

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу**

«На правах рукопису»
УДК 004.932.4

«До захисту допущено»

Завідувач кафедри

_____ О.Л.Тимошук

«__» _____ 20__ р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності 124 Системний аналіз

на тему: «Інтелектуальна система фотореалістичного перенесення стилів між зображеннями»

Виконала:
студентка II курсу, групи КА-61м
Піпко Анна Сергіївна

Керівник:
доцент, к.т.н., доц.
Дідковська М.В.

Рецензент:
доцент кафедри програмного забезпечення
комп'ютерних систем ФПМ,
к.т.н., доц.
Заболотня Т.М.

Засвідчую, що у цій магістерській дисертації
немає запозичень з праць інших авторів без
відповідних посилань.

Студентка _____

Київ
2018

РЕФЕРАТ

Магістерська дисертація: 72 с., 16 рис., 23 табл., 2 додатки та 26 джерел.

Об'єктом дослідження є методи технічного зору та перенесення стилів між зображеннями.

Метою даної роботи є розробка системи фотореалістичного перенесення стилів між знімками, а також реалізація додатку для демонстрації роботи системи. У роботі проаналізовано методи технічного зору, проведено огляд існуючих підходів перенесення стилів та підвищення реалістичності зображень.

Методом дослідження є модель, що поєднує техніки комп'ютерного зору та глибокого навчання.

Результати роботи:

- запропоновано алгоритм реалістичного перенесення стилю між фотознімками;
- реалізовано запропоновану конфігурацію для системи перетворення похмурої погоди на фотознімках на сонячну;
- створено додаток із зручним для користувача графічним інтерфейсом.

Результати даної роботи рекомендується використовувати у фоторедакторах та інших сервісах для обробки зображень. При подальших дослідженнях у цій області доцільно покращити точність перенесення стилю та підвищити швидкість обробки кадрів.

ТЕХНІЧНИЙ ЗІР, МАШИННЕ НАВЧАННЯ, РОЗПІЗНАВАННЯ ОБРАЗІВ, СЕГМЕНТАЦІЯ, ГЛИБОКА ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, ФОТОРЕАЛІСТИЧНЕ ПЕРЕНЕСЕННЯ СТИЛІВ, ЕКРАНОВАНЕ РІВНЯННЯ ПУАССОНА.

ABSTRACT

The master's thesis: 72 p., 16 fig., 23 tabl., 2 appendices and 26 sources.

The theme of this thesis is “Intelligent system of photorealistic style transfer between images”.

Object of study: computer vision and style transfer techniques.

The purpose of this thesis is to develop system of photorealistic style transfer between images. Several methods of computer vision and style transfer were analyzed in the thesis.

The method of observation: a model that combines computer vision and deep learning techniques.

The results of the thesis:

- an algorithm for photorealistic style transfer between images was proposed;
- the proposed configuration was implemented for transferring sunny weather to the photo;
- an application with graphical user interface to demonstrate the system's work was developed.

The results of this thesis are recommended for use in the photoeditors and another applications for editing images. In further research it is reasonable to improve the faithfulness and increase the speed of image processing.

COMPUTER VISION, MACHINE LEARNING, IMAGE RECOGNITION, SEGMENTATION, DEEP CONVOLUTIONAL NEURAL NETWORK, PHOTOREALISTIC STYLE TRANSFER, SCREENED POISSON EQUATION

ЗМІСТ

ПЕРЕЛІК ПРИЙНЯТИХ СКОРОЧЕНЬ.....	8
ВСТУП.....	9
РОЗДІЛ 1 МЕТОДИ ПЕРЕНЕСЕННЯ СТИЛІВ МІЖ ЗОБРАЖЕННЯМИ	11
1.1 Аналіз актуальності задачі	11
1.2 Дослідження існуючих підходів.....	12
1.2.1 Налаштування фотокамери	12
1.2.2 Використання фоторедакторів.....	13
1.2.2 Інтелектуальні методи	14
1.3. Системний підхід до задачі.....	18
1.4. Формалізація постановки задачі.....	19
Висновки за розділом	19
РОЗДІЛ 2 АЛГОРИТМИ ФОТОРЕАЛІСТИЧНОГО ПЕРЕНЕСЕННЯ СТИЛІВ.....	20
2.1 Представлення зображення у глибоких згорткових нейронних мережах	20
2.2 Алгоритми перенесення стилів.....	20
2.2.1 Алгоритм NeuralStyle.....	22
2.2.2 Алгоритм StyleSwap.....	24
2.3 Семантична сегментація зображень.....	27
2.3.1 Метод DeepLab	29
2.4 Методи підвищення фотореалістичності	30
2.4.1 Алгоритм на основі Matting Laplacian	31
2.4.2 Алгоритм на основі екранованого рівняння Пуассона	32
2.5 Система оцінки якості генератора.....	34
2.6 Архітектура системи фотореалістичного перенесення часу доби.....	34
2.6.1 Параметри методу StyleSwap.....	35
2.6.2 Використання сегментації.....	38

Висновки за розділом	40
РОЗДІЛ 3 АРХІТЕКТУРА ПРОГРАМНОГО ПРОДУКТУ ТА АНАЛІЗ ПРАКТИЧНИХ РЕЗУЛЬТАТІВ	41
3.1 Вибір інструментів розробки	41
3.2 Аналіз архітектури продукту	42
3.3 Інтерфейс користувача	45
3.4 Аналіз практичних результатів.....	46
Висновки за розділом	49
РОЗДІЛ 4 РОЗРОБЛЕННЯ СТАРТАП ПРОЕКТУ	50
4.1 Опис ідеї проекту	51
4.2 Технологічний аудит ідеї проекту.....	53
4.3 Аналіз ринкових можливостей запуску стартап-проекту.....	54
4.4 Аналіз ринкової стратегії проекту.....	60
4.5 Розроблення маркетингової програми стартап-проекту.....	63
Висновки за розділом	67
ВИСНОВКИ.....	68
ПЕРЕЛІК ПОСИЛАНЬ	70
ДОДАТОК А ЛІСТИНГ ПРОГРАМИ.....	73
ДОДАТОК Б ІЛЮСТРАТИВНІ МАТЕРІАЛИ ДОПОВІДІ.....	80

ПЕРЕЛІК ПРИЙНЯТИХ СКОРОЧЕНЬ

ANN – Artificial Neural Network, штучна нейронна мережа

CNN – Convolutional Neural Network, згорткова нейронна мережа

SPE – Screened Poisson Equation, екрановане рівняння Пуассона

ПП – програмний продукт

ВСТУП

Однією зі значних сучасних тенденцій є стрімкий розвиток соціальних мереж в мережі Інтернет. Користувачі прагнуть максимально повно поділитися своїми враженнями з друзями і найпростіше це зробити за допомогою фотографій. Зазвичай знімки зроблені фотокамерою смартфона і тому втрачають виразність через брак освітленості. У похмуру погоду чи в час сутінок усі об'єкти на зображеннях стають сірими та тьмяними і композиція стає значно менш вражаючою, ніж при сонячній погоді чи у світлі ліхтарів. За допомогою фоторедакторів можна відретушувати фотографію, проте цей кропіткий процес потребує достатньо спеціалізованого досвіду роботи з низькорівневими характеристиками зображення (баланс білого, яскравість, теплота тонів тощо) і тому недоступний більшості користувачів фотоапаратів. Отже, виникає потреба у створенні інструменту, який дозволяв би відредагувати фотографію за допомогою зрозумілих кожному високорівневих характеристик: сонячна чи похмура погода, день чи ніч.

Таким чином, метою роботи є розробка системи фотореалістичної зміни часу доби на зображеннях.

Для досягнення такої мети були вирішені наступні задачі:

- аналіз існуючих алгоритмів зміни часу доби, погоди, сезону;
- розробка системи перенесення стилів між фотознімками;
- практична реалізація алгоритму фотореалістичної зміни часу доби на фотографії;
- реалізація додатку із зручним для користувача графічним інтерфейсом для демонстрації роботи системи.

Об'єктом дослідження є методи технічного зору та перенесення стилів між зображеннями.

Предметом дослідження є методи збереження реалістичності зображення при перенесенні стилів між фотографіями.

Було запропоновано алгоритм зміни часу доби на фотографії, який базується на перенесенні стилю з використанням семантичної сегментації.

Практичними результатами роботи є реалізація системи фотореалістичної зміни похмурої погоди на фотознімках на сонячну.

Робота складається з 4 розділів. В першому розділі розглядається актуальність проблеми та існуючі підходи до її розв'язання, формалізується постановка задачі дослідження. Другий розділ присвячений математичним основам методів, які було застосовано у дослідженні, та вибору критеріїв оцінки результатів роботи. У третьому розділі здійснено огляд технологій та аргументовано вибір інструментів, що використовуються в роботі, описано процес розробки програмного продукту та проведено дослідження отриманих результатів його роботи. У четвертому розділі розглядаються можливості комерціалізації розробки у вигляді стартап-проекту.

РОЗДІЛ 1 МЕТОДИ ПЕРЕНЕСЕННЯ СТИЛІВ МІЖ ЗОБРАЖЕННЯМИ

1.1 Аналіз актуальності задачі

З розвитком технологій відстань між людьми значно скоротилася, активність віртуального спілкування зростає з кожним днем. Для багатьох користувачів соціальних мереж дуже важливо, наскільки сподобались їх записи та фотографії іншим, для деяких це навіть є шляхом заробітку. Все частіше можна дізнатись про життя своїх друзів не від них самих, а через ті матеріали, які вони викладають на своїх сторінках після свят, подорожей та й просто кожного дня під час повсякденних справ. І, звісно, кожен хоче створити для своїх друзів найбільш заманливу видимість свого життя через яскраві зображення та захопливі історії.

Під час зйомок у місті погодні умови часто стають для професійних фотографів перешкодою для отримання необхідних кадрів, що зумовлює додаткові часові та матеріальні затрати. А під час подорожі похмура погода може завадити збереженню тих яскравих вражень на фото. В цих випадках доводиться ретушувати знімки, щоб надати їм більшої виразності та відновити втрачені кольори.

На сьогодні існує безліч фоторедакторів різного рівня складності та вартості, за допомогою яких можна змінювати зображення за власним задумом. Проте для досягнення гарного результату необхідно провести різноманітну обробку яскравості, теплоти, різкості, балансу білого тощо, часто ще й поділивши зображення на певні частини та обробивши їх окремо. Таке низькорівневе редагування фотографій надто складне для більшості користувачів-аматорів.

Рішенням такої проблеми міг би бути інструмент, який дозволяв би працювати з характеристиками більш високого рівня, інтуїтивно зрозумілими кожному: час доби, пора року, погодні явища. Ці ознаки в певній мірі задають стиль, атмосферу зображення, проте ще не несуть його основну семантику. Отже, задача полягає в тому, щоб дозволити змінювати стиль зображення, зберігаючи при цьому його семантику.

1.2 Дослідження існуючих підходів

Фінальна якість зображення залежить від низки факторів, таких як зовнішні умови, технічні характеристики камери, якість обробки. І якщо на перший фактор вплинути достатньо важко, то останні два цілком придатні для регулювання.

1.2.1 Налаштування фотокамери

Сучасні камери дозволяють провести у ручному режимі достатньо точне налаштування спалаху, витримки, експозиції, значення діафрагми, оптимальні значення яких дуже залежать від зовнішніх умов, характеру зйомки та навіть ракурсу. Користування цими налаштуваннями потребує значного досвіду та глибокого розуміння їх значення та впливу на кінцевий результат.

Автоматичні режими визначають основні класи сцен (портрет, нічна зйомка, сутінки, пейзаж, макро, рух, погана освітленість, прожектор тощо) та встановлюють певні комбінації значень параметрів, проте вони лише наближені до оптимальних і тому не завжди є вдалими.

Налаштування фотокамери дозволяють зменшити ймовірність нечітких, засвічених чи занадто темних кадрів. Проте процес зйомки зазвичай досить динамічний, потребує зосередженості також і на композиції кадру, тому врахувати всі фактори для оптимальних налаштувань дуже важко.

1.2.2 Використання фоторедакторів

На сьогоднішній день доступно багато фоторедакторів, що дозволяють змінювати різноманітні характеристики зображення. Професійні програми надають користувачу велику кількість налаштувань та інструментів, на вивчення яких можна витратити декілька місяців. Сьогодні є досить популярними курси та школи обробки фотографій різного рівня професійності. Також у вільному доступі наявні уроки користування цими інструментами для досягнення безлічі ефектів на фотографіях, в тому числі і зміни часу доби. Проте для такої трансформації необхідно здійснити декілька кроків, також підбравши оптимальні параметри індивідуально для кожної фотографії. Для дійсно якісного результату необхідно мати достатньо досвіду та часу, тому такі маніпуляції більше відносяться до професійної діяльності, ніж до повсякденних розваг.

Останнім часом дуже популярним стало використання різноманітних ефектів або фільтрів (сепія, ранок, портрет, акцент, димка, вінтаж, нуар тощо), які являють собою певну комбінацію значень налаштувань, накладання градієнтів чи автоматичне застосування інших художніх перетворень. Такі трансформації зазвичай виконують досить грубу обробку, яка виявляється чи недостатньою, чи недоцільною, адже варіативність серед зображень дуже висока. Тому такі інструменти врятувати невдалу фотографію навряд чи зможуть, а лише наведуть трохи лоску.

1.2.2 Інтелектуальні методи

Розвиток інформаційних технологій та методів штучного інтелекту активно впливає на все більше нових галузей. Методи комп'ютерного зору вже менше характеризують лише спеціалізовані продукти для фахівців, а все більше інтегруються в інструменти для масового споживача. Сьогодні ведуться активні дослідження можливості використання методів машинного навчання та глибоких нейронних мереж для високорівневого аналізу та обробки зображень [1]. Кожен з них має свої переваги та недоліки, зазвичай балансує між ефективністю та швидкістю, адже більшість методів потребують певного обчислювального рівня.

Для опису зображення в області комп'ютерного зору часто застосовують поняття стилю та семантики (змісту). Семантика виражена конкретними об'єктами дійсності, які представлені на сцені зображення. Стиль зображення – сукупність низькорівневих ознак, текстур, характерних патернів, за якими можна суб'єктивно визначити схожість зображень, не враховуючи їх семантику. Обидва поняття інтуїтивно зрозумілі, проте майже не формалізовані.

В процесі ретушування фотографії зазвичай мають за мету зберегти семантику зображення, проте змінити другорядні ознаки зображення, тобто змінити стиль. Інтуїтивно стиль найпростіше описати, вказавши на інше зображення з бажаним стилем, адже це комплексне поняття, його складові тісно пов'язані та впливають одна на одну, тому проявляють свою характерність лише у сукупності. Таким чином, ретушування фотографії можна виразити через перенесення нового стилю зі збереженням початкової семантики.

Задача відділення стилю від семантики досі є дуже складною, їх комбінування та перенесення може розглядатися як задача синтезу та перенесення текстур, проте значно ускладнюється врахуванням змісту зображення. Проте нещодавні досягнення глибоких загорткових нейронних мереж дозволили виділяти високорівневу семантичну інформацію із зображень. Було показано, що згорткові нейронні мережі, натреновані для задачі класифікації об'єктів на значній вибірці розмічених

зображень, здатні виділяти високорівневі ознаки [2], які залишаються інформативними і на інших вибірках, і застосовно до інших задач комп'ютерного зору.

Загалом розроблені методи перенесення стилю можна поділити на дві групи: більш ранні, засновані на низькорівневих характеристиках зображення, та більш сучасні, з використанням глибоких згорткових нейронних мереж.

Методи першої групи дозволяють виконувати нескладні перетворення зображення.

Алгоритми перенесення глобального стилю перетворюють зображення шляхом застосування просторово-інваріантної функції трансформації. Ці методи ефективні та можуть застосовуватись для таких простих стилів, як глобальне зміщення кольору (наприклад, сепія) та кривих тону (наприклад, висока та низька контрастність). Наприклад, у [3] автори співставляли середнє та стандартне відхилення вхідного та стильового зображень після їх переведення до декорельованого простору кольорів. У [4] було описано алгоритм перенесення повної тривимірної гистограми кольорів, використовуючи послідовність одновимірних гістограм. Проте ці методи виявились обмеженими у їх можливості передавати складні стилі.

Алгоритми перенесення локального стилю, засновані на просторовому співставленні кольорів, більш виразні та можуть застосовуватись для широкого класу задач, таких як зміна часу доби, художньої манери, погоди та сезону.

Виділяється розробка, що дозволяла змінювати погоду, час доби, пору року [5]. Для цього було створено підбірку фотографій із 101 камери протягом довгого періоду та проанотовано її відповідно до присутніх на зображеннях кліматичних явищ та характеристик. Таким чином, для кожної сцени можна було виділити власне її характерні особливості (незмінну композицію об'єктів) та перехідні характеристики (ознаки об'єктів, що змінюються з часом) та їх вплив на вигляд об'єктів. До таких характеристик і відносилися погода, час доби, пора року, які визначали такі атрибути як літо, осінь, зима, весна, день, ніч, дощ, туман, сонячно, холодно, сніг, спекотно, світанок, сутінки тощо. Користувачу пропонувалося

збільшити чи зменшити міру прояву того чи іншого атрибуту вхідного зображення, тобто задати бажану зміну фотографії шляхом опису високорівневих, інтуїтивно зрозумілих ознак. Спочатку алгоритм знаходив у базі схоже зображення на вихідне, а потім визначав необхідну трансформацію. В рамках роботи було натреновано класифікатори, які визначали ступінь прояву кожного з атрибутів, що можна було також використовувати для впорядкування баз зображень та пошуку в них за атрибутами. Недоліком такого підходу була залежність від репрезентативності фотобаз та абсолютно не враховувалася семантика зображень.

Нова хвиля розробок почалася після дослідження [2] представлення зображення на різних шарах глибокої згорткової нейронної мережі, а саме тренуваної на дуже великій вибірці для класифікації ImageNet (задачі розпізнавання 1000 класів об'єктів). Тоді було встановлено, що нейрони ранніх шарів активуються на простих патернах (наприклад, границі, градієнти, кути тощо), а на пізніх шарах здатні виділяти більш змістовні ознаки (круглі предмети, обличчя людей, видовжені предмети тощо). Тобто можна казати про наближене відділення текстури від змісту.

Методи другої групи так чи інакше виникли під впливом роботи з перенесення стилів художників на зображення [6]. Основна ідея полягала в тому, щоб відокремити стиль зображення від його змісту за допомогою представлень зображення на початкових та останніх згорткових шарах глибокої нейронної мережі. Основним недоліком таких методів була втрата реалістичності навіть при перенесенні стилю з фотографії на фотографію. Згенероване зображення мало викривлені лінії, часто текстури переносились у невідповідні місця. Також обробка зображення була дуже довгою. Тому основними напрямками досліджень в цій групі є пошук шляхів підвищення фотореалістичності зображень та збільшення швидкодії алгоритмів.

Перенесення стилю базується на розв'язанні складної задачі оптимізації, яка може потребувати від декількох хвилин до годин обчислень на сучасному комп'ютері. Розповсюдженим шляхом пришвидшення цієї процедури є тренування іншої нейронної мережі, яка буде давати наближений розв'язок задачі оптимізації за один прохід вперед. Така обробка зображення значно швидша, проте втрачає

гнучкість відносно різноманітності стилів, тобто для кожного стильового зображення необхідно тренувати власну нейронну мережу. Відповідно до цього серед методів перенесення стилю можна виділити дві основні групи: повільні для довільного стилю (засновані на оптимізації) та швидкі для скінченної кількості певних стилів (засновані на апроксимуючій нейронній мережі).

У [6] було сформульовано перенесення стилю як задачу оптимізації, яка комбінує синтезування текстури з реконструкцією структури. Формулювання комбінує адитивні функції похибки, розташовані на декількох шарах попередньо натренованої згорткової нейронної мережі, з певними функціями похибки синтезу текстур стильового зображення та реконструкції семантики зображення. Стилізоване зображення отримується внаслідок оптимізації градієнтним методом, градієнти обчислювалися методом зворотнього розповсюдження похибки. Альтернативний підхід використовує схожість фрагментів стильового та вхідного зображень. Наприклад, у [7] було побудовано фрагментарну функцію похибки, де кожен згенерований фрагмент має найближчого сусіда серед вхідних фрагментів, якому він має відповідати. Цей тип фрагментарної функції похибки був скомбінований з формулюванням [6]. Ці методи дозволяють переносити стиль довільного зображення, проте процес оптимізації робить процес досить повільним.

В декількох роботах було натреновано обернені до попередньо тренованих згорткових нейронних мереж для візуалізації. В інших дослідженнях вони були треновані як частина автоенкодера. Для тренування було використано зображення та функції похибки у просторі RGB.

У [8] було запропоновано метод безпосереднього конструювання вихідних активацій певного шару попередньо тренованої CNN, знаходячи найближчі фрагменти у просторі активацій. Ця процедура детерміністична, тому може бути застосована до обробки відео без використання оптичного потоку для гарантування консистентності. Сформований шар активації можна перевести у вихідне зображення як шляхом оптимізації, так і використанням оберненої CNN. Даний алгоритм може бути застосований для довільного стильового зображення, а на відміну від інших методів з оберненою CNN, використовується функція похибки не

на основі рівня пікселів, а на рівні активацій. За умови певного попереднього налаштування, обернена мережа навіть може обробляти активації, які не належать початковій множині.

Отже, в області комп'ютерного зору сьогодні активно розвивається напрям інтелектуального редагування фотографій, у кожного класу методів є свої переваги та недоліки, які ще не в повній мірі досліджені, проте все одно дозволяють отримати придатні до використання результати і можуть бути застосовані до користувацького редагування фотографій. Дослідження шляхів покращення цих методів є актуальним як з наукової, так і з практичної точки зору.

1.3. Системний підхід до задачі

Задача створення системи фотореалістичного перенесення стилю в першу чергу характеризується суперечливістю вимог: необхідно досягти максимальної повноти перенесення стилю, зберігаючи при цьому фотореалістичність зображення. Також одразу виникає невизначеність, зумовлена неформалізованістю понять фотореалістичності та відповідності стилю.

Зображення слід розглядати як систему, представлену змістом та стилем, тісно пов'язаних між собою. Тому навіть при зміні лише однієї зі складових необхідно враховувати комплексний вплив перетворення.

Також, зважаючи на кількість параметрів моделі представлення зображення у просторі активацій нейронів глибокої мережі, дану задачу можна віднести до класу трансобчислювальної складності. Тому, для забезпечення раціональності розв'язку задачі, допускається наближене представлення замість пошуку глобального оптимуму.

Такі характеристики задачі обумовлюють використання системного підходу та відповідних методологій.

1.4. Формалізація постановки задачі

Метою даної роботи є розробка системи для зміни часу доби на фотографії за допомогою фотореалістичного перенесення стилів між зображеннями.

У даній роботі виділяємо такі підзадачі:

- аналіз існуючих алгоритмів зміни часу доби, погоди, сезону;
- побудова алгоритму фотореалістичного перенесення стилів між зображеннями;
- розробка системи для зміни часу доби на фотографії;
- практична реалізація алгоритму фотореалістичного перенесення часу доби між фотознімками;
- реалізація додатку із зручним для користувача графічним інтерфейсом для демонстрації роботи системи;
- проведення практичних досліджень роботи створеної системи.

Висновки за розділом

У першому розділі було обґрунтовано актуальність задачі фотореалістичного перенесення стилів між зображеннями та необхідність системного підходу.

Було проведено ознайомлення з основними існуючими підходами до підвищення якості фотознімків та їх редагування. Проаналізовано методи перенесення стилів між зображеннями, можливість підвищення їх фотореалістичності та труднощі при їх використанні. Також було формалізовано постановку задачі відповідно до мети даної роботи.

РОЗДІЛ 2 АЛГОРИТМИ ФОТОРЕАЛІСТИЧНОГО ПЕРЕНЕСЕННЯ СТИЛІВ

2.1 Представлення зображення у глибоких згорткових нейронних мережах

Згорткова нейронна мережа – нейронна мережа прямого поширення, що містить згорткові шари [9]. Ця архітектура була запропонована у 1988 році Я.Лекуном і показувала значно кращі результати у розпізнаванні рукописних цифр на маленьких зображеннях. З розвитком обчислювальних можливостей та появою дуже великих навчальних вибірок глибокі згорткові мережі показали значний прорив у задачі класифікації об'єктів на зображенні, значно випередивши всі попередні розробки на щорічному змаганні ImageNet, і вже за декілька років випередили людину в точності розпізнавання. Подальші дослідження показали, що згорткові нейромережі можна успішно застосовувати у задачах розпізнавання образів комп'ютерного зору.

Особливу увагу привернуло представлення зображення на різних згорткових шарах. Кожен рівень мережі представляє собою набір фільтрів, при чому початкові рівні активуються на простих патернах (наприклад лінії, градієнти), а з глибиною складність фільтрів росте і вони можуть виділяти колеса автомобілів, тварин, обличчя людей тощо. Тобто початкові рівні передають текстуру, а глибокі – зміст зображення (рис.2.1) [2].

2.2 Алгоритми перенесення стилів

Перенесення стилів між зображеннями є активною сферою досліджень. Більшість підходів базуються на глибоких загорткових мережах (Convolutional Neural Networks – CNN), що відрізняються методом оптимізації та функцією

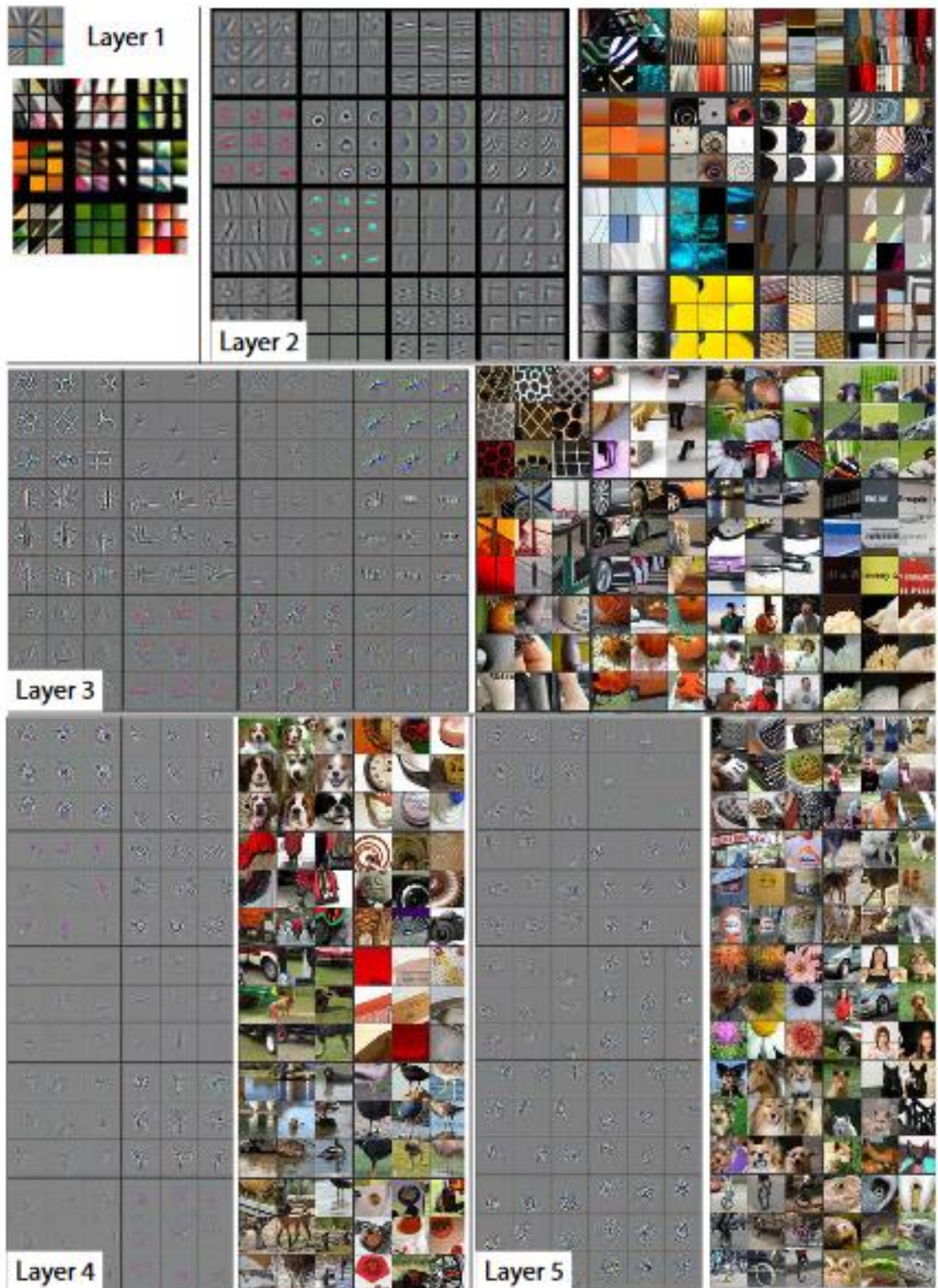


Рисунок 2.1 – Активації нейронів

похибки. Методи, що не використовують згорткові мережі, також були запропоновані, проте все ж не були кращими.

Алгоритми можна поділити на декілька категорій.

Методи, що створюють зображення шляхом прямої генерації зображення, дозволяють переносити стиль з довільного зображення на довільне інше, проте обробка одного зображення займає декілька хвилин[6].

Значно швидшими є попередньо треновані нейронні мережі для перенесення одного чи кількох фіксованих стилів, проте вони значно зменшують варіативність можливих згенерованих зображень.

Деякі поступаються у швидкості обробки, проте дозволяють переносити стиль з довільного зображення методами, що формують представлення вихідного зображення на певному шарі згорткової нейронної мережі з найбільш близьких фрагментів до початкових зображень та проводять зворотній хід по мережі для отримання результуючого зображення.

Також виділяються методи, які переносять агрегований стиль з деякої вибірки зображень[10,11].

2.2.1 Алгоритм NeuralStyle

Цей метод запропонували Gatys, Ecker та Bethge [6], тим самим поклавши початок цілому ряду досліджень та модифікацій методів перенесення стилів.

За основу було взято мережу VGG [12], яка була натренована на класифікацію об'єктів. Ваги мережі було нормалізовано так, щоб середнє значення активацій кожного згорткового фільтра по всім зображенням та позиціям було рівне 1. Було використано лише 16 згорткових та 5 агрегувальних шарів, повнозв'язні шари були відкинуті.

Кожен шар мережі представляє собою набір фільтрів, складність яких зростає зі збільшенням глибини шару. Довільне зображення \vec{x} може бути закодоване відповідними відкликами фільтрів. Для візуалізації інформації, яку кодує певний шар, можна застосувати до білого шуму метод градієнтного спуску, який наблизить до зображення з відповідним відкликом шару.

Нехай \vec{p} та \vec{x} – вихідне та згенероване зображення відповідно, P^l та F^l – їх представлення на шарі l . Можна задати квадратичну похибку між цими представленнями:

$$L_{content}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2 \quad (2.1)$$

Застосувавши таку похибку до достатньо глибокого шару, можна описати відхилення змісту згенерованого зображення від вихідного.

Для представлення стилю вихідного зображення використовується простір для опису текстур, що складається з кореляцій між відкликами різних фільтрів на будь-якому шарі нейронної мережі. Такі кореляції можна отримати у вигляді матриці Грама G^l :

$$G_{ij}^l = \sum_k F_{ik}^l F_{jk}^l \quad (2.2)$$

Агрегувавши кореляцію відкликів для декількох шарів, можна отримати постійне, масштабове представлення вихідного зображення, яке містить інформацію про його текстуру, але не про глобальну структуру.

Нехай \vec{a} – стильове зображення, A^l - стильове представлення шару l . Тоді похибка стилю визначається наступним чином:

$$L_{style}(\vec{a}, \vec{x}) = \sum_{l=0}^L w_l \frac{1}{4N_l^2 M_l^2} \sum_{i,j} (G_{ij}^l - A_{ij}^l)^2 \quad (2.3)$$

де w_l – вага вкладу шару l ;

N_l – кількість фільтрів шару l ;

M_l – кількість пікселів зображення на шарі l .

Тоді загальну похибку можна визначити як

$$L_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{content}(\vec{p}, \vec{x}) + \beta L_{style}(\vec{a}, \vec{x}) \quad (2.4)$$

де α – вага врахування змісту;

β – вага врахування стилю.

Таким чином, можна згенерувати зображення з бажаним стилем та семантикою шляхом мінімізації отриманої функції похибки.

2.2.2 Алгоритм StyleSwap

Chen та Schmidt [8] запропонували метод безпосереднього конструювання вихідних активацій певного шару попередньо тренованої CNN, знаходячи найближчі фрагменти у просторі активацій стильового зображення. Ця процедура детерміністична, тому може бути застосована до обробки відео без використання оптичного потоку для гарантування консистентності. Сформований шар активації можна перевести у вихідне зображення як шляхом оптимізації, так і використанням оберненої CNN. Даний алгоритм може бути застосований для довільного стильового зображення, а на відміну від інших методів з оберненою CNN, використовується функція похибки не на основі рівня пікселів, а на рівні активацій. За умови певного попереднього налаштування, обернена мережа навіть може обробляти активації, які не належать початковій множині.

Нехай C та S – представлення у просторі RGB вихідного та стильового зображень відповідно, та $\Phi(\bullet)$ – перетворення, що здійснюється повністю

згортковою частиною попередньо тренованої нейронної мережі до певного внутрішнього простору активацій.

Процедура заміни стилю полягає в наступному. З активацій $\Phi(C)$ та $\Phi(S)$ формуються множини фрагментів $\{\phi_i(C)\}$ та $\{\phi_j(S)\}$, які містять усі канали активацій та мають суттєво накладатися. Для кожного вихідного фрагменту визначається найближчий стильовий фрагмент на основі нормалізованої крос-кореляційної міри:

$$\phi_i^{SS}(C, S) = \operatorname{argmax}_{\phi_j(S), j=1 \dots n_s} \frac{\langle \phi_i(C), \phi_j(S) \rangle}{\|\phi_i(C)\| \|\phi_j(S)\|}, \quad (2.5)$$

де n_s – кількість фрагментів, виділених із стильового зображення.

Далі відбувається заміна кожного вихідного фрагменту $\phi_i(C)$ на найближчий стильовий $\phi_i^{SS}(C, S)$ та відновлюється весь шар активації $\Phi^{SS}(C, S)$ шляхом усереднення областей, що перекриваються.

Таким чином отримали внутрішній шар активацій, що відповідає змісту вихідного зображення та текстурам стильового.

Варто зазначити, що процедура заміни стилю може бути представлена у вигляді мережі, що використовує три операції: двовимірний згортковий шар, поканальний аргмаксимум та двовимірний обернений згортковий шар. А отже, можливе використання їх ефективних реалізацій.

Піксельне представлення стилізованого зображення може бути обчислене за допомогою введення функції похибки у просторі активацій з цільовими активаціями $\Phi^{SS}(C, S)$. Аналогічно до інших методів перенесення стилю використовується квадратична помилка:

$$I_{stylized}(C, S) = \operatorname{argmin}_I \|\Phi(I) - \Phi^{SS}(C, S)\|_F^2 + \lambda l_{TV}(I), \quad (2.6)$$

де $\|\cdot\|_F$ – норма Фробеніуса;

$l_{TV}(\cdot)$ – регуляризуючий доданок загальної варіації.

Він визначений наступним чином:

$$l_{TV}(I) = \sum_{i=1}^{h-1} \sum_{j=1}^w \sum_{k=1}^d (I_{i+1,j,k} - I_{i,j,k})^2 + \sum_{i=1}^h \sum_{j=1}^{w-1} \sum_{k=1}^d (I_{i,j+1,k} - I_{i,j,k})^2 \quad (2.7)$$

де h – висота зображення;

w – ширина зображення;

d – глибина зображення (у випадку RGB $d=3$).

Оскільки $\Phi(\bullet)$ містить декілька максимізуючих агрегувальних шарів, які зменшують розмір зображення, використання такої регуляризації дозволяє отримати просторово гладкіші результати відновленого зображення.

Оскільки $\Phi(\bullet)$ є частиною попередньо тренованої нейронної мережі та хоча б один раз субдиференційовна, то розв'язок задачі (2.6) можна отримати за допомогою стандартного субградієнтного методу оптимізації.

Для пришвидшення відновлення стилізованого зображення можна застосувати апроксимацію оптимального розв'язку використовуючи іншу нейронну мережу. Один раз натренована, така мережа може застосовуватись до будь-яких вихідних та стильових зображень. Для тренування мережі формується вибірка із зображень до і після заміни стилю та відповідних шарів активацій.

У випадку використання такої мережі майже весь час виконання алгоритму належить процесу обміну фрагментів і лінійно залежить від розміру стильового зображення. Інакше більшість обчислень припадає на задачу оптимізації та залежить від розміру вихідного зображення.

Варто зазначити, що навіть у випадку, коли стильове зображення є фотографією, отримане стилізоване зображення вже втрачає фотореалістичність (рис. 2.2).

Проте не зважаючи на це StyleSwar не поступається іншим методам перенесення стилю, адже достатньо гнучкий щодо зміни стильового зображення та потребує відносно небагато часових та обчислювальних ресурсів.



Рисунок 2.2 – Стилiзоване зображення

2.3 Семантична сегментація зображень

Однією з проблем при перенесенні стилю було неправильне співставлення текстур на вхідному та вихідному зображеннях. Наприклад, якщо на одному зображенні було багато будівель, а на другому – багато неба, то на вихідному зображенні можна було отримати відображення вікон на небі. Ефективним шляхом позбавлення від таких артефактів стало використання семантичної сегментації, яка дозволяла встановити відповідність між регіонами зображення, між якими коректно переносити текстуру.

Семантична сегментація - задача встановлення до якого класу об'єктів належить кожний піксель зображення, це один з кроків розуміння сцени. Очевидно, що ця задача складніша за класичну сегментацію, оскільки необхідно не тільки виділити однорідні частини зображення, а ще й вказати клас об'єктів, які вони визначають. Також вона і складніша за задачу класифікації, оскільки окрім класу

необхідно визначити точне просторове представлення об'єкту, його структури та частин.

Задачу семантичної сегментації можна розглядати декількома способами. В одному випадку вона зводиться до задачі класифікації кожного пікселя окремо. В іншому – до задачі визначення класів сегментів, виділених на зображенні.

Перші вдалі алгоритми семантичної сегментації базувалися на виділенні певних заданих характеристик та тренуваних на їх основі класифікаторів, наприклад бустингу, випадкового лісу чи машини опорних векторів. Суттєве покращення досягалося за допомогою врахування інформації з контексту зображення та технік структурованого передбачення, проте все одно вони обмежувалися малою виразністю та інформативністю обраних характеристик.

Досягнення глибокого навчання у класифікації зображень вплинули і на прогрес задачі семантичної сегментації, основне питання в тому, як скомбінувати задачі класифікації та сегментації. Алгоритми можна умовно поділити на декілька категорій.

Перша група методів застосовує до каскаду агрегуючих сегментацій класифікацію регіонів, яка вже буде також враховувати і форму сегменту. Деякі методи в якості сегментів використовують суперпкселі.

Друга група методів базується на використанні характеристик, отриманих за допомогою глибокої згорткової нейронної мережі, для класифікації кожного пікселя окремо. Для згладження передбачень може використовуватись ієрархія агрегуючих сегментацій.

Третя група методів застосовує глибокі згорткові нейронні мережі до всього зображення напряду для класифікації кожного пікселя, тобто повністю ігноруючи задачу сегментації. Для загорткових нейронних мереж з кожним шаром характерне збільшення глибини та зменшення розміру представлення зображення. Оскільки для семантичної сегментації необхідно, щоб на виході був початковий розмір зображення, то зазвичай використовуються архітектура кодувальник-декодувальник. Заміна повнозв'язних шарів на згорткові дозволяє обробляти зображення будь-якого розміру.

2.3.1 Метод DeepLab

На відміну від більшості архітектур кодувальник-декодувальник, у [13] було запропоновано інший підхід до семантичної сегментації. Нова архітектура DeepLab дозволяє контролювати зменшення розмірності з допустимою втратою інформації та виділення мультимасштабних контекстуальних ознак.

DeepLab використовує архітектуру ResNet як основу для екстрагування ознак, проте в останньому блоці звичайні згортки замінено на розширені з різним кроком. Це дозволяє врахувати контекст різного масштабу. Після цього блоку застосовується піраміда розширеної агрегації з різним кроком, що дозволяє класифікувати регіони будь-якого масштабу.

ResNet – дуже популярна архітектура глибоких згорткових нейронних мереж, основною перевагою якої є можливість створювати дуже глибокі мережі, не зазнаючи затухання градієнта. Оригінальна робота [14] представляє два типи залишкових модулів: базовий та вузький (рис. 2.3).

Пряма передача входу у нелінійний блок дозволяє отримати сильніший градієнт від глибших шарів до верхніх.

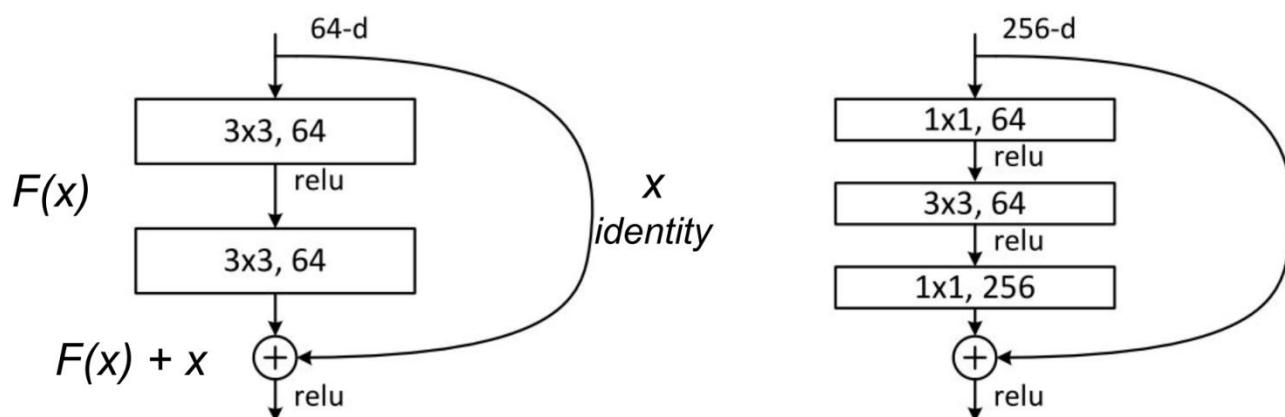


Рисунок 2.3 – Базовий та вузький ResNet модулі

Розширена згортка дозволяє врахувати ширший контекст, проте не збільшуючи при цьому кількість параметрів та обсяг обчислень (рис.2.4).

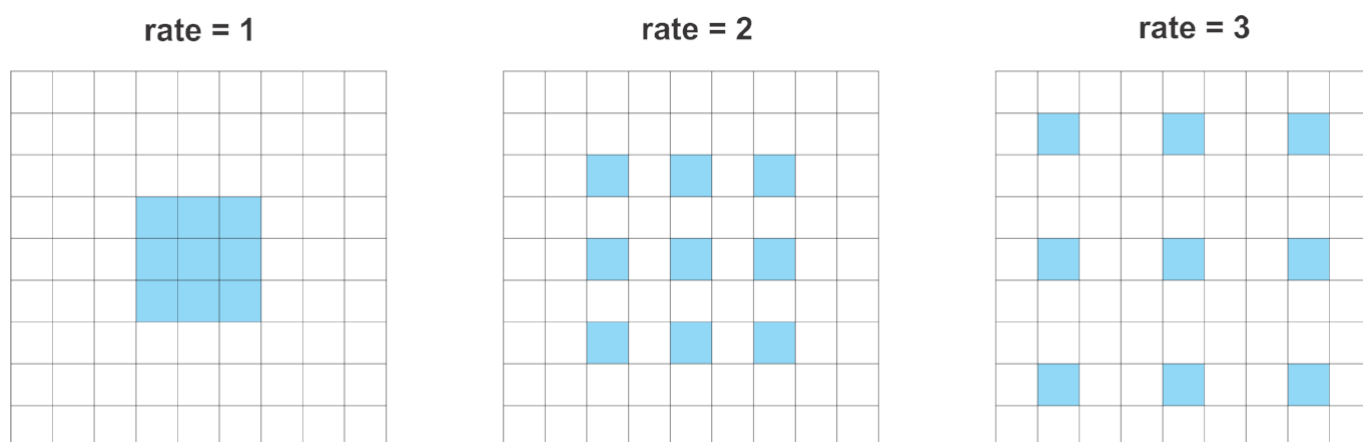


Рисунок 2.4 – Розширена згортка

2.4 Методи підвищення фотореалістичності

Нещодавно запропоновані методи перенесення стилю показали вражаючі результати, проте для них характерний брак фотореалістичності. Зображення дещо нечіткі, чим нагадують живопис, і тому є непридатними для багатьох випадків фоторедагування. Мета у обробці фотознімків дуальна: необхідно досягнути бажаного художнього ефекту, зберігаючи при цьому природність та реалістичність зображення. Більшість методів добре справляється з першою частиною, проте потерпає невдачу на другій.

2.4.1 Алгоритм на основі Matting Laplacian

У [15] запропонували регуляризувати цільову функцію для збереження структури вхідного зображення та отримання фотореалістичного результату. Принцип полягає в тому, щоб виражати це обмеження не напряму для вихідного зображення, а для перетворень, які застосовуються до вхідного зображення. Опис простору фотореалістичних зображень є важко формалізованою і досі не вирішеною задачею. Проте, зважаючи на те, що вже відомо, що вхідне зображення задовольняло умову реалістичності, достатньо слідкувати за тим, щоб не втратити цю характеристику під час його обробки. Тому можливі трансформації обмежили до локально афінних у просторі кольорів, тобто таких, що для кожного фрагменту вихідного зображення існує афінне перетворення, яке переводить вхідні значення RGB у відповідні вихідні значення. Кожний фрагмент може мати своє власне перетворення, що дозволяє просторову варіативність.

Метод спирається на Matting Laplacian, запропонований у [16], які показали вираження матування градацій сірого як локальну афінну комбінацію вхідних RGB каналів. В дослідженні було описано штрафну функцію у формі найменших квадратів, яку можна мінімізувати за допомогою стандартної системи лінійних рівнянь, представленої матрицею, що залежить лише від вхідного зображення.

$$L_m = \sum_{c=1}^3 V_c[O]^T M_I V_c[O] \quad (2.8)$$

де c – номер каналу;

$V_c[O]$ - векторизована форма каналу c вхідного зображення;

M_I -матуюча матриця;

Оскільки M_I - симетрична, то

$$\frac{dL_m}{dV_c[O]} = 2M_I V_c[O] \quad (2.9)$$

Це дозволяє використовувати градієнтні методи оптимізації.

У [15] запропонували розв'язати цю проблему шляхом двоступеневої оптимізації. На першому кроці відбувається перенесення стилю за допомогою NeuralStyle [6]. На другому кроці трансформація обмежується до локально афінної у просторі кольорів. Як зазначено у дослідженні, така двоступенева оптимізація краща за розв'язання вихідного функціоналу напряму, оскільки запобігає пригніченню локального перетворення сильною регуляризациєю фотореалістичності.

2.4.2 Алгоритм на основі екранованого рівняння Пуассона

Під час перенесення стилю з реальної фотографії, отримані зображення мають певні артефакти, які призводять до втрати реалістичності. Можна виділити їх три основних типи. Перший – поява чужорідних текстур на однорідних частинах зображення. Другий – викривлення структур, наприклад, хвилясті лінії замість прямих. Третій – втрата чіткості деталей, що призводить до ефекту мультфільму. Проте всі ці три типи об'єднуються тим, що виникають через спотворення градієнтів зображення. Однорідні області повинні мати невеликі градієнти, лінії мають відповідати градієнтам певного напрямку, а загальна фотореалістичність характеризується належністю певному статистичному розподілу градієнтів[17].

Для корекції градієнтів у [18] було запропоновано наближувати їх до градієнтів початкового зображення, зберігаючи при цьому новий стиль. Для цього цільова функція має два доданки: доданок відповідності, який контролює збереження нового стилю (характеристики стилю), та градієнтний доданок (2.10), який контролює відхилення від початкових градієнтів (характеристики реалістичності).

$$\int_{\Omega} \|\nabla O - \nabla C(x, y)\|^2 dx dy \quad (2.10)$$

$$\int_{\Omega} \|O - C_S\|^2 + \lambda \|\nabla O - \nabla C(x, y)\|^2 dx dy \rightarrow \min \quad (2.11)$$

$$\begin{aligned} L &= \|O - C_S\|^2 + \lambda \|\nabla O - \nabla C(x, y)\|^2 = \\ &= \|O - C_S\|^2 + \lambda(O_x - C_x)^2 + \lambda(O_y - C_y)^2 \end{aligned} \quad (2.12)$$

де $\nabla C(x, y)$ – градієнтне поле початкового зображення;

$O(x, y)$ – бажане відредагзоване зображення;

$C_S(x, y)$ - стилізоване зображення;

λ – параметр, що контролює відносну вагу доданку відповідності та градієнтного доданку.

$O(x, y)$, що мінімізує (2.11) має задовольняти рівняння Ейлера-Лагранжа:

$$\frac{\partial L}{\partial O} - \frac{\partial}{\partial x} \frac{\partial L}{\partial O_x} - \frac{\partial}{\partial y} \frac{\partial L}{\partial O_y} = 0 \quad (2.13)$$

Звідки маємо екрановане рівняння Пуассона:

$$O - \lambda \nabla^2 O = C_S - \lambda \nabla^2 C(x, y), \quad (2.14)$$

$$(O - C_S) - \lambda(O_{xx} - C_{xx}) - \lambda(O_{yy} - C_{yy}) = 0. \quad (2.15)$$

Це рівняння визначає систему лінійних рівнянь, яку можна розв'язати методом найменших квадратів, перетворенням Фур'є[19] або за допомогою згортки[20].

Зображення представлені у просторі кольорів Lab, SPE розв'язується для кожного каналу окремо з $\lambda=5$ для каналу L та $\lambda=1$ для каналів a та b.

Метод базується на екранованому рівнянні Пуассона (Screened Poisson Equation - SPE), вперше запропонованому у [19] для фільтрування зображення. Пізніше у [21] його використовували для відновлення зображень та суміщення погано узгоджених зображень. У їх роботах кольори та градієнти генерувалися окремо та потім об'єднувалися за допомогою SPE. В іншій роботі [22] SPE

застосовується для підвищення контрастності. У [23] було запропоновано використовувати схожі ідеї для внесення фотореалістичності у перетворення зображень.

2.5 Система оцінки якості генератора

Основні характеристики зображення, отриманого за допомогою генератора, - реалістичність, повнота перенесення стиля і збереження змісту – суто суб'єктивні, тому коректніше їх оцінювати за допомогою опитування користувачів.

У попередніх дослідженнях перенесення стилів між зображеннями використовували декілька форматів опитування. Користувачам пропонували оцінити реалістичність за шкалою від 1 до 4 або визначити, чи було оброблене фото, для серії зображень, оцінити схожість стилів для серії пар зображень.

В даній роботі кожному опитуваному надавалося дві серії зображень. В першій серії необхідно оцінити фотореалістичність зображення за шкалою від 1 до 4, при чому серед зображень як трансформовані зразки, так і реальні фотографії. В другій серії надаються пари зображень – до і після обробки, для яких потрібно оцінити, чи відповідає результат режиму обробки. З отриманих оцінок визначався середній показник фотореалістичності та ефективності.

2.6 Архітектура системи фотореалістичного перенесення часу доби

Для вирішення задачі зміни часу доби на фотознімках пропонується перенесення стилю методом StyleSwap з використанням сегментації DeepLab та подальшою обробкою методом екранованого рівняння Пуассона.

2.6.1 Параметри методу StyleSwap

Для використання методу StyleSwap необхідно було визначити певні гіперпараметри (тут було використано назви відповідних аргументів скрипта, що програмно реалізовує метод):

- patchSize: розмір фрагментів, між якими відбувається обмін (чим більші фрагменти, тим агресивніше переноситься стиль), тут було обрано мінімальне значення – 3;

- patchStride: крок, з яким обираються фрагменти (збільшення параметру дозволяє зменшити кількість необхідної для обчислень пам'яті), тут було обрано мінімальне значення – 1;

- pooling: вид агрегувальних шарів (максимізаційні чи усереднюючі), тут було обрано максимізаційні;

- decoder: чи використовувати попередньо треновану зворотню нейронну мережу для декодування перетвореного шару активацій, тут було обрано декодування шляхом оптимізації;

- learningRate: крок навчання під час оптимізації, тут було обрано 0.05;

- init: спосіб ініціалізації зображення для оптимізації (випадковими числами чи початковим зображенням), тут було обрано заповнення значеннями початкового зображення;

- layer: шар нейронної мережі, на якому буде відбуватися заміна фрагментів, тут було обрано relu3_1 (при використанні більш ранніх шарів відбувається перенесення дрібніших текстур, проте значно збільшується кількість необхідної для обчислень пам'яті);

- optimIter: кількість ітерацій оптимізації, тут було обрано 100.

Для встановлення необхідних гіперпараметрів було проведено низку експериментів, з яких можна зробити наступні висновки.

По-перше, варто зазначити, що очікувати прийнятний результат можна лише за умови коректного вхідного зображення, тобто в даному випадку це має бути фотографія міського пейзажу.

Оскільки система має на меті перетворювати погоду фотографії на сонячну, то необхідно обрати відповідне стильове зображення (рис 2.5).



Рисунок 2.5 – Стильове зображення

Проте стилізовані зображення (тут і надалі мається на увазі зображення вже після відновлення фотореалістичності) можуть бути дуже різні, хоч і стильові фотографії виглядають схожими – зображають сонячне місто. Аналогічно результат може бути нестабільний для різних вхідних зображень. Це викликано тим, що варіативність вхідних зображень не відповідає варіативності одного стильового зображення, тобто серед фрагментів стильового зображення не завжди вдається підібрати вдалий відповідник. Для підвищення робустності моделі можна створити базу фрагментів, зібраних з набору стильових зображень, розділених на основні семантичні групи, всередині яких і будуть шукатися відповідники для сегментів вхідного зображення.

Тривалість процесу заміни стилю, а саме пошуку відповідних фрагментів, лінійно залежить від розміру стильового зображення, тому звуження множини пошуку за рахунок сегментації дозволяє підвищити швидкодію.

Для декодування перетворених активацій можна використовувати метод оптимізації (рис. 2.6) чи застосувати попередньо треновану обернену нейронну мережу (рис. 2.7), яка надасть наближене до оптимального зображення.



Рисунок 2.6 – Стилізоване зображення з використанням методу оптимізації

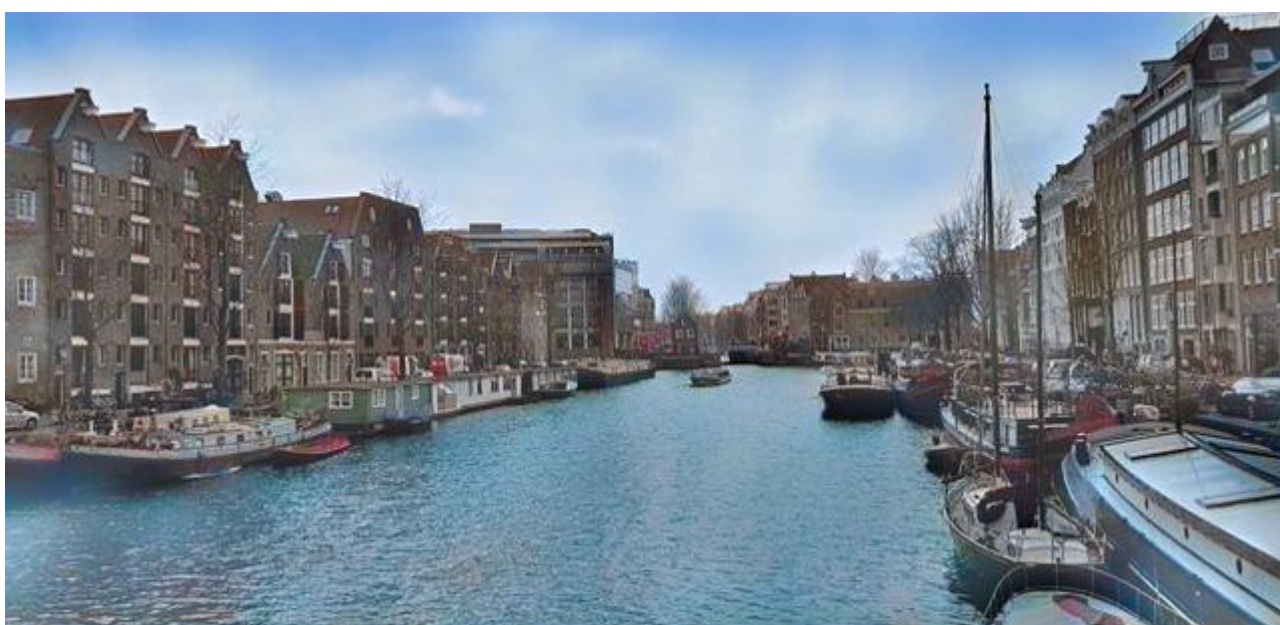


Рисунок 2.7 – Стилізоване зображення з використанням декодера

Хоч перший варіант і значно повільніший (5 секунд проти 1,5 при обчисленні на GPU), проте в другому варіанті часто з'являються кольорові плями та відблиски. Проте варто зазначити, що використаний тут декодер з оригінального дослідження [8] був тренований на вибірці з фотографій та художніх картин, тому відновлює зображення з врахуванням особливостей цих видів зображень. Можливим шляхом покращення оберненої трансформації може бути тренування декодера виключно на фотографіях з цільової множини. Проте в межах даного дослідження було обрано декодування шляхом оптимізації.

2.6.2 Використання сегментації

В деяких випадках сегментація не впливає на кінцевий результат. Такий ефект може спостерігатися, коли на вихідному та стильовому зображеннях кожному семантичному сегменту вже відповідають різні текстури, а на зображеннях схожі сцени. В цьому випадку навіть можна обробляти вихідне зображення у градаціях сірого – необхідні відповідні фрагменти визначаються лише за рахунок яскравості і стилізоване зображення набуде кольорів стильового.

Проте у випадку, коли на вихідному зображенні було кілька різних семантичних сегментів з однаковою текстурою, а на стильовому лише один, то на стилізованому зображенні текстури перенесуться некоректно. Наприклад, якщо на зображенні були сегменти з поверхнею води та піщаним пляжем (рис. 2.8), то на стилізованому зображенні отримаємо голубий пісок (рис 2.9).



Рисунок 2.8 – Приклад однакових текстур різних семантичних сегментів



Рисунок 2.9 – Хибне перенесення текстури без використання сегментації

При використанні сегментації вдається запобігти хибному перенесенню текстури води на пляж (рис 2.9).



Рисунок 2.9 – Стилізоване зображення з використанням сегментації

Висновки за розділом

В даному розділі було проаналізовано алгоритми та методи перенесення стилю між зображеннями та шляхи підвищення фотореалістичності, а саме використання сегментації та відновлення градієнтів.

Сформовано систему критеріїв оцінки якості генератора на основі експертного оцінювання для подальшого аналізу роботи програми. Було запропоновано архітектуру системи для зміни часу доби на фотографіях, досліджено вплив параметрів на стилізоване зображення та визначено їх оптимальні значення.

РОЗДІЛ 3 АРХІТЕКТУРА ПРОГРАМНОГО ПРОДУКТУ ТА АНАЛІЗ ПРАКТИЧНИХ РЕЗУЛЬТАТІВ

3.1 Вибір інструментів розробки

Для попередньої обробки зображень було обрано OpenCV (Open Source Computer Vision Library) – кросплатформену бібліотеку алгоритмів комп'ютерного зору з відкритим кодом, вільну до використання у наукових та комерційних цілях [24]. Вона надає засоби для обробки зображень та їх аналізу, починаючи від таких простих перетворень як фільтрація зображень, закінчуючи алгоритмами із застосуванням методів машинного навчання та підтримкою фреймворків для глибоких нейронних мереж. Бібліотека написана на оптимізованому C та C++, розроблена з акцентом на високу продуктивність та використання у додатках реального часу. Також доступне використання можливостей багатопоточного виконання та прискорення на основі GPU. Бібліотека працює під Linux, Windows та Mac OS X, доступні версії для мобільних пристроїв на базі iOS або Android, підходить для програмування мовами Python, C++, Matlab. В роботі використано версію бібліотеки OpenCV 2.4.11 для програмування мовою Python в операційній системі Ubuntu.

Для частини стилізації зображення було використано Torch — відкриту бібліотеку для машинного навчання з підтримкою використання GPU [25]. Бібліотека орієнтована на максимальну гнучкість та швидкість під час розробки користувацьких алгоритмів, зберігаючи при цьому простоту самого процесу, що досягається використанням мови сценаріїв LuaIT та реалізацією мовою C в основі. Torch базується на популярних бібліотеках нейронних мереж та методів оптимізації, які прості у використанні та максимально гнучкі у реалізації складних топологій нейронних мереж. Користувач може будувати довільні графи нейронних мереж та ефективно виконувати паралельні обчислення за допомогою CPU та GPU. Доступна велика система пакетів машинного навчання, комп'ютерного зору, обробки сигналів, паралельних обчислень на основі мови програмування Lua. Бібліотека

активно розвивається, використовується такими великими компаніями як Google, Facebook, Twitter та багатьма дослідницькими лабораторіями.

Модуль підвищення реалістичності реалізований мовою Python з використанням бібліотеки NumPy [26], що надає можливість ефективної обробки великих багатовимірних масивів та використання великої кількості оптимізованих алгоритмів для роботи з ними, які за швидкістю наближені до відповідного коду мовою C. Доступні пакети методів оптимізації, лінійної алгебри, інтегрування, інтерполяції, перетворення Фур'є, обробки зображень та сигналів, розв'язання звичайних диференціальних рівнянь тощо. За допомогою C API можна створювати мовою C власні оптимізовані модулі для Python. Внутрішньо NumPy базується на бібліотеці LAPACK, призначеної для вирішення основних задач лінійної алгебри. Також є інструменти візуалізації даних. Таким чином, NumPy можна розглядати як гарну вільну альтернативу MATLAB.

3.2 Аналіз архітектури продукту

Для демонстрації роботи системи було необхідно створити додаток, який надавав би можливість завантажити фотографію та відредагувати у режимі «перетворити на сонячний день».

Запропонована система перенесення стилю має достатньо високі вимоги до технічних характеристик обчислювального пристрою. Прийнятна швидкість обробки зображення можлива з використанням GPU певної потужності. Сьогодні на ринку наявний вибір моделей, які задовольняють цим вимогам, проте все одно вони недостатньо поширені серед цільової аудиторії. Тому доцільно використовувати клієнт-серверну архітектуру сервісу: приймати користувачьке зображення, обробляти його на достатньо потужному віддаленому сервері та відправляти назад.

Зв'язок з сервером відбувається за допомогою мережі Інтернет, доступ до якої сьогодні наявний практично у кожного в будь-якому місці.

Сьогодні одним з розповсюджених і зручних для користувача та розробника методів організації клієнт-серверної архітектури сервісу є використання бота месенджера Telegram. Бот фактично є клієнтом месенджера, за допомогою відповідного API відправляє запити та отримує на них відповіді, тобто може відправляти іншим користувачам повідомлення та отримувати надіслані йому.

Таким чином, користувач відправляє боту повідомлення з фотознімком, бот обробляє отримане зображення та відправляє результат назад. Взаємодія з сервісом відбувається швидко та зручно через інтерфейс месенджера. Оскільки Telegram доступний у версіях як для мобільних пристроїв, так і для стаціонарних комп'ютерів, то і для створеного сервісу доступні ці пристрої, єдиною вимогою залишається підключення до мережі інтернет.

За допомогою спеціального бота BotFather можна отримати унікальний токен-ідентифікатор для нового користувацького бота, за допомогою якого буде відбуватися аутентифікація на сервері.

Програмна реалізація бота написана мовою Python. Оскільки частина перенесення стилю реалізована мовою Lua, то це перетворення відбувається шляхом виклику команди `bash` з основного скрипту бота.

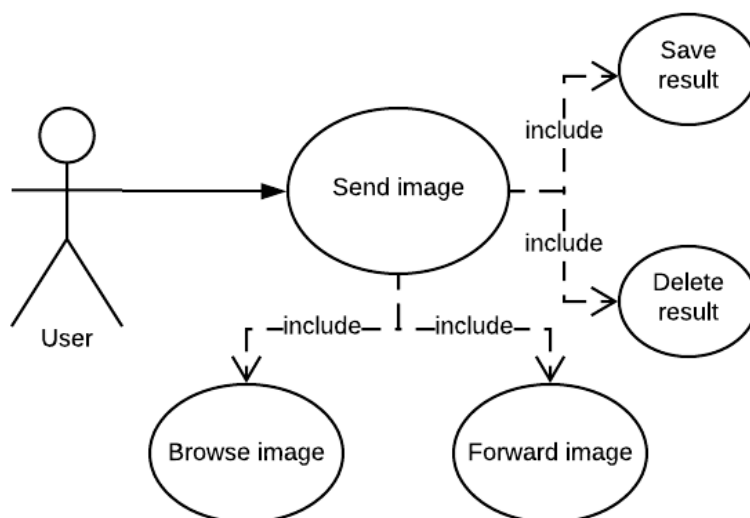


Рисунок 3.1 – UML-діаграма використання програмного продукту

Розглянемо алгоритм обробки запиту користувача (рис.3.2). Спочатку відбувається семантична сегментація зображення для коректної обробки кожної частини. Далі для кожного сегмента генерується його стилізована версія та результати об'єднуються у єдине зображення, після чого проводиться відновлення фотореалістичності.



Рисунок 3.2 – Алгоритм обробки повідомлення

3.3 Інтерфейс користувача

Інтерфейс користувача було реалізовано у вигляді бота месенджера Telegram.

Для початку користування необхідно знайти бота через опцію пошуку по користувачам за його назвою та почати діалог з ним, натиснувши кнопку Start, або надіславши команду /start (цей крок необхідний лише при першому зверненні).

Для того, щоб обробити зображення, необхідно надіслати його повідомленням боту або переслати з іншого діалогу (будь-який текст та інші види повідомлень будуть проігноровані). У відповідь бот надішле відредагований знімок (рис. 3.3).



Рисунок 3.3 – Графічний інтерфейс користувача

Таким чином, розроблений інтерфейс дозволяє зручно змінювати час доби на фотографіях та оцінити роботу розробленої системи.

3.4 Аналіз практичних результатів

Для оцінки якості роботи системи у функціонал бота було додано тестовий режим з опцією оцінки: разом з обробленим зображенням користувачу відправлялося запитання, наскільки фотореалістичним було зображення та наскільки відповідає режиму «сонячний день» з варіантами оцінки від 1 до 4. Також на сервер було завантажено підбірку попередньо оброблених алгоритмом чи популярними фільтрами зображень (рис 3.4) і користувач, викликавши команду /test, міг отримати випадкове фото з бази та оцінити його реалістичність та сонячність.



Рисунок 3.4 – Оригінальне зображення

Для порівняння було обрано два фільтри, наближених за своїм ефектом до режиму «сонячний день»: фільтр «Morning» (рис 3.5) та фільтр «Сонячне світло» (рис 3.6).



Рисунок 3.5 – Зображення оброблене фільтром «Morning»



Рисунок 3.6 – Зображення оброблене фільтром «Сонячне світло»

Таким чином було отримано оцінки 47 зображень. Для порівняння алгоритмів було використано частку фотографій, оцінених на 3 та 4, тобто як «скоріше фотореалістично» та «фотореалістично», «скоріше сонячно» та «сонячно» відповідно.

Таблиця 3.1 – Оцінка методів фоторедагування за обраними критеріями

Алгоритм	Фотореалістичність	Сонячність
SunnyWeatherBot	64%	78%
Фільтр «Morning»	83%	69%
Фільтр «Сонячне світло»	77%	65%

З отриманих значень можна зробити висновок, що все ж зі збільшенням ефекту «сонячний день» все ще погіршується фотореалістичність (рис. 3.7). Можна відмітити, що такий ефект може виникати за рахунок неприродності отриманих відтінків, іноді відбувається хибне перетворення легких хмарок на небі і вони навпаки стають темнішими за небо.



Рисунок 3.7 – Зображення оброблене за допомогою SunnyWeatherBot

Висновки за розділом

В даному розділі описано структурну схему розробленого програмного продукту, обґрунтовано вибір інструментів розробки, надано коротку інструкцію з експлуатації додатку. Програмний продукт надає можливість відредагувати зображення в режимі «перетворити на сонячний день» та зберегти результат.

За отриманими результатами роботи можна зробити висновок, що програмний продукт досягає поставлені цілі у достатній мірі та значно переважає доступні існуючі інструменти.

Перспективами розвитку програмного продукту є його адаптація до використання на пристрої клієнта, підвищення швидкості обробки зображення, дослідження додаткової обробки класичними автоматичними методами фоторедагування.

РОЗДІЛ 4 РОЗРОБЛЕННЯ СТАРТАП ПРОЕКТУ

Стартап як форма малого ризикового (венчурного) підприємництва впродовж останнього десятиліття набула широкого розповсюдження у світі через зниження бар'єрів входу в ринок (із появою Інтернету як інструменту комунікацій та збуту стало простіше знаходити споживачів та інвесторів, займатись пошуком ресурсів, перетинати кордони між ринками різних країн), і вважається однією із наріжних складових інноваційної економіки, оскільки за рахунок мобільності, гнучкості та великої кількості стартап-проектів загальна маса інноваційних ідей зростає.

Проте створення та ринкове впровадження стартап-проектів відзначається підвищеною мірою ризику, ринково успішними стає лише невелика частка, що за різними оцінками складає від 10% до 20%. Ідея стартап-проекту, взята окремо, не вартує майже нічого: головним завданням керівника проекту на початковому етапі його існування є перетворення ідеї проекту у працюючу бізнес-модель, що починається із формування концепції товару (послуги) для визначеної клієнтської групи за наявних ринкових умов.

Розроблення та виведення стартап-проекту на ринок передбачає здійснення низки кроків, в межах яких визначають ринкові перспективи проекту, графік та принципи організації виробництва, фінансовий аналіз та аналіз ризиків і заходи з просування пропозиції для інвесторів. Далі наведено маркетинговий аналіз стартап проекту. В межах цього етапу:

а) розробляється опис самої ідеї проекту та визначаються загальні напрями використання потенційного товару чи послуги, а також їх відмінність від конкурентів;

б) аналізуються ринкові можливості щодо його реалізації;

в) на базі аналізу ринкового середовища розробляється стратегія ринкового впровадження потенційного товару в межах проекту.

4.1 Опис ідеї проекту

В межах підпункту було проаналізовано і подано у вигляді таблиць:

- а) зміст ідеї (що пропонується);
- б) можливі напрямки застосування;
- в) основні вигоди, що може отримати користувач товару (за кожним напрямком застосування);
- г) чим відрізняється від існуючих аналогів та замінників.

Перші три пункти подані у вигляді таблиці (таблиця 4.1) і дають цілісне уявлення про зміст ідеї та можливі базові потенційні ринки, в межах яких потрібно шукати групи потенційних клієнтів.

Таблиця 4.1 - Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
Розроблену систему можна використовувати для швидкого інтелектуального редагування фотознімків	1. Користувачами фотокамер для виправлення невдалих кадрів 2. Професійними фотографами	Додавання виразності фотографіям шляхом зміни погоди Зменшення часу на обробку кадрів

Аналіз потенційних техніко-економічних переваг ідеї (чим відрізняється від існуючих аналогів та замінників) порівняно із пропозиціями конкурентів передбачає:

- а) визначення переліку техніко-економічних властивостей та характеристик ідеї;
- б) визначення попереднього кола конкурентів (проектів-конкурентів) або товарів-замінників чи товарів-аналогів, що вже існують на ринку, та проводиться збір інформації щодо значень техніко-економічних показників для ідеї власного проекту та проектів-конкурентів відповідно до визначеного вище переліку;

в) проводиться порівняльний аналіз показників: для власної ідеї визначаються показники, що мають гірші значення (W, слабкі), аналогічні (N, нейтральні) значення, кращі значення (S, сильні) (табл. 4.2).

Таблиця 4.2 - Визначення сильних, слабких та нейтральних характеристик ідеї проекту

Техніко-економічні характеристики ідеї	Товари/концепції конкурентів		W	N	S
	Мій проект	Photoshop	(слабка сторона)	(нейтральна сторона)	(сильна сторона)
Кросплатформеність	Можливість використання на різних ОС	Тільки Windows			+
Зручність використання/орієнтованість на кінцевого споживача	Зручний інтерфейс з підтримкою укр, англ та рос мов.	Тільки англійська мова.			+
Універсальність	Обмежені функції: зміна часу доби	Наявні інші функції	+		

Визначений перелік слабких, сильних та нейтральних характеристик та властивостей ідеї потенційного товару є підґрунтям для формування його конкурентоспроможності.

4.2 Технологічний аудит ідеї проекту

В межах даного підрозділу було проведено аудит технології, за допомогою якої можна реалізувати ідею проекту (технології створення товару). Визначення технологічної здійсненності ідеї проекту передбачає аналіз таких складових (таблиця 4.3):

- а) за якою технологією буде виготовлено товар згідно ідеї проекту;
- б) чи існують такі технології, чи їх потрібно розробити/доробити;
- в) чи доступні такі технології авторам проекту.

Таблиця 4.3 - Технологічна здійсненність ідеї проекту

Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
Створення програмного забезпечення для інтелектуального редагування фотографій	Алгоритми перенесення стилів Нейронні мережі Web-платформа	Наявна Наявна Наявна	Доступна Доступна Доступна

За результатами аналізу таблиці зроблено висновок щодо можливості технологічної реалізації проекту. Технологічним шляхом реалізації проекту було обрано такі технології, як Tensorflow та OpenCV через їх доступність та безкоштовність.

4.3 Аналіз ринкових можливостей запуску стартап-проекту

Визначення ринкових можливостей, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє спланувати напрями розвитку проекту із урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів.

Спочатку було проведено аналіз попиту: наявність попиту, обсяг, динаміка розвитку ринку (таблиця 4.4).

Таблиця 4.4 - Попередня характеристика потенційного ринку стартап-проекту

Показники стану ринку (найменування)	Характеристика
Кількість головних гравців, од	1
Загальний обсяг продаж, грн/ум.од	150000
Динаміка ринку (якісна оцінка)	Зростає
Наявність обмежень для входу (вказати характер обмежень)	-
Специфічні вимоги до стандартизації та сертифікації	-
Середня норма рентабельності в галузі (або по ринку), %	18

Середню норму рентабельності в галузі було порівняно із банківським відсотком на вкладення. Останній є меншим, тому є сенс вкладати гроші саме у цей проект.

За результатами аналізу таблиці 4.4 було зроблено висновок, що ринок є привабливим для входження.

Надалі були визначені потенційні групи клієнтів, їх характеристики та зформовано орієнтовний перелік вимог до товару для кожної групи (табл. 4.5).

Таблиця 4.5 - Характеристика потенційних клієнтів стартап-проекту

Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
Програмне забезпечення для інтелектуального редагування фотографій	Компанії, приватні підприємства, специфіка роботи яких пов'язана з обробкою фотографій.	Відмінності у сферах діяльності компаній та приватних підприємств, підходи до створення контенту.	Зручний інтерфейс, швидкість і надійність у використанні.

Після визначення потенційних груп клієнтів було проведено аналіз ринкового середовища: складено таблиці факторів, що сприяють ринковому впровадженню проекту, та факторів, що йому перешкоджають (табл. 4.6, 4.7).

Визначено можливу реакцію компанії на конкуренцію, зміну потреб користувачів та появу нових методів з врахуванням можливостей ринку та вітчизняних особливостей.

Таблиця 4.6 - Фактори загроз

Фактор	Зміст загрози	Можлива реакція компанії
Конкуренція	Вихід на ринок продуктів з кращими характеристиками	Передбачити додаткові переваги власного програмного продукту (ПП) для того, щоб повідомити про них саме після виходу на ринок конкурентів. Вдосконалення технічних моментів власного продукту. Обрати нову цільову аудиторію і зосередитися на ній: зниження цін.
Зміна потреб користувачів	Користувачам необхідне програмне забезпечення з іншим функціоналом	Передбачити можливість додавання нового функціоналу до створюваного ПП

Таблиця 4.7 - Фактори можливостей

Фактор	Зміст можливості	Можлива реакція компанії
Конкуренція	Відсутність аналогічного продукту для вітчизняного користувача.	Адаптація програмного продукту до вітчизняних особливостей.
Поява нових методів обробки фотографій	З'являться нові методи, що будуть швидше та ефективніше обробляти фотографії	Покращити ПП додаванням нового функціоналу, розширення можливостей

Надалі було проведено аналіз пропозиції: визначили загальні риси конкуренції на ринку (таблиця 4.8).

Таблиця 4.8 - Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1. Вказати тип конкуренції - монополія	На ринку присутні декілька компаній-конкурентів, але їх товар дещо відрізняється між собою.	Підтримка якості продукту та постійні нововведення, вдосконалення.
2. За рівнем конкурентної боротьби - міжнародний	Компанії-конкуренти з інших країн	Створити основу ПП таким чином, щоб можна було легко переробити даний ПП для використання у галузях інших країн.
3. За галузевою ознакою - міжгалузева	Продукт може використовуватись для різних галузей	Постійне вдосконалення продукту, що не має прив'язки до сфери
4. Конкуренція за видами товарів: - товарно-видова	Конкуренція між видами ПП, їх особливостями.	Створити ПП, враховуючи недоліки конкурентів
5. За характером конкурентних переваг - нецінова	Вдосконалення технології створення ПП, щоб собівартість була нижчою	Удосконалення моделі. Використання більш дешевих технологій для розробки, ніж використовують конкуренти, але тільки якщо ці технології відповідають необхідним вимогам якості.
6. За інтенсивністю - не марочна	Бренд присутній, але його роль незначна	Реклама, участь у конференціях, семінарах.

Було проведено аналіз конкуренції у галузі за моделлю М. Портера (табл. 4.9).

Таблиця 4.9 - Аналіз конкуренції в галузі за М. Портером

Складові аналізу	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
	Навести перелік прямих конкурентів	Визначити бар'єри входження в ринок	Визначити фактори сили постачальників	Визначити фактори сили споживачів	Фактори загроз з боку замінників
	Photoshop	Наявність вже існуючих рішень	-	Контроль якості продукту	Наявність більш широкого функціоналу, зручнішого інтерфейсу та авторитет (перевірена якість)
Висновки:	Досить інтенсивна конкурентна боротьба з вже закріпивши мися на ринку гравцями	Є можливості виходу на ринок, але є і конкуренти. Строки – 18 місяців.	-	Клієнти диктують умови роботи на ринку: зручний інтерфейс, надійний, швидкий, точний та достовірний ПП для побудови моделей і прогнозів.	Необхідно випускати ПЗ не гірше, ніж у конкурентів та розширяти функціонал.

За результатами аналізу табл. 4.9 було зроблено висновок про можливість роботи на ринку з огляду на конкурентну ситуацію. Також було зроблено висновок щодо характеристик, які повинен мати проект, щоб бути конкурентноспроможним на ринку.

Цей висновок був врахований при формулюванні переліку факторів конкурентноспроможності у наступному пункті. На основі аналізу конкуренції, проведеного в табл. 4.9, а також із урахуванням характеристик ідеї проекту (табл.

4.2), вимог споживачів до товару (табл. 4.5) та факторів маркетингового середовища (таблиці 4.6, 4.7) визначається та обґрунтовується перелік факторів конкурентоспроможності. Аналіз оформлено у табл. 4.10.

Таблиця 4.10 - Обґрунтування факторів конкурентоспроможності

Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
Ціна	Більш доступна ціна збільшує кількість потенційних клієнтів
Кросплатформність ПП	Можливість використання програмного забезпечення на будь-якій платформі.
Орієнтованість на кінцевого споживача	Продукт орієнтований на взаємодію з клієнтом

За визначеними факторами конкурентоспроможності (табл. 4.10) проведено аналіз сильних та слабких сторін стартап-проекту (табл. 4.11).

Таблиця 4.11 - Порівняльний аналіз сильних та слабких сторін проекту

Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні						
		-3	-2	-1	0	1	2	3
Ціна	15	*						
Кросплатформність ПП	20			*				
Орієнтованість на кінцевого споживача	7						*	

Фінальним етапом ринкового аналізу можливостей впровадження проекту є складання SWOT-аналізу (матриці аналізу сильних (Strength) та слабких (Weak) сторін, загроз (Troubles) та можливостей (Opportunities) (таблиця 4.12) на основі виділених ринкових загроз та можливостей, та сильних і слабких сторін (таблиця 4.11). Перелік ринкових загроз та ринкових можливостей було складено на основі аналізу факторів загроз та факторів можливостей маркетингового середовища.

Ринкові загрози та ринкові можливості є наслідками (прогнозованими результатами) впливу факторів, і, на відміну від них, ще не є реалізованими на ринку та мають певну ймовірність здійснення. Наприклад: зниження доходів потенційних споживачів – фактор загрози, на основі якого можна зробити прогноз щодо посилення значущості цінового фактору при виборі товару та відповідно, – цінової конкуренції (а це вже – ринкова загроза).

Таблиця 4.12 - SWOT-аналіз стартап-проекту

Сильні сторони: Ціна Орієнтованість на кінцевого споживача	Слабкі сторони: Кросплатформність ПП
Можливості: Конкуренція Поява нових методів обробки фотографій	Загрози: Зміна потреб користувачів

На основі SWOT-аналізу було розроблено альтернативи ринкової поведінки (перелік заходів) для виведення стартап-проекту на ринок та орієнтовний оптимальний час їх ринкової реалізації з огляду на потенційні проекти конкурентів, що можуть бути виведені на ринок (див. таблицю 4.9, аналіз потенційних конкурентів). Визначені альтернативи були проаналізовані з точки зору строків та ймовірності отримання ресурсів (таблиця 4.13).

Таблиця 4.13 - Альтернативи ринкового впровадження стартап-проекту

Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
Безкоштовне розповсюдження створеного ПП	45%	18 місяців
Створення ПП з подальшим розповсюдженням за певну оплату	85%	18 місяців
Створення вебсайту, на якому можна буде користуватися ПП	75%	16 місяців

Після аналізу було обрано альтернативу №2.

4.4 Аналіз ринкової стратегії проекту

Розроблення ринкової стратегії першим кроком передбачає визначення стратегії охоплення ринку: було проведено опис цільових груп потенційних споживачів (таблиця 4.14).

Таблиця 4.14 - Вибір цільових груп потенційних споживачів

Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
Компанії (українські та міжнародні) діяльність яких пов'язана з обробкою фотографій	Висока	Високий	Сильна	Складно
Приватні підприємства міського та міжнародного рівня, діяльність яких пов'язана з обробкою фотографій	Висока	Високий	Сильна	Складно
Приватні підприємства, обласного рівня.	Помірна	Помірний	Помірна	Середня складність
Підприємства регіонального характеру	Помірна	Слабкий	Слабка	Просто

За результатами аналізу потенційних груп споживачів було обрано цільові групи, для яких буде запропоновано даний товар, та визначено стратегію охоплення ринку - стратегію диференційованого маркетингу (компанія працює з декількома сегментами). Як цільові групи обрано: 1,2,3.

Для роботи в обраних сегментах ринку сформовано базову стратегію розвитку (таблиця 4.15).

Таблиця 4.15 - Визначення базової стратегії розвитку

Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку
	Визначити потреби кожної з груп, розробити відповідно до них стратегії приваблення клієнтів та маркетингової комунікації	Цінова політика, універсальність продукту (миттєве практичне застосування), орієнтованість на кінцевого споживача	Стратегія диференціації

Наступним кроком обрано стратегію конкурентної поведінки (таблиця 4.16).

Таблиця 4.16 - Визначення базової стратегії конкурентної поведінки

Чи є проект «першопрохідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки
«Прешопроходець»	Шукати нових	Ні	Стратегія заняття конкурентної ніші

На основі вимог споживачів з обраних сегментів до постачальника (стартап-компанії) та до продукту (див. таблицю 4.5), а також в залежності від обраної базової стратегії розвитку (таблиця 4.15) та стратегії конкурентної поведінки (таблиця 4.16) розроблено стратегію позиціонування (таблиця 4.17), що полягає у формуванні ринкової позиції (комплексу асоціацій), за яким споживачі мають ідентифікувати торгівельну марку/проект.

Таблиця 4.17 - Визначення стратегії позиціонування

Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
Легкість розуміння, зручний інтерфейс, надійний, швидкий, точний та достовірний ПП для інтелектуальної обробки фотографій	Стратегія диференціації	Позиція на основі порівняння фірми з товарами конкурентів; Відмінні особливості споживача	Економія часу; Зручність застосування; Практичність та точність результату

Результатом виконання підрозділу стала узгоджена система рішень щодо ринкової поведінки стартап-компанії, яка визначає напрями роботи стартап-компанії на ринку.

4.5 Розроблення маркетингової програми стартап-проекту

Сформовано маркетингову концепцію товару, який отримає споживач. Для цього у таблиці 4.18 підсумовано результати попереднього аналізу конкурентоспроможності товару. Концепція товару - письмовий опис фізичних та інших характеристик товару, які сприймаються споживачем, і набору вигод, які він обіцяє певній групі споживачів.

Таблиця 4.18 - Визначення ключових переваг концепції потенційного товару

Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
Швидкість отримання результату	Швидка обробка фотографії	Відсутність необхідності звертатися до сторонньої особи/компанії для ретушування фотографій.
Зручність застосування	Не потрібно мати глибоких знань, для того щоб обробити фотографію	Не потрібно мати глибоких знань у обробці фотографій для того, щоб користуватися ПП. Не потрібні особливі обчислювальні можливості, лише доступ до мережі Інтернет.
Практичність та точність результату	Користувач отримує фотореалістичні результати.	Користувач на виході роботи ПП отримує фото реалістичне зображення зі зміною часу доби та збереженням змісту.

Розроблено трирівневу маркетингову модель товару: уточнюється ідея продукту та/або послуги, його фізичні складові, особливості процесу його надання (таблиця 4.19).

1-й рівень: при формуванні задуму товару вирішується питання щодо того, засобом вирішення якої потреби або проблеми буде даний товар, яка його основна

вигода. Дане питання безпосередньо пов'язаний з формуванням технічного завдання в процесі розробки конструкторської документації на виріб.

2-й рівень: цей рівень являє рішення того, як буде реалізований товар в реальному, включає в себе якість, властивості, дизайн, упаковку, ціну.

3-й рівень: товар з підкріпленням (супроводом) - додаткові послуги та переваги для споживача, що створюються на основі товару за задумом і товару в реальному виконанні (гарантії якості, доставка, умови оплати та ін).

Таблиця 4.19 - Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Зручність та швидкість отримання практичного результату щодо редагування фотографій.		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. функція генерації зображення		
	Якість: фотореалістичність зображення		
	Пакування: відсутнє		
	Марка: SunnyPhoto		
III. Товар із підкріпленням	До продажу: відсутнє		
	Після продажу: персональна підтримка в обслуговуванні за додаткову плату.		
Вихідний код закритий. На ідею зареєстровано патент.			

Після формування маркетингової моделі товару слід відмітити, що проект буде захищено від копіювання за допомогою ноу-хау. Наступним кроком є визначення цінових меж, якими необхідно керуватись при встановленні ціни на потенційний товар (остаточне визначення ціни відбувається під час фінансово-економічного аналізу проекту), яке передбачає аналіз ціни на товари-аналоги або товари субститути, а також аналіз рівня доходів цільової групи споживачів (таблиця 4.20). Аналіз проведено експертним методом.

Таблиця 4.20 - Визначення меж встановлення ціни

Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1800\$	3500\$	У всіх трьох груп високий рівень доходів	Базова покупка 1000\$ Подальша персональна підтримка в обслуговуванні 150\$/місяць

Наступним кроком є визначення оптимальної системи збуту, в межах якого було прийняте рішення (таблиця 4.21)

Таблиця 4.21 - Формування системи збуту

Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
Цільові клієнти – компанії, які бажають впровадити у своїй роботі сучасні засоби, які допоможуть отримати вигоду та покращити дохідність. Вони цікавляться сучасними розробками та інноваційними рішеннями, тому відвідують конференції, інтернет-конференції, семінари.	Встановлення контактів із споживачами і підтримання їх. Формування попиту і стимулювання збуту. Дослідницька робота зі збору маркетингової інформації. Доробка товару, виходячи з потреб конкретного покупця.	Один (від виробника одразу споживачу)	Прямий канал збуту до споживача, мінімізувати збутові витрати розвиток маркетингового спілкування із споживачем

Останньою складовою маркетингової програми є розроблення концепції маркетингових комунікацій, що спирається на попередньо обрану основу для позиціонування, визначену специфіку поведінки клієнтів (таблиця 4.22).

Таблиця 4.22 - Концепція маркетингових комунікацій

Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
Цільові клієнти – компанії, які бажають впровадити у своїй роботі сучасні засоби, які допоможуть отримати вигоду та покращити дохідність. Вони цікавляться сучасними розробками та інноваційними рішеннями, тому відвідують конференції, інтернет-конференції, семінари.	Конференції, інтернет-конференції, семінари, огляд професійної літератури, інтернет, періодичні видання у різноманітних (профільних) галузях.	Позиція на основі порівняння фірми з товарами конкурентів; Відмінні особливості споживача	Створення репутації фірми — виробнику чи посереднику; збільшення чистого прибутку та рентабельності фірми; збільшення потоків покупців та обсягів продажу; стабілізація обсягів продажу в період зменшення попиту та загального спаду ділової активності.	Не чекайте вдалої погоди. Змінійте її!

Результатом підрозділу стала ринкова (маркетингова) програма, що включає в себе концепції товару, збуту, просування та попередній аналіз можливостей ціноутворення, спирається на цінності та потреби потенційних клієнтів, конкурентні переваги ідеї, стан та динаміку ринкового середовища, в межах якого впроваджено проект, та відповідну обрану альтернативу ринкової поведінки.

Висновки за розділом

В даному розділі було проведено аналіз програмного продукту у якості стартап проекту. Можна зазначити, що у проекту є можливість комерціалізації, оскільки ринок потребує якісний продукт, що надає можливість інтелектуального редагування фотографій.

На ринку наявна монополістична конкуренція, існує декілька фірм-конкурентів, але їх товар дещо відрізняється, тому вихід на ринок не буде легким і потребує грамотної стратегії виходу. Для впровадження ринкової реалізації проекту слід обрати альтернативу, яка передбачає розробку програмного продукту з подальшим розповсюдженням за певну плату.

Можна сказати, що подальший розвиток проекту є доцільним, оскільки він знайде свою цільову аудиторію.

ВИСНОВКИ

Стрімкий розвиток інформаційних технологій з кожним днем робить взаємодію людини з технікою все простішою та більш інтуїтивною. Незважаючи на те, що технічні характеристики камер вже досягли дуже високої планки, залишається актуальним питання редагування фотографій через семантичні недоліки, які не залежать від апаратури (наприклад, невдалі погодні умови). Проте цей процес потребує чималих навиків роботи з низькорівневими характеристиками зображення, що унеможлиблює його для аматорів та залишається кропітким для професіоналів. Тому виникає потреба у розробках, які не потребують специфічного досвіду та великих затрат часу. Користувачі часто застосовують складні маніпуляції під час обробки фотографій. Вони комбінують декілька ефектів, таких як експозиція, зміна тону та насиченості, накладання фільтрів. Загальне перетворення комплексне та складне до відтворення. Методи перенесення стилю спрямовані на подолання цієї проблеми, надаючи користувачу простий спосіб контролювати стиль зображення, автоматично переносячи стиль цільового зображення на інше. Нещодавні роботи мали вражаючий успіх у перенесенні стилів живопису на зображення. Проте, ці методи досить слабкі у фотореалістичності. Навіть якщо вхідне та цільове зображення – фотографії, результат має дефекти, через які нагадує живопис.

Дана робота присвячена проблематиці задачі фотореалістичного перенесення стилів між зображеннями. В результаті роботи розглянуто такі питання:

- проведено дослідження існуючих підходів до вирішення задачі зміни часу доби на фотознімках, проаналізовано методи перенесення стилів та особливості їх використання для отримання фотореалістичних зображень;

- запропоновано алгоритм фотореалістичної зміни часу доби на фотознімках з використанням сегментації та глибоких згорткових нейронних мереж;

- досліджено процедуру навчання генератора та отримано такі його параметри, що забезпечують достатню трансформацію зображення зі збереженням його реалістичності, також розглянуто шляхи збільшення швидкодії обробки зображень;

- проведено порівняння генератора з іншими доступними інструментами за запропонованими в роботі критеріями оцінки якості, отримані результати свідчать, що реалізований генератор розв'язує поставлену задачу у достатній якості;

- практичним результатом роботи є архітектура інтелектуальної системи фотореалістичної зміни похмурої погоди на фотознімках на сонячну та відповідний програмний продукт із графічним інтерфейсом користувача.

ПЕРЕЛІК ПОСИЛАНЬ

1. Лукьяница А.А. Обработка видеоизображений / А.А.Лукьяница, А.Г.Шишкин; Под ред. Поздняков С.А. — М.: Ай-Эс-Эс Пресс, 2009. — 764 с.
2. Zeiler M.D. Visualizing and Understanding Convolutional Networks [Электронный ресурс] / M.D. Zeiler, R. Fergus. – Режим доступа: <http://www.matthewzeiler.com/wp-content/uploads/2017/07/axiv2013.pdf>
3. Color Transfer between Images [Электронный ресурс] / [E. Reinhard, M. Ashikhmin, B. Gooch, P Shirley]. – Режим доступа: <https://www.cs.tau.ac.il/~turkel/imagepapers/ColorTransfer.pdf>
4. Pitie F. N-Dimensional Probability Density Function Transfer and its Application to Colour Transfer [Электронный ресурс] / F. Pitie, A. C. Kokaram, R. Dahyot. – Режим доступа: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.524.6154&rep=rep1&type=pdf>
5. Transient Attributes for High-Level Understanding and Editing of Outdoor Scenes [Электронный ресурс] / [P.-Y. Laffont, Z. Ren, X. Tao, et al.] – Режим доступа: <http://transattr.cs.brown.edu/files/TransientAttributes-paper.pdf>
6. Gatys L. A. Image Style Transfer Using Convolutional Neural Networks [Электронный ресурс] / L. A. Gatys, A. S. Ecker, M. Bethge. – Режим доступа: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Gatys_Image_Style_Transfer_CVPR_2016_paper.pdf
7. Li C. Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis [Электронный ресурс] / C. Li, M. Wand. – Режим доступа: <https://arxiv.org/pdf/1601.04589.pdf>
8. Chen T. Q. Fast Patch-based Style Transfer of Arbitrary Style [Электронный ресурс] / T. Q. Chen, M. Schmidt. – Режим доступа: <https://arxiv.org/pdf/1612.04337.pdf>
9. Вьюгин В.В. Математические основы теории машинного обучения и прогнозирования / В.В. Вьюгин. – М.: МЦНМО, 2013. – 387 с.

10. Neural Color Transfer between Images [Электронный ресурс] / [М. He, J. Liao, L. Yuan, et al.] – Режим доступа: <https://arxiv.org/pdf/1710.00756.pdf>
11. Liu Y. AutoStyle: Automatic Style Transfer from Image Collections to Users' Images [Электронный ресурс] / [Y. Liu, M. Cohen, M. Uyttendaele, S. Rusinkiewicz]. – Режим доступа: http://gfx.cs.princeton.edu/pubs/Liu_2014_AAS/egsr14.pdf
12. Simonyan K. Very Deep Convolutional Networks for Large-Scale Image Recognition [Электронный ресурс] / К. Simonyan, А. Zisserman. – Режим доступа: <https://arxiv.org/pdf/1409.1556.pdf>
13. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs [Электронный ресурс] / [L.-C. Chen, G. Papandreou, I. Kokkinos, et al.] – Режим доступа: <https://arxiv.org/pdf/1606.00915.pdf>
14. He K. Deep Residual Learning for Image Recognition [Электронный ресурс] / [К. He, X. Zhang, S. Ren, J. Sun]. – Режим доступа: <https://arxiv.org/pdf/1512.03385.pdf>
15. Luan F. Deep Photo Style Transfer [Электронный ресурс] / [F. Luan, S. Paris, E. Shechtman, K. Bala.] – Режим доступа: <https://arxiv.org/pdf/1703.07511.pdf>
16. Levin A. A Closed-Form Solution to Natural Image Matting [Электронный ресурс] / A. Levin, D. Lischinski, Y. Weiss. – Режим доступа: <http://webee.technion.ac.il/people/anat.levin/papers/Matting-Levin-Lischinski-Weiss-CVPR06.pdf>
17. Weiss Y. Deriving intrinsic images from image sequences [Электронный ресурс] / Y. Weiss. – Режим доступа: http://www.ai.mit.edu/courses/6.899/papers/13_02.PDF
18. Mechrez R. Photorealistic Style Transfer with Screened Poisson Equation [Электронный ресурс] / R. Mechrez, E. Shechtman, L. Zelnik-Manor. – Режим доступа: <https://arxiv.org/pdf/1709.09828.pdf>
19. Fourier Analysis of the 2D Screened Poisson Equation for Gradient Domain Problems [Электронный ресурс] / [P. Bhat, B. Curless, M. Cohen et al.] – Режим доступа: http://grail.cs.washington.edu/projects/screenedPoissonEq/screenedPoissonEq_files/screenedPoissonEq.pdf

20. Farbman Z. Convolution Pyramids [Электронный ресурс] / Z. Farbman, R. Fattal, D. Lischinski. – Режим доступа: <http://leibniz.cs.huji.ac.il/tr/1289.pdf>
21. Image Melding: Combining Inconsistent Images using Patch-based Synthesis [Электронный ресурс] / [S. Darabi, E. Shechtman, C. Barnes et al.] // Proceedings of SIGGRAPH 2012 ACM Transactions on Graphics (TOG). – 2012. - vol. 31, № 4, art. 82. – pp. 1-10. – Режим доступа: <http://www.ece.ucsb.edu/~psen/melding>
22. Morel J.-M. Screened Poisson Equation for Image Contrast Enhancement [Электронный ресурс] / J.-M. Morel, A.-B. Petro, C.Sbert // Image Processing On Line. – 2014. - vol. 4. - pp.16–29. – Режим доступа: http://www.ipol.im/pub/art/2014/84/article_lr.pdf
23. Mechrez R. Saliency Driven Image Manipulation [Электронный ресурс] / R. Mechrez, E. Shechtman, L. Zelnik-Manor. – Режим доступа: <https://arxiv.org/abs/1612.02184>
24. OpenCV library [Электронный ресурс]. – Режим доступа: <https://opencv.org/>
25. Torch [Электронный ресурс]. – Режим доступа: <http://torch.ch/>
26. NumPy [Электронный ресурс]. – Режим доступа: <http://www.numpy.org/>