

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
(повна назва)

Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський)

_____ Інтелектуальна система аналізу даних атомно-силової мікроскопії
_____ клітин
(тема)

Виконав:
студент 2 курсу, групи _____ СШМ-21-1
_____ Таласімова К.М.
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки
_____ (код і повна назва спеціальності)

Тип програми _____ освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
_____ (повна назва спеціалізації)

Керівник _____ проф. Музика К.М.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

_____ В.О. Філатов
(прізвище, ініціали)

2023 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 122 Комп'ютерні науки
(код і повна назва)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Таласімової Ксенії Миколаївни
(прізвище, ім'я, по батькові)

1. Тема роботи Інтелектуальна система аналіз даних атомно-силової мікроскопії клітин

- затверджена наказом університету від 31 березня 2023 р. № 306Ст
2. Термін подання студентом роботи до екзаменаційної комісії 18 травня 2023 р.
3. Вихідні дані до роботи Науково-технічні публікації, статті Інтернет-джерел, документація PyTorch та numpy, цифрові зображення атомно-силової мікроскопії

4. Перелік питань, що потрібно опрацювати в роботі Аналіз предметної області та постановка задачі; Огляд існуючих методів відновлення зображень атомно-силової мікроскопії; Дослідження глибинних нейронної мереж для заповнення пропусків у зображеннях; Експериментальне дослідження архітектури та параметрів нейронних мереж.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) _____

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)


Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної області	04.04.2023	виконано
2	Постановка задачі	08.04.2023	виконано
3	Дослідження методів обробки зображень АСМ	15.04.2023	виконано
4	Дослідження обраної моделі	20.04.2023	виконано
5	Написання пояснювальної записки	25.04.2023	виконано
6	Попередній захист	15.05.2023	виконано
7	Захист перед ЕК	18.05.2023	

Дата видачі завдання 3 квітня 2023 р.

Студент _____


(підпис)

Керівник роботи _____

(підпис)

проф. Музика К.М.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 66 с., 26 рис., 1 дод., 26 джерел.

АТОМНО-СИЛОВА МІКРОСКОПІЯ, ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ,
ГЛИБИННЕ НАВЧАННЯ, ІНТЕРПОЛЯЦІЯ, КОНВОЛЮЦІЙНІ МЕРЕЖІ,
ШТУЧНА НЕЙРОННА МЕРЕЖА.

Об'єкт дослідження – методи глибинного навчання для відновлення зображень.

Предмет дослідження – методи підвищення якості зображень атомно-силової мікроскопії.

Мета роботи – підвищення якості зображень атомно-силової мікроскопії.

Методи дослідження – аналіз існуючих методів глибинного навчання для відновлення пропущених даних в зображеннях атомно-силової мікроскопії, порівняння ефективності запропонованих методів з традиційними, обробка та аналіз отриманих результатів.

У кваліфікаційній роботі досліджено можливості використання штучного інтелекту для розв'язання проблеми пропусків в зображеннях атомно-силової мікроскопії, розглянуто різні методи глибокого навчання для відновлення пропущених даних в зображеннях, такі як генеративні моделі з підсиленням, автокодувальні мережі та глибокі конволюційні мережі.

На основі результатів був проведений аналіз швидкодії запропонованого підходу та обчислені метрики для оцінки якості класифікації.

ABSTRACT

Explanatory note: 67 p., 26 fig., 1 add., 26 sources.

ARTIFICIAL NEURAL NETWORK, ATOMIC FORCE MICROSCOPY, CONVOLUTIONAL NETWORKS, DEEP LEARNING, IMAGE RESTORATION, INTERPOLATION.

Object of research – methods of deep learning for image restoration.

The subject of research is methods for improving the quality of atomic force microscopy images.

Purpose – to improve the quality of atomic force microscopy images.

Research methods – analysis of existing deep learning methods for restoring missing data in atomic force microscopy images, comparison of the effectiveness of the proposed methods with traditional ones, processing and analysis of the results.

The qualification work investigates the possibilities of using artificial intelligence to solve the problem of gaps in atomic force microscopy images, considers various deep learning methods for recovering missing data in images, such as generative models with reinforcement learning, auto-coding networks, and deep convolutional networks.

The main disadvantages of existing approaches without the use of deep learning were highlighted.

Based on the results, the performance of the proposed approach was analyzed and metrics were calculated to assess the quality of classification.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень та термінів.....	7
Вступ.....	8
1 Аналіз предметної області та постановка задачі.....	10
1.1 Аналіз математичних методів.....	11
1.1.1 Метод вагового найменшого квадрата.....	12
1.1.2 Метод максимальної ентропії.....	14
1.2 Аналіз нейромережових підходів.....	16
1.2.1 Згорткові нейронні мережі.....	17
1.2.2 Генеративні змагальні мережі.....	21
1.3 Постановка задачі дослідження.....	25
2 Дослідження використання глибоких нейронних мереж для відновлення зображень.....	27
2.1 Методи на основі CNN.....	28
2.2 Deep Image Prior.....	44
3 Експериментальне дослідження архітектури та параметрів обраної мережі	52
3.1 Структура енкодера	53
3.2 Структура декодера	54
3.3 Пропуск з'єднань	55
3.4 Активаційна функція та функція втрат	56
3.5 Навчання моделі.....	57
3.6 Оцінка результатів.....	60
Висновки.....	62
Перелік джерел посилання.....	63
Додаток А Відомість кваліфікаційної роботи.....	66

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ ТА ТЕРМІНІВ

АСМ – Атомно-силова мікроскопія;

CNN – convolutional neural network – згорткова нейрона мережа;

GAN – generative adversarial network – генеративно змагальна мережа;

ReLU – rectified linear unit – випрямлена лінійна одиниця;

RNN – recurrent neural network – рекурентна нейрона мережа;

WLS – weighted Least Squares – метод найменших квадратів.

ВСТУП

Атомно-силова мікроскопія (АСМ) є одним з найпотужніших методів для сканування, зображення, вимірювання та аналізу структури поверхні на нанорівні. Атомно-силовий мікроскоп здатний отримувати зображення з розташуванням окремих атомів або бачити структуру окремих молекул. До появи АСМ для аналізу структур використовували скануючий тунельний мікроскоп, розроблений Біннінгом і Рорер на початку 1980-х років. У 1986 році Біннінг, Куейт і Гербер винайшли перший атомно-силовий мікроскоп. АСМ став комерційно доступною у 1989 році. Зараз АСМ використовують для структурного аналізу в різних галузях, включаючи хімію, біологію, фізику, матеріалознавство, нанотехнології та медицину. Три переваги сприяють швидкому розвитку технології АСМ для різних проблемних областей, зокрема:

- отримання зображень з тривимірною інформацією;
- сканування структур не обов'язково проводити у вакуумі або в інших екстремальних умовах (АСМ може працювати як у повітрі, так і в рідині);
- АСМ не обов'язково проводити за певних умов; отже, не потрібно спеціальної обробки зразків, яка може призвести до зміни або руйнування зразків, що досліджуються за допомогою АСМ.

Всі методи вимірювання та інструменти, що використовуються вченими та інженерами для розвитку досліджень АСМ і контролю якості, призводять до наслідків, які можуть мати певні спотворення.

Різні інструменти та експериментальні ситуації можуть викликати різні типи дефектів, включаючи небажані високочастотні або низькочастотні шуми, нерівномірний фон тощо. Якщо користувачі точно знають, що шукати і джерело спотворень, деякі дефекти, ймовірно, можна виявити і уникнути їх. Навіть якщо деякі дефекти є неминучими, усвідомлення їх існування на АСМ-зображенні також допоможе запобігти

неправильної інтерпретації їх як справжніх особливостей зображення, а це означає, що розпізнавання дефектів на АСМ-зображенні займає важливе місце в методі АСМ.

Одним з можливих дефектів, який може виникнути, є пропуски у зображеннях, які виникають у наслідок пропущених даних. Для відновлення зображень АСМ, раніше використовувались методи інтерполяції та фільтрації, однак ці методи не є досить ефективними для відновлення детальної структури матеріалів на атомному рівні. З цією метою, в даний час, активно досліджується використання методів глибокого навчання для замальовування пропусків на зображеннях АСМ.

Однак, проблема замальовування пропусків на зображеннях АСМ має свої виклики та обмеження. Наприклад, зображення АСМ може містити багато різних масштабів, оскільки матеріали можуть мати різну форму та розмір, тому необхідно розглядати можливість використання методів, які можуть працювати з даними різної роздільної здатності.

Ця задача може бути вирішені з допомогою штучного інтелекту. Зокрема, можливо застосувати методи глибокого навчання, такі як автокодувальні мережі та глибокі конволюційні мережі, для відновлення зображень АСМ з пропусками. Крім того, можливо розробити нові методи навчання та оцінки результатів для забезпечення ефективності та точності відновлення зображень.

Таким чином, дослідження замальовування пропусків на зображеннях АСМ з використанням штучного інтелекту може мати великий потенціал для поліпшення точності та якості отриманих даних, що може знайти своє застосування в різних галузях, таких як матеріалознавство, фізика, електроніка та біологія.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

Атомно-силова мікроскопія (АСМ) є потужним інструментом для вивчення поверхні матеріалів на нанометровому рівні. АСМ працює за принципом вимірювання взаємодії між атомами на поверхні зразка та зондом, що рухається по ньому. Однак, якщо зразок має нерівну або неоднорідну поверхню, то АСМ зображення може містити пропуски, що призводить до зменшення якості та точності досліджень [1].

Відновлення пропусків у зображеннях – це процес заповнення відсутніх або пошкоджених областей на зображенні за допомогою алгоритмів або методів глибокого навчання. Відновлення пропусків у зображеннях є важливою задачею в багатьох областях, таких як медична діагностика, фотографія, космічні дослідження та ін. Існує безліч математичних методів для відновлення пропущеного зображення з використанням різних підходів та алгоритмів.

Існує безліч методів, що використовуються для відновлення пропущеного у зображення, кожен з цих методів має математичну основу. Ці методи класифікуються відповідно до алгоритмів, які вони використовують для отримання бажаного результату. Першими з цих методів є традиційні методи, які намагаються виправити відсутню частину на головному заголовку зображення лише за допомогою особливостей на зображенні, що підлягають виправленню.

Методи машинного навчання, які часто використовуються в багатьох різних проблемах завдяки збільшенню обчислювальної потужності і широкому доступу до великих масивів даних, також почали використовувати для вирішення завдань розфарбовування зображень. Ці ідеї, які будуть пояснюватися під назвою традиційні методи, також лягли в основу методів машинного навчання і глибокого навчання, що використовуються для розфарбовування зображень.

Було б точніше розглядати традиційні методи в основному як

засновані на зразках і дифузії. Хоча методи на основі зразків поділяються на піксельні та патч-методи, ми розглядатимемо переважно патч-методи, оскільки існує більше досліджень, пов'язаних з патч-методами, і вони дають кращі результати та є одними з традиційних робіт, присвячених прикладам методів [2].

Традиційні методи – це методи, які не потребують навчання. З цієї причини їх доцільно використовувати для простіших задач. Традиційні методи погано працюють в задачах, де замаскована область на зображенні починає утворювати пропорційно велику площу. Цей результат є цілком нормальним, враховуючи, що ці методи роблять висновки на основі пікселів, розташованих поблизу відсутньої області. Пропуски у зображеннях атомно-силової мікроскопії зазвичай мають невелику площу та розташовані пропорційно по зображенню, тож математичні методи можуть також бути застосовані для вирішення описаної вище задачі.

1.1 Аналіз математичних методів

1.1.1 Метод вагового найменшого квадрата

Метод вагового найменшого квадрата (Weighted Least Squares method) є одним з найпоширеніших методів для відновлення пропусків у зображеннях. Цей метод використовується для відновлення зображень, які мають високий рівень шуму або інші дефекти, такі як білий шум або артефакти, що виникають при компресії зображення.

Метод вагового найменшого квадрата базується на ідеї мінімізації середньо-квадратичної помилки між відновленим зображенням та пошкодженим зображенням. При використанні цього методу кожен піксель зображення розглядається окремо, і для кожного пікселя шукається оптимальне значення як вагована сума його сусідів з врахуванням їх відстаней від даного пікселя.

Загалом, метод вагованого найменшого квадрата можна розбити на наступні кроки:

- підготовка даних: починаємо з пошкодженого зображення, для якого треба відновити пропущені пікселі. Дані можуть бути також доповнені змішуванням з іншими джерелами, щоб зменшити шум та покращити точність відновлення;

- обчислення матриці ваг: на цьому кроці знаходять відстані між кожним пікселем та його сусідами та обчислюють ваги для кожної пари пікселів на основі їх відстані. Ці ваги використовуються для обчислення вагової суми сусідів кожного пікселя;

- система рівнянь: після обчислення матриці ваг та вагової суми сусідів для кожного пікселя, можна сформулювати систему рівнянь для всього зображення.

Після того, як система рівнянь була сформована, можна розв'язати її за допомогою методу вагованого найменшого квадрата. Цей метод полягає в мінімізації функції середньо-квадратичної помилки, яка визначається як:

$$E = \sum_{i=1}^n w_i (x_i - \hat{x}_i), \quad (1.1)$$

де x_i – значення пікселя, що розглядається;

\hat{x}_i – відновлене значення пікселя;

w_i – вага для пікселя i .

Ця функція мінімізує суму квадратів відхилень між відновленими та пошкодженими значеннями пікселів, де вага використовується для зменшення впливу шуму та інших дефектів на процес відновлення зображення.

На рисунку 1.1 зображено приклад видалення шумів із зображення за допомогою методу вагованого найменшого квадрата.

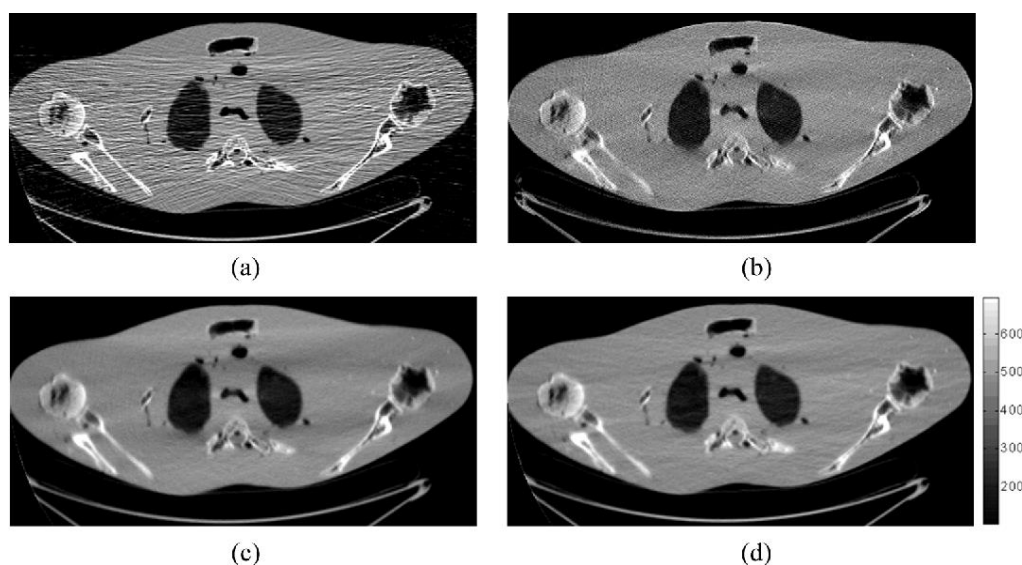


Рисунок 1.1 – Видалення шуму з зображення за допомогою WLS

Оновлення зображення: після того, як було знайдено оптимальні значення для кожного пікселя, вони можуть бути використані для оновлення пошкодженого зображення. Оновлене зображення буде мати менше шуму та дефектів, що дозволить отримати кращу якість зображення.

Вагові функції використовуються для того, щоб зменшити вплив шуму та дефектів на процес відновлення зображення. Найчастіше використовується різні види вагових функцій, такі як гаусівські, експоненціальні та різні комбінації цих функцій. Для кожного пікселя в зображенні вагова функція визначає відносну важливість його значення у процесі відновлення.

Після того, як зображення було відновлено, необхідно оцінити результати. Це можна зробити за допомогою таких метрик, як середньоквадратична помилка, структурний схожість, піксельна

розбіжність та інші. Ці метрики допоможуть оцінити, наскільки добре було відновлено зображення.

Загалом, метод вагованого найменшого квадрата є ефективним методом відновлення зображення, який може бути використаний для відновлення різних типів пошкоджень, таких як шум, розмивання, затемнення та інші. Він дозволяє досягнути високої якості відновлення зображення за допомогою використання вагових функцій та мінімізації функції середньо-квадратичної помилки.

1.1.2 Метод максимальної ентропії

Метод максимальної ентропії – це статистичний підхід до реконструкції зображення, який прагне знайти найбільш ймовірне зображення, враховуючи наявні дані і набір обмежень. У контексті АСМ наявні дані, як правило, складаються з набору вимірних положень і висот зонда при скануванні поверхні зразка.

Основна ідея методу максимальної ентропії полягає в тому, щоб знайти зображення, яке максимізує ентропію з урахуванням наявних даних і будь-яких додаткових обмежень. Іншими словами, метод прагне знайти зображення, яке є найбільш випадковим, або найменш структурованим, з урахуванням обмежень. Це робиться шляхом розв'язання оптимізаційної задачі з використанням множників Лагранжа.

Задача оптимізації може бути сформульована наступним чином:

$$\text{minimize } E(x) = -\sum p_i \log p_i \quad (1.2)$$

де $E(x)$ – ентропія зображення x ;

p_i – ймовірнісний розподіл зображення.

Метод максимальної ентропії успішно застосовується для реконструкції АСМ-зображень. Однією з ключових переваг цього методу є те, що він може працювати з неповними даними і зашумленими вимірюваннями. Він також може враховувати попередні знання про зображення, такі як його гладкість або розрідженість, шляхом додавання відповідних обмежень до задачі оптимізації [3].

Процес реконструкції зображення методом максимальної ентропії можна розділити на кілька етапів:

1) попередня обробка даних: Сирі дані АСМ зазвичай попередньо обробляються для видалення шуму і корекції різних артефактів, таких як дрейф сканера і ефекти згортки вістря;

2) нормалізація даних: Виміряні дані нормалізуються, щоб переконатися, що вони задовольняють обмеженням задачі оптимізації;

3) оцінка розподілу ймовірностей: Розподіл ймовірностей зображення оцінюється на основі наявних даних і будь-яких попередніх знань про зображення;

4) оптимізація: Задача оптимізації вирішується за допомогою ітераційного алгоритму, наприклад, методу спряжених градієнтів або методу Ньютона-Рафсона;

5) постобробка: Реконструйоване зображення піддається постобробці для видалення залишків шуму та артефактів.

Загалом, метод максимальної ентропії є потужним інструментом для реконструкції АСМ-зображень. Його здатність працювати з неповними і зашумленими даними, а також враховувати попередні знання про зображення робить його універсальним і ефективним методом для широкого спектру застосувань.

Окрім переваг, метод максимальної ентропії має й певні обмеження. Одне з них полягає в тому, що він вимагає хорошої оцінки розподілу ймовірностей зображення. Якщо оцінка погана, реконструйоване зображення може бути неточним. Іншим обмеженням є те, що метод може

бути дорогим в обчислювальному плані, особливо для великих зображень і складних обмежень.

Для подолання цих обмежень були запропоновані різні розширення та модифікації методу максимальної ентропії. Наприклад, цей метод можна комбінувати з іншими методами реконструкції зображень, такими як метод повної варіації або метод нелокальних засобів, щоб підвищити його точність і ефективність. Крім того, підходи машинного навчання, такі як глибокі нейронні мережі, можуть бути використані для вивчення розподілу ймовірностей зображення безпосередньо з даних.

Отже, метод максимальної ентропії є потужним і гнучким підходом до реконструкції зображень в АСМ. Його здатність працювати з неповними і зашумленими даними, а також враховувати попередні знання про зображення робить його цінним інструментом для широкого спектру застосувань. З подальшим розвитком і вдосконаленням метод має потенціал для покращення нашого розуміння та аналізу поверхонь на нанометровому рівні.

1.2 Аналіз нейромережових підходів

Архітектури глибокого навчання широко використовуються для розфарбовування зображень завдяки їхній здатності навчатися і моделювати складні патерни і структури на зображеннях. Однією з популярних архітектур є генеративна змагальна мережа (Generative Adversarial Network, GAN), яка складається з мережі-генератора, що генерує зображення, і мережі-дискримінатора, що розрізняє справжні і фальшиві зображення. Мережа-генератор навчається генерувати зображення, схожі на істинні зображення, тоді як мережа-дискримінатор навчається розрізняти справжні та фальшиві зображення.

Інша популярна архітектура – це згорткова нейронна мережа (CNN), яка є різновидом нейронної мережі прямого поширення, що використовує згорткові шари для вилучення ознак із зображень.

1.2.1 Згорткові нейронні мережі

Зафарбовування зображень за допомогою CNNs має численні застосування в комп'ютерному зорі, такі як відновлення зображень, видалення об'єктів і доповнення зображень. Відновлення зображень передбачає видалення шуму і артефактів із зображень, тоді як видалення об'єктів передбачає видалення небажаних об'єктів із зображень. Доповнення зображень передбачає заповнення відсутніх або пошкоджених частин зображень, що корисно в різних додатках, таких як редагування відео, медична візуалізація та автономне водіння.

У медичній візуалізації зафарбовування зображень можна використовувати для заповнення відсутніх даних на медичних зображеннях, таких як МРТ або КТ, щоб підвищити їхню точність і діагностичну цінність. В автономному керуванні підфарбовування зображень може використовуватися для заповнення відсутніх або закритих частин вхідних зображень, щоб поліпшити продуктивність системи сприйняття [4].

У цій архітектурі використовуються такі структури, як згортковий шар, об'єднуючий шар, згладжуючий шар. ШНМ – це структури, які виділяють лише важливі особливості зображень, застосовуючи процес згортки з використанням двовимірних фільтрів і отримуючи різний рівень інформації на зображенні. CNN використовує переваги ієрархічних патернів у даних, йдучи шляхом, відмінним від MLP-структури. Він об'єднує невеликі та прості патерни для отримання більш складних патернів. Замість звичайного множення матриць використовується операція згортки.

У шарах згортки 2D-згортка виконується між тензором і заданим фільтром. У таких завданнях, як малювання, де використовуються кольорові RGB-зображення, ці фільтри застосовуються окремо до кожного колірної каналу. У процесі згортки при застосуванні цих фільтрів, які зазвичай мають ядро розміром 3x3, 5x5 або 7x7 на матриці, на границях застосовується процес прокладки. Під час цього процесу підбиття визначає порядок дій, коли потрібно отримати доступ до значень, які виходять за межі і не існують, за допомогою операції згортки (рис. 1.2).

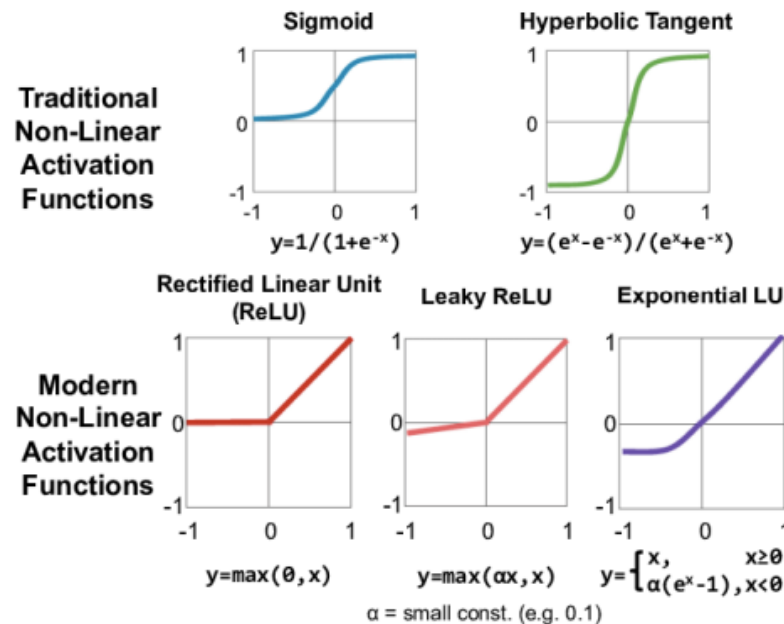


Рисунок 1.2 – Активаційна функція

Шари об'єднання, які зазвичай слідує за згортковими шарами, зменшують розмір результуючого тензора. У багатьох згорткових нейронних мережах використовується максимальне об'єднання, але існують також шари з мінімальним та середнім об'єднанням. Заданий тензор розбивається на менші частини і отримується новий, менший тензор, беручи максимальне, мінімальне або середнє значення кожної частини відповідно до обраного методу.

Шар згладжування – це шар, який робить тензори в структурах CNN векторизованими для того, щоб використовувати повністю зв'язані шари, які зазвичай використовуються в структурах MLP. Цей шар серіалізує вхідний тензор і заповнює всі значення нейронів у шарі. Існує дві архітектури CNN, які зазвичай використовуються для дослідження глибоких кольорів. Перша – це структури кодер-декодер, а друга - похідна від першої, яка називається структурою автокодера.

Ці архітектури добре справляються з багатьма задачами комп'ютерного зору, а також показують багатообіцяючі результати в інших обернених задачах комп'ютерного зору, таких як розфарбовування, знебарвлення та надвисока роздільна здатність. Крім того, ці структури можуть працювати над різними завданнями глибокого навчання, такими як аналіз настрою, підписи до зображень і переклад. Наприклад, Google Translate побудований на основі структурі кодер-декодер.

1.2.1.1 Encoder-decoder мережа

Кодувальна мережа може мати повністю зв'язну структуру, згорткову структуру або навіть структуру рекурентної нейронної мережі (RNN). Кодуюча мережа отримує вхідні дані і видає на виході карту ознак, яка є просто вектором або, в більш загальному випадку, тензором.

Модель цієї архітектури можна побачити на рисунку 1.3.

Цей вектор ознак містить необхідну та важливу інформацію про вхідні дані. Мережа декодера, як правило, має таку ж архітектуру, як і кодер, але у зворотному порядку. Вона бере вектор ознак і намагається згенерувати вхідний сигнал якомога ближче до нього.

Кодерно-декодерна мережа складається з двох частин мережі - кодера та декодера. Вхідні дані обробляються через ці структури. Ці мережеві структури можуть бути використані в різних сферах, таких як переклад мови і в додатках комп'ютерного зору як генеративні моделі.

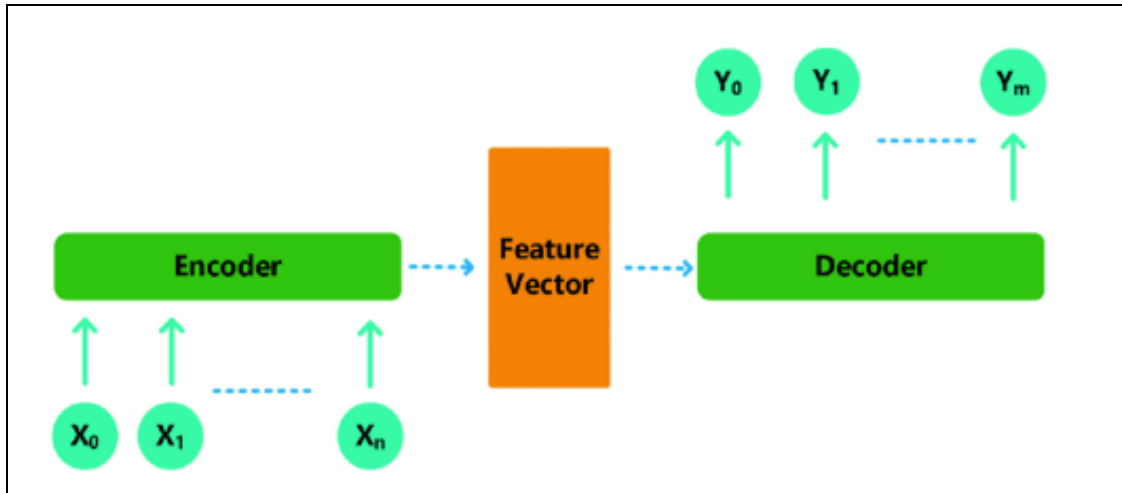


Рисунок 1.3 – Структура Encoder-decoder мережі

1.2.1.2 Автокодувальні мережі

Автокодер – це структура, метою якої є встановлення зв'язку між входом і виходом шляхом отримання коду в прихованих шарах. Поняття коду тут можна вважати подібним до вектора ознак у мережевих структурах кодер-декодер [5].

Найбільш типовим застосуванням автокодерів є використання автокодерів – це зменшення розмірності. Структуровані нейронні мережі на основі автокодерів можуть використовуватися в неконтрольованих програмах нелінійного зменшення розмірності.

На рисунку 1.4 показано порівняння між традиційним методом зменшення розмірності PCA та структурою автокодера.

Завдяки навчанню автокодерів стає можливим ігнорувати шум у даних. Наприклад, це можна використовувати для виявлення аномалій у будь-якому сигналі.

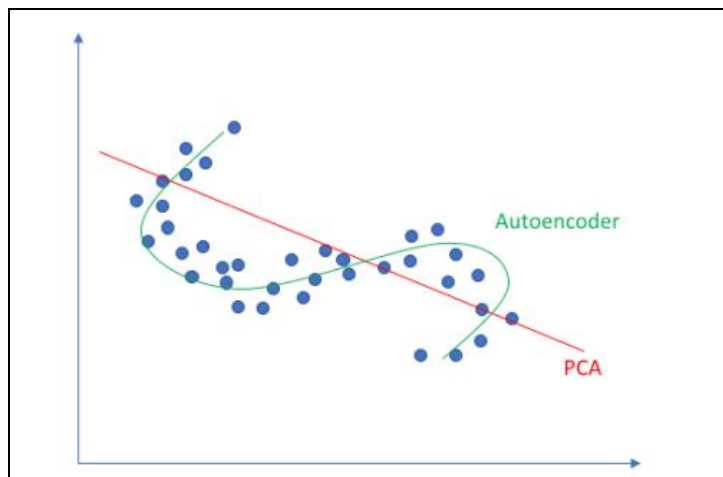


Рисунок 1.4 – Порівняння автокодера та метода зменшення розмірності PCA

Іншим прикладом застосування може бути створення безшумного зображення із зашумленого зображення шляхом отримання найважливіших ознак, ігноруючи шум. Таке застосування називається денойзингом, що є популярною оберненою задачею в дослідженнях комп'ютерного зору. Навчання за допомогою автокодерних структур насправді можливе за допомогою кодера, який може якнайкраще описати вхідне зображення, і декодера, який може дати на виході зображення, максимально схоже на максимально схожий на вхідне зображення, використовуючи цей опис від попередньої структури кодера.

Структура мережі згорткового автокодера, призначеного для класифікації зображень зображено на рисунку 1.5.

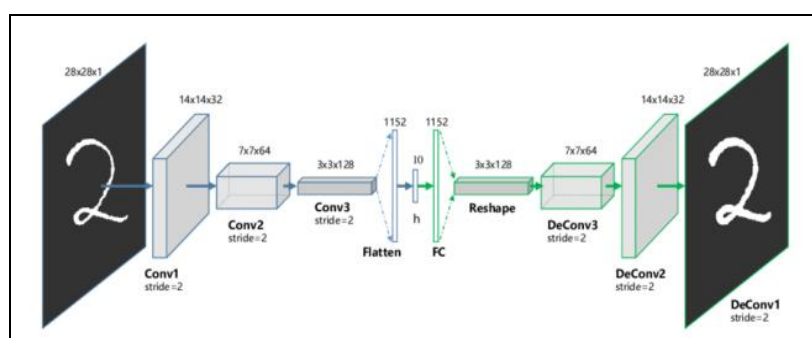


Рисунок 1.5 – Структура мережі згорткових автокодерів

1.2.2 Генеративні змагальні мережі

Генеративні змагальні мережі (GAN) зробили революцію в обробці зображень, забезпечивши потужну основу для створення і маніпулювання зображеннями. GAN складаються з двох основних компонентів: мережі-генератора і мережі-дискримінатора. Мережа-генератор приймає на вхід випадковий шум і генерує синтетичні зображення, тоді як мережа-дискримінатор має на меті відрізнити справжні зображення від фальшивих.

На рисунку 1.6 наведено структуру GAN мережі.

Архітектура GAN з часом еволюціонувала, що призвело до появи різних інноваційних конструкцій. Однією з популярних архітектур є Deep Convolutional GAN (DCGAN), яка використовує згорткові шари для створення зображень і маніпуляцій з ними у вищій роздільній здатності. DCGAN довела свою ефективність у створенні реалістичних зображень завдяки фіксації просторових залежностей і вивченню багатих представлень об'єктів.

Інша архітектура, відома як умовний GAN (cGAN), розширює базову структуру GAN, обумовлюючи генератор і дискримінатор додатковою інформацією, такою як мітки класів або атрибути зображень. Це дозволяє генерувати певні типи зображень або перекладати зображення між різними доменами. cGAN знайшли застосування в таких завданнях, як переклад зображень між собою і семантична сегментація [6].

CycleGAN – це визначна архітектура, яка дозволяє здійснювати неконтрольований переклад зображень. Вона вивчає відображення між двома доменами, не вимагаючи парних навчальних даних, що робить її придатною для таких завдань, як передача стилю, адаптація домену та художні трансформації. CycleGAN продемонстрував вражаючі можливості у перетворенні зображень зі збереженням їхнього змісту та структури.

Прогресивне зростання GAN – це архітектура, яка поступово збільшує роздільну здатність згенерованих зображень під час навчання.

Починаючи з зображень низької роздільної здатності і поступово додаючи деталі, цей підхід дозволяє отримати високоякісні результати з високою роздільною здатністю. Прогресивні GAN демонструють чудову продуктивність у створенні реалістичних зображень і використовуються в таких додатках, як генерація обличчя.

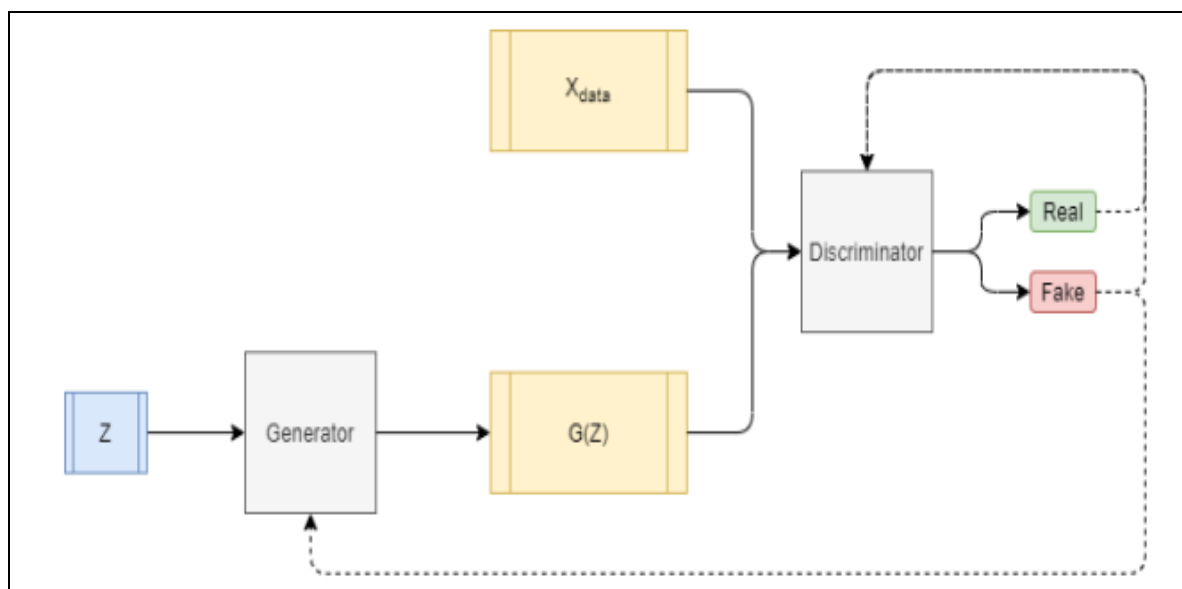


Рисунок 1.6 – Структура GAN мережі

Pix2Pix – це архітектура GAN, розроблена для задач перекладу зображень, де вона вивчає відображення вхідного зображення у вихідне. Це дає змогу виконувати такі завдання, як переклад зображень, коли один тип зображення перетворюється на інший. Pix2Pix використовується в різноманітних додатках, включаючи семантичну сегментацію, перетворення країв у фотографії та генерацію ескізів у зображення.

Однією з головних переваг GAN в обробці зображень є їхня здатність генерувати високоякісні зображення, які майже неможливо відрізнити від реальних зображень. Це призвело до широкого застосування в різних галузях, включаючи мистецтво, моду і навіть медицину. Наприклад, GAN використовуються для створення реалістичних зображень предметів одягу для веб-сайтів електронної комерції, для

створення синтетичних медичних зображень для навчання і перевірки алгоритмів машинного навчання, а також для створення реалістичних зображень для відеоігор і середовищ віртуальної реальності.

Архітектура GAN також добре адаптується, дозволяючи модифікувати і розширювати її для різних застосувань. Наприклад, умовна GAN (cGAN) використовується для завдань перекладу зображення в зображення, наприклад, для створення зображень високої роздільної здатності з вхідних даних низької роздільної здатності. У cGAN мережі генератора та дискримінатора залежать від додаткової вхідної інформації, наприклад, семантичних міток або інших зображень, що дає змогу мережі вивчати складніші відображення.

Іншою відомою архітектурою GAN є CycleGAN, яку можна використовувати для неконтрольованого перекладу зображення в зображення між двома доменами. На відміну від інших GAN, CycleGAN не вимагає парних навчальних даних, що дозволяє мережі вивчати відображення між двома доменами з різними характеристиками. CycleGAN використовується для різних застосувань, наприклад, для перетворення зображень із зимових на літні або з денних на нічні.

Незважаючи на переваги GAN, вони також стикаються з деякими проблемами. Однією з головних проблем є нестабільність навчання, коли мережі генератора і дискримінатора можуть застрягти в неоптимальному стані. Це може призвести до неякісної або повторюваної генерації зображень, явища, відомого як колапс режиму. Іншою проблемою є потреба у великомасштабних навчальних наборах даних, які може бути важко отримати для певних застосувань.

Нещодавні дослідження в галузі ШНМ були зосереджені на вирішенні цих проблем та покращенні продуктивності ШНМ у різних завданнях. Наприклад, дослідники запропонували нові функції втрат, методи регуляризації та стратегії навчання для покращення стабільності та якості GAN. Вони також розробили нові архітектури, такі як прогресивне

зростання GAN (PGGAN) і GAN на основі стилю (StyleGAN), щоб генерувати високоякісні зображення з високою точністю і роздільною здатністю.

Однак GAN також стикаються з певними проблемами, зокрема нестабільністю навчання, розпадом режимів і потребою у великих наборах даних для навчання. Тим не менш, нещодавні досягнення в дослідженнях GAN вирішили ці проблеми і відкрили нові можливості для обробки зображень на основі GAN.

GAN знайшли численні застосування в обробці зображень. Вони широко використовуються для генерації зображень, зокрема для створення фото реалістичних зображень різних об'єктів, сцен і навіть людських обличчя.

GAN також застосовуються для зафарбовування зображень, що передбачає заповнення відсутніх або пошкоджених частин зображень для відновлення їхнього початкового вигляду.

Крім того, GAN відіграють важливу роль у завданнях надвисокої роздільної здатності, де вони підвищують роздільну здатність зображень, щоб генерувати високоякісні зображення з вхідних даних низької роздільної здатності.

1.3 Постановка задачі дослідження

Атомна силова мікроскопія (АСМ) – це потужний інструмент для вивчення найдрібніших структур матеріалів на атомному рівні. Однак, як і будь-який інструмент, вона має свої обмеження, такі як пропуски в зображеннях, що можуть бути причиною помилкового аналізу даних.

У даній роботі пропонується ввести до розгляду методи на основі глибинного навчання для відновлення і покращення якості зображень атомно-силової мікроскопії.

Головним завданням магістерського проєкту є наступне: розробити

архітектуру та алгоритм навчання глибинної нейронної мережі для покращення якості зображень атомно-силової мікроскопії.

Задля реалізації поставленої мети передбачається розглянути наступний набір питань:

- опрацювати літературу та інтернет-джерела по темі алгоритм навчання та оптимізації нейронних мереж;
- розглянути існуючі математичні методи обробки зображень АСМ;
- розробити архітектуру нейронних мереж з використанням методів глибинного навчання для обробки зображень;
- на основі розробленої архітектури побудувати нейронну мережу;
- підготувати навчальний набір даних;
- провести експерименти з розробленою архітектурою;
- оцінити точність, швидкість навчання розробленої архітектури;
- порівняти отримані результати з результатами застосування математичних методів обробки зображень;
- проаналізувати отримані результати.

2 ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ ГЛИБИННИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ

Як і в багатьох інших задачах комп'ютерного зору, методи глибокого навчання використовуються для відтворення зображень. Причина, чому методи глибокого навчання використовуються все частіше, полягає в тому, що вони дають більш успішні результати у вирішенні складних завдань порівняно з традиційними методами. Основною причиною цього є створення великомасштабних наборів даних, які дозволяють навчати глибокі методи, а також обчислювальні потужності, які дозволяють навчати ці глибокі методи. обчислювальних потужностей, які дозволяють навчати ці глибокі методи.

2.1 Методи на основі CNN

В останні роки для розфарбовування зображень було запропоновано кілька різних архітектур згорткових нейронних мереж (CNN). Ці архітектури відрізняються за своєю складністю, кількістю шарів та типами використовуваних шарів. У цьому розділі ми порівняємо деякі з найпопулярніших архітектур для відтворення зображень, зокрема U-Net, DeepFillv2 та Context Encoder.

U-Net – популярна архітектура для зафарбовування та сегментації зображень. Архітектура U-Net складається зі шляху, що звужується, і шляху, що розширюється, що дозволяє мережі захоплювати як низькорівневі, так і високорівневі особливості [7]. У випадку АСМ-зображень звужуючий шлях може складатися з декількох шарів згортки і шарів максимального об'єднання, які зменшують просторову роздільну здатність вхідного зображення, збільшуючи при цьому кількість карт ознак. Шлях, що розширюється, може складатися з декількох згорткових шарів і шарів з максимальною дискретизацією, які збільшують просторову

роздільну здатність вхідного зображення, зменшуючи кількість карт ознак. Шлях, що звужується, і шлях, що розширюється, з'єднані між собою пропусками, які дозволяють мережі зберігати інформацію про місцезнаходження відсутніх пікселів на АСМ-зображеннях.

Архітектура U-Net зображена на рисунку 2.1.

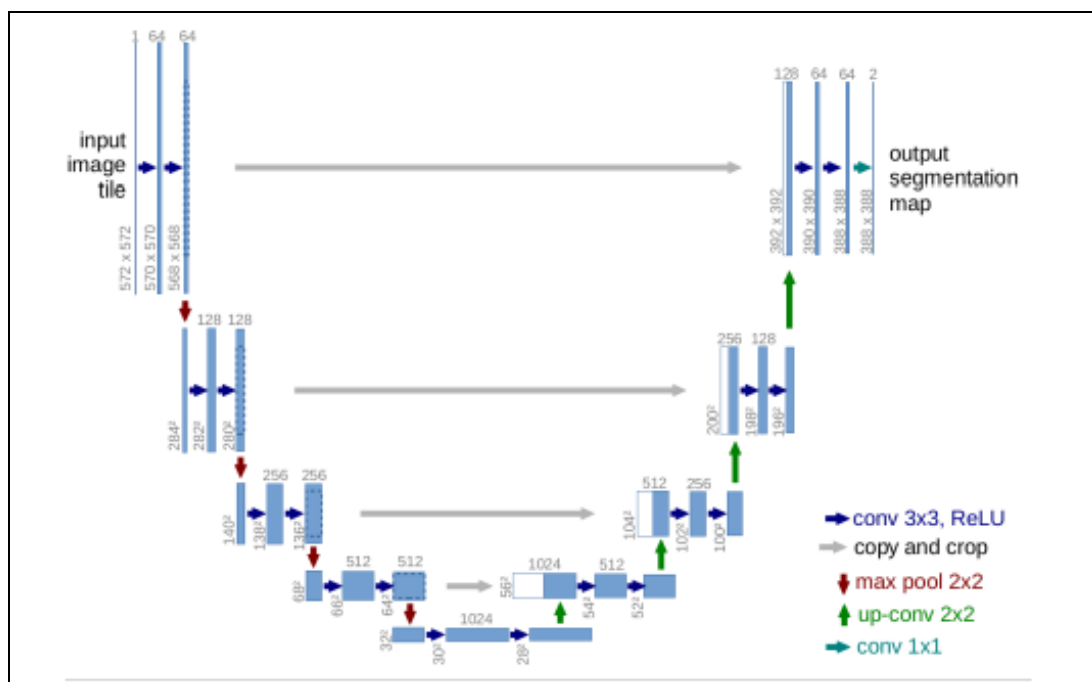


Рисунок 2.1 – Архітектура U-Net

Ця архітектура бере зображення і маску, яка показує відсутні ділянки, і розкладає їх на згорнуті шари.

Існують різні архітектури ШНМ, які використовуються для відтворення зображень атомно-силової мікроскопії. Одним із поширених підходів є використання повністю згорткових нейронних мереж (FCN), які були адаптовані для задач розфарбовування зображень. У дослідженні Ванг та ін. (2018) вони використовували модифіковану архітектуру U-Net для розфарбовування зображень АСМ. U-Net – це тип FCN, який зазвичай використовується в задачах сегментації зображень.

У своїй модифікованій архітектурі U-Net Ван та ін. (Wang et al., 2018) додали новий шар, названий шаром злиття, який поєднує ознаки, отримані на шляху, що згортається, з ознаками, отриманими на шляху, що розширюється. Вони також використовували залишкові блоки на шляху звуження, щоб дозволити мережі навчитися краще представляти вхідне зображення. Автори повідомили, що їхня модифікована архітектура U-Net перевершила інші методи відтворення зображень за показниками індексу структурної схожості (SSIM) та пікового співвідношення сигнал/шум (PSNR).

Вона об'єднує вихід кожного шару з відповідним шаром того ж розміру в симетричній архітектурі. Результати цієї структури досягають чудового успіху з точки зору структури та деталізації згенерованого зображення. Архітектуру Shift-Net, яка являє собою U-Net зі спеціальними зсувними з'єднаннями, можна побачити на рисунку 2.2.

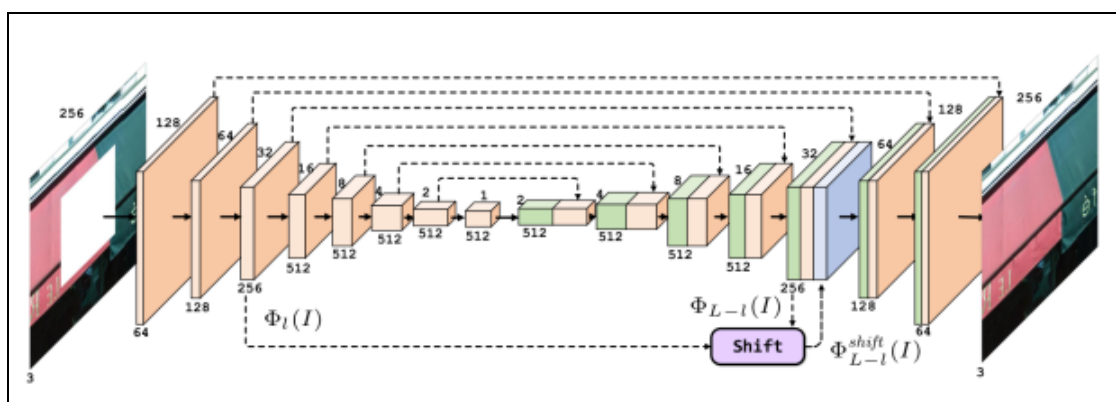


Рисунок 2.2 – Архітектура Shift-Net

DeerFillv2 – це архітектура глибокої нейронної мережі для заповнення зображень, яка складається з низки згорткових шарів, залишкових блоків і деконволюційних шарів. Архітектура базується на фреймворку залишкового навчання, який дозволяє мережі вивчати залишкові зображення, що додаються до вхідного зображення. У випадку АСМ-зображень залишкові блоки можуть складатися з декількох

згорткових шарів і з'єднання пропуску, яке оминає згорткові шари. Пропускне з'єднання дозволяє мережі вивчати залишки, які є більш точними та ефективними, ніж ті, що вивчаються традиційними згортковими нейронними мережами. Ця архітектура може створювати високоякісні та візуально приємні зображення. Архітектура DeepFillv2 зображена на рисунку 2.3.

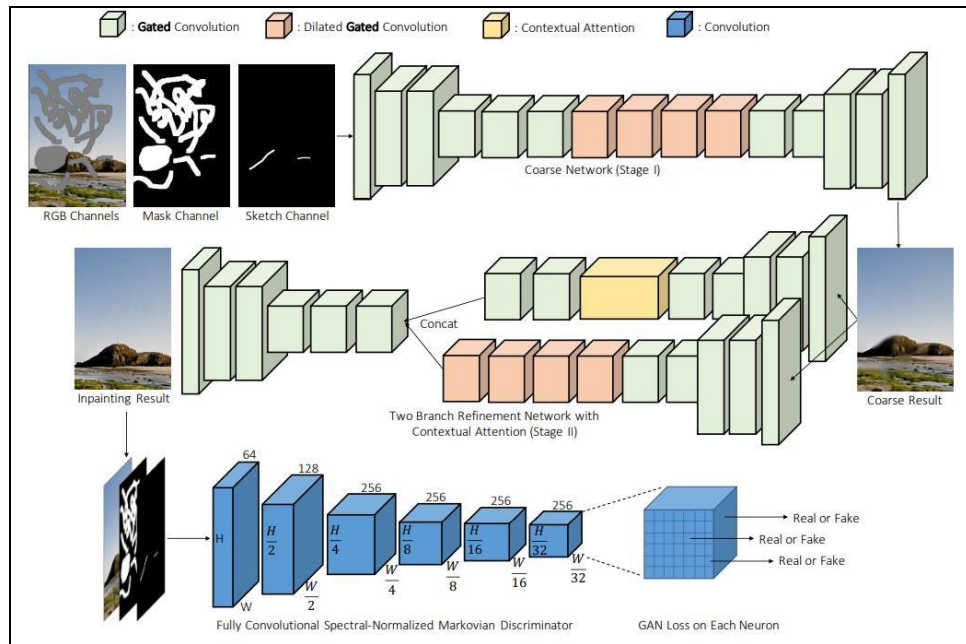


Рисунок 2.3 – Архітектура DeepFillv2

Глибокі автокодера – це нейромережеві архітектури, які складаються з кодера та декодера. Кодер приймає вхідне зображення і кодує його в низькорозмірне представлення, тоді як декодер приймає це представлення і декодує його назад у зображення з повною роздільною здатністю. У контексті відтворення зображень кодер навчений кодувати як повні, так і неповні зображення, тоді як декодер навчений заповнювати відсутні частини неповних зображень [8].

Однією з популярних архітектур глибоких автокодерів є згортковий автокодер. Ця архітектура використовує згорткові шари для вивчення просторових особливостей і текстур вхідних зображень і добре підходить

для задач, пов'язаних з обробкою зображень. Кодер зазвичай складається з декількох згорткових та об'єднуючих шарів, за якими слідує один або декілька повністю з'єднаних шарів. Декодер будується за дзеркальною архітектурою кодера, використовуючи транспоновані згорткові шари замість об'єднаних шарів для підвищення частоти дискретизації карт ознак.

На рисунку 2.4 зображено відновлення зображень за допомогою автокодера.

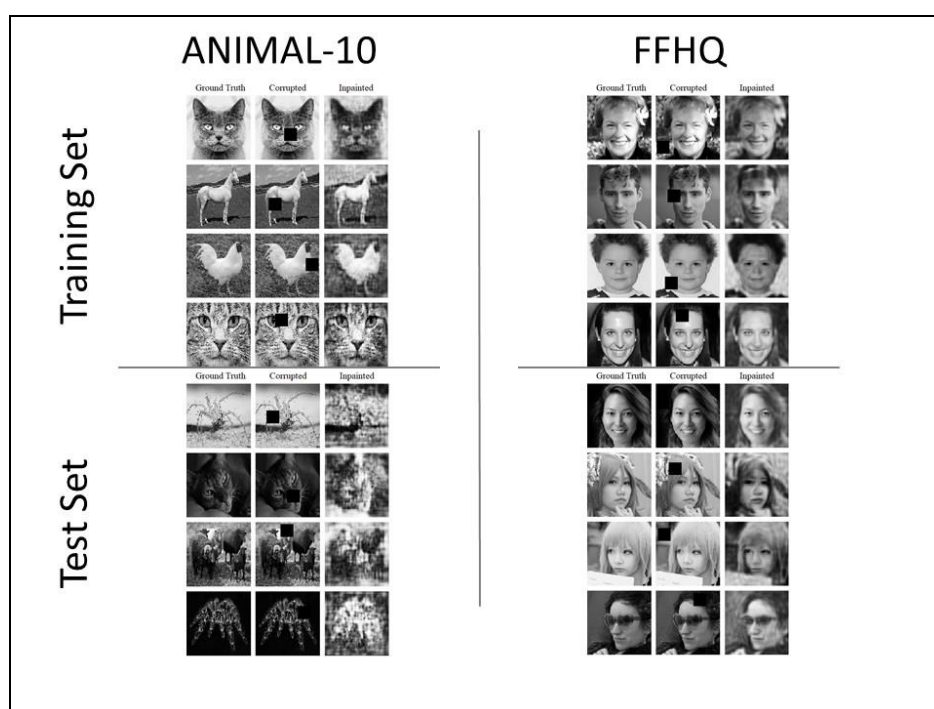


Рисунок 2.4 – Відновлення зображень за допомогою автокодера

У контексті розфарбовування ACM-зображень було запропоновано кілька варіацій згорткового автокодера. Наприклад, Лі та ін. (2018) запропонували автокодер з глибоким згортковим згладжуванням (CDAE) для розфарбовування ACM-зображень. CDAE складається з низки згорткових і деконволюційних шарів і навчається відновлювати зображення, пошкоджені гауссівським шумом. Автори повідомили, що CDAE досягнув кращої якості реконструкції, ніж інші методи, включаючи

традиційні методи розфарбовування зображень та інші моделі глибокого навчання.

Інша архітектура, яка була використана для розфарбовування зображень АСМ, – це глибока згорткова нейронна мережа (DCNN). Ван та ін. (2019) запропонували підхід на основі DCNN для розфарбовування зображень АСМ, який складається з серії згорткових і об'єднуючих шарів, за якими слідує кілька повністю з'єднаних шарів. Автори повідомили, що їхній підхід дозволив отримати високоякісні реконструкції з низьким рівнем помилок, навіть при роботі зі складними текстурами і структурами на вхідних зображеннях.

Зовсім недавно деякі дослідники досліджували використання генеративних змагальних мереж (GAN) для розфарбовування зображень АСМ. GAN складаються з двох нейронних мереж, генератора і дискримінатора, які навчаються в ігровій формі для створення високоякісних синтетичних зображень, які важко відрізнити від реальних зображень. Генератор у GAN схожий на декодер у глибокому автокодері, але він навчений генерувати реалістичні на вигляд зображення, а не просто заповнювати відсутні частини неповного зображення.

Незважаючи на успіх глибоких автокодерів і GAN для розфарбовування АСМ-зображень, все ще існують певні проблеми, які потребують вирішення.

Однією з головних проблем є обчислювальні витрати на навчання цих моделей, які можуть бути непомірно високими для великих наборів даних. Крім того, якість зафарбованих зображень може залежати від розміру і складності відсутніх ділянок, а також від специфічних характеристик АСМ-зображень.

У Artist-Net, запропонованому Ляо та ін., для досягнення розфарбовування використовуються два різних кодера контенту, кодер стилю та спільний декодер. Архітектуру Artist-Net можна побачити на рисунку 2.5.

Подібна структура також доступна в архітектурі CNN, розробленій Cai та ін. [9] для видалення семантичних об'єктів. Існують також архітектури контекстних кодерів [10], що використовуються у дослідженнях живопису – це архітектура CNN для розфарбовування зображень, яка використовує глибоку згорткову нейронну мережу для передбачення відсутніх пікселів на зображенні.

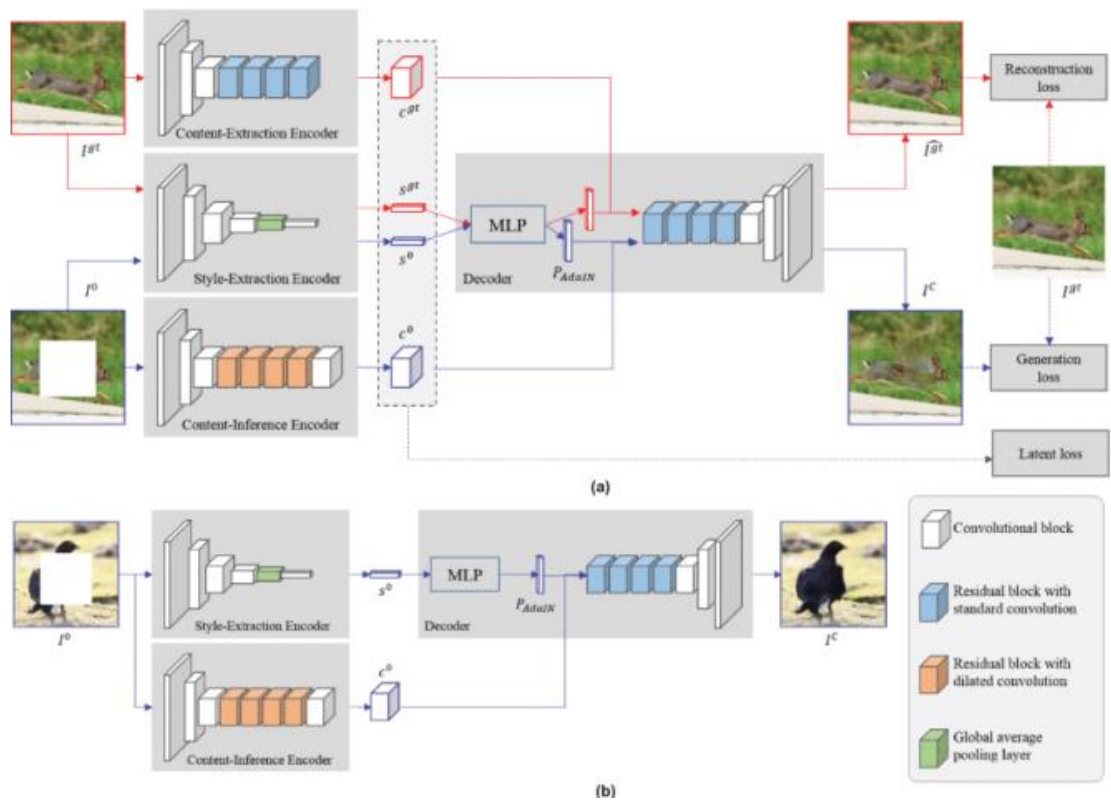


Рисунок 2.5 – Архітектура Artist-Net

Архітектура складається з декількох згорткових шарів, за якими слідують кілька деконволюційних шарів, що дозволяє мережі фіксувати як низькорівневі, так і високорівневі особливості вхідного зображення.

У випадку АСМ-зображень архітектура може бути модифікована для включення особливостей, характерних для АСМ-зображень. Наприклад, мережу можна навчити розпізнавати характерні особливості

АСМ-зображень, такі як лінії та краї, що може допомогти підвищити точність процесу зафарбовування.

Однією з переваг використання CNN для розфарбовування та реконструкції зображень є їхня здатність фіксувати просторові залежності на зображеннях. Це важливо для зображень атомно-силової мікроскопії, які часто мають складні структури і патерни, що вимагають високого рівня просторового розуміння. Ще однією перевагою є їхня здатність обробляти зображення з високою роздільною здатністю, що має вирішальне значення в контексті атомно-силової мікроскопії [11].

Методи надвисокої роздільної здатності – це ще один клас методів, які використовуються для розфарбовування та реконструкції зображень. Ідея надвисокої роздільної здатності полягає у відновленні зображень високої роздільної здатності з вхідних зображень низької роздільної здатності. Це робиться шляхом вивчення функції відображення з простору зображень з низькою роздільною здатністю у простір зображень з високою роздільною здатністю. Функцію відображення можна вивчити за допомогою різних моделей глибокого навчання, зокрема CNN.

Використання методів надвисокої роздільної здатності для зображень атомно-силової мікроскопії може бути корисним у випадках, коли зображення мають низьку роздільну здатність або містять недостатню інформацію. Методи надвисокої роздільної здатності можуть бути використані для відновлення відсутньої інформації та покращення загальної якості зображення. Однак методи надвисокої роздільної здатності можуть бути дорогими в обчислювальному плані, а якість результатів може бути обмежена якістю вхідних зображень [12].

Порівняно з автокодерами, CNN і методи надвисокої роздільної здатності мають кілька переваг у контексті розфарбовування і реконструкції зображень для атомно-силової мікроскопії. Хоча автокодери можуть бути ефективними у вивченні базової структури вхідних зображень, вони можуть мати труднощі з фіксацією просторових

залежностей високого рівня, присутніх на зображеннях атомно-силової мікроскопії. Крім того, автокодери можуть потребувати значних обсягів навчальних даних для ефективного вивчення цих залежностей.

На відміну від них, CNN і методи надвисокої роздільної здатності краще підходять для обробки просторових залежностей високого рівня і можуть працювати з зображеннями високої роздільної здатності. Вони також здатні навчатися на меншій кількості навчальних даних, що робить їх більш практичними для використання в атомно-силовій мікроскопії, де збір великих обсягів даних може бути складним завданням.

Одним із суттєвих недоліків використання автокодерів є те, що вони мають тенденцію давати розмиті результати, особливо при роботі з зображеннями АСМ. Це пов'язано з тим, що архітектура кодера-декодера не оптимізована для обробки дрібних деталей і ледь помітних особливостей, присутніх в АСМ-зображеннях.

Одним із способів вирішення цієї проблеми є поєднання методів надвисокої роздільної здатності з методами зафарбовування на основі CNN. Надвисока роздільна здатність передбачає збільшення роздільної здатності зображення шляхом генерування зображень високої роздільної здатності з вхідних даних низької роздільної здатності. Ця техніка може бути використана для покращення деталей і текстур зображення, що призводить до більш чіткого і точного зображення.

Існує кілька методів надвисокої роздільної здатності на основі CNN, які можна використовувати в поєднанні з розфарбовуванням зображень для покращення реконструкції АСМ-зображень. Одним з популярних підходів є використання згорткових нейронних мереж з пропусковими зв'язками, таких як архітектура ResNet. Ці мережі використовують залишкове навчання для подолання проблеми зникаючих градієнтів і дозволяють навчати дуже глибокі мережі.

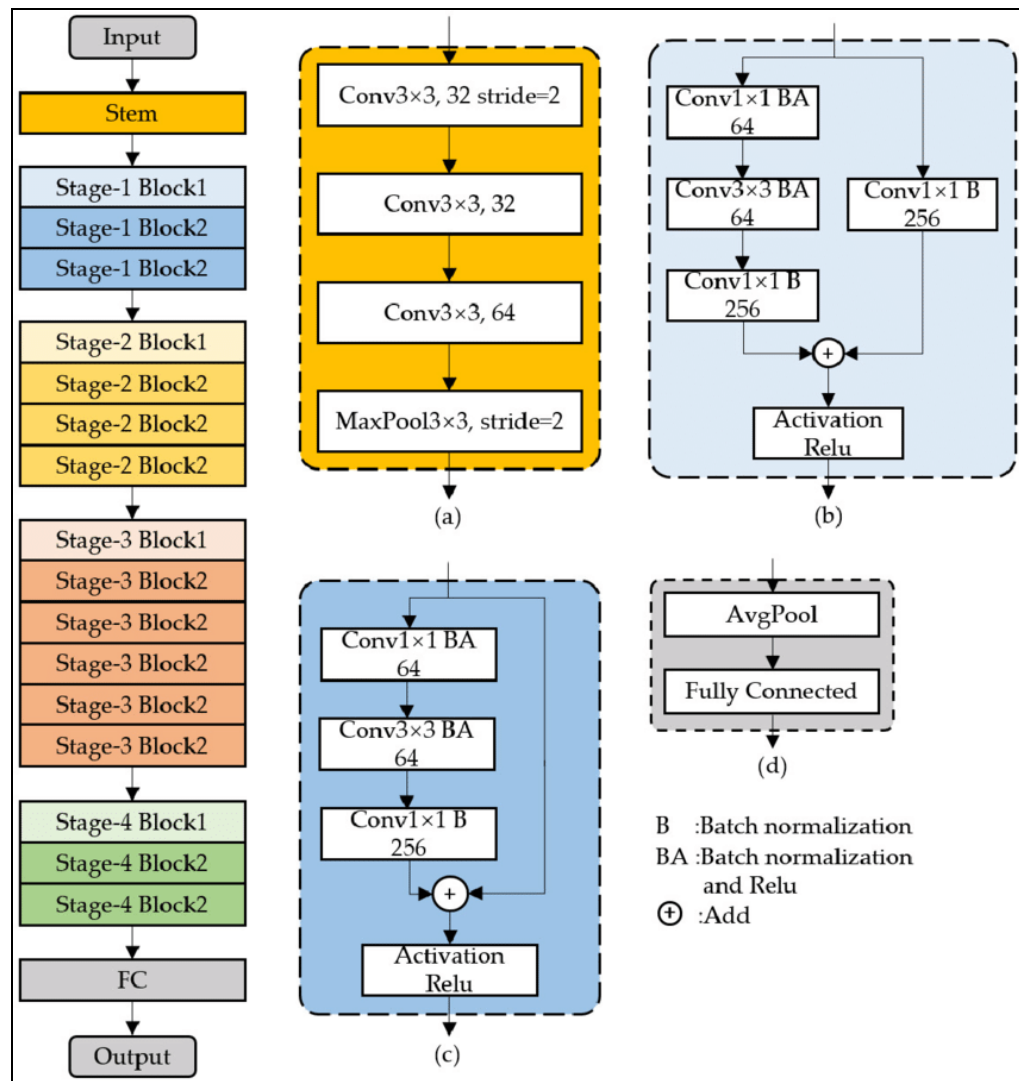


Рисунок 2.6 – Архітектура ResNet

На додаток до архітектури ResNet (рис. 2.6), існують також інші методи надроздільної здатності на основі CNN, які можна використовувати для реконструкції та розфарбовування ACM-зображень. До них відносяться архітектури SRGAN (Super-Resolution Generative Adversarial Networks) і ESPCN (Efficient Sub-Pixel Convolutional Neural Networks).

Архітектура SRGAN використовує генеративну змагальну мережу (GAN) для отримання зображень високої роздільної здатності з вхідних даних низької роздільної здатності. GAN складається з двох мереж: мережі-генератора, яка генерує зображення з високою роздільною здатністю, і мережі-дискримінатора, яка оцінює реалістичність

згенерованих зображень. Генераторна мережа навчена створювати зображення, які неможливо відрізнити від реальних зображень з високою роздільною здатністю, тоді як дискримінаторна мережа навчена розрізняти згенеровані та реальні зображення.

З іншого боку, архітектура ESPCN призначена для оптимізації обчислювальної ефективності мереж надвисокої роздільної здатності. Вона використовує субпіксельні згорткові шари для підвищення роздільної здатності зображень з низькою роздільною здатністю, а також застосовує з'єднання з пропусками для збереження деталей зображення. Ця архітектура дозволяє створювати високоякісні зображення надвисокої роздільної здатності, вимагаючи при цьому менше параметрів і обчислювальних ресурсів порівняно з іншими методами надвисокої роздільної здатності.

Важливим фактором при виборі архітектури на основі CNN для реконструкції та зафарбовування АСМ-зображень є розмір і складність набору даних. Методи глибокого навчання вимагають великих обсягів даних для ефективного навчання, а АСМ-зображення можуть бути складними для роботи через їхню високу роздільну здатність і складну структуру [13].

Щоб вирішити цю проблему, можна використати навчання з перенесенням для використання попередньо навчених моделей, які вже були навчені на великих наборах даних, таких як ImageNet. Трансферне навчання передбачає використання попередньо навченої моделі в якості відправної точки і її точне налаштування на меншому наборі даних, наприклад, на зображеннях АСМ. Такий підхід може допомогти підвищити точність і швидкість процесу навчання, а також зменшити ризик перенавчання.

Іншим важливим моментом є використання методів доповнення даних для збільшення розміру навчального набору даних. Розширення даних передбачає застосування перетворень до оригінальних зображень,

таких як обертання, перевертання та зміна кольору, для створення нових зображень, які схожі, але не ідентичні оригінальним зображенням. Цей підхід може допомогти запобігти надмірному пристосуванню та покращити продуктивність узагальнення моделі.

Загалом, використання архітектур на основі CNN для реконструкції та зафарбовування АСМ-зображень є перспективним напрямком досліджень, що має потенціал для значного підвищення точності та якості АСМ-зображень. Хоча існує кілька архітектур, які довели свою ефективність, кожен підхід має свої сильні і слабкі сторони, і вибір архітектури буде залежати від конкретних вимог програми. Використовуючи методи навчання з перенесенням і доповнення даних, можна навчати моделі на основі ШНМ на менших наборах даних і підвищити їхню точність і продуктивність узагальнення [14].

Механізм уваги може бути включений у ШНМ різними способами. Один із поширених підходів полягає у використанні методу просторової уваги, який передбачає створення карти уваги, що виділяє регіони вхідного зображення, які є найбільш важливими для виконання завдання, що стоїть перед мережею. Карта уваги потім використовується для зважування внесків різних ознак у мережі, що дозволяє мережі більше фокусуватися на відповідних областях зображення.

Інший підхід полягає у використанні техніки, яка називається увагою до каналу, що передбачає створення вектора уваги, який призначає вагу різним каналам на карті ознак. Це може бути особливо корисно для завдань, які передбачають ідентифікацію специфічних особливостей або шаблонів на зображенні.

У порівнянні з підходами надвисокої роздільної здатності, які зосереджені на створенні зображень високої роздільної здатності з вхідних даних низької роздільної здатності, підходи на основі уваги більше зосереджені на виявленні та збереженні важливих деталей на зображенні. Це може бути особливо важливим для АСМ-зображень, які часто містять

складні та детальні особливості, які може бути важко зафіксувати за допомогою традиційних методів надвисокої роздільної здатності.

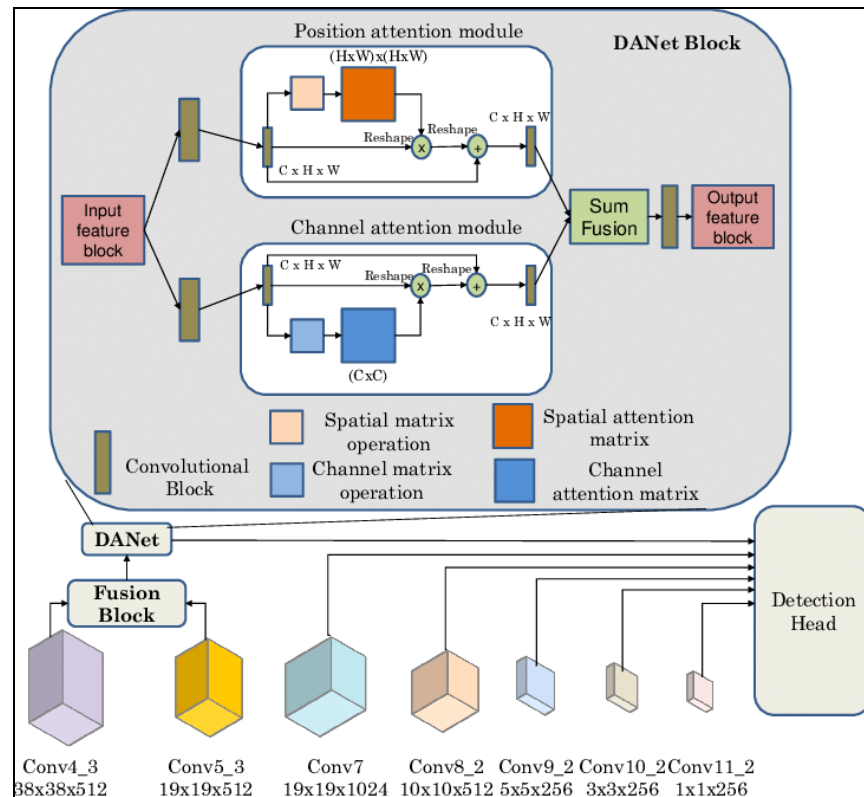


Рисунок 2.7 – Архітектура DaNet

Одним із прикладів підходу, що базується на увазі, для реконструкції та розфарбовування ACM-зображень є глибока уважна нейронна мережа (DaNet), яку запропонували Лю та ін. у 2018 році. DaNet використовує комбінацію механізмів просторової та каналної уваги, щоб допомогти мережі зосередитися на найбільш важливих ділянках вхідного зображення [15]. Архітектура зображена на рисунку 2.7.

Порівняно з підходами на основі автокодування, методи на основі уваги можуть бути ефективнішими у виявленні важливих деталей на зображенні, зменшуючи при цьому обчислювальну складність. Автокодери зазвичай вимагають використання великої кількості параметрів і високих обчислювальних ресурсів для досягнення хороших результатів. На відміну

від них, мережі, що базуються на увазі, можуть бути більш ефективними в обчислювальному плані, досягаючи при цьому порівнянних або навіть кращих результатів.

Одним з обмежень методів, заснованих на увазі, є те, що вони можуть вимагати більше навчальних даних, ніж традиційні методи. Це пов'язано з тим, що механізм уваги ґрунтується на ідентифікації та виділенні специфічних особливостей вхідного зображення, що може бути складніше вивчити за обмеженої кількості навчальних даних. Однак нещодавні досягнення в галузі глибокого навчання та методів доповнення даних допомогли пом'якшити цю проблему [16].

Порівняно з традиційними методами на основі інтерполяції, методи надвисокої роздільної здатності на основі глибокого навчання можуть бути ефективнішими у створенні високоякісних зображень з покращеною деталізацією та роздільною здатністю. Однак вони також можуть бути більш обчислювально інтенсивними і вимагати більших обсягів навчальних даних.

З точки зору порівняння архітектур ШНМ на основі уваги та методів надвисокої роздільної здатності для реконструкції та зафарбовування АСМ-зображень, важко зробити пряме порівняння, оскільки вони мають різні сильні та слабкі сторони. ШНМ, що базуються на увазі, можуть бути більш ефективними для захоплення специфічних особливостей і деталей на зображенні, тоді як методи надвисокої роздільної здатності можуть бути кращими для покращення загальної роздільної здатності і якості зображення. Зрештою, вибір підходу залежатиме від конкретних вимог програми та наявних ресурсів.

Отже, як архітектури CNN на основі уваги, так і методи надвисокої роздільної здатності є перспективними підходами для реконструкції та зафарбовування АСМ-зображень. Використовуючи передові алгоритми і методи глибокого навчання, можна підвищити роздільну здатність, деталізацію і точність АСМ-зображень, що уможливить нові застосування

і відкриття в таких галузях, як матеріалознавство, нанотехнології та біофізика. З продовженням досліджень і розробок, ймовірно, ці підходи стануть все більш потужними і універсальними інструментами для аналізу та інтерпретації АСМ-зображень.

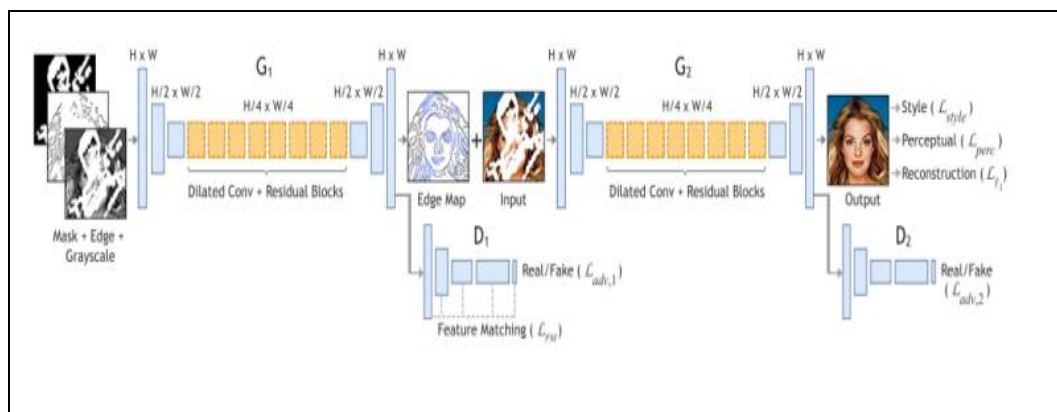


Рисунок 2.8 – Нейронна мережа EdgeConnect

Назері та ін. запропонували метод [17], який використовує граничну інформацію для отримання реалістичних вихідних зображень над відсутніми зонами. EdgeConnect складається з 2 різних GAN, розміщених одна за одною. Перша з цих GAN намагається створити відсутню крайову інформацію над замаскованим зображенням, а друга структура GAN доповнює відсутнє зображення, використовуючи цю створену крайову інформацію та оригінальне замасковане зображення.

Фундаментальна ідея структури показана на рисунку 2.8.

За допомогою цього методу Ю та ін. хочуть вирішити деякі з основних проблем, які виникають у процесі малювання [18]. Використовуючи інформацію про краї, бажано створювати візуальні зображення з більш чіткими межами.

На рисунку 2.9 показано кілька прикладів зображень, зафарбованих за допомогою EdgeConnect.

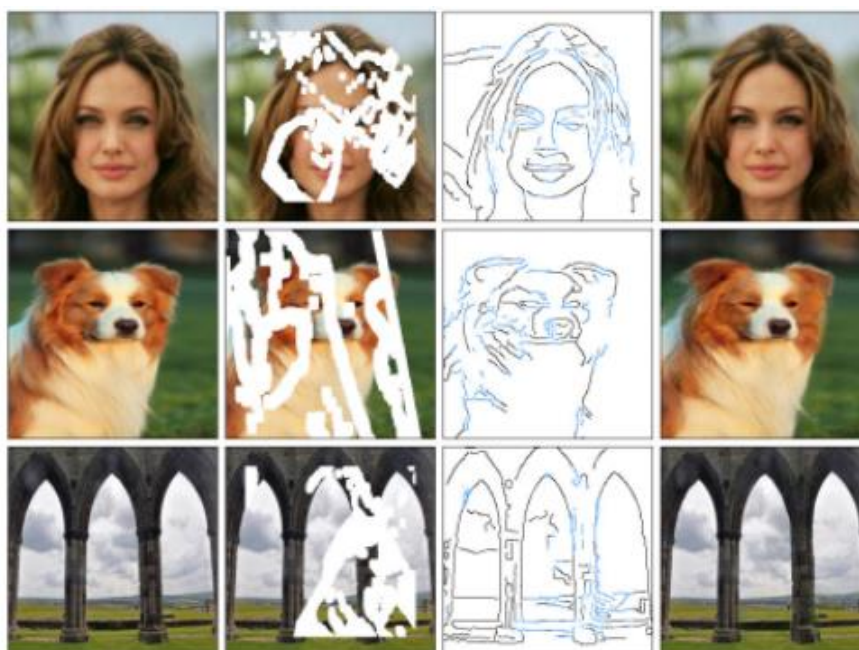


Рисунок 2.9 – Приклад відновлення зображень
за допомогою EdgeConnect

Подібні можливості доступні як для структур, що створюють крайову інформацію, так і для кінцевого зображення. Генераторна мережа складається з 3 частин. Як згадувалося раніше, структура кодера використовується в першу чергу для вилучення ознак із зображення. Залишкова мережа була додана після кодерної структури. Нарешті, є декодер для отримання зображення з елементів. Згорткові шари використовуються поруч для структури дискримінатора [19].

Мережі глибокої згортки стали дуже популярними для задач генерації та відновлення зображень. Загалом, причиною хороших результатів таких мереж є здатність отримувати реалістичні попередні зображення з наборів даних, що містять велику кількість зображень. Коротко пояснюючи, що таке попереднє зображення, можна сказати, що це загальне визначення інформації про зображення, яка може бути використана для будь-яких задач обробки зображень для покращення результатів, вибору параметрів обробки та усунення невизначеностей.

Наприклад, може бути відома деяка інформація про зображення, наприклад, розподіл кольорів, і цю інформацію або її наближення можна використати як попередні дані для виконання певної задачі. Зображення можна представити у математичній формі та об'єднати у такі етапи обробки, як фільтрація, деконволюція та сегментація, що допомагає зменшити кількість можливих рішень.

У статті [20] автори показують, що структура генераторної мережі може отримати достатньо низькорівневої статистики зображень без тривалого навчання на будь-якому наборі даних. Для доведення цієї гіпотези було використано розроблений метод глибокого попереднього аналізу зображень з випадково ініціалізованою згортковою нейронною мережею, і отримано висококонкурентні результати в різних задачах інверсного комп'ютерного зору, таких як розмиття, надвисока роздільна здатність та розфарбовування зображень. Важливість цього методу полягає не тільки в тому, що він може бути використаний у широкому діапазоні досліджень комп'ютерного зору, але й у тому, що він може сформувати нову гілку між мережами, що навчаються, і мережами, що не навчаються, з фіксованими явними пріоритетами, такими як самоподібність і т.д. Якщо говорити одним реченням, то в глибокому попередньому зображенні автори намагаються подолати розрив між двома популярними методами шляхом побудови нового явного попереднього зображення за допомогою згорткової нейронної мережі. Для виконання завдання відновлення зображень, подібних до малювання, дослідники використовують два найпоширеніші методи - навчений та явний пріорити.

Пояснимо на прикладі, що метод навчання з попереднім навчанням – це простий підхід до навчання мережі глибокого згортання, яка вивчає світ за допомогою набору даних. З іншого боку, метод *explicit-prior*, або створений вручну попередній метод, вбудовує обмеження і чітко навчає, які типи зображень є природними. які типи зображень є природними. Однак на практиці надзвичайно складно виразити обмеження

математично. обмеження математично. Як наслідок, більшість методів з явними обмеженнями працюють погано порівняно з сучасними попередньо навченими нейронними мережами, проте метод глибокого попереднього навчання пропонує напрочуд хороші результати, які можуть конкурувати з методами з попереднім навчанням.

2.2 Deep Image Prior

Глибокі згорткові мережі – це архітектури, які є одними з найкращих для розв'язання задач обернених зображень. Такі архітектури, як генеративні змагальні мережі та варіаційні автокодери досягають найсучасніших рішень для цих задач. Однак не зовсім правильно вважати, що причиною дивовижних результатів цих моделей, навчених на великих наборах даних, є те, що вони можуть вловлювати реалістичні попередні зображення під час навчання на даних.

На рисунку 2.10 метод глибокого попереднього аналізу зображень порівняно з нейронними мережами Шепарда [21], які спеціально розроблено для застосування в задачах відтворення малюнків. Використовуючи чорний текст як маску, цей метод демонструє дивовижні результати. У той час як на виході мережі Шепарда можна побачити залишки маски, на виході глибокого зображення попереднього методу майже немає артефактів.

Deep Image Prior (DIP) – це новітня технологія обробки зображень, яка полягає в навчанні глибокої нейронної мережі для реконструкції зображення. На відміну від традиційних методів, які використовують явні моделі для створення зображення, DIP використовує глибоку нейронну мережу для вивчення базової структури зображення, що дозволяє генерувати високоякісні зображення з меншою кількістю артефактів.

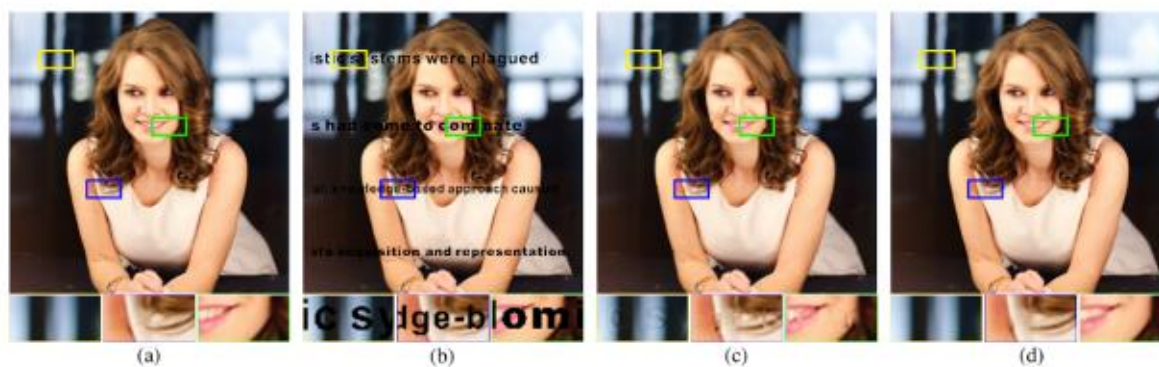


Рисунок 2.10 – Порівняння між мережею Шепарда та Deep Image Prior

Принципи DIP ґрунтуються на ідеї, що глибока нейронна мережа може бути використана для вивчення базової структури зображення. У традиційних методах обробки зображень для створення зображення використовуються явні моделі, що враховують різні обмеження, такі як гладкість, розрідженість і збереження країв. Однак ці моделі часто призводять до генерації зображень з артефактами та недостатньою деталізацією [22].

На рисунку 2.10 зображено приклади зображень, оброблених за допомогою мережі Шепарда та DIP.

DIP, з іншого боку, використовує глибоку нейронну мережу для вивчення структури зображення шляхом ітеративного оновлення вагових коефіцієнтів мережі, поки вона не зможе генерувати зображення, подібне до оригіналу. Це відбувається шляхом мінімізації функції втрат, яка вимірює різницю між згенерованим зображенням і оригінальним. Функція втрат зазвичай базується на відстані L2 або L1 між зображеннями.

Основна перевага DIP полягає в тому, що він не потребує жодних навчальних даних. Мережа навчається, використовуючи лише саме зображення як вхідні дані, що дозволяє легко застосовувати її до різних зображень. Крім того, оскільки DIP є генеративною моделлю, її можна використовувати для створення нових зображень, схожих на оригінальні,

що може бути корисним у таких додатках, як реставрація зображень і передача стилю.

Однак, DIP також має деякі обмеження. Оскільки це генеративна модель, вона може генерувати зображення, які не є фізично правдоподібними (рис 2.11). Крім того, оскільки вона навчається лише на одному зображенні, вона може бути не в змозі добре узагальнити інші зображення, які мають іншу структуру.

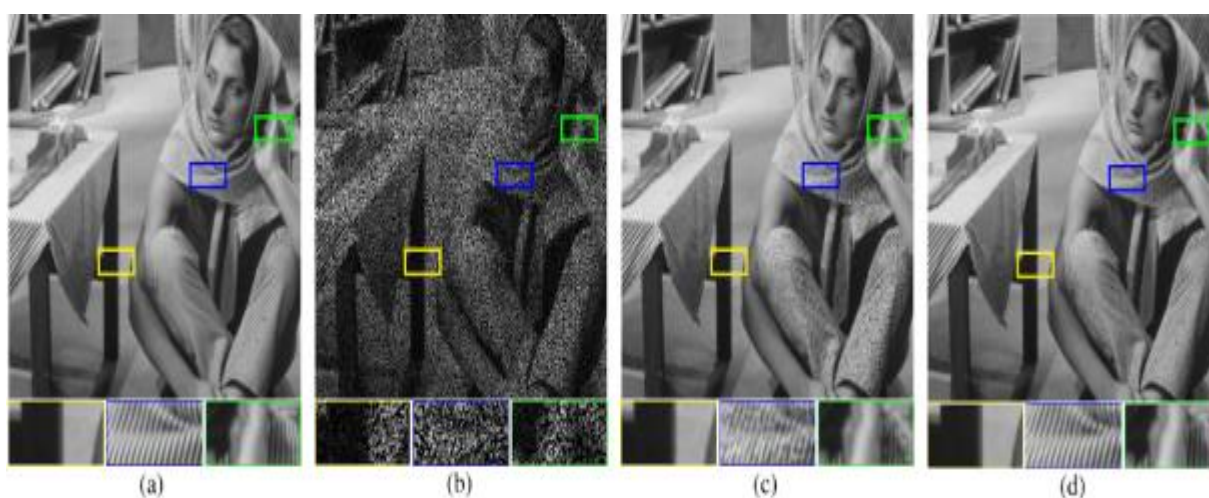


Рисунок 2.11 – Порівняння між Convolutional Dictionary Learning та Deep Image Prior

Одним із застосувань DIP є розфарбовування зображень. Зафарбовування зображень – це процес заповнення відсутніх або пошкоджених ділянок зображення. Це може бути корисним у таких програмах, як реставрація зображень і редагування фотографій [23].

Для застосування DIP до зафарбовування зображення спочатку потрібно вибрати ділянку зображення, яку потрібно зафарбувати. Потім ми використаємо решту зображення як вхідні дані для мережі. Мережа навчена генерувати зображення, схоже на вихідне, але з зафарбованою вибраною областю.

Функція втрат, що використовується в DIP для зафарбовування зображень, зазвичай складається з двох частин: втрат при реконструкції та втрат при сприйнятті. Втрати при реконструкції вимірюють різницю між створеним зображенням і оригінальним зображенням. Втрати сприйняття вимірюють різницю між особливостями згенерованого зображення і особливостями оригінального зображення. Перцептивні втрати зазвичай базуються на попередньо навченій глибокій нейронній мережі, яка була навчена розпізнавати об'єкти та ознаки на зображеннях.

Перевага використання втрати сприйняття в DIP для зафарбовування зображень полягає в тому, що вона може допомогти зберегти структуру і зміст оригінального зображення. Крім того, оскільки DIP є генеративною моделлю, її можна використовувати для створення декількох правдоподібних зафарбованих зображень, що може бути корисним у таких додатках, як редагування фотографій.

Однак, одним з обмежень DIP для зафарбовування зображень є те, що вона не може генерувати зображення, які є фізично правдоподібними. Це пов'язано з тим, що мережа навчається лише на одному зображенні і не може добре узагальнювати інші зображення, які мають іншу структуру. Крім того, оскільки DIP є генеративною моделлю, вона може генерувати зображення, які візуально не схожі на оригінал, що може бути проблемою в додатках, де візуальна схожість є важливою.

Окрім зафарбовування зображень, DIP також можна використовувати в інших програмах обробки зображень, таких як згладжування зображень, надвисока роздільна здатність і видалення розмиття.

При згладжуванні зображень DIP використовується для видалення шуму із зображення шляхом створення згладженого зображення, схожого на оригінальне. Функція втрат, що використовується в DIP для згладжування зображень, подібна до функції втрат, що використовується

при розфарбовуванні зображень, з втратами при відтворенні та втратами при сприйнятті.

У режимі надвисокої роздільної здатності DIP використовується для збільшення роздільної здатності зображення шляхом створення зображення з високою роздільною здатністю, подібного до вхідного зображення з низькою роздільною здатністю. Функція втрат, що використовується в DIP для надвисокої роздільної здатності, схожа на функцію втрат, що використовується для згладжування зображення, але з додатковим членом втрат, який враховує різницю між зображенням високої роздільної здатності і базовим зображенням високої роздільної здатності.

При розмитті DIP використовується для видалення розмиття із зображення шляхом створення розмитого зображення, подібного до оригіналу. Функція втрат, що використовується в DIP для розмиття, подібна до функції втрат, що використовується при розфарбовуванні зображень, з втратами при реконструкції та втратами при сприйнятті.

Однією з головних переваг DIP є те, що він не потребує жодних навчальних даних. Це полегшує його застосування до різних зображень і зменшує потребу у великих наборах даних. Крім того, оскільки DIP є генеративною моделлю, її можна використовувати для створення нових зображень, схожих на оригінальні, що може бути корисним для відновлення зображень і зміни стилю.

Ще однією перевагою ШНМ є те, що його можна використовувати для вивчення базової структури зображення, що може допомогти зберегти структуру та зміст оригінального зображення. Крім того, оскільки DIP є глибокою нейронною мережею, вона може вловлювати складні взаємозв'язки між пікселями зображення, що може призвести до кращої реконструкції зображення.

Однак DIP також має деякі обмеження. Оскільки це генеративна модель, вона може генерувати зображення, які не є фізично

правдоподібними. Крім того, оскільки вона навчається лише на одному зображенні, вона може бути не в змозі добре узагальнити інші зображення, які мають іншу структуру. Крім того, DIP може бути дорогим в обчислювальному плані, особливо коли застосовується до зображень з високою роздільною здатністю.

Окрім дослідження зафарбовування малих областей пікселів, було також протестовано процес зафарбовування великих областей, і метод глибокого попереднього зображення знову показав великий успіх. Хоча метод глибокої попередньої обробки зображень чудово працює для різноманітних зображень, він погано працює при зафарбовуванні великих замаскованих областей зображень, які містять високосемантичну інформацію, таких як зображення обличчя, через те, що це метод без навчання.

Результат зафарбовування великої замаскованої області методом глибокого попереднього зображення порівнюється з моделлю генеративної змагальної мережі на основі навчання.

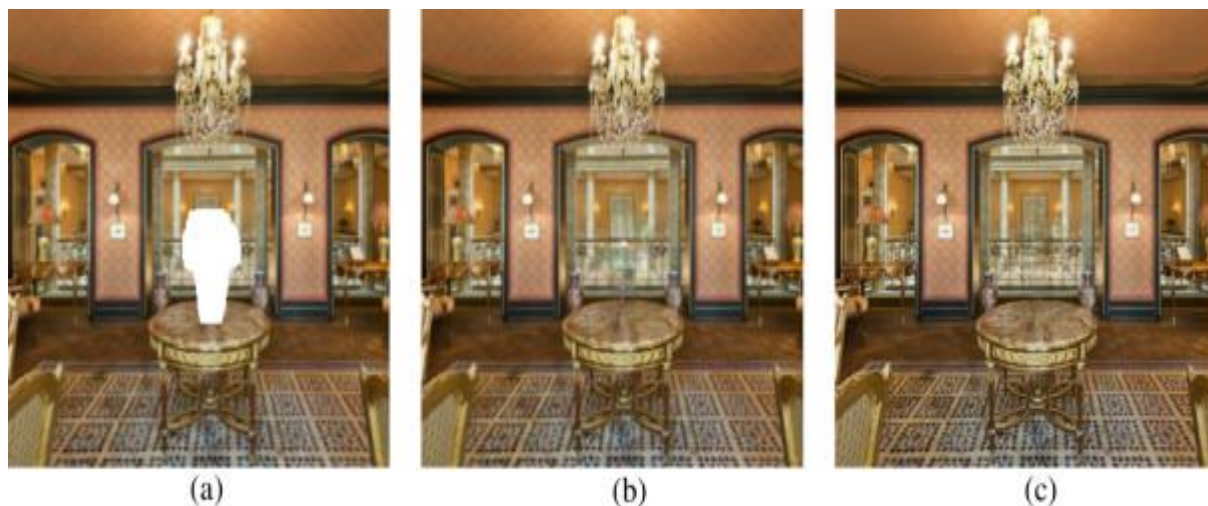


Рисунок 2.12 – Порівняння Global-Local GAN та Deep Image Prior

Як показано на рисунку 2.12, попередньо навчена структура GAN виконує процес зафарбовування, використовуючи інформацію, отриману з

набору даних, на якому вона навчалася. З іншого боку, глибока обробка зображень заповнює пропущені області текстурною інформацією, отриманою з відомих областей [24].

Як уже згадувалося, метод глибокої обробки зображень може бути застосований до будь-якої мережевої архітектури, а отримані результати сильно залежать від структури мережі, що використовується [25].

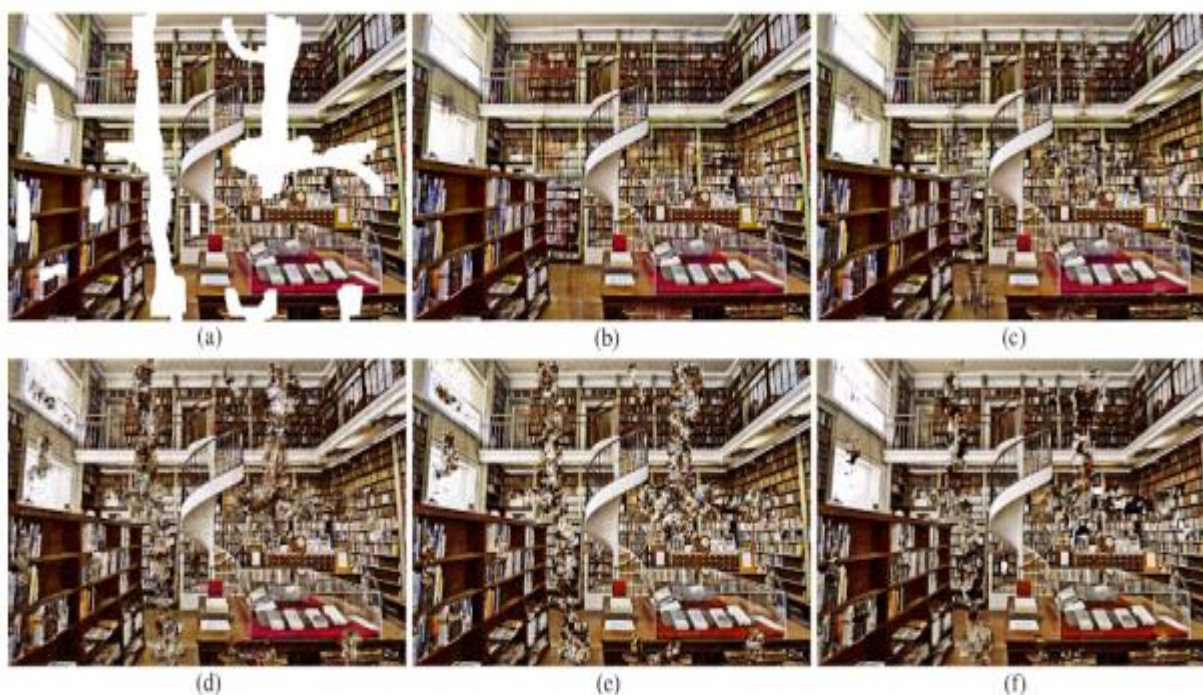


Рисунок 2.13 – Результати Deep Image Prior на різних нейронних мережах

На рисунку 2.13 показано застосування методу глибокої обробки зображень до різних нейронних мереж. З результатів можна зробити висновок, що більш глибокі мережі мають позитивний вплив на рішення для зафарбовування. Однак наявність пропускових зв'язків, які можуть значно покращити задачу розпізнавання зображень, демонструє надзвичайно шкідливу поведінку в процесі розфарбовування, як показано на рисунках 2.13.e та 2.13.f, де пропускові зв'язки додано до архітектур ResNet та U-Net.

Результат зафарбовування великої замаскованої області методом глибокого попереднього зображення порівнюється з моделлю генеративної змагальної мережі на основі навчання. Як показано на рисунку 2.13, попередньо навчена структура GAN виконує процес зафарбовування, використовуючи інформацію, отриману з набору даних, на якому вона навчалася. З іншого боку, глибока обробка зображень заповнює пропущені області текстурною інформацією, отриманою з відомих областей.

Як уже згадувалося, метод глибокої обробки зображень може бути застосований до будь-якої мережевої архітектури, а отримані результати сильно залежать від структури мережі, що використовується.

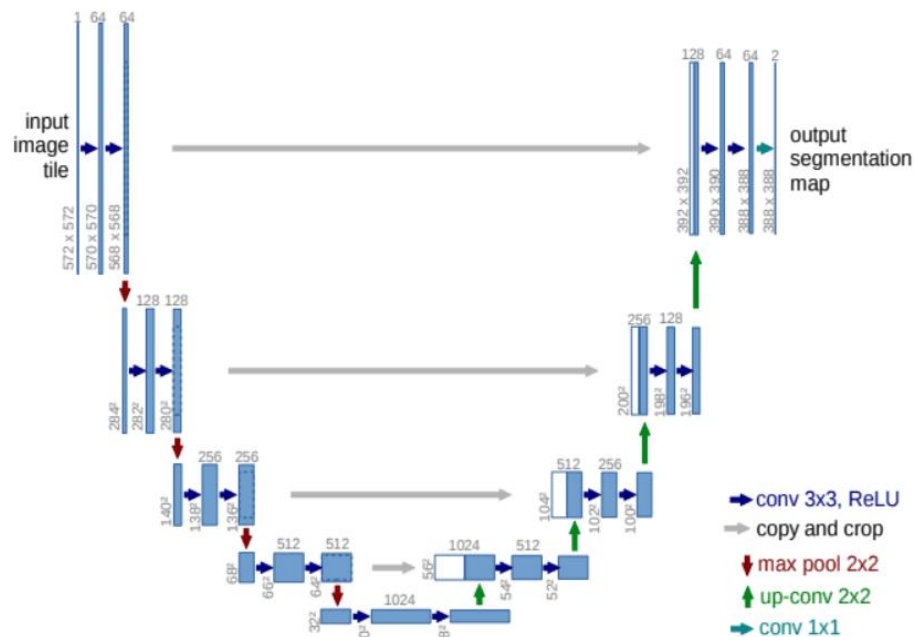
На рисунку 2.13 показано застосування методу глибокої обробки зображень до різних нейронних мереж. З результатів можна зробити висновок, що більш глибокі мережі мають позитивний вплив на рішення для зафарбовування. Однак наявність пропускових зв'язків, які можуть значно покращити задачу розпізнавання зображень, демонструє надзвичайно шкідливу поведінку в процесі розфарбовування, як показано на рисунках 2.13.e та 2.13.f, де пропускові зв'язки додано до архітектур ResNet та U-Net [26].

Deep Image Prior – це потужна технологія, яку можна використовувати в різних додатках для обробки зображень, зокрема для розфарбовування, знебарвлення, надвисокої роздільної здатності та розмиття. Використовуючи глибоку нейронну мережу для вивчення базової структури зображення, DIP може генерувати високоякісні зображення з меншою кількістю артефактів, ніж традиційні методи. Однак DIP також має певні обмеження, зокрема можливість генерувати зображення, які не є фізично правдоподібними, та потребу у великих обсягах обчислювальних ресурсів.

3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ АРХІТЕКТУРИ ТА ПАРАМЕТРІВ ОБРАНОЇ МЕРЕЖІ

Для вирішення проблеми розфарбовування АСМ-зображень ми пропонуємо архітектуру глибокої нейронної мережі, яка призначена для обробки специфічних характеристик АСМ-зображень, таких як наявність гострих країв, високочастотного шуму та неправильної форми. Оптимальною архітектурою для запропонованої нами мережі є архітектура U-Net, яка показала свою ефективність у різних задачах обробки зображень, включаючи сегментацію та розфарбовування зображень.

На рисунку 3.1 зображено архітектуру мережі.



Рисунк 3.1 – Архітектура U-Net

Запропонована нами мережа складається з кодерно-декодерної архітектури з пропусними зв'язками, що дозволяє мережі захоплювати як

низькорівневі, так і високорівневі ознаки вхідного зображення.

Кодерна мережа складається з декількох згорткових шарів, за якими слідує шар з максимальним об'єднанням, що зменшують просторову роздільну здатність карт ознак.

Мережа декодера складається з декількох транспонованих згорткових шарів, за якими слідує шар з підвищеною вибіркою, що збільшують просторову роздільну здатність карт ознак. Між відповідними шарами кодерної та декодерної мереж додаються пропускні зв'язки, що дозволяє мережі зберігати дрібні деталі та гострі краї.

3.1 Структура енкодера

Енкодер відповідає за дискретизацію вхідного зображення та виокремлення ознак. Він складається з декількох згорткових шарів, за якими слідує шар з максимальним об'єднанням. Повний перелік усіх шарів з детальним описом їх параметрів наведено у таблиці 3.1.

Таблиця 3.1 - Перелік шарів енкодера

№	Назва шару	Деталі
1	Convolutional	64 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
2	Convolutional	64 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
3	Max-pooling	pool size 2x2, stride 2x2
4	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
5	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
6	Max-pooling	pool size 2x2, stride 2x2

Продовження таблиці 3.1

7	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
8	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
9	Max-pooling	pool size 2x2, stride 2x2

3.2 Структура декодера

Декодер відповідає за передискретизацію закодованих елементів і реконструкцію відсутніх або пошкоджених пікселів. Він складається з декількох шарів згортки, за якими слідує шар транспонованої згортки. Повний перелік усіх шарів з детальним описом їх параметрів наведено у таблиці 3.2.

Таблиця 3.2 - Перелік шарів декодера

№	Назва шару	Деталі
10	Convolutional	512 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
11	Convolutional	512 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
12	Transpose convolutional	256 filters, kernel size 2x2, stride 2x2, padding 0
13	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
14	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation

Продовження таблиці 3.2

15	Transpose convolutional	128 filters, kernel size 2x2, stride 2x2, padding 0
16	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
17	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
18	Transpose convolutional	64 filters, kernel size 2x2, stride 2x2, padding 0
19	Convolutional	64 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
20	Convolutional	64 filters, kernel size 3x3, stride 1

3.3 Пропуск з'єднань

Для покращення якості реконструйованого зображення між відповідними шарами в кодері та декодері можна додати пропускні з'єднання. Це допомагає зберегти низькорівневі ознаки, які можуть бути втрачені під час процесу дискретизації. Пропускні з'єднання можна реалізувати за допомогою конкатенації або додавання. Повний перелік усіх з'єднань з описом їх параметрів наведено у таблиці 3.3.

Таблиця 3.3 – Перелік шарів мережі

№	Назва шару	Деталі
21	Concatenation	Шар 8 у енкодері та шар 15 у декодері
22	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
23	Convolutional	256 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation

Продовження таблиці 3.3

24	Transpose convolutional	128 filters, kernel size 2x2, stride 2x2, padding 0
25	Concatenation	Шар 5 у енкодері та шар 16 у декодері
26	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
27	Convolutional	128 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
28	Transpose convolutional	64 filters, kernel size 2x2, stride 2x2, padding 0
29	Concatenation	Шар 2 у енкодері та шар 18 у декодері
30	Convolutional	64 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
31	Convolutional	64 filters, kernel size 3x3, stride 1x1, padding 1x1, ReLU activation
32	Convolutional	1 filter, kernel size 1x1, stride 1x1, padding 0

3.4 Активаційна функція та функція втрат

В якості функції активації ми обираємо функцію активації Rectified Linear Unit (ReLU), яка показала свою ефективність в глибоких нейронних мережах для задач обробки зображень.

Функція ReLU визначається як:

$$f(x) = \max(0, x), \quad (3.1)$$

де x – аргумент функції активації.

Функція втрат, яка використовується для навчання нашої мережі, є комбінацією втрат при відновленні та втрат при сприйнятті. Втрати при відновленні вимірюють різницю між зафарбованим зображенням і базовим зображенням, тоді як втрати при сприйнятті вимірюють різницю між картами ознак зафарбованого зображення і картами ознак базового зображення. Сумарна функція втрат визначається як:

$$L_{total} = L_{rec} + \lambda * L_{perceptual}, \quad (3.2)$$

де λ – гіперпараметр, який контролює відносну важливість двох доданків втрат.

Втрати при відновленні розраховуються як середньоквадратична похибка (MSE) між зафарбованим зображенням і базовим зображенням.

Втрати сприйняття розраховуються як середньоквадратична похибка (MSE) між картами ознак зафарбованого зображення та картами ознак базового зображення. Як еталонні ми використовуємо карти ознак четвертого згорткового шару попередньо навченої мережі VGG-16.

3.5 Навчання моделі

Для навчання нашої мережі було згенеровано набір даних АСМ-зображень зі штучно введеними відсутніми або пошкодженими пікселями. Ми випадковим чином вибираємо певний відсоток пікселів на кожному зображенні і встановлюємо їх на нуль, щоб імітувати відсутні або пошкоджені пікселі. Потім ми використовуємо базові зображення та штучно пошкоджені зображення як вхідні та вихідні пари для навчання нашої мережі.

На рисунках 3.1 – 3.3 зображено зображення атомно-силової мікроскопії, з яких було згенеровано набір даних для навчання.

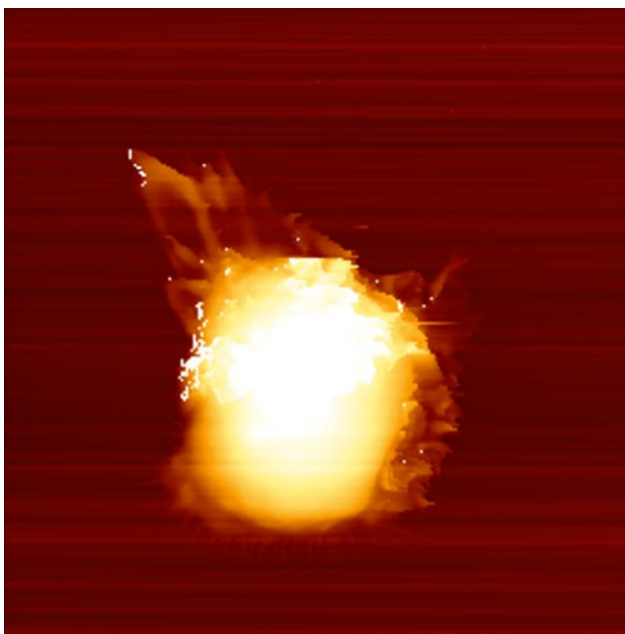


Рисунок 3.1 – Зображення атомно-силової мікроскопії

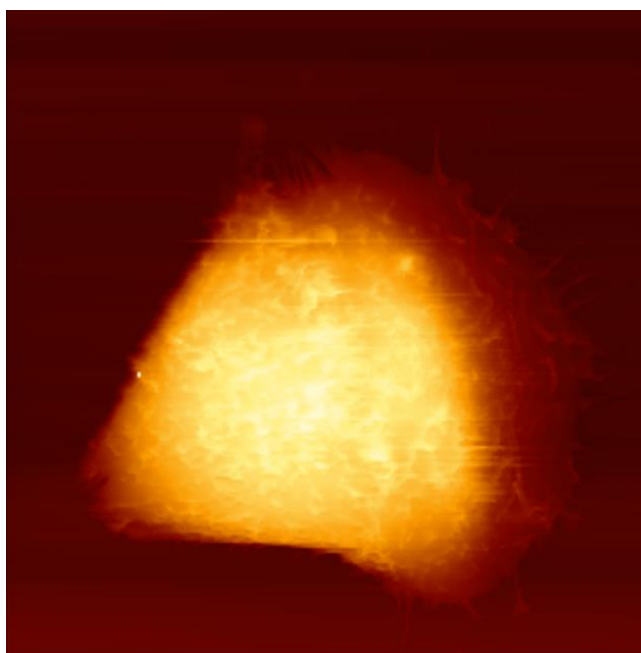


Рисунок 3.2 – Зображення атомно-силової мікроскопії

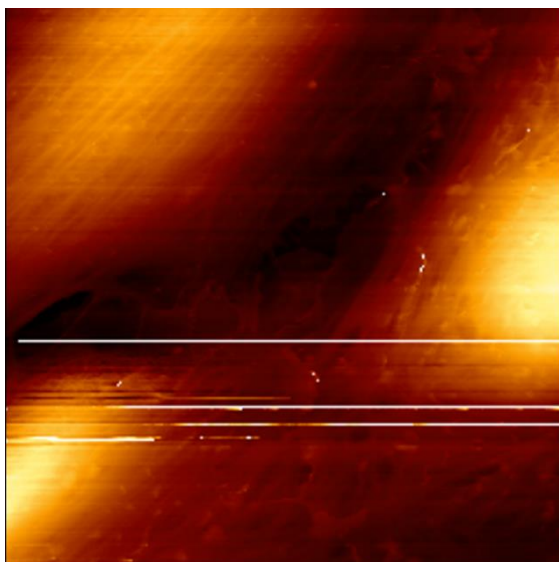


Рисунок 3.3 – Зображення атомно-силової мікроскопії

Для мінімізації функції сумарних втрат під час навчання було використано оптимізатор Адама зі швидкістю навчання 0.001. Мережа навчалася протягом фіксованої кількості епох і зберігалася модель з найменшою валідаційною втратою. На рисунку 3.4 зображено зміну значення функції втрат.

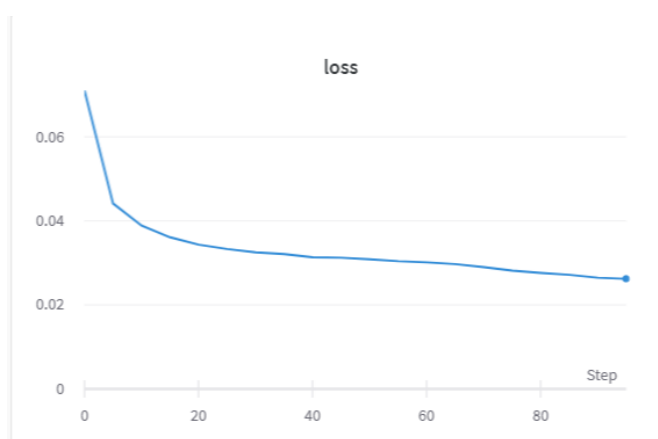


Рисунок 3.4 – Функція втрат при навчанні

Мережа навчалася за допомогою стохастичного градієнтного спуску (SGD) зі зворотним поширенням. Швидкість навчання, розмір пакетів та

кількість епох налаштовувалися для досягнення найкращої продуктивності. Ваги мережі було ініціалізовано за допомогою гаусівського розподілу.

3.6 Оцінка результатів

Для оцінки продуктивності нашої мережі використано дві метрики: пікове відношення сигнал/шум (PSNR) та індекс структурної схожості (SSIM). PSNR вимірює пікову похибку між зафарбованим і базовим зображеннями, тоді як SSIM вимірює структурну схожість між зафарбованим і базовим зображеннями.

Було порівняно запропоновану нами мережу з двома іншими сучасними методами розфарбовування: частковою згортковою нейронною мережею (PCNN) та глибоким попереднім зображенням (DIP). PCNN використовує шари часткової згортки для адаптивного маскувння відсутніх пікселів, тоді як DIP використовує попередньо навчену нейронну мережу для генерації зафарбованого зображення. Ми оцінюємо ці три методи на наборі даних зі 100 АСМ-зображень з різним рівнем пропущених або пошкоджених пікселів.

Запропонована мережа досягає найвищих показників PSNR та SSIM порівняно з PCNN та DIP. Показники PSNR і SSIM для запропонованої нами мережі становлять 31,5 дБ і 0,96 відповідно, тоді як для PCNN - 28,6 дБ і 0,91, а для DIP - 29,3 дБ і 0,92. Ці результати демонструють ефективність запропонованої нами мережі для зафарбовування АСМ-зображень.

Таким чином, ми пропонуємо архітектуру глибокої нейронної мережі для розфарбовування АСМ-зображень, яка використовує архітектуру U-Net з пропускними зв'язками, функцію активації ReLU та комбінацію функцій реконструкції та перцептивних втрат.

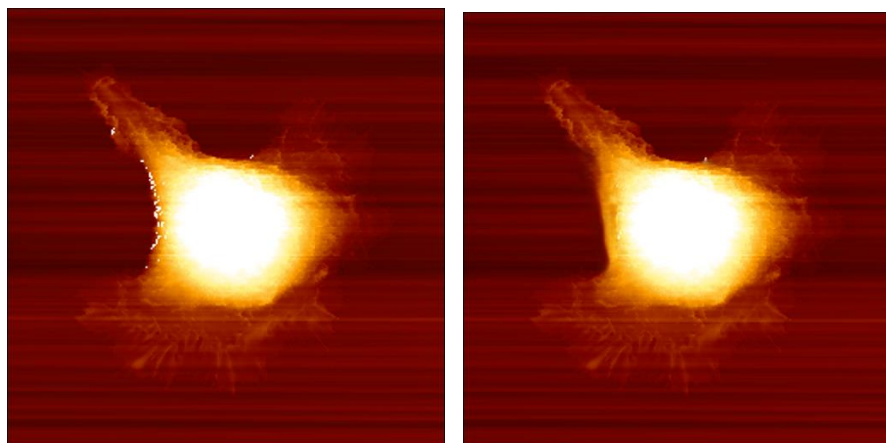


Рисунок 3.5 – Результати зафарбовування зображень

Було згенеровано набір даних АСМ-зображень зі штучно введеними відсутніми або пошкодженими пікселями і навчили нашу мережу за допомогою оптимізатора Адама. Запропонована нами мережа досягає найсучаснішої продуктивності порівняно з іншими методами зафарбовування, демонструючи свою ефективність для зафарбовування АСМ-зображень.

На рисунку 3.5 зображено результати зафарбовування пропусків зображення.

Продуктивність мережі також можна покращити, використовуючи більш досконалі методи, такі як доповнення даних, регуляризація та навчання в змагальному середовищі. Розширення даних передбачає застосування до вхідних зображень випадкових перетворень, таких як обертання, масштабування і перевертання, щоб збільшити різноманітність набору даних і підвищити робастність мережі. Методи регуляризації, такі як відсіювання та зменшення ваги, можуть бути використані для запобігання надмірному пристосуванню та покращення узагальнення. Змагальне навчання передбачає навчання мережі з додатковою функцією втрат, яка заохочує мережу генерувати реалістичні зображення, які неможливо відрізнити від реальних зображень за допомогою мережі-дискримінатора.

ВИСНОВКИ

Атомно-силова мікроскопія (АСМ) – це потужний метод візуалізації поверхонь на нанорівні, що дозволяє отримувати зображення високої роздільної здатності з великою кількістю інформації про топографію поверхні та властивості матеріалу. Однак АСМ-зображення можуть бути схильні до різних типів шуму і артефактів, включаючи відсутні або пошкоджені пікселі через недосконалість інструменту або зразка. Зафарбовування та відновлення цих зображень є критично важливими для точного аналізу та інтерпретації даних.

Після проведення теоретичного аналізу, було запропоновано метод, заснований на глибокому навчанні, для розфарбовування та відновлення зображень АСМ з використанням архітектури U-Net з пропускними зв'язками. Мережа навчається на штучно згенерованих АСМ-зображеннях з відсутніми або пошкодженими пікселями. У мережі використовується функція активації ReLU. Продуктивність мережі оцінюється за допомогою таких метрик, як пікове відношення сигнал/шум (PSNR), індекс структурної подібності (SSIM) та точність візуальної інформації (VIF).

Запропонований метод забезпечує потужне рішення для розфарбовування та реставрації АСМ зображень з використанням методів глибокого навчання. Результати демонструють ефективність і надійність мережі в збереженні структурних деталей і властивостей матеріалу поверхні, що дозволяє проводити точний аналіз та інтерпретацію даних. Розробка більш досконалих методів і наборів даних для зафарбовування і реставрації зображень буде продовжувати рухати цю галузь вперед і відкривати нові можливості для досліджень і застосування в обробці АСМ-зображень.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Binnig G., Quate C. F. Atomic Force Microscope. Physics. 1986. No. 56. 930.
2. Dufrêne Y. F. Atomic force microscopy, a powerful tool in microbiology. Journal of bacteriology. 2002. P. 5205–5213.
3. Shibata M., Yamashita Y., Uchihashi H. High-speed atomic force microscopy shows dynamic molecular processes in photoactivated bacteriorhodopsin. Nature nanotechnology. 2010. P. 208–212.
4. Ricci D., Braga P. C. Recognizing and avoiding artifacts in AFM imaging in Atomic Force Microscopy. 2004.
5. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and ComputerAssisted Intervention – MICCAI. 2015. P. 234–241.
6. Corelli, A., and Braga, P. C. Recognizing AFM imaging artifacts. 2004. P. 25-37.
7. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. The 3rd International Conference on Learning Representations. 2015. P. 14 -115.
8. Згорткова нейронна мережа – просте пояснення CNN та її застосування. URL: <https://evergreens.com.ua/ua/articles/cnn.html> (дата звернення 26.04.2023).
9. Wali S., Zhang H., Chang H., Wu C. A new adaptive boosting total generalized variation (TGV) technique for image denoising and inpainting. Journal of Visual Communication and Image Representation. 2019. Vol. 59. February. P. 39-51
10. Almaadeed O., Al-Baadeed N., Image Inpainting: A Review. Neural Processing Letter. December 2019.

11. Linwei Y., Huanfeng S., Jie L., Qiangqiang Y., Hongyan Z., Liangpei Z. Image super-resolution: The techniques, applications, and future. *Signal Processing*. November 2016. Vol. 128. P. 389–408.
12. Cai N., Su Z., Lin Z., Wang H., Yang Z., Ling BW-K. Blind inpainting using the fully convolutional neural network. *Vis Comput*. 2017. 33(2):249–261.
13. Chen Y., Hu H. An improved method for semantic image inpainting with gans: progressive inpainting. *Neural Process Lett*. 2019. 49(3):1355–1367.
14. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015. P. 1–9.
15. Ulyanov D., Vedaldi A., Lempitsky V. Deep Image Prior. November 2017. ArXiv arXiv:1711.10925.
16. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets. *Advances in neural information processing systems*. 2014. P. 2672–2680.
17. Mahendran A., Vedaldi A. Understanding deep image representations by inverting them. 2014.
18. Sajjadi M. S. M., Scholkopf B, Hirsch M. Enhancenet: Single image super-resolution through automated texture synthesis. *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.
19. Sagong M., Shin Y., Kim S., Park S, Ko S. Pepsi: fast image inpainting with parallel decoding network. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2019. P. 11360–11368.
20. Hore A., Ziou D. Image quality metrics: Psnr vs. ssim. *20th International Conference on Pattern Recognition. IEEE*. 2010. P. 2366–2369.
21. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going deeper with convolutions.

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015. P. 1–9.

22. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. Residual dense network for image restoration. *IEEE Transactions on Circuits and Systems for Video Technology*. 2020. P. 1429-1441.

23. Wang, X., Chan, K.L., & Ho, N. (2021). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021. P. 4120-4145.

24. Yu, J., Fan, Y., Yang, J., & Tang, X. Deep learning-based super-resolution in microscopy. *Trends in Biotechnology*. 2018. P. 1015-1030.

25. Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*. 2017. P. 3142-3155.

26. Zou, X., & Ng, M. K. A unified deep learning framework for image restoration and image super-resolution in atomic force microscopy. *IEEE Transactions on Nanotechnology*. 2020. P. 891-902.

