



(особистий підпис)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ЗАЛІКОВА КНИЖКА № kll-0119
ЗДОБУВАЧА ВИЩОЇ ОСВІТИ

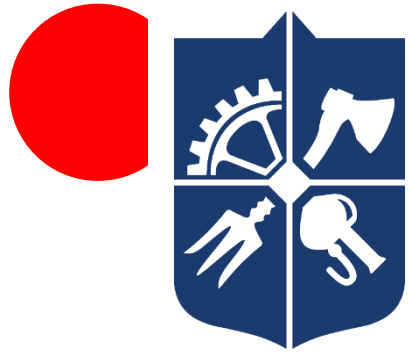
Прізвище, ім'я, по батькові Поліщук Сергійович
Інститут, факультет Інститут математики
Спеціальність ІЗ прикладної математики
Освітня програма Наука про дані та математичне моделювання
Рівень вищої освіти бакалавр
Форма навчання денна
Вступи 6 на 1 курс з 21 09 2016 року

Директор інституту,
декан факультету

30.09.2016
(дата видачі залікової книжки)



Власне ім'я та ПРІЗВИЩЕ
Поліщук Сергій



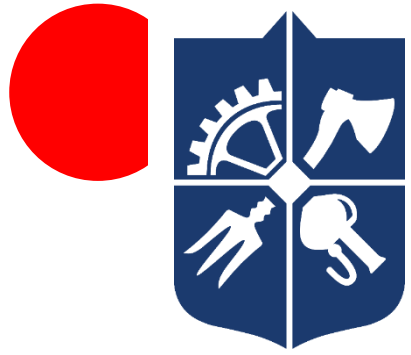
«Математичне та програмне забезпечення чат боту з покращення якості зображення на межах кольорів»

Керівник:
Асистент ПМА ФПМ
Ковальчук-Химюк
Людмила Олександрівна

Виконав:
Студент групи КМ-01
Романецький Микита
Сергійович

Вступ	00-01
Вибір теми. Актуальність	02
Постановка задачі	03-04
Огляд існуючих рішень	05-06
Огляд обраних алгоритмів	07-11

Набір даних	12-16
Навчання моделі	17-19
Верифікація та валідація	20-26
Висновки	27-28
Q&A	29



Вибір теми. Актуальність

1. Розширення можливостей наукових досліджень, наприклад, комп'ютерної томографії, мікроскопії та астрономії;
2. Покращення чіткості зображень сприяє підвищенню безпеки в відеоспостереженні та військовій сфері;
3. Використання у побутових речах.



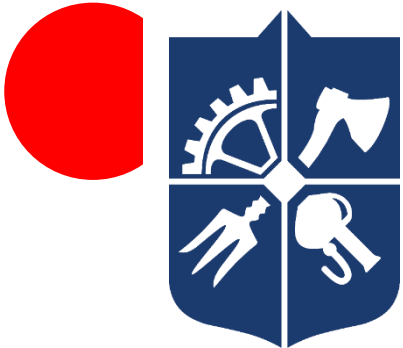
Постановка задачі

Об'єктом дослідження є цифрові зображення, зокрема їхні границі кольорів, де спостерігаються переходи між різними кольоровими областями.

Предметом дослідження є математичне та програмне забезпечення у вигляді чат боту, алгоритми та методи машинного навчання для обробки цифрових зображень, які спрямовані на покращення якості зображень на межах кольорових переходів, включаючи деталізацію та різкість.

Метою цього дослідження є покращення якості зображень на межах кольорів за допомогою відповідних алгоритмів машинного навчання. Покращення деталізації та різкості на межах кольорових переходів. Мінімальне збільшення PSNR має бути не менше за 0.3 dB.

Кінцевим результатом роботи є розробка ефективних алгоритмів машинного навчання для покращення якості зображень на границях кольорів. Розроблений чат бот, який надає користувачам можливість використовувати процес покращення якості зображень.



Порівняння альтернатив для інтерфейсу взаємодії з користувачем

Альтернативи Критерії	Чат бот	Веб-сайт	Мобільний додаток
Складність реалізації	Середня	Висока	Висока
Зручність використання	Висока	Середня	Висока
Доступність	Висока	Висока	Середня
Підтримка платформи	Багато платформний	Багато платформний	Обмежено платформою
Інтерактивність	Висока	Середня	Висока



Огляд існуючих комерційних
рішень покращення зображень
та їх функціоналу

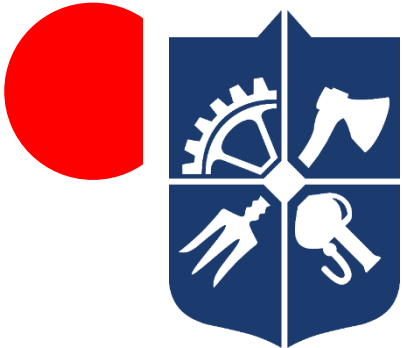


Огляд існуючих комерційних рішень покращення зображень

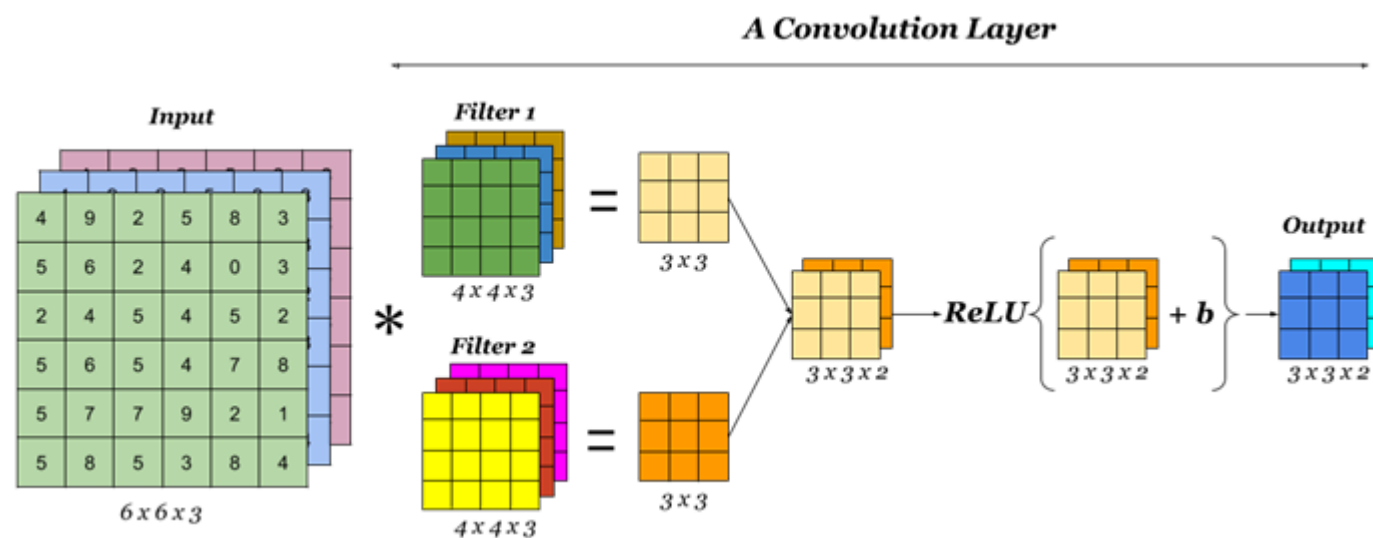
Критерії \ Рішення	Media.io	PicWish.com	Topaz Gigapixel AI	Let's Enhance	Remini
Платформи	Веб	Веб	Десктоп	Веб	Мобільний додаток
Вартість	Безкоштовно / платно	Безкоштовно / платно	Платно	Безкоштовно / платно	Безкоштовно / платно
Якість вихідних зображень	Висока	Середня	Дуже висока	Висока	Висока
Простота використання	Висока	Висока	Середня	Висока	Висока
Швидкість обробки	Висока	Висока	Середня	Висока	Висока
Підтримка форматів	Різноманітні	Обмежені	Різноманітні	Різноманітні	Обмежені

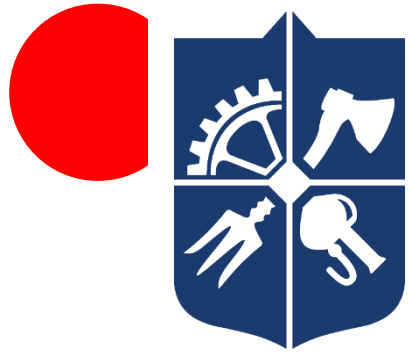


Огляд обраних алгоритмів машинного навчання

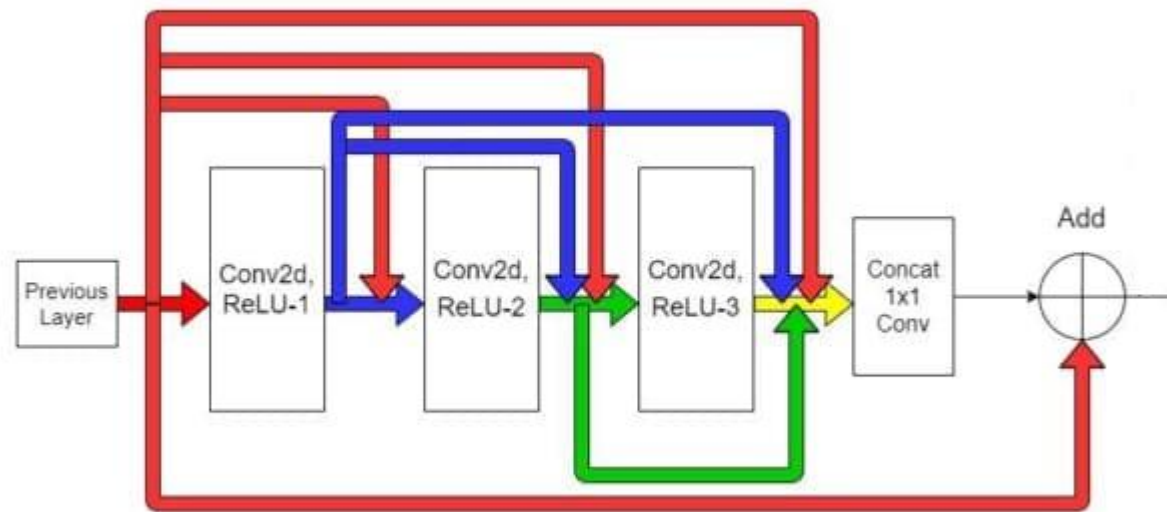


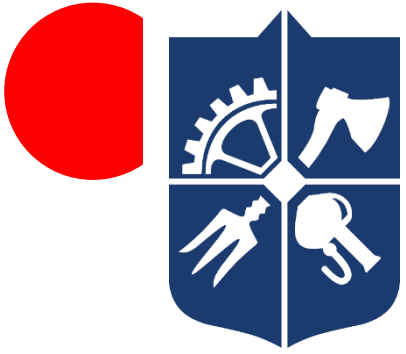
Огляд обраних алгоритмів машинного навчання



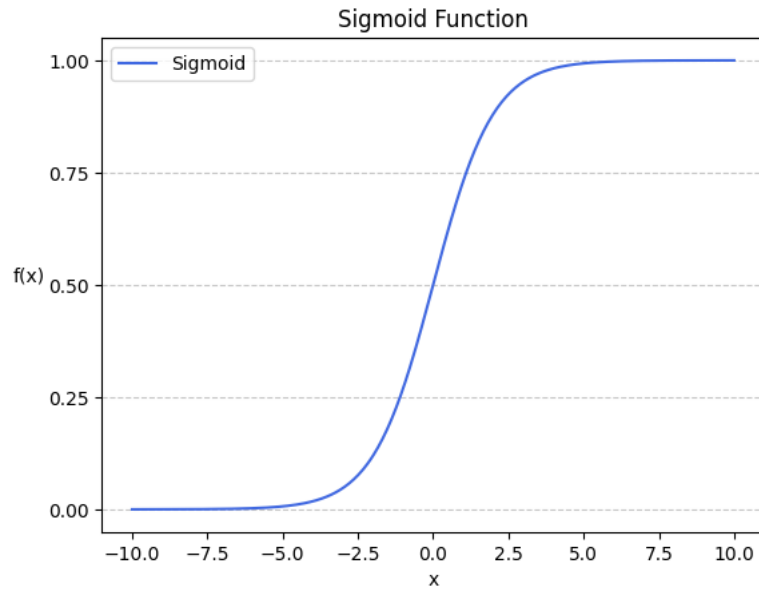


Огляд обраних алгоритмів машинного навчання

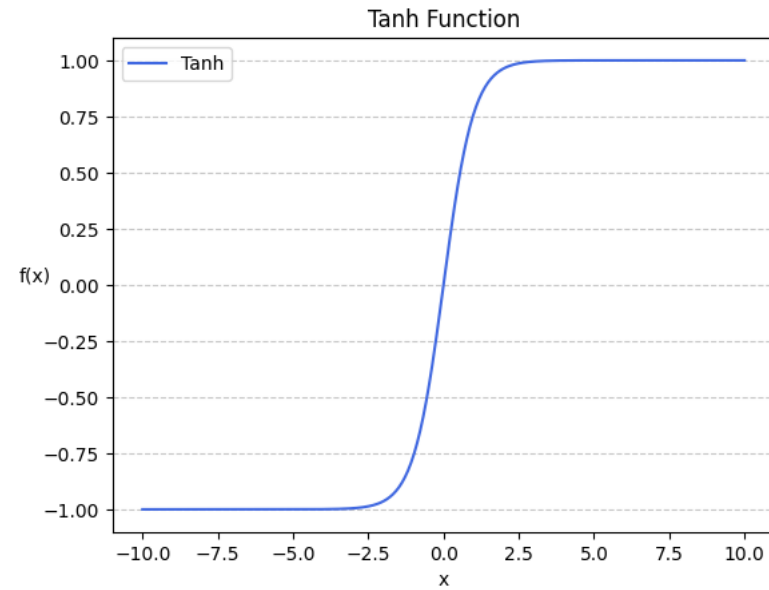




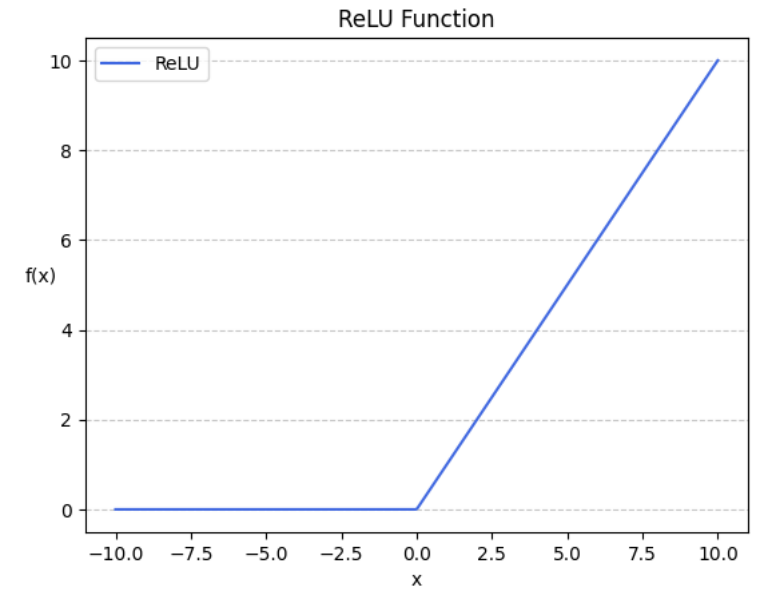
Огляд обраних алгоритмів машинного навчання



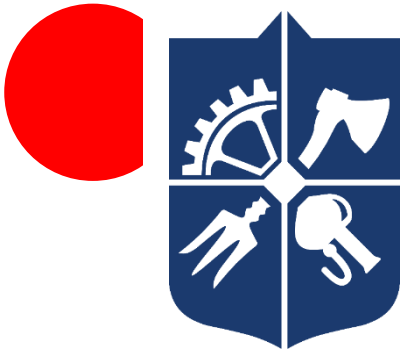
$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}$$



$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



$$\text{ReLU}(x) = \max(0, x)$$



Опис архітектури моделі машинного навчання

```
def Model(channels, upscale_factor):  
    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 * (upscale_factor**2), 3, padding='same', activation='relu', kernel_initializer='Orthogonal')(X)  
  
    outputs = tf.nn.depth_to_space(X, upscale_factor)  
    return keras.Model(inputs, outputs)
```



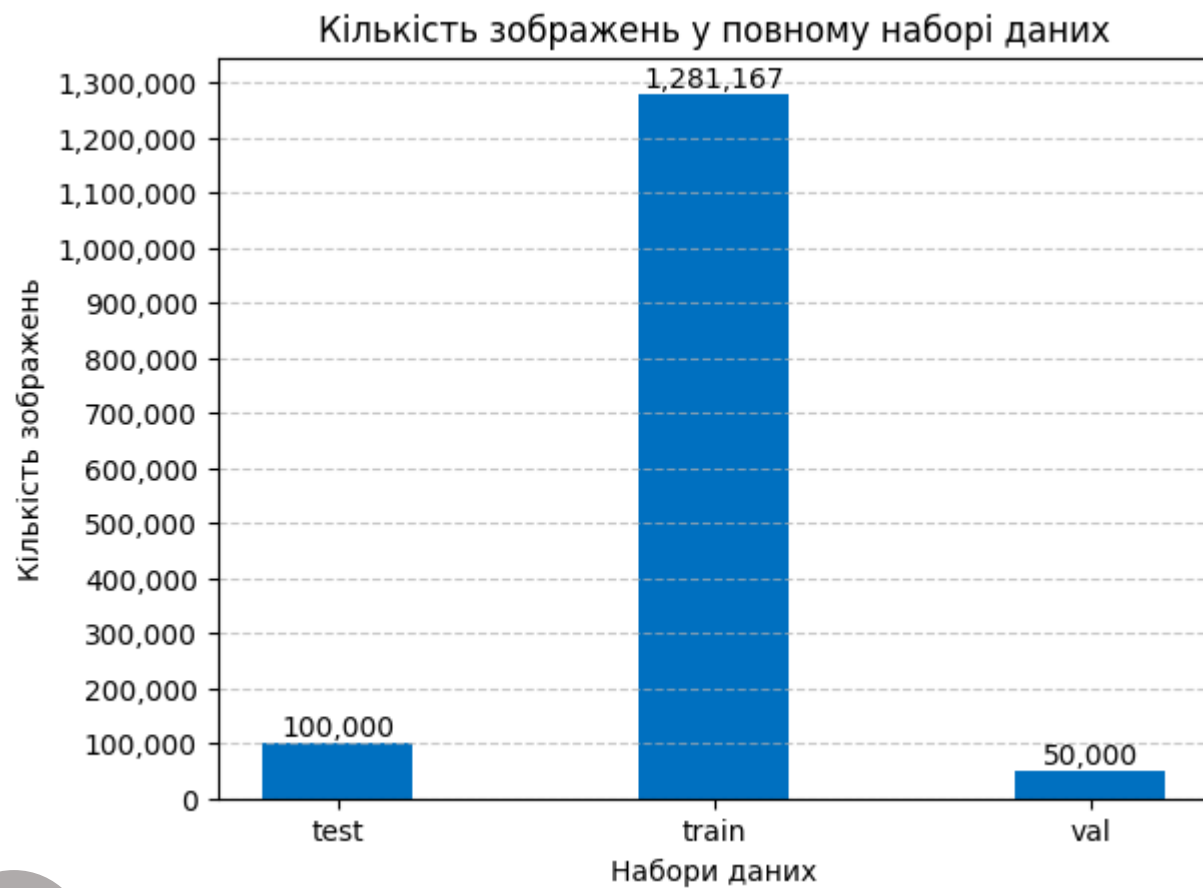
Набір даних



Походження та огляд набору даних

IMAGENET

kaggle



Нормалізація набору даних

Нехай вхідна матриця (A):

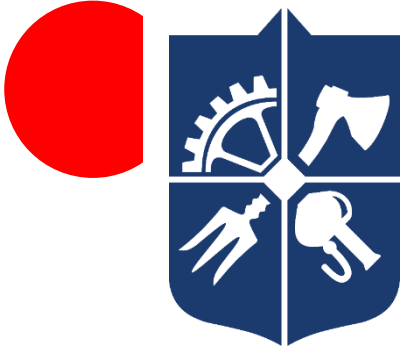
34	87	123
45	190	230
210	12	65

Ділення кожного елементу
матриці A на 255

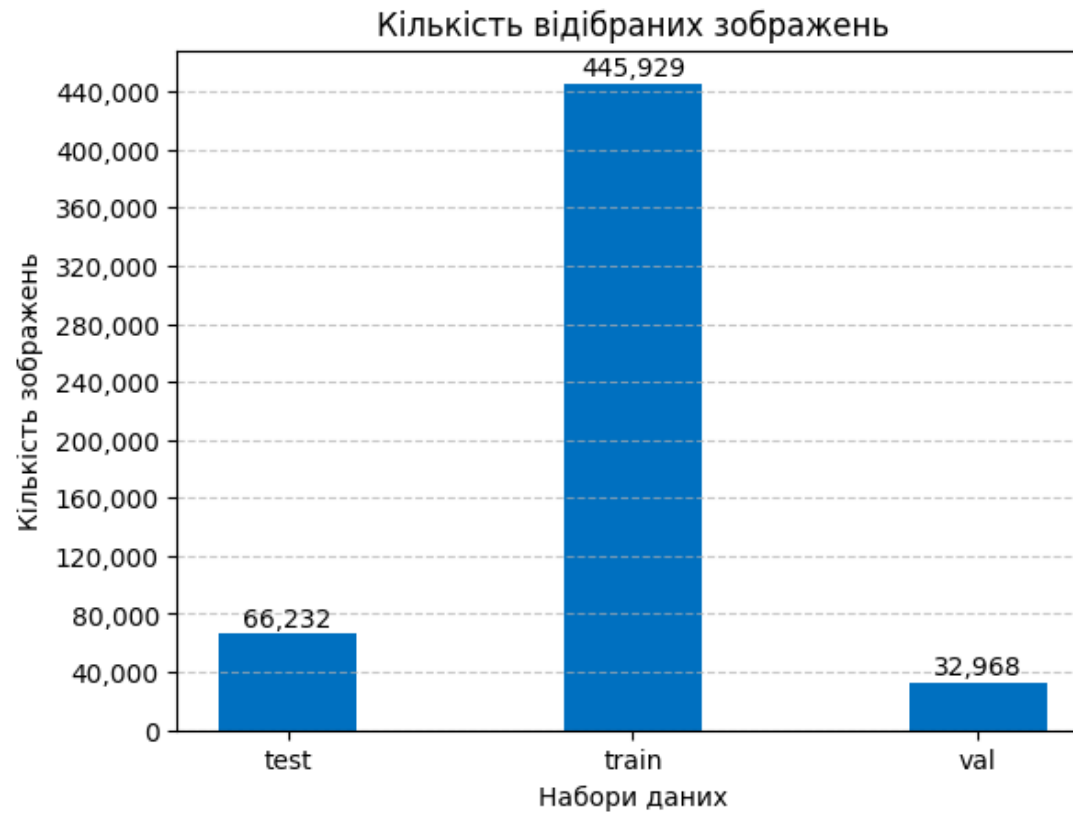
Вихідна матриця (B):

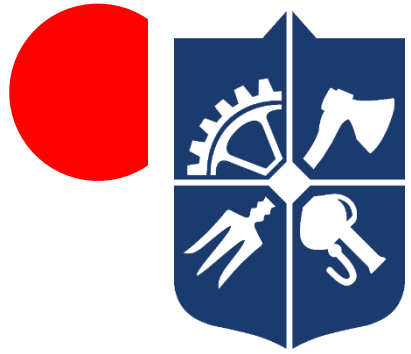
0.13	0.34	0.48
0.18	0.75	0.9
0.82	0.05	0.25

*255 – це максимально можливе значення інтенсивності кольору для 8-ми бітних зображень



Відбір даних





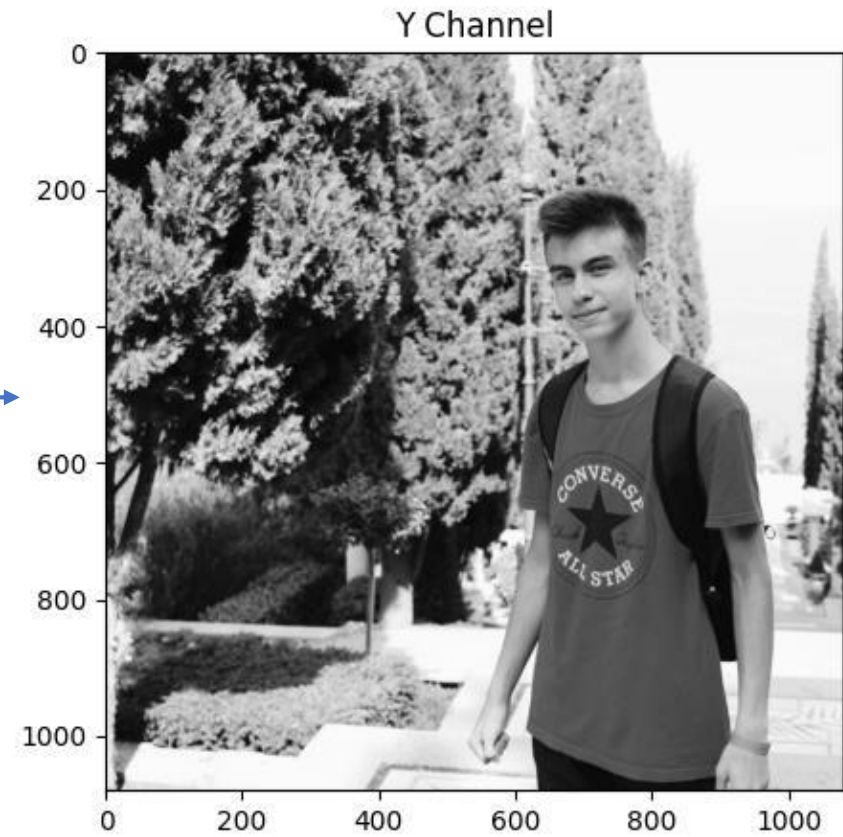
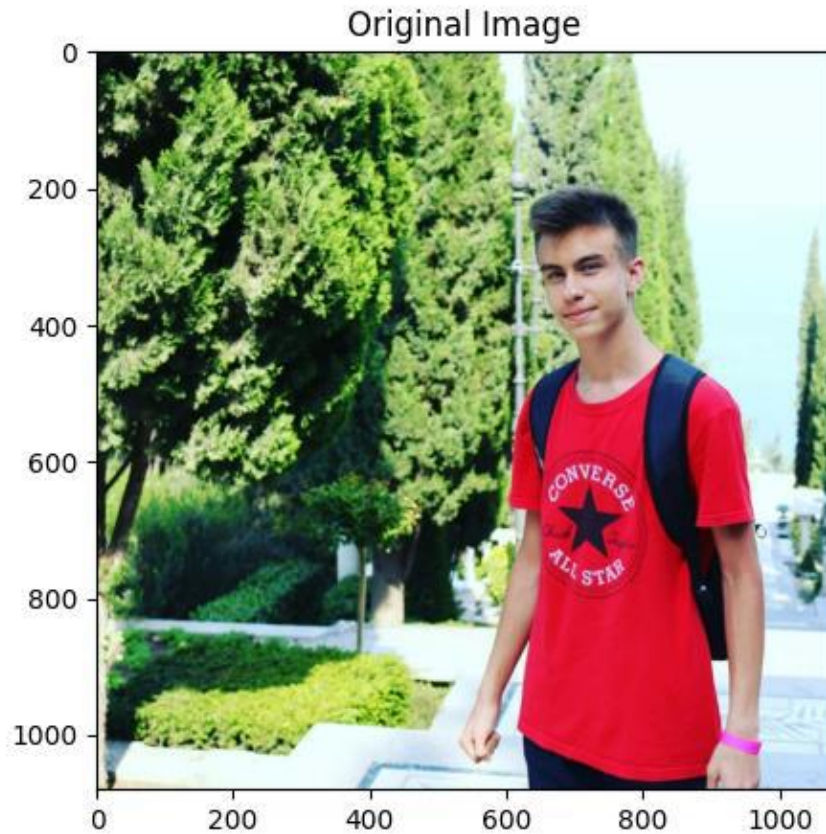
Передобробка набору даних

RGB \rightarrow YUV

$$Y = 0.299 * R + \\ + 0.587 * G + \\ + 0.114 * B$$

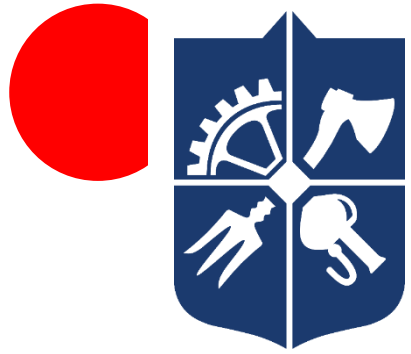
$$U = 0.492 * (B - Y)$$

$$V = 0.877 * (R - Y)$$

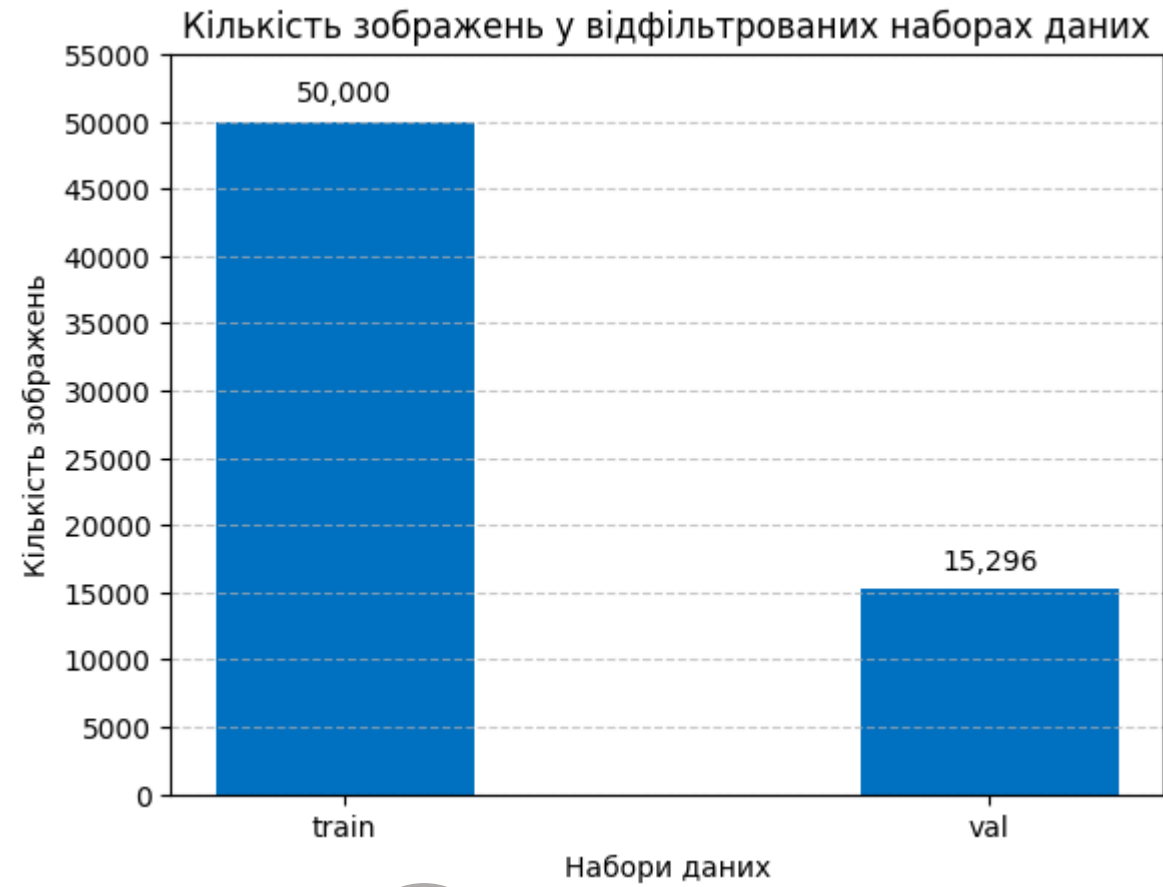


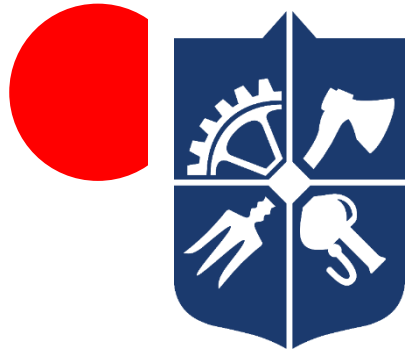


Навчання моделі



Навчання моделі

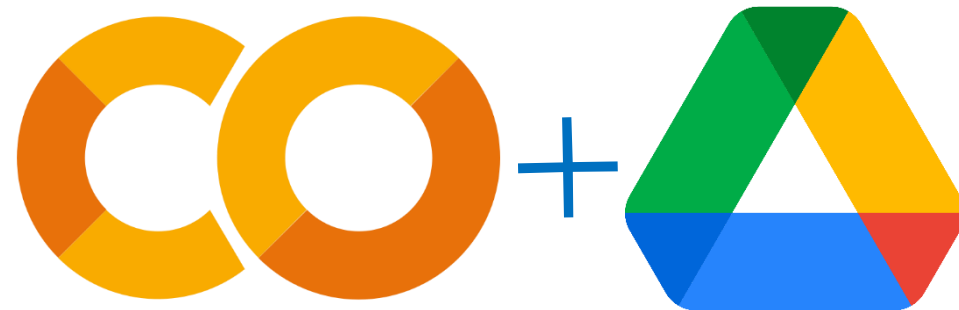




Навчання моделі



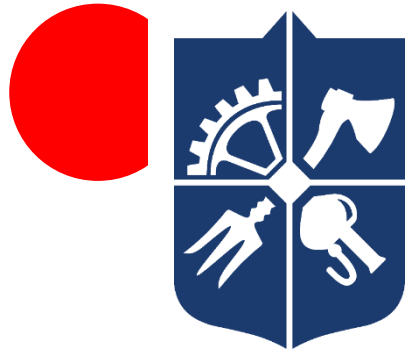
Варіант №1



Варіант №2



Верифікація та валідація



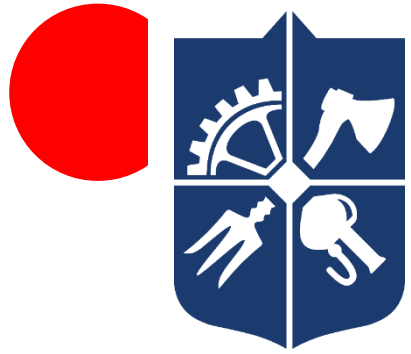
Метрики оцінки результатів

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

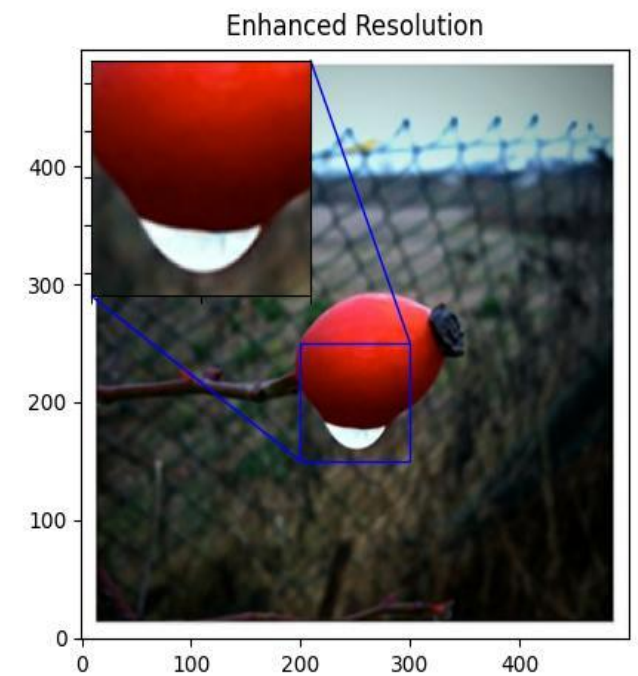
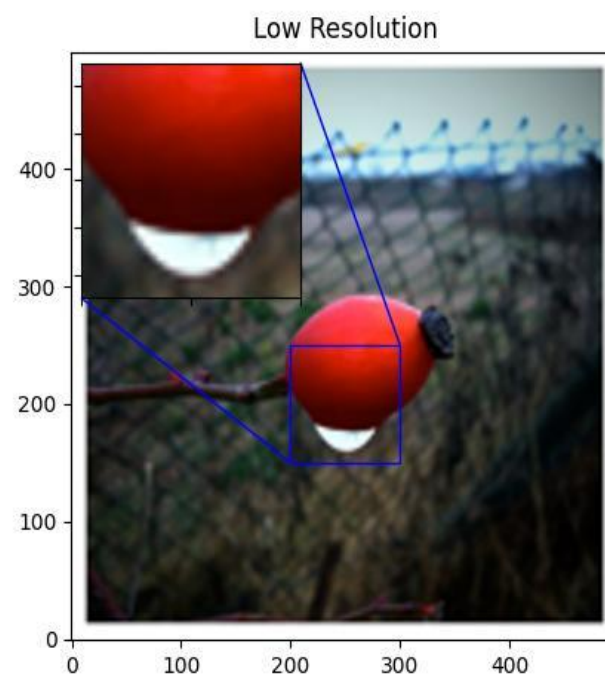
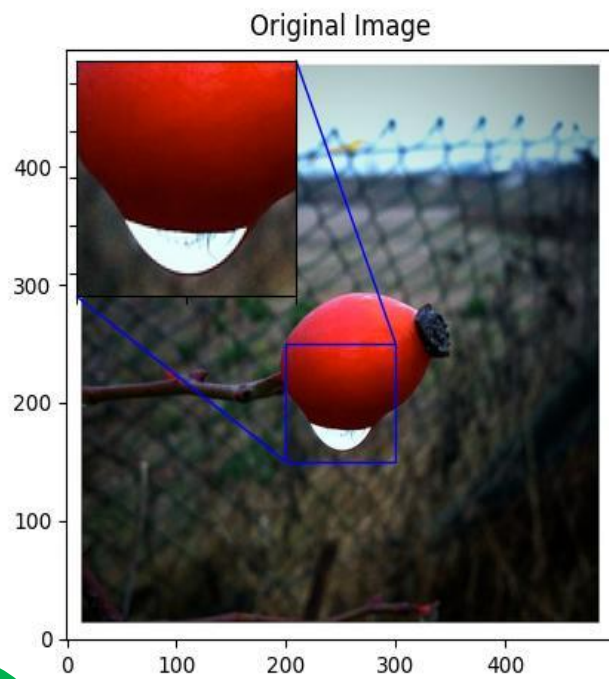
$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I(i,j) - K(i,j))^2$$

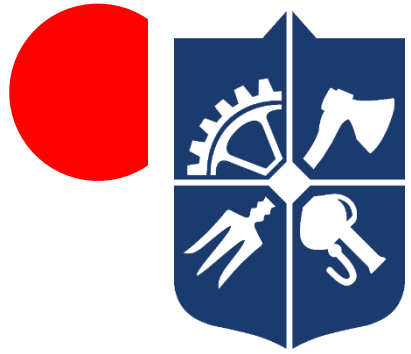
$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |I(i,j) - K(i,j)|$$

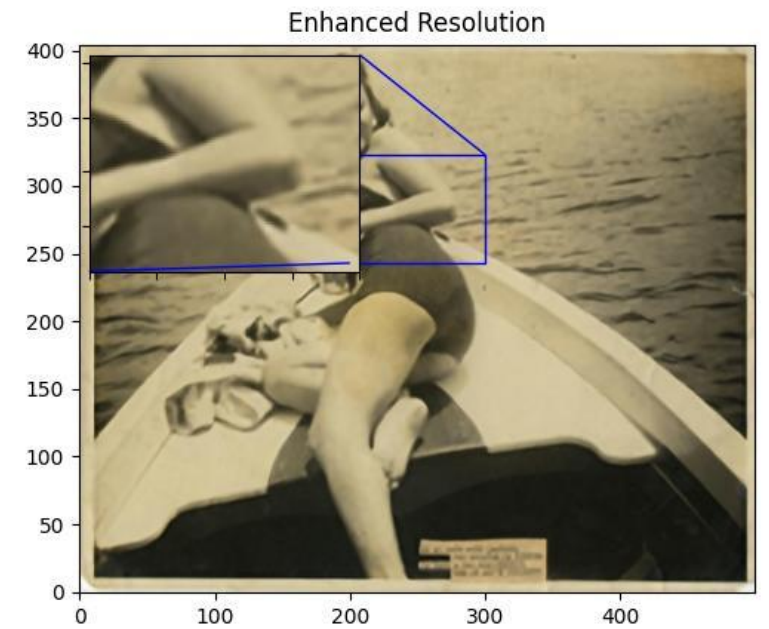
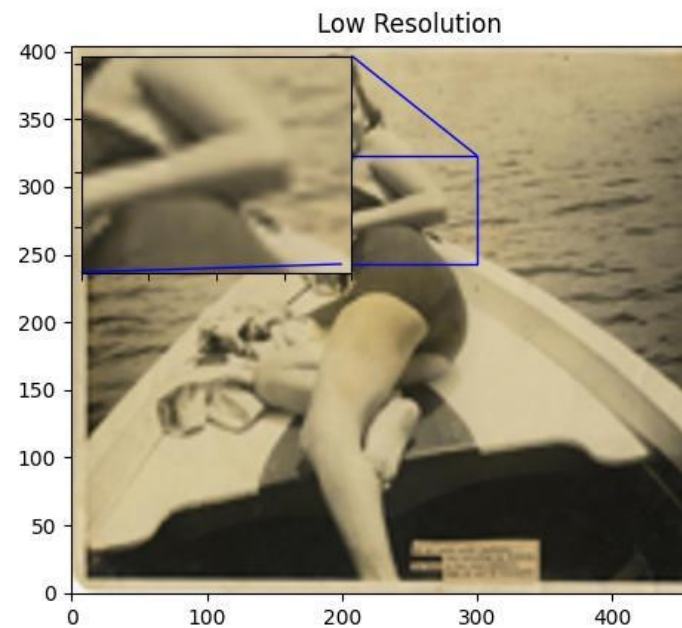
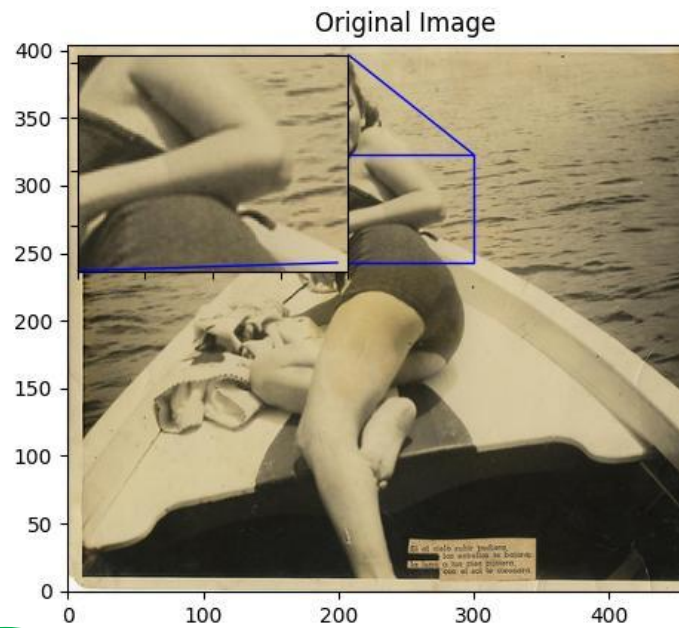


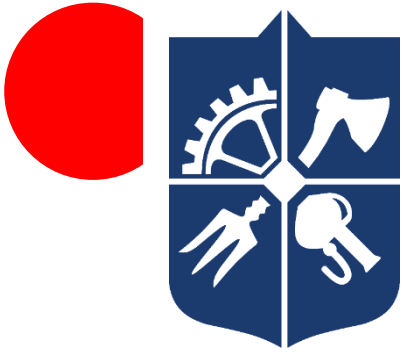
Демонстрація результатів



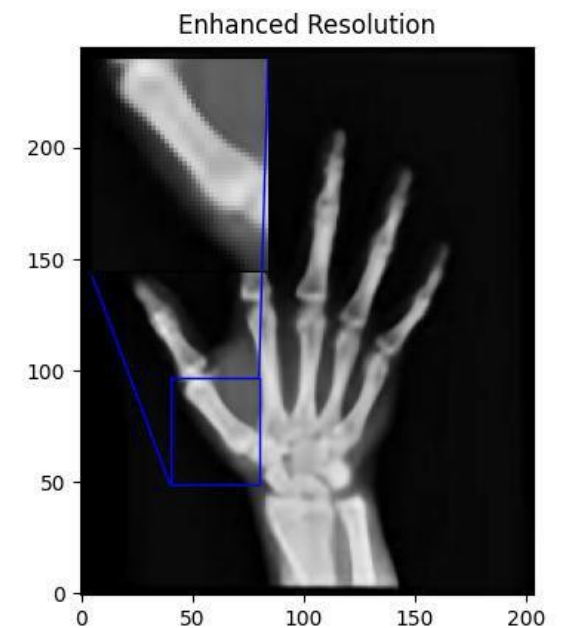
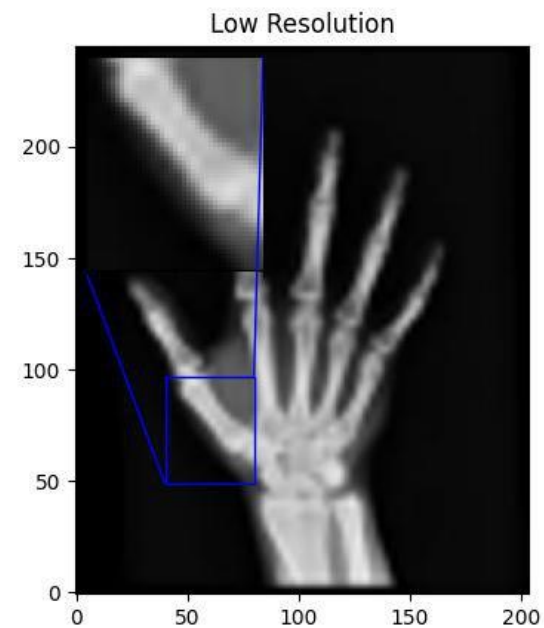
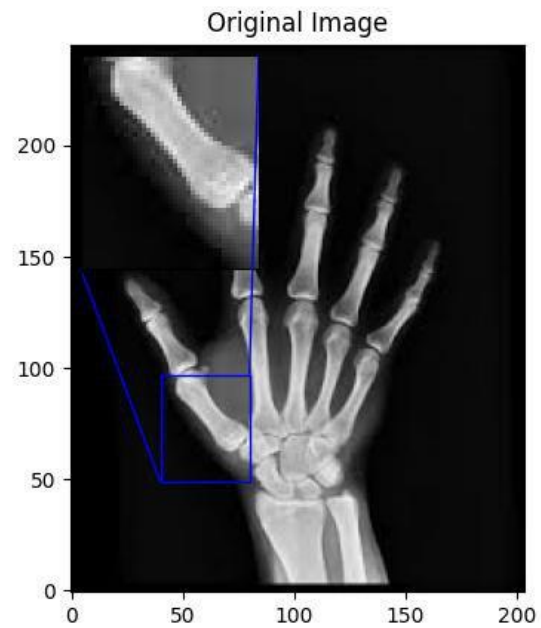


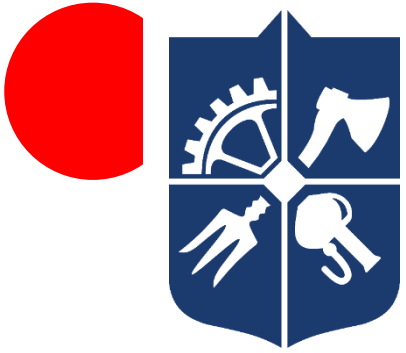
Демонстрація результатів



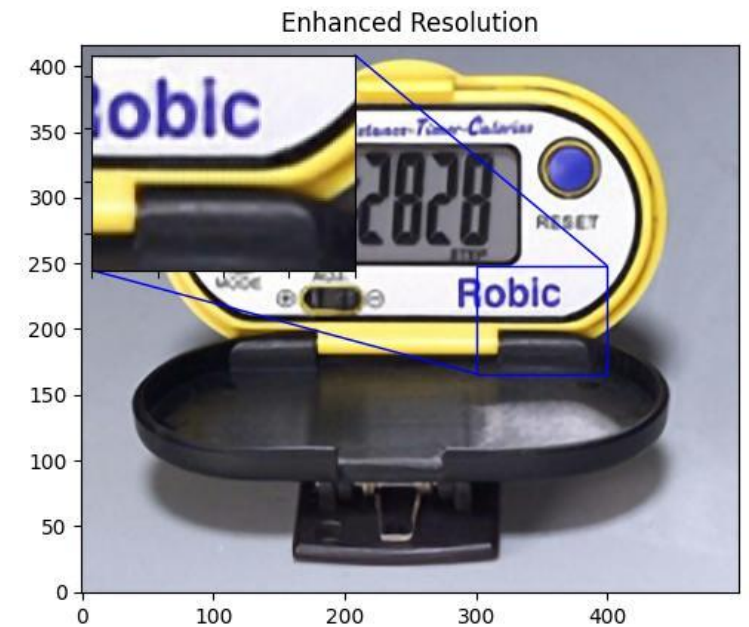
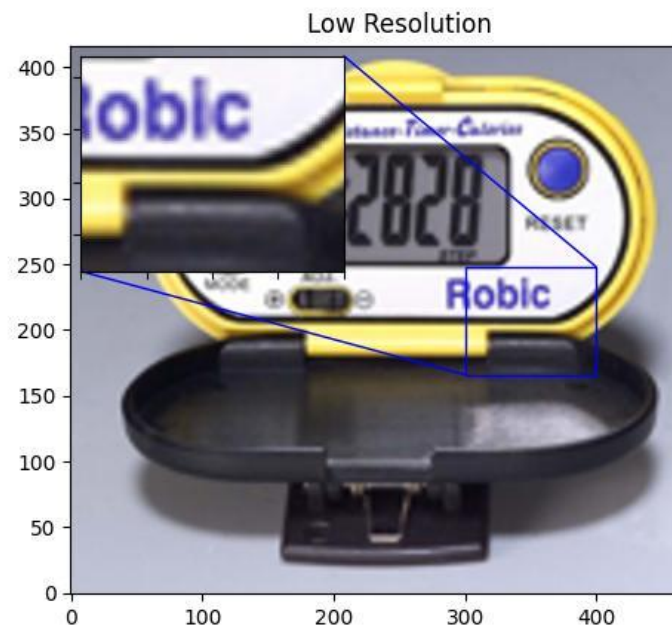
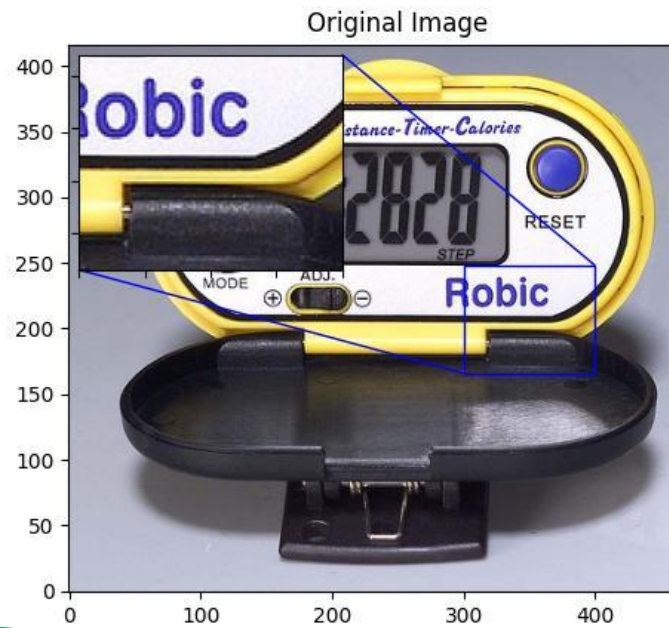


Демонстрація результатів

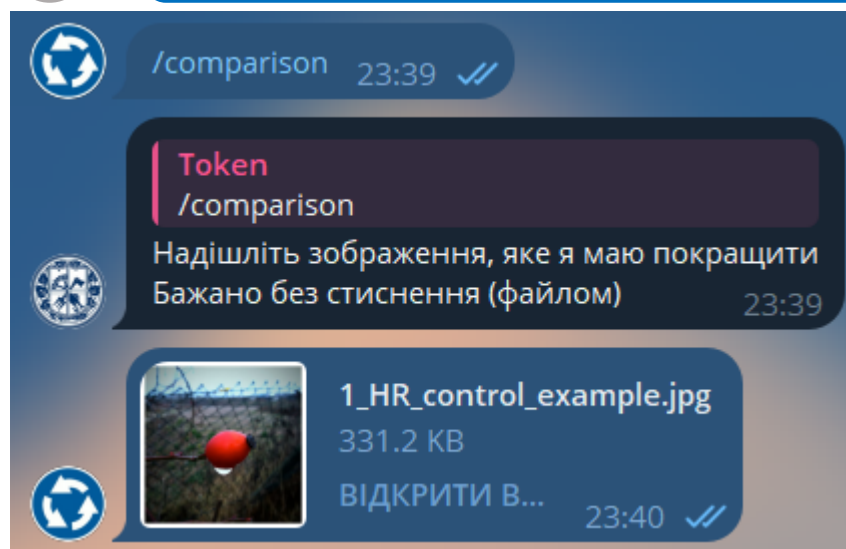




Демонстрація результатів

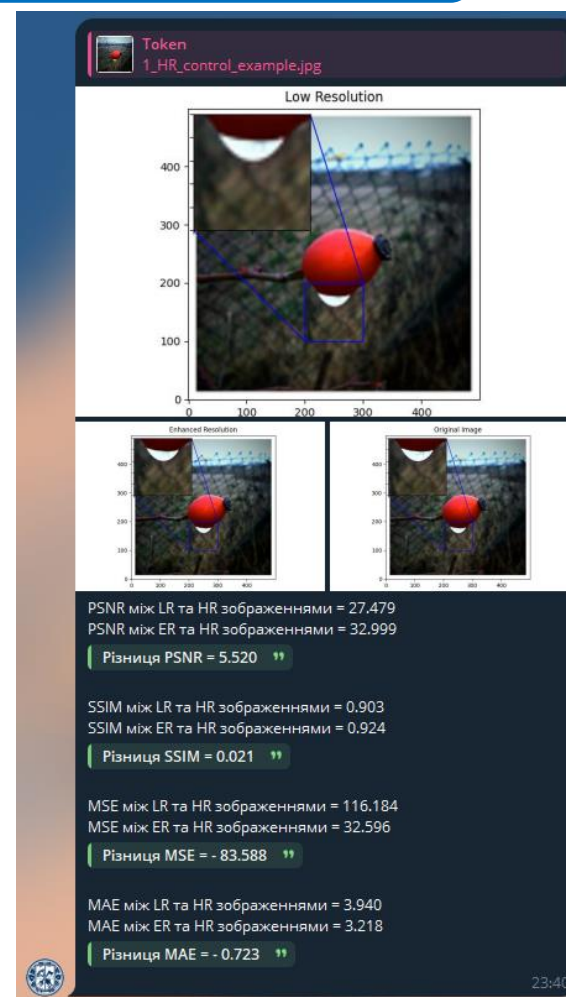


Демонстрація результатів



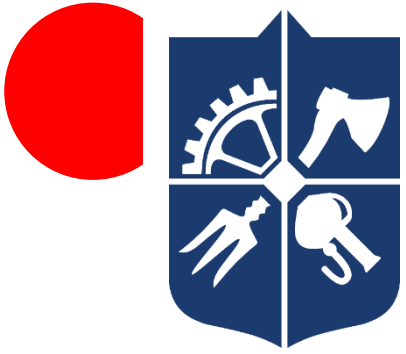
Середній PSNR між LR та HR зображеннями = 25.704
Середній PSNR між ER та HR зображеннями = 26.351

$$26.351 - 25.704 = 0.647 \text{ dB}$$





Висновки



Висновки

У рамках дипломної роботи:

1. Здійснено аналіз наявних комерційних рішень для розв'язку поставленої задачі;
2. Було розглянуто алгоритми реалізації задачі покращення якості зображень. Серед алгоритмів на основі виконаного порівняльного аналізу було обрано метод згорткової нейронної мережі;
3. Спроектовано математичне забезпечення та підібрано архітектуру моделі;
4. Навчено власну модель машинного навчання, яка здатна покращити якість зображення в середньому на 0.647 dB;
5. Розроблено чат бота для взаємодії користувача та моделі.



Q&A