



BÁO CÁO MÔN HỌC DEEP LEARNING

Đề tài: "Phát hiện và cảnh báo xâm nhập sử dụng YOLOv8"

Nhóm thực hiện : Nhóm 4

Thai Nguyen University of Information
Technology and Communication

Ngày 7 tháng 5 năm 2024



Nội dung

- 1. Tổng quan về bài toán phát hiện và cảnh báo xâm nhập sử dụng YOLOv8**
- 2. Tổng quan về Deep Learning và mạng nơ-ron tích chập**
- 3. Phát hiện đối tượng và tổng quan về mô hình YOLO**
- 4. Ứng dụng cài đặt YOLOv8 cho bài toán nhận diện và cảnh báo xâm nhập**



Tổng quan về bài toán phát hiện và cảnh báo xâm nhập sử dụng YOLOv8



1.1 Giới thiệu về đề tài

- Là một ứng dụng trong lĩnh vực an ninh và giám sát.
- ập trung vào việc sử dụng mô hình YOLOv8 để phát hiện và định vị người xâm nhập trong hình ảnh hoặc video.
- Kết nối trực tuyến với ứng dụng telegram để nhận cảnh báo.



1.2 Mục tiêu của đề tài

- Vận dụng, triển khai cài đặt được mô hình YOLOv8 cho bài toán phát hiện người xâm nhập.
- Áp dụng trong các hệ thống an ninh như giám sát, bảo vệ tài sản cho cá nhân, doanh nghiệp v.v.



Tổng quan về Deep Learning và mạng nơ-ron tích chập

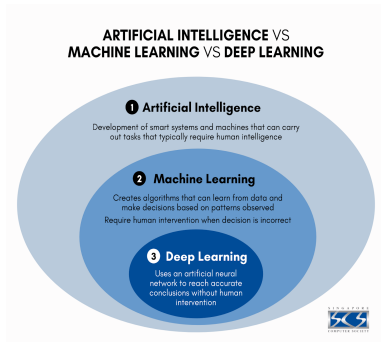


2.1 Khái niệm về Deep Learning

- Là một kỹ thuật học máy tiên tiến dựa vào việc học đại diện -> cho phép máy móc tự động học các biểu diễn tính năng cấp cao từ dữ liệu.
- Các thuật toán học sâu sử dụng mạng nơ-ron nhân tạo học các tính năng cấp cao từ dữ liệu bằng việc tăng độ sâu trong mạng.
- mạng lưới thần kinh được lấy cảm hứng từ một phần mạng lưới thần kinh sinh học, nơi có các tế bào kết nối và hoạt động cùng nhau.



2.1 Khái niệm về Deep Learning



Hình: 1. Deep Learning



2.2 Cách hoạt động của Deep Learning

- Bắt chước bộ não con người thông qua sự kết hợp của dữ liệu đầu vào, trọng số và bias.
- Bao gồm nhiều lớp kết nối với nhau, mỗi lớp được xây dựng dựa trên lớp trước đó để tinh chỉnh và tối ưu hóa.
- Sử dụng phương pháp lan truyền ngược, tính toán lỗi trong các dự đoán để điều chỉnh trọng số và độ lệch -> khắc phục sai số giúp mô hình đạt độ chính xác cao nhất.



2.3 Ứng dụng và lĩnh vực của học sâu

a. Lĩnh vực

- Xử lý ngôn ngữ tự nhiên Thị giác máy tính
- Học tăng cường

b. Ứng dụng

- Phân loại hình ảnh
- Nhận dạng giọng nói
- Phát hiện gian lận
- Trong y tế



2.4 Lịch sử mạng nơ-ron tích chập

- 1988, Yan Lecun sử dụng mạng tích chập trong tác vụ phân loại chữ số viết tay.
- 2009, bộ dữ li Image được giới thiệu là một trong những bộ dữ liệu đư sử dụng phổ biến cho những domain khác nhau.
- 2012 mạng Alexnet ra đời sử dụng tích chập CNN.
- Xuất hiện lần lượt các kiến trúc CNN mới VGGNet, GoogleNet, ResNet, DenseNet.



2.5 Đặc trưng chung của mạng nơ-ron tích chập

- Sử dụng tích chập
- Kiến trúc phân tầng
- Huấn luyện trên những bộ dữ liệu lớn
- Kích thước layer giảm dần
- Độ sâu tầng layer tăng dần
- Sử dụng các fully connected

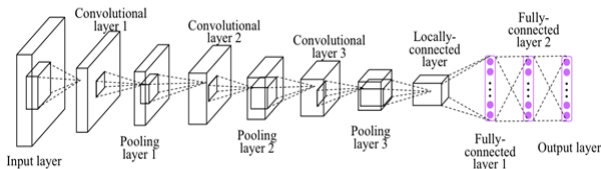


2.6 Khái niệm tích chập

- Tích chập là một khái niệm trong xử lý tín hiệu số nhằm biến đổi thông tin đầu vào thông qua phép tích chập đầu ra là một tín hiệu mới
- Các tín hiệu đầu ra sẽ làm giảm những đặc trưng mà bộ lọc không quan tâm và chỉ giữ những đặc trưng chính.



2.7 Kiến trúc chung của mạng nơ-ron tích chập



Hình: 2. Kiến trúc chung của mạng nơ-ron tích chập

- Quá trình tích chập
- Quá trình tổng hợp
- Quá trình kết nối hoàn toàn



Phát hiện đối tượng và tổng quan về mô hình YOLO



3.1 Khái niệm và tầm quan trọng của phát hiện đối tượng

a. Khái niệm

- Là một khía cạnh quan trọng của thị giác máy tính để xác định và định vị các đối tượng trong hình ảnh hoặc video.
- Dữ liệu thường ở dạng thời gian thực.

a. Tầm quan trọng của phát hiện đối tượng

- Cho phép nhận dạng và định vị tức thời các đối tượng trong môi trường động.
- Đưa ra các quyết định kịp thời.



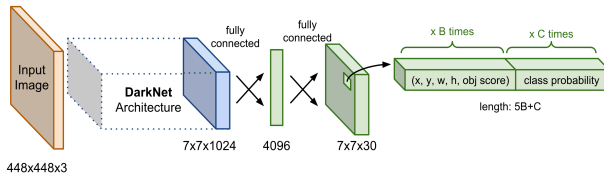
3.2 Định nghĩa về mô hình YOLO

- YOLO viết tắt của You Only Look Once.
- Phát hiện đối tượng trong một lần chuyển tiếp duy nhất của mạng.
- Cho phép xử lý và dự đoán đồng thời nhiều hình ảnh cực nhanh có độ chính xác và tốc độ cao.
- Phù hợp với các ứng dụng thời gian thực.



3.2 Kiến trúc mạng YOLO

- base network: là các mạng nơ ron làm nhiệm vụ trích xuất đặc trưng
- extra layers: giúp phát hiện các các vật thể trên feature map của base network.
- fully conected: giúp dự đoán nhãn và tọa độ bounding box của vật thể.



Hình: 3. Kiến trúc YOLOv3



Lịch sử các phiên bản YOLO

- Mới đây nhất đã có thêm phiên bản YOLOv9 đầu năm 2024.



Hình: 4. Lịch sử các phiên bản YOLO

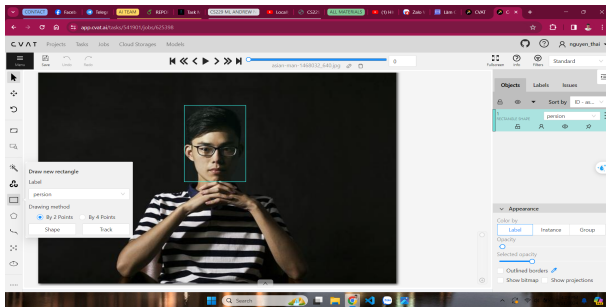


Ứng dụng cài đặt YOLOv8 cho bài toán nhận diện và cảnh báo xâm nhập



4.1 Thu thập dữ liệu và huấn luyện mô hình

- Thu thập 2000 ảnh người trên internet và tiến hành gán nhãn cho từng ảnh.

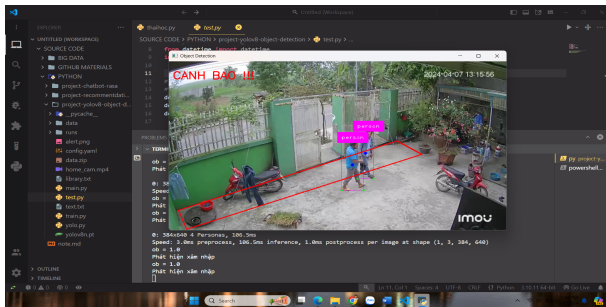


Hình: 5. Thu thập và gán nhãn dữ liệu



4.2 Demo chương trình

- Thực hiện huấn luyện mô hình YOLOv8 với tập dữ liệu trên 20 epochs.



Hình: 6. Demo chương trình sau khi huấn luyện



4.3 Cách thức hoạt động của mô hình YOLO

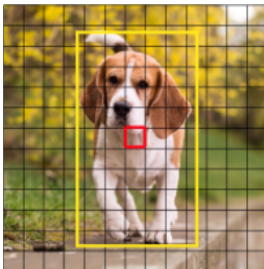
a. Các bước hoạt động

- Từ hình ảnh đầu vào, chia toàn bộ hình ảnh thành các ô lưới
- Đối với mỗi ô tiến hành dự đoán các hộp bao quanh vật thể (các ô là t ủa vật thể)
- Dự đoán lớp của đối tượng, dự đoán hộp ranh giới, xác suất các lớp có thể có trong hộp.
- Chọn ra các hộp ranh giới tốt nhất dựa vào ngưỡng và một số thuật toán như NMS để làm dự đoán cuối cùng.



4.4 Cách thức hoạt động của mô hình YOLO

b. Xác định các anchor box để ước lượng bounding box

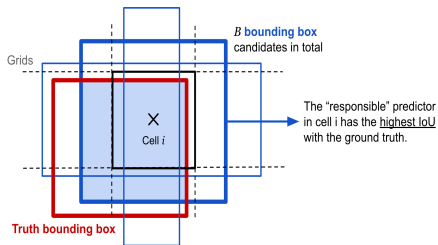


Hình: 7. Xác định tâm cho vật thể



4.4 Cách thức hoạt động của mô hình YOLO

b. Xác định các anchor box để ước lượng bounding box



Hình: 8. Ước lượng bounding box



4.4 Cách thức hoạt động của mô hình YOLO

c. Thuật toán Non Maximum Suppression và chỉ số IoU – Hiện tượng overlap



Hình: 9. Hiện tượng overlap



4.4 Cách thức hoạt động của mô hình YOLO

c. Thuật toán Non Maximum Suppression và chỉ số IoU

- Thuật toán tìm cách giảm bớt số lượng các bounding box bằng cách lọc các bounding box có xác suất vật thể nhỏ hơn một ngưỡng nào đó.



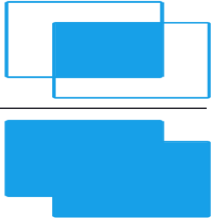
Hình: 9. Hiện tượng overlap



4.4 Cách thức hoạt động của mô hình YOLO

c. Thuật toán Non Maximum Suppression và chỉ số IoU

- Đối với các bounding box giao nhau, non-max sẽ lựa chọn ra một bounding box có xác suất chứa vật thể là lớn nhất và sau đó tính chỉ số IoU với các bounding box còn lại.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


Hình: 10. Chỉ số IoU



4.5 Đánh giá mô hình

a. Confusion matrix

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision Value $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

Hình: 11. Ma trận nhầm lẫn



4.5 Đánh giá mô hình

b. Recall và Precision

- Recall (độ phủ) đo lường tỷ lệ chính xác các trường hợp dương tính trên toàn bộ các mẫu thuộc nhóm dương tính:

$$recall = \frac{TP}{TP + FN}$$

- Precision dùng để xác định xác trường hợp là đúng trong các trường hợp được dự báo là dương tính:

$$precision = \frac{TP}{TP + FP}$$



4.5 Đánh giá mô hình

c. Average precision

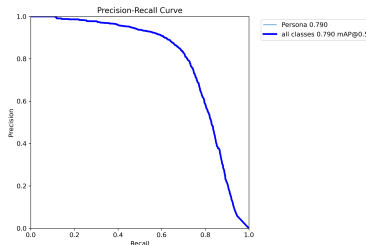
- Average precision (AP) hay còn gọi là Precision Recall Curve đơn giản là một biểu đồ có giá trị precision trên trục y và recall trên trục x.
- Độ chính xác ở đây còn được gọi là PPV (Positive Predictive Value), độ nhạy là TPR (True Positive Rate).



4.6 Đánh giá mô hình

d. mAP

- Mean Average Precsion (mAP) chính là trung bình cộng của các AP.
- Chỉ số này càng tiến gần tới 1 thì mô hình càng tốt.



Hình: Độ đo AP và mAP



Cảm ơn thầy cô và các bạn đã lắng nghe !

Nhóm thực hiện : Nhóm 4

Thai Nguyen University of Information
Technology and Communication

Ngày 7 tháng 5 năm 2024