RDITT - wpł. 08.11.2023
M. Ikony

Kraków, 31.10.2023

Prof. dr hab. inż. Bogusław Cyganek
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji
Instytut Elektroniki

Recenzja rozprawy doktorskiej *„Multi-image super-resolution reconstruction using deep graph neural networks"* Pana magistra inżyniera Tomasza Tarasiewicza

Wstęp

Recenzja rozprawy doktorskiej Pana magistra inżyniera Tomasza Tarasiewicza pt. „*Multi-image super-resolution reconstruction using deep graph neural networks*[1]". Praca powstała w roku 2023 na Wydziale Automatyki Elektroniki i Informatyki Politechniki Śląskiej. Promotorem pracy jest Pan dr hab. inż. Michał Kawulok, profesor Politechniki Śląskiej. Przygotowanie recenzji zostało wykonane na zlecenie Przewodniczącego Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja, Pana prof. dr hab. inż. Andrzeja Polańskiego. Praca została napisana w języku angielskim. Recenzja sporządzona została w postaci odpowiedzi na pytania dotyczące rozprawy doktorskiej Pana magistra inżyniera Tomasza Tarasiewicza.

1. *Jakie zagadnienie naukowe jest rozpatrzone w pracy i czy zostało ono dostatecznie jasno sformułowane prze autora? Jaki charakter ma rozprawa?*

Główne zagadnienie naukowe podjęte w pracy doktorskiej Pana magistra inżyniera Tomasza Tarasiewicza dotyczy *rekonstrukcja obrazów o dużej rozdzielczości z reprezentacji wielo-obrazowych z użyciem głębokich grafowych sieci neuronowych.*

Obecnie każdy z nas wykonuje niezliczone zdjęcia, często wykorzystując aparat cyfrowy wbudowany w nasz telefon komórkowy. Zapewne jest więc naszym wspólnym doświadczeniem, że niektóre zdjęcia po prostu „nie wychodzą", albo nie taka eskpozycja, albo odbicie światła, albo zdjęcie jest poruszone, albo na przykład chcemy odczytać numer rejestracji, którego jednak nie widać ze względu na zbyt niską rozdzielczość naszego zdjęcia. Problem taki oczywiście możemy naprawić wykonując kolejne zdjęcie, o ile nasz obiekt już się nie oddalił, czy też sięgając po lepszy aparat lub droższą „komórkę". Tym niemniej istnieją pewne ograniczenia fizyczne związane między innymi z długością ogniskowej, wielkością sensora optycznego, jasnością obiektywu, zniekształceniami optycznymi, itd. które ani

[1] Rekonstrukcja obrazów o dużej rozdzielczości z reprezentacji wielo-obrazowych z użyciem głębokich grafowych sieci neuronowych.

B. Cy

w tym, ani w innych bardziej wyrafinowanych przypadkach, takich jak wykonywanie zdjęć satelitarnych, nie umożliwią uzyskania dowolnie dużej rozdzielczości. Jeśli nawet istnieją metody sprzętowego obejścia tego, czy podobnych problemów to przeważnie są one albo niedostępne, albo po prostu zbyt drogie. Okazuje się jednak, że możliwa jest poprawa jakości obrazu, a w szczególności zwiększenie jego rozdzielczości już po wykonaniu zdjęcia, poprzez wykorzystanie odpowiedniego algorytmu. Korzystając z wcześniej zgromadzonych informacji, jak również pewnej redundancji informacyjnej w naszym zdjęciu algorytm taki zmieni zdjęcie w taki sposób, że będzie możliwe odczytanie wcześniej niewidocznych przykładowych numerów rejestracji. Wymienione metody bazują na tzw. algorytmach super-rozdzielczości, czy też nad-rozdzielczości (ang. *super-resolution*; SR), które na podstawie albo pojedynczego zdjęcia niskiej rozdzielczości (ang. *low resolution*; LR), albo też ich sekwencji, obliczą wersję SR. Problem jest jednak bardzo trudny i nie istnieje jedno uniwersalne rozwiązanie. Przeciwnie, wersji algorytmów jest bardzo wiele i ciągle powstają nowe i lepsze. Istnieje również bardzo wiele dziedzin, które praktycznie potrzebują dobrze działających metod obliczania SR – obrazy satelitarne, obrazowanie medyczne, kinematografia, nadzór i bezpieczeństwo, to tylko niektóre z nich. Praca naukowa wraz z metodami opisanymi w dysertacji doktorskiej Pana magistra inżyniera Tomasza Tarasiewicza wpisuje się właśnie w ten nurt – zadanie, które postawił przed sobą Pan Tomasz to opracowanie nowatorskich metod otrzymywania reprezentacji SR na podstawie sekwencji obrazowych, charakteryzujących się najlepszą możliwą jakością. Zadanie to, podzielone na szereg bardziej szczegółowych, zostało przez niego skutecznie rozwiązane, o czym będzie mowa w dalszej części recenzji.

Podsumowując – problemy badawcze podjęte przez Pana magistra inżyniera Tomasza Tarasiewicza mają znaczenie zarówno w obszarze badań teoretycznych, jak i eksperymentalnych i mogą znaleźć liczne zastosowania w dziedzinach pokrewnych. Praca ma charakter teoretyczno-doświadczalny.

2. *Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle, czy świadczą one o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonywujący?*

Zagadnienia badawcze, jak i metody ich rozwiązania zaprezentowane w pracy Pana magistra inżyniera Tomasza Tarasiewicza są bardzo istotne zarówno z naukowego punktu widzenia, jak również od strony aplikacyjnej. Każdego dnia przybywają nowe prace na temat szeroko rozumianej poprawy jakości obrazów, w tym w szczególności odtworzenia większej rozdzielczości, czy to na podstawie obrazu pojedynczego, czy też kilku obrazów tej samej sceny. Niemniej istotne są tutaj narzędzia naukowe w postaci głębokich architektur neuronowych, a w szczególności ich nowej odmiany – grafowych sieci głębokich. Cała dziedzina tzw. ML/AI wykorzystująca głębokie sieci neuronowe jest bardzo dynamicznie rozwijana, czego rezultatem są wręcz setki publikacji zaledwie w przeciągu ostatnich kilku lat. Choćby pobieżne śledzenie osiągnięć i prac, jak również próba usystematyzowania tak wielkiego strumienia prac i publikacji wymaga od badaczy olbrzymiej pracy. W tym aspekcie Pan mgr inż. Tomasz Tarasiewicz bardzo dobrze sobie poradził. W obszernej rozprawie doktorskiej odwołuje się on aż do stu siedemdziesiąt pozycji literaturowych, z których w kilku przypadkach jest również jednym z autorów. Świadczy to o bardzo dobrym przygotowaniu naukowym Doktoranta.

3. *Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?*

Główne zagadnienia naukowe zaprezentowane przez Pana magistra inżyniera Tomasza Tarasiewicza w Jego rozprawie doktorskiej dotyczą opracowania nowych metod rekonstrukcji obrazów o dużej rozdzielczości na podstawie reprezentacji wielo-obrazowych z użyciem głębokich grafowych sieci neuronowych. Autor sformułował następujące tezy pracy (str. 7, w tłumaczeniu):

1. Zestaw obrazów LR z przesunięciami sub-pikselowymi reprezentowany za pomocą grafu, może być następnie przetworzony za pomocą neuronowej sieci grafowej (GNN) w celu uzyskania super-rozdzielczości prowadząc do wyników porównywalnych lub lepszych do tych uzyskanych przez wiodące architektury oparte na sieciach konwolucyjnych i z wykorzystaniem wielu-obrazów (ang. *multi-image super-resolution*; MISR).
2. Sieci GNN mogą poprawić swoje wyniki w zakresie MISR poprzez zastosowanie technik inspirowanych istniejącymi, najnowocześniejszymi modelami MISR opartymi na CNN. Techniki te obejmują indywidualną ekstrakcję cech dla każdego obrazu LR, wykorzystanie mechanizmów uwagi (ang. *attention*) oraz dynamiczną, i dającą się wytrenować, rejestrację danych wejściowych.
3. Sieci GNN mogą zrekonstruować scenę z określonego punktu w czasie, wybierając konkretny obraz referencyjny z zestawu wejściowych obrazów niskiej rozdzielczości LR, a inne obrazy służą jako dodatkowe źródła informacji w celu zwiększenia dokładności super-rozdzielczości. Takie podejście może pomóc w redukcji artefaktów wizualnych w obszarach o dużej zmienności czasowej.

Tak postawione tezy zostały przez Pana magistra inżyniera Tomasza Tarasiewicza wykazane, zarówno teoretycznie, ale przede wszystkim zweryfikowane eksperymentalnie z wykorzystaniem zarówno danych syntetycznych, jak i otrzymanych z faktycznych pomiarów. Co więcej, opracowane metody zostały opublikowane w znanych materiałach konferencyjnych oraz czołowych periodykach naukowych. Analiza publikacji Doktoranta zawarta została w dalszej części recenzji.

Realizując plan badawczy, którego zarys prezentują powyższe tezy naukowe, Pan magister inżynier Tomasz Tarasiewicz, wykazał się dogłębną wiedzą dotyczącą przede wszystkim przetwarzania obrazów o różnych spektrach, jak również znajomością budowy oraz projektowania nowoczesnych architektur głębokich sieci neuronowych, w tym sieci grafowych. Użyte przez Pana magistra inżyniera Tomasza Tarasiewicza założenia oraz użyte metody, które zostały przedstawione w rozprawie uważam za uzasadnione oraz w pełni odpowiadające poziomowi oraz osiągnięciom światowym w tej dziedzinie. Świadczą one o głębokiej wiedzy Doktoranta w dyscyplinie Informatyka Techniczna i Telekomunikacja.

4. *Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?*

Jak już wspomniano, główne zagadnienia naukowe zaprezentowane przez Pana magistra inżyniera Tomasza Tarasiewicza w Jego rozprawie doktorskiej dotyczą opracowania nowych metod rekonstrukcji obrazów o dużej rozdzielczości na podstawie reprezentacji wielo-obrazowych z użyciem głębokich

15.9.15

grafowych sieci neuronowych. Główne osiągnięcia Doktoranta Pana Tomasza Tarasiewicza w tych obszarach badawczych i zaprezentowane w rozprawie są następujące.

1. Opracowanie szeregu architektur głębokich neuronowych sieci grafowych przystosowanych do obliczania super-rozdzielczości na podstawie szeregu obrazów wejściowych o niskiej rozdzielczości. Są to architektury nazwane przez Autora w następujący sposób (Rozdział 3):
   - **MagNet** – jest to podstawowa architektura grafowa do problemu super-rozdzielczości. Obraz wejściowy reprezentowany jest za pomocą grafu, po którym następują warstwy obliczeń splotów (konwolucji) za pomocą funkcji sklejanych (ang. *splines*). Sygnał przetwarzany jest przez kolejne warstwy redukujące (ang. *max pooling*) oraz przetwarzające kolejne wersje grafu wejściowego, aż do warstwy próbkowania-w-górę (ang. *upsampling*), po której otrzymujemy obraz o zwiększonej rozdzielczości.
   - **MagNet++** – jest to wersja MagNet ze zmodyfikowanym modułem skalowania-w-górę, która zapewnia lepsze zachowanie informacji wizyjnej.
   - **MagNet$_{enc}$** – jest to wersja MagNet ze zmodyfikowanym modułem ekstrakcji cech poprzez m.in. zaczepienie węzłów grafu poprzez dopasowanie do pozycji pikseli medianowych obliczanych z wejściowej serii obrazów niskiej rozdzielczości. Współbieżne przetwarzanie obrazu wraz z obrazem referencyjnym umożliwia uwypuklenie różnic pomiędzy nimi, czyli detekcję cech odpowiadających wysokim częstotliwościom w sygnale, które są charakterystyczne z występowaniem szczegółów obrazowych (np. krawędzie, narożniki, itd.), tak istotnymi przy odtwarzaniu super-rozdzielczości.

   Kolejna grupa metod opracowanych przez Doktoranta to:

   - **MagNAt** – zaawansowana architektura bazująca na sieciach grafowych. Charakterystyczną cechą jest zaproponowana nowa wersja mechanizmu uwagi (ang. *attention*), opisana w Rozdziale (3.5.1). Kolejna cecha to adaptacyjny i trenowany mechanizm rejestracji (dopasowania przestrzennego) obrazów wejściowych. Zaproponowany został mechanizm dynamicznego umieszczania węzłów w modelu, który dostosowuje relację obrazów LR w stosunku do obrazu referencyjnego przy każdym przeliczeniu przez sieć (ang. *forward pass*). Różni się to od modeli konwencjonalnych choćby tym, że relacje te są zwykle ustawiane raz podczas fazy ładowania danych i następnie pozostają niezmienne przez cały czas treningu. Mechanizm dynamicznego dopasowania (ang. *dynamic alignment*) ma wiele zalet, takich jak m.in. odporność na błędy początkowe w ułożeniu obrazów, możliwość dynamicznej adaptacji modelu podczas uczenia, jak również rozszerzone możliwości ekstrakcji istotnych cech obrazowych, czy też zapobieżenie tzw. przeuczeniu się modelu (ang. *overfitting*).
   - **MagNAt$_{no_reg}$** – wersja MagNAt z mechanizmem automatycznej rejestracji (dopasowania) obrazów, tj. brakiem warstwy rejestracyjnej RegNet, co pozwala na porównanie z metodami wykorzystującymi wcześniej obliczone wektory przesunięć obrazów (ang. *shift vectors*).
   - **MagNAt$_{lead}$** – zmodyfikowana wersja MagNAt pod względem rejestracji, kodowania oraz próbkowania-w-górę obrazów wejściowych. Metoda ta zakłada wybór początkowego obrazu jako referencyjnego, co ma znaczenie w zapewnieniu spójności czasowej, tzn. w przypadku obrazów będących sekwencją czasową, tak jak w obrazach wideo.

2. Przeprowadzenie badań eksperymentalnych z wykorzystaniem wszystkich wyżej wymienionych nowych metod, oryginalnie opracowanych przez p. mgra inż. Tomasza Tarasiewicza, jak również innych metod w celu porównania wyników ich działania. Analiza otrzymanych wyników oraz wykazanie skuteczności zaproponowanego przez siebie rozwiązania.

Wszystkie wyżej wymienione osiągnięcia świadczą o dojrzałości naukowej Pana magistra inżyniera Tomasza Tarasiewicza. Doktorant swobodnie porusza się różnorodnych zagadnieniach analizy i przetwarzania obrazów wielospektralnych, jak również w aspektach budowy i modyfikacji wielu typów architektur głębokich sieci neuronowych, a w szczególności sieci opartych na grafach. Na pochwałę zasługuje również Jego warsztat eksperymentalny – przeprowadzone eksperymenty wraz z analizą wyników są dobrze zaprojektowane, wszechstronne, metodologicznie poprawne i wręcz wzorowe.

5. Jaka jest umiejętność poprawnego i przekonywującego przedstawienia uzyskanych wyników oraz jaki jest dorobek publikacyjny Doktoranta

Rozprawa doktorska Pana magistra inżyniera Tomasza Tarasiewicza jest obszerna i liczy 147 strony. Napisana jest bardzo starannie, w języku angielskim. Praca podzielona jest na 7 rozdziałów głównych, po których zamieszczono dodatki. Bibliografia liczy 170 pozycje literaturowe. W spisie tym znajdują się również publikacje p. T. Tarasiewicza. W pracy niestety brakuje krótkiego abstraktu, który w zwięzły sposób opisywałby treść pracy i główne osiągnięcia autora.

Lista publikacji, w których Pan mgr inż. Tomasz Tarasiewicz jest współautorem liczy **8 pozycji**, z których na szczególną uwagę zasługują następujące:

1. **Tomasz Tarasiewicz**, Jakub Nalepa, Reuben A. Farrugia, Gianluca Valentino, Mang Chen, Johann A. Briffa, "*Multitemporal and Multispectral Data Fusion for Super-Resolution of Sentinel-2 Images*," IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-19, **2023**, (IF=8,2).
2. Pawel Kowaleczko, **Tomasz Tarasiewicz**, Maciej Ziaja, Daniel Kostrzewa, Jakub Nalepa, Przemyslaw Rokita, Michal Kawulok "*A Real-World Benchmark for Sentinel-2 Multi- Image Super-Resolution*". Scientific Data, Vol. 10, 2023, p. 644 (IF=8,5).

W pierwszej publikacji, mimo iż autorstwa siedmiu autorów, Pan magister Tomasz Tarasiewicz jest autorem głównym. W kolejnym z wymienionych – jest on drugim z autorów. Obydwie te publikacje mają wysoki współczynnik wpływu (ang. *impact factor*; IF), powyżej 8 punktów.

W czterech z pozostałych publikacjach p. mgr inż. Tomasz Tarasiewicz jest pierwszym z autorów. Są to też bardzo wartościowe publikacje konferencyjne.

Liczba publikacji Doktoranta znacznie przekracza minimalną wartość wymaganą w przewodach doktorskich. Również jakość publikacji świadczy o aktywności i dojrzałości naukowej Pana magistra inżyniera Tomasza Tarasiewicza, jak również o jego rozpoznawalności w środowisku naukowym.

B 975

6. *Jakie są słabe strony rozprawy i jej główne wady?*

Rozprawa doktorska Pana Tomasza Tarasiewicza w znaczącym stopniu bazuje na wcześniejszych publikacjach Doktoranta wraz z innymi współautorami z zespołu Pana Profesora Michała Kawuloka. Wszystkie one bardzo dobrze wpisują się w najnowszy nurt badawczy nad super-rozdzielczością, o których była mowa w poprzedniej części recenzji. Praca jest bardzo konstruktywna – zostały zaproponowane liczne i nowatorskie metody obliczania super-rozdzielczości za pomocą grafowych sieci neuronowych osiągając przy tym wyniki, które dorównują, a niekiedy przewyższają te osiągnięte za pomocą najnowszych światowych rozwiązań (tzw. *state-of-the-art*; SOTA). Mimo tak znaczącej liczby opisanych metod i wynikającej z tego objętości pracy, nie znalazłem poważniejszych błędów, czy też uchybień merytorycznych.

Poniżej przedstawiam uwagi i pytania o naturze polemicznej, czy też rozszerzającej materiał zawarty w rozprawie.

1. Główna metodyka przyjęta przez Doktoranta polega na zbudowaniu sieci grafowej o strukturze, która umożliwia obliczenia super-rozdzielczości na podstawie serii obrazów, tak jak to jest opisane w Rozdziale (3.3) poświęconemu architekturze MagNet++. Jednym z kluczowych etapów jest wstępny krok mapowania oryginalnego wejściowego obrazu, albo serii obrazów na pewną mniej lub bardziej skomplikowaną strukturę grafową. I niewątpliwie jest to etap kluczowy i najbardziej chyba oryginalny w każdej z zaprezentowanych metod. Co więcej, wydaje się że jest to jedyny etap w którym występują „czyste" reprezentacje grafowe, gdyż już kolejny krok (np. opisany na str. 49) to transformacja tej reprezentacji grafowej do tensora wejściowego o wymiarowości $S{\cdot}W{\times}S{\cdot}H$, gdzie S jest współczynnikiem próbkowania w górę (ang. *up-sampling*), natomiast W i H oznaczają odpowiednio szerokość oraz wysokość obrazów wejściowych. W kolejnych akapitach Autor pisze więc o „standard 2D convolutional layers" itd. Nasuwa się więc pytanie o miejsce i rolę grafu w całej tej konstrukcji, albo innymi słowy, na ile podejście oparte o struktury grafowe jest tu kluczowe i jak bardzo odróżnia sieci grafowe od innych sieci warstwowych?
2. Mimo dość ogólnych schematów architektur, np. Rysunek 3.6, brakuje bardziej szczegółowych schematów kolejnych warstw opracowanych sieci neuronowych. Dla przykładu, na str. 49 jest mowa o połączeniach rezydualnych, czy jest to więc rodzaj sieci typu ResNet? Temat ten powinien być bardziej rozwinięty i powinny być przedstawione szczegóły architektur, być może z dostępem do kodu.
3. Również pobieżnie opisany, czy też wyłącznie wspomniany jest mechanizm „uwagi" (attention) w Rozdziale (3.5). Jak czytamy, jest on odpowiedzialny za zbadanie relatywnego wpływu każdego sąsiadującego węzła („*By examining the relative importance of each neighbouring node within a graph representation* [...]") – ale brak szczegółów tego mechanizmu. Na ile jest oryginalny, a na ile powiela ideę i konstrukcje ze znanych transformerów wizyjnych?
4. Proces zwiększania rozdzielczości generuje „nowe informacje", ponieważ chcemy uzyskać wyższą rozdzielczość tam gdzie jej nie było. Skąd jednak możemy mieć pewność, że nie jest to halucynacja sieci, czyli wprowadzenie czegoś „ładnie wyglądającego", ale niekoniecznie prawdziwego. Problem ten występował już wielokrotnie w dziedzinie widzenia komputerowego, np. przy morfologicznych przestrzeniach skal, itd.
5. Pan mgr inż. Tomasz Tarasiewicz dość dokładnie zmierzył czasy działania opracowanych przez siebie metod. Jednakże nie znalazłem żadnych informacji na temat ich implementacji – jak

została wykonana, czy jest dostępna dla innych autorów i czy jest w niej możliwość dalszej optymalizacji pod kątem przyspieszenia działania?

6. Prace eksperymentalne zostały wykonane bardzo precyzyjnie i szczegółowo. Jednakże Autor w różnych miejscach różnie definiuje partycje używane do trenowania oraz do testowania wyników. Niekiedy jednak nie ma mowy o danych 'testowych' (np. str. 65), a jedynie o 'walidacyjnych' (np. str. 60). Czy Doktorant mógłby doprecyzować jakie były używane, a jakie powinny być używane, i jak powinny być otrzymywane wspomniane partycje danych?
7. Niemniej istotnym aspektem usprawniania obrazów są (i) redukcja szumów różnego pochodzenia oraz (ii) próba zwiększania dynamiki sygnału (ang. *high dynamic range imaging*; HDR). Czy opracowane metody mogą być jednocześnie zastosowane/przystosowane do wyżej wymienionych problemów?
8. Autor opracował szereg różnorodnych metod do obliczania super-rozdzielczości, każda o nieco innych właściwościach. Czy jest jakiś sposób połączenia tych metod w rodzaj zespołu (ang. *ensemble*), którego wyniki działania byłyby lepsze niż metod składowych? Rozwiązania takie są znane m.in. w dziedzinie klasyfikacji wzorców.
9. Na ile sposób trenowania, czy też użyty dataset wpływa na późniejszą predykcję? Tzn. jak ściśle-domenowa jest każda metoda (jest to ogólniejszy problem tzw. *tail distribution problem in classification*).
10. Istnieje wiele tematów pokrewnych, które warto by przeanalizować oraz uwzględnić jako ewentualne tematy przyszłych badań, między innymi:
    - Najnowsze trendy w SR: dyfuzyjne metody SR, transformery SR, sieci GAN (ang. *generative adversarial network*; GAN) w zastosowaniach SR;
    - Rola tzw. różnych metod próbkowania w górę (np. *learning based upsampling, attention based upsampling*), bardziej efektywne funkcje kosztu;
    - Nowe i lepsze metodyki ewaluacji algorytmów i danych testowych;
    - Zastosowanie metody *deep image prior* do SR z wykorzystaniem struktur grafowych? Czy jest możliwe zastosowanie takiego podejścia w metodach Doktoranta? A jeśli tak, to jak struktura grafu mogłaby przenosić informację o zadaniu SR.

Bardziej szczegółowe pytania i uwagi natury edycyjnej są następujące:

11. Rozdział 2.6.3 – „*The mechanics of Grpah Neural Networks*", czy w grafach można mówić o „mechanice"?
12. Figure 3.1 oraz otaczający tekst – pewnego wyjaśnienia wymaga mapowanie wartości kolorów RGB, czy też ogólniej – poszczególnych składowych spektralnych do struktury grafu. Na rysunku są dwa wymiary X-Y, podczas gdy wspomniane składowe mogłyby być reprezentowane kolejnym wymiarem Z. Problem ten wymaga pogłębionej analizy, tym bardziej w świetle komentarza na str. 42.
13. Równanie (3.1) zakłada jedynie liniowe przesunięcia współrzędnych. Czy jest to zawsze uzasadnione i czy nie powinno być rozszerzone na inne przekształcenia? Na ile konieczne jest uwzględnienie dyskretnego charakteru współrzędnych obrazowych?
14. Czy w równaniu (3.3) przyjęcie wartości 1 dla $\mathbf{A}(i,j)$, a następnie 1 dla wartości r to nie są zbytnie uproszczenia? Czy zwiększenie wartości r nie dodałoby mechanizmu „uwagi" w lokalnych sąsiedztwach pikseli?

R. 975

15. Str. 45 – wspomniane są funkcje sklejane B-spline, ale jakie dokładnie są to funkcje?
16. Str. 46 – Po stwierdzeniu „*has transformed the graph into a regular, grid-like structure*" można ponowić pytanie, na ile proponowana architektura grafowa istotnie różni się od „klasycznej" architektury warstwowej głębokiej sieci neuronowej, takiej jak np. AlexNet, czy ResNet?
17. W rozdziale (4.1.1) jest mowa o znacznym ograniczeniu obrazów wejściowych, zarówno do skali szarości, jak i do wymiarów zaledwie 222×222 piksele. Jak się to ma do opisywanych zalet grafów, które mogą potencjalnie operować na obrazach o dowolnej głębi i rozmiarach? Z jeszcze mniejszymi łatami obrazowymi rzędu 64×64 mamy do czynienia w eksperymentach, których wyniki opisane są w Tabeli 5.1. Autor powinien skomentować te ograniczenia.
18. Rozdział (4.2.3) – jednym z opisanych tu ograniczeń metod super-rozdzielczości jest brak możliwości radzenia sobie z miejscami o brakującej, czy też ograniczonej dostępności wartości pikseli w pewnych fragmentach/obszarach obrazu. Są jednak metody tzw. *inpaintingu*, które radzą sobie z tym problemem – czy mogłyby być one użyte wraz z zaprezentowanymi metodami super-rozdzielczości?
19. Przyjęcie na stałe aż 1000 epok dla każdego eksperymentu (str. 72) jest chyba wartością nieco przesadzoną i nieuzasadnioną w epoce globalnego ocieplenia. Czy opracowane metody rzeczywiście muszą być uczone aż tak długo? Czy wszystkie? Czy można ten proces jakoś skrócić?
20. W Rozdziale (5.3.3) jest mowa o „ostrości krawędzi" – czy zjawisko to można zmierzyć w jakiś bardziej ilościowy sposób, np. wykorzystując transformację Fouriera? Jak bardzo jesteśmy pewni, że sieć nie halucynuje (zjawisko znane np. z modeli językowych), tzn. nie generuje sztucznych, ale błędnych bo w rzeczywistości nieistniejących cech?
21. Tabele 6.1 – 6.3 są trudne do analizy choćby ze względu na to, że lepiej działające metody dla niektórych miar oznaczają małe wartości, a inne wręcz przeciwnie – duże wartości. Być może oznaczenie kolejności (ang. *rank*) pomiędzy metodami byłoby lepszym rozwiązaniem?
22. Rozdział (6.2.1) – co dokładnie oznacza słowo „performance" w tym kontekście? Jakość odpowiedzi, czy szybkość odpowiedzi, czy też obydwa te czynniki?
23. Testy statystyczne opisane w Rozdziale (6.2.3) i przedstawione na Rysunku 6.11 oraz 6.12 wymagają pogłębionej analizy (szczególnie przypadek TR-MISR z cLPIPS).

Powyższe uwagi wymagają pewnej analizy, czy też ustosunkowania się do nich Doktoranta. Być może będą stanowiły rozszerzenia pewnych zagadnień naukowych poruszonych w rozprawie i przyczynią się do kontynuacji badań nad nowoczesnymi metodami SR. W żadnym stopniu nie umniejszają jednak wysokich walorów naukowych recenzowanej rozprawy.

7. *Jaka jest przydatność rozprawy dla nauk technicznych?*

Pan magister inżynier Tomasz Tarasiewicz zaprezentował nowatorskie rozwiązania do bardzo istotnego problemu poprawy jakości rozdzielczości zarówno pojedynczych, jak i serii obrazów. Metody te bazują na nowoczesnym podejściu wykorzystującym głębokie neuronowe struktury grafowe. Otrzymane wyniki świadczą o bardzo dużej skuteczności opracowanych metod. Dlatego uważam, że są one bardzo przydatne dla nauk technicznych, o czym świadczą między innymi liczne publikacje Doktoranta wraz z zespołem Pana Profesora Michała Kawuloka w uznanych periodykach naukowych. Metody te stanowią i będą stanowić źródło odniesienia dla innych naukowców pracujących nad metodami SR, a szczególnie takich które również będą wykorzystywały głębokie sieci grafowe. Co więcej, zaprezentowane metody mają olbrzymi potencjał aplikacyjny, gdyż problem poprawy rozdzielczości

obrazów spotykany jest w wielu praktycznych zadaniach przetwarzania obrazów, o czym była już mowa.

Wnioski końcowe

Postawione przez Pana magistra inżyniera Tomasza Tarasiewicza tezy naukowe zostały w pełni wykazane. Opracowane przez niego metody są oryginalne i na najwyższym poziomie światowym. Stanowią one oryginalny własny wkład Doktoranta Pana magistra inżyniera Tomasza Tarasiewicza w rozwój dyscypliny naukowej Informatyka Techniczna i Telekomunikacja. Przy ich realizacji Pan magister inżynier Tomasz Tarasiewicz wykazał się dogłębną znajomością najnowszych metod i algorytmów przetwarzania obrazów wielospektralnych za pomocą różnorodnych architektur głębokich sieci neuronowych, w tym sieci opartych na grafach, jak również dobrym opanowaniem warsztatu badawczego, a w rezultacie również wysoką dojrzałością naukową.

Recenzowaną pracę oceniam jako spełniającą z nadmiarem wymagania stawiane rozprawom doktorskim. Wnioskuję o jej przyjęcie oraz o dopuszczenie Pana magistra inżyniera Tomasza Tarasiewicza do publicznej obrony.

Ponadto, ze względu na nowatorski charakter opracowanych metod, świetne wyniki eksperymentalne, jak również liczne publikacje naukowe, w tym w czołowych czasopismach światowych, przedkładam wniosek o **wyróżnienie rozprawy** Pana magistra inżyniera Tomasza Tarasiewicza.

Bogusław Cyganek