

УДК 004.93

CLIPPEROK: ПЛАТФОРМА АВТОМАТИЗАЦІЇ СТВОРЕННЯ КОРОТКИХ ВІДЕО

Ольховський Б.Д.

e-mail: bohdan.olkhovskiy@nure.ua

Науковий керівник – ст. викл. Сінельнікова Т.Ф.

Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This work introduces ‘Clipperok’, an AI platform for automated vertical video production. The system provides SMM teams with a unified ecosystem for transcription, scene detection, and social media posting. It utilizes OpenAI Whisper for speech-to-text and FFmpeg for high-quality 9:16 rendering with dynamic subtitles. The microservice architecture, powered by FastAPI and React, ensures integration flexibility via REST API. The project demonstrates the high efficiency of AI-driven analytics in modern media automation workflows.

Сучасна трансформація медіасфери під впливом штучного інтелекту призвела до необхідності автоматизації процесів, які раніше потребували значних ресурсів ручного монтажу. Ця тенденція є критично важливою для SMM-команд, що працюють з адаптацією довгих відео під вертикальні форми, такі як TikTok, Instagram Reels та YouTube Shorts. Платформа Clipperok пропонує інтегроване рішення, яке об’єднує AI-транскрибацію, інтелектуальну нарізку фрагментів на основі семантичного аналізу тексту та автоматизовану публікацію готового контенту через офіційні API соціальних мереж.

В основі розробленої вебплатформи лежить наскрізний пайплайн, що забезпечує автоматичну транскрибацію тексту, виділення найбільш значущих епізодів та фінальний рендеринг відео. Архітектура системи, побудована на React SPA та FastAPI, дозволяє інтегрувати сервіс в існуючі контури автоматизації (наприклад, n8n) через REST API. Питання безпеки вирішуються шляхом впровадження механізмів JWT-авторизації та використання протоколу OAuth 2.0 для прямої взаємодії з соціальними мережами.

Технологічний стек проєкту орієнтований на максимальну ефективність обробки відеоряду. Модуль рендерингу на базі FFmpeg реалізує складний single-pass алгоритм: одночасне кадрування (cropping) центральної області під формат 9:16, корекцію кольору та накладання стилізованих динамічних субтитрів. Пошук таймкодів для кожного слова дозволяє синхронізувати візуальні ефекти з мовленням з мілісекундною точністю. Для забезпечення зворотного зв’язку з користувачем у реальному часі впроваджено протокол WebSocket, через який передаються статуси прогресу рендерингу.

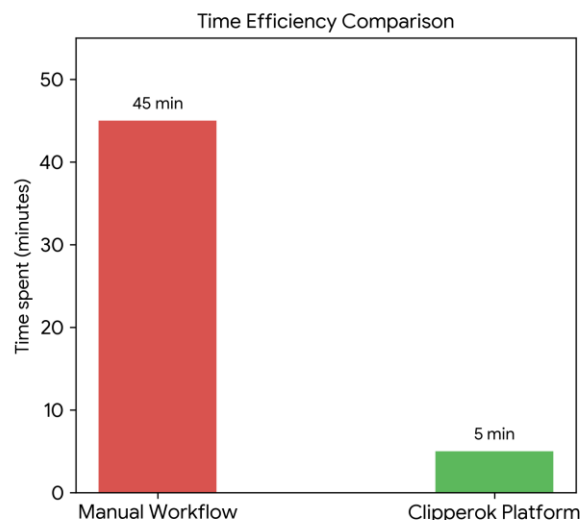


Рисунок 1 – Результати порівняння часової ефективності підготовки контенту

Практична апробація MVP-версії системи підтвердила її високу ефективність. Середній час повної генерації кліпу з 10-хвилинного вихідного файлу становить 4–5 хвилин при використанні GPU-прискорення (NVIDIA CUDA). Точність транскрибації за допомогою Whisper Large v3 досягає 98% для основних європейських мов. Автоматизація постингу через OAuth 2.0 контур дозволила скоротити загальні часові витрати користувача на 80% порівняно з традиційним робочим процесом у відеоредакторах. Таким чином, впровадження подібних AI-платформ радикально підвищує ефективність.

Подальший розвиток проєкту Clipperok передбачає впровадження модулів комп'ютерного зору для автоматичного відстеження обличчя в кадрі (face tracking), що дозволить динамічно змінювати область кадрування. Також планується інтеграція великих мовних моделей (LLM) для автоматичної генерації віральних описів та тегів до відео, що ще більше знизить когнітивне навантаження на контент-мейкерів та прискорить виведення медіа-продуктів на ринок.

Список використаних джерел:

1. Кобилін О. А. Методи цифрової обробки зображень : навч. посіб. Харків : ХНУРЕ, 2021. 124 с.
2. Radford A. et al. Robust Speech Recognition via Large-Scale Weak Supervision. OpenAI, 2022. URL: <https://arxiv.org/abs/2212.04356>.
3. Voron T. Building Data Science Applications with FastAPI. Birmingham : Packt Publishing, 2023. 326 p.
4. FFmpeg Documentation. URL: <https://ffmpeg.org/documentation.html>
5. Субботін С. О. Подання й обробка знань у інтелектуальних системах : навч. посіб. Запоріжжя : ЗНТУ, 2010. 228 с.