

Dr hab. inż. Maciej Kusy, prof. PRz
Katedra Podstaw Elektroniki
Wydział Elektrotechniki i Informatyki
Politechnika Rzeszowska im. Ignacego Łukasiewicza
al. Powstańców Warszawy 12
35-959 Rzeszów
e-mail: mkusy@prz.edu.pl

Rzeszów, 16 września 2024 r.

Recenzja rozprawy doktorskiej
mgr. inż. Macieja Żelaszczyka
pt. „Deep Representation Learning in Varied Settings”

Prawną podstawą przygotowania recenzji rozprawy doktorskiej mgr. inż. Macieja Żelaszczyka jest uchwała Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja (RNDITiT) Politechniki Warszawskiej (PW) nr 697/2024 z dnia 21 maja 2024 r., w której Rada Dyscypliny powołuje recenzentów. O treści uchwały zostałem poinformowany pismem Przewodniczącego RNDITiT – prof. dr. hab. inż. Jarosława Arabasa, które otrzymałem w dniu 14 czerwca 2024 r.

Niniejsza recenzja została przygotowana na podstawie rozprawy doktorskiej. Przedmiotowa praca została zrealizowana na Wydziale Matematyki i Nauk Informacyjnych PW, pod kierunkiem prof. dr. hab. inż. Jacka Mańdziuka.

Rozprawa, zawierająca streszczenie w języku polskim i angielskim, wprowadzenie, zagadnienia wstępne, pogrupowaną w cztery rozdziały analizę: (i) odzyskiwania cech w kontekście adversarialnym, (ii) tłumaczenia dźwięku na obraz za pomocą wspólnej reprezentacji, (iii) wykorzystania wspólnych komponentów do uczenia się interpretowalnych reprezentacji zmiennych, (iv) narzucania ograniczeń na komponenty reprezentacji poprzez uczenie adversarialne oraz bibliografię, liczy łącznie 140 stron. Zawiera łącznie 35 rysunków, 22 tabele i 7 algorytmów. Literatura obejmuje 11 stron. W tej grupie znajduje się 5 prac, których Doktorant jest współautorem.

1. Tematyka pracy

Praca koncentruje się na zagadnieniu uczenia się reprezentacji, które jest kluczowe dla projektowania modeli sztucznej inteligencji zdolnych do osiągnięcia ludzkiego poziomu percepcji i rozwiązywania problemów. W przedmiotowej rozprawie mgr inż. Maciej Żelaszczyk analizuje różne przypadki wykorzystania procesu uczenia w różnych zadaniach i kontekstach. W jednym z przypadków badane jest nienadzorowane odzyskiwanie surowych obrazów z adversarialnych (uszkodzonych) danych wejściowych, gdzie wykorzystuje się pośrednie reprezentacje do rekonstrukcji zarówno obrazów, jak i informacji potrzebnych do klasyfikacji. Kolejny przypadek dotyczy konwersji sygnałów dźwiękowych na obrazy przy użyciu wspólnych reprezentacji audiowizualnych, co pozwala na generowanie wizualnie spójnych obrazów, zawierających istotne informacje. Doktorant porusza również temat osadzania zmiennych, które są poddane ograniczeniom strukturalnym, zarówno w odniesieniu do samych reprezentacji, jak i mechanizmu uwagi używanego do ich połączenia. Odpowiada także na pytanie, jak ograniczyć reprezentacje, skłaniając klasyfikator do nauki takich reprezentacji, dla których trudno jest przewidzieć brakującą część na podstawie pozostałej reprezentacji.

Uważam, że podjęta w rozprawie doktorskiej mgr. inż. Macieja Żelaszczyka tematyka badawcza jest z naukowego punktu widzenia ważna i aktualna.

2. Cel i zakres pracy

Celem pracy była analiza mocy uczenia się reprezentacji w czterech odmiennych przypadkach, różniących się zadaniami, poziomem nadzoru oraz domeną danych. W szczególności Doktorant: (i) zaprezentował sposób odzyskiwania surowych obrazów z adversarialnych danych wejściowych przy użyciu nienadzorowanych reprezentacji pośrednich; (ii) wykorzystał wspólne reprezentacje audiowizualne do generowania obrazów z sygnałów; (iii) na podstawie współdzielonych reprezentacji zaprezentował sposób nadzorowanej klasyfikacji danych tabelarycznych w podejściu uczenia wielozadaniowego; (iv) przeprowadził analizę adversarialnych metod restrykcji reprezentacji w odniesieniu do nadmiarowości informacji w wysokopoziomowych reprezentacjach.

3. Hipoteza badawcza

Mgr inż. Maciej Żelaszczyk nie stawia w swojej rozprawie tezy. Formułuje natomiast następującą hipotezę badawczą: *Uczenie się reprezentacji umożliwia rozwiązywanie różnych zadań w odmiennych dziedzinach*. Zakłada ona, iż różnorodność analizowanych w pracy konfiguracji, architektur i metod uczenia modeli sztucznej inteligencji, w ramach których możliwe jest uzyskanie istotnych wewnętrznych reprezentacji, potwierdza, iż uczenie się reprezentacji jest metodą o charakterze ogólnym. Oznacza to, że metoda ta znajduje zastosowanie w szerokim zakresie kontekstów.

Realizacja rozprawy doktorskiej, w której Autor nie formułuje i nie wykazuje poprawności (prawdziwości) tezy nie jest podejściem powszechnie stosowanym, ale nie ma wymogów formalnych, aby tezę do pracy wprowadzać. Uważam podejście Autora za jak najbardziej poprawne.

4. Struktura pracy

Praca napisana jest w języku angielskim. Zaprezentowane w dysertacji zagadnienia zostały opisane w 7 rozdziałach.

Rozdział 1 – *Introduction* składa się z trzech sekcji. Pierwsza z nich przedstawia ideę reprezentacji danych i ich transformacji z postaci surowej do nowej formy na podstawie wiedzy o analizowanym problemie (ang. feature engineering) oraz krótko charakteryzuje technikę uczenia się reprezentacji. W sekcji tej umieszczone są ponadto cele badawcze Autora. W drugiej części tego rozdziału zaprezentowana jest hipoteza badawcza, która zostanie zweryfikowana, aby zrealizować zamierzone cele. Ta część prezentuje również spis publikacji, których mgr inż. Maciej Żelaszczyk jest współautorem. Należy zaznaczyć, iż na 6 podanych prac, 3 nie zostały opublikowane. W ostatniej sekcji Autor omawia strukturę rozprawy i nakreśla jak treści poszczególnych rozdziałów odnoszą się do własnych publikacji.

Rozdział 2 – *Preliminaries* przedstawia przegląd podejść istotnych dla architektur i metod zaprezentowanych w rozprawie. Na samym początku tej części pracy Autor wyjaśnia termin modelu generatywnego. Podaje jego interpretację w z punktu widzenia dwóch definicji oraz wskazuje różnicę w odniesieniu do modelu dyskryminacyjnego. Dokonuje opisu formalnego dwóch architektur: autoenkodera wariacyjnego oraz generatywnej sieci adwersarialnej. Następnie omawia zagadnienie ataków adwersarialnych z punktu widzenia analizy obrazów, wyszczególniając w kolejnych sekcjach ich rodzaje, a w szczególności: *L-BFGS*, metodę *fast gradient sign method* czy *projected gradient descent*. Rozdział kończy się krótkim

przedstawieniem istoty zagadnienia osadzania słów; wprowadzono oznaczenia dla zmiennych reprezentujących słowo, słownictwo czy otoczenie. Podana jest również interpretacja pojęcia otoczenia w odniesieniu do kontekstu konceptu w przypadku przetwarzania języka naturalnego.

Rozdział 3 – *Adversarial defenses via a mixture of generators* koncentruje się na procedurze przetwarzania danych adversarialnych za pomocą kombinacji generatorów w celu odzyskania poprawnej klasy sprzed ataku. Doktorant przedstawia metodę obrony przed atakami adversarialnymi za pomocą zestawu konkurujących sieci generatorów w konfiguracji typu generative adversarial network (GAN). Każdy generator przetwarza dane z zakłóceniami adversarialnymi, tworząc nową reprezentację, która usuwa skutki ataku, bez potrzeby korzystania z etykiet obrazów lub ataków. To z kolei umożliwia aktualizację parametrów dyskriminatora na podstawie oryginalnych obrazów i przeprowadzanie finalnej klasyfikacji. Metoda dobrze radzi sobie z wieloma atakami, zachowując spójność obrazu i przydatność do klasyfikacji. Wariant, w którym każdy generator jest uczony na konkretnych atakach, poprawia dokładność, a zastosowanie zespołu mniejszych generatorów jest bardziej efektywne i modułowe niż jedna duża sieć. Podejścia testowane są na zbiorze danych MNIST.

Rozdział 4 – *Audio-to-image cross-modal generation* prezentuje sposób tworzenia obrazów z sygnałów dźwiękowych za pomocą cross-modalnej architektury autoencodera wariacyjnego (VAE) oraz jej modyfikacji. Autor wprowadza do analizy dwa zbiory danych konwertujących sygnał na obraz: wiele obrazów na jeden sygnał oraz unikalne pary sygnał-obraz; ocenia, jak te zbiory wpływają na generowanie obrazów. Po procesie uczenia obydwu architektur na tych zbiorach, dane sygnałowe można reprezentować jako cechy audio-wizualne, przez co modele mogą wygenerować obrazy, które wizualnie pasują do obrazów rzeczywistych. W rozdziale omówione są również: (i) dostosowanie straty rekonstrukcji do kontroli dopasowania między rzeczywistymi a generowanymi obrazami, (ii) oceny jakościowe możliwości generowania obrazów z danych dźwiękowych, oraz (iii) potencjał generowania obrazów z sygnałów dźwiękowych w dziedzinie uczenia cross-modalnego.

Rozdział 5 – *Interpretable multi-task learning with shared variable embeddings* wprowadza do zagadnienia wspólnego osadzania zmiennych dla modeli predykcyjnych w problemach wielozadaniowych. Autor proponuje metodę, która rozszerza tradycyjne podejście do osadzania zmiennych. Umożliwia ona ponowne wykorzystanie osadzeń między zmiennymi i zbiorami danych za pomocą wspólnej macierzy osadzeń i mechanizmu uwagi. Metoda osiąga porównywalną do oryginalnego podejścia dokładność klasyfikacji, ale bez konieczności dostrajania modelu. Eksperymenty pokazują, że wariant z rzadką uwagą zwiększa

dokładność i skraca czas uczenia, kosztem interpretowalności, podczas gdy podstawowa metoda przewyższa inne pod względem wyjaśnialności.

Rozdział 6 – *Restricting Representations* omawia nową procedurę uczenia klasyfikatorów, w której wysokopoziomowe reprezentacje wewnątrz klasyfikatora są wykorzystywane zarówno jako dane wejściowe, jak i cele dla predyktorów, a oba te elementy są uczone w sposób adwersarialny. Klasyfikator dąży do minimalizacji swojej funkcji straty oraz maksymalizacji straty predyktorów, podczas gdy predyktory minimalizują własną funkcję straty. Poprzez wzajemną rywalizację klasyfikatora i predyktorów model uczy się bardziej odpornych i ogólnych cech. Na podstawie przeprowadzonych eksperymentów, Autor pokazuje, że wykorzystanie wysoko skorelowanych komponentów reprezentacji pomaga predyktorom generalizować poza dane uczące.

Ostatni rozdział stanowi podsumowanie rozprawy. W tej części pracy, na podstawie uzyskanych wyników i wniosków, Autor potwierdza zasadność postawionej hipotezy.

5. Ocena rozprawy doktorskiej

5.1. Uwagi krytyczne

Uwagi ogólne:

1. Mimo wysokiego poziomu użytego języka, w tekście rozprawy brakuje prostszych, inżynierskich interpretacji, definicji i opisów (jednym z wielu przykładów jest kluczowe pojęcie uczenia reprezentacji). Tekst jest płynny, lecz brakuje mu precyzyjnych, technicznych wyjaśnień. Styl charakteryzuje się nadmierną rozwlekłością, zamiast rzeczowej precyzji a zdania często niepotrzebnie komplikują proste zagadnienia. Lepsze byłoby zwięzłe i konkretne wykorzystanie języka, co jest istotne w pracach naukowych, gdzie wymagana jest jasność i jednoznaczność kluczowych zagadnień.
2. Algorytmy przedstawione w pracy mają raczej charakter opisowy niż proceduralny, co bardziej odpowiada formatowi pseudo-kodu niż formalnemu zapisowi algorytmu. W przypadku algorytmów oczekuje się połączenia syntaktycznej formy językowej, zapisu formalnego oraz zwięzłych opisów tekstowych. Dodatkowo, dla zwiększenia przejrzystości, poszczególne kroki algorytmów powinny być ponumerowane. Umożliwiłoby to odniesienia do konkretnych linii w tekście, co ułatwiłoby analizę i interpretację ich działania. W przypadku algorytmów takich jak 4.1 i 4.2, które są stosunkowo krótkie i intuicyjne, takie podejście nie jest konieczne, jednak w odniesieniu do bardziej złożonych algorytmów, jak 3.1, 4.3, 4.4 oraz 6.1, numeracja

i szczegółowy opis kroków są, moim zdaniem, niezbędne. Dobrze ilustruje to zdanie na stronie 107: „The overall training procedure is presented as Algorithm 6.1. It includes one modification to Equation 6.3.” Algorytm 6.1 zajmuje całą stronę 108, jednak brakuje odniesienia do konkretnego kroku, w którym następuje uczenie klasyfikatora w oparciu o formułę (6.4) rozszerzającą (6.3). Numeracja kroków oraz bardziej formalny zapis algorytmów znacząco zwiększyłyby ich czytelność i zrozumiałość dla odbiorcy.

3. Czy dokładności sieci LeNet5 w omawianych symulacjach w rozdziałach 3 oraz 4 uzyskano na podstawie pojedynczych symulacji? Jeśli tak, to wyniki mogą być przypadkowe. Przy licznych powtórzeniach, wykresy pudełkowe wraz ze średnimi i odchyleniami standardowymi mogłyby potwierdzić stabilność i powtarzalność rezultatów. Podobne pytanie należy zadać w przypadku wyznaczania dokładności klasyfikacji dla metod osadzania zmiennych przedstawionych w rozdziale 5.
4. Z uwagi na to, że rozdziały *Adversarial Defenses via a Mixture of Generators* oraz *Audio-to-Image Cross-Modal Generation* są tytułami artykułów opublikowanych w recenzowanych materiałach z międzynarodowych konferencji naukowych, odpowiednio: ICONIP (Bali, 2021) oraz IJCNN (Padwa, 2022), Doktorant powinien szczegółowo wyjaśnić różnice między opublikowanymi pracami a treścią dysertacji. Wynika to z faktu, iż w ocenianych rozdziałach umieszczono liczne rysunki, tabele, algorytmy i formuły pochodzące z wcześniej opublikowanych prac. Od Autora oczekuje się również właściwego zastosowania odnośników do tych publikacji.

Uwagi do poszczególnych rozdziałów

1. Uwagi do Rozdziału *Introduction*
 - Rozdział nie przedstawia jasno określonych celów badawczych. Mimo, iż na stronie 4 rozprawy Doktorant wyjaśnia nad czym skoncentruje się w rozprawie oraz wymienia dwa cele, przedmiotowa treść nie jest wyrazista w kontekście całości pracy. Rozdział wstępny wymaga także dokładnego przedstawienia **ogólnych celów pracy** (które można uszczegółowić, co zrobiono na str. 24-25, 47-48, 74-75 i 103), **motywacji** oraz **szczegółowego opisu problematyki badawczej**, czego w obecnej wersji brakuje. W pracy Autor wprowadza *podzielenie celów i motywacji na rozdziały*, co wynika z faktu, iż rozdziały te bazują w dużej mierze na opublikowanych pracach. Wydaje się to być słusznym podejściem ze względu na zróżnicowanie tematyczne poszczególnych części rozprawy; niestety

przedstawiona motywacja nie jest wystarczająca. Przykładowo jako czytelnik chciałbym znać odpowiedź na pytanie, dlaczego zagadnienie generowania obrazów z sygnałów dźwiękowych jest ważne (Rozdział 4). Po co je stosować? Czy jest pożądane, słuszne a może czy w ogóle ma sens?

- Przedstawienie na stronie 5 publikacji [1]-[4] weryfikujących hipotezę H1 nie jest konieczne; może być nawet nieco mylące, zwłaszcza że Doktorant wymienia dwie publikacje [5,6] niezwiązane z przedmiotową hipotezą. Autorskie prace umieszczone są w spisie literatury i jest to w mojej opinii wystarczające.
- Ogólnie sekcja 1.3. *Outline* wprowadza zamieszanie. Autor, z jednej strony, opisuje, co zostanie przedstawione w poszczególnych rozdziałach, a z drugiej, wprowadza nagłówki, które nie są zgodne z faktycznymi tytułami rozdziałów. Na przykład: na stronie 6 pojawia się nagłówek *Chapter 3 – feature recovery in an adversarial setting*, podczas gdy rozdział 3 nosi tytuł ADVERSARIAL DEFENSES VIA A MIXTURE OF GENERATORS. Podobnie, na stronie 7 doktorant wprowadza nagłówek *Chapter 4 – translating audio into an image through a common representation*, podczas gdy rzeczywisty tytuł rozdziału to AUDIO-TO-IMAGE CROSS-MODAL GENERATION. Jako recenzent i czytelnik, zupełnie nie rozumiem, co kryje się za takim podejściem. Dlaczego nie przedstawić struktury pracy w prosty i przejrzysty sposób, wyjaśniając jasno, z jakich rozdziałów się składa oraz w jaki sposób omówione w nich zagadnienia są ze sobą powiązane? Takie podejście znacząco ułatwiłoby odbiór i zrozumienie rozprawy. Przykładowo, jak – po przeczytaniu rozdziałów 1 i 2 – odnieść do całości rozprawy zawartość rozdziału trzeciego?

2. Uwagi dotyczące *Rozdziału 3*

- Jak zdefiniowany jest wskaźnik sukcesu w Tabeli 3.2?
- Jak jest uzasadnienie wyboru hiperparametru h ?
- Czy ze względu na wykorzystane architektury i zastosowanie modeli wstępnie wyuczonych złożoność obliczeniowa jest niewielka (zwłaszcza dla generatora o większej strukturze)?

3. Uwagi dotyczące *Rozdziału 4*

- Jak praktycznie zrealizowany jest proces uczenia przy funkcjach straty zdefiniowanych w (4.2) i (4.3) w kontekście algorytmów 4.3 i 4.4? Nie jest jasne, czy Autor zastosował własną implementację, czy też oparł się na istniejących

rozwiązaniach. Zaleca się doprecyzowanie tego aspektu, aby umożliwić pełne zrozumienie przyjętej metodyki oraz ocenę, na ile implementacja jest oryginalna.

- W sekcji 2.2.3 α jest hiperparametrem ataku. Jakie znaczenie α ma w optymalizacji (4.3)? Dlaczego w eksperymentach przyjęto wartości: $\alpha = \{0.2, 0.5, 1, 2\}$?
- Pytanie: czy można by zwizualizować rezultat Algorytmu 4.4 (np. dla jednej przykładowej cyfry)? Chodzi mi tu o graficzną prezentację tej cyfry po usuwaniu kolejnych cech (elementów) z krokiem (np.) 10; zarówno postać sygnału jaki i obrazu. Umożliwiłoby to zaobserwowanie krokowego niszczenia cyfry.

4. Uwagi dotyczące *Rozdziału 5*

Jak wyznaczana jest dokładność klasyfikacji dla porównywanych w Tabeli 5.1 (a także Tabeli 5.2) metod osadzania zmiennych? Co oznacza najlepsza („best” w opisie Tabeli 5.1) dokładność w tym kontekście?

5. Uwagi dotyczące *Rozdziału 6*

Jak dobrze rozumiem, liczba 512-stu składowych wektora wynika ze struktury modyfikowanej sieci ResNet18. Czy w związku z tym można coś powiedzieć o tych składowych? Które są najczęściej razy wybierane do predykcji w trakcie procesu uczenia? Co reprezentują? Czy da się je jakoś zinterpretować?

5.2. Uwagi drobne:

1. *Statistical metrics* – słowo *metric* powinno być wykorzystywane jako funkcja, służąca do mierzenia odległości w zadanej przestrzeni.
2. Prace nieopublikowane (submitted: [3,6]) oraz niezłożone (work in progres: [4]) nie powinny być wyszczególnione w sekcji 1.2 i 1.3.
3. Strona 3, zdanie: “In the neural network context, this means that the representations obtained in the final layers of a neural network are reused for tasks or datasets other than the ones on which the network was initially trained”. Nie do końca jest jasne, co oznacza: “final layers”. Ponadto, lepiej użyć słowo: „adjusted” zamiast „reused”, co jest wyjaśnione w kolejnym zdaniu.
4. W rozdziale *Preliminaries* Autor powinien umieścić informację o zastosowaniu poszczególnych, opisanych w sekcjach zagadnień.
5. Strona 20: Poprzez zbiór V oznaczone jest słownictwo; następnie w formule (2.14) słownictwo jest reprezentowane przez macierz E . Czy nie chodzi tu o otoczenie?

6. Wykorzystane w Tabeli 3.1 skróty do nazw parametrów generatorów i dyskryminatorów nie zostały wyjaśnione; skróty wszystkich nazw ataków, wraz z odniesieniami do pozycji literaturowych wymienionych w Tabeli 3.2 (str. 31), zostały wprowadzone dopiero na stronie 35. W Tabelach 4.2, 4.3 i 4.4 skróty parametrów modeli są niejasne. W pracach naukowych wymaga się wcześniejszego wyjaśnienia skrótów wykorzystywanych w treści.
7. Nazwa parametru dokładności powinna być ujednolicona w rozprawie. Autor wykorzystuje różne nazwy: „ACCURACY”, „ACC”, „Accuracy”.
8. Błędy językowe: “training such a system allows to recover representations” (str. 6); powinno być: “allows for recovering”; Podobnie: “We show that training such a system allows to recover”; Czy w zdaniu: “Area to which a given dataset and, consequently, a variable belongs, according to the delineation introduced by the UCI Machine Learning Repository.” nie brakuje przypadkiem orzeczenia? Skomplikowane konstrukcje językowe, np.: “The adversarial attack and defense interplay in our setting for image classification begins with a classifier trained on a train subset”; czy nie lepiej: “The interplay between adversarial attacks and defenses in our image classification setting begins ...”

Nie jest konieczne odnoszenie się przez Doktoranta do zaprezentowanych powyżej uwag drobnych.

5.3. Ocena ogólna

Formułując ocenę ogólną, należy przede wszystkim podkreślić, iż mgr inż. Maciej Żelaszczyk bardzo dobrze rozumie zagadnienia związane z uczeniem się reprezentacji, które odgrywa kluczową rolę w projektowaniu zaawansowanych modeli sztucznej inteligencji. Posiada dogłębną wiedzę na temat różnych metod przetwarzania danych, takich jak nienadzorowane odzyskiwanie obrazów z uszkodzonych danych oraz konwersja sygnałów dźwiękowych na obrazy przy użyciu wspólnych reprezentacji audiowizualnych. Doskonale zna procesy wykorzystywane w klasyfikacji wielozadaniowej, w tym osadzanie zmiennych w modelach predykcyjnych, co umożliwi poprawę wyników klasyfikacji oraz ich interpretację. Ponadto, posiada umiejętność analizy nadmiarowości informacji w zadaniach klasyfikacji obrazów, co świadczy o jego zaawansowanym rozumieniu zarówno nadzorowanego, jak i adwersarialnego uczenia modeli.

Praca doktorska mgr. inż. Macieja Żelaszczyka ma charakter eksperymentalny, co wyraźnie widać w obszernych sekcjach prezentujących wyniki badań w rozdziałach 3–6. Takie podejście jest niezbędne ze względu na istotny wkład algorytmiczny (Autor zaproponował 7 algorytmów), parametryzację modeli głębokich, liczbę danych oraz propozycje architektur neuronowych. Mimo eksperymentalnego charakteru, praca zachowuje wymagany formalizm matematyczny, prezentując zarówno podstawowe informacje, jak i równania niezbędne do reprezentacji danych, formułowania funkcji celu oraz realizacji zmodyfikowanych procesów uczenia. Rozprawa jest przez to spójna i dojrzała.

Pomimo niewątpliwej wartości merytorycznej, praca posiada kilka obszarów wymagających dopracowania, na które zwróciłem już uwagę we wcześniejszych komentarzach. Obejmują one zarówno kwestie kluczowe, jak i drobniejsze detale. W ujęciu ogólnym, można dostrzec niepotrzebne komplikowanie prostych zagadnień, co wynika z nadmiernie wyszukanego języka i stylu. Brakuje bardziej formalnego, inżynierskiego opisu zaproponowanych algorytmów, jak również podstawowej analizy statystycznej wyników, co może rodzić trudności w powtarzalności uzyskanych rezultatów. Ponadto, praca nie zawiera kompleksowego przedstawienia celów i motywacji w szerszym kontekście rozprawy. Chociaż poruszane zagadnienia mają wspólny naukowo-badawczy mianownik, Autor powinien, moim zdaniem, zaprezentować cele i motywację w sposób bardziej całościowy, szczególnie biorąc pod uwagę aktualność i wagę podejmowanej tematyki.

Niezależnie od opisanych przeze mnie niedociągnięć, praca posiada ogromny potencjał, który można jeszcze bardziej wyeksponować poprzez uwzględnienie przedstawionych sugestii. Odpowiedzi na zawarte w recenzji komentarze i pytania niewątpliwie utwierdzą mnie w przekonaniu, iż opracowany materiał jest kompletny, a przede wszystkim unikalny.

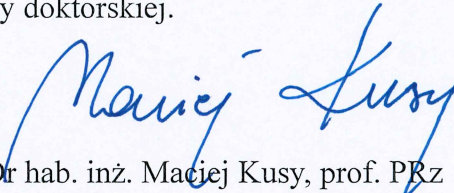
Elementami wyróżniającymi dorobek Autora są bez wątpienia: artykuł naukowy opublikowany w renomowanym międzynarodowym czasopiśmie *Information Fusion* (IF = 14,7) a także prace przedstawione na prestiżowych konferencjach międzynarodowych: *International Joint Conference on Neural Networks* oraz *International Conference on Neural Information Processing*, które następnie zostały opublikowane w recenzowanych materiałach z tych konferencji. W roku publikacji artykuły te, w ostatecznej formie, znalazły się w wykazie sporządzonym zgodnie z przepisami wydanymi na podstawie art. 267 ust. 2 ustawy z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (Dz. U. 2018 poz. 1668). Należy podkreślić, iż w pracach tych mgr inż. Maciej Żelaszczyk jest autorem pierwszym.

6. Podsumowanie

Rozprawa doktorska mgr. inż. Macieja Żelaszczyka, przygotowana pod opieką prof. dr. hab. inż. Jacka Mańdziuka, stanowi oryginalne rozwiązanie zagadnienia naukowego oraz potwierdza ogólną wiedzę teoretyczną i praktyczną Doktoranta w danej dyscyplinie naukowej. Na podstawie przedmiotowego przeglądu literatury, opisu zagadnień teoretycznych, sformułowania problemu, wykonanej analizy porównawczej i prezentacji wyników badań, stwierdzam, iż wiedzę Autora w dyscyplinie *informatyka techniczna i telekomunikacja* należy ocenić bardzo wysoko. Opieka Promotora i odpowiednie kierowanie rozwojem naukowym Doktoranta w trakcie realizowanego przewodu doktorskiego przyczyniły się do opublikowania wysoko jakościowych wspólnych prac w prestiżowych czasopismach i punktowanych konferencjach naukowych, co jest bezdyskusyjnie osiągnięciem wybitnym.

Stwierdzam zatem, że recenzowana przeze mnie rozprawa doktorska mgr. inż. Macieja Żelaszczyka pt. „Deep Representation Learning in Varied Settings” spełnia kryteria określone w ustawie z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce” (Dz. U. 2018 poz. 1668). Na tej podstawie wnioskuję do Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej o dopuszczenie Doktoranta do dalszych etapów przewodu doktorskiego.

Uwzględniając aktualność tematyki badawczej, jej szeroki zakres, wysoką jakość prezentowanych wyników oraz ich istotny wkład w rozwój dyscypliny, jak również znaczącą aktywność naukową mgr. inż. Macieja Żelaszczyka, udokumentowaną licznymi publikacjami, wnoszę o wyróżnienie recenzowanej rozprawy doktorskiej.


Dr hab. inż. Maciej Kusy, prof. PRz