

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів

Кафедра комп'ютерних наук та інформаційних технологій

Пояснювальна записка
до кваліфікаційної роботи
(тип кваліфікаційної роботи)

магістра

(освітній ступінь)

на тему «Дослідження генеративних дифузійних моделей для вирішення задачі
заміни обличчя на зображеннях»

XAI.302.366.24O.122. 203105 ПЗ

Виконав: здобувач 6 курсу групи № 366
Спеціальність 122 «Комп'ютерні науки»

(код та найменування)

Освітня програма Комп'ютеризація
обробки інформації та управління

Мельник С.В.

(прізвище й ініціали студента)

Керівник: доц. каф. 302, к.т.н.

Малєєва Ю.А.

(прізвище й ініціали)

Рецензент: Литвинов А.Л.

(прізвище й ініціали)

Харків – 2024

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів
 Кафедра комп'ютерних наук та інформаційних технологій
 Рівень вищої освіти другий (магістерський)
 Галузь знань 12 «Інформаційні технології»
 (код та найменування)
 Спеціальність 122 «Комп'ютерні науки»
 Освітня програма Комп'ютеризація обробки інформації та управління

ЗАТВЕРДЖУЮ

Завідувач кафедри

Олег ФЕДОРОВИЧ

(підпис)

(ім'я та прізвище)

«08» листопада 2024 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Мельник Сергію Віталійовичу

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи Дослідження генеративних дифузійних моделей для вирішення задачі заміни обличчя на зображеннях

керівник кваліфікаційної роботи Малєєва Ю.А., к.т.н., доц. каф. 302,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету № 2050-уч. від « 06 » листопада 20 24 року

2. Термін подання здобувачем кваліфікаційної роботи 15.12.2024

3. Вихідні дані до проекту (роботи) дослідження нейромережових моделей для завдань заміни осіб на зображеннях та відео

4. Зміст пояснювальної записки (перелік завдань, які потрібно розв'язати) провести аналіз задачі заміни обличчя на зображеннях та існуючих підходів до її вирішення; вивчити архітектуру генеративних моделей, зокрема дифузійної моделі DiffFace та Stable Diffusion; розробити концептуальні положення та запропонувати шляхи покращення дифузійної моделі для заміни обличчя; реалізувати прототип дифузійної моделі та провести серію експериментів із заміни обличчя на зображеннях; оцінити результати експериментів та запропонувати рекомендації для подальшого вдосконалення моделей

5. Перелік графічного матеріалу зміна обличчя як сфера комп'ютерного зору; ціль та задачі дослідження; аналіз існуючих підходів; дифузійна модель DiffFace; LDM (Stable Diffusion); шляхи покращення дифузійної моделі DiffFace; опис рішення для взаємодії архітектури DiffFace та моделі LDM; експерименти

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада Консультанта	Підпис, дата	
		завдання видав	завдання прийняв
з основної частини	Малєєва Ю.А. доц. каф. 302		
з спеціального розділу	Малєєва Ю.А. доц. каф. 302		
з економіки та управління	Малєєва Ю. А. доц. каф. 302		

Нормоконтроль _____ Олександр ПРОХОРОВ «__» _____ 20__ р.
(підпис) (імя та прізвище)

7. Дата видачі завдання « 06 » вересня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	АНАЛІЗ ОСОБЛИВОСТЕЙ ТА ІСНУЮЧИХ ПІДХОДІВ ДО ПЕРЕНЕСЕННЯ ОСІБ ЛЮДЕЙ НА ФОТО ТА ВІДЕО	06.09 – 20.09	
2	КОНЦЕПТУАЛЬНІ ПОЛОЖЕННЯ ТА ОПИС МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ДЛЯ РЕАЛІЗАЦІЇ ГЕНЕРАТИВНИХ ЗАВДАНЬ ДИФУЗІЇ	21.09 – 20.10	
3	ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ З ЗАМІНИ ОСІБ НА ЗОБРАЖЕННЯХ	21.10 – 20.11	
4	ЕКОНОМІЧНІ РОЗРАХУНКИ	20.11 – 25.11	

Здобувач _____ Сергій МЕЛЬНИК
(підпис) (імя та прізвище)

Керівник кваліфікаційної роботи _____ Юлія МАЛЄЄВА
(підпис) (імя та прізвище)

УДК 004.8

РЕФЕРАТ

Дослідження генеративних дифузійних моделей для вирішення задачі заміни обличч на зображеннях – Мельник Сергій, кваліфікаційна робота магістра, Харків, Національний аерокосмічний університет ім. М. Є. Жуковського «ХАІ», кількість сторінок, враховуючи додатки 68, кількість таблиць 16, кількість рисунків 35, кількість джерел літератури 15.

КЛЮЧОВІ СЛОВА: комп'ютерний зір, глибокого навчання, перенесення обличчя в зображеннях, заміна осіб, генеративні моделі, дифузійні моделі, латентний простір.

ЦІЛЬ: дослідження та розробка дифузійної моделі для вирішення завдань заміни обличч на зображеннях.

ОБ'ЄКТ ДОСЛІДЖЕННЯ: процес генерації та заміни обличч на зображеннях із використанням генеративних дифузійних моделей.

ПРЕДМЕТ ДОСЛІДЖЕННЯ: методи та алгоритми генеративних дифузійних моделей для заміни обличч на зображеннях.

ЗАДАЧІ:

1. Провести аналіз задачі заміни обличч на зображеннях та існуючих підходів до її вирішення.
2. Вивчити архітектуру генеративних моделей, зокрема дифузійної моделі DiffFace та Stable Diffusion.
3. Розробити концептуальні положення та запропонувати шляхи покращення дифузійної моделі для заміни обличч.
4. Реалізувати прототип дифузійної моделі та провести серію експериментів із заміни обличч на зображеннях.
5. Оцінити результати експериментів та запропонувати рекомендації для подальшого вдосконалення моделей.

МЕТОДИ ДОСЛІДЖЕННЯ: методи штучного інтелекту та машинного навчання, глибокі нейронні мережі, генеративні дифузійні моделі.

ОТРИМАНІ РЕЗУЛЬТАТИ: розроблено нейромережеву модель для вирішення завдань заміни обличч на зображеннях.

UDC 004.8

Research on generative diffusion models for solving the problem of replacing faces in images – Serhii Melnyk, Master Thesis, Kharkiv, National Aerospace University “KhAI”, amount of pages 27, amount of tables 16, amount of figures 35, amount of literature’s sources 15.

KEYWORDS: computer vision, deep learning, face swapping in images, face replacement, generative models, diffusion models, latent space.

PURPOSE OF RESEARCH: to study and develop a diffusion model for solving face replacement tasks in images.

OBJECT OF RESEARCH: the process of generating and replacing faces in images using generative diffusion models.

SUBJECT OF RESEARCH: methods and algorithms of generative diffusion models for face replacement in images.

TASKS: analyze the problem of face replacement in images and the existing approaches to its solution; study the architecture of generative models, specifically the diffusion models DiffFace and Stable Diffusion; develop conceptual foundations and propose ways to improve the diffusion model for face replacement; implement a prototype of the diffusion model and conduct a series of experiments on face replacement in images; evaluate the experimental results and propose recommendations for further model improvement.

METHODS OF RESEARCH: artificial intelligence and machine learning methods, deep neural networks, generative diffusion models.

RESULTS: a neural network model for solving face replacement tasks in images has been developed.

ЗМІСТ

ВСТУП.....	7
1 АНАЛІЗ ОСОБЛИВОСТЕЙ ТА ІСНУЮЧИХ ПІДХОДІВ ДО ПЕРЕНЕСЕННЯ ОСІБ ЛЮДЕЙ НА ФОТО ТА ВІДЕО	9
1.1 Завдання заміни осіб на зображенні.....	9
1.2 Аналіз існуючих підходів.....	9
1.2.1 Ідея автоенкодерів.....	9
1.2.2 Підходи, засновані лише на одному фото	13
1.2.3 FaceShifter.....	14
1.2.4 SimSwap та SimSwap++	16
1.2.5 HifiFace	20
1.2.6 MegaFS	23
Висновки за розділом.....	26
2 КОНЦЕПТУАЛЬНІ ПОЛОЖЕННЯ ТА ОПИС МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ ДЛЯ РЕАЛІЗАЦІЇ ГЕНЕРАТИВНИХ ЗАВДАНЬ ДИФУЗІЇ	27
2.1 Діфузійна модель DiffFace	27
2.2 Stable Diffusion.....	30
2.3 Шляхи покращення діфузійної моделі DiffFace	34
2.4 Опис рішення.....	35
Висновки за розділом.....	42
3 ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ З ЗАМІНИ ОСІБ НА ЗОБРАЖЕННЯХ	44
3.1 Експерименти заміни осіб на зображеннях.....	44
Висновки за розділом.....	48
4 ЕКОНОМІЧНИЙ РОЗДІЛ	49
4.1 Мета розділу	49
4.2 Визначення тривалості та розрахунок собівартості виконання проекту (перший варіант).....	49
4.3 Визначення тривалості та розрахунок собівартості виконання проекту (другий варіант).....	54
Висновки за розділом.....	58
ВИСНОВКИ.....	59
ПЕРЕЛІК ПОСИЛАНЬ	60
ДОДАТОК А. Графічні матеріали.....	62

ВСТУП

Актуальність теми дослідження.

Зміна обличчя стала захоплюючою сферою комп'ютерного зору та глибокого навчання, уможливлуючи безперебійне перенесення обличчя однієї людини на тіло іншої в зображеннях і відео.

Таким чином, заміна осіб – це генеративне завдання штучного інтелекту, де потрібно замінити обличчя однієї людини на обличчя іншої, зберігаючи при цьому важливі характеристики зображення, такі як вираз обличчя, поза та фон.

Використання генеративно-змагальних мереж є класичним підходом у цій області. У ранніх роботах, таких як FSLSD та MegaFS, для досягнення цієї мети застосовувалася попередньо навчена модель Style-GAN. Однак останнім часом дифузійні моделі стали більш популярною альтернативою генеративно-змагальних мереж, пропонуючи більш гнучкі можливості та кращу якість.

Є методи, що стали вже класичними в цій галузі, такі як DDIM та Stable Diffusion.

Ми проведемо дослідження щодо застосування такого підходу в галузі дифузії для вирішення задачі заміни осіб.

Мета дослідження – дослідження та розробка дифузійної моделі для вирішення завдань заміни облич на зображеннях.

Для досягнення мети необхідно виконати наступні **завдання**:

1. Провести аналіз задачі заміни облич на зображеннях та існуючих підходів до її вирішення.
2. Вивчити архітектуру генеративних моделей, зокрема дифузійної моделі DiffFace та Stable Diffusion.
3. Розробити концептуальні положення та запропонувати шляхи покращення дифузійної моделі для заміни облич.
4. Реалізувати прототип дифузійної моделі та провести серію експериментів із заміни облич на зображеннях.
5. Оцінити результати експериментів та запропонувати рекомендації для подальшого вдосконалення моделей.

Об'єкт дослідження – процес генерації та заміни облич на зображеннях із використанням генеративних дифузійних моделей.

Предмет дослідження – методи та алгоритми генеративних дифузійних моделей для заміни облич на зображеннях.

Практичне значення отриманих результатів. Розроблено нейромережеву модель для вирішення завдань заміни облич на зображеннях.

Структура та обсяг кваліфікаційної роботи магістра. Кваліфікаційна робота магістра складається зі вступу, 4 розділів, висновків, додатків та переліку

посилань. Повний обсяг кваліфікаційної роботи магістра становить 68 сторінок, в тому числі: 35 рисунків; 15 таблиць; 1 додаток на 7 сторінках; перелік посилань з 15 найменувань на 2 сторінках.