**Mô tả giải pháp**

**Tiếp cận bài toán**: Ban đầu em nghĩ đây sẽ là một challenge text-video nhưng sau khi khảo sát và xem qua cấu trúc của dữ liệu mà ban tổ chức cung cấp thì em xác định đây sẽ là bài toán text-image và image-image. Đó là dự tính ban đầu cho vòng sơ loại này của em vì BTC cũng không cung cấp tập test và ground truth gì ngoài một số file có object detection nên cũng khó khăn cho việc train model gì nên định hướng của em sẽ tập trung tìm hiểu một số model pretrained tốt cho bài toán này và dành thời gian làm một giao diện đẹp hơn sau khi vào vòng sơ loại.

**Xử lí dữ liệu:** Em chỉ sử dụng phần KeyFrames cho bài toán này. Các dữ liệu ảnh sẽ được đi qua một model để trích xuất đặc trung và lưu dưới dạng các vector vào milvus. Một dạng database giúp lưu trữ và truy xuất nhanh dữ liệu vector.

**Xây dựng hệ thống:** search engine API còn đang trong quá trình làm còn các phần khác đều đã hoàn thành.