

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

KHOA CÔNG NGHỆ THÔNG TIN 1

BỘ MÔN THỰC TẬP CƠ SỞ



ĐỀ CƯƠNG THỰC HIỆN ĐỀ ÁN

Đề án: Xây dựng hệ thống mô hình YOLO tối ưu hóa khả năng nhận diện và phân loại đầu sách trên bộ dữ liệu hình ảnh bìa tự xây dựng

Giảng viên hướng dẫn

Họ và tên sinh viên

Mã sinh viên

Lớp

Kim Ngọc Bách

Nguyễn Trung Tín

B23DCVT416

D23CQCE04-B

Hà Nội – 2026

Mục Lục

I. Giới thiệu dự án	3
1. Lý do chọn đề tài	3
2. Ý nghĩa và ứng dụng	3
3. Giá trị học thuật	3
II. Cơ sở lý thuyết và công nghệ sử dụng	4
1. Thuật toán YOLO (You Only Look Once)	4
2. Công nghệ sử dụng	4
III. Phân tích yêu cầu dự án	4
1. Yêu cầu chức năng (Functional Requirements)	4
2. Yêu cầu phi chức năng (Non-functional Requirements)	5
IV. Kế hoạch thực hiện dự án	6
References	7

I. Giới thiệu dự án

1. Lý do chọn đề tài

- Trong kỷ nguyên số, các thư viện và nhà sách truyền thống đang phải đối mặt với thách thức lớn về tốc độ và độ chính xác trong công tác kiểm kê. Phương pháp nhận diện thủ công không chỉ tiêu tốn nhiều nhân lực mà còn tiềm ẩn nguy cơ sai sót. Chính vì vậy, việc ứng dụng Computer Vision trở thành giải pháp tối ưu, giúp tự động hóa quy trình, giải phóng sức lao động cũng như nâng cao hiệu quả quản lý và giảm thiểu chi phí vận hành.

2. Ý nghĩa và ứng dụng

- **Ý nghĩa:** Chuyển đổi mô hình quản lý truyền thống sang nền tảng tự động hóa thông minh, nâng cao hiệu suất và giảm thiểu sai sót
- **Tính ứng dụng:** Ứng dụng có thể được triển khai trong các thư viện, nhà sách cũng như trên nền tảng di động, hỗ trợ người dùng tra cứu thông tin sách nhanh chóng thông qua camera

3. Giá trị học thuật

- Dự án góp phần nghiên cứu khả năng thích ứng của thuật toán YOLO trên bộ dữ liệu tùy chỉnh (Custom Dataset) có độ nhiễu cao. Đồng thời, nghiên cứu các tối ưu hóa các tham số (Hyperparameters) để giải quyết bài toán nhận diện vật thể trong nhiều điều kiện ánh sáng cũng như góc độ

II. Cơ sở lý thuyết và công nghệ sử dụng

1. Thuật toán YOLO (You Only Look Once)

- YOLO là mô hình Object Detection theo kiến trúc “one-stage”, cho phép dự đoán vị trí (Bounding Box) và phân loại lớp (Class) chỉ qua một lần quét mạng neural.
- Theo Redmon et al (2016), YOLO vượt trội về tốc độ so với các mô hình “two-stage” như Faster R-CNN (Redmon, J et al., 2016,)

2. Công nghệ sử dụng

- **Ngôn ngữ:** Python 3.10+
- **Thư viện chính:**
 - + Ultralytics(YOLO)
 - + OpenCV
 - + Pytorch
- **Công cụ gán nhãn:** Roboflow hoặc CVAT để chuẩn hóa dữ liệu đầu vào

III. Phân tích yêu cầu dự án

1. Yêu cầu chức năng (Functional Requirements)

- **Thu thập dữ liệu:**
 - + **Quy mô:** Tối thiểu 400-500 ảnh bìa sách
 - + **Đa dạng hóa dữ liệu:**
 - Màu sắc
 - Kích thước
 - Góc chụp
 - Điều kiện sáng
 - Thể loại sách
 - + **Tiền xử lý:**
 - Chuẩn hóa kích thước ảnh
 - Gắn nhãn cho từng bìa sách
 - Chia dữ liệu thành các tập train, test, validation

- **Huấn luyện:**
 - + **Mục tiêu:** Mô hình nhận diện được ít nhất 5-10 loại đầu sách phổ biến
 - + **Quy trình huấn luyện:**
 - Chọn kiến trúc mô hình
 - Thiết lập tham số epoch, batch size, learning rate
 - Huấn luyện trên các tập và theo dõi các chỉ số loss, precision, recall, mAP để đánh giá tiến trình
 - + **Tối ưu hóa:**
 - Dùng các kỹ thuật tăng cường dữ liệu như xoay, lật, thay đổi độ sáng,...
 - Điều chỉnh cách tham số để cải thiện độ chính xác
- **Nhận diện:**
 - + **Đầu ra của mô hình:**
 - Khung bao (bounding box) quanh vùng chứa bìa sách
 - Tên sách được nhận diện
 - Độ tin cậy (confidence score) > 0.7 để đảm bảo độ tin cậy của kết quả
 - + **Đánh giá hiệu quả:**
 - So sánh kết quả nhận diện với nhãn gốc
 - Tính toán các chỉ số: Accuracy, Precision, Recall,...
 - Kiểm tra mAP để đánh giá tổng thể khả năng nhiều loại sách

2. Yêu cầu phi chức năng (Non-functional Requirements)

- **Độ chính xác:** Mô hình cần đạt chỉ số mAP ở ngưỡng IoU 0.5 tối thiểu từ (70% - 85%), nhằm đảm bảo khả năng nhận diện chính xác đối tượng trong đa số trường hợp
- **Thời gian đáp ứng:** Quá trình suy luận của máy phải có tốc độ xử lý phù hợp với mỗi trường hợp ảnh khi chạy trên GPU, đảm bảo tính khả dụng trong thời gian thực
- **Tính ổn định:** Hệ thống phải duy trì hiệu năng ổn định tốt trong các điều kiện phổ biến cũng như phức tạp đặc biệt là các môi trường ánh sáng không đạt tiêu chuẩn

IV. Kế hoạch thực hiện dự án

Giai đoạn	Nội dung công việc	Kết quả
Giai đoạn 1	Khảo sát, thu thập dữ liệu bìa sách thực tế	Tập hợp ảnh thô (Raw Images)
Giai đoạn 2	Gán nhãn (Annotation) và tiền xử lý dữ liệu	Dataset chuẩn YOLO (.txt)
Giai đoạn 3	Huấn luyện mô hình cơ bản và tinh chỉnh (fine-tuning)	File trọng số best.pt
Giai đoạn 4	Kiểm tra, đánh giá sai số và tối ưu hóa	Báo cáo đánh giá mô hình
Giai đoạn 5	Xây dựng giao diện demo và hoàn thiện báo cáo	Sản phẩm hoàn chỉnh

References

Jocher, G., et al. (2023). *ultralytics/ultralytics: Ultralytics YOLO* .

GitHub. Retrieved February 10, 2026, from

<https://github.com/ultralytics/ultralytics>

Redmon, J, Divvala, S, Girshick, R, & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. 10.1109/CVPR.2016.91