywność p. dr Zielińskiego. Lecz część z prac wchodzących w osiągnięcie habilitacyjne jest relatywnie świeża i może wymaga jeszcze trochę czasu na ich dostrzeżenie przez środowisko naukowe. Mam nadzieję, że to nastąpi w kolejnych latach - czego życzę Habilitantowi.

Ponadto pan dr B.Zieliński brał udział jako recenzent we wsparciu komitetów programowanych konferencji międzynarodowych (zestawienie w autoreferacie jest dość krótkie, lecz spojrzenie na własną stronę WWW wskazuje nowe działania - cyt "conferences AAAI, ECCV, WACV, ECML PKDD) i czasopisma (TIP, AIR, CSBJ, CBM, T-BME, TRENDS MICROBIOL, PloS one)"). Czyli jest to dobra aktywności, choć mogłaby być szersza widząc inne aktywności p. Dr Zielińskiego.

Pan dr Bartosz Zieliński ma *duże doświadczenie w realizacji różnorodnych projektów - grantów*. Ich łączna liczba - osiem - wydaje się być wysoką jak na okres zatrudnienia na UJ. Wśród nich dwukrotnie pełnił rolę kierownika (w tym pozyskał Sonata NCN - który jest bezpośrednio związany z częścią dorobku habilitacyjnego w zakresie topologicznej analizy danych). Niektóre projekty, w których był albo ciągle jest wykonawcą, są powiązane z projektami jego macierzystego zakładu (np. bardzo trudny do zdobycia FNP Team Net). Brał też udział w grantach stosowanych NCBiR przy współpracy przemysłowej. Według strony WWW p. Dr Zieliński pracuje także jako tzw. Lead Data Scientist w firmie Ardigen, gdzie kieruje zespołem ML w zakresie rozpoznawania obrazów. W przeszłości współpracował z innymi firmami, np. realizował projekty dla Samsung. Ten obszar działalności badawczo-wdrożeniowej oceniam *bardzo wysoko* (ponad przeciętnie).

Na podstawie autoreferatu widać, że brał on udział w stażach i wyjazdach zagranicznych, m.in. Vienna University of Technology (2015), St Polten University of Applied Science (2017), University of Zagreb (2013) oraz Instituto Superior Técnico in Lisbon (ponad miesiąc 2019), w większości obejmujących wspólne badania w zakresie analizy obrazów, widzenia komputerowego, jak i topologicznej analizy danych - co także przełożyło się na wspólne publikacje. Ponadto p. dr Zieliński bardzo dobrze skorzystał z możliwości udziału w międzynarodowych szkołach doktorskich, letnich (wiele z nich dotyczyło wizji komputerowej oraz przetwarzania danych obrazowych, głębokich sieci neuronowych lub uczenia maszynowego). Widać także umiejętności pana dr Zielińskiego w nawiązywaniu współpracy z naukowcami z innych ośrodków (także krajowych), które doprowadzała do wielu wspólnych publikacji (patrz lista 11 osób punkcie 5.3 autoreferat). Ta część aktywności pana dr. Zielińskiego jest mojej ocenie *ponad przeciętna* dla typowych dorobków habilitacyjnych.

Habilitant przedstawił obszerną dokumentację na temat dzielności dydaktycznej (patrz punkt 6.1. autoreferatu lub własna strona WWW). Prowadzi lub prowadził wiele przedmiotów na kierunku Informatyka, w tym w zakresie rozpoznawania obrazów, widzenia komputerowego a także liczne ćwiczenia lub zajęcia laboratoryjne (od kilkunastu last - także historycznie wiele z podstawowych przedmiotów z informatyki w różnych paradygmatach programowania). Na uwagę zasługują też specjalistyczne wykłady dla doktorantów z zakresu głębokich sieci neuronowych. Ponad standardowe jest jego zaangażowanie w promowanie prac dyplomowych - był promotorem ponad 20 magistrantów (tematy niektórych prac są bardzo interesujące, patrz zestawienie na własnej stronie WWW). Ponadto w jakieś formie (wspólnie z promotorami)-opiekuje się się trójką aktualnych doktorantów. Brał także udział w dodatkowych aktywnościach -

wykładach popularyzujących naukę (w ramach Studenckich Festiwali Informatycznych) a że w komitetach organizacyjnych np. tzw. Krakowska Konferencja Kognitywistyczna (w latach 2008-2010). W mojej opinii ta część dydaktycznej aktywności jest *bardzo dobra* i *zdecydowania lepsza* w stosunku do znanych mnie typowych wniosków habilitacyjnych.

Ku mojemu pozytywnemu zaskoczeniu pan dr Bartosz Zieliński jest, lub był mocno, zaangażowany w różne działania organizacyjne na swoim Wydziale (patrz punkt 6.2 zestawienia w autoreferacie). Wymieńmy przykładowo zostanie członkiem (z wyboru) Senatu UJ, Rady Wydziału, Rady Instytutu Informatyki i Matematyki, Komisji Wyborczych oraz współpracy w kierownictwie 1 stopnia studiów informatycznych. Wszystko to jest wręcz wyróżniającą się i modelową działalnością tzw. młodego pracownika nauki, znacznie przekraczającego typowe oczekiwania. Widać także z zestawień w dokumentacji, iż pan dr Bartosz Zieliński był kilkukrotnie nagradzany przez Rektora UJ.

4. Konkluzja końcowa

Kończąc niniejszą recenzję osiągnięcia habilitacyjnego oraz pozostałego dorobku naukowego, dydaktycznego i organizacyjnego Pana dra Bartosza Zielińskiego stwierdzam jednoznacznie, że moja opinia jest **pozytywna**.

Przedstawiony cykl publikacji zawiera wiele ciekawych propozycji złożonych systemów uczących wykorzystujących różne rozwiązania głębokiego uczenia się oraz adaptacji metod z innych obszarów uczenia maszynowego dla rozpoznawania i klasyfikacji danych obrazowych. Ponadto niektóre zastosowania, zwłaszcza biologiczne lub medyczne są interesujące i wartościowe. Dostrzegłem pomysłowość i pracowitość Pana dra Bartosza Zielińskiego oraz duży wysiłek w tworzeniu interesujących publikacji naukowych, oraz realizacji wielu projektów badawczych oraz wdrożeniowych. W mojej ocenie publikacje przedstawione w tzw. osiągnięciu wpisują się dobrze do problematyki uczenia maszynowego, w tym aktualnych rozwiązań głębokich sieci neuronowych w odniesieniu do zastosowań rozpoznawania i klasyfikacji obrazów. Mieszczą się one w ramach dyscypliny naukowej informatyka techniczna i telekomunikacja w zakresie nauk inżynieryjno-technicznych. Ponadto bardzo dobrze oceniam inne obszary aktywności naukowej, dydaktycznej i organizacyjnej Pana dra Bartosza Zielińskiego - w wielu elementach szczegółowych te osiągnięcia są wyższe od znanych mi z innych wniosków w postępowaniu habilitacyjnych.

W związku z powyższym wnoszę o dopuszczenie Pana dra Bartosza Zielińskiego do dalszych, przewidzianych odnośnymi przepisami etapów postępowania w sprawie nadania stopnia doktora habilitowanego.

Jerzy Stefanowski

prof. Jerzy Stefanowski