

Phân tích và trực quan hóa không gian cảnh

1 Mô tả tổng quát

Trong bài tập lớn này, sinh viên làm việc theo nhóm (3–4 sinh viên) để xây dựng một hệ thống Computer Vision nhằm phân tích một cảnh thực (indoor hoặc outdoor) từ ảnh hoặc video. Mục tiêu là hiểu cấu trúc hình học, nội dung và đặc trưng thị giác của cảnh, đồng thời trực quan hóa các kết quả xử lý một cách rõ ràng và có ý nghĩa.

Đề tài được thiết kế theo hướng mở: sinh viên được tự lựa chọn dữ liệu, phạm vi bài toán và mức độ phức tạp, miễn là có vận dụng tổng hợp các kiến thức Computer Vision đã học.

Đây là bài tập lớn mang tính tổng hợp, xuyên suốt môn học, vì vậy, nhóm sinh viên được yêu cầu xây dựng kế hoạch và thực hiện sớm ngay từ đầu kỳ.

2 Mục tiêu học tập

Sau khi hoàn thành bài tập, sinh viên cần:

- Thiết kế và hiện thực một pipeline Computer Vision hoàn chỉnh.
- Vận dụng tổng hợp nhiều kỹ thuật CV khác nhau trong cùng một bài toán.
- Rèn luyện kỹ năng phân tích, đánh giá và trình bày kết quả thực nghiệm.
- Làm việc nhóm hiệu quả với phân công nhiệm vụ rõ ràng.

3 Yêu cầu cốt lõi

Nhóm sinh viên cần hiện thực một hệ thống bao gồm **tối thiểu các thành phần** sau:

3.1 Thu thập và tiền xử lý dữ liệu

- Dữ liệu đầu vào là ảnh hoặc video do nhóm tự thu thập hoặc lấy từ nguồn công khai phù hợp.
- Thực hiện các bước tiền xử lý cần thiết, ví dụ:
 - Resize, crop ảnh.
 - Chuyển đổi không gian màu (RGB, grayscale, ...).
 - Giảm nhiễu hoặc cân bằng ánh sáng (nếu cần).

3.2 Phân tích hình học và cấu trúc cảnh

Áp dụng **ít nhất một** trong các hướng sau:

- Phát hiện biên, đường thẳng hoặc các đặc trưng hình học (edges, lines, corners).
- Ước lượng các phép biến đổi hình học (affine, projective).

- Ghép ảnh (panorama) hoặc căn chỉnh nhiều ảnh của cùng một cảnh.
- Ước lượng mặt phẳng hoặc các vùng bìa mặt chính trong cảnh.

3.3 Phân tích nội dung hoặc đối tượng

Áp dụng **ít nhất một** kỹ thuật phân tích nội dung, ví dụ:

- Phát hiện đối tượng.
- Theo dõi đối tượng trong video.
- Phân vùng ảnh ở mức cơ bản.
- Phát hiện vùng quan tâm (salient regions).

Sinh viên có thể sử dụng các phương pháp truyền thống hoặc kết hợp mô hình học sâu đã huấn luyện sẵn, nhưng cần giải thích rõ vai trò, giả định và hạn chế của phương pháp sử dụng.

3.4 Trực quan hóa kết quả

- Hiển thị rõ ràng các kết quả xử lý như bounding boxes, mask, keypoints, đường thẳng, mặt phẳng, ...
- So sánh trực quan giữa dữ liệu đầu vào và kết quả đầu ra.
- Với video, trình bày kết quả theo trình tự thời gian.

4 Yêu cầu mở rộng (khuyến khích)

Nhóm có thể lựa chọn thêm một hoặc nhiều hướng sau để nâng cao chất lượng bài làm:

- So sánh nhiều phương pháp khác nhau cho cùng một bài toán.
- Đánh giá kết quả bằng các tiêu chí định tính hoặc định lượng.
- Tối ưu pipeline để cải thiện tốc độ hoặc độ ổn định.
- Xây dựng demo tương tác đơn giản.

5 Yêu cầu báo cáo và sản phẩm

- Báo cáo PDF theo hướng dẫn chung của môn học.
- Báo cáo cần trình bày:
 - Mô tả bài toán và dữ liệu sử dụng.
 - Pipeline tổng thể của hệ thống.
 - Phân tích và thảo luận kết quả.
- Chỉ đưa vào báo cáo các đoạn code chính; code đầy đủ được nộp qua repository (public access).

6 Kết quả mong đợi

Sau khi hoàn thành bài tập, nhóm sinh viên cần chứng minh được:

- Khả năng thiết kế và hiện thực một hệ thống Computer Vision hoàn chỉnh.
- Sự hiểu biết và kết nối giữa các kiến thức CV đã học.
- Năng lực phân tích, đánh giá và trình bày kỹ thuật một cách logic và rõ ràng.
- Mỗi thành viên cần có phần việc và đóng góp kỹ thuật rõ ràng. Phân công nhiệm vụ cần được mô tả ngắn gọn trong báo cáo.