Łódź, dn. 22 lipca 2022 r.

Prof. dr hab. inż. **Anna Fabijańska**
Politechnika Łódzka
Wydział Elektrotechniki, Elektroniki,
Informatyki i Automatyki
Instytut Informatyki Stosowanej
Ul. Stefanowskiego 18
90-537 Łódź

RECENZJA

wniosku w postępowaniu o nadanie
dr. Bartoszowi Zielińskiemu
stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych
w dyscyplinie informatyka techniczna i telekomunikacja

1. Podstawa sporządzenia recenzji

Podstawę opracowania niniejszej recenzji stanowiło pismo Przewodniczącego Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja Politechniki Wrocławskiej Pana prof. dr hab. inż. Michała Woźniaka z dnia 24 maja 2022 r., wystosowane w związku z Uchwałą nr 312/16/RDND03/2021-2024 Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja PWr z dnia 18 maja 2022 r., powołującą mnie w skład komisji habilitacyjnej w postępowaniu w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja dr. Bartoszowi Zielińskiemu i powierzającą mi funkcję recenzenta w tym postępowaniu.

Recenzja sporządzona została na podstawie przekazanej mi dokumentacji przygotowanej przez Habilitanta, która obejmowała:

- kopię wniosku o przeprowadzenie postępowania habilitacyjnego,
- dane wnioskodawcy,
- kopię dokumentu potwierdzającego posiadanie stopnia doktora,
- autoreferat,
- wykaz osiągnięć naukowych,
- oświadczenia współautorów publikacji,
- kopie prac składających się na cykl publikacji stanowiących osiągnięcie naukowe,
- dokumenty potwierdzające istotną działalność naukową.

2. Sylwetka habilitanta

Pan dr Bartosz Zieliński ukończył w 2007 roku studia magisterskie na Wydziale Matematyki i Informatyki Uniwersytetu Jagiellońskiego w Krakowie, uzyskując tytuł zawodowy magistra. Od 2009 roku jest związany z tym wydziałem zawodowo, gdzie aktualnie pracuje na stanowisku adiunkta (od marca 2012 roku). Stopień doktora nauk technicznych w dyscyplinie informatyka dr

WPŁYNĘŁO
28-07-2022
RDN IT.T/217/2022

A. Fab.

Bartosz Zieliński uzyskał w Instytucie Podstawowych Problemów Techniki Polskiej Akademii Nauk w Warszawie w czerwcu 2012 roku, po przedłożeniu i obronie rozprawy nt.: *„Techniki wykrywania wybranych objawów chorób reumatycznych na podstawie zdjęć rentgenowskich rąk"*. W okresie po uzyskaniu stopnia doktora, Habilitant kontynuował prace naukowe w obszarze przetwarzania i analizy obrazów z wykorzystaniem metod klasycznych (nadzorowanych i nienadzorowanych), jak również bazujących na ogólnie pojętym uczeniu głębokim – głównie w konkretnych problemach z obszaru wspomagania medycznej diagnostyki obrazowej. Zajmował się również problemem nadzorowanej klasyfikacji danych. Te obszary tematyczne zaliczane są m.in. dyscypliny informatyka techniczna i telekomunikacja, w której dr Bartosz Zieliński ubiega się o stopień doktora habilitowanego.

Pan dr Bartosz Zieliński jest współautorem 27 publikacji zaindeksowanych w bazie Scopus, z czego większość (17) to artykuły w materiałach konferencyjnych. Publikował również w czasopismach posiadających IF takich (*Artificial Intelligence Review, Journal of Thermal Biology, Biochimica et Biophysica Acta - Molecular Basis of Disease, Expert Systems with Applications, Computer Vision and Image Understanding, PLoS ONE, Biocybernetics and Biomedical Engineering, Pattern Recognition*). Wg. bazy Scopus prace dra Bartosza Zielińskiego były cytowane łącznie 266 razy, a wskaźnik Hirsha Habilitanta wynosi 9. Zdecydowana większość tych prac wpisuje się w obszar tematyczny *computer science* (wg. bazy Scopus). Habilitant brał czynny udział w 26 konferencjach o zasięgu międzynarodowym. Ponadto, dr Bartosz Zieliński uczestniczy lub uczestniczył w realizacji ośmiu projektów finansowanych ze środków krajowych przez NCN, NCBiR, FNP, POIG oraz POKL. W dwóch projektach pełnił rolę kierownika (NCN oraz POKL).

W ramach obowiązków dydaktycznych na Uniwersytecie Jagiellońskim, Pan dr Bartosz Zieliński prowadził wykłady oraz seminaria z zakresu rozpoznawania obrazów, widzenia komputerowego oraz programowania, a także szereg zajęć laboratoryjnych dla studentów kierunku informatyka. Wypromował dwudziestu magistrów, z których dwoje jest aktualnie studentami doktorantami. Prowadził również wykłady dla doktorantów.

Dr Bartosz Zieliński był członkiem Senatu UJ w kadencji 2016-2020 jako przedstawiciel nauczycieli akademickich nieposiadających tytułu naukowego lub stopnia doktora habilitowanego. Tą grupę pracowników reprezentuje również w Radzie Wydziału Matematyki i Informatyku UJ od 2012 roku (trzecia kadencja). W latach 2016-2020 był członkiem Rady Instytutu Informatyki i Matematyki Komputerowej UJ. Uczestniczy w pracach komitetów organizacyjnych dwóch konferencji. Podejmował także pewne aktywności popularyzujące naukę.

Realizuje również współpracę z otoczeniem gospodarczym, w szczególności od 2017 roku kieruje zespołem ukierunkowanym na analizę obrazów biomedycznych przy użyciu głębokich sieci neuronowych w firmie Ardigen S.A. Ponadto, zrealizował również projekt dla Samsung R&D.

3. Stwierdzenie spełnienia przesłanki, o której mowa w art. 219 ust. 1, pkt. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce dotyczącej posiadania stopnia doktora

Pan Bartosz Zieliński posiada stopień doktora nauk technicznych w dyscyplinie informatyka, nadany uchwałą Rady Naukowej Instytutu Podstawowych Problemów Techniki Polskiej Akademii Nauk z dnia 28 czerwca 2012 r. Kopia dyplomu stanowi załącznik do recenzowanego wniosku.

A. Fab

4. Ocena osiągnięcia naukowego wskazanego przez Habilitanta zgodnie z art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce

4.1. Temat i publikacje wchodzące w skład osiągnięcia

Jako osiągnięcie naukowe wynikające z art. 219, pkt. 2 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce Pan dr Bartosz Zieliński przedstawił cykl powiązanych tematycznie artykułów naukowych w czasopismach naukowych lub w recenzowanych materiałach z konferencji międzynarodowych nt.: ***Wyjaśnialne i interpretowalne uczenie maszynowe z zastosowaniami w analizie obrazów medycznych***. Na cykl ten składa się 9 artykułów naukowych opublikowanych w latach 2019-2021, czyli w okresie po uzyskaniu przez Habilitanta stopnia doktora.

W skład ww. cyklu publikacji wchodzą następujące artykuły naukowe (w nawiasach podano deklarowany wkład Habilitanta w powstanie poszczególnych publikacji, potwierdzony oświadczeniami współautorów załączonymi do recenzowanego wniosku):

[A1] Dominika Basaj, Witold Oleszkiewicz, Igor Sieradzki, Michał Górszczak, Barbara Rychalska, Tomasz Trzciński, **Bartosz Zieliński (25%)**. *Explaining Self-Supervised Image Representations with Visual Probing*. International Joint Conference on Artificial Intelligence (IJCAI), DOI:10.24963/ijcai.2021/82, p. 592–598, 2021. (MEiN$_{2021}$: 200, Core rank: A*).

[A2] **Bartosz Zieliński (70%)**, Michał Górszczak. *What Pushes Self-Supervised Image Representations Away?* International Conference on Neural Information Processing (ICONIP), Communications in Computer and Information Science 1516, DOI:10.1007/978-3-030-92307-5_60, p. 514–521, 2021. (MEiN$_{2021}$:140, Core rank: B).

[A3] **Bartosz Zieliński (55%)**, Anna Plichta, Krzysztof Misztal, Przemysław Spurek, Monika Brzychczy-Włoch, Przemysław Spurek, Dorota Ochońska. *Deep learning approach to bacterial colony classification*. PLoS ONE 12 (9), DOI:10.1371/journal.pone.0184554, p. 1–14, 2017. (MEiN$_{2021}$: 100, IF: 3.24).

[A4] **Bartosz Zieliński (70%)**, 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), DOI:10.1371/journal.pone.0234806, p. 1–16, 2020. (MEiN$_{2021}$:100, IF: 3.24).

[A5] Adriana Borowa, Dawid Rymarczyk, Dorota Ochońska, Monika Brzychczy-Włoch, **Bartosz Zieliński (40%)**. *Deep learning classification of bacteria clones explained by persistence homology*. International Joint International Joint Conference on Neural Networks (IJCNN), DOI:10.1109/IJCNN52387.2021.9533954, p. 1–8, 2021. (MEiN$_{2021}$: 140, Core rank: B).

[A6] Dawid Rymarczyk, Łukasz Struski, Jacek Tabor, **Bartosz Zieliński (40%)**. *ProtoPShare: Prototypical Parts Sharing for Similarity Discovery in Interpretable Image Classification*. ACM SIGKDD Conference on Knowledge Discovery & Data Mining, DOI:10.1145/3447548.3467245, p. 1420–1430, 2021. (MEiN$_{2021}$: 200, Core rank: A*).

[A7] Dawid Rymarczyk, Adriana Borowa, Jacek Tabor, **Bartosz Zieliński (40%)**. *Kernel Self-Attention for Weakly-supervised Image Classification using Deep Multiple Instance Learning*. IEEE Winter Conference on Applications of Computer Vision (WACV),

A. Fab.

DOI:10.1109/WACV48630.2021.00176, p. 1720–1729, 2021. (MEiN$_{2021}$: 140, Core rank: A).

[A8] **Bartosz Zieliński (60%)**, 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 (IJCAI), DOI:10.24963/ijcai.2019/624, p. 4489–4495, 2019. (MEiN$_{2021}$: 200, Core rank: A*).

[A9] **Bartosz Zieliński (55%)**, Michał Lipiński, Mateusz Juda, Matthias Zeppelzauer, Paweł Dłotko. *Persistence codebooks for topological data analysis.* Artificial Intelligence Review 54 (3), DOI:10.1007/s10462-020-09897-4, p. 1969–2009, 2021. (MEiN$_{2021}$: 140, IF: 8.139).

4.2. Charakterystyka ilościowa osiągnięcia

Skład publikacji przedstawionych przez Habilitanta do oceny w ramach powiązanego tematycznie cyklu kształtuje się więc następująco:

- 3 artykuły w czasopismach posiadających współczynnik IF i uwzględnionych w bazie Journal Citation Reports (JCR);
- 6 referatów na konferencjach międzynarodowych uwzględnionych w bazie CORE Conference Ranking i posiadających rankingi A* (3 artykuły), A (1 artykuł) oraz B (2 artykuły).

Zgodnie z informacją zawartą w autoreferacie, sumaryczna liczba punktów MEiN przypisanych do publikacji wchodzących w skład cyklu wynosi ok. 1360, przy czym punktacja poszczególnych prac waha się w zakresie od 100 do 200, osiągając średnią na poziomie ok. 151 punktów. Punktowo, jest to wynik bardzo dobry. Niemniej jednak, Habilitant lokował prace wchodzące w skład cyklu głównie w recenzowanych materiałach z konferencji międzynarodowych, co uważam za pewną drogę „na skróty". Wprawdzie są to konferencje o wysokiej punktacji MEiN i renomie wśród badaczy reprezentujących m.in. dyscyplinę informatyka techniczna i telekomunikacja (*Int. Joint Conf. on Artificial Intelligence, Int. Conf. on Neural Information Processing, Int. Joint Conf. on Neural Networks, ACM Conf. on Knowledge Discovery & Data Mining, IEEE Winter Conf. on Applications of Computer Vision*), jednak z uwagi na charakter referatów konferencyjnych, wyniki przedstawione w tego typu publikacjach są zazwyczaj mniej wnikliwe i szczegółowe niż artykuły w czasopismach.

Sumaryczny współczynnik *impact factor* (IF) publikacji w czasopismach wchodzących w skład osiągnięcia z uwzględnieniem wartości z roku 2021 wynosi 17.09. W szczególności, Habilitant publikował w dwóch czasopismach: *PLoS ONE* (IF: 3.752, 2 prace) oraz *Artificial Intelligence Review* (IF: 9.588). Tylko to ostatnie jest czasopismem jednoznacznie identyfikowanym z informatyką. *PLoS ONE* jest czasopismem multitematycznym, publikującym prace z wielu obszarów. Zatem tylko jedna z prac w cyklu przedstawionym przez Habilitanta została opublikowana w czasopiśmie *stricte* związanym z dyscypliną, w której prowadzone jest postępowanie i odnotowana została w bazie DBLP. Uważam to za słaby punkt bibliometryczny osiągnięcia przedstawionego przez Habilitanta do oceny.

Przynależność osiągnięcia naukowego przedstawionego do oceny przez dra Bartosza Zielińskiego do dyscypliny informatyka techniczna i telekomunikacja nie budzi jednak wątpliwości. Wyjaśnialne i interpretowalne uczenie maszynowe, analiza obrazów oraz klasyfikacja danych stanowią aktualny przedmiot badań m.in. w tej dyscyplinie. Ponadto, poza

AFab.

wspomnianymi wcześniej publikacjami w czasopiśmie *PLoS ONE*, pozostałe publikacje w cyklu zostały odnotowane w bazie bibliograficznej DBLP indeksującej publikacje z obszaru informatyki oraz zaliczone przez bazę Scopus do obszaru tematycznego *computer science*.

Pan dr Bartosz Zieliński **nie jest samodzielnym autorem żadnej z publikacji wchodzących w skład cyklu, co również uważam za duży niedostatek.** Wprawdzie jest pierwszym autorem w pięciu z tych prac, a ostatnim w czterech, jednakże liczba autorów waha się od 2 do 7 (średnio 5), co utrudnia określenie indywidualnego wkładu Habilitanta. Jego zadeklarowany udział w poszczególnych artykułach waha się w zakresie 25-70%, osiągając średnią na poziomie 46%. Z załączonych do wniosku oświadczeń Habilitanta wynika, że zdefiniował problem badawczy (lub uczestniczył w tym procesie) rozważany w większości prac wchodzących w skład osiągnięcia. Był autorem kluczowych idei w nich przedstawionych (we wszystkich na poziomie koncepcji, w niektórych na poziomie implementacji). Zaplanował i samodzielnie przeprowadził lub nadzorował eksperymenty. W zdecydowanej większości kierował pracami opisanymi w pracach wchodzących w skład cyklu. **Indywidualny wkład merytoryczny dra Bartosza Zielińskiego w prace wchodzące w skład osiągnięcia wydaje się więc wiodący.**

4.3. Podsumowanie i ocena osiągnięcia

W przedstawionym do oceny cyklu publikacji, Pan dr Bartosz Zieliński podejmuje ciekawy i obecnie bardzo aktualny temat wytłumaczalnego AI (ang. *eXplainable AI*, XAI). W ogólnym ujęciu, prace w tym obszarze mają na celu rozwój metod, które pomogą wyjaśnić decyzje podjęte przez modele AI/ML (postrzegane dotychczas jako przysłowiowe czarne skrzynki), a tym samym zwiększyć ich wiarygodność w oczach użytkowników końcowych. Habilitant deklaruje, iż rozważa problem w dwóch aspektach: **interpretowalności** [A1]-[A5] oraz **wytłumaczalności** [A6]-[A9]. W szczególności, interpretowalność modelu uczenia maszynowego definiuje jako bierną cechę opisującą poziom jego transparentności i zdolność do podawania wyjaśnień swoich decyzji, natomiast wytłumaczalność jako cechę czynną wymagającą od modelu posiadania dodatkowego interfejsu odzwierciedlającego w zrozumiały sposób działanie modelu.

Tematyka prac wchodzących w skład cyklu jest dość różnorodna, a deklarowany w temacie osiągnięcia problem interpretowalności i wytłumaczalności decyzji modeli AI/ML nie zawsze jest jednoznaczny i często poboczny. Fraza *„z zastosowaniami w analizie obrazów medycznych"* jest myląca i nie oddaje treści publikacji wchodzących w skład osiągnięcia, gdyż część rozwiązań proponowanych przez Habilitanta testowanych było jedynie na obrazach scen naturalnych.

Zasadnicze wątki, rozważone w pracach wchodzących w skład osiągnięcia obejmują:

- opracowanie samonadzorujących się metod reprezentacji obrazów [A1, A2];
- klasyfikację obrazów mikrobiologicznych z wykorzystaniem zagregowanego wektora cech pochodzących z głębokich sieci neuronowych oraz klasycznych deskryptorów [A3, A4, A5];
- wykorzystanie części prototypowych [A6] oraz mechanizmu atencji [A7] do zwiększenia interpretowalności modeli AI/ML;
- zastosowanie teorii persystentnej homologii w problemie klasyfikacji danych w tym danych obrazowych [A5], chmur punktów i danych reprezentowanych grafami [A8, A9].

Praca [A1] wprowadza wizyjne klasyfikatory diagnostyczne (ang. *visual probing tasks*) przeznaczone do analizowania reprezentacji dostarczonych przez metody samonadzorujące się

A. Fab.

(ang. *self supervised*). W szczególności, Habilitant zaproponował przeniesienie metod dedykowanych do analizy reprezentacji tekstu (NLP), na grunt informacji obrazowej. W tym celu wprowadził analogię pomiędzy tekstem a obrazem, w której obraz jest odpowiednikiem zdania, klastry superpikseli są odpowiednikami słów, natomiast pojedyncze piksele wchodzące w skład superpikseli – odpowiednikami liter w wyrazie. Słowa wizualne otrzymane zostały w drodze grupowania superpikseli SLIC z wykorzystaniem metody ACE (ang. *Automatic Concept-Based Explanations*) i poddawane analizie przez klasyfikatory diagnostyczne zaadaptowane wg. pomysłu Habilitanta z NLP. W szczególności, były to *word content* (WC) służący do wykrywania zadanych słów wizualnych w obrazie, *sentence length* (SL) służący do określania liczby słów wizualnych w obrazie, *character bin* (CB) pozwalający na diagnozowanie rozmiaru superpiksela oraz *Semantic Odd Man Out* (SOMO) pozwalający na sprawdzanie, czy obraz został zmodyfikowany poprzez zamianę superpiksela, tj. zaburzenie informacji semantycznej. Powyższe klasyfikatory diagnostyczne wykorzystane zostały w połączeniu z popularnymi metodami samonadzorującego się znajdowania reprezentacji obrazu bazującymi na kontrastowej funkcji kosztu (ang. *contrastive loss*), takimi, jak MoCo, SimCLR, BYOL (ang. *Bootstrap Your Own Latent*) oraz SwAV. W szczególności reprezentacja będąca wynikiem działania ww. metod stanowiła dane wejściowe do klasyfikatorów diagnostycznych. Skuteczność podejścia sprawdzono w odniesieniu do obrazów ze zbioru ImageNet reprezentujących 55 klas przynależnych do 5 kategorii z uwzględnieniem 100 najistotniejszych słów wizualnych wybranych z zastosowaniem metody TCAV (ang. *Testing with Concept Activation Vector*) jako miary istotności słowa. Eksperymenty wykazały, że takie podejście pozwala uzyskać informację semantyczną odpowiadającą złożoności i spójności obrazu. Ujawniły też różnice jakości reprezentacji dostarczonej przez rozważone metody samonadzorujące się, z wyraźną przewagą na korzyść metody SimCLR połączonej z każdym z zaproponowanych klasyfikatorów diagnostycznych.

W pracy [A2] Habilitant kontynuował prace nad metodą przedstawioną w [A1], badając znaczenie wybranych słów wizualnych dla wynikowej reprezentacji obrazu. W tym celu zaproponował analizę reprezentacji obrazu opartą na zniekształceniach i teorii percepcji wzrokowej. W szczególności, porównywał reprezentacje dostarczone przez metody samonadzorujące się (MoCo, SimCLR, BYOL oraz SwAV) dla obrazu oryginalnego oraz jego zniekształconej wersji, w którym pojedyncze superpiksele przypisane do wybranego słowa wizualnego zastąpione zostały wartością średnią barwy uzyskaną dla obrazu. Rozważone zostały superpiksele o różnych właściwościach wizualnych – od obszarów o jednorodnej barwie, poprzez superpiksele zawierające teksturę do takich, które odpowiadały kształtom 2D i formom 3D lub ich fragmentom. Przeprowadzone eksperymenty wykazały, że proste (jednorodne w kontekście barwy i tekstury) słowa wizualne nie mają istotnego wpływu na reprezentację obrazu uzyskaną metodami samonadzorujacymi się, podczas gdy w największym stopniu wpływają na nią skomplikowane słowa przedstawiające formy i kształty. Nie zauważono natomiast istotnych różnic pomiędzy rozważonymi metodami samonadzorującymi się.

Obydwie prace [A1, A2] w mojej ocenie wpisują się w obszar XAI (ang. *eXplainable AI*) pozwalając wskazać fragmenty obrazu istotne z punktu widzenia interpretacji jego treści przez modele ML. Szczególnie ciekawym pomysłem jest zaproponowana analogia pomiędzy obrazem a językiem naturalnym, która pozwoliła przenieść część rozwiązań z NLP na grunt informacji obrazowej. Skuteczność proponowanych metod nie została jednak przetestowana na obrazach medycznych, które mają inne właściwości i poziom szczegółowości, niż rozważone obrazy scen naturalnych.

M. Fab.

Praca [A3] rozpoczyna drugi wątek poświęcony klasyfikacji obrazów mikroskopowych w zastosowaniach mikrobiologicznych. W szczególności, klasyfikacji poddawane są obrazy przedstawiające bakterie. Habilitant zaproponował oraz zaimplementował metodę klasyfikacji gatunków bakterii względem zagregowanej „głębokiej" oraz „płytkiej" reprezentacji. Reprezentacje płytkie pozyskane zostały z fragmentów obrazu w wyniku zastosowania deskryptora DSIFT. Z kolei do uzyskania reprezentacji głębokich zastosowane zostały architektury AlexNet, VGG-M, VGG-VD wykorzystywane jako ekstraktory cech. Cechy „płytkie" i „głębokie" zostały następnie zagregowane z wykorzystaniem metody FV (ang. *Fisher Vector*), stanowiąc dane wejściowe dla metody SVM z różnymi jądrami oraz lasów losowych (porównawczo). Optymalny wybór parametrów przeprowadzony został w wyniku optymalizacji Bayesowskiej. Eksperymenty zostały przeprowadzone na bazie DIBaS (ang. *Digital Images of Bacteria Species dataset*) zawierającej obrazy 33 gatunków bakterii. Najlepsze wyniki uzyskano stosując architekturę VGG-M oraz metodę FV odpowiednio do ekstrakcji i agregacji cech (redukcji wymiarowości) oraz metodę SVM do klasyfikacji. Elementem nowości w tej pracy jest agregacja „płytkiej" i „głębokiej" reprezentacji obrazów mikroskopowych bakterii. Nie jest jednak jasne, jak proponowana metoda wiąże się z wytłumaczalnym i interpretowalnym AI.

Podobną metodologię Habilitant zastosował w problemie klasyfikacji gatunków grzybów rozważonym w pracy [A4]. W tym przypadku jako ekstraktory cech głębokich wykorzystał nowsze architektury CNN (tj. DenseNet169, Inception V3 oraz Resnet). Na etapie agregacji cech „płytkich" i „głębokich", oprócz metody FV sprawdził również metodę worka słów (ang. *Bag of Words*, BoW), wykazując, że pierwsza metoda pozwala na uzyskanie lepszych wyników. W kontekście wyjaśnialności decyzji modelu, Habilitant zaproponował wizualną macierz błędów, w której przykłady błędnie sklasyfikowane zestawiono z przykładami reprezentującymi poprawnie sklasyfikowane próbki. Wykorzystał również klastry worka słów do znalezienia, a następnie wizualizacji fragmentów obrazu reprezentujących każdą z rozważanych klas i fragmentów obrazów znajdujących się najbliżej w przestrzeni cech od środków klastrów. Pozwoliło to zdefiniować pewne wizualne, morfologiczne oraz przestrzenne cechy odróżniające gatunki grzybów od siebie. Ponadto, wyznaczono rozkłady wartości wybranych cech dla każdego klastra w worku słów i skonfrontowano je z wiedzą domenową.

Praca [A5] podejmuje problem rozpoznawania (klasyfikacji) trzech klonów bakterii *klebsiella pneumoniae* na podstawie obrazów mikroskopowych. Z uwagi na dużą rozdzielczość obrazów, przetwarzanie zrealizowane zostało w oparciu o łaty wycięte z obrazu, których indywidualne reprezentacje, uzyskane w wyniku zastosowania bloku reprezentacji sieci ResNet-18 zostały zagregowane przy użyciu metody AbMIL (ang. *Attention Based Multiple Instance Learning*), tj. wielokrotnej instancji opartej na atencji. Uzyskana zagregowana reprezentacja stanowiła podstawę klasyfikacji przeprowadzonej względem trzech rozważanych rodzajów klonów bakterii. Takie podejście pozwoliło na osiągnięcie dokładności na poziomie 65% vs. 55% uzyskane dla klasycznego zastosowania CNN. W kontekście wyjaśnialności decyzji modelu, Habilitant zaproponował dwa dedykowane podejścia. Pierwsze z nich sięga po topologiczne metody analizy danych, wykorzystane w tym przypadku do opisu przestrzennego ułożenia bakterii. W szczególności, z centrów bakterii wykrytych metodami segmentacji obrazów (i traktowanych jako chmura punktów) utworzony został diagram persystencji (ang. *persitence diagram*, PD) opisujący kształt danych, poprzez kodowanie informacji o czasie powstania (narodzin) i zaniku (śmierci) poszczególnych struktur w obrazie w wyniku systematycznego pogrubiania środków komórek. Tak przygotowany diagram został następnie przekazywany do

A. Fab.

metody persystentnego worka słów (ang. *Persistence Bag of Words*, PBoW) [A8, A9] generującego wektor osadzenia o stałej długości. Średnie wektory dla każdego klonu pozwoliły na analizę różnic w przestrzennym ułożeniu. Druga z zaproponowanych metod zakłada wykorzystanie wyniku segmentacji obrazów do określenia i wyświetlenia statystyk dotyczących kształtu, wielkości i barwy komórek bakterii, w każdym z trzech rozważanych klonów. Metoda PBoW pozwoliła zaobserwować różnice w odległościach pomiędzy komórkami w poszczególnych klonach bakterii. Druga z rozważanych metod wykazała na statystycznie znaczące różnice w parametrach kształtu komórek. Analogicznie jak w pracy [A4], zastosowano również jakościowe macierze pomyłek, które w tym przypadku nie pozwoliły na uzyskanie znaczących informacji.

Ideę metody persystentnego worka słów wykorzystanej w pracy [A5], szerzej przedstawiają prace [A8, A9], dotyczące zastosowania persystentnej homologii w problemie klasyfikacji danych. W pracach tych Habilitant wprowadził persystentne słowniki, tj. zwektoryzowane reprezentacje diagramów persystencji o ustalonej długości. W autoreferacie Habilitant określa proponowane metody mianem interpretowalnych metod topologicznych, chociaż prace [A8, A9] nie odnoszą się bezpośrednio do problemu wyjaśnialności decyzji modeli ML. W szczególności, w pracy [A8] Habilitant zaproponował metodę persystentnego worka słów. Metoda jest inspirowana workiem słów popularnym w reprezentacji informacji tekstowej oraz obrazowej. Wejściowe diagramy persystencji danych transformowane są na współrzędne 2D (czas narodzin, persystencja, tj. czas życia obiektu) i konsolidowane w jedną przestrzeń danych wykorzystywaną do utworzenia słownika. Habilitant zaproponował dwa podejścia do tworzenia słownika wykorzystujące odpowiednio (i) metodę k-średnich oraz (ii) mieszaninę rozkładów Gaussa (ang. *Gaussian Mixture Modelling*, GMM). W pierwszym przypadku punkt danych jest zaliczany do słowa opisanego najbliższym centroidem. Wynikowy histogram reprezentuje natomiast liczbę wystąpień danego słowa. W drugim podejściu, które Habilitant określa mianem stabilnego persystentnego worka słów (ang. *Stable Persistence Bag of Words*, sPBoW), rozkład punktów w diagramie skonsolidowanym modelowany jest przez mieszaninę rozkładów Gaussa (GMM), gdzie liczba rozkładów wyznacza liczbę słów. W konsekwencji, dla nowego diagramu, histogram słów wyznaczany jest z uwzględnieniem, prawdopodobieństwa, że punkt danych należy do słowa opisanego przez dany rozkład Gaussa. Dodatkowo, w metodzie uwzględniono, że punkty diagramu o dużej persystencji powinny mieć większy wpływ na wyznaczanie klastrów (a co za tym idzie słów). Z tego powodu Habilitant zaproponował, aby zastosować taki podzbiór skonsolidowanej przestrzeni diagramów względem trwałości (długości życia), który spowoduje, że punkty o większej persystencji z większym prawdopodobieństwem będą mogły zostać środkami klastrów. W pracy [A9] Habilitant zaproponował dodatkowo, aby przy tworzeniu histogramu słów w słowniku uwzględniać się sumę składowych persystencji w danym worku, a nie liczbę elementów. Ten wariant metody określił mianem ważonego persystentnego worka słów (ang. *Weighted Persistence Bag of Words*, wPBoW). Dodatkowo, zaproponował persystentne adaptacje metod VLAD (ang. *Vectors of Locally Aggregated Descriptions*) oraz FV (ang. *Fisher Vector*), znajdujących swoje źródła w metodzie worka słów. Powyższe metody przetestowane zostały w problemie klasyfikacji na siedmiu zbiorach danych, reprezentujących dane o różnym charakterze (głównie chmur punktów 3D, powierzchni 3D, grafów sieci społecznych, oraz sekwencji 3D ruchów stawów), a także porównane z metodami konkurencyjnymi opartymi na jądrach oraz metodami wektorowymi. Metoda persystentnego worka słów okazała się wielokrotnie szybsza od metod konkurencyjnych przy nie gorszej dokładności. Okazała się też bardziej skuteczna, niż persystentne wersje VLAD oraz FV. W kontekście wytłumaczalności

AFab.

decyzji modeli, Habilitant wspomina zdawkowo, że można w tym celu wykorzystać wizualizację średnich wektorów reprezentacji uzyskanych metodą PBoW. Wytłumaczalność decyzji modelu nie jest jednak wątkiem przewodnim prac [A8, A9], które głównie koncentrują się na optymalizacji czasu wykonania klasyfikacji danych reprezentowanych przez diagramy persystencji.

Do wątku interpretowalnego uczenia maszynowego Habilitant powrócił w pracy [A6], proponując metodę ProtoPShare (ang. *Prototypical Part Shared*) dedykowaną do uzasadnienia predykcji przy użyciu tzw. części prototypowych (prototypów) posiadanych przez obiekt. Prototyp jest w tym przypadku rozumiany jako reprezentacja istotnego fragmentu obrazu, przedstawiającego cechy/elementy wspólne dla danej klasy. Metoda rozszerza sieć części prototypowych ProtoPNet (ang. *Prototypical Part Network*). Model sieci bazowej składa się z warstw splotowych służących jako ekstraktory cech, warstwy prototypowej oraz warstwy w pełni połączonej dokonującej klasyfikacji. Cechy obrazu źródłowego są wyodrębniane przez warstwy splotowe jako części prototypowe i porównywane z obrazami treningowymi, aby stworzyć mapę aktywacji pokazującą podobieństwo między nimi, i tym samym podkreślić obszar prototypowy szczególnie istotny z punktu widzenia podjętej decyzji. Podobieństwo uzyskane z warstwy prototypowej jest też podstawą do przypisania obrazu do danej klasy. W porównaniu do metody bazowej, zasadniczą nowością metody ProtoPShare jest współdzielenie prototypów pomiędzy klasami. Osiągnięcie tego możliwe było dzięki autorskiemu pomysłowi Habilitanta dot. znajdowania prototypów podobnych semantycznie, łączenia ich w jeden, przy jednoczesnym przypisywaniu „scalonego prototypu" do kilku klas, w których był obecny. W celu znajdowania prototypów podobnych semantycznie nawet, gdy są odległe w przestrzeni ukrytej, Habilitant zaproponował autorską miarę odległości. Jak wykazały eksperymenty na zbiorach CUB-200-2011 zawierającym obrazy 200 gatunków ptaków oraz StanfordCars prezentującym 196 modeli samochodów, dzięki współdzieleniu prototypów między wieloma klasami w warstwie prototypowej możliwe było kilkukrotne zmniejszenie liczby wymaganych prototypów, przy jednoczesnym zachowaniu dokładności metody bazowej. Zwiększa to interpretowalność modelu. Ponadto, Habilitant wyeliminował drugie ograniczenie metody bazowej, jakim jest odsuwanie od siebie prototypów różnych klas o podobnej semantyce. Podobne wnioski uzyskane zostały dla czterech różnych wariantów modelu ProtoPShare wykorzystującego w warstwie splotowej modele ResNet34, ResNet152, DenseNet121 oraz DenseNet161, ustalając limit 10 prototypów na klasę. Podejście nie zostało jednak sprawdzone na obrazach medycznych.

Z kolei w pracy [A7], Habilitant wykorzystał mechanizm atencji do zwiększenia wytłumaczalności modeli w problemach uczenia wielokrotnej instancji (ang. *Multiple Instance Learning*, MIL) będącego formą uczenia słabo nadzorowanego (ang. *Weakly Supervised Learning*). W szczególności, w tego typu problemach dane (instancje) uczące zgrupowane są w tzw. workach, a etykieta odnosi się do całej zawartości worka, a nie do pojedynczych instancji (jak w klasycznym uczeniu nadzorowanym). Negatywne worki zawierają wyłącznie negatywne przypadki, natomiast worki z pozytywną etykietą zawierają przynajmniej jedną pozytywną instancję. Habilitant rozszerzył metodę *Attention-based Deep Multiple Instance Learning* dedykowaną temu problemowi poprzez dodanie mechanizmu uwagi (ang. *self attention*) pozwalającego na modelowanie zależności pomiędzy instancjami danego worka. W szczególności, moduł ten wprowadzony został do architektury źródłowej pomiędzy warstwami splotowymi dokonującymi ekstrakcji cech, a procesem agregacji instancji danego worka poprzez wyznaczenie średniej reprezentacji ważonej wagami uzyskanymi z sieci neuronowej (agregacji). Warstwa agregująca

A.Fab.

przekształca znalezione zależności pomiędzy instancjami oraz informacje o ich kontekście na wektor o ustalonej długości, który następnie poddawany jest klasyfikacji przez warstwy gęste. Wynikowa metoda o akronimie SA-AbMIL (ang. *Self-Attention Attention-Based MIL*) została przetestowana w pięciu scenariuszach obrazowych, na zbiorach MNIST, dwóch rodzajach obrazów histopatologicznych (w problemach wykrywania raka piersi oraz raka jelita grubego), mikrobiologicznym zbiorze DIFaS oraz na zbiorze Messidor (zawierającym obrazy z fundus kamery w problemie detekcji cukrzycy). Eksperymenty wykazały, że metoda autorska jest w większości przypadków dokładniejsza, niż bazowa. Na etapie tworzenia map atencji, dodatkowo sprawdzona została możliwość zastąpienia iloczynu skalarnego jądrami: *Radial Basis Functions*, *Inverse Quadratic*, *Laplace* oraz *Module*, których zastosowanie korzystnie wpływa na dokładność klasyfikacji. Uzyskane mapy atencji stanowią natomiast wyjaśnienie decyzji modelu dla eksperta.

***

Podsumowując, **oryginalne osiągnięcia dra Bartosza Zielińskiego** przedstawione w ocenianym cyklu publikacji **obejmują** (zgodnie z załączonymi oświadczeniami Habilitanta):

1) **w zakresie rozwoju metod wytłumaczalnego i interpretowalnego uczenia maszynowego**:
   a) adaptację klasyfikatorów diagnostycznych do problemów obrazowych, poprzez zaproponowanie mapowania pomiędzy informacją tekstową a informacją obrazową [A1] oraz wykazanie w drodze analizy reprezentacji obrazu opartej na zniekształceniach i teorii percepcji wzrokowej Marra słów wizualnych o kluczowym oraz ograniczonym znaczeniu [A2],
   b) opracowanie metody ProtoPShare (ang. *Prototypical Part Shared*) dedykowanej do uzasadnienia predykcji przy użyciu części prototypowych posiadanych przez obiekt [A6],
   c) opracowanie metody SA-AbMIL (ang. *Self-Attention Attention-Based Multiple Instance Learning*) wykorzystującej mechanizm atencji do zwiększenia wytłumaczalności modeli w problemach uczenia wielokrotnej instancji [A7],
   d) wykorzystanie średnich wektorów reprezentacji uzyskanych metodą persystentnego worka słów (PBoW) oraz wizualizacja zawartości poszczególnych klastrów [A4],
2) **w zakresie rozwoju metod klasyfikacji danych reprezentowanych przez diagramy persystencji:**
   a) opracowanie metody persystentnego worka słów (PBoW) oraz jego stabilnej (sPBoW) i ważonej (wPBoW) wersji, a także persystentnych wersji metod VLAD oraz FV [A8, A9],
3) **w zakresie rozwoju metod klasyfikacji obrazów, w tym mikrobiologicznych:**
   a) opracowanie metody klasyfikacji działającej w oparciu o zagregowane „płytkie" oraz „głębokie" cechy obrazowe [A3, A4],
   b) opracowanie metody klasyfikacji agregującej reprezentację fragmentów obrazu przy użyciu wielokrotnej instancji opartej na atencji [A5].

Oryginalne rozwiązania wchodzące w skład recenzowanego osiągnięcia w większości posiadają również **przekonywującą weryfikację eksperymentalną**, która wskazuje na przewagę proponowanych rozwiązań w stosunku do wybranych metod konkurencyjnych.

A. Fab.

Za słabe strony recenzowanego osiągnięcia uważam:

- brak samodzielnych prac Habilitanta opisujących elementy składowe osiągnięcia,
- ograniczoną liczbę publikacji w czasopismach z dyscypliny informatyka techniczna i telekomunikacja opisujących elementy składowe osiągnięcia,
- tytuł jedynie częściowo oddający zagadnienia rozważane przez Habilitanta,
- dość heterogeniczną i nie zawsze jednoznacznie powiązaną tematykę publikacji wchodzących w skład cyklu.

Pomimo ww. uwag krytycznych, moja **ocena osiągnięcia naukowego dra Bartosza Zielińskiego** pt. *Wyjaśnialne i interpretowalne uczenie maszynowe z zastosowaniami w analizie obrazów medycznych* **jest pozytywna**, gdyż proponowane przez Habilitanta rozważania **stanowią ciekawy i oryginalny wkład naukowy**, poszerzający aktualny stan wiedzy.

5. Ocena aktywności naukowej Habilitanta zrealizowanej w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej zgodnie z art. 219 ust. 1 pkt. 3 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce

W roku 2019 dr Bartosz Zieliński odbył miesięczny staż podoktorski w Instituto Superior Tecnico Uniwersytetu Lizbońskiego w Portugalii. Nie jest jednak jasne, jaką istotną aktywnością naukową wykazał się w czasie tego pobytu oraz jakimi efektami naukowymi pobyt ten zaowocował. W dorobku Habilitanta nie znalazłam bowiem publikacji wspólnych z pracownikami Uniwersytetu Lizbońskiego.

Dr Bartosz Zieliński odbył również trzy pięciodniowe wizyty odpowiednio w Uniwersytecie Nauk Stosowanych w St. Polten w Austrii (w maju 2017 roku), w Uniwersytecie Technicznym w Wiedniu w Austrii (we wrześniu 2015 roku) oraz w Uniwersytecie w Zagrzebiu w Chorwacji (w kwietniu 2013 roku). Wizyty w ww. uniwersytetach austriackich zaowocowały łącznie czterema współautorskimi publikacjami z prof. Matthiasem Zeppelzauerem poświęconymi wykorzystaniu deskryptorów topologicznych w analizie powierzchni 3D oraz zastosowaniem metody worka słów i jej rozszerzeń jako deskryptorów diagramów persystencji. Jedna z tych prac opublikowana została w uznanym czasopiśmie *Artificial Intelligence Review*. Ponadto, Habilitant kierował grantem NCN w programie Sonata z udziałem prof. Zeppelzauera realizowanym w latach 2016-2020. Autoreferat nie wskazuje natomiast, jaką aktywnością wykazał naukową się Habilitant w Uniwersytecie w Zagrzebiu.

Kandydata prowadzi współpracę naukową (potwierdzoną wspólnymi publikacjami oraz projektami) z naukowcami reprezentującymi jednostki naukowe i badawcze w kraju (AGH, Instytut Matematyczny PAN, Politechnika Warszawska, Uniwersytet Wrocławski, Politechnika Krakowska) oraz zagranicą (m.in.: University Hospital St. Poelten, St. Poelten University of Applied Sciences, Northeastern University, Czech Academy of Sciences, Swansea University). W szczególności, wg. bazy Scopus (narzędzie Scival) od roku 2012 blisko 29% (tj. sześć) publikacji Kandydata powstało we współpracy międzynarodowej, natomiast ponad 33% (tj. siedem) publikacji – we współpracy z innymi instytucjami naukowymi w Polsce.

Uważam, że ww. aktywności **są wystarczające, aby uznać, że dr Bartosz Zieliński wykazuje się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.**

A. Fab.

6. Podsumowanie

W podsumowaniu niniejszej recenzji stwierdzam, że zarówno osiągnięcie naukowe dra Bartosza Zielińskiego w postaci cyklu powiązanych tematycznie artykułów naukowych, jak i Jego aktywność naukowa zrealizowana w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej **spełniają wymagania stawiane kandydatom do stopnia naukowego doktora habilitowanego** nauk inżynieryjno-technicznych w dyscyplinie informatyka techniczna i telekomunikacja. W związku z powyższym, wniosek **dra Bartosza Zielińskiego o nadanie stopnia doktora habilitowanego uznaję za zasadny.**

Anna Fabijańska

A. Fab.