

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра Медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий (магістерський)
(позначення документа)

Комплексна тема. Використання нейромереж у роботі з фото та відео.
Генерація матеріалів.
(тема)

Виконав:

студент 2 курсу, групи СТМм-22-1
Єгор САМОЙЛИК

(прізвище, ініціали)

Спеціальність 171 – Електроніка

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Системи, технології і комп'ютерні засоби мультимедіа

(повна назва освітньої програми)

Керівник ст. викл. Тетяна ВАСИЛЕНКО

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри

(підпис)

Володимир КАРТАШОВ

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформації

Кафедра Медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 171 – Електроніка

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма "Системи, технології і комп'ютерні засоби мультимедіа "

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« ____ » _____ 20 ____ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студентові Самойлику Єгору Сергійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи: **Комплексна тема. Використання нейромереж у роботі з фото та відео. Генерація матеріалів.**

затверджена наказом по університету від " 20.11.2023"р. № 1371Ст

2. Термін подання студентом роботи 08.01.2024 р.

3. Вихідні дані до проекту (роботи) _____

1. Огляд та дослідження принципів роботи нейромереж для роботи з медіаконтентом

2. Огляд програмного забезпечення

3. Практичне застосування нейромереж у створенні сценарію, фото та відео

4. Перелік питань, що потрібно опрацювати в роботі

ВСТУП

1.ОГЛЯД ПРИНЦИПІВ СТВОРЕННЯ МУЛЬТИМУДІА НЕЙРОМЕРЕЖЕЮ

2.АНАЛІЗ ТА ПОРІВНЯННЯ СЕРВІСІВ, ЩО ВИКОРИСТОВУЮТЬ НЕЙРО-МЕРЕЖІ ДЛЯ СТВОРЕННЯ ФОТО- ТА ВІДЕОКОНТЕНТУ

3.ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖ У СТВОРЕННІ СЦЕНАРІЮ ,ФОТО, ТА ВІДЕО

ВИСНОВКИ

ПЕРЕЛІК ПОСИЛАННЯ

ДОДАТКИ

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій:

1. Вступ; 2. Основні відомості про роботу; 3. Генерація відео нейромережею; 4. Процес створення фото нейромережею; 5. Аналіз та порівняння сервісів, що використовують нейромережі для створення фото- та відеоконтенту; 6. Практичне застосування нейромереж у створенні контенту; 7. Результати генерації медіаконтенту.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термин виконання етапів роботи	Примітка
1.	Отримання теми дипломної роботи	21.11.23-27.11.23	
2.	Підготовка теоретичної частини	28.11.23-02.12.23	
3.	Робота з технічною частиною	04.12.23-09.12.23	
4.	Оформлення дипломної роботи	11.12.23-20.12.23	
5.	Графічна частина роботи	21.12.23-23.12.23	
6.	Перевірка керівником	24.12.23-26.12.23	
7.	Перевірка на академічний плагіат	27.12.23-28.12.23	
8.	Перевірка завідувачем кафедри, рецензу-	29.12.23-02.01.24	

Дата видачі завдання _____ 20.11.2023 р. _____



Студент _____
(підпис)

_____ Єгор САМОЙЛИК _____

Керівник роботи _____
(підпис)

_____ ст. викл. Тетяна ВАСИЛЕНКО _____

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи магістра має: 58 сторінки, 22 джерел, 16 рисунка.

НЕЙРОННІ МЕРЕЖІ, ГЛИБОКЕ НАВЧАННЯ, МУЛЬТИМЕДІЙНИЙ ВМІСТ, ШТУЧНИЙ ІНТЕЛЕКТ, ОБРОБКА МУЛЬТИМЕДІА, СЦЕРІЙ, ВИБІР КАДРІВ, ФОТО, ВІДЕО.

Об'єкт дослідження - використання нейронних мереж, зокрема глибокого навчання, для обробки та творчого застосування мультимедійного контенту. Робота охоплює широкий спектр завдань, включаючи створення сценаріїв, тексту, фото та відео в різних стилях.

Предмет дослідження - вивчення та впровадження передових технологій у сфері обробки мультимедійного контенту. Робота зосереджена не лише на аналізі сучасних методів маніпулювання мультимедійним контентом, але й на розробці нових підходів та використанні нейронних мереж для підвищення якості та ефективності обробки.

Мета дослідження - виявлення нових можливостей для мультимедійних професіоналів та знаходження широкого застосування у різних галузях. Робота виходить за межі простого аналізу та огляду, спрямовуючись на розробку інноваційних підходів і використання нейронних мереж.

У кваліфікаційній роботі використовуються методи глибокого аналізу та експериментальні підходи для вивчення функціонування нейронних мереж у мультимедійній обробці. Дослідження охоплює застосування нейронних мереж у створенні мультимедійного контенту та їхні практичні застосування у сценарію, фото та відео.

Структура магістерської роботи включає в себе вступ, три основні розділи, висновки та список використаної літератури. Дана кваліфікаційна робота має велике теоретичне та практичне значення, а її результати можуть внести вагомий внесок у галузь нейронних мереж у мультимедійному сегменті та розвиток вдосконалених методологій навчання.

Кваліфікаційна робота має велике теоретичне та практичне значення. Її результати можуть бути важливим внеском у подальші дослідження в галузі нейронних мереж у мультимедійному сегменті та розробки вдосконалених методологій навчання для підвищення кваліфікації у цьому напрямку.

ABSTRACT

Explanatory note for the master's qualification work includes: 58 pages, 22 sources, 16 figures.

NEURAL NETWORKS, DEEP LEARNING, MULTIMEDIA CONTENT, ARTIFICIAL INTELLIGENCE, MULTIMEDIA PROCESSING, SCENARIOS, TEXT, PHOTO, VIDEO.

Research Object - the use of neural networks, particularly deep learning, for processing and creative application of multimedia content. The work covers a wide range of tasks, including the creation of scenarios, text, photos, and videos in various styles.

Research Subject - studying and implementing advanced technologies in the field of multimedia content processing. The work focuses not only on analyzing modern methods of manipulating multimedia content but also on developing new approaches and using neural networks to enhance the quality and efficiency of processing.

Research Objective - identifying new opportunities for multimedia professionals and finding broad applications in various industries. The work goes beyond simple analysis and review, aiming at the development of innovative approaches and the utilization of neural networks.

In the master's thesis, deep analysis methods and experimental approaches are employed to study the functioning of neural networks in multimedia processing. The research covers the application of neural networks in creating multimedia content and their practical use in scenarios, photos, and videos.

The structure of the master's thesis includes an introduction, three main sections, conclusions, and a list of references. This qualification work has significant theoretical and practical significance, and its results can make a substantial contribution to the field of neural networks in the multimedia segment and the development of advanced teaching methodologies.

The master's qualification work holds great theoretical and practical significance. Its results can be a crucial contribution to further research in the field of neural networks in the multimedia segment and the development of enhanced teaching methodologies to improve qualifications in this direction.

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

- NN - Neural Network - Нейромережа
- ML - Machine Learning - Машинне навчання
- DL - Deep Learning - Глибоке навчання
- CNN - Convolutional Neural Network - Згорткова Нейромережа
- RNN - Recurrent Neural Network - Рекурентна Нейромережа
- GAN - Generative Adversarial Network - Генеративно-антиагентська мережа
- LSTM - Long Short-Term Memory - Довга Короткочасна Пам'ять
- ReLU - Rectified Linear Unit - Лінійна функція активації з випрямленим прямокутником
- ANN - Artificial Neural Network - Штучна нейромережа
- AI - Artificial Intelligence - Штучний Інтелект
- DNN - Deep Neural Network - Глибока Нейромережа
- GPU - Graphics Processing Unit (often used for accelerated neural network training)
- Відеокартковий Графічний Процесор (часто використовується для прискореної тренування нейромереж)
- IoT - Internet of Things (applications of neural networks in IoT) - Інтернет речей (застосування нейромереж в IoP)
- BERT - Bidirectional Encoder Representations from Transformers -
- PCA - Principal Component Analysis -
- SVM - Support Vector Machine (not exclusive to neural networks but relevant in machine learning contexts) - Побудова взаємодії Кодера за допомогою Перетворень (англ. Bidirectional Encoder Representations from Transformers)
- MSE - Mean Squared Error (a common loss function in neural network training) - Середньоквадратичне Відхилення (звичайна функція втрат при тренуванні нейромереж)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ.....	8
ВСТУП.....	11
1 ОГЛЯД ПРИНЦИПІВ СТВОРЕННЯ МУЛЬТИМУДІА НЕЙРОМЕРЕЖЕЮ.....	12
1.1 Генерація відео нейромережею	13
1.1.1 Створення відео нейромережею на основі фото.....	15
1.1.2 Створення відео нейромережею на основі текстового запиту	17
1.1.3 Редагування відео нейромережею завдяки вхідним тестовим запитам .	18
1.2 Процес створення фото нейромережею.....	19
1.2.1 Процес створення фото нейромережею завдяки текстового запиту	20
1.2.2 Огляд архітектур GANs та VAEs.....	21
1.3 Створення сценарію штучним інтелектом	24
1.3.1 Використання Нейромереж та Алгоритмів Глибокого Навчання.	25
2 АНАЛІЗ ТА ПОРІВНЯННЯ СЕРВІСІВ, ЩО ВИКОРИСТОВУЮТЬ НЕЙРОМЕРЕЖІ ДЛЯ СТВОРЕННЯ ФОТО- ТА ВІДЕОКОНТЕНТУ	27
2.1 Сервіси та програмні рішення на основі штучного інтелекту для створення фото.....	27
2.1.1 Сервіс Midjourney.....	28
2.1.2 Сервіс Blue Willow	30
2.1.3 Онлайн сервіс Stable Diffusion Online	31
2.2 Сервіси та програмні рішення на основі штучного інтелекту для створення відео	34
2.2.1 Онлайн сервіс Runway	34
2.2.2 Онлайн сервіс Pictory	37
2.3 Нейромережі для створення сценарію	39

	10
2.3.1 Онлайн ресурс GPT	40
2.3.2 Сервіс SoCreate	43
3 ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖ У СТВОРЕННІ СЦЕНАРІЮ ,ФОТО, ТА ВІДЕО	46
3.1 Створення сценарію завдяки ChatGPT	47
3.2 Створення фото завдяки сервісу Stable Diffusion Online	50
3.3 Створення відео завдяки онлайн сервісу Runway.....	51
ВИСНОВКИ.....	55
ПЕРЕЛІК ПОСИЛАНЬ	56
Додатки.....	59
Додаток А.....	60
Додаток Б	Ошибка! Закладка не определена.

ВСТУП

Сучасний етап технологічного розвитку визначає необхідність поєднання мистецтва та техніки для створення вражаючих та високоякісних фото- та відеопроєктів. Завдяки стрімкому розвитку і широкому застосуванню нейромереж, відкриваються нові перспективи в області обробки та творення візуального контенту.

Об'єднуючи можливість нейромереж з творчістю у сфері фотографії та відеозйомки, ми можемо досягти надзвичайних результатів у створенні сценаріїв та композицій для фото- та відеорядів. Данна кваліфікаційна робота спрямована на вивчення та реалізацію можливостей використання нейромереж у творчому процесі створення візуальних шедеврів.

У кваліфікаційній роботі магістра розглянуто вплив нейромереж на обробку фотографій та відеозаписів, а також їхнє використання у створенні унікальних сценаріїв та композицій. Зосереджуючись на злитті технологій та творчості, ми розглянемо можливості автоматизації та покращення процесів обробки візуального контенту за допомогою нейромереж.

У кваліфікаційній роботі ставить перед собою завдання розкрити потенціал нейромереж у сфері фотографії та відеомистецтва, дослідити їхні можливості та внести нові підходи до створення сценаріїв та фото-, відеорядів. Результати даного дослідження можуть виявити великий вплив на розвиток сучасних технологій у галузі творчості та візуального мистецтва..

1 ОГЛЯД ПРИНЦИПІВ СТВОРЕННЯ МУЛЬТИМЕДІА НЕЙРОМЕРЕЖЕЮ

Створення мультимедійних нейромереж включає в себе використання глибокого навчання та нейромереж для обробки та генерації мультимедійного вмісту, такого як зображення, відео, аудіо, або комбінації елементів. Основними етапами створення мультимедійних нейромереж є навчання, архітектура та застосування для конкретних задач. Збір та підготовка даних.

Збір даних. Отримання великої кількості мультимедійних даних, які будуть використовуватися для тренування нейромережі. Це може бути фотографії, відео, аудіозаписи, текстові описи і так далі.

Анотація даних. Розмітка даних інформацією, яка визначає класи об'єктів на зображеннях або надає інші необхідні параметри для навчання моделі. Наприклад, у випадку класифікації зображень, кожне зображення може бути позначене певним класом.

Тип нейромережі. Визначення, яка архітектура нейромережі буде використовуватися в залежності від задачі. Наприклад, для обробки зображень часто використовуються згорткові нейромережі (CNN), а для обробки послідовностей (таких як аудіо або відео) - рекурентні нейромережі (RNN).

Використання даних для тренування. Подача підготовлених даних на вхід нейромережі з метою налаштування її ваг для виконання конкретної задачі.

Вибір функції втрат. Визначення метрики, яка визначає, наскільки точно модель виконує свою задачу. Це може бути крос-ентропійна втрата для класифікації або середньоквадратична втрата для регресії.

Оптимізація ваг. Використання алгоритмів оптимізації для налаштування ваг нейромережі. Це може включати в себе стохастичний градієнтний спуск (SGD) або більш розширені методи, такі як Adam.

Валідація результатів. Використання окремого набору даних для оцінки продуктивності навчання. Це допомагає визначити, чи модель вчиться коректно, чи можливо виникає перенавчання.

Модифікація гіперпараметрів. виправлення гіперпараметрів, таких як швидкість навчання або розмір пакету, для поліпшення результатів. Це процес ітеративного налаштування моделі.

Тестування на тестовому наборі. Використання окремого набору даних для оцінки загальної ефективності моделі, яка не брала участі в тренуванні або валідації.

Впровадження моделі. Застосування навченої моделі для реальних завдань. Наприклад, використання моделі для обробки зображень у фотокамері або розпізнавання мови в голосовому асистенті.

Адаптація до нових даних. Розгляд можливостей для постійного навчання, яке дозволяє моделі адаптуватися до нових даних та покращувати свої навички з часом.

Даний процес представляє собою ітеративний цикл, де кожен етап дозволяє моделі навчатися більше і адаптуватися до конкретної задачі обробки мультимедіа.

1.1 Генерація відео нейромережею

Створення відео за допомогою нейромережі - це складний та цікавий процес, який включає в себе кілька етапів. Давайте розглянемо їх детально.

Збір та підготовка даних. Перший етап включає в себе збір великого обсягу відеоданих, які будуть використовуватися для навчання нейромережі. Дані можуть включати в себе відеозаписи різного типу та стилів. Перед тим, як нейромережа розпочне навчання, дані потрібно підготувати, а це включає в себе обрізку, зміну роздільної здатності, конвертацію формату тощо.

Вибір архітектури нейромережі для генерування відео. Вибір правильної архітектури нейромережі визначається завданням, яке потрібно виконати. Для створення відео можуть використовуватися різні архітектури, такі як глибокі конволюційні мережі (CNN), рекурентні нейронні мережі (RNN) або комбінації обох.

Навчання моделі. Після вибору архітектури необхідно провести процес навчання нейромережі на підготовлених даних. Під час даного етапу модель навчається розпізнавати патерни та особливості відеоданих. Навчання вимагає значної обчислювальної потужності та може займати тривалий час, в залежності від обсягу даних та складності завдання.

Тестування та оцінка. Після завершення навчання модель перевіряється на тестових даних для оцінки її точності та ефективності. Це включає в себе аналіз результатів, виявлення можливих проблем та виправлення їх.

Постобробка та виробництво. Після успішного навчання та тестування модель може бути застосована для створення відео. Даний етап включає в себе ряд операцій, таких як генерація відеокадрів, їх об'єднання в кінцевий відеоряд та можливий постобробка для покращення якості відео.

Оптимізація та вдосконалення. Після випуску відео може виникнути необхідність у подальших оптимізаціях та вдосконаленнях. Це може включати в себе налаштування параметрів моделі, удосконалення алгоритмів або додавання нових даних для покращення загальної продуктивності.

Обробка та постобробка Відео. Після створення відео нейромережею, може знадобитися додатковий постобробка для поліпшення якості та візуальної привабливості. Це може включати в себе корекцію кольорів, розфокусування, підвищення чіткості та інші ефекти. Також можуть застосовуватися техніки аугментації для різноманітності та створення унікальних візуальних ефектів.

Валідація та застосування. Після обробки відео важливо провести валідацію результатів. Це включає перевірку на наявність артефактів, аномалій або неочікуваного змісту. Після успішної валідації відео може бути використане для різних цілей, таких як розваги, навчання, медицина, дизайн тощо.

Ітеративний процес. Розробка відео за допомогою нейромережі - це ітеративний процес. Після валідації та застосування результатів можуть виникнути нові ідеї або виявитися області для вдосконалення. Це може призвести до додаткових раундів навчання, оптимізації та покращень моделі.

Масштабування та впровадження. Якщо відео успішно пройшло всі етапи розробки та валідації, може виникнути необхідність у масштабуванні та впровадженні. Це включає в себе оптимізацію для використання на різних платформах та пристроях, а також врахування можливих обмежень щодо швидкості та ресурсів.

Супровід та підтримка. Після впровадження важливо забезпечити належний супровід та підтримку. Це включає в себе виявлення та вирішення можливих проблем, а також взаємодію з користувачами для отримання зворотнього зв'язку та вдосконалення продукту з часом.

Створення відео нейромережею - це складний та багатоетапний процес, який вимагає поєднання експертизи в області машинного навчання, обробки відео та творчого підходу для досягнення оптимальних результатів.

1.1.1 Створення відео нейромережею на основі фото

Процес створення відео на основі фотографій за допомогою нейромережі може включати в себе використання генеративних моделей, зокрема глибоких генеративних моделей (GAN) або автоенкодерів. В цьому контексті розглянемо основний підхід, що використовує GAN:

Вибір та Підготовка Даних. Починаючи зі створення відео на основі фотографій, потрібно зібрати набір фотографій, який буде використовуватися для навчання моделі. Фотографії можуть представляти послідовність подій або сценарії, які ви хочете перетворити в відео.

Вибір Генеративної Моделі (GAN). Глибокі генеративні моделі, такі як GAN, виявилися ефективними для генерації нових зображень на основі навчальних даних. Нейромережа GAN складається з двох компонентів - генератора та дискримінатора, які конкурують між собою. Генератор створює нові фрейми відео на основі вхідного зображення, а дискримінатор визначає, наскільки дані фрейми виглядають правдоподібно.

Навчання Генератора та Дискримінатора. В процесі навчання генератора і дискримінатора взаємодіють між собою. Генератор намагається створити відтворення фотографій у вигляді відео, тоді як дискримінатор намагається визначити, чи є фрейми реальними чи створеними генератором. Даний процес триває до того моменту, коли генератор стає досить добре обманювати дискримінатор.

Генерація відео. Після навчання можна використовувати генератор для створення нових кадрів, які утворюють відео. Вхідні фотографії подаються в генератор, який перетворює їх у відеошлях.

Постобробка та обробка. Отримане відео може потребувати додаткового постобробку для поліпшення якості та естетики. Це може включати в себе корекцію кольорів, зменшення шуму, підвищення чіткості та інші ефекти для досягнення бажаного візуального ефекту.

Валідація та оцінка. Після генерації важливо провести валідацію та оцінку отриманого відео. Це включає в себе перевірку наявності артефактів, реалістичності та відповідності вихідних фотографій.

Оптимізація та виробництво. Залежно від результатів валідації може знадобитися оптимізація параметрів генеративної моделі або додатковий розвиток для досягнення кращих результатів. Також важливо підготувати модель для використання в реальних умовах та на різних пристроях.

Створення відео на основі фотографій залежить від глибокого розуміння та оптимізації генеративних моделей, а також від творчого підходу до обробки та підготовки даних.

1.1.2 Створення відео неймережею на основі текстового запиту

Створення відео на основі текстового запиту за допомогою неймережі включає в себе використання технік генерації відео з текстового опису, або так званих Conditional Video Generation моделей. Нижче подано основні етапи даного процесу:

- Збір та підготовка даних. Починаємо зі збору та підготовки даних для навчання моделі. Це може включати в себе велику базу відео, яка відповідає текстовим описам. Дані потрібно анотувати так, щоб кожне відео було пов'язане з конкретним текстовим описом, що буде використовуватися для навчання моделі;
- Вибір та побудова моделі. Обираємо або створюємо модель, здатну генерувати відео на основі текстового опису. Можливі варіанти включають глибокі генеративні моделі, такі як GANs або варіації автоенкодерів. Модель повинна бути навчена взаємозв'язку між текстом та відео;
- Навчання моделі. Під час навчання модель використовує пари текст-відео для встановлення зв'язку між текстовим описом і його візуальним вираженням у вигляді відео. Модель у процесі навчання вчитиметься відтворювати візуальні елементи, що відповідають текстовому опису;
- Генерація відео. Після навчання модель може використовуватися для генерації відео на основі текстових запитів. Текстовий опис введений в модель в якості вхідних даних, і модель створює відео, яке відображає зміст даного опису;
- Валідація та тюнінг. Важливо провести валідацію результатів, перевірити, наскільки відео відповідає текстовому запиту. Якщо результати не задовільняють, може знадобитися додатковий тюнінг параметрів моделі або розширення та оптимізація даних для навчання;
- Постобробка виправлення. Отримане відео може потребувати додаткового постпроцесингу для покращення якості, такого як зменшення шуму, корекція кольорів або додавання спеціальних ефектів;

– Виробництво та впровадження. Після успішного навчання та валідації модель може бути впроваджена для генерації відео на основі текстових запитів в реальних умовах.

Даний процес вимагає глибокого розуміння як текстових, так і візуальних даних, а також використання складних генеративних моделей для ефективної генерації відео на основі текстового контексту.

1.1.3 Редагування відео нейромережею завдяки вхідним тестовим запитам

Введення та обробка додаткової інформації. Після створення відео нейромережею можна вводити додаткові елементи, такі як ефекти або текстові вставки. Це може бути здійснено шляхом додавання додаткового візуального або текстового контенту до вихідного відео.

Використання алгоритмів Обробки Відео. Нейромережі можуть використовувати алгоритми обробки відео для додавання різних ефектів. Наприклад, використання алгоритмів розпізнавання обличчя для додавання фільтрів або віртуальних об'єктів до кадрів.

Вставка тексту. Нейромережі можуть бути програмовані для вставки текстової інформації у відео. Це може включати в себе додавання заголовків, підписів або будь-якого іншого тексту, що відобразиться на екрані під час відтворення.

Редакція звукового супроводу. Нейромережі можуть також застосовувати редагування до аудіосупроводу. Це може бути звукові ефекти, фонові музика чи будь-які інші аудіоелементи, які доповнюють візуальний зміст відео.

Аналіз та застосування ефектів. Нейромережі можуть проводити аналіз відео та використовувати ефекти для покращення візуального враження. Це може бути підвищення контрастності, корекція кольорів або введення стилізаційних ефектів.

Спеціалізовані моделі для редагування. Існують спеціалізовані моделі для редагування відео, які можуть використовувати глибоке навчання для автоматизованого видалення об'єктів, трекінгу рухомих об'єктів, або навіть автоматизованого вирізання та об'єднання сцен.

Валідація та доналаштування. Після редагування важливо провести валідацію результатів, перевірити, наскільки добре додані ефекти чи текстові вставки інтегруються з вихідним відео. При необхідності можуть застосовуватися доналаштування для досягнення оптимального результату.

Збереження та вивід. Після внесення необхідних змін відео може бути збережене у відповідному форматі або виведене для подальшого використання.

1.2 Процес створення фото нейромережею

Створення фотографій за допомогою нейромереж є складним процесом, який включає кілька етапів. Ось докладний опис кожного кроку:

Збір даних. Перший крок полягає у зборі великої кількості фотографій для навчання нейромережі. Дані можуть бути розмічені (з вказанням об'єктів або сцен на фото) або нерозмічені, залежно від задачі. Наприклад, для генерації фотографій обличчя людей, можуть використовуватися зображення з баз даних обличчя, таких як CelebA.

Передобробка даних. Зображення піддаються передобробці, щоб зменшити розмір файлів і полегшити їх обробку. Даний етап також може включати нормалізацію, розширення даних (зміна розміру, обрізка), а також інші техніки для покращення якості даних.

Створення архітектури нейромережі. Вибір архітектури нейромережі є ключовим етапом. Для генерації фотографій часто використовують глибокі генеративні моделі, такі як GANs (Generative Adversarial Networks) або VAEs (Variational Autoencoders). ГАН-моделі включають у себе генератор і дискримінатор, які взаємодіють для навчання мережі створювати вірогідно реалістичні зображення.

Навчання моделі. Під час даного етапу модель навчається на основі набору даних, щоб вона могла вивчити кореляції та особливості об'єктів на фотографіях. Генератор намагається створити реалістичні зображення, тоді як дискримінатор оцінює, наскільки вони схожі на реальні.

Оптимізація та налаштування. Модель піддається оптимізації для покращення її продуктивності та здатностей до генерації високоякісних зображень. Це може включати в себе налаштування параметрів, вибір оптимальних гіперпараметрів та вдосконалення алгоритмів.

Генерація фотографій. Після завершення навчання, нейромережа готова генерувати нові фотографії. Даний процес включає в себе введення випадкового вектора або зображення, яке потрібно згенерувати, у генератор, який повертає створене зображення.

Оцінка та вдосконалення. Фінальний етап включає оцінку якості згенерованих фотографій та, в разі необхідності, вдосконалення моделі. Оцінка може проводитися за допомогою різних метрик, таких як SSIM (Structural Similarity Index), або оцінки експертів.

Кожен з кроків відіграє важливу роль у процесі створення фотографій за допомогою нейромережі, і їх взаємодія допомагає досягти високоякісних результатів.

1.2.1 Процес створення фото нейромережею завдяки текстового запиту

Процес запиту на створення фотографії зазвичай включає наступні кроки:

Вибір архітектури і моделі. Користувач вибирає певну архітектуру нейромережі та модель для генерації фотографій. Це може бути попередньо навчена модель або нова модель, яку користувач хоче навчити на власних даних.

Підготовка вхідних даних. Зазвичай користувач подає вхідні дані, які можуть бути випадковим вектором або початковим зображенням, яке слід покращити або змінити. Це може бути частиною інтерфейсу користувача або вхідним параметром API.

Виклик нейромережі. Користувач або програма викликає нейромережу з вхідними даними. Це може включати передачу вектора або зображення через API, яке пов'язане з навченою моделлю.

Обробка запиту. Нейромережа обробляє вхідні дані за допомогою своєї архітектури. Генеративна модель намагається створити вірогідно реалістичне зображення відповідно до вхідних даних.

Генерація фотографії. Модель генерує нове зображення на основі вхідних даних. Це може відбуватися за допомогою генеративного процесу, наприклад, за допомогою глибоких згенеративних моделей, де генератор створює нове зображення.

Вивід результату. Згенероване зображення повертається користувачеві або використовується відповідно до вказівок. Це може бути представленням новоствореного зображення або може включати додаткові кроки, такі як збереження на сервері або відображення у відповідному інтерфейсі.

Оцінка результату (опційно). Користувач може оцінити результат і визначити, чи задоволений він створеним зображенням. Також може бути застосована автоматична оцінка якості за допомогою метрик, таких як SSIM або інші.

Налаштування та повторення. Залежно від результату, користувач може внести корективи до вхідних параметрів або повторити процес з іншими параметрами для отримання кращого результату.

Даний процес відображає основні кроки взаємодії користувача з нейромережею для створення фотографій. Нейромережа використовує свої внутрішні ваги і знання, набуте під час навчання, для генерації нового та оригінального контенту відповідно до вхідних даних.

1.2.2 Огляд архітектур GANs та VAEs

Generative Adversarial Networks (GANs) — це інноваційний підхід у галузі штучного інтелекту та глибокого навчання, який використовує два нейронні мережі, генератор і дискримінатор, для створення нових, надзвичайно реалістичних зображень.

Variational Autoencoders (VAEs) - це клас генеративних моделей, які поєднують у собі елементи автокодерів та ймовірнісних методів для генерації нових даних.

Використання. Широкий спектр завдань: GANs(рис. 1.1) використовуються для генерації реалістичних облич, створення арт-зображень, анімацій, а також у різних візуальних та креативних задачах.

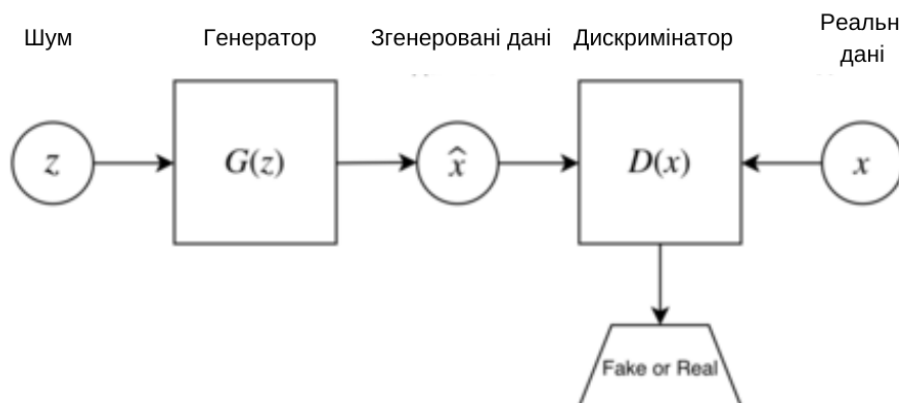


Рисунок 1.1 - Архітектура GANs [1]

Генератор. Генератор відповідає за створення нових зображень. Він приймає на вхід випадковий шум або вектор і трансформує його в зображення. Задача генератора полягає в тому, щоб навчитися створювати зображення, які надзвичайно схожі на ті, які можна знайти в навчальному наборі.

Дискримінація. Дискримінація служить як "суддя", який визначає, наскільки справжнім є дане зображення. Він отримує на вхід як зразок справжніх даних, так і зображення, створене генератором, намагається відрізнити між цими двома типами зображень, присвоюючи їм ймовірність бути справжніми або згенерованими.

Процес навчання. Генератор і дискримінація навчаються шляхом конкуренції між собою. Генератор намагається створити такі реалістичні зображення, щоб дискримінація був неспроможний їх відрізнити від справжніх.

Тренування. Процес тренування триває, поки генератор не досягне рівня, коли дискримінація не може надмірно відрізнити справжні дані від згенерованих. Навчання на реальних зразках: Дискримінація отримує навчання на реальних даних, тим самим покращуючи свою спроможність визначати справжність зображень.

GANs відіграють ключову роль у сучасному глибокому навчанні, надаючи ефективний підхід до створення вражаючих та реалістичних штучно згенерованих зображень.

Основна ідея VAEs полягає в тому, щоб забезпечити не тільки генерацію, а й контроль над простором у високорозмірному просторі. Давайте розглянемо подрібний аналіз архітектури VAEs.

Кодувальна Архітектура (Encoder):

Автокодер використовує вхідні дані і зменшує їх розмірність до латентного простору.

Закодоване Представлення (Mean and Variance). Закодований шар представляє латентні параметри, такі як середнє значення (mean) та дисперсія (variance), що визначають нормальний розподіл в цьому просторі.

Спроба Семплування (Sampling). Використовуючи отримані параметри, здійснюється випадкове семплування з нормального розподілу для отримання точок у латентному просторі.

Декодувальна Архітектура (Decoder). Відновлення Оригінального Зображення. Отримані латентні представлення використовуються для відновлення оригінального зображення вихідного простору.

Збереження Інформації. Архітектура старається зберігати якомога більше інформації під час декодування.

Функція Втрат (Loss Function). Втрата Повернення (Reconstruction Loss): Це міра того, наскільки точно вихід моделі відтворює оригінальний вхід.

Дивергенція Кульбака-Лейблера. Враховується для вимірювання відстані між отриманим латентним розподілом і стандартним нормальним розподілом. Це дозволяє створити просторово ефективніший та структурований латентний простір.

Регуляризація та Гіперпараметри. Коштовна Регуляризація (KL Divergence Regularization). Застосовується для забезпечення розумного розподілу в латентному просторі.

Гіперпараметри. Існує необхідність налаштування гіперпараметрів, таких як коефіцієнт регуляризації та інші, для оптимальної продуктивності.

Простір Зображень. Латентний простір VAEs дозволяє виражати різні варіанти вхідних зображень

1.3 Створення сценарію штучним інтелектом

Збір та розуміння даних. Штучний інтелект розпочинає свій процес, збираючи великий обсяг даних, пов'язаних із сценаріями фото та відео. Це можуть бути тексти, сценарії, відгуки глядачів, технічні аспекти кіно та інші релевантні джерела.

Відбір теми та параметрів. Користувач вводить параметри для сценарію, такі як тема, жанр, настрій, обрані персонажі та ключові елементи. Інтелект аналізує параметри, визначаючи основні напрямки для генерації.

Використання глибокого навчання для семантичного розуміння. Інтелект використовує глибокі нейромережі для семантичного аналізу текстових та візуальних даних, отриманих на попередньому етапі. Це допомагає розуміти сутність ключових слів, а також визначати специфіку персонажів та подій.

Генерація структурного каркасу. На цьому етапі, штучний інтелект створює структурний каркас сценарію, визначаючи початок, розвиток та кульмінацію. Це може включати ключові сцени, діалоги, зміни обстановки тощо.

Генерація текстових описів та діалогів. Інтелект використовує генеративні моделі для написання текстових описів сцен, діалогів персонажів та розвитку сюжету. Враховуються вимоги користувача, а також стилістика та лінгвістичні особливості.

Візуалізація та сценарій для відео. Інтелект може використовувати технології візуалізації для конвертації текстового сценарію відразу в фото- та відеоматеріали. Це включає відтворення сцен, анімацію персонажів, та врахування камерних кутів.

Автоматична корекція та вдосконалення. Інтелект використовує механізми самокорекції на основі введених відгуків користувача. Він може автоматично адаптувати сценарій відповідно до побажань та невеликих корекцій.

Вивід готового сценарію. Штучний інтелект представляє готовий сценарій у зручному форматі для користувача. Це може бути текстовий документ, що включає діалоги, описи та рекомендації для режисера, чи ж фінальний відеоматеріал для редагування.

Даний процес об'єднує аналіз даних, творчий підхід та використання глибокого навчання для створення вражаючих та індивідуальних сценаріїв фото та відео.

1.3.1 Використання Нейромереж та Алгоритмів Глибокого Навчання.

Прогнозування реакції глядачів. Глибокі нейромережі використовуються для аналізу реакцій глядачів на різні аспекти сценарію. Моделі отримують вхідні дані з попередніх переглядів, враховуючи вирази обличчя, емоційні реакції та реакції на ключові події. Це дозволяє системі прогнозувати, які елементи сценарію можуть викликати найсильніші враження та взаємодії.

Оптимізація сценарію. Алгоритми глибокого навчання використовуються для оптимізації сценарію, враховуючи зібрані дані про реакції глядачів. Моделі автоматично адаптують структуру сценарію, підлаштовуючи його під преференції та емоційні реакції аудиторії. Це сприяє створенню більш інтригуючого та захопливого контенту.

Персоналізація взаємодії. Глибокі нейромережі використовуються для створення персоналізованих елементів сценарію, враховуючи індивідуальні вподобання та історію взаємодії кожного глядача. Данна індивідуалізація дозволяє кожному глядачеві відчувати, що контент створений саме для нього, підвищуючи залученість.

Аналіз динаміки сприйняття. Алгоритми глибокого навчання використовуються для аналізу та розуміння динаміки сприйняття глядачів на різних етапах сценарію. Це охоплює виявлення ключових моментів, які привертають особливу увагу аудиторії, а також реакції на емоційно насичені епізоди. Аналіз даних дозволяє вдосконалювати подальший розвиток сценарію.

Даний підхід до використання нейромереж та алгоритмів глибокого навчання не лише розширює можливості генерації сценаріїв, але і забезпечує інноваційність та адаптивність контенту, що відповідає унікальним потребам глядачів.,

Висновки по розділу: були висвітлені ключові аспекти застосування штучного інтелекту у генерації мультимедійного контенту. Генерація відео нейромережею, розгляд процесу створення фото за допомогою нейромереж та вивчення можливостей створення сценарію штучним інтелектом сприяли глибшому розумінню інноваційних можливостей цих технологій. Застосування нейромереж для творчого процесу в мультимедіа відкриває широкі перспективи для розвитку та вдосконалення якісного контенту. Отже, даний розділ став важливим кроком у дослідженні та розумінні новаторських підходів до створення мультимедійного контенту за допомогою штучного інтелекту та нейромереж.

2 АНАЛІЗ ТА ПОРІВНЯННЯ СЕРВІСІВ, ЩО ВИКОРИСТОВУЮТЬ НЕЙРОМЕРЕЖІ ДЛЯ СТВОРЕННЯ ФОТО- ТА ВІДЕОКОНТЕНТУ

Даний розділ присвячений науковому дослідженню та вивченню ролі та потенціалу, який нейромережі мають в контексті створення фотографій та відеозаписів. Нейромережі для обробки зображень і відео відкривають широкі перспективи в галузі творчості, виробництва та наукового дослідження.

Вивчення використання нейромереж у створенні фотографій та відео визначається не лише технічними аспектами їхньої реалізації, але і важливим внеском у розвиток образотворчого мистецтва, аналіз візуальної інформації та оптимізацію технологічних процесів. В даному дослідженні намічено профундний аналіз та оцінку ефективності нейромереж, які застосовуються у відтворенні та обробці зображень і відео, розглядаючи їхні можливості, переваги та потенційні обмеження.

Дана тема в контексті наукового дослідження дозволить зрозуміти не тільки сучасні досягнення в галузі, але і визначити перспективи подальшого розвитку, створюючи фундамент для надалі вдосконаленого використання нейромереж у створенні фото та відео.

2.1 Сервіси та програмні рішення на основі штучного інтелекту для створення фото

В сучасному інформаційному віці розквіт технологій нейромереж відкриває широкі можливості в галузі комп'ютерного зору та обробки зображень. Даний розділ присвячений вивченню та аналізу використання нейромереж у контексті створення фотографій, що є актуальною темою для вивчення та дослідження. Нейромережі, здатні до глибокого навчання, виявляються потужним інструментом для генерації, реставрації та покращення фотографій, а також для творчого перетворення та стилізації зображень.

Даний відділ розглядає різноманітні аспекти використання нейромереж у сфері фотографії, включаючи технічні аспекти їхньої реалізації, методи навчання, архітектури та результати їхньої застосовності. За допомогою наукового підходу та аналітики, ми докладемо зусиль для вивчення і розуміння впливу нейромереж на творення та обробку фото, розглядаючи їхню ефективність, можливості та перспективи подальшого розвитку в даній галузі.

2.1.1 Сервіс Midjourney

Midjourney представляє інноваційну платформу, спроектовану на основі передових технологій глибокого навчання для синтезу візуального контенту. Система забезпечує користувачам можливість взаємодії з нейромережею через сервер Discord, використовуючи команди для формулювання запитів на генерацію зображень. Технологічна основа системи покладена на комплексну архітектуру глибокого навчання, що дозволяє досягти вражаючої якості синтезованих візуальних робіт.

Критичним етапом взаємодії з платформою є ознайомлення з внутрішніми правилами спільноти, яке передбачає виконання користувачами заданих процедурних вимог. Після даного вони отримують доступ до каналу #newbies, де можуть викликати нейромережу, використовуючи команду /imagine та висловлюючи свій запит англійською мовою.

Система надає користувачам чотири варіанти зображень відповідно до запиту, кожен із яких може бути підданий подальшій обробці для збільшення роздільної здатності або внесення інших аспектів. Це визначається адаптивною природою нейромережі, яка включає в себе техніки глибокого навчання для стилізації та деталізації зображень в різних контекстах.

Нейромережа Midjourney продемонструвала вражаючі здатності створення різноманітних серій зображень, охоплюючи широкий спектр стилів та рівнів деталізації. Це досягається завдяки глибокому навчанню, що сприяє абстрагуванню стилізованого представлення та обробці зображень.

Окрім індивідуальних графічних творів, система також володіє функцією об'єднання декількох зображень у складні композиції. Данна можливість розширює креативні горизонти користувачів, дозволяючи їм створювати унікальні візуальні аранжування. На (рис. 2.1) показан Diskord сервер нейромережі .

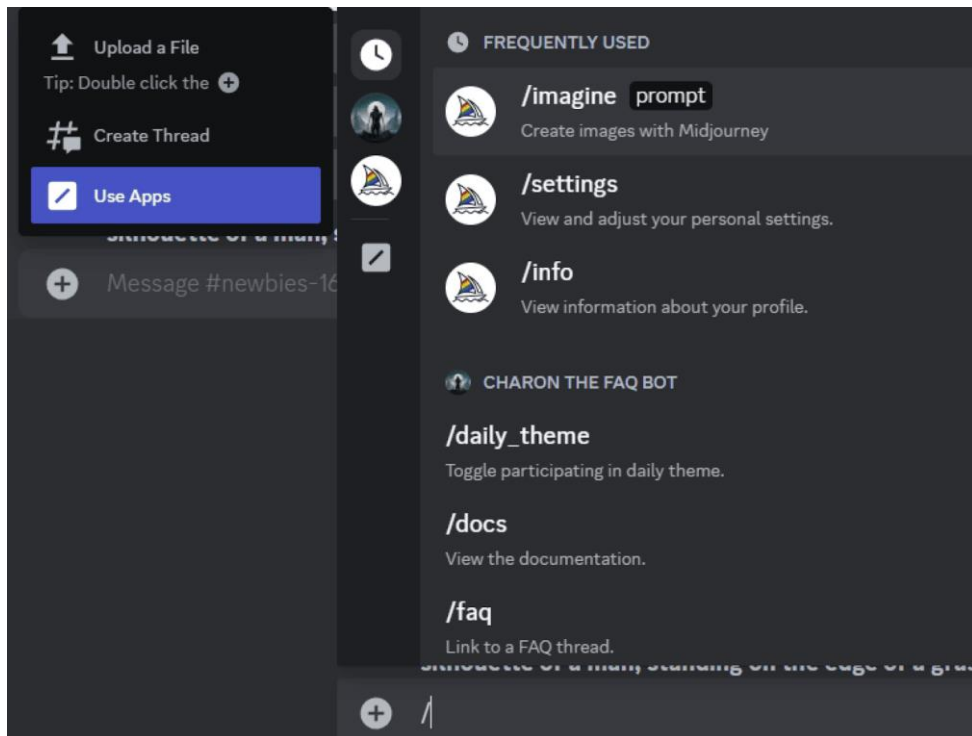


Рисунок 2.1 - Diskord сервер нейромережі [2]

Позитивні аспекти :

- Можливість налаштувати алгоритм генерації;
- Генерація нового зображення на основі вже наявного;
- Генерація 4-х варіантів на вибір;
- Швидкість генерації;
- Збільшення і варіації;
- Якість генерації;
- Відсутність обмежень деталізації prompt.

Негативні аспекти :

- Ігнорування деталізації (наприклад, колір або тип одягу);
- Труднощі з відтворенням потрібних емоцій, виразу обличчя або дії;

- Відсутність можливості змінювати певний параметр;
- Витрати часу;
- Обмеженість функціоналу для редагування;
- Відсутність бета-періоду;
- Непрогнозований результат.

2.1.2 Сервіс Blue Willow

Blue Willow – це нейромережа, спеціалізована на генерації зображень за текстовим описом (prompt). Вона відкриває можливість творення унікальних зображень в різних жанрах та стилях.

Основні особливості цієї нейросеті варто відзначити:

Навчання на обширних обсягах даних: Blue Willow пройшла тренування на великій кількості вхідних даних, що сприяло її високій продуктивності та здатності до творчого синтезу.

Інтерфейс через Discord: Нейросеть взаємодіє з користувачами через платформу Discord, надаючи зручний інтерфейс для створення та отримання зображень, що робить її доступною для використання безпосередньо в чаті . Сервіс (рис. 2.2) зображен нижче .

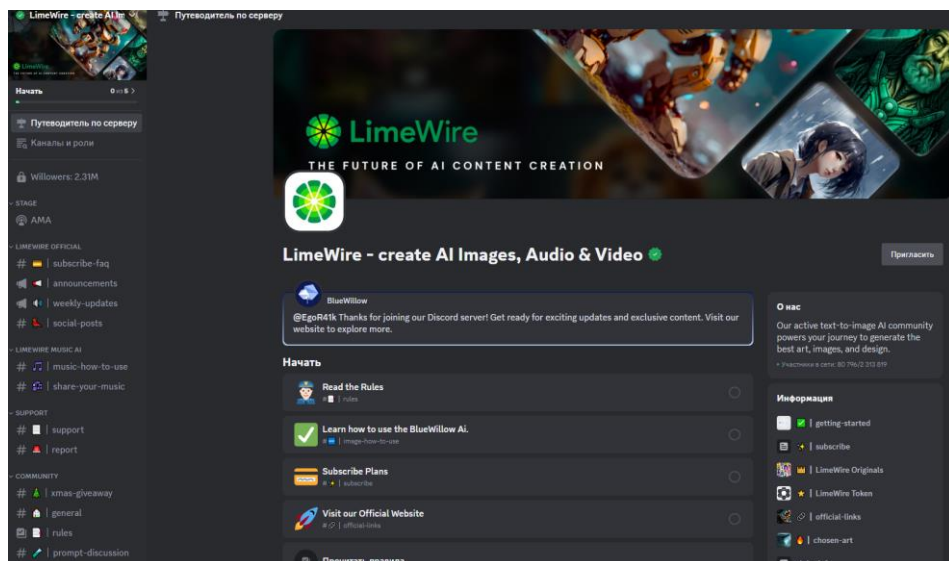


Рисунок 2.2 - Сервіс нейромережі Blue Willow

Відсутність власної генеративної моделі: Blue Willow виступає як агрегатор інших використовуваних моделей, дозволяючи отримувати максимально різноманітні результати.

Широкі налаштування для роботи з готовими варіантами: Нейросеть надає користувачам значний контроль над процесом генерації, включаючи різноманітні налаштування для роботи з різними типами зображень, такими як портрети, пейзажі, абстракції та інші.

Незважаючи на те, що Blue Willow наразі офіційно не вийшла з бета-тестування, вже зараз вона завоювала довіру серед митців, ілюстраторів, дизайнерів, творців логотипів та представників інших творчих професій.

Позитивні аспекти :

- Відмінна деталізація;
- Можливість вибору формату;
- Безлімітне та безкоштовне використання;
- Одна з найкращих безкоштовних нейромереж для генерації зображень.

Негативні аспекти :

- Іноді потрібно надсилати запити довжиною 3-4 рядки, щоб нейромережа правильно їх розуміла;
- Проблема з генерацією рук, яка практично вирішена в п'ятій версії Midjourney;
- Загальна якість зображень відстає від Midjourney.

2.1.3 Онлайн сервіс Stable Diffusion Online

Stable Diffusion Online - це швидкий та точний інструмент обробки зображень на основі штучного інтелекту, який дозволяє швидко обробляти великі обсяги зображень. Він пропонує ряд функцій, що автоматизують та оптимізують процес обробки зображень. Робочий інтерфейс Stable Diffusion Online(рис. 2.3) наведено нижче

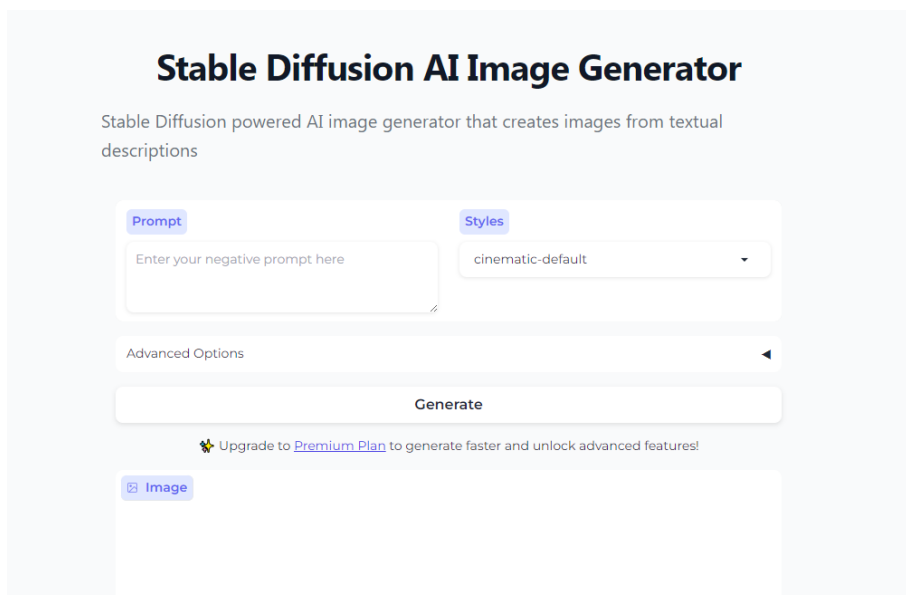


Рисунок 2.3 – Робочий інтерфейс сервісу

Основні можливості Stable Diffusion Online

Автоматичне обрізання зображень: Stable Diffusion Online може автоматично обрізати зображення, щоб залишити лише необхідну частину.

Автоматична корекція кольору: За допомогою Stable Diffusion Online можна автоматично коригувати кольори зображень, щоб вони виглядали якнайбільш натурально.

Видалення шуму та підвищення різкості: Stable Diffusion Online може видаляти шуми та підвищувати різкість зображень, забезпечуючи більш чітке та деталізоване зображення.

Обробка кількох зображень одночасно: Stable Diffusion Online дозволяє обробляти кілька зображень одночасно, що значно скорочує час обробки.

Stable Diffusion Online - це швидкий та точний інструмент обробки зображень на основі штучного інтелекту, який забезпечує автоматизацію та оптимізацію виробничого процесу. Він підходить для компаній і організацій, які працюють з великою кількістю зображень, а також для фотографів і дизайнерів, які бажають прискорити процес обробки своїх зображень. Stable Diffusion Online дозволяє швидко та точно обробляти великі обсяги зображень, зменшуючи час обробки та підвищуючи ефективність виробництва.

Позитивні аспекти :

- Швидкий та точний процес обробки: Stable Diffusion Online швидко та точно обробляє великі обсяги зображень, скорочуючи час обробки та підвищуючи ефективність виробництва;
- Автоматизація процесу обробки: Stable Diffusion Online автоматизує процес обробки зображень, що дозволяє зменшити кількість помилок та підвищити якість обробки;
- Доступність: Stable Diffusion Online доступний в Інтернеті, що дозволяє використовувати його в будь-якому місці та в будь-який час;
- Оптимізація виробничих процесів: Stable Diffusion Online дозволяє оптимізувати виробничі процеси, зменшуючи витрати на обробку зображень.

Негативні аспекти :

- Залежність від Інтернет-підключення: Використання Stable Diffusion Online вимагає постійного доступу до Інтернету, що може бути обмежуючим у випадках низької швидкості підключення або відсутності зв'язку;
- Обмежені можливості безкоштовної версії: Незважаючи на те, що є безкоштовна версія, деякі розширені функції та можливості можуть бути доступні лише у платній версії, що може обмежити функціональність користувачів;
- Відсутність локального контролю: Оскільки сервіс базується на хмарних технологіях, користувачі не мають повного контролю над обробкою даних, що може викликати певні обмеження в забезпеченні конфіденційності;
- Необхідність аплікації сторонніх засобів захисту даних: З міркувань безпеки, користувачам може знадобитися використовувати додаткові засоби для захисту своїх даних, оскільки даний сервіс може не мати вбудованих механізмів захисту від певних загроз;

- Потенційні проблеми з конфіденційністю даних: При передачі зображень для обробки онлайн існує ризик збереження або використання даних сторонніми, що може порушити конфіденційність користувача.

2.2 Сервіси та програмні рішення на основі штучного інтелекту для створення відео

В еру стрімкої технологічної революції нейромережі виступають ключовим елементом у сфері створення відеоконтенту. Даний розділ має на меті здійснити огляд використання нейромереж у процесі створення відео, проаналізувати їхній вплив та визначити потенційні перспективи в динамічній області.

Науковий аналіз ролі нейромереж у створенні відео охопить технічні особливості їхнього застосування, переваги та обмеження. Дослідження охопить також взаємодію нейромереж із сучасними відомостями щодо відеопродукції, а також їхній внесок у розвиток різних галузей, включаючи розваги, науку та бізнес.

За допомогою систематичного огляду ми розкриємо ключові аспекти використання нейромереж для створення відео, враховуючи інноваційні підходи та тенденції, що визначають майбутнє цієї захоплюючої області.

2.2.1 Онлайн сервіс Runway

В нейромережі Runwayml реалізовані два алгоритми для генерації контенту: Gen-1 і Gen-2. Перший алгоритм в основному працює з зображеннями, тоді як другий може обробляти відео. Gen-1 надає можливість створювати анімації лише на основі двох зображень: початкового та кінцевого. Алгоритм Gen-2 ще знаходиться на етапі розробки і працює у тестовому режимі, але за його допомогою можна генерувати відео з нуля або на основі готового матеріалу. Інтерфейс Runway(рис. 2.4) зображен далі.

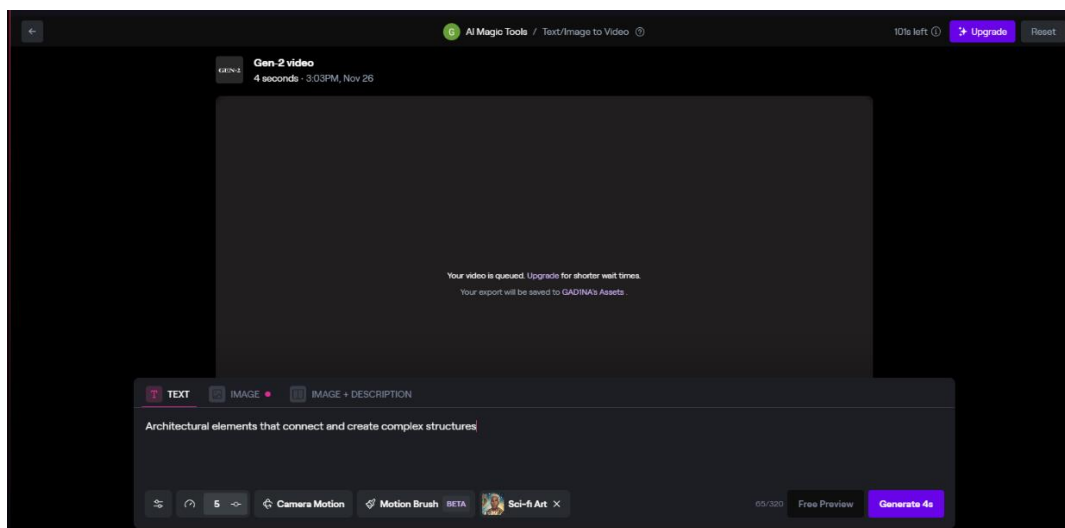


Рисунок 2.4 - Інтерфейс Runway

У рамках нейромережі Runwayml розроблено два алгоритми для генерації відеоконтенту: Gen-1 та Gen-2, кожен з яких має свої унікальні особливості та обмеження. Gen-1, який переважно взаємодіє зі зображеннями, надає можливість створювати анімації, базуючись на двох зображеннях - початковому та кінцевому. З іншого боку, алгоритм Gen-2 перебуває на етапі розробки і функціонує у тестовому режимі, проте його потенціал полягає в здатності генерувати відео як з нуля, так і на основі наявного матеріалу.

Основні функціональності, які пропонує штучний інтелект, включають синтез за текстовим описом, генерацію відео на основі тексту та зображення, створення відео на основі одного зображення, стилізацію готового кліпу за допомогою завантаженого зображення, накладання маски на готове відео за текстовою підказкою, рендерінг відео без текстур за зображенням або текстом, а також створення анімації на основі двох вхідних зображень.

Важливо відзначити, що алгоритм Gen-2, як ще нестабільний етап розробки, може виявити обмеження та непередбачені труднощі. Зазначені функціональності відкривають широкий спектр можливостей для творчого використання, проте враховуючи невизначені межі алгоритмів та обмеження пробної версії, користувачеві слід ретельно планувати та використовувати дані інструменти.

Runwayml, як інструмент для генерації відеоконтенту за допомогою нейромереж, має свої переваги та недоліки, які варто розглянути.

Позитивні аспекти :

- Інноваційні Алгоритми. Runwayml включає інноваційні алгоритми, такі як Gen-1 та Gen-2, що дозволяє користувачам створювати відеоконтент з високим рівнем творчості.
- Широкий Функціонал Штучного Інтелекту. Інструмент надає широкий набір можливостей, таких як синтез за текстовим описом, генерація відео на основі тексту та зображень, стилізація готових кліпів, накладання масок та інші, роблячи його високофункціональним.
- Творчий Процес. Runwayml надає користувачам можливість творити відеоконтент, використовуючи інтуїтивний та креативний підхід, що робить його привабливим для художників та дизайнерів.

Негативні аспекти :

- Стадія Розробки Gen-2: Наявність алгоритму Gen-2 на стадії розробки може призводити до нестабільності та обмежених можливостей у порівнянні з Gen-1.
- Обмеженість Пробної Версії. Пробна версія обмежена кількістю кредитів та має обмежені можливості експорту, що може обмежити користувачів у тестуванні та використанні.
- Невизначені Межі Алгоритмів. Існуючі невизначені межі щодо точності та здатності розпізнавання об'єктів, особливо для Gen-2, який перебуває на етапі розробки.

2.2.2 Онлайн сервіс Pictory

Платформа Pictory представляє собою інноваційний інструмент, який використовують автори YouTube-каналів, маркетологи, менеджери соціальних мереж, блогери та творці освітніх курсів. Здійснення генерації готових відеокліпів з підписами стає можливим завдяки широкій базі медіаматеріалів, яка становить основу цієї платформи. Переважаючи в сфері створення відеоконтенту, система Pictory також ефективно працює з зображеннями та анімаціями. Інтерфейс Pictory (рис. 2.5) наведено нижче.

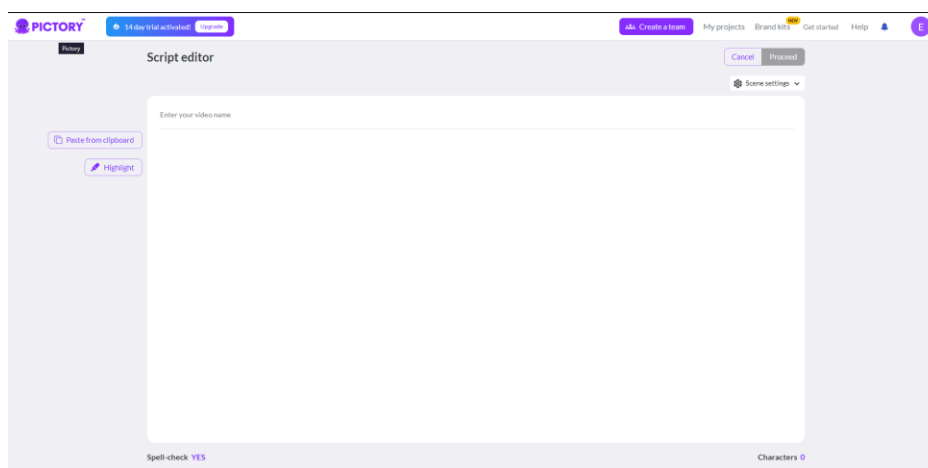


Рисунок 2.5 – Інтерфейс онлайн сервісу Pictory

Створення відеоролика за готовим текстовим сценарієм. Користувачі можуть створювати відеоконтент, використовуючи заздалегідь розроблений текстовий сценарій, що спрощує процес та забезпечує стабільність створеного контенту.

Генерація відеоконтенту на основі посилання на статтю чи блог. Pictory дозволяє автоматично генерувати відеоматеріал на основі зазначених посилань, роблячи процес більш ефективним та зручним.

Автоматичне додавання субтитрів. Система автоматично впроваджує субтитри до створеного відеоконтенту, що поліпшує доступність та розуміння для аудиторії.

Можливість обрізати частину кліпу. Користувачі можуть здійснювати редагування та обрізання відеоконтенту згідно з власними вимогами.

Додавання логотипу, інтро чи аудіо. Платформа надає можливість вставляти логотип, введення або аудіо, що робить контент більш індивідуалізованим та брендним.

Додавання візуальних ефектів на основі готових матеріалів. Pictory пропонує користувачам можливість використовувати візуальні ефекти, базуючись на наявних матеріалах, що сприяє творчості та різнообразності створюваного вмісту.

Узагальнюючи, платформа Pictory визначається великою функціональністю та можливостями для творчого створення відеоконтенту, спрощуючи процес для широкого спектру користувачів.

Позитивні аспекти :

- Інноваційність та Творчий Підхід. Pictory відзначається використанням інноваційних алгоритмів, що дозволяє користувачам застосовувати творчий підхід при створенні відеоконтенту. Інтеграція готових матеріалів робить процес більш гнучким та доступним для різних груп користувачів, включаючи авторів YouTube, маркетологів, менеджерів соціальних мереж, блогерів і творців освітніх курсів;
- Розширені Можливості Створення Вмісту. Система дозволяє створювати відеоконтент за готовим текстовим сценарієм, використовувати статті або блоги як основу для генерації відеоконтенту, а також автоматично додавати субтитри. Це забезпечує велику різноманітність та гнучкість при створенні відеоматеріалів;
- Редагування та Кастомізація. Користувачі отримують можливість вирізати частину відеокліпу, додавати власний логотип, інтро чи аудіо, що робить контент більш індивідуалізованим та відповідним їхнім потребам;
- Візуальні Ефекти та Креативність. Платформа дозволяє додавати візуальні ефекти на основі готових матеріалів, що сприяє творчості та різнообразності створюваного вмісту.

Негативні аспекти :

- Стадія Розробки та Нестабільність Gen-2. Наявність алгоритму Gen-2 на стадії розробки може призводити до нестабільності та обмежених можливостей у порівнянні з Gen-1, що потребує уважного врахування при використанні;
- Невизначені Межі Алгоритмів. Існуючі невизначені межі щодо точності та здатності розпізнавання об'єктів можуть виникнути через стадію розробки та тестування алгоритмів.

2.3 Нейромережі для створення сценарію

Створення сценаріїв визначається великою мірою творчістю та уявою, але сучасні досягнення в галузі нейромереж відкривають нові можливості для автоматизації даного процесу. У цьому розділі дипломного проекту розглянутья приклади використання нейромереж для створення сценаріїв, детально висвітлюючи їхні можливості та внесок у творчий процес сценарного письма.

Приклади Використання Нейромереж для Створення Сценарію:

Синтез Тексту за Заданим Сценарієм. Нейромережі можуть використовуватися для автоматизованого створення тексту, відповідного заданому сценарію. Моделі можуть генерувати діалоги, описи сцен та важливі моменти сюжету, спрощуючи процес написання сценарію.

Адаптація до Жанру та Тону. Нейромережі можуть навчатися різним жанрам та тонам, адаптуючи стилістику сценарію відповідно до вибраного жанру. Це може бути особливо корисним для авторів, які експериментують з різноманітним жанровим підходом.

Створення Життєвих Ситуацій. Нейромережі можуть генерувати реалістичні життєві ситуації та діалоги, сприяючи реалізму та емоційності сценарію.

Автоматична Перевірка Консистентності. Моделі можуть використовуватися для автоматичної перевірки консистентності сценарію, допомагаючи уникнути логічних або характерних суперечностей у тексті.

Можливості та Виклики Використання Нейромереж у Сценарному Письмі. Збереження Авторського Стилю. Важливо враховувати, як нейромережі можуть адаптуватися до стилю автора та як зберегти унікальний креативний внесок письменника.

Ліміти Генерації. Потрібно уважно враховувати обмеження генерації нейромереж, зокрема в уникненні занадто стандартизованих або передбачуваних елементів в сценарії.

Потенціал для Інтерактивності. Розгляд можливостей використання нейромереж для створення інтерактивних елементів сценарію, що може взаємодіяти з читачем або глядачем.

Роль Автора у Контролі Процесу. Автор повинен зберігати активну роль у контролі над процесом генерації, визначати параметри та вносити креативні корекції за необхідності.

Даний розділ буде присвячений детальному розгляду аспектів, надаючи аналіз прикладів та висновків щодо ефективного використання нейромереж у сценарному письмі.

2.3.1 Онлайн ресурс GPT

GPT, або Generative Pre-trained Transformer, здатний трансформувати процес генерації сценаріїв завдяки його вражаючій креативності та високому рівню розуміння семантики. Даний інтелектуальний інструмент розроблений для створення текстів, що мають природну мову, і виокремлюється своєю здатністю адаптуватися до різних стилів та жанрів. Зокрема, версії GPT-3 і GPT-4 представляють сучасні досягнення у сфері генерації тексту та творчого написання. Інтерфейс платформи(рис. 2.6) зображен нижче .

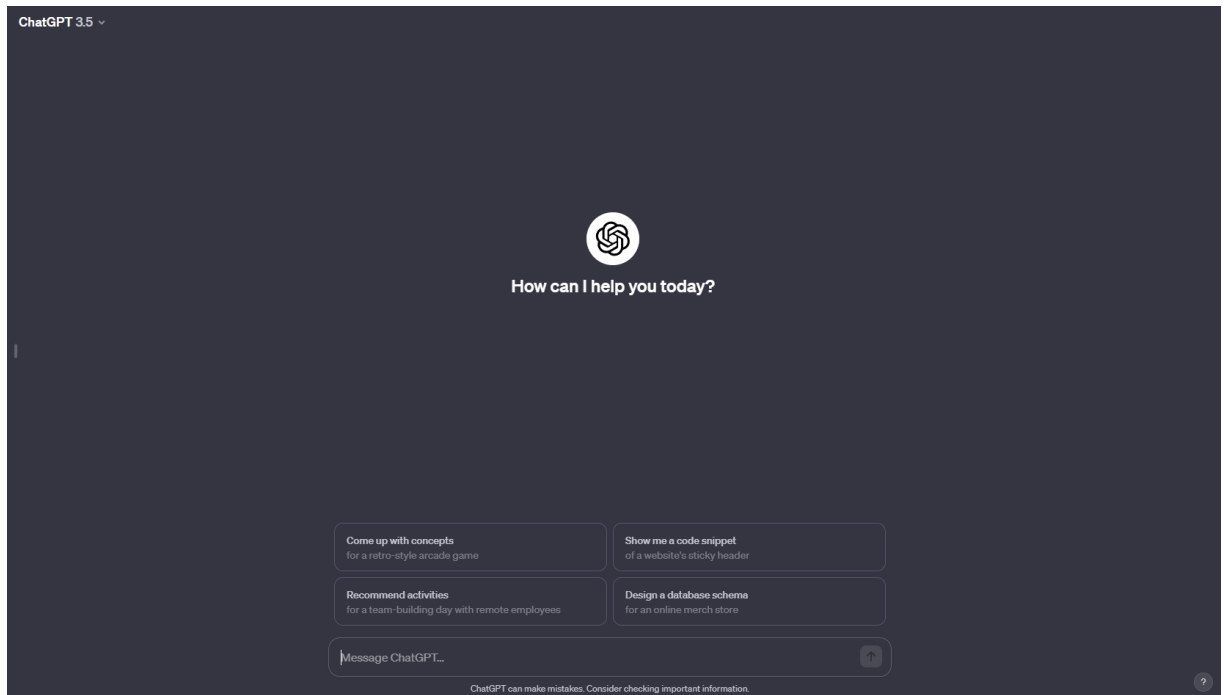


Рисунок 2.6 - Інтерфейс платформи GPT

Однією з головних переваг є високий рівень креативності, який може дозволити авторам генерувати оригінальний та захоплюючий контент. Велика масштабованість GPT-моделей, зокрема GPT-3 і GPT-4, відкриває можливості для обробки широкого контексту та створення тексти, що адаптуються до різних сценарійних вимог.

Позитивні аспекти використання:

- Креативність і Натуральність Мови. GPT вражає своєю здатністю генерувати тексти, які виглядають природним чином, зберігаючи при цьому високий рівень креативності;
- Адаптація до Різних Стилів і Жанрів. Можливість адаптуватися до різних стилів та жанрів робить GPT вельми універсальним для вирішення різноманітних сценарійних завдань;
- Велика Масштабованість. GPT-моделі, особливо GPT-3 і GPT-4, мають велику кількість параметрів, що дозволяє їм ефективно обробляти широкий контекст;

- Гнучкість у Використанні. GPT може бути застосований для різних сценаріїв, включаючи написання сценаріїв для фільмів, телевізійних шоу, відеоігор та багатьох інших.

Негативні аспекти використання:

- Можливість Слідування Стандартам. У деяких випадках GPT може генерувати тексти, які слідують суворо стандартам, що може виглядати менш креативним;
- Необхідність Уважного Контролю. Автор повинен уважно контролювати процес генерації, оскільки GPT може створювати непередбачувані вирази, які можуть не відповідати інтенціям;
- Потреба у Додатковій Обробці. Тексти, згенеровані GPT, можуть вимагати додаткової обробки та адаптації для конкретного використання;
- Масштабованість та Ресурси. Велика масштабованість GPT може вимагати значних обчислювальних ресурсів, особливо при великому обсязі генерації текстів;
- Інтеграція GPT у генерацію сценаріїв може бути дуже результативною, якщо враховувати його переваги та обмеження. З вірним контролем та творчим підходом, даний інструмент може допомогти авторам та сценаристам створювати вражаючі та захоплюючі тексти для різних медійних форматів.

2.3.2 Сервіс SoCreate

Платформа, представлена в даному контексті, визначається як інструмент, спрямований на упрощення процесу написання сценаріїв для відеороликів, призначених для соціальних мереж, таких як Reels, Shorts або TikTok. Досягнення цієї мети здійснюється через використання нейромереж та інших технологічних засобів.

У першу чергу, користувачеві пропонується обрати формат відео, для якого необхідно розробити сценарій. Після даного визначається назва локації, де розгортається подія, і сервіс автоматично створює сторіборд для початкової сцени. Даний підхід забезпечує визначення візуальної структури для подальшого викладення історії.

Платформа дозволяє генерувати персонажів для відповідної сцени, формулювати текст для них та описувати розвиток подій. Важливою можливістю є зміна локації, впровадження спеціальних ефектів чи редагування тексту з метою уточнення деталей. Крім того, користувач має можливість визначити параметри, такі як переходи між відеокамерами, вміст тексту на екрані та сюжетні переходи.

Зазначено, що платформа надає можливість працювати над проектом кільком особам, що підкреслює його потенціал для командної співпраці. Зазначається, що платформа пропонує безкоштовний 3-місячний пробний період, після чого використання її функціоналу доступне за умови обрання одного з платних тарифів. Наводиться інформація, що інтерфейс платформи реалізований англomовним, але за допомогою піктограм користувач легко розбереться в його функціоналі.

Важливо наголосити, що, не зважаючи на істотну полегшеність процесу написання сценаріїв за допомогою нейромереж, аналогічні сервіси не можуть повністю замінити творчість та інтуїцію людини. Тому рекомендується розглядати їх як додаткові інструменти, а не як повноцінні заміники, проводячи аналіз отриманих результатів та адаптуючи їх під особливості цільової аудиторії відеороликів. Інтерфейс платформи SoCreate(рис. 2.7) зображен нижче



Рисунок 2.7- Інтерфейс платформи SoCreate

Позитивні аспекти :

- Ефективність та Швидкість: Платформа значно прискорює процес написання сценаріїв, що дозволяє ефективно використовувати час та ресурси;
- Автоматизація: Автоматичне створення сторінки на основі обраних параметрів допомагає уникнути витрат часу на рутинні завдання;
- Колаборація: Можливість спільної роботи декількох користувачів над проектом сприяє творчій взаємодії та розвитку ідей;
- Бібліотека Зображень та Ілюстрацій: Інтегрована бібліотека сприяє візуалізації ідей та надає додаткові ресурси для написання сценаріїв.

Негативні аспекти :

- Залежність від Технологій: Робота з платформою вимагає доступу до Інтернету та використання технічних засобів, що може бути незручним у деяких ситуаціях;
- Необхідність Адаптації: Інтерфейс, хоч і зрозумілий за допомогою піктограм, все ж може вимагати адаптації для користувачів, що не володіють англійською мовою;
- Платні Тарифи: Після завершення пробного періоду виникає необхідність вибору платного тарифу, що може стати фінансовим обмеженням для деяких користувачів;

- Обмежена Творчість Нейромереж: Незважаючи на автоматизацію, нейромережі не можуть повністю відтворити творчий потенціал та інтуїцію людини;
- Неуніверсальність: Хоча платформа може бути ефективною для певних форматів відео (Reels, Shorts, TikTok), вона може не відповідати всім потребам та вимогам різних платформ соціальних мереж.

Висновки по розділу: вивчення та оцінка різноманітних сервісів та програмних рішень, що базуються на штучному інтелекті для генерації фото- та відеоматеріалів. Детальний розгляд сервісів для створення фотоконтенту та відеоконтенту розкрив значущі технологічні особливості, переваги та недоліки кожного із них.

Аналіз сервісів для генерації сценаріїв штучним інтелектом підкреслив їхню роль у творчому процесі та можливості для автоматизації та поліпшення створення контенту. Розгляд цих аспектів дозволив визначити оптимальні рішення для конкретних завдань та вирішення проблем, пов'язаних із генерацією фото- та відеоматеріалів.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖ У СТВОРЕННІ КОНТЕНТУ

У сучасному цифровому епохальному віці, характеризованому стрімким розвитком технологій, використання нейромереж в сфері творчості та медіа набуває надзвичайної важливості. Застосування штучних нейронних мереж у створенні фотографій, відеоконтенту та сценаріїв відкриває нові можливості для творців і змінює парадигму творчого процесу. Даний розділ присвячений розгляданню впливу нейромереж на процеси творення та їх застосування в сфері візуальних мистецтв та кінематографу.

Впровадження нейромереж у творчий процес відзначається не тільки технічною складністю, але й відкриттям нових можливостей для артистичного виразу та експериментації. Сполучаючи в собі потужність обчислювальних алгоритмів та архітектур нейромереж, творці отримують доступ до інструментів, що значно полегшують та збагачують творчий процес у створенні фото-, відеоконтенту та сценаріїв.

В даному розділі роботи розглядатиметься різноманітність застосувань нейромереж у вищезазначених галузях, а також вивчатиметься взаємодія між людським творчим потенціалом та штучним інтелектом. Поглиблене розуміння даного взаємодії може відкрити нові перспективи для технологічного зростання та творчого розвитку.

Зокрема, аналізуватимуться алгоритми глибокого навчання, що використовуються для генерації зображень та відео, розглядатиметься їхній вплив на реалізацію творчих задумів та емоційне сприйняття аудиторією. Велика увага буде приділена аналізу роботи нейромереж у контексті написання сценаріїв, виявляючи їхню спроможність допомагати та розширювати творчий потенціал сценаристів.

Дане дослідження також має на меті розкриття переваг та обмежень використання нейромереж у творчому процесі. Порівняльний аналіз результатів, отриманих за допомогою нейромереж, та традиційних методів може висвітлити ефективність та можливості використання новітніх технологій.

Загальна мета даного розділу полягає у висвітленні сучасного стану використання нейромереж у творчому процесі створення візуального контенту та сценаріїв. Досягнення цієї мети передбачає розкриття технічних аспектів роботи нейромереж, а також їхній вплив на естетичні та творчі аспекти створення візуальних творів.

У роботі використовуватимуться аналітичні, порівняльні та експериментальні методи дослідження. Вивчення аспектів сприятиме розширенню знань у галузі мистецтва, кінематографії та технологічного розвитку, сприяючи таким чином новітній еволюції в сфері творчості та технологій.

3.1 Створення сценарію завдяки ChatGPT

Для розробки сценарію використовувався сервіс ChatGPT, що є потужним інтелектуальним мовним моделювальником, розробленим OpenAI на основі архітектури GPT-3.5. Вибір даного сервісу обумовлений його здатністю вирішувати завдання штучного інтелекту пов'язані із зрозумінням та генерацією тексту відповідно до запитань користувачів.

Важливим етапом було проведення розуміння запиту, адже ChatGPT вміє аналізувати контекст і взаємодіяти з користувачем. Це було досягнуто завдяки високій якості обробки природної мови, що дозволило ефективно взаємодіяти з текстовим запитом та отримувати адекватні відповіді.

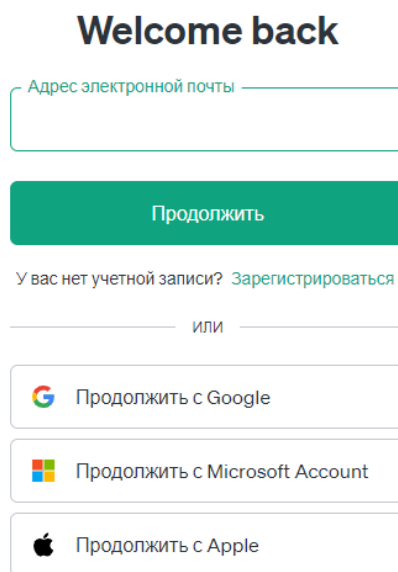
Під час створення сценарію велика увага була приділена процесу редагування, оскільки важливо було впевнитися в логічній послідовності і точності висловленого. Використання ChatGPT дозволило легко вносити зміни та коригувати текст, щоб досягти оптимального результату.

Мовний моделювальник також демонструє високу гнучкість у розумінні різних стилів інтеракції, що було особливо корисно при створенні сценарію. Даний аспект дозволив досягти бажаного тону та відчуття в тексті, щоб відповідати конкретним потребам аудиторії.

Завдяки здатності ChatGPT генерувати текстові відповіді на запитання, отримання змістовного та адекватного сценарію стало ефективним та продуктивним процесом. Використання даного сервісу виявилось важливим кроком у вдосконаленні та прискоренні процесу написання текстів, забезпечуючи високу якість і легкість використання.

Загалом, обрання сервісу ChatGPT для створення сценарію було обгрунтовано його здатністю до розуміння запитів, гнучкістю у взаємодії та високою якістю генерації тексту. Використання даного інструменту сприяло вдосконаленню та швидкості процесу написання сценаріїв.

Для початку роботи з сервісом ChatGPT потрібно авторизуватися(рис 3.1) на сайті , де є змога зареєструватися шляхом авторизації завдяки акаунтам в Google, Microsoft та Apple



The image shows a login interface for ChatGPT. At the top, it says "Welcome back". Below that is a text input field with the placeholder "Адрес электронной почты". Underneath the field is a green button labeled "Продолжить". Below the button, there is a link that says "У вас нет учетной записи? Зарегистрироваться". Below this is a horizontal line with the word "или" in the center. Underneath the line are three buttons for social login: "Продолжить с Google" (with the Google logo), "Продолжить с Microsoft Account" (with the Microsoft logo), and "Продолжить с Apple" (with the Apple logo).

Рисунок 3.1 – Поле авторизації сервісу ChatGPT

Після отримання запиту від користувача, ChatGPT, базуючись на своїй архітектурі GPT-3.5, розпочинає процес генерації текстової відповіді. Модель аналізує синтаксичну та семантичну структури тексту для коректного розуміння запитання.

Ключовим етапом є генерація тексту, де модель використовує свої знання та контекст для створення змістовної відповіді. Механізми глибокого навчання дозволяють зберігати логічну послідовність та контекстуальний зв'язок.

ChatGPT виявляє гнучкість у реагуванні на різні стилі взаємодії, а редагування тексту використовується для уточнення відповіді. В результаті отримується адекватна та адаптована до контексту відповідь, що відповідає стандартам штучного інтелекту. (рис. 3.2)



Рисунок 3.2 – Процес генерації відповіді ChatGPT

У редакційному процесі можна користуватися лапками(рис. 3.3), що дозволяє виділити конкретну частину тексту та використовувати їх для цитування та посилань на відповідний контекст.

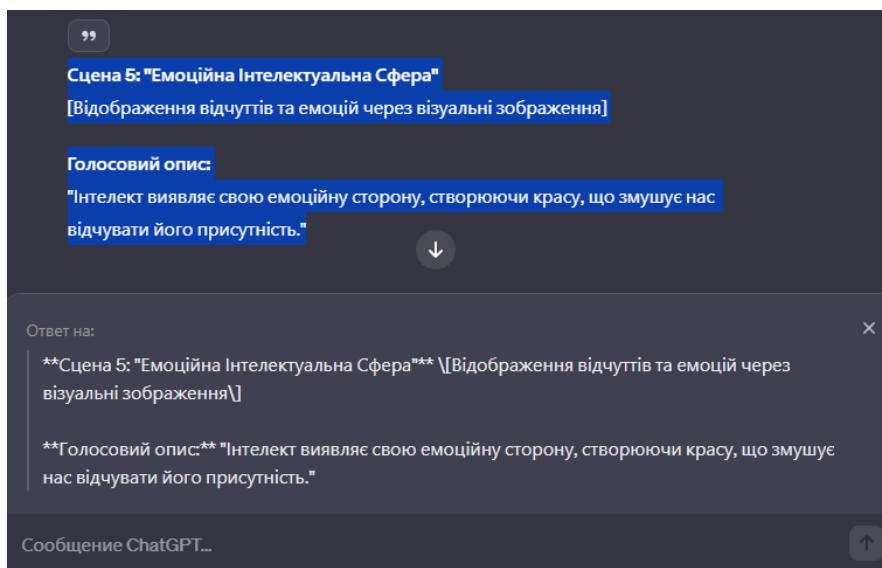


Рисунок 3.3. – Користування лапками

Дана функція відкриває можливість акцентування уваги на окремих елементах висловлювання та спрощує посилання на визначені частини тексту під час обговорення або аналізу.

3.2 Створення фото завдяки сервісу Stable Diffusion Online

Початок використання сервісу Stable Diffusion Online(рис 3.4) передбачає проведення ініціації процесу генерації зображень шляхом чіткої формалізації текстових запитів, відомих як промти. У контексті даного сервісу, промт слугує ключовим структурним елементом, на основі якого система виробляє візуальні представлення. Процес введення промтів визначає концептуальний каркас для подальшого генерування контенту.

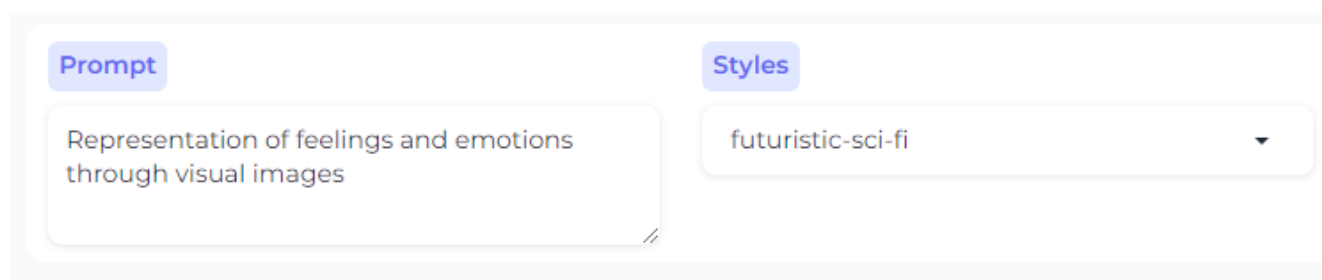


Рисунок 3.4 – Надання промту

Користувачеві також доступно вибір різних стилів(рис. 3.5), які надають можливість впливати на художні та стилістичні аспекти зображення. Дана можливість представляє собою значущий аспект, що дозволяє індивідуалізувати вихідні результати та адаптувати їх до конкретних вимог чи естетичних уподобань користувача.

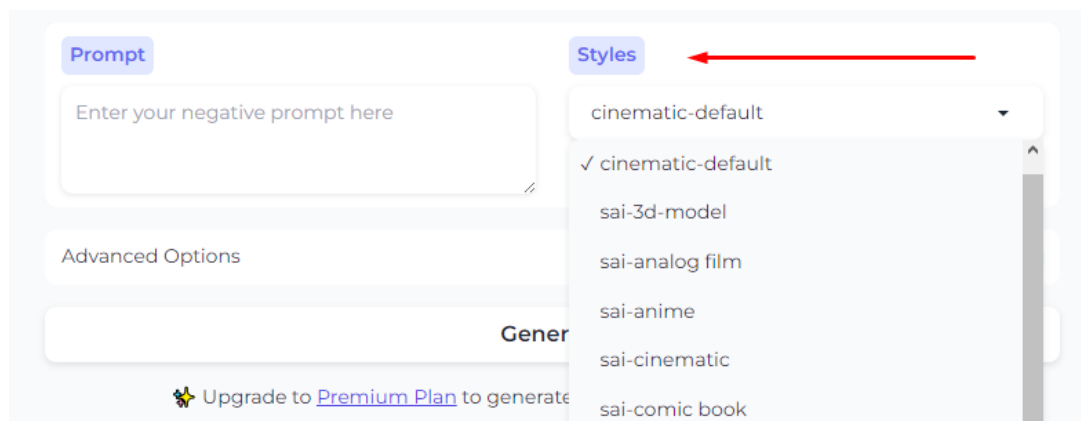


Рисунок 3.5 – приклад стилів

Основною технічною основою Stable Diffusion Online є застосування алгоритмів глибинного навчання, які базуються на концепціях генеративних моделей. Дані алгоритми, зокрема, можуть включати в себе варіації генеративно-зіставних мереж (GANs), які дозволяють створювати нові зображення, зберігаючи при цьому структуру та контент згідно із заданими промтами та обраними стилями.

Науковий аспект використання Stable Diffusion Online визначається не лише природою алгоритмів, але й можливістю налаштування параметрів, що впливає на візуальну якість та естетичні характеристики отриманих зображень. Вибір оптимальних конфігурацій сприяє досягненню бажаного результату, а саме - генерації високоякісних та стилізованих зображень у відповідь на запит користувача.

3.3 Створення відео завдяки онлайн сервісу Runway

У процесі початку роботи з платформою Runway, перед використанням її функціональностей, обов'язковою є авторизація користувача (рис. 3.6). Для виконання

цієї операції відкрита можливість використання облікових записів у сервісах Google, Microsoft або за допомогою принципу single sign on (SSO). Важливо відзначити, що даний етап є необхідним для взаємодії з платформою та отримання доступу до її ресурсів.

Welcome to Runway


Don't have an account? [Sign up for free](#)


Username or Email

Password

Log in

OR

 **Log in with Google**

 **Log in with Apple**

Use Single Sign-On (SSO)

[Forgot Password](#)

Рисунок 3.6 – Поле авторизації

Після успішної авторизації користувач має можливість розпочати процес створення відео, обираючи одну з двох доступних опцій: «Start with image» або «Start with text»(рис. 3.7), відповідно до власних вподобань та потреб. Критично визначити, що вибір даного етапу визначає подальші можливості для користувача в реалізації його творчого підходу.

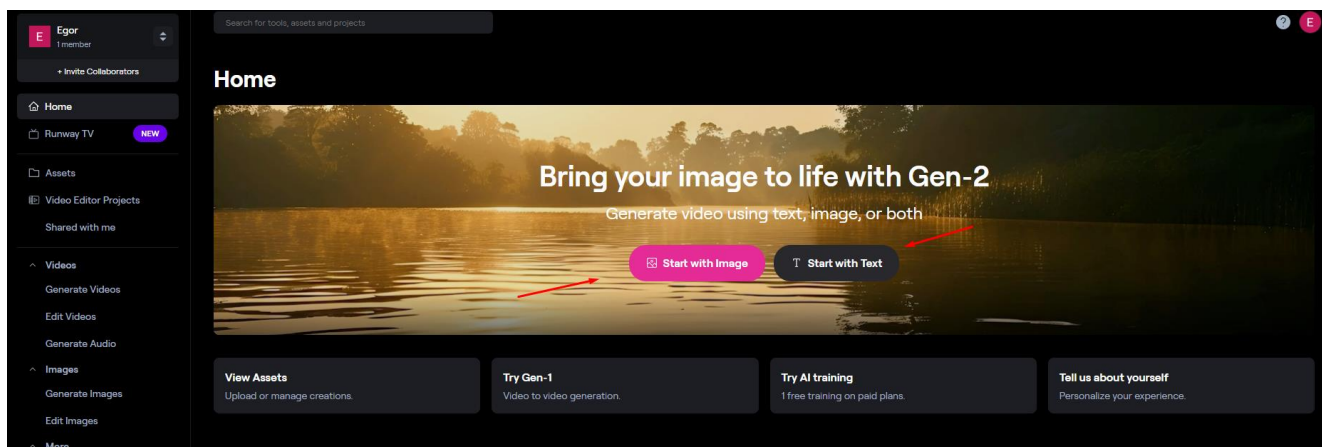


Рисунок 3.7 – Поле вибору функціоналу

У випадку обрання «Start with image», користувач отримує можливість створити запит, використовуючи фотографії. Зазначене програмне забезпечення, застосовуючи технології штучного інтелекту, генерує відеоконтент на основі введених фотографій та обраного стилю. Альтернативно, в разі обрання «Start with text», користувач може створити текстовий запит для подальшого перетворення його в відео.

Додатково, існує можливість комбінування фото з текстовим запитом. Користувач має можливість вказати очікуваний результат, використовуючи це взаємодіюче інтерфейсне середовище для втілення своїх ідей та концепцій в мультимедійний контент.

Готовий відеофайл після генерації стає доступним для перегляду безпосередньо в інтерфейсі Runway(рис. 3.8).

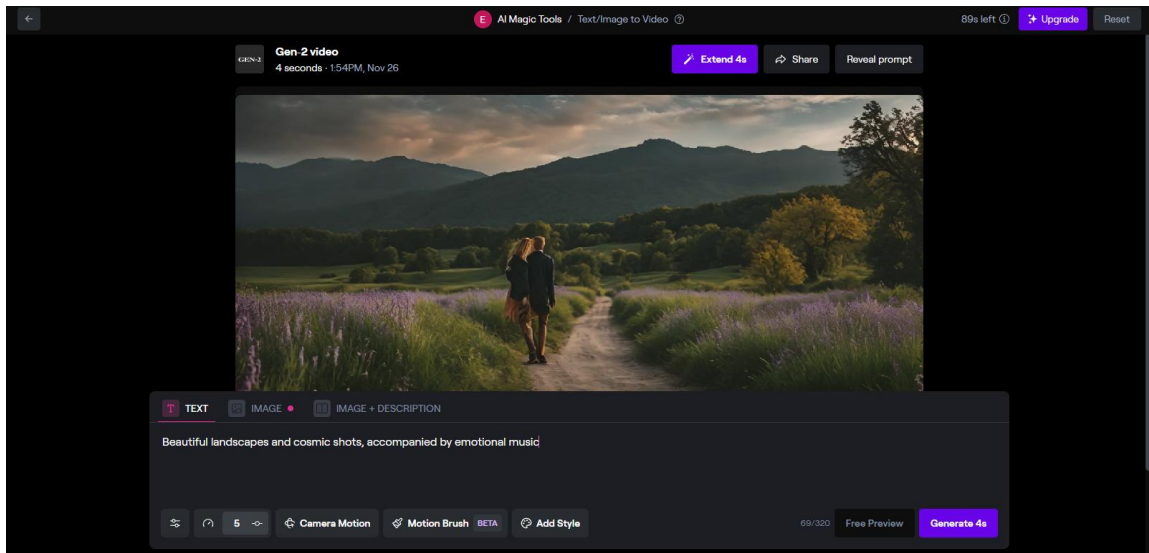


Рисунок 3.8 – Поле з готовим відео

Користувач може оцінити результат та впевнитися, що він відповідає його очікуванням та творчим вимогам. У разі задовільної реакції користувача, готове відео можна завантажити на носій та використовувати його у подальших етапах

/*

ВИСНОВКИ

В даному дипломному проєкті було вивчено та розглянуто можливості використання неймереж для створення сценаріїв, фотографій та відеоматеріалів. В ході досліджень було виявлено, що застосування неймереж, зокрема ChatGPT для сценаріїв, Stable Diffusion Online для генерації фотографій та Runway для створення відео, дозволяє досягти високого рівня творчості та реалізму у виробництві медійних контентів.

Створений сценарій за допомогою ChatGPT вражає своєю натуральністю та креативністю, що робить його відмінним інструментом для написання сценаріїв у різних жанрах та форматах. Stable Diffusion Online дозволяє генерувати фотографії високої якості з вражаючою деталізацією та естетичністю. Використання Runway для створення відео дозволяє легко комбінувати різні візуальні ефекти та створювати вражаючі відеоматеріали.

Завдяки новаторським інструментам та технологіям, можливості для творчого виявлення в сфері медійного виробництва розширюються, відкриваючи нові перспективи для розробки та створення унікальних та захоплюючих контентів. У проєкті було продемонстровано, що поєднання неймереж із сучасними технологіями дозволяє досягти вражаючих результатів у сфері медійного мистецтва .

ПЕРЕЛІК ПОСИЛАНЬ

1. Режим доступу до ресурсу:
https://neerc.ifmo.com/wiki/index.php?title=Generative_Adversarial_Nets_%28GAN%29 (дата звернення: 22.12.2023)
2. Режим доступу до ресурсу: <https://incrypted.com/top-nejrosetej-dlja-generacii-kartinok-nft-po-opisaniju/> (дата звернення: 20.12.2023)
3. Goodfellow, I., Bengio, Y., & Courville, A. Deep Learning. MIT Press. - 2016, pp. 235–245.
4. Szeliski, R. Computer Vision: Algorithms and Applications. Springer. - 2010, pp. 102–112.
5. Foster, D. Generative Deep Learning. O'Reilly Media. - 2019, pp. 451–461.
6. Géron, A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media. - 2019, pp. 789–799.
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
8. Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer.
9. Foster, D. (2019). Generative Deep Learning. O'Reilly Media.
10. Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media.
11. Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image Style Transfer Using Convolutional Neural Networks. Режим доступу до ресурсу: [arXiv:1508.06576](<https://arxiv.org/abs/1508.06576>) (дата звернення: 05.12.2023).
12. Goodfellow, I., et al. (2014). Generative Adversarial Nets. Режим доступу до ресурсу: [arXiv:1406.2661](<https://arxiv.org/abs/1406.2661>) (дата звернення: 01.11.2023).

13. Radford, A., et al. (2016). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. Режим доступа до ресурсу: [arXiv:1511.06434](https://arxiv.org/abs/1511.06434) (дата звернення: 15.11.2023).
14. Vaswani, A., et al. (2017). Attention is All You Need. Режим доступа до ресурсу: [arXiv:1706.03762](https://arxiv.org/abs/1706.03762) (дата звернення: 20.12.2023).
15. Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A Neural Algorithm of Artistic Style. Режим доступа до ресурсу: [arXiv:1508.06576](https://arxiv.org/abs/1508.06576) (дата звернення: 10.12.2023).
16. GitHub - Awesome Deep Learning. Режим доступа до ресурсу: [GitHub - Awesome Deep Learning](https://github.com/ChristosChristofidis/awesome-deep-learning) (дата звернення: 08.12.2023).
17. Ng, A. (Coursera). Convolutional Neural Networks. Режим доступа до ресурсу: [Coursera - Convolutional Neural Networks](https://www.coursera.org/learn/convolutional-neural-networks) (дата звернення: 12.11.2023).
18. Ng, A. (Coursera). Sequence Models. Режим доступа до ресурсу: [Coursera - Sequence Models](https://www.coursera.org/learn/nlp-sequence-models) (дата звернення: 03.12.2023).
19. Kartashov, V.M., Oleynikov V.N, Zubkov, O.V., Korytsev I.V., Babkin, S. I., Sheiko, S.A., Kolendovskaya, M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles; Telecommunications and Radio Engineering, 2020. Vol. 79, Iss, 9, pp.769–780.
20. Kartashov, V.M., Sidorov, G.I., Sheiko, S.A., Kolendovska, M.M., Sergienko O.Yu. Principles of Construction and Assessment of technical Characteristics of multi-Frequency atmospheric Sodar in the Humidity Measurement Mode / Telecommunications and Radio Engineering. – New York, 2020. – Vol. 79, №4, pp. 323–333. DOI: 10.1615/TelecomRadEng.v79.i4.50 (стаття).
21. V. M. Kartashov, V. N. Oleynikov, S. A. Sheyko, I. V. Koryttsev, S. I. Babkin, O. V. Zubkov, «Peculiarities of small unmanned aerial vehicles detection and

recognition». *Telecommunications and Radio Engineering*, 2019, V. 78, Iss. 9, pp. 771–781. DOI: 10.1615/TelecomRadEng.v78.i9.30.

22. V.M. Semenets, V.M. Kartashov, V.I. Leonidov. Features of Acoustic Noise of Small Unmanned Aerial Vehicles // *Telecommunications and Radio Engineering*. – New York, 2020.– Vol. 79, №11, pp. 985–995. DOI: 10.1615/TelecomRadEng.v79.i11.80 (статья).