

Recenzja dorobku naukowego, dydaktycznego i organizacyjnego w przewodzie habilitacyjnym dr inż. Tomasza Trzcíńskiego

Recenzja dotyczy oceny jednotematycznego cyklu publikacji pt. „Metody estymacji subiektywnych atrybutów charakteryzujących materiały wizualne z wykorzystaniem głębokich sieci neuronowych”, jak również dorobku naukowego, dydaktycznego i organizacyjnego w przewodzie habilitacyjnym dr inż. Tomasza Trzcíńskiego.

Niniejsza recenzja została przygotowana na zlecenie Przewodniczącego Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Warszawskiej, w piśmie z dnia 10 lutego 2020 roku.

Podstawę do opracowania recenzji stanowi „Wniosek o przeprowadzenie postępowania w sprawie nadania stopnia doktora habilitowanego”, autorstwa dr inż. Tomasza Trzcíńskiego. Wniosek zawiera komplet materiałów umożliwiających jego recenzję, m.in. autoreferat, informację o aktywności naukowej, dydaktycznej i organizacyjnej Habilitanta oraz o jego współpracy z otoczeniem społecznym i gospodarczym. Dołączono do niego również wydruki publikacji stanowiących cykl jednotematyczny, oświadczenia o wielkości udziału w pracach naukowych i płytę CD zawierającą wersje elektroniczne wymienionych wyżej dokumentów.

Dr inż. Tomasz Trzcíński doświadczenia zawodowe i naukowe zdobywał na trzech uczelniach we Włoszech, Hiszpanii i Szwajcarii. Studiował na Politechnice Turyńskiej, gdzie uzyskał tytuł zawodowy inżyniera w zakresie Telekomunikacji (2008 r.) i tytuł magistra inżyniera w zakresie Inżynierii Elektronicznej (2011 r.). W 2010 roku tytuł magistra otrzymał również na Politechnice Katalońskiej, w zakresie Badań nad Technologiami Informacyjnymi i Komunikacją. Wynikiem badań doktorskich w Federalnej Politechnice w Lozannie jest stopień naukowy doktora w zakresie Informatyki (2014 r.). Temat rozprawy doktorskiej pt. „Uczenie i dopasowanie binarnych lokalnych deskryptorów cech” jest z zakresu ucznia maszynowego, które do tej pory znajduje się w polu zainteresowań Habilitanta. W trakcie pracy nad doktoratem był on zatrudniony na stanowisku asystenta w Federalnej Politechnice w Lozannie. Od 2015 r. pracuje na stanowisku adiunkta w Zakładzie Grafiki Komputerowej Instytutu Informatyki Politechniki Warszawskiej.

1. Ocena dorobku naukowego dr inż. Tomasza Trzcíńskiego

1.1. Ocena osiągnięcia naukowego

Przedstawiony do oceny jednotematyczny cykl publikacji „Metody estymacji subiektywnych atrybutów charakteryzujących materiały wizualne z wykorzystaniem głębokich sieci neuronowych” składa się z 8 prac naukowych (indeksowanych od P1 do P8) opublikowanych w latach 2017-2018, w tym 5 prac należy do Journal Citation Report i posiada

sumaryczny Impact Factor równy 14,581; wg wkładu Habilitanta 7,315. Pozostałe 3 prace są recenzowanymi materiałami konferencyjnymi indeksowanymi w Web of Science.

Do oceny przedstawiono również publikację konferencyjną (W1) z 2017 roku, opisującą realizowane przez Habilitanta osiągnięcie konstrukcyjne będące wdrożeniem metod opisanych w jednotematycznym cyklu publikacji.

We wszystkich wymienionych wyżej publikacjach dr inż. Tomasz Trzciniński jest współautorem.

Merytoryczna ocena osiągnięcia naukowego

Główna dziedzina pracy badawczej Habilitanta, przedstawiona również w jednotematycznym cyklu publikacji, to uczenie maszynowe, w tym wykorzystujące głębokie sieci neuronowe. Obecnie algorytmy uczenia maszynowego i głębokiego uczenia są powszechnie używane, począwszy od prostych botów na stronach internetowych, do zarządzania lotem i śledzenia celów w nowoczesnych myśliwcach. Wykorzystuje się je w systemach analityki biznesowej i systemach bezpieczeństwa IT do automatyzacji podejmowania decyzji, wskazywania właściwych działań, wykrycia zagrożeń typu zero-day. W marketingu stosowane są do personalizacji treści reklamowych, śledzenia zachowań konsumenckich, proponowania produktów danemu klientowi, prognozowania popularności reklamy online.

Prawdziwym wyzwaniem jest tworzenie reklamy bez słów, więc materiały marketingowe są zwykle multimodalne. Głębokie sieci neuronowe są szczególnie przydatne gdy trzeba analizować duże zbiory złożonych wielomodalnych danych, na przykład reklama banerowa i wideo w sieciach społecznościowych, czy dane produktów w systemach e-commerce dużych sieci handlowych, zawierające obraz/obrazy i tekst. Ich stosowanie generuje dwa istotne problemy. Po pierwsze materiały wizualne posiadają trudne do zmierzenia atrybuty subiektywne wpływające na ocenę tych materiałów przez ludzi. Po drugie głębokie sieci neuronowe wymagają znacznej mocy obliczeniowej, w związku z czym pożądane są szybkie algorytmy, w tym do obliczeń równoległych na jednostkach graficznych komputerów (tzn. GPU – Graphics Processing Unit).

W dobie Internetu i ciągle rosnących zbiorów danych wizualnych istotne znaczenie, zarówno z naukowego jak i aplikacyjnego punktu widzenia, ma automatyzacja analizy tych danych. W badania te wpisują się prace Habilitanta nad automatyczną estymacją subiektywnych atrybutów charakteryzujących materiały wizualne. Koncentrują się one na:

- przewidywaniu popularności filmów opublikowanych w sieciach społecznościowych,
- estymacji podobieństwa stylistycznego materiałów opisujących elementy wystroju wnętrza oraz ubrań,
- klasyfikacji emocji na podstawie obrazów twarzy.

Metody przewidywania popularności są szczególnie ważne dla marketingowych domów medialnych publikujących reklamy online, ponieważ umożliwiają podjęcie działań dostosowujących ich treści. W przypadku promocji w mediach społecznościowych istotna jest nie tylko precyzja lecz także szybkość predykcji popularności. Zdarza się, że zabawny spot reklamowy rozprzestrzenia się jak wirus, uzyskując w ciągu kilku dni setki tysięcy lub miliony wyświetleń. Zwykle odzew jest pozytywny, lecz jeśli reklama wywoła krytykę, to na odrobienie

strat jest za późno. Klasyczne testowanie A/B polegające na określeniu czy zmiana B w oryginale A wpływa na współczynnik konwersji, nie pozwala jednocześnie testować wielu zmian i badać korelacji między nimi. Ponadto wymaga odpowiednio dużej grupy testerów.

Habilitant opracował (i opisał w publikacji P1) metodę *Popularity-SVR* - przeznaczoną dla automatycznej predykcji popularności traktowanej jako zadanie regresji, czyli predykcję liczby wyświetleń klipów wideo w czasie tt , w oparciu o dane z pierwszych tr godzin/dni po publikacji. Wykorzystał model regresji oparty o maszynę wektorów nośnych SVR (tzn. Support Vector Machine) z jądrem przekształceń RBF (tzn. Radial Basis Functions). Umożliwiło to mapowanie wektorów cech w przestrzeni nieliniowej, w której łatwiej jest uchwycić relacje między wzorcami ewolucji popularności filmów. Habilitant porównał metodę *Popularity-SVR* z innymi metodami opartymi na modelu regresji, tzn. UL (tzn. Univariate Linear), ML (tzn. Multivariate Linear), MRBF (kombinacja regresji ML z RBF). We wszystkich metodach czas treningu rośnie wraz ze wzrostem zbioru treningowego, natomiast czas predykcji jest stabilny. Metoda *Popularity-SVR* ma najkrótszy czas treningu dla małego zbioru treningowego, potem krótszy czas osiąga metoda UL. Czas predykcji metodą *Popularity-SVR* jest krótszy tylko od MRBF. Reasumując, zaprezentowana metoda *Popularity-SVR* jest oryginalna, z punktu widzenia czasu obliczeń porównywalna z innymi metodami, natomiast odznacza się większą precyzją predykcji.

Ważnym osiągnięciem habilitanta jest analiza jak poszczególne atrybuty wideo, wizualne i czasowe, wpływają na popularność. Dla twórców wideo szczególnie pożądana jest wiedza na temat atrybutów wizualnych, ponieważ umożliwia wstępne oszacowanie popularności wideo jeszcze przed jego umieszczeniem w sieci. W pracy P1 habilitant podał współczynniki korelacji Spearmana dla dziewięciu analizowanych atrybutów wizualnych. Dobrą korelację wykazały trzy atrybuty wizualne, które moim zdaniem warto brać pod uwagę przy szacowaniu popularności danego wideo. Najlepiej skorelowane są tzw. głębokie cechy wideo (współczynnik korelacji równy 0,1361) otrzymane w wyniku przetworzenia przez sieć neuronową ResNet-152 miniaturek reprezentujących sceny wideo. Dobrą korelację wykazał również atrybut clutter (0,1201), będący miarą ewentualnego „bałaganu” obecnego na ramkach wideo. Ramka powinna mieć wyeksponowany obiekt/obiekty pierwszoplanowe, a zbyt dużo szczegółów (np. w tle) przeszkadza widzowi w odbiorze filmu. Dynamika sceny (0,0822) może mieć również wpływ na popularność, ponieważ widz jest przyzwyczajony do dynamicznych klipów wideo, w których ujęcia łączone przez cięcia mają 2 lub 3 sekundy, natomiast nie lubi długich statycznych ujęć. Zaskakujący jest niski współczynnik korelacji uzyskany dla koloru (0,0385), który uważany jest za główną cechę obrazu. Być może wynika to z przyjęcia dla klasyfikacji koloru dziesięciu kolorów będących współrzędnymi modelu kolorów HSV (Hue Saturation Value). Na obrazie z rysunku 1 widzimy głównie kolor brązowy, natomiast komputer obliczył kolor pomarańczowy. Brąz jest słabo nasyconą barwą pomarańczowo-czerwoną, więc teoretycznie wynik jest poprawny, lecz z punktu widzenia percepcji widza brązowy i pomarańczowy są zdecydowanie różnymi kolorami. Warto sprawdzić czy w klasyfikacji koloru poprawniejsze wyniki da pięć klas używanych w projektowaniu: kolory ciepłe jasne (np. żółty, pomarańczowy, czerwony), kolory ciepłe ciemne (np. złoty, czerwono-

brązowy, brązowy), kolory zimne jasne, kolory zimne ciemne, kolory neutralne (biały, szary, czarny). Nie zaskakuje ujemny współczynnik korelacji dla tekstu (-0,0157), który na rysunku 1 jest napisem „nadrukowanym” na ramkę wideo, ponieważ widz nie lubi czytać tekstu na filmie jeśli nie musi. Dlatego w filmie obcojęzycznym lektor zastąpił teksty dialogów. Nie oznacza to, że tekst na obrazie filmowym zawsze jest zbędny. Reklama bez słów nie istnieje, lecz tekst jest w niej integralną częścią kompozycji obrazu.

Pozytywną cechą charakterystyczną działalności badawczej Habilitanta jest poszukiwanie nowych obszarów badawczych i nowych zastosowań opracowanych metod i modeli, tym bardziej że z reguły wymagają one nowych rozwiązań lub rozszerzenia. Rozszerzoną metodę Popularity-SVR wykorzystał on w metodzie Comixify (P4) automatycznie przekształcającej film w komiks. Zadaniem Habilitanta była ekstrakcja klatek kluczowych filmu. W jednym z jej etapów użył on wartość popularności oszacowanej metodą opartą o regresję wektorów nośnych (SVR) jako miarę wartości estetycznej. Alternatywną miarą wartości estetycznej była jakość obrazu. W publikacji nie analizowano wyników uzyskanych tymi metodami (czy obie zwracały podobne wyniki, którą stosować i kiedy).

Generalnie, ocena estetyki jest trudna, ponieważ na końcu zawsze mamy do czynienia z subiektywną oceną widzów. Oglądając obrazki umieszczone w publikacji P4 uważam, że lepsze wyniki z punktu widzenia wartości estetycznej daje metoda Habilitanta i jego zespołu. Niemniej, w przypadku badań, w których estetyka jest jednym z kluczowych wymagań, oprócz oceny obiektywnej (np. poprawność i jakość odwzorowanie krawędzi i koloru) celowa byłoby też ocena ekspercka i/lub subiektywne testy z użytkownikami (widzami) dla porównania estetycznej wartości komiksu otrzymanego przez autorów z komiksami innych badaczy.

W autoreferacie Habilitant napisał, że ograniczeniem Popularity-SVR jest rozdzielenie algorytmu przewidywania popularności na dwa kroki: ekstrakcję metadanych oraz predykcję popularności na ich podstawie. Rozwiązaniem tego problemu jest opisana w autoreferacie i publikacji C4 metoda Popularity-LRCN (Long-term Recurrent Convolutional Network), która opiera się na wytrenowaniu modelu głębokiej sieci neuronowej w celu binarnej klasyfikacji materiałów wideo na popularne i niepopularne w oparciu o predykcję popularności. Habilitant napisał, że zastosowanie rekurencyjnej architektury neuronowej opartej o elementy LSTM (Long Short-Term Memory) w połączeniu z elementami konwolucyjnymi umożliwiło zwiększenie precyzji predykcji wideo o 7%, przy wzroście współczynnika korelacji Spearmana o ponad 30%. Umieszczone w autoreferacie wyniki wskazują, że metoda Popularity-LRCN wykorzystuje w obliczeniach zarówno informacje o wizualnej kompozycji sceny jak również o jej kolorze. Prawie wszystkie miniatury filmów zaliczonych do popularnych mają ciepłą kolorystykę, natomiast w filmach o niskim współczynniku popularności dominują kolory zimne.

Habilitant ze współautorami zaproponował w publikacji C5 architekturę sieci neuronowej, opartą o dwukierunkowe elementy LSTM, przeznaczoną do predykcji popularności wideo wyłącznie w oparciu o tekst tytułu. Na wejściu sieci podawane są słowa tytułu przetworzone do postaci wektorów o stałej długości przy pomocy pre-trenowanych wektorów GloVe. Na wyjściu otrzymuje się przynależność wideo do klasy popularnej lub niepopularnej. Wyniki

klasyfikacji są o 15% precyzyjniejsze w porównaniu z metodami płytkimi i o kilka procent precyzyjniejsze w porównaniu z innymi metodami głębokimi. Zaletą dwukierunkowej sieci neuronowej jest możliwość interpretacji wpływu danego słowa na finalną predykcję.

W zasadzie publikacje C4 i C5, w których przedstawiono wymienione wyżej metody (i na które Habilitant powołuje się w autoreferacie) powinny zostać włączone do jednotematycznego cyklu publikacji. Z drugiej strony ich elementy pojawiły się w publikacjach P2 i P3, w których opisano sieci z dodanym mechanizmem uwagi i model predykcji popularności na podstawie fuzji obu modalności: obrazu wideo i tekstu tytułu. Kombinacja tych modalności zwiększyła o 17% wartość współczynnika korelacji Spearmana i o 4% precyzję.

Znacznym osiągnięciem w zakresie badań nad popularnością z użyciem głębokich sieci neuronowych są aplikacyjne dokonania Habilitanta. Webowa aplikacja metody Comixify została udostępniona w Internecie wraz z kodem źródłowym. Wynikiem badań zaprezentowanych w P1, P2 i P3 jest zestaw narzędzi informatycznych (W1) dedykowanych twórcom treści multimedialnych publikowanych w mediach społecznościowych, które zostały wdrożone przez Group Nine Media - jedną z największych na świecie firm na rynku wydawców treści dla mediów społecznościowych:

- responsywny bot umożliwiający interaktywną rozmowę na temat zgromadzonych w serwisie danych,
- internetowa aplikacja wykorzystująca metodę Popularity-LRCN w klasyfikacji filmu i znajdująca miniaturę optymalną w kontekście przyszłej popularności wideo,
- interfejs analizujący popularność filmu na podstawie jego tytułu, oparty na sieci neuronowej opisanej w publikacjach P2 i P3 i wizualizujący wagę każdego słowa tytułu w oparciu o mechanizm uwagi,
- aplikacja umożliwiająca przewidywanie popularności filmu na podstawie jego klatek (P2 i P3) i wizualizująca przy pomocy mapy ciepła obszary, które przyciągają wzrok widza.

Kolejnym obszarem zainteresowań Habilitanta, posiadającym duży potencjał naukowy i praktyczny, jest stosowanie głębokich sieci neuronowych w e-commerce. Już dawno firmy przekonały się, że handel prowadzony w Internecie może przynieść znaczne korzyści nawet wtedy, gdy firma oferuje także tradycyjne formy sprzedaży, ponieważ klienci często przeglądają ofertę produktów na stronie internetowej przed pójściem do sklepu. Ważne jest, aby potencjalny klient znalazł to czego poszukuje, co przy dużym asortymencie może być trudne. W niektórych branżach kluczowe dla wyszukiwarki są tekstowe parametry techniczne (np. AGD, komputery). W innych branżach parametrami głównymi z punktu widzenia klienta są cechy wizualne produktów (np. ubrań, butów, elementów wystroju wnętrz). Dlatego w niektórych sklepach internetowych klientka może wybrać z list wizualne parametry produktu, takie jak kolor, styl i dekolot sukienki (Zalando) lub kolor i strukturę płytek łazienkowych (modnydom24.pl). Firma Plastream oferuje branży e-commerce platformę technologiczną w celu rozpoznawania produktów widocznych na zdjęciach i w materiałach wideo (np. sukienki celebrytki), aby wskazać identyczne lub najbardziej podobne produkty w bazie zdjęć sklepów internetowych.

Habilitant podjął nowatorską próbę wyszukiwania produktów na podstawie ich podobieństwa stylistycznego, które nie jest tożsame z podobieństwem wizualnym i jest zadaniem znacznie trudniejszym. W autoreferacie napisał, że głównym wyzwaniem naukowym była odpowiedź na pytanie jak definiować podobieństwo stylistyczne korzystając z reprezentacji różnych modalności: wizualnej (zdjęcie produktu) i tekstowej (opis produktu). W P5 zdefiniowano je jako „charakterystyczny sposób, który pozwala na grupowanie obiektów w pokrewne kategorie”. Aby wynik wyszukiwania nie był kompilacją jednomodalnych wyników, Habilitant zaproponował koncepcję głębokiej sieci neuronowej wykorzystującej mechanizm sieci syjamskiej, która przyjmuje dane z pary obraz-tekst (zamiast pojedynczych danych). Umożliwia to znalezienie lepiej dopasowanych produktów, ponieważ klient może rozszerzyć zapytanie wizualne o zapytanie tekstowe, w którym dostarczy nowych informacji. O tym, że badania nad podobieństwem stylistycznym mają nie tylko duży potencjał naukowy, lecz także praktyczny świadczy wykorzystanie ich w Style Search Engine, webowej aplikacji pokazującej detekcję produktu i wyszukanie podobnych produktów w bazie danych IKEI (P5). Habilitant napisał w autoreferacie, że aplikacja zostanie wykorzystana w silniku rekomendacji „jednego z największych dostawców rozwiązań e-commerce w Polsce”. Domyślam się, że jest to IKEA.

Trzy ostatnie publikacje (P6, P7, P8) są wynikiem badań Habilitanta nad rozpoznawaniem obrazów. Choć pierwsze badania w tej dziedzinie podjęto w latach sześćdziesiątych ubiegłego wieku, także obecnie ma ona olbrzymi potencjał naukowy. Zastosowanie w niej głębokich sieci neuronowych zrewolucjonizowało technologię rozpoznawania wizualnego. Na przykład w medycynie umożliwiło automatyczne wykrywanie, klasyfikowanie i przewidywanie ewolucji tkanki nowotworowej. Po wprowadzeniu RODO, Facebook za zgodą użytkownika rozpoznaje jego twarz na profilach innych użytkowników i zawiadamia o tym. W Chinach „płacenie twarzą” jest popularniejsze niż płacenie gotówką lub kartą bankową i choć firmy ubezpieczeniowe bronią się przed autonomicznymi samochodami, coraz bliższy jest czas gdy będziemy nimi jeździli.

Kolejnym ważnym osiągnięciem Habilitanta są opracowane przez niego nowe modele lokalizacji i śledzenia twarzy oraz rozpoznawania emocji na twarzach znajdujących się na zdjęciach lub wideo. Możliwość prawidłowej oceny emocji jest istotna nie tylko dla psychiatrów, psychologów i socjologów, lecz także dla automatycznego wykrywania zagrożeń przez systemy monitoringu i dla twórców antropomorficznych botów (w komunikacji człowiek-komputer). Komputerowo generowana animacja twarzy jest stosowana od postaci z kreskówek po realistyczne modele 3D, które można wykorzystać w filmach zamiast prawdziwych aktorów.

Komputerowa klasyfikacja emocji jest zadaniem bardzo trudnym, ponieważ nawet człowiek nie zawsze jest pewien jakie emocje rysują się na twarzy, którą widzi. Ponadto obraz emocji zależy od kontekstu. Niemniej, każdą twarz można opisać używając cech charakterystycznych określonych za pomocą punktów lub aktywnych konturów. Od specyfiki problemu badawczego zależy sposób opisania na twarzy zmian wywołanych emocją. Jedną z metod jest słowny opis, np. „oczy zamknięte i usta lekko otwarte”. Szeroko stosowany

(np. przez psychologów i grafików komputerowych) jest FACS (Facial Action Coding System), który koduje emocje przy pomocy jednostek akcji (AU). Innym sposobem jest opis ilościowy w odniesieniu do zmian geometrycznych twarzy. Takie zmiany geometrii można opisać za pomocą standardu MPEG-4 z jego parametrami animacji twarzy (FAP).

Habilitant używa punkty charakterystyczne i system FACS. Do ekstrakcji punktów charakterystycznych ze zdjęcia lub materiału wideo zaproponował oryginalną architekturę DAN (Deep Alignment Network) konwolucyjnej sieci neuronowej (P6), inspirowaną frameworkiem Cascade Shape Regression (CSR). DAN w przeciwieństwie do CSR w każdym etapie obliczeń używa do ekstrakcji cały obraz twarzy (zamiast jego część) i normalizuje położenie punktów. Czyny to obliczenia bardzo odporne na duże różnice, zarówno w inicjalizacji jak i ułożeniu głowy oraz na okluzję.

Rozszerzeniem metody lokalizacji i śledzenia głowy z zastosowaniem DAN jest metoda EmotionalDAN (P7), która pozwala sklasyfikować emocje. Z punktu widzenia wielu zastosowań dobrym pomysłem były dwa poziomy emocji: wysokopoziomowy (*emocje pozytywne*, negatywne i neutralne) i niskopoziomowy (*szczęście, pogarda, strach, smutek, obrzydzenie, złość, neutralna*), przy czym każda emocja z poziomu niskiego, może należeć tylko do jednej klasy z poziomu wysokiego. Dzięki dodaniu do funkcji kosztu sieci DAN składowej obliczanej jako wynik entropii krzyżowej bazującej na klasach emocji, osiągnięto lepsze wyniki klasyfikacji niż znane z literatury. Podsumowaniem analizy emocji jest tabela, która dla każdej emocji podaje jednostki akcji AU (razem z ich opisem słownym) i odpowiadające im punkty charakterystyczne (P8). Wartościowa jest również możliwość wizualizacji (w postaci mapy ciepła) obszarów obrazu, które są analizowane przez sieć.

Habilitant znalazł także tym razem aktualny cel aplikacyjny. Model rozpoznawania emocji został przez niego zaimplementowany w samochodowym systemie analizy w celu wspomagania decyzji podejmowanych przez system autonomiczny, na które wpływ ma obraz z kamery zewnętrznej (rejestrującej sytuację na drodze) i wewnętrznej (rejestrującej emocje kierowcy i pasażera).

Jakość redakcyjna cyklu publikacji

Każda publikacja z ocenianego cyklu jest starannie i jasno napisana oraz właściwie zilustrowana. Jeśli nie ograniczają jej ramy konferencyjne zawiera nie tylko odniesienia do literatury przedmiotu, lecz także opisy metod lub modeli konkurencyjnych lub tych z których czerpał inspiracje. Znaczną część każdej publikacji stanowi rzetelne omówienie testów i ich wyników. Natomiast mimo długich list literatury w wielu publikacjach brakuje szerszego spojrzenia i wnikliwszej dyskusji zagadnień pojawiających się w rozdziałach „Related work”.

Podsumowanie

Przedstawiony do oceny cykl publikacji świadczy o dużej wiedzy, intuicji, dociekliwości i dojrzałości naukowej Habilitanta. Opracowane przez Habilitanta nowe metody i architektury sieci neuronowych oraz zbiory danych, stanowią znaczny twórczy wkład w dyscyplinę informatyka techniczna. Należy podkreślić, że dotyczą dynamicznie rozwijającego się

głębokiego uczenia, które jest obecnie polem intensywnych badań na całym świecie. Główne osiągnięcia Habilitanta to:

- Popularity-SVR - nowa metoda predykcji popularności materiałów opublikowanych w sieciach społecznościowych w oparciu o trend wzrostu popularności,
- nowe architektury sieci neuronowych, w tym architektura do wizualizacji decyzji podejmowanych przez sieć neuronową przy predykcji popularności, architektura lokalizacji i śledzenia twarzy Deep Alignment Network (DAN), rozszerzająca ją architektura EmotionalDAN do klasyfikacji emocji, multimodalna architektura DeepStyle do wyszukiwania podobnych stylistycznie materiałów,
- zgromadzone i opublikowane zbiory danych, umożliwiające dalszy rozwój metod predykcji popularności materiałów publikowanych w sieciach społecznościowych i metod wyszukiwania podobieństw stylistycznych.

Istota badań Habilitanta leży nie tylko w warstwie naukowej, lecz także w warstwie praktycznej. Wiele firm poszukuje oryginalnych nowatorskich rozwiązań, ponieważ daje to im przewagę konkurencyjną. O wysokiej ocenie rozwiązań zaproponowanych przez Habilitanta świadczy zainteresowanie nimi dużych światowych firm:

- Group Nine Media – zestaw narzędzi informatycznych do predykcji popularności, wskazany jako osiągnięcie konstrukcyjne (W1,
- IKEA - wyszukiwanie produktów na podstawie obliczania podobieństwa stylistycznego.

1.2. Ocena dorobku publikacyjnego

Dorobek publikacyjny przed doktoratem

Do istotnego dorobku naukowego dr inż. Tomasza Trzcíńskiego, uzyskanego przed doktoratem, należy zaliczyć publikacje, które są wynikiem jego badań związanych z doktoratem pt. „Uczenie i dopasowanie binarnych lokalnych deskryptorów cech” :

- 3 publikacje w czasopismach z bazy Journal Citation Report, IF = 15,581,
- 3 recenzowane publikacje konferencyjne.

Dorobek publikacyjny po doktoracie

Do istotnego dorobku naukowego dr inż. Tomasza Trzcíńskiego, uzyskanego po doktoracie, należy zaliczyć:

1) publikacje z jednotematycznego cyklu pt. „Metody estymacji subiektywnych atrybutów charakteryzujących materiały wizualne z wykorzystaniem głębokich sieci neuronowych” (2017 – 2018):

- 5 publikacji w czasopismach z bazy Journal Citation Report, IF = 14,581; wg wkładu Habilitanta IF = 7,315,
- 3 recenzowane publikacje konferencyjne indeksowane w Web of Science,

2) publikacje z badań nad tematem jak wyżej, niezaliczone do jednotematycznego cyklu publikacji:

- 1 publikacja w czasopiśmie z listy B,
- 6 recenzowanych publikacji konferencyjnych,

- 3) publikacje z badań nad rozwojem algorytmów uczenia maszynowego do analizy danych w eksperymencie ALICE na Wielkim Zderzaczu Hadronów w CERN (2015 – 2019):
 - 37 publikacji w czasopiśmie z bazy Journal Citation Report, IF = 177,001
 - 1 publikacja w czasopiśmie z listy B,
 - 4 publikacje w recenzowanych materiałach konferencyjnych,
- 4) publikacje z badań nad rozwojem metod uczenia maszynowego dla metod jednoczesnej lokalizacji i mapowania na podstawie obrazu (2017 – 2018):
 - 2 publikacje w czasopiśmie z bazy Journal Citation Report, IF = 12,236,
 - 6 publikacji w recenzowanych materiałach konferencyjnych,
- 5) publikacje na temat analizy zdjęć medycznych z użyciem sieci neuronowych (2018 - 2019):
 - 1 publikacja w czasopiśmie z bazy Journal Citation Report, IF = 0,964
 - 3 publikacje w recenzowanych materiałach konferencyjnych.
- 6) publikacje, które Habilitant nie zakwalifikował do wyżej wymienionych grup:
 - 1 publikacja w czasopiśmie z bazy Journal Citation Report, IF = 1,066 (2019)
 - 6 publikacji w recenzowanych materiałach konferencyjnych (2017 – 2020)

W sumie po doktoracie 76 publikacji: 46 publikacji z listy A, dwie publikacje z listy B, 28 recenzowanych publikacji konferencyjnych. Sumaryczny IF = 205,848. Liczby te różnią się od podanych przez Habilitanta, który w zestawieniu sumarycznym podał 72 publikacje po doktoracie.

Ocena liczby cytowań

O jakości prowadzonych badań i wpływie osiągnięć naukowych dr inż. Tomasza Trzcńskiego na dyscyplinę Informatyka Techniczna świadczy liczba cytowań jego prac. Parametry te zostały określone na podstawie baz Web of Science, Scopus i Google Scholar:

Baza Web of Science:

- Całkowita liczba cytowań: 709
- Liczba cytowań bez autocytowań: 679
- Indeks Hirscha: 9

Baza Scopus:

- Całkowita liczba cytowań: 1026
- Liczba cytowań bez autocytowań: 950
- Indeks Hirscha: 11

Baza Google Scholar:

- Całkowita liczba cytowań: 2055
- Indeks Hirscha: 19

Biorąc pod uwagę okres pięciu lat po doktoracie uważam, że dorobek publikacyjny dr inż. Tomasza Trzcńskiego jest imponujący. Wszystkie publikacje są wprawdzie współautorskie (od dwóch do czterech autorów), lecz nawet uwzględniając tylko udział dr inż. Tomasza Trzcńskiego, wartość Impact Factor jest wysoka. Ponadto liczba cytowań świadczy

o zauważalności prac dr inż. Tomasza Trzcińskiego w środowisku naukowym. Są to wartości wysokie i wystarczające do uzyskania stopnia doktora habilitowanego w dyscyplinie Informatyka Techniczna i Telekomunikacja.

1.3. Ocena aktywności w prowadzeniu projektów po doktoracie

Po doktoracie dr inż. Tomasz Trzciński był kierownikiem i wykonawcą w następujących projektach badawczych:

1) Granty Narodowego Centrum Nauki:

- pt. „Opracowanie metod uczenia maszynowego do monitorowania jakości danych o dużej objętości oraz interaktywnych metod ich wizualizacji na przykładzie eksperymentu ALICE na Wielkim Zderzaczu Hadronów w CERN” (SONATA 11/ST6, 2016-2019) – kierownik projektu,
- pt. „Tracking ALICE – Szybka Symulacja z Zastosowaniem Metod Uczenia Maszynowego w ALICE CERN” (PRELUDIUM 16/ST6, 2019-2022) – opiekun naukowy.

2) Granty Narodowego Centrum Badań i Rozwoju:

- 3 granty w których Habilitant był kierownikiem projektu realizowanego jako podwykonawca w grantie NCBiR „Szybka Ścieżka”,

3) Projekty badawcze finansowane w ramach współpracy z firmą Google:

- 3 granty badawcze we współpracy z Google Zürich (2016-2020) – kierownik projektu,
- 2 Google Cloud Platform Grant: dotacje do wykorzystania na obliczenia.

4) 3 granty dziekańskie dla młodych naukowców – kierownik projektu.

Od 2015 roku dr inż. Tomasz Trzciński jest również kierownikiem naukowym ponad 20 badaczy, inżynierów i naukowców, odpowiedzialnym za badania R&D w warszawskiej firmie Toolploom, produkującej oprogramowanie dla firm technologicznych. W latach 2017-2018 był kierownikiem i wykonawcą ośmiu projektów w których wykorzystał wiedzę i wyniki swoich badań w dziedzinie uczenia maszynowego i wizji komputerowej do opracowania algorytmów i ich implementacji:

- 5 projektów dla firm zagranicznych: Group Nine (USA), NakedLabs (USA), MindMaze (USA), June Life (USA), Voyage (USA),
- 3 projekty dla firm polskich: Grupa Pracuj, Sanok Rubber, TomTom.

Biorąc pod uwagę okres pięciu lat od doktoratu, liczba projektów badawczych realizowanych przez dr inż. Tomasza Trzcińskiego jest znaczna i wystarczająca do uzyskania stopnia doktora habilitowanego.

1.4. Ocena pozostałej aktywności naukowej po doktoracie

Dr inż. Tomasz Trzciński znalazł się wśród 100 liderów Europy Centralnej i Wschodniej, przewodzących globalnej innowacji, wybranych w konkursie New Europe 100 Challenger 2016. Jest to najlepsze podsumowanie jego potencjału naukowego, który od tamtego czasu znacznie się zwiększył.

Brał aktywny udział w życiu naukowym. Po doktoracie uczestniczył w 11 konferencjach (8 zagranicznych, 3 międzynarodowe konferencje polskie), był członkiem komitetów

technicznych i komitetów naukowych (lub Rady Naukowej). Wygłosił 5 wykładów plenarnych i 5 wykładów towarzyszących referatowi. Wykonał 25 recenzji materiałów konferencyjnych i 24 recenzje dla czasopism z listy JCR. Jest członkiem komitetu redakcyjnego czasopisma IEEE Acces i członkiem kilku stowarzyszeń naukowych: IEEE, Computer Vision Foundation, Marie Curie Alumni Association, Polskiego Stowarzyszenia Sztucznej Inteligencji.

W 2017 roku odbył staż naukowy na Uniwersytecie Stanforda (USA) w ramach europejskiego projektu Horizon 2020 Maria Skłodowska-Curie, natomiast w 2019 roku w Nanyang University of Technology (Singapur).

Jest członkiem grupy roboczej przy Ministerstwie Cyfryzacji ds. edukacji w ramach przygotowania Polskiej Strategii AI oraz członkiem wielu zespołów oceniających projekty badawcze:

- członek panelu oceniającego wnioski grantowe kierowane do Narodowego Centrum Nauki (Call OPUS i PRELUDIUM) w dziedzinie ST6 (Informatyka i technologie informacyjne),
- recenzent wniosków o finansowanie badań w ramach wielu programów: SONATA NCN, Maria-Skłodowska-Curie EDGE Fellowship, ESMERALDA Project (European SMEs Robotics Applications), Composable Models and Software for Robotic Systems,
- członek jury konkursowego programu POTENCJOMETR (ogólnopolskiego konkursu na studencki projekt naukowy o potencjale wdrożeniowym).

Jako członek zespołu eksperckiego wykonał trzy ekspertyzy zamówione przez firmy w Politechnice Warszawskiej w 2018 roku.

Powyższy dorobek jest znaczny i świadczy o bardzo dużej aktywności naukowej dr inż. Tomasza Trzcieskiego oraz o pozytywnej ocenie ze strony międzynarodowego środowiska naukowego.

2. Ocena dorobku dydaktycznego, organizacyjnego i popularyzującego naukę

Dorobek dydaktyczny i organizacyjny dr inż. Tomasza Trzcieskiego obejmuje:

1) wypromowanych absolwentów w Politechnice Warszawskiej:

- studia magisterskie: 8,
- studia inżynierskie: 13,

2) opiekę naukową nad doktorantami w Politechnice Warszawskiej: 1

3) zajęcia dydaktyczne na Wydziale Elektroniki i Technik Informatycznych Politechniki Warszawskiej:

- Podstawy sztucznej inteligencji
- Przetwarzanie Cyfrowe Obrazów
- Analiza Algorytmów

Dorobek organizacyjny:

1) współtworzenie kierunku studiów „Biznes.AI. Technologia, Prawo, Zastosowanie Sztucznej Inteligencji” na Akademii Leona Koźmińskiego w Warszawie,

- 2) organizacja warsztatu „Discovering Brain Structure with Machine Learning” podczas konferencji Applied Machine Learning Days na Ecole Polytechnique Federale de Lausanne,
- 3) organizacja sesji specjalnej pt. „Recent advancements in computer vision and image processing with special focus on deep learning methods” podczas konferencji Microwave Week.

Jako swój dorobek popularyzujący naukę Habilitant podał:

- uczestnictwo w konferencjach studenckich: 2,
- uczestnictwo w konferencjach branżowych: 3,
- zaproszone wykłady: 22,
- prowadzone szkolenia: 2.

Uważam, że są to istotne osiągnięcia dr inż. Tomasza Trzcíńskiego w zakresie dydaktyki i popularyzowania nauki.

3. Konkluzja

Ocena osiągnięcia naukowego dr inż. Tomasza Trzcíńskiego, którym jest jednotematyczny cykl publikacji pt. „Metody estymacji subiektywnych atrybutów charakteryzujących materiały wizualne z wykorzystaniem głębokich sieci neuronowych”, jest pozytywna. Cykl ten stanowi twórczy wkład Habilitanta w dyscyplinę Informatyka Techniczna i Telekomunikacja, w szczególności w dziedzinie uczenia maszynowego z użyciem głębokich sieci neuronowych.

Pozytywna jest także ocena istotnej aktywności naukowej dr inż. Tomasza Trzcíńskiego. Oceniany dorobek naukowy, dydaktyczny i organizacyjny po uzyskaniu stopnia naukowego doktora jest wystarczający do uzyskania stopnia naukowego doktora habilitowanego i świadczy o dojrzałości naukowej dr inż. Tomasza Trzcíńskiego.

W mojej ocenie dr inż. Tomasz Trzcínski spełnia wszystkie ustawowe i zwyczajowe kryteria wymagane do uzyskania stopnia doktora habilitowanego. Wnoszę zatem o dopuszczenie dr inż. Tomasza Trzcíńskiego do kolejnych etapów przewodu habilitacyjnego, zmierzających do nadania mu stopnia doktora habilitowanego w dyscyplinie Informatyka techniczna i telekomunikacja.

M. Piętnuszka