

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерних наук _____

Кафедра _____ програмної інженерії _____

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____

Дослідження методів та моделей штучного інтелекту для покращення якості
зображень _____

(тема)

Виконав: студент 2 курсу, групи ПЗМ-22-1 _____

Тукало Ростислав Андрійович _____
(прізвище, ініціали)

Спеціальність 121 – Інженерія програмного _____
забезпечення _____
(код і повна назва спеціальності)

Тип програми освітньо-наукова _____
(повна назва освітньої програми)

Керівник проф. Лесна Н.С. _____
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри

(підпис)

З.В. Дудар _____
(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет _____ КН _____

Кафедра _____ програмної інженерії _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 121 – Інженерія програмного забезпечення _____
(код і повна назва)

Тип програми _____ освітньо-наукова _____

Освітня програма _____ Інженерія програмного забезпечення _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. Кафедри _____
(підпис)

« _____ » _____ 20 ____ р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

студентові _____ Тукалу Ростиславу Андрійовичу _____
(прізвище, ім'я, по-батькові)

1. Тема роботи: _____ Дослідження моделей та методів штучного інтелекту для покращення якості зображень _____

Затверджена наказом університету від «29» березня 2024 р. №250Ст

2. Термін подання студентом роботи до екзаменаційної комісії: «17» червня 2024 р.

3. Вихідні дані до роботи: виконати аналіз предметної галузі, навести принципи роботи та особливості застосування обраних моделей та методів ШІ, практично розробити та застосувати певну модель для покращення якості зображень _____

4. Зміст пояснювальної записки: вступ, аналіз предметної галузі, постановка задачі, огляд і практичне застосування методів ШІ, висновки, додатки. _____

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи (проекту)	Термін виконання етапів проекту (роботи)	Позначка про виконання
1	Аналіз предметної галузі	20.04.2024	<i>виконано</i>
2	Постановка задачі	21.04.2024	<i>виконано</i>
3	Огляд методів ШІ	15.05.2024	<i>виконано</i>
4	Підготовка даних та тренування мережі для практичного застосування	28.05.2024	<i>виконано</i>
5	Валідація результатів	30.05.2024	<i>виконано</i>
6	Оформлення пояснювальної записки	03.06.2024	<i>виконано</i>
7	Нормоконтроль, рецензування	12.06.2024	<i>виконано</i>
8	Підготовка презентації та доповіді	12.06.2024	<i>виконано</i>
9	Попередній захист	13.06.2024	<i>виконано</i>
10	Здача роботи у електронний архів	14.06.2024	<i>виконано</i>
11	Допуск до захисту у зав. кафедри	15.06.2024	<i>виконано</i>

Дата видачі завдання «29» березня 2024 р.

Студент _____ Тукало Ростислав Андрійович
(підпис)

Керівник роботи _____ проф. Лєсна Н.С.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ / ABSTRACT

Звіт: 70 стор., 38 рис., 16 джерел, 5 додатків.

ЗОБРАЖЕННЯ, НЕЙРОННА МЕРЕЖА, РОЗДІЛЬНА ЗДАТНІСТЬ,
ПОКРАЩЕННЯ ЯКОСТІ, ШТУЧНИЙ ІНТЕЛЕКТ.

Об'єкт дослідження – моделі та методи штучного інтелекту для покращення якості зображень.

Мета дослідження – огляд історії, принципів роботи, випадків застосування, а також безпосередньо практичне застосування обраних моделей та методів ШІ.

У результаті роботи було визначено модифікації алгоритмів, що більше пристосовані до тих чи інших сценаріїв використання, експериментальним шляхом випробувано використання згорткової нейронної мережі для підвищення роздільної здатності зображень.

ARTIFICIAL INTELLIGENCE, IMAGE, NEURAL NETWORK, RESOLUTION,
QUALITY ENHANCEMENT.

The object of study are models and methods of artificial intelligence for improving image quality.

The research goal is to review the history, principles of operation, cases of application, as well as the direct practical application of selected AI models and methods.

As a result of the work, modifications of algorithms that are more adapted to certain usage scenarios were identified, and the use of a convolutional neural network for image quality enhancement was experimentally tested.

Я, Тукало Ростислав Андрійович, студент гр. ІПЗм-22-1, здобувач вищої освіти на другому (магістерському) рівні кафедри програмної інженерії, заявляю: моя кваліфікаційна робота на тему «Дослідження методів та моделей штучного інтелекту для покращення якості зображень» виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу EIAg KhNURE. Усі запозичення з друкованих та електронних джерел мають відповідні посилання.

Я ознайомлений із діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови до допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

ЗМІСТ

Вступ.....	7
1 Аналіз предметної галузі	9
1.1 Аналіз предметної галузі.....	9
1.2 Огляд наукових матеріалів.....	17
1.3 Постановка задачі.....	22
2 Огляд методів штучного інтелекту для покращення якості зображень	25
2.1 Згорткові нейронні мережі	25
2.2 Генеративні змагальні мережі	30
3 Практичне застосування методів штучного інтелекту для покращення якості зображень	35
3.1 Підготовка даних та тренування мережі.....	35
3.2 Перевірка результатів.....	42
Висновки	48
Перелік джерел посилання	50
Додаток А Результати апробації дослідження.....	52
Додаток Б Результат перевірки роботи на академічну доброчесність	56
Додаток В Результати перевірки роботи на нормоконтроль	57
Додаток Г Перелік джерел посилання за науковими напрямками кафедри програмної інженерії	59
Додаток Д Слайди презентації до кваліфікаційної роботи	60

ВСТУП

Зображення є важливим засобом комунікації та передачі інформації, оскільки вони використовуються в усіх сферах нашого життя, від медіа та реклами до науки та медицини. Візуальна інформація має унікальну здатність передавати складні концепції та емоції, що робить її незамінним інструментом у сучасному суспільстві. Однак зображення часто можуть бути пошкоджені або знижені в якості через різні фактори, такі як шум, низька освітленість або деформація, спричинена, наприклад, рухом об'єктів під час зйомки. Ці проблеми негативно впливають на здатність зображень передавати інформацію чітко і ефективно.

Для покращення якості зображень можна використати такий потужний і перспективний інструмент, як штучний інтелект (ШІ). Цей напрям є однією з ключових галузей використання ШІ, зокрема в контексті комп'ютерного зору та візуальних технологій. Сучасні методи ШІ, включаючи нейронні мережі та глибоке навчання, здатні навчатися розпізнавати візуальні особливості в зображеннях, а потім використовувати ці знання для усунення шуму, підвищення контрастності та відновлення зображень. Ці алгоритми не лише відновлюють зображення до їх первісного стану, але й можуть покращити їхню якість до рівня, що перевищує початковий.

Застосування моделей ШІ для покращення якості зображень є інноваційним підходом, який відкриває нові можливості у багатьох сферах. Глибоке навчання, зокрема, дозволяє створювати алгоритми, здатні адаптуватися до різних типів шуму та деформацій, що робить їх універсальними інструментами для обробки зображень.

Головна мета цієї роботи полягає у вивченні історії виникнення методів штучного інтелекту, які можна використати для покращення якості або відновлення зображень. Також буде охоплено аналіз принципів роботи цих методів, їх еволюцію та визначення найбільш відповідних випадків і ситуацій застосування для зазначених

засобів штучного інтелекту. Розглядаючи історичний контекст і сучасні тенденції, ми спробуємо окреслити перспективи подальшого розвитку цієї галузі.

Оскільки використання штучного інтелекту для покращення якості зображень є відносно новим напрямом, існує багато можливостей для використання результатів досліджень такого роду. Це відкриває нові горизонти для науковців та інженерів, сприяє розвитку нових технологій та підходів, які можуть знайти застосування у різноманітних сферах, від медичних зображень та супутникових даних до розважальних медіа та візуальних ефектів. Вивчення цієї тематики є важливим кроком до створення інноваційних рішень, які зможуть значно покращити якість візуальної інформації, що оточує нас у повсякденному житті.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

1.1 Аналіз предметної галузі

Покращення якості зображень – це процес підвищення чіткості, деталізації та загальної якості зображень, що включає в себе усунення артефактів, покращення контрасту та відновлення втрачених елементів. Цей процес має велике значення у багатьох галузях, включаючи медицину, науку, мистецтво, безпеку та розваги. Це явище стало дуже важливим інструментом у багатьох галузях, зокрема завдячуючи розвитку штучного інтелекту, технологій глибокого навчання (deep learning) і машинного навчання (machine learning) [1]. Завдяки цим технологіям з'явилися нові методи та алгоритми, що дозволяють значно підвищити якість зображень, автоматизуючи і вдосконалюючи процеси обробки зображень. Різні галузі й надалі вивчають підходи комп'ютерного зору, а також впроваджують їх у продукти та пропонують результати для комерційного користування [2].

Першою з таких галузей є фото- і відеографія. Можливо, ця галузь є найближчою і найбільш використовуваною для звичайного необізнаного користувача: за статистикою 2019 року, смартфони використовувало 69% населення [3]. Смартфон зазвичай обладнаний камерою, здатною робити знімки та записувати відеороліки, а отже майже 70 відсотків українців (як, певно, і кожен власник смартфона у світі) регулярно камерою користуються – як для побутових задач (накшталт сфотографувати лічильник, зафіксувати важливі документи в цифровому форматі, зберегти цікаву друковану інформацію тощо), так і для збереження вражень, фотографування пейзажів, архітектури, інших пам'яток, або ж близьких людей.

Власне, для поліпшення досвіду фотографії, одним із застосувань алгоритмів штучного інтелекту є покращення якості аматорських фотографій. Для звичайного користувача без великого досвіду фотографії на мобільному телефоні найбільш частий варіант використання камери такий: дістати телефон і, не особливо занурюючись у просунуті налаштування зображення, натиснути кнопку зйомки. Так само

відбувається у випадку, якщо хочеться сфотографувати швидкоплинну композицію, спортивну ситуацію або динамічну обстановку тощо.

Покращення якості зображення дозволяє не замислюючись, у фоновому режимі виправити більшу частину недоліків, що могли з'явитися внаслідок швидкого використання камери. До таких недоліків належать «змазані» фотографії, що відбувається через рух смартфона під час зйомки, нечіткі текстури або краї об'єктів на фото, несфокусовані зображення, а також зображення, що містять шум, мають низький контраст або не чітко зафіксовані деталі. Варто також зауважити, що в цілому рівень камер у смартфоні з точки зору можливостей обладнання не є високим. Обмеженнями тут виступають розмір апаратного модулю, а також вартість якісного обладнання. Особливо корисними методи покращення якості стають для нічних фотографій, у випадку яких апаратна складова без застосування програмних засобів демонструє відчутно гірші результати. Комплекс програмних підходів, алгоритмів і методів для загального покращення якості зображень і компенсації недоліків фізичного обладнання отримав назву «обчислювальна фотографія» («computational photography») [4]. До цього ж поняття входять, зокрема, і методи штучного інтелекту, що розглядаються в цій роботі.

Іншим застосуванням є відновлення якості старих фотографій. Старі фотографії часто мають величезну емоційну та історичну цінність. Їхнє відновлення дозволяє зберегти ці спогади та важливі записи для майбутніх поколінь. Використання штучного інтелекту для повернення якості та фіксації старих фотографій надає унікальну можливість відновити зображення, що піддалися дії часу, були пошкоджені фізично або втратили свою початкову чіткість. Повернення якості та фіксація старих фотографій за допомогою штучного інтелекту надає можливість ділитися цими фотографіями в цифровому форматі та гарантувати, що вони не загубляться з часом або не зазнають подальшого фізичного пошкодження. Крім того, цифрові версії фотографій можуть бути використані в історичних дослідженнях,

музеях, архівах та інших контекстах, де важливим є збереження та доступ до візуальних матеріалів.

Наступною великою галуззю, в якій можуть мати широке застосування методи покращення якості зображень, є медицина [5]. Наприклад, одним з найпоширеніших типів медичної візуалізації у діагностиці є рентгенівське зображення. Рентгенівські знімки є важливим і цінним джерелом досліджень та діагностики захворювань, оскільки вони є недорогими та широко доступними. Рентгенівські промені є видом електромагнітного хвильового випромінювання, яке створює монохромне зображення внутрішніх органів людського тіла, в різних відтінках чорного та білого. Розрізняють декілька видів діагностичних процедур, включаючи мамографію, комп'ютерну томографію (КТ) та флюорографію. Мамографія може допомогти лікарям виявити рак молочної залози, а КТ-сканування використовує послідовність рентгенівських знімків, зроблених під різними кутами, щоб надати лікарям більш детальні зображення. Рентгенівські знімки грудної клітки, або ж флюорографія, використовуються для діагностики таких захворювань, як COVID-19, рак легенів і пневмонія.

Більшість отриманих медичних зображень мають низький контраст тканин, що ускладнює розрізнення між різними видами цих тканин. До видів діагностики, що потребують високої якості зображень, можна зарахувати і магнітно-резонансну томографію, яка використовується для визначення онкологічних діагнозів, а також для діагностики захворювань опорно-рухового апарату і центральної нервової системи. В результаті, покращення якості результатів таких досліджень є, мабуть, ключовим етапом на шляху до коректної діагностики та встановлення обставин захворювання [6]. Покращуючи якість результатів перелічених досліджень, наприклад, підвищуючи контраст, загальну яскравість або різкість, лікарі отримують можливість побачити деталі знімків, які без використання додаткової обробки залишилися б непоміченими (рис. 1.1).

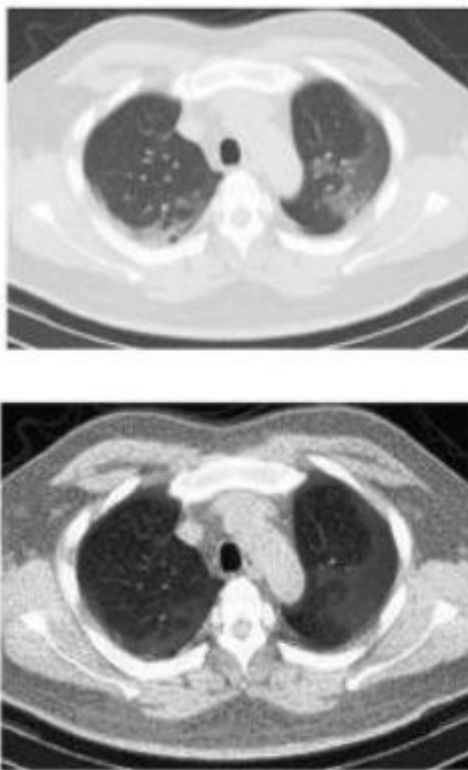


Рисунок 1.1 – Приклад покращення якості зображення – КТ-сканування легенів пацієнта з підозрою на COVID

Ще одним вагомим результатом використання штучного інтелекту для покращення якості зображень є підвищення безпеки пацієнтів і лікарів та зниження впливу діагностичних апаратів на тіла досліджуваних. Це має особливе значення в контексті рентгенологічних досліджень, оскільки ці методи використовують іонізуюче випромінювання, яке необхідно застосовувати в суворо обмежених обсягах через його потенційну шкідливість для здоров'я.

Покращення якості результуючих знімків за допомогою штучного інтелекту може суттєво знизити необхідні дози випромінювання під час діагностики. Завдяки здатності алгоритмів глибокого навчання ефективно обробляти і підвищувати чіткість зображень, можливо отримувати високоякісні діагностичні зображення при зменшенні інтенсивності випромінювання. Це, в свою чергу, знижує ризики, пов'язані з радіаційним опроміненням, для пацієнтів, що є критично важливим для їхнього

здоров'я та безпеки. Завдяки швидкому та точному покращенню якості зображень, можна досягти бажаних результатів за менший час, що зменшує тривалість впливу іонізуючого випромінювання на пацієнтів.

Останньою галуззю, більш розлого розглянутою в цьому переліку, стане безпека і спостереження. Однак, авжеж, існує набагато більше можливих випадків застосувань покращення якості зображень – це і електронна комерція, і створення контенту, і галузь розваг тощо.

Покращення якості зображень за допомогою штучного інтелекту може мати широке застосування у відеоспостереженні. Всеохопна мережа камер і пристроїв для нагляду є одним із основних компонентів безпеки та запобігання злочинам у містах. Таким чином, в будь-який момент часу важливо мати якісний потік зображень для фіксування потенційних злочинів, а отже для забезпечення невідворотності покарання. На заваді цьому можуть стати як обмеження апаратної складової систем відеоспостереження, так і важкі умови для зйомки (зазвичай темний час доби); загалом, дещо повторюються проблеми використання камер у смартфонах.

Освітлення відіграє надзвичайно важливу роль у якості відео з камер відеоспостереження. Для ефективного використання цих записів у розслідуваннях злочинів необхідно підтримувати середній рівень освітлення (рис. 1.2). Як надто яскраве, так і недостатнє світло можуть значно знизити інформативність відео, ускладнюючи ідентифікацію підозрюваних та спостереження за їхньою діяльністю. Загалом, криміналісти покликані гарантувати достовірність цифрових доказів. Проте, іноді відеозаписи з камер відеоспостереження низької якості відображають номерний знак, логотип або текст, які важко прочитати. У такому випадку для отримання ключової інформації для подальшого розгляду справи використовується безпечний та ефективний засіб покращення якості відео з камер відеоспостереження [7].



Рисунок 1.2 – Приклад покращення якості зображення на сцені з низьким освітленням

Наостанок, супутникова зйомка є важливим інструментом для моніторингу навколишнього середовища, управління ресурсами та реагування на надзвичайні ситуації. Однак якість супутникових зображень може бути обмежена факторами, такими як атмосферні умови, обмеження сенсорів та обмеження роздільної здатності. На супутникових знімках може виникнути шум через різні електронні відхилення в сенсорах. Невдалий дизайн або обмежений бюджет можуть призвести до дефектів у знімках: наприклад, через обмеження технічних характеристик сенсорів, яке іноді призводить до низької роздільної здатності зображень, що може бути недостатнім для певних науково-дослідницьких або комерційних потреб. Атмосферні ефекти також спричиняють дефекти на знімках. Їх причинами є розсіювання світла, вологість, аерозолі та інші атмосферні фактори, які можуть викривити якість знімка, змінивши яскравість та колірну гаму. Для вирішення цієї проблеми використовують алгоритми атмосферної корекції та видалення туману, щоб компенсувати атмосферні ефекти

(приклад подано на рисунку 1.3). Ці методи моделюють атмосферні умови та коригують знімки відповідно до них.



Рисунок 1.3 – Приклад усунення атмосферних ефектів із супутникового знімку

Відповідно, випадки застосування технологій покращення якості зображень дещо повторюються і тут. Основними аспектами є підвищення роздільної здатності зображень, зменшення шуму і корекція кольорової гами. Такі вдосконалення зображень можуть мати значний вплив на різні сфери, зокрема на моніторинг змін у навколишньому середовищі. Наприклад, результати обробки зображень можна використовувати для виявлення змін у структурі та поширенні лісів, льодовиків або водойм. Це дозволяє більш точно відстежувати екологічні зміни, що є критично важливим для оцінки впливу глобального потепління та інших екологічних проблем. Крім того, покращені зображення можуть бути використані для виявлення незаконної вирубки лісів, браконьєрства або забруднення води. Це надає можливість більш ефективно боротися з екологічними злочинами та забезпечувати охорону природних ресурсів. Також технології покращення якості зображень можуть допомогти виявляти

людей або пошкодження та руйнування після катаклізмів чи стихійних лих. Це має велике значення для оперативного реагування на надзвичайні ситуації та надання допомоги постраждалим.

Особливо важливо зазначити можливість застосування технологій покращення якості супутникових знімків у розвідувальних і військових цілях (рис. 1.4). У цій сфері вдосконалені зображення можуть бути використані для виявлення, підрахунку та аналізу наявних сил і засобів супротивника, фіксації переміщення та розгортання ворожих угруповань тощо. Це надає значну перевагу у військових операціях та забезпечує більш точну і надійну інформацію для прийняття стратегічних рішень.



Рисунок 1.4 – Приклад покращення якості супутникового знімку з авіаційною технікою

Таким чином, вище було подано декілька галузей застосування технологій покращення якості зображень, для яких дане дослідження може мати вагомні практичні наслідки. Ці галузі включають, але не обмежуються, медичну візуалізацію, де підвищена якість зображень може значно покращити діагностичну точність та ефективність лікування; супутникову та аерофотозйомку, де висока якість зображень

є критичною для точного картографування та моніторингу змін навколишнього середовища.

Очевидно, список можливих випадків використання не є вичерпним. Інші перспективні напрями можуть включати відновлення історичних та мистецьких артефактів, де покращення якості зображень може суттєво допомогти в реставрації та збереженні культурної спадщини. Відновлені зображення стародавніх документів, картин та інших артефактів можуть стати більш доступними для дослідників та широкої публіки, сприяючи збереженню культурної спадщини для майбутніх поколінь.

Крім того, технології покращення якості зображень можуть знайти своє застосування у наукових дослідженнях у таких галузях, як астрономія або мікробіологія, де висока деталізація є важливою для отримання точних наукових даних. В астрономії, наприклад, покращення якості зображень може допомогти виявити нові небесні тіла або деталі космічних об'єктів, які раніше були недоступні для спостереження. У мікробіології це може сприяти точнішому аналізу зразків та відкриттю нових біологічних явищ.

Враховуючи широкий спектр можливих застосувань, дане дослідження має потенціал значно вплинути на різні галузі, сприяючи підвищенню точності, ефективності та інформативності зображень. Подальші дослідження та розробки у цьому напрямі можуть відкрити нові можливості для використання технологій покращення якості зображень, що матиме позитивний вплив на різні аспекти науки, техніки та повсякденного життя. Це, в свою чергу, підкреслює важливість інвестування в дослідження та розвиток у галузі штучного інтелекту та комп'ютерного зору, що може призвести до значних покращень у різних сферах людської діяльності.

1.2 Огляд наукових матеріалів

Штучний інтелект (ШІ) набуває все більшого значення в різних галузях науки і техніки, демонструючи свій потенціал у вирішенні складних завдань, які раніше

вимагали значних людських ресурсів. Однією з таких галузей є обробка та покращення зображень, де застосування алгоритмів ШІ дозволяє досягти високої якості результатів, які важко досягти традиційними методами. Сьогодні існує значна кількість наукових матеріалів, робіт, статей щодо застосування штучного інтелекту для покращення якості зображень. Ці дослідження, створені як українськими науковцями, так і закордонними, охоплюють широкий спектр методів, таких як глибоке навчання, нейронні мережі, машинне навчання та інші передові технології. Важливим аспектом досліджень є також етичні та правові питання, пов'язані із застосуванням ШІ, зокрема конфіденційність даних та авторські права на алгоритми та результати їх роботи.

До прикладу, можна навести статтю під назвою «Adaptive Image Enhancement Model for the Robot Vision System» (адаптивна модель покращення якості зображень), що була подана викладачами кафедри ШІ К.С. Смеляковим та А.С. Чуприною спільно з литовськими колегами на 14 Міжнародну науково-практичну конференцію в Резекне, Латвія [4]. У цій роботі розкривається питання покращення якості зображень для роботизованих систем, що обладнані камерами. Робототехніка є одним із важливих напрямків сучасного розвитку науки і технологій. Більшість сучасних роботів і дронів мають власну систему візуалізації, включаючи відеокамеру, яку вони використовують для зйомки цифрових фотографій та відеопотоків. Ці дані використовуються для аналізу ситуації у полі зору камери робота, а також для визначення алгоритму поведінки робота в режимі реального часу. У контексті новизни у статті розглядається спеціальна поліноміальна математична модель і методи адаптивної градаційної корекції зображення. Запропонована модель і метод дозволяють самостійно підлаштовуватися до шкал яскравості та форматів зображення і оптимально виконувати градаційну корекцію зображення за різних умов освітлення. Таким чином, забезпечується ефективність усього наступного циклу аналізу зображень у системі візуалізації робота.

Схожій тематиці присвячена стаття «Image Enhancement Guided Object Detection in Visually Degraded Scenes» (розпізнавання об'єктів на основі покращення якості зображень у візуально погіршених сценах), опублікована у журналі «IEEE Transactions on Neural Networks and Learning Systems» [8]. Точність виявлення об'єктів значно погіршується у візуально погіршених сценах. Логічним рішенням є спочатку покращити зображення, а потім виконати виявлення об'єктів. Однак, це не є рішенням оптимальним і не обов'язково призводить до покращення виявлення об'єктів через розділення завдань покращення зображення та виявлення об'єктів. Для вирішення цієї проблеми дослідники пропонуємо метод виявлення об'єктів, керований покращенням зображення, який вдосконалює мережу виявлення шляхом додавання гілки покращення. Зокрема, гілка покращення та гілка виявлення організовані паралельно, і розроблено модуль керування характеристиками для з'єднання двох гілок, який оптимізує поверхневі характеристики вхідного зображення у гілці виявлення, щоб вони були максимально узгоджені з характеристиками покращеного зображення. Оскільки гілка покращення зафіксована під час навчання, така конструкція дозволяє використовувати характеристики покращених зображень для керування навчанням гілки виявлення об'єктів, щоб зробити навчений модуль виявлення об'єктів обізнаним як про якість зображення, так і про виявлення об'єктів. Під час тестування гілка покращення та модуль керування характеристиками видаляються, і тому не додається жодних додаткових обчислювальних витрат на виявлення. Розширені експериментальні результати на наборах даних з виявлення об'єктів під водою, в умовах туману та при слабкому освітленні демонструють, що запропонований метод може значно покращити продуктивність популярних мереж виявлення (YOLO v3, Faster R-CNN, DetectoRS) у візуально погіршених сценах.

Наступним розглянутим дослідженням є стаття «Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images» (дослідження впливу технік покращення якості зображень на виявлення COVID-19 за

рентгенологічними знімками грудної клітини), опублікована в журналі «Computers in Biology and Medicine» [9]. Комп'ютеризована діагностика для надійного та швидкого виявлення коронавірусної хвороби (COVID-19) стала необхідністю для запобігання поширенню вірусу під час пандемії та для зниження навантаження на систему охорони здоров'я. Рентгенографія грудної клітки (CXR) має кілька переваг над іншими методами візуалізації та виявлення. Було повідомлено про численні роботи щодо виявлення COVID-19 на основі невеликого набору оригінальних рентгенівських знімків. Однак у літературі до цього не було повідомлень про вплив покращення зображень та сегментації легень великого набору даних на виявлення COVID-19. Була запропонована нова модель U-Net і порівняна зі стандартною моделлю U-Net для сегментації легень. Нова модель U-Net показала точність на рівні 98.63% для сегментації легень. Запропонований підхід сприятиме швидкому та надійному виявленню COVID-19 за допомогою рентгенівських знімків грудної клітини.

У науковому просторі є й інші дослідження у галузі охорони здоров'я, що присвячені не тільки хворобі, що спричинила останню пандемію. Оглянемо статтю «Image Enhancement for Tuberculosis Detection Using Deep Learning» (покращення якості зображень для виявлення туберкульозу за допомогою глибокого навчання), що її опубліковано в журналі «IEEE Access» [10]. Останнє дослідження Всесвітньої організації охорони здоров'я (ВООЗ) за 2018 рік показує, що щороку близько 1,5 мільйона людей помирають та близько 10 мільйонів людей інфікуються туберкульозом. Крім того, більше ніж 4000 людей щодня помирають від туберкульозу. Частину цих смертей можна було б запобігти, якби хворобу виявили раніше. У сучасній літературі можна знайти важливі роботи з автоматизації діагностики шляхом застосування методів глибокого навчання до медичних зображень. Незважаючи на те, що глибоке навчання дало обнадійливі результати у багатьох галузях, всебічні діагностичні дослідження туберкульозу залишаються обмеженими. Глибоке навчання вимагає великої кількості високоякісних навчальних зразків для досягнення кращих

результатів. Через низький контраст рентгенограм грудної клітини, вони часто мають низьку якість. Це дослідження оцінює вплив покращення зображень на ефективність техніки глибокого навчання для вирішення цієї проблеми. Використаний алгоритм покращення зображень зміг виділити загальні або локальні характеристики зображень, включаючи деякі цікаві особливості. Покращені зразки зображень потім подавалися на попередньо навчені моделі ResNet та EfficientNet для трансферного навчання. У наборі даних зображень (рентгенограм пацієнтів, що хворіють на туберкульоз) дослідники досягли 89,92% точності класифікації.

Цікавими є дослідження щодо покращення якості медичних зображень, що не стосуються конкретних захворювань, а радше методів діагностики. Наприклад, оглянемо статтю «TXI: Texture and Color Enhancement Imaging for Endoscopic Image Enhancement» (покращення якості текстур і кольорів для покращення якості зображень ендоскопічних досліджень) із журналу «Journal of Healthcare Engineering» [11]. Виявлення уражень зі слабкими морфологічними та/або кольоровими змінами під час ендоскопії з використанням білого світла (WLI) залишається складним завданням. Часто ендоскопічні зображення страждають від нерівномірного освітлення через вигини в просвіті та напрямок світла ендоскопа. Науковці пропонують технологію покращення зображень, що вирішує вищезазначені недоліки, під назвою покращення текстури та кольору (TXI). TXI розроблений для покращення трьох факторів зображення в WLI (текстури, яскравості та кольору) з метою чіткого визначення слабких відмінностей тканин. У цьому методі покращення на основі ретінексу використовується у ланцюгу обробки ендоскопічних зображень. Покращення на основі ретінексу поєднується з покращенням кольору для значного акцентування тонових відмінностей слизових поверхонь. TXI застосовується до ендоскопічних зображень тварин; його ефективність оцінюється порівняно з традиційними технологіями покращення ендоскопічних зображень, загальноживаними методиками обробки зображень та новітніми методиками для підвищення якості хірургічних

ендоскопічних зображень. Експериментальні результати показують, що ТХІ може вибірково підвищувати яскравість у темних ділянках ендоскопічного зображення та покращувати слабкі відмінності тканин, такі як незначні морфологічні або кольорові зміни, водночас запобігаючи надмірному підсиленню. Ці експериментальні результати демонструють потенціал запропонованого алгоритму ТХІ як майбутнього клінічного інструменту для виявлення уражень шлунково-кишкового тракту зі складними для розпізнавання відмінностями тканин.

Таким чином, аналіз наукових джерел з покращення якості зображень за допомогою штучного інтелекту підтверджує значний прогрес у цій галузі, демонструючи численні переваги застосування ШІ у різноманітних аспектах обробки зображень. Зокрема, технології на основі глибокого навчання показують високу ефективність у зменшенні шуму, підвищенні роздільної здатності та покращенні контрасту, що є критично важливими для медичних і наукових зображень.

Розглянуті дослідження підкреслюють, що моделі ШІ, такі як згорткові нейронні мережі та генеративно-змагальні мережі, здатні не лише автоматизувати процеси обробки зображень, але й досягати рівня точності, який перевищує традиційні методи. Це особливо актуально у медичній діагностиці, де висока якість зображень може суттєво вплинути на точність і своєчасність постановки діагнозу.

Окрім цього, впровадження ШІ дозволяє зменшити обчислювальні витрати та підвищити швидкість обробки, що є важливим для реального застосування у клінічних умовах. Аналіз показав, що використання ШІ для покращення якості зображень є перспективним напрямком, який відкриває нові можливості для медичної візуалізації, а також інших областей, де висока якість зображень є критично важливою.

1.3 Постановка задачі

Отже, зокрема базуючись на проведеному аналізі предметної галузі, у ході виконання завдання на кваліфікаційну роботу необхідно буде провести огляд можливих методів і засобів для покращення якості зображень за допомогою штучного

інтелекту. У цьому контексті важливо здійснити систематичний огляд наукової літератури та сучасних досліджень, присвячених різноманітним підходам до обробки зображень із застосуванням технологій машинного навчання та нейронних мереж. Буде вивчено та визначено найбільш відповідні випадки застосування (use cases) для алгоритмів, перелічених у цій роботі. Водночас, слід ретельно оцінити переваги та недоліки кожного з підходів, беручи до уваги їхню ефективність, обчислювальну складність та можливість масштабування. Крім того, доцільно розглянути практичні аспекти реалізації цих методів у конкретних умовах, а також їхню інтеграцію у вже існуючі системи обробки зображень. Результати проведеного дослідження сприятимуть глибшому розумінню потенціалу штучного інтелекту у галузі покращення якості зображень та нададуть обґрунтовані рекомендації щодо вибору оптимальних алгоритмів для різних сценаріїв використання.

Окрім наведення теоретичних та оглядових відомостей про методи та моделі штучного інтелекту, необхідно виконати експериментальне випробування обраних методів. Такий підхід забезпечує всебічне розуміння ефективності та практичної застосовності запропонованих алгоритмів. Практичне застосування моделей штучного інтелекту дозволить не лише закріпити опрацювання теоретичних матеріалів, але й оцінити реальні можливості та обмеження цих технологій.

Експериментальне випробування передбачає тестування моделей у реальних умовах, що включає обробку різноманітних наборів зображень та аналіз результатів на предмет якості відновлення та покращення. Особливу увагу слід приділити тестуванню в умовах обмежених обчислювальних ресурсів, характерних для персональних комп'ютерів. Це дозволить оцінити ефективність алгоритмів у середовищах, де доступ до потужних обчислювальних серверів є обмеженим, і тим самим розширити можливості їх практичного застосування.

Важливим аспектом експериментального дослідження є визначення показників якості оброблених зображень. Використання стандартних метрик, таких як

співвідношення пікового рівня сигналу до шуму (PSNR), дозволить об'єктивно оцінити покращення якості зображень. Також необхідно враховувати суб'єктивні оцінки, що надаються користувачами, для повнішого розуміння впливу застосованих методів на сприйняття якості зображень.

Таким чином, поєднання теоретичних досліджень із практичним випробуванням методів штучного інтелекту є ключовим підходом для оцінки їх ефективності та можливостей. Це не лише сприяє глибшому розумінню технологій, але й демонструє їх реальне застосування, зокрема в умовах обмежених ресурсів. Такий комплексний підхід дозволяє сформувати більш повну картину можливостей сучасних технологій обробки зображень та їх потенціалу для вирішення практичних завдань у різних галузях.

2 ОГЛЯД МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ

2.1 Згорткові нейронні мережі

Згорткові або конволюційні нейронні мережі (convolutional neural networks, CNN) – це один з класів нейронних мереж, що найбільше застосовується для обробки саме зображень. Це нейронна мережа прямого поширення, тобто мережа, в якій сигнали поширюються лише в одному напрямку, від вхідного шару через приховані шари і аж до вихідного, де отримується результат опрацювання вхідних даних. Зворотні зв'язки у таких мережах відсутні.

Принцип згорткової нейронної мережі (ЗНМ) наслідує зорові механізми живих організмів. Історія ЗНМ бере свій початок у 1960-х роках, коли дослідники вперше почали використовувати згортки для обробки зображень. Одним із піонерів у цій галузі був Алоїс Хохштадтер, який у 1962 році запропонував використовувати згортки для виявлення кордонів на зображеннях. Також роботи Кухініко Фукусіма започаткували використання згорткових операцій для моделювання рецептивних полів нейронів у зоровій корі головного мозку (розроблена ним система мала назву Neocognitron, призначалася для розпізнавання японських рукописних символів і стала одним із базових алгоритмів для згорткових нейронних мереж). У 90-х роках ХХ століття ЗНМ стали все більш популярними, оскільки технологія обчислювальних систем покращилася, що дозволило їм обробляти більші набори даних. У 2010-х роках ЗНМ досягли значного прогресу, що призвело до революції в галузі комп'ютерної візуалізації. У цей період були розроблені кілька нових архітектур ЗНМ, включаючи ResNet, Inception і YOLO. Загалом, згорткові нейронні мережі вважаються однією з найпотужніших та найуніверсальніших архітектур машинного навчання.

Згорткові нейронні мережі складаються з декількох видів шарів (layers), які можуть бути застосовані попарно і почергово декілька разів [12].

Перший шар – це згортковий шар (convolution layer), який є основним типом шарів у згорткових нейронних мережах (ЗНМ). У цьому шарі використовуються невеликі фільтри, які зазвичай мають розмір, що є меншим за роздільну здатність вхідного зображення. Ці фільтри, також відомі як ядра, виконують згортку по ширині та висоті вхідних даних протягом прямого проходу.

Під час згортки кожен фільтр здійснює обчислення скалярного добутку між даними фільтру та відповідними ділянками вхідного зображення. Це обчислення проводиться для кожного можливого положення фільтру на зображенні, в результаті чого формується двовимірна карта активацій, або карта збудження, для кожного фільтру. Кожен елемент цієї карти активацій відповідає результату згортки фільтру з відповідною ділянкою вхідного зображення.

Фільтри в згортковому шарі здатні виявляти локальні особливості зображення, такі як контури, кути та текстури, на основі значень ваги всередині ядра. Кожен фільтр спеціалізується на виявленні певних характеристик зображення, і по мірі проходження через декілька згорткових шарів мережа може навчитися розпізнавати все більш складні ієрархічні особливості.

Позначимо вхідний сигнал як I , ядро як K . Виміри вихідних елементів наступні:

$$I = m_1 \times m_2 \times m_c,$$

$$K = n_1 \times n_2 \times n_c,$$

де m_1, n_1 – це висота зображення або ядра,

m_2, n_2 – ширина зображення або ядра,

m_c, n_c – кількість каналів зображення або ядра (кількості каналів мають дорівнювати одне одному).

Відтак, виміри результуючої карти збудження фільтру наступні:

$$F = (m_1 - n_1 + 1) \times (m_2 - n_2 + 1) \times f_c$$

де f_c – це кількість каналів, що дорівнює m_c , n_c .

Операцію згортки (вирахування ij -ого елемент матриці карти збудження) можна подати наступним чином:

$$f[i, j] = \sum_x^{m_1} \sum_y^{m_2} \sum_z^{m_c} K_{[x,y,z]} I_{[i+x-1, j+y-1, z]}$$

де x, y, z – це індекси для ітерації за висотою, шириною і каналами зображення відповідно.

Другий шар – агрегувальний шар (pooling layer) – відіграє важливу роль у структурі згорткових нейронних мереж. Ці шари призначені для зменшення розмірів карти ознак, створених згортковими шарами, шляхом зменшення просторових розмірів зображення, при цьому зберігаючи ключову інформацію. Це зменшення розмірів досягається завдяки операціям агрегування, що допомагають знизити кількість параметрів і обчислювальні витрати, а також запобігають перенавчанню моделі.

Між двома шарами згорткової нейронної мережі використовується нелінійна функція активації, яка додає складність і підвищує виразність отриманих особливостей. Нелінійні функції активації, такі як ReLU (Rectified Linear Unit), Sigmoid або Tanh, допомагають моделі навчитися складним шаблонам, додаючи нелінійність до обчислень і тим самим покращуючи здатність моделі до класифікації та розпізнавання складних структур у даних.

Існує декілька методів реалізації агрегування, серед яких найпоширенішим є максимізаційне агрегування (max pooling). Цей метод поділяє вхідне зображення на набір прямокутних областей без перекриття, і для кожної такої підобласті обчислює максимальне значення. Основна ідея полягає в тому, що точне положення ознаки не настільки важливе, як її наявність і розташування відносно інших ознак. Таким чином, max pooling допомагає зменшити просторові розміри карти ознак, зберігаючи при цьому суттєву інформацію про наявність важливих особливостей.

Використання агрегувальних шарів також допомагає контролювати обчислювальні витрати, оскільки менша карта ознак вимагає менше обчислювальних ресурсів для подальшої обробки. Крім того, цей підхід сприяє загальній стабільності та ефективності моделі, оскільки зменшення розмірів карти ознак запобігає перенаванчання, забезпечуючи узагальнення моделі на нові, невідомі дані.

Формула максимізаційного агрегування наступна:

$$P[i, j] = \max_{k \in \{0, \dots, m_1 - 1\}, l \in \{0, \dots, m_2 - 1\}} I_{[i*s+k, j*s+l, :]},$$

де k, l – індекси для ітерації за висотою і шириною відповідно,

s – крок виконання операції,

$:$ - означає, що виконання відбувається за всіма значеннями каналів вхідної карти збудження.

Також у якості функції активації може застосовуватися ReLU. Ця функція відкидає від'ємні значення та замінює їх на нуль:

$$R(x) = \max(0, x).$$

Насамкінець, після однієї або кількох пар згорткових та агрегувальних шарів, використовуються заключні, повноз'єднані шари (fully-connected layers). Ці кінцеві

шари об'єднують витягнуті особливості з попередніх шарів і роблять кінцеві прогнози або класифікацію. Нейрони у повноз'єднаному шарі мають з'єднання з усіма збудженнями попереднього шару, як це можна бачити у звичайних нейронних мережах. Принципова схема роботи згорткових нейронних мереж подана нижче, на рисунку 2.1 [13].

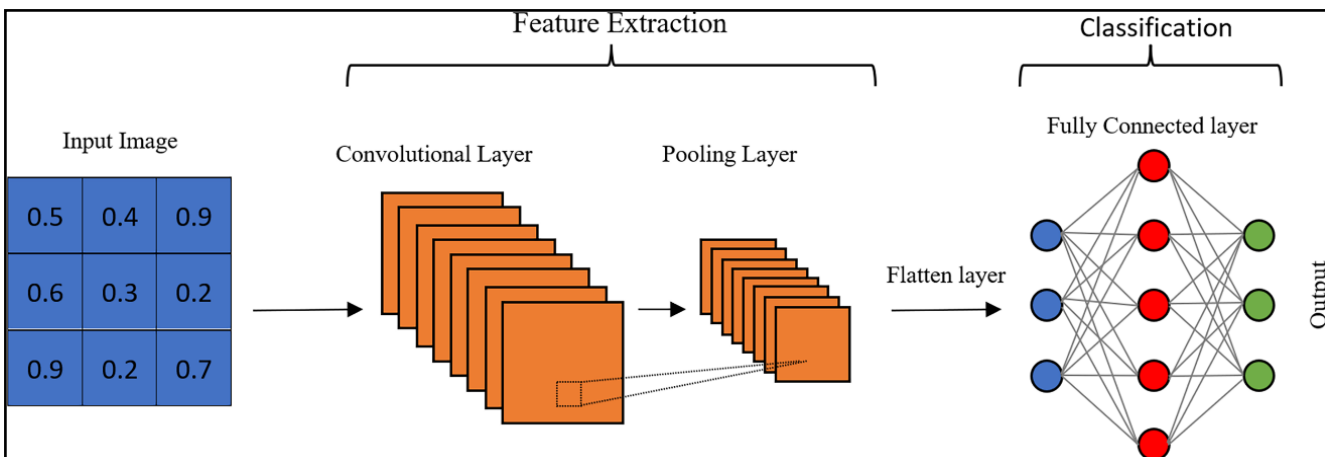


Рисунок 2.1 – Схема роботи згорткових нейронних мереж

Згорткові нейронні мережі використовуються для широкого спектру використань, пов'язаних із зображеннями і комп'ютерним зором. До застосувань належать розпізнавання об'єктів і сцен на зображеннях, наприклад, класифікація фотографій тварин або розпізнавання різних типів ландшафтів; або ж виявлення та обведення конкретних об'єктів на зображенні, наприклад, виявлення обличчя на фотографіях або ідентифікація автомобілів на дорозі. Також у згорткових нейронних мереж є інші застосування, які не стосуються обробки зображень: обробка природної мови (natural language processing), пошук нових ліків (використовувались для передбачення взаємодії між хімічними молекулами ліків та біологічними білками), до того ж, ЗМН навчали грати в різні настільні ігри (го та шашки).

Щодо задачі, що розглядається у цій роботі, то ЗНМ можна використовувати за допомогою різних методів, як-от:

- реконструкція зображення з низькою роздільною здатністю до вищої: моделі ЗНМ навчаються на зображеннях з високою роздільною здатністю, потім ці моделі використовуються для відновлення низькороздільних зображень;
- зменшення шуму на зображеннях (наприклад, застосовується для фотографій, виконаних в умовах слабкого освітлення);
- корекція кольору зображень.

2.2 Генеративні змагальні мережі

Генеративні змагальні мережі (generative adversarial networks, GAN) – це клас алгоритмів штучного інтелекту, що використовує навчання без учителя. ГЗМ були розроблені Яном Гудфеллоу і групою його колег-дослідників у 2014 році в рамках дослідницької роботи в Університеті Монреалю, Канада, про що було опубліковано статтю на міжнародній конференції з систем нейронної обробки інформації (The International Conference on Neural Information Processing Systems, NIPS 2014).

Особливість цього підходу в тому, що алгоритми реалізуються системою двох штучних нейронних мереж, які «змагаються» одна з одною в рамках гри з нульовою сумою: так, перемога однієї моделі стає поразкою і втратою іншої моделі. Перша модель називається генератором, а друга – дискримінатором. Генератор намагається створювати фальшиві зображення, які виглядають як справжні, тоді як дискримінатор намагається відрізнити справжні від згенерованих.

На початковому етапі розвитку GAN, основною проблемою було забезпечення стабільності навчання, оскільки обидві мережі повинні навчатися збалансовано. Ці труднощі були частково вирішені шляхом вдосконалення архітектур мереж та алгоритмів навчання. Зокрема, були запропоновані такі варіанти GAN, як Deep Convolutional GAN (DCGAN), який використовує глибокі згорткові нейронні мережі,

та Wasserstein GAN (WGAN), який вводить нові способи обчислення функції втрат для покращення стабільності.

Головна ідея ГЗН базується на «непряму» навчанні через дискримінатор, яка намагається визначити, наскільки «реалістичним» є вхідне зображення, яке також динамічно оновлюється. Це означає, що генератор навчається не мінімізувати відстань до конкретного зображення, а радше змусити дискримінатор ухвалювати неправильні рішення. Це дозволяє моделі навчатися без учителя.

Генератор і дискримінатор навчаються один від одного, покращуючи свою продуктивність і ефективність. Результатом є генератор, який може виробляти зображення високої якості та високої роздільної здатності з вхідних даних низької якості та низької роздільної здатності (схема алгоритму подана на рисунку 3.2 [14]). Відповідно, для досягнення цілі дослідження, покращення якості зображень, у цій системі є важливим саме генератор, а дискримінатор має допоміжну роль у навчанні генератора.

Генератор, як правило, є деконволюційною нейронною мережею, а дискримінатор – навпаки ж, згортковою нейронною мережею, яка розглядалася вище.

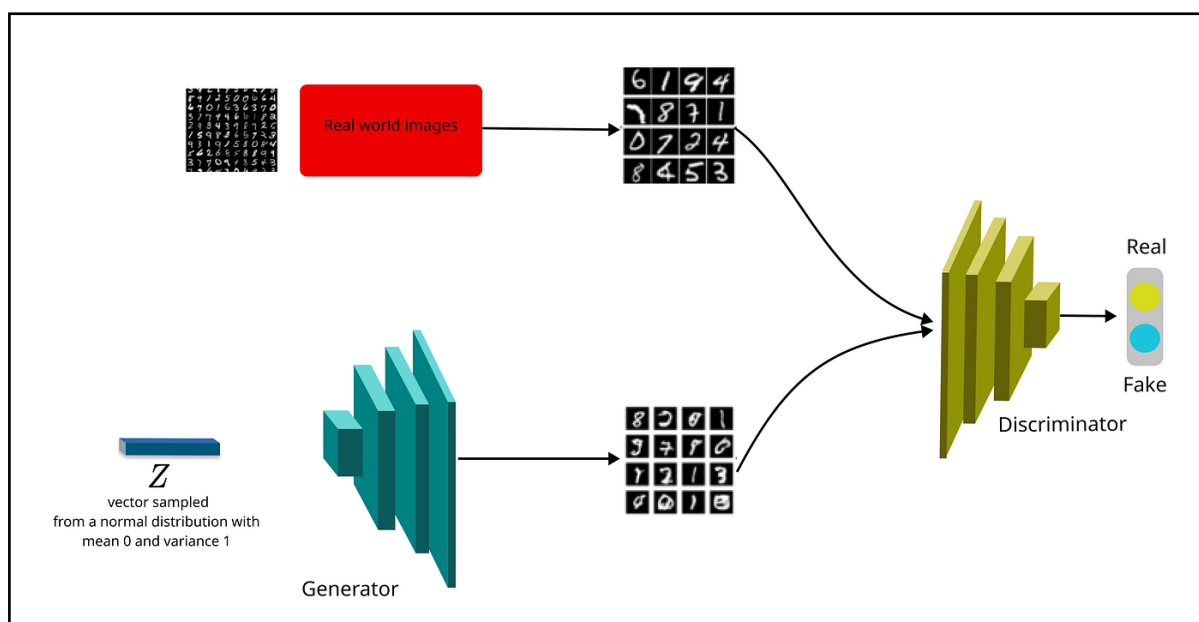


Рисунок 2.2 – Графічний опис принципу роботи генеративних змагальних мереж

У загальному вигляді алгоритм можна описати математично наступним чином, за допомогою правила мінімакс [15]:

$$\min_G \max_D V(D, G) = E_x[\log D(x)] + E_z[\log(1 - D(G(z)))].$$

У формулі x позначає справжні дані, z – згенерований шум, тобто несправжнє зображення. G – це модель генератора у алгоритмі, D – модель дискримінатора. $D(x)$ є ймовірністю дискримінатора коректно класифікувати справжнє зображення, а $D(G(x))$ – ймовірністю дискримінатора некоректно позначити шум, згенерований генератором, як реальне зображення. Наостанок, $E(x)$ – це очікування щодо розподілу реальних даних (зображення), а $E(z)$ – очікування щодо розподілу хибних даних.

Формулу можна інтерпретувати наступним чином: перший додатак описує намагання дискримінатора максимально покращити власну здатність коректної класифікації поданих даних; другий же додатак позначає намагання генератора генерувати складніші хибні зображення, аби дискримінатор сприйняв їх неправильно. Разом це мінімакс правило стимулює обидві моделі із часом покращувати свої спроможності. Таким чином, генеративні змагальні мережі є подібними до еволюційної біології, з еволюційними «перегонами озброєнь» між двома моделями.

Генеративні змагальні мережі (ГЗМ) мають кілька переваг для завдань покращення зображення, таких як надвисока роздільна здатність, усунення розмиття, усунення шумів, домальовування та розфарбовування. По-перше, ГЗМ здатні генерувати реалістичні та природні зображення без внесення артефактів чи спотворень.

По-друге, ГЗМ можуть навчатися на нерозмічених даних, зменшуючи потребу у дорогих і працемістких операціях із визначення класів зображень. Це особливо важливо в умовах, коли отримання великої кількості розмічених даних є складним або неможливим. Використовуючи методи самонавчання, ГЗМ можуть виявляти

приховані закономірності у даних і використовувати їх для генерації високоякісних зображень.

По-третє, ГЗМ можуть працювати з різноманітними та складними сценаріями, такими як різні умови освітлення, пози, фони та стилі. Це забезпечує гнучкість у застосуванні ГЗМ до широкого спектра задач.

Окрім цього, ГЗМ демонструють високу ефективність у таких завданнях, як автоматичне домальовування відсутніх частин зображення або розфарбовування чорно-білих фотографій.

Висока адаптивність ГЗМ до різноманітних задач покращення зображення робить їх потужним інструментом, який здатен значно покращити якість візуальної інформації. Це, у свою чергу, сприяє розвитку нових технологій і методів візуалізації, що мають велике значення для науки, техніки та мистецтва. Таким чином, генеративні змагальні мережі продовжують відігравати ключову роль у сучасних дослідженнях і практичних застосуваннях, забезпечуючи інноваційні підходи до обробки та покращення зображень.

Для різноманітних сценаріїв використання генеративних змагальних мереж (ГЗМ) існують спеціалізовані алгоритми, які адаптовані до конкретних завдань, забезпечуючи оптимальні результати. Наприклад, SRGAN (Super-Resolution Generative Adversarial Network) – це ГЗМ, яка може збільшити роздільну здатність зображень до 4 разів, зберігаючи при цьому деталі та текстури. Ця модель є надзвичайно корисною для покращення якості зображень у різних галузях, таких як медична діагностика або спостереження.

Іншим прикладом є DeblurGAN, яка може усунути розмитість із зображень, що є важливим для фотографій, знятих рухомою камерою або в умовах слабкого освітлення. Цей підхід знаходить застосування в галузях, де чіткість зображення має вирішальне значення, наприклад, у судово-медичній експертизі або відеоспостереженні.

CycleGAN – це модифікація ГЗМ, яка дозволяє передавати стиль і колір однієї обстановки зображення в іншу. Це включає перетворення фотографій на картини або зміну пори року на фотографії, що є корисним для художньої творчості та віртуальної реальності. Ця модель забезпечує високу якість стилізації, зберігаючи при цьому природність і реалістичність зображень.

Pix2Pix – це ще одна ГЗМ, яка може створювати реалістичні зображення з ескізів або мап. Наприклад, ця модель здатна перетворювати фасади на об'ємні будівлі або контури на повноцінні обличчя, що має значний потенціал у сфері архітектури та дизайну. Така технологія полегшує процес проектування та візуалізації, дозволяючи швидко отримувати високоякісні результати.

Ці приклади не є вичерпними, адже окрім перелічених, існує понад 30 підходів до реалізації генеративних змагальних мереж, кожен з яких має свої унікальні особливості та області застосування. Кожен з цих підходів демонструє значні досягнення у відповідних завданнях, забезпечуючи гнучкість та ефективність використання ГЗМ у різних наукових та практичних контекстах. Таким чином, генеративні змагальні мережі продовжують розширювати свої можливості, сприяючи інноваціям і прогресу в багатьох галузях.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ МЕТОДІВ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ

3.1 Підготовка даних та тренування мережі

У цьому розділі буде розглянуто код створення моделі згорткової нейронної мережі та її застосування на зображеннях, для яких штучно було знижено роздільну здатність (РЗ). Таким чином, експериментально буде випробувано один із методів штучного інтелекту для покращення якості зображень. Крім того, у розділі буде описано процес підготовки та попередньої обробки даних, включаючи методи зниження роздільної здатності вихідних зображень та техніки доповнення даних для підвищення стійкості моделі до різних варіацій вхідних зображень. Таким чином, експериментально буде випробувано один із методів штучного інтелекту для покращення якості зображень, що дозволить оцінити ефективність застосування ЗМН для цієї задачі.

Виконання експерименту умовно можна поділити на три етапи:

- етап попередньої обробки тренувальних даних – на цьому етапі буде використано набір даних, що складається із зображень високої роздільної здатності. Для цього етапу використовуватимемо набір даних ImageNet [16], який загалом складається з 1500000 зображень, але у цьому експерименті використаємо лише 50000 зображень;

- етап формування і тренування моделі;

- етап валідації результатів – декілька зразків роботи коду буде виведено для користувача, і перевірити результати роботи можна буде візуальним порівнянням зображення зі штучно зниженою роздільною здатністю, зображення після покращення РЗ за допомогою нейронної мережі і оригінальної версії.

Отже, у цьому експерименті для тренування моделі буде використано набір звичайних фотографій із повноцінною, високою роздільною здатністю. Після нормалізації даних (видалення зображень із занадто високою та занадто низькою

роздільними здатностями, а також чорно-білих зображень), всі зображення приводяться до однакових роздільної здатності і розмірів (рис. 3.1). Код нормалізації буде подано нижче.

Нормалізація і фільтрація зображень із надто високою РЗ відбувається, до прикладу, з тої причини, що навіть після зниження якості зображення роздільна здатність залишається надто високою, через що використання такого зображення в експерименті є неефективною.



Рисунок 3.1 – Приклад уніфікованого зображення з високою РЗ

```
def Remove1DImages (path) :  
    os.chdir (path)  
  
    for _ in os.listdir() :  
        with PIL.Image.open(_) as image:  
            array = img_to_array(image)  
            channels = array.shape[-1]
```

```

        if channels != 3:
            os.remove(_)

def RemoveHDImages(path):
    os.chdir(path)

    i = os.listdir()

    for _ in i:
        with PIL.Image.open(_) as image:
            width = image.size[0]
            height = image.size[1]

            if width >= 800 or height >= 800:
                os.remove(_)

def RemoveLowres(path):
    os.chdir(path)

    i = os.listdir()

    for _ in i:
        with PIL.Image.open(_) as image:
            width = image.size[0]
            height = image.size[1]

            if width < 500 or height < 300:
                os.remove(_)

```

Варто звернути увагу на константи, що використовуватимуться протягом експерименту. У обробці буде залучено лише один канал яскравості для підвищення точності, тож ця змінна дорівнює 1. Коефіцієнт зміни (підвищення/зниження) роздільної здатності для зображення зі зниженою РЗ дорівнюватиме 3. Розмір же нормалізованого зображення буде 498x300 пікселів (очевидно, напрошується значення 500 по висоті, однак воно не ділиться націло на коефіцієнт підвищення РЗ).

```

channels = 1
updated_width = 498
updated_height = 300
enhancement_coeff = 3

input_width = updated_width // enhancement_coeff
input_height = updated_height // enhancement_coeff

```

Далі відбувається завантаження зображень із директорії та їхній запис у змінні. Формується дві підмножини даних: навчальний та тестовий набори. Розділення масиву даних відбувається автоматично у пропорції 80/20, тому вивантаження можна здійснити із загальної директорії.

```
training_data = image_dataset_from_directory(
    path,
    image_size=(updated_width, updated_height),
    validation_split=0.2,
    subset='training',
    seed=240,
    batch_size = 32,
    label_mode=None)

test_data = image_dataset_from_directory(
    path,
    image_size=(updated_width, updated_height),
    validation_split=0.2,
    subset='validation',
    seed=240,
    batch_size = 32,
    label_mode=None)
```

Після цього формуються набори вхідних даних (feature) і цільових вихідних даних (target) для моделі. У тренуванні моделі буде використовуватися лише канал яскравості (luminance channel), кольорові канали буде тимчасово видалено. Щоб досягти цього (а зображення перебувають у форматі RGB), зображення перетворюються у формат YUV, що складається з каналу яскравості Y і двох каналів, що визначають колір, U та V. Відповідно, надалі два останніх канали нового формату тимчасово ігноруються і до обробки не залучаються.

Для вхідних даних знижується роздільна здатність (а отже ми отримуємо зображення із низькою роздільною здатністю), а для цільових зображень РЗ не змінюється, проте залишається також один канал яскравості (тож це стане оригіналом). Для обох варіантів виводиться візуальний результат для підтвердження коректності. Ці приклади результатів обробки подано на рисунках 3.2 та 3.3, а код операції подано нижче.



Рисунок 3.2 – Приклад зображення для вхідних даних



Рисунок 3.3 – Приклад цільових даних

```
def prepare_features(input, new_width, new_height):  
    input = tf.image.rgb_to_yuv(input)  
    last_axis = len(input.shape) - 1  
    y, u, v = tf.split(input, 3, axis=last_axis)  
    return tf.image.resize(y, [new_width, new_height], method="area")  
  
def prepare_target(input):  
    input = tf.image.rgb_to_yuv(input)  
    last_axis = len(input.shape) - 1  
    y, u, v = tf.split(input, 3, axis=last_axis)  
    return y
```

```

training_data = training_data.map(lambda x: (prepare_features(x, input_width,
input_height),prepare_target(x)))

test_data = test_data.map(lambda x: (prepare_features(x, input_width,
input_height),prepare_target(x)))

for batch in training_data.take(1):
    for img in batch[0]:
        display(array_to_img(img))
    for img in batch[1]:
        display(array_to_img(img))

```

На цьому етап підготовки даних завершено, і можна перейти до наступного – етапу формування і тренування моделі.

Перший крок – необхідно сконструювати функцію залишкового щільного блоку (residual dense block, RDB). Мета використання цих блоків у моделі полягає в видобуття ієрархічних ознак із вихідних зображень з низькою роздільною здатністю. Залишкові щільні блоки видобувають складні та високорівневі особливості за допомогою видобутку якомога більше ієрархічних ознак. Це досягається шляхом обміну виходом кожного шару всередині як зворотного зв'язку.

```

def rdb_block(inputs, numLayers):
    channels = inputs.get_shape()[-1]
    storedOutputs = [inputs]

    for _ in range(numLayers):
        localConcat = tf.concat(storedOutputs, axis=-1)
        out = Conv2D(filters=channels, kernel_size=3, padding="same",
            activation="relu")(localConcat)
        storedOutputs.append(out)

    finalConcat = tf.concat(storedOutputs, axis=-1)
    finalOut = Conv2D(filters=channels, kernel_size=1,
        padding="same", activation="relu")(finalConcat)

    finalOut = Add()([finalOut, inputs])

    return finalOut

```

Після цього необхідно зібрати, власне, фінальну модель. Згортова нейронна мережа в основному складається з згорткових шарів. У моделях класифікації також

використовуються ще й агрегувальні шари у поєднанні шарами вирівнювання. Однак, оскільки ця модель є моделлю регресії, використовуються лише згорткові шари. Гіперпараметри для цих шарів та кількість використовуваних шарів протягом розробки зазнали різних змін і налаштувань для досягнення найкращого можливого результату.

Можна помітити, що вхідний шар нашої моделі має форму (None, None, channels). Причина цього полягає в тому, що ми не знаємо форми зображень, які будуть використовуватися для перевірки даних і після тренування моделі, оскільки різні зображення, швидше за все, матимуть різні, відмінні роздільні здатності.

Також останній шар моделі використовує відмінну кількість фільтрів, ніж попередні шари. Оскільки вхідне зображення з низькою роздільною здатністю необхідно буде масштабувати, щоб воно відповідало вихідній, високій РЗ, використовується метод Pixel Shuffling. Роздільна здатність помножується на той самий коефіцієнт масштабування, який використовувався для пониження роздільної здатності з оригінального зображення. Таким чином, вихід моделі і буде зображенням з покращеним роздільною здатністю, тобто очікуваним результатом роботи моделі.

Також для тренування моделі застосовується рання зупинка (early stopping). Призначення цієї функції – запобігання перенавчанню моделі, а відповідно і збереження обчислювальних ресурсів, які не витрачатимуться без особливої користі для експерименту. Умовою ранньої зупинки є десять підряд епох навчання із незначними результатами покращення. Код моделі подано нижче.

```
def Model(channels, enhancement_coeff):
    inputs = keras.Input(shape=(None, None, channels))

    X = Conv2D(64, 5, padding='same', activation='relu',
kernel_initializer='Orthogonal')(inputs)

    X = Conv2D(64, 3, padding='same', activation='relu',
kernel_initializer='Orthogonal')(X)
    X = rdb_block(X, numLayers=3)
```

```

X = Conv2D(32, 3, padding='same', activation='relu',
kernel_initializer='Orthogonal')(X)

X = rdb_block(X, numLayers=3)

X = Conv2D(channels * (enhancement_coeff**2), 3, padding='same',
activation='relu', kernel_initializer='Orthogonal')(X)

outputs = tf.nn.depth_to_space(X, enhancement_coeff)

return keras.Model(inputs, outputs)

early_stopping = EarlyStopping(monitor='loss', patience=10, min_delta=0.0001)
model = Model(channels, enhancement_coeff)
model.compile(optimizer='adam', loss='MSE')

model.fit(training_data, epochs=40, callbacks=[early_stopping],
validation_data = test_data)
model.save('model.h5')

```

Останні рядки ділянки коду, поданої вище, запускають процес навчання і зберігають створену модель у окремий файл, який можна використовувати пізніше. Відповідно, на цьому процес підготовки даних і середовища до покращення якості завершено, і можна перейти безпосередньо до випробування моделі і порівняння результатів експерименту.

3.2 Перевірка результатів

Як було зазначено раніше, цей етап складатиметься із двох компонентів: візуального порівняння роботи моделі користувачем, що наочно дозволить якісно порівняти результати обробки, та обчислення метрики, що виявить кількісну різницю між роздільною здатністю оригіналу, зображення зі зниженою роздільною здатністю та результату обробки попарно.

Для візуального порівняння застосуємо окрему функцію, що дозволить виводити усі зображення у однаковому форматі. До того ж, ця функція виділятиме та збільшуватиме фрагменти зображень для демонстрації роботи, легшого розрізнення і порівняння результатів.

```

def plot_results(img, prefix, title, z1=200, z2=300, z3=150, z4=250):
    img_array = img_to_array(img)
    img_array = img_array.astype("float32") / 255.0

    figure, parent = plt.subplots()

    parent.imshow(img_array[:, :-1], origin="lower")

    plt.title(title)

    inset = zoomed_inset_axes(parent, 2, loc='upper left')

    inset.imshow(img_array[:, :-1], origin="lower")

    x1, x2, y1, y2 = z1, z2, z3, z4

    inset.set_xlim(x1, x2)
    inset.set_ylim(y1, y2)

    plt.yticks(visible=False)
    plt.xticks(visible=False)

    mark_inset(parent, inset, loc1=1, loc2=3, fc="none", ec="blue")

    plt.show()

```

Тепер можна виконати фінальну частину експерименту, а саме покращення якості зображень за допомогою створеної моделі згорткової нейронної мережі. Це відбувається, перебираючи колекцію зображень із високою роздільною здатністю, за наступними кроками:

- зберегти оригінал у виділену змінну для подальшого порівняння;
- штучно зменшити роздільну здатність оригіналу, таким чином умовно утворюючи зображення із низькою РЗ;
- перетворити зображення із низькою роздільною здатністю, залишивши лише канал яскравості (повторюючи механізм тренування моделі);
- запустити «передбачення», тобто покращення якості зображення за допомогою моделі;
- нормалізувати значення яскравості у діапазоні $[0, 255]$, забираючи значення поза цим діапазоном (аби зберегти коректність кольорів), поєднати назад канали

яскравості та кольору і зрештою трансформувати результуюче зображення у початковий RGB-формат;

- вивести оригінал, зображення зі зменшеною роздільною здатністю і результат за допомогою функції, що її було описано вище.

Використаний код у цілому наведено нижче.

```
for index, _ in enumerate(visualize_results):
    original_image = PIL.Image.open(_)

    descaled_image = original_image.resize((original_image.size[0] //
enhancement_coeff,
                                           original_image.size[1] // enhancement_coeff),
                                           PIL.Image.BICUBIC)

    ycbcr = descaled_image.convert("YCbCr")
    y, cb, cr = ycbcr.split()
    y = img_to_array(y)
    y = y.astype("float32") / 255.0

    input = y.reshape(1, y.shape[0], y.shape[1], y.shape[2])

    output = model.predict(input)
    output = output[0]
    output *= 255.0
    output = output.clip(0, 255)
    output = output.reshape((output.shape[0], output.shape[1])) # (Width,
Height)
    output = PIL.Image.fromarray(np.uint8(output))
    output = output.resize(original_image.size, PIL.Image.Resampling.NEAREST)

    cb = cb.resize(output.size, PIL.Image.Resampling.BICUBIC)
    cr = cr.resize(output.size, PIL.Image.Resampling.BICUBIC)

    enhanced_image = PIL.Image.merge("YCbCr", (output, cb, cr))
    enhanced_image = enhanced_image.convert("RGB")

    descaled_image = descaled_image.resize(enhanced_image.size,
PIL.Image.Resampling.BICUBIC)

    plot_results(descaled_image, index, 'Low Resolution')
    plot_results(enhanced_image, index, 'Enhanced Resolution')
    plot_results(original_image, index, 'Original Image')
```

Скомбіновані підсумки роботи моделі на деяких прикладах зображень подано на рисунках 3.4-3.6.

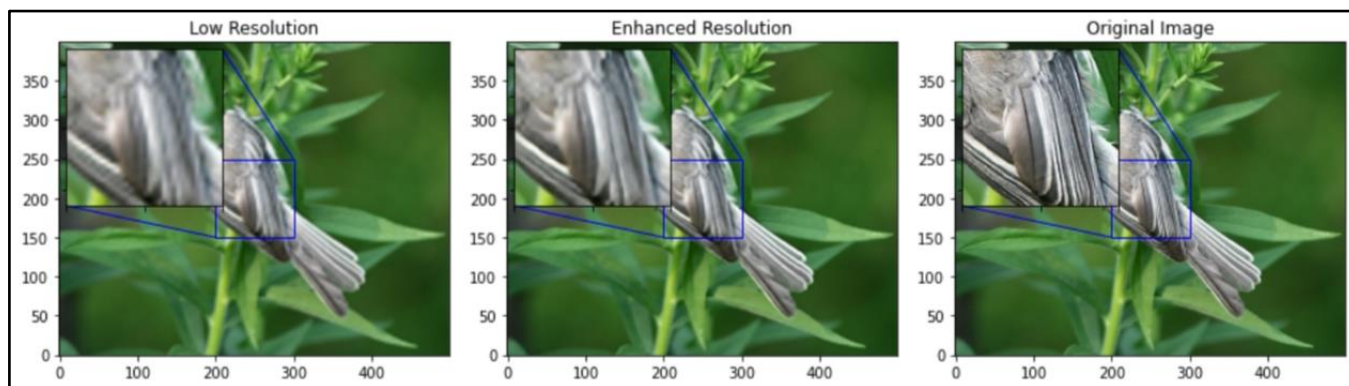


Рисунок 3.4 – Приклад підвищення якості зображення на зображенні птаха

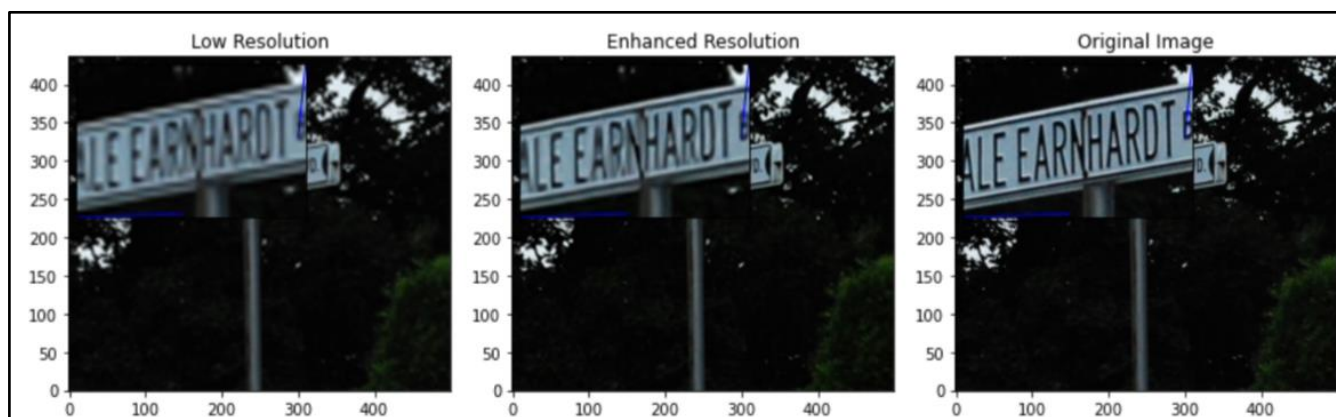


Рисунок 3.5 – Приклад підвищення якості зображення на зображенні вуличного вказівника

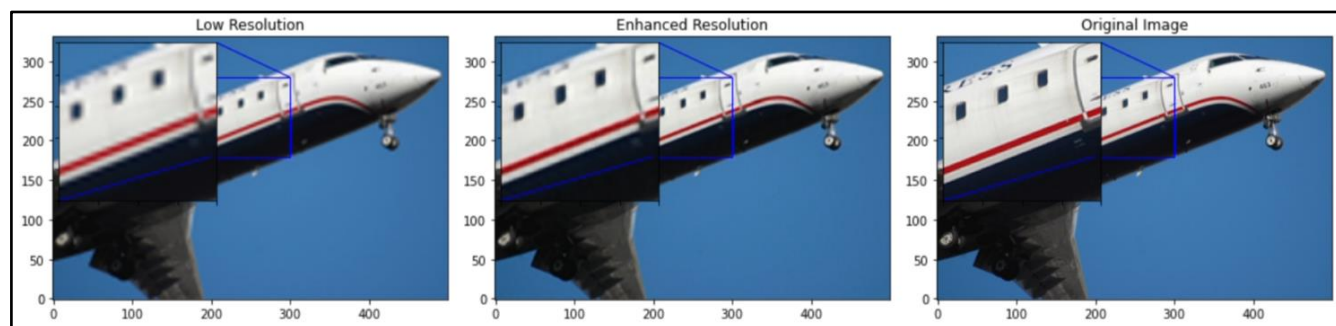


Рисунок 3.6 – Приклад підвищення якості зображення на зображенні літака

Друга частина валідації результатів обробки – кількісне обчислення результатів за допомогою визначеної метрики. Обрана метрика для оцінки якості зображень – це

співвідношення пікового рівня сигналу до шуму (peak signal-to-noise ratio, PSNR). PSNR є мірою якості реконструйованих або стиснутих зображень у порівнянні з оригінальними зображеннями і широко використовується для оцінки ефективності різних алгоритмів обробки зображень, таких як зменшення шуму, стиснення без втрат або з втратами, а також покращення якості зображень. Це робить PSNR особливо підходящою метрикою для оцінки результатів даного дослідження.

Обчислення PSNR базується на середньому значенні квадратів різниць (mean squared error, MSE) між відповідними пікселями оригінального та обробленого зображень. Формула для обчислення PSNR включає логарифмічне перетворення MSE, що дозволяє виразити співвідношення у децибелах (dB). Високі значення PSNR вказують на високу якість реконструйованого зображення, оскільки вони свідчать про низький рівень спотворень або шуму порівняно з оригіналом.

Метрика PSNR є корисною не тільки через свою здатність кількісно оцінювати відмінності між оригінальним і обробленим зображенням, але й через свою простоту та загальноприйнятність у науковій спільноті. Використання PSNR дозволяє дослідникам легко порівнювати результати різних алгоритмів і технологій обробки зображень, забезпечуючи об'єктивну основу для таких порівнянь.

Для обчислення метрики використовуємо методи, що надаються бібліотекою Tensorflow. Обчислення відбуватиметься попарно: PSNR зображення з низькою роздільною здатністю та оригіналу, PSNR обробленого зображення та оригіналу. Приклад результату обчислення подано на рисунку 3.7. Код операції також подано нижче.

```
descaled_image_arr = img_to_array(descaled_image)
original_image_arr = img_to_array(original_image)
enhanced_image_arr = img_to_array(enhanced_image)

bicubic_psnr = tf.image.psnr(descaled_image_arr, original_image_arr,
                             max_val=255)
test_psnr = tf.image.psnr(enhanced_image_arr, original_image_arr,
                           max_val=255)
```

```
print("PSNR of LR image and HR image is: %f" % bicubic_psnr)  
print("PSNR of ER image and HR image is: %f" % test_psnr)
```

```
PSNR of LR image and HR image is: 25.208151  
PSNR of ER image and HR image is: 25.868135
```

Рисунок 3.7 – Приклад виведення результату обчислення метрики

ВИСНОВКИ

У ході роботи було проведено дослідження методів і засобів штучного інтелекту для покращення якості зображень. Роботу було розпочато з аналізу предметної галузі, у процесі якого було наведено важливість теми дослідження та охарактеризовано різні поля застосування методів ШІ для покращення якості зображень. Зокрема, підкреслено актуальність використання штучного інтелекту в таких галузях, як медицина, розваги, безпека та збереження культурної спадщини.

Далі було детально проаналізовано два важливих методи штучного інтелекту – згорткові нейронні мережі (ЗНМ) та генеративні змагальні мережі (ГЗМ), які в свою чергу використовують ЗНМ як один із компонентів системи алгоритмів. Описано принципи роботи цих методів, зокрема їх здатність навчатися на великих наборах даних та виявляти складні патерни, що недоступні традиційним алгоритмам. Також було наведено потенціал використання обох засобів, зокрема у завданнях сегментації, класифікації та генерації зображень.

Розглянуто модифікації алгоритмів, що існують, та наведено найбільш відповідні випадки використання для них. Наприклад, згорткові нейронні мережі знайшли широке застосування у завданнях класифікації об'єктів на зображеннях, виявлення об'єктів та їх сегментації, тоді як генеративні змагальні мережі продемонстрували високу ефективність у завданнях генерації реалістичних зображень, підвищення роздільної здатності та усунення артефактів.

У ході роботи можна зробити висновок, що перелічені методи демонструють значну ефективність у виконанні різнопланових задач, зокрема і в обробці зображень, але застосування засобів штучного інтелекту не обмежується лише візуальними даними. Їх використання можливе і в аналізі текстових, звукових та інших типів даних, що відкриває широкі перспективи для розвитку міждисциплінарних досліджень.

Насамкінець, було на практиці імплементовано та використано згорткові нейронні мережі для одного з варіантів покращення якості зображень – підвищення

роздільної здатності. Використання методів штучного інтелекту дозволило значно покращити якість зображень за умови відносної «дешевизни» та ефективності використання обмежених ресурсів персонального комп'ютера. Результати експериментів підтвердили високу ефективність обраного методу, що дозволяє рекомендувати його для практичного застосування у різних галузях, де висока якість зображень є критичною.

Дослідження було апробовано на XXVI Міжнародній науково-практичній конференції «Modern scientific research: theoretical and practical aspects», що була проведена у Оттаві, Канада, 5-7 червня 2024 року. Результати участі у конференції наведено у додатку А.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Smelyakov, K., Savulioniene, L., Chupryna, A., Sakalys, P., Sandrkin, D. Adaptive Image Enhancement Model for the Robot Vision System. Vide. Tehnologija. Resursi - Environment, Technology, Resources. 2023. №3. С. 246-251.
2. Image Quality Analysis, Enhancement and Optimization Techniques for Computer Vision // Edge AI + Vision Alliance. URL: <https://www.edge-ai-vision.com/2017/02/image-quality-analysis-enhancement-and-optimization-techniques-for-computer-vision> (дата звернення: 14.01.2024).
3. Ukraine – Mobile Ownership // Media Landscapes. URL: <https://medialandscapes.org/country/ukraine/telecommunications/mobile-ownership> (дата звернення: 14.01.2024).
4. Kyrychenko, I., Nazarov, O., Huliiev, N., Avdieiev, O. Selection of Artificial Neural Networks for Disease Prediction. CEUR Workshop Proceedings. 2023. №3387. С. 236-248.
5. Computational photography part I: What is computational photography? // Digital Photography Review. URL: <https://www.dpreview.com/articles/9828658229/computational-photography-part-i-what-is-computational-photography> (дата звернення: 16.01.2024).
6. A new X-ray images enhancement method using a class of fractional differential equation // National Library of Medicine. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10310480/> (дата звернення: 15.01.2024).
7. How To Enhance CCTV Footage Easily And Fast // HitPaw. URL: <https://www.hitpaw.com/video-tips/cctv-video-enhancement.html> (дата звернення 18.01.2024).
8. Hongmin Liu, Fan Jin, Hui Zeng, Huayan Pu, Bin Fan. Image Enhancement Guided Object Detection in Visually Degraded Scenes. IEEE Transactions on Neural Networks and Learning Systems. 2023. С. 1-14.

9. T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. Bin Abul Kashem, M. Tariqul Islam, S. Al Maadeed, S. M. Zughailer, M. Salman Khan, M. E.H. Chowdhury. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Computers in Biology and Medicine*. 2021. №132. С. 1-16.
10. K. Munadi, K. Muchtar, N. Maulina, B. Pradhan. Image Enhancement for Tuberculosis Detection Using Deep Learning. *IEEE Access*. 2020. №8. С. 217897-217907.
11. T. Sato. TXI: Texture and Color Enhancement Imaging for Endoscopic Image Enhancement. *Journal of Healthcare Engineering*. 2021. №1.
12. Convolutional Neural Networks, Explained // Medium. URL: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939> (дата звернення: 18.01.2024).
13. Ahmad Alsaleh, Cahit Perkgoz. A space and time efficient convolutional neural network for age group estimation from facial images // PeerJ. URL: <https://peerj.com/articles/cs-1395/> (дата звернення: 01.04.2024).
14. Generative Adversarial Networks (GAN): Introduction and Example // Medium. URL: <https://chingisoinar.medium.com/generative-adversarial-networks-gan-introduction-and-example-3b66f5f235e9> (дата звернення: 10.04.2024).
15. Mini-Max Optimization Design of Generative Adversarial Networks (GAN) // Towards Data Science. URL: <https://towardsdatascience.com/mini-max-optimization-design-of-generative-adversarial-networks-gan-dc1b9ea44a02> (дата звернення: 09.04.2024).
16. ImageNet Object Localization Challenge // Kaggle. URL: <https://www.kaggle.com/competitions/imagenet-object-localization-challenge/data> (дата звернення: 30.04.2024).