

TRUY HỒI KHOẢNH KHẮC TRONG VIDEO

Phạm Huỳnh Long Vũ - 23521813
Nguyễn Huy Phước - 23521234

Tóm tắt

Nguyễn Huy Phước - 23521234



Phạm Huỳnh Long Vũ - 23521813

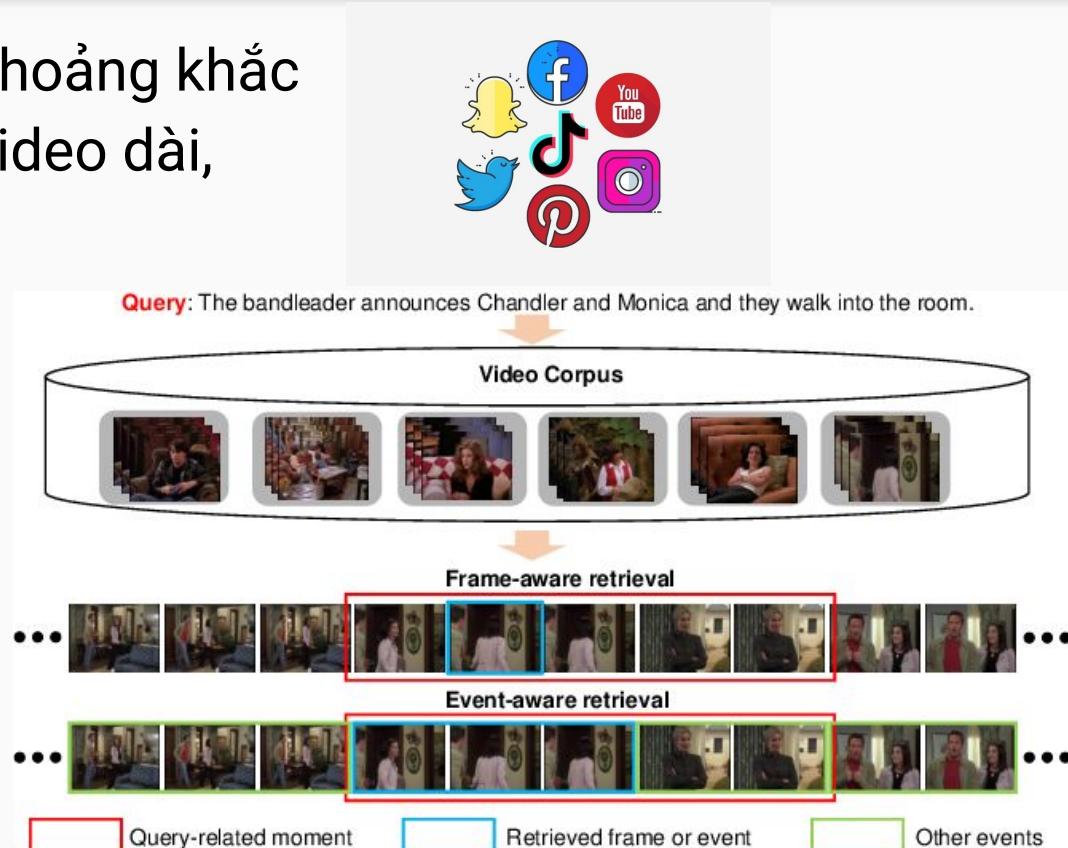


Link YouTube video: <https://www.youtube.com/watch?v=TxUOXc1c9U4>

Link Github của nhóm: <https://github.com/23521813/CS519.P11>

Giới thiệu

- Nhu cầu: Tìm kiếm theo khoảnh khắc trong một tập video lớn, video dài, khoảnh khắc rất ngắn.
- Input: truy vấn văn bản + bộ sưu tập video
- Output: video id+ đoạn thời gian xuất hiện khoảnh khắc đó.

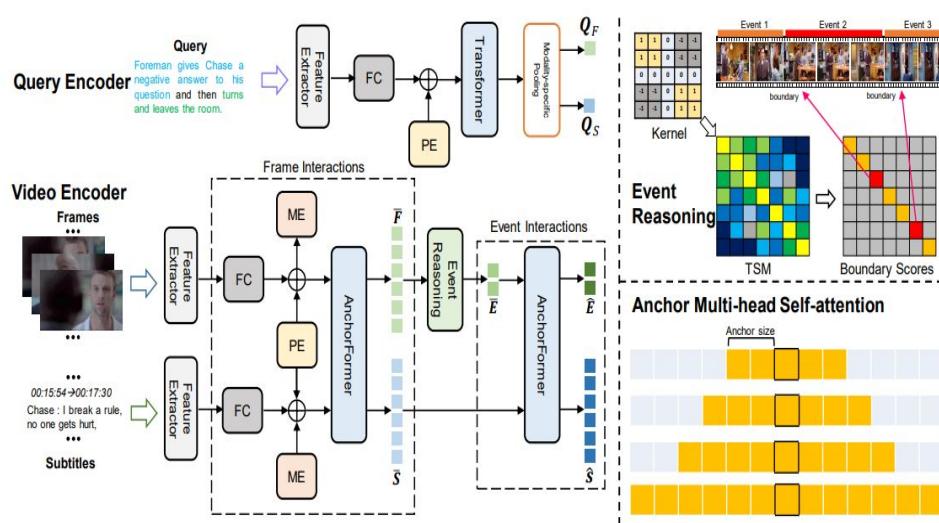


Mục tiêu

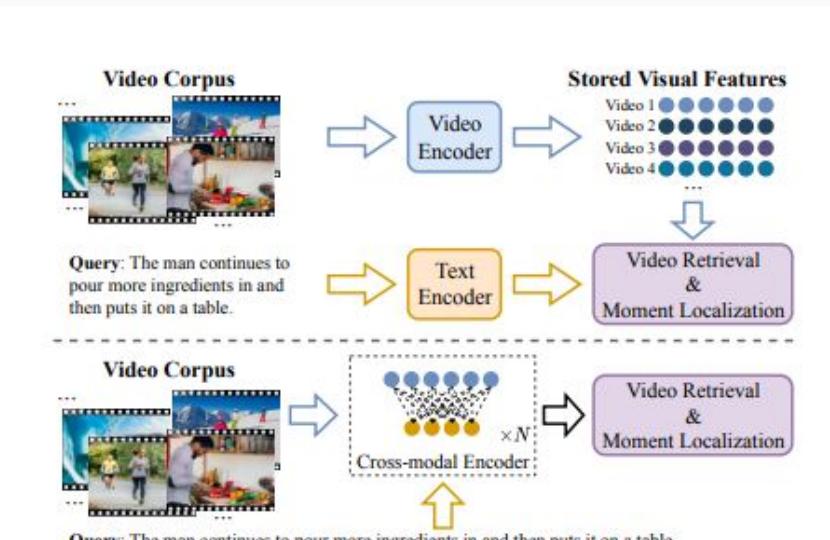
1. Khảo sát và phân tích các hướng tiếp cận sẵn có phân tích ưu và nhược điểm chung.
2. Thiết kế và hiện thực một mô hình baseline truy hồi và định vị nhẹ hơn về mặt huấn luyện, truy vấn nhanh.
3. Mô hình cần được làm rõ các tiêu chí đánh giá.

Nội dung và Phương pháp

Mô hình hai giai đoạn



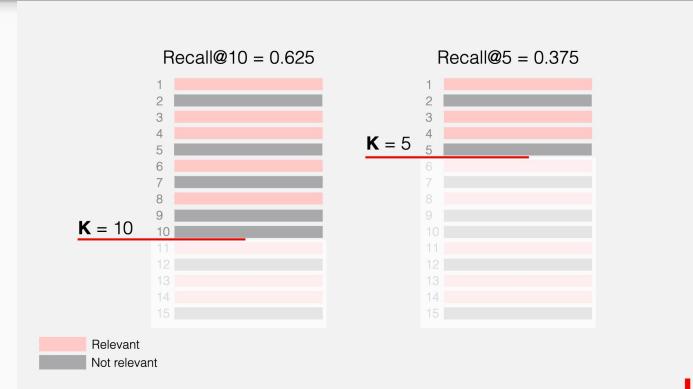
Mô hình đầu cuối



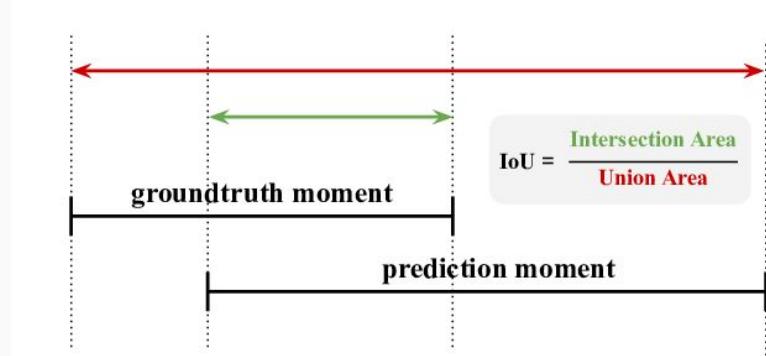
Nội dung và Phương pháp

Độ đo:

1. Recall@K & Temporal IoU threshold
2. Query time response



$$\text{Average Response Time} = \frac{\text{Sum of all response times}}{\text{Total number of tickets}} \times 100$$



Nội dung và Phương pháp

Đề xuất giải pháp

1. Trích xuất keyframe từ video và sàng lọc các frame giống nhau (Giai đoạn chuẩn bị dữ liệu).
2. Tách truy vấn thành các sự kiện đơn.
3. Lọc các video tìm năn, mỗi sự kiện xuất hiện ít nhất 1 lần trong video.
4. Sử dụng kỹ thuật truy hồi quy hoạch động để lọc các bộ đáp án từ tập video tiềm năng.

Kết quả dự kiến

Một kho dữ liệu keyframe cho toàn bộ tập video được khảo sát.

Xây dựng hệ thống cho phép người dùng nhập câu truy vấn, hệ thống trả ra video id kèm theo thời gian tương ứng trong kho video khảo sát.

Thu được bộ kết quả thực nghiệm đánh giá hệ thống trên tiêu chí

- i. Mức độ chính xác.
- ii. Tốc độ truy vấn.

Làm cơ sở đề xuất các cải tiến ở các nghiên cứu tiếp theo.

Tài liệu tham khảo

- [1]. Meng Liu, Liqiang Nie, Yunxiao Wang, Qi Tian, Tat-Seng Chua:
A Survey on Video Moment Localization. *ACM Computing Surveys*, 2023
- [2] Hao Zhang, Aixin Sun, Wei Jing, Guoshun Nan, Liangli Zhen, Joey Tianyi Zhou, Rick Siow Mong Goh:
Video Corpus Moment Retrieval via Contrastive Learning. SIGIR 2021: 1552–1561.
- [3] Shubhashis Roy Dipta, Francis Ferraro
Query-to-Event (Q2E) Decomposition for Zero-Shot Text-to-Video Retrieval. NAACL 2025.
- [4]. Danyang Hou, Liang Pang, Huawei Shen, Xueqi Cheng:
Event-Aware Video Corpus Moment Retrieval. 2024.
- [5] Junyu Gao, Changsheng Xu
Fast Video Moment Retrieval. ICCV 2021: 11651–11660.