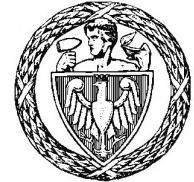


Politechnika Warszawska

W Y D Z I A Ł E L E K T R Y C Z N Y



Instytut Sterowania i Elektroniki Przemysłowej

Praca dyplomowa magisterska

na kierunku Informatyka Stosowana

w specjalności Inżynieria danych

Transfer stylu z wykorzystaniem sieci głębokiego uczenia

mgr. inż. Michał Bogacz

numer albumu 279086

promotor

prof. dr hab. inż. Marcin Iwanowski

WARSZAWA 2021

Transfer stylu z wykorzystaniem sieci głębokiego uczenia

Streszczenie

Celem pracy jest analiza oraz badania metod transferu stylu malarskiego ze zdjęcia obrazu na zdjęcie na przykładzie dzieł polskich artystów. Praca skupia się na rozwiązaniu dwóch podstawowych problemów przeniesienia stylu. Pierwszym jest zaproponowanie i zimplementowanie nowej metody pozwalającej połączenie i przeniesienie stylu wielu obrazów źródłowych na jeden docelowy, metoda opisana jako "wiele-do-jednego". Drugim rozwiązany problemem jest poprawa zaburzeń kolorów które często pojawiają się przy przenoszeniu stylu między obrazami. W pracy zaproponowano serię metod które rozwiązują oba te problemy i przedstawiono wyniki działania metod. Przedstawione rozwiązania bazują na standardowych metodach poprawy działania struktur sieci neuronowych w zagadnieniach wizji komputerowej. Pierwsza procedura oparta została na augmentacji użytych obrazów, druga na wzmacnieniu możliwości generalizacyjnych algorytmu, natomiast ostatnia na uproszczeniu problemu do przeniesienia stylu czarno-białego i wykonaniu sztucznego kolorowania wynikowego obrazu. Wykazano równocześnie, że zaproponowane metody poprawiają jakość transferu stylu względem metody przedstawionej jako bazowa.

Słowa kluczowe: transfer stylu, splotowe sieci neuronowe, uczenie głębokie

Style transfer algorithms

Abstract

The purpose of this thesis is to analyze and research the methods of transferring the painting style from a picture to a picture on the example of works by Polish artists. The work focuses on solving two basic problems of style transfer. The first is to propose and implement a new method that allows you to combine and transfer the style of multiple source images to one target, a method described as "many-to-one". The second problem solved is to correct the color distortions that often occur when style is transferred between images. The paper proposes a series of methods that solve both these problems and presents the results of the methods. The presented solutions are based on standard methodologies for improving the operation of neural network structures in terms of computer vision. The first procedure was based on augmentation of the images used, the second on enhancing the generalization capabilities of the algorithm, and the last on simplifying the problem of transferring the black and white style and performing artificial colorization of the resulting image. It was also shown that the proposed methods improve the quality of style transfer in relation to the method presented as the baseline.

Keywords: style transfer, convolutional neural networks, deep learning



.....
miejscowość i data
mgr. inż. Michał Bogacz....

.....
imię i nazwisko studenta
279086.....

.....
numer albumu
Informatyka Stosowana.....

kierunek studiów

OŚWIADCZENIE

Świadomy/-a odpowiedzialności karnej za składanie fałszywych zeznań oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie, pod opieką kierującego pracą dyplomową.

Jednocześnie oświadczam, że:

- niniejsza praca dyplomowa nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994 roku o prawie autorskim i prawach pokrewnych (Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.) oraz dóbr osobistych chronionych prawem cywilnym,
- niniejsza praca dyplomowa nie zawiera danych i informacji, które uzyskałem/-am w sposób niedozwolony,
- niniejsza praca dyplomowa nie była wcześniej podstawą żadnej innej urzędowej procedury związanej z nadawaniem dyplomów lub tytułów zawodowych,
- wszystkie informacje umieszczone w niniejszej pracy, uzyskane ze źródeł pisanych i elektronicznych, zostały udokumentowane w wykazie literatury odpowiednimi odnośnikami,
- znam regulacje prawne Politechniki Warszawskiej w sprawie zarządzania prawami autorskimi i prawami pokrewnymi, prawami własności przemysłowej oraz zasadami komercjalizacji.

Oświadczam, że treść pracy dyplomowej w wersji drukowanej, treść pracy dyplomowej zawartej na nośniku elektronicznym (płycie kompaktowej) oraz treść pracy dyplomowej w module APD systemu USOS są identyczne.

.....
czytelny podpis studenta

Spis treści

1 Wstęp	11
1.1 Wstęp	11
1.2 Cel i zakres pracy	12
1.3 Zawartość pracy	13
2 Metodyka prowadzenia projektu	15
2.1 Zbiory obrazów	15
2.2 Metryki jakości	17
2.3 Użyte środowisko pracy	18
3 Wariant „jeden do jednego”	23
3.1 Opis metody	23
3.1.1 Funkcja jakości wzoru	24
3.1.2 Macierz Grama	25
3.1.3 Funkcja jakości stylu	26
3.1.4 Procedura transferu	26
3.2 Implementacja metody	27
3.3 Wyniki	30
4 Wariant „wiele do jednego”	39
4.1 Transfer wiele do jednego	39
4.2 Opis metody augmentacji obrazu stylu	41
4.3 Wyniki	44
4.4 Opis metody generalizacji reprezentacji stylu	45
4.5 Wyniki	46
5 Wariant ze sztucznym kolorowaniem	55
5.1 Transfer bez koloru	55
5.2 Wyniki	57
5.3 Sztuczne kolorowanie obrazu	57

5.4 Wyniki	58
6 Podsumowanie	65
Bibliografia	67
Spis rysunków	69

Rozdział 1

Wstęp

1.1 Wstęp

Ludzkość jest od wielu tysiącleci silnie związana ze sztuką. Pomimo niezwykle szerokiej i nieprecyzyjnej definicji tego co uważane jest za sztukę, miała ona ogromny wpływ na pojmowanie rzeczywistości przez człowieka. Percepcja kolejnych generacji dzieł artystycznych zmieniał się przez kolejne lata, a sama sztuka staje się coraz bardziej abstrakcyjna i nie ma na celu zaledwie odzwierciedlania prostej rzeczywistości. Sztuka stała się silnym narzędziem wpływającym na decyzje dużych grup ludzi, była używana jako narzędzie propagandy i przekazu opinii politycznych. Niezależnie od tego co sądzimy o sztuce i jak podchodzimy do jej analizy, wpływa ona niezwykle na nasze życie. To właśnie ta fascynacja sztuką, jak również próba udostępnienia i upowszechnienia narzędzi którymi dotąd umieli posługiwać się wyłącznie artyści była siłą napędową dla rozwoju zagadnień opisanych w tej pracy. Sztuka może objawiać się przez wiele różnych mediów, takich jak obraz, muzyka czy rzeźba. Jednak w tej pracy autor skupi się wyłącznie na analizie metod odtwarzania stylu malowanych obrazów.

Za prekursorów zagadnienia transferu stylu malarstwego można uznać dwa równolegle rozwijające się pomysły. Pierwszy z nich polegał na stworzeniu metody która umożliwiłaby teksturowanie dowolnej powierzchni przy użyciu odpowiednio małych wycinków oryginalnego zdjęcia. Metoda ta została szerzej opisana już na przełomie 2000-nego roku w artykule [4][17]. Drugi natomiast [9] przypominał w swoich założeniach już znacznie bardziej zagadnienie z którym mamy do czynienia w tej pracy. Autorzy artykułu zaproponowali rozwinięcie metody teksturowania przeprowadzone w taki sposób, aby używając odpowiednich filtrów stworzyć relatywnie wierne odwzorowanie charakterystycznych struktur obrazu, takich jak rozmycia czy kształt pociągnięć pędzla.

Zagadnienie Transferu Stylu Malarskiego jest zagadnieniem pochodnym od tych opisanych powyżej. Głównym celem procesu transferu stylu jest stworzenie nowego dzieła sztuki

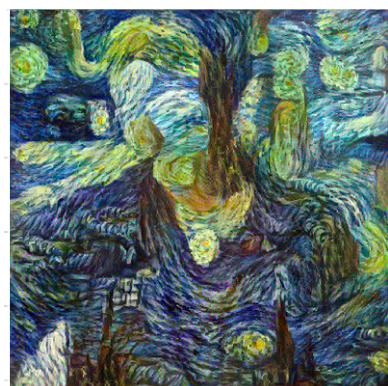
bazującego na cechach dwóch obrazów wejściowych, obrazu wzorcowego oraz obrazu stylu. Ważną cechą prawidłowego przeprowadzenia transferu jest pozostawienie najważniejszych cech wzorca nienaruszonymi przy równoczesnym przeniesieniu jak największej ilości cech stylowych. Ogromnym przełomem który pozwolił na znaczne poprawienie jakości wyników było zastosowanie głębokich konwolucyjnych sieci neuronowych, zaproponowane przez Leona Gatysa et al. w artykule "*A Neural Algorithm of Artistic Style*" [5]. Dzięki użyciu wysoko-poziomowych cech pobranych z różnych warstw pre-trenowanej sieci neuronowej możliwym stało się łatwe zautomatyzowanie doboru odpowiednich najważniejszych cech definiujących zarówno styl jak i wzorzec obu obrazów. Przykład takiego transferu został pokazany na bardzo popularnym w tym zagadnieniu obrazie "Gwieździsta noc" autorstwa Vincent-a van Gogh-a Rys. 1.



(a) Olga Bogacz



(b) "Gwieździsta noc"



(c) Zdjęcie stylizowane

Rysunek 1. Przykładowy wynik transferu stylu z wykorzystaniem metody opisanej w rozdziale 3.

1.2 Cel i zakres pracy

Rozwój głębokich sieci neuronowych oraz coraz zwiększone możliwości sprzętowe kart graficznych pozwoliły na zastosowanie dużych architektur głębokiego uczenia do wielu nigdy wcześniej nie eksplorowanych zastosowań. Jednym z takich zastosowań było wydobywanie głęboko ukrytych cech z obrazów i używanie tych cech w procesie transferu tekstur między obrazami. Przełomową pracą która opisała dokładnie procedurę ekstrakcji cech stylu z jednego obrazu i przeniesieniu tych właśnie cech na inny obraz była wspomniana w poprzedniej sekcji praca Leona Gatysa [5]. Opisano w niej bardzo dokładnie wyniki oraz procedurę wykonywania transferu który można określić jako transfer *jeden-do-jednego*. Styl pojedynczego zdjęcia wybranego jako zdjęcie stylowe jest przenoszony na jedno wybrane zdjęcie wzorca. Pomimo, że metoda skonstruowana w taki sposób jest bardzo efektowna, a otrzymane wyniki są zdecydowanie bardzo ciekawe, obciążona jest dwoma dużymi problemami.

Pierwszy z tych problemów jest związany wyłącznie z jakością otrzymanych wyników. Jak zostanie to zaprezentowane w dalszej części pracy, wyniki obciążone są pewnymi błędami. Problemami z odwzorowaniem prawidłowych kolorów oraz pojawiającymi się niechcianymi artefaktami transferu.

Druga łatwo zauważalna wada oryginalnego algorytmu to singularny charakter transformacji, gdzie pobieramy cechy tylko z jednego zdjęcia. Bardzo często jeden obraz źródłowy dla stylu nie jest wystarczający, aby zapewnić odpowiednio generalną reprezentację stylu. Na zadanym obrazie mogą nie występować konkretne cechy przedmiotów czy osób, które łatwo utożsamiamy z tworzącym je artystą. Każdy twórca ma bardzo specyficzny sposób przedstawiania różnych elementów rzeczywistości i warto wziąć pod uwagę wiele różnych możliwych źródeł stylu, aby wiarygodnie przedstawić ideę artysty na nowym obrazie.

Cele poniższej pracy można przedstawić w dwóch kategoriach. Poprawienie działania podstawowej metody *jeden-do-jednego* przez zmniejszenie ilości pojawiających się artefaktów oraz zmniejszenie zaszumienia kolorów na obrazach wynikowych. Zaproponowanie działającej metody zmieniającej zagadnienie na transfer stylu *wiele-do-jednego*. Nowo zdefiniowana forma zagadnienia ma na celu umożliwienie generalizacji stylu pochodzącego z wielu obrazów i aplikowaniu tej uogólnionej reprezentacji na obrazie wzorcowym.

Realizacja pracy wymaga zaimplementowania metody podstawowej, umożliwiającej porównanie i ocenę otrzymanych wyników. Następnie musi zostać zdefiniowana metoda *wiele-do-jednego*, na której bazować będą wszystkie kolejne przedstawione rozwiązania. Przedstawione muszą zostać procedury augmentacji pojedynczego zdjęcia stylowego, generalizacja stylu malarza oraz metoda sztucznego przenoszenia oryginalnych kolorów obrazu wzorcowego.

1.3 Zawartość pracy

Praca składa się z sześciu rozdziałów.

Rozdział drugi poświęcony zostanie przedstawieniu metodyki prowadzenia projektu. Po przedstawieniu podstawowych informacji o użytym środowisku pracy, pokazane zostaną metryki pomiaru jakości wykonanego transferu. Użyte metryki będą podzielone na te użyte w procesie treningu oraz na takie które używane są przy porównywaniu wyników metod. Ostatni podrozdział poświęcony zostanie opisowi wybranych zbiorów zdjęć i obrazów na których bazować wyniki działania zaproponowanych metod.

Rozdział trzeci posłuży opisowi metody podstawowej (ang. *baseline*) na bazie której przeprowadzony zostanie opis wszystkich elementów algorytmu. Wyniki pokazane w tym rozdziale użyte zostaną do przeprowadzenia ostatecznej analizy porównawczej i określeniu,

Rozdział 1. Wstęp

czy zaproponowane algorytmy spełniają postawione im założenia oraz czy udało się dzięki nim poprawić działanie algorytmu.

Następny, czwarty rozdział zawierać będzie przedstawianie procedury *wiele-do-jednego*, na bazie której wyprowadzone będą dwie metody pochodne. Pierwsza metoda będzie bezpośrednim ulepszeniem metody podstawowej, gdzie przy pomocy zestawu zaproponowanych augmentacji obrazu stylowego poprawiona zostanie jakość wykonanego transferu. Druga część tego rozdziału zostanie poświęcona przedstawieniu metody generalizacji cech stylu zawartych w zestawie obrazów i przeprowadzeniu procesu transferu stylu dla tej nowej reprezentacji.

Piąty rozdział poświęcony zostanie na przeprowadzenie kolejnej iteracji ulepszeń rozwiązań przedstawionych w poprzedzających rozdziałach. Jak zostanie pokazane w dalszej części pracy, wspólnym problemem wszystkich metod przedstawionych w rozdziałach 3 oraz 4 jest niewłaściwa reprezentacja kolorów na obrazach. Większa część stylu malarskiego jest zawarta w warstwie jasności obrazu, dlatego w tym rozdziale zaproponowana zostanie metoda przeniesienia oryginalnych kolorów na wystylizowany czarno-biały obraz.

Ostatnia część pracy składa się z całościowego wybiórczego porównania wyników wszystkich przedstawionych metod oraz podsumowania całości wykonanych prac. Pracę kończy bibliografia oraz spis rysunków.

Rozdział 2

Metodyka prowadzenia projektu

2.1 Zbiory obrazów

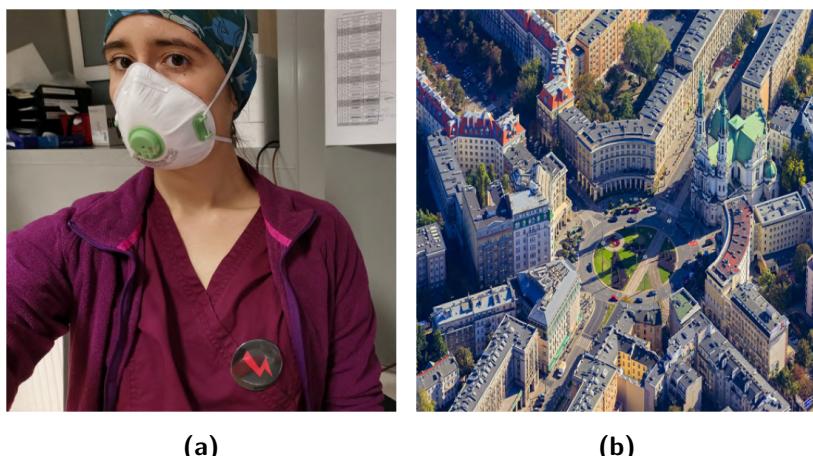
Wszystkie użyte w eksperymetach zdjęcia należały do jednego z dwóch gatunków obrazów. Zarówno zdjęcia wzorcowe jak i zdjęcia stylu są portretami lub pejzażami. Te dwa nurty w sztuce są klasyczne i bardzo często używane, dlatego zdecydowano się właśnie na zastosowanie tych obrazów. Równocześnie gatunki portretu i pejzażu są od siebie zupełnie różne pod względem zawartości obrazów. Wykorzystanie zupełnie różnych typów obiektów na obrazie stylu i wzorca powinno pozwolić na sprawdzenie czy metoda rzeczywiście rozróżnia te dwie cechy od siebie. Przeniesienie stylu z namalowanego budynku na rzeczywistą osobę, dla idealnej metody, nie powinno sprawić żadnego problemu i wyniki powinny pozostać naturalne.

Pozycje używane jako obrazy wzorcowe, Rys. 2, są klasycznie wykonanymi zdjęciami cyfrowymi, reprezentacjami opisanych powyżej dwóch gatunków. Wybrano je tak, aby były zgodne pod względem zawartości z opowiadającymi im wersjami obrazów stylu. Taki dobór pozwoli zbadać czy zawartości zdjęć wpływają na jakość transferu samego stylu.

Obrazy stylu, tak jak obrazy wzoru, zostały dobrane w taki sposób aby reprezentować szeroką gamę stylów malarstw oraz założone wcześniej typy gatunków malarstw.

Akwarele malowane przez Tytusa Brzozowskiego przedstawione na Rys. 3, są jednym z przykładów użytych pejzaży. Są z reguły bardzo jasne, a autor w większości wykorzystuje przy ich tworzeniu dużo pastelowych barw. Natura wykorzystanego medium malarstw wymusza tworzenie obrazu o dużych jednolitych plamach koloru, które łagodnie przechodzą jedne w drugie. Równocześnie na obrazach nie ma zaznaczonych mocno konturów, a używany światłocień jest bardzo naturalny. Warto jednak zauważyć, że autor czasem wykorzystuje bardzo kontrastujące kolory, szczególnie do tworzenia silnego światłocienia.

Wybrane obrazy Edwarda Dwurnika, przedstawione na Rys. 4, to częściowo ikonografie, a częściowo obrazy olejne. Można zaobserwować pewne niewielkie różnice między obrazami,



(a)

(b)

Rysunek 2. Zdjęcia używane jako zdjęcia wzorcowe w dalszej części pracy: (a) Olga Bogacz - siostra autora pracy, (b) "Warszawski Plac Zbawiciela z lotu ptaka" [16].

a ikonogramiemi, te drugie są mniej jaskrawe i wydają się nieco wyblakłe. Jednak pomijając niewielkie różnice w nasyceniu kolorów wiele z pozostałych cech stylu jest stałych, wszystkie obrazy są wykonane stylem przypominającym impresjonizm. Duża wariancja występująca w używanych kolorach oraz reprezentacja płaszczyzn przy pomocy wielu małych pociągnięć pędzla, gdzie jednolita barwa powstaje dopiero po spojrzeniu na obraz z większej odległości. Duża ilość kolorów i ich niejaka losowość pomaga w zasłonięciu niewielkich niedokładności. Taki charakter obrazów z reguły lepiej sprawdza się w metodach transferu stylu, niż obrazy złożone z jednolitych plam.

Prace Tamary Łempickiej, Rys. 5, reprezentują drugi z używanych gatunków malarskich. Autorka tych olejnych obrazów jest rozpoznawalna na całym świecie, a jej niepowtarzalny sposób przedstawiania rzeczywistości nie może być pomylony z żadnym innym. Lekka geometryzacja postaci człowieka na dziełach jest bardzo ważną cechą prezentowanego stylu. Równocześnie autorka używa wysokiego kontrastu między kolorami, a światłocień bardziej przypomina kontury postaci niż rzeczywiste cienie. Styl przedstawiony na Rys. 5 jest równocześnie najmniej oparty na konkretnych kolorach, mogłoby się wydawać, że dowolne zastosowanie kolorów na wszystkich obrazach nie zmieniłoby wcale odbioru stylu przez osobę oglądającą. Nie jest to prawdą za obrazów przedstawionych na Rys. 3 i Rys. 4, na tych pracach zmiana koloru wiążałaby się ze nieodwracalną utratą ważnej informacji, możliwym jest, że tak nowo zbudowany styl byłby nadal interesujący, ale w opinii autora może nie być już bliźniaczy do oryginalnego.

Zbiór treningowy, zarówno zdjęć wzoru jak i stylu, mógłby być nieskończonym powiększaniem o kolejne pozycje. Dowolne wykonane zdjęcie mogłoby stać się nowym zdjęciem wzorcowym, a ilość różnych artystów, znanych jak i niszowych jest ogromna. Przygotowany i przedstawiony w

tym rozdziale zbiór danych wydaje się jednak być złotym środkiem, pozwala na przetestowanie wszystkich przedstawionych hipotez oraz na zrealizowanie wszystkich celów. Niewielka liczba zdjęć pozwoli również na znacznie prostsze i dokładniejsze przedstawienie wyników wykonanych prac.

2.2 Metryki jakości

Istotnym elementem każdego zagadnienia badawczego jest określenie odpowiednich metryk porównawczych, które zostaną użyte do oceny jakości samego transferu. Dodatkowo w tym rozdziale przedstawione zostaną standardowe metryki używane w procesie uczenia maszynowego, jako optymalizowana metryka błędu. Są one arbitralne względem samej metody transferu stylu, dlatego wybór odpowiedniej zostanie opisany tutaj a nie w kolejnych rozdziałach.

W opisanych poniżej metrykach błędu przyjęto następujące oznaczenia zmiennych, gdzie:

- y_i - wartość wyznaczona w procesie stylizacji
- \hat{y}_i - oryginalna wielkość reprezentująca cechę stylu
- n - liczba porównywanych elementów

MAE (en. Mean Absolute Error) - podstawowa i jedna z najczęściej używanych metryk błędu. Niestety nie jest to norma idealna, nie karze modelu za dużą odległość badanych próbek od siebie. Ze względu na tę niedoskonałość nie będzie ona użyta w procesie przekształcania obrazów.

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (1)$$

MSE (en. Mean Square Error) - rozwinięcie metryki przedstawionej wcześniej, gdzie zamiast wartości bezwzględnej różnicy liczb brana jest pod uwagę różnica podniesiona do kwadratu. Dzięki takiej zmianie jesteśmy w stanie reagować silniej na wartości znacznie odstające od oczekiwanych.

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (2)$$

Ze względu na znacznie lepsze wyniki otrzymane przy analizie wyników z użyciem obu metryk, zastosowana zostanie jedynie ta druga. Jest ona używana wewnątrz algorytmu przenoszenia tekstu między obrazami, jednak jej jednoznaczny wybór jest bardzo ważny z punktu widzenia poprawnie przeprowadzonych badań.

Porównanie ze sobą wyników pochodzących z różnych wersji algorytmów może zostać wykonane przy pomocy numerycznych funkcji jakości używanych w procesie przenoszenia stylu, jednak, takie podejście jest bardzo naiwne i nie gwarantuje prawidłowych wyników. Styl składa się z wielu niekwantyfikowalnych cech których nie można łatwo uchwycić przy

pomocy pojedynczej metryki błędu, szczególnie takiej którą można przenieść pomiędzy obrazami o całkowicie różnych stylach. Równocześnie często pojawiającym się problemem jest przetrenowanie transformowanie obrazu, przez co zamienia się on w szum który w żaden sposób nie przypomina ani obrazu wzorcowego ani stylu, jednak z punktu widzenia metryki jest lepszy od wcześniejszych wersji obrazu wynikowego. Dokładne przykłady takich problemów zostaną opisane w kolejnych rozdziałach, jednak ważne jest zdefiniowanie sposobu oceny jakości transferu jeszcze przed rozpoczęciem jakichkolwiek testów. Przekształcenia będą na potrzeby tej pracy oceniane słownie przez autora, subiektywna opinia będzie dokonywana na podstawie trzech najważniejszych aspektów stylu. Tekstury, czyli mikroskopowych kształtów z których składa się powstał obraz, form w jakie układają się pociągnięcia pędzla. Płaszczyznom, czyli dużym kształtom z jakich składają się całe obiekty na obrazach. Koloru oraz jego rozłożeniu na obrazie, kolory nie koniecznie muszą być takie same, aby reprezentować ten sam styl, często chodzi o łączenie ich w odpowiedni sposób, lub zachowanie przejść pomiędzy zmieniającymi się kolorami. Opisane wyżej różnice bardzo dobrze widać na Rys. 6, obrazy przedstawionych tam malarzy zostały wykonane zupełnie innymi technikami, przez co wszystkie z wymienionych cech są tam przeciwwstawne. Styl E. Dwurnika przypomina styl impresjonistyczny, nie ma w nim dużych płaszczyzn jednolitego koloru, natomiast same pociągnięcia pędzla są niewielkie. Natomiast T. Brzozowski maluje przy użyciu nowoczesnego stylu z użyciem akwareli, przez co na jego obrazie często widoczne są duże płaszczyzny jednolitych barw, a ślady pędzla są prawie nieistniejące. Obrazem wzorcowym natomiast jest zdjęcie tej samej lokacji, co oba pozostałe obrazy, na tym zdjęciu krawędzie nie są rozmyte, a równocześnie mamy do czynienia z dużymi różnicami w jasności i kontraste.

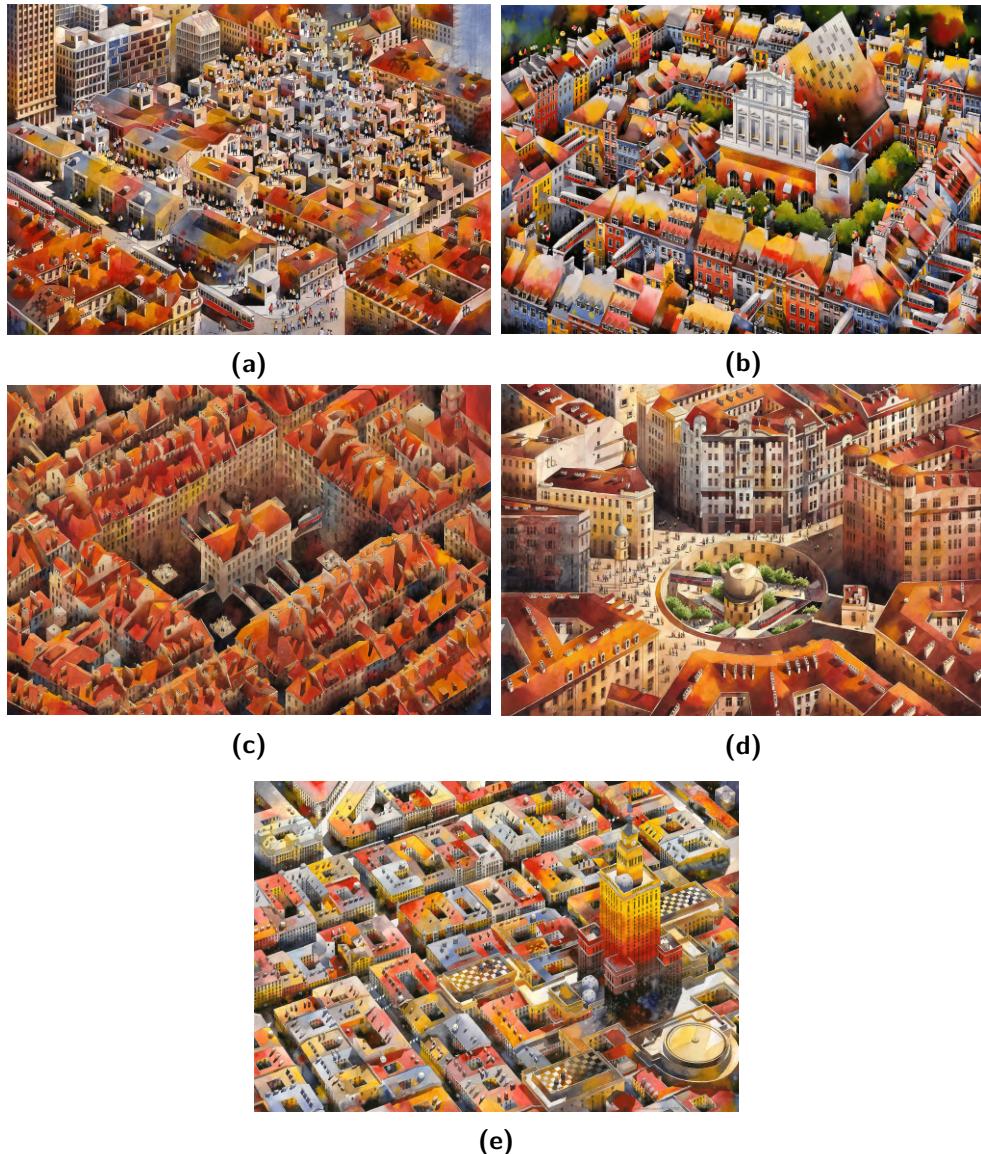
2.3 Użyte środowisko pracy

Wszystkie skrypty oraz programy użyte w trakcie przeprowadzonych badań i testów zostały napisane przy pomocy języka programistycznego Python. Język ten jest niezwykle często wykorzystywany do rozwiązywania problemów z obszaru nauk ścisłych i technicznych, szczególnie związanych z analizą danych. Szeroka gama gotowych bibliotek oraz modułów jest zdecydowanie zaletą tego środowiska i pozwala na znaczne przyśpieszenie pracy nad rozwiązywanym zagadnieniem. Dodatkowo większość artykułów i prac które można uważać za najlepsze (ang. *state of the art*) implementowanych jest również w języku Python, dzięki czemu łatwo znaleźć rozwiązania referencyjne.

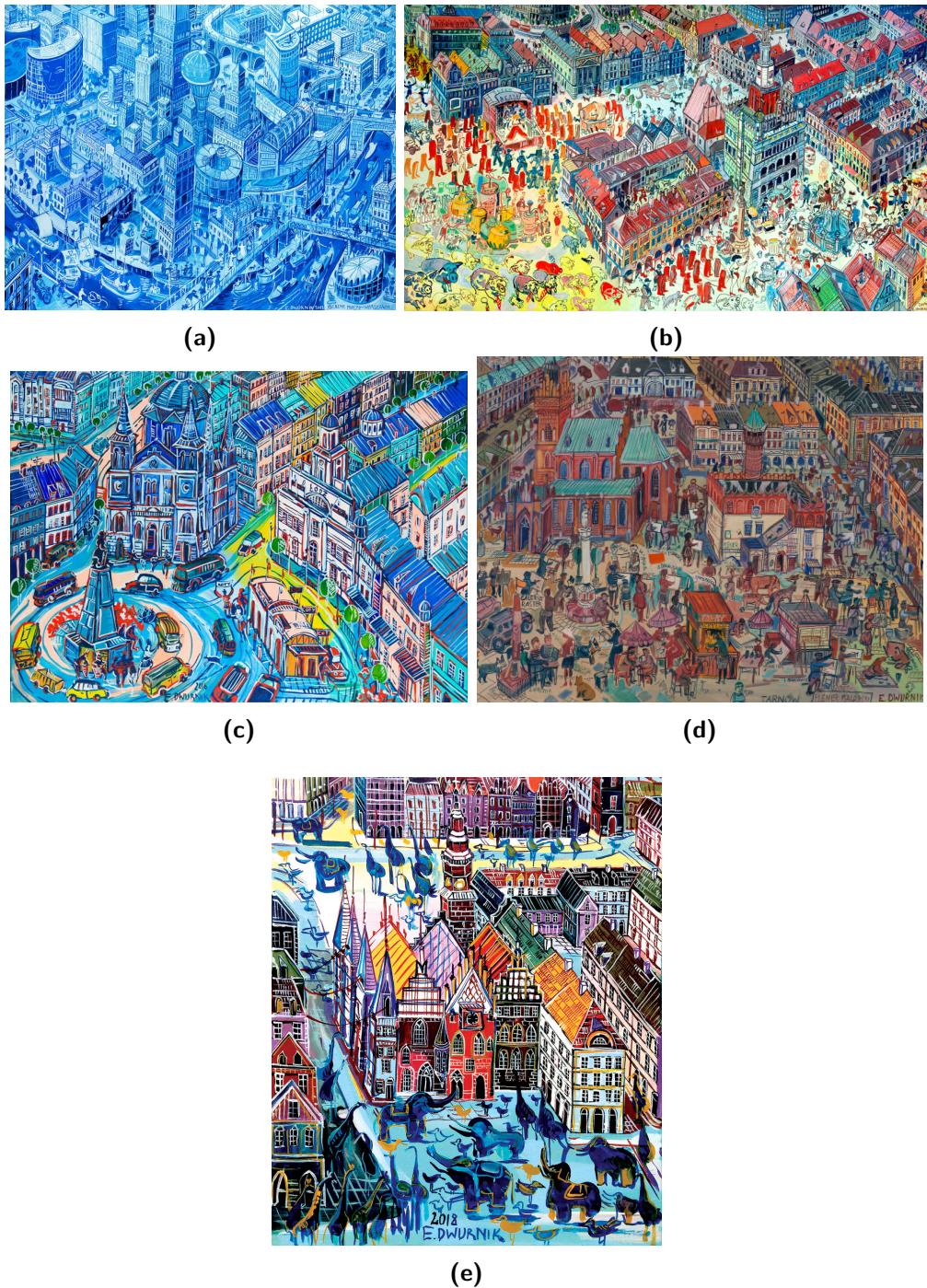
Alternatywnymi rozwiązaniami które mogłyby zastąpić język Python są środowisko Matlab oraz język programistyczny Julia. Oba te rozwiązania są często używane w zagadnieniach związanych z analizą danych. Niestety Julia ma mniejsze wsparcie dla bibliotek ułatwiających

analizę danych obrazowych, co mogłoby znacznie utrudnić przeprowadzenie badań. Natomiast Matlab jest świetnie wyposażony do prowadzenia tego typu eksperymentów, natomiast jest mniej popularny na aktualnym rynku pracy, dlatego ostateczny wybór padł na język Python.

Całkowite przeprowadzenie badań wymaga dodatkowo odpowiednich bibliotek, które pozwolą na implementację samej sieci neuronowej oraz odpowiednich modułów które pozwolą na ułatwienie manipulacji obrazami. Jako framework do nauczania maszynowego wybrano TensorFlow2.0 z doinstalowanymi sterownikami CUDA, które umożliwiają przeniesienia obliczeń sieci na kartę graficzną. Dla ułatwienia obsługi relatywnie skomplikowanego produktu jakim jest TensorFlow wykorzystano natywnie współpracujące API Keras. Przetwarzanie obrazów umożliwiły dwa moduły wykorzystywane zamiennie w zależności od potrzeb Pillow (skrót PIL) oraz OpenCV. PIL pozwala na przetwarzanie obrazów w postaci specjalnych obiektów natywnie zgodnych z oryginalnymi formatami plików wejściowych, natomiast OpenCV ułatwia manipulowanie obrazami w postaci załadowanych do pamięci wielowymiarowych macierzy.



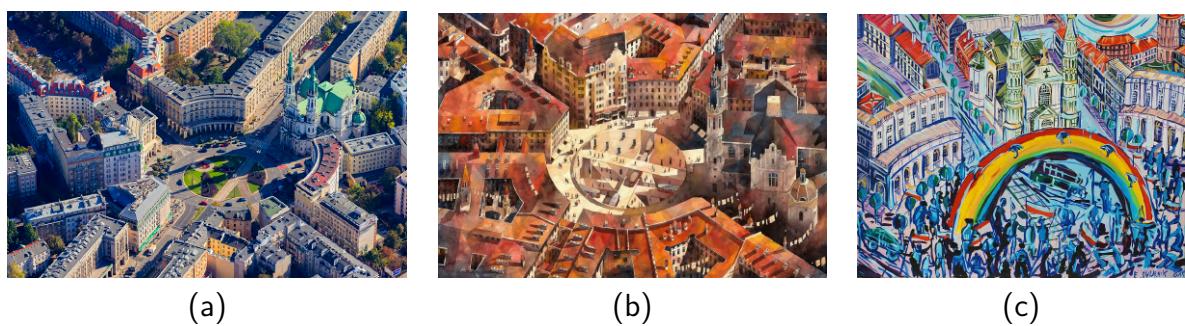
Rysunek 3. Obrazy Tytusa Brzozowskiego [1](używane w pracy jako źródła stylu): (a) "Norblin", (b) "Archikatedra w zieleni", (c) "Rynek", (d) "Plac Unii Lubelskiej", (e) "Plac Defilad"



Rysunek 4. Obrazy Edwarda Dwurnika [3](używane w pracy jako źródła stylu): (a) "Idealne miasto", (b) "Poznań 2", (c) "Pl. Zbawiciela", (d) "Tarnów", (e) "Wrocław".



Rysunek 5. Obrazy Tamary Łempickiej [11](używane w pracy jako źródła stylu): (a) "Autoportret w Zielonym Bugatti", (b) "Dziewczęta", (c) "Portret Dr. Boucard", (d) "Les Deux Amies", (e) "Dziewczyna z Mandoliną".



Rysunek 6. Zdjęcia i obrazy reprezentujące skrajnie różne cechy stylu: (a) rzeczywiste zdjęcie, (b) "Plac Zbawiciela" T. Brzozowski[1], (c) "Pl. Zbawiciela" E. Dwurnik[3]

Rozdział 3

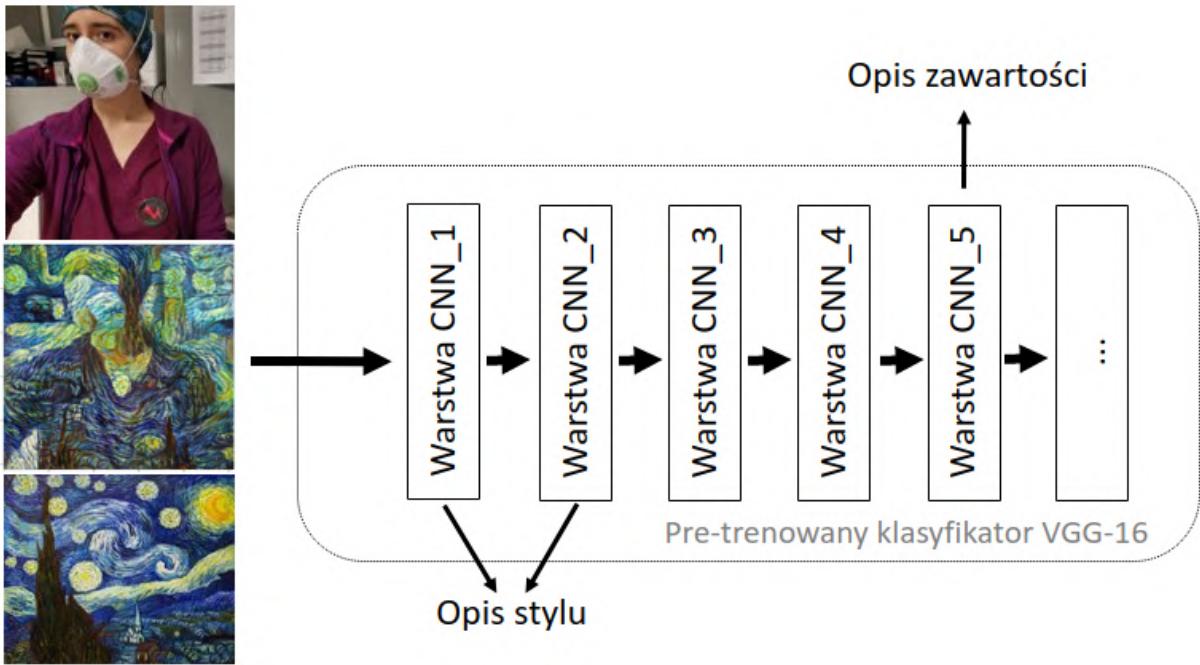
Wariant „jeden do jednego”

3.1 Opis metody

Pierwsze użycie konwolucyjnych sieci neuronowych w zagadnieniu transferu stylu zostało wykonane przez Leona Gatysa [5] na przełomie lat 2015 i 2016. Oryginalny algorytm został zbudowany na bazie sieci VGG-16, w swoim czasie jednym z najlepiej działających głębokich klasyfikatorów, użyta sieć była trenowana na zbiorze danych ImageNet. Procedura transferu stylu lub tekstury jest procesem optymalizacji który przeprowadzany jest na trzech obrazach, stylu, wzorcu i obrazie transformowanym. Iteracyjnie zmniejszana jest sumaryczna różnica między stylem oryginalnym, a stylem zdjęcia transformowanym i między zawartością oryginalnego wzorca, a zawartością zdjęcia transformowanego. Obie te wielkości zdefiniowane są przy pomocy map cech pobranych z odpowiednich warstw głębokiej sieci CNN. Założeniem głównym metody jest możliwość uniezależnia od siebie wzorca i stylu, taka operacja wykonywana jest przez pobranie tych cech z różnych poziomów sieci bazowej (ang. *backbone*).

Użyta struktura VGG [14] jest typem szeregowej struktury, gdzie podobne zestawy warstw ukrytych występują w kolejnych blokach, a informacja przechodzi bezpośrednio z wyjścia jednego bloku do wejścia kolejnego. Taka struktura świetnie nadaje się do zagadnienia które realizowane jest w poniższej pracy, informacje wyciągnięte z różnych głębokości sieci nie są ze sobą bezpośrednio połączone i reprezentują inny typ informacji. Płytkie warstwy przedstawiają całe obiekty i duże ich wycinki, obiekty przedstawione są relatywnie realistycznie, głębokie warstwy skupiają się na abstrakcyjnych cechach obiektów. Proces pobierania danych przedstawiony jest na Rys. 7.

Pobrane elementy stylu i wzoru są łączone w ważonej funkcji jakości, gdzie \mathcal{L} jest stratą, odpowiednio całkowitą, wzoru oraz stylu. Współczynniki α oraz β decydują o ważności wzoru i stylu w procesie optymalizacji. Przedmiotami optymalizacji są trzy obrazy \vec{p} , \vec{s} oraz \vec{x} które odpowiadają kolejno zdjęciom: wzorcowemu, stylu i obrazowi w danej chwili poddawanemu



Rysunek 7. Ekstrakcja stylu i wzorca, na bazie których generowany jest opis cech obrazu w algorytmie Neural Style Transfer. VGG-19 to wersja sieci użyta we własnej implementacji autora, oryginalna praca [5] proponuje użycie sieci VGG-16. Pierwszy obraz od góry to obraz wzorcowy, drugi to aktualnie przetwarzana wersja obrazu, natomiast ostatni obraz to obraz źródłowy stylu. W procesie optymalizacji używane są odpowiednie cechy stylu i zawartości wszystkich trzech obrazów.

procesowi optymalizacji (transformowanemu). Składowe tej funkcji jakości zostaną dokładnie zdefiniowane w kolejnych podrozdziałach. Całkowita strata powinna reprezentować normę odległości wzorca oraz stylu zdjęcia transformowanego od odpowiednio zawartości zdjęcia wzorcowego i stylu zdjęcia stylowego.

$$\mathcal{L}_{\text{całkowite}}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{\text{wzoru}}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{\text{stylu}}(\vec{a}, \vec{x}). \quad (3)$$

3.1.1 Funkcja jakości wzoru

Zdefiniowanie funkcji straty wzoru zdjęcia jest prostsze niż zdefiniowanie funkcji jakości stylu. Zawartość zdjęcia pobierana jest tylko z jednej konkretnej warstwy konwolucyjnej, dzięki czemu nie ma potrzeby łączenia ze sobą wielu dwuwymiarowych map cech. Z reguły wybierane są warstwy leżące głębiej w strukturze sieci, ponieważ lepiej odwzorowują one zawartość zdjęcia.

Przyjmując strukturę sieci neuronowej jak ta opisana w poprzednim rozdziale, każdy obraz \vec{x} jest przetwarzany przez kolejne N_1 filtrów w N_1 map cech, gdzie każda taka mapa ma rozmiar M_1 . Więc odpowiedź konkretnej warstwy 1 możemy zapisać jako macierz $F_1 \in \mathbb{R}^{N_1 \times M_1}$, a wartość aktywacji i^{tego} filtra na j^{ej} w warstwie 1 jako F_{ij}^1 .

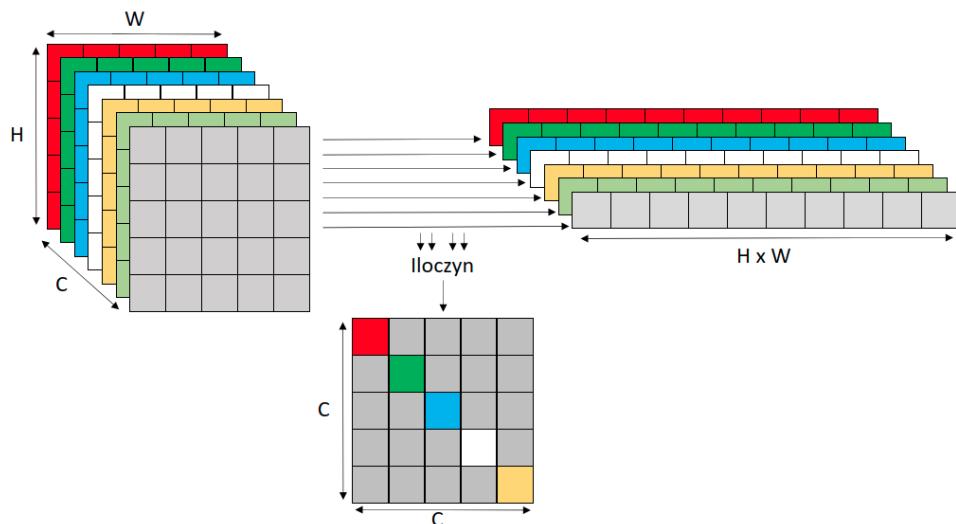
Jeśli zdjęcie transformowane oznaczone zostanie jako \vec{x} , a obraz wzorcowy jako \vec{p} . Natomiast wyżej opisane cechy wzoru pobrane z warstwy l jako F_l i P_l , to funkcję jakości przeniesienia zawartości można zapisać następującym równaniem,

$$\mathcal{L}_{\text{wzoru}}(\vec{p}, \vec{x}, l) = 1/2 * \sum_{i,j} (F_{i,j}^l - P_{i,j}^l)^2. \quad (4)$$

Gradient tak zapisanej funkcji może następnie zostać użyty w dowolnej procedurze optymalizacji, w celu upodobnienia tych konkretnych cech pomiędzy dwoma obrazami.

3.1.2 Macierz Grama

Ze względu na zaproponowany sposób reprezentacji stylu, przez połączenie ze sobą wyników pochodzących z kilku, najczęściej płytowych, warstw pojawiła się potrzeba wybrania metody łączenia ze sobą informacji pochodzących z kilku równoległych map cech. Rozwiązaniem zaproponowanym w oryginalnej pracy jest wykorzystanie Macierzy Gram'a [6]. Wykorzystanie tego matematycznego konceptu pozwala na połączenie filtrów, gdzie każdy z nich może zawierać inny typ informacji, takich jak kolor, kształt czy tekstura. Koncept wykonania operacji opisywanej jako wyliczenie Macierzy Grama został przedstawiony na Rys. 8. Jeżeli



Rysunek 8. Wyliczenie Macierzy Grama odpowiadającej zestawowi map cech. Każdy z elementów macierzy jest tak naprawdę współczynnikiem korelacji pomiędzy aktywacjami danego filtru.

tak utworzoną przestrzeń, czyli Macierz Grama oznaczmy jako $G_l \in \mathbb{R}^{N_l \times N_l}$, a odpowiednio spłaszczone i^{te} oraz j^{te} mapy cech pochodzące z warstwy l , to całą macierz można zapisać

jako:

$$G_{l,i,j} = \sum_k F_{i,k}^l F_{j,k}^l. \quad (5)$$

Dzięki zastosowaniu Macierzy Grama dodatkowo zmniejszamy wpływ informacji o przestrzennym położeniu piksela w danym filtrze. Jeśli zmienimy położenie pikseli w filtrze, wartości zapisane w przestrzeni Grama powinny pozostać takie same. Przestrzeń zapisuje statystyczną reprezentację stylu i nie powinna brać pod uwagę zawartości, dzięki czemu jest świetnym narzędziem do przedstawiania reprezentacji stylu.

3.1.3 Funkcja jakości stylu

Dzięki użyciu przestrzeni opisanej w poprzednim podrozdziale do połączenia cech pochodzących z wielu warstw możliwym jest wygenerowanie głębszej reprezentacji tekstury, kolorów i kształtów. Zakładając, że obraz zawierający oryginalny styl oznaczany będzie przez \vec{a} , natomiast obraz na który dany styl jest przenoszony oznaczony zostanie przez \vec{x} . Macierze Grama wygenerowane dla tych obrazów na warstwie l zostaną oznaczone odpowiednio jako A_l oraz G_l , wtedy stratę stylu dla cech wygenerowanych na tej konkretnej warstwie możemy zapisać jako:

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{i,j}^l - A_{i,j}^l)^2, \quad (6)$$

natomiast całkowitą różnicę stylu między takimi dwoma obrazami zapiszemy jako:

$$\mathcal{L}_{stylu}(\vec{a}, \vec{x}) = \sum_{l=0}^L E_l. \quad (7)$$

3.1.4 Procedura transferu

Poniżej przedstawiono schemat iteracyjnego przeprowadzenia transferu stylu Alg. 1, którego składowe zostały wyprowadzone w poprzednich podrozdziałach tej sekcji pracy. Przeprowadzenie algorytmu pod wieloma względami przypomina proces uczenia sieci neuronowej, w przypadku transferu używamy dwóch funkcji strat. Jedna związana jest ze stylem natomiast druga z zawartością obrazów. Dodatkowo jak można zauważyć poniżej w procedurze możemy zdefiniować dowolne zbiory warstw których użyjemy do rekonstrukcji stylu i wzoru. Dokładne porównanie różnych zestawów, ich zalet i wad zostało przedstawione w kolejnym rozdziale pracy. Po obliczeniu całkowitej metryki jakości można przeprowadzić proces propagacji wstecznej (eng. back-propagation) którego wyników używamy do przetransformowania obrazu wyjściowego w pewnym stopniu tak, aby bardziej przypominał dwa obrazy wejściowe, odpowiednio pod względem stylu i wzorca. Samą optymalizację pikseli wykonać można

dowolną techniką iteracyjnej optymalizacji, najlepiej sprawdzą się oczywiście metody z pakietów takich jak Keras lub PyTorch, ponieważ umożliwiają automatyczną propagację wsteczną błędu. Po wykonaniu n kroków optymalizacji, lub epok optymalizacji, można przerwać proces przekształcania i uzyskać obraz stylowo podobny do źródła stylu, a zawartością do źródła wzorcowego.

Algorytm 1: Algorytm optymalizacji Neural Style Transfer

Data: \vec{p} , \vec{a} - obraz wzorcowy, obraz stylu

Result: \vec{x}_{out} - obraz \vec{x} z przeniesionym stylem i zawartością

begin

```

/* pobranie danych cech wzoru i stylu */  

cechy_wzoru ← PobierzCechy( $\vec{p}$ )  

cechy_stylu ← PobierzCechy( $\vec{a}$ )  

/* Inicjalizacja zdjęcia transformowanego */  

 $\vec{x}_{\text{wyj}}$  ←  $\vec{p}$   

/* Proces iteracyjnej optymalizacji */  

for  $i \leftarrow 1$  to  $n_{\text{iteracji}}$  do  

    cechy_wyjsciowe ← PobierzCechy( $\vec{x}_{\text{wyj}}$ )  

     $\mathcal{L}_{\text{stylu}} \leftarrow 0$   

     $\mathcal{L}_{\text{wzoru}} \leftarrow 0$   

    for  $l \in \{\text{warstwa\_wzoru}, \text{warstwy\_stylu}\}$  do  

        if  $l \in \text{warstwa\_wzoru}$  then  

             $\mathcal{L}_{\text{wzoru}} \leftarrow \mathcal{L}_{\text{wzoru}} + \text{strata\_wzoru}(\text{cechy\_wyjsciowe}, \text{cechy\_wzoru})$   

        if  $l \in \text{warstwa\_stylu}$  then  

             $\mathcal{L}_{\text{stylu}} \leftarrow \mathcal{L}_{\text{stylu}} + \text{strata\_stylu}(\text{cechy\_wyjsciowe}, \text{cechy\_stylu})$   

         $\mathcal{L}_{\text{całkowite}} \leftarrow \frac{\alpha \mathcal{L}_{\text{wzoru}}}{n_{\text{warstw wzoru}}} + \frac{\beta \mathcal{L}_{\text{stylu}}}{n_{\text{warstw stylu}}}$   

        PropagacjaWsteczna( $\mathcal{L}_{\text{całkowite}}$ )  

    Zmien( $\vec{x}_{\text{wyj}}$ )  

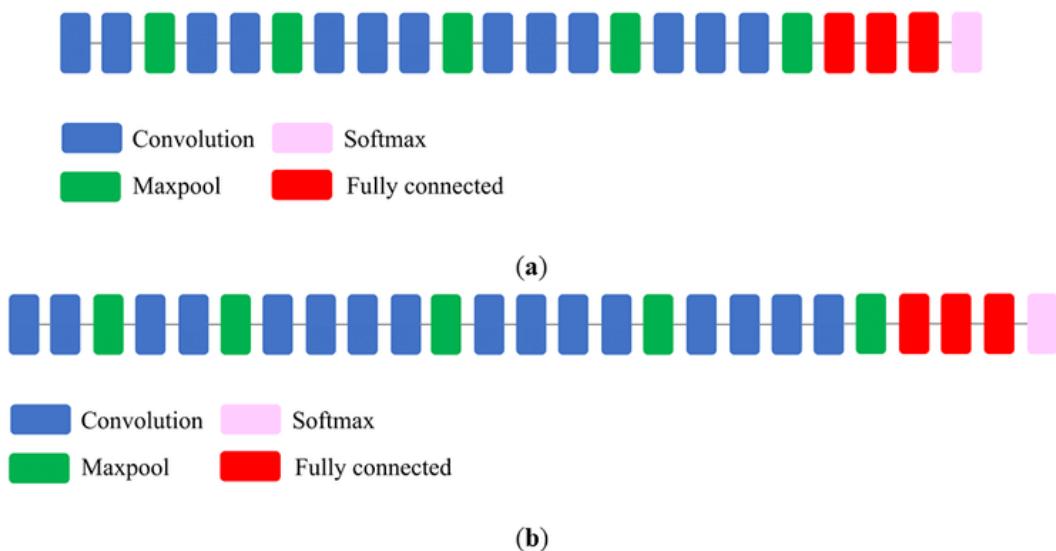
    Wynik( $\vec{x}_{\text{wyj}}$ )

```

3.2 Implementacja metody

Wszystkie elementy użytego algorytmu transferu stylu zgadzają się z tymi opisanymi z rozdziałem 3.1.4. Kolejne kroki oraz sposób pobierania danych są takie same jak te opisane w oryginalnym artykule [5] przytoczonym kilkukrotnie w tym rozdziale. Na bazie wstępnych doświadczeń wybrano jednak inną strukturę CNN od tej używanej w oryginalnej pracy.

Pierwszym wyborem odmiennym od tego który został wykonany w oryginalnym artykule [5] była użyta struktura głębszej sieci neuronowej. Leon Gatys zaproponował wykorzystanie sieci VGG-16, jest to mniejsza wersja sieci użytej przez autora tej pracy dyplomowej, VGG-19. Obie sieci mają podobną budowę, co przedstawiono na Rys. 9, jedyna różnica to ilość ukrytych warstw konwolucyjnych w odpowiednich blokach. Struktura VGG-19 została wybrana ze względu na możliwość wyciągnięcia większej ilości informacji z wybranego bloku. Równocześnie niewielkie spowolnienie przeprowadzania procesu optymalizacji nie jest problemem, więc większa elastyczność i możliwości są ważniejsze niż niewielka poprawa w działaniu. Obie sieci zostały przetestowane dla zestawu sprzętowego: Intel i9-9900K i Nvidia GTX 1080 Ti, średnie wyniki czasu przetworzenia obrazu zostały przedstawione w Tab. 1.



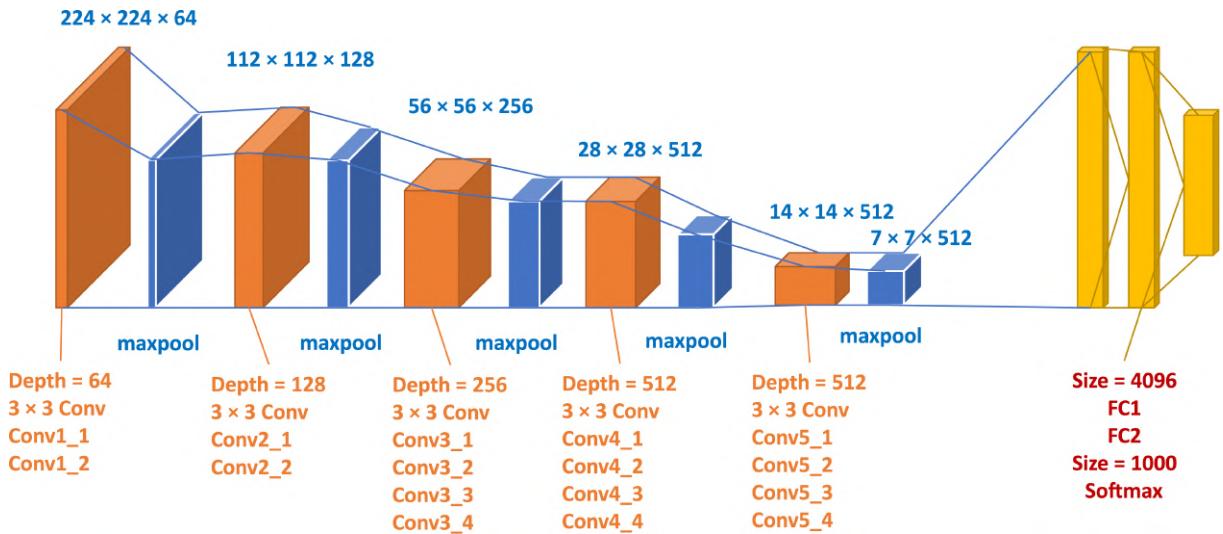
Rysunek 9. Porównanie struktur (a) VGG-19 i (b) VGG-16.

Kolejnym krokiem po wybraniu odpowiedniej struktury sieci, było dobranie odpowiednich warstw z których dokonywana jest ekstrakcja cech. Przyjmując oznaczenia takie jak na Rys. 10, autor pracy wzorcowej proponował użycie warstwy "Conv5_1" jako źródła dla informacji o wzorze, ta sama warstwa została użyta badanej implementacji.

Wybranie odpowiednich warstw z których można pobrać informacje dotyczące stylu musi zostać wykonane przez wykonanie odpowiednich prób i porównanie wyników. Ze względu na

Struktura	Średni czas dla n=100 prób [ms]
VGG-16	128.62
VGG-19	147.32

Tabela 1. Czasy przetworzenia obrazu przy pomocy badanych sieci.



Rysunek 10. Schemat używanej struktury głębokiej sieci neuronowej VGG-19 [15]. Najważniejsze są zaznaczone na pomarańczowo bloki konwolucyjne, to z nich pobierane są cechy stylu i zawartości. Żółta warstwa wyjścia została usunięta i nie jest używana.

iteracyjny proces transformacji ważnym było wybranie stałej ilości kroków optymalizacji, w celu określenia najlepiej działających warstw posłużyono się ilością kroków $n = 50$. Moment ten definiuje obraz wystarczająco przekształcony, ale równocześnie w większości ustawień nie pojawiają się na nim artefakty. Wyniki porównania zostały przedstawione na Rys. 11 oraz Rys. 12, na bazie tych wyników podjęto decyzję o użyciu wielu różnych bloków konwolucyjnych, gdzie z każdego wybrana została jedna konkretna warstwa aktywacji. Wybrane miejsca z których pobrano informacje to "Conv1_1", "Conv2_1", "Conv3_1" oraz "Conv4_1". Dla takiego zestawu uzyskano najlepszy transfer, połączono wiele cech pochodzących z różnej głębokości warstw, najprawdopodobniej różne typy reprezentacji stylu zostały skompresowane, co pozwoliło na wytworzenie efektywnej reprezentacji i uniknięto dużego zasumienia. Zdecydowanie mniejsze znaczenie ma odpowiedni dobór warstw w przypadku transferu stylu przeprowadzanego na pejzażach, mniejsze szczegóły i dużo więcej obiektów pozwala łatwiej zgubić niedoskonałości i zakłócenia.

Ostatnią zmianą wykonaną względem oryginalnej implementacji jest inne zdefiniowanie zdjęcia startowego. Leon Gatys proponował użycia trójwymiarowej ramki białego szumu, o wymiarach zdjęcia, jako punktu startowego dla algorytmu. Natomiast po pierwszych próbach okazało się, że znacznie lepsze wyniki można uzyskać stosując oryginalne zdjęcie wzorcowe jako punkt wyjściowy. Zabieg taki zmniejszał ilość pojawiających się artefaktów i znacznie polepszał ogólny wygląd przetwarzanego zdjęcia.

3.3 Wyniki

Bardzo duża liczba wygenerowanych obrazów nie pozwala na to żeby pokazać wszystkie możliwe kombinacje obrazów. W tym rozdziale autor skupi się na przedstawianiu najważniejszych cech wygenerowanych obrazów, zalet oraz wad metody opisanej w poprzednich podrozdziałach. Pierwszymi zaprezentowanymi przykładami przeniesienia stylu będą obrazy, na których zawartość między zdjęciami źródłowym i docelowym jest zgodna. Będą to przypadki transferu z pejzażu do pejzażu lub z portretu do portretu. Obrazy zostały przedstawione w różnych stadiach procesu optymalizacji, dzięki temu można zaobserwować w jaki sposób algorytmy zmieniają stopniowo obraz wejściowy w stylizowany, przy równoczesnym bardzo wydajnym zachowaniu stylu.

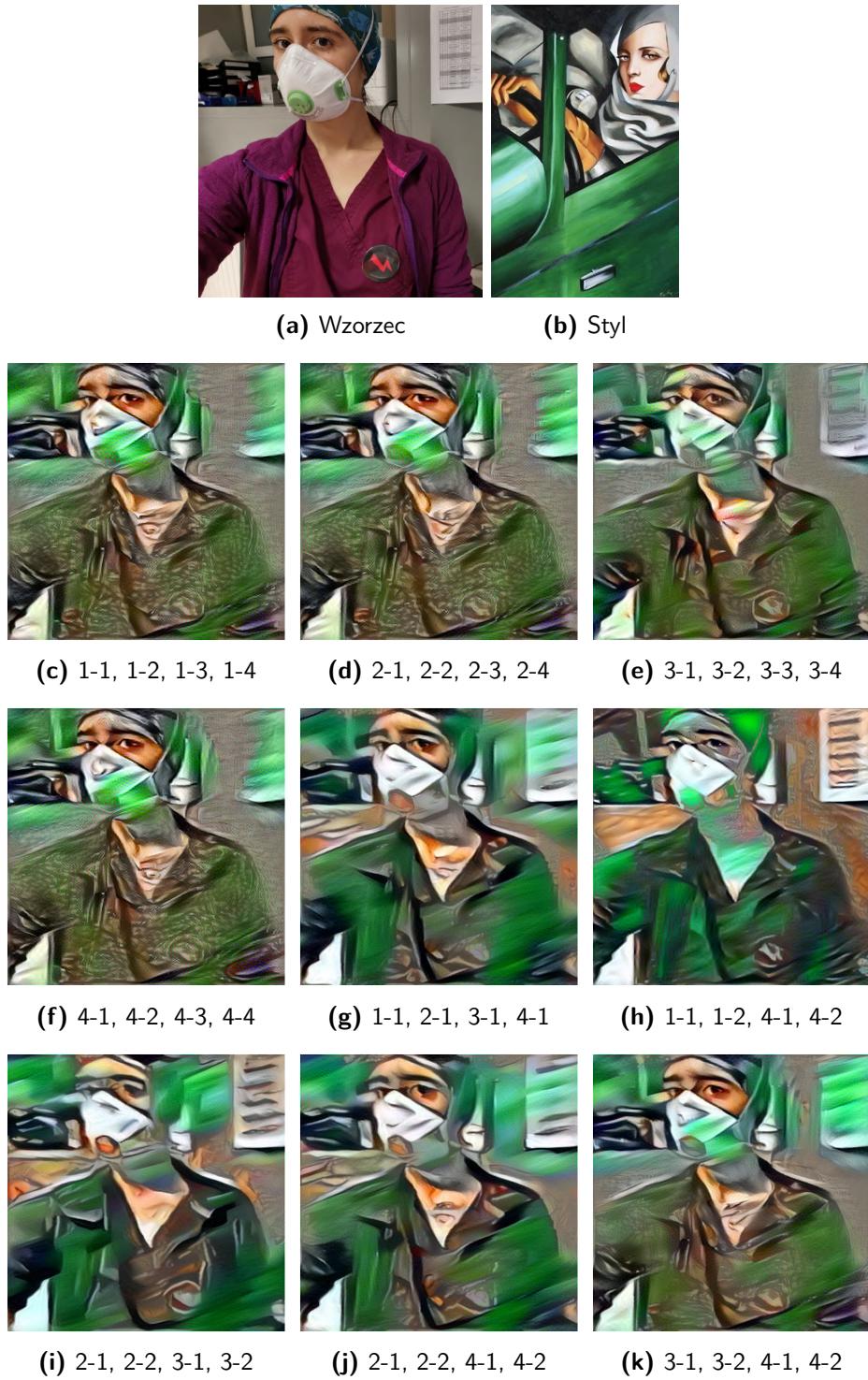
Rys. 13 przedstawia porównanie transferu stylu dla różnych obrazów typu portret. Przykłady z tego rysunku najlepiej pokazują jak metody radzą sobie z przenoszeniem tekstury z pomiędzy dużymi płaszczyznami. Zdjęcie wzorcowe jak i obrazy Tamary Łempickiej są świetnymi przykładami właśnie takiego gatunku, gdzie relatywnie jednolity kolor zaaplikowany jest na dużych obszarach. Światłoń przenoszony jest bardzo dobrze, kontury się wyostrzają, a nawetauważalna jest niewielka zmiana w kształtach części twarzy. Niestety kolory nie są aplikowane we właściwych miejscach, przez co wiele obszarów pokolorowanych jest nienaturalnie i może to psuć nieco odbiór przetworzonego dzieła.

Rys. 14 oraz Rys. 15 prezentują przykłady przeniesienia stylu między obrazami typu pejzaż. Przypadki te ponownie zachowują ogólne podobieństwo zawartości na zdjęciach oraz kompozycji. W obu przypadkach jednak znacznie wyraźniej widać pojawiające się artefakty, rondo nie zawsze przypomina rondo, natomiast na prawie żadnym zdjęciu nie widać prawidłowo zarysowanych dróg. Ilość drobnych szczegółów działa jednak pozytywnie na ogólny wygląd obrazu, w porównaniu z obrazami portretów, znacznie rzadziej można zaobserwować kolory które całkowicie nie powinny się na zdjęciu pojawić, jest to w dużej mierze zasługa typu stylu oryginalnych prac.

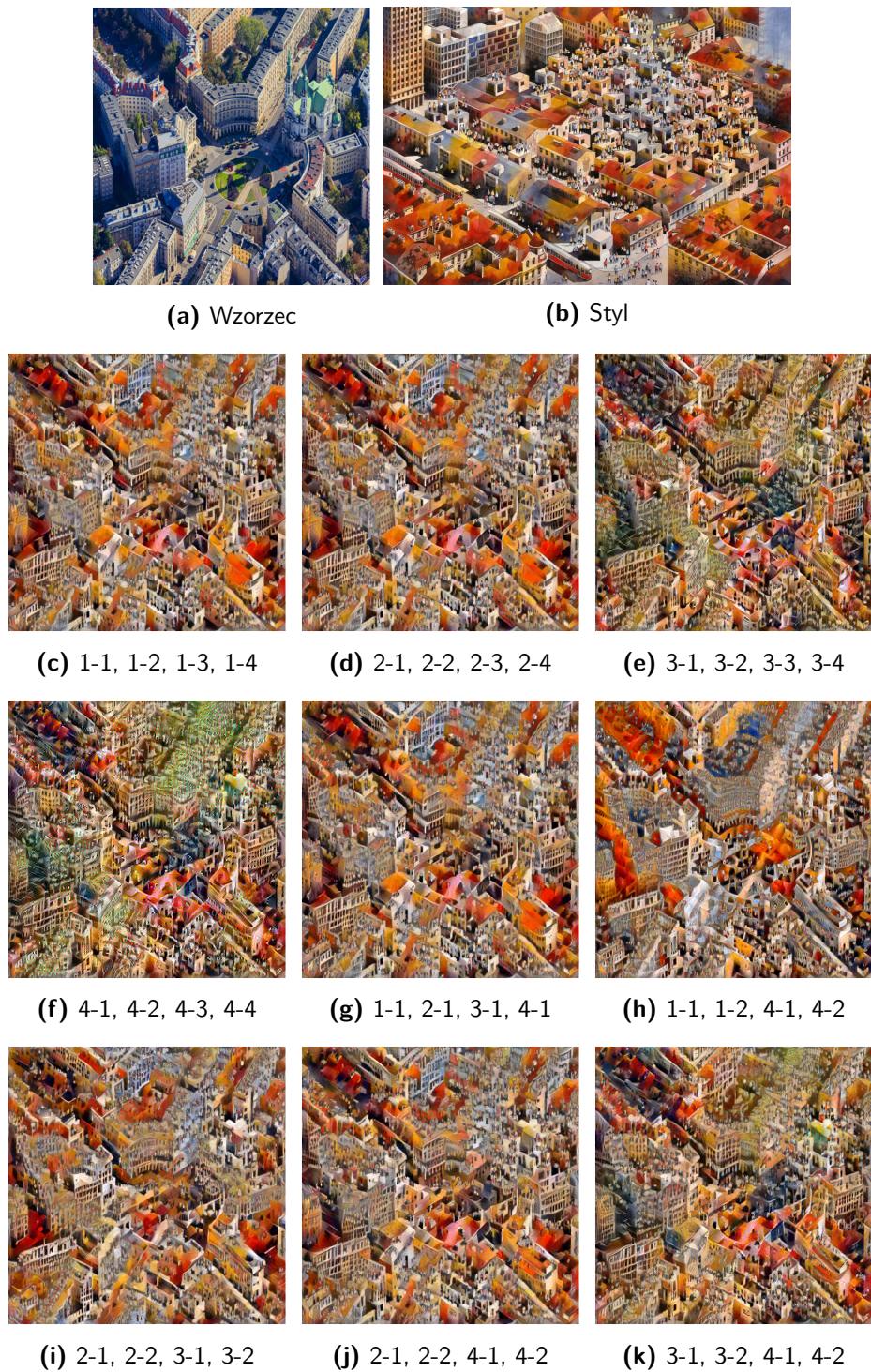
Rys. 16 jest zestawieniem wybranych wyników pozyskanych przy pomocy zmieszania gatunków użytych obrazów, styl pejzażu przeniesiony na portret i odwrotnie, styl portretu przeniesiony na pejzaż. Główny problem który można w tym przypadku zaobserwować jest całkowicie nieudany transfer wykonany między obrazami przedstawiającymi plac Zbawiciela, a stylem Tamary Łempickiej. Kształty zostały całkowicie zmienione, cała zawartość zdjęcia jest zniekształcona, niestety, w tym miejscu pojawia się główna wada metody zaproponowanej w oryginalnym artykule Leona Gatys'a [5], jednak całkowite rozłączenie stylu i zawartości nie jest możliwe. Ten sam problem jest widoczny w przypadku transferu wykonanego w odwrotnym kierunku, mianowicie, z obrazu pejzażu na obraz portretu. Przykłady te zawierają

bardzo wyraźne artefakty w postaci przeniesionych całych elementów budynków, oczekiwany byłoby, aby tylko niewielkie elementy koloru zostały przeniesione z obrazu na obraz, powodem takiego zmieszania zawartości i stylu jest ponownie niemożliwość całkowitego rozgraniczenia tych dwóch elementów obrazu.

Kolejne rozdziały pracy skoncentrowane będą na zaproponowaniu zmian które mogą poprawić wszystkie problemy opisane w tym rozdziale.



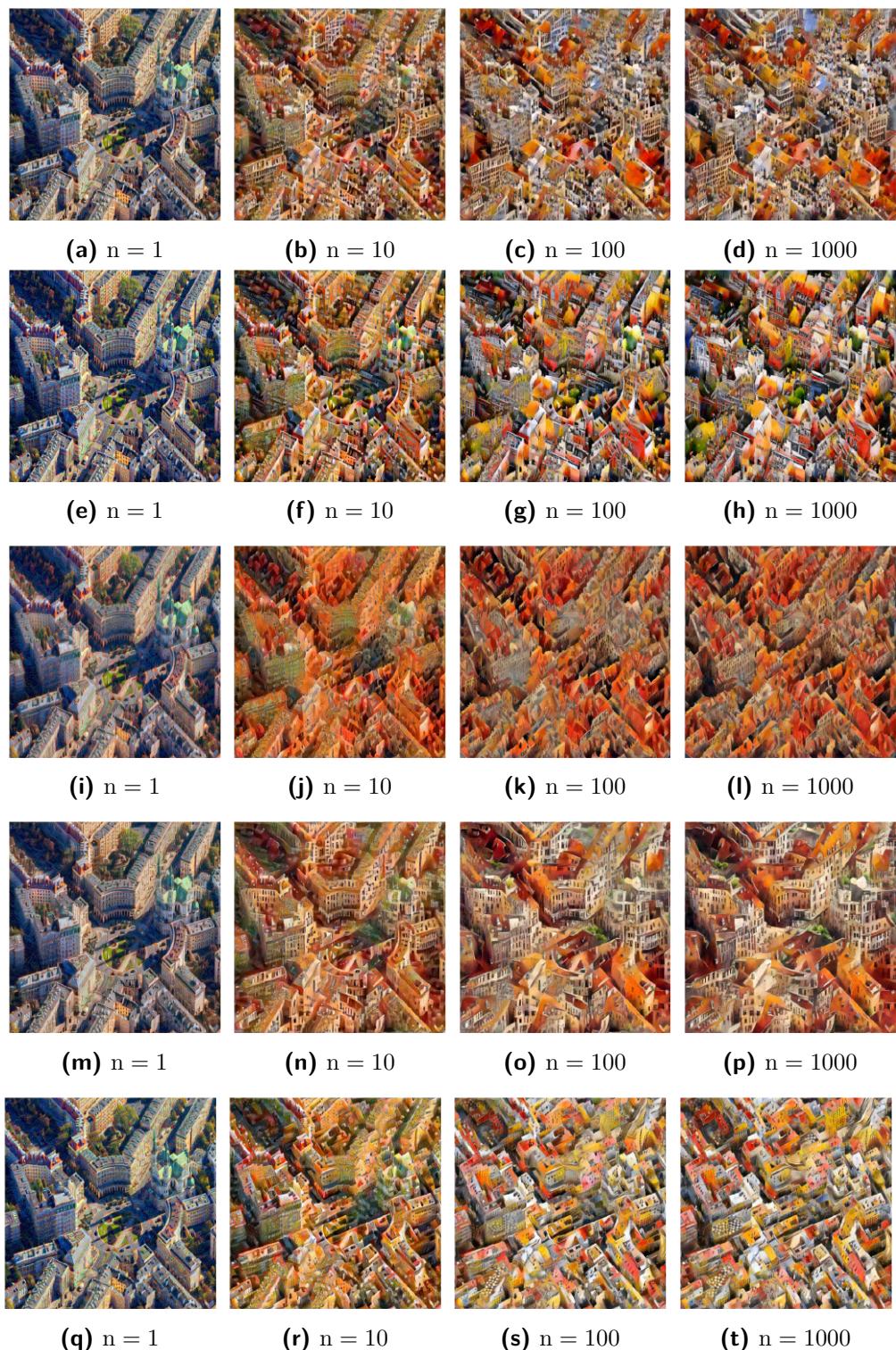
Rysunek 11. Prezentacja wyników transferu stylu obrazów typu portret. Reprezentacja stylu pochodząca z różnych zestawów warstw sieci VGG-19.



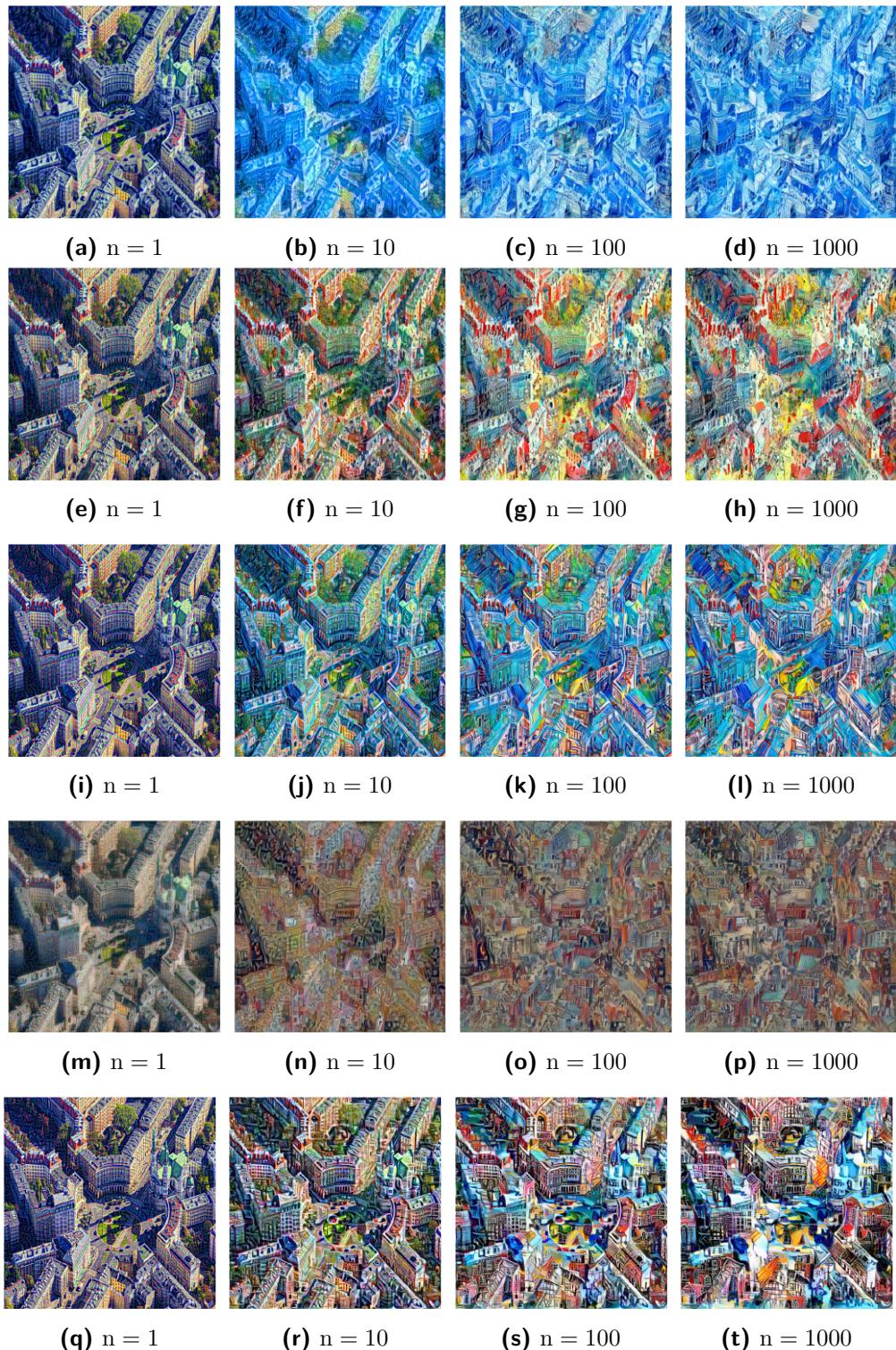
Rysunek 12. Prezentacja wyników transferu stylu obrazów typu pejzaż. Reprezentacja stylu pochodząca z różnych zestawów warstw sieci VGG-19.



Rysunek 13. Wyniki zostały przygotowane przy użyciu Rys. 2a jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 5.



Rysunek 14. Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 3.



Rysunek 15. Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 4.



Rysunek 16. Wyniki przeniesienia stylu przy użyciu zdjęć o różnym początkowym wzorcu.

Rozdział 4

Wariant „wiele do jednego”

4.1 Transfer wiele do jednego

Wyniki przedstawione w poprzednim rozdziale są obiecujące, natomiast zdecydowanie można zauważać pewne niedoskonałości opisane szerzej wcześniej. Wiele osób pracowało nad poprawieniem działania metody Neural Style Transfer, w tym sam autor oryginalnego artykułu [7]. Przedstawił on możliwość sterowania podstawowymi parametrami procesu przekształcenia, takimi jak położenie pojawiających się detali oraz poprawienie przenoszenia kolorów. Praca Leona Gatysa koncentruje się jednak nadal na transferze wykonywanym pomiędzy tylko dwoma obrazami. Podejście *jeden-do-jednego* znacznie ogranicza możliwości poprawy jakości transferu. Dzięki zdefiniowaniu procedury biorącej pod uwagę transfer *wiele-do-jednego* możliwym będzie zastosowanie metod augmentacji obrazów używanych w standardowych zagadnieniach uczenia maszynowego, jak również stworzenie propozycji metody generalizującej styl pokazany pochodzący z wielu różnych obrazów.

Zakładając oznaczenia zgodne z tymi przedstawianymi w poprzednich rozdziałach, gdzie dodatkowo n jest liczbą używanych zdjęć stylu, równanie przedstawione poniżej opisuje porządną funkcję straty używaną w zagadnieniu optymalizacji dla wielu obrazów.

$$\mathcal{L}_{\text{calkowite}}(\vec{p}, \vec{a}, \vec{x}) = \sum_n (\alpha \mathcal{L}_{\text{wzoru}}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{\text{stylu}}(\vec{a}_n, \vec{x})), \quad (8)$$

Zaprezentowana funkcja jest funkcją docelową, natomiast ze względu na charakter prowadzonej transformacji, została ona zaimplementowana przy pomocy prostej heurystyki. Implementacja jest przedstawiona w Alg. 2, gdzie w stosunku do implementacji przedstawionej w Alg. 1 zmienia się forma podawania obrazów stylu. Obrazy są podawane kolejno w pętli co pozwala na sumaryczne zredukowanie błędu, implementacja działa jako dobra heurystyka

Rozdział 4. Wariant „wiele do jednego”

przedstawionego problemu. Wykresy przedstawione na Rys. 17 zbiegają się do wartości minimalnej, niezależnie od ilości obrazów stylu sumaryczna wartość błędu spada.

Algorytm 2: Algorytm optymalizacji *wiele-do-jednego*

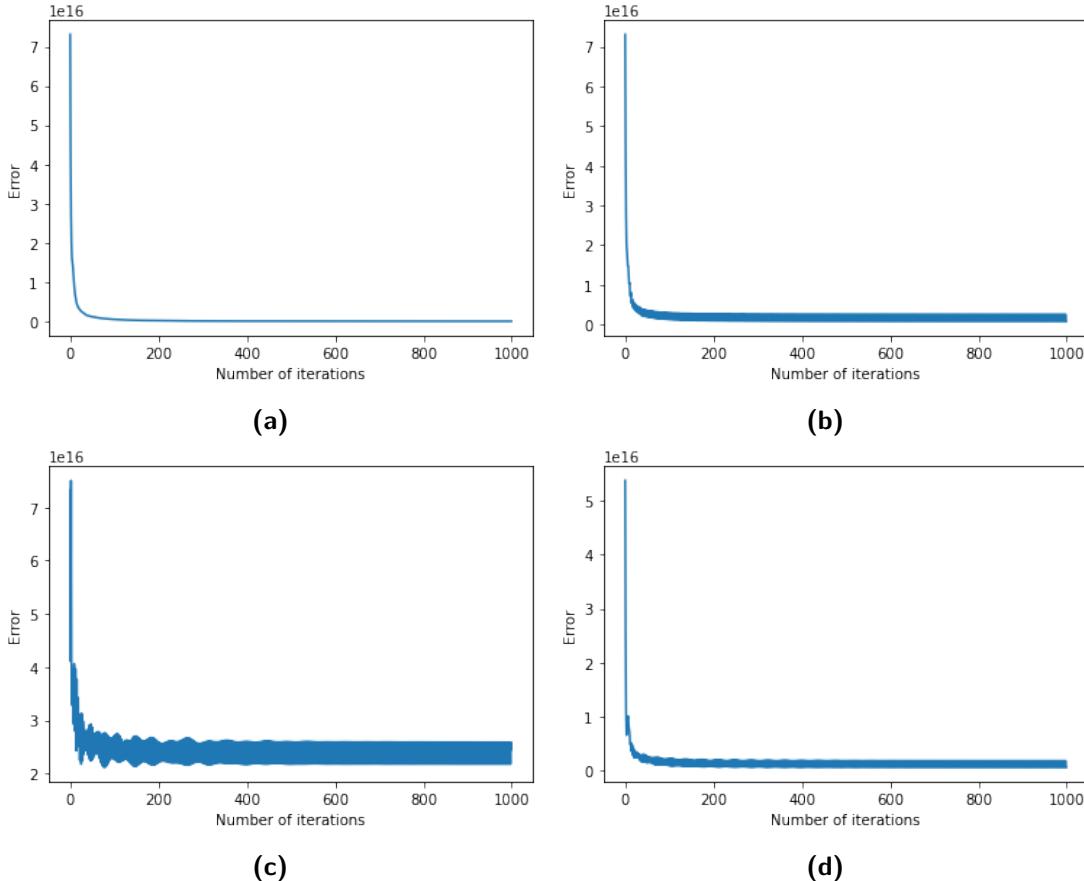
Data: \vec{p} , \vec{a} - obraz wzorcowy, obraz stylu

Result: \vec{x}_{out} - obraz \vec{x} z przeniesionym stylem i zawartością

begin

```
/* pobranie danych cech zawartości i stylu */  
cechy_wzoru ← PobierzCechy( $\vec{p}$ )  
cechy_stylu ← PobierzCechy( $\vec{a}$ )  
/* Inicjalizacja zdjęcia transformowanego */  
 $\vec{x}_{\text{wyj}} \leftarrow \vec{p}$   
/* Proces iteracyjnej optymalizacji */  
for  $i \leftarrow 1$  to  $n_{\text{iteracji}}$  do  
    for  $n \leftarrow 1$  to  $n_{\text{obrazu\_stylu}}$  do  
         $\text{cechy\_wyjsciowe} \leftarrow \text{PobierzCechy}(\vec{x}_{\text{wyj}})$   
         $\mathcal{L}_{\text{stylu}} \leftarrow 0$   
         $\mathcal{L}_{\text{wzoru}} \leftarrow 0$   
        for  $l \in \{\text{warstwa\_wzoru}, \text{warstwy\_stylu}\}$  do  
            if  $l \in \text{warstwa\_wzoru}$  then  
                 $\mathcal{L}_{\text{wzoru}} \leftarrow \mathcal{L}_{\text{wzoru}} + \text{strata\_wzoru}(\text{cechy\_wyjsciowe}, \text{cechy\_wzoru})$   
            if  $l \in \text{warstwa\_stylu}$  then  
                 $\mathcal{L}_{\text{stylu}} \leftarrow \mathcal{L}_{\text{stylu}} + \text{strata\_stylu}(\text{cechy\_wyjsciowe}, \text{cechy\_stylu})$   
             $\mathcal{L}_{\text{calkowite}} \leftarrow \frac{\alpha \mathcal{L}_{\text{wzoru}}}{n_{\text{warstw wzoru}}} + \frac{\beta \mathcal{L}_{\text{stylu}}}{n_{\text{warstw stylu}}}$   
            PropagacjaWsteczna( $\mathcal{L}_{\text{calkowite}}$ )  
        Zmien( $\vec{x}_{\text{wyj}}$ )  
    Wynik( $\vec{x}_{\text{wyj}}$ )
```

Głównym problem który pojawia się w przypadku takiej implementacji tego zagadnienia jest widoczny na Rys. 17c. Pomimo odpowiedniego trendu zmniejszania się błędu optymalizacji widoczna jest pewna rozbieżność δ wyników. Ponieważ zdjęcie podawane są kolejno i dla każdego z nich uzyskany błąd jest inny widoczne są oscylacyjne zmiany w wielkości błędu, zaproponowane rozwiązanie widoczne na Rys. 17d. Zdjęcia podawane w losowej kolejności znacznie zmniejszyły wartość rozbieżności δ , dlatego wszystkie kolejne metody zawierające transfer *wiele-do-jednego*, będą wykonywane przy pomocy losowego podstawiania obrazów stylu.



Rysunek 17. Przebiegi błędu procesu optymalizacji, dla różnych zestawów wielu źródeł stylu: (a) jeden obraz stylu, (b) trzy obrazy będące augmentacjami jednego obrazu źródłowego (bardzo podobne), (c) trzy różne obrazy tego samego autora, (d) obrazy identyczne jak w przypadku (c), ale poddawane transferowi w losowej kolejności.

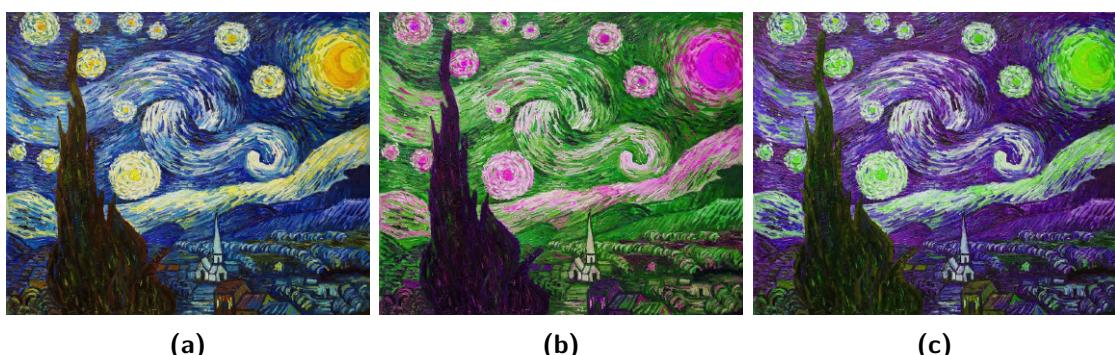
4.2 Opis metody augmentacji obrazu stylu

Metody augmentacji obrazu są szeroko stosowane w procesie uczenia sieci neuronowych [8]. Niewielkie zamiany w obrazie wykonywane na etapie treningu pozwalają na sztuczne zwiększenie ilości obrazów dowolnej klasy, co prowadzi do powiększenia możliwości generalizacyjnych modelu. Zestaw możliwych do wykonania augmentacji [2] umożliwia stworzenie praktycznie nieskończonego zbioru obrazów pochodnych. Popularnie używane augmentacje można podzielić na dwie główne grupy, przekształcenia geometryczne i przekształcenia kolorystyczne. Pierwsza grupa składa się z prostych przekształceń, takich które przychodzą na myśl od razu, obroty o zadany kąt, powiększenia i pomniejszenia czy translacje. Równocześnie opisywane są możliwości korzystania z przekształceń geometrycznych opartych o homografię lub transformacje elastyczne, augmentacje tego typu mogą znacznie zmienić obraz względem

obrazu początkowego. Augmentacje oparte na zmianie koloru nie muszą oznaczać jedynie zmian w natężeniu konkretnych barw ale również zmiany w jasności, kontraste czy całkowite usunięcie kolorów i transformowanie obrazu do formy czarno-białej. Wszystkie te i inne operacje mogą poprawić działanie algorytmu lub je pogorszyć, szczególnie jeśli nie używamy ich w procesie klasycznego treningu ważnym jest odpowiednie dobranie transformacji które nie pogorszą działania algorytmu i co najważniejsze nie wprowadza nieporządnego nowych informacji do procesu przetwarzania.

Dzięki odpowiedniemu zdefiniowaniu procedury przekazania stylu z kilku zdjęć źródłowych do jednego zdjęcia docelowego, możliwym stało się użycie klasycznych metod augmentacji. Najważniejszym krokiem było odpowiednie dobranie właściwych augmentacji w taki sposób, aby nie zmienić zbyt mocno stylu reprezentowanego przez przetransformowane obrazy.

Pierwszy relatywnie łatwy do odrzucenia przypadek augmentacji są to wszystkie augmentacje oparte o transformacje koloru. Przykłady takich transformacji zostały pokazane na Rys. 18. Łatwo można zaważyć, że te przedstawione ilustracje są do siebie podobne, ale tylko pod względem tekstu i pociągnięć pędzla. Kolorystycznie są t zupełnie różne od siebie dzieła, w przypadku gdzie algorytm bierze pod uwagę kolorystyczne cechy stylu wprowadzana jest nieporządną informację która może tylko zaburzyć działanie procedury. Natomiast przypadek w którym nie bierzemy pod uwagę koloru, a operujemy jedynie na kanale luminescencji dla wszystkich trzech przedstawionych przykładów nie ulega zmianie, więc z punktu widzenia obrazu czarno-białego, a w szczególności kanału luminescencji, te trzy zdjęcia są identyczne. Kolory powinny być zgodne ze zdjęciem stylu lub ze zdjęciem wzorcowym, kontr produktywnym jest tworzenie nowej palety barw dla powstającego zdjęcia. Wady tych transformacji są na tyle duże, że zmian kolorów zostały odrzucone już we wczesnym momencie testowania metody.



Rysunek 18. Przykłady obrazów na których wykonano augmentacje przez przesunięcie koloru, (a) obraz oryginalny, (b) dodatnie przesunięcie HUE oraz (c) ujemne przesunięcie HUE.

Drugi typ rozpatrywanych transformacji, czyli wycięcie, został pokazany na Rys. 19. Przekształcenie takie jest przeprowadzonym równocześnie operacjami wycięcia oraz powiększenia

wyciętego rozmiaru do wielkości oryginalnego obrazu. Niestety pomimo faktu, że ten zestaw operacji nie zmienia bezpośrednio koloru ani tekstuury, w trakcie takiego przekształcenia pośrednio zmieniana jest forma przedstawionej na obrazie stylu. Rys. 19c świetnie pokazuje ideę tego problemu, przez wycięcie i powiększenie znacznie zmienia się bilans kolorów przedstawionych na obrazie, a równocześnie zmienia się wielkość i częstotliwość pociągnięć pędzla, czyli tekstuра. Zmiana tekstuury całkowicie zmienia styl obrazu, z impresjonizmu, na styl oparty o większe jedno-kolorowe płaszczyzny. Podobnie do augmentacji przedstawionych w poprzednim akapicie, te pokazane na Rys. 19 nie są obiecujące i nie zostaną użyte w dalszej części badań.



Rysunek 19. Przykłady obrazów na których wykonano augmentacje przez zmianę rozmiaru i wycięcie, (a) obraz oryginalny, (b) oraz (c) wycięta część obrazu oryginalnego.

Ostatnim typem augmentacji które były rozważane w kontekście poprawy działania zaproponowanej wcześniej metody były obroty oraz przerzuty (ang. *flips*). Początkowo można uznać, że metoda reprezentacji stylu przy pomocy macierzy Grama powinna wyeliminować potrzebę augmentacji obrazów względem położenia elementów na zdjęciu. Jednak, co zostanie wykazane dokładniej w kolejnym rozdziale macierz Grama eliminuje wpływ położenia elementów względem pojedynczych filtrów w warstwach konwolucyjnych, ale mimo wszystko jakość transferu zmienia się wraz z dodaniem kolejnych przerzuconych i obróconych zdjęć. Przerzucenia zdjęcia w osi poziomej i pionowej zostały pokazane na Rys. 20, jak łatwo można zauważyć, styl zdjęcia nie zmienia się wraz z augmentacją, przy równoczesnej zmianie położenia elementów względem siebie, co pozwala na zmianę reprezentacji stylu pomimo użycia macierzy Grama. Niestety obroty nie zostały zaimplementowane w ostatecznej wersji rozwiązania, ponieważ w przeciwieństwie do klasycznych zagadnień uczenia maszynowego, w przypadku transferu stylu nie można pozwolić sobie na użycie uzupełnienia obrazu (ang. *padding*). Niezależnie od wybranej wartości czy algorytmu uzupełnienia brakujących miejsc, każde jego użycie znacznie zmieniłoby styl przedstawiony na obrazie. Wycięcie obrazu po

wykonaniu operacji rotacji, w taki sposób aby pominąć uzupełnianie brakujących pikseli, powoduje problemy opisane w poprzednim rozdziale.



Rysunek 20. Przykłady obrazów na których wykonano augmentację przez zmianę położenia względnego elementów na obrazie, (a) obraz oryginalny, (b) obraz przerzucony w pionie oraz (c) obraz przerzucony w poziomie.

Cechy przekształceń opisane we wcześniejszych akapitach wymuszają użycie konkretnych typów augmentacji. Zaimplementowane transformacje to przerzuty w różnych osiach, ze względu na relatywnie długi czas generowania wyników skupiono się tylko na tych konkretnych transformacjach. Wyniki tych przekształceń zaprezentowano w kolejnych rozdziale.

4.3 Wyniki

Przedstawione obrazy wynikowe zostały ułożone i pogrupowane w taki sam sposób jak w rozdziale 3.3. Dokładnie zaprezentowano transfer pomiędzy takimi samymi gatunkami obrazów, a w przypadku mieszanego transferu skupiono się głównie na najbardziej reprezentatywnych próbkach.

Rys. 21 przedstawia takie same przekształcenia jak Rys. 13 z tą różnicą, że do stworzenia tego pierwszego użyto augmentacji opisanych wcześniej w tym rozdziale. Obrazy na pierwszy rzut oka wydają się bardzo podobne i w przypadku stylu Tamary Łempickiej nie występują znaczące różnice. Jednak dokładniej porównując przykłady widać, że w przypadku augmentowanych obrazów transfer przebiegał szybciej już dla $n = 100$ widać dobrze przedstawione szczegóły i wyraźny zanik szumów. Równocześnie kolory są bardziej przejrzyste a kontury wyraźne, wydaje się również, że obrazy są bardziej przejrzyste. Niestety nie udało się uniknąć całkowicie artefaktów kolorów, które nadal obecne są na każdym zdjęciu w postaci zmieniających kolor ścian.

Rys. 22 przedstawia analogicznie transfer przy użyciu augmentowanych wersji obrazów Tytusa Brzozowskiego. Ponownie zauważalna jest poprawa prędkości i wydajności przeniesienia

stylu. Dodatkowo można wyraźnie zaobserwować redukcję ilości artefaktów, szczególnie w okolicach ronda, dla argumentowanych przykładów wszystkie ulice są wyraźne i dobrze odcięte na tle budynków. Bilans kolorów również uległ poprawie, jasność obrazu wydaje się być rozłożona naturalniej w porównaniu z nieaugmentowaną wersją obrazów.

Rys. 23 ponownie przedstawia proces transferu dla augmentowanych obrazów źródłowych stylu. Na stylizowanych obrazach widać podobną poprawę jakości co w przypadku obrazów Brzozowskiego, ulice i rondo są znacznie bardziej widoczne i wyraźne. Przeciwnie do obrazów Brzozowskiego, Edward Dwurnik używa niewielkich pociągnięć pędzla, co zostało uwydatnione dzięki użyciu dodatkowych wzbogacających transformacji. Kolory są bardziej przejrzyste, a światłocienie wyraźne.

Rys. 24 przedstawia miks gatunków, gdzie portretów i pejzaży użyto krzyżowo w procesie transferu, wszystkie przedstawione przypadki są wersjami wzbogaconymi oryginalnych dzieł malarstwa. Niestety transfer stylu Tamary Łempickiej na zdjęcia wzorcowe zawierające pejzaż nie poprawił się bardzo. Zawartość zdjęcia ponownie całkowicie zanika i niestety nie jest rozpoznawalna na zdjęciu. Jednak wydaje się, że na obrazach pomimo złego odwzorowania zawartości pojawia się mniej szumu. Poprawę można natomiast zaobserwować w przypadku transferu stylu pejzażu na obrazy portretowe. Podobnie jak na Rys. 22 i Rys. 22 widoczna jest poprawa rozkładu kolorów i światłocienia. Kontury twarzy są wyraźniej widoczne i szczegóły takie jak książki w tle są mniej zaszumione.

Głównym wnioskiem który można wyciągnąć jest fakt, że augmentacje znacznie lepiej działają w przypadku obrazów o wielu drobnych szczegółach. Wydaje się, że dzięki zaimplementowanym obrotom umożliwione jest lepsze dopasowanie się krótkich pociągnięć pędzla do kształtów na zdjęciu wzorcowym, w przypadku dużych i jednolitych płaszczyzn koloru taki zabieg nie wprowadza wielu nowych informacji.

4.4 Opis metody generalizacji reprezentacji stylu

Problem generalizacji stylu jest bardzo skomplikowany i można podjąć próby rozwiązania go przy pomocy wielu różnych metod. Świetnym przykładem pracy która podejmuje polemikę na temat tego zagadnienia jest "A Style-Aware Content Loss for Real-time HD Style Transfer" [12]. Praca przedstawia próbę wykorzystania struktury auto-enkodera jako elementu nadającego styl przetwarzanemu zdjęciu, a odpowiednio wytrenowany klasyfikator używany jest jako dyskryminator który zastępuje metrykę pomiaru stylu jako reprezentacji wielu obrazów równocześnie. Takie podejście jest często widoczne w artykułach i literaturze, jednak nie jest ono idealne, obie struktury, enkoder i dyskryminator wymagają długiego procesu uczenia i są przystosowane tylko do określonej puli obrazów. Algorytmy tego typu przenoszą obciążenie

obliczeniowe do fazy uczenia, taki zabieg może być porządnego w zależności od przedstawionych wymagań projektu. Warto jednak zdawać sobie sprawę z tego, że utworzenie wielu takich modeli działających równocześnie jest bardzo trudne i czasochłonne.

Metoda która zostanie przedstawiona w tym rozdziale nie będzie wymagała wykonania długiej i ciężkiej obliczeniowo fazy uczenia. Dzięki wykorzystaniu przedstawionej wcześniej heurystyki możliwym jest użycie szeregu obrazów tego samego artysty do przeniesienia stylu obrazów zawartych w szeregu na obraz docelowy. Głównym problemem pojawiającym się w trakcie przenoszenia stylu przy pomocy tej techniki jest duża rozbieżność błędu uczenia Rys. 17c, która powodowana jest przez relatywnie duże różnice w lokalnych reprezentacjach tego samego stylu na kolejnych przedstawianych jako źródła stylu zdjęciach. W tym przypadku znacznie ważniejszym jest zastosowanie opisanego wcześniej rozwiązania tego problemu 17d, w celu zmniejszenia powtarzalnych różnic które mogą wpływać na jakość procesu optymalizacji zdjęcia przedstawiane są w losowej kolejności. Zastosowanie takiego zabiegu było subiektywnym odczuciem autora opartym o doświadczenia zdobyte w trakcie wykonywanych treningów, okazało się bardzo dobrze zmniejszać problem rozbiegania się błędu, rozwiązując problem i umożliwiając wykorzystanie heurystyki poza przykładami bardzo podobnych obrazów.

4.5 Wyniki

Przedstawione wyniki zostały podzielone względem gatunku obrazu wzorcowego. Metoda generalizacji stylu powinna jeszcze lepiej poradzić sobie z przeniesieniem stylu pomiędzy obrazami o różnej zawartości niż metody przedstawione wcześniej. Na każde zdjęcie wzorcowe przetransferowano style obrazów wszystkich trzech artystów.

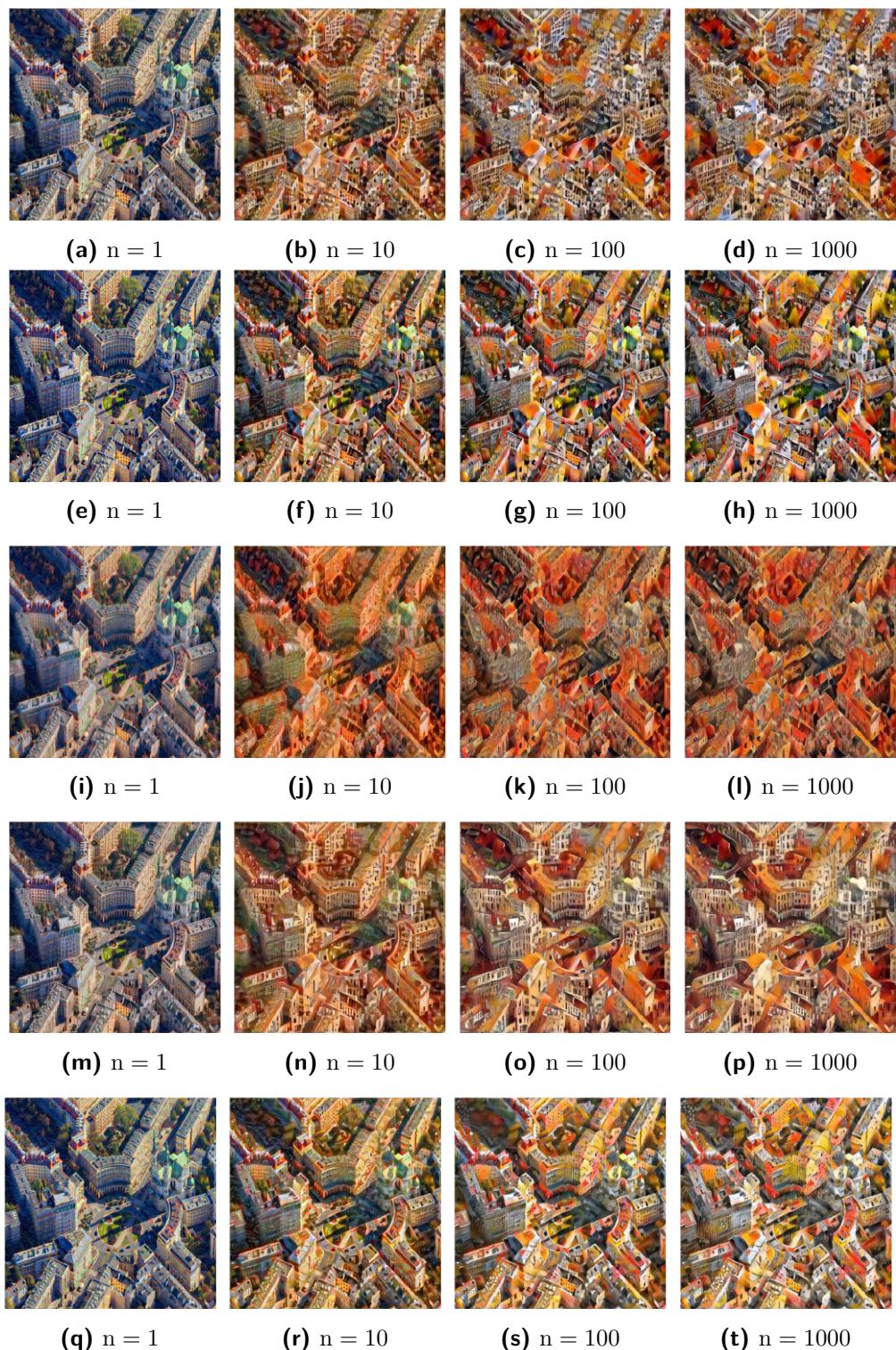
Rys. 25 przedstawia portret na bazie którego został wykonany transfer stylu artysty. Bazą stylu były obrazy przedstawione w poprzednich rozdziałach. Łatwo można zaobserwować, że transfer stylu rzeczywiście odbył się prawidłowo, na obrazach wynikowych widać cechy pochodzące ze wszystkich obrazów, przy równoczesnym braku artefaktów pochodzących z konkretnych pojedynczych dzieł. Kolory dopasowały się zaskakująco dobrze do zdjęcia wzorcowego pobierając wartości kolorów z różnych miejsc różnych obrazów. Widać to szczególnie na polarze i bluzce, tam cienie są wyraźnie ciemniejsze a czerwony kolor pozostał w większości bez zmian. Równocześnie dzięki użyciu przedstawionych wcześniej augmentacji uniknięto problemów z zaszumieniem kolorów, co znacznie poprawia wizualny obraz stworzonego obrazu.

Rys. 26 przedstawia analogicznie wykonaną prezentację, gdzie obrazem wzorcowym jest pejzaż. Wiele pozytywnych cech i spostrzeżeń jest takich samych jak w przypadku portretu. Augmentacja nie przynosi tutaj spektakularnych efektów, ponieważ na zdjęciu naturalnie wy-

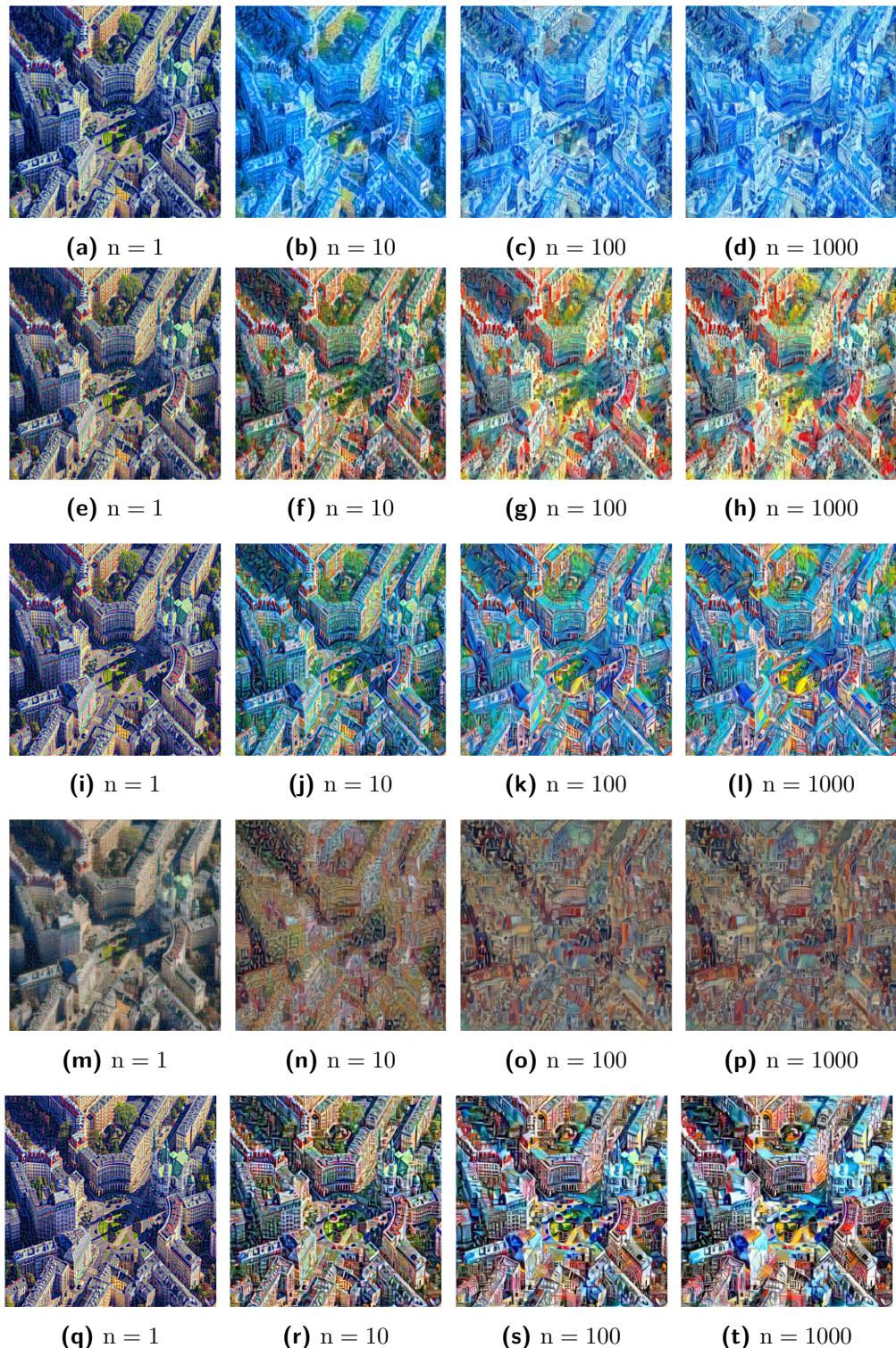
stępuje dużo obiektów o częstych zmianach koloru, dlatego transfer naturalniej przeprowadzany jest bez zaszumienia. Wydaje się, że wyniki wyglądają jeszcze lepiej niż w przypadku portretów. Kolory są różnorodne, ale pasujące idealnie do palety artysty. Równocześnie nie pojawiają się prawie żadne artefakty i zawartość oryginalnego zdjęcia jest bez problemu wykrywalna. Niestety jak w poprzednich próbach transfer stylu Tamary Łempickiej jest niemożliwy, ogromne różnice w formie użycia kolorów i geometrii na obrazie powoduje całkowity brak rozpoznawalności oryginalnej zawartości.



Rysunek 21. Wyniki zostały przygotowane przy użyciu Rys. 2a jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 5.



Rysunek 22. Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 3.



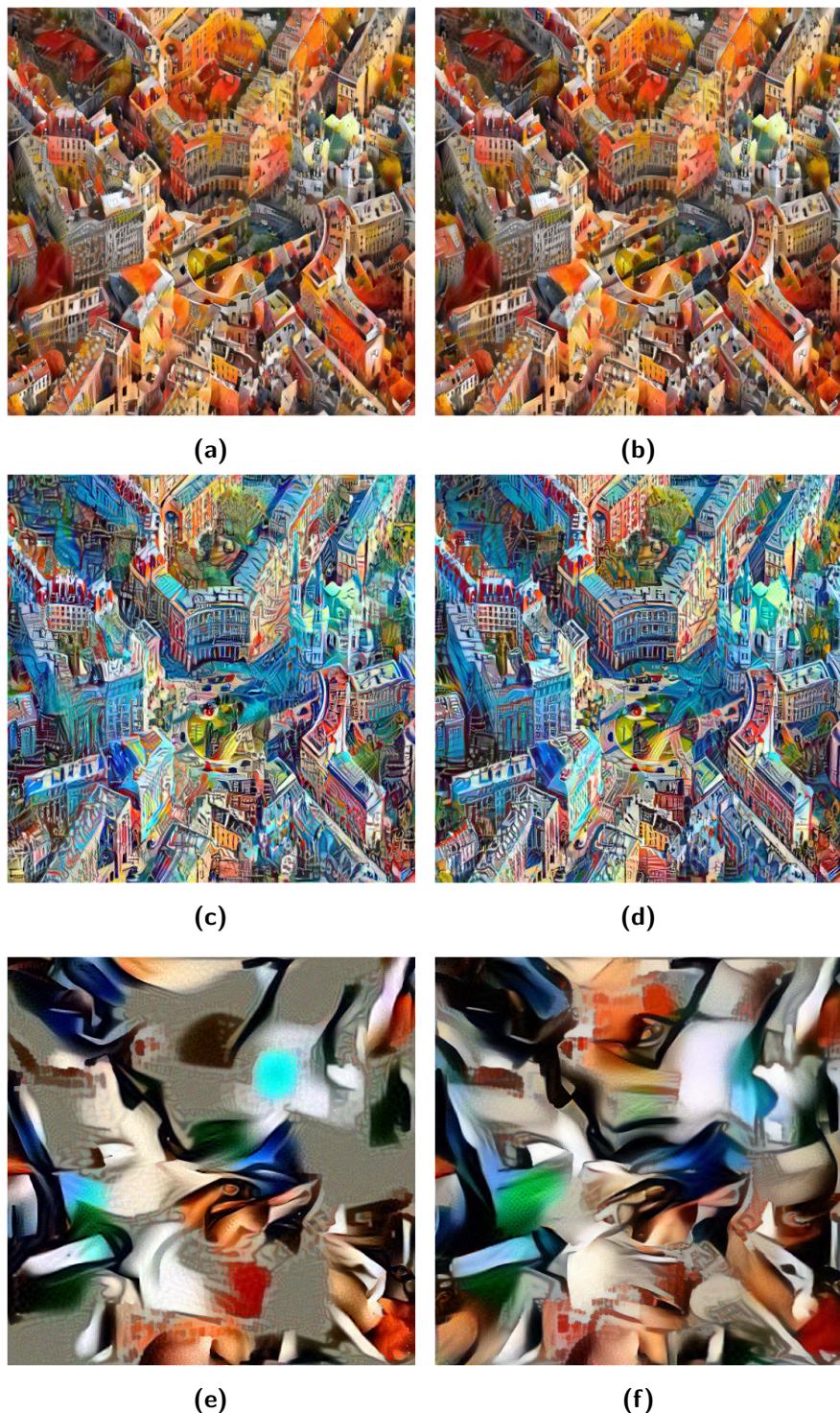
Rysunek 23. Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 4.



Rysunek 24. Wyniki przeniesienia stylu przy użyciu zdjęć o różnej początkowej zawartości.



Rysunek 25. Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.



Rysunek 26. Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza: (a),(b) Edward Dwurnik, (c),(d) Tytus Brzozowski oraz (e),(f) Tamara Łempicka. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.

Rozdział 5

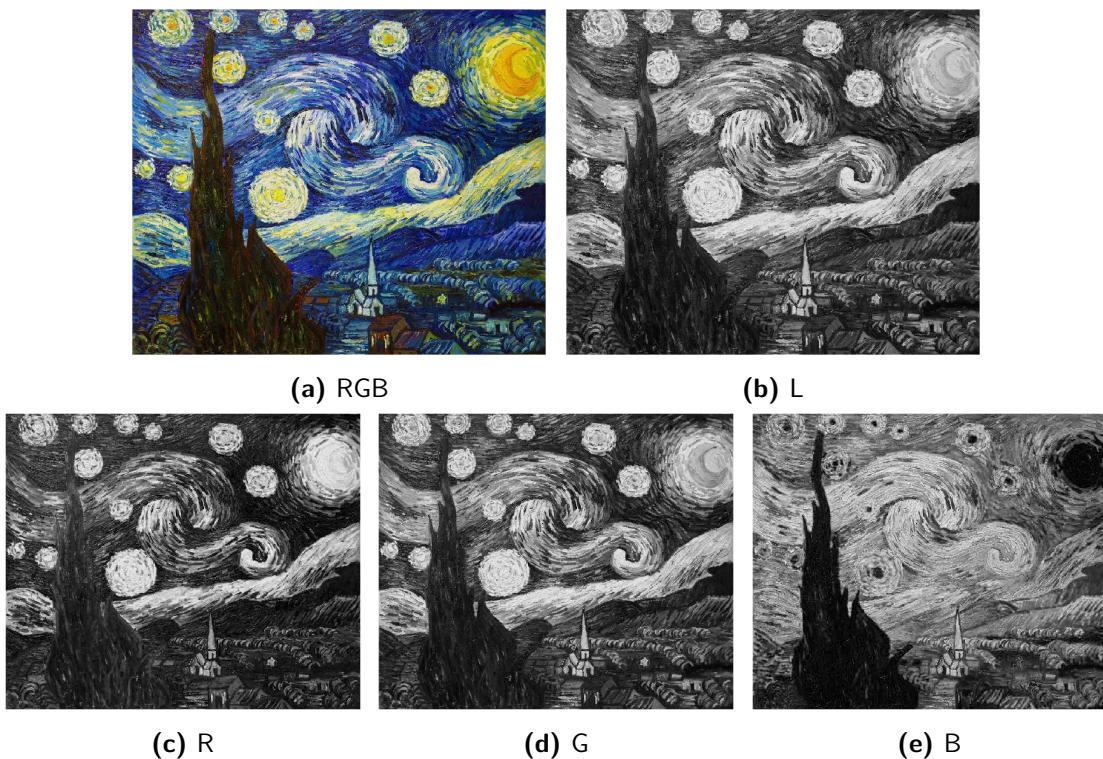
Wariant ze sztucznym kolorowaniem

5.1 Transfer bez koloru

Informacja o kolorze obrazu jest ważnym aspektem percepcyjnym jego stylu. Jednocześnie jest to w dużej mierze niezależne od innych aspektów stylu, takich jak rodzaj pędzla użyte obrysów lub dominujące kształty geometryczne. Dlatego pożądanym jest możliwość samodzielnego decydowania o kolorach zawartych na obrazie wyjściowym. Natywne metoda przedstawiona przez Leona Gatysa kopiuje również rozkład kolorów obrazu stylu, który może być w wielu przypadkach niepożądany. Na przykład jak na Rys. 1, portret przedstawiający osobę w czerwonych ubraniach na białym tle został zamieniony w obraz którego kolory świetnie reprezentują tylko jeden konkretny obraz Van Gogha. Transformacja znacznie lepiej wyglądałaby jeśli kolory oryginalnego obrazu wzorcowego pozostałyby bez zmian, natomiast jedynymi cechami stylu jakie zostałyby zaaplikowane to kształty i pociągnięcia pędzla.

Umożliwienie swobodnej kontroli koloru w trakcie przenoszenia stylu wymaga przedstawienia metody, która pozwoli całkowicie wyeliminować informacje o kolorze zawarte na obrazie. Oryginalna metoda 'Neural Style Transfer' wymusza użycia trzech płaszczyzn kolorów, w przypadku użycia obrazu kolorowego domyślną przestrzenią barw jest skala RGB (ang. red-green-blue). Niestety jak widać na Rys. 27c, d i e każda z warstw RGB reprezentuje nieco inny styl, przez co wybranie pojedynczej przetworzonej warstwy nie może dać prawidłowych wyników. Powyższy problem wymaga rozwiązania przez wybranie innego typu przestrzeni barw, takiego który zawiera informacje dotyczące tylko jasności. Takie informacje pozwalają na wyodrębnienie konturów, kształtów geometrycznych i cieni, które dają bardzo dobrą reprezentację stylu.

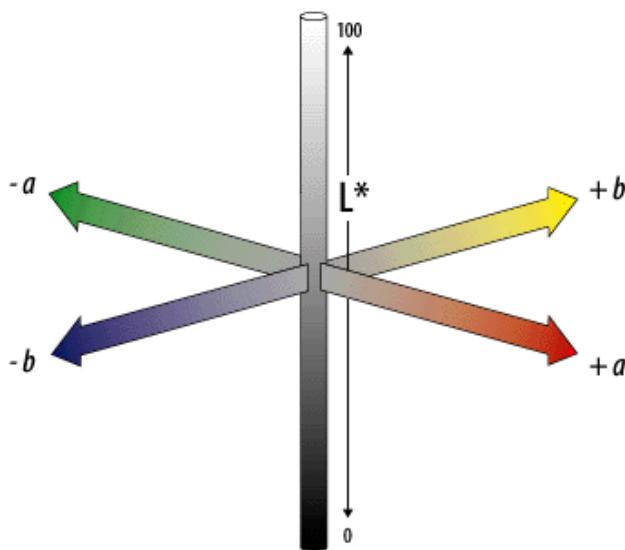
Przestrzeń barw L,a,b Rys. 28, podobnie do przestrzeni RGB przedstawia każdy piksel obrazu przy pomocy trzech koordynatów. Inaczej jednak niż w przypadku RGB, nie każda z trzech wartości reprezentuje wielkość barwy. Zgodnie z tym co zostało postawione jako założenie dla



Rysunek 27. Przedstawienie obrazu w różnych skalach barw: (a) przestrzeń trzy-kanałowa RGB, (b) kanał luminescencji wyciągnięty z reprezentacji w przestrzeni L,a,b, (c) kanał czerwony - RGB, (d) kanał zielony - RGB oraz (e) kanał niebieski - RGB.

wybieranej przestrzeni barw L, odpowiada za wartość jasności, wartości te przyjmują wielkości między 0, a 100. Obie wielkości a i b przyjmują wartości -128 do 128. Jednak osie a i b nie dla każdej wielkości L rozciągają się na całym swoim możliwym zakresie, ponieważ są nieliniowo skompresowanymi reprezentacjami przestrzeni CIEXYZ. Wyeksportowana warstwa L pozwoli na analizę jedynie kanału jasności i na przeniesienie cech stylu z pominięciem barw.

Pre-trenowana głęboka sieć neuronowa VGG-19 na której oparta jest architektura algorytmu przenoszenia stylu nie jest natywnie dostosowana do przetwarzania obrazów czarno-białych. Pierwsza warstwa konwolucyjna wymusza trzy kanały koloru na każdym obrazie wejściowym. Sieć w tym przypadku nie jest używana do zagadnienia klasyfikacji, a jedynie jako element dobierający cechy, dlatego nie ma potrzeby tworzenia i uczenia nowej sieci, przystosowanej specjalnie do jednowymiarowego wejścia. Zaproponowanym rozwiązaniem jest konkatenacja trzech identycznych obrazów czarno-białych w jeden trój-wymiarowy, nie jest to idealne rozwiązanie, jednak ze względu na to, że zarówno zdjęcia stylu i zdjęcia wzorcowe poddawane są tym samym zabiegom ekstrakcji cech, takie podejście nie powinno powodować złych efektów. Tezę tą potwierdzają wyniki przedstawione w kolejnym rozdziale, gdzie cechy obrazów, światłocień i kontury, zostały świetnie wyodrębnione z całości dzieła.



Rysunek 28. Przedstawienie graficznego opisu przestrzeni L,a,b , gdzie L - kanał jasności (luminescencji), a,b - kanały reprezentujące kolory[13].

5.2 Wyniki

Obrazy przedstawione w tym podrozdziale są wynikami pośrednimi. Przeciwnie do wyników pokazywanych wcześniej nie zostaną tutaj przedstawione wszystkie możliwe konfiguracje przeprowadzonych transferów. Wybrana została reprezentatywna próbka badawcza, na bazie której opisane zostaną odpowiednie wnioski. Rys. 29 przedstawia transfery oparte na zestawie augmentowanych zdjęć jednego obrazu źródłowego stylu. Łatwo można zauważyć, że czarno-białe obrazy nie są najlepszą reprezentacją cech stylu artysty. Kolor gra ogromną rolę w odbiorze dzieła, nie musi być on zgodny z oryginalnym dziełem, ale niestety luminescencja nie przekazuje odpowiednio dużej ilości informacji, aby łatwo powiązać reprezentację stylu z konkretnym artystą. Można zaobserwować jednak, że obrazy o dużej ilości drobnych szczegółów znacznie gorzej prezentowane są przy pomocy jedynie luminescencji. Duża ilość światłocienia zamieniana jest w szum, natomiast duża część obiektów na obrazie nie ma odzwierciedlenia w kanale jasności a w kanałach kolorów.

5.3 Sztuczne kolorowanie obrazu

Metody sztucznego kolorowania obrazów czarno-białych są rozwijane od wielu lat i dzięki użyciu głębokich sieci neuronowych możliwym było uzyskanie zaskakująco realistycznych wyników. Doskonałym przykładem rozwiązania tego problemu jest prezentująca niesamowite wyniki praca 'Colorization Transformer' [10], można zobaczyć jak doskonale z takim zagadnieniem

radzą sobie struktury transformatorów. Niestety pomimo, że zastosowanie takiej struktury mogłoby przynieść ciekawe wyniki, niestety nie ma możliwości użycia tego rozwiązania w procedurze transferu stylu. Wszystkie pre-trenowane metody nie zdadzą egzaminu w połączeniu z algorytmem stylizacji ponieważ modele nie są przygotowane do obróbki stylizowanych zdjęć. Algorytmy oczekują bardzo specyficznego typu obrazów wejściowych, takich które zostały wykonane zwykłym aparatem i przypominają natywne zdjęcia wzorcowe. Dodatkowym napotkanym problemem jest pewna losowość dobieranych kolorów, w kontekście całego obrazu tworzą one spójną całość, ale w porównaniu z oryginalnym obrazem źródłowym, barwy mogą nie być takie same. W kontekście transferu stylu ważna jest spójność barw z obrazem wzorcowym lub stylu, w przeciwnym wypadku, powstały obraz będzie sprawiał wrażenie niekonsekwentnego. Wymienione powody wymuszają użycie prostszej metody, takiej którą łatwiej kontrolować i dla której możliwym jest całkowite i wierne odwzorowanie barw obrazu źródłowego zawartości.

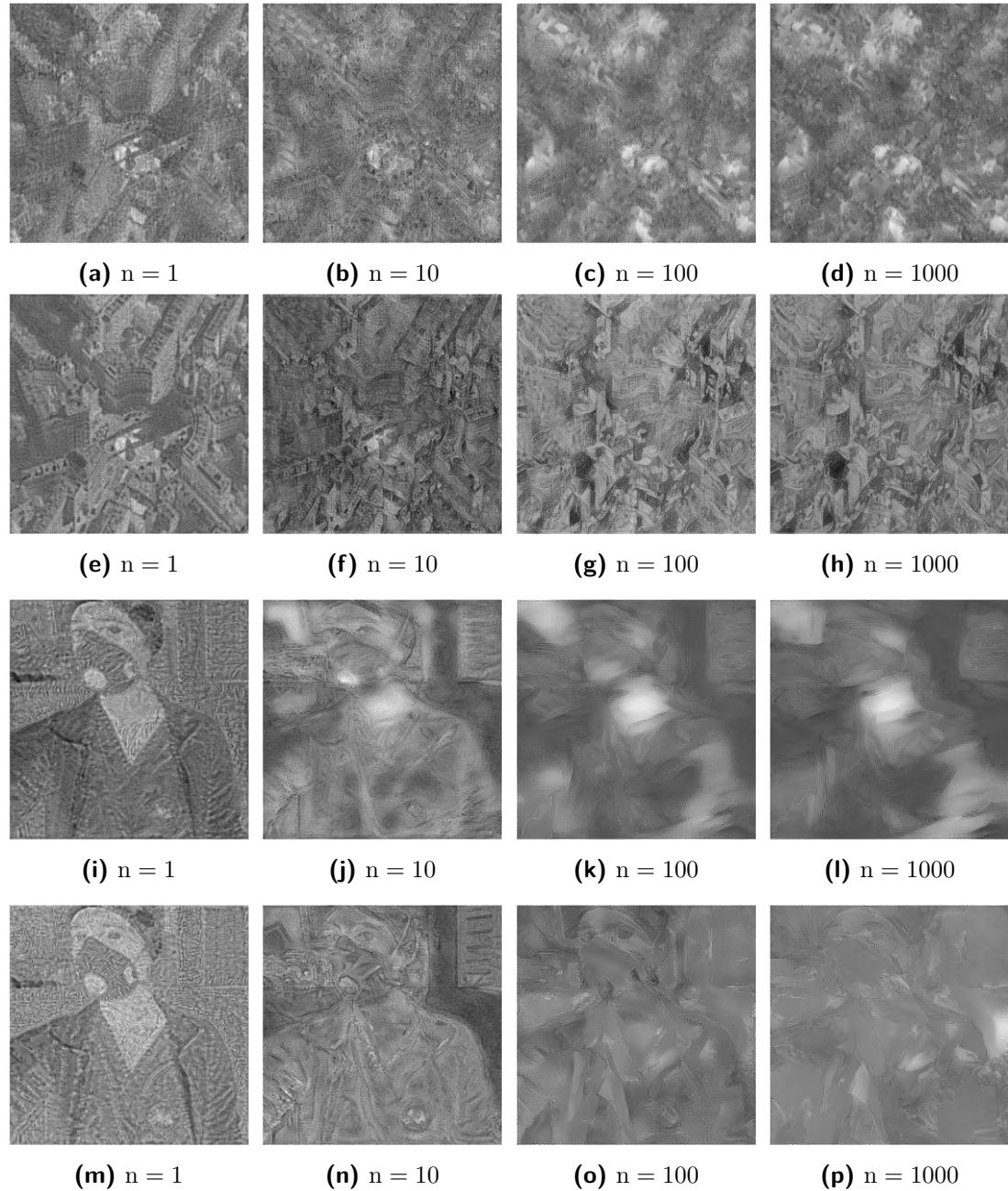
Zdecydowano się, że ze względu na pewną niedokładność i naiwność ludzkiej percepcji najlepszym rozwiązaniem będzie naiwne podmienienie warstw barwnych w przestrzeni L,a,b Rys. 30. Zgodnie z tym co można zaobserwować na przykładach zabieg taki nie jest dobry dla każdej konfiguracji obrazów, sam charakter zapisu koordynatów w przestrzeni L,a,b wiele barw całkowicie ginie i nie jest widocznych jeżeli poziomy jasności nie są odpowiednio dobrane. Szczęśliwe w poprzednim rozdziale pokazane zostało, że istnieje możliwość przeniesienia kanałów luminescencji z bardzo dobrym skutkiem i dokładnością, a przeniesione kontury i światłocień dobrane zakrywają niedokładności samego sztucznego kolorowania. Procedura zabiegu sztucznego kolorowania została przedstawiona na Rys. 31. Zapisane warstwy a,b są przenoszone bez zmian i łączone z wystylizowaną warstwą jasności. Proces przenoszenia stylu między obrazami składającymi się tylko z czarno-białych punktów jasności został przedstawiony dokładniej w poprzednim rozdziale.

5.4 Wyniki

Przedstawione w tym rozdziale wyniki zostały ograniczone do wyników bazujących na generalizowanej reprezentacji stylu połączonej z przedstawioną wcześniej procedurą sztucznego kolorowania obrazu. Rys. 32 przedstawia portrety, już tutaj wyraźnie widać, że metoda kolorowania działa znacznie lepiej przypadkach gdzie wyraźnie pojawia się duża ilość czarnych elementów. Świecące odwzorowane są cienie i kontury, czyli elementy bardzo prominentne w stylu malarstwa Tamary Łempickiej. Znacznie gorzej odwzorowane są drobne punkty i małe pociągnięcia pędzla które utożsamiamy ze stylem Dwurnika. Podobny problem widać na obrazach bazujących na stylu Brzozowskiego gdzie jasność nie odwzorowała dobrze łagodnych

przejść między kolorami. Podobne wnioski można wyciągnąć przyglądając się Rys. 33, tutaj również pomimo pewnych kolorystycznych mankamentów najlepiej wygląda wersja pejzażu bazująca na obrazach Tamary Łempickiej. Niestety bez użycia odpowiednich barw i kolorów style Tytusa Brzozowskiego czy Edwarda Dwurnika są niezwykle ciężkie do odtworzenia.

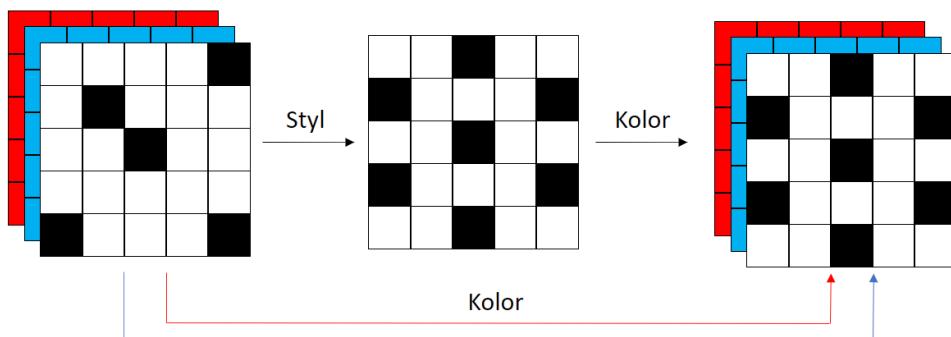
Oglądając przedstawione wyniki można przypuszczać, że im mniej małych kolorowych detali na obrazie źródłowym stylu, tym lepiej przebiegnie koloryzacja. Obrazy o nienaturalnych kolorach po transformacji nie będą wyglądały dobrze i nie można liczyć, że proces transferu przebiegnie poprawnie. Warto zauważyć jednak, że ponownie zabieg augmentacji pomógł w zmniejszeniu ilości zakłóceń kolorystycznych i jest to bardzo przydatna metoda poprawy jakości transferu.



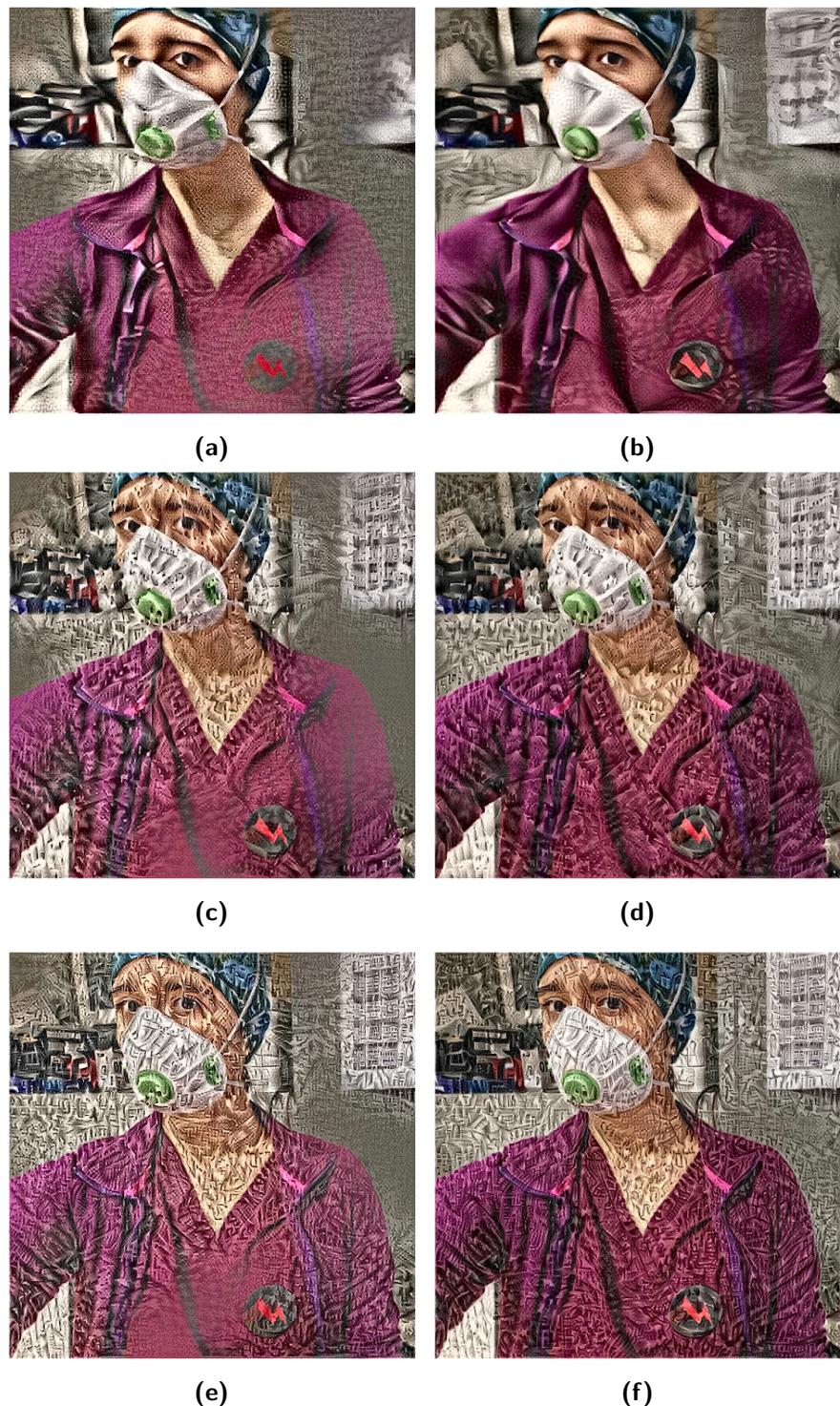
Rysunek 29. Przykłady czarno-białych obrazów wygenerowanych na podstawie augmentowanych obrazów stylu: (a)-(d) Tytus Brzozowski "Norblin", (e)-(h) Edward Dwurnik "Idealne miasto", (i)-(l) Tamara Łempicka "Autoportret w Zielonym Bugatti" oraz (m)-(p) Tamara Łempicka "Portret Dr. Boucard".



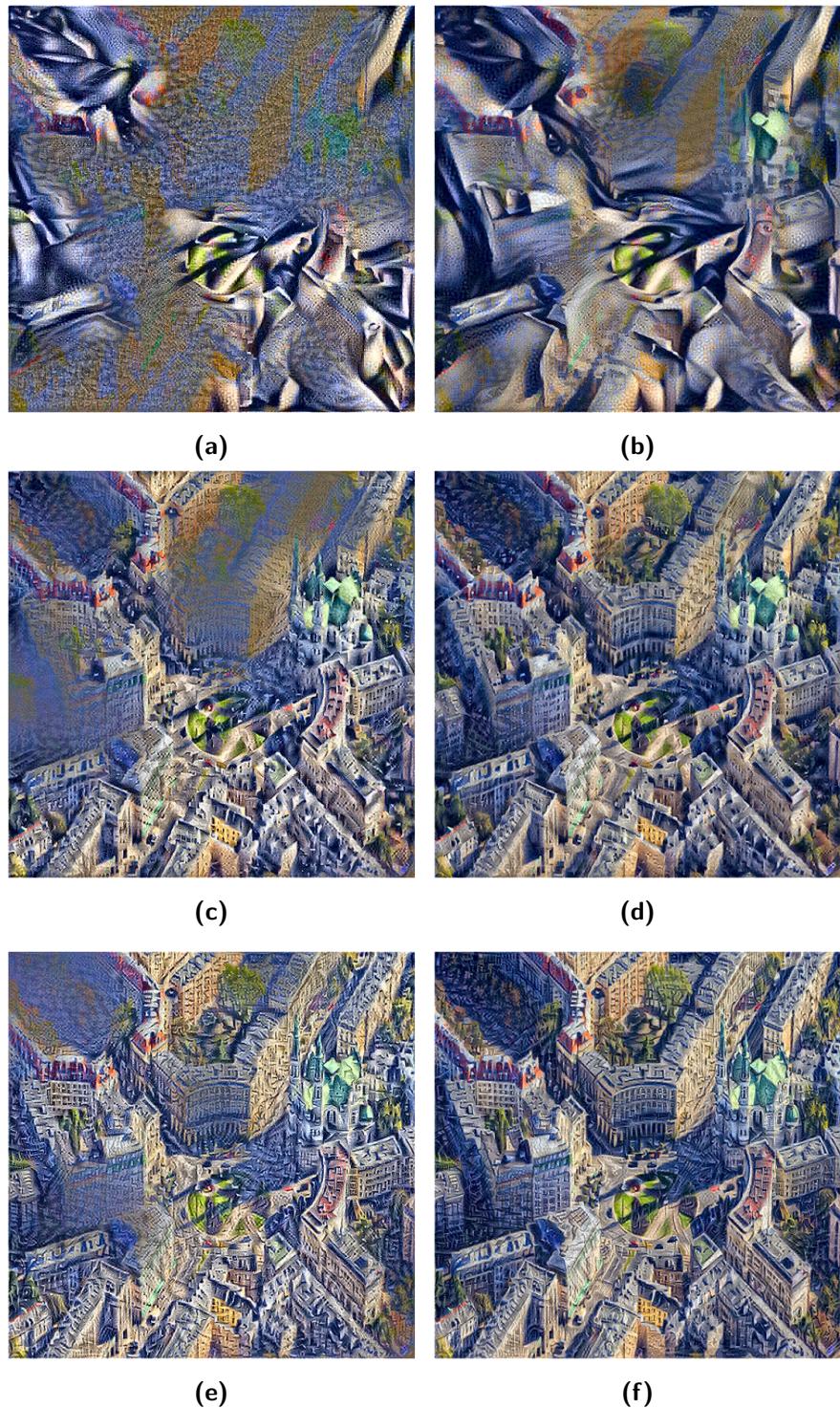
Rysunek 30. Przykłady koloryzacji obrazów w przestrzeni L,a,b: (a),(d) obrazy oryginalne, odpowiednio stylu i wzorcowy, (b),(e) odpowiednio tylko warstwa luminescencji, (c),(f) sztuczne kolorowanie obrazów, odpowiednio warstwy a,b wzorca nałożone na warstwę L obrazu stylu i warstwy a,b obrazu stylu nałożone na warstwę L wzorca, ostatni obraz (g) to warstwa L poddana procesowi transferu stylu z nałożonymi warstwami a,b oryginalnego obrazu wzorcowego.



Rysunek 31. Przedstawienie graficznego opisu przestrzeni L,a,b , gdzie L - kanał jasności (luminencji), a,b - kanały reprezentujące kolory[13].



Rysunek 32. Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza oraz koloryzacji: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.



Rysunek 33. Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza oraz koloryzacji: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.

Rozdział 6

Podsumowanie

Głównymi celami pracy było poprawienie działania metody "*Neural Style Transfer*" oraz stworzenie procedury która umożliwi zaaplikowanie generalnego stylu malarza na zdjęciu. Wszystkie cele pracy zostały zrealizowane, przedstawiono trzy nowe metody, bazujące na implementacji *wiele-do-jednego*. Augmentacja oraz sztuczne kolorowanie pozwoliły na zmniejszenie ilości artefaktów i poprawienie jakości kolorów na obrazie wynikowym. Natomiast realizacja generalizacji stylu pozwoliła na przeniesienie całkowitego stylu artysty.

Wyniki przedstawione w pracy nie są idealnie jednoznaczne, niektóre metody powodowały bardziej zauważaną poprawę dla konkretnych zestawów obrazów. W niektórych przypadkach natomiast wyniki prawie w ogóle się nie różniły od siebie. Przypadki wykonywanej augmentacji pozwoliły na poprawę i zmniejszenie ilości występujących artefaktów, dla obrazów o dużej ilości szczegółów, natomiast powodowały znacznie mniejszą poprawę dla obrazów które składały się głównie z pojedynczych płaszczyzn. Metody augmentacji nie były jednak główną częścią pracy, zdecydowanie można powiedzieć, że zaproponowana metoda *wiele-do-jednego* działała prawidłowo i pozwoliła na generalizację informacji zawartych na obrazach źródłowych. Takie wyniki pozwoliły z kolei na zaproponowanie kolejnych metod które umożliwiły tworzenie reprezentacji całego stylu pojedynczego autora, jak również na mieszanie różnych stylów, co pozwoliło na uzyskanie bardzo ciekawych i satysfakcjonujących wyników. Metody przenoszenia stylu między obrazami czarno-białymi, jak i zaproponowana metoda kolorowania obrazów, pozwoliły na odkrycie zupełnie nowych możliwości, które nigdy wcześniej nie były szeroko zbadane, a uzyskane wyniki w wielu przypadkach całkowicie przekroczyły oczekiwania autora.

Ilościowa ocena jakości transferu jest bardzo trudna, zdefiniowanie jednoznacznej metryki jakości dla tego konkretnego zagadnienia jest praktycznie niemożliwe. Jedyna możliwość to ocena obrazów jakościowo indywidualnie w przypadku każdego transferu. Każdy krok wykonany w pracy był następnie oceniony na bazie konkretnych przykładów, żadna z zaproponowanych

Rozdział 6. Podsumowanie

metod nie spowodowała pogorszenia, a w większości można zaobserwować znaczną ogólną poprawę.

Wyniki opisane w pracy są obiecujące i zdecydowanie warto spróbować dalej je rozbudować. Równolegle z przedstawionymi badaniami rozwija się wiele innych metod transferu stylu opartych o zupełnie nowe struktury głębokich sieci neuronowych. Tak jak zostało to opisane we wcześniejszych rozdziałach transformatory okazały się niezwykle efektywne w rozwiązywaniu problemu transferu stylu. Dodatkowo połączenie struktur generator i dyskryminator w jedną strukturę przenoszącą pozwala na zmniejszenie potrzeby obecności człowieka w procesie transformacji. Pierwszymi krokami które należałyby podjąć w celu dalszego rozwinięcia jakości zaproponowanych metod jest przetestowanie innych rodzajów augmentacji obrazów źródłowych stylu, proponując niestandardowe metody przetwarzania takie jak zniekształcenia homologiczne czy elastyczne można mieć nadzieję na uzyskanie jeszcze lepszych wyników. Kolejnym krokiem byłoby zastosowanie przedstawionych zmian we współpracy z inną procedurą transformacji, świetnym punktem startowym mogłyby być metoda przedstawiona w artykule "A Style-Aware Content Loss for Real-time HD Style Transfer"[12]. Umożliwia ona na połączenie wszystkich opisanych w tej pracy metod i dodatkowo wprowadza strukturę generator-dyskryminator, takie połączenie różnych metodyk może dać bardzo dobre wyniki, uzyskiwane w czasie rzeczywistym.

Bibliografia

- [1] Brzozowski, T., *Różne dzieła*, Accessed: 2021-06-11. adr.: <https://tytusbrzozowski.pl/pl/obrazy/>.
- [2] Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M. i Kalinin, A. A., „Albumentations: Fast and Flexible Image Augmentations”, *Information*, t. 11, nr. 2, s. 125, lut. 2020, ISSN: 2078-2489. DOI: [10.3390/info11020125](https://doi.org/10.3390/info11020125). adr.: <http://dx.doi.org/10.3390/info11020125>.
- [3] Dwurnik, E., *Różne dzieła*, Accessed: 2021-06-11. adr.: <https://desamodern.pl/artyści-edward-dwurnik>.
- [4] Efros, A. A. i Leung, T. K., „Texture Synthesis by Non-parametric Sampling”, w *IEEE International Conference on Computer Vision*, Corfu, Greece, wrz. 1999, s. 1033–1038.
- [5] Gatys, L. A., Ecker, A. S. i Bethge, M., „A Neural Algorithm of Artistic Style”, *CoRR*, t. abs/1508.06576, 2015. arXiv: [1508.06576](https://arxiv.org/abs/1508.06576). adr.: <http://arxiv.org/abs/1508.06576>.
- [6] ——, *Texture Synthesis Using Convolutional Neural Networks*, 2015. arXiv: [1505.07376 \[cs.CV\]](https://arxiv.org/abs/1505.07376).
- [7] Gatys, L. A., Ecker, A. S., Bethge, M., Hertzmann, A. i Shechtman, E., „Controlling Perceptual Factors in Neural Style Transfer”, *CoRR*, t. abs/1611.07865, 2016. arXiv: [1611.07865](https://arxiv.org/abs/1611.07865). adr.: <http://arxiv.org/abs/1611.07865>.
- [8] Géron, A., *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media, 2017, ISBN: 978-1491962299.
- [9] Hertzmann, A., Jacobs, C., Oliver, N., Curless, B. i Salesin, D. H. a., „Image Analogies”, Association for Computing Machinery, Inc., sierp. 2001. adr.: <https://www.microsoft.com/en-us/research/publication/image-analogies/>.
- [10] Kumar, M., Weissenborn, D. i Kalchbrenner, N., „Colorization Transformer”, *CoRR*, t. abs/2102.04432, 2021. arXiv: [2102.04432](https://arxiv.org/abs/2102.04432). adr.: <https://arxiv.org/abs/2102.04432>.

Bibliografia

- [11] Łempicka, T., *Różne dzieła*, Accessed: 2021-06-11. adr.: <https://www.wikiart.org/en/tamara-de-lempicka>.
- [12] Sanakoyeu, A., Kotovenko, D., Lang, S. i Ommer, B., „A Style-Aware Content Loss for Real-time HD Style Transfer”, *CoRR*, t. abs/1807.10201, 2018. arXiv: 1807.10201. adr.: <http://arxiv.org/abs/1807.10201>.
- [13] Simões, N., *Soil colors – what more could you want?*, Accessed: 2021-06-11. adr.: <https://blogs.egu.eu/divisions/ssss/2014/04/15/soil-colors-what-more-could-you-want/>.
- [14] Simonyan, K. i Zisserman, A., *Very Deep Convolutional Networks for Large-Scale Image Recognition*, 2015. arXiv: 1409.1556 [cs.CV].
- [15] *Struktura sieci VGG-19*, Accessed: 2021-06-11. adr.: <https://peerj.com/articles/cs-364/>.
- [16] *Warszawski Plac Zbawiciela z lotu ptaka*, Accessed: 2021-06-11. adr.: <https://polandonair.com/produkt/plac-zbawiciela/>.
- [17] Wei, L.-Y. i Levoy, M., „Fast Texture Synthesis Using Tree-Structured Vector Quantization”, *Computer Graphics (Proceedings of SIGGRAPH'00)*, t. 34, maj 2000. DOI: 10.1145/344779.345009.

Spis rysunków

1	Przykładowy wynik transferu stylu z wykorzystaniem metody opisanej w rozdziale 3.	12
2	Zdjęcia używane jako zdjęcia wzorcowe w dalszej części pracy: (a) Olga Bogacz - siostra autora pracy, (b) "Warszawski Plac Zbawiciela z lotu ptaka" [16]. . .	16
3	Obrazy Tytusa Brzozowskiego [1](używane w pracy jako źródła stylu): (a) "Norblin", (b) "Archikatedra w zieleni", (c) "Rynek", (d) "Plac Unii Lubelskiej", (e) "Plac Defilad"	20
4	Obrazy Edwarda Dwurnika [3](używane w pracy jako źródła stylu): (a) "Idealne miasto", (b) "Poznań 2", (c) "Pl. Zbawiciela", (d) "Tarnów", (e) "Wrocław". .	21
5	Obrazy Tamary Łempickiej [11](używane w pracy jako źródła stylu): (a) "Autoportret w Zielonym Bugatti", (b) "Dziewczęta", (c) "Portret Dr. Boucard", (d) "Les Deux Amies", (e) "Dziewczyna z Mandoliną".	22
6	Zdjęcia i obrazy reprezentujące skrajnie różne cechy stylu: (a) rzeczywiste zdjęcie, (b) "Plac Zbawiciela" T. Brzozowski[1], (c) "Pl. Zbawiciela" E. Dwurnik[3]	22
7	Ekstrakcja stylu i wzorca, na bazie których generowany jest opis cech obrazu w algorytmie Neural Style Transfer. VGG-19 to wersja sieci użyta we własnej implementacji autora, oryginalna praca [5] proponuje użycie sieci VGG-16. Pierwszy obraz od góry to obraz wzorcowy, drugi to aktualnie przetwarzana wersja obrazu, natomiast ostatni obraz to obraz źródłowy stylu. W procesie optymalizacji używane są odpowiednie cechy stylu i zawartości wszystkich trzech obrazów.	24
8	Wyliczenie Macierzy Grama odpowiadającej zestawowi map cech. Każdy z elementów macierzy jest tak naprawdę współczynnikiem korelacji pomiędzy aktywacjami danego filtru.	25
9	Porównanie struktur (a) VGG-19 i (b) VGG-16.	28

Spis rysunków

10	Schemat używanej struktury głębokiej sieci neuronowej VGG-19 [15]. Najważniejsze są zaznaczone na pomarańczowo bloki konwolucyjne, to z nich pobierane są cechy stylu i zawartości. Żółta warstwa wyjścia została usunięta i nie jest używana.	29
11	Prezentacja wyników transferu stylu obrazów typu portret. Reprezentacja stylu pochodząca z różnych zestawów warstw sieci VGG-19.	32
12	Prezentacja wyników transferu stylu obrazów typu pejzaż. Reprezentacja stylu pochodząca z różnych zestawów warstw sieci VGG-19.	33
13	Wyniki zostały przygotowane przy użyciu Rys. 2a jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 5.	34
14	Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 3.	35
15	Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 4.	36
16	Wyniki przeniesienia stylu przy użyciu zdjęć o różnym początkowym wzorcu.	37
17	Przebiegi błędu procesu optymalizacji, dla różnych zestawów wielu źródeł stylu: (a) jeden obraz stylu, (b) trzy obrazy będące augmentacjami jednego obrazu źródłowego (bardzo podobne), (c) trzy różne obrazy tego samego autora, (d) obrazy identyczne jak w przypadku (c), ale poddawane transferowi w losowej kolejności.	41
18	Przykłady obrazów na których wykonano augmentacje przez przesunięcie koloru, (a) obraz oryginalny, (b) dodatnie przesunięcie HUE oraz (c) ujemne przesunięcie HUE.	42
19	Przykłady obrazów na których wykonano augmentacje przez zmianę rozmiaru i wycięcie, (a) obraz oryginalny, (b) oraz (c) wycięta część obrazu oryginalnego.	43
20	Przykłady obrazów na których wykonano augmentację przez zmianę położenia względnego elementów na obrazie, (a) obraz oryginalny, (b) obraz przerzucony w pionie oraz (c) obraz przerzucony w poziomie.	44
21	Wyniki zostały przygotowane przy użyciu Rys. 2a jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 5.	48
22	Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 3.	49
23	Wyniki zostały przygotowane przy użyciu Rys. 2b jako zdjęcia wzorcowego, oraz kolejnych obrazów z Rys. 4.	50
24	Wyniki przeniesienia stylu przy użyciu zdjęć o różnej początkowej zawartości.	51

25	Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.	52
26	Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza: (a),(b) Edward Dwurnik, (c),(d) Tytus Brzozowski oraz (e),(f) Tamara Łempicka. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.	53
27	Przedstawienie obrazu w różnych skalach barw: (a) przestrzeń trzy-kanałowa RGB, (b) kanał luminescencji wyciągnięty z reprezentacji w przestrzeni L,a,b, (c) kanał czerwony - RGB, (d) kanał zielony - RGB oraz (e) kanał niebieski - RGB.	56
28	Przedstawienie graficznego opisu przestrzeni L,a,b, gdzie L - kanał jasności (luminescencji), a,b - kanały reprezentujące kolory[13].	57
29	Przykłady czarno-białych obrazów wygenerowanych na podstawie augmentowanych obrazów stylu: (a)-(d) Tytus Brzozowski "Norblin", (e)-(h) Edward Dwurnik "Idealne miasto", (i)-(l) Tamara Łempicka "Autoportret w Zielonym Bugatti" oraz (m)-(p) Tamara Łempicka "Portret Dr. Boucard".	60
30	Przykłady koloryzacji obrazów w przestrzeni L,a,b: (a),(d) obrazy oryginalne, odpowiednio stylu i wzorcowy, (b),(e) odpowiednio tylko warstwa luminescencji, (c),(f) sztuczne kolorowanie obrazów, odpowiednio warstwy a,b wzorca nałożone na warstwę L obrazu stylu i warstwy a,b obrazu stylu nałożone na warstwę L wzorca, ostatni obraz (g) to warstwa L poddana procesowi transferu stylu z nałożonymi warstwami a,b oryginalnego obrazu wzorcowego.	61
31	Przedstawienie graficznego opisu przestrzeni L,a,b, gdzie L - kanał jasności (luminescencji), a,b - kanały reprezentujące kolory[13].	62
32	Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza oraz koloryzacji: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.	63
33	Przykłady wyników procesu stylizacji z użyciem generalizowanej reprezentacji stylu malarza oraz koloryzacji: (a),(b) Tamara Łempicka, (c),(d) Tytus Brzozowski oraz (e),(f) Edward Dwurnik. Pierwsza kolumna została uzyskana przy pomocy nieaugmentowanego zestawu zdjęć, druga, każde zdjęcie było augmentowane.	64