

ДОДАТОК А

Слайди презентації

Моделі генерації анімаційного зображення з врахуванням аудіо- поточку

Коміна О.О., група ІПЗм-23-3
Науковий керівник: доцент Турута О.П.



__ червня 2025

Дослідження

Актуальність дослідження зумовлена поширенням використання ШІ в останні роки, стрімким розвитком технологій галузі GenAI, розповсюдженням його використання та появі його в щоденному побуті.

Метою дослідження є дослідити основні існуючі open-source пропозиції для Talking face генерації, вивчити їх документацію, визначити основні риси і напрями застосування, визначити практично можливість їх використання на прикладі звичайного користувача ПК і проаналізувати результати.

Об'єктом дослідження є моделі генерації анімаційного зображення з врахуванням аудіо-поточку, які також відомі як Talking Face.



Огляд літератури та аналогів

Моделі для генерації відеоряду (Wav2Lip, DInet, MakeItTalk, SadTalker)

Моделі для оцінювання якості штучно згенерованого відеоряду (SyncNet, Q-Align, vBench)

Моделі для відновлення обличчя (VQFR, GFPGAN)

Додаткові програми (OpenFace, Ffmpeg, pip, virtualenv, git)



Постановка задачі

Досліджувана проблема: дослідити основні існуючі open-source пропозиції для Talking face генерації, вивчити їх документацію, визначити основні риси і напрями застосування, визначити практично можливість їх використання з Windows та Linux на прикладі Ubuntu з використанням VirtualBox, проаналізувати легкість встановлення і застосування на прикладі звичайного користувача ПК, проаналізувати результати генерації на різних прикладах зображень-першоджерел.

Очікуваний результат: розуміння технології Talking face і основних технологій, що використовуються в генерації відеоряду, розуміння принципу їх встановлення і застосування, аналіз основних переваг і недоліків різних застосунків, а також основні вимоги до вихідних зображень.



Методологія

Використані методи дослідження:

- аналіз літературних джерел
- порівняльний аналіз
- розробка прототипу
- експериментальне дослідження
- аналіз результатів експериментів
- опитування та експертні оцінки



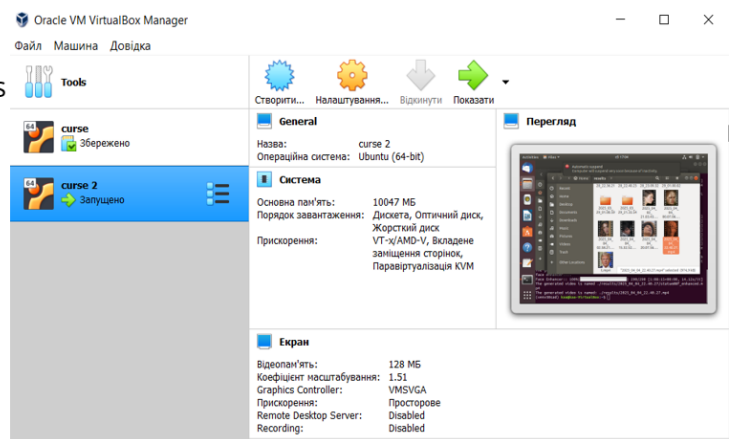
5

Архітектура системи для проведення експериментального дослідження

Для дослідження було вирішено використати Ubuntu 16.04 (використовуючи Virtualbox) та Windows 10

Також для дослідження було встановлено:

- Ffmpeg
- Git
- Python (including pip and virtualenv)



6

Опис програмного забезпечення, що було використано у дослідженні

Моделі для генерації відеоряду

- Wav2Lip
- DINet
- MakeItTalk
- SadTalker



7

Зміст проведеного експерименту



Оригінал



Wav2Lip

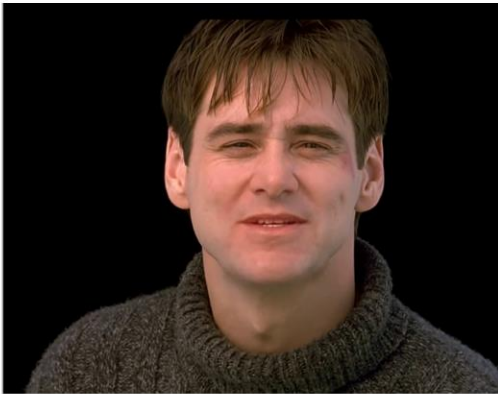


SadTalker

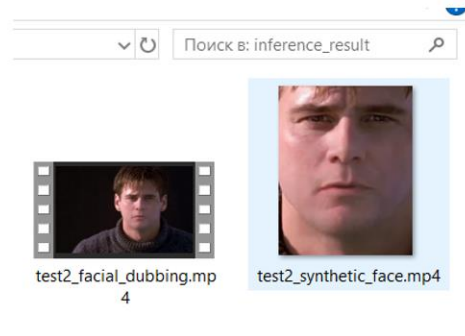


8

Зміст проведеного експерименту



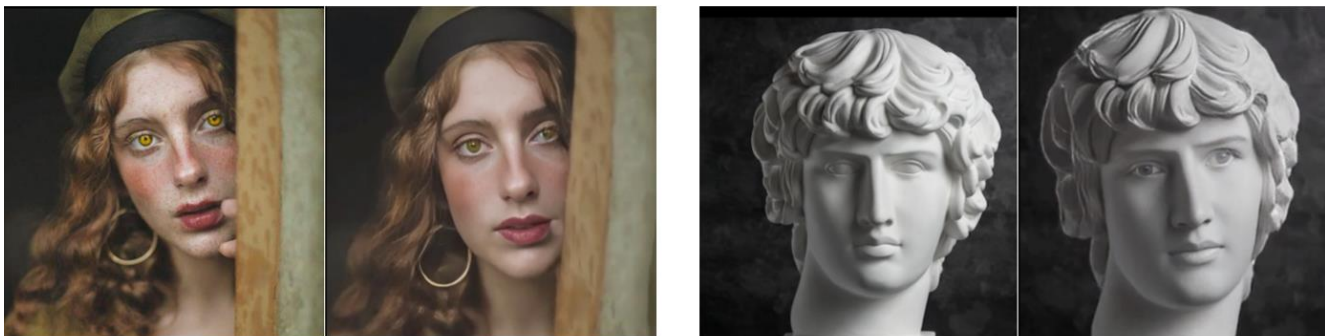
Дослідження DINet



Зміст проведеного експерименту



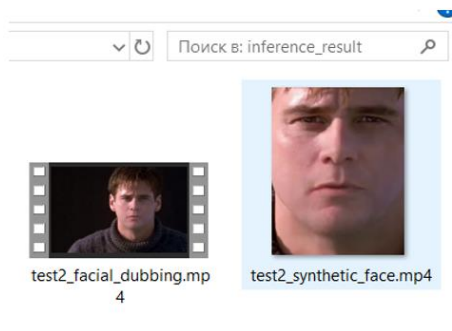
Зміст проведеного експерименту



Зміст проведеного експерименту



Дослідження [DINet](#)



Результати експерименту

	Якість і ефективність	Доступність і легкість використання	Часова ефективність	Вимогливість до ресурсів ПК	Об'єм зайнятої пам'яті	Багатофункціональність
Wav2Lip	1	3	4	3	5	3
Wav2Lip-GAN	3	3	4	1	5	3
Wav2Lip-HD	3	2	4	1	4	3
SadTalker	5	5	1	3	1	3
DINet	5	2	3	5	2	2
MakeItTalk	1	3	4	1	1	2

Аналіз отриманих результатів

Альтернативи	Сума
Wav2Lip	3,12
Wav2Lip - GAN	3,12
SadTalker	3,21
DINet	3,02

Аналіз отриманих результатів

- Для отримання кращого результату слід обирати фотографії, зроблені в умовах м'якого, розсіяного освітлення без чітко виражених світлових плям і глибоких тіней.
- Оптимальним варіантом є зображення у фронтальному ракурсі.
- Необхідно використовувати фото без будь-яких об'єктів, що частково перекривають обличчя.
- Важливо враховувати, що під час генерації напрям погляду визначається за орієнтацією голови, а не положенням очей на оригінальному фото.
- При роботі з неживими або стилізованими зображеннями слід зважати на можливу появу ефекту «зловісної долини», коли згенероване відео виглядає неприродно.



Аналіз отриманих результатів

- Слід звертати увагу на дрібні індивідуальні особливості шкіри, як-от родимки чи веснянки, особливо якщо відомо, що вони наявні у людини на фото.
- Повна нерухомість тіла або, навпаки, надмірно активні рухи можуть свідчити про штучність відео.
- Присутність яскравого та контрастного освітлення з глибокими тінями часто вказує на реальність зйомки.
- Зображення, зроблені під нестандартними ракурсами (наприклад, зверху або знизу), складніше згенерувати, тому такі ракурси можуть підтверджувати справжність медіа.
- Наявність об'єктів, що частково закривають обличчя, як-от рука, волосся або сторонні предмети, ускладнює генерацію та слугує додатковим критерієм справжності відео.
- Зміна або спотворення геометрично правильних форм, наприклад, сережок, може вказувати на синтетичне походження відео.
- Варто уважно оцінювати сталість рис обличчя та загальних пропорцій як порівняно з оригіналом, так і протягом усього відео — їхня варіативність свідчить про можливу генерацію.





Підсумки

- В результаті виконано такі завдання як:
- Проведено огляд предметної галузі і визначено основні моделі подібної синхронної генерації.
- Проаналізовано обрані моделі, визначено їх слабкі і сильні риси, визначено вимоги для їх використання.
- Створено системи-прототип для успішного запуску протестованих застосунків і тестування їх на різних операційних системах.
- Виконано експериментальне дослідження результатів виконання різних застосунків, проаналізовано результати виконання на різних типах вихідних даних.


В майбутньому можна продовжити дослідження вивченням інших моделей, перевірці їх на інших конфігураціях техніки і дослідження генерації в залежності від аудіоряду

ДОДАТОК Б

Результати перевірки на антиплагіат

Дата звіту 6/12/2025
Дата редагування ---


Звіт не був оцінений

Звіт подібності

метадані

Назва організації
Kharkiv National University of Radio Electronics


Заголовок
2025_M_PI_IP3-23-3_Коміна_О_О_скорочений

Автор Науковий керівник / Експерт
Коміна Олександра Олександрівна Турута О.П./Нечволод В.Ю.


підрозділ
каф. ПІ

Обсяг знайдених подібностей


Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



25
Довжина фрази для коефіцієнта подібності 2






7554
Кількість слів



56346
Кількість символів

Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв		0
Інтервали		0
Мікропробіли		0
Білі знаки		0
Парафрази (SmartMarks)		3

Подібності за списком джерел

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз		Колір тексту
ПОРЯДКОВИЙ НОМЕР	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	https://openarchive.nure.ua/bitstream/document/14498/1/2020_M_IMI_Maksymovsky_AV.docx	28 0.37 %
2	https://arxiv.org/abs/2303.03988	13 0.17 %
3	https://openarchive.nure.ua/bitstreams/39da9559-5870-4e28-a844-bad3121070be/download	6 0.08 %

з бази даних RefBooks (0.00 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-----------	--

ДОДАТОК В

Експертний висновок результатів перевірки кваліфікаційної роботи на
відповідність оформлення вимогам ДСТУ 3008: 2015

Експертний висновок результатів перевірки кваліфікаційної роботи

студент
(посада)

програмної інженерії
(кафедра)

ІПЗм-23-3
(група)

Коміна Олександра Олександрівна

(прізвище, ім'я, по батькові)

Зауваження

Пункт ДСТУ 3008-2015	Зміст пункту	Сторінка кваліфікаційної роботи
1	2	3
	7.1 Загальні положення	
	7.3 Нумерація сторінок звіту	
	7.4 Нумерація розділів, підрозділів, пунктів, підпунктів	
	7.5 Рисунки	
	7.6 Таблиці	
	7.7 Переліки	
	7.8 Примітки	
	7.9 Виноски	
	7.10 Формули та рівняння	
	7.11 Посилання	
	7.13 Список авторів	
	7.14 Скорочення та умовні позначки	
	7.15 Додатки	

зауважень немає

Експерт

(підпис)

Олена ОЛІЙНИК

(прізвище, ініціали)

12.06.2025