

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Програмної інженерії

**КВАЛІФІКАЦІЙНА РОБОТА**

**Пояснювальна записка**

\_\_\_\_\_ другий (магістерський)

(рівень вищої освіти)

Дослідження методів штучного інтелекту для збільшення масштабу  
зображень

Виконав:

студент 2 курсу, групи ШЗМ-20-2

\_\_\_\_\_ Нечаєв М.Ю.

(прізвище, ініціали)

Спеціальність 121 – Інженерія програмного

\_\_\_\_\_ забезпечення

(код і повна назва спеціальності)

Тип програми \_\_\_\_\_ Освітньо-наукова

Керівник \_\_\_\_\_ к.т.н., доц. Мельнікова Р.В.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. Кафедри \_\_\_\_\_

3.В. Дудар

2022 р.

## Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук  
 Кафедра Програмної інженерії  
 Рівень вищої освіти другий (магістерський)  
 Спеціальність 121 - Інженерія програмного забезпечення  
 (код і повна назва)  
 Тип програми освітньо-наукова програма  
 Освітня програма Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_

(підпис)

«\_\_» \_\_\_\_\_ 202\_ р.

### ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Нечаєву Максиму Юрійовичу  
 (прізвище, ім'я, по батькові)

1. Тема роботи «Дослідження методів штучного інтелекту для збільшення масштабу зображень»

затверджена наказом університету від «24» березня 2022 р. №412Ст

2. Термін подання студентом роботи до екзаменаційної комісії «20» травня 2022 р.

Термін здачі студентом закінченої роботи „20” травня 2022 р.

3. Вихідні дані до роботи: технічне завдання, календарний план, методичні вказівки

4. Перелік питань, що потрібно опрацювати в роботі: аналіз предметної галузі, постановка задачі, методи дослідження, математична модель, проведення експерименту, опис розробленої програмної системи, результати роботи.

## КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної галузі	24.02.2022	<i>Виконано</i>
2	Огляд існуючих методів	31.02.2022	<i>Виконано</i>
3	Формування вимог, проектування та розробка ПЗ	26.03.2022	<i>Виконано</i>
4	Підготовка пояснювальної записки	13.05.2022	<i>Виконано</i>
5	Підготовка доповіді	19.05.2022	<i>Виконано</i>
6	Попередній захист	20.05.2022	<i>Виконано</i>
7	Нормоконтроль та перевірка на плагіат	13.05.2022	<i>Виконано</i>
8	Рецензування	18.05.2022	<i>Виконано</i>
9	Занесення диплому в електронний архів	18.05.2022	<i>Виконано</i>
10	Допуск до захисту у зав. кафедри	20.05.2022	<i>Виконано</i>

Дата видачі завдання 17 січня 2022 р.

Студент \_\_\_\_\_ Нечаєв М.Ю.  
(підпис)

Керівник роботи \_\_\_\_\_ доц. Мельнікова Р.В.  
(підпис) (посада, прізвище, ініціали)

## РЕФЕРАТ / ABSTRACT

Кваліфікаційна робота магістра містить: 64 стор., 30 рис., 4 табл., 13 джерел.  
МАСШТАБУВАННЯ ЗОБРАЖЕНЬ, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ,  
СУПЕР РОЗДІЛЬНА ЗДАТНІСТЬ, ESRGAN, PYTHON, MATLAB, UPSCALE

Об'єктами дослідження є згорткові нейронні мережі створення зображень супер роздільної здатності.

Метою роботи є порівняння експериментальним засобом алгоритмів та моделей масштабування для визначення підходящих сценаріїв їх використання.

У результаті роботи було порівняно експериментальним методом ефективність алгоритмів та моделей штучного інтелекту для збільшення роздільної здатності зображень. Спроектвана та реалізована програмна система для порівняння та оцінки масштабованих зображень.

IMAGE SCALING, CONVOLVED NEURAL NETWORKS, SUPER RESOLUTION, ESRGAN, PYTHON, MATLAB, UPSCALE

The objects of research are convolutional neural networks for creating super-resolution images.

The aim of the work is to compare algorithms and scaling models by experimental means to determine suitable scenarios for their use.

As a result of the work, the efficiency of algorithms and models of artificial intelligence to increase the resolution of images was compared by an experimental method. Designed and implemented software system for comparison and evaluation of scalable images.

Я, Нечаєв Максим Юрійович, студент групи ІІЗм-20-2, здобувач вищої освіти на другому (магістерському) рівні кафедри «Програмна інженерія», заявляю: моя кваліфікаційна робота на тему «Дослідження методів штучного інтелекту для збільшення масштабу зображень», що буде представлена до ЕК для публічного захисту, виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу EIArKhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання. Я ознайомлений (а) з діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

**ЗМІСТ**

ВСТУП.....	7
1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ.....	10
1.1 Огляд.....	10
1.2 Аналіз існуючих рішень.....	11
1.3 Аналіз проблем та актуалізація рішень.....	14
2 ПОСТАНОВКА ЗАДАЧІ.....	17
3 МЕТОДИ ДОСЛІДЖЕННЯ.....	18
4 МАТЕМАТИЧНА МОДЕЛЬ.....	19
4.1 Метрики для оцінки алгоритмів.....	19
5 ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ.....	26
5.1 Методологія експерименту.....	26
5.2 Підготовка вхідних зображень.....	26
5.3 Алгоритм ESRGAN.....	29
5.4 Алгоритм Real-ESRGAN.....	33
6 ОПИС РОЗРОБЛЕНОЇ ПРОГРАМНОЇ СИСТЕМИ.....	47
7 РЕЗУЛЬТАТИ РОБОТИ.....	48
ВИСНОВКИ.....	50
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	51
ДОДАТОК А.....	53
ДОДАТОК Б.....	54
ДОДАТОК В.....	55
ДОДАТОК Г.....	56
ДОДАТОК Д.....	64

## ВСТУП

Безперечно, зір – чи не найцінніший дар природи, адже за його допомогою ми орієнтуємося у навколишньому світі і здатні вести якісне, повноцінне життя. Зір є найважливішим сенсором людини, який здатний передати велике різноманіття життєво необхідної або прекрасної інформації. Зір дозволяє бачити явища, що нас оточують, у достатку кольорів та деталізації, а пара очей спрощує вимірювання відстані до спостережуваного предмету. Нам не доводиться замислюватися над тим, що таке зір, зображення, хоча насправді це найскладніший орган, який намагаються штучно відтворити останні двісті років.

Терміном зображення описують відтворення форми, кольору і деталей предмета при залученні світла. Здатність бачити предмет визначається здатністю цього предмета відбивати промені світла, які потрапляють на сітківку ока та створюють уявлення предмету після обробки мозком.

Тим часом колір це лише уявлення, яким чином мозок інтерпретує зображення. Під кольором розуміється електромагнітне випромінювання з певною довжиною хвилі, від якої залежить інтерпретований колір, або його відсутність через невидимість для людського зору. При випромінюванні всіх видимих довжин хвиль у електромагнітному спектрі ми бачимо білий колір, тоді як певні предмети можуть мати різноманітні характеристики по здатності поглинати одну частину видимого спектру, та віддзеркалювати другу, тим самим придбаючи певне фарбування для стороннього спостерігача.

З можливістю спостерігати в реальному часі навколишнє середовище виникла потреба зберегти спостереження, бо людська пам'ять неідеальна та через бажання поділитися своїм спостереженням з іншими людьми. Людина довгий час вдосконалювала свої навички малювання, принаймні 30 тисяч років, починаючи з примітивних печерних малюнків. При вдосконаленні людина знаходила нові джерела фарби, сірий камінь замінювався світлішим і портативним папірусом, а потім і білосніжним полотном, а на заміну грубому інструменту малювання приходила точніша кисть і перо. І ось на 19-те сторіччя припадає зачаток створення

штучного аналогу зору у вигляді камери. За останні 200 років стрімких досліджень камера пройшла дорогу від чорно-білої низькоякісної фотографії з тривалою витримкою за кілька годин та недовговічним засобом зберігання фото, до цифрових професійних камер з деталізацією та збільшенням кращими ніж може отримати людське око. Більше того, спеціальні камери взагалі переймають найкращі здібності зору тварин – бачення в непроглядній темряві та панорамне зображення у 360 градусів.

Цифрове зображення – масив даних, отриманий шляхом перетворення з аналогового об'єктиву фотокамери. Таке зображення має растровий тип, що представляє собою двовимірний масив пікселів, з чого випливає що є кінцева кількість інформації (пікселів, бітів), яка зберігається у файлі. Кількість інформації або роздільна здатність впливає на деталізацію зображення, одночасно збільшуючи розмір файлу. Якщо взяти до уваги, що сховища даних і трафік не нескінченні, та взагалі мало хто володіє гідним фотоапаратом насамперед, то можна припустити що існує величезна кількість зображень у незадовільної якості.

Збільшення масштабу зображення є проблемою, з якою доводиться стикатися багатьом компаніям, особливо в світі, де людям потрібно більше місця для зберігання своїх файлів, ніж будь-коли раніше. Всі фотографії, які щодня роблять користувачі смартфонів, створюють попит на сховище з боку цих користувачів. Ефективний спосіб вирішення цієї проблеми — зберігати зображення з нижчою роздільною здатністю та збільшувати їх щоразу, коли користувачеві потрібно отримати доступ до цього файлу, що значно полегшує зберігання кількох наборів зображень у базі даних.

Реставрація старих медіа, таких як фільми 20-го століття, старих фотографій та ретро-відеоігор – все це є важливими задачами для методів збільшення масштабу. Без збільшення масштабу старі медіа, які відтворюються на сучасних телевізорах і моніторах, матимуть низьку чіткість. Штучний інтелект (ШІ) може різко підвищити ефективність масштабування та підвищити роздільну здатність медіа за малу частку часу у порівнянні з ручним масштабуванням. Масштабування ШІ не є ідеальним і часто потребує втручання людини, коли бажані результати не

досягаються. Різні типи медіа вимагають різних алгоритмів, тому не існує універсального штучного інтелекту, який міг би точно масштабувати все.

Таким чином у даній роботі будуть розглянуті різні види масштабування, які використовують штучний інтелект. Буде експериментально порівняно між собою алгоритми масштабування та перевірена доцільність їх використання на різних прикладах зображень.

За результатами досліджень були опубліковані тези доповіді на тему «Дослідження методів штучного інтелекту для збільшення масштабу зображень» у дванадцятій міжнародній науково-технічній конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління» [1].

## 1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

### 1.1 Огляд

Перш ніж перейти до розгляду III методів масштабування, треба розглянути традиційні методи інтерполяції, які існують давно, виконуються за замовченням у будь-якому редакторі, що пов'язано з їх простотою у впровадженні та виконанні. Це найпростіший спосіб розтягнути зображення з нижчою роздільною здатністю на більший дисплей. Пікселі із зображення з нижчою роздільною здатністю копіюються та повторюються, щоб заповнити всі пікселі дисплея з вищою роздільною здатністю.

Фільтрування застосовується, щоб згладити зображення та заокруглити небажані нерівні краї, які можуть стати видимими через розтягування. В результаті виходить зображення, яке вміщується на дисплеї 4К, але часто може виглядати приглушеним або розмитим.

Інтерполяція зображення (масштабування зображення) відноситься до зміни розміру цифрових зображень і широко використовується програмами, пов'язаними з зображеннями [2]. До традиційних методів належать інтерполяція найближчого сусіда, лінійна, білінійна, бікубічна інтерполяція тощо.

Інтерполяція найближчого сусіда (NNI) є простим і інтуїтивно зрозумілим алгоритмом. Він вибирає значення найближчого пікселя для кожної позиції для інтерполяції незалежно від будь-яких інших пікселів.

Білінійна інтерполяція (BLI) спочатку виконує лінійну інтерполяцію на одній осі зображення, а потім виконує її на іншій осі [3]. Оскільки це призводить до квадратичної інтерполяції з розміром сприйнятливого поля  $2 \times 2$ , вона показує набагато кращу продуктивність, ніж інтерполяція найближчого сусіда, зберігаючи при цьому відносно високу швидкість.

Аналогічно, бікубічна інтерполяція (BCI) виконує кубічну інтерполяцію по кожній із двох осей. У порівнянні з BLI, BCI враховує  $4 \times 4$  пікселі, що призводить до більш гладких результатів з меншою кількістю артефактів, але значно меншою швидкістю. Зверніться до цього для детального обговорення.

Методи, засновані на інтерполяції, часто викликають деякі побічні ефекти, такі як складність обчислень, посилення шуму, результати розмиття тощо. Дані методи не можуть убирати шум, а масштабування не додає нових деталей.

Процес збільшення масштабу складається з отримання невеликого зображення з низькою роздільною здатністю та перетворення його у великий файл високої роздільної здатності. Це означає, що нове зображення матиме набагато більше пікселів, ніж попереднє, і для цього знадобиться алгоритм, щоб передбачити, яким буде колір нових пікселів у зображенні з високою роздільною здатністю. Існує багато типів алгоритмів, і більшість із них використовують мережі згортки, щоб спробувати точно передбачити колір нових пікселів, які будуть вставлені в збільшене зображення.

Згорткові нейронні мережі (CNN) використовуються для виявлення закономірностей у зображеннях. Зображення вставляються у вхідний шар, які потім надсилаються в приховані шари згортки, які виконують визначення шаблону за допомогою фільтрів. Фільтр — це матриця випадкових ініціалізованих значень, які проходять через кожен піксель зображення. Ці фільтри можуть розрізняти форми, текстури, краї тощо. Оскільки зображення проходить через все більше і більше згорткових шарів, фільтри стають більш розвиненими до точки, коли вони можуть розпізнавати більш складні структури, як-от автомобілі чи тварини. Потім зображення передаються на вихідний шар, коли ШІ вирішить, як класифікувати зображення.

## 1.2 Аналіз існуючих рішень

Метод глибокого навчання для супер-роздільної здатності одного зображення (SRCNN) безпосередньо вивчає наскрізне зіставлення між зображеннями з низькою/високою роздільною здатністю. Відображення представлено у вигляді глибокої згорткової нейронної мережі (CNN), яка приймає зображення з низькою роздільною здатністю на вхід і виводить зображення з високою роздільною здатністю [4]. На відміну від традиційних методів, які обробляють кожен компонент окремо, даний метод спільно оптимізує всі шари, має

легку структуру та демонструє найсучаснішу якість відновлення досягаючи високої швидкості для практичного використання в режимі онлайн. На рис. 1 зображено дію SRCNN методу у порівнянні з традиційними математичними методами масштабування. З наведеного малюнка ми бачимо, що за допомогою SRCNN отримано пікове співвідношення сигналу до шуму у розмірі в 27,58 дБ, що набагато краще, ніж класичне бікубічне та розріджене кодування (SC), яке було і залишається дуже актуальною темою дослідження.

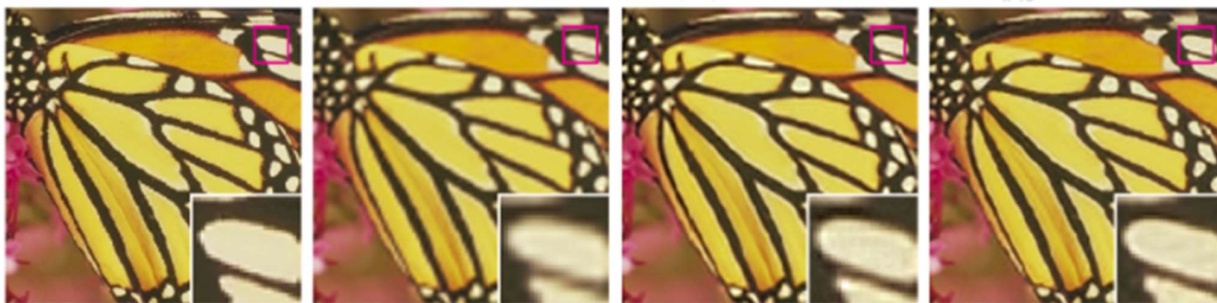


Рисунок 1 – Порівняння мережі SRCNN з традиційними засобами [5]

Другий з методів ШІ масштабування є Waifu2x, що спеціалізується на розширенні зображень в мультиплікаційному стилі з використанням глибоких згорткових нейронних мереж. Даний метод використовує модифіковану згортку нейронної мережі супер-роздільної здатності (m-SRCNN) для ілюстрацій та художніх робіт.

Наступний метод масштабування, також відомий як генеративна супер роздільна мережа, SRGAN є методом збільшення одного зображення, який вирішує головну проблему CNN, яка полягає у втраті більш дрібних деталей із збільшенням роздільної здатності. SRGAN використовує генеративну змагальну мережу (GAN), яка складається з двох нейронних мереж: генеративної та дискримінаційної мережі.

Генеративна мережа створює зразки для аналізу дискримінаційної мережі. Метою дискримінаційної мережі є точне розпізнавання, чи є зображення вже масштабованими (підробка), чи природно фотореалістичними (реальні). Навчання таким чином дозволяє ШІ точніше відновлювати дрібніші деталі під час процесу

масштабування. На рисунку 2 зображено порівняння мереж з двома індикаторами якості – пікове співвідношення сигналу до шуму (більше-краще), та індекс структурної відмінності (менше-краще). З наведених результатів видно, що краще з усіх впорався метод ESRGAN, о якому і піде мова далі.

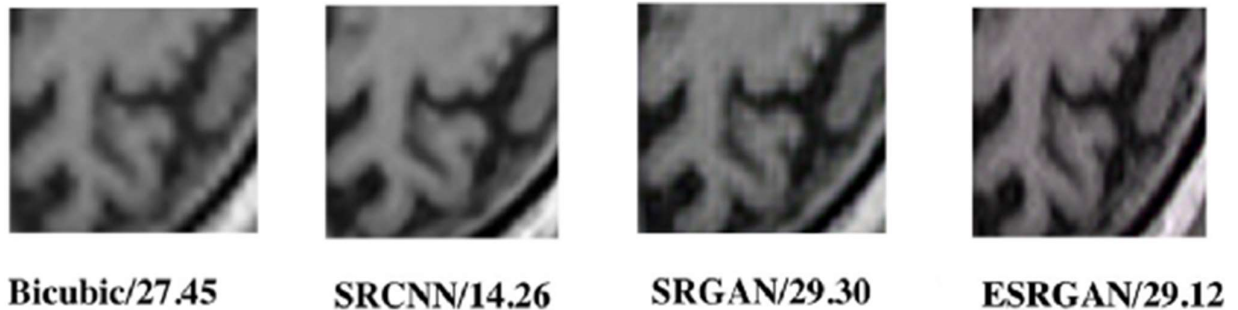


Рисунок 2 – Порівняння мереж масштабування [5]

Розширена версія SRGAN являє собою новий метод ESRGAN, що був створений шляхом покращення функції втрати сприйняття. Він являє собою висококласне програмне забезпечення для зображень, яке використовує технологію штучного інтелекту для збільшення розміру та якості зображення, не впливаючи на його пікселі. Платформа надає алгоритм, який спирається на попередньо визначені моделі для виконання всього процесу збільшення зображення. З рисунку 3 можна оцінити еволюцію методу ESRGAN у порівнянні зі своїм попередником. З цього порівняння можна зробити висновок, що ESRGAN дійшов набагато ближче до оригіналу через додавання різкості та більшої кількості деталей, але при детальному перегляді помітна різниця в орнаментах, коли масштабувальник «додумує» невідомі деталі, що не співпадають з дійсністю.

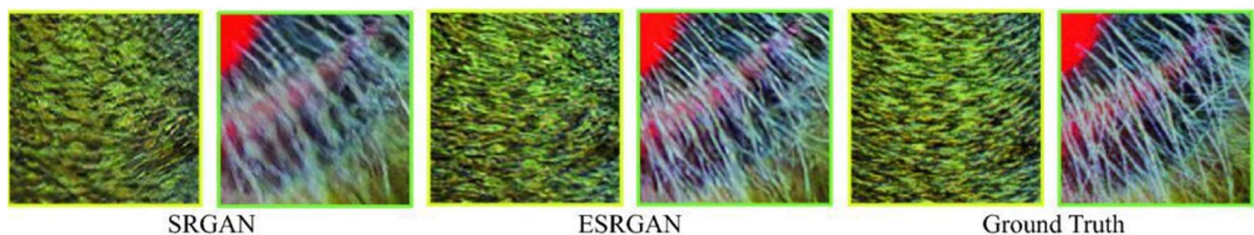


Рисунок 3 – Порівняння SRGAN та ESRGAN з оригіналом [6]

Усе збільшення роздільної здатності залежить від близькості попередньо визначеної моделі до бажаного входу або виходу. Платформа також дозволяє тренувати власні моделі за допомогою великих наборів зображень і великої обчислювальної потужності для підвищення роздільної здатності будь-якої графіки відеоігор. Рішення дозволяє редагувати фотографії, після чого заповнює пікселі в порожніх місцях. ESRGAN дозволяє об'єднати результати двох моделей, щоб створити нову модель [7]. Більше того, він не обмежує користувачів одним розміром і налаштовує зображення відповідно до нього, але користувачі можуть налаштовувати свої моделі скільки завгодно.

Хоча масштабування AI не ідеально, воно значно перевершує традиційні методи збільшення. ESRGAN, наприклад, виділяє дрібні деталі, які в іншому випадку були б втрачені при збільшенні розміру зображення. Оскільки ШІ піддається впливу більших варіацій навчальних зразків, він стає швидшим і точнішим.

Такі компанії, як Samsung і Sony, вже починають впроваджувати AI в телевізори, щоб динамічно покращувати якість зображення та роздільну здатність. Шанувальники старих відеоігор використовували штучний інтелект, щоб покращити якість зображення до рівня, якого раніше не було. А у сучасних комп'ютерних іграх ШІ масштабування дозволяє підтягнути якість зображення з меншої розподільної здатності, економлячи обчислювальні ресурси. Усі смартфони роблять ставку на фото алгоритми, які згладжують застій у прогресі мобільної оптики. Оскільки старі технології стають надзвичайно застарілими, ШІ, ймовірно, відіграє важливу роль у збереженні медіа.

### 1.3 Аналіз проблем та актуалізація рішень

Збільшення масштабу одиничного зображення є активною темою дослідження, яка спрямована на реконструкцію зображення високої роздільної здатності від його аналога з низькою роздільною здатністю. Починаючи з новаторської роботи SRCNN, підходи нейронної мережі глибокої згортки принесли процвітаючий розвиток у сфері супер-роздільної здатності. Проте більшість

підходів масштабування припускають, що зображення ідеально стиснуто бікубічним методом зниження дискретизації, що відрізняється від реальних шляхів деградації зображення. Ця невідповідність методів стиснення робить ці підходи непрактичними в реальних сценаріях.

Існують деякі проблеми, які перешкоджають ефективному застосуванню алгоритмів на реальних прикладах. Зокрема поява невідомих та складних деградацій зображень у результаті складних комбінацій різних процесів погіршення якості, таких як особливості камер, редагування зображень та передача через Інтернет. Наприклад, коли ми робимо фотографію за допомогою мобільних телефонів, фотографії можуть бути погіршені розмиттям камери, шумом сенсору, артефактами різкості та стиснення JPEG. Потім ми робимо деяке редагування та завантажуюмо в застосунок для соціальних мереж, яка вносить додаткове стиснення та непередбачувані шуми. Описаний вище процес ускладнюється, якщо зображення поширюється кілька разів в Інтернеті.

До таких деградацій належать артефакти дзвону та перельоту. Артефакти дзвону часто з'являються у вигляді помилкових країв біля різких переходів зображення. Вони візуально виглядають як смуги або «привиди» біля країв. Артефакти перельоту зазвичай поєднуються з артефактами дзвону, які проявляються як посилений стрибок на кінці краю.

Приклади артефактів дзвону та перельоту зображені на рисунку 4.



Рисунок 4 – Приклади артефактів дзвону та перельоту [8]

Навчання моделей це дуже трудомісткий процес який займає часи, навіть на десятці вхідних зображень та при використанні потужної відеокарти. У поєднанні з загальною втратою стабільності моделі, якщо вона перевантажена різноманітними стилями та сценаріями зображень, це майже унеможлиблює розробку універсальної моделі, яка б могла ефективно себе показувати при любых сценаріях.

Ці та інші проблеми будуть супроводжувати нас у спробах масштабування зображень за допомогою досліджуваних алгоритмів та їх моделей у частині проведення експериментів.

## 2 ПОСТАНОВКА ЗАДАЧІ

У результаті огляду та аналізу предметної галузі були виявлені найбільш проблемні місця існуючих аналогів, були сформовані вимоги щодо дослідження та розроблюваної програмної системи та сформульована постановка задачі.

Метою роботи є створення програмної системи, яка буде призначена для аналізу вхідних еталонних та масштабованих зображень з наданням оцінки о якості масштабування конкретної моделі або алгоритму на конкретному зображенні.

Програма має надавати наглядний звіт по проведеному експерименту – для кожного еталонного зображення має бути створена група, до якої входять стиснене та масштабовані зображення, які належать до цього еталонного. Після виконання аналізу програма має надати звіт в якому, зображення розташовуються у одну лінію для наочності порівняння, користувач має змогу перед створенням звіту обрізати усі зображення по однаковим координатам для перегляду малої ділянки.

Програма має подавати результат у вигляді метрик NIQE, PIQE, PSNR та SSIM. Також для кожної окремої метрики створюється таблиця у вигляді матриці, де кожному зображенню у рядках протиставляється кожний алгоритм у стовпчиках, а на перетині має знаходитись значення метрики. Програма має надати можливість відобразити для кожного зображення монохромну мапу структурної схожості (SSIM Map), на якій наглядно будуть позначені ділянки які сильно відрізняються від еталону.

Для тестування алгоритмів масштабування мають бути взяті зображення великого розміру та стиснені до маленького, який вже подається на вхід до мережі, після чого результат порівнюється з оригіналом для визначення достовірності. Щоб додатково зменшити емпіричність порівнянь, зображення мають бути оцінені математичними функціями якості, які дозволяють оцінити та порівняти зображення за допомогою числових індексів.

### 3 МЕТОДИ ДОСЛІДЖЕННЯ

У даній дослідницькій роботі використовуються як емпіричний і теоретичний методи дослідження на основі експериментів.

Теоретична частина дослідження полягає в використанні математичних оцінок реалістичності зображення. Для цього будуть обчисленні індекси сприйняття, природності та шуму результуючого зображення.

У даному дослідженні одних тільки теоретичних методів не вистачить, через неможливість об'єктивно оцінити зображення однією математикою. Емпіричне дослідження знадобиться для оцінки таких найскладніших завдань як масштабування живих істот, зокрема обличчя. При масштабуванні мультфільмів важливим також буде збереження оригінального стилю.

Перш ніж віддати нейромережовому алгоритму зображення на масштабування, цей алгоритм треба спочатку натренувати на вибірці зображень, щоб створити модель. А від обраної вибірки і залежить спеціалізація даної моделі, тому різні моделі і можуть надавати різний результат в одному і тому ж сценарії. Таким чином для всеосяжного експерименту треба навчити моделі для кожного можливого сценарію або створити універсальну модель.

Для порівняння існуючих алгоритмів масштабування, а також їх тренуваних моделей у різних сценаріях треба звернутися до експериментального методу дослідження. Буде порівняно алгоритми як з використанням нейронних мереж, так і примітивні математичні алгоритми.

## 4 МАТЕМАТИЧНА МОДЕЛЬ

### 4.1 Метрики для оцінки алгоритмів

Для визначення ефективності алгоритмів та моделей, наведених у цій роботі, треба надати математичну модель. Математична модель буде основним, але не останнім орієнтиром при визначенні характеристик і можливостей кожного з алгоритмів. Для даного дослідження математична модель має сенс у якості попереднього аналізу результату роботи досліджуваних алгоритмів.

Для оцінки ефективності алгоритмів та моделей будуть застосовані метрики оцінки якості зображень, отриманих після масштабування, які можна обчислити математично. Їх можна умовно поділити на дві групи.

До першої належать показники якості зображення без еталону. Важливою метою дослідження проблеми сліпої оцінки якості зображень (IQA) є розробка моделей, які можуть передбачити якість спотворених зображень, маючи якомога менше попередніх знань про зображення або їх спотворення. Сучасні сліпі алгоритми загального призначення вимагають знань про очікувані викривлення зображень у формі навчальних прикладів та відповідних оцінок людської думки. Однак нещодавно було виведено абсолютно сліпі моделі без тренування на людських оцінках. Будуть застосовані наступні дві метрики без посилання.

Перша метрика називається Naturalness Image Quality Evaluator (NIQE) – метод сліпої оцінки якості зображень. Це повністю сліпий аналізатор якості зображень, який використовує лише вимірювані відхилення від статистичних закономірностей, що спостерігаються в природних зображеннях, без навчання на оцінених людиною спотворених зображеннях, і без будь-якого впливу спотворених зображень. Аналізатор порівнює зображення з моделлю, розрахованою на основі зображень природних сцен. Метод повертає бал з плаваючою комою у діапазоні від 0 до 100, де менший бал свідчить про кращу якість сприйняття.

Друга метрика має назву Perception based Image Quality Evaluator (PIQE) – це теж сліпий метод оцінки якості зображень реального світу на основі сприйняття.

Він використовує середній коефіцієнт нормалізації контрасту віднімання для обчислення показника якості зображення. Являє собою неконтрольований метод, який не вимагає моделі навчання. Також повертає оцінку у діапазоні від 0 до 100, де менший бал означає кращу якість.

Для обчислення NIQE попередньо тренується модель, використовуючи ідентичні передбачувані статистичні характеристики, які називаються статистикою природної сцени або Natural Scene Statistics (NSS). NSS базується на нормалізованих коефіцієнтах яскравості на просторах зображення та моделюються як багатовимірний гаусовий розподіл. Спотворення з'являються як збурення гаусового розподілу (див. рис 5).

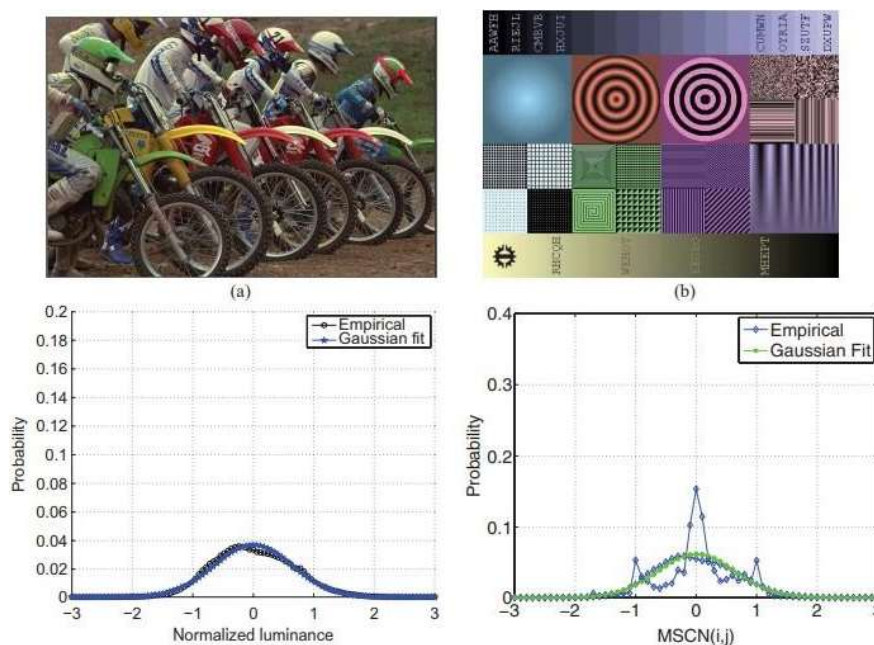


Рисунок 5 – Гаусовий розподіл NSS для природного та штучного зображень

Через те що NSS модель тренується тільки на реальних зображеннях, не має сенсу її застосовувати для аналізу штучно створених зображень, тому що усе зображення виявиться неприродним та буде вважатися артефактом, що помітно на графіку гаусового розподілу.

Покрокове виконання NIQE методу виглядає наступним чином:

- отримуються ознаки NSS зі статистично значущих блоків у спотвореному зображенні (див. рис. 6);

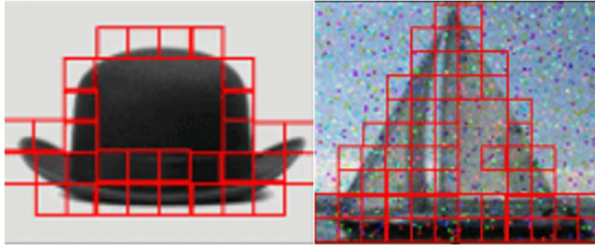


Рисунок 6 – Статистично значущі блоки

- функція підганяє багатовимірне розподілення Гауса до характеристик NSS зображення;
- показник якості обчислюється як відстань між гауссовими розподілами: еталонного (природного) та вхідного (спотвореного).

PIQE обчислюється за допомогою оцінки поблочного спотворення, використовуючи такі кроки:

- для кожного пікселя на зображенні розраховується коефіцієнт з нормованим контрастом із відніманням середнього значення MSCN (Mean-Subtracted Contrast-Normalized) для подальшого аналізу їх розподілу. Відомо, що для природних зображень такий розподіл поводить себе нормально, а різного роду спотворення порушують цю закономірність. Дані коефіцієнти дослівно обчислюються через віднімання середнього значення від вхідного зображення та нормалізацією контрасту, що виражається формулою нижче:

$$\hat{I}(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + c}, \quad (1)$$

де  $(i, j)$  - це координати пікселя;

$\hat{I}$  – MSCN коефіцієнти;

$I$  – вхідне зображення;

- $\mu$  – локальне середнє від зображення;
- $\sigma$  – поле локальної дисперсії;
- $C$  – константа.

Графічне пояснення до формули 1 зображено на рисунку 7. Далі дані коефіцієнти застосовуються для подальшого аналізу їх розподілу. Відомо, що для природних зображень такий розподіл поводитья нормально, а різного роду спотворення порушують цю закономірність;

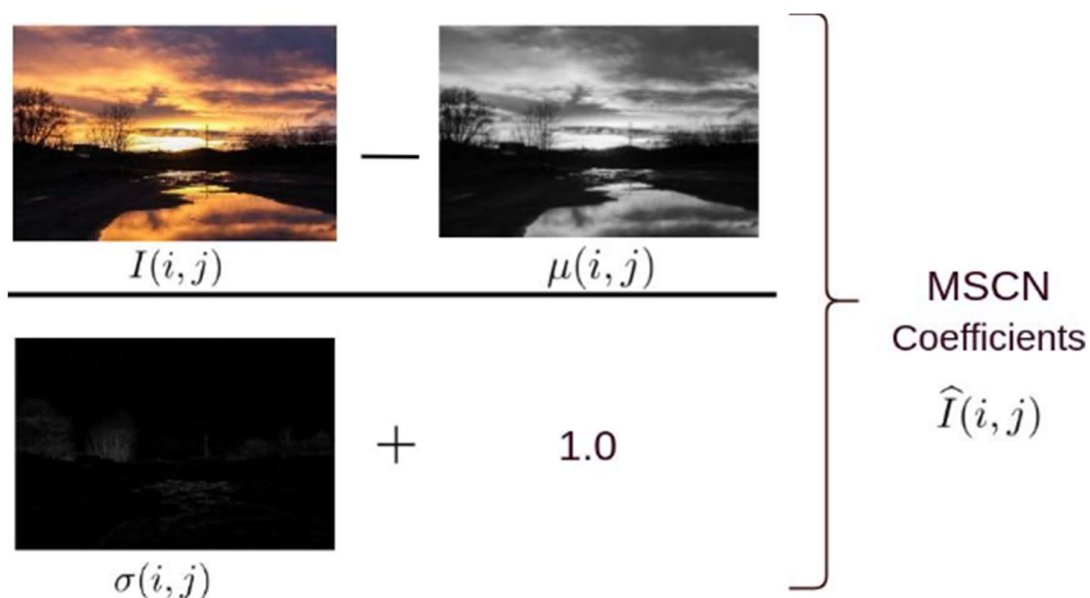


Рисунок 7 – Обчислення MSCN коефіцієнтів

- вхідне зображення розділяється на блоки, що не перекриваються, розміром 16 на 16;
- визначаються високоактивні у просторі блоки на основі дисперсії коефіцієнтів MSCN;
- генерується маска активності (див. рис. 8), використовуючи ідентифіковані високоактивні у просторі блоки;
- у кожному блоці оцінюється спотворення через артефакти блокування та шум за допомогою коефіцієнтів MSCN;

- використовуються порогові критерії, щоб класифікувати блоки як спотворені блоки з блочними артефактами, спотворені блоки з гаусовим шумом і неспотворені блоки;
- створюється маска помітних артефактів із спотворених блоків із блочними артефактами та маска шуму із спотворених блоків із гаусівським шумом (див. рис. 8);



Рисунок 8 – Вхідне спотворене зображення, маска активності, маска помітних артефактів та маска шуму.

- обчислюється оцінка PIQE для вхідного зображення як середнє значення балів у спотворених блоках.

До другої групи належать метрики які роблять оцінку на основі порівняння еталонного зображення з масштабованим.

Перша метрика з посиланням на еталон називається PSNR (peak signal-to-noise ratio) – це пікове співвідношення сигналу до шуму означає співвідношення між максимумом можливого значення сигналу та потужністю шуму, що спотворює значення сигналу [9]. Для його визначення використовується середньоквадратичне відхилення (MSE) яке обчислюється наступним чином:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2,$$

де  $I$  та  $K$  – монохромні зображення розміру  $m \times n$ , одне з яких є зображення-оригінал, а друге є масштабована версія з шумом.

PSNR обчислюється за наступною формулою:

$$PSNR = 10 \log_{10} \left( \frac{MAX_I^2}{MSE} \right) = 20 \log_{10} \left( \frac{MAX_I}{\sqrt{MSE}} \right),$$

де  $MAX_I$  – відповідає за максимальне значення пікселя зображення, що дорівнює  $2^8 - 1$  для 8-бітного зображення.

Друга метрика SSIM (structure similarity) – індекс структурної подібності вважається складнішою, та її треба застосовувати для більш точного визначення різниці двох зображень. Даний показник якості зображення оцінює візуальний вплив трьох характеристик зображення: яскравості, контрасту та структури. Він є розвитком традиційних методів, описаних вище, та він, на відміну від попередніх є сумісним з фізіологією людського сприйняття. Особливістю є, що він завжди лежить у проміжку від -1 до 1, причому при його значенні рівному 1 вважається, що ми маємо дві однакові картинки. На рисунку 9 зображено мапу SSIM оцінки кожного пікселя, де білому кольору відповідає значення пікселю наближене до 1, а чорному – наближене до -1.



Рисунок 9 – SSIM мапа, де чорним кольором позначено значну різницю між еталонним та спотвореним зображеннями

Індекс оцінки якості SSIM базується на обчисленні трьох характеристик, а саме терміну яскравості, терміну контрасту та структурному терміну.

Загальний індекс являє собою мультиплікаційну комбінацію трьох термінів:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma,$$

$$\text{де } l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_{2x} + \mu_{2y} + C_1},$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_{2x} + \sigma_{2y} + C_2},$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3},$$

де  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$ ,  $\sigma_y$ , та  $\sigma_{xy}$  – локальні середні значення, стандартні відхилення та перехресна коваріація для зображень  $x$ ,  $y$ .

Для підсумкової оцінки моделей алгоритмів буде розраховано середнє арифметичне з оцінок метрик, отриманих з 9-ти зображень, у одну за наступною формулою:

$$\overline{Metric_{model}} = \frac{\sum_{i=1}^n Metric_i}{n},$$

де *Metric* – це метрики NIQE, PIQE, PSNR, SSIM;

*model* – це модель, до якої належить ця метрика;

*i* – порядковий номер від 1 до 9 по кожному зображенню;

*n* – загальна кількість зображень.

За цією формулою буде визначена найбільш універсальна модель з досліджуваних у експерименті.

## 5 ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ

### 5.1 Методологія експерименту

Використовуючи попередньо наведену методологію можна сформуванати структуру та послідовність ходу експерименту. Метою експерименту є порівняння існуючих алгоритмів масштабування та їх моделей на наборі вхідних зображень. Зображення представляють собою приклади реальних сценаріїв використання, для кожного з котрих треба визначити найбільш ефективний засіб масштабування, пояснити загальні принципи його роботи та обґрунтувати отриманий результат.

Для проведення експерименту взято 9 початкових зображень у категоріях природа, обличчя, малюнок, архітектура та тварина. Дані зображення будуть стиснені до однакового розміру у двох екземплярах з довжиною сторони у 128 пікселів – для їх подальшого масштабування за допомогою алгоритмів, та 512 пікселів – для порівняння масштабованого зображення з еталоном. Таким чином для кожного алгоритму будуть використовуватися однакові набори даних. Алгоритми будуть проаналізовані на цих зображеннях, будуть надані пояснення щодо принципів роботи.

В експерименті будуть задіяні 2 найбільш прогресивних алгоритми, які в сумарній кількості мають 6 моделей.

Отримані дані з результатів експерименту будуть занесені до підсумкової порівняльної таблиці. Дані для порівняння ефективності масштабування являють собою чотири метрики, що були наведені в попередній частині, а саме: оцінки природності, сприйняття, шуму та структурної подібності до оригіналу.

Після проведення експерименту на усьому обсязі прикладів з використанням усіх моделей алгоритмів та оформлення порівняльної таблиці, останньою частиною буде підведення підсумків, наведення пропозицій щодо тренування моделей.

### 5.2 Підготовка вхідних зображень

Для всеосяжного аналізу алгоритмів стояла мета визначитися з конкретними сценаріями їх використання, а для кожного сценарію відібрати підходящі зображення. Було відібрано такі сценарії: архітектура, природа, обличчя, тварина

та мультиплікація. За даними сценаріями відібрано зображення та зображено на рисунку 10.

Архітектура	Природа	Обличчя	Тварина	Мульти.
				
				

Рисунок 10 – Колекція вхідних зображень

Усі відібрані зображення у високій якості, а для експерименту потрібно навпаки їх зменшити. Тому кожне зображення стискається до розміру 128 пікселів однієї з сторін, а інша теж зменшується із збереженням пропорцій. Для масштабування будуть використовуватися 4x моделі, як найоптимальніші, тобто зображення буде масштабовано у 4 рази. З приводу того, що в нас мають дві метрики які потребують еталонне зображення, а точніше еталонне має дорівнювати масштабованому до пікселя, нам також треба заздалегідь зробити версію зображення розміром у 4 рази більше, ніж найбільш стиснене у 128 пікселів, тобто 512 пікселів. Таким чином на модель алгоритму подається зображення із стороною в 128 пікселів, його масштабує до 512, і тоді ми обчислюємо метрики PSNR та SSIM у порівнянні з еталоном. Але на цьому моменті у абсолютної більшості зображень може не зійтись розмір масштабованого та еталонного зображення. Виявляється, що при зменшенні зображення треба слідкувати, щоб пропорційна сторона еталону після зменшення була ділима на 4 без остатку. У разі якщо еталон отримав розмір

512x333, тоді випробуване зображення матиме розмір 128x83.25, що округлюється у меншу сторону, після чого алгоритм масштабування поверне зображення розміром 512x332, яке очевидно не збігається.

Наступним розглянутим питанням було застосування методу ресемплінга, тобто зменшення зображення, адже маються щонайменше 4 засоби: білінійний, найближчий сусід, бікубічний звичайний та бікубічний гостріший розташовані у порядку на рисунку 11.



Рисунок 11 – Методи ресемплінгу

З цих зображень миттєво неозброєним оком виділяється друге, піддане зменшенню методом найближчого сусіду, і це не має бути дивним, тому що цей метод використовується в більшості випадків для піксель-артів та ретро ігор. Щоб знайти кращий метод з трьох зображень, що залишилися, вдамося до масштабування.

У результаті масштабування одною з моделей алгоритму ESRGAN та використовуючи розроблену програму на платформі MATLAB отримаємо не тільки графічне порівняння, а й всі оціночні метрики (див. рис. 12). Деталі алгоритмів програм буде наведено в наступній частині.

У даному ряду картинок зліва направо розташовані по порядку: еталон розміром у 512 пікселів, зменшене зображення розміром у 128 пікселів, масштабування білінійно, гостро бікубічно, автоматично та просто бікубічно зменшених зображень. Під автоматичним варіантом мається на увазі стандартний метод зміни розміру зображення у фотошопі, та й він повністю збігається з гострим

бікубічним зменшенням, тому можна зробити висновок, що автоматичний режим просто дублює існуючі методи, тому його недоцільно використовувати.

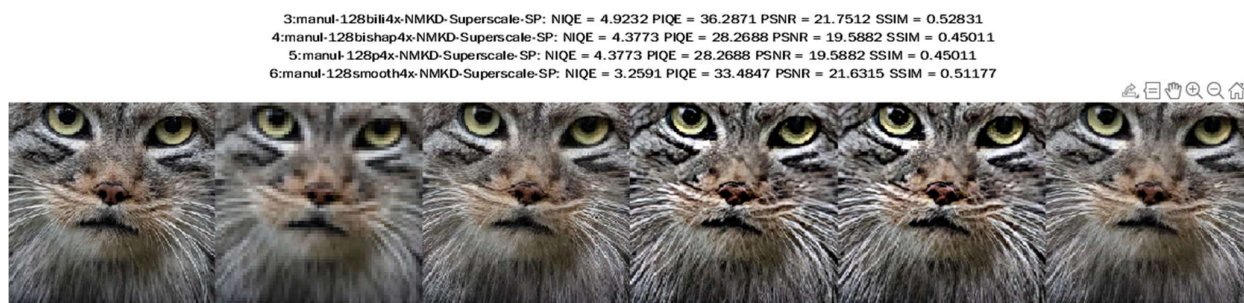


Рисунок 12 – Масштабування зменшених різними засобами зображень

Бікубічний гострий метод виділяється на фоні надмірною різкістю та товстими лініями на шерсті. Цікаво що метрики з використанням еталону в цілому погоджуються з тезою, що передостанні зображення значно відрізняються від оригіналу як по шуму, так і по структурній схожості. Проте сліпі метрики мають різне бачення щодо надмірно різкого манула. Індекс сприйняття PIQE дуже високо оцінив різке зображення через велику кількість деталей, що спливли, на противагу розмиттю, яке б погіршило оцінку. У той же час індекс природності NIQE віддає перевагу саме бікубічному методу, через відсутність малопомітних блочних спотворень на зображенні.

За сумарними плюсами бікубічного методу, надалі буде використовуватись тільки він.

### 5.3 Алгоритм ESRGAN

ESRGAN являє собою доволі мейнстрімний та зрілий алгоритм масштабування, а завдяки його безкоштовності та відкритості вихідного коду він сильно обріс користувацькими моделями на будь-який смак.

Базовий алгоритм за яким здійснюється масштабування та аналіз результатів: – завантажені моделі з сумісною архітектурою розміщуються у папці “models”;

- зображення у поганій якості заносимо до папки, з якої їх буде зчитувати алгоритм;
- запускаємо алгоритм з надаванням назви моделі у параметрі;
- після виконання алгоритму, виконуємо ще раз для інших моделей за потреби. Отримані результати та еталонні зображення переносимо до платформи MATLAB, розміщуючи у відповідних папках;
- програма проходить по кожному еталонному зображенню “А” та збирає до нього за назвою усі масштабовані зображення різних моделей “А-х”, “А-у” і так далі;
- для кожного окремого зображення обчислюються сліпі метрики NIQE та PIQE у один крок та без параметрів;
- обчислення PSNR відбувається дещо складніше зі зверненням до моделі “psnr.slx”, а при обчисленні SSIM можна вказати другим параметром змінну для отримання монохромної мапи відмінностей;
- останнім кроком циклу є друк отриманих метрик та відображення зображень однієї групи у один ряд.

Для тестування алгоритму було задіяно 4 моделі, кожна з яких має свою сферу застосування, тому не для кожного зображення підійде кожна модель. Наступне зображення дозволяє оцінити масштабування у сценаріях природи та тварин (див. рис. 13).

```
3:patron-128blsmooth4x-NMKD-Superscale-SP: NIQE = 2.3461 PIQE = 27.2371 PSNR = 25.42 SSIM = 0.85778
4:patron-128blsmooth4x-OLDIES: NIQE = 6.414 PIQE = 82.0005 PSNR = 20.8379 SSIM = 0.82733
5:patron-128blsmooth4x-UltraSharp.pth: NIQE = 2.0947 PIQE = 23.7888 PSNR = 25.0282 SSIM = 0.84981
6:patron-128blsmooth4x-cat-patch: NIQE = 3.8072 PIQE = 43.7205 PSNR = 27.13 SSIM = 0.90936
```



Рисунок 13 – Масштабування зображення patron.jpg за допомогою 4-ьох моделей

На цьому ряді зображень перші два знову представляють собою еталонне та зменшене зображення, а наступні уособлюють такі моделі:

- SuperScale, який призначений для збільшення масштабу реалістичних зображень/фотографій із шумом і артефактами стиснення;
- OLDIES – модель для масштабування мультфільмів та аніме;
- UltraSharp – універсальна модель для створення багато деталей, залишає гарну текстуру на зображеннях;
- Cat-patch – модель тренувана на домашніх улюбленцях.

У даному прикладі одразу стикаємося з тим, що моделі для мультфільмів (четверте зображення) не можна використовувати для серйозного масштабування реальних зображень, але можна розглядати як фільтр для створення мультяшних зображень із реальних. Метрики наведено у таблиці 5.1.

Таблиця 5.1 – Результати проведення експерименту для алгоритму

Модель	Метрика			
	NIQE (менше-краще)	PIQE (менше-краще)	PSNR (більше-краще)	SSIM (більше-краще)
SuperScale	2.3461	27.2371	25.42	0.85778
OLDIES	6.414	82.0005	20.8379	0.82733
UltraSharp	2.0947	23.7888	25.0282	0.84981
Cat-patch	3.8072	43.7205	27.13	0.90936

До цього також додати, що сліпі метрики NIQE та PIQE, які навчалися виключно на реальних зображеннях, також завжди будуть недооцінювати мультяшні зображення через відсутність в них природних та сприятливих ознак. В той же час інші метрики будуть також знижувати оцінку через велику розбіжність з еталоном. Таким чином неправильно обрана модель може тільки погіршити зображення.

Стосовно інших реалістичних моделей, можна виділити останню як надслабку, продукуючу занадто мильну картинку, що також відображається у метриках.

Останні дві моделі, на рисунку 13 на 3-тій і 5-тій позиціях, дуже близькі одна до іншої як візуально так і за метриками. Можна позначити що, спираючись на метрики, модель SuperScale створює менш гарне зображення, але воно найбільш наближене до оригіналу, тоді як UltraSharp навпроти – отримує кращі оцінки по природності та сприйняттю, але трохи відстає від еталону. Таким чином можна зробити попередній висновок, що перша модель більш реалістична, а друга додає зайві неіснуючі деталі.

Повернемося до зображення манула, бо для даного сценарію є підходяща модель cat-patch, яка і змогла значно обійти усі інші моделі за реалістичністю за метриками PSNR та SSIM (див. табл. 5.2).

Таблиця 5.2 – Результати проведення експерименту для алгоритму

Модель	Метрика			
	NIQE (менше-краще)	PIQE (менше-краще)	PSNR (більше-краще)	SSIM (більше-краще)
SuperScale	3.2591	33.4847	21.6315	0.51177
OLDIES	5.4899	81.0585	19.8019	0.45996
UltraSharp	3.1849	21.5706	22.5715	0.573
Cat-patch	3.7137	22.8625	24.0021	0.6669

Якщо спробуємо емпірично оцінити останнє зображення у порівнянні до попередніх, то одразу впадає в очі як гармонічно та натурально виглядає остання тварина, на відміну від попередніх, де присутня зайва різкість, жорсткі контури морди, через що видно багато нерівностей та виділяючихся пікселів (див. рис. 14).



Рисунок 14 - Масштабування зображення manul.jpg, зображено тільки моделі

Таким чином, вузькоспеціалізовані моделі мають більше шансів проявити найкращу ефективність при належному сценарії.

#### 5.4 Алгоритм Real-ESRGAN

Real-ESRGAN – це суттєва переробка попереднього алгоритму ESRGAN, яка ставить ціллю боротьбу з різноманітними артефактами стискання.

Дана мережа намагається вирішити насущні проблеми свого попередника в області тренування всеосяжних моделей, з ціллю навчити прогнозувати усі можливі спотворення найменшими задіяними ресурсами.

Перша проблема стосується неможливості навчити попередній метод усім можливим комбінаціям стиснення. Оскільки простір деградації зображень набагато більше, ніж можуть охопити моделі методу ESRGAN, навчання також стає більш складним. Для вирішення цієї проблеми потрібні парні навчальні дані з деградаціями подібними до реальних сценаріїв для тренування мереж.

Існує декілька методів для навчання мережі, які можна умовно розділити на три категорії:

- зйомка парних даних за допомогою певних камер з подальшим вирівнюванням;
- вивчення розподілу деградацій зображень напряму з непарних даних з подальшою синтезацією парних даних для тренування;
- синтезування парних даних з класичними операторами, такими як шум, розмиття тощо, а потім узагальнити навчені моделі до реальних погіршень якості.

Перевага даного засобу в тому, що нема необхідності вести зйомку реальних прикладів та аналізувати їх, а навпаки дозволяє синтезувати ще неіснуючі дані штучно, що робить вибірку більш різноманітною. Єдиний виклик в тому, щоб синтезовані дані були максимально наближені до природних даних.

Друга проблема полягає в тому, як обробляти різні деградовані зображення в одній уніфікованій мережі. Наприклад, сліпі методи супер-роздільної здатності (Blind Super-Resolution)[10] зазвичай складаються з двох гілок, де одна прогнозує

деградації зображень, а друга, на основі першої робить умовне відновлення зображень

Метод сліпої супер-роздільної здатності має на меті відновити зображення з низькою роздільною здатністю, які страждають від невідомих і складних деградацій. Існуючі підходи можна приблизно розділити на явне моделювання та неявне моделювання відповідно до процесу деградації, що лежить в основі.

Класична модель деградації зображення, яка складається з розмиття, зниження дискретизації, додавання шуму та стиснення методом JPEG, широко використовується в методах явного моделювання. Однак реальні деградації зазвичай занадто складні, щоб їх можна було моделювати за допомогою простої комбінації кількох деградацій. Таким чином, ці методи легко зазнають невдачі в реальних зразках. Методи неявного моделювання використовують навчання розподілу даних за допомогою генеративної змагальної мережі (GAN) для отримання моделі деградації зображення [11]. Тим не менш, вони обмежуються деградацією в навчальних наборах даних і не можуть добре узагальнити нерозповсюдженні зображення.

Ключовою відмінністю нового алгоритму є спрощення проблеми відновлення загальних реальних зображень низької роздільної здатності шляхом створення навчальних пар із більш реалістичним процесом деградації.

Це спонукає ускладнити класичну модель деградації зображень таким чином, щоб вона моделювалась за допомогою кількох повторюваних процесів деградації (деградація високого порядку), кожен з яких є класичною моделлю деградації. Для гарного балансу між простотою та ефективністю було використано процес деградації другого порядку, тобто кількість повторюваних процесів дорівнюватиме двом. Інновацією стало також використання стратегії випадкового перемішування для синтезу більш практичних деградацій, однак ця стратегія не виключає фіксовану кількість процесів деградації, і неясно, чи всі перемішані деградації корисні чи ні. Для створення цих деградацій, задіяно фільтр sinc, що являє собою ідеалізований фільтр, який відсікає високі частоти для синтезу артефактів дзвону (ringing) і перельоту (overshoot) для навчальних пар [12].

Фільтр відрізання високих частот для запобігання артефактів sinc виражається за наступною формулою:

$$k(i, j) = \frac{\omega_c}{2\pi\sqrt{i^2 + j^2}} J_1(\omega_c\sqrt{i^2 + j^2}),$$

де  $(i, j)$  – координата ядра;

$\omega_c$  – частота зрізу, графічну візуалізацію якого можна побачити на рис. 15;

$J_1$  – це функція Бесселя першого порядку першого роду.

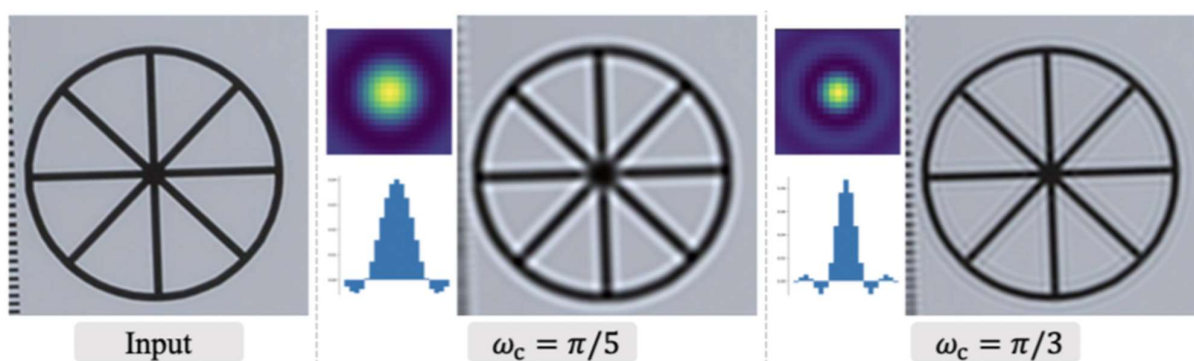


Рисунок 15 – Приклади впливу частоти зрізу  $\omega_c$  на синтезоване зображення

Із радикальним збільшенням тренувальних пар, дискримінатор вимагає більш потужної можливості розпізнавання реальності від складних тренувальних результатів, та він має бути більш точним для локального покращення деталей. Тому було поліпшено дискримінатор VGG, що застосовується в ESRGAN, до більш сучасного U-Net дизайну.

По-друге, структура U-Net і ускладнення навчальних сетів надскладними деградаціями також збільшують нестабільність навчання. Для протистоянь цьому використано регуляризацію спектральної нормалізації (Spectral Normalization) для стабілізації динаміки навчання [13]. Завдяки спеціальним удосконаленням ESRGAN може легко тренуватися і досягається гарний баланс локального покращення деталей і подавлення артефактів.

Підводячи підсумок, у Real-ESRGAN є наступні відмінності від оригінального ESRGAN:

- використовується процес деградації високого порядку для моделювання практичної деградації зображення та використовуються фільтри «sinc» для моделювання артефактів дзвону та перельоту;
- використано кілька істотних модифікацій (наприклад, дискримінатор U-Net зі спектральною нормалізацією), щоб збільшити можливості дискримінатора та стабілізувати динаміку навчання;
- використовуються чисті синтетичні дані у навчанні, що дозволяє відновлювати більшість реальних зображень і досягати кращої візуальної продуктивності, ніж попередні роботи, що робить метод більш практичним у реальному світі.

Даний алгоритм ще у розробці, тому на нього є лише обмежена кількість моделей, тому буде експеримент буде обмежений використанням однієї загальної та однієї мультиплікаційної моделей. Результати аналізу моделей Real-ESRGAN зображені на рисунку 16.



Рисунок 16 – Масштабування patron.jpg алгоритмом Real-ESRGAN

Перше, що одразу кидається у очі – невідмінність зображень у мініатюрі незважаючи на використання зовсім протилежних моделей. Це частково обумовлено відсутністю жовтого фільтру, а також стиль не намагається походити під малюнок олівцем. Навпроти – реальне зображення здебільшого перетворюється до зліпленого з пластиліну. Хоча японські анімації не виглядають саме таким

чином, модель не намагається нав'язати якийсь стиль, а лише робить кольори більш плоскими, а лінії більш рівними, прибираючи пухнастість. Це також спостерігається по метрикам, майже всі з котрих мають дуже близькі оцінки (див. табл. 5.3).

Таблиця 5.3 – Результати експерименту алгоритму на зображенні patron.jpg

Модель	Метрика			
	NIQE (менше-краще)	PIQE (менше-краще)	PSNR (більше-краще)	SSIM (більше-краще)
Anime-model	3.6645	42.7223	24.9186	0.86073
RealESRGAN x4plus	3.4008	29.861	25.7616	0.87682

Якщо порівнювати реалістичну модель цього алгоритму з попереднім ESRGAN, то видно що дива не сталося, метрики десь у тому ж діапазоні.

Застосуємо масштабування до зображення манула (див. рис. 17)

3:manul-128blsmooth-anime.jpg: NIQE = 3.6978 PIQE = 28.1911 PSNR = 22.4532 SSIM = 0.57512  
4:manul-128blsmooth: NIQE = 4.0983 PIQE = 21.0437 PSNR = 22.1124 SSIM = 0.56776

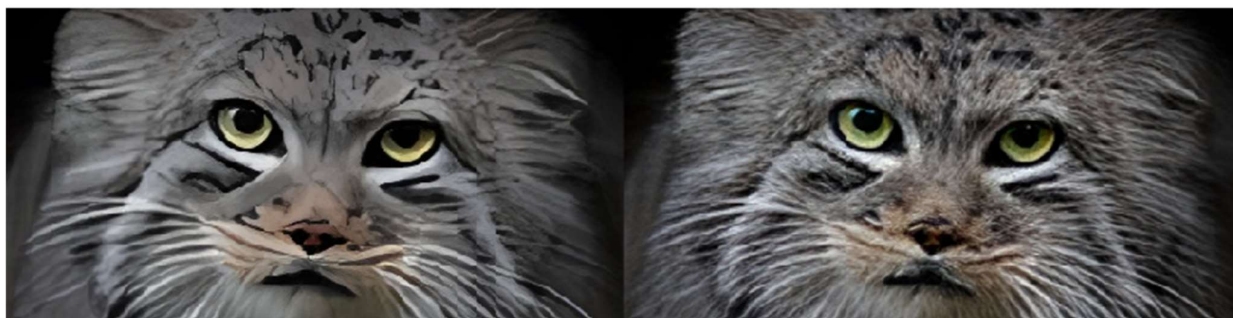


Рисунок 17 – Масштабування manul.jpg алгоритмом Real-ESRGAN

При спробі масштабувати зображення манула через дані моделі, отримується дуже парадоксальний результат де мультиплікаційна версія виграє одразу по трьом метрикам у реалістичної, значно відрізняючись від еталонного, та не виглядаючи дуже природньо (див. табл. 5.4). Це зовсім ставить під сумнів використання будь-якої метрики.

Таблиця 5.4 – Результати експерименту алгоритму на зображенні manul.jpg

Модель	Метрика			
	NIQE (менше- краще)	PIQE (менше- краще)	PSNR (більше- краще)	SSIM (більше- краще)
Anime-model	3.6978	28.1911	22.4532	0.57512
RealESRGAN x4plus	4.0983	21.0437	22.1124	0.56776

На наступному рисунку 18 зображено масштабування мультиплікаційної картини.

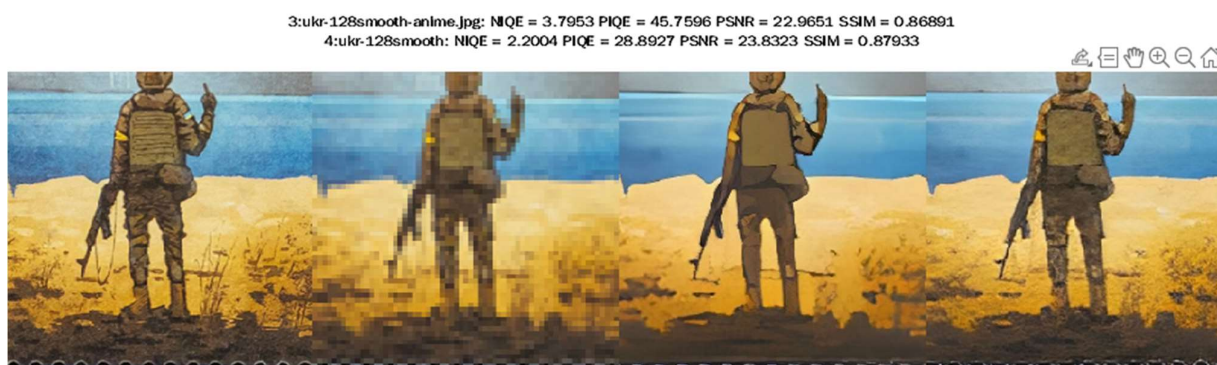


Рисунок 18 – Масштабування марки Укрпошти

Можна прослідкувати що використання реалістичної моделі загального призначення дуже позитивно та по усіх метриках виграє у аніме моделі, та візуально помітно у чому справа.

Якщо ціллю масштабування є вилучення втрачених деталей із стисненого зображення, то з цією задачею реалістична модель впорається краще, на відміну від мультиплікаційної, яка ставить ціллю навпаки спростити деталі.

### 5.5 Порівняння результатів алгоритмів

Для формування уявлення про те, яка модель підходить найкраще до якого зображення, у програмній системі було реалізовано формування чотирьох таблиць по однієї для кожної з метрик, де стовпчики репрезентують застосовану модель, а рядки – оброблене зображення. На перетині рядка і стовпчика отримаємо значення індексу (див. рис. 19).

NIQE	4x-NMKD-Superscale-SP	4x-OLDIES	4x-UltraSharp.pth	4x-cat-patch	anime-esrgan	real-esrgan
"banner"	"3.9783"	"5.1557"	"2.7671"	"3.1283"	"4.0786"	"2.3624"
"delmonico"	"4.6153"	"6.5314"	"3.9345"	"3.7404"	"4.3005"	"3.1328"
"dry"	"3.4551"	"4.3194"	"2.2365"	"3.3775"	"3.2733"	"2.2525"
"einstein"	"7.1575"	"5.6677"	"3.0155"	"3.8742"	"2.4667"	"2.3772"
"faces"	"4.5523"	"7.8981"	"4.6192"	"4.5069"	"6.4486"	"4.3082"
"landscape"	"3.7418"	"7.6419"	"2.1515"	"5.0276"	"3.3953"	"2.4025"
"manul"	"3.2591"	"5.4899"	"3.1849"	"3.7137"	"3.6978"	"4.0983"
"patron"	"2.3461"	"6.414"	"2.0947"	"3.8072"	"3.6645"	"3.4008"
"ukr"	"3.3282"	"6.3023"	"3.2044"	"3.2859"	"3.7953"	"2.2004"

Рисунок 19 – Перехресна таблиця метрики NIQE (менше-краще)

Рамками означені найкращі або одні з найкращих результатів для конкретної моделі, а колір приблизно відображає наскільки даний показник високий у порівнянні з усіма зібраними метриками. З даної таблиці можна зрозуміти, що перша модель SuperScale краще всього впоралась з зображенням пса, за думкою метрики природності NIQE. Тим часом друга модель впоралась поганіше усіх інших з усіма можливими зображеннями, а найлегше їй далось зображення сухого дерева (див. рис. 20).



Рисунок 20 – Масштабування крон дерев

Фаворитами даної метрики стали реалістичні загальні моделі UltraSharp та Real-ESRGAN. Вони впевнено обійшли майже усі моделі, але між ними є деякі відмінності – так перша модель має перевагу у масштабуванні пухнастих тварин, а друга має перевагу при обробці двовимірних та монохромних зображень.

На рисунку 21 зображена таблиця з метрикою сприйняття PIQE.

PIQE	4x-NMKD-Superscale-SP	4x-OLDIES	4x-UltraSharp.pth	4x-cat-patch	anime-esrgan	real-esrgan
"banner"	"34.6912"	"81.3919"	"36.9888"	"40.6129"	"51.2174"	"35.3703"
"delmonico"	"37.0609"	"83.5936"	"37.0605"	"51.5557"	"42.1047"	"37.3564"
"dry"	"32.0889"	"79.1307"	"23.9647"	"29.6691"	"34.5189"	"30.213"
"einstein"	"40.7613"	"82.7346"	"12.3998"	"15.7883"	"29.7455"	"27.5423"
"faces"	"44.387"	"82.0314"	"47.9311"	"55.8347"	"45.1651"	"42.6655"
"landscape"	"40.9634"	"84.9296"	"26.7496"	"35.2544"	"35.9481"	"37.5041"
"manul"	"33.4847"	"81.0585"	"21.5706"	"22.8625"	"28.1911"	"21.0437"
"patron"	"27.2371"	"82.0005"	"23.7888"	"43.7205"	"42.7223"	"29.861"
"ukr"	"22.8987"	"83.7429"	"29.884"	"24.8396"	"45.7596"	"28.8927"

Рисунок 21 – Перехресна таблиця метрики PIQE (менше-краще)

Для більшості комбінацій зображень та алгоритмів оцінка виявилася не дуже задовільною. Для деяких зображень є схильність до регулярного отримання кращих оцінок, ніж до інших, які гарних оцінок жодного разу не отримали. Це може бути пов'язано з тим, що погано оцінені зображення й справді мають складну архітектуру, нерозбірливий текст та спотворенні обличчя людей. З іншого боку, зображення, що являють собою портрет людини або тварини, простіші для алгоритмів через краще збереження рис після стискання. Загальний вердикт, що знову дві моделі UltraSharp та Real-ESRGAN розділяють перші місця за оцінками метрики.

У наступній таблиці на рисунку 22 оцінки надаються метрикою шуму PSNR з посиланням на еталон.

PSNR	4x-NMKD-Superscale-SP	4x-OLDIES	4x-UltraSharp.pth	4x-cat-patch	anime-esrgan	real-esrgan
"banner"	"18.9737"	"17.4177"	"18.9532"	"19.8393"	"18.5377"	"18.7788"
"delmonico"	"18.6263"	"17.3523"	"17.9243"	"19.4055"	"18.3785"	"17.5894"
"dry"	"15.7814"	"15.2005"	"16.6849"	"16.9318"	"15.8278"	"15.7118"
"einstein"	"23.1749"	"20.159"	"25.2412"	"25.1581"	"24.5225"	"24.867"
"faces"	"24.4614"	"18.9116"	"22.816"	"25.4299"	"20.5167"	"21.6482"
"landscape"	"21.3771"	"19.612"	"22.152"	"23.4866"	"22.2059"	"20.5997"
"manul"	"21.6315"	"19.8019"	"22.5715"	"24.0021"	"22.4532"	"22.1124"
"patron"	"25.42"	"20.8379"	"25.0282"	"27.13"	"24.9186"	"25.7616"
"ukr"	"24.483"	"19.6364"	"23.6863"	"25.2128"	"22.9651"	"23.8323"

Рисунок 22 – Перехресна таблиця метрики PSNR

Одразу спостерігається велика перевага по оцінюванню останніх шести зображень, однак оцінки не так сильно і відрізняються друг від друга, що може свідчити про відсутність шуму як такого. Можна зауважити, що найвищу сукупну

оцінку отримала вузькоспеціалізована модель cat-patch, хоча вона не призначена для масштабування більшості представлених зображень, тому скоріш ця модель краще виключно тільки при маскуванні шумів.

І остання таблиця – рисунок 23. Маємо в наявності метрики структурної схожості з еталоном SSIM. Одразу перейдемо до 4-тої моделі cat-patch, яка на кожному зображенні методично випереджує усі інші моделі та алгоритми з помітним відривом.

SSIM	4x-NMKD-Superscale-SP	4x-OLDIES	4x-UltraSharp.pth	4x-cat-patch	anime-esrgan	real-esrgan
"banner"	"0.60851"	"0.55146"	"0.61897"	"0.65629"	"0.60806"	"0.6288"
"delmonico"	"0.65959"	"0.61406"	"0.62664"	"0.6702"	"0.62758"	"0.59881"
"dry"	"0.47752"	"0.35896"	"0.48991"	"0.48586"	"0.46893"	"0.47491"
"einstein"	"0.50321"	"0.46376"	"0.61601"	"0.65096"	"0.62252"	"0.63148"
"faces"	"0.87642"	"0.78379"	"0.86394"	"0.90231"	"0.79734"	"0.82206"
"landscape"	"0.73655"	"0.68822"	"0.74986"	"0.79896"	"0.76353"	"0.70317"
"manul"	"0.51177"	"0.45996"	"0.573"	"0.6669"	"0.57512"	"0.56776"
"patron"	"0.85778"	"0.82733"	"0.84981"	"0.90936"	"0.86073"	"0.87682"
"ukr"	"0.87796"	"0.81886"	"0.86671"	"0.89933"	"0.86891"	"0.87933"

Рисунок 23 – Перехресна таблиця метрики SSIM

Щоб розібратись з SSIM результатами наведено мапи відмінностей від еталону на рисунку 24. На даній мапі наносяться темні пікселі, якщо існують відмінності, тому чим мапа світліше, тим краща схожість. Взагалі можна помітити наскільки темніші мапи до зображень manul.jpg та einstein.jpg, що свідчить про загальну складність точного відтворення зображення з маленької роздільної здатності, наприклад, через пишну шерсть або волосся, які втрачаються при стисненні та їх не можливо відтворити з точністю до волосу.

У порівнянні до зображення patron.jpg маємо більш просту задачу по відтворенню, через коротку шерсть, яку і не видно в деталях з такої відставні – тому масштабувати простіше, а мапа SSIM світліше. Якщо тепер переглянути саме мапу для моделі cat-patch у 4-тому стовпчику рисунка 24, то знову помітна менша кількість чорних пікселів. Це у свою чергу має говорити про те, що начебто ця модель найкраще усіх впорюється з усіма зображеннями, хоча насправді дана модель просто дуже консервативна та акуратна при покращенні деталей.

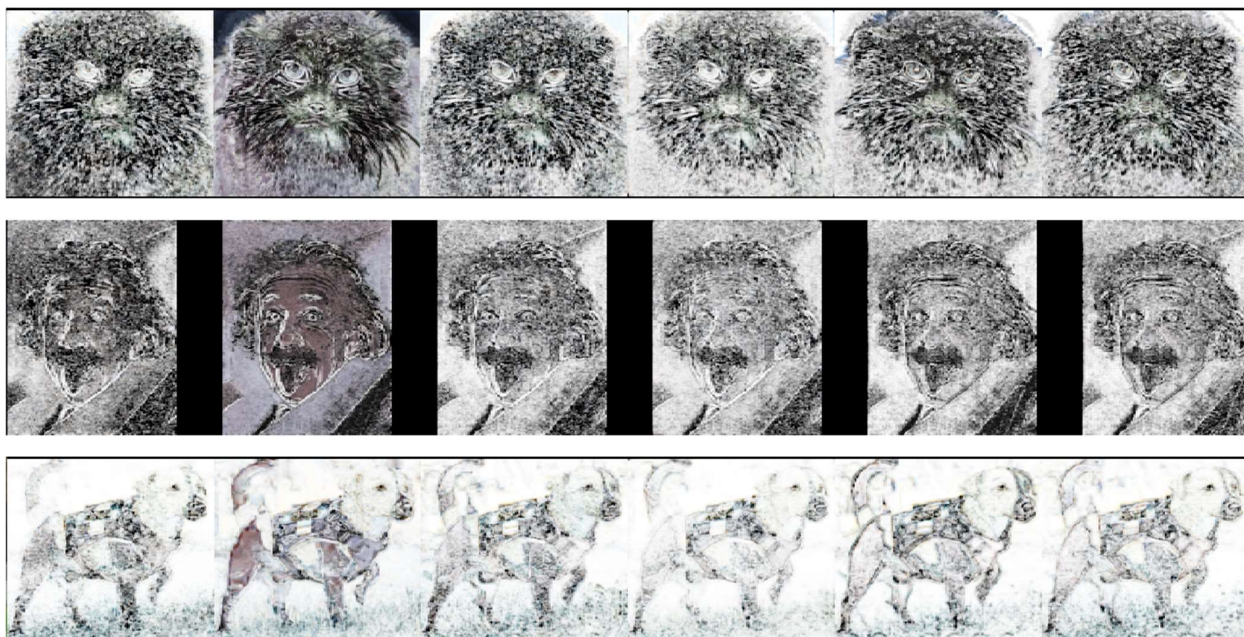


Рисунок 24 – Порівняння SSIM мап зображень манулу, Ейнштейна, собаки для усіх моделей

Якщо потужні реалістичні моделі намагаються відтворити кожну деталь, створюють свої текстури для поверхонь, та й взагалі можуть створити результат з кращою чіткістю ніж був еталон, то дана модель синтезує доволі розмите зображення, але без видимих артефактів та зайвих деталей, тому воно ближче до еталону, але не отримує гарних оцінок в метриках NIQE та PIQE через нечіткість.

Для порівняння моделей зобразимо графіки, обчисливши середнє арифметичне для кожної з метрик. Для NIQE та PIQE метрик з ціллю наочності графіки було перегорнуто, а позитивну частину графіків пофарбовано у зелений, щоб спростити уявлення терміну «менше-краще» (див. рис. 25). Таким чином індекс NIQE для всіх моделей однаково гарний, навіть найневдаліша модель OLDIES не сильно погано себе показує.

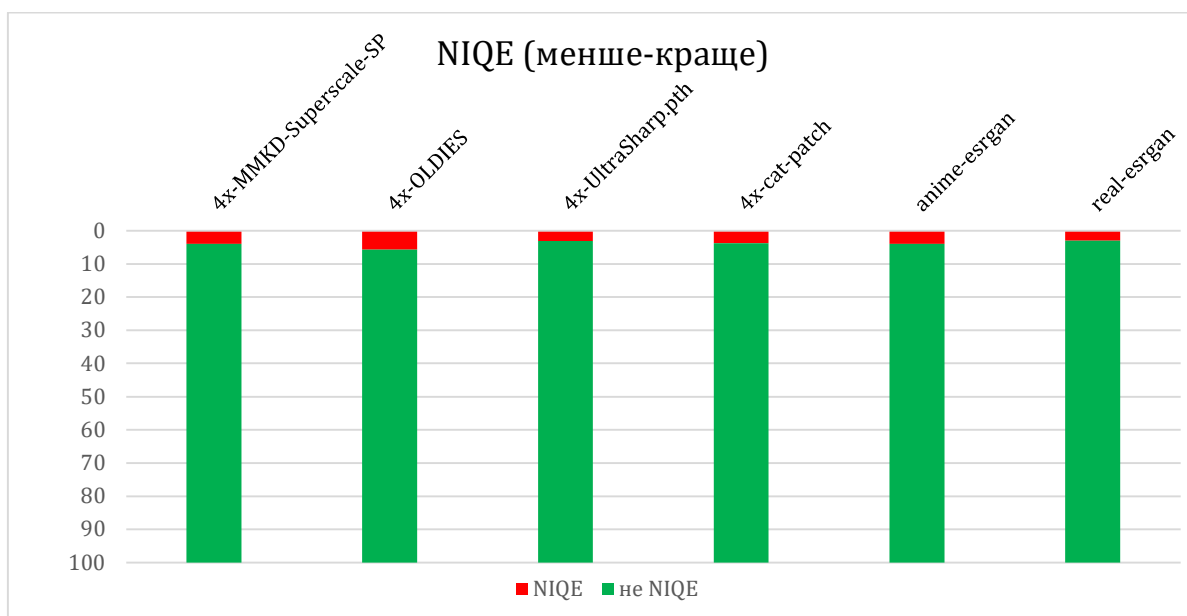


Рисунок 25 – Перегорнута діаграма індексу NIQE серед усіх моделей

На наступному рисунку 26, вже у всіх метрик значно погіршилися результати, але аніме модель OLDIES випередила всіх і стала найгіршою по індексу сприятливості зображення.

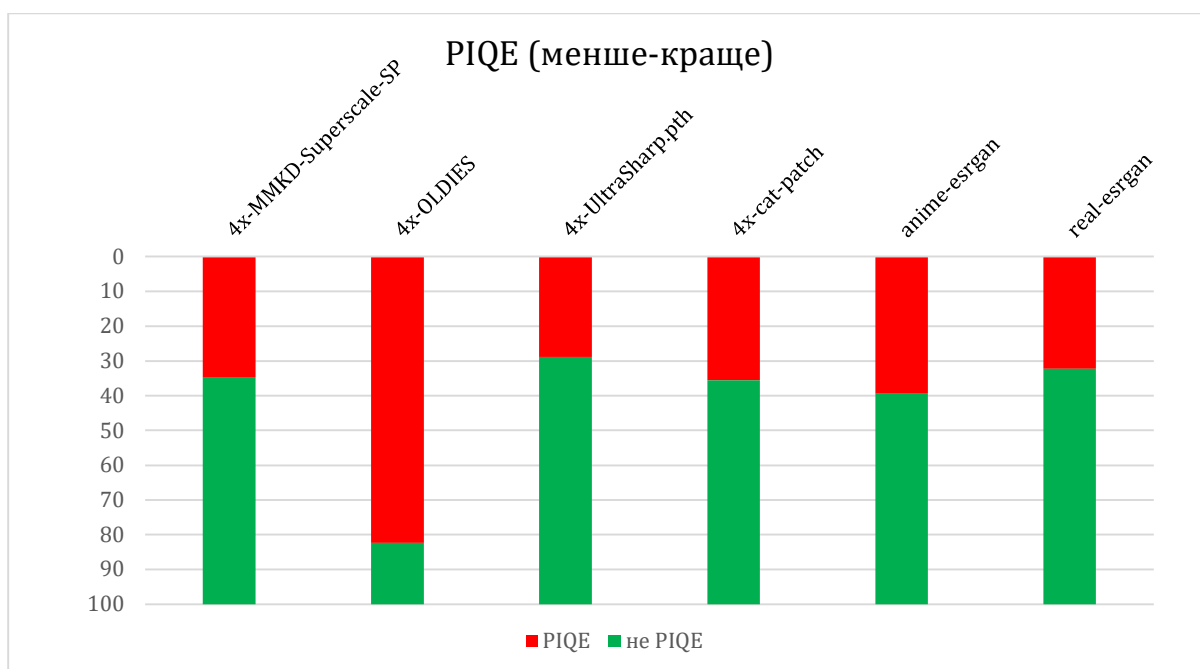


Рисунок 26 – Перегорнута діаграма індексу PIQE серед усіх моделей

Таким чином за цією діаграмою тут можна обрати пару фаворитів – моделі UltraSharp та стандартна модель Real-ESRGAN. На рисунку 27 зображена діаграма порівняння по метриці шуму, а кольором відділені один від одного різні мережі.

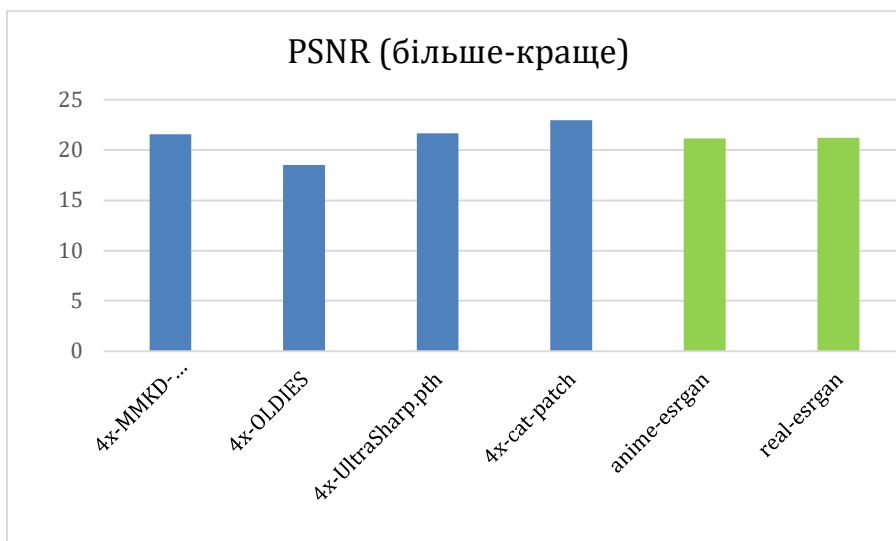


Рисунок 27 – Діаграма порівняння метрики PSNR серед усіх моделей

Маємо невелику перевагу моделі cat-patch над іншими. Остання діаграма метрики SSIM зображена на рисунку 28. Спостерігається відчутна перевага четвертої моделі.

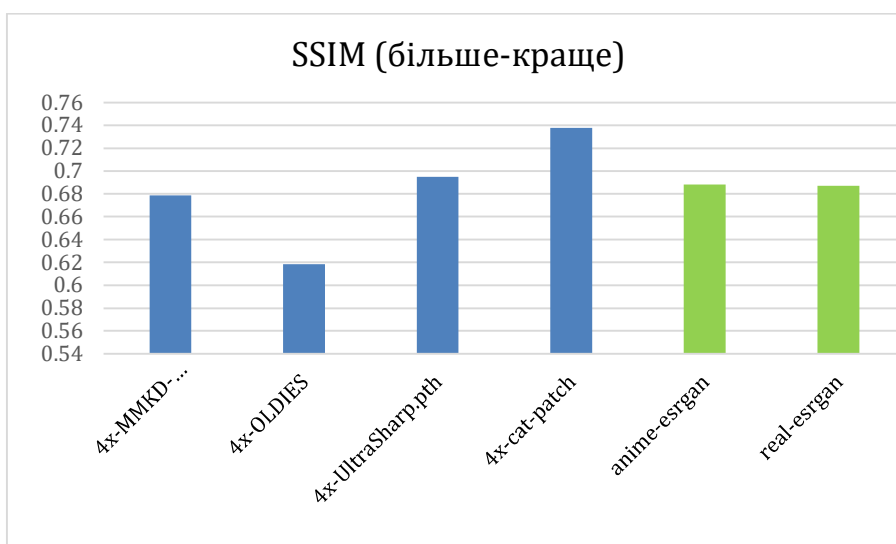


Рисунок 28 – Діаграма порівняння метрики SSIM серед усіх моделей

У наступній таблиці на рисунку 29 зображено узагальнення попередніх чотирьох перехресних таблиць (рис. 19-23) для визначення найкращого універсального алгоритму.

Середнє арифметичне	4x-MMKD-Superscale-SP	4x-OLDIES	4x-UltraSharp	4x-cat-patch	anime-esrgan	real-esrgan
NIQE (min)	4.06856	5.762817	3.1392	3.7235	3.902289	3.021963
PIQE (min)	34.83258	82.29041	28.92643	35.57086	39.48586	32.27211
PSNR (max)	21.5477	18.51178	21.67307	22.95512	21.14733	21.21124
SSIM (max)	0.678812	0.618489	0.694983	0.737797	0.68808	0.687016

Рисунок 29 – Таблиця середнього арифметичного метрик з усіх дев'яти зображень

Зеленим кольором зафарбовані найкращі результати для кожної метрики, а жовтим – другі після найкращих. Таким чином можна обрати фаворитом третю модель UltraSharp алгоритму ESRGAN як найбільш універсальну. Модель cat-sharp може посісти друге місце, хоча з великим застереженням, що вона найкраща тільки для обробки зображень тварин, а усі інші будуть виходити нечіткими, хоча завжди з високими рейтингами схожості до еталону. Real-ESRGAN має право посісти друге місце у рейтингу універсальних алгоритмів.

Однак ніякі метрики не зможуть замінити емпіричний аналіз, якщо переглянемо порівняння масштабованих зображень наочно (див. рис. 30).



Рисунок 30 – Масштабування моделями SuperScale, UltraSharp та Real-ESRGAN

Виявиться, що високі оцінки UltraSharp обумовлені його бажанням наносити текстуру каміння на усі плоскі поверхні, а самі метрики до того ж не розпізнають живої істоти та оцінюють детальну текстуру краще, ніж нечітку шерсть. Отримуємо ситуацію, коли кращій за усіма метриками засіб, виявився найгіршим при візуальній оцінці. Перша модель SuperScale для відновлення зображення зробила набагато більше ніж решта – додала більше шерсті, змогла розпізнати око, однак має погану різкість.

Здебільшого тільки через різкість і отримала дуже слабку оцінку з метрик, а ще через те, що метрики з еталоном негативно оцінюють додавання зайвих деталей до зображення. Остання стандартна модель Real-ESRGAN справилась дещо краще з різкістю, та не додавала зайвих деталей, але все одно програє через те, що в еталоні є багато шерсті. Таким чином зостається UltraSharp, яка має величезну перевагу у різкості, але зовсім не призначена для масштабування живих істот.

## 6 ОПИС РОЗРОБЛЕНОЇ ПРОГРАМНОЇ СИСТЕМИ

Було розроблено програму для аналізу якості масштабування за допомогою існуючих алгоритмів та моделей. Програма складається з двох частин, одна з котрих призначена для безпосереднього застосування алгоритмів, а друга для обчислювання якісних метрик та візуального порівняння результатів.

Перша частина написана мовою програмування Python та має функціонал масштабування зображень за допомогою обраної моделі та алгоритму. Також є можливість навчити свою модель за наявності великого обсягу тренувальних зображень.

Друга частина розроблена на базі кросплатформової обчислювальної платформи MATLAB. При запуску цієї програми будуть перевірені папки на наявність результатів масштабування та оригінальних зображень. Після чого для кожної групи зображень оригінал-результати будуть обчислені якісні метрики NIQE, PIQE та PSNR, SSIM. У сформованому звіті зображення однієї групи будуть розміщені поспіль для наочності та відображені обчислені метрики для кожного зображення, з яких робиться висновок о перевазі одного алгоритму над іншим у окремому сценарії, будуть створені перехресні таблиці для кожної метрики.

Перша частина має назву `upsampling-algorithms`, де один файл викликається командою `python upscale.py <назва моделі>` для процесу масштабування, а інший файл `train.py` потрібен для навчання своїх моделей. Для функціонування потрібен Python 3, з модулями Pytorch та opencv-python.

Друга частина називається `AnalyzeUpscaled`, запускається за допомогою однойменного файлу у форматі `«.msx»`. Для запуску цієї частини достатньо мати доступ до веб-версії MATLAB Online, або встановленого пакету прикладних програм MATLAB.

## 7 РЕЗУЛЬТАТИ РОБОТИ

За результатами дослідження, було проаналізовано 2 методи штучного інтелекту для масштабування із сумарною кількістю моделей 6. Було експериментально порівняно між собою алгоритми масштабування та перевірена доцільність їх використання на різних сценаріях зображень. Було виділено кілька сценаріїв, для яких ці моделі підходять найкраще всього.

Для зображень із шерстяними тваринами підходять одразу дві моделі: SuperScale та Cat-patch. Але на цьому їх подібність і закінчується, тому що вони досить різних масштабів, коли перша це більш універсальна модель для усього, тоді як друга вузькоспеціалізована. Вони також мають різні підходи до реконструкції зображення – перша суттєво видозмінює зображення, що воно далеко відходить від еталону, у спробі додати деталей, тоді як друга навпаки дуже стримувана та результує у найбільш наближеному до реальності зображенню, хоча і розмитому.

Для неживих предметів та архітектури краще підходить модель UltraSharp, утворююча дуже рівні та чіткі грані, додає до усього текстуру, що унеможливує використання на живих істотах.

Для аніме, мультиплікації, тексту підходять однойменна модель з Real-ESRGAN та Oldies, які мають дуже різних підхід до реставрації. Перша модель аніме доволі універсальна для любого стилю анімації, тому що робить тільки дві речі – виводить лінії та спрощує палітру кольорів. Тоді як друга нав'язує старий стиль та жовтий фільтр.

Таким чином залишається лише одна універсальна модель і це стандартна модель від мережі Real-ESRGAN. Вона дуже збалансована - не додає зайвих деталей, результує у достатньо чіткому зображенні. Вона не зможе позмагатися з вузькоспеціалізованими моделями у їх сценаріях, але виконує масштабування достатньо якісно. Також алгоритм дуже ефективно впорюється з шумом та артефактами, хоча спеціалізована для цього метрика PSNR не може оцінити його по гідності, тому що в першу чергу оцінює на відмінність від еталону.

На жаль, в цій роботі ми не проводили експерименти з дуже прогресивним пропрієтарним програмним забезпеченням Toraz GigaPixel AI. Зокрема через закритість авторів щодо деталей роботи алгоритмів, мережі, тренування моделей. Але даний програмний засіб має постійно оновлюванні моделі, пришвидшену обробку зображення, а також унікальну функцію реконструкції обличчя.

Також для покращення дослідження можна було б спробувати створити свою власну модель із зображень, хоча й тренування найпростішої моделі з десяти зображень займе кілька годин.

Всі методи, що використовувались в цій роботі, виявились не здатні реконструювати обличчя малої роздільної здатності. А найпоширеніші метрики виявились вузькоспеціалізованими, та підходять більше для оцінки неживих предметів та реалістичних моделей.

## ВИСНОВКИ

У результаті роботи було проведено аналітичний огляд традиційних методів масштабування та виділено їхні недоліки у порівнянні до методів штучного інтелекту масштабування зображень. Було проаналізовано існуючі методи ШІ, їх переваги та користь при використанні у реальному світі.

Для проведення дослідницької роботи було створено програмну систему, яка отримує на вхід усі масштабовані та еталонні зображення, після чого аналізує існуючі алгоритми за чотирма метриками: індексом природності, індексом сприятливості, співвідношенню сигналу до шуму та структурної схожості. У результаті роботи створює звіт з перехресними таблицями оцінок метрик, генерує мапи структурної схожості SSIM.

За результатами роботи можна сказати, що універсальна модель тільки одна, яка ставиться до всіх зображень нейтрально, не додаючи зайвих деталей, та має солідний рівень різкості. Інші методи більш вузькоспеціалізовані, та краще всього показують себе у належних сценаріях, які і було досліджено.

Всі зазначенні моделі не покривають усіх сценаріїв використання, тому для вдосконалення масштабування на неохопленому сценарії можна натренувати свою модель на специфічних зображеннях.

Для подальшого дослідження можна було б застосувати бібліотеку розпізнавання зображень, щоб програмна система мала змогу передбачувати яку модель треба застосовувати. Ще більш досконаліше була б реалізація розпізнавання одразу усіх сценаріїв на зображенні з точними координатами, що давало б змогу використання конкретних моделей тільки для окремих частин зображення.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Мельнікова Р.В., Нечаєв М.Ю. Дослідження методів штучного інтелекту для збільшення масштабу зображень. Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління. Тези доповідей дванадцятої міжнародної науково-технічної конференції. Том 2: секція 5. 27-28 квітня 2022, м. Харків, Україна, с. 145.
2. He K. Вивчення глибокої згорткової мережі для зображення супер-роздільної здатності / Chao Dong. – Цюрих: ECCV, 2014. – 14 с.
3. Liu Y.-Q. Оцінка узагальнених ядер розмиття за Гаусом для усунення розмитості зображення, що не в фокусі. Транзакції IEEE щодо схем і систем для відеотехнологій [Електронний ресурс] / Xin Du. – 2020. – 15 с. – URL: <https://ieeexplore.ieee.org/document/9079550> (дата звернення: 02.02.2022)
4. Gu J. Слепа супер-роздільна здатність з ітераційною корекцією ядра / Hannan Lu. – Каліфорнія, CVPR, 2019. – 10 с.
5. Tsang S.-H. Review: SRCNN (Super Resolution) [Електронний ресурс] – 2018. – URL: <https://medium.com/coinmonks/review-srcnn-super-resolution-3cb3a4f67a7c> (дата звернення: 08.02.2022)
6. Iriondo R. ESRGAN — Enhanced Super-Resolution Generative Adversarial Networks (Super Resolution & GAN) [Електронний ресурс] – 2021. – URL: <https://towardsai.net/p/machine-learning> (дата звернення: 16.02.2022)
7. Agustsson E. Окреме зображення та відео в режимі реального часу із супер роздільною здатністю з використанням ефективної субпіксельної згорткової нейронної мережі / Timofte Radu. – Гонолулу, CVPRW, 2017. – 12 с.
8. Elad M. Відновлення одного зображення з надвисокою роздільною здатністю з кількох вимірних зображень із розмитими, шумними та недостатньо дискретизованими. Транзакції IEEE щодо обробки зображень [Електронний ресурс] / Feuer Arie. – 1997. – 12 с. – URL: <https://ieeexplore.ieee.org/document/650118> (дата звернення: 24.02.2022)

9. Смеляков К.С., Критерий и метод оценивания качества фильтрации шума. Системы озброєння і військова техніка. 2013, с. 124-126..
10. Fritsche M. Розділення частот для реальної супер-роздільної здатності / Gu Shuhang. – Сеул, ICCVW, 2019. – 10 с.
11. Glasner D. Супер роздільна здатність з одного зображення / Bagon Shai. – Кіото, ICCV, 2009. – 8 с.
12. Kim J. Точна супер-роздільна здатність зображення з використанням дуже глибоких згорткових мереж / Jung Kwon Lee. – Лас Вегас, CVPR, 2016. – 9 с.
13. Zhang K. Вивчення єдиної згорткової мережі із супер роздільною здатністю для кількох деградацій / Wangmeng Zuo. – Солт-Лейк-Сіті, CVPR, 2018 – 10 с.