

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
(повна назва)

Кафедра _____ Штучного інтелекту _____
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____

«Відновлення пошкоджених зображень за допомогою
глибинної нейронної мережі»
(тема)

Виконав:

студент 2 курсу, групи _____ СІШМ-19-1

Омельченко С.О.

(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту

(повна назва спеціалізації)

Керівник _____ проф. Кулішова Н.Є.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри

_____ (підпис)

В.О. Філатов

(прізвище, ініціали)

2020 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-професійна
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту (СШІ)
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 20 __ р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУ

студентові Омельченко Софії Олегівні
(прізвище, ім'я, по батькові)

1. Тема роботи Відновлення пошкоджених зображень за допомогою глибинної нейронної мережі

затверджена наказом університету від 30 жовтня 20 20 р. № 1497ст

2. Термін подання студентом роботи до екзаменаційної комісії _____ 20 __ р.

3. Вихідні дані до роботи Науково-технічні публікації, статті Інтернет-джерел, документація PyTorch та numpy, цифрові зображення різного походження та змісту

4. Перелік питань, що потрібно опрацювати в роботі Аналіз предметної області та постановка задачі; Види задач відновлення зображення; Методи відновлення пошкоджених зображень; Дослідження глибинних нейронної мереж для видалення шумів; Експериментальне дослідження архітектури та параметрів обраної мережі

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Рисунок 1 – Види задач відновлення зображення, Рисунок 2 – Мета роботи, Рисунок 3 – Постановка задачі, Рисунок 4 – Відновлення зображення за допомогою deep image prior підходу, Рисунок 5 – Архітектури мереж, Рисунок 6 – Експериментальне дослідження архітектури та параметрів обраної мережі, Рисунок 7 – Результати відновлення, Рисунок 8 – Висновки .

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
Аналіз предметної області та постановка задачі	проф Кулішова Н.Є.		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної області	9.11.20	виконано
2	Постановка задачі	10.11.20	виконано
3	Дослідження технологій обробки зображень	18.11.20	виконано
4	Експериментальні дослідження	23.11.20	виконано
5	Написання пояснювальної записки	25.11.20	виконано
6	Попередній захист	8.12.20	виконано
7	Захист перед ЕК	16.12.20	

Дата видачі завдання 01 листопада 2020 р.

Студент _____
(підпис)

Керівник роботи _____
(підпис) проф. Кулішова Н.Є.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 85 с., 5 табл., 30 рис., 2 дод., 40 джерел.

АЛГОРИТМ, ВДОСКОНАЛЕННЯ, ВИДАЛЕННЯ ШУМІВ, ЗОБРАЖЕННЯ, ГЛИБИННЕ НАВЧАННЯ, ОПТИМІЗАТОР, ФУНКЦІЯ АКТИВАЦІЇ.

Мета роботи – підвищення якості зображень, пошкоджених шумом, спотвореннями в ході реєстрації та передачі по мережам і каналам зв'язку.

Головна задача полягає у оптимізації структури глибинної мережі шляхом встановлення найбільш оптимальних компонентів, які у комплексі дадуть кращий результат.

Об'єкт дослідження – процеси формування та обробки цифрових зображень в інформаційних системах.

Предмет дослідження – нейромережевий підхід, зокрема, нейромережі з глибинним навчанням.

У різних сферах діяльності існує безліч цифрових зображень, які містять цінну інформацію, наприклад, медична візуалізація, дистанційне зондування, військові системи, робототехніка і штучний інтелект. Пошкодження цих зображень безповоротно руйнує їх змістовність. Саме тому методи видалення шумів на зображенні знаходять широке застосування у багатьох галузях.

Глибинні згорткові мережі вже давно стали популярним інструментом для створення та відновлення зображень. Звичайно, що для кожної задачі будь-яка обрана мережа повинна бути попередньо налаштована, проте це виконується для всіх спостережень одночасно. Саме тому, у цій роботі було проведено налаштування мережі таким чином, що для кожного зображення видалення шумів буде виконано із достатньою якістю.

РЕФЕРАТ

Пояснительная записка: 85 с., 5 табл., 30 рис., 2 доп., 40 источников.

АЛГОРИТМ, ГЛУБОКОЕ ОБУЧЕНИЕ, ИЗОБРАЖЕНИЕ, ОПТИМИЗАТОР, УДАЛЕНИЕ ШУМОВ, УСОВЕРШЕНСТВОВАНИЕ, ФУНКЦИЯ АКТИВАЦИИ.

Цель работы – повышение качества изображений, поврежденных шумом, искажениями в ходе регистрации и передачи по сетям и каналам связи.

Главная задача заключается в оптимизации структуры глубинной сети путем установления наиболее оптимальных компонентов, которые в комплексе дадут лучший результат.

Объект исследования – процессы формирования и обработки цифровых изображений в информационных системах.

Предмет исследования – нейросетевой подход, в частности, нейросети с глубинным обучением.

В различных сферах деятельности существует множество цифровых изображений, которые содержат ценную информацию, например, медицинская визуализация, дистанционное зондирование, военные системы, робототехника и искусственный интеллект. Повреждение этих изображений безвозвратно разрушает их содержательность. Именно поэтому методы удаления шумов изображений находят широкое применение во многих отраслях.

Глубинные сверточные сети уже давно стали популярным инструментом для создания и восстановления изображений. Конечно, для каждой задачи любая выбранная сеть должна быть предварительно настроена, однако это выполняется для всех наблюдений одновременно. Именно поэтому в данной работе была проведена настройка сети таким образом, что для каждого изображения удаление шумов будет выполнено с достаточным качеством.

ABSTRACT

Explanatory note: 85 p., 5 tab., 30 fig., 2 additional, 40 sources.

ACTIVATION FUNCTION, ALGORITHM, DEEP LEARNING, IMAGE, IMPROVEMENT, NOISE REMOVAL, OPTIMIZER.

The purpose of work is to improve the quality of images damaged by noise, distortions during registration and transmission over networks and communication channels.

The main task is to optimize the structure of the deep network by establishing the most optimal components that together will give the best results.

The object of research is the processes of generation and processing of digital images in information systems.

The subject of the study is a neural network approach, in particular, a neural network with depth learning.

In different fields of activity there are many digital images that contain valuable information, such as medical imaging, remote sensing, military systems, robotics and artificial intelligence. Damage to these images irrevocably damages their content. That is why image noise removal methods are widely used in many industries.

Deep convolution networks have long been a popular tool for image creation and restoration. Of course, any selected network must be pre-configured for each task, but this is done for all observations simultaneously. That is why this work has been carried out in such a way that for each image noise removal will be performed with sufficient quality.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів.....	8
Вступ.....	9
1 Аналіз предметної області та постановка задачі.....	10
1.1 Види задач відновлення зображення.....	10
1.2 Методи відновлення пошкоджених зображень.....	12
1.2.1 Image denoising.....	12
1.2.2 Задача супер роздільної здатності.....	19
1.2.3 Image inpainting.....	22
1.3 Переваги та недоліки методів відновлення зображень.....	25
1.4 Постановка задачі.....	20
2 Дослідження глибинних нейронних мереж для видалення шумів.....	31
2.1 Глибинні мережі для очищення зображень від шумів.....	32
2.1.1 Архітектури мереж.....	32
2.1.2 Оптимізатори мереж.....	36
2.1.3 Функції активації.....	42
2.2 Опис методу deep image priors.....	47
3 Експериментальне дослідження архітектури та параметрів обраної мережі.....	53
3.1 Оптимізатори мережі.....	58
3.2 Функції активації.....	61
3.3 Перевірка швидкодії.....	64
Висновки.....	66
Перелік джерел посилання.....	68
Додаток А Програмний код.....	72
Додаток Б Відомість атестаційної роботи магістра.....	84

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

Adam – adaptive moment estimation – метод адаптивної оцінки моментів, adaptive moment estimation;

CNN – convolutional neural network – згорткова нейрона мережа;

DCGAN – deep convolutional generative adversarial network – глибинна згорткова генеративна змагальна мережа;

DL – deep learning – глибинне навчання;

ELU – exponential linear unit – експоненціальна лінійна одиниця;

GAN – generative adversarial network – генеративно змагальна мережа;

HR – high resolution – зображення з високою роздільною здатністю;

L-BFGS – limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm – модифікація алгоритму Бройде-Флетчера-Гольдфарба-Шанно з використанням обмеженого обсягу пам'яті комп'ютера;

LR – low resolution – зображення з низькою роздільною здатністю;

LReLU – leaky rectified linear unit – лінійна активація з випрямленою дією;

MPT – магнітно-резонансна томографія;

PSNR – peak signal-to-noise ratio – пікове співвідношення сигналу до шуму;

ReLU – rectified linear unit – випрямлена лінійна одиниця;

ResNet – residual neural network – залишкова мережа,;

SR – super resolution – задача супер роздільної здатності;

SSIM – structure similarity – індекс структурного походження ;

VGG – згорткова нейрона мережа, розроблена Visual Geometry Group.

ВСТУП

У різних сферах діяльності існує безліч цифрових зображень, які містять цінну інформацію, наприклад, медична візуалізація, дистанційне зондування, військові системи, робототехніка і штучний інтелект. Пошкодження цих зображень безповоротно руйнує їх змістовність. Саме тому методи видалення шумів на зображенні знаходять широке застосування у багатьох галузях.

Оскільки, за своєю природою та походженням, пошкодження різьчє відрізняються, то й методи та підходи для їх усунення не можуть бути універсальними, та придатними для вирішення цих проблем. На жаль, не існує такого одного засобу для вирішення усіх проблем, а якщо й був, то він однозначно не був би найоптимальнішим для будь-якого виду пошкоджень.

Мета роботи – підвищення якості зображень, пошкоджених шумом та спотвореннями в ході реєстрації та передачі по мережам і каналам зв'язку.

Головна задача полягає у оптимізації алгоритмів шляхом модифікації стадій побудови мережі, а саме стадій ініціалізації та навчання, налаштування параметрів та вибору оптимізатора.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

Серед інформації, яку сучасне людство споживає та генерує щодня, зображення становлять значну та невід'ємну частку усього обсягу. Це не дивно, адже наш мозок переважно сприймає та обробляє візуальну інформацію краще і швидше за дані у будь-якому іншому форматі. Недарма для опису складних систем, явищ та структур використовують презентації з схемами, умовними позначеннями, діаграмами і графіками для полегшення сприйняття та засвоєння матеріалу. Тому закономірно, що кількість графічних файлів, різноманіття їх форматів та способів обробки значно збільшилося за останні роки.

Всі ці засоби розроблені з однією метою – полегшити роботу із зображенням на різних стадіях їх життєвого циклу, починаючи від створення до трансформування у різні формати. Попри таке різноманіття інструментів, вони не завжди можуть допомогти з обробкою та відтворенням пошкоджених файлів.

Пошкодження можуть виникнути у будь-який момент та носити різний характер, починаючи від пошкоджень самого файлу на носії, до застосування навмисних ушкоджень фільтрами, масками чи алгоритмами. Тільки алгоритмами Інстаграму можна змінити фотографію до невпізнавального стану, що вже казати про професійні програми. Процес створення, модифікації зображень та їх результати можна побачити без проблем, але зворотні маніпуляції викликають більше питань та задач для розв'язання.

1.1 Види задач відновлення зображення

Суть задачі відновлення зображення полягає у знаходженні невідомого справжнього зображення з його наданої пошкодженої версії. Як зазначалось раніше, погіршення зображення може відбуватися під час

формування, передачі та зберігання зображення (рис. 1.1). Найбільшу частку зображень, які потребують відновлення становлять саме ті, що були отримані з пошкодженнями. Саме тому, ця задача стала найбільш актуальною та широко використовується для поліпшення якості супутникових знімків, фотографій, зроблених при слабкому освітленні, тощо. Завдяки прогресу в цифрових, обчислювальних та комунікаційних технологіях відновлення деградованого зображення набуло суттєвого значення, а отже, перетворилося на окреме поле досліджень, яке перетинається з обробкою зображень, комп'ютерним зором та обчислювальними зображеннями.

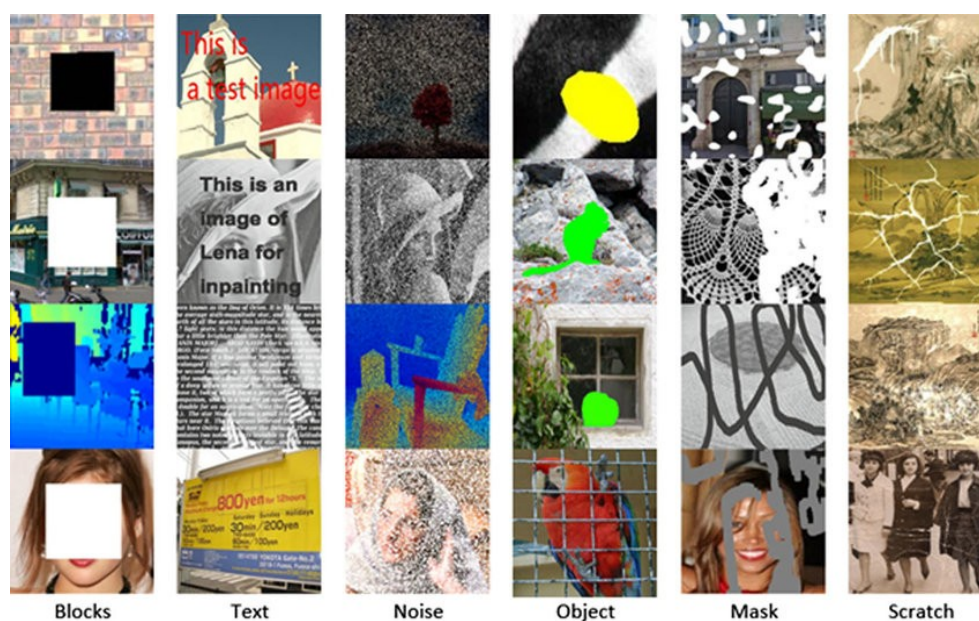


Рисунок 1.1 – Типи пошкоджень зображень

Виділяють три основні завдання відновлення зображення:

- видалення шумів на зображенні (image denoising) – відновлення зображення, пошкодженого адитивними шумами. Це найпростіше завдання реставрації зображень, і тому воно було ретельно вивчене;
- супер-роздільна здатність (image super resolution) – процес створення зображення (або послідовності зображень) з високою

роздільною здатністю із набору зображень з низькою роздільною здатністю;

– домальовування зображень (image inpainting) – це процес відновлення втрачених, пошкоджених частин зображень.

У класичному вигляді процес погіршення зображення описується наступним чином [1]. Нехай y – спостережуване шумне і розмите зображення, H оператор, що описує систему деградації, x вхід в систему і n шум, доданий до вихідного зображення, тоді маємо наступне рівняння:

$$y = Hx + n, \quad (1.1)$$

де y, x, n – вектори, які представляють спостережуване, оригінальне та шумове зображення;

H – матриця згортки, яка моделює деформацію зображення, наприклад, розмиття.

Задача відновлення полягає у побудові деякого наближення \hat{y} до оригінального зображення з використанням пошкодженого спостережуваного зображення y , інформації про матрицю деформації H та про шум n .

1.2 Методи відновлення пошкоджених зображень

1.2.1 Image denoising

У сучасну епоху комп'ютеризованого аналізу, в зв'язку з масовим збільшенням кількості цифрових зображень методи відновлення зображень стали незамінними інструментами. У різних областях діяльності використовують безліч цифрових зображень, які містять цінну інформацію, але завади можуть безповоротно зруйнувати їх зміст. Саме тому методи

видалення шумів на зображенні знаходять широке застосування у багатьох галузях.

Насамперед через те, що найчастіше зображення отримуються у несприятливих умовах, наприклад, при умовах поганого освітлення. З численних видів шумів, які переважають у різних видах зображень, найбільш відомими типами шумів є:

- адитивний білий гаусів шум (additive white Gaussian noise);
- імпульсний шум (salt and pepper);
- квантовий шум;
- дробовий або пуассоновий;
- спекл-шум.

У роботі [2] відмічають, що переважання інших типів шумів відбувається головним чином через дефекти виробництва пристроїв реєстрації зображень, бітові помилки та недостатню кількість фотонів під час отримання зображення.

Коли пошкодження зображення обумовлене виключно наявністю шуму, рівність (1.1) набуває вигляду:

$$s_k = x_{k+1} - x_k, \quad (1.2)$$

де x_0 – пошкоджене зображення;

x – чисте зображення;

ε – зображення, яке містить білий адитивний Гаусовий шум, з стандартним відхиленням σ .

Доданок, ε що описує шум, невідомий, тому просто відняти його з функції x_0 неможливо.

Методи видалення шумів умовно розділяють на п'ять категорій, а саме:

- методи просторових доменів (spatial domain methods);
- методи перетворення доменів (transform domain methods);

- методи в статистичних доменах (methods in statistical domains);
- гібридні методи;
- методи розрідженого уявлення (sparse representation methods);
- методи вивчення словників (dictionary learning methods).

Основним та найбільш простим підходом, яким користуються для знешумлення, та який не потребує алгоритмів машинного навчання чи нейронних мереж, є фільтрація. Фільтрація сигналів має фундаментальне значення для базової обробки зображення і вже давно застосовується для згладжування, підвищення різкості, виявлення країв і поліпшення контрастності зображень. Основним принципом алгоритму компенсації шумів зображення є те, що шум за своєю природою є некорельованим між пікселями, а інтенсивність справжніх пікселів сигналу корелюється між собою як це показано у роботах [3], [4], [5].

Залежно від способу відбору пікселів, який було обрано у процесі роботи, види фільтрів можна розділити на локальні та нелокальні. Фільтр вважається локальним, коли усереднення пікселів проводиться не по всій зоні визначення, а на порівняно невеликому радіусу навколо кожної точки зображення. Нелокальний фільтр використовує кореляцію між усім діапазоном пікселів на зображенні.

Популярними локальними фільтрами, призначеними для зменшення шуму, є:

- гаусівський фільтр;
- фільтр найменшого середнього квадрату(least mean square);
- білатеральний фільтр(Bilateral filter);
- фільтр Вінера;
- фільтр SUSAN;
- анізотропний дифузійний фільтр;
- регресія ядра керування (steering kernel regression);
- метрична регресія ядра керування(metric steering kernel regression).

Медіанна фільтрація, зважений медіанний фільтр та ранговий фільтр є основними та найбільш тривіальними прикладами нелінійних фільтрів, які були розроблені та офіційно задокументовані в літературі.

Зазвичай зображення сильно різняться за кольоровими, структурними, частотними властивостями, тому застосовність фільтрів, підготовлених до зображень з дуже різними характеристиками, значно обмежена. За таких обставин ефективний підхід з використанням концепції так званої нечіткої однорідної групи (fuzzy peer group). У статті [6] представлена архітектура кольорового фільтра цифрових зображень, який базується на нечіткій однорідній групі та використовує гнучкі двовимірні функції належності. Це дозволяє зробити фільтрування більш гнучким в умовах дії спотворень різних типів на зображення з різною структурою. Таким чином, запропонований фільтр на основі нечітких однорангових груп може ефективно придушити шум, властивий більшості зображень, зроблених цифровими фотокамерами.

На відміну від попередніх підходів, методи трансформації використовують властивість розрідженості пікселів. Тобто сигнал у таких методах може бути представлений меншою кількістю ненульових коефіцієнтів, або зображення може бути представлене як лінійне розширення кількох коефіцієнтів вищого порядку. Ця властивість зробила їх надзвичайно зручним інструментом цифрової обробки сигналу, як в одновимірному, так і в двовимірному домені. Репрезентативні атрибути, такі як локалізація, ізотропія, мульти-роздільна здатність та орієнтація базових функцій у різних напрямках являються основними властивостями методів перетворення зображення.

У цій категорії засобів було запропоновано велику кількість підходів, наприклад, таких як: перетворення Фур'є, швидке перетворення Фур'є, дискретне косинусне перетворення, вейвлети, курвлети, керовані вейвлети, ріджлети, ріплети, контурлети, веджлети, бандлети, дірекшнлети, ширлети і банки фільтрів спрямованої дії (рис. 1.2).

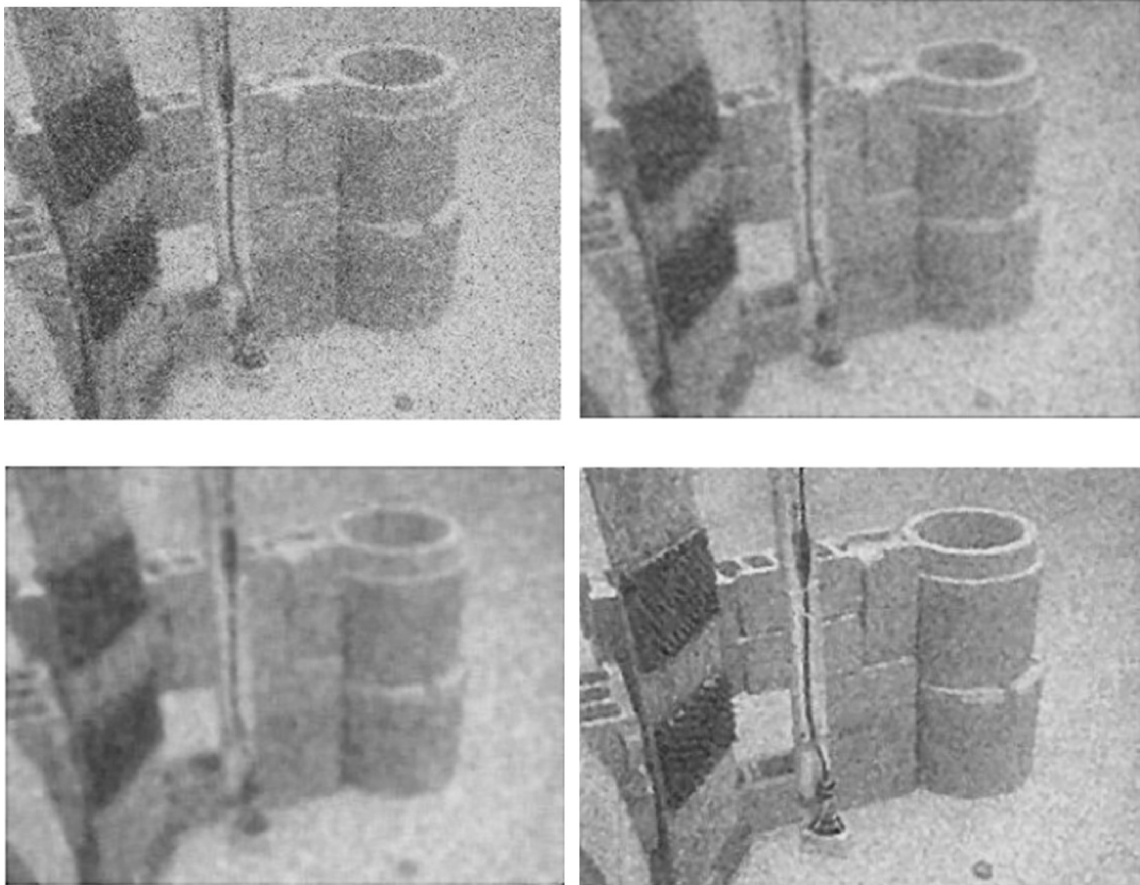


Рисунок 1.2 – Підводні знімки з шумами; зняті зображення після використання фільтра усереднення (перший рядок L–R); медіанного фільтра; вейвлет-метода

Окрім просторових методів і методів трансформації доменів, дослідники розробили методи видалення шумів на зображенні на основі статистичних моделей багатомасштабних (multi-scale) коефіцієнтів та випадкових полів. У схемах статистичних моделей коефіцієнти безшумності оцінюються на основі первинної інформації та на базі байєсової оцінки. Ці методи спрямовані на побудову просторово однорідних моделей, адаптивних до різних рівнів сигналів, за допомогою параметризації змінних, які мають самовипадковий (self-random) характер.

Методи, засновані на випадкових полях, становлять ще один дуже популярний домен, не тільки застосовується в області видалення шумів на

зображенні, а також для інших задач, наприклад, класифікації та сегментації зображень. У цих методах береться до уваги той факт, що значення інтенсивності пікселя залежить від сусідніх пікселів. Ці методи засновані на спостереженні, що глобальне представлення зображення може бути отримане з його локальних фізичних структур, що здійснюється за допомогою функції умовного розподілу ймовірностей, яку називають Марковським випадковим полем. Далі за байєсівським фреймворком визначається максимальна апіорна ймовірність, яка стверджує, що дане чисте та чітке зображення могло би бути шумним, щоб отримати певне погіршене зображення. Ці моделі встановлюють взаємозв'язок між сусідніми пікселями та між спостережуваними з оціненими значеннями на кожному пікселі [7].

Серед інших методів для вирішення задачі очищення зображення від шуму достатньо ефективними виявилися і алгоритми глибинного навчання. Так, CNN мають широке застосування не лише в області обробки зображень, а й виявилися особливо продуктивними і для задач шумозаглушення. Існує кілька методів глибинного навчання для видалення шумів, включаючи застосування CNN та їх поєднання з методами вилучення загальних ознак, поєднання методу оптимізації та CNN. Переважно, виокремлюють два основні типи методів глибинного навчання для зменшення шумів зображення: одинарний наскрізний (end-to-end) CNN та CNN у поєднанні з попередніми знаннями.

Було виявлено, що зміна архітектури мережі також є ефективним способом усунення шуму для пошкодженого зображення. Багатомасштабні мережі не менш ефективні для зменшення шуму зображення. Наприклад, ReLU та RL мережі використовують різні фазові особливості для посилення здатності моделі до відтворення виразності деталей на зображеннях при слабкому освітленні [8].

У реальному світі зображення легко пошкоджуються, а шум за своєю будовою комплексний через те, що на зображення впливає складне

середовище. Тому методи, які вирішують задачі пошкодження сліпим шумом, виявилися вкрай важливими для відновлення таких зображень [9]. У роботі [10] для FFDNet рівень шуму та шум був використаний як вхід до CNN для підготовки денойзера для невідомих шумних зображень. Згодом було запропоновано кілька методів для вирішення проблеми сліпого шуму.

Подібно до серійних зображень, відео розкладається на кожен кадр. Отже, техніки глибинного навчання для адитивного шумозаглушення білого шуму, реального шуму, сліпого шуму, гібридного шумозаглушення зображення також придатні для видалення шумів з відео. Рекурентні мережі використовують наскрізні CNN для видалення шуму з пошкодженого відео.

Ще один гібридний підхід запропоновано у статті [11], де пропонується розширена архітектура неофузі-нейрона, яка є узагальненням стандартного неофузі-нейрона у випадку нечіткого порядку виведення вище нуля. Вважається, що NFN має покращені апроксимаційні властивості, що характеризуються високою швидкістю навчання, має просту чисельну реалізацію. Метою цієї роботи є розробка адаптивного фільтра, який може компенсувати шум на цифрових зображеннях без значного зниження їх якості.

За останні кілька років техніки видалення шуму Гауса досягли великих успіхів, особливо у випадках, коли він регулярний. Однак у реальному світі таке трапляється рідко, бо шум зазвичай складний і нерегулярний. Удосконалення апаратних пристроїв з метою кращого придушення шуму для отримання якісного зображення є дуже важливим завданням для розробників. Більше того, отримане зображення може бути розмитим, із низькою роздільною здатністю та пошкодженим. Тому дуже важливо визначити, як ефективно відновити чисте зображення із зображення з накладеними шумами. Більше того, хоча використання методів глибинного навчання для вивчення особливостей вимагає знань про початкові дані, отримані шумні зображення найчастіше не мають

непошкодженого істинного варіанту. Це нагальні проблеми, для яких дослідники та вчені ще мають знайти вирішення.

1.2.2 Задача супер роздільної здатності

Існує чимало випадків, коли не вдається отримати картинку задовільної якості та розміру. Перешкоди найчастіше являються саме апаратними, адже деякі камери не здатні зробити зображення з високою роздільною якістю через ті чи інші причини. Трапляються також випадки, коли об'єкт знаходиться на значній відстані чи малий за розміром, але його деталі, написи тощо вкрай необхідно бачити чітко. Всі ці приклади описують наступне завдання обробки зображень – задачу супер роздільної здатності (SR).

Зазначена задача широко використовується у наступних випадках:

- спостереження: для виявлення, ідентифікації та розпізнавання обличчя на зображеннях із низькою роздільною здатністю, наприклад, отриманих із камер безпеки;
- медицина: у багатьох випадках отримати МРТ знімок із високою роздільною здатністю може виявитися складно. Завдяки методам вирішення цієї задачі можна подолати цю проблему, отримавши знімки з кращою роздільною здатністю (рис. 1.3);
- медіа: задача супер-роздільної здатності може бути використана для зменшення витрат на потужності серверів, оскільки зображення може надсилатися з меншою роздільною здатністю та збільшуватися у пункті призначення на льоту.



Рисунок 1.3 – МРТ зображення коліна до та після збільшення роздільної здатності

Задача супер роздільної здатності зображення полягає у підвищенні роздільної здатності зображення з низькою роздільною здатністю (LR) з певним коефіцієнтом підвищення до зображення з високою (HR). Вирішення оберненої задачі встановлюється як:

$$E(x; x_0) = \|d(x) - x_0\|^2, \quad (1.3)$$

де $E(x; x_0)$ – відповідний до задачі терм даних;

x_0 – LR-зображення;

x – HR-зображення;

$d(\cdot)$ – оператор пониження, який змінює розмір зображення у t -разів.

З технічної точки зору, залежно від виду вхідних даних SR задача може бути класифікована як багатокadroва або однокадрова. Якщо можна отримати декілька зображень однієї сцени зі здвигом субпікселів, то додаткову інформацію між ними можна використовувати для реконструкції зображення з більш високою роздільною здатністю або послідовності таких зображень. Однак, іноді декілька LR-зображень можуть бути

недоступні для реконструкції, отже, HR-зображення потрібно відновити за допомогою обмеженої інформації низької якості. Така постановка задачі визначається як однокадрова задача SR [12].

На самому початку пошуку методів для вирішення даної проблеми зображень більшість методів зосереджувались на частотній області. Алгоритми частотної області можуть використовувати взаємозв'язок між HR-зображенням та LR-спостереженнями на основі простої теоретичної основи та мають високу обчислювальну ефективність. Однак ці методи мали очевидні обмеження, такі як чутливість до помилок та труднощі в обробці більш складних моделей руху, що перешкоджало їх подальшому розвитку. Саме через недоліки алгоритмів частотної області методи просторової області стали основною тенденцією [13]. До популярних методів просторової області належать нерівномірна інтерполяція, ітераційна зворотна проекція (IBP), проекція на опуклі множини (POCS), регуляризовані методи та ряд гібридних алгоритмів. Серед них найбільш популярними є методи регуляризації завдяки своїй ефективності та гнучкості. Тому більшість нещодавніх репрезентативних статей про SR зосереджувались на регуляризованих системах.

Методи глибинного навчання досить успішно вирішили проблему надвисокої роздільної здатності зображення та відео. Для вирішення цього завдання було запропоновано чимало методів, які базуються на певних типах мереж, серед них:

- супер-роздільна здатність з попередньою повторною вибіркою (pre-upsampling super resolution): SRCNN, Very Deep Super Resolution (VDSR);
- супер-роздільна здатність з наступною повторною вибіркою (post-upsampling super resolution): FSRCNN, ESPCN;
- залишкові мережі (residual networks): EDSR, MDSR, CARN;
- багатоетапні залишкові мережі (multi-stage residual networks): BTSRN;

- рекурсивні мережі: Deep Recursive Convolutional Network (DRCN), Deep Recursive Residual Network (DRRN);
- Progressive Reconstruction Networks: LAPSRN;
- Multi-Branch Networks: Cascaded Multi-Scale Cross-Network (CMSC), Information Distillation Network (IDN);
- Attention-Based Networks: SelNet, Residual Channel Attention Networks (RCAN);
- Generative Models: SRGAN, EnhanceNet.

1.2.3 Image inpainting

Раніше для старих фото застосовували техніки доповнення зображень, щоб видалити подряпини та покращити пошкоджені зображення. Тепер цей підхід використовується для видалення небажаних артефактів, втрачених фрагментів підчас копіювання чи сканування документів, зображень, тощо. Доповнення зображень найчастіше використовується для видалення будь-яких спотворень, включаючи текст, блоки, шум, подряпини, лінії або різні за формою маски, застосовані до них.

При вирішенні цієї задачі на початкове зображення x_0 дається зображення з відсутніми пікселями відповідно до двійкової маски $V_k = I - \rho_k y_k s_k^T$. Таким чином, мета – відновити відсутні дані. Відповідний термін даних задано як:

$$E(x; x_0) = \|(x - x_0) \odot m\|^2, \quad (1.4)$$

де x – оригінальне зображення;

x_0 – початкове зображення;

m – маска.

Використовуючи нещодавно розроблені алгоритми, доповнення зображень може відновлювати цілісні як текстурні, так і структурні компоненти зображення. Отримані результати демонструють, що ці методи можуть видаляти небажані об'єкти із зображень, не залишаючи слідів та заповнюючи пропуски таким чином, що вони стають майже непомітними без детального вивчення (рис. 1.4).



Рисунок 1.4 – Відновлення зображення за допомогою NVIDIA

До теперішнього часу існуючі підходи для виконання цієї задачі розділяють на три підкатегорії: послідовні, засновані на згорткових (CNN) та генеративно змагальних мережах (GAN) [14].

Послідовні підходи можна розділити на дві категорії: на основі патча (patch-based) та на основі дифузії (diffusion-based). Методи, засновані на патчах, базуються на техніках, які заповнюють відсутні області шляхом пошуку підходящих заміників (тобто патчів-кандидатів) у неушкодженій частині зображення, та їх копіюванням у відповідні місця. Було запропоновано багато методів, заснованих на цій техніці, серед них метод

заснований на вимірюванні подібності між патчами за допомогою суми квадратних різниць (SSD) [15], метод із використанням загальної узагальненої варіації (TGV) [16]. Методи, засновані на дифузії, заповнюють відсутню область, плавно поширюючи вміст зображення від кордону до внутрішньої частини відсутньої області.

CNN використовуються в цій області спеціально для поліпшення очікуваних результатів з використанням великомасштабних даних. Послідовні методи успішно застосовуються в деяких частинах живопису, таких як заповнення текстурних деталей, однак проблема захоплення глобальної структури все ще залишається складною задачею.

Запропоновано декілька методів для доповнення зображень з використанням згорткових мереж. Shift-Net на основі архітектури UNet є одним з таких методів, який відновлює відсутній блок з хорошою точністю з точки зору структури і дрібнодисперсної текстури [17].

Деякі алгоритми для домальовування засновані на базі CNN:

- VORNet – доповнення відео з видаленням об'єктів [18];
- PEN-NET – доповнення зображень з високою роздільною здатністю [19];
- VICNN – метод сліпого доповнення зображень [20].

Генеративні змагальні мережі – найбільш широко використовувана на сьогоднішній день техніка для вирішення задач генерації зображень [21]. GAN – це структура, яка містить дві мережі зворотного зв'язку, генератор G і дискримінатор D. Існуючих методів домальовування зображень, заснованих на GAN, насправді не так вже й багато. Один з них описаний у роботі Чен і Ху [22], які запропонували метод нанесення семантичного зображення на основі GAN, названий прогресивним нанесенням зображень, де для відновлення зображення застосовується пірамідна стратегія відновлення від низької роздільної здатності до більш високої.

Нижче наведено частковий список мереж даного типу, які були розроблені для вирішення даної задачі:

- PEPSI – Fast Image Inpainting with Parallel Decoding Network [23];
- PEPSI++ – розширена версія PEPSI [24];
- deep convolutional generative adversarial network (DCGAN) [25] для відтворення відсутніх частин зображення;
- SRGAN – Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [26];
- EnhanceNet – Single Image Super-Resolution Through Automated Texture Synthesis [27].

Нещодавно було запропоновано декілька підходів, які поєднують деякі підходи GAN та CNN. Наприклад, у своїй роботі Цзяо та ін. [28] поєднав кодер-декодер та багатопарові шари згортки для відновлення зображень.

1.3 Переваги та недоліки методів відновлення зображень

Оскільки, за своєю природою та походженням, пошкодження на зображеннях різко відрізняються, то й методи та підходи для їх усунення не можуть бути універсальними, придатними для вирішення цих проблем. На жаль, не існує такого одного засобу, а якщо й був, то він однозначно не був би найоптимальнішим для кожного виду пошкоджень.

Кожен із перелічених методів для обробки зображення поступається чи переважає над іншим за певними критеріями. Для виявлення та опису особливостей кожного із них необхідно витратити чимало часу та мати чималий практичний досвід роботи з ними. Тому, у цьому розділі, будуть наведені найпоширеніші переваги та недоліки методів для кожного із видів задач, які стали відомі на даний час.

Дослідниками алгоритмів знешумлення було встановлено, що ідеальний алгоритм знешумлення повинен мати представлення з

переповненням (overcomplete representation), нелокальне групування, адаптивні основні функції та бути обчислювально ефективним. Окремий алгоритм компенсації шумів не може показати бажаний рівень застосовності, тому частіше застосовують їх поєднання у ансамблях.

Для методів, заснованих на використанні мереж, було виявлено, що завдяки зменшенню кількості шарів глибинних нейронних мереж, вдалося досягти великих успіхів у підвищенні ефективності видалення шумів. Таким чином було доведено, що зменшення глибини або ширини глибинних нейронних мереж не тільки впливає на складність цих мереж спрощуючи їх, але й не заважає якості видалення шуму зображення. Крім того, використання малого згорткового ядра та групової згортки може зменшити кількість параметрів, тим самим прискорюючи швидкість навчання. Злиття методів зменшення розмірів, таких як аналіз основних компонентів (PCA) та CNN, також може призвести до поліпшення ефективності дії алгоритмів знешумлення. Для покращення складних шумних зображень поетапна обробка виявилась дуже популярним методом. Наприклад, використання двоступеневого механізму – це спосіб боротьби з шумами та з низькою роздільною здатністю зображення.

Незважаючи на те, що розробники методів глибинного навчання досягли великих успіхів у розв'язанні задачі видалення шумів у різних сценаріях, все ще існують проблеми в цій області досліджень. До них належать наступні:

- глибинні мережі вимагають більше ресурсів пам'яті за інші методи;
- навчання нейронних мереж не є стабільним рішенням для справжніх зашумлених зображень та зображень з різним ступенем погіршенням якості;
- справжні шумні зображення не легко зафіксувати, що призводить до використання не завжди репрезентативних навчальних вибірок;

– глибоким CNN важко вирішити завдання відновлення без попереднього навчання;

– для зображення з шумом потрібно знайти більш точні показники. PSNR та SSIM є популярними показниками для завдання відновлення зображення. PSNR страждає від надмірного згладжування, яке дуже важко розпізнати схожими зображеннями. SSIM залежить від яскравості, контрасту та структури, і тому не може точно оцінити якість сприйняття зображення.

З детального аналізу випливає, що кожна категорія методів для поліпшення зображень є несумнівно важливою та має власний набір застосувань за певних умов. Наприклад, програми, які потребують миттєвих результатів, застосовують методи локальної фільтрації або методи перетворення доменів, оскільки нелокальне групування потребує значного обчислювального навантаження.

Більше того, було помічено, що більшість методів, як правило, створюють залишковий шум і не здатні підтримувати свою ефективність при більш високих рівнях шуму. Вивчення знешумлення, здається, наближається до своїх меж. Однак, для існуючих методів, все ще є напрямки для досліджень, наприклад, з метою:

- вдосконалення стабільності роботи;
- забезпечення високої візуальної якості та збереження дрібних особливостей деталей;
- пом'якшення артефактів;
- знаходження більш адаптивних базових функцій [2].

Для вирішення деяких з цих проблем вже було запропоновано методи. Наприклад, робота [29] повністю присвячена розробці методу сліпої деконволюції кольорових цифрових зображень за наявності лінійних спотворень та шуму з максимальним збереженням контурів, текстур та плавних градацій тонів для онлайн обробки зображень.

Завдяки тому, що технології глибокого навчання прогресують все далі, процес доповнення зображень став настільки автоматизованим, що в наші дні він взагалі не потребує втручання людини. Все що вимагається від користувача – просто подати пошкоджене зображення на вхід в нейронну мережу і отримати виправлений результат. Сучасні розробники надають таку можливість, наприклад, з веб-майданчиком NVIDIA, який демонструє, як мережа заповнює відсутні частини для будь-якого зображення. Але це можливо лише при попередньому довготривалому навчанні мережі та великому об'єму даних для цього.

Простіші засоби, наприклад послідовні методи, надійні для простих зображень, але коли зображення складне чи містить багато текстур та об'єктів, або зниклий об'єкт займає велику область зображення, то пошук подібного патча для відновлення пошкодженої частини може бути тривалим, а кінцевий результат – незадовільним.

Методи, засновані на GAN, показують гарну ефективність серед алгоритмів доповнення зображень, але їх швидкість навчання нижча і потребує дуже високої апаратної потужності. Це пояснюється вимогами до параметрів мережі та операції згортки, точніше, через їх значну кількість.

Для методів доповнення, які використовують CNN, як описано в попередньому розділі, ефективність кожного підходу пов'язана з розміром і типом використовуваних даних та реалізованою архітектурою.

Час обчислень – критичний показник для багатьох завдань комп'ютерного зору, особливо для програм, працюючих у режимі реального часу. Крім того, зі швидкістю розвитку методів глибокого навчання (тобто від CNN до GAN), час навчання, швидкість навчання та час виведення стають проблемою для методів обробки зображень/відео.

Поряд із багатообіцяючою продуктивністю, яку алгоритми глибокого навчання досягли в задачах SR, все ще залишається декілька важливих невирішених досі проблем. Наразі, актуальною постає потреба в більш легких архітектурах глибоких мереж для ефективної обробки. Хоч

для SR була досягнута висока точність вдосконалених моделей, все ще важко застосовувати ці моделі до реальних сценаріїв, що в основному пов'язане з масивними параметрами та обчисленнями. Іншою потребою вважається розробка більш ефективних алгоритмів DL для широкомасштабних задач низької роздільної здатності та задач з невідомим пошкодженням.

1.4 Постановка задачі

Наразі немає єдиного універсального варіанту для вирішення усіх перелічених видів задач – існують тільки засоби призначені для окремих випадків чи класів задач. У попередніх розділах були розглянуті різновиди задач, причину їх виникнення, та галузі, які потребують таких рішень. Для кожного випадку були вказані найбільш ефективні та популярні методи, перелічені їх основні особливості та обмеження, адже їх детальний аналіз заслуговує окремих досліджень і не є головною метою цієї роботи.

Підсумовуючи попередній розділ, виявлені переваги та недоліки, можна зробити наступні висновки:

- сучасні методи відрізняються своєю складністю та застосуванням, що робить пошук оптимального варіанту для певної проблеми можливим;
- універсального рішення на всі випадки не існує;
- дослідження окремих задач досягли своїх меж, що залишає дослідникам тільки працювати над оптимізацією раніше запропонованих алгоритмів;
- методи, засновані на глибинних мережах, дають найкращі результати для великого діапазону прикладів, але й потребують значних затрат, як часових, так і обчислювальних.

Враховуючи перелічені пункти, було вирішено зупинитися на вирішенні останньої, та найбільш актуальної проблеми.

Мета роботи – підвищення якості зображень, пошкоджених шумом, спотвореннями в ході реєстрації та передачі по мережам і каналам зв'язку.

Головна задача полягає у оптимізації структури глибокої мережі шляхом встановлення найбільш оптимальних компонентів, які у комплексі дадуть кращий результат.

Об'єкт дослідження – процеси формування та обробки цифрових зображень в інформаційних системах.

Предмет дослідження – нейромережевий підхід, зокрема, нейромережі з глибоким навчанням.

2 ДОСЛІДЖЕННЯ ГЛИБИННИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИДАЛЕННЯ ШУМІВ

У попередньому розділі були розглянуті види задач реконструкції зображень, методи та алгоритми для їх вирішення, головні недоліки та переваги, області та умови застосування. Для вирішення кожної із задач, крім алгоритмів різної ступені складності, дослідники неодмінно застосовували різні види глибинних мереж.

При порівнянні результатів відновлення різними методами, використання мереж часом виявляється надмірним та не надто оптимальним варіантом. Справді, незважаючи на кінцевий результат, якість та задовільність, витрати для його досягнення перекареслюють всі позитивні фактори.

Труднощі, які виникають під час пошуку репрезентативної навчальної вибірки, час навчання мережі, та кількість параметрів для налаштування роблять процес відтворення зображення складним у порівнянні з іншими алгоритмічними методами. Звичайно, після того, як всі ці етапи пройдені, а параметри налаштовані, для видалення шумів, чи збільшення роздільної здатності, чи доповнення певного зображення необхідно лише декілька мілісекунд, але чи варто воно того? Попри такий миттєвий результат, не кожен обере такий шлях для задоволення своїх потреб.

Таким чином, для того, аби можна було використати всю потужність глибинних мереж, на практиці необхідною та достатньою умовою буде оптимізація процесу навчання мережі, чи його вилучення, як обов'язкового етапу.

2.1 Глибинні мережі для очищення зображень від шумів

Безумовно, на кінцевий результат може значним чином вплинути багато параметрів. Неможливо проаналізувати всі можливі комбінації та варіанти усіх мереж щоб порівняти їх. Проте, можна сказати напевно, від зміни яких компонентів вихід мережі залежить найбільше. Тож, знаючи, що головним чином це залежить насамперед від архітектури мережі та вибору структурних елементів, починаючи від вибору функцій активації, закінчуючи оптимізатором, можна передбачити, якими будуть показники на виході.

2.1.1 Архітектури мереж

Однією з найперших мереж, яка мала достатньо компактну архітектуру та надавала можливість швидко обробляти зображення стала мережа VGG [30]. Її головною особливістю, яка стала новаторською на той час, була вперше застосована блокова структура для побудови. Використання блоків призводить до дуже компактних уявлень про визначення мережі, що дозволяє ефективно проектувати складні структури, зосереджуючись на більших елементах, що дозволяє збільшити масштаби та складність розроблювальних мереж.

Подібним чином, дизайн архітектур нейронних мереж поступово став більш абстрактним, адже дослідники перешли від мислення з точки зору окремих нейронів до цілих шарів, а тепер й до блоків, повторюючи шаблони шарів (рис. 2.1).



Рисунок 2.1 – Архітектура VGG

Після успішного створення більш глибоких мереж, дослідження перейшли до наступного логічного кроку – збільшення їх ширини. Таким чином, у наступній мережі GoogLeNet [31] було збільшено ширину для покращення продуктивності графічних програм. Більше того, GoogLeNet перетворив велике згорткове ядро на два менших ядра згортки, щоб зменшити кількість параметрів та обчислювальні витрати (рис. 2.2).

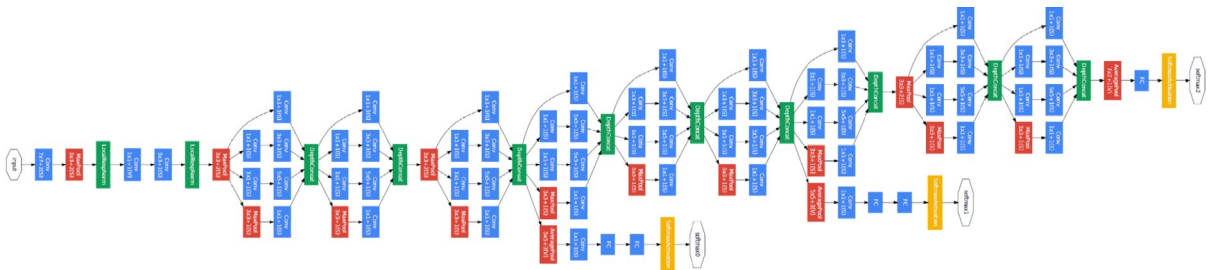


Рисунок 2.2 – Архітектура GoogLeNet

Хоча методи застосовані у мережах VGG і GoogLeNet підтвердили свою ефективність для обробки зображень, вони мають два значних недоліки: якщо мережа дуже глибока, це може призвести до зникнення або вибуху градієнтів; а якщо мережа дуже широка, вона може зазнати явища перенавчання. Для подолання цих проблем у 2016 році було запропоновано ResNet [32]. У цій мережі кожен блок був включений

функцію втрат, щоб згенероване зображення було ближче до чіткого, зберігаючи більше деталей кольору та вмісту.

Широкі експерименти доводять, що навчена мережа ефективна не лише для зображень із низькою роздільною здатністю в тестовому наборі, але також для інших зображень, що не входять в навчальний діапазон, таких як: розмивання зображень, знешкодження зображень медичної КТ, розмивання рухів тощо. Архітектура застосованих генератора та дискримінатора наведені на рисунку 2.4.

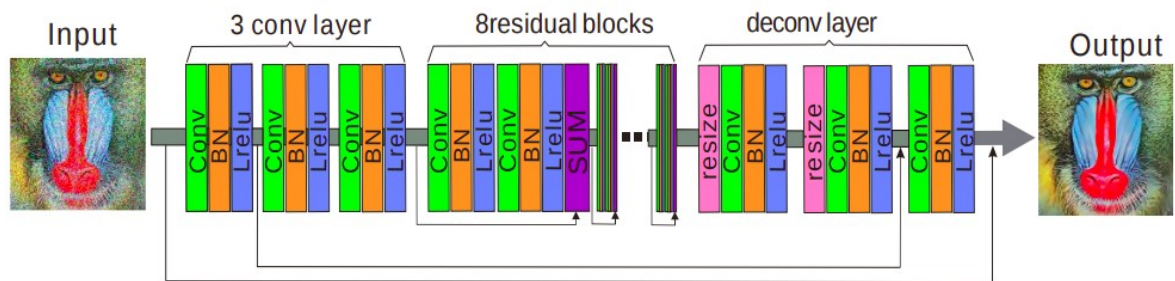


Figure 1. Generator Network

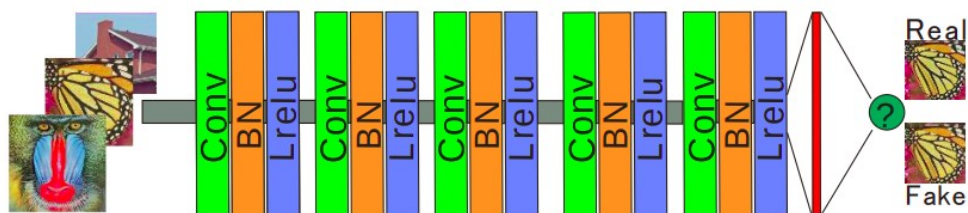


Figure 2. Discriminator Network

Рисунок 2.4 – Архітектура Deep Residual GAN

Автор наступної гібридної мережі [32] використовує глибинні згорткові генеративні змагальні мережі (DCGAN) для виконання різних завдань з обробки зображень, таких як надвисока роздільна здатність, шумозаглушення та деконволюція. DCGAN дозволяє використовувати єдину архітектуру для виконання різних завдань з обробки зображень та досягнення конкурентних балів PSNR. Хоча результати мережі показують трохи нижчий PSNR порівняно з традиційними методами, зображення, створені DCGAN, є більш привабливими та природними за своїм виглядом.

DCGAN може вчитися на великих наборах даних і автоматично додавати високочастотні деталі та функції до зображень, тоді як традиційні методи не можуть виконати цього. Саме архітектура генератора–дискримінатора DCGAN (рис. 2.5) стала причиною створення більш реалістичних та привабливих зображень.

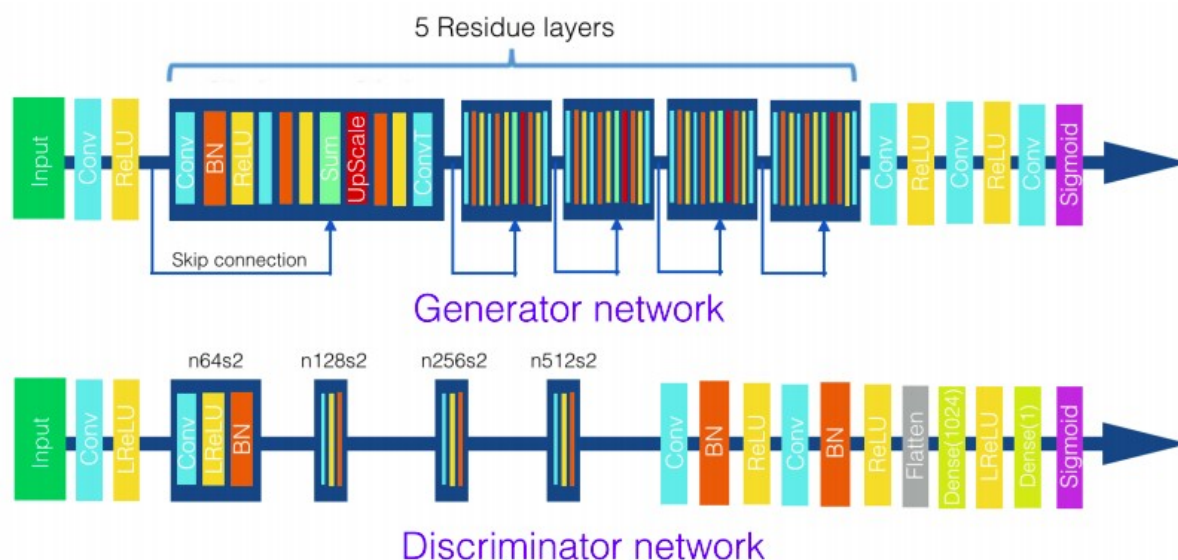


Рисунок 2.5 – Архітектура DCGAN

2.1.2 Оптимізатори мереж

Теоретичні дослідження з першого розділу підтвердили, що задача видалення шумів вже давно стала класичною і має велику кількість розв'язків. Проте, самих нейронних мереж часом може виявитися недостатньо, тож для кращого результату та з метою зменшення часу розробки, науковцями були запропоновані додаткові засоби, а саме, оптимізатори для мереж. Оптимізація стає нагальною при великій кількості параметрів для мереж, тож зазвичай використовується у поєднанні з глибокими мережами.

Глибинне навчання – це ітераційний процес. З такою кількістю параметрів, яку потрібно налаштувати, або методів, які потрібно спробувати, важливо мати можливість швидко навчати моделі, щоб швидко завершити ітераційний цикл. Звідси впливає важливість таких алгоритмів оптимізації, як стохастичний градієнтний спуск, мінімальний пакетний градієнтний спуск, градієнтний спуск з імпульсом та оптимізатор Adam.

Для покращення швидкості шумозаглушення метод оптимізації з використанням CNN був ефективним інструментом для швидкого пошуку оптимального рішення. Метод оптимізації Adam [35], безумовно, є одним з найкращих алгоритмів оптимізації для глибинного навчання, і його популярність певний час зростала дуже швидко. Він поєднує в собі і ідею накопичення руху, і ідею слабшого поновлення вагових коефіцієнтів для типових ознак.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (2.1)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \quad (2.2)$$

де m – середне значення градієнта;

v – квадрат градієнта;

g – градієнт поточної міні-партії;

β – гіперпараметри алгоритму.

Гіперпараметри мають справді хороші значення за замовчуванням 0,9 та 0,999 відповідно. Майже ніхто ніколи не змінює цих значень. Вектори ковзних середніх ініціалізуються нулями на першій ітерації.

Щоб побачити, як ці значення співвідносяться з моментом, визначеним як у рівняннях (2.1 – 2.2), розглянемо очікувані значення ковзних середніх. Оскільки m і v – оцінки першого та другого моментів, необхідно забезпечити таку властивість:

$$d_k = -H_k g_k, \quad (2.3)$$

$$E[v_t] = E[g_t^2]. \quad (2.4)$$

Очікувані значення повинні дорівнювати оцінюваному параметру, параметром у даному випадку є також очікуване значення.

Чим далі розширюється значення v , тим менше перші значення градієнтів впливають на загальне, оскільки вони множаться на дедалі менші бета-версії. Таким чином, можна переписати формулу для ковзного середнього:

$$v_t = (1 - \beta_2) \sum_{i=0}^t \beta_2^{t-i} g_i^2. \quad (2.5)$$

Тепер давайте подивимось на очікуване значення v , щоб побачити, як воно співвідноситься із справжнім першим моментом, щоб ми могли виправити невідповідність двох. Беручи очікування щодо лівої та правої сторін рівняння з (2.3) маємо (2.6).

$$E[v_t] = E[(1 - \beta_2) \sum_{i=0}^t \beta_2^{t-i} g_i^2] = E[g_t^2] (1 - \beta_2) \sum_{i=0}^t \beta_2^{t-i} + \zeta = E[g_t^2] (1 - \beta_2^t) + \zeta, \quad (2.6)$$

де ζ – константа, яка дорівнює 0, якщо справжній другий момент нерухомий.

Інакше ζ можна визначати малим значенням, оскільки експоненціальну швидкість занепаду β_1 можна (і потрібно) вибирати таким чином, щоб експоненціальна ковзна середня призначала невеликі ваги градієнтам занадто далеко в минулому.

За умови, що $t = 0$, отримуємо нескінченну послідовність однакових оновлень (2.7), (2.8).

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad (2.7)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad (2.8)$$

де – \hat{m}_t, \hat{v}_t оцінки, скориговані на зміщення.

Враховуючи попередні оцінки, можна описати правило Adam для оновлення параметрів (2.9).

$$\Theta_{t+1} = \Theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \hat{m}_t, \quad (2.9)$$

де Θ – параметри моделі для оптимізації;

η – крок ітерацій.

Хоч було не раз помічено деякі проблеми в певних областях з його використанням, дослідники продовжують працювати над рішеннями, щоб привести результати Adam в один ряд із SGD.

Найчастіше у протизвагу до алгоритму Adam пропонують використовувати Limited-memory BFGS (L-BFGS) [36] – алгоритм оптимізації з сімейства квазіньютонівських методів, який апроксимує алгоритм Бroyдена-Флетчера-Голдфарба-Шанно (BFGS) із використанням обмеженого обсягу пам'яті комп'ютера. Як і оригінальний BFGS, L-BFGS використовує оцінку зворотної матриці Гесса для керування пошуком у змінному просторі, але у випадку коли BFGS зберігає щільне $n \times n$ наближення до зворотного гессіана (n – кількість змінних у задачі), L-BFGS зберігає лише кілька векторів, які неявно представляють апроксимацію.

Нехай відоме x_k – положення на k -й ітерації, тоді значення градієнту для даної ітерації дорівнює (2.10).

$$g_k \equiv \nabla f(x_k) , \quad (2.10)$$

де f – деяка мінімізована, не опукла функція, яка має безперервні другі похідні.

Всі вектори є векторами стовпців. Ми також припускаємо, що ми зберегли останні m оновлень у наступному вигляді (2.11) та (2.12):

$$s_k = x_{k+1} - x_k , \quad (2.11)$$

$$y_k = g_{k+1} - g_k . \quad (2.12)$$

Алгоритм базується на рекурсії BFGS для зворотного гессіану як (2.13):

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T , \quad (2.13)$$

де $\rho_k = \frac{1}{y_k^T s_k}$ початкове наближення оберненого гессіана H_k^0 , з якого

починається пошук на ітерації k ;

V_k – визначається за виразом (2.14).

$$V_k = I - \rho_k y_k s_k^T . \quad (2.14)$$

Алгоритм описується наступним чином:

- 1) обрати значення x_0 , m , $0 < \beta' < 1/2$, $\beta' < \beta < 1$ та симетричну початкову матрицю H_0 ,
- 2) обчислити градієнт та вектор за (2.15) та (2.16),

$$d_k = -H_k g_k , \quad (2.15)$$

$$x_{k+1} = x_k + \alpha_k d_k . \quad (2.16)$$

де α_k – коефіцієнт, який задовольняє умові Вулфа (2.17) та (2.18),

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \beta' \alpha_k g_k^T d_k , \quad (2.17)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq f(x_k) + \beta g_k^T d_k . \quad (2.18)$$

3) визначити значення $\hat{m} = \min(k, m-1)$, оновити H_0 $\hat{m}+1$ разів використовуючи пари значень $\{y_j, s_j\}_{j=k-m}^k$,

4) оновити значення $k=k+1$ та перейти до кроку 2.

Завдяки своїй лінійній вимозі до пам'яті, метод L-BFGS особливо добре підходить для задач оптимізації з багатьма змінними. Попри це, однією з проблем методу Ньютона була негарантована збіжність до мінімуму для неопуклих функцій, яка притаманна також й L-BFGS.

Було опубліковано чимало наукових робіт, які порівнювали різні методи оптимізації для глибинних нейронних мереж. З їх результатів можна зробити наступні висновки, що:

- для більшості випадків Adam працював краще, ніж інші алгоритми;
- програмна реалізація для цього методу була набагато простішою порівняно з методами другого порядку;
- серед наукових робіт більш поширеною практикою є використання методів першого порядку при навчанні нейронних мереж;
- незважаючи на те, що є приклади методів другого порядку, які працюють краще за методи першого порядку, методи першого порядку добре працюють і є послідовними, що становить значну перевагу.

2.1.3 Функції активації

Як ми знаємо, у штучних нейронах з даних, які подаються на вхід, та вагових коефіцієнтів обчислюється зважена сума, яка потім передається функції активації, що й перетворює її у вихідний результат. Отже, в основному, функція активації використовується для зіставлення вхідних даних із вихідними. Саме ця функція допомагає нейронній мережі встановлювати складні взаємозв'язки та закономірності в даних.

Тепер питання полягає в тому, що трапиться, якщо ми не використовуємо жодну функцію активації і не дозволимо нейрону подавати зважену суму входів, у такому ж вигляді якою вона є на виході? У цьому випадку обчислення будуть дуже складними, оскільки зважена сума вхідних даних не має визначеного діапазону, тобто залежно від вводу вона може приймати будь-яке значення. Звідси й впливає одне важливе використання функції активації, що полягає у обмеженні виходу заданим діапазоном.

Іншим призначенням функції активації є додавання елемента нелінійності даних. Саме для цього в якості функцій активації зазвичай обирають нелінійні функції. Нелінійність означає, що вихід не можна згенерувати з лінійної комбінації входів.

Нелінійність важлива в нейронних мережах, оскільки лінійних функцій активації недостатньо для формування універсального апроксиматора функцій. Якщо ми використовуємо функції лінійної активації в глибинній нейронній мережі, з якої б кількості шарів вона не складалася, це буде еквівалентно звичайній лінійній нейронній мережі без прихованих шарів, оскільки ці функції лінійної активації можуть бути об'єднані, щоб сформувати ще одну єдину лінійну функцію. Отже, в такому випадку вся мережа буде зведена до одного нейрона з комбінованою лінійною функцією як функцією активації, і цей єдиний нейрон не зможе вивчити складні взаємозв'язки в даних. Оскільки більшість реальних

проблем дуже складні, на практиці, потрібні нелінійні функції активації в нейронній мережі, адже, нейронна мережа без нелінійних функцій активації буде простою моделлю лінійної регресії.

На практиці було визначено та випробувано чимало функцій активації, кожна з яких може бути використана для побудови будь-якої за складністю мережі. Найпопулярніші із них такі:

- лінійна;
- сигмоїдна;
- softmax;
- ReLU;
- Leaky ReLU;
- exponential linear unit (ELU);
- swish.

У даному розділі будуть розглянуті лише деякі з цих функцій, які були застосовані у мережах для вирішення задач видалення шумів із зображень, а саме: Leaky ReLU, ELU, Swish.

Leaky ReLU або нещільна ReLU [37] – це покращена версія функції ReLU. Функція ReLU може на кожній ітерації перевести деякі нейрони в такий стан, що вони стають неактивними для всіх вхідних значень, це явище також відоме як мертві зони. LReLU може подолати цю проблему, замість того, щоб давати на вихід 0 для всіх негативних значень, вона використовуватиме відносно невеликий компонент вводу для обчислення результату. Отже, ця функція ніколи не «вб'є» жодного нейрона.

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0.01x & \end{cases}, \quad (2.19)$$

де x – вхідне значення нейрона.

Тож, переваги LReLU полягають у тому що вона:

- проста для обчислення;

- не викликає проблеми зникаючого градієнта;
- не викликає проблеми мертвих зон.

Проте вона також стає причиною проблеми вибуху градієнта.

На практиці значення функції матимуть такий вигляд, як показано на рисунку 2.6.

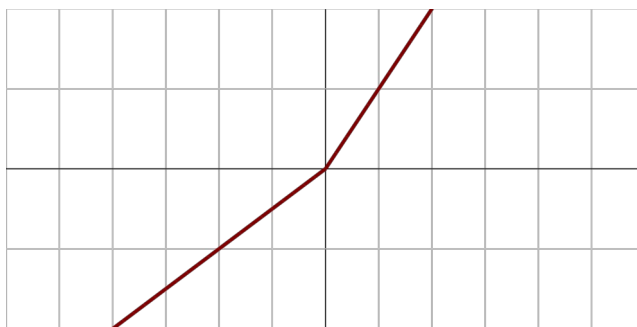


Рисунок 2.6 – Графік функції активації Leaky ReLU

Експоненціально-лінійна ReLU (ELU) – являє собою ще один різновид ReLU, який намагається зробити значення активації ближчими до нуля, що у свою чергу прискорює навчання. Такий підхід показує кращу точність класифікації, ніж ReLU. Функція може приймати негативні значення, які зміщують середнє значення ближче до нуля:

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ \alpha(e^x - 1), & \text{if } x \leq 0 \end{cases}, \quad (2.20)$$

де x – вхідне значення нейрона;

α – гіперпараметр, який налаштовується у діапазоні $[0, +\infty)$.

На відміну від попередньої функції активації, ця не спричиняє проблеми мертвих зон, але серед її недоліків можна відзначити наступні:

- обчислювально дорога функція;
- функція не уникає проблеми вибуху градієнта;
- необхідно заздалегідь визначити значення гіперпараметра альфа.

Значення функції складатимуть такий графік, як показано на рисунку 2.7.

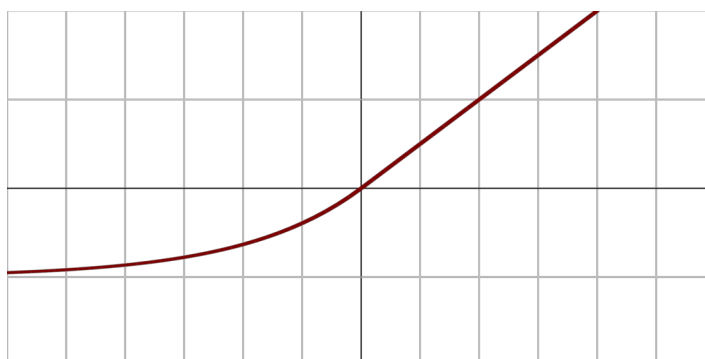


Рисунок 2.7 – Графік функції активації ELU

Swish – це плавна, немонотонна функція, яка послідовно відповідає значенням або перевершує ReLU у глибоких мережах різних доменів [38]:

$$f(x) = \frac{x}{1 + e^{-\beta x}}, \quad (2.21)$$

де β – може бути або константою, визначеною перед тренуванням, або параметром, який можна тренувати під час навчання.

Для $\beta = 1$ функція стає еквівалентною функції сигмовидно-зваженої лінійної (SiL), яка використовується для навчання підкріплення, тоді як для $\beta = 0$, функція перетворюється на масштабовану лінійну функцію $f(x) = x / 2$. При $\beta \rightarrow \infty$ – стає подібним до функції ReLU. Таким чином, її можна розглядати як згладжувальну функцію, яка нелінійно інтерполює між лінійною та функцією ReLU (рис. 2.8).

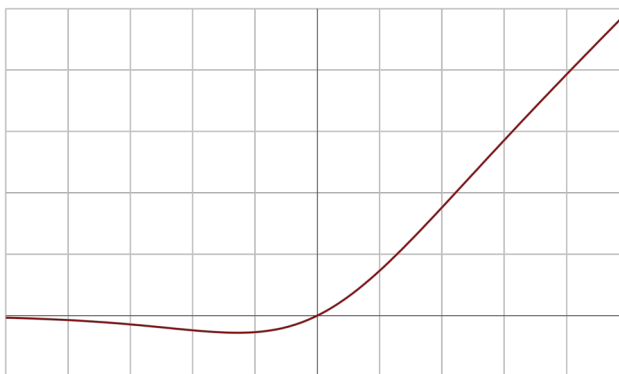


Рисунок 2.8 – Графік функції активації Swish

Плюси такого рішення наступні:

- Swish безперервна і диференційована у всіх точках;
- проста у використанні;
- на відміну від ReLU, не страждає від проблеми відмираючих нейронів;
- не потерпає від проблем вибуху або зникнення градієнтів.

Попри це Swish має такі недоліки:

- швидкість обчислення менша порівняно з ReLU та його модифікаціями, такими як Leaky ReLU та Parameterized ReLU, через використання сигмовидної функції, яка бере участь у обчисленні виходів;
- функція активації swish нестабільна і не може бути передбачена апріорі.

Розглянувши лише обмежену вибірку функцій активації, їх переваги та недоліки варто також зазначити, що це найпоширеніші у використанні на практиці функції.

Науковцями були визначені рекомендації щодо вибору найкращих функцій залежно від їх застосування:

- ReLU слід віддавати перевагу у використанні на прихованих шарах. Якщо це спричиняє проблему мертвих зон, тоді слід використовувати такі модифікації, як LReLU, ELU, тощо;

- для глибинних мереж swish працює краще, ніж ReLU;
- для останнього шару у випадку регресії лінійна функція є правильним вибором, для бінарної класифікації – сигмоїд, а для багатокласової класифікації – softmax.

2.2 Опис методу deep image priors

Глибинні згорткові мережі вже давно стали популярним інструментом для створення та відновлення зображень. Як правило, їх відмінна продуктивність обумовлена їх здатністю засвоювати реалістичні апіорні зображення (image priors) з великої кількості прикладів. У своїй роботі [39] автори наголошують на тому, що, навпаки, структура генераторної мережі є достатньою, щоб охопити велику кількість статистики низького рівня зображень до початку навчання. Для цього вони стверджували, що випадково ініціалізована нейронна мережа може бути використана як вручну налаштований апіорі і показати чудові результати у стандартних зворотних задачах, таких як видалення шумів, збільшення роздільної здатності та доповнення зображень.

Таким чином стало відомо, що, на відміну від загального переконання, навчання не є обов'язковим для створення хороших зображень, адже велика частина статистики з них фіксується згортковою структурою генератора, незалежно від цього етапу. Це справедливо та особливо помітно для статистичних даних, необхідних для вирішення різних проблем відновлення зображень, де попереднє зображення потрібно для інтеграції інформації, втраченої в процесах деградації.

Головна мета цього підходу – дослідити неявно захоплене апіорі, вибором певної структури мережі, перш ніж будь-який з її параметрів буде визначений навчанням. Для цього автори використовують максимальний апостеріорний розподіл для оцінки непоміченого значення на основі емпіричних даних:

$$x^* = \arg \max p(x_0 \vee x), \quad (2.22)$$

де x^* – відновлене зображення;

x_0 – пошкоджене зображення;

x – оригінальне непошкоджене зображення.

Тоді, використовуючи теорему Байєса, ми можемо висловити це як апіорну вірогідність:

$$p(x \vee x_0) = \frac{p(x_0 \vee x)p(x)}{p(x_0)} \propto \max p(x_0 \vee x)p(x). \quad (2.23)$$

Замість того, щоб працювати з розподілами окремо, ми можемо сформулювати рівняння як задачу оптимізації:

$$x^* = \operatorname{argmin} -\log p(x_0 \vee x) - \log p(x) = \operatorname{argmin} E(x \vee x_0) + R(x) \quad (2.24)$$

де $E(x \vee x_0)$ – термін даних, який є негативним логарифмом вірогідності;

$R(x)$ – апіорний термін зображення, регуляризатор.

Тепер завдання полягає в мінімізації рівняння (2.24) для зображення x . Звичайний підхід полягає в тому, щоб ініціалізувати x випадковим шумом, а потім обчислити градієнт функції щодо x і пройти по простору зображення, доки не буде досягнуто сходження у певній точці (рис. 2.9). Інший підхід полягає у побудові функції g , яка ініціалізується випадковим значенням θ , вихід якого з іншого простору може бути зіставлений із зображенням x , з подальшим оновленням значення θ , використовуючи градієнтний спуск (рис. 2.10).

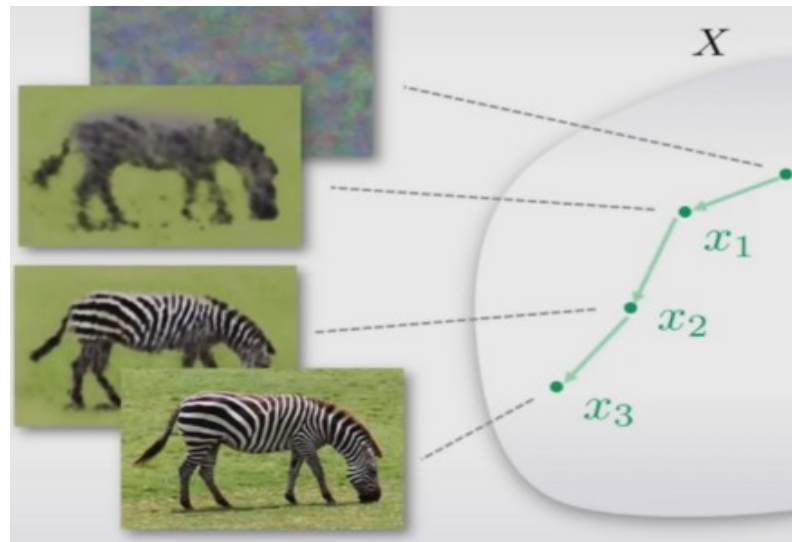


Рисунок 2.9 – Візуалізація класичного підходу

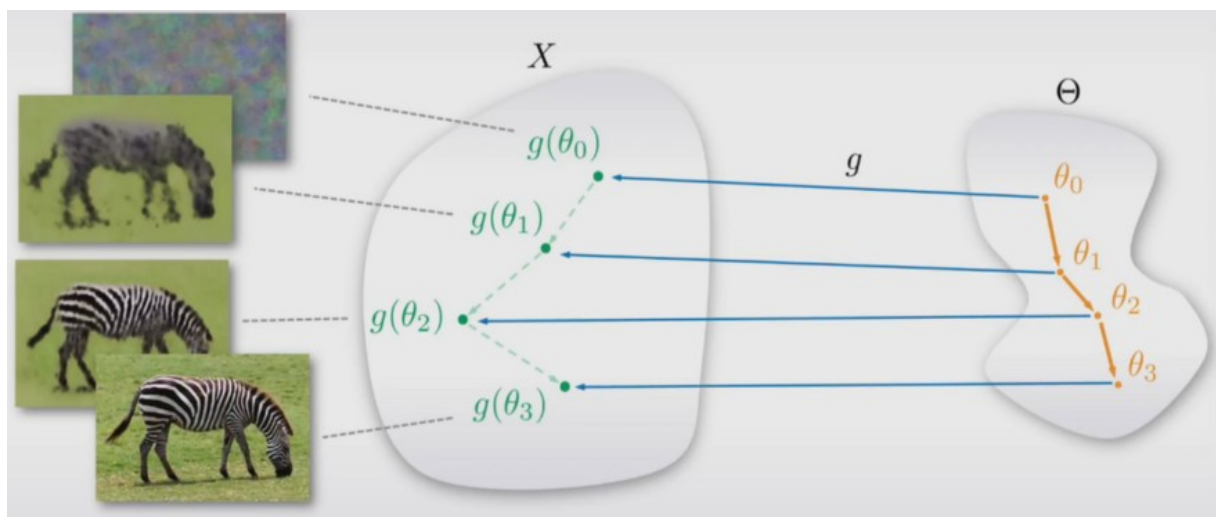


Рисунок 2.10 – Візуалізація параметризованого підходу

Отже, замість оптимізації по простору зображення, може бути виконана оптимізація по θ :

$$x^* = \arg \min E(g(\theta), x_0) + R(g(\theta)). \quad (2.25)$$

Такий підхід можливий, оскільки теоретично, якщо g сюр'єктивне відображення $g: \theta \rightarrow x$ (якщо принаймні одне θ відображається на зображенні x), то ця задача оптимізації рівнозначна, тобто має однакові рішення. Але на практиці функція g різко змінюється щодо того, як метод оптимізації здійснює пошук у просторі зображень. Ми можемо насправді розглядати g як гіперпараметр і налаштовувати його. За таких умов спостерігаємо, що $g(\theta)$ діє як ап'єрі, який допомагає вибрати хороше відображення, яке вже надалі дає бажане вихідне зображення та перешкоджає використанню неправильних зображень.

Отже, замість оптимізації суми двох компонентів можна виконати оптимізацію лише першого:

$$g(\theta) \equiv f_{\theta}(z), \quad (2.26)$$

$$\arg \min_{\theta} E(f_{\theta}(z), x_0), \quad (2.27)$$

де z – випадкове фіксоване вхідне зображення;

θ – випадково ініціалізовані ваги, які будуть оновлені за допомогою градієнтного спуску, щоб отримати бажане вихідне зображення.

У роботі [39] автори провели експеримент, який показав, що при використанні градієнтного спуску для оптимізації мережі, вона неохоче сприймає шумні зображення та набагато швидше і легше опускається до природних зображень (рис. 2.11). Зображення природного вигляду призводять до набагато швидшої конвергенції, тоді як шум відхиляється.

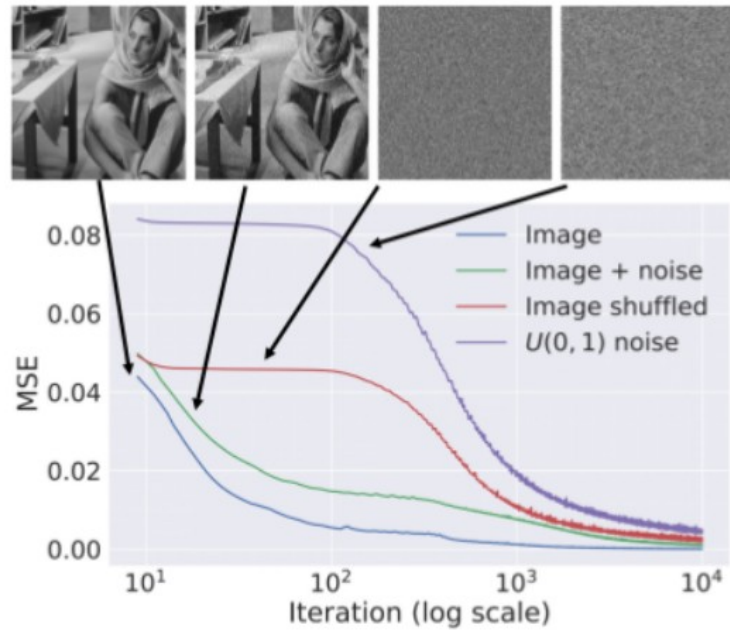


Рисунок 2.11 – Вивчення кривих для завдання реконструкції зображень з використанням різних зображень

Спираючись на приведені факти та формули можна описати алгоритм Deep Image Prior наступним чином. Нехай x_0 – пошкоджене зображення, тоді для його відновлення необхідно виконати наступні кроки:

1. ініціалізувати вхідні параметри: заповнити вхід z рівномірним шумом або будь-яким іншим випадковим зображенням;
2. вирішити та оптимізувати функцію градієнтним методом за (2.27) та (2.28);

$$\theta^{k+1} = \theta^k - \alpha \frac{\delta E(f_\theta(z), x_0)}{\delta \theta}. \quad (2.28)$$

3. Знаходимо оптимальний θ і отримуємо оптимальне зображення, передаючи фіксований вхід z до мережі з параметрами θ^* (формула 2.7).

$$x^* = f_{\theta^*}(z). \quad (2.29)$$

У весь процес відновлення зображення у загальному виді представлений на рисунку 2.12.

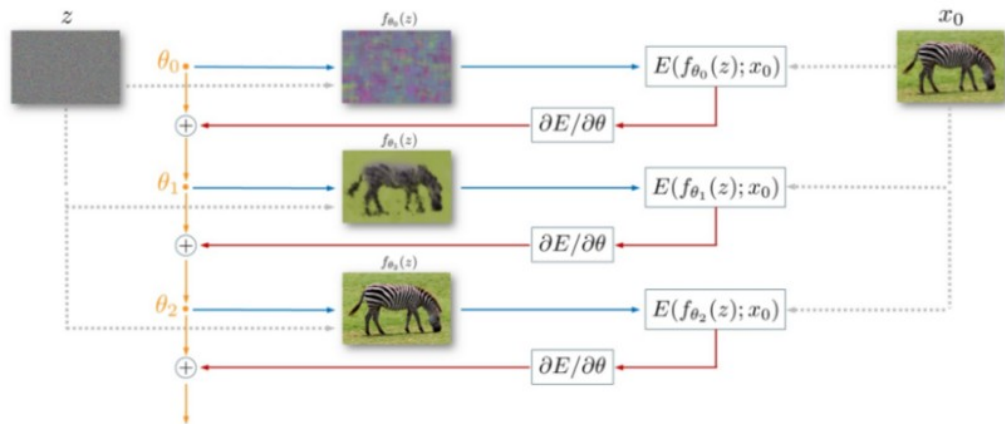


Рисунок 2.12 – Відновлення зображення за допомогою deep image prior підходу

3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ АРХІТЕКТУРИ ТА ПАРАМЕТРІВ ОБРАНОЇ МЕРЕЖІ

Розробники методу deep image prior [39] продемонстрували його переваги для низки прикладів, результати відновлення яких, не можуть не вражати, враховуючи відсутність стадії навчання мережі. Проте, їх метод все ж таки не можна вважати легким для застосування. Головні труднощі полягають у тому, що для кожного із наведених прикладів необхідно виконати попереднє індивідуальне налаштування мережі. Тобто, для нового прикладу конфігурація попереднього може виявитися не підходящою та мережа видасть на виході не оптимальний результат (рис. 3.1).

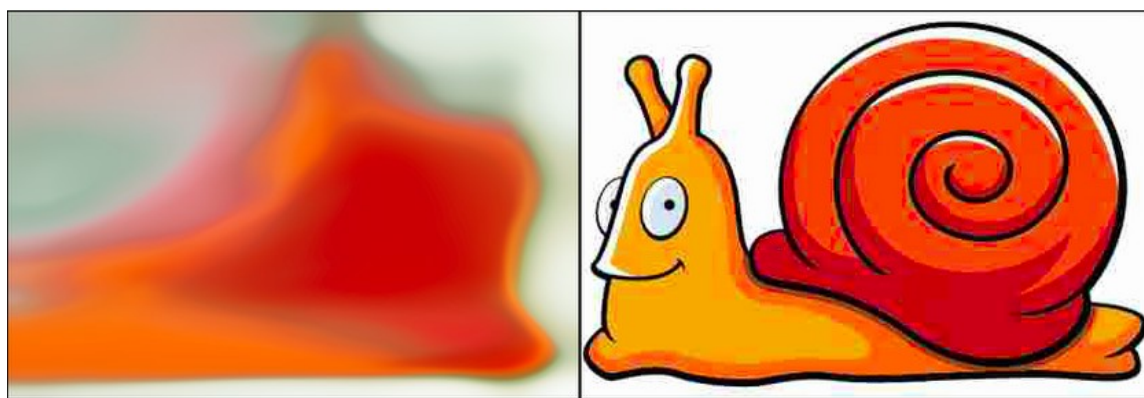


Рисунок 3.1 – Приклад видалення шумів без спеціального налаштування (результат та початкове зображення)

Серед параметрів, які мають бути встановлені, присутні наступні:

- тип мережі: у деяких випадках модифікується архітектура мережі; наприклад, додаються нові шари;
- метод оптимізації – автори використовують Adam та L-BFGS;
- функція активації (LeakyReLU, Swish, ELU);
- додаткові показники: рівень шумів, кількість ітерацій, кількість каналів в тензорі, тощо.

Звичайно, що для кожної задачі будь-яка обрана мережа повинна бути попередньо налаштована, проте це виконується для всіх спостережень одночасно. Саме тому, у цій роботі, буде виконана спроба налаштування мережі таким чином, що для кожного зображення видалення шумів буде виконано із достатньою якістю, яке буде оцінюватися за обраними кількісними показниками.

Для встановлення впливу зміни перелічених факторів були запропоновані та проведені наступні експерименти, у ході яких, за однакових попередньо встановлених та описаних умов, були визначені показники шумів для зображень до та після їх видалення.

У якості зображень для відновлення було обрано кілька різних зображень, які відрізняються рівнем присутніх шумів, діапазоном кольорів, деталізацією та складністю зображень, усі вони наведені на рисунках 3.2 – 3.6.



Рисунок 3.2 – Зображення для відновлення «Gary»



Рисунок 3.3 – Зображення для відновлення «Plane»

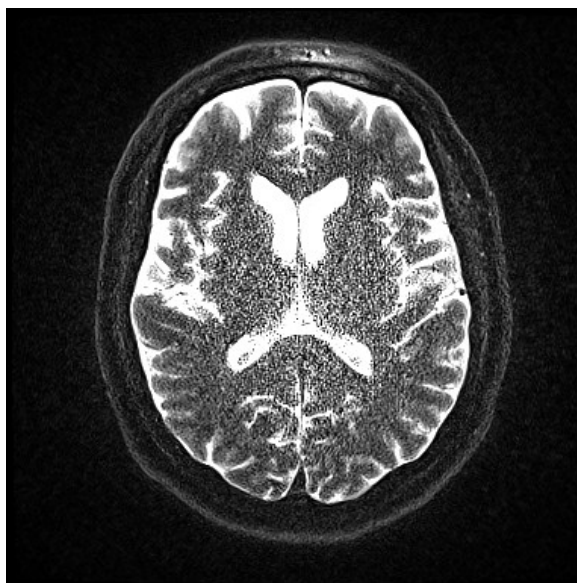


Рисунок 3.4 – Зображення для відновлення «Brain»



Рисунок 3.5 – Зображення для відновлення «Rocket»

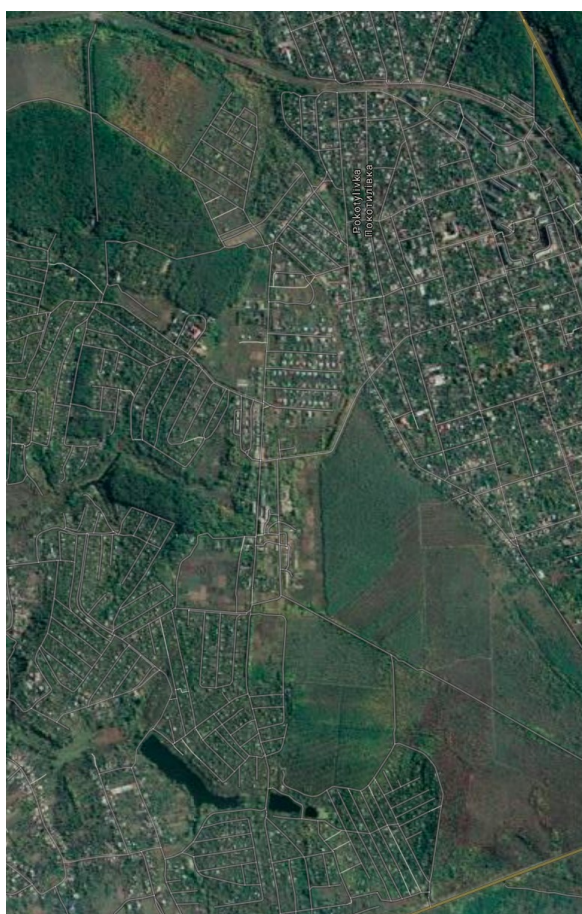


Рисунок 3.6 – Зображення для відновлення «Land»

Для проведення експериментів було вирішено використовувати CNN мережу, яка була запропонована авторами deep image prior [39] для проведення їх досліджень. Їх мережа налічує 572755 вагових коефіцієнтів та має конфігурацію, яку показано на рисунку 3.7.

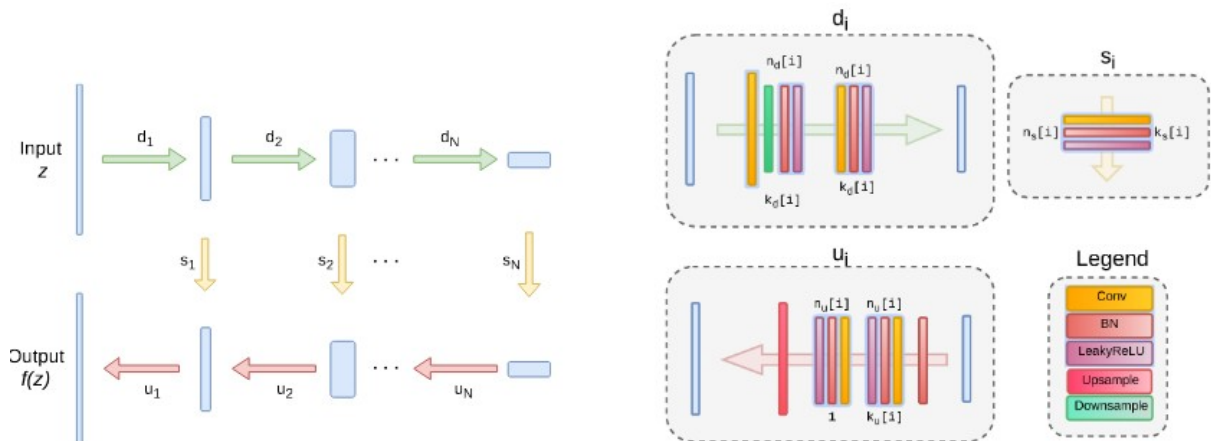


Рисунок 3.7 – Архітектура мережі

Для кількісної оцінки шумів зазвичай використовують значення пікового співвідношення сигнал/шум (PSNR) [40] для перевірки ефектів шумозаглушення. Для оцінки PSNR зображення необхідно порівняти його з ідеально чистим зображенням з максимально можливою потужністю. PSNR визначається наступним чином:

$$PSNR = 10 \log_{10} \left(\frac{(L-1)^2}{MSE} \right) = 20 \log_{10} \left(\frac{L-1}{RMSE} \right), \quad (3.1)$$

де L – кількість максимально можливих рівнів інтенсивності (мінімальний рівень інтенсивності припустимо дорівнює 0) на зображенні;

MSE – середньоквадратична помилка;

$RMSE$ – середня квадратична помилка.

MSE визначається як:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (O(i, j) - D(i, j))^2, \quad (3.2)$$

де O – представляє матричні дані вихідного зображення;

D – матричні дані деградованого зображення;

m – кількість рядків пікселів;

i – індекс цього рядка зображення;

n – кількість стовпців пікселів;

j – індекс цього стовпця зображення.

Усі експерименти були проведені для моделі, запропонованої авторами статті [39] з використанням їх кодової бази (додаток А). Були виконані самостійні налаштування та оптимізація алгоритмів для проведення наступних перевірок.

Реалізація моделей була виконана засобами мови програмування Python, з використанням бібліотек для створення нейронних мереж – PyTorch та numpy, у інтерактивному середовищі colab.

3.1 Оптимізатори мережі

Для визначення впливу оптимізаторів для досліджуваного підходу були проведені експерименти за однакових умов для алгоритмів Adam та L-BFGS з наступними налаштуваннями:

- функція активації – Leaky ReLU;
- архітектура мережі CNN;
- кількість ітерацій дорівнює 3500;
- зображення – gary, brain, plane, rocket;
- рівні шумів sigma 25 та 50 для кожного із варіантів.

Перед кожним запуском до зображень було штучно додано шуми відповідно до встановленого рівня. Після пригнічення шумів за допомогою

пропонованої мережі було обчислено значення функції втрат та пікове співвідношення сигнал/шум PSNR, результати надані в таблицях 3.1 та 3.2.

Таблиця 3.1 – Результати знешумлення зображення для різних оптимізаторів з рівнем шумів $\sigma=25$

Оптиміза тор	Оцінка	Snail	Brain	Plane	Rocket	Land
Adam	Loss	0.0074	0.0159	0.0089	0.0098	0.011
	PSNR	30.3212	17.9897	30.3382	32.7357	26.5483
L-BFGS	Loss	0.0736	0.0738	0.0167	0.01854	0.0167
	PSNR	11.3306	11.1528	21.1745	20.2697	21.1918

Таблиця 3.2 – Результати знешумлення зображення для різних оптимізаторів з рівнем шумів $\sigma=50$

Оптиміза тор	Оцінка	Snail	Brain	Plane	Rocket	Land
Adam	Loss	0.0244	0.0303	0.0309	0.0349	0.0331
	PSNR	25.1671	18.7839	27.6255	31.0372	25.1225
L-BFGS	Loss	0.0775	0.0728	0.0571	0.0395	0.0385
	PSNR	11.3613	11.4309	15.0629	22.4334	21.4148

Експериментальним шляхом було встановлено, що, незважаючи на наявні недосконалості після «видалення шумів», оптимізатор Adam дає кращий результат за L-BFGS за однакових умов та налаштувань, хоч його зображення мають дещо розмиті грані та контури. Результати можна оцінити візуально на рисунках 3.8 – 3.10.

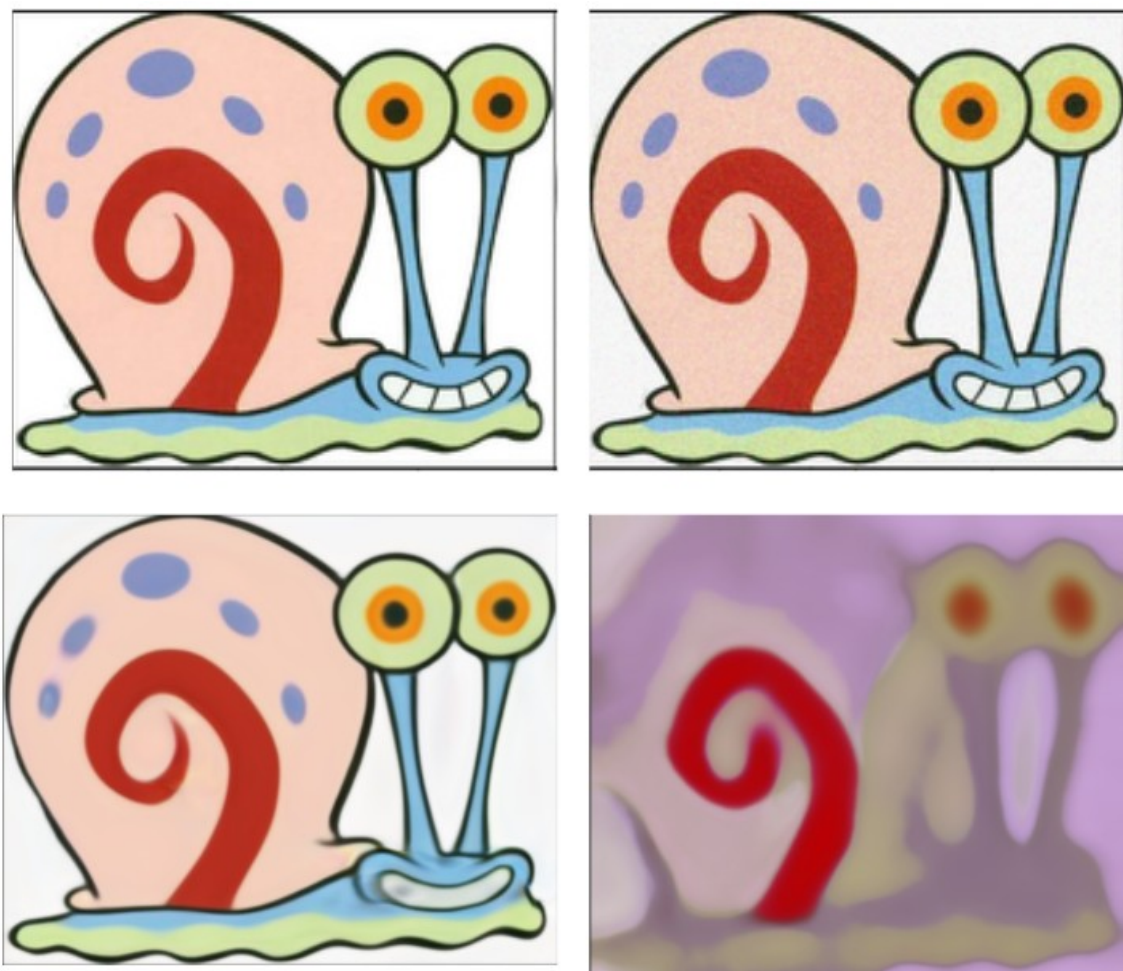


Рисунок 3.8 – Оригінальне, зашумлене та відновлені зображення «Gary» з оптимізатором Adam та L-BFGS

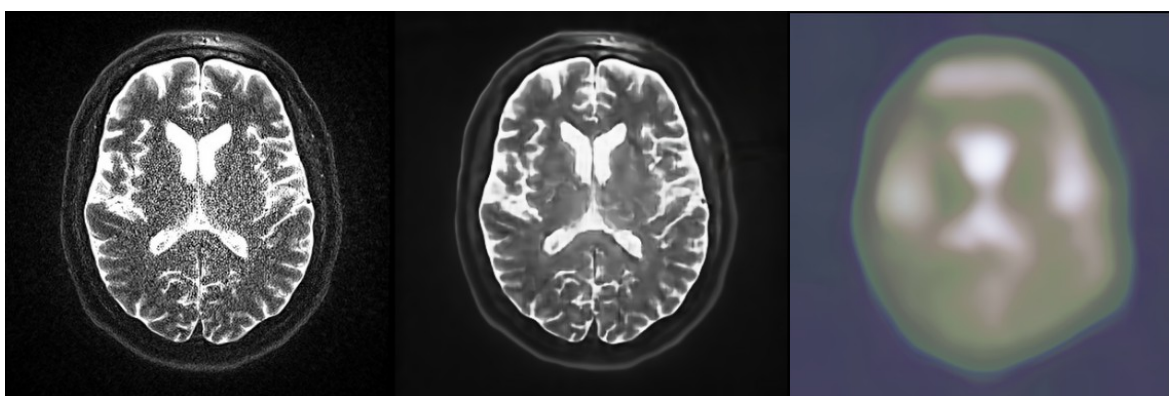


Рисунок 3.9 – Оригінальне та відновлені зображення «Brain» з оптимізатором Adam та L-BFGS



Рисунок 3.10 – Оригінальне та відновлені зображення «Land» з оптимізатором Adam та L-BFGS

3.2 Функції активації

Для визначення впливу функції активації для досліджуваного підходу були проведені експерименти за однакових умов для функцій Leaky ReLU, ELU та Swish з наступними налаштованими параметрами:

- оптимізатор – Adam;
- кількість ітерацій дорівнює 3500;
- архітектура мережі CNN;
- зображення – «Gary», «Rocket»;
- рівні шумів σ 50.

Перед кожним запуском до зображення було штучно додано шуми та визначено їх рівень. Отриманий показник рівня шумів та кінцеве значення після видалення шумів знаходиться у таблиці результатів 3.3 у стовпчику «PSNR noisy» та «PSNR». Він був використаний для обчислення різниці PSNR показника до та після знешумлення «delta».

Таблиця 3.3 – Результати знешумлення «Gary» для різних функцій активації

Функція	Loss	PSNR noisy	PSNR	delta
Leaky ReLU	0.007462952	21.27089311	29.73836895	8.46747584
ELU	0.007539869	21.22636156	29.39488701	8.16852545
Swish	0.007348333	21.33811172	30.15028539	8.81217367

Таблиця 3.4 – Результати знешумлення «Rocket» для різних функцій активації

Функція	Loss	PSNR noisy	PSNR	delta
Leaky ReLU	0.009881	20.051944	32.579374	12.527429
ELU	0.009864	20.059689	32.612652	12.552963
Swish	0.009648	20.155713	34.010993	13.855279

За результатами можна побачити, що використання функції Swish дає найкращий результат, наступною йде Leaky ReLU, а останньою ELU. Таким чином, окрім переваги у швидкості обчислення значень за допомогою Swish, маємо ще й кращі показники серед обраних функцій активації.

Початкові дані представлені на рисунку 3.8, а результати видалення шумів для кожної із перерахованих функцій активації можна побачити на рисунках 3.11 – 3.14. Результати можна вважати задовільними, та прийнятними для подальшого використання.

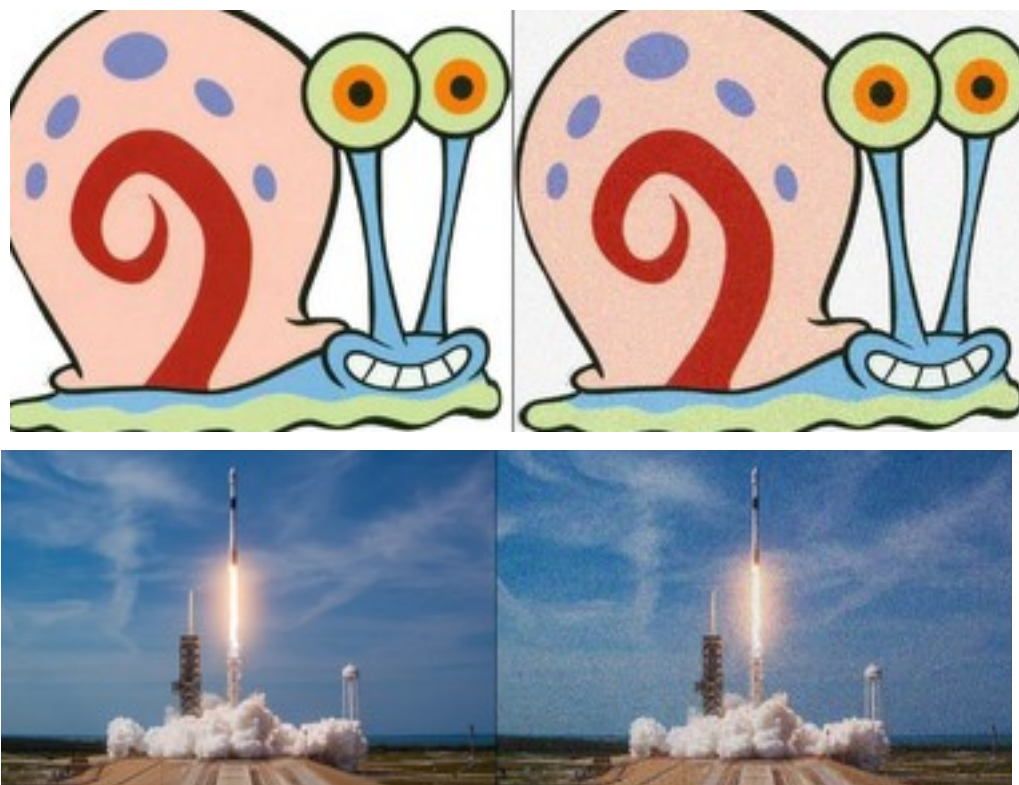


Рисунок 3.11 – Оригінальні та зашумлені зображення



Рисунок 3.12 – Відновлене зображення з функціями Leaky ReLU

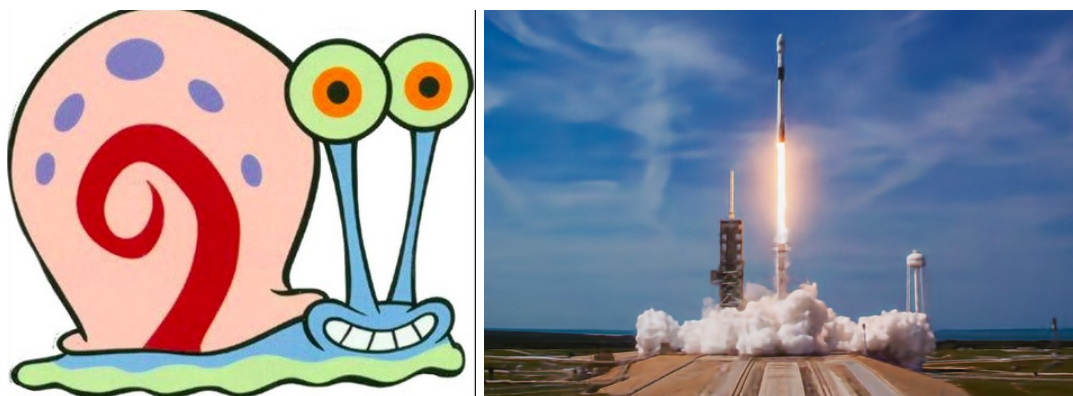


Рисунок 3.13 – Відновлене зображення з функцією ELU

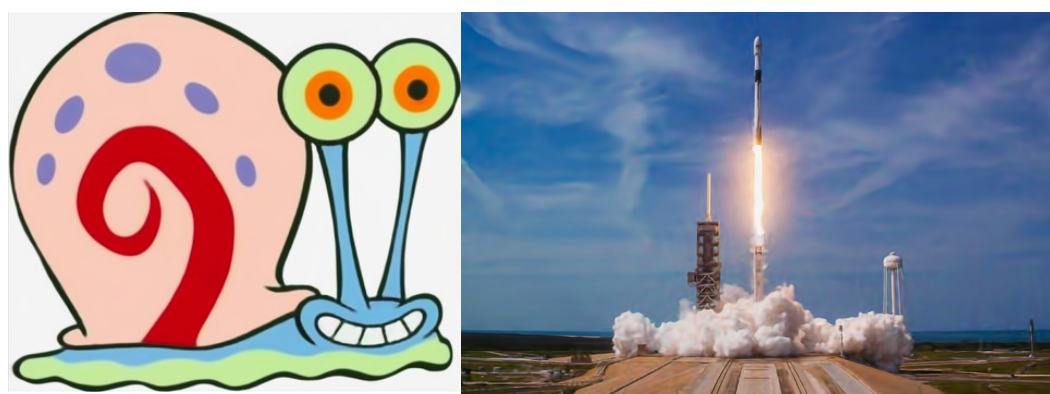


Рисунок 3.14 – Відновлене зображення з функцією Swish

3.3 Перевірка швидкодії

Головною перевагою deep image prior є все ж таки швидкість виконання обробки зображень. Аби перевірити чи це справді так, та скільки дійсно необхідно часу для відновлення одного зображення було і проведено наступний експеримент.

Оскільки тестові зображення представляють різні за складністю приклади, було вирішено провести їх відновлення та визначити час, який було витрачено для цього за наступних умов:

- оптимізатор – Adam, L-BFGS;
- функція активації – Leaky ReLU;

- кількість ітерацій дорівнює 3500;
- архітектура мережі CNN;
- рівні шумів σ 50.

Тривалість знешумлення обраних зображень для даних умов представлена в таблиці 3.5

Таблиця 3.5 – Результати швидкодії алгоритму

Зображення	Тривалість відновлення зображення, мкс	
	Adam	L-BFGS
Gary	429801	735319
Brain	621255	617267
Plane	319873	461370
Rocket	880845	542547
Land	155199	690545

Всі результати наведені в мікросекундах. Порівнявши їх можемо побачити, що оптимізатор Adam не лише виявився кращим для вирішення задачі видалення шумів, а й у декілька разів швидшим за L-BFGS. Можливо, за інших умов та організації шарів глибокої мережі можна було б досягти кращих результатів. Наразі можна напевно стверджувати, що для простих зображень, наприклад, «Gary», алгоритм видає найкращі результати.

ВИСНОВКИ

За останні кілька років засоби видалення шумів досягли великих успіхів, особливо в сценаріях, коли він регулярний. Однак у реальному світі таке трапляється рідко, бо шум зазвичай складний і нерегулярний. Тому дуже важливо визначити, як ефективно відновити чисте зображення із зображення з накладеними шумами. Більше того, хоч використання методів глибинного навчання для вивчення таких особливостей вимагає знань про початкові дані, отримані шумні зображення найчастіше не мають непошкодженого істинного варіанту. Це нагальні проблеми, для яких дослідники та вчені ще мають знайти вирішення.

Глибинні згорткові мережі вже давно стали популярним інструментом для створення та відновлення зображень. Як правило, їх відмінна продуктивність обумовлена їх здатністю засвоювати реалістичні апріорні зображення з великої кількості прикладів.

У цій роботі було розглянуто метод «Deep image priors» для відновлення зображень, автори якого наголошують на тому, що випадково ініціалізована нейронна мережа може бути використана як вручну налаштована апріорі і показати чудові результати у стандартних зворотних задачах, таких як видалення шумів, збільшення роздільної здатності та доповнення зображень. В роботі були виконані самостійні налаштування та оптимізація алгоритмів для проведення практичних досліджень.

Метою експериментальних досліджень стало визначення оптимальної для заданої мережі конфігурації, перевірка результатів роботи мережі з різними прикладами. Таким чином було виявлено, що найкращі результати можуть бути отримані при використанні оптимізатора Adam з функцією активації Swish.

Алгоритм було перевірено на швидкість обробки зображень. Результати демонструють, оптимізатор Adam потребує більше часу, але надає кращі результати у порівнянні з L-BFGS.

Реалізація моделей була виконана засобами мови програмування Python, з використанням бібліотек для створення нейронних мереж – PyTorch та numpy, у інтерактивному середовищі colab.

У підсумку можна зазначити, що дані розробки будуть корисними для швидкого відновлення не дуже складних зображень. Надалі можуть бути виконані додаткові вдосконалення та оптимізації для покращення результатів.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Banham M. R., Katsaggelos A. K. Digital Image Restoration. *IEEE Signal Processing Magazine*. 1997. Vol. 14, No. 2. P. 24–41.
- 2 Goyala B., Dograb A., Agrawala S., Sohic B.S., Sharmaa A. Image denoising review: From classical to state-of-the-art approaches. *Information Fusion*. 2020. Vol. 55. P. 220–244.
3. Lebrun M., Colom M., Buades A., Morel J.M. Secrets of image denoising cuisine. *Acta*. 2012. Vol. 21. P. 475–576.
4. Gonzalez R.C., Woods R.E. Image processing Digit. *Image Process*. 2007. P. 2.
5. Yan R., Shao L., Liu L., Liu Y. Natural image denoising using evolved local adaptive filters. *Signal Process*. 2014. No. 103, P. 36–44.
6. Bodyanskiy Ye.V., Kulishova N.Ye. Noisy distorted images recovery using blind deconvolution // *Radioelectronics. Computer Science. Control*. № 1(28), 2013. - С. 109 – 112.
7. Rangarajan A., Chellappa R. Markov random field models in image processing. *The Handbook of Brain Theory and Neural Networks* / Arbib M.A. Cambridge; MA: MIT Press. 1995. P. 564–567.
8. Chatterjee P., Milanfar P. Clustering-based denoising with locally learned dictionaries. *IEEE Trans. Image Process*. 2009. July 18, P. 1438–1451.
9. Liu J., Wang Y., Su K., He W. Image denoising with multidirectional shrinkage in directionlet domain. *Signal Process*. 2016. Vol. 125, P. 64–78.
10. Frosio I., Kautz J. Statistical nearest neighbors for image denoising *IEEE Trans. Image Process*. 2018. Vol. 28. P. 723–738.
11. Bodyanskiy Ye.V., Kulishova N.Ye. Extended neo-fuzzy neuron in the task of images filtering// *Radioelectronics. Computer Science. Control*. № 1(32), 2014. С. 112 – 119.

12. 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.
13. Park S.C, Park M.K., Kang M.G. Super-resolution image reconstruction: a technical overview. *IEEE Signal Process*. 2003. P. 21–36.
14. Elharrouss O., Almaadeed N., Al-Maadeed S., Akbari Y. Image Inpainting: A Review. *Neural Processing Letter*. December 2019.
15. Fan Q., Zhang L. A novel patch matching algorithm for exemplar-based image inpainting. *Multimed Tools Appl*. 2018. September 2. P. 10807–10821.
16. 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.
17. Yan Z., Li X., Li M., Zuo W., Shan S. Shift-net: Image inpainting via deep feature rearrangement. *Proceedings of the European conference on computer vision (ECCV)*. 2018. P. 1–17.
18. Chang Y-L., Yu Liu Z., Hsu W. VORNet: spatio-temporally consistent video inpainting for object removal. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2019.
19. Zeng Y., Fu J., Chao H., Guo B. (2019) Learning pyramid-context encoder network for high-quality image inpainting. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1486–1494.
20. 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.
21. 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.
22. Chen Y., Hu H. An improved method for semantic image inpainting with gans: progressive inpainting. *Neural Process Lett*. 2019. 49(3):1355–1367.

23. 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.
24. Shin YG., Sagong MC., Yeo YJ., Kim SW., Ko SJ. Pepsi++: fast and lightweight network for image inpainting. 2019. arXiv preprint arXiv:1905.09010
25. Dong J., Yin R., Sun X., Li Q., Yang Y., Qin X. Inpainting of remote sensing sst images with deep convolutional generative adversarial network. *IEEE Geosci Remote Sens Lett*. 2018. 16(2):173–177.
26. Ledig C., Theis L, Huszar F., Caballero J, Cunningham A., Acosta A., Aitken A., Tejan A.i, Totz J., Wang Z., Shi W. Photo-realistic single image super-resolution using a generative adversarial network. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
27. 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.
28. Jiao L., Wu H., Wang H., Bie R. Multi-scale semantic image inpainting with residual learning and GAN. *Neurocomputing*. 2019 331:199–212.
29. N. Kulishova and Y. Bodyanskiy, "Flexible 2D membership functions for images filtering using fuzzy peer group approach," *2015 Xth International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT)*, Lviv, 2015, pp. 82-84, doi: 10.1109/STC-CSIT.2015.7325437.
30. Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. ArXiv preprint arXiv:1409.1556.
31. 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.

32. He, K., Zhang, X., Ren, S., Sun, J., Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016. P. 770–778.
33. Ziyuan W., Lidan W., Shukai D., Yunfei L. An Image Denoising Method Based on Deep Residual GAN. *Journal of Physics: Conference Series*. Vol. 1550.
34. Qiaojing Y. DCGANs for image super-resolution, denoising and deblurring. 2017. URL http://stanford.edu/class/ee367/Winter2017/yan_wang_ee367_win17_report.pdf
35. Diederik P. Kingma, Ba J. Adam: A method for stochastic optimization. 2014. arXiv:1412.6980v9.
36. Liu D., Nocedal J. On the limited memory BFGS method for large-scale optimization *Math. Program.*, 45 .1989. P. 503-528.
37. Andrew L. Maas, Awni Y. Hannun, Andrew Y. Ng. Rectifier Nonlinearities Improve Neural Network Acoustic Models. 2014. URL: http://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf
38. Ramachandran P. Zoph B., Le Q. Swish: A Self-Gated Activation Function v1. Google Brain. arXiv:1710.05941v1.
39. Ulyanov D., Vedaldi A., Lempitsky V. Deep Image Prior. November 2017. ArXiv arXiv:1711.10925.
40. Hore A., Ziou D. Image quality metrics: Psnr vs. ssim. *20th International Conference on Pattern Recognition. IEEE*. 2010. P. 2366–2369.