

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних наук
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр

(назва освітнього ступеня)

на тему: Дослідження застосування технологій штучного інтелекту для відеомонтажу

Виконав: студент IV курсу, групи СН-43

спеціальності 122 Комп'ютерні науки

(шифр і назва спеціальності)

(підпис)

Шипка В.П.

(прізвище та ініціали)

Керівник

(підпис)

Небесний Р.М.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Шимчук Г.В.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Боднарчук І.О.

(прізвище та ініціали)

Рецензент

(підпис)

Чихіра І.В.

(прізвище та ініціали)

Тернопіль
2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра комп'ютерних наук
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
Боднарчук І.О.
(підпис) (прізвище та ініціали)
« 27 » січня 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Бакалавр
(назва освітнього ступеня)
за спеціальністю 122 Комп'ютерні науки
(шифр і назва спеціальності)
Студенту Шипці Вячеславу Петровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження застосування технологій штучного інтелекту для відеомонтажу

Керівник роботи Небесний Руслан Михайлович, доктор філософії, доцент
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від « 07 » травня 2025 року № 4/7-444

2. Термін подання студентом завершеної роботи 28 червня 2025 р.

3. Вихідні дані до роботи Літературні та інтернет джерела

4. Зміст роботи (перелік питань, які потрібно розробити) Вступ. 1. Аналіз предметної області та постановка завдання. 1.1 Сучасний стан відеомонтажу та його автоматизація. 1.2 Застосування штучного інтелекту у відеомонтажі. 1.3 Огляд програмних рішень. 1.4 Огляд сучасних моделей ШІ для відеомонтажу. 1.5 Актуальні напрями застосування відеомонтажу на основі ШІ. 1.6 Порівняння традиційного та AI-орієнтованого відеомонтажу. 1.7 Тренди застосування ШІ у відеоіндустрії. 1.8 Висновок до першого розділу. 2. Реалізація монтажу з використанням ШІ. 2.1 Концепція відеопроєкту та загальна ідея. 2.2 Обґрунтування вибору інструментів та моделей ШІ. 2.3 Загальна архітектура та етапи реалізації відеопроєкту. 2.4 Порівняльний аналіз альтернативних інструментів для реалізації відеопроєкту. 2.5 Ризики, обмеження та етичні аспекти використання ШІ у відеомонтажі. 2.6 Оцінка технічних вимог та ресурсів для реалізації. 2.7 Варіанти масштабування та адаптація методу. 2.8 Висновок до другого розділу. 3. Створення відео з використанням ШІ-технологій. 3.1 Формування сценарного плану та підказок для генерації відео. 3.2 Генерація відеофрагментів у RunwayML (Gen-4). 3.3 Озвучення голосом головної героїні за допомогою ElevenLabs. 3.4 Генерація фонові музики та звуків оточення за допомогою ШІ. 3.5 Створення субтитрів за допомогою ШІ. 3.6 Фінальний монтаж у DaVinci Resolve. 3.7 Тестування якості відео: візуальні, аудіо та емоційні параметри. 3.8 Висновок до третього розділу. 4. Безпека життєдіяльності, основи охорони праці. 4.1 Ризики психоемоційного вигорання користувачів ПК. 4.2 Характеристика життєдіяльності людини у системі «людина – машина – середовище». 4.3 Висновок до четвертого розділу. Висновки. Перелік джерел. Додатки
5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи охорони праці	Сенчишин Віктор Степанович		

7. Дата видачі завдання 27 січня 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	27.01.2025	
2.	Підбір та опрацювання літературних джерел по темі кваліфікаційної роботи	08.05.2025–15.05.2025	
3.	Виконання дослідження щодо застосування штучного інтелекту для відеомонтажу	19.05.2025–23.05.2025	
	Розроблення дипломного проєкту, міні-фільм		
4.	Оформлення розділу «Аналіз предметної області та постановка завдання»	26.05.2025–02.06.2025	
5.	Оформлення розділу «Реалізація монтажу з використанням штучного інтелекту»	26.05.2025–02.06.2025	
6.	Оформлення розділу «Створення відео з використанням ШІ»	26.05.2025–02.06.2025	
7.	Виконання завдання до підрозділу «Безпека життєдіяльності»	02.06.2025–05.06.2025	
8.	Виконання завдання до підрозділу «Основи охорони праці»	02.06.2025–05.06.2025	
9.	Оформлення кваліфікаційної роботи	06.06.2025–10.06.2025	
10.	Нормоконтроль	11.06.2025–13.06.2025	
11.	Перевірка на плагіат	18.06.2025	
12.	Попередній захист кваліфікаційної роботи	20.06.2025	
13.	Захист кваліфікаційної роботи	29.06.2025	

Студент

(підпис)

Шипка В.П.

(прізвище та ініціали)

Керівник роботи

(підпис)

Небесний Р.М.

(прізвище та ініціали)

АНОТАЦІЯ

Дослідження застосування технологій штучного інтелекту для відеомонтажу // Кваліфікаційна робота освітнього рівня «Бакалавр» // Шипка Вячеслав Петрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра комп'ютерних наук, група СН-43 // Тернопіль, 2025 // С. 57, рис. – 7, табл. – 10, додат. – 3, бібліогр. – 41.

Ключові слова: штучний інтелект, відеомонтаж, автоматизація, комп'ютерний зір, нейронні мережі, машинне навчання, обробка відео, порівняльний аналіз.

Кваліфікаційна робота присвячена дослідженню застосування сучасних технологій штучного інтелекту у процесі відеомонтажу.

В першому розділі кваліфікаційної роботи описано основні поняття, типові завдання відеомонтажу та сучасний стан розвитку інструментів штучного інтелекту, що використовуються для обробки відео. Висвітлено принципи роботи таких систем. Розглянуто особливості ручного та автоматизованого монтажу. Проаналізовано існуючі програмні засоби.

В другому розділі кваліфікаційної роботи визначено вимоги до інструментів, обґрунтовано вибір середовищ, описано архітектуру та принципи роботи обраних технологій. Досліджено можливості сучасних рішень. Подано модель експерименту для порівняння результатів.

В третьому розділі кваліфікаційної роботи описано процес реалізації монтажу вручну та з використанням ШІ, проаналізовано результати, проведено оцінювання ефективності методів.

ANNOTATION

Research on the Application of Artificial Intelligence Technologies for Video Editing
// Qualification work of the educational level «Bachelor» // Shypka Viacheslav
Petrovych // Ternopil Ivan Pulyu National Technical University, Computer and
Information Systems and Software Engineering Faculty, Computer Sciences
Department, group SN-43 // Ternopil, 2025 // P. 57, fig. – 7, tabl. – 10, annexes. – 3,
references – 41.

Keywords: artificial intelligence, video editing, automation, computer vision,
neural networks, machine learning, video processing, comparison.

The qualification work is dedicated to the study of the application of artificial intelligence technologies in the process of video editing.

The goal of the work is to investigate the effectiveness of AI-based tools for automating video editing and to compare them with manual editing methods.

The first section of the qualification paper considered the basic concepts of video editing and artificial intelligence, as well as an overview of existing tools and technologies.

In the second section of the qualification work, the structure and functionality of selected AI-based systems were examined, and a methodology for comparative analysis was proposed.

In the third section of the qualification work, a practical implementation of both manual and AI-based video editing was carried out, followed by performance evaluation and result analysis.

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ І ТЕРМІНІВ

AI (англ. Artificial Intelligence) – штучний інтелект.

CV (англ. Computer Vision) – комп'ютерний зір; напрям штучного інтелекту, що дозволяє комп'ютерам «бачити» й аналізувати зображення та відео.

DL (англ. Deep Learning) – глибинне навчання; підгалузь машинного навчання, що використовує штучні нейронні мережі для аналізу великих обсягів даних.

ML (англ. Machine Learning) – машинне навчання; технологія, що дозволяє системам навчатися на основі даних без явного програмування.

NLP (англ. Natural Language Processing) – обробка природної мови; галузь, яка займається взаємодією між комп'ютерами та людською мовою.

TTS (англ. Text-to-Speech) – перетворення тексту в мовлення; технологія, що забезпечує синтез голосу з текстової інформації.

ШІ – Штучний інтелект.

ЗМІСТ

ВСТУП.....	8
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАВДАННЯ	9
1.1 Сучасний стан відеомонтажу та його автоматизації.....	9
1.2 Застосування штучного інтелекту у відеомонтажі	10
1.3 Огляд програмних рішень	11
1.4 Огляд сучасних моделей штучного інтелекту для відеомонтажу	13
1.5 Актуальні напрями застосування відеомонтажу на основі ШІ	14
1.6 Порівняння традиційного та AI-орієнтованого відеомонтажу	15
1.7 Тренди застосування ШІ у відеоіндустрії	17
1.8 Висновок до першого розділу	18
РОЗДІЛ 2. РЕАЛІЗАЦІЯ МОНТАЖУ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ	20
2.1 Концепція відеопроєкту та загальна ідея	20
2.2 Обґрунтування вибору інструментів та моделей штучного інтелекту	21
2.3 Загальна архітектура та етапи реалізації відеопроєкту	23
2.4 Порівняльний аналіз альтернативних інструментів для реалізації відеопроєкту	25
2.5 Ризики, обмеження та етичні аспекти використання ШІ у відеомонтажі.....	28
2.6 Оцінка технічних вимог та ресурсів для реалізації.....	29
2.7 Варіанти масштабування та адаптація методу	31
2.8 Висновок до другого розділу	33
РОЗДІЛ 3. СТВОРЕННЯ ВІДЕО З ВИКОРИСТАННЯМ ШІ-ТЕХНОЛОГІЙ.....	33
3.1 Формування сценарного плану та підказок для генерації відео	34
3.2 Генерація відеофрагментів у RunwayML (Gen-4)	35
3.3 Озвучення голосом головної героїні за допомогою ElevenLabs	37

3.4 Генерація фонові музики та звуків оточення за допомогою ШІ	38
3.5 Створення субтитрів за допомогою ШІ.....	39
3.6 Фінальний монтаж у DaVinci Resolve.....	41
3.7 Тестування якості відео: візуальні, аудіо та емоційні параметри	43
3.8 Висновок до третього розділу	45
РОЗДІЛ 4. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ	46
4.1 Ризики психоемоційного вигорання користувачів ПК	46
4.2 Характеристика життєдіяльності людини у системі «людина – машина – середовище існування»	48
4.3 Висновок до четвертого розділу.....	50
ВИСНОВКИ	51
ПЕРЕЛІК ДЖЕРЕЛ	53
ДОДАТКИ	

ВСТУП

Актуальність теми. У сучасну епоху цифрових технологій відеоконтент став одним із найпопулярніших способів передачі інформації, самовираження, реклами та розваг. Одним із ключових напрямів таких інновацій є застосування технологій штучного інтелекту. Алгоритми комп'ютерного зору, машинного навчання та нейромереж відкривають нові можливості для аналізу, обробки та редагування відео, що дозволяє значно зменшити часові витрати, підвищити ефективність і креативний потенціал відеомонтажу. Зважаючи на стрімкий розвиток ШІ та його інтеграцію в цифрові творчі процеси, дослідження ефективності таких рішень у сфері відеомонтажу є надзвичайно актуальним.

Мета і задачі дослідження. Метою даної кваліфікаційної роботи освітнього рівня «Бакалавр» є дослідження можливостей застосування сучасних технологій штучного інтелекту для автоматизації відеомонтажу, а також порівняння їх ефективності з традиційними методами редагування. Для досягнення поставленої мети потрібно виконати ряд завдань, зокрема:

- провести аналіз сучасного стану розвитку технологій ШІ, що застосовуються у відеомонтажі;
- визначити найбільш популярні програмні засоби, які реалізують такі технології;
- узагальнити результати дослідження та сформулювати висновки щодо ефективності застосування технологій ШІ.

Практичне значення одержаних результатів. Результати дослідження можуть бути використані для підвищення ефективності виробництва відеоконтенту в рекламній, освітній, творчій та комерційній сферах. Реалізований порівняльний аналіз дозволяє краще зрозуміти переваги та обмеження ШІ-інструментів у порівнянні з ручним монтажем, а також сформулювати рекомендації щодо їх раціонального використання.

РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАВДАННЯ

1.1 Сучасний стан відеомонтажу та його автоматизації

У сучасному інформаційному просторі відео є одним з основних форматів передачі даних – для реклами, освіти, розваг, соціальних медіа. З огляду на це, процес відеомонтажу набуває особливого значення, адже якісний монтаж безпосередньо впливає на сприйняття кінцевого продукту.

Традиційний відеомонтаж є трудомістким та креативно-залежним процесом, який вимагає значного досвіду оператора, вміння працювати з великою кількістю матеріалу, дотримуватись логічної структури, таймінгу, ритму тощо. Для оптимізації цих процесів з кожним роком зростає кількість інструментів, що використовують автоматизовані алгоритми та технології штучного інтелекту [1].

Ще одним важливим аспектом, що впливає на відеомонтаж, є збільшення обсягів відеоконтенту, який щодня створюється та завантажується користувачами. Згідно зі статистикою, щохвилини на YouTube завантажується понад 500 годин відео, і ця цифра постійно зростає [7]. Така кількість матеріалу вимагає швидкої обробки, а ручне редагування в таких масштабах є малоефективним, що й стимулює розвиток технологій автоматизованого монтажу.

Активне впровадження мобільних додатків і хмарних платформ зробило відеомонтаж доступним для широкого кола користувачів – навіть тих, хто не має професійних навичок. Це сприяє популяризації інструментів, які дозволяють створювати привабливі ролики за лічені хвилини, завдяки попередньо налаштованим алгоритмам та шаблонам. У цьому контексті штучний інтелект виступає не лише як інструмент для професіоналів, а й як

помічник для звичайних користувачів, що бажають створити якісний відеоконтент без складного програмного забезпечення [8].

1.2 Застосування штучного інтелекту у відеомонтажі

Технології штучного інтелекту активно впроваджуються у сферу відеомонтажу, відкриваючи нові можливості для автоматизації та підвищення якості обробки відеоматеріалів. Одним із основних напрямів застосування ШІ є автоматичне вирізання непотрібних фрагментів відео, що дозволяє значно зменшити тривалість монтажу та уникнути рутинної роботи. Також поширеним є використання алгоритмів розпізнавання ключових моментів у відео – за змінами голосу, мімікою або рухами в кадрі – що забезпечує логічне структурування відеоряду [2].

Прикладами застосування ШІ у відеомонтажі є такі програмні рішення, як RunwayML – платформа з багатьма готовими моделями для роботи з відео, Adobe Sensei – інтегрована система ШІ у Premiere Pro, яка надає інструменти Smart Cut, Auto Reframe тощо, а також онлайн-сервіси на зразок Magisto, Wisecut і Pictory. Вони орієнтовані як на професійних користувачів, так і на новачків, що прагнуть швидко та якісно створити відеопродукт [6].

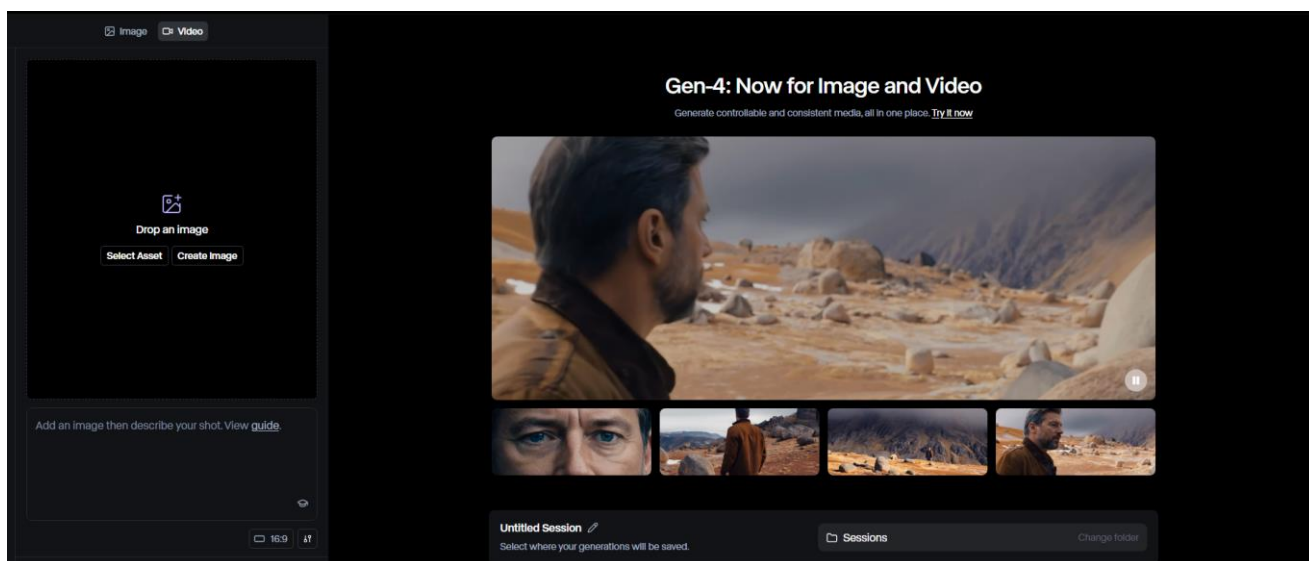


Рисунок 1.1 – Головний екран роботи в програмі RunwayML

III також застосовується для стабілізації зображення, покращення кольорокорекції, узгодження відеоряду з аудіодоріжкою, а також для генерації субтитрів і титрів. Ці функції реалізуються за допомогою складних моделей комп'ютерного зору (computer vision), які аналізують зміст кадрів, та глибинного навчання (deep learning), що дозволяє виявляти контекст сцен і передбачати логіку монтажу. Крім того, за допомогою методів обробки природної мови (NLP) можлива автоматична генерація описів, анотацій і навіть сценаріїв до відео [3].

Особливої популярності набули генеративні моделі, зокрема GAN (Generative Adversarial Networks) та diffusion-моделі, які дозволяють створювати новий відеоконтент або редагувати існуючий на основі заданих параметрів. Такі моделі вже використовуються не лише в наукових дослідженнях, а й у комерційних продуктах, що підтверджує їхню практичну цінність.

1.3 Огляд програмних рішень

У таблиці 1.1 подано огляд популярних III-інструментів для відеомонтажу.

Таблиця 1.1 – Популярні програмні засоби для відеомонтажу з використанням III

Назва інструменту	Основні можливості	Використання технології
RunwayML	Автоматичне кадрування, заміна фону	CV, DL
Adobe Sensei	Smart Cut, Color Match, Auto Reframe	ML, CV
Magisto	Автоматичне створення кліпів з шаблонів	ML, шаблонна логіка
Wisecut	Видалення пауз, автоматичні субтитри	NLP, TTS, CV
Pictory	Генерація відео з тексту	NLP, DL

Одним із найпопулярніших інструментів на базі штучного інтелекту для відеомонтажу є RunwayML. Ця платформа надає доступ до десятків моделей

комп'ютерного зору та генеративного навчання, які працюють у режимі реального часу. Серед найбільш затребуваних функцій – автоматичне видалення фону, стильова трансформація відео, синтез мови, генерація субтитрів. Платформа RunwayML дозволяє інтегрувати II-можливості у креативні процеси без необхідності програмування, що робить її доступною навіть для користувачів без технічного бекграунду [4].

Adobe Sensei – це високотехнологічна платформа штучного інтелекту, інтегрована в продукти Adobe, зокрема Premiere Pro та After Effects. Вона забезпечує такі функції, як Auto Reframe (автоматична адаптація відео під різні формати), Scene Edit Detection (виявлення місць склеювання сцен), Color Match (автоматичне узгодження кольору між кліпами) та інші. Завдяки тісній інтеграції в професійне програмне забезпечення Adobe, Sensei орієнтований передусім на професіоналів у сфері відео та дизайну [5].

Онлайн-сервіс Magisto дозволяє створювати відеоролики автоматично на основі завантажених фото, відео та обраної музики. Користувачу потрібно лише вказати тип відео (наприклад, реклама або привітання), а система сама виконує монтаж за заданими шаблонами. Magisto використовує алгоритми III для вибору найкращих моментів, вставлення переходів, візуальних ефектів і синхронізації з музикою. Цей інструмент орієнтований переважно на бізнес-цілі та SMM-сегмент.

Wisecut спеціалізується на автоматичному монтажі відео за аудіотреком. Алгоритм визначає паузи, «слова-паразити», невпевнені інтонації й автоматично видаляє їх, зберігаючи динамічний темп. Крім того, Wisecut додає субтитри, фонову музику та може синхронізувати текст з відео. Завдяки цьому сервіс особливо корисний для створення навчальних відео, інтерв'ю, подкастів, відеоблогів [6].

Pictory – це хмарна платформа, яка генерує відео з тексту. Користувач вводить сценарій або блог-пост, і система автоматично підбирає візуальний контент, озвучення, фонову музику та титри. Pictory активно використовується маркетологами, освітянами та блогерами для створення коротких відео на

основі статей чи текстових звітів. Платформа поєднує алгоритми NLP, генеративного синтезу зображення та аудіо.

1.4 Огляд сучасних моделей штучного інтелекту для відеомонтажу

Розвиток генеративного штучного інтелекту відкрив новий етап у відеомонтажі – створення відео або його елементів без участі традиційних інструментів редагування. Одними з найперспективніших напрямів стали моделі трансформації відео за допомогою текстових підказок, стилізації, заміни об'єктів, зміни фону та автоматичної генерації відеофрагментів з тексту або зображення.

Однією з перших загальнодоступних моделей цього типу була Gen-1 від RunwayML – модель, яка дозволяє трансформувати відео на основі вхідного зображення або анімаційної маски. Наприклад, користувач може застосувати стиль малюнка до відео або замінити об'єкти у сцені. Gen-1 базується на глибинних генеративних моделях та diffusion-підходах, працює в хмарі та має зручний веб-інтерфейс [4].

Продовженням цієї технології стала Gen-2, яка підтримує створення відео з тексту (text-to-video), а також зображення (image-to-video). Gen-2 суттєво розширила можливості: користувач може ввести опис на кшталт «A drone shot of a forest during sunset», і модель створить кількaseкундний кліп з відповідною динамікою, кольорами та перспективою. Відео створюється на основі навчання на великій множині даних і враховує контекст, стиль, камеру, рух [4].

Ще більш потужною моделлю є Sora від OpenAI – експериментальна система, здатна створювати відео тривалістю до хвилини з високим ступенем реалістичності. На відміну від попередніх моделей, Sora розуміє не лише текстовий опис, але й просторову взаємодію об'єктів, що дозволяє створювати складні сцени з кількома персонажами, рухами камери та зміною середовища [9]. Система перебуває на етапі обмеженого доступу, однак уже

демонструє потенціал повністю замінити частину функцій традиційного відеомонтажу.

Також варто згадати VEED AI – платформу для автоматизованого відеомонтажу, що поєднує функції розпізнавання мови, генерації субтитрів, створення коротких роликів для соцмереж, а також автоматичного кадрування. VEED AI більше орієнтована на практичне застосування для відеомаркетингу, YouTube, TikTok та освітніх проектів [10].

Кожна з цих моделей має свої переваги. Наприклад, Gen-1 і Gen-2 – для стилізації й модифікації наявного відео, Sora – для генерації відео «з нуля», а VEED AI – для швидкої адаптації контенту під платформу чи аудиторію.

1.5 Актуальні напрями застосування відеомонтажу на основі ШІ

Технології штучного інтелекту у відеомонтажі дедалі активніше застосовуються в різних галузях, де відео відіграє ключову роль у комунікації, навчанні, рекламі або розвагах.

Однією з найактивніших сфер є маркетинг і реклама. Компанії використовують ШІ для створення рекламних відео з персоналізованим контентом, який адаптується до кожної аудиторії. Наприклад, за допомогою AI-алгоритмів можна автоматично змінювати мову, стиль, графіку залежно від країни чи демографії глядача. Платформи на кшталт Pictory чи VEED AI дозволяють маркетологам створювати відео з блогів, сценаріїв чи підказок – за лічені хвилини [10].

У журналістиці та медіа ШІ допомагає в автоматизації репортажів і новинних сюжетів. Наприклад, деякі медіакомпанії використовують інструменти, які аналізують відеозаписи подій, автоматично виділяють ключові моменти та формують короткі оглядові відео. Такі можливості застосовуються в реальному часі під час трансляцій подій, як-от спортивні матчі або політичні дебати [2].

У освіті відео стало основним інструментом для дистанційного навчання. Інструменти на базі ШІ допомагають викладачам автоматично генерувати навчальні ролики, створювати субтитри, синхронізувати лекції з презентаціями, а також адаптувати відео до потреб людей з особливими освітніми потребами. Наприклад, AI може автоматично розбивати лекцію на тематичні блоки, створювати тести або узагальнення [6].

В кіновиробництві II використовується для прискорення постпродакшену. Автоматичне вирізання непотрібних сцен, заміна фону без використання хромакею, кольорокорекція та створення спецефектів – усе це частково або повністю автоматизується. Генеративні моделі, як-от Gen-2 чи Sora, відкривають нові горизонти для експериментального кіно, де сцени можуть створюватись «з нуля» без зйомки в традиційному розумінні [9].

У соціальних мережах і сфері контент-креації ШІ став незамінним помічником. Креатори можуть швидко створювати динамічний контент, що адаптується під алгоритми TikTok, YouTube Shorts або Instagram. Автоматичне кадрування, додавання субтитрів, динамічний монтаж – усе це дозволяє прискорити виробництво контенту і зробити його привабливішим для глядача [4].

Застосування ШІ у відеомонтажі сприяє цифровій трансформації творчих індустрій, що відповідає сучасним тенденціям розвитку бізнес-структур [23]. Використання ШІ поширюється в усіх сферах, де потрібна швидкість, масштабованість та адаптивність відеоконтенту.

1.6 Порівняння традиційного та AI-орієнтованого відеомонтажу

Відеомонтаж як процес формування аудіовізуального контенту має довгу історію, що включає як аналогові, так і цифрові етапи розвитку. До недавнього часу монтаж вимагав значного обсягу ручної роботи: від зйомки відеоматеріалу до багатогодинного редагування, синхронізації звуку, кольорокорекції та постобробки. У свою чергу, сучасні інструменти на базі штучного інтелекту

суттєво змінюють підхід до створення відео: багато процесів автоматизуються або спрощуються, що дає змогу зосередитися на креативній складовій.

Нижче представлено порівняльну таблицю 1.2 ключових характеристик традиційного та AI-орієнтованого відеомонтажу. Це дозволяє наочно оцінити сильні та слабкі сторони кожного підходу, а також виявити сценарії, в яких використання ШІ є найбільш доцільним.

Таблиця 1.2 – Порівняння характеристик двох типів відеомонтажу

Критерій	Традиційний монтаж	AI-орієнтований монтаж
Джерело відео	Реальні зйомки (камера, сцени, актори)	Генерація відео за описом, стокові дані
Час виконання	Від кількох днів до тижнів	Від кількох годин до 1–2 днів
Вартість	Висока (обладнання, персонал, локації)	Середня або низька (доступ до платформ ШІ)
Потреба в команді	Обов'язкова (оператор, режисер, монтажер)	Можливість виконання одним користувачем
Гнучкість сценарію	Залежить від зйомок, обмежена	Висока: можна змінювати prompts і сцени
Точність і контроль	Повний контроль на всіх етапах	Часткова залежність від генеративної моделі
Вимоги до навичок	Професійні знання в монтажі	Базові навички та креативне мислення
Оцінка якості результату	Висока стабільність при хорошій підготовці	Варіативність якості, залежно від генерації

У більшості випадків AI-орієнтований монтаж дозволяє значно зекономити час і ресурси, особливо для малобюджетних проєктів, навчальних матеріалів, соціальних відео або блогінгу. Проте варто враховувати, що повний контроль за результатом частково втрачається, особливо при використанні генеративних моделей, які працюють за статистичними принципами, а не за чіткими правилами.

Наприклад, для зйомки художнього відео з акторами потрібно забезпечити реквізит, локацію, знімальну групу, а також монтаж у професійному середовищі. У той час як генеративні моделі, такі як Gen-4 від RunwayML, дають змогу створити сцену за описом, без жодної зйомки. Той

самий підхід можна застосувати і до голосу: замість студійного запису можна використати ElevenLabs або PlayHT.

Загалом, ШІ не повністю витісняє традиційний монтаж, але створює альтернативну парадигму – швидку, доступну, персоналізовану. У випадках, коли важливий результат, а не процес, саме AI-орієнтований підхід має значні переваги. У наступному підрозділі буде розглянуто, як ці технології трансформують відеоіндустрію глобально.

1.7 Тренди застосування ШІ у відеоіндустрії

Швидкий розвиток технологій штучного інтелекту докорінно змінює сучасну відеоіндустрію. Якщо раніше створення відео вимагало багаторівневої роботи з великою командою фахівців, то сьогодні завдяки AI-інструментам цей процес може виконуватись навіть однією людиною за кілька годин. Це сприяє демократизації відеовиробництва й відкриває нові можливості як для незалежних творців, так і для великих компаній. Сучасні ШІ-технології, такі як RunwayML чи Sora, базуються на адаптивних інформаційних системах, подібних до тих, що застосовуються для персоналізованих рішень [32].

Одним із найпомітніших трендів є автоматична генерація коротких відеоформатів для TikTok, YouTube Shorts, Instagram Reels. За допомогою платформ на зразок Pictory, Lumen5 або Synthesia, користувачі можуть створювати динамічні відео з текстів, презентацій або сценаріїв. ШІ не лише монтує відеоряд, а й добирає музику, накладає субтитри, додає ефекти – без участі людини у творчому процесі.

Іншим трендом є персоналізовані AI-трейлери, які адаптуються до інтересів користувача. Наприклад, Netflix використовує штучний інтелект для створення різних варіантів прев'ю одного фільму – залежно від історії переглядів, регіону та переваг глядача [31]. Це підвищує залучення та релевантність відеоконтенту, одночасно знижуючи витрати на постпродакшн.

ШІ також активно застосовується для створення deepfake-контенту, зокрема у кіно, рекламі та розважальних проєктах. Інструменти як D-ID, Reface або Synthesia Studio дозволяють змінювати обличчя, накладати міміку, синтезувати мовлення та зображення – що відкриває нові можливості у візуальному сторітелінгу. Водночас це викликає етичні виклики, які вже стали предметом обговорення в індустрії.

Важливим напрямом є і автоматизація редагування відео. Програмне забезпечення, наприклад Adobe Premiere Pro з функціями Sensei AI, автоматично обрізає відео під різні формати, стабілізує кадри, виконує колірну корекцію або видаляє паузи в мові. Це значно знижує навантаження на монтажера, особливо при роботі з великими обсягами матеріалу [33].

Можна відокремити п'ять ключових трендів застосування ШІ у відеоіндустрії, дані тренди показано в таблиці 1.3.

Таблиця 1.3 – Ключові тренди використання ШІ у відеоіндустрії

№	Тренд	Суть та приклади застосування
1	Автоматичне створення відео з тексту або даних	Генерація відеороликів з текстових сценаріїв (Pictory, Lumen5, Synthesia)
2	Індивідуалізація контенту для аудиторії	AI-трейлери, адаптовані під глядача (Netflix, YouTube)
3	Генерація візуального й аудіо контенту з нуля	Створення відео, голосу, облич, міміки без реальних акторів (RunwayML, ElevenLabs)
4	Швидке масштабування відеовиробництва	Можливість створювати десятки роликів одночасно через API, шаблони автогенерацію
5	Гібридні моделі співтворчості ШІ і людини	Людина задає сценарій і стиль, ШІ виконує технічну реалізацію з частковим контролем результату

1.8 Висновок до першого розділу

У першому розділі кваліфікаційної роботи проведено комплексний аналіз предметної області відеомонтажу з акцентом на можливості його автоматизації за допомогою сучасних технологій штучного інтелекту. Проаналізовано еволюцію інструментів відеомонтажу – від класичних професійних середовищ

до сучасних AI-платформ, які здатні самостійно генерувати відео, аудіо, субтитри й навіть сценарій на основі текстового опису. Особливу увагу приділено принципам роботи ШІ у сфері комп'ютерного зору, обробки природної мови та генеративних моделей, що стали основою для нового формату створення відеоконтенту.

У розділі також розглянуто програмні рішення, які активно використовуються як професіоналами, так і новачками – зокрема, RunwayML, Adobe Sensei, Magisto, Wisecut, Pictory, Gen-2, Sora, VEED AI. Детально проаналізовано особливості та функціональні можливості моделей типу Gen-1/2 та новітніх diffusion-моделей, що дозволяють створювати відео з нуля. У межах окремого підрозділу охарактеризовано ключові напрями застосування ШІ у відеомонтажі – від освіти й маркетингу до журналістики та соціального контенту.

Порівняння традиційного і AI-орієнтованого монтажу дозволило виокремити переваги та обмеження обох підходів. Зокрема, ШІ забезпечує значну економію часу та ресурсів, надає гнучкість, масштабованість і доступність навіть для користувачів без досвіду. Разом з цим, у розділі окреслено ключові тренди розвитку відеоіндустрії, серед яких: автоматизоване створення відео з тексту, персоналізація контенту, генерація аудіовізуального матеріалу без реальних зйомок, швидке масштабування та гібридна взаємодія людини й ШІ. Отже, теоретичне підґрунтя першого розділу підтверджує актуальність обраної теми дослідження та створює чітку базу для подальшої проєктної і практичної реалізації відеопроєкту в наступних розділах дипломної роботи.

РОЗДІЛ 2. РЕАЛІЗАЦІЯ МОНТАЖУ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

2.1 Концепція відеопроєкту та загальна ідея

У межах дослідження було спроектовано відеопроєкт у форматі короткометражного художнього фільму під назвою «The Last Letter from Paris». Відео є драматичним монологом дівчини, яка читає останній лист від зниклого коханого. Концепція проєкту передбачає створення повноцінного відеоконтенту з використанням інструментів штучного інтелекту на кожному з ключових етапів: від генерації візуального матеріалу до озвучення та субтитрування.

Проєкт має на меті дослідити можливості сучасних ШІ-сервісів у реалізації творчих задумів без залучення знімального обладнання, акторів або студійних ресурсів. Відео створюється шляхом генерації окремих сцен за текстовими сценарними підказками, додавання синтетичного голосу, автоматично згенерованої музики та субтитрів. Це дозволяє оцінити, наскільки генеративні моделі можуть замінити традиційні етапи відеовиробництва.

Таким чином, у рамках проєктної частини сформульовано наступні завдання:

- визначити формат відео та його сюжетну структуру;
- підібрати ШІ-інструменти, які можуть реалізувати кожен з етапів (генерація відео, озвучення, музика, субтитри);
- обґрунтувати вибір кожного з інструментів;
- побудувати загальну архітектуру процесу створення відео на основі ШІ;
- підготувати технічні параметри відео та вимоги до якості.

2.2 Обґрунтування вибору інструментів та моделей штучного інтелекту

Для реалізації проєкту «The Last Letter from Paris» було обрано набір інструментів та моделей штучного інтелекту, які забезпечують повний цикл створення відеоконтенту: генерацію відео, озвучення, супровідну музику, субтитрування та фінальний монтаж. Вибір здійснювався на основі таких критеріїв: функціональність, якість результату, інтеграція з іншими компонентами, простота використання та доступність для некомерційного застосування.

Основою візуального ряду стала модель GEN-4, доступна через платформу RunwayML. Це одна з найсучасніших генеративних моделей типу text-to-video, яка дозволяє створювати реалістичні відеофрагменти на основі коротких текстових описів. Gen-4 демонструє високий рівень деталізації, стилістичну варіативність та глибоку кінематографічну композицію, що робить її придатною для реалізації сюжетного відео з емоційним забарвленням [4].

Для озвучення тексту було обрано ШІ-сервіс зі синтетичною генерацією голосу. Популярними платформами такого типу є ElevenLabs, PlayHT, LOVO AI. Вони надають можливість обрати емоційний стиль, інтонацію, швидкість та навіть акцент диктора. Така гнучкість дозволяє створити глибоко персоналізований голос, який звучить природно й артистично, що важливо для драматичного контенту [5].

Генерація музичного супроводу здійснюється за допомогою платформ, які працюють на основі штучного інтелекту у сфері композиції. Одним із варіантів є Soundraw – сервіс, що створює музичні треки з заданим настроєм, жанром і динамікою. Іншим подібним рішенням є AIVA (Artificial Intelligence Virtual Artist), яка спеціалізується на симфонічних композиціях і дає змогу створювати мелодії з кінематографічним звучанням [12]. Ці інструменти дають змогу забезпечити стилістичну єдність між візуальним і аудіальним рядом.

Для створення субтитрів було передбачено використання сервісів на базі розпізнавання мовлення (ASR) та перекладу з NLP–моделями, як–от DeepL Translate, Veed.io або інструменти, вбудовані в Descript чи Kapwing. Вони забезпечують автоматичне транскрибування англійського тексту, синхронізацію з відео та переклад українською із збереженням смислового та емоційного навантаження [13].

Фінальний монтаж планувалося здійснити у DaVinci Resolve – функціональному відеоредакторі, що має безкоштовну версію та забезпечує професійну кольорокорекцію, роботу зі звуком, ефектами, субтитрами й високоякісним експортом. Його перевага полягає в гнучкому таймлайні, можливості точної синхронізації й професійному контролі фінальної збірки [14].

В таблиці 2.1 наведено вибрані технології, сервіси та інструменти для створення та реалізації даного проєкту.

Таблиця 2.1 – Вибрані інструменти III та їх функціональне призначення

Інструмент/сервіс	Призначення	Тип технології
RunwayML (Gen–4)	Генерація відео з тексту	Text–to–Video (Diffusion)
ElevenLabs / PlayHT	Генерація синтетичного голосу	Speech Synthesis (TTS) [17]
Soundraw / AIVA	Генерація фонові музики	AI Composition
DeepL, Descript	Субтитри, переклад, розпізнавання	ASR + NLP
DaVinci Resolve	Фінальний відеомонтаж	NLE (Non–linear editing)

Як видно з таблиці, кожен із обраних інструментів виконує чітко визначену функцію в рамках проєкту, формуючи злагоджену екосистему для створення відео. Такий підхід дозволяє оптимізувати робочий процес, забезпечити креативну гнучкість і досягти високої якості без потреби у традиційному кіновиробництві.

2.3 Загальна архітектура та етапи реалізації відеопроєкту

Процес створення короткометражного відео «The Last Letter from Paris» охоплює кілька етапів, які логічно послідовні й взаємопов'язані. Кожен етап реалізовано з використанням відповідних інструментів на базі штучного інтелекту, що забезпечує повну автономність проєкту від традиційного знімального процесу.

На першому етапі формується сценарна структура відео. Вона включає послідовність подій (експозиція, кульмінація, розв'язка), емоційний настрій сцен, їхня візуальна атмосфера. Визначається, скільки фрагментів потрібно створити та яка їхня тривалість. Для кожної сцени прописуються текстові підказки (prompts) англійською мовою, які описують візуальний зміст: розташування персонажів, стиль зображення, настроїв, погода, час доби тощо.

Далі, на другому етапі, виконується генерація відеофрагментів за допомогою моделі Gen-4 на платформі RunwayML. Кожен текстовий запит опрацьовується як інструкція для створення сцени тривалістю до 10 секунд. Згенеровані відео зберігаються окремими файлами у високій якості (наприклад, 1080p, 24fps) для подальшого монтажу [4].

Після цього реалізується озвучення. Підготовлений сценарний текст монологу героїні розбивається на репліки, які озвучуються за допомогою синтетичного голосу через інструменти на зразок ElevenLabs або PlayHT. Кожен аудіофайл відповідає окремій сцені. В процесі генерації враховується емоційний тембр голосу, інтонація, паузи, що забезпечує більш виразне звучання [5].

Наступним етапом є генерація фонові музики. Залежно від сцени, добирається відповідний настрій: сум, напруження, спокій. Через сервіси AIVA або Soundraw створюється музичний супровід із урахуванням заданої тривалості, динаміки, інструментального складу. У деяких сценах додаються звуки оточення (дощ, міський шум), які також генеруються або підбираються з бібліотек [12].

Окремим важливим етапом є створення субтитрів. Згенерований голос обробляється через систему розпізнавання мовлення (наприклад, у Descript або Karwing) із подальшим перекладом українською мовою. Отримані субтитри редагуються вручну та синхронізуються з аудіо, після чого експортуються у форматі .SRT або .ASS [6].

Фінальним етапом є збірка всіх компонентів у DaVinci Resolve. Усі відеофрагменти розташовуються на часовій шкалі відповідно до сценарію, синхронізуються з озвученням та музикою, додаються субтитри. Здійснюється кольорокорекція, стилізація сцен, додавання переходів і візуальних ефектів. Завдяки професійним інструментам монтажу забезпечується плавність, гармонійність та художня завершеність проєкту [14]. На рисунку 2.1 можна побачити загальну архітектуру створення даного проєктного відео.

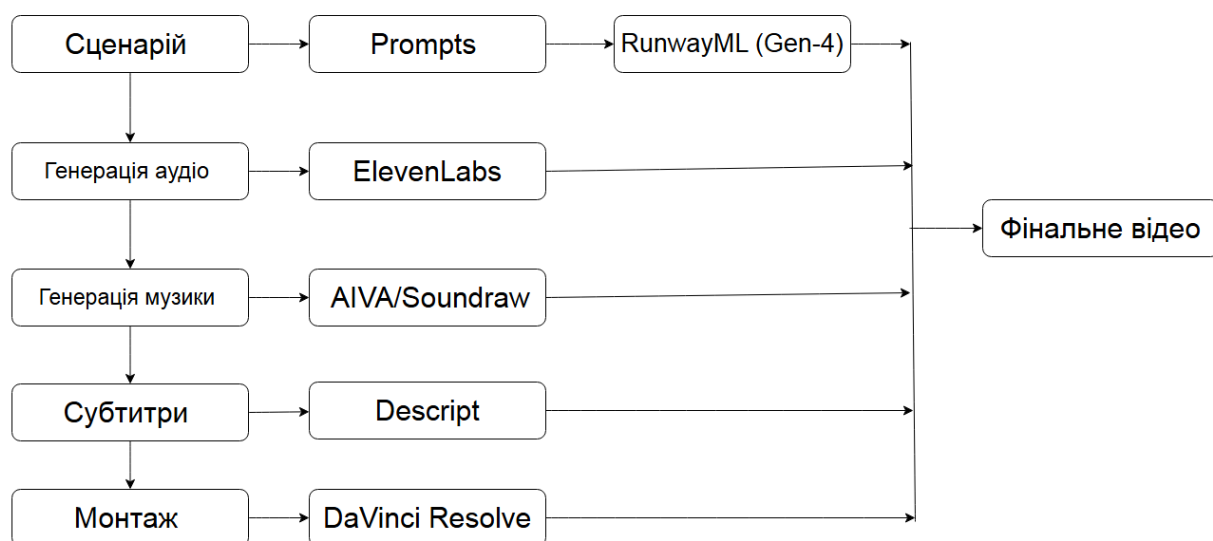


Рисунок 2.1 – Загальна архітектура створення відео з використанням ШІ

Варто зазначити, що кожен з етапів створення відео взаємопов'язаний і вимагає ретельної підготовки – від точного формулювання текстових підказок до фінального зведення усіх елементів у єдине відео. Узгодженість між візуальним рядом, аудіо та субтитрами забезпечує високу якість сприйняття, а

також дозволяє досягти цілісності наративу. Нижче наведено узагальнену схему реалізації проєкту.

2.4 Порівняльний аналіз альтернативних інструментів для реалізації відеопроєкту

У процесі розробки проєкту було обрано конкретний набір інструментів штучного інтелекту: RunwayML (для генерації відео), ElevenLabs (для озвучення), AIVA / Soundraw (для музики), DeepL + Descript (для субтитрів) та DaVinci Resolve (для фінального монтажу). Проте ринок ШІ-продуктів стрімко розвивається, й існує багато альтернатив, які можна було б використати для реалізації аналогічного проєкту. Нижче наведено детальний порівняльний аналіз альтернативних сервісів за кожним функціональним блоком.

Генерація відео з текстових описів – одна з найінтенсивніше досліджуваних сфер. Моделі на кшталт Gen-4, Sora та Pika демонструють різні підходи до генерації сцен, як за якістю, так і за швидкістю. У таблиці 2.2 нижче наведено порівняння найпоширеніших платформ, які можуть бути використані для генерації художнього відео.

Таблиця 2.2 – Порівняння ШІ-інструментів для створення відео

Сервіс	Модель	Особливості	Переваги	Обмеження
RunwayML	Gen-4	Висока стабільність, 10-секундні сцени	Зручний інтерфейс, стабільний вивід	Не підтримує довгі сцени
Pika Labs	Pika 1.0	Фокус на швидкості, генерація коротких кліпів	Простота, інтеграція з Discord	Менша деталізація зображень
Sora (OpenAI)	Sora	Генерація складних сцен до 1 хвилини	Якість, динаміка, деталі	Обмежений доступ, тестовий режим [9]
Synthesia	Пропрієтарна	Фокус на AI-аватарах для освітнього або бізнес контенту	Швидке створення презентацій	Менш художній стиль

Музика створена штучним інтелектом активно використовується у маркетингу, кіно й навчальних відео. Генеративні сервіси, такі як AIVA, Soundraw чи Ecrett, дозволяють створювати інструментальні композиції, адаптовані до настрою сцени. У таблиці 2.3 наведено порівняння їхніх основних характеристик.

Таблиця 2.3 – Порівняння ШІ-інструментів для створення музики

Сервіс	Алгоритм	Призначення	Переваги	Обмеження
AIVA	GAN + класичні правила	Симфонічна/емоційна музика	Ідеально підходить для фільмів	Менше опцій для електроніки [19]
Soundraw	Автоматизований шаблон	Сценічна, електронна, поп	Вибір стилю, настрою	Обмежена свобода редагування
Ecrett Music	Імовірнісна генерація	Швидкий трек на задану тему	Інтуїтивність	Прості композиції

Створення субтитрів – важливий етап забезпечення доступності контенту. Сервіси, що поєднують ASR та переклад, значно спрощують процес. У таблиці 2.4 подано короткий огляд найпоширеніших рішень.

Таблиця 2.4 – Порівняння ШІ-інструментів для створення субтитрів

Сервіс	Можливості	Переваги	Обмеження
Descript	Транскрипція, субтитри, редагування	Професійне середовище, точність	Доступність певних функцій лише у Pro версії [21]
Veed.io	Субтитри, переклад, AI-редактор	Зручне онлайн-редагування	Обмеження у безкоштовній версії
Kapwing	Генерація субтитрів + переклад	Відкритість, шаблони, креативність	Менш точна синхронізація
DeepL	Точний переклад текстів	Висока лексична адекватність	Не створює субтитри, лише переклад [6]

У завершальному етапі реалізації відео застосовуються редактори, які підтримують багат шаровий монтаж, синхронізацію аудіо, субтитрів і візуальних ефектів. У таблиці 2.5 порівнюються інструменти, які найбільше підходять для різних форматів відео.

Таблиця 2.5 – Порівняння програм для відеомонтажу

Редактор	Платформа	Призначення	Переваги	Обмеження
DaVinci Resolve	Windows/macOS/Linux	Професійний монтаж	Кольорокорекція, субтитри, безкоштовно	Вища крива навчання [14]
Adobe Premiere Pro	Windows/macOS	Індустріальний стандарт	Плагіни, інтеграція з АЕ	Платна, вища системна вимогливість
CapCut Desktop/Web	Web/IOS/Android	Легкий монтаж для соцмереж	Простота, автогенерація субтитрів	Обмежена точність на рівні студії

Вибір конкретного набору інструментів для реалізації відеопроєкту «The Last Letter from Paris» був обґрунтований необхідністю досягти оптимального балансу між якістю результату, швидкістю реалізації, доступністю та можливістю креативного контролю. Аналіз альтернатив підтверджує, що сучасні ШІ-платформи пропонують широкий спектр функцій – від базової генерації до глибокого налаштування стилю, емоційності й темпоритму. Водночас жоден окремий сервіс не є універсальним, і ефективне поєднання декількох спеціалізованих рішень дає кращий результат, ніж спроба використати одну платформу для всього циклу.

Таким чином, здійснене порівняння дозволяє не лише обґрунтувати прийняті проєктні рішення, а й сформувану потенційну базу для майбутніх досліджень. У випадку масштабування проєкту або його адаптації до інших сфер – освіти, комерційної реклами, відеоблогінгу чи соціальних кампаній – вибір альтернативних або гібридних комбінацій сервісів може бути ключовим фактором ефективності. З огляду на це, далі доцільно проаналізувати ризики та етичні обмеження, пов’язані з інтеграцією ШІ у творчий процес, що буде розглянуто у наступному підрозділі.

2.5 Ризики, обмеження та етичні аспекти використання ШІ у відеомонтажі

Хоча використання технологій штучного інтелекту у відеомонтажі відкриває широкі можливості для автоматизації, творчого експерименту й економії ресурсів, воно також супроводжується низкою ризиків, технічних обмежень і етичних викликів. Їхнє врахування є необхідною умовою для безпечної, відповідальної та правомірної інтеграції ШІ у творчі процеси.

Один з головних ризиків – неконтрольованість результату генеративних моделей. Наприклад, відео, створене з текстового опису, може містити артефакти, неочікувані інтерпретації або візуальні елементи, які суперечать задуму. Це обумовлено внутрішньою непрозорістю навчальних даних та ймовірнісним характером генерації у таких моделях, як Gen-4 або Sora [9]. Аналогічно, синтетичний голос, згенерований, наприклад, у ElevenLabs, може іноді мати інтонаційні зміщення або паузи, які впливають на емоційне сприйняття [5].

Крім технічних обмежень, важливим є питання етики використання ШІ в роботі з персональними ознаками – зокрема, голосами. Використання голосів, що імітують реальних людей, навіть без порушення авторських прав, може спричинити етичні дилеми, пов'язані з маніпуляцією емоцією або створенням фейкових наративів. Саме тому сервіси як ElevenLabs або Resemble AI вимагають підтвердження прав на використання кастомних голосів [18].

Ще один виклик – правова невизначеність щодо авторства результатів, створених ШІ. У більшості країн правова база не передбачає чіткої системи авторських прав на контент, повністю створений штучним інтелектом. Це особливо актуально для відео, де всі візуальні, звукові та текстові компоненти згенеровані без прямої участі людини як автора. За відсутності правових норм виникають ситуації, коли неможливо захистити чи передати права на згенерований твір.

Важливо також зазначити, що деякі генеративні інструменти можуть мати приховану упередженість, вбудовану в їхні моделі. Наприклад, тренування на переважно англomовних або західноцентричних датасетах може призвести до недостатнього відображення мультикультурних елементів, стереотипізації персонажів або стилів. Це потребує уважного підходу при створенні наративів у контексті різних культур, щоб уникнути спотворення реальності або образливих інтерпретацій [24].

З етичного погляду викликає питання і використання ШІ в емоційно чутливих контекстах – наприклад, у відео, що стосуються тем втрати, депресії чи особистої трагедії. Глядач може не відрізнити, що історія, голос чи сцени є повністю згенерованими, що відкриває простір для маніпуляцій. Тому важливо чітко маркувати контент, створений із застосуванням ШІ, та зберігати прозорість перед аудиторією [25].

У процесі планування та реалізації відеопроєкту «The Last Letter from Paris» всі етапи були спроектовані з урахуванням зазначених обмежень. Озвучення не імітувало реальний голос, всі інструменти використовувались відповідно до ліцензій, а стиль сценарію не передбачав соціально чутливих тем. Окрему увагу було приділено естетичній та етичній нейтральності сцен, щоб уникнути неоднозначного трактування з боку глядача.

2.6 Оцінка технічних вимог та ресурсів для реалізації

Реалізація відеопроєкту з використанням інструментів штучного інтелекту вимагає не лише креативного підходу, але й наявності відповідних технічних ресурсів, які забезпечують стабільну та ефективну роботу з сучасними ШІ-сервісами. Вибір інструментів ШІ для відеомонтажу базувався на аналізі компромісів між якістю, доступністю та швидкістю, подібно до методів підтримки прийняття рішень у розробці програмного забезпечення [26]. У цьому підрозділі здійснено оцінку апаратного забезпечення, програмного

середовища та інтернет-ресурсів, необхідних для успішного впровадження проєкту «The Last Letter from Paris».

Більшість генеративних інструментів, задіяних у проєкті (RunwayML, ElevenLabs, AIVA, Soundraw, Descript), працюють у хмарному середовищі, тому основним технічним фактором є стабільне високошвидкісне інтернет-з'єднання. Мінімально рекомендована швидкість – 20 Мбіт/с для завантаження і не менше 5 Мбіт/с для вивантаження даних, що особливо важливо при роботі з відеофайлами великого розміру (від 100 до 500 МБ на кожен фрагмент). Крім того, браузер має підтримувати сучасні вебтехнології (Chrome 90+, Firefox 100+, Edge Chromium) для коректної взаємодії з платформами [26].

З точки зору апаратного забезпечення, критично важливим є етап фінального монтажу у DaVinci Resolve, який є ресурсомістким програмним забезпеченням. Для комфортної роботи рекомендується комп'ютер із процесором Intel Core i7 / AMD Ryzen 7 або кращим, 16–32 ГБ оперативної пам'яті, SSD-диском і відеокартою не нижче NVIDIA GTX 1060 з 4–6 ГБ VRAM. DaVinci Resolve активно використовує апаратне прискорення GPU, тому для плавного відтворення відео, обробки кольору й рендерингу потрібна дискретна графіка [14]. Робота на системах із нижчими параметрами можлива, однак значно знижує продуктивність і час експорту.

У таблиці 2.6 наведено орієнтовну оцінку технічних вимог для кожного з ключових етапів.

Таблиця 2.6 – Оцінка технічних вимог етапів реалізації

Етап реалізації	Мінімальні вимоги	Рекомендовані вимоги
Генерація відео (RunwayML)	Інтернет 10 Мбіт/с, браузер Chrome	Інтернет 30+ Мбіт/с, SSD, RAM 8 ГБ
Генерація голосу (ElevenLabs)	Інтернет, підтримка Web Audio API	RAM 4–8 ГБ, браузер Chrome/Edge
Музика, субтитри (Soundraw, Descript)	Будь-який сучасний ПК або ноутбук	RAM 8 ГБ, стабільний інтернет
Фінальний монтаж (DaVinci Resolve)	CPU i5, RAM 8 ГБ, відеокарта 2 ГБ VRAM	CPU i7, RAM 16–32 ГБ, GPU 4–8 ГБ VRAM

Додатково слід враховувати місце на диску. Середній обсяг проєкту з 7–10 фрагментами, звуком і субтитрами в форматі 1080p може сягати 1,5–2 ГБ до експорту. Рендер готового відео займає від 3 до 15 хвилин залежно від системи.

Реалізація подібного проєкту не вимагає надпотужного обладнання – за умови використання хмарних ІІІ–сервісів. Однак на етапі монтажу доцільно мати середньопродуктивну систему або ноутбук, що забезпечує комфортну роботу з відео у Full HD–форматі. Це дає змогу зосередитись на творчості без технічних обмежень.

2.7 Варіанти масштабування та адаптація методу

Побудований у проєкті робочий процес може бути масштабований або адаптований під інші тематичні, навчальні чи комерційні завдання. Більш того, описаний підхід має універсальний характер, оскільки не вимагає спеціалізованого обладнання, акторського складу чи студійних ресурсів. Це відкриває широкі перспективи його впровадження в суміжні сфери [30].

Одним із напрямів масштабування є створення освітнього відеоконтенту. Наприклад, викладачі або студенти можуть створювати навчальні мініфільми, інфографічні ролики чи історичні реконструкції з голосовим супроводом і візуальними сценами, згенерованими за сценарієм. Університети та онлайн–платформи (Coursera, EdX, Prometheus) уже використовують анімовані лекції, які можуть бути значно здешевлені та пришвидшені завдяки генеративним інструментам [30].

Інша потенційна сфера – відеомаркетинг та соціальні медіа. Маркетологи можуть створювати персоналізовані рекламні ролики або сторітелінгові відео на основі описів продуктів чи брендів. Завдяки AI–озвученню та автоматичному субтитруванню можливе швидке створення локалізованих відео для різних ринків. Наприклад, платформи на кшталт Pictory чи Lumen5 уже впроваджують подібні функції для підприємств малого бізнесу [31].

Метод також може бути адаптований у сфері документального відео, блогінгу та контенту для YouTube і TikTok. Блогери можуть генерувати унікальні візуальні історії або візуалізувати свої подкасти без потреби у камері чи студії. Особливо це актуально для початківців, які обмежені у ресурсах. Набір інструментів, описаних у цьому проєкті, є ідеальною основою для створення відео за моделлю «один автор – повний цикл».

На рівні командної роботи масштабування можливе через розділення процесів між різними користувачами. Один спеціаліст може відповідати за генерацію відео, інший – за текст, озвучення або монтаж. У таких випадках стає актуальним створення шаблонів сценаріїв, стилістичних промптів, а також повторне використання напрацьованих голосових моделей і аудіотреків.

Варто також зазначити можливість інтеграції автоматизованих процесів через API або спеціальні інструменти, які підтримують пакетну генерацію, наприклад, Runway API або PlayHT API. Це дозволяє організувати масове створення відео, наприклад, для новинних дайджестів, оглядів або освітніх серій із мінімальним втручанням людини [33].

Таким чином, розроблений у межах проєкту підхід може служити модульною платформою, яку легко адаптувати під нові задачі, змінювати масштаб, додаючи нові компоненти або автоматизуючи окремі етапи. Це робить його актуальним не лише для індивідуальних творців, а й для командних, комерційних і освітніх ініціатив.

2.8 Висновок до другого розділу

У другому розділі кваліфікаційної роботи було детально сформульовано концепцію реалізації експериментального відеопроєкту «The Last Letter from Paris», який передбачає комплексне використання сучасних технологій штучного інтелекту на всіх етапах відеовиробництва. Було обґрунтовано вибір інструментів і моделей ШІ для генерації відео (RunwayML Gen-4), озвучення (ElevenLabs), створення музики (Soundraw, AIVA), генерації субтитрів (Descript, DeepL) та фінального монтажу (DaVinci Resolve). На основі цих інструментів побудовано архітектуру проєкту, описано логіку реалізації й технічні аспекти синхронізації мультимедійних компонентів.

Особливу увагу було приділено порівняльному аналізу альтернативних сервісів, що дало змогу оцінити сильні та слабкі сторони кожного інструмента. Це сприяло більш усвідомленому підходу до вибору технологій у відповідності до вимог проєкту. Також розглянуто ризики та етичні аспекти використання ШІ у творчій роботі: від правової невизначеності щодо авторства згенерованого контенту до можливих маніпуляцій емоціями через синтетичні голоси. Аналіз показав, що навіть на етапі проєктування слід враховувати не лише технічні, а й соціальні та етичні наслідки.

РОЗДІЛ 3. СТВОРЕННЯ ВІДЕО З ВИКОРИСТАННЯМ ШІ–ТЕХНОЛОГІЙ

3.1 Формування сценарного плану та підказок для генерації відео

Першим етапом реалізації відеопроєкту «The Last Letter from Paris» стала розробка сценарного плану, який слугував основою для створення сюжетних сцен із використанням генеративних моделей. Сценарій був побудований у формі монологу головної героїні, яка емоційно згадує коханого, що зник, залишивши лише один лист. Такий формат дозволив зробити відео емоційно наповненим, ліричним і кінематографічним, що добре поєднується з можливостями сучасних ШІ–моделей для візуальної розповіді сюжету [15].

Сценарій було умовно розбито на 5 логічних сцен, що показані в таблиці 3.1, кожна з яких мала свою емоційну домінанту:

Таблиця 3.1 – Логічні сцени для створення відео

№	Сцена	Настрій/Емоція
1	Вигляд на Париж	Надія
2	Самотність під дощем	Меланхолія
3	Спогади про щастя	Ностальгія
4	Читання листа	Смуток, тепло
5	Прийняття та прощання	Спокій, завершеність

Для кожної сцени було складено текстовий опис (prompt), який передавав візуальне й емоційне навантаження епізоду, що показано на рисунку 3.1. Наприклад:

- «A young woman stands alone under a grey sky, Paris streets wet from the rain. Cinematic lighting, soft depth of field, dramatic tone.»
- «A vintage wooden desk with a handwritten letter on it. Candlelight flickers in the background, Paris window view.»
- «Close-up of a tear rolling down her cheek as she reads the letter. Warm lighting, slow motion.»

Ці підказки враховували ключові параметри для генеративної моделі Gen-4: освітлення, композиція, емоційність, час доби, колірна гама та об'єкти в кадрі. У результаті RunwayML інтерпретував такі підказки з урахуванням вбудованих тренувальних даних і створював послідовні фрагменти, які логічно поєднувалися в єдину історію [4].

Було враховано загальні рекомендації з формулювання промптів для генеративного відео, зокрема: уникати неоднозначних фраз, конкретизувати дієслова та уникати описів, які містять протиріччя (наприклад, “bright night”). Врахування цих рекомендацій дало змогу покращити якість виводу і скоротити кількість невдалих генерацій [16].

Усі промпти були складені англійською мовою, оскільки RunwayML Gen-4 підтримує лише англійські запити. Це також дало змогу краще контролювати стилістику сцен і уникнути автоматичних перекручень при перекладі.

3.2 Генерація відеофрагментів у RunwayML (Gen-4)

Після створення сценарної структури й формулювання текстових підказок розпочався процес генерації відеофрагментів за допомогою моделі Gen-4, доступної на платформі RunwayML. Це одна з найбільш передових diffusion-моделей, здатна створювати відео з коротких описів (text-to-video) з вражаючим рівнем стилістичної та композиційної точності. На момент створення відео Gen-4 була останньою версією, яка підтримувала покадрову плавність, рух камери, кінематографічну композицію та глибину сцени [4].

Інтерфейс RunwayML є інтуїтивно зрозумілим: користувач вводить текстову підказку (prompt), обирає тривалість відео (до 10 секунд), бажаний стиль або режим анімації, а також, за потреби, завантажує зображення для стилізації. Для даного проєкту використовувались суто текстові запити, що робить експеримент показовим саме для сценарного генеративного відео.

Кожен запит опрацьовувався системою впродовж 1–3 хвилин. Після обробки модель створювала короткий відеофайл у форматі .mp4 (H.264), зазвичай з частотою 24 кадри на секунду та роздільною здатністю 1024×576 або 1280×720. Для подальшого монтажу відео було масштабоване до 1080p у DaVinci Resolve.

Більшість сцен генерувались з першого або другого запиту. У деяких випадках (наприклад, з емоційними виразами обличчя чи точним фоном Парижа) доводилося переформулювати prompt, додавати уточнення, прибирати суперечності. Згідно з порадами PromptHero, найкращий результат дає короткий, конкретний опис без надмірного ускладнення [16].

На рисунку 3.1 показано створення відеофрагмента за допомогою чітко продуманого промπτу.

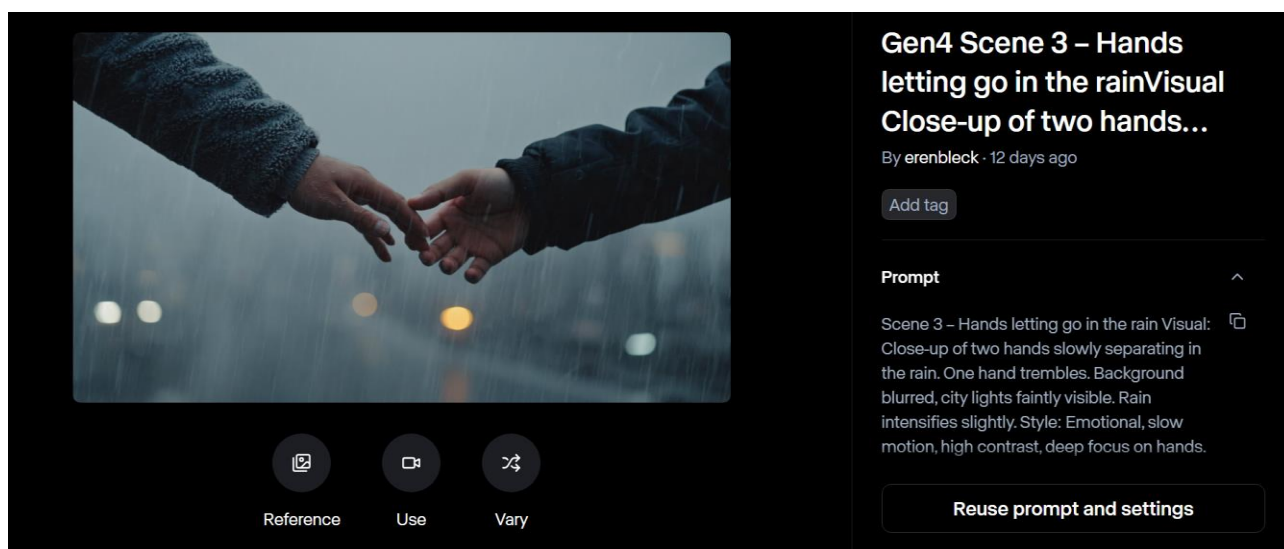


Рисунок 3.1 – Створення відеоролику в RunwayML з використанням промπτу

Підсумково було створено 13 відеофрагментів, які повністю покривають сценарну структуру. Усі вони були збережені у локальному сховищі для подальшої обробки, без необхідності ручного редагування кожного кадру, що значно пришвидшило процес виробництва.

3.3 Озвучення голосом головної героїні за допомогою ElevenLabs

Озвучення є однією з ключових складових у створенні емоційного відеоконтенту. У короткометражному фільмі «The Last Letter from Paris» головна героїня виступає оповідачкою, яка через монолог передає особисту історію втрати, любові та спогадів. Для досягнення необхідного емоційного ефекту було вирішено використати штучно згенерований голос, що дозволяє передати заданий настрій без участі реальної акторки.

Для озвучення було обрано сервіс ElevenLabs – одну з найсучасніших платформ штучного інтелекту для синтезу мовлення. Сервіс підтримує багатомовні моделі, детальне налаштування емоційного тону та має широкий вибір голосів. У проєкті було використано голос «Rachel» на базі моделі Eleven Multilingual v2, яка дозволяє відтворювати англomовний текст із високим рівнем натуралізму [5].

Оповідь складалась із 7 основних сцен, кожна з яких мала відповідний текстовий блок (монолог). Для кожної частини було згенеровано окремий аудіофайл, з урахуванням наступних параметрів (як показано на рис. 3.3):

- Stability (Стабільність) – близько 85%: забезпечує плавність мовлення без різких змін тембру;
- Style Exaggeration (Емоційність) – 40%: дозволяє підкреслити ліричність і драматизм;
- Speed (Швидкість мовлення) – збережено на стандартному рівні;
- Speaker Boost – активовано: надає більшої чіткості та гучності голосу.

У результаті було отримано емоційно глибокий, м'який жіночий голос, який точно передавав настрій відео. Наприклад, фраза «You spilled your coffee... I spilled my heart.» прозвучала із характерною паузою та інтонацією жалю, що неможливо досягти у звичайному TTS без стилістичного налаштування.

На рисунку 3.2 показано робота та налаштування голосу головної героїні та текст озвучення з інтонаційними паузами і знаками.

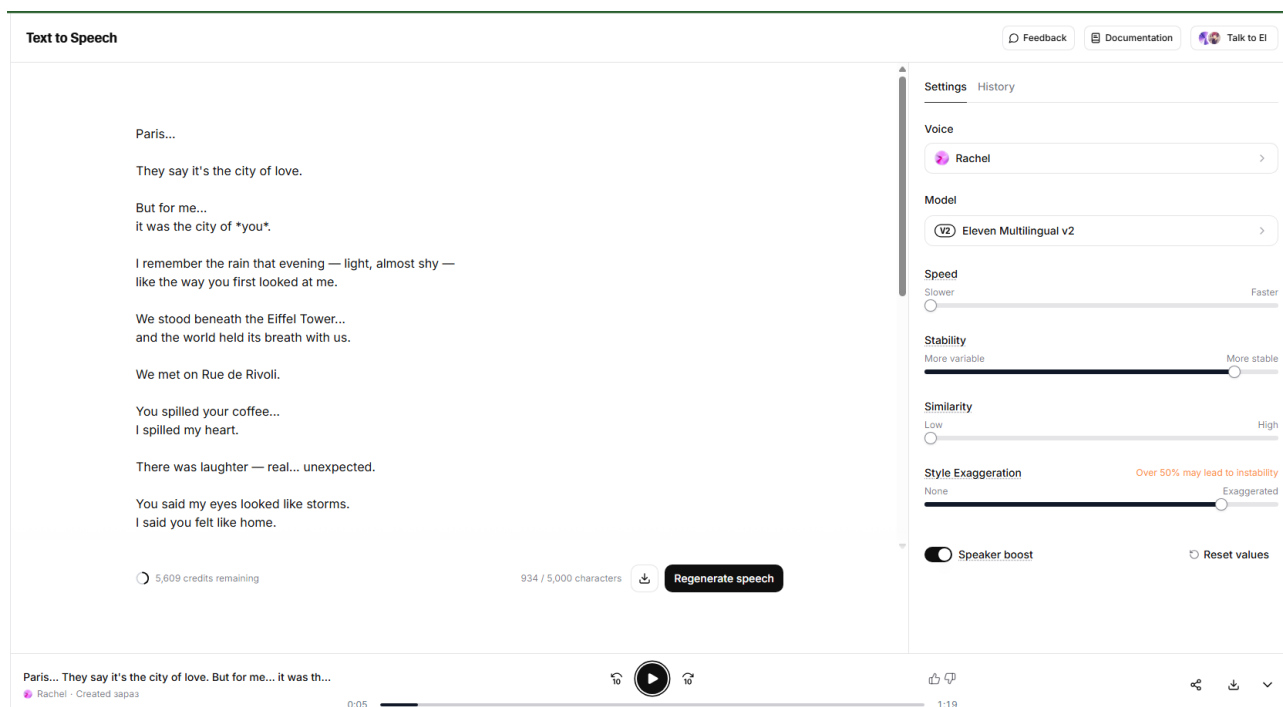


Рисунок 3.2 – Генерування голосу героїні в сервісі ElevenLabs

Готові звукові файли експортувались у форматі .mp3 та були синхронізовані з відповідними сценами під час фінального монтажу у DaVinci Resolve. Службовий текст не потребував жодної корекції вручну, оскільки точність вимови та інтонація були високого рівня.

3.4 Генерація фоновій музики та звуків оточення за допомогою ШІ

Музичний супровід відіграє вирішальну роль у формуванні емоційного сприйняття відео. У проєкті музика повинна була не лише доповнити візуальні образи, а й посилити настрій сцен – меланхолію, спогади, напруження або спокій. Враховуючи обмеження щодо авторських прав і бажання забезпечити повну автономність процесу, для створення звукового супроводу було вирішено скористатися генеративними AI-сервісами.

Для реалізації музичної частини використовувалися платформи Soundraw та AIVA:

Soundraw.io – сервіс, який дозволяє створювати музику на основі вибору жанру, настрою, темпу та інструментального складу. Користувач може вказати

параметри, і ШІ створює кілька варіантів треків, які можна редагувати за тривалістю та структурою [18].

AIVA (Artificial Intelligence Virtual Artist) – система, орієнтована на кінематографічну та симфонічну музику, яка генерує мелодії відповідно до сценарної структури або емоційного опису сцени. Інтерфейс дозволяє обирати стиль (наприклад, Romantic, Ambient, Sad Piano) та темп, а також редагувати партитуру вручну [19].

Процес відбувався так:

1. Для кожної сцени описувався бажаний настрій: наприклад, «calm sadness with piano and strings» або «soft ambient with rain undertone».
2. На основі цих описів було згенеровано кілька варіантів музики, які прослуховувалися і відбиралися вручну.
3. Обрані фрагменти експортувались у форматі .mp3 або .wav і синхронізувались з відео.

Окремо було створено атмосферні звуки дощу, які використовувалися у перехідних сценах. Ці звуки або підбирались з бібліотек AI-платформ, або генерувались із вказаним параметром («light rain in Paris street at night»). Звуковий ефект допоміг підкреслити настрій самотності й очікування.

Комбінування фонового треку та атмосферних шумів дозволило створити глибину, характерну для кінематографа. Усе це додавало реалізму сценам і допомагало глядачу емоційно зануритись у сюжет.

3.5 Створення субтитрів за допомогою ШІ

Оскільки озвучення відео «The Last Letter from Paris» було виконане англійською мовою, виникла потреба забезпечити доступність для україномовної аудиторії. Для цього було створено субтитри з автоматичним перекладом. Цей процес було реалізовано з використанням інструментів на базі штучного інтелекту, які поєднують розпізнавання мовлення (ASR), транскрибування, переклад та синхронізацію з відеорядом.

На першому етапі було використано Descript – ефективну AI-платформу для транскрибування аудіо та відео. Вона підтримує автоматичне розпізнавання англomовного мовлення й створює точний текстовий скрипт з розбивкою по таймінгах [21]. Аудіофайли озвучення були завантажені в систему, яка упродовж кількох секунд формувала транскрипцію, зручно подану для редагування.

Другим етапом став переклад транскрипту на українську мову. Для цього застосовувався сервіс DeepL Translator, який відзначається високою точністю перекладу художнього тексту, з урахуванням контексту та емоційного забарвлення [6]. Після автоматичного перекладу всі фрази були перевірені вручну, щоб зберегти логіку, ритм та інтонаційну відповідність оригіналу.

Отримані українські репліки були вставлені у шаблон субтитрів у форматі .SRT. Для цього використовувались інструменти Karwing або вбудовані редактори в Descript, які дозволяють синхронізувати субтитри з аудіо – як автоматично, так і з ручною корекцією таймінгів [22].

Особливу увагу приділялось:

- читабельності тексту (довжина фраз не перевищувала 2 рядків);
- контрастності й стилю (білі субтитри з легким тіншовим контуром на затемненому фоні);
- часу появи (1,5–4 секунди на екран, залежно від довжини репліки).

На рисунку 3.3 показано роботу в сервісі Descript для налаштування та перекладу субтитрів в проєкт.

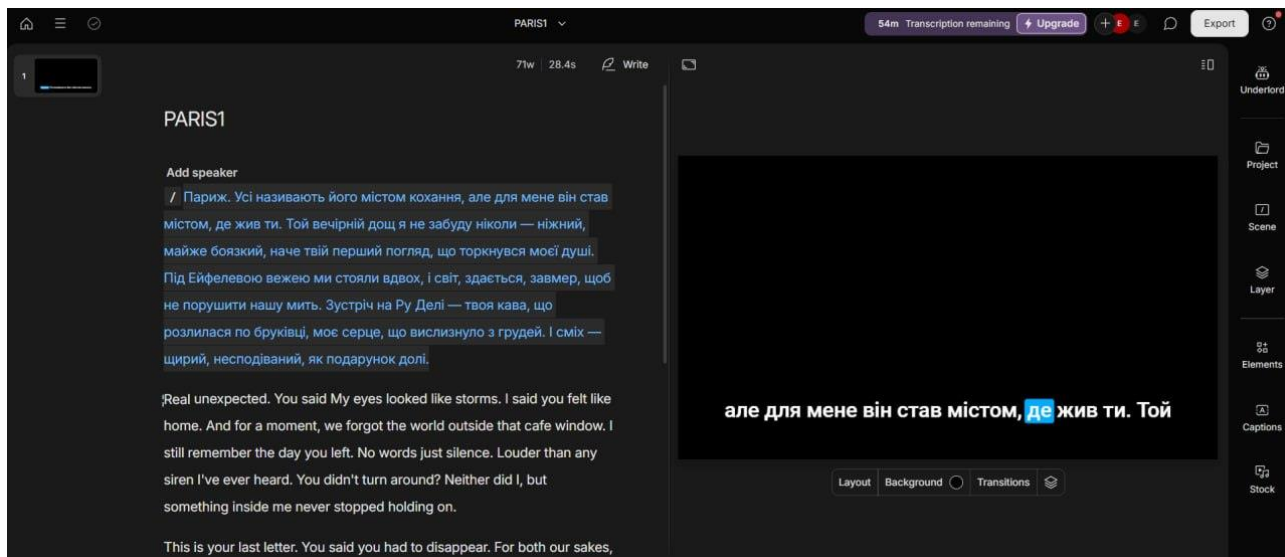


Рисунок 3.3 – Створення та налаштування українських субтитрів

Фінальні субтитри були інтегровані у відео під час монтажу у DaVinci Resolve, після чого перевірені на синхронність і відповідність емоційному ритму відео. Таким чином, завдяки ШІ-сервісам вдалося оперативно та якісно реалізувати двомовну доступність, що значно підвищує сприйняття відео широкою аудиторією.

3.6 Фінальний монтаж у DaVinci Resolve

Останнім етапом реалізації відеопроєкту «The Last Letter from Paris» став фінальний монтаж, виконаний у професійному відеоредакторі DaVinci Resolve. Це програмне забезпечення було обране завдяки своїм розширеним можливостям для роботи з відео, аудіо, субтитрами та кольорокорекцією, а також через наявність повністю функціональної безкоштовної версії [14].

Процес монтажу розпочався з імпорту згенерованих відеофрагментів, які були створені у RunwayML (Gen-4). Всі 13 сцен були розміщені у монтажному таймлайні відповідно до структури сценарію. Для досягнення плавності переходів між сценами використовувалися fade-in / fade-out ефекти, а також крос-розчинення (cross dissolve) у місцях емоційного акценту.

Наступним кроком стало додавання озвучення, згенерованого у ElevenLabs. Аудіофайли були синхронізовані з відповідними сценами, у деяких місцях відредаговано гучність і темп, щоб досягти ідеального поєднання з відеорядом. У сценах із внутрішнім напруженням застосовувався легкий ревербераційний ефект, що додав глибини.

Після цього було накладено фонову музику та звукові ефекти, створені у Soundraw та AIVA. Для кожної сцени використовувались окремі звукові шари, щоб створити збалансовану звукову композицію [28]:

- у сценах спогадів переважали струнні та фортепіанні акорди;
- у сценах самотності – фоновий дощ, легкий вітер, шум міста.

На наступному етапі у відео було додано українські субтитри у форматі .srt, які були попередньо згенеровані та синхронізовані. DaVinci Resolve підтримує імпорт субтитрів з можливістю редагування розміру, розміщення, кольору тексту, а також експорту відео з вбудованими субтитрами.

Крім того, було застосовано кольорокорекцію, щоб усі згенеровані сцени мали єдиний стиль. Зокрема:

- зменшено насиченість кольорів (для створення «приглушеної» атмосфери);
- підвищено контраст у ключових моментах (наприклад, під час прочитання листа);
- додано віньєтування (легке затемнення країв кадру) для створення кінематографічного ефекту.

Рисунок 3.4 наглядно показує роботу з даним матеріалом в програмі.

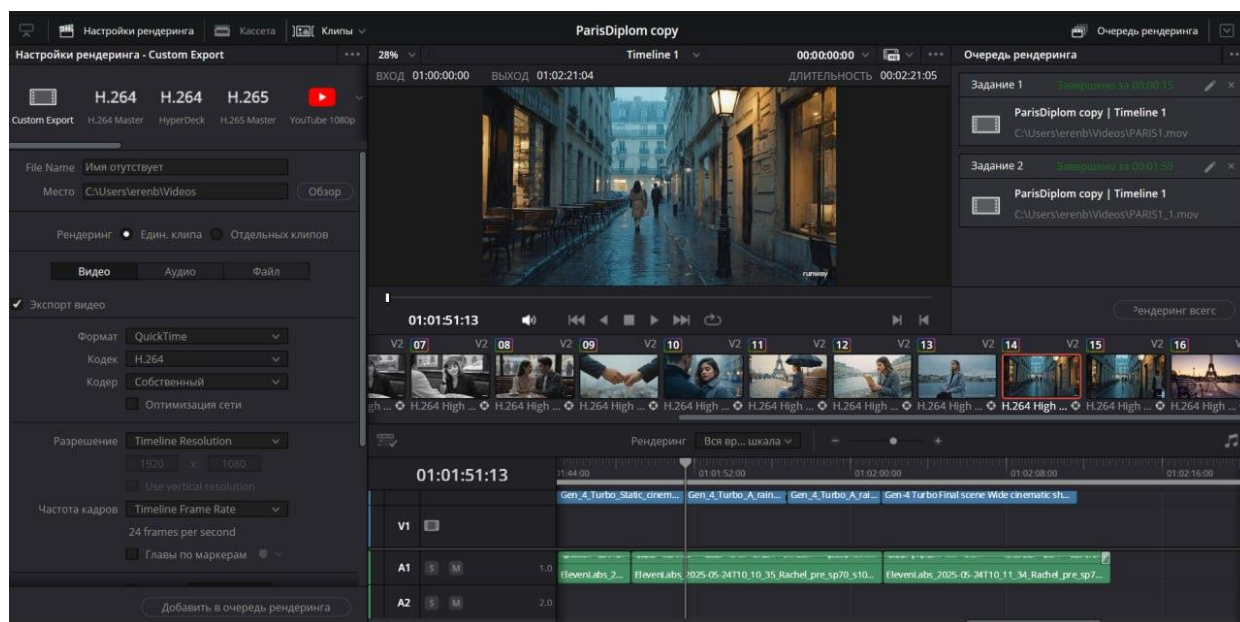


Рисунок 3.4 – Процес фінального монтування в програмі DaVinci Resolve

У результаті був сформований фінальний відеофайл тривалістю близько 1 хвилини 20 секунд у форматі .mp4, Full HD (1920×1080), 24 кадри на секунду. Відео було збережено з високим бітрейтом (15 Мбіт/с), що гарантує чіткість зображення навіть у темних сценах.

3.7 Тестування якості відео: візуальні, аудіо та емоційні параметри

Після завершення створення відеопроєкту «The Last Letter from Paris» важливою складовою стало оцінювання його якості – як з технічної, так і з емоційної точки зору. Тестування якості аудіо– та відеокомпонентів проводилося з урахуванням періодичних характеристик сигналів, подібних до моделювання циклічних процесів [20]. Також це охоплювало візуальні аспекти, якість звуку, узгодженість усіх елементів та сприйняття фінального продукту потенційною аудиторією. Для цього було сформовано просту методику оцінювання, що поєднує об'єктивні технічні критерії та суб'єктивне враження.

У межах технічного тестування було перевірено:

- чіткість відеоряду (відсутність розмиття або артефактів);
- стабільність кадрів, особливо в сценах з рухом;

- відповідність таймінгу озвучення і субтитрів;
- баланс гучності між голосом, музикою та звуковими ефектами;
- колірну гармонію між сценами (особливо після кольорокорекції в DaVinci Resolve).

Ці параметри оцінювались при перегляді відео на різних пристроях: ноутбук із FullHD-дисплеєм, мобільний телефон, а також у навушниках і через вбудовані динаміки. Для тестування використовувались такі формати експорту, як .mp4 (H.264) у якості 1080p, 24 fps, із середнім бітрейтом 10–15 Мбіт/с, що забезпечує оптимальний компроміс між якістю й вагою файлу.

Окрему увагу було приділено емоційному впливу відео. Для цього проєкт було протестовано в невеликій фокус-групі з 5 глядачів віком 20–30 років, які не брали участі в розробці. Глядачам запропонували переглянути фінальне відео та заповнити коротку анкету з 5 запитаннями:

1. Чи вдалося зрозуміти суть історії?
2. Наскільки добре передано емоційний настрій?
3. Як оцінюєте якість озвучення?
4. Чи було зручно сприймати субтитри?
5. Загальне враження від відео (оцінка за 10-бальною шкалою).

За результатами анкетування:

- 4 з 5 глядачів дали загальну оцінку 9 або 10 балів;
- 100% зазначили, що емоційна складова «відчутна» або «дуже сильна»;
- 80% визнали голос озвучення «природним»;
- 1 особа зауважила, що субтитри у 2 місцях з'являлись дещо пізно, це було виправлено.

Візуальна оцінка підтвердила, що генерація сцен за допомогою RunwayML Gen-4 була достатньо якісною, хоч і в кількох фрагментах помітні незначні спотворення на рівні текстур заднього фону. Також було підтверджено вдалий вибір музики й атмосферних звуків – шум дощу створював відповідний настрій, не перевантажуючи голос.

Виявлені дрібні зауваження були оперативно враховані, що дозволило досягти фінального результату, повністю відповідного задуму.

3.8 Висновок до третього розділу

В третьому розділі кваліфікаційної роботи було детально описано реалізацію експериментального відеопроєкту «The Last Letter from Paris», який став практичним прикладом застосування сучасних технологій штучного інтелекту в галузі відеомонтажу. Робота охоплювала всі ключові етапи створення відео: від формування сценарію та генерації відеофрагментів до озвучення, підбору музики, створення субтитрів і фінального зведення в DaVinci Resolve.

Під час створення візуального ряду було використано генеративну модель Gen-4 на платформі RunwayML, яка дозволила створити кінематографічні сцени лише на основі текстових описів. Озвучення героїні здійснено у сервісі ElevenLabs з налаштуванням емоційного тону, що забезпечило високу якість синтетичного голосу. Музичний супровід і атмосферні ефекти було згенеровано через Soundraw та AIVA, що забезпечило емоційне занурення у сюжет. Завдяки використанню Descript і DeepL, вдалося оперативно створити точні українські субтитри [29]. Усі компоненти були інтегровані в професійне середовище DaVinci Resolve, де відбулася остаточна обробка й кольорокорекція.

Реалізація цього відео довела, що за допомогою відкритих ШІ-інструментів можливо повністю реалізувати креативний аудіовізуальний продукт без участі знімальної групи, студійного обладнання чи акторів. Практична частина не лише підтвердила теоретичні припущення, а й продемонструвала реальні переваги ШІ у відеомонтажі: економію часу, гнучкість, доступність і високу якість результату.

РОЗДІЛ 4. БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1 Ризики психоемоційного вигорання користувачів ПК

У сучасному цифровому середовищі відеомонтаж із використанням штучного інтелекту (ШІ) став важливою складовою медіавиробництва. Програми, які використовують машинне навчання, комп'ютерний зір, автоматичне кадрування, розпізнавання обличчя та об'єктів, значно пришвидшують процес постпродакшену. Проте інтенсивне використання таких технологій супроводжується новими психоемоційними ризиками для фахівців. На рисунку 4.1 можна побачити основні чинники виникнення вигорання.

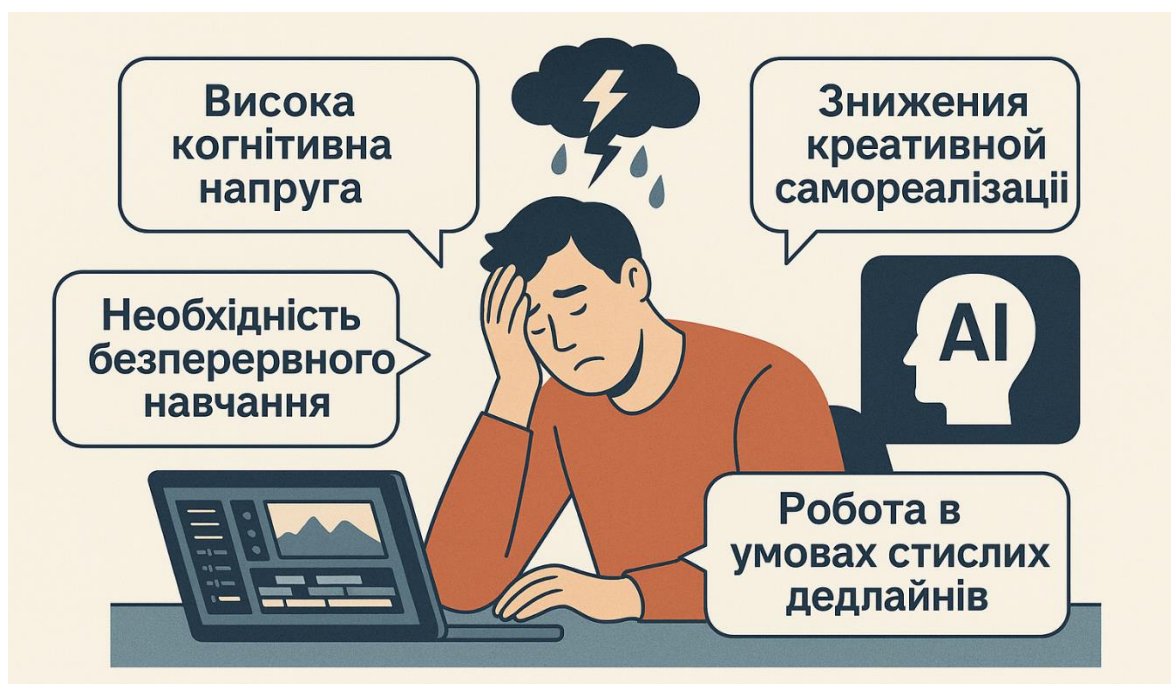


Рисунок 4.1 – Психоемоційне вигорання операторів ПК

Інформаційні технології стали невід'ємною частиною професійної діяльності більшості працівників. Особливо це стосується спеціалістів, які працюють з персональним комп'ютером (ПК) тривалий час. Постійна робота за комп'ютером без належного відпочинку та організації робочого процесу може призвести до розвитку психоемоційного вигорання.

Психоемоційне вигорання – це стан емоційного виснаження, що супроводжується втратою інтересу до роботи, зниженням мотивації та професійної ефективності. Основними чинниками, що сприяють виникненню такого стану в користувачів ПК, є тривала статична робота, високе розумове навантаження, багатоетапні завдання, необхідність багатозадачності та обмежена соціальна взаємодія.

Згідно з дослідженнями, одним із поширених джерел стресу є тривале перебування в одному положенні перед екраном комп'ютера без перерв. Це призводить не лише до фізичної втоми, а й до психологічного виснаження. При цьому користувач може не одразу помічати ознаки вигорання, що робить ситуацію ще більш небезпечною. Серед основних симптомів виділяють втрату концентрації, роздратованість, апатію, а також часті головні болі та порушення сну [34].

Високий рівень когнітивного навантаження, властивий професіям, пов'язаним з обробкою інформації, збільшує ризики вигорання. Дослідження показують, що понад 60% офісних працівників, які працюють з ПК понад 6 годин на добу, мають ті чи інші ознаки стресу або виснаження.

Причинами такого стану є не лише перевантаження, але й відсутність чіткої організації праці, перерв, недостатня фізична активність, монотонність та відсутність зворотного зв'язку. Як зазначає дослідниця В. С. Говоруха, «тривала робота за ПК без регламентованих перерв і зміни діяльності порушує природний ритм психофізіологічної активності людини» [35].

Крім того, високі очікування з боку керівництва, обмежені терміни виконання завдань і часті позапланові доручення створюють додатковий тиск. У ситуаціях, коли працівник не має достатнього контролю над власною роботою або не бачить її результатів, виникає відчуття безпорадності. Це, у свою чергу, є передумовою для розвитку професійного вигорання.

Ще одним ризиком для психічного здоров'я є стирання меж між робочим і особистим життям. Особливо це актуально для фахівців, які працюють дистанційно. В умовах, коли немає чіткого розмежування між роботою та

відпочинком, працівник постійно перебуває в напруженому стані, що негативно позначається на його психоемоційному стані. Крім психоемоційних чинників, важливим аспектом, що впливає на самопочуття працівників, є умови освітлення робочого місця. Належне освітлення не тільки зберігає зір, але й підвищує продуктивність праці. Зокрема, у методичці йдеться: «Правильне освітлення сприяє безпечному виконанню робіт, зберігає зір і підвищує працездатність працівника та продуктивність праці. Недостатнє або нерівномірне освітлення погіршує умови праці, внаслідок чого працівник передчасно втомлюється, можливі випадки травматизму та очні хвороби» [36]. Це свідчить про те, що організація комфортного і безпечного освітлення на робочому місці є одним з важливих способів профілактики вигорання.

З метою запобігання вигоранню рекомендується дотримуватись правил гігієни праці, зокрема організовувати регулярні перерви, чергувати види діяльності, створити комфортне робоче місце, а також забезпечити емоційне розвантаження – фізичну активність, спілкування, хобі. Згідно з методичними рекомендаціями МОЗ України, кожна година роботи за комп'ютером має супроводжуватись п'ятихвилинною перервою, під час якої бажано виконувати вправи для очей та розминку [37].

4.2 Характеристика життєдіяльності людини у системі «людина – машина – середовище існування»

Сучасне виробниче середовище дедалі більше інтегрує людину в складні технічні системи, де вона виступає не лише виконавцем, а й активним учасником управлінського процесу. Система «людина – машина – середовище існування» (СЛМС) являє собою взаємозв'язок між трьома ключовими компонентами: людиною як оператором, технічним засобом (машиною) і зовнішніми умовами, у яких здійснюється діяльність.

Упродовж історії розвитку праці функції людини суттєво змінювалися. На початкових етапах людина переважно використовувала свою фізичну силу

для керування технікою. Проте із появою складніших технічних засобів, таких як автомобілі чи літаки, виникла необхідність у врахуванні психологічних особливостей оператора – швидкості реакції, концентрації уваги, стану емоцій та пам'яті [38].

Операторська діяльність, яка є основною формою взаємодії людини з машиною, передбачає переробку великої кількості інформації. Людина–оператор повинна не лише сприймати сигнали, але й аналізувати їх, приймати рішення та реалізовувати їх у вигляді конкретних дій. Як зазначає Я.І. Бедрій, ця діяльність вимагає «значних витрат нервово–емоційної та розумової енергії», що потребує врахування психофізіологічних можливостей людини під час проектування технічних систем [38].

Особливістю роботи оператора є те, що він не має прямого контакту з об'єктом керування, а працює з інформаційною моделлю процесу. Це висуває високі вимоги до когнітивних функцій: уваги, пам'яті, здатності до аналізу, прийняття рішень. Як підкреслюється у роботі [39], «у високотехнологічних системах безпека залежить не стільки від техніки, скільки від ефективності психофізіологічного узгодження між людиною та машиною».

У високотехнологічному середовищі значно зростають вимоги до точності, швидкості прийняття рішень і стійкості до стресових ситуацій. Важливим є те, що оператор уже не взаємодіє з реальним об'єктом напряду, а працює з інформаційною моделлю, представленою на екранах приладів або дисплеїв. Це суттєво ускладнює когнітивне навантаження та вимагає високого рівня підготовки.

Згідно з класичною ергономічною моделлю, діяльність оператора складається з чотирьох послідовних етапів: прийняття інформації, її аналіз та переробка, прийняття рішення і реалізація дій. Такий підхід дозволяє детально дослідити особливості взаємодії людини з машиною в контексті впливу навколишнього середовища [40].

Комп'ютеризація призвела до значних змін у структурі професій, поставивши перед користувачем складні завдання, які вимагають точності,

швидкості, адаптивності. Згідно з дослідженнями [41], «ефективність функціонування ЛМС прямо залежить від відповідності технічних параметрів обладнання можливостям людини».

4.3 Висновок до четвертого розділу

У розділі було розглянуто взаємодію людини з комп'ютерною технікою у системі «людина – машина – середовище існування» та проаналізовано основні психоемоційні ризики, пов'язані з інтенсивною працею за ПК. Встановлено, що тривала робота в інформаційному середовищі без дотримання гігієнічних та ергономічних норм сприяє розвитку психоемоційного вигорання, зниженню продуктивності праці, появі тривожних та депресивних станів. Ці наслідки мають як індивідуальні, так і соціально–економічні ризики, що потребують системного підходу до профілактики.

Водночас ефективна побудова взаємодії у системі «людина – машина» вимагає врахування фізіологічних, когнітивних і психологічних можливостей користувача. Особливо важливо забезпечити ергономічність робочого середовища, адаптацію інтерфейсів, чітку регламентацію режиму праці та відпочинку. Комплексний підхід до організації безпечної роботи з комп'ютером дозволяє не лише знизити рівень навантаження на організм, але й підвищити загальну ефективність діяльності в умовах цифрової трансформації.

ВИСНОВКИ

У даній кваліфікаційній роботі освітнього рівня «Бакалавр» виконано комплексне дослідження застосування технологій штучного інтелекту для автоматизації відеомонтажу, а також проведено порівняльний аналіз ефективності таких рішень із традиційними методами редагування відео.

У першому розділі кваліфікаційної роботи:

- Подано аналіз сучасного стану відеомонтажу та його автоматизації, висвітлено розвиток технологій ШІ у цій сфері.

- Розглянуто принципи роботи систем комп'ютерного зору, машинного навчання та нейронних мереж, що застосовуються для обробки відео.

- Висвітлено особливості ручного та автоматизованого монтажу, а також проаналізовано популярні програмні засоби, такі як RunwayML, Adobe Sensei, Magisto, Wisecut і Pictory.

- Проаналізовано сучасні моделі ШІ: Gen-1, Gen-2, Sora та їхні можливості у створенні й редагуванні відеоконтенту.

У другому розділі кваліфікаційної роботи:

- Досліджено концепцію відеопроєкту «The Last Letter from Paris» та обґрунтовано вибір інструментів ШІ для його реалізації, зокрема RunwayML Gen-4, ElevenLabs, Soundraw, AIVA, Descript і DaVinci Resolve.

- Обґрунтовано архітектуру робочого процесу створення відео, включаючи генерацію відеофрагментів, озвучення, створення музики, субтитрів і фінальний монтаж.

- Сформовано порівняльний аналіз альтернативних інструментів, розглянуто ризики, етичні аспекти та технічні вимоги до реалізації проєкту.

У третьому розділі кваліфікаційної роботи:

- Розроблено сценарний план відеопроєкту, сформульовано текстові підказки для генерації відеофрагментів у RunwayML Gen-4.

- Запропоновано та реалізовано озвучення за допомогою ElevenLabs, генерацію музики та звукових ефектів через Soundraw і AIVA, а також створення субтитрів у Descript і DeepL.

- Спроектовано та виконано фінальний монтаж у DaVinci Resolve із застосуванням кольорокорекції та синхронізації всіх компонентів.

- Протестовано якість відео за візуальними, аудіо та емоційними параметрами, підтвердивши високу ефективність ШІ-інструментів у створенні кінематографічного контенту.

У розділі «Безпека життєдіяльності, основи охорони праці»:

- Висвітлено психоемоційні ризики, пов'язані з тривалою роботою за комп'ютером, зокрема ризик вигорання, та запропоновано заходи для їх профілактики.

- Проаналізовано взаємодію людини з технічними системами в контексті системи «людина – машина – середовище існування», підкреслено важливість ергономічного підходу до організації праці.

Результати дослідження підтвердили, що технології штучного інтелекту значно спрощують і прискорюють процес відеомонтажу, забезпечуючи високу якість результату навіть за відсутності професійного обладнання чи знімальної групи. Практична реалізація відеопроєкту продемонструвала переваги ШІ у створенні доступного, гнучкого та креативного контенту, що має потенціал для масштабування в освітній, комерційній та творчій сферах. Водночас виявлено етичні та правові виклики, пов'язані з використанням ШІ, що потребують подальшого вивчення та регулювання.

ПЕРЕЛІК ДЖЕРЕЛ

1. BigVu. (2023). The Role of AI in Video Production and Editing – A Game Changer. – 2023. – Режим доступу: <https://bigvu.tv/blog/the-role-of-ai-in-video-production-and-editing>.
2. Adobe Inc. (2025). Introducing new AI-powered features and workflow enhancements in Premiere Pro and After Effects 25.2. – 2025. – Режим доступу: <https://blog.adobe.com/en/publish/2025/04/02/introducing-new-ai-powered-features-workflow-enhancements-premiere-pro-after-effects>.
3. Adobe Inc. (2024). Bringing generative AI to video editing workflows in Adobe Premiere Pro. – 2024. – Режим доступу: <https://blog.adobe.com/en/publish/2024/04/15/bringing-gen-ai-to-video-editing-workflows-adobe-premiere-pro>.
4. RunwayML. (2024). Official Documentation.. – 2024. – Режим доступу: <https://runwayml.com>.
5. Adobe Sensei. (2024). Generative AI – Adobe Sensei. – 2024. – Режим доступу: <https://www.adobe.com/uk/sensei/generative-ai.html>.
6. Wisecut. (2024). Wisecut – AI Video Editor. – 2024. – Режим доступу: <https://www.wisecut.ai/>.
7. Statista. (2022). Hours of video uploaded to YouTube per minute. – 2022. – Режим доступу: <https://soax.com/research/how-many-hours-of-video-are-uploaded-to-youtube-every-minute>.
8. Kapwing. (2024). Auto-subtitle any video or audio file with Kapwing's AI-powered Subtitle Generator. – 2024. – Режим доступу: <https://www.kapwing.com/subtitles>.
9. OpenAI. (2024). Introducing Sora – OpenAI's text-to-video generation model. – 2024. – Режим доступу: <https://openai.com/sora/>.
10. VEED.io. (2025). Free AI Video Generator – Create Videos from Text. – 2025. – Режим доступу: <https://www.veed.io/tools/ai-video>.

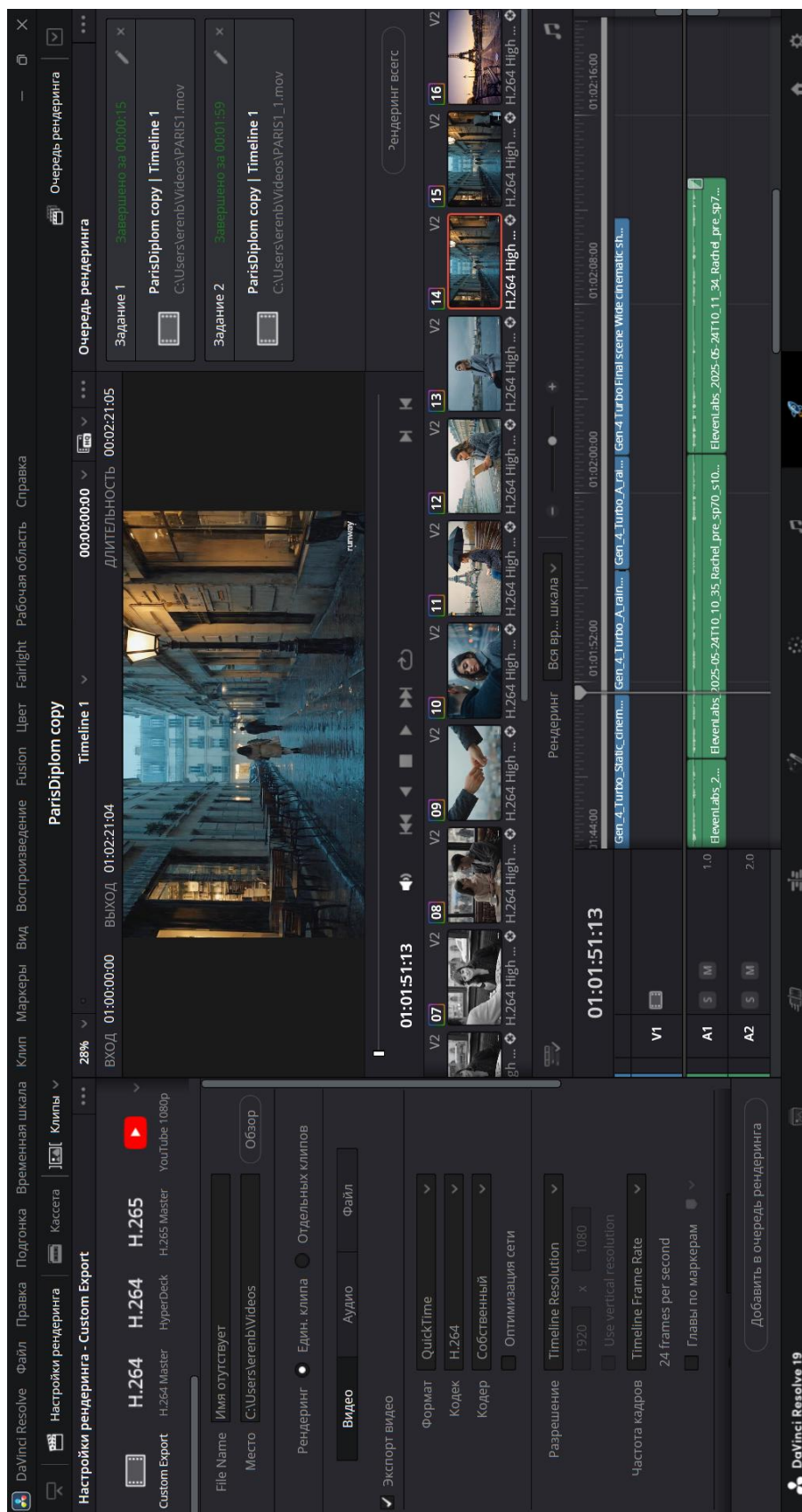
11. Lyu, S. (2023). Recent Advances in Text-to-Video Generation using Diffusion and Transformer Models. In: AI Media Journal, Vol. 4(1).. – 2023. – Режим доступу: <https://arxiv.org/html/2505.07652v1>.
12. Soundraw.io. (2024). Your Comprehensive Guide to SOUNDRAW's AI Music Licensing. – 2024. – Режим доступу: <https://blog.soundraw.io/post/soundraw-licensing-explained>.
13. How AI Helps Generate Subtitles / Kapwing. – 2024. – Режим доступу: <https://www.kapwing.com/help/add-captions>
14. DaVinci Resolve 18 Reference Manual / Blackmagic Design. – 2024. – Режим доступу: <https://documents.blackmagicdesign.com/UserManuals/DaVinci-Resolve-18-Colorist-Guide.pdf>
15. He, Z., Zhou, Y., Ma, S., Zhou, J., Chen, W., Xu, Y., Liu, Z., Chen, Z., Qin, T., Liu, T. Animate-A-Story: Storytelling with Retrieval-Augmented Video Generation/ Z. He, Y. Zhou, S. Ma та ін. // arXiv preprint. – 2023. – № arXiv:2307.06940. – Режим доступу: <https://arxiv.org/abs/2307.06940>
16. PromptHero. (2024). Best Practices for Writing Prompts for Text-to-Video Models.. – 2024. – Режим доступу: <https://prompthero.com/blog/text-to-video-prompts>.
17. PlayHT. (2024). AI Voice Generator with Custom Emotions.. – 2024. – Режим доступу: <https://play.ht>.
18. Soundraw.io. (2024). AI Music Generator for Creators.. – 2024. – Режим доступу: <https://soundraw.io>.
19. AIVA. (2024). Compose Emotional Music with AI.. – 2024. – Режим доступу: <https://www.aiva.ai>.
20. Lytvynenko, I., Lupenko, S., Kunanets, N., Nazarevych, O., Shymchuk, G., & Hotovych, V. (2021). Simulation of gas consumption process based on the mathematical model in the form of cyclic random process considering the scale factors. 1st International Workshop on Information Technologies: Theoretical and Applied.

21. Descript. (2024). All-in-one AI Audio and Video Editor.. – 2024. – Режим доступу: <https://www.descript.com>.
22. Kapwing. (2024). Generate and Edit Subtitles with AI.. – 2024. – Режим доступу: <https://www.kapwing.com>.
23. Strutynska, I., Kozbur, H., Dmytrotsa, L., Bodnarchuk, I., & Hlado, O. (2019). Small and medium business structures clustering method based on their digital maturity. 2019 IEEE International Scientific–Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), 278–282.
24. Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. New Haven; London: Yale University Press. – 2021. – Режим доступу: <https://yalebooks.yale.edu/book/9780300252392/atlas-ai>.
25. UNESCO. (2022). Recommendation on the Ethics of Artificial Intelligence. – 2022. – Режим доступу: <https://www.unesco.org/en/artificial-intelligence/ethics>.
26. Kharchenko, A., Bodnarchuk, I., & Halay, I. (2013). Decision support system of software architect. 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 1, 265–269.
27. Google Developers. (2023). Performance for a content-driven web app frontend. – 2023. – Режим доступу: <https://developers.google.com/solutions/content-driven/frontend/performance>.
28. Soundraw.io. (2024). Your Comprehensive Guide to SOUNDRAW's AI Music Licensing. – 2024. – Режим доступу: <https://blog.soundraw.io/post/soundraw-licensing-explained>.
29. Descript. (2024). Descript system requirements. – 2024. – Режим доступу: <https://help.descript.com/hc/en-us/articles/10503411779213-Descript-system-requirements>.
30. Nebesnyi, R., Kunanets, N., Vaskiv, R., & Veretennikova, N. (2021). Formation of an IT project team in the context of PMBOK requirements. 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT).

31. Pictory AI. (2024). Video Generation for Social Media Campaigns.. – 2024. – Режим доступу: <https://pictory.ai>.
32. Kunanets, N., Nebesnyi, R., Nazaruk, M. V., & Pasichnyk, V. V. (2018). Information technology of personalized choice of profession in smart cities. Інформаційні технології і засоби навчання, 277–290.
33. RunwayML. (2024). API Documentation: Automating Video Generation.. – 2024. – Режим доступу: <https://docs.runwayml.com>.
34. Шевченко Т. В. Психоемоційне вигорання в умовах інтенсивної інформаційної діяльності // Психологія і суспільство. – 2021. – № 1. – С. 114–119.
35. Говоруха В. С. Оцінка рівня психоемоційного навантаження у працівників, що працюють за комп'ютером // Охорона праці. – 2020. – № 3. – С. 39–43.
36. Сенчишин В. С., Гурик О. Я., Король О. І. Методичні вказівки до лабораторної роботи № 8 з дисципліни «Основи охорони праці». Дослідження факторів, які впливають на освітленість робочих місць у приміщенні [Електронний ресурс] / ТНТУ ім. Івана Пулюя. – Тернопіль, 2013. – Режим доступу: <http://elartu.tntu.edu.ua/handle/lib/20725>
37. Наказ МОЗ № 518 від 30.12.2002 р. «Про затвердження Державних санітарних норм і правил при роботі з ПК».
38. Бедрій Я. І. Безпека життєдіяльності. – Тернопіль : Навчальна книга – Богдан, 2010. – С. 32–35.
39. Шевчук Р. М. Ергономіка сучасного виробництва. – Київ : Либідь, 2021. – С. 44–48.
40. Сидоренко В. П. Основи ергономіки та безпеки праці. – Харків : ХНУРЕ, 2019. – С. 59–62.
41. Інститут кібернетики ім. В.М. Глушкова НАН України. Наукова доповідь з ергономіки. – Київ, 2020.

ДОДАТКИ

Кадри створення відеопроєкту «The Last Letter form Paris»



Image

Video

Chat Mode

Parisian Romance in Rain ...

Gen-4 Turbo

11 credits

Upgrade



A wide cinematic shot of the Eiffel Tower at dusk. The sky is painted in soft gra...

1

2

3

4

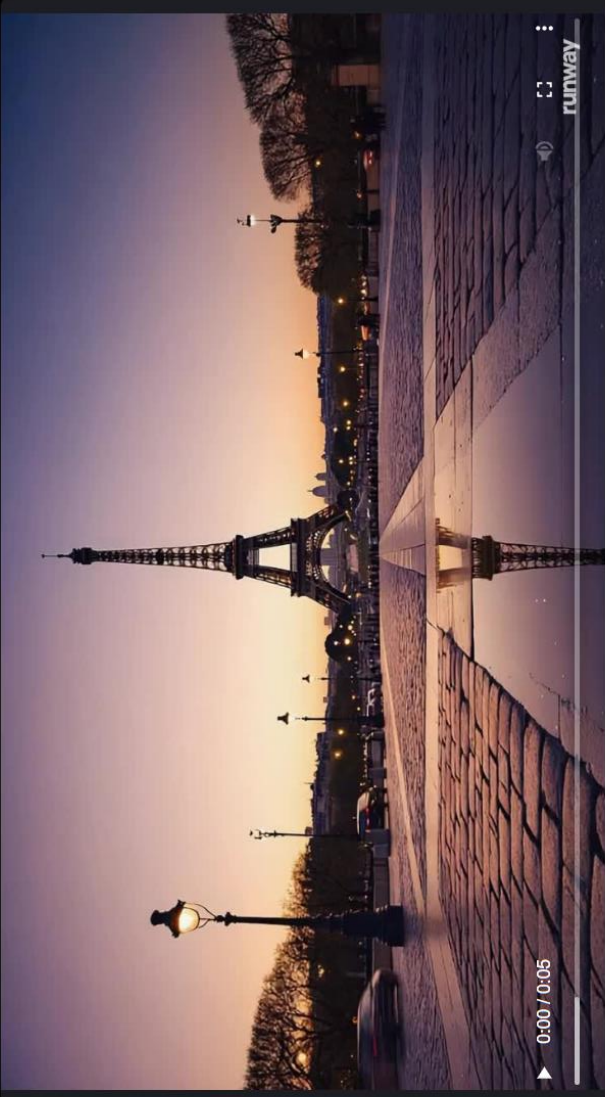
5

6

7

8

9



0:00 / 0:05

runway

Restyle

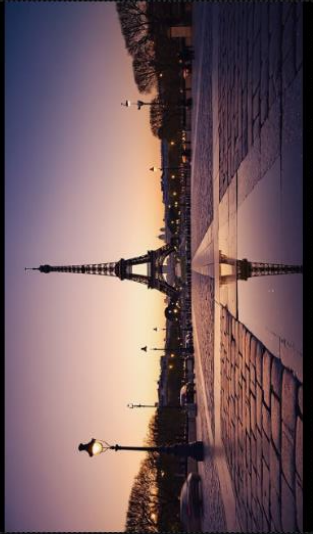
Use frame

4K

Share

Like

Download



begins with a low angle close up of a person's face. The Eiffel Tower in the water. Soft evening lights gradually turn on along the streetlamps. Cars gently pass by in the background. Trees frame the scene on both sides, swaying slightly in the evening breeze. The ambiance is calm, romantic, and peaceful.

16:9

5s

Gen-4 Turbo

Upgrade (out of credits)