ArtemisSearch: A Multimodal Search Engine for Efficient Video Log-Life Event Retrieval Using Time-Segmented Queries and Vision Transformer-based Feature Extraction

Hoang-Phuc Nguyen, Thuy-Nga Ho, Minh-Dai Tran-Duong, The-Luan Nguyen, Duc-Hao Truong, Quyen Nguyen Huu, Duy Phan The, Van-Hau Pham

Introduction

In today's era of data explosion, users seek tools for fast, accurate searches. ArtemisSearch is a web application designed to empower users to create temporally-aware queries. It allows them to seamlessly explore top results, and view precise video previews at key timestamps with ease.

Method

ArtemisSearch utilizes **CLIP-ViT-H/14** and **BEiT3** models to extract semantically rich features from videos, enabling precise search capabilities. Powered by **Milvus**, an open-source vector database, the system efficiently stores and retrieves feature vectors, ensuring scalable and fast processing of complex queries across large video datasets.

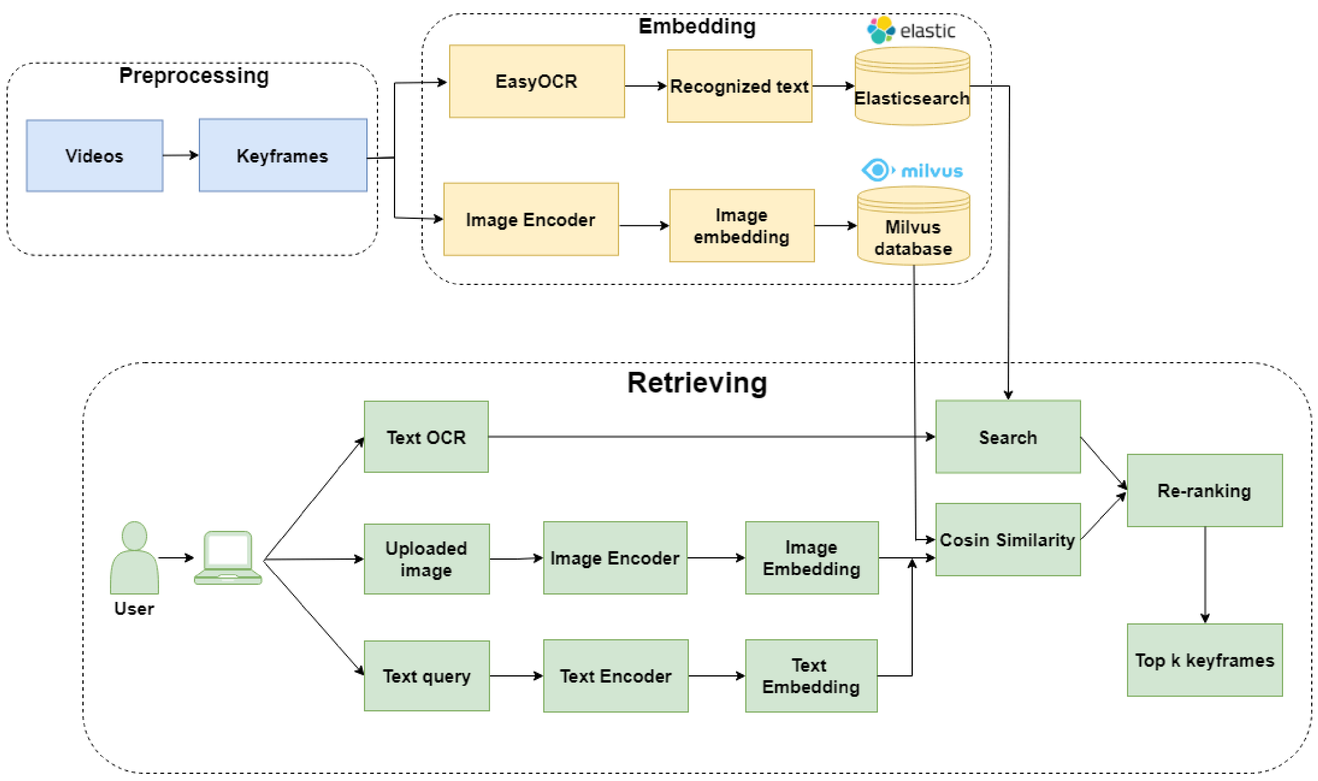
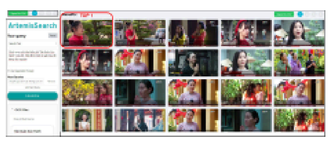


Fig. 1: The Architecture overview of ArtemisSearch System.

(a) Without OCR: Target image at rank 28

(b) Frame 7131 contains the OCR text, which is a previous frame of the top 28st - frame 7253.



(c) With OCR: Target image at rank 1

Fig. 2: Comparison of search results with and without OCR integration.

To enhance retrieval accuracy, we integrate **OCR** capabilities using **EasyOCR** and store the extracted results in **Elasticsearch**, enabling efficient indexing and retrieval of textual information from video frames.

To integrate scores from multiple event queries effectively, we implement a temporal scoring mechanism that establishes relationships between consecutive frames during the extraction process. This approach enhances the search for frames containing closely occurring events, delivering more accurate and context-aware retrieval results.

Results

ArtemisSearch presents a powerful and efficient solution for fast, reliable video retrieval. Its performance was rigorously tested during the AI Challenge Ho Chi Minh City 2024, where it successfully and accurately addressed all queries in the final round, demonstrating exceptional speed and precision in real-world applications.

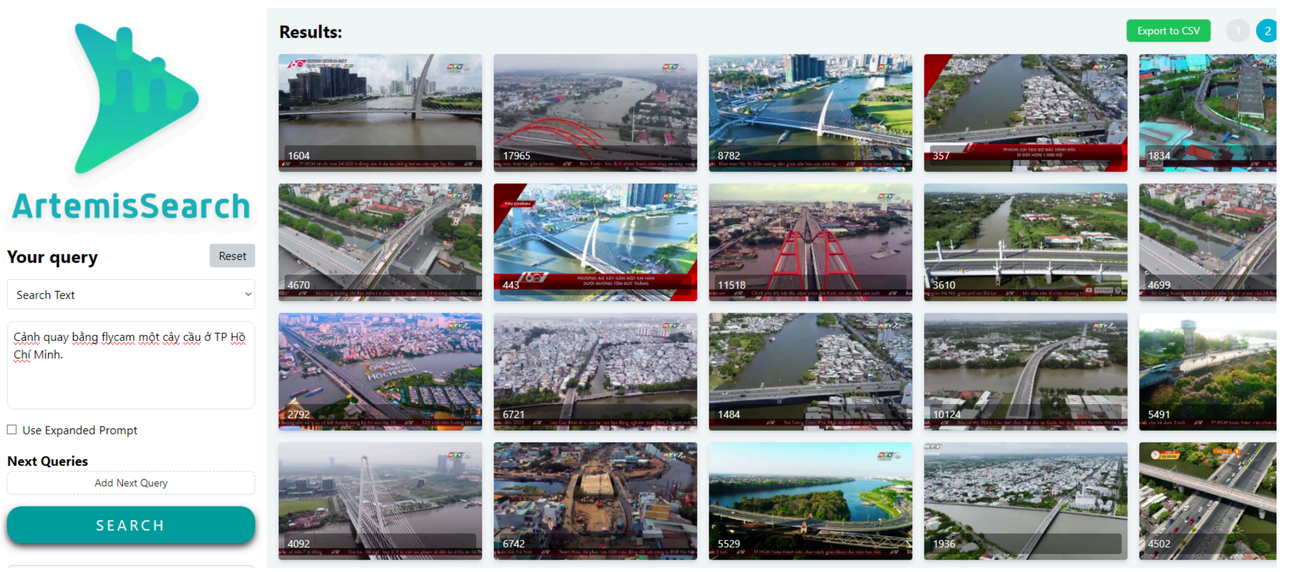


Fig. 3: The User Interface of ArtemisSearch.