

Kraków, 21 marca 2025r.



Marcin Kurdziel

Instytut Informatyki

Wydział Informatyki

Akademia Górniczo-Hutnicza im. St. Staszica w Krakowie

al. Mickiewicza 30, 30-059 Kraków

kurdziel@agh.edu.pl

Recenzja w postępowaniu w sprawie nadania stopnia doktora habilitowanego doktorowi Dawidowi Połapowi.

1 Sylwetka kandydata

Dr Dawid Połap posiada stopień naukowy doktora nauk inżynieryjno-technicznych nadany uchwałą Rady Wydziału Inżynierii Mechanicznej i Informatyki Politechniki Częstochowskiej z dnia 19 września 2019 r. Podstawą uzyskania stopnia była rozprawa doktorska zatytułowana: *Hybrydowa metoda ekstrakcji i klasyfikacji cech dla systemów człowiek-maszyna*. W okresie od 1 lutego 2017 r. do 31 października 2019 r. dr Połap pracował na stanowisku asystenta na Wydziale Matematyki Stosowanej Politechniki Śląskiej. Od dnia 1 listopada 2019 r. jest zatrudniony w tym Wydziale na stanowisku adiunkta.

W dniu 24 lipca 2024 r. dr Połap złożył wniosek o przeprowadzenie postępowania w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk inżynieryjno-technicznych, w dyscyplinie Informatyka Techniczna i Telekomunikacja. Kandydat nie ubiegał się uprzednio o nadanie stopnia doktora habilitowanego.

2 Ocena osiągnięcia naukowego

2.1 Kryteria oceny obowiązujące na dzień wszczęcia postępowania w sprawie nadania stopnia doktora habilitowanego

Przedłożone osiągnięcie naukowe jak również pozostałe przesłanki nadania stopnia doktora habilitowanego oceniałem zgodnie przepisami obowiązującymi na dzień wszczęcia postępowania habilitacyjnego, a w szczególności kryteriami wymienionymi w Art. 219 ustawy *Prawo o szkolnictwie wyższym i nauce* z dnia 20 lipca 2018 r., z późniejszymi zmianami.

2.2 Osiągnięcie naukowe

Osiągnięciem naukowym przedstawionym przez dra Dawida Połapa jest cykl powiązanych tematycznie artykułów naukowych, zatytułowany:

NOWE METODY ANALIZY OBRAZÓW DWUWYMIAROWYCH Z UŻYCIEM GŁĘBOKICH
KASKADOWYCH I WIELOATENCYJNYCH SIECI NEURONOWYCH

Na cykl składa się dziesięć artykułów w czasopismach naukowych i cztery artykuły w materiałach pokonferencyjnych:

- [A1] Dawid Połap, Natalia Wawrzyniak, Marta Włodarczyk-Sielicka. *Side-scan sonar analysis using ROI analysis and deep neural networks*. IEEE Transactions on Geoscience and Remote Sensing 60 (2022), strony 1-8.
- [A2] Dawid Połap, Gautam Srivastava. *Neural image reconstruction using a heuristic validation mechanism*. Neural Computing and Applications 33.17 (2021), strony 10787-10797.
- [A3] Dawid Połap. *Neuro-heuristic analysis of surveillance video in a centralized IoT system*. ISA transactions 140 (2023), strony 402-411.
- [A4] Dawid Połap, Marcin Woźniak. *Meta-heuristic as manager in federated learning approaches for image processing purposes*. Applied Soft Computing 113 (2021), strony 1-14.

- [A5] Dawid Połap, Marcin Woźniak. *A hybridization of distributed policy and heuristic augmentation for improving federated learning approach*. Neural Networks 146 (2022), strony 130- 140.
- [A6] Dawid Połap. *Automatic fuzzy parameter tuning for neural network models*, Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2022), strony 1-5.
- [A7] Dawid Połap, Marta Włodarczyk-Sielicka, Natalia Wawrzyniak. *Automatic ship classification for a riverside monitoring system using a cascade of artificial intelligence techniques including penalties and rewards*. ISA Transactions 121 (2022), strony 232-239.
- [A8] Dawid Połap. *An adaptive genetic algorithm as a supporting mechanism for microscopy image analysis in a cascade of convolution neural networks*. Applied Soft Computing 97 (2020), 1-11.
- [A9] Dawid Połap, Marcin Woźniak. *Acceleration of data handling in neural networks by using cascade classification model*, Proceedings of IEEE Symposium Series on Computational Intelligence (SSCI 2019), strony 917-923.
- [A10] Dawid Połap. *Hybrid image analysis model for hashtag recommendation through the use of deep learning methods*. Expert Systems with Applications 229 (2023), strony 1-9.
- [A11] Dawid Połap, Antoni Jaszcz. *Sonar digital twin layer via multi-attention networks with feature transfer*. IEEE Transactions on Geoscience and Remote Sensing (2024).
- [A12] Dawid Połap, Antoni Jaszcz, Katarzyna Prokop, Gautam Srivastava. *Lightweight CNN based on Spatial Features for a Vehicular Damage Detection System*. Proceedings of IEEE International Conference on Big Data (IEEE BigData 2023), strony 5097-5101.
- [A13] Antoni Jaszcz, Dawid Połap. *Semantic Segmentation for Moon Rock Recognition Using U-Net with Pyramid-Pooling-Based SE Attention Blocks*. Proceedings of 16th International Conference on Agents and Artificial Intelligence (ICAART 2024), strony 965-971.
- [A14] Dawid Połap, Antoni Jaszcz. *Decentralized medical image classification system using dual-input CNN enhanced by spatial attention and heuristic support*. Expert Systems with Applications (2024), strony 1-12.

Ideą przewodnią przedstawionego osiągnięcia jest rozwój algorytmów klasyfikacji obrazów wykorzystujących sieci neuronowe. Habilitant badał w szczególności hybrydy sieci neuronowych i heurystycznych algorytmów optymalizacyjnych, modele zbudowane w formie kaskad klasyfikatorów oraz nowe warstwy atencyjne w sieciach konwolucyjnych. Oceniając przedłożone osiągnięcie naukowe odstąpię od kolejności zaproponowanej w autoreferacie, w pierwszej kolejności prezentując wyniki, które, w mojej ocenie, stanowią jego najmocniejszy element.

Wielogałęziowe sieci konwolucyjne i nowe moduły atencyjne

W pracach [A10]–[A14] Habilitant badał architektury konwolucyjne z więcej niż jednym wejściem lub wyjściem oraz nowe moduły atencyjne dla warstw konwolucyjnych.

Praca [A10] przedstawia koncepcję wielogałęziowej sieci konwolucyjnej służącej do przypisywania obrazom tzw. hashtagów. Sieć ta przyjmuje jako wejście dwa obrazy utworzone w procesie segmentacji oryginalnego przykładu wejściowego: obraz z pierwszoplanowym obiektem oraz obraz z usuniętym pierwszoplanowym obiektem. Obrazy te są przetwarzane przez równoległe gałęzie konwolucyjne rozpoznające wybrane klasy, np. gałąź rozpoznającą gatunek zwierzęcia i gałąź rozpoznającą jego otoczenie. Etykiety klas są następnie wiązane z zestawem hashtagów (za pomocą osadzeń wyrazów skonstruowanym algorytmem Skip-gram). Autorzy argumentują, że wykorzystanie sieci wielogałęziowej umożliwi rozszerzenie rozpoznawanego zestawu klas bez konieczności trenowania całego modelu – wystarczy wytrenować kolejną gałąź. Propozycja ta wychodzi naprzeciw istotnemu problemowi w praktycznych zastosowaniach sieci neuronowych, jakim jest wysoki koszt obliczeniowy związany z ich trenowaniem. Architektura wielogałęziowa może, potencjalnie, ograniczyć koszt douczania sieci, zapewniając zarazem wysoką skuteczność wytrenowanego modelu. W pracy nie dostrzegłem jednak porównania do naturalnego punktu odniesienia, jakim byłoby dodanie dodatkowego neuronu (klasy) w warstwie klasyfikacyjnej sieci z rysunku 3 i dotrenowanie wyłącznie jego wag (przy zamrożeniu pozostałych parametrów modelu). Warto również zwrócić uwagę, że do trenowania modelu wykorzystano maski wygenerowane za pomocą klasycznych algorytmów przetwarzania obrazów. Zaproponowany system mógłby potencjalnie uzyskać lepsze wyniki, gdyby do konstrukcji masek wykorzystano pre-trenowany model segmentacji obrazów.

W pracy [A11] podjęto problem trenowania sieci neuronowej w sytuacji, gdy liczba dostępnych

przykładów uczących jest stosunkowo niewielka. Dotyczyło to w szczególności klasyfikacji obiektów na obrazach z sonaru. Klasycznym rozwiązaniem problemu uczenia sieci przy ograniczonej liczbie etykietowanych danych jest wykorzystanie mechanizmu transferu wiedzy. Wymaga to jednak, by obrazy z nowej domeny były do pewnego stopnia zbliżone do obrazów wykorzystanych w trenowaniu modelu bazowego. W [A11] zaproponowano więc rozwiązanie alternatywne. Polega ono na wytrenowaniu modelu generatywnego (sieci GAN) na dostępnych przykładach uczących, a następnie przeniesienie części jego wag do sieci klasyfikującej. W praktyce pomysł ten został zrealizowany poprzez zaproponowanie nowego modułu atencyjnego, wykorzystywanego w generatorze sieci GAN. Wagi tego modułu są następnie przenoszone do analogicznych modułów atencyjnych w sieci klasyfikującej. Ułatwia to trenowanie sieci klasyfikującej i poprawia jej metryki skuteczności. Pomysł adaptacji części wag z sieci generatywnej do powiązanej sieci klasyfikującej uważam za bardzo ciekawy. Jak pokazano w [A11], jest on przydatny w sytuacji, gdy dla przetwarzanych danych nie ma dostępnych dobrych modeli pre-trenowanych.

Moduły atencyjne były również przedmiotem badań w pracach [A12] i [A13]. W [A12] warstwę uwagi zbudowano wykorzystując konwolucje z filtrem 1×1 . Podobne warstwy konwolucyjne są szeroko wykorzystywane w istniejących architekturach neuronowych, aczkolwiek z reguły nie jako multiplikatywne warstwy służące zwiększeniu wagi pewnych obszarów zdjęcia. Architektura zaproponowana w [A12] została wykorzystana do klasyfikacji uszkodzeń na zdjęciach pojazdów mechanicznych, uzyskując dokładność lepszą od kilku większych modeli neuronowych. W [A13] moduł atencyjny zbudowano na bazie tzw. bloków *Squeeze-and-Excitation*, które są wykorzystywane w niektórych architekturach konwolucyjnych. W [A13] bloki te zmodyfikowano poprzez zastąpienie klasycznej warstwy zbierającej (ang. *pooling layer*) piramidową warstwą zbierającą (ang. *pyramid pooling layer*) oraz zastąpienie warstw w pełni połączonych konwolucjami z filtrem 1×1 . Zbudowane w ten sposób warstwy atencyjne użyto do konstrukcji sieci segmentującej wzorowanej na architekturze U-Net. Sieć tą przetestowano na publicznie dostępnym zbiorze symulowanych zdjęć z powierzchni księżyca.

Praca [A14] ponownie podejmuje temat sieci wielogałęziowej. Zaproponowano w niej mianowicie sieć konwolucyjną, która w jednej gałęzi przyjmuje obraz oryginalny, zaś w drugiej obraz uproszczony, wygenerowany heurystyczną techniką superpikseli. Cechy wykryte przez obie gałęzi sieci są scalane i przetwarzane w końcowych warstwach klasyfikujących. W modelu

tym również wykorzystano moduły atencyjne zbudowane z konwolucji z filtrem 1×1 . Zaproponowaną architekturę wykorzystano do klasyfikacji zbioru zdjęć medycznych, uzyskując wyniki lepsze od kilku opublikowanych modeli referencyjnych.

Pewnym niedostatkiem prac [A12]–[A14] jest, w mojej ocenie, brak wyczerpującej analizy ablacji (ang. *ablation study*). Analiza taka pozwala ocenić jaką rolę w końcowej skuteczności modelu uczącego się pełnią jego poszczególne elementy. W kontekście ocenianych prac, analiza ablacji pozwoliłaby zweryfikować zasadność wszystkich proponowanych modyfikacji architektury sieci. Prace przedstawiają natomiast głównie końcową skuteczność modelu (niekiedy w wielu konfiguracjach). I choć uzyskane rezultaty są lepsze od opublikowanych wyników referencyjnych, to sam końcowy wynik nie daje przesłanek by sądzić, iż każda zaproponowana modyfikacja architektury sieci jest w istocie niezbędna do osiągnięcia tego wyniku.

Kaskady klasyfikatorów

W pracach [A7]–[A9] badano tzw. *kaskady klasyfikatorów*. W pracy [A7] kaskada zbudowana jest jako kolejka priorytetowa grupująca kilka odmiennych klasyfikatorów rozwiązujących ten sam problem klasyfikacyjny. W pracy wykorzystano w szczególności sieci w pełni połączone, sieci konwolucyjne oraz klasyfikator k-najbliższych sąsiadów operujący w przestrzeni klasycznych cech obrazu. W zaproponowanym algorytmie w pierwszej kolejności ewaluowany jest klasyfikator o największym priorytecie. Jeśli zwróci on decyzję z wysokim oszacowanym prawdopodobieństwem wybranej klasy, pozostałe klasyfikatory nie są ewaluowane. W przeciwnym razie algorytm ewaluuje każdy klasyfikator w kolejce i zwraca klasę z największym oszacowanym prawdopodobieństwem. Po wykonaniu procedury klasyfikacji, zaproponowany algorytm aktualizuje priorytety w kolejce, biorąc pod uwagę oszacowane prawdopodobieństwa klas. Tak skonstruowana kaskada klasyfikatorów stanowi pewną formę uczenia zespołowego, łączącą kilka odmiennych klasyfikatorów obrazu. Mechanizm priorytetów pozwala zarazem uniknąć ewaluacji całej kaskady, w przypadku, gdy klasyfikator z największym priorytetem zwróci wysokie oszacowanie prawdopodobieństwa wybranej klasy. Zaproponowany algorytm wykorzystano w problemie klasyfikacji statków na zdjęciach, uzyskując wyniki wskazujące potencjalną przydatność mechanizmu kaskady klasyfikatorów.

Podobny schemat klasyfikacji zaproponowano również w pracy [A8], tym razem wykorzystując

jednak kaskadę klasyfikatorów binarnych. W szczególności, każdy z elementów kaskady wytrenowany został do rozpoznawania jednej, przypisanej mu klasy (w odróżnieniu od pozostałych klas). Zrezygnowano również z mechanizmu priorytetów w kaskadzie, zastępując je funkcją dopasowania. Funkcja ta wybiera klasyfikatory mające zespołowo podjąć decyzję o rozpoznanej klasie. Parametry funkcji dopasowania zoptymalizowano algorytmem genetycznym. Jako bazowe klasyfikatory w kaskadzie wykorzystano trzy architektury sieci konwolucyjnych. Zaproponowany algorytm zastosowano do klasyfikacji kształtów komórek na zdjęciach mikroskopowych.

Kaskada klasyfikatorów binarnych była również przedmiotem badań w pracy [A9]. Podobnie jak w pracy [A7], wykorzystano tu mechanizm priorytetów. W szczególności, klasyfikatory były ewaluowane w kolejności oszacowanej dokładności predykcji. Kaskadę zbudowano w dwóch wariantach: z sieciami w pełni połączonymi i z sieciami konwolucyjnymi. Oba warianty zastosowano do klasyfikacji zdjęć zwierząt domowych. Uzyskane wyniki sugerują, że kaskada poprawia dokładność klasyfikacji względem pojedynczej sieci neuronowej.

Pomysły przedstawione w pracach [A7]–[A9] można uznać za pewną formę uczenia zespołowego. W obu tych podejściach końcowa predykcja konstruowana jest poprzez ewaluację wielu klasyfikatorów i uzgodnienie ich decyzji. Zwiększa to dokładność predykcji, kosztem większego nakładu obliczeniowego. W mojej ocenie, algorytmy przedstawione w [A7]–[A9] warto by więc porównać nie tylko z pojedynczymi klasyfikatorami bazowymi, lecz również z innymi metodami uczenia zespołowego o podobnym koszcie obliczeniowym.

W pracy [A7] zasadnym byłoby również szersze przedyskutowanie sposobu, w jaki kaskada wykorzystuje prawdopodobieństwa oszacowane przez sieci neuronowe. W szczególności, klasycznie wytrenowana sieć neuronowa nie musi być modelem skalibrowanym – zwracane przez nią prawdopodobieństwa nie muszą korelować z faktycznym oszacowaniem niepewności predykcji. Co istotne, problem ten wydaje się dotyczyć między innymi sieci konwolucyjnych.¹

Hybrydy sieci neuronowych i heurystycznych algorytmów optymalizacyjnych

W pracach [A1]–[A6] Habilitant badał metody hybrydowe, łączące sieci neuronowe i

¹M. Minderer, i in. (2021). *Revisiting the calibration of modern neural networks*. *Advances in neural information processing systems*, 34, 15682-15694.

heurystyczne algorytmy optymalizacyjne.

Praca [A1] podejmuje zagadnienie wykrywania obiektów na zdjęciach o bardzo dużej rozdzielczości. Choć zdjęcia takie można przetwarzać sieciami w pełni konwolucyjnymi (które nie ograniczają wejścia do określonego rozmiaru), to jednak wiąże się to z dużymi wymaganiami odnośnie pamięci wykorzystanych układów obliczeniowych. W [A1] zaproponowano algorytm, który dzieli obraz wejściowy na mniejsze fragmenty, a następnie decyduje (wykorzystując analizę histogramów), które fragmenty stanowią region zainteresowania (ang. *region of interest*, ROI). Tak skonstruowany ROI jest następnie klasyfikowany siecią konwolucyjną uczoną techniką transferu wiedzy. W [A1] algorytm ten wykorzystano dla analizy obrazów z sonaru, uzyskując wyniki lepsze od bezpośredniej klasyfikacji siecią konwolucyjną. Co warto podkreślić, w ramach ewaluacji wykorzystano przygotowany w tym celu zestaw zdjęć sonarowych obrazujących dno rzeki w Szczecinie.

W pracy [A2] przedstawiono algorytm łączący heurystyki optymalizacyjne inspirowane naturą z sieciami w architekturze auto-koderów. W szczególności, heurystyki optymalizacyjne wykorzystano do wskazania istotnych punktów na wejściowych obrazach. W tym celu zaproponowano dość złożoną funkcję celu dla algorytmu optymalizującego, wykorzystującą między innymi filtrowanie obrazu. Punkty wskazane przez algorytm heurystyczny stanowią wejście do funkcji kosztu auto-kodera. W odróżnieniu od klasycznego błędu kwadratowego, funkcja kosztu zaproponowana w [A2] dopasowuje entropię rekonstrukcji obrazu w punktach wskazanych przez heurystykę optymalizacyjną. W pracy wykonano obszerną eksperymentalną ewaluację zaproponowanego algorytmu hybrydowego. Uwzględniono w niej 6 różnych heurystyk inspirowanych naturą i 3 architektury auto-koderów. Sposób wykonania eksperymentów sugeruje, że istotną motywacją badań była chęć zmniejszenia kosztu obliczeniowego w stosunku do zwykłych auto-koderów. Celu tego nie udało się w pełni zrealizować – o ile zaproponowane rozwiązanie okazało się zmniejszać sam czas trenowania sieci, to zysk ten był w dużej mierze negowany kosztem obliczeniowym heurystyki optymalizacyjnej.

W pracy [A3] heurystyki inspirowane naturą wykorzystano jako element algorytmu przetwarzającego strumień wideo z kamer monitoringu. W szczególności, heurystykę optymalizacyjną wykorzystano do selekcji punktów kluczowych na ramkach ze strumienia

wideo. Zmiana intensywności pikseli w tych punktach interpretowana była jako zmiana zawartości kadru, co skutkowało wysłaniem ramki do dalszej analizy. Ramki wybrane w ten sposób były klasyfikowane za pomocą sieci konwolucyjnych wytrenowanych do wykrywania sytuacji niebezpiecznych. Co warto podkreślić, sieci te trenowane były w schemacie uczenia federacyjnego, w którym grupa klientów uśrednia prywatnie wytrenowane modele. Pozwala to wytrenować sieć na danych pochodzących z wielu strumieni wideo, zachowując zarazem prywatność tych danych. W pracy wykonano ewaluację kilku heurystyk optymalizacyjnych i architektur sieci konwolucyjnych. Zbadano również wyniki uczenia z transferem wiedzy. Ostatecznie osiągnięto rezultaty lepsze od kilku opublikowanych wcześniej prac korzystających z tego samego zbioru danych. Autorzy zaznaczają jednak, że modele referencyjne analizują strumienie wideo, podczas gdy w omawianym artykule klasyfikowano wybrane ramki. W mojej ocenie ciekawym punktem odniesienia, którego zabrakło w [A3], byłoby również porównanie wyników z wariantem algorytmu, w którym heurystyka wykrywająca punkty kluczowe została zastąpiona szybkim algorytmem detekcji obiektów, np. YOLO (zmiana zestawu wykrytych obiektów byłaby wówczas przesłanką do dalszej analizy ramki). Porównanie rezultatów uzyskanych w ten sposób z wynikami zaprezentowanymi w [A3] pozwoliłoby odpowiedzieć na pytanie, czy heurystyczny algorytm wyboru punktów kluczowych faktycznie wpływa w sposób decydujący na skuteczność zaproponowanego rozwiązania.

W [A4] zaproponowano wykorzystanie heurystyk inspirowanych naturą w algorytmie uczenia federacyjnego. Scenariusz rozważany w tej pracy zakłada, że na niektórych węzłach obliczeniowych dane uczące mogły zostać zatrute, tj. wprowadzono do nich przykłady z błędnie przypisanymi klasami. Hybryda zaproponowana w [A4] wykorzystuje algorytmy heurystyczne na kilku etapach: do lokalnej optymalizacji modeli przesłanych przez węzły obliczeniowe, do detekcji przykładów z błędnymi etykietami oraz do wykrywania węzłów obliczeniowych, w których znaczna część danych uczących jest zatruta. Rezultaty uzyskane w przeprowadzonych eksperymentach pokazują, że zaproponowany schemat uczenia federacyjnego poprawia wyniki względem rezultatów opublikowanych w literaturze. Wykonano również eksperymenty z symulowanym zatruciem zbioru uczącego, które pokazały odporność zaproponowanej metody. Warto jednak podkreślić, że w pracy zbadano tylko jedną formę ataku zatrutymi danymi – atak poprzez modyfikację etykiet przykładów uczących. W praktyce możliwe są również inne formy ataku, np. zatrucie modelu użytego do transferu wiedzy ([A4]

bada uczenie w mechanizmie transferu wiedzy), czy też modyfikacja cech przykładów uczących. W mojej ocenie cenne byłoby również szersze przedyskutowanie zaproponowanego kryterium detekcji zmanipulowanych etykiet. Bazuje ono na śledzeniu funkcji kosztu w kolejnych cyklach uczenia federacyjnego. Zmiana kosztu przekraczająca zmianę średnią traktowana jest jako detekcja zmanipulowanej etykiety. Na pierwszy rzut oka, tak postawione kryterium (zmiana powyżej średniej) może wyłączyć z uczenia znaczną część przykładów trenujących, niezależnie od faktycznego poziomu zatrucia zbioru uczącego.

W pracy [A5] heurystyka inspirowana naturą została użyta jako element algorytmu augmentacji danych uczących. Zasadnicza idea zaproponowanej metody polega na wykorzystaniu heurystyki optymalizacyjnej do znalezienia zestawu kluczowych obszarów na obrazie z wybranej klasy. Obszary te są następnie przenoszone do innego obrazu z tej samej klasy i potencjalnie poddawane dalszym przekształceniom klasycznymi metodami przetwarzania obrazu. Utworzony w ten sposób nowy przykład uczący jest weryfikowany i dodawany do zbioru uczącego. Wykonane eksperymenty pokazują, że zaproponowana metoda augmentacji danych daje istotnie lepsze wyniki w uczeniu sieci neuronowych, niż klasyczne algorytmu augmentacji (takie jak losowe skalowanie, modyfikacja jasności czy dodanie szumu). Pomysł tworzenia nowych przykładów poprzez kombinację istniejących elementów zbioru danych jest bardzo ciekawy. Warto jednak zwrócić uwagę, że istnieją już wcześniejsze prace proponujące algorytmy tego typu. Szczególnie ważny jest tu algorytm *MixUp* zaproponowany w 2018r.² Algorytm ten tworzy nowe przykłady uczące przez uśrednianie przykładów istniejących. Ten prosty, a zarazem bardzo wydajny algorytm okazuje się być zaskakująco skuteczną metodą augmentacji danych. W mojej ocenie praca [A5] znacząco zyskałaby, gdyby *MixUp* uwzględnić w zestawie metod referencyjnych.

Tematyka pracy [A6] odbiega od przedstawionych powyżej metod hybrydowych wykorzystujących algorytmy inspirowane naturą. Zaproponowano w niej technikę optymalizacji współczynnika uczenia (ang. *learning rate*) za pomocą wnioskowania rozmytego i klasycznej metody symulowanego wyżarzania. Praca przedstawia również pewne rezultaty w zakresie wpływu współczynnika uczenia na dokładność wytrenowanego modelu. Badania ograniczono do jednego zbioru danych (CIFAR-10). [A6] pomija niestety kluczowy aspekt uczenia

²H. Zhang i in. (2018). *MixUp: Beyond Empirical Risk Minimization*. W: 2018 International Conference on Learning Representations.

współczesnych modeli ML, jakim jest zmiana wartości współczynnika uczenia w trakcie procesu trenowania (ang. *learning rate schedule*). Optymalizacja pojedynczej wartości współczynnika uczenia nie ma istotnego przełożenia na współczesne metody trenowania sieci neuronowych – w tym kontekście należało by optymalizować funkcję opisującą jak zmienia się współczynnik uczenia w kolejnych epokach. W pracy nie dostrzegłem również porównania do innych metod adaptacji współczynnika uczenia.

2.3 Ocena osiągnięcia

Przedstawione osiągnięcie naukowe obejmuje wiele technik i algorytmów mających na celu usprawnienie modeli neuronowych do przetwarzania obrazów. Wyniki te można pogrupować w trzy wątki badawcze: modyfikacje architektur sieci neuronowych, kaskady klasyfikatorów neuronowych oraz algorytmy hybrydowe łączące sieci neuronowe z heurystykami inspirowanymi naturą. Najciekawsze wyniki uzyskano, w mojej ocenie, w pracach poświęconych nowym modułom atencyjnym oraz sieciom wielogałęziowym. Szczególnie interesującym jest tu pomysł transferu wiedzy z modelu generatywnego poprzez przeniesienie wag modułów atencyjnych. Ciekawe są również propozycje modyfikacji nowszych architektur sieci konwolucyjnych (takich jak sieci z blokami *Squeeze-and-Excitation*) oraz architektury wielogałęziowe integrujące różne reprezentacje obrazu. Badania poświęcone kaskadom klasyfikatorów mogą potencjalnie prowadzić do efektywnych obliczeniowo modeli zespołowych. Taką rolę mogą w szczególności pełnić kaskady klasyfikatorów ewaluowane według priorytetu zależnego od pewności klasyfikacji. W mojej ocenie to również jest ciekawy wątek badawczy w przedłożonym osiągnięciu naukowym. Ostatni wątek badawczy, dotyczący metod hybrydowych, skupia się na kilku różnych zagadnieniach związanych z przetwarzaniem obrazów sieciami neuronowymi. Najciekawszą propozycją jest tu, w mojej ocenie, algorytm augmentacji danych tworzący nowe przykłady uczące poprzez kombinacje przykładów już istniejących. Interesujące są również badania nad zastosowaniem algorytmów hybrydowych do detekcji zatrutych danych.

Omawiając osiągnięcie habilitacyjne wskazałem również kilka aspektów, które, w mojej ocenie, warto by szerzej przedyskutować w przedłożonym cyklu prac. W szczególności, w pracach poświęconych modyfikacjom architektur sieci neuronowych eksperymenty warto by rozszerzyć o szczegółową analizę ablacji. W niektórych pracach warto by również uwzględnić dodatkowe

metody referencyjne, np. *MixUp* w [A5], detekcję kluczowych klatek za pomocą YOLO w [A3], czy też algorytmy uczenia zespołowego w wątku poświęconym kaskadom klasyfikatorów. W kontekście proponowanych kaskad klasyfikatorów warto by również szerzej przedyskutować kwestię kalibracji modelu, która może wpływać na wiarygodność oszacowanych prawdopodobieństw.

Cykl prac stanowiący osiągnięcie naukowe został opublikowany w większości w czasopismach o wysokich współczynnikach wpływu i górnych progach punktacji na wykazie czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych. Cztery prace opublikowano na konferencjach ujętych w tym wykazie, w tym jednej wysoko punktowanej. Kierując się współczynnikiem wpływu, za wiodące spośród periodyków w których opublikowano osiągnięcie naukowe można uznać czasopisma *Expert Systems with Applications* i *Applied Soft Computing*. W cyklu nie ma publikacji na najważniejszych konferencjach informatycznych dotyczących widzenia maszynowego (CVPR, ICCV, ECCV) lub sieci neuronowych i uczenia maszynowego (NeurIPS, ICLR, ICML). Zgodnie z bazą Scopus cykl prac cytowany był łączne 311 razy. Habilitant jest pierwszym autorem w 13 pracach cyklu. W autoreferacie precyzyjnie wskazano wkład Habilitanta do każdej pracy wieloautorskiej. W mojej ocenie jest to wkład kluczowy.

Podsumowując, oceniam, że przedłożone osiągnięcie naukowe stanowi znaczny wkład w dyscyplinę naukową Informatyka Techniczna i Telekomunikacja.

3 Istotna aktywności naukowa Kandydata i uzyskane wskaźniki naukometryczne

Doktor Połap odbył trzy starze naukowe w Uniwersytecie w Katanii, we Włoszech (okresy: marzec–wrzesień 2017, maj–wrzesień 2019, 16.07.2021–14.08.2021). W mojej opinii spełniają one warunek istotnej aktywności naukowej realizowanej w więcej niż jednej jednostce, sformułowany w Art. 219 pkt. 1 par. 3) ustawy *Prawo o szkolnictwie wyższym i nauce* z dnia 20 lipca 2018 r.

Na dzień złożenia wniosku habilitacyjnego, dr Połap był współautorem 79 (37)³ publikacji w czasopiśmie naukowych, 27 (16) wystąpień konferencyjnych oraz jednego rozdziału w monografii naukowej. Sumaryczny współczynnik wpływu tych prac to 306,383 zaś sumaryczna punktacja zgodnie z wykazem MNiSW to 8863 (6520). Zgodnie z bazą Scopus, w dniu złożenia wniosku habilitacyjnego prace dra Połapa były cytowane 3274 razy (w tym 2905 razy bez autocytowań), dając indeks Hirscha równy 32.

4 Osiągnięcia dydaktyczne, organizacyjne i popularyzujące naukę

Doktor Połap ma szerokie doświadczenie dydaktyczne. W latach 2020-2024 r. prowadził 7 wykładów oraz zajęcia laboratoryjne w 6 przedmiotach. Był promotorem prac magisterskich, prac licencjackich oraz projektów inżynierskich. Był również opiekunem trzech studentów w ramach prowadzonego na Politechnice Śląskiej programu *Rozwiń Skrzydła*.

Doktor Połap posiada również osiągnięcia w zakresie popularyzacji nauki. Był między innymi członkiem zespołu organizującego ogólnopolski konkurs *Algorytmion*. Przygotowywał również zadania na konkurs organizowany przez *Stowarzyszenie Studentów BEST Gliwice*. Warto również wspomnieć, że dr Połap pełnił funkcję przewodniczącego sesji w 5 konferencjach międzynarodowych. Ponadto brał udział w komitetach redakcyjnych 8 czasopism i komitetach programowych 4 konferencji. Wykonał również recenzje ponad setki artykułów nadesłanych do międzynarodowych czasopism naukowych. Dr Połap był również wykonawcą w trzech projektach badawczo-rozwojowych finansowanych ze źródeł zewnętrznych.

5 Konkluzja

W mojej ocenie przedstawione osiągnięcia naukowe oraz pozostała aktywność dra Połapa spełniają warunki uzyskania stopnia doktora habilitowanego zawarte w Art. 219 Ustawy *Prawo o szkolnictwie wyższym i nauce*. Wnoszę o dopuszczenie dra Połapa do dalszych etapów postępowania w sprawie nadania stopnia doktora habilitowanego w dziedzinie nauk

³W nawiasach podano wskaźniki naukometryczne po uzyskaniu stopnia doktora.

inżynieryjno-technicznych, w dyscyplinie Informatyka Techniczna i Telekomunikacja.

Z poważaniem,
Marcin Kurdziel