Politechnika Wrocławska

Wrocław, 29 lipca 2024

Prof. dr hab. inż. Urszula Markowska-Kaczmar
Katedra Sztucznej Inteligencji, Politechnika Wrocławska

Recenzja rozprawy doktorskiej

sporządzona w postępowaniu o nadanie stopnia doktora magistrowi Marcinowi Przewięźlikowskiemu w dyscyplinie informatyka techniczna i telekomunikacja.

Dane formalne

- **Autor rozprawy doktorskiej:** mgr Marcin Przewięźlikowski
- **Tytuł rozprawy:** *Data-Efficient Representation Learning and Adaptation*
- **Promotor:** dr hab. Marek Śmieja
- **Jednostka prowadząca postępowanie:** Wydział Matematyki i Informatyki Uniwersytetu Jagiellońskiego
- **Podstawa sporządzenia recenzji:** uchwała Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Uniwersytetu Jagiellońskiego z dnia 21 maja 2026, informacja zawarta w piśmie z dnia 25.05.2026 podpisanym przez Zastępcę Przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja dr hab. Barbarę Strug, prof. uczelni.
- **Dyscyplina:** informatyka techniczna i telekomunikacja

1. Charakterystyka rozprawy

Recenzowana rozprawa doktorska oparta jest na cyklu publikacji. Autoreferat napisany w języku angielskim składa się z pięciu rozdziałów i spisu literatury. W dodatku A załączono oryginały artykułów tworzących cykl publikacji. Autoreferat obejmuje wprowadzenie, rozdział odnoszący się do metod efektywnej adaptacji modelu przy małej liczbie danych (*Few-shot Learning methods for data-efficient adaptation*), rozdział trzeci poświęcony analizie i doskonaleniu reprezentacji uczenia samonadzorowanego (*Understanding and improving Self-supervised representations*) oraz rozdział czwarty przedstawiający sylwetkę kandydata. Całość zamykają podsumowanie, spis literatury oraz dodatek z oryginałami artykułów wchodzących w skład cyklu.

Tematyka pracy obejmuje adaptację metod uczenia z niewielką liczbą przykładów oraz samonadzorowane uczenie reprezentacji obrazów. Wspólnym mianownikiem obu nurtów tematycznych jest ograniczona dostępność zaetykietowanych (zaanotowanych) danych. Praca podejmuje problem istotny z punktu widzenia rozwoju dyscypliny informatyka techniczna i telekomunikacja, w szczególności w obszarze uczenia maszynowego i wizji komputerowej, dotyczący tworzenia systemów zdolnych do działania przy niewielkiej liczbie zaanotowanych (etykietowanych) danych oraz uczenia reprezentacji z wykorzystaniem uczenia samonadzorowanego i federacyjnego dla ograniczonej liczby zaetykietowanych. Zakres rozprawy obejmuje: opracowanie i analizę metod *fewshot* opartych na hipersieciach, opracowanie metod reprezentacji uczenia samonadzorowanego (w tym MIM – Masked Image Modeling oraz federacyjnych metod SSL – SelfSupervised Learning) oraz ich eksperymentalną ocenę pod kątem jakości reprezentacji i efektywnej adaptacji modeli.

W skład cyklu publikacji wchodzą następujące prace:

[I] Marcin Przewięźlikowski*, Przemysław Przybysz*, Jacek Tabor, Maciej Zięba, Przemysław Spurek. *HyperMAML: Few-shot adaptation of deep models with hypernetworks,* Neurocomputing, Volume 598, 2024, 128179, ISSN 0925-2312 (140 pkt).

[II] Marcin Sendera*, Marcin Przewięźlikowski*, Konrad Karanowski, Maciej Zięba, JacekTabor, Przemysław Spurek. *HyperShot: Few-shot Learning by Kernel HyperNetworks,* Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, (WACV), 2023, pp. 2469-2478 (CORE A, 140 pkt).

[III] Marcin Sendera*, Marcin Przewięźlikowski*, Jan Miksa, Mateusz Rajski, Konrad Karanowski, Maciej Zięba, Jacek Tabor, and Przemysław Spurek. *The general framework for Few-shot Learning by kernel HyperNetworks,* Machine Vision and Applications 34, no. 4 (2023): 53. (100pkt).

[IV] Marcin Przewięźlikowski, Mateusz Pyla, Bartosz Zieliński, Bartłomiej Twardowski, Jacek Tabor, Marek Śmieja. *Augmentation-aware Self-supervised Learning with Conditioned Projector*, Knowledge-based Systems 2024, 112572, ISSN 0950-7051 (200pkt)

[V] Marcin Przewięźlikowski, Randall Balestriero, Wojciech Jasiński, Marek Śmieja, Bartosz Zieliński. *Beyond [cls]: Exploring the True Potential of Masked Image Modeling Representation,* Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2025, pp. 23442-2345, (CORE A*, 200 pkt).

[VI] Marcin Przewięźlikowski, Marcin Osial, Bartosz Zieliński, Marek Śmieja. *A deep cut into Split Federated Self-supervised Learning,* Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD), 2024/8/22, 444-459 (CORE A, 140 pkt).

Dwa wspomniane wyżej nurty tematyczne omawiane w rozprawie znajdują odzwierciedlenie w następujących publikacjach: prace [I, II, III] dotyczą efektywnej adaptacji modeli przy małej liczbie danych, natomiast trzy kolejne prace [IV, V, VI] odnoszą się do uczenia samonadzorowanego (SSL) oraz uczenia federacyjnego przy tym samym założeniu.

Rozpoczynając omawianie rozprawy od nurtu efektywnej adaptacji modeli przy niewielkiej liczbie zanotowanych danych, praca [I] wprowadza HyperMAML – metodę *fewshot*, w której zamiast klasycznego uczenia gradientowego stosowanego w MAML (Model-Agnostic Meta-Learning) do adaptacji modelu wykorzystuje się trenowalną hipersieć generującą zadaniowo specyficzne aktualizacje parametrów. Pozwala to szybciej i skuteczniej dostosowywać głębokie modele do nowych zadań na podstawie niewielkich zbiorów danych. Praca [II] opisuje HyperShot – metodę *fewshot*, w której hipersieć oparta na jądrach (kernel hypernetwork) na podstawie zbioru wsparcia generuje parametry klasyfikatora dla konkretnego zadania bez użycia uczenia gradientowego, dzięki czemu model może skutecznie adaptować się do nowych, nawet skrajnie różnych zadań przy bardzo małej liczbie przykładów. Praca [III] rozwija HyperShot do ramowego podejścia *fewshot* z wykorzystaniem hipersieci opartych na jądrach, w którym hipersieć na podstawie reprezentacji zbioru

wsparcia w przestrzeni jąder generuje parametry klasyfikatora dla wielu zadań, a rozszerzenie BayesHyperShot umożliwia dodatkowo modelowanie niepewności predykcji.

Przechodząc do omówienia prac poświęconych uczeniu samonadzorowanemu i federacyjnemu, publikacja [IV] wprowadza CASSLE – metodę uczenia samonadzorowanego z warunkowanym modułem projekcji, w której klasyczne podejście łącznej reprezentacji w SSL zostaje rozszerzone o świadomość augmentacji: projektor otrzymuje informację o zastosowanych przekształceniach, przez co wymusza na ekstraktorze cech zachowanie w reprezentacji tych własności danych, które są modyfikowane przez augmentacje, co poprawia transfer na zadania zależne m.in. od koloru czy lokalnych struktur. W pracy [V] autorzy analizują, dlaczego reprezentacje Vision Transformerów trenowanych z wykorzystaniem Masked Image Modeling (MIM) mają niską jakość bez dodatkowego dostrajania. Pokazują, że problemem jest nieskuteczna agregacja informacji przez token [cls], a następnie proponują lekką metodę *selective aggregation*, która dynamicznie wybiera i scala istotne tokeny, znacząco poprawiając jakość reprezentacji MIM bez potrzeby dalszego dostrajania. Artykuł [VI] bada, jak głęboko można dzielić sieć między klientami a serwerem w *split federated selfsupervised learning*, pokazuje, że dotychczasowe metody (np. MocoSFL) tracą jakość przy minimalnej komunikacji, oraz wprowadza *MomentumAligned contrastive SFL* (MonAcoSFL), które utrzymuje dopasowanie (*alignment*) modeli klienckich (tj. sprawia, że uczą się one w zbliżonej przestrzeni cech i nie „ciągną" modelu globalnego w sprzeczne strony mimo różnych danych), dzięki czemu uzyskuje porównywalną dokładność przy znacznie mniejszym narzucie komunikacyjnym i lepszej ochronie danych.

2. Cel pracy i problem badawczy

Celem rozprawy jest opracowanie i zbadanie metod uczenia reprezentacji oraz metod szybkiej adaptacji do nowych zadań przy ograniczonej liczbie zaetykietowanych danych. Praca łączy rozwój metod *Fewshot Learning* opartych na zadaniowo warunkowanej generacji parametrów z analizą i udoskonalaniem nowoczesnych metod uczenia samonadzorowanego (SSL), tak aby lepiej rozumieć i wykorzystywać właściwości powstających reprezentacji w zadaniach praktycznych. Problem badawczy został sformułowany jasno i odpowiada aktualnym wyzwaniom badawczym, szczególnie w zastosowaniu uczenia maszynowego w takich dziedzinach, gdzie zaanotowanych danych jest niewiele, na przykład w wielu zastosowaniach medycznych. Został poprawnie powiązany z postawionymi pytaniami badawczymi.

3. Aktualność i znaczenie podejmowanego zagadnienia na tle literatury przedmiotu

Przedstawiony w rozprawie stan wiedzy dotyczący rozwiązywanych problemów jest aktualny i rzetelnie osadzony w literaturze dotyczącej głębokiego uczenia efektywnego względem małej liczby danych. Autor jasno identyfikuje kluczowy problem współczesnych modeli głębokich – silną zależność od dużych, kosztownych w anotacji zbiorów danych – oraz pokazuje, w jaki sposób paradygmaty *fewshot learning* i *selfsupervised learning* odpowiadają na te ograniczenia. Omówienie głównych tematów badawczych (metalearning, metody metryczne, hipersieciowe podejścia do adaptacji oraz nowoczesne podejścia SSL) oraz powiązanie ich z własnymi pracami świadczą o bardzo dobrym zrozumieniu stanu wiedzy, umiejętności jego krytycznej analizy oraz jasnym określeniu niszy badawczej rozprawy.

4. Metodyka badań

Zastosowana w cyklu publikacji metodyka badań jest spójna i adekwatna do postawionych celów naukowych. W pracach poświęconych HyperMAML i kernel HyperNetworks wykorzystano klasyczne protokoły metauczenia *fewshot* (zadania 1shot/5shot, standaryzowane benchmarki typu MiniImageNet, CUB, eksperymenty *crossdomain*), co zapewnia porównywalność wyników z literaturą oraz wiarygodną ocenę jakości proponowanych rozwiązań względem metod referencyjnych. Metodyka ma charakter twórczy – łączy dobrze ugruntowane modele z nowymi elementami, takimi jak hipersieć generująca aktualizacje wag zamiast

klasycznego gradientu w HyperMAML czy warunkowanie wprowadzone w projektor w CASSLE i systematycznie bada ich wpływ na skuteczność, stabilność oraz efektywność obliczeniową modeli.

Eksperymenty są rzetelnie zaprojektowane: obejmują szerokie spektrum ustawień (różne liczby przykładów na klasę, różne zbiory danych, scenariusze crossdomain, porównania z licznymi metodami *stateoftheart*), a protokoły treningowotestowe oraz dobrane miary jakości są jasno opisane, co umożliwia powtórzenie wyników i krytyczną ocenę wniosków. W pracach dotyczących uczenia samonadzorowanego uwzględniającego augmentacje i federacyjnego uczenia reprezentacji metodyka konsekwentnie łączy analizę jakości reprezentacji z testami w zadaniach docelowych – *downstream* (klasyfikacja, detekcja itp., bez wcześniejszego dostrajania), co pozwala ocenić nie tylko poprawność teoretyczną, lecz również praktyczną użyteczność proponowanych rozwiązań. Tym samym zastosowana metodyka badań spełnia wysokie standardy w dziedzinie głębokiego uczenia, a zarazem pokazuje umiejętność projektowania złożonych, wieloetapowych eksperymentów przez kandydata.

5. Wyniki i ich interpretacja

Wyniki przedstawione w omawianym cyklu publikacji są zaprezentowane w sposób przejrzysty i zrozumiały – autorzy stosują czytelne tabele i wykresy. Konsekwentnie opisują konfiguracje eksperymentów oraz jasno wskazują, które metody pełnią rolę odniesienia, co znacząco ułatwia interpretację uzyskanych rezultatów.

Poprawa jakości względem metod referencyjnych (m.in. klasycznych wariantów MAML oraz standardowych podejść uczenia samonadzorowanego) jest nie tylko raportowana liczbowo, lecz także omawiana w kontekście właściwości proponowanych modeli, takich jak możliwość generowania większych aktualizacji wag przez hipersieci czy zwiększona wrażliwość reprezentacji na istotne cechy danych.

We wszystkich pracach wyniki są systematycznie porównywane z metodami odniesienia – zarówno z wiodącymi podejściami *fewshot*, jak i popularnymi metodami uczenia samonadzorowanego – co pozwala wiarygodnie ocenić, na ile proponowane rozwiązania stanowią realny postęp względem istniejących metod odniesienia. Wnioski formułowane przez autorów wynikają bezpośrednio z zaprezentowanych danych i są poparte analizą zachowania modeli w różnych ustawieniach eksperymentalnych (różne zbiory, liczby przykładów).

Do mocnych stron wyników należy zaliczyć konsekwentnie lepsze wyniki względem metod referencyjnych na wielu benchmarkach. Jednocześnie autorzy świadomie sygnalizują ograniczenia proponowanych metod, takie jak złożoność modeli wykorzystujących hipersieci czy zależność od jakości dobranych augmentacji, co stanowi dobrą bazę dla dalszych badań i rozwijania przedstawionych koncepcji.

Wyniki są raportowane zgodnie z przyjętą praktyką w badaniach paradygmatów uczenia *fewshot* jako średnie z wielu epizodów/zadań wraz z odchyleniami standardowymi, co w pewnym stopniu pozwala ocenić stabilność metod oraz zróżnicowanie ich działania w różnych zadaniach. W pracach dotyczących uczenia samonadzorowanego podawane są jednak jedynie średnie wyniki, bez odchylenia standardowego, co nie daje możliwości oceny stabilności uzyskanych rezultatów. Formalne testy statystyczne (np. testy istotności porównujące metody) nie występują w analizie.

6. Oryginalność i wkład naukowy

Przedstawiony cykl publikacji stanowi oryginalne rozwiązanie problemu naukowego związanego z adaptacją modeli głębokich w warunkach *fewshot* oraz uczeniem reprezentacji z niewielkich, rozproszonych i słabo oznaczonych danych. Wkład naukowy obejmuje w szczególności opracowanie nowych algorytmów i modeli opartych na hipersieciach (HyperMAML, HyperShot oraz ogólny schemat kernel HyperNetworks), a także nowych metod uczenia samonadzorowanego (uczenie samonadzorowane oparte na augmentacjach) i federacyjnego uczenia reprezentacji, które umożliwiają bardziej efektywną i stabilną adaptację modeli do nowych zadań w praktycznych zastosowaniach. Cykl prac rozwija i pogłębia istniejące koncepcje uczenia w paradygmacie *fewshot*, hipersieci oraz uczenia samonadzorowanego i federacyjnego, proponując ich nowe ujęcia i integracje, poparte obszernym zestawem eksperymentów.

Publikacje jednoznacznie potwierdzają umiejętność samodzielnego prowadzenia badań naukowych kandydata. Obejmują pełny cykl prac – od zidentyfikowania istotnych luk w istniejących podejściach, poprzez projekt nowych modeli i dobór zestawów danych, po krytyczną interpretację wyników oraz wskazanie możliwych kierunków dalszego rozwoju metod. Spójność tematyczna, konsekwentne rozwijanie własnych badań oraz wysoka ranga czasopism i konferencji, w których opublikowano wyniki, dodatkowo wzmacniają ocenę zdolności kandydata do samodzielnego i twórczego uprawiania badań naukowych.

7. Znaczenie praktyczne pracy

Metody ujęte w rozprawie mają wyraźne znaczenie praktyczne. Dostarczają zaawansowanych, a zarazem stosunkowo uniwersalnych narzędzi do adaptacji modeli w ustawieniach *few-shot* i samonadzorowanego uczenia, które mogą być stosowane w rzeczywistych systemach, gdzie danych jest mało, są rozproszone lub zaetykietowane w niewielkiej liczbie. Przykładami takich zastosowań są systemy diagnostyczne w medycynie i przemyśle, systemy bezpieczeństwa oraz klasyfikacja danych tekstowych i multimedialnych.

8. Uwagi krytyczne i dyskusyjne

Wiele słabych stron proponowanych podejść jest już sygnalizowanych w publikacjach przez samych autorów. Ze swojej strony wskazałabym także:

- Brak oceny na trudnych, rzeczywistych danych. Dotychczasowe prace są oceniane głównie na klasycznych benchmarkach wizji komputerowej, a brakuje systematycznych eksperymentów w bardziej wymagających, praktycznych scenariuszach (np. dane medyczne, silny *domain shift*, niestandardowe dane kliniczne). Ogranicza to możliwość pełnej oceny użyteczności proponowanych metod w szerszym kontekście zastosowań.
- Ograniczone uzasadnienie teoretyczne i formalny opis działania metod. Mimo wprowadzenia wielu interesujących konstrukcji (hipersieci, kernel HyperNetworks, zadaniowo warunkowane SSL, selektywna agregacja w MIM, federacyjne SSL), analizy mają w przeważającej mierze charakter empiryczny. Brakuje pogłębionych wyników teoretycznych dotyczących stabilności, zbieżności i własności reprezentacji, które wzmocniłyby sformułowane wnioski oraz ułatwiły transfer metod do nowych dziedzin.
- Wysoka złożoność i koszty treningu przy ograniczonej analizie skalowalności. Metody oparte na hipersieciach (HyperMAML, HyperShot, kernel HyperNetworks) oraz federacyjne SSL (MonAcoSFL) są architektonicznie złożone, wrażliwe na inicjalizację i dobór hiperparametrów, a ich trening bywa bardzo kosztowny. Prace jedynie częściowo dyskutują, w jaki sposób te koszty i złożoność skalują się do większych modeli, obrazów 3D czy zastosowań przemysłowych, co utrudnia ocenę łatwości ich wykorzystania w praktyce.
- Brak analizy statystycznej wyników oraz pełnych wyników estymacji. W publikacjach brak jest analizy statystycznej wyników a w artykułach dotyczących samouczenia i uczenia federacyjnego nie ma nawet odchylenia standardowego rezultatów, co ogranicza możliwość oceny wiarygodności uzyskanych wyników.

Mam również uwagę dotyczącą wkładu Autora w poszczególne publikacje. Prace wchodzące w skład rozprawy mają charakter wieloautorski (od 5 do 8 współautorów). W artykułach opublikowanych w czasopismach wskazano wkład autorów, jednak w większości przypadków niemal wszyscy współautorzy uczestniczyli w każdym etapie powstawania publikacji. W publikacjach konferencyjnych takie informacje nie są podawane. Utrudnia to rzetelną ocenę indywidualnego wkładu Doktoranta w powstałe prace. Z punktu widzenia oceny dorobku naukowego interesujące byłoby doprecyzowanie, jaki był udział pozostałych współautorów w publikacjach, w których Autor figuruje obok – na przykład – siedmiu innych autorów.

9. Ocena spełnienia wymagań ustawowych

Na podstawie analizy Autoreferetu i sześciu publikacji składających się na rozprawę (HyperMAML, HyperShot, The General Framework for Few-shot Learning by Kernel HyperNetworks, Augmentation-aware Self-supervised Learning with Conditioned Projector, Beyond [CLS] oraz A Deep Cut into Split Federated Self-supervised Learning) stwierdzam, że:

- rozprawa prezentuje szeroką, aktualną i pogłębioną wiedzę teoretyczną Doktoranta w zakresie uczenia reprezentacji, metod *few-shot*, hipersieci, samouczenia oraz uczenia federacyjnego, co znajduje odzwierciedlenie w publikacjach w wysoko punktowanych czasopismach (m.in. Neurocomputing, Machine Vision and Applications, Knowledge-Based Systems) oraz w referatach na konferencjach najwyższej rangi (ICCV, ECML PKDD).
- rozprawa potwierdza umiejętność samodzielnego prowadzenia pracy naukowej, obejmując projektowanie oryginalnych architektur (hipersieci oparte na jądrach, HyperMAML, MonAcoSFL), przygotowanie zaawansowanych procedur eksperymentalnych oraz krytyczną analizę zachowania nowoczesnych modeli głębokich w licznych zadaniach komputerowego widzenia i uczenia federacyjnego.
- rozprawa stanowi oryginalne rozwiązania problemów naukowych w kilku podobszarach dyscypliny: wprowadza nowe uogólnienia metod typu MAML z wykorzystaniem hipersieci, ogólne ramy uczenia *few-shot* z hipersieciami opartymi na jądrach, metodę *augmentation-aware self-supervised learning* z warunkowanym projektorem oraz federacyjną metodę MonAcoSFL, zapewniającą korzystny kompromis między jakością reprezentacji, kosztem komunikacji a prywatnością danych, co łącznie stanowi istotny wkład w rozwój współczesnych metod uczenia głębokiego oraz ich potencjalnych zastosowań w praktyce.

10. Konkluzja

Podsumowując, stwierdzam, że rozprawa doktorska magistra Marcina Przewięźlikowskiego, zatytułowana *Data-Efficient Representation Learning and Adaptation*, spełnia wszystkie wymagania stawiane pracom doktorskim w dyscyplinie informatyka techniczna i telekomunikacja: prezentuje szeroką, aktualną wiedzę teoretyczną kandydata, jednoznacznie potwierdza jego umiejętność samodzielnego prowadzenia badań naukowych oraz zawiera szereg oryginalnych, wysoko ocenionych rozwiązań problemów badawczych. Tym samym rozprawa spełnia kryteria określone w art. 187 ustawy – Prawo o szkolnictwie wyższym i nauce. Wnoszę o dopuszczenie rozprawy do dalszych etapów przewodu doktorskiego.

Ze względu na wieloosobowy charakter publikacji składających się na rozprawę (do ośmiu współautorów) ocena indywidualnego wkładu Doktoranta w poszczególne prace jest utrudniona. Biorąc jednak pod uwagę całokształt jego działalności naukowej, rangę i spójność cyklu publikacji oraz rolę Autora w opracowaniu kluczowych metod, uważam, że rozprawa zasługuje na ocenę wyróżniającą.

Urszula Markowska-Kaczmar