

Politechnika Wrocławska

Wydział Elektroniki, Fotoniki i Mikrosystemów

KIERUNEK: Automatyka i Robotyka (AIR)

PRACA DYPLOMOWA INŻYNIERSKA

TYTUŁ PRACY:
Aplikacja webowa zwiększająca
rozdzielczość obrazów

AUTOR:
Eryk Wójcik

PROMOTOR:
dr hab. inż. Andrzej Rusiecki,
Katedra Informatyki Technicznej

Spis treści

1	Wstęp	3
2	Podstawy teoretyczne	5
2.1	Definicja super-rozdzielczości	5
2.2	Przegląd metod powiększania obrazów	6
2.3	Wprowadzenie do głębokiego uczenia się w przetwarzaniu obrazów	8
2.4	Wstęp do funkcji falkowych	9
3	DWSR: Deep Wavelet Super Resolution	11
3.1	Architektura DWSR	11
3.2	Kluczowe cechy i innowacje	11
3.3	Proces treningu i implementacji	11
3.4	Przykłady zastosowań i rezultaty	11
4	ESRGAN	13
4.1	Architektura ESRGAN	13
4.2	Kluczowe cechy i innowacje	13
4.3	Proces treningu i implementacji	13
4.4	Przykłady zastosowań i rezultaty	13
5	Porównanie algorytmów ESRGAN i DWSR	15
5.1	Kryteria porównawcze	15
5.2	Analiza wydajności	15
5.3	Jakość odtwarzania obrazów	15
5.4	Ograniczenia i wyzwania	15
6	Aplikacja webowa do powiększania rozdzielczości obrazów	17
6.1	Projektowanie aplikacji	17
6.2	Wybór narzędzi i technologii	17
6.3	Implementacja aplikacji	17
6.4	Integracja algorytmów DWSR i ESRGAN	17
6.5	Wdrożenie i utrzymanie aplikacji	17
7	Podsumowanie i wnioski	19
7.1	Dyskusja wyników	19
7.2	Rekomendacje i kierunki dalszych badań	19
	Bibliografia	20

Rozdział 1

Wstęp

Cel pracy

Opis celu badań, czyli stworzenia aplikacji webowej służącej do zwiększania rozdzielczości obrazów z użyciem algorytmów ESRGAN i DWSR oraz analiza i porównanie tych algorytmów.

Zakres pracy

Przedstawienie koncepcji i zagadnień, które zostaną omówione w pracy, w tym wybrane metody i technologie.

Rozdział 2

Podstawy teoretyczne

Celem rozdziału jest przedstawienie podstawowych definicji, wytłumaczenie aparatu matematycznego oraz metod wykorzystywanych w algorytmach na których skupia się praca. Dodatkowo ma on na celu ułatwienie dalszego czytania poprzez zapoznanie czytelnika z przyjętymi konwencjami, oznaczeniami oraz symbolami, które mogą pojawić się w kolejnych rozdziałach.

2.1 Definicja super-rozdzielczości

Super-rozdzielczość (ang. Super-Resolution) odnosi się do procesu poprawy rozdzielczości obrazu lub sekwencji obrazów. W kontekście cyfrowym, super-rozdzielczość jest często realizowana za pomocą algorytmów komputerowych, które mają na celu odtworzenie wysokiej rozdzielczości obrazu [Rys 2, 3] z jednego lub wielu obrazów o niskiej rozdzielczości [Rys 1].



Rys 1. Obraz oryginalny



Rys 2. Obraz powiększony czterokrotnie



Rys 3. Obraz powiększony szesnastokrotnie

Przykłady zastosowań super-rozdzielczości

W kontekście praktycznym, techniki super-rozdzielczości znalazły zastosowanie w wielu dziedzinach. Przykładowo, w medycynie, poprawa rozdzielczości obrazów rentgenowskich

może znacząco wpływać na dokładność diagnoz. W bezpieczeństwie publicznym, super-rozdzielczość pozwala na uzyskanie wyraźniejszych obrazów z kamer monitoringu, co jest kluczowe dla identyfikacji osób lub pojazdów. W branży rozrywkowej, technika ta umożliwia remastering starych filmów i materiałów wideo do standardów HD czy 4K.

W ostatnich miesiącach popularne stało się generowanie obrazów z użyciem sztucznej inteligencji na podstawie podanego opisu tekstowego. Generowane w ten sposób obrazy są określonej wielkości i popularne modele takie jak **DALL-E** czy **Midjourney** nie pozwalają na zwiększenie jej. W tym miejscu z pomocą przychodzi super-rozdzielczość, która pozwala na powiększenie obrazu.

2.2 Przegląd metod powiększania obrazów

Istnieje wiele metod powiększania rozdzielczości obrazów. Najprostszą z nich jest **interpolacja najbliższego sąsiada**, która polega na powieleniu pobliskich pikseli w celu zwiększenia rozdzielczości obrazu.

Metoda ta jest bardzo prosta w implementacji, jednakże nie daje ona zadowalających rezultatów. Obraz powiększony w ten sposób wygląda jak obraz o niskiej rozdzielczości z większymi pikselami [Rys 5].



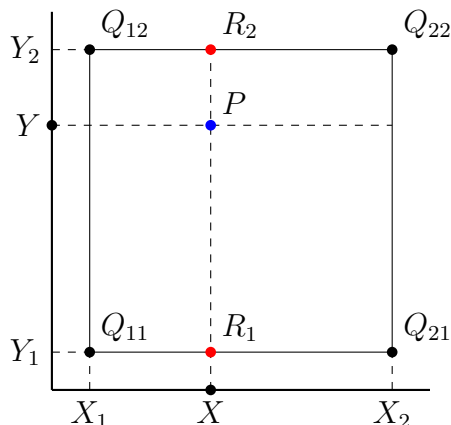
Rys 4. Obraz oryginalny



Rys 5. Obraz powiększony metodą najbliższego sąsiada

Aby poprawić jakość obrazu, można zastosować **interpolację dwuliniową**. Metoda ta rozszerza interpolację liniową na interpolację funkcji dwóch zmiennych [Rys 6].

W efekcie polega to na wyznaczeniu średniej ważonej pikseli sąsiadujących z pikselem, który chcemy powielić. Współczynniki wag są wyznaczane na podstawie odległości od piksela, który chcemy powielić.



Rys 6. Wizualizacja interpolacji dwuliniowej

Kroki algorytmu:

1. Przeprowadzana jest interpolacja liniowa wzdłuż osi OX :

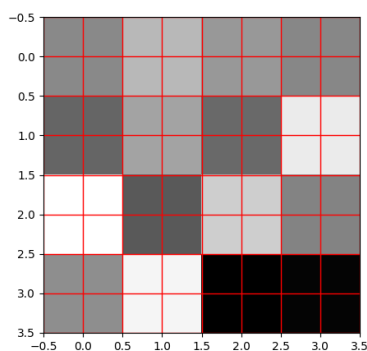
$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}) \quad \text{gdzie} \quad R_1 = (x, y_1),$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}) \quad \text{gdzie} \quad R_2 = (x, y_2).$$

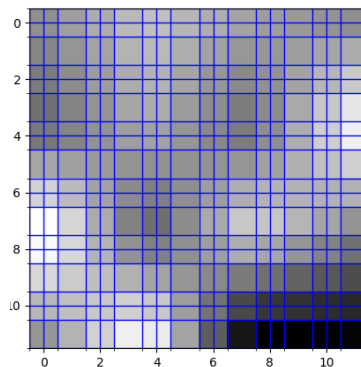
2. Następnie przeprowadzana jest interpolacja wzdłuż osi OY :

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2).$$

W efekcie otrzymujemy obraz wyglądający następująco [Rys 8, 10].



Rys 7. Obraz wejściowy



Rys 8. Interpolacja dwuliniowa

Metoda ta daje lepsze rezultaty niż interpolacja najbliższego sąsiada, jednakże wprowadziła ona duże rozmycie, które jest szczególnie widoczne na krawędziach obiektów i obszarach wysokiej częstotliwości.

W teorii informacji istnieje koncepcja zwana **nierównością przetwarzania danych**. Zgodnie z nią niezależnie od sposobu przetwarzania danych, **nie można dodać informacji, której nie ma w oryginale**. Oznacza to, że brakujących danych nie można odzyskać poprzez dalsze przetwarzanie. Czy to oznacza, że super-rozdzielczość jest teoretycznie niemożliwa?

Nie, jeśli mamy dodatkowe źródło informacji.



Rys 9. Obraz wejściowy



Rys 10. Obraz powiększony metodą interpolacji dwuliniowej

2.3 Wprowadzenie do głębokiego uczenia się w przetwarzaniu obrazów

Głębokie uczenie rewolucjonizuje przetwarzanie obrazów, wprowadzając modele zdolne do uczenia się cech z serii danych. W przetwarzaniu obrazów, głębokie sieci neuronowe są wykorzystywane do zadań takich jak detekcja obiektów, segmentacja, klasyfikacja obrazów, czy właśnie super-rozdzielczość.

Sieć neuronowa może nauczyć się odtwarzać szczegóły obrazów na podstawie pewnych informacji, które zbiera z dużego zbioru obrazów.

Szczegóły dodawane do powiększanego obrazu w przy użyciu modelu uczenia maszynowego nie naruszają nierówności przetwarzania danych, ponieważ wykorzystywane informacje są w zbiorze treningowym, nawet jeśli nie ma ich na obrazie wejściowym.



Rys 11. Obraz wejściowy



Rys 12. Obraz powiększony algorytmem DWSR



Rys 13. Obraz powiększony algorytmem ESRGAN

2.4 Wstęp do funkcji falkowych

Omówienie funkcji falkowych, opis do czego to narzędzie służy począwszy od transformaty Fouriera i jej ograniczeń, w jaki sposób funkcje falkowe rozwijają FFT, przedstawienie działania.

Rozdział 3

DWSR: Deep Wavelet Super Resolution

[1]

3.1 Architektura DWSR

Dokładne przedstawienie struktury i funkcjonowania sieci DWSR, podkreślając jej unikalne cechy i mechanizmy.

3.2 Kluczowe cechy i innowacje

Dyskusja na temat głównych innowacyjnych rozwiązań zastosowanych w DWSR i ich wpływu na efektywność metody.

3.3 Proces treningu i implementacji

Wyjaśnienie procedur związanych z treningiem DWSR, z uwzględnieniem specyfikacji danych, procesu uczenia i kwestii implementacji.

3.4 Przykłady zastosowań i rezultaty

Ilustracja praktycznych zastosowań DWSR oraz ocena i interpretacja osiągniętych dzięki niemu wyników.

Rozdział 4

ESRGAN

[2]

4.1 Architektura ESRGAN

Szczegółowy opis architektury sieci ESRGAN, w tym jej głównych komponentów i zasady działania.

4.2 Kluczowe cechy i innowacje

Omówienie innowacji wprowadzonych w ESRGAN i w jaki sposób różnią się one od wcześniejszych podejść.

4.3 Proces treningu i implementacji

Opis procesu treningu sieci ESRGAN, w tym zbierania danych, uczenia oraz wyzwań implementacyjnych.

4.4 Przykłady zastosowań i rezultaty

Prezentacja przykładów, gdzie ESRGAN został użyty oraz analiza wyników, jakie osiągnięto dzięki tej technologii.

Rozdział 5

Porównanie algorytmów ESRGAN i DWSR

5.1 Kryteria porównawcze

Ustalenie kryteriów, które będą stosowane do oceny i porównania skuteczności i efektywności algorytmów super rozdzielczości.

5.2 Analiza wydajności

Bezpośrednie porównanie wydajności obu metod w różnych warunkach, bazujące na ustalonych kryteriach.

5.3 Jakość odtwarzania obrazów

Ocena jakości obrazów generowanych przez oba algorytmy, uwzględniając różne aspekty jakości wizualnej.

5.4 Ograniczenia i wyzwania

Dyskusja na temat ograniczeń obu metod i potencjalnych wyzwań w ich stosowaniu.

Rozdział 6

Aplikacja webowa do powiększania rozdzielczości obrazów

6.1 Projektowanie aplikacji

Wy tłumaczenie wyboru określonych technologii i narzędzi użytych do stworzenia aplikacji webowej.

Projekt interfejsu użytkownika

Omówienie procesu projektowania interfejsu użytkownika, w tym wytycznych ergonomii i użyteczności.

6.2 Wybór narzędzi i technologii

6.3 Implementacja aplikacji

Opis technicznego procesu integracji wybranych algorytmów z aplikacją, wraz z napotkanymi wyzwaniami.

6.4 Integracja algorytmów DWSR i ESRGAN

6.5 Wdrożenie i utrzymanie aplikacji

Omówienie procesu wdrożenia gotowej aplikacji oraz planów dotyczących jej przyszłego utrzymania i aktualizacji.

Rozdział 7

Podsumowanie i wnioski

7.1 Dyskusja wyników

Krytyczna analiza uzyskanych wyników w kontekście celów pracy oraz istniejących badań i literatury w dziedzinie.

7.2 Rekomendacje i kierunki dalszych badań

Sugestie dotyczące potencjalnych ulepszeń i obszarów, które wymagają dalszych badań, w oparciu o obserwacje i wyniki badań.

Literatura

- [1] T. Guo, H. S. Mousavi, T. H. Vu, V. Monga. Deep wavelet prediction for image super-resolution. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017.
- [2] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C. C. Loy. Esrgan: Enhanced super-resolution generative adversarial networks. *The European Conference on Computer Vision Workshops (ECCVW)*, September 2018.