

Báo cáo tiến độ tuần 7 (13/05 – 19/05)

1. Bổ sung tính năng xử lý nhiều ảnh cùng lúc

- Đã nâng cấp hệ thống từ xử lý ảnh đơn sang khả năng xử lý ảnh hàng loạt.
- Giao diện được bổ sung nút “Chọn nhiều ảnh” để người dùng có thể nhập vào một thư mục chứa nhiều ảnh.
- Hệ thống lần lượt thực hiện:
 - Phát hiện đối tượng trong từng ảnh bằng mô hình **YOLOv8** đã huấn luyện trước.
 - Sinh mô tả nội dung ảnh bằng mô hình **BLIP (Bootstrapping Language-Image Pretraining)**.
 - Dịch mô tả sang tiếng Việt bằng **Google Translate API**.
- Kết quả được hiển thị đồng thời dưới dạng danh sách ảnh kèm mô tả tương ứng.

2. Cải thiện phần mô tả ngữ cảnh bằng cách kết hợp nhiều ảnh

- Đã thử nghiệm phương pháp **tổng hợp mô tả từ nhiều ảnh liên tiếp** (ví dụ ảnh cùng một cảnh quay hoặc sự kiện).
- Triển khai logic nhóm các ảnh cùng thư mục hoặc tên gần giống nhau thành một “ngữ cảnh”.
- BLIP được gọi để sinh caption từng ảnh → gom lại thành một đoạn văn mô tả bối cảnh.
- Dịch đoạn mô tả chung đó sang tiếng Việt, giúp người dùng hiểu tổng thể nội dung trong nhiều ảnh.

3. Viết chương “Kết quả thực nghiệm” trong báo cáo

- Đã tiến hành viết phần đầu tiên của chương “Kết quả thực nghiệm”:
 - Mô tả cách tiến hành thực nghiệm với mô hình YOLOv8 + BLIP.
 - Nêu rõ tiêu chí đánh giá như: thời gian xử lý, độ chính xác nhận diện, mức độ hợp lý của mô tả ngữ cảnh.
 - Trình bày một số ví dụ minh họa kèm ảnh gốc, caption tiếng Anh và bản dịch tiếng Việt.

