

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ**

Факультет автоматизації і інформаційних технологій

Кафедра інформаційних технологій

**ПОЯСНЮВАЛЬНА ЗАПИСКА
ДО КВАЛІФІКАЦІЙНОЇ ВИПУСКНОЇ РОБОТИ
НА ЗДОБУТТЯ ОСВІТНЬОГО СТУПЕНЯ МАГІСТР**

на тему:

Розробка GAN моделі для підвищення якості зображень

Осокін Арсен Станіславович

(прізвище, ім'я та по батькові здобувача повністю)

Київ 2024 р.

Презентація Кваліфікаційної випускної роботи



Кваліфікаційна випускна робота

РОЗРОБКА GAN МОДЕЛІ ДЛЯ ПІДВИЩЕННЯ ЯКОСТІ ЗОБРАЖЕНЬ

ВИКОНАВ Студент КНм-23 Осокін А.С. КЕРІВНИК к.т.н., доц. Горда О.В.

Слайд 1 – Розробка GAN моделі для підвищення якості зображень

АКТУАЛЬНІСТЬ ТА МЕТА

- На сьогоднішній день проблема підвищення якості зображень є критично важливою для багатьох професійних сфер застосування. В умовах постійно зростаючих вимог до якості візуального контенту та поширення пристроїв з екранами високої роздільної здатності, завдання ефективного масштабування та покращення якості зображень набуває особливої актуальності.
- Методи глибокого навчання, зокрема генеративні змагальні мережі (**GAN**), відкривають нові можливості для вирішення цієї задачі, демонструючи значний потенціал у відновленні дрібних деталей та текстур при збільшенні роздільної здатності зображень.
- Метою роботи є розробка та тренування власної **GAN** моделі для підвищення якості зображень з одночасним збільшенням їх роздільної здатності. Особлива увага приділяється оптимізації архітектури моделі для забезпечення ефективного відновлення деталей при збереженні реалістичності результатів.

2

Слайд 2 – Актуальність та мета

ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

- Виходячи з поставленої мети, об'єктом дослідження визначено процес підвищення якості цифрових зображень з використанням методів машинного навчання. Методи глибокого навчання, зокрема генеративні змагальні мережі (**GAN**), відкривають нові можливості для вирішення цієї задачі, демонструючи значний потенціал у відновленні дрібних деталей та текстур при збільшенні роздільної здатності зображень.
- В якості предмету дослідження обрано методи та алгоритми на основі генеративних змагальних мереж, як найбільш перспективного напрямку вирішення поставленої задачі.
- Методологічною основою дослідження виступає комбінація онтологічного підходу до аналізу предметної області та експериментального методу при розробці системи. Онтологічний підхід дозволив створити формальну специфікацію концептуалізації в області обробки зображень, що стало основою для подальшого проектування системи. Експериментальна складова включала серію досліджень з різними параметрами моделі для визначення оптимальної конфігурації.
- Додатково в роботі були використані методи системного аналізу для декомпозиції проблеми та визначення критеріїв якості, а також статистичні методи для оцінки результатів експериментів. Це забезпечило комплексний підхід до вирішення поставленої задачі.

3

Слайд 3 – Об'єкт, предмет та методи дослідження

ДЕРЕВО ЦІЛЕЙ



- Базуючись на визначених об'єкті, предметі та методах дослідження, а також враховуючи актуальність проблеми, було сформовано детальне дерево цілей проекту.
- Кожен етап має основну ціль, яка безумовно підпорядкована головній меті, а також підцілі - невеликі завдання, виконання яких разом приводить до досягнення результату. Ієрархічний спуск відображає розкладання цілей на прості завдання, які їх складають.

4

Слайд 4 – Дерево цілей

АНАЛІЗ ІСНУЮЧИХ МЕТОДІВ

➤ Проведений аналіз показав, що традиційні методи, такі як бікубічна інтерполяція та різні види фільтрації, мають суттєві обмеження при відновленні дрібних деталей зображення. Сучасні підходи на основі нейронних мереж демонструють значно кращі результати, особливо методи на базі **GAN**.

➤ Нижче наведено приклади приклади їхньої роботи :



➤ Оригінал



➤ Бікубічна інтерполяція



➤ Глибокі згорткові нейронні мережі (CNN)



➤ Метод найближчого сусіда



➤ Інтерполяція на основі перетворення Фур'є



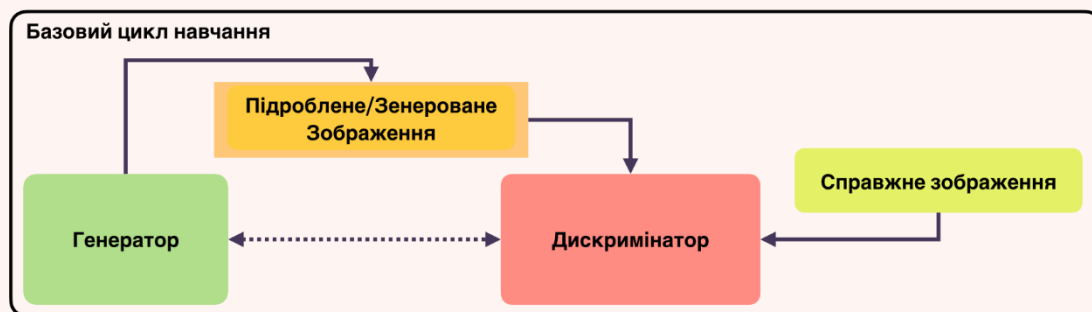
➤ Глибокі згорткові нейронні мережі з використанням перцептивної втрати (GAN)

5

Слайд 5 – Аналіз існуючих методів

ЩО ТАКЕ GAN ?

➤ Генеративно-змагальна мережа (**GAN**) - це архітектура, де дві нейромережі змагаються між собою. Генератор створює підроблені зображення. Дискримінатор намагається відрізнити підроблені зображення від справжніх. У процесі такого змагання обидві мережі постійно вдосконалюється, навчаючись створювати все більш реалістичні зображення.

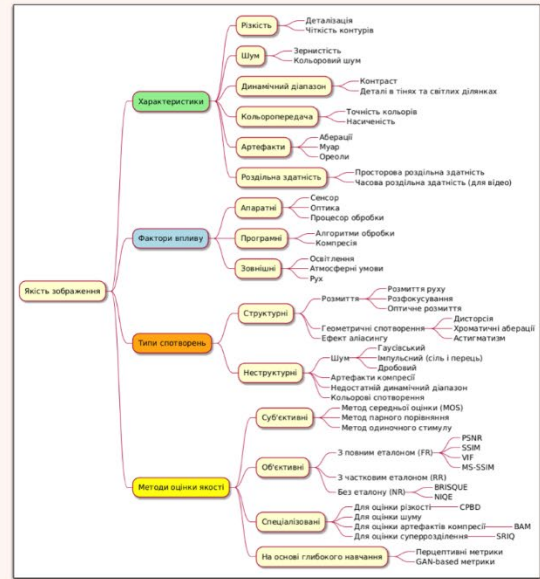


6

Слайд 6 – Що таке GAN ?

ОНТОЛОГІЧНИЙ ПІДХІД ДО ПОНЯТТЯ ЗОБРАЖЕННЯ

➤ Для систематизації знань про якість зображень та факторів впливу на неї було розроблено комплексну онтологію. Вона включає чотири ключові компоненти: характеристики зображень, фактори впливу, типи спотворень та методи оцінки якості. Такий структурований підхід дозволив чітко визначити всі аспекти, які необхідно врахувати при розробці системи. Створена онтологія стала фундаментом для подальшого проектування моделі та вибору методів оцінки.

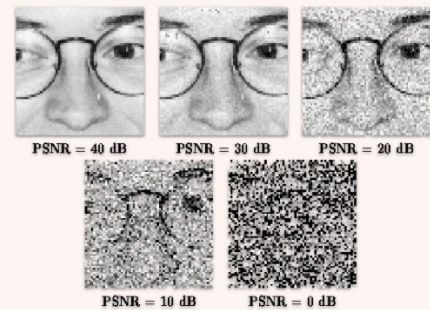
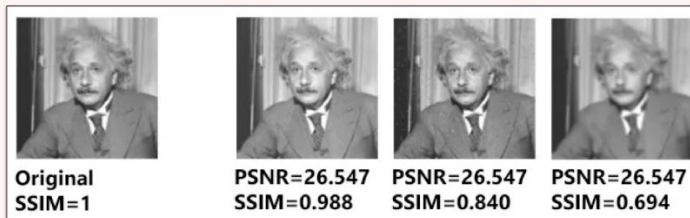


7

Слайд 7 – Онтологічний підхід до поняття зображення

ВИЗНАЧЕННЯ КРИТЕРІЇВ ОЦІНКИ

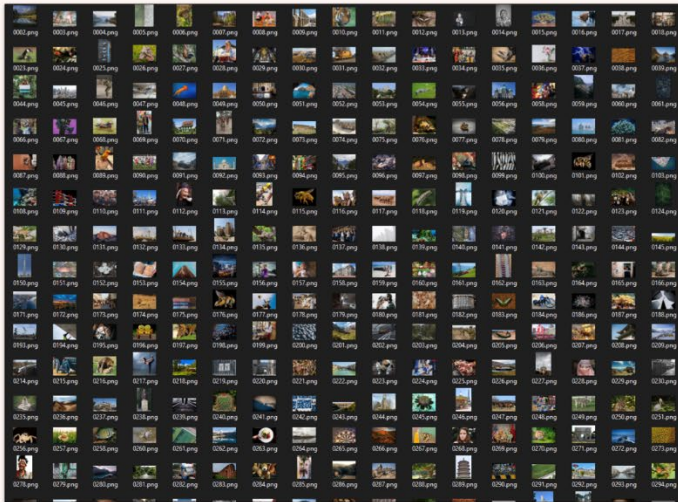
➤ На основі розробленої онтології визначено ключові критерії оцінки якості роботи системи. Основними метриками є **PSNR (Peak Signal-to-Noise Ratio)** та **SSIM (Structural Similarity Index)**, які дозволяють об'єктивно оцінювати якість результатів. Крім того, важливим критерієм є візуальне сприйняття, оскільки воно визначає реальну якість зображення з точки зору користувача.



8

Слайд 8 – Визначення критеріїв оцінки

ОБРАННЯ НАБОРУ ДАНИХ ДЛЯ НАВЧАННЯ



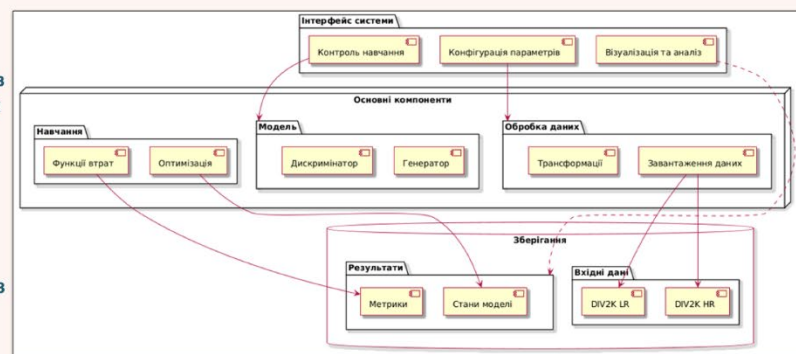
► На основі визначених цілей та критеріїв було обрано набір даних **DIV2K**, який містить високоякісні зображення для тренування моделі. Цей набір даних включає **800** зображень високої роздільної здатності разом з їх версіями нижчої якості. Важливою перевагою **DIV2K** є різноманітність контенту та типів зображень, що дозволяє моделі навчитися працювати з різними сценаріями. Для навчання дані було розділено на тренувальну та валідаційну вибірки у співвідношенні **80/20**.

9

Слайд 9 – Обрання набору даних для навчання

► На основі проведеного аналізу існуючих рішень та результатів онтологічного підходу, для нашої системи було обрано та адаптовано архітектуру **Real-ESRGAN Compact (SRVGGNet)**. Вибір цієї архітектури обумовлений її оптимальним балансом між якістю результатів та вимогами до обчислювальних ресурсів. В процесі адаптації особлива увага була приділена оптимізації компонентів для роботи в умовах обмежених ресурсів, зберігаючи при цьому високу якість генерації. Імплементація моделі виконана з використанням фреймворку **PyTorch**, що забезпечило необхідну гнучкість у налаштуванні та оптимізації архітектури.

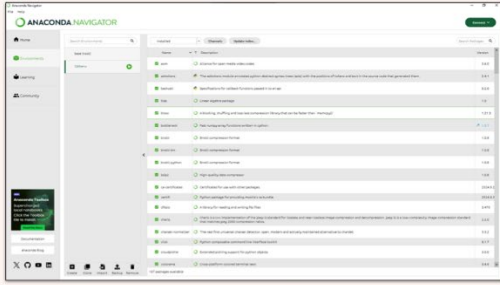
ПРОЕКТУВАННЯ МОДЕЛІ



10

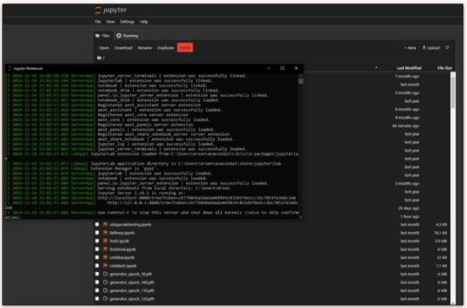
Слайд 10 – Проектування моделі

ТЕХНІЧНІ ПЛАТФОРМИ



➤ Для реалізації системи обрано комплекс сучасних технологічних інструментів, що забезпечують ефективну розробку та впровадження AI-рішень. В якості основного середовища розробки використовується **Jupyter Notebook**, який надає інтерактивний інтерфейс для розробки та тестування компонентів системи.

- Управління залежностями та віртуальним середовищем здійснюється через **Anaconda Navigator**, де створено спеціалізоване середовище **GANenv** з усіма необхідними бібліотеками.
- Основою системи є фреймворк **PyTorch**, який забезпечує гнучкі можливості для розробки та навчання нейронних мереж. Для оптимізації обчислень використовується технологія **CUDA**, що дозволяє ефективно задіяти потужності графічного процесора. Система розроблена з урахуванням можливості масштабування та підтримки різних апаратних конфігурацій, від локальних робочих станцій до хмарних середовищ розробки.

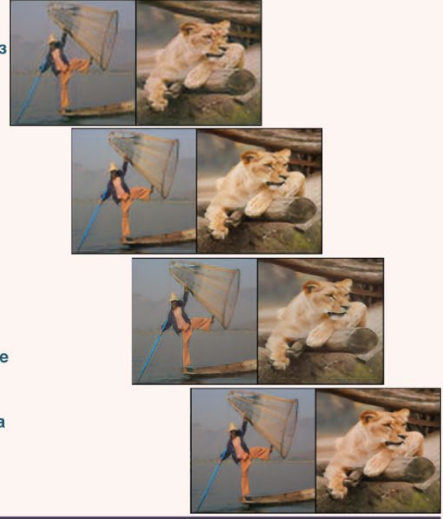


13

Слайд 13 – Технічні платформи

НАВЧАННЯ МОДЕЛІ

- В ході реалізації проекту було проведено навчання моделі протягом **200 епох** на обраному наборі даних. Для оптимізації використовувався алгоритм **Adam** з початковою швидкістю навчання **0.0002** та коефіцієнтами **beta1=0.5**, **beta2=0.999**. Стабільність навчання забезпечувалась використанням механізму накопичення градієнтів з параметром **accumulation_steps=4**.
- Для ефективного використання обчислювальних ресурсів було застосовано технологію змішаної точності обчислень (**mixed precision training**) з використанням **GradScaler**. Це дозволило оптимізувати використання відеопам'яті при збереженні точності обчислень. Розмір батчу встановлено на рівні **4** зображень, що забезпечило оптимальний баланс між швидкістю навчання та використанням пам'яті.
- Кожні **10 епох** проводилась оцінка якості на валідаційному наборі з обчисленням метрик **PSNR** та **SSIM**, а також збереженням проміжних результатів для візуального аналізу. Процес навчання демонстрував стабільне покращення як об'єктивних метрик якості, так і візуальних результатів.
- Реалізована система логування дозволила відстежувати динаміку навчання та вчасно виявляти можливі проблеми. Додатково було реалізовано механізм збереження стану моделі, що забезпечило можливість відновлення навчання після можливих перерв або збоїв.



14

Слайд 14 – Навчання моделі

ТЕХНІЧНІ ВИМОГИ ТА ОБМЕЖЕННЯ

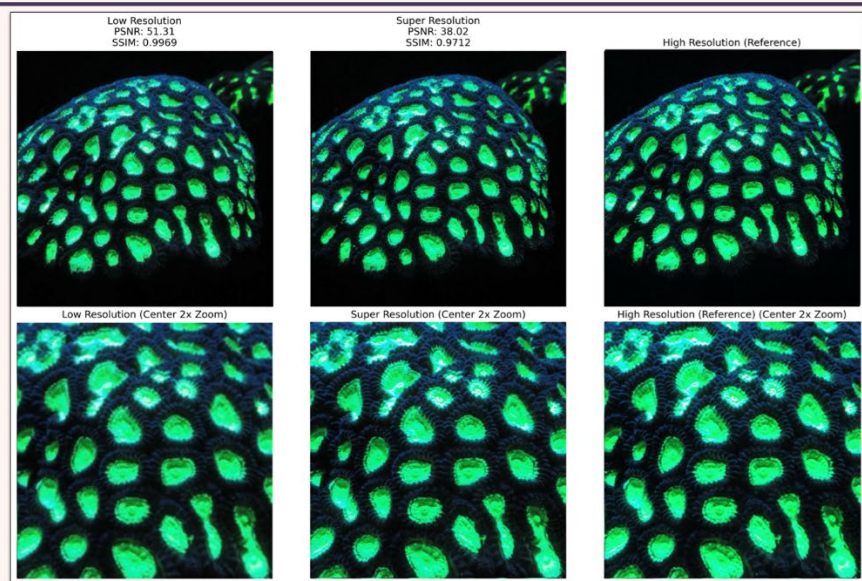
- При розробці та тестуванні системи було проведено серію експериментів на різних апаратних конфігураціях, де ключовим фактором виявився обсяг доступної відеопам'яті (**VRAM**). Експерименти показали, що обсяг **VRAM** безпосередньо впливає на розмір батчу та швидкість навчання: при **8 ГБ VRAM** система стабільно працює з батчем розміром **4** зображення, тоді як **16 ГБ** дозволяють збільшити батч до **12-16** зображень, прискорюючи навчання в **2.5-3** рази.
- Тестування на різних **GPU** підтвердило, що саме обмеження **VRAM**, а не обчислювальна потужність, є головним фактором, що впливає на продуктивність при роботі з високороздільними зображеннями. Оптимальна конфігурація включає **GPU з 12-16 ГБ VRAM, 32 ГБ** оперативної пам'яті та швидкий **SSD** накопичувач.
- Експерименти з різними об'ємами системної пам'яті та типами накопичувачів показали їх меншу, але все ж важливу роль у загальній продуктивності. При такій конфігурації система стабільно обробляє зображення до **2K** роздільної здатності, забезпечуючи оптимальний баланс між якістю результатів та швидкістю обробки.

15

Слайд 15 – Технічні вимоги та обмеження

ПРИКЛАД РОБОТИ

- Кількісна оцінка демонструє високу ефективність системи: для вхідного зображення **PSNR** складає **51.31** та **SSIM** **0.9969**, тоді як для обробленого зображення отримано **PSNR 38.02** та **SSIM 0.9712**. Незважаючи на дещо нижчі числові показники (очікуємо) у обробленого зображення, візуальний аналіз збільшених фрагментів показує, що система успішно відновлює дрібні деталі та текстури, зберігаючи природний вигляд зображення. Особливо помітно покращення чіткості контурів флуоресцентних структур та відновлення тонких градієнтів яскравості.



16

Слайд 16 – Приклад роботи

ВИСНОВКИ

- У цій роботі було створено власну генеративну змагальну модель. Аналіз актуальності теми підтвердив, що якість зображень є важливим аспектом для різних сфер застосування, таких як медицина, мультимедіа, машинне зорове сприйняття та інші.
- Реалізована модель на базі **Real-ESRGAN Compact (SRVGGNet)** продемонструвала високу ефективність у відновленні дрібних деталей та текстур, особливо в складних випадках з різними типами спотворень. Практична реалізація підтвердила стабільність навчання та візуальне покращення якості зображень.
- Застосування онтологічного підходу дозволило структурувати знання про спотворення та методи їх корекції, виявивши нові взаємозв'язки між ними. Це стало основою для оптимізації архітектури генератора та дискримінатора, що значно підвищило ефективність моделі.
- Подальша оптимізація параметрів навчання та архітектури моделі може покращити як традиційні метрики оцінки, так і візуальну якість результатів. Розширення онтології сприятиме адаптації моделі до нових типів спотворень та специфічних сценаріїв використання.
- Загалом, розробка була успішною та має значний потенціал для подальшого розвитку та масштабування.

17

Слайд 17 – Висновки

ДЯКУЮ ЗА УВАГУ



Слайд 18 – Дякую за увагу