

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
Кафедра Програмної інженерії

АТЕСТАЦІЙНА РОБОТА (ПРОЕКТ)

Пояснювальна записка

другий (магістерський)

«Дослідження методів генерування ілюстрацій на основі тексту»

Виконала: студент 2 курсу, групи ПЗСм-18-1
спеціальності

121 - Інженерія програмного забезпечення
освітньо-професійної програми

Програмне забезпечення систем

Пидоренко Д. О.

Керівник:

доцент каф. ПІ, к.т.н., доцент Турута О.П.

Допускається до захисту

Зав. кафедри, проф.

_____ Дудар З. В.
(підпис)

2019 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Програмної інженерії

Рівень вищої освіти – другий (магістерський)

Спеціальність 121 - Інженерія програмного забезпечення

Тип програми освітньо-професійна програма

Освітня програма Програмне забезпечення систем

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«___» _____ 2019 р.

ЗАВДАННЯ

НА АТЕСТАЦІЙНУ РОБОТУ

студентові Пидоренко Дар'ї Олександрівні

1. Тема роботи проекту «Дослідження методів генерування ілюстрацій на основі тексту» затверджена наказом по університету від «___» _____ 2019 р. № _____.
2. Термін подання студентом роботи (проекту): 16 грудня 2019 р.
3. Вихідні дані до роботи: алгоритми роботи з текстом, роботи із зображеннями, пояснювальна записка. Використовувати ОС Windows, текстовий редактор VS Code.
4. Перелік питань, що потрібно опрацювати в роботі: мета роботи, аналіз проблемної галузі і постановка задачі, огляд методів скорочення тексту, створення анотацій до тексту, пошуку ключових слів у тексті, огляд нейронних мереж для генерації та перетворення зображень з тексту або із зображень.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи (проекту)	Термін виконання етапів проекту (роботи)	Примітка
1.	Аналіз предметної галузі	10 листопада 2019 р.	
2.	Огляд існуючих методів	17 листопада 2019 р.	
3.	Методи роботи з текстом	22 листопада 2019 р.	
4.	Методи роботи із зображеннями	29 листопада 2019 р.	
5.	Підготовка пояснювальної записки	5 грудня 2019 р.	
6.	Підготовка презентації та доповіді	8 грудня 2019 р.	
7.	Попередній захист	10 грудня 2019 р.	
8.	Нормоконтроль, рецензування	12 грудня 2019 р.	
9.	Занесення диплома в електронний архів	13 грудня 2019 р.	
10.	Допуск до захисту у зав. кафедри	14 грудня 2019 р.	

Дата видачі завдання: « ____ » _____ 2019 р.

Студент _____ Пидоренко Д. О.

(підпис)

Керівник роботи _____ доцент каф. ПІ, к.т.н., доцент Турута О. П.

(підпис)

РЕФЕРАТ

Атестаційна робота магістра містить: 92 с., 4 рис., 14 табл., 22 джерел.

ГЕНЕРАЦІЯ ЗОБРАЖЕНЬ, КЛЮЧОВІ СЛОВА, ПЕЙЗАЖ,
ПЕРЕТВОРЕННЯ ЗОБРАЖЕННЯ НА ЗОБРАЖЕННЯ, ПЕРЕТВОРЕННЯ
ТЕКСТУ НА ЗОБРАЖЕННЯ, ПРИРОДА, СТИСНЕННЯ ТЕКСТУ

Мета роботи – створити ілюстрації до текстів. Для перетворення тексту у зображення є два етапи. По-перше, аналіз тексту, в ході якого отримується ключова інформація (скорочення) тексту. Для цього можуть використовуватися різні способи, такі як алгоритм RAKE, NLP або створення текстових анотацій. По-друге, генерування зображення з міток (ключових слів). Для створення зображень існує безліч GAN. Вдосконалення умовних GAN, таких як BigGAN, AttnGAN використовують мітки як основу для створення зображення. Зображення можна перетворювати з одного на інше за допомогою StyleGAN, CycleGAN, Pix2Pix тощо. Також існують методи для створення зображень за конкретними темами, такими як природа (GauGAN) або люди (PoseGAN). SinGAN дозволяє створювати безліч зображень за допомогою моделі, натренованої на лише одному зображенні. Ці зображення є варіаціями вихідного зображення з додаванням «шуму».

Отримані в ході роботи зображення були оцінені на відповідність тексту. Також були запропоновані шляхи для подальших досліджень.

Для розробки використовується мова програмування Python, фреймворк PyTorch, бібліотека NLTK.

Текстовий редактор VSCode був обраний як середовище розробки.

IMAGE GENERATION, IMAGE-TO-IMAGE TRANSLATION, KEYWORDS,
LANDSCAPE, NATURE, TEXT COMPRESSION, TEXT-TO-IMAGE
TRANSLATION

The aim of this work is creating illustrations for texts. There are two steps for translation text to image. First, analyzing a text to obtain key information from the text. Different ways can be used for that like RAKE algorithm, NLP or text summarizers. Second, generating an image from labels (keywords). There are a lot of GANs used for

image generation. Improvements of conditional GANs like BigGAN, AttnGAN uses labels as a base for image generation. Images can be translated from one to another using StyleGAN, CycleGAN or Pix2Pix, etc. Also, there are methods to generate images for specific topics like nature (GauGAN) or people (PoseGAN). SinGAN allows creating a lot of images using a model trained on only one image. These images are variations of the original image with the «noise».

The images obtained in the work was evaluated for their consistency with the text. Also, ways for further researches were suggested.

Python programming language, PyTorch framework, NLTK library are used for development.

The VSCode text editor is selected as the development environment.

ЗМІСТ

Вступ.....	7
1 Аналіз предметної галузі	9
2 Постановка задачі.....	14
2.1 Стиснення тексту	14
2.2 Генерація зображень.....	16
2.3 Планування експерименту	18
3 Стиснення тексту.....	21
4 Генерація зображень	27
5 Оцінка отриманого результату	38
Висновки	50
Перелік джерел посилання	53
Додаток А Тези «A survey of methods of text-to-image translation»	55
Додаток Б Стаття «A survey of methods of text-to-image translation»	58
Додаток В Стаття «A text-to-image translation: nature illustrations».....	65
Додаток Г Слайди презентації	71
Додаток Д Відгук та рецензії	89

ВСТУП

У час інформаційних технологій самвидав є дуже поширеним. Найпростіше, що можна зробити для розповсюдження своїх думок – це публікація їх в Інтернеті. Поширення власних творів перестало бути недоступною для більшості людей. Доступ до Інтернету вирішує майже кожен проблему, яка може виникнути. Такі публікації людина може робити самотужки без залучення спеціалістів та великих грошових витрат.

Ще однією перевагою онлайн-публікацій є можливість вибору формату, стилю, шрифту, розташування тексту, ілюстрацій до нього та інших факторів. Ці чинники здебільшого естетичні, але саме оформлення публікації в інформаційну епоху може привернути до (або відвернути від) неї увагу.

Розташувати текст певним чином не так важко, як знайти до нього ілюстрації. Не кожен може створити свій унікальний малюнок або придбати права на чужий. Зображення у публічному доступі часто повторюються і швидко стають нудними. Автор може знаходитися у творчому пошуку або зовсім не мати гарних ідей для ілюстрації. Як же тоді можна створити зображення, яке буде приманювати читачів, без грошей та мистецьких навичок?

Існує багато текстів, які взагалі не мають ілюстрацій або втратили їх з плином часу. Велика частина текстів потребує нових обкладинок та яскравих сторінок. Але скільки часу та людських зусиль потрібно, щоб проілюструвати їх усі, якщо кожен день їх кількість може навіть зростати?

Одним із варіантів розв'язання цієї проблеми може бути створення зображень за допомогою програмного забезпечення.

Проблема генерації зображень є достатньо відомою в науковому світі – друкуються численні статті, проводяться виступи. І поступово зі світу наукових статей вона переходить до звичайних людей, популяризується – за допомогою зручних онлайн-сервісів, де кожен може відчувати себе творцем та використати результат їх роботи у практичних цілях. Зараз можна згенерувати практично все що завгодно – обличчя людей, яких ніколи не існувало, фантастичні пейзажі, реалістичні інтер'єри чи архітектуру невідомих домівок, квіти, птахів чи

персонажів аніме. Ці зображення бувають настільки схожі на справжні, що інколи не важко й помилитися.

Генерація зображень може базуватися на певних мітках (тегах) – словах чи словосполученнях, які будуть характеризувати зображення, допомагати контролювати його створення. Але повністю текст, який призначений для публікації, не може стати основою для власних ілюстрацій. Його потрібно певним чином переробити – отримати ключову інформацію, яка буде характеризувати його найкращим чином. Цю ключову інформацію можна отримати із текстів за допомогою різних методів обробки тексту.

В ході роботи будуть проаналізовані, реалізовані та оцінені методи, які дозволяють згенерувати ілюстрацію до літературного твору у два етапи: перший з яких це скорочення (виділення ключової інформації) тексту, другий – власне генерація ілюстрації на основі отриманої інформації.

1 АНАЛІЗ ПРЕДМЕТНОЇ ГАЛУЗІ

Вирішуючи проблему створення ілюстрацій, ми стикаємося з двома задачами: скороченням тексту та генерацією зображень.

Існують нейронні мережі, такі як GAN. GAN – це генеративно-змагальна мережа, яка базується на поєднанні двох нейронних мереж, одна з яких генерує кандидатів (generator), а інша намагається відрізнити правильних кандидатів від неправильних (discriminator).

Кожен з варіантів архітектури GAN виконує свою роботу.

Наприклад, зараз дуже популярна генерація обличчя. Працює навіть веб-сайт, який створює обличчя вигаданих людей. Цей процес генерації зображень заснований на StyleGAN. Архітектура StyleGAN призводить до автоматично засвоєного, непідконтрольного розділення атрибутів високого рівня (наприклад, пози та ідентичності під час тренувань на обличчях людини) та стохастичних змін у створених зображеннях (наприклад, веснянки, волосся), а це дозволяє інтуїтивно зрозуміти масштабно-специфічний контроль синтезу [1].

Також викликає інтерес алгоритм Neural Style, який може застосовувати стиль одного зображення до іншого [2]. Зазвичай такі нейронні мережі використовуються для застосування стилю деяких митців. Але не має значення, яке зображення буде використовуватиметься як вхідні дані – витвір мистецтва, фотографії чи щось інше.

Існує велика кількість способів перетворення зображення на зображення, наприклад, Pix2Pix і CycleGAN.

Pix2Pix навчає перетворювати зображення на основі пар зображень $\{A, B\}$, де A і B – це два різних зображення однієї і тієї ж базової сцени. Ця архітектура має приклади перетворення силуетів у фасади будівель, контурів у фотографії, зміни дня на ніч або чорно-білого зображення на кольорове [3].

CycleGAN – це непарна версія архітектури Pix2Pix. Вона може перетворити зображення з вихідного домену X у цільовий домен Y за відсутності парних прикладів. Ця архітектура використовується для зміни погоди або пори року на фотографії або для застосування стилю конкретного художника [4].

Є багато GAN, які можуть генерувати зображення на основі конкретної додаткової інформації (міток), що призводить до створення якісніших зображень та можливості контролювати якість зображення. Їх називають умовними GAN.

Серед них DCGAN – перше покращення GAN. ProGAN, яка поступово додають нові шари високої роздільної здатності під час тренувань та створюють неймовірно реалістичні зображення. InfoGAN здатна кодувати значущі функції зображення в частині шуму вектору Z без нагляду (наприклад, кодування чергування цифри). Wasserstein GAN, яка змінює функцію втрат, щоб включити відстань Вассерштейна, вона працює досить добре при генерації приміщень, навіть якщо спробувати дивні архітектури [5].

Детальніше варто розглянути BigGAN. Ця мережа створює дуже реалістичні образи для конкретного класу. Наприклад, її можна навчити за допомогою складних наборів даних, таких як ImageNet. Одна зі спостережуваних помилок частково натренованих моделей – це class leakage (протікання класу), де зображення з одного класу містять властивості іншого. Отримані зображення можна з легкістю перекладати з одного класу в інший [6].

Ще одне вдосконалення умовного GAN – це Attentional Generative Adversarial Network (AttnGAN), яка була запропонована групою розробників з Ліхайського, Раттерського та Дюкського університетів США і компанії Microsoft. Ця нейрона мережа являє собою поліпшення традиційного підходу і дозволяє поступово змінювати сформоване згенероване зображення відповідно до зміни окремих слів в текстовому описі [7].

AttnGAN використовує варіаційний автокодувальник.

Автокодувальники – це нейронні мережі, які копіюють свої входи на свої виходи. Вони працюють, стискаючи вхідні дані у латентне представлення простору, а потім реконструюючи вихідні дані із цього подання. Така мережа складається з двох частин: кодера і декодера.

Варіативні автокодувальники (на відміну від ванільних автокодувальників) дозволяють створювати дані, подібні до раніше поданих даних, або навіть змінювати їх у певному напрямку. Ці автокодувальники навчаються розширювати

свій прихований простір, генеруючи можливі вхідні дані і подаючи їх до декодера для створення нових зразків вихідних даних [8].

Зазвичай генеровані зображення не такі гарні, як ті, які можуть намалювати люди. Чим більший набір зображень нейрона мережа може генерувати (натренована на такому наборі) – тим менш зрозумілим є малюнок.

Тож як зробити мережі для створення ілюстрацій для дійсно великої, але все ж конкретної теми, як природа?

NVIDIA презентувала мережу GauGAN у березні 2019 року та надала інтерактивну програму, яка генерує реалістичні зображення пейзажу з малювання користувачами макета [9]. Ця мережа дозволяє створити певний тип ілюстрацій (природу), якщо користувач може надати для неї макет (або сегментацію).

Існує велика кількість алгоритмів для семантичної сегментації зображень, більшість з яких – це вдосконалення повністю конволюційної мережі (FCN).

Один з них – маска R-CNN. Маска R-CNN – це швидше R-CNN з 3 вихідними гілками: перша обчислює координатні рамки, друга – обчислює пов'язаний клас, а остання обчислює двійкову маску для сегментації об'єкта. Двійкова маска має фіксований розмір, і вона генерується FCN для даного RoI.

Швидша архітектура R-CNN для виявлення об'єктів використовує мережу пропозицій регіону (RPN), щоб запропонувати кандидатам обмежувального поля. RPN витягує область інтересів (RoI) та рівень RoIPool обчислює особливості цих пропозицій для висновку координат обмежувального поля та класу об'єкта. Маска R-CNN також використовує шар RoIAlign замість RoIPool, щоб уникнути розбіжностей через квантування координат RoI.

Особливістю моделі Mask R-CNN є її багатозадачні втрати, що поєднують втрати координатних обмежувальних коробок, передбачуваний клас та маску сегментації [10].

Ці сегментації можна використовувати як макет у мережі GauGAN.

SinGAN – це безумовна генеративна модель, яка може бути натренована на одному реальному зображенні. Вона може генерувати високоякісні, різноманітні зразки, що містять той же візуальний вміст, що і вихідні зображення.

SinGAN містить піраміду повністю згорнутих GAN, кожна з яких відповідає за вивчення розподілу латок в іншому масштабі зображення. Це дозволяє генерувати нові реалістичні зразки довільного розміру та співвідношення сторін. Ці зразки мають значну мінливість, але зберігають як глобальну структуру, так і тонкі текстури зображення, на якому тренується мережа.

На відміну від інших GAN, які тренуються з одним зображенням, підхід SinGAN не обмежується текстурними зображеннями і не є умовним (тобто він генерує зразки, базуючись на шумі) [11].

Таким чином існує маса способів генерування зображень на основі вхідних зображень або опису тексту. Але як можна отримати цю інформацію з великого тексту?

Для знаходження ключових слів у тексті існує декілька варіантів. Наприклад, алгоритм швидкого автоматичного вилучення ключових слів (RAKE). Він базується на побудові графу спільної появи (co-occurrence graph) зі слів, які зустрічаються у тексті найчастіше [12]. Тобто RAKE намагається визначити ключові фрази тексту шляхом аналізу частоти появи слова та його співзвучності з іншими словами в тексті.

У мові Python RAKE має декілька реалізацій з різними можливостями налаштування.

Інший спосіб – це обробка природних мов (NLP). NLP надає багато різних способів роботи з текстом, його аналізу та перетворення. Наприклад, токенізація (процес розбиття тексту на слова, пропозиції, абзаци чи інші типи лексем) є одним з основних компонентів практично будь-якого завдання NLP, і це може бути першим кроком для підготовки тексту до обробки.

Існує декілька графічних методів підсумку тексту, таких як Lex Rank або Text Rank. А також Luhn – алгоритм, який оцінює речення на основі частоти найважливіших слів. Латентний семантичний аналіз (LSA) – метод узагальнення без вчителя, який поєднує термінні частотні методи з сингулярним розкладанням значень для узагальнення текстів [13].

Також можливо використати тегування частинами мови (POS-tag), ще відоме як класи класів або лексичні категорії. Завдання POS-тегування – позначити слова речення відповідними частинами мови (іменники, займенники, дієслова, прикметники тощо) [14].

Таким чином вибір методів залежить від конкретної задачі та має базуватися на попередньому аналізі.

Задача створення ілюстрацій до тексту не є вирішеною на даний момент та може бути досягнута через поєднання декількох наявних методів з певними припущеннями та налаштуваннями.

Проблема дослідження у створенні ілюстрацій до великої кількості різноманітних текстів, із яких попередньо потрібно витягти ключову інформацію для більшої відповідності та конкретності ілюстрацій.

В ході роботи потрібно проаналізувати вищевказані методи, визначити їх переваги та недоліки, можливість їх використання для досягнення поставленої мети, продумати декілька варіантів створення ілюстрацій, відтворити ці варіанти та оцінити отримані результати.

2 ПОСТАНОВКА ЗАДАЧІ

Задачею дослідження є створення ілюстрацій до текстів. Цю задачу можна розділити на дві частини. Перша з них – стиснення тексту до відповідних даних для введення (отримання ключової інформації). Друга – це перетворення тексту на зображення або зображення на зображення (власне генерація зображення).

В ході виконання роботи будуть використані наступні методи:

- теоретичні: дослідження наявних методів розв’язання даної задачі (їх аналіз та класифікація, систематизація), формальне моделювання;
- емпіричні методи: науковий експеримент (перевірка різних методів та їх поєднань на практиці), аналіз результатів експерименту.

Таким чином, дослідження буде проведено як і на фундаментальному, так і на прикладному рівні.

Розглянемо етапи створення ілюстрацій до тексту докладніше.

2.1. Стиснення тексту

В рамках першого етапу потрібно:

- дослідити різні методи скорочення літературного твору;
- дослідити проблеми використання методів скорочення літературного твору на практиці;
- оцінити якість скорочення літературного твору різними методами (швидкість, відповідність вихідним даним, зручність для подальшої обробки та ін.).

Повний текст літературного твору може бути занадто великий для перетворення його на конкретне зображення, тому його потрібно певним чином скоротити. Існують різні способи роботи з текстом: генерування анотацій, отримання ключових фраз або вибір певних слів або комбінацій слів.

Наприклад, з тексту можна виділити лише іменники, оскільки іменники характеризують текст найбільше. Або можна вибрати поєднання іменника та дієслова. Ці комбінації слів показують, які актори є в текстах та які дії вони виконують – чи не основна ідея тексту. Для цього можна використати методи NLP. Або знайти ключові слова за допомогою алгоритму RAKE.

Результат першого кроку необхідно контролювати, щоб він був придатним для другого кроку. Якщо отримані ключові слова занадто абстрактні, будь-яка GAN не зможе генерувати зображення, незалежно від того, на якому наборі даних вона навчалася. У цьому випадку процес перетворення даних можна змінити, побудувавши залежність у зворотному напрямку. Ключові слова повинні залежати від того, на якому наборі даних тренувалася модель GAN, тобто які зображення вона може генерувати.

Але як можна контролювати, що саме отримано?

Наприклад, є набір даних із зображеннями з мітками. Отже, список усіх можливих міток можна отримати. Ці мітки можна назвати словником.

Тоді текст можна перевірити на наявність слів зі словника і слід вибрати лише ті речення чи абзаци (текстові одиниці), які їх містять. І лише після цього кроку отриманий текст слід стискати. Отримані ключові фрази також можна проаналізувати на наявність слів зі словника.

Кількість відповідних слів або відсоток таких слів у текстовій одиниці можна обчислити, і слід вибрати лише ті одиниці, кількість яких перевищує певне порогове значення.

Абзаци можна аналізувати окремо, або суміжні абзаци можна комбінувати в один.

Використовуючи WordNet, можна перевірити не лише наявність слів зі словника, а і їх синоніми (виберіть певний відсоток подібності). WordNet – це велика лексична база даних англійської мови. Іменники, дієслова, прикметники та прислівники згруповані в набори пізнавальних синонімів (синсетів), кожен з яких виражає окреме поняття [15].

Залежно від теми діалоги можуть бути враховані, а можуть і не рахуватися. Наприклад, описи природи взагалі можна знайти без діалогів.

«Зайві» розділові знаки також можна видалити.

Також при виконанні цього етапу може виникнути проблема мови твору – тобто тексти, написані англійською мовою, будуть мати більше можливостей для власного опрацювання.

2.2. Генерація зображень

В рамках другого етапу потрібно:

- дослідити різні мережі та їх конфігурації, які можуть бути використані для генерації зображень;
- дослідити наявні датасети (набори даних) з урахуванням можливої мови літературних творів та різноманітності зображень;
- дослідити проблеми використання цих мереж на практиці;
- оцінити якість генерації зображення різними мережами (швидкість, відповідність вихідним даним (як скороченому тексту, так і твору взагалі) та ін.).

Зображення повинні створюватися на основі ключових слів, отриманих одним зі способів, описаних у підрозділі 2.1.

Для цього можна використовувати мережі GAN. Наприклад, AttnGAN.

Також на основі комбінації іменників та дієслів з пункту 2.1 можуть утворюватися зображення людей. Існує багато способів генерування облич, але краще генерувати цілу людину, яка виконує конкретну дію (має конкретну позу).

Наприклад, можна використати мережу з генерування зображень особи PoseGAN, що генерує позу, яка дозволяє синтезувати зображення людини у довільних позах на основі зображення цієї людини та романної пози. [16].

Поєднання іменника та дієслова можна отримати з тексту. Зображення можна вибрати зі спеціального набору даних за допомогою цих комбінацій слів, які показують, які актори є у текстах (перша група зображень із зовнішнім виглядом людей) та які дії вони виконують (друга група зображень із позиціями людей). На основі цього можна створювати зображення людей.

Інший спосіб – не генерувати зображення (оскільки результат може бути непередбачуваним), а знайти конкретні зображення за мітками у наборах даних.

На основі ключових слів, отриманих у пункті 2.1, слід вибрати зображення з набору даних.

Набір даних повинен мати не лише зображення, але й мітки (теги), які описують кожне зображення. Зображення можна вибрати за допомогою цих міток. Теги зображень повинні максимально відповідати вхідним ключовим словам.

Прикладом таких набору даних є COCO та ImageNet.

COCO – це масштабний набір даних (зображень) для виявлення, сегментації та тегування об'єктів [17]

ImageNet – це база даних зображень, організована відповідно до ієрархії WordNet, в якій кожен вузол ієрархії зображений сотнями і тисячами зображень. В середньому мається понад п'ятсот зображень на вузол [18].

Також можна створити власний датасет, знайти зображення та власноруч позначити їх певними тегами.

Далі отримані зображення можна певним чином перетворити. Існує декілька способів зробити це.

По-перше, можуть використовуватися різні автокодувальники. Автокодувальники копіюють свій вхід на вихід, а варіативні автокодувальники можуть генерувати новий вихід на основі входу.

По-друге, GauGAN. Першим кроком буде отримання сегментації зображення, а потім генерування нового зображення на основі сегментації.

По-третє, SinGAN. Моделі тренуються для кожного можливого зображення. Потім обираються відповідні моделі та генеруються нові зображення з додаванням довільного шуму.

По-четверте, StyleGAN. У цьому випадку буде створено нове зображення на основі двох обраних зображень: буде прийнята сегментація першою та застосований другий стиль. Або навіть спочатку GauGAN можна застосувати до одного із зображень, а потім стилі інших зображень можна застосувати до отриманого зображення.

Таким чином, можна отримати велику кількість нових зображень, майже унікальних і досить різноманітних.

При виконанні цього етапу може виникнути проблема з вибором датасету (мова міток датасету, кількість зображень, які взагалі варіанти зображень існують, їх різноманітність), а також проблема зі швидкістю навчання нейронних мереж.

2.3 Планування експерименту

Експеримент буде проводитися з написанням програмного забезпечення та оцінкою отриманих результатів наступним чином:

- розробка плану експерименту;
- обрання критеріїв оцінки вимірювань та засобів для проведення експерименту;
- власне проведення експерименту – окреме виконання двох етапів наукового дослідження, поєднання їх результатів та спільне послідовне виконання обох етапів;
- обробка і аналіз експериментальних даних, встановлення адекватності отриманого зображення.

Розроблені методи будуть реалізуватися на мові Python.

На рисунку 2.1 представлена формальна модель, яка буде використана в роботі. Експеримент буде проведений у декілька етапів (скорочення тексту, генерація зображення). Спочатку будуть проведена попередня оцінка якості результату для кожного з етапів, та тільки після отримання задовільного результату етапи будуть поєднані у єдину експериментальну систему.

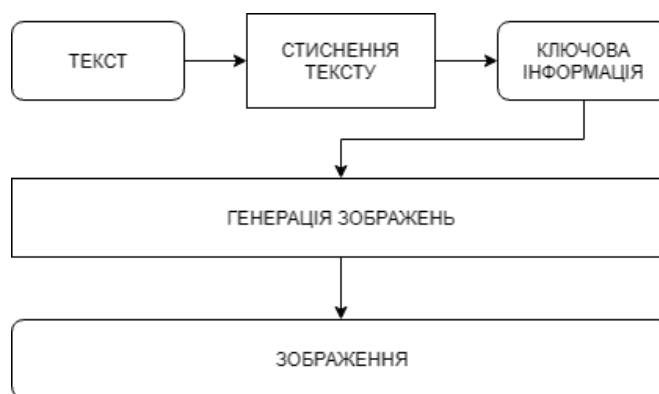


Рисунок 2.1 – Формальна модель дослідження

Розберемо етап створення зображень детальніше. На рисунку 2.2 можна побачити різні варіанти проведення цього етапу: генерація зображення за

допомогою AttnGAN та вибір зображень із подальшим перетворенням за допомогою SinGAN, GauGAN або StyleGAN.

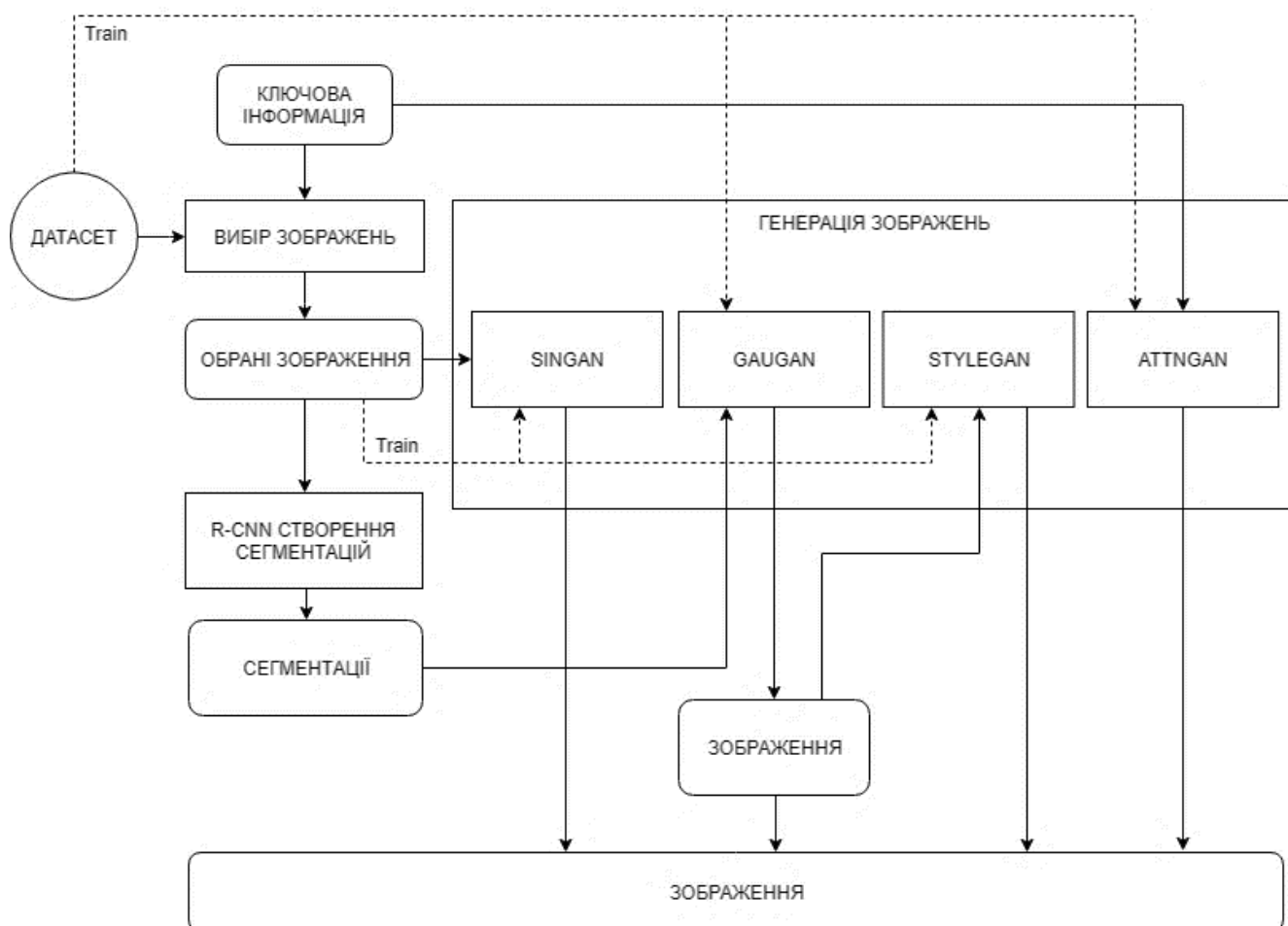


Рисунок 2.2 – Формальна модель етапу створення зображення

Всі вищевказані алгоритми роботи з текстом представлені Python-бібліотеками, RAKE навіть декількома, які відрізняються між собою певними особливостями реалізації.

Для роботи з нейронними мережами Python має декілька фреймворків, серед яких TensorFlow, Keras (надбудова над фреймворками DeepLearning4j, TensorFlow і Theano) та PyTorch. Ці фреймворки реалізують велику кількість алгоритмів для машинного навчання та наукових розрахунків. Для генерації зображень у рамках дослідження буде використаний PyTorch, який будує динамічні графи обчислень на відміну від Keras та TensorFlow [19].

Помилки можуть виникнути як і на етапі скорочення тексту, так і на етапі перетворення тексту на зображення через недосконалість використаних методів або їх реалізації, тому при проведенні експерименту будуть додані перевірки для оцінювання якості перетворення. Ці перевірки допоможуть проаналізувати результати випробувань (тестів).

Перевірки якості будуть проводитися як людиною, так і автоматично (з написанням програмного забезпечення на мові Python).

Також буде вимірюватися час проходження кожного з етапів.

Результати дослідження будуть оцінюватися наступним чином:

- швидкість проведення усіх перетворень (як взагалі, так і окремо), з урахуванням витрат на тренування моделей;
- оцінка якості (адекватності) скорочення тексту людиною;
- оцінка якості отриманого зображення, його відповідності тексту людиною та автоматично;

Отриману систему можна буде використати при створенні веб-сервісу, який буде ілюструвати літературні твори.

3 СТИСНЕННЯ ТЕКСТУ

Перший етап роботи – це стиснення тексту. Приклад роботи деяких методів для знаходження ключових слів або створення анотацій з англійськими текстами приведені у таблиці 3.1. Англійська мова була вибрана, оскільки не усі методи мають можливість роботи з текстами на інших мовах.

Для аналізу були використані тексти «Останній листок» (О. Генрі) та «Старий і море» (Е. Хемінгуей).

Таблиця 3.1 – Аналіз методів роботи з текстом (англійська мова)

Метод	«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)	
Rake	de world mit der foolishness	19.5	portuguese man-of-war floating dose	27.67
	dot poor leetle miss yohnsy	18.0	white tipped wide pectoral fins	17.05
	elegant horseshow riding trousers	16.0	portuguese men-of-war:	16.34
	useless woollen shoulder scarf	16.0	long deadly purple filaments trailing:	14.93
	pen-and-ink drawing	14.25	purple pectoral fins set wide	14.19
	sue found behrman smelling strongly	13.7	two-decker metal container	13.5
	red eyes plainly streaming	13.34	projecting green sticks dip sharply	13.33
	especial mastiff-in-waiting	13.0	ordinary pyramid-shaped teeth	12.6
	lone ivy leaf clinging	12.48	man-of-war bird	12.18
	wide-open eyes staring	12.34	thrusting all-swallowing jaws	12.03
	small dutch window-panes	12.06	high dorsal fin knifing	11.99
	three-story brick sue	11.63	great erect tail slicing	11.72
Multi-Rake	michael angelo's moose beard curling	25.0	white tipped wide pectoral fins	17.01
	de world mit der foolishness	19.5	long deadly purple filaments trailing	14.8
	dot poor leetle miss yohnsy	18.0	purple pectoral fins set wide	14.12
	elegant horseshow riding trousers	16.0	portuguese man-of-war floating dose	14.0
	useless woollen shoulder scarf	16.0	projecting green sticks dip sharply	13.23
	red eyes plainly streaming	14.34	high dorsal fin knifing	11.99
	sue found behrman smelling strongly	13.7	great erect tail slicing	11.62
	lone ivy leaf clinging	12.48	he'll weigh ten pounds	10.76
	brick house twenty feet	11.6	small delicate dark terns	10.7
	busy doctor invited sue	11.23	great long white spine	10.45
	ravager strode boldly	9.0	shark's yellow cat-like eyes	10.4
	feet trod slowly	9.0	small tuna's shivering pull	10.1
Rake NLTK	dot poor leetle miss yohnsy	27.05	good luck old man good luck	35.03
	miss yohnsy shall lie sick	22.17	know others better que va	33.82
	brush without getting near enough	20.17	beg keep warm old man	31.5
	de world mit der foolishness	20.0	know sleep well old man	26.21
	art people soon came prowling	17.6	sorry qua va	26.15
	brick house twenty feet away	16.6	get four fresh ones one	25.95
	useless woollen shoulder scara	16.0	white cumulus built like friendly piles	23.25
	elegant horseshow riding trousers	16.0	old man saw flying fish spurt	20.83
	allow dot silly pusiness	16.0	old man said happily	20.13
	sue found behrman smelling strongly	15.9	head let pedrico chop	19.76
	last one said johnsy	15.8	hands well old man	19.29
	brick wall one ivy leaf	15.46	day ends let us hope	19.16

Кінець таблиці 3.1

Метод	«Останній листок» (О. Генрі)	«Старий і море» (Е. Хемінгуей)
Lex Rank	"It is the last one," said Johnsy. And then they found a lantern, still lighted, and a ladder that had been dragged from its place, and some scattered brushes, and a palette with green and yellow colours mixed on it, and – look out the window, dear, at the last ivy leaf on the wall.	He cannot know that it is only one man against him, nor that it is an old man. But there was not much of it.
Luhn	You may bring a me a little broth now, and some milk with a little port in it, and – no; bring me a hand-mirror first, and then pack some pillows about me, and I will sit up and watch you cook." And then they found a lantern, still lighted, and a ladder that had been dragged from its place, and some scattered brushes, and a palette with green and yellow colours mixed on it, and – look out the window, dear, at the last ivy leaf on the wall.	After that he began to dream of the long yellow beach and he saw the first of the lions come down onto it in the early dark and then the other lions came and he rested his chin on the wood of the bows where the ship lay anchored with the evening off-shore breeze and he waited to see if there would be more lions and he was happy. He took all his pain and what was left of his strength and his long gone pride and he put it against the fish's agony and the fish came over onto his side and swam gently on his side, his bill almost touching the planking of the skiff and started to pass the boat, long, deep, wide, silver and barred with purple and interminable in the water.
LSA	In a little district west of Washington Square the streets have run crazy and broken themselves into small strips called "places." "Johnsy, dear," said Sue, bending over her, "will you promise me to keep your eyes closed, and not look out the window until I am done working?"	How would you like to see me bring one in that dressed out over a thousand pounds?" "I'll get the cast net and go for sardines. No matter what passes I must gut the dolphin so he does not spoil and eat some of him to be strong.
Text Rank	And then, with the coming of the night the north wind was again loosed, while the rain still beat against the windows and pattered down from the low Dutch eaves. And then they found a lantern, still lighted, and a ladder that had been dragged from its place, and some scattered brushes, and a palette with green and yellow colours mixed on it, and – look out the window, dear, at the last ivy leaf on the wall.	He saw the phosphorescence of the Gulf weed in the water as he rowed over the part of the ocean that the fishermen called the great well because there was a sudden deep of seven hundred fathoms where all sorts of fish congregated because of the swirl the current made against the steep walls of the floor of the ocean. After that he began to dream of the long yellow beach and he saw the first of the lions come down onto it in the early dark and then the other lions came and he rested his chin on the wood of the bows where the ship lay anchored with the evening off-shore breeze and he waited to see if there would be more lions and he was happy.

Для реалізації алгоритму RAKE були обрані наступні властивості: мінімум три символи у слові та максимум сім слів у словосполученні. Були вибрані дванадцять найчастіших словосполучень.

Наведемо приклади роботи з українськомовними та російськомовними текстами різних реалізацій алгоритму RAKE, які можуть працювати з різними мовами, у таблиці 3.2. Були використані ті ж самі тексти, що і при англomовному аналізі.

Таблиця 3.2 – Аналіз методів роботи з текстом (українська та російська мови)

Метод	Мова	«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)	
Rake	Укр.	які не могли платити натурникові-професіоналу	32.56	завжди кишла сила-силенна всілякої риби	32.08
		знущаючись із таких ідіотських вигадок	24.5	згадай-но про великого ді маджо	30.53
		подумай хоч би про мене	24.5	він вправно орудував однією рукою	30.47
		тарабанячи по низьких голландських карнизах	24.5	я все-таки наловлю тобі сардин	28.44
		знайшла привід вийти в коридор	24.43	з-під нігтів і в нього	26.9
		навряд чи була тією дичиною	24.3	наче їй кінця-краю не було	26.88
		схожим на звук натягнутої струни	24.3	що він от-от пуститься духу	26.26
		усе на світі стає чуже:	24.3	й от-от мало зійти сонце	26.1
		і вирішили найняти спільну студію	24.3	а вітер нам усе-таки друг	25.4
		то другого своїми крижаними пальцями	24.17	і все-таки я здолаю її	25.03
		знов у вікна перішив дощ	24.1	дальшим краєм дороги пробіг кіт	24.5
		від якого дуже тхнуло ялівцівкою	23.9	великий ді маджо знову розігрався	24.45
Rake	Рос.	готовое принять первые штрихи шедевра	25.0	позади которой вырисовывались серо-голубые холмы	33.5
		склоняясь усталой головой к подушке	24.67	раскинув светло-сиреневые крылья грудных плавников	30.5
		упорный дождь пополам со снегом	24.5	чтобы не совершить какую-нибудь глупость	30.03
		одна приехала из штата мэйн	24.25	еще каких-нибудь два-три круга	29.08
		старый-старый плющ с узловатым скатываясь с низкой голландской кровли	24.14	тогда-то и начнется настоящая работа	29.06
		с удовольствием довязывая ярко-синий	24.14	на какой-то миг движение прекратилось	29
		вся наша фармакопеея теряет смысл	24.0	что рыба свернула на северо-восток	28.58
		насмехаясь над такими идиотскими фантазиями	24.0	чтобы выманить из-под нее акулу	27.64
		становится чуждой всему на свете	23.89	когда пойдет по-настоящему большая рыба	27.63
		ему было уже за шестьдесят	23.84	акула по-прежнему держалась за рыбу	27.32
		пока та не размокла окончательно	23.77	вот этого-то мы и ждали	26.69
				поэтому когда он что-нибудь делал	26.3
Multi-Rake	Укр.	широко розплющених очей із опущеної зеленої штори	47.7	зліплена з цупких брунькових щитків королівської пальми	47.52
		ці проїзди утворюють химерні кути й повороти:	46.25	була припнута петлею до гнучкого зеленого вудлища	45.48
		надворі вперто йшов холодний дощ із снігом	46.2	звичайно стоїть найкраща в цілому році погода	44.98
		яка лаштується вирушити в далеку таємничу подорож	46.13	ті райдужні бульби вабили око своєю красою	44.97
		ковбоя із штату айдахо в елегантних штанах	45.83	заразом пробуючи рукою швидкість зустрічного потоку води	44.81
		до половини тієї стіни видряпався старезний плющ	44.84	як удар у п'яту шпорою бойового півня	44.67
		крім якоїсь мазанини — вивісок та реклам	44.01	зловлю що-небудь підживитися й посмоктати замість води	44.57
		холодний подих осені струсив з нього листя	43.0	зі свистом розтинаючи повітря своїми твердими крильми	44.32
		сорок років тримав він у руках пензель	42.89	побачу світло котрогось із нових прибережних селищ	44.0
		три дні тому їх було майже сто	42.75	чому їх створено такими маленькими й кволими	43.92
		що збирач боргів з рахунком за фарби	42.65	що жадібно хапають зі столу останні крихти	43.78
		не хочу позувати для вашого йолопа відлюдька	42.28	між рибалок вважалося негожим базікати у морі	43.65

Продовження таблиці 3.2

Метод	Мова	«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)		
Multi-Rake	Рос.	какого фасона рукава будут носить этой зимой	49.0	связка тонкого троса лежала в круглой корзине	47.22	
		специально приставленного для охраны двух молодых художниц	48.5	переливаясь чистым золотом в последних лучах солнца	46.37	
		по восточной стороне этот душегуб шагал смело	48.0	полосы эти были шире растопыренной мужской руки	44.93	
		цикорный салат и модные рукава вполне совпадают	47.04	отравление наступало быстро и пронзало острой болью	44.86	
		холодное дыхание осени сорвало листья с лозы	46.92	прощался с жизнью под судорожные удары аккуратного	44.77	
		широко раскрытых глаз со спущенной зеленой шторы	46.75	но на большей части лодок царило молчание	44.61	
		после проливного дождя и резких порывов ветра	45.37	легко прорезая воду своим высоким спинным плавником	44.56	
		бумагу и холст повстречает там самого себя	44.79	пятна спускались по щекам до самой шеи	44.55	
		когда люди начинают действовать в интересах гробовщика	44.61	большинству рыбаков вкус этого жира казался отвратительным	44.39	
		не получив ни единого цента по счету	44.31	они просидели целые сутки друг против друга	44.24	
		сборщик из магазина со счетом за краски	44.22	я потеряю двести саженей крепкой каталонской веревки	43.9	
		при хорошем уходе вы одержите победу	43.49	которыми были унижены острие и полукружие крючка	43.86	
Rake NLTK	Укр.	вона дуже хвора й зовсім знесилена, а від високої температури наче й схибнулася	219.4	на темних стінах із розрівняних і щільно припасованих один до одного шоретких волокнистих щитків висіло кольорове зображення святого серця господнього	449.6	
		ні за що нетриматися й полетіти як один з отих нещасних	197	і ще одне — мідної богоматері		
		вулиці показалися й розбилися на вузькі смужки	188	я людина не побожна, одначе ладен 10 разів проказати отче наш і 10 разів богородицю, аби подужати цю рибину	384.4	
		біле мишенятко сьогодні в лікарні від запалення легенів помер містер берман	172.1	а вільний кінець при потребі можна було приєднати до запасних мотків і пустити рибу від човна більш як на 300 сажнів		313.9
		що, кричав він з жахливим німецьким акцентом, хіба ще є такі дурні	163.4	а він тим часом простяг ліву руку й міцно прив'язав вільний кінець двох запасних мотків до петлі другої снасті	299.4	
		щось примусило цей останній листок залишитись там	160.7	він також пив щодня по кухлю жиру з акулячої печінки — велика залізна бочка з тим жиром стояла в хижці		281.3
		сюю опустила штору аж до підвіконня й знаком веліла пройти в другу кімнату	156.6	а третій і четвертий зависли в голубій воді за 100 й 125 сажнів од човна	229	
		краще я побуду біля тебе відповіла сюю, до того ж я не хочу	149.3	та й на поверхні води старий не бачив ознак риби, лише плавали жовті жмутки збляклої на сонці саргасової трави		214.4
		тепер ти можеш дати мені трохи бульйону й молока з портвейном або ні	142.8	треба б скласти до купи обоє весел і прилаштувати до корми, вони ще дужче придержать рибину вночі, сказав старий	213.3	
		він занадто багато пив і ще не облишив балачок про свій майбутній шедевр	135.5	аж поки старий не торохнув її кийком по лискуцій золотавій голові, лиш тоді вона востаннє сіпнулась і затихла		211.7
		дивлячись крізь невеличкі шибки голландського вікна на глуху стіну сусіднього мурованого будинку сюю розповіла старому про химери й про свої побоювання відносно того	134.6	коли розвидніє, треба якось добутися на корму до тієї снасті й також перерізати її солоний піт затікав йому в очі й боляче обпікав їх і так само боляче обпікав подряпини на лобі	203.5	
			125.1	а ця майже зовсім не солодка й ще не втратила ані крихти поживності ... треба викинути з голови все		202.9
						194.9

Кінець таблиці 3.2

Метод	Мова	«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)	
Rake NLTK	Рос.	чтобы она могла сбыть его редактору и купить вина для своей больной девочки и свиных котлет для себя	240.8	одна из приезжих заметила среди пустых пивных жестянок и дохлых медуз длинный белый позвоночник с огромным хвостом на конце	257.6
		едва ли могла считаться достойным противником для дюжего старого тупицы с красными кулачищами и одышкой набрасывая для рассказа фигуру ковбоя из айдахо в элегантных бриджах и с моноклем в глазу	197.1	я бы целый день поливал водой нос лодки и давал ей высохнуть — к вечеру у меня была бы соль	215.5
		сью спустила штору до самого подоконника и сделала берману знак пройти в другую комнату однажды утром озабоченный доктор одним движением косматых седых бровей вызвал сью в коридор	186.2	только летучие рыбы поднимались из моря и разлетались в обе стороны от носа его лодки да желтели островки водорослей	199.9
		сью выбежала в мастерскую и плакала в японскую бумажную салфеточку	159.9	только бы мне больше не сражаться, но в полночь он сражался с акулами снова а вот теперь он по многу раз повторял свои мысли вслух — ведь они никому не могли быть в тягость	197.9
		затем они перевезли туда с бой авеню несколько оловянных кружек и одну	149	но я прочту десять раз отчете наш и столько же раз богородицу, чтобы поймать эту рыбу	195.8
		утром первого дня швейцар нашел бедного старика на полу в его комнате	147.1	я должен беспрерывно ослаблять тягу своим телом и быть готов в любую минуту опустить бечеву обеими руками	193.5
		глядя сквозь мелкий переплет голландского окна на глухую стену соседнего кирпичного дома	128.1	что каждый толчок акулы об лодку означал кусок оторванного мяса и что рыба теперь оставляет в море след	179.9
		для молодых художников путь в искусство бывает вымощен иллюстрациями к рассказам	128	на этот раз он попал в кость у основания черепа и ударил снова по тому же самому месту	176.4
		сью рассказала старику про фантазию джонси и про свои опасения насчет него	125.9	но что может сделать с ними человек в темноте голыми руками, все его тело ломило и саднило	175.1
		одиноким лист плюща держится на своем стебельке на фоне кирпичной стены	125.8	я прочту сто раз отчете наш и сто раз богородицу	171.1
			122.9	а левой рукой привязал свободный конец двух запасных мотков к петле двух запасных мотков второй удочки	167.9
			115.6		166.5

Для кожного словосполучення наводиться оцінка слова, яка розраховується як відношення спільної появи кандидата (словосполучення, яке розглядається) з іншими кандидатами до частоти появи цього кандидата.

Можна побачити, що не тільки різні реалізації алгоритму дають різні результати, а і різні мови. Хоча деякі словосполучення (або фрази) збігаються для отриманих результатів незалежно від обраної мови.

В ході роботи були проаналізовані різні реалізації алгоритму RAKE, а також TextRank, LexRank, Luhn та LSA. Були розглянуті англійська, українська та російська мови.

Для подальших досліджень був обраний RAKE алгоритм та його реалізація Multi-Rake, яка призначена саме для роботи з текстами, написаними на різних мовах. Ця реалізація дає найвідповідніші та найадекватніші результати та має найбільші можливості для налаштування [20].

Наприклад, для української мови Multi-Rake знаходить фразу «широко розплющених очей із опущеної зеленої штори», а для російської відповідну «широко раскрытых глаз со спущенной зелёной шторы», «збирач боргів з рахунком за фарби» – «сборщик из магазина со счетом за краски». Та для всіх трьох мов Multi-Rake знаходить слово плющ: «lone ivy leaf clinging (одинокий листок плюща, що чіпляється)», «до половини тієї стіни видряпався старезний плющ» та «старый-старый плющ с узловатым (старий-старий плющ з вузлуватим)».

Переваги Multi-Rake:

- можливість використовувати різні мови;
- налаштування кількості слів у словосполученнях, найменшого розміру слів, частоти та ін.;
- розрахування частоти (кількісної характеристики) для кожного обраного словосполучення;
- пошук окремих словосполучень замість зв'язного тексту, що дозволяє генерувати окремі ілюстрації для кожного образу у тексті.

Недоліками методу може бути:

- неконтрольований «обрив» словосполучень (не завжди за правилами вибраної мови);
- «зайві» слова, які ніяк не характеризують майбутню ілюстрацію.

Для подальшого використання результати методу можна обробляти, обираючи певну кількість необхідних словосполучень та вилучаючи з них беззмістовні слова.

4 ГЕНЕРАЦІЯ ЗОБРАЖЕНЬ

Другий етап роботи – це генерація зображень. Для цього можна використати нейронні мережі GAN, які можуть як створювати зображення, так і перетворювати наявні (рис. 4.1).



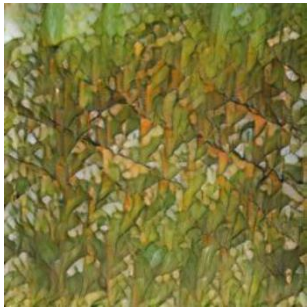
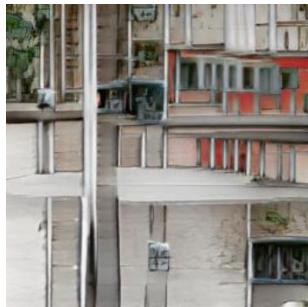
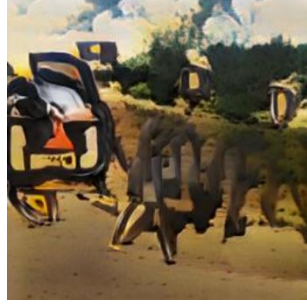
Рисунок 4.1 – Деякі види GAN для генерації зображень

Почнемо з AttnGAN. Як і у таблицях 3.1-3.2 аналізуються два тексти: «Останній лист» та «Старий і море» (англійською мовою). Як текст на вході, беруться результати роботи алгоритму Multi-Rake із таблиці 3.1. Отриманні результати показані у таблиці 4.1.

Таблиця 4.1 – Аналіз роботи AttnGAN для генерації зображень

«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)	
Вхідний текст	Отримане зображення	Вхідний текст	Отримане зображення
brick wall one ivy leaf		small tuna's shivering pull	

Кінець таблиці 4.1

«Останній листок» (О. Генрі)		«Старий і море» (Е. Хемінгуей)	
Вхідний текст	Отримане зображення	Вхідний текст	Отримане зображення
lone ivy leaf clinging		purple pectoral fins set wide	
In a little district west of Washington Square the streets have run crazy and broken themselves into small strips called "places."		After that he began to dream of the long yellow beach and he saw the first of the lions come down onto it in the early dark and then the other lions came and he rested his chin on the wood of the bows where the ship lay anchored with the evening off-shore breeze and he waited to see if there would be more lions and he was happy.	
		long yellow beach	

Для покращення результату виберемо тему генерації зображень і перевіримо результат. Наприклад, оберемо тему природи.

Спочатку застосуємо алгоритм RAKE для отримання ключових фраз тільки певної теми. Як було перевірено вище, RAKE може бути застосований до текстів на різних мовах (включаючи українську або російську).

Результат роботи алгоритму з російськомовним текстом «Старий і море» (з урахуванням теми «природа») продемонстрований у таблиці 4.2.

Таблиця 4.2 – Приклад роботи алгоритму RAKE (Multi-Rake)

Метод	Ключові слова – Старий і море (Е. Хемінгуей)
Rake	[('самый что ни на есть невезучий', 36.0)] [('луна волнует море как женщину', 25.0), ('— думал старик', 9.0)] [('в море не принято разговаривать без особой нужды', 64.0)] [('— интересно какой вид с такой высоты имеет море', 81.0)] [('разве можно заблудиться в море', 25.0)] [('близился вечер и кругом было видно лишь небо да море', 100.0)] [('мне жалко, что я ушел так далеко в море', 81.0)] [('— просто я слишком далеко ушел в море', 64.0)]

Виберемо ключову фразу з найвищим коефіцієнтом та згенеруємо для неї зображення за допомогою AttnGAN, використовуючи або фразу повністю, або окремі ключові слова. Наприклад, оберемо тільки іменники за допомогою NLP та бібліотеки NLTK (POS-тегування).

Отриманий результат можна подивитися у таблиці 4.3.

Таблиця 4.3 – Аналіз роботи AttnGAN для генерації зображень

Вхідний текст	Отримані зображення	
Близился вечер и кругом было видно лишь небо да море		
Вечер небо море		

При роботі алгоритму порогове значення присутності «природи» у тексті дорівнювало 2. Послідовні задовільні параграфи поєднувалися у один.

Отримані зображення досить абстрактні, нечіткі. Важко розпізнати, які об'єкти присутні. Інформація передається більш кольором, ніж формою. Проблема

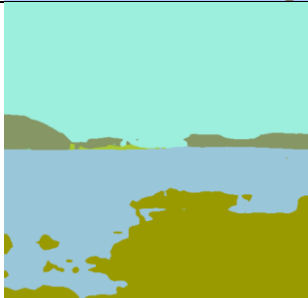
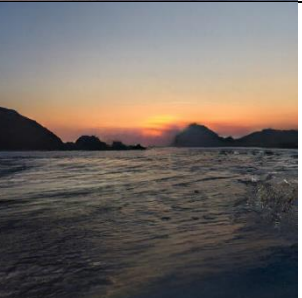
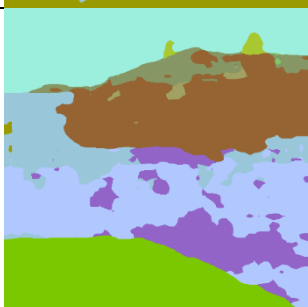
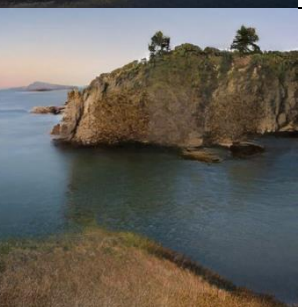
моделі в тому, що вона натренована недостатньо, щоб генерувати такі зображення, які потрібно, виходячи з тексту. Коло можливих зображень занадто велике. Чим більше класів зображень треба вміти генерувати, тим довше тренування моделі.

Так що зображення можна не генерувати, а обирати із набору даних, а потім трансформувати певним чином для отримання унікальних зображень.

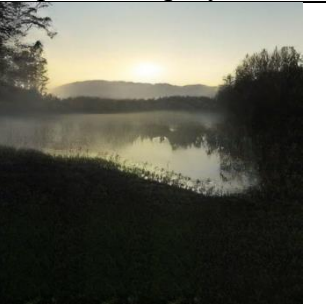
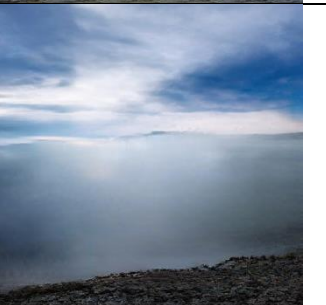
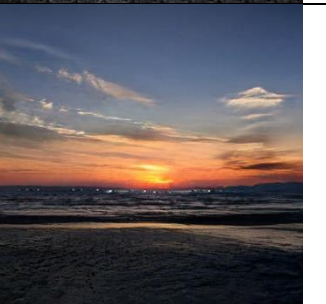
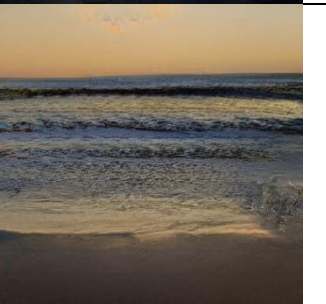
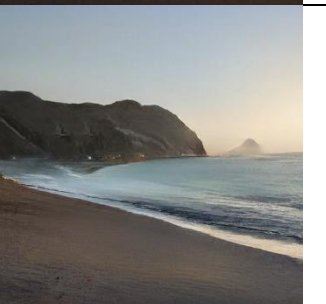
Приклад перетворення зображень за допомогою GauGAN можна подивитися у таблиці 4.4. Обраний текст – «Старий і море» (Е. Хемінгуей). Зображення були обрані на основі ключових слів, отриманих за допомогою бібліотеки Multi-Rake та приведених у таблиці 4.2 (вечір, небо, море).

Набір даних, із яких були обрані зображення нижче, називається Earth Porn і був узятий з Reddit [21]. Для зображень були власноруч проставлені відповідні мітки (теги) для трьох мов – англійської, української та російської.

Таблиця 4.4 – Аналіз роботи GauGAN для створення зображень

Знайдене зображення	Сегментація	Отриманий результат
		
		
		

Кінець таблиці 4.4

Знайдене зображення		Сегментація	Отриманий результат
			
			
			
			
			
			

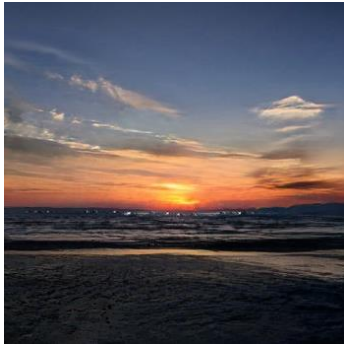
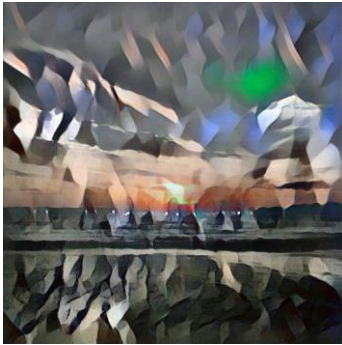
До отриманих зображень може бути застосований стиль інших зображень. Результат продемонстрований у таблиці 4.5.

Таблиця 4.5 – Приклад роботи алгоритму Neural Style

Вхідне зображення	Джерело стилю	Вихідне зображення	Подібність
			3 vs 1 MSE: 1546.63 SSIM: 0.63
			3 vs 2 MSE: 4497.75 SSIM: 0.36
			2 vs 1 MSE: 4685.83 SSIM: 0.31

Або стиль може бути застосований одразу для вибраних з датасету зображень до подальшої трансформації. Приклад застосування стилю відомих картин наведений у таблиці 4.6.

Таблиця 4.6 – Приклад роботи Neural Style зі стилем відомих картин

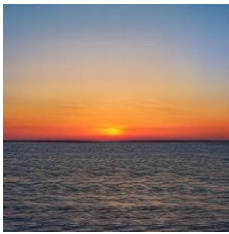
Вхідне зображення	Джерело стилю	Вихідне зображення	Подібність
			3 vs 1 MSE: 1630.56 SSIM: 0.59
			3 vs 2 MSE: 6217.69 SSIM: 0.16
			2 vs 1 MSE: 6080.97 SSIM: 0.22
			3 vs 1 MSE: 2482.25 SSIM: 0.25
			3 vs 2 MSE: 7118.99 SSIM: 0.04
			2 vs 1 MSE: 7699.41 SSIM: 0.08

Кінець таблиці 4.6

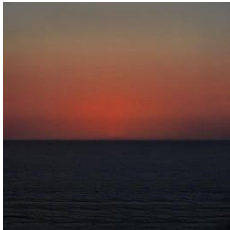
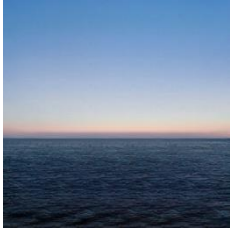
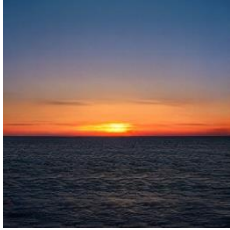
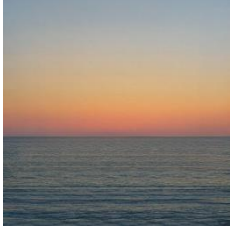
Вхідне зображення	Джерело стилю	Вихідне зображення	Подібність
			3 vs 1 MSE: 4232.52 SSIM: 0.20
			3 vs 2 MSE: 7017.57 SSIM: 0.03
			2 vs 1 MSE: 8542.09 SSIM: 0.11
			3 vs 1 MSE: 3128.65 SSIM: 0.44
			3 vs 2 MSE: 7945.17 SSIM: 0.06
			2 vs 1 MSE: 6737.61 SSIM: 0.15

Також у GauGAN може бути використана шаблонна сегментація, а знайдені зображення – як джерело стилю. Отриманий результат показаний у таблиці 4.7.

Таблиця 4.7 – Застосування стилю зображень до шаблонної сегментації

			
Вхідне зображення	Отримане зображення	Вхідне зображення	Отримане зображення
			

Кінець таблиці 4.7

			
Вхідне зображення	Отримане зображення	Вхідне зображення	Отримане зображення
			
			
			
			

На основі кожного зі знайдених зображень може бути згенеровано нове із додаванням шуму (та використанням попередньо тренованих моделей) за допомогою SinGAN. Отримані результати показані у таблиці 4.8.

У даному випадку (як і вище) подібність розраховується через середній квадрат помилки (MSE) та індекс структурної подібності (SSIM). Детальніше методи знаходження подібності будуть розглянуті у 5 розділі.

Таблиця 4.8 – Генерація зображень за допомогою SinGAN

«Старий і море» (Е. Хемінгуей)			
Вхідне зображення	Отримане зображення	Масштаб	Подібність
		Scale = 5	MSE: 1600.13 SSIM: 0.40
		Scale = 3	MSE: 550.87 SSIM: 0.64
		Scale = 3	MSE: 1213.44 SSIM: 0.40
		Scale = 4	MSE: 476.72 SSIM: 0.65
		Scale = 4	MSE: 1004.14 SSIM: 0.49

Кінець таблиці 4.8

«Старий і море» (Е. Хемінгуей)			
Вхідне зображення	Отримане зображення	Масштаб	Подібність
		Scale = 3	MSE: 178.28 SSIM: 0.73
		Scale = 5	MSE: 551.59 SSIM: 0.56

Як можна побачити, зображення, які були створені на основі інших зображень, більш реалістичні, ніж згенеровані на основі лише ключової інформації.

Для роботи з зображеннями (створення чи перетворення) використовувалися нейронні мережі GAN, такі як: AttnGAN, GauGAN, StyleGAN та SinGAN. Порівняємо ці нейронні мережі та наведемо результати порівняння у таблиці 4.9.

Таблиця 4.9 – Порівняння нейронних мереж для роботи із зображеннями

Критерії	AttnGAN	GauGAN	StyleGAN	SinGAN
Вхідні дані	Слово або словосполучення, або речення	Зображення або його сегментація	Зображення	Натренована модель зображення
Тренування	Одна модель	Одна модель	Модель для кожного стилю	Модель для кожного зображення
Достатній час тренування	Декілька днів	Декілька днів	Декілька годин	Декілька годин
Датасет для тренування	Зображення з описом	Зображення з мітками	Одне зображення	Одне зображення
Результат	Зображення	Зображення	Зображення	Зображення

Кінець таблиці 4.9

Критерії	AttnGAN	GauGAN	StyleGAN	SinGAN
Стиль результату	Не має чіткої відповідності	На основі одного із входів	На основі одного із входів	На основі входу
Сегментація результату	Не має чіткої відповідності	На основі входу	На основі іншого із входів	На основі входу

Таким чином кожна з нейронних мереж має свої переваги або на етапі тренування, або на етапі використання тренованої моделі. Вони мають різні входні та вихідні дані, різний спосіб роботи з ними.

Результат роботи цих нейронних мереж потрібно правильно оцінити. Саме реалістичність та відповідність входним даним є важливими складовими якості отриманого результату.

5 ОЦІНКА ОТРИМАНОГО РЕЗУЛЬТАТУ

Проведемо оцінку отриманого результату. Оцінити зображення можна власними очима (чи опитати інших людей) або автоматичним способом.

Проблеми, які можуть з'явитися при оцінюванні людиною:

- необхідність організувати опитування;
- важкість опитати значну кількість людей;
- можлива суб'єктивність думки.

Проблеми автоматичного оцінювання:

- складність реалізації;
- неможливість оцінити все одразу, необхідність виділення чітких критеріїв та обмежень.

Наскільки реалістичне зображення – достатньо абстрактне питання, тому на нього у першу чергу може відповісти саме людина.

Наприклад, три зображення на рисунку 5.1 згенеровані (перші два за допомогою GauGAN, третє – SinGAN). Вони є більш-менш реалістичними, хоча мають певні патерни, які повторюються. Зображення, згенеровані AttnGAN та наведені вище у 3 розділі, – менш реалістичні, тому що згенеровані без опору на наявну структуру зображення, присутні на зображенні об'єкти мають незвичні та дивні форми.

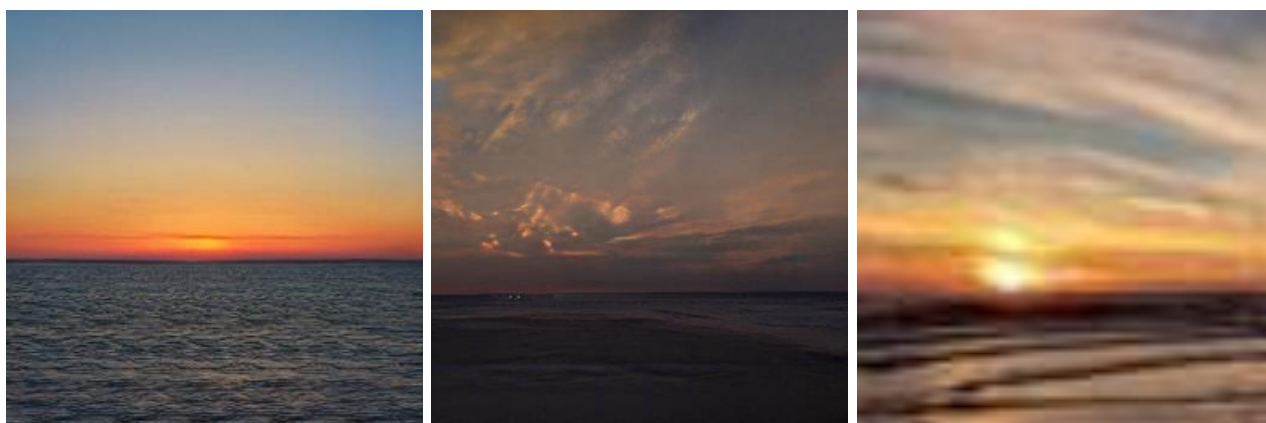


Рисунок 5.1 – Реалістично згенеровані зображення

Автоматичну оцінку можна отримати різними способами. Виділимо декілька факторів, які потребують оцінювання при створенні ілюстрації.

По-перше, треба оцінити наскільки отримане зображення подібно до вхідного. Це важливо, оскільки задача створення ілюстрації передбачає унікальність отриманого зображення.

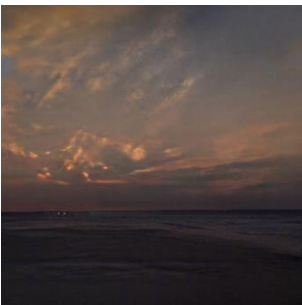
Наприклад, можна використати автокодувальники або вірогідний метод зниження розмірів багатовимірних даних (locality-sensitive hashing, LSH), або такі алгоритми, як середній квадрат помилки (MSE) та структурний індекс подібності (SSIM). SSIM намагається моделювати сприйману зміну структурної інформації зображення, тоді як MSE фактично оцінює сприйняття помилки [22].

Розмір порівнюваних зображень повинен бути однаковий.

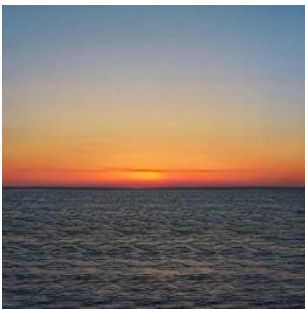
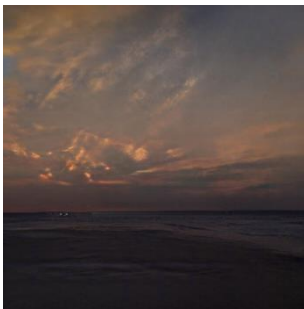
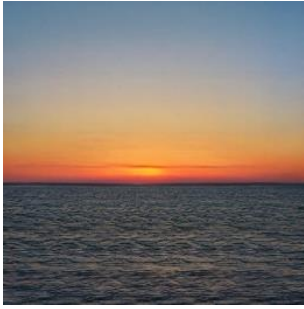
Саме за допомогою алгоритмів MSE та SSIM отримувалася оцінка у 3 розділі при стилізації зображення за допомогою StyleGAN та генерації зображення за допомогою SinGAN.

Наведемо ще один приклад у таблиці 5.1. Використовуються оригінальні та згенеровані зображення (які були показані на рисунку 5.1).

Таблиця 5.1 – Порівняння подібності двох зображень

Зображення 1	Зображення 2	Отримана оцінка	Характеристика
		MSE: 0.0 SSIM: 1.0	Однакові зображення
		MSE: 5410.42 SSIM: 0.36	Зображення 2 згенероване за допомогою GauGAN.

Кінець таблиці 5.1

Зображення 1	Зображення 2	Отримана оцінка	Джерело
		MSE: 2502.06 SSIM: 0.54	Зображення 1-2 згенеровані за допомогою GauGAN.
		MSE: 2131.86 SSIM: 0.41	Зображення 2 згенероване за допомогою GauGAN.
		MSE: 821.08 SSIM: 0.50	Зображення 2 згенероване за допомогою SinGAN.

Можна побачити, що алгоритм працює коректно на прикладі однакових зображень. Зображення згенеровані за допомогою SinGAN або GauGAN мають подібність менше ніж 50% та більш подібні між з собою, ніж із вхідним зображенням.

По-друге, можна оцінити, наскільки отримане зображення відповідає тексту, для якого воно було створено. Можна отримати мітки для зображень і порівняти їх з текстом, на основі якого вони були створені.

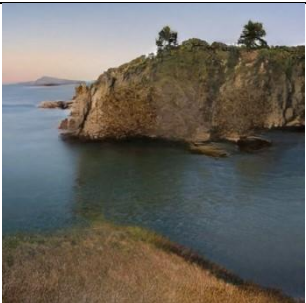
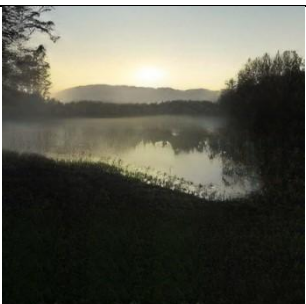
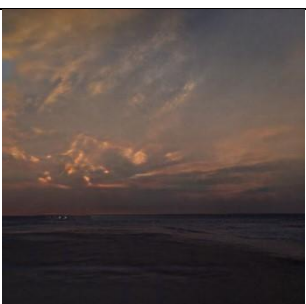
Класифікатор використовує власний багатозначний завантажувач даних та згорткову нейронну мережу (CNN). Цей класифікатор був натренований на тому ж датасеті, що і генератори (англомовна версія).

Приклад роботи класифікатору наведений у таблиці 5.2. Були оцінені усі вище згенеровані зображення. Також була введена суб'єктивна метрика реалістичності, яка показує оцінку зображень з моєї точки зору (5-бальна шкала).

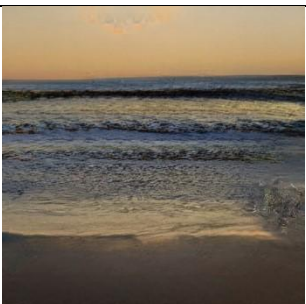
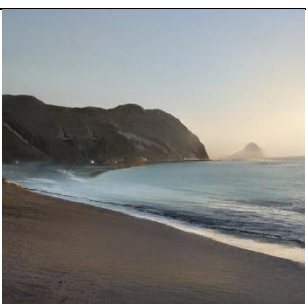
Таблиця 5.2 – Класифікація та оцінка зображень

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Близилася вечір і кругом було видно лише небо да море		river water sea snow valley moon ice mountain island winter dawn cloud night rock evening fog beach lake	Зображення згенероване за допомогою AttnGAN	Присутні зайві мітки	2
Близилася вечір і кругом було видно лише небо да море		river water sea snow moon ice mountain dark island winter dawn night cloud rock evening fog beach lake	Зображення згенероване за допомогою AttnGAN	Присутні зайві мітки	2
Вечір небо море		river water sea dark cloud rock evening lake ocean	Зображення згенероване за допомогою AttnGAN	Висока відповідність	3
Вечір небо море		river water sea road flower island black night evening	Зображення згенероване за допомогою AttnGAN	Присутні зайві мітки	2

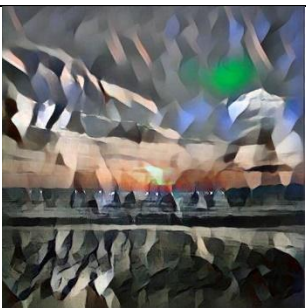
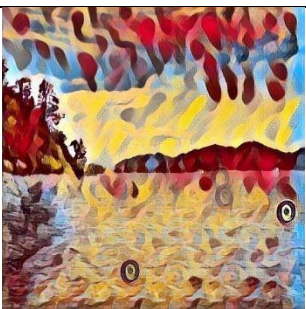
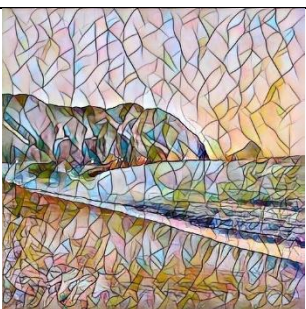
Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea snow orange dark night cloud dawn evening beach lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	3
Вечер небо море		river water sea orange mountain dark island black dawn rock evening lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	4
Вечер небо море		river water sea mountain island forest rock evening	Зображення згенероване за допомогою GauGAN	Висока відповідність	5
Вечер небо море		river water sea valley dark island black green dawn forest mountain fog evening lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	5
Вечер небо море		river water sea cloud orange mountain dark valley fog dawn evening ocean lake beach sand	Зображення згенероване за допомогою GauGAN	Присутні зайві мітки	5

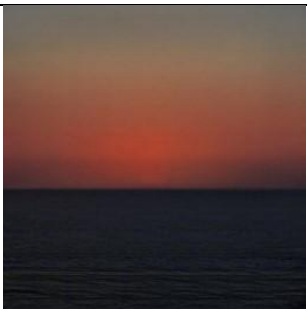
Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea mountain island green black cloud forest rock lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	4
Вечер небо море		river water sea snow ice moon cloud winter rock dark lake	Зображення згенероване за допомогою GauGAN	Присутні зайві мітки	4
Вечер небо море		river water sea orange ice black night cloud dawn rock evening beach lake	Зображення згенероване за допомогою GauGAN	Присутні зайві мітки	5
Вечер небо море		river water sea snow orange ice dark cloud dawn evening beach lake	Зображення згенероване за допомогою GauGAN	Присутні зайві мітки	4
Вечер небо море		river water sea valley mountain dark cloud dawn rock evening beach	Зображення згенероване за допомогою GauGAN	Висока відповідність	5

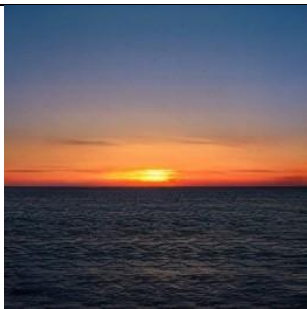
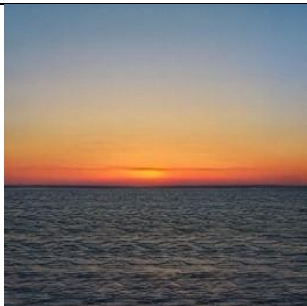
Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea valley mountain dark green forest dawn	Зображення згенероване за допомогою GauGAN та StyleGAN	Висока відповідність	5
Вечер небо море		river water sea red orange mountain lava island green dawn evening beach lake waterfall	Зображення згенероване за допомогою GauGAN та StyleGAN	Погана відповідність	4
Вечер небо море		river water sea flower snow valley orange leaf mountain dark island green rock palm beach lake coast	Зображення згенероване за допомогою GauGAN та StyleGAN	Погана відповідність	2
Вечер небо море		river water sea flower valley orange lava ground mountain island red dawn rock beach lake evening	Зображення згенероване за допомогою GauGAN та StyleGAN	Погана відповідність	3
Вечер небо море		river water sea snow flower valley orange island green dawn moss forest rock fog cloud palm beach	Зображення згенероване за допомогою GauGAN та StyleGAN	Погана відповідність	4

Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea orange dark black night red evening beach sand	Зображення згенероване за допомогою GauGAN	Висока відповідність	3
Вечер небо море		river water sea orange sun dark black red winter dawn night evening	Зображення згенероване за допомогою GauGAN	Висока відповідність	4
Вечер небо море		river water sea orange dark red dawn fog beach evening	Зображення згенероване за допомогою GauGAN	Висока відповідність	4
Вечер небо море		river water sea sun orange dark red black dawn evening beach lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	5
Вечер небо море		river water sea orange dark dawn fog evening beach lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	3

Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea sun orange dark black red night dawn evening beach lake ocean sand	Зображення згенероване за допомогою GauGAN	Висока відповідність	4
Вечер небо море		river water sea dark black dawn cloud fog evening beach lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	5
Вечер небо море		river water sea sun orange dark red black night dawn evening beach lake	Зображення згенероване за допомогою GauGAN	Висока відповідність	3
Вечер небо море		river water sea black dark night cloud fog evening beach sand	Зображення згенероване за допомогою GauGAN	Висока відповідність	3
Вечер небо море		river water sea orange red night dark black dawn evening beach	Зображення згенероване за допомогою GauGAN	Висока відповідність	3

Продовження таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea valley mountain dark island dawn fog rock evening beach	Зображення згенероване за допомогою SinGAN	Висока відповідність	3
Вечер небо море		river water sea sun beach orange lava dark dawn night evening	Зображення згенероване за допомогою SinGAN	Присутні зайві мітки	3
Вечер небо море		river water sea purple valley orange mountain dark black red island rock cloud evening dawn	Зображення згенероване за допомогою SinGAN	Присутні зайві мітки	3
Вечер небо море		river water sea mountain dark island dawn night cloud fog evening beach lake	Зображення згенероване за допомогою SinGAN	Висока відповідність	4
Вечер небо море		river water sea flower valley purple moon forest moss mountain dark green black night dawn evening fog	Зображення згенероване за допомогою SinGAN	Присутні зайві мітки	3

Кінець таблиці 5.2

Вхідний текст	Зображення	Вихідний текст	Джерело	Оцінка відповідності	Моя оцінка
Вечер небо море		river water sea orange mountain valley beach lava dark island dawn evening	Зображення згенероване за допомогою SinGAN	Присутні зайві мітки	3
Вечер небо море		river water sea sun orange dark red night dawn black evening mountain rock	Зображення згенероване за допомогою SinGAN	Висока відповідність	5
Вечер небо море		river water sea sun valley orange mountain dark island black dawn night evening cloud rock beach fog	Зображення згенероване за допомогою SinGAN	Присутні зайві мітки	4

Мітки розташовані від найбільш вірогідної до найменш.

«River (річка)», «water (вода)», «sea (море)» – є найчастіше передбаченими мітками для зображень, які повинні ілюструвати твір «Старий і море» Ернеста Хемінгуея. Також для деяких зображень високу ймовірність мали мітки «valley (долина)» та «orange (жовтогарячий)». Жовтогарячий – це характеристика сонця та вечірнього, передзахідного неба, а море на недостатньо чітких зображеннях може нагадувати долину. Таким чином зображення достатньо точно, хоч і не ідеально, показують текст, до якого вони були створені.

Більшість ілюстрацій, отримані за допомогою StyleGAN, мають велику кількість зайвих міток, як «lava (лава)» чи «flower (квітка)». Це очікувано, оскільки як джерело стилю використовувалися не тільки реальні зображення, а й не дуже реалістичні малюнки художників.

Наведемо результати оцінювання для кожної з використаних нейронних мереж у таблиці 5.3.

Таблиця 5.3 – Оцінювання різних нейронних мереж

Критерії	AttnGAN	GauGAN	StyleGAN	SinGAN
Моя оцінка (середнє значення)	2.25	4	3.6	3.5
Унікальність (метрики MSE та SSIM)	Висока унікальність (нема базового зображення)	MSE: 3348.113 SSIM: 0.44	MSE: 2604.122 SSIM: 0.422	MSE: 796.45 SSIM: 0.55
Відповідність тексту	Відповідає на достатньому рівні	Відповідає на високому рівні	Відповідає на поганому рівні	Відповідає на достатньому рівні

Як бачимо, найкращі результати генерації ілюстрацій на тему природи показала GauGAN. Але інші нейронні мережі також можуть застосовуватися, наприклад, для інших тем з певними обмеженнями. Все залежить від того, яку саме ілюстрацію потрібно згенерувати.

ВИСНОВКИ

В ході виконання дипломної роботи була проаналізована проблема та предметна область створення ілюстрацій до тексту, розглянуті методи роботи з текстом та генерації зображень, деякі з них були реалізовані, отримані результати були продемонстровані та оцінені.

Було запропоновано створювати ілюстрації до тексту у два етапи: стиснення тексту (отримання ключової інформації) та генерація зображення на основі результату першого етапу. Була поставлена задача дослідити методи для виконання кожного з етапів.

Для задачі стиснення тексту були перевірені реалізації алгоритму RAKE – Rake, Multi-Rake, Rake NLTK, а також алгоритми TextRank, LexRank, Luhn та LSA. Були розглянуті англійська, українська та російська мови.

Алгоритм RAKE був визнаний найбільш відповідним поставленій задачі, оскільки знаходить ключові слова, а не генерує цілісні анотації, а також має велику кількість налаштувань. Найкращий результат для роботи з декількома мовами показала бібліотека Multi-Rake, яка і була використана у подальшому дослідженні.

Перевага даного методу у тому, що він дозволяє отримувати ключові фрази для текстів на різних мовах, задавати параметри аналізу (найбільша кількість слів у фразі, найменша кількість букв у слові, найменша частота слова та ін.).

Результати роботи алгоритмів наведені у розділі 3.

Для задачі генерації зображень використовувалися різні GAN. Спочатку була обрана AttnGAN, яка показала посередній результат. Створені зображення були занадто абстрактні та неконтрольовані, оскільки був відсутній конкретний об'єкт для генерації, а коло можливих зображень було занадто велике.

Було з'ясовано, що ключова інформація тексту не завжди може бути проілюстрована. Можливість генерування будь-якого зображення ускладнювала навчання мережі та її подальше використання. Тому було запропоновано вибрати конкретну тему для створення ілюстрації та знайти ключові слова, опираючись на неї, щоб зробити генеровані зображення менш абстрактними.

Також було припущено, що результати можна покращити, використовуючи методи перетворення зображення на зображення, коли зображення генерується на основі зображення, а не тексту. Таким чином свобода створення GAN-зображень може поєднуватися з легким і точним підбором картинок. Якщо GAN отримує певний шаблон для створення зображення, отримані зображення будуть більш логічними та реалістичними.

Для перетворення зображення на зображення була вибрана GauGAN та тема природи. Тема природи була обрана, оскільки нейрона мережа GauGAN має підготовлені датасети та попередньо натреновану модель.

Поєднання перетворення тексту на зображення та перетворення зображення на зображення за допомогою GauGAN забезпечує кращі результати зображення з генерування тексту, ніж просто перетворення тексту на зображення за допомогою AttnGAN. З іншого боку обмеження датасету дозволяє контролювати AttnGAN та створювати чіткі зображення певного об'єкта (наприклад, квіти або птиці) та змінювати ці зображення у межах цього об'єкта (наприклад, кількість пелюсток або колір).

Також були показані приклади використання StyleGAN, де стиль одного зображення застосовується до іншого.

Наступною була розглянута SinGAN, яка дозволяє створювати зображення на основі лише однієї картинки. Модель SinGAN можна достатньо швидко натренувати, але для кожного зображення цю модель потрібно тренувати окремо. Було показане, що SinGAN генерує досить реалістичні зображення.

Результати роботи нейронних мереж наведені у розділі 4.

Отримані результати демонструють можливість створювати зображення для конкретно обраної теми з використанням попередньо підготовленого розміченого датасета.

У рамках подальших досліджень зображення, створені для різних тем, можна об'єднати. Наприклад, можна окремо створювати зображення людей за допомогою PoseGAN, об'єкти (наприклад, літак чи якусь тварину) за допомогою AttnGAN та фонове зображення (наприклад, пейзаж чи інтер'єр) за допомогою GauGAN.

Зображення людей можна генерувати з наявних, обираючи з тексту опис людини та дії, яка вона виконує. Тоді ці зображення можна поєднувати між собою, щоб отримати одне спільне зображення. З одного зображення можна буде зробити декілька варіацій, щоб користувач обрав, більш реалістичне на його думку, за допомогою SinGAN та StyleGAN.

Для кожної теми (типу зображення) мережі можна навчати окремо з певними параметрами та додатковими обмеженнями або умовами. У такому разі основною проблемою буде знайти відповідний датасет.

Результати дослідження можна використовувати як у науковій діяльності, продовжуючи дослідження, так і у практичній – створюючи ілюстрації до публікацій.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Karras T. A Style-Based Generator Architecture for Generative Adversarial Networks / T. Karras, S. Laine, T. Aila // IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – 2019. – P. 4401-4410.
2. Gatys L. A. A Neural Algorithm of Artistic Style / L. A. Gatys, A. S. Ecker, M. Bethge [Electronic resource] // arXiv e-prints, arXiv:1508.06576v2 – 2015. – P. 1.
3. Image-to-Image Translation with Conditional Adversarial Networks / P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros // IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – 2017. – P. 1125-1134.
4. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks / J.-Y. Zhu, T. Park T., P. Isola, A. A. Efros // IEEE International Conference on Computer Vision (ICCV) – 2017. – P. 2223-2232.
5. Perarnau G. Fantastic GANs and where to find them [Electronic resource] / G. Perarnau // Guim Perarnau Blog – 2017. – Regime of access: <http://guimperarnau.com/blog/2017/03/Fantastic-GANs-and-where-to-find-them/>
6. Brock A. Large Scale GAN Training for High Fidelity Natural Image Synthesis / A. Brock, J. Donahue, K. Simonyan [Electronic resource] // arXiv e-prints, arXiv:1809.11096v2 – 2019. – P. 8.
7. AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks / T. Xu, P. Zhang, Q. Huang, et al. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – 2018. – P. 1316-1324.
8. Shafkat I. Intuitively Understanding Variational Autoencoders [Electronic resource] / I. Shafkat // Towards Data Science – 2018. – Regime of access: <https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf/>
9. Semantic Image Synthesis with Spatially-Adaptive Normalization / T. Park, M.-Y. Liu, T.-C. Wang, J.-Y. Zhu // IEEE Conference on Computer Vision and Pattern Recognition (CVPR) – 2019. – P. 2337-2346.
10. Ouaknine A. Review of Deep Learning Algorithms for Image Semantic Segmentation [Electronic resource] / A. Ouaknine // Medium – 2018. – Regime of access:

https://medium.com/@arthur_ouaknine/review-of-deep-learning-algorithms-for-image-semantic-segmentation-509a600f7b57/

11. Shaham T. R., Dekel T., Michaeli T. SinGAN: Learning a Generative Model from a Single Natural Image // IEEE International Conference on Computer Vision (ICCV) – 2019. – P. 4570-4580.

12. Automatic Keyword Extraction from Individual Documents / S. Rose, D. Engel, N. Cramer, W. Cowley // M. W. Berry & J. Kogan (Eds.), Text Mining: Theory and Applications – John Wiley & Sons, Hoboken, 2010. – P. 1-20.

13. Ondeniyi E. Extractive Text Summarization Techniques With sumy [Electronic resource] / E. Ondeniyi // Medium. – 2017. – Regime of access: <https://medium.com/@ondeniyi.eric/extractive-text-summarization-techniques-with-sumy-3d3b127a0a32>

14. Naskar A. Extract Custom Keywords using NLTK POS tagger in python sumy [Electronic resource] / A. Naskar // ThinkInfi. – 2018. – Regime of access: <https://www.thinkinfi.com/2018/10/extract-custom-entity-using-nltk-pos.html>

15. WordNet [Electronic resource] // Princeton University – Regime of access: <https://wordnet.princeton.edu/>

16. Pose Guided Person Image Generation / L. Ma, X. Jia, Q. Sun, et al. // Advances in Neural Information Processing Systems. – 2017. – P. 405-415.

17. COCO [Electronic resource] – Regime of access: <http://cocodataset.org/>

18. ImageNet [Electronic resource] – Regime of access: <https://image-net.org/>

19. PyTorch [Electronic resource] – Regime of access: <https://pytorch.org/>

20. Multi-Rake [Electronic resource] // The Python Package Index – Regime of access: <https://pypi.org/project/multi-rake/>

21. r/EarthPorn [Electronic resource] // Reddit – Regime of access: <https://www.reddit.com/r/EarthPorn/>

22. Rosebrock A. How-To: Python Compare Two Images [Electronic resource] / A. Rosebrock // PyImageSearch. – 2014. – Regime of access: <https://pyimagesearch.com/2014/09/15/python-compare-two-images/>