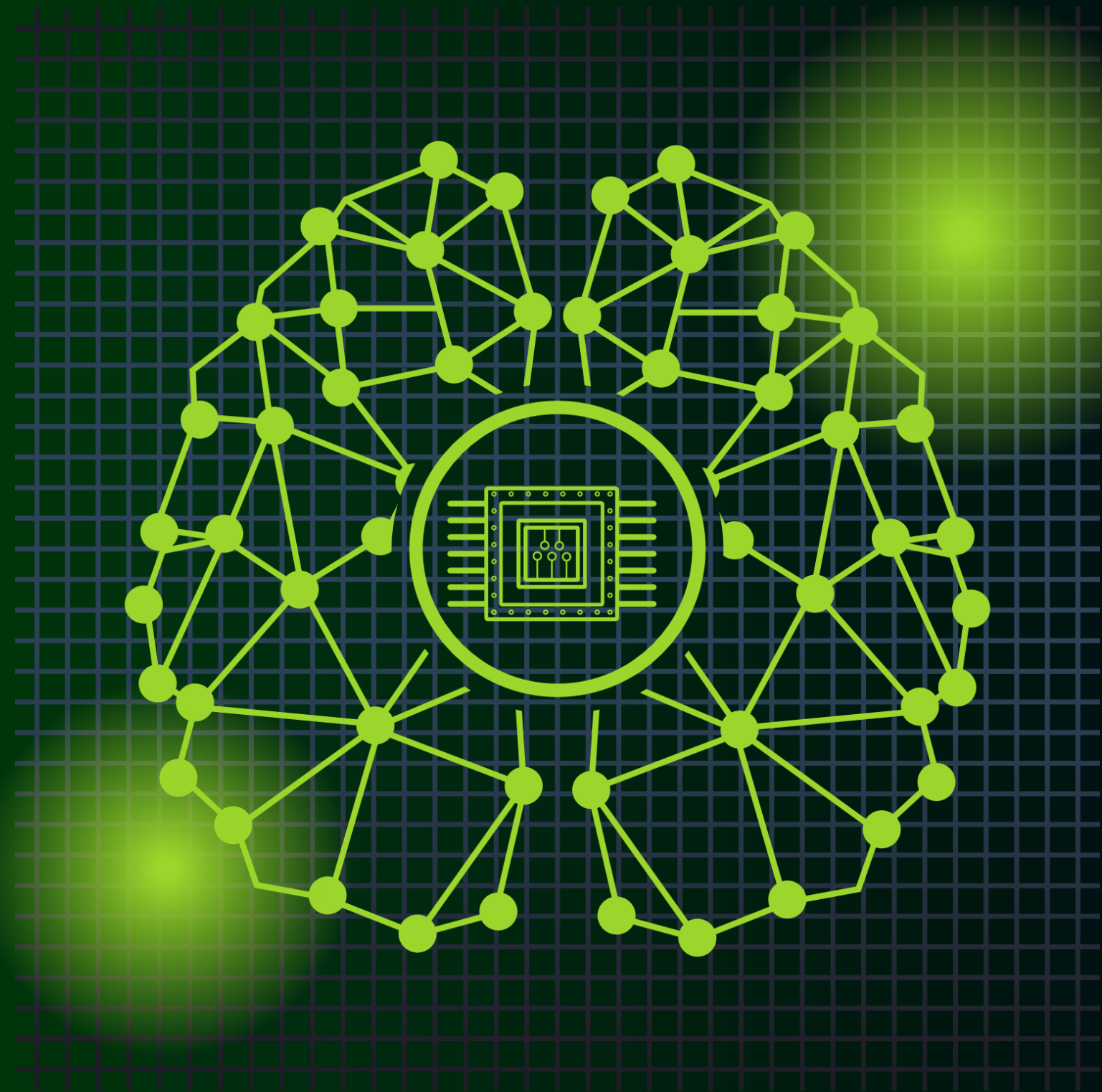


NHÓM 2 - IT3

XỬ LÝ ẢNH VÀ THỊ GIÁC MÁY TÍNH



Giảng viên hướng dẫn: Lương Thị Hồng Lan

PHÂN CÔNG NHIỆM VỤ

Họ và tên	Nhiệm vụ
Nguyễn Thế Mạnh	Code nhận diện đối tượng, viết báo cáo, nội dung slide
Lương Thị Lan Anh	Kiểm thử, cài đặt thư viện, viết báo cáo, thuyết trình
Hồ Anh Nam	Code giao diện ứng dụng, viết báo cáo, nội dung slide
Bùi Quang Trung	Code đếm đối tượng, viết báo cáo, nội dung slide
Trần Quang Minh	Thiết kế giao diện, viết báo cáo, thiết kế slide

CÁC NỘI DUNG CHÍNH

01 Phân tích bài toán

02 Xây dựng bài toán với yolo

03 Các bước thực hiện bài toán với mô hình Yolo

04 Thử nghiệm và demo sản phẩm

PHÂN TÍCH BÀI TOÁN

Bài toán xây dựng hệ thống nhận diện và đếm đối tượng trong ảnh nhằm tạo ra một hệ thống thông minh, tự động nhận diện và đếm đối tượng bằng mô hình học sâu YOLO, một thuật toán phát hiện nhanh và hiệu quả.

01 Thu thập và Chuẩn bị Dữ liệu

Thu thập ảnh: Tìm kiếm ảnh đa dạng về góc độ, ánh sáng, vị trí và kích thước đối tượng.

Đánh dấu dữ liệu: Gắn nhãn và tạo bounding box cho từng đối tượng trong ảnh.

Chia dữ liệu: Phân chia ảnh thành tập huấn luyện, kiểm định và thử nghiệm để đánh giá mô hình.

02 Xây dựng Mô Hình YOLO

Chọn kiến trúc: Chọn phiên bản YOLO phù hợp với yêu cầu bài toán.

Cấu hình mô hình: Tùy chỉnh hyperparameter (số lớp, kích thước ảnh, learning rate,...) để tối ưu hiệu suất.

03 Huấn luyện Mô Hình

Đưa dữ liệu vào: Huấn luyện mô hình với dữ liệu đã đánh dấu.

Tối ưu hóa: Dùng thuật toán tối ưu hóa để tối ưu tham số.

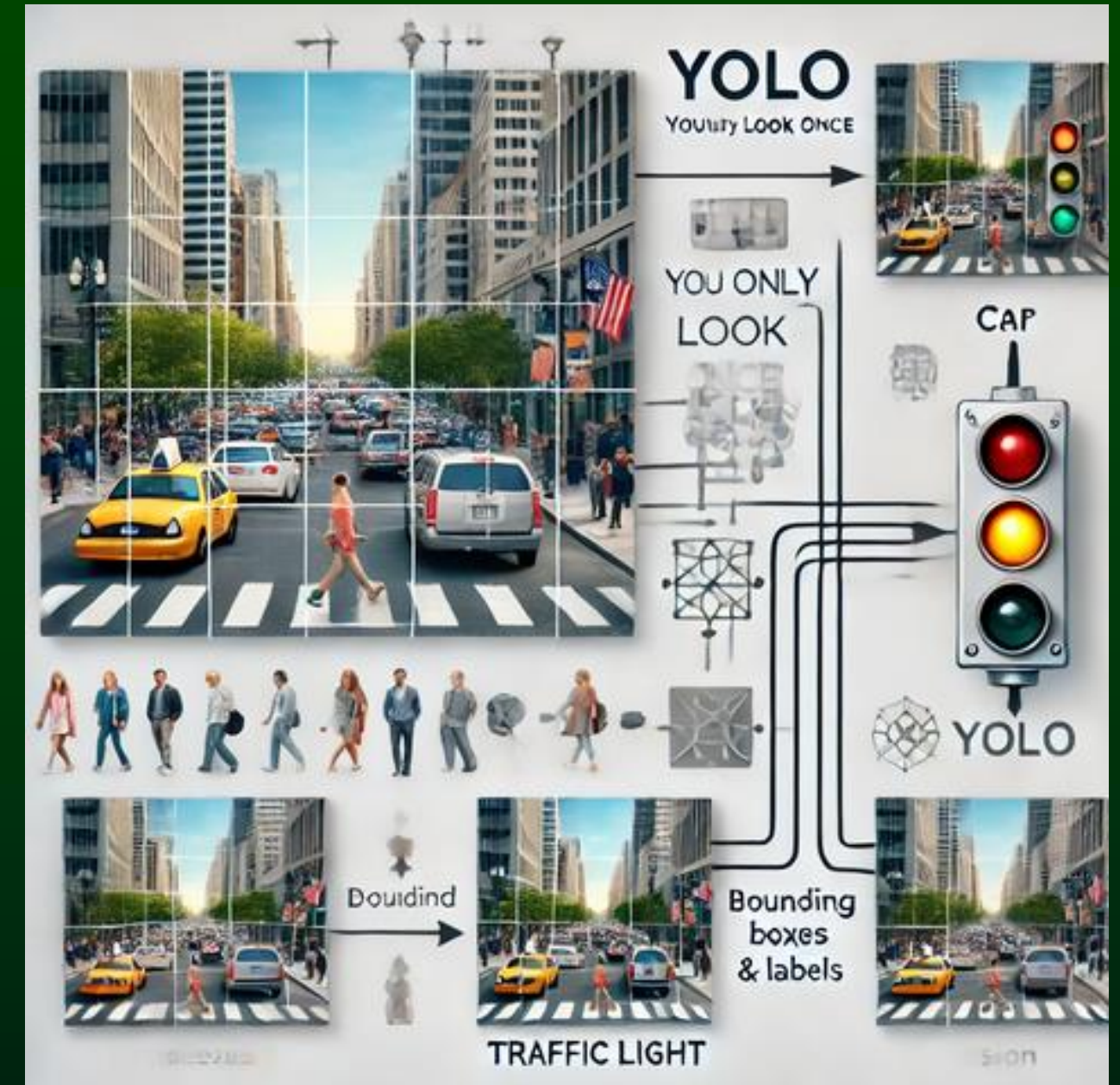
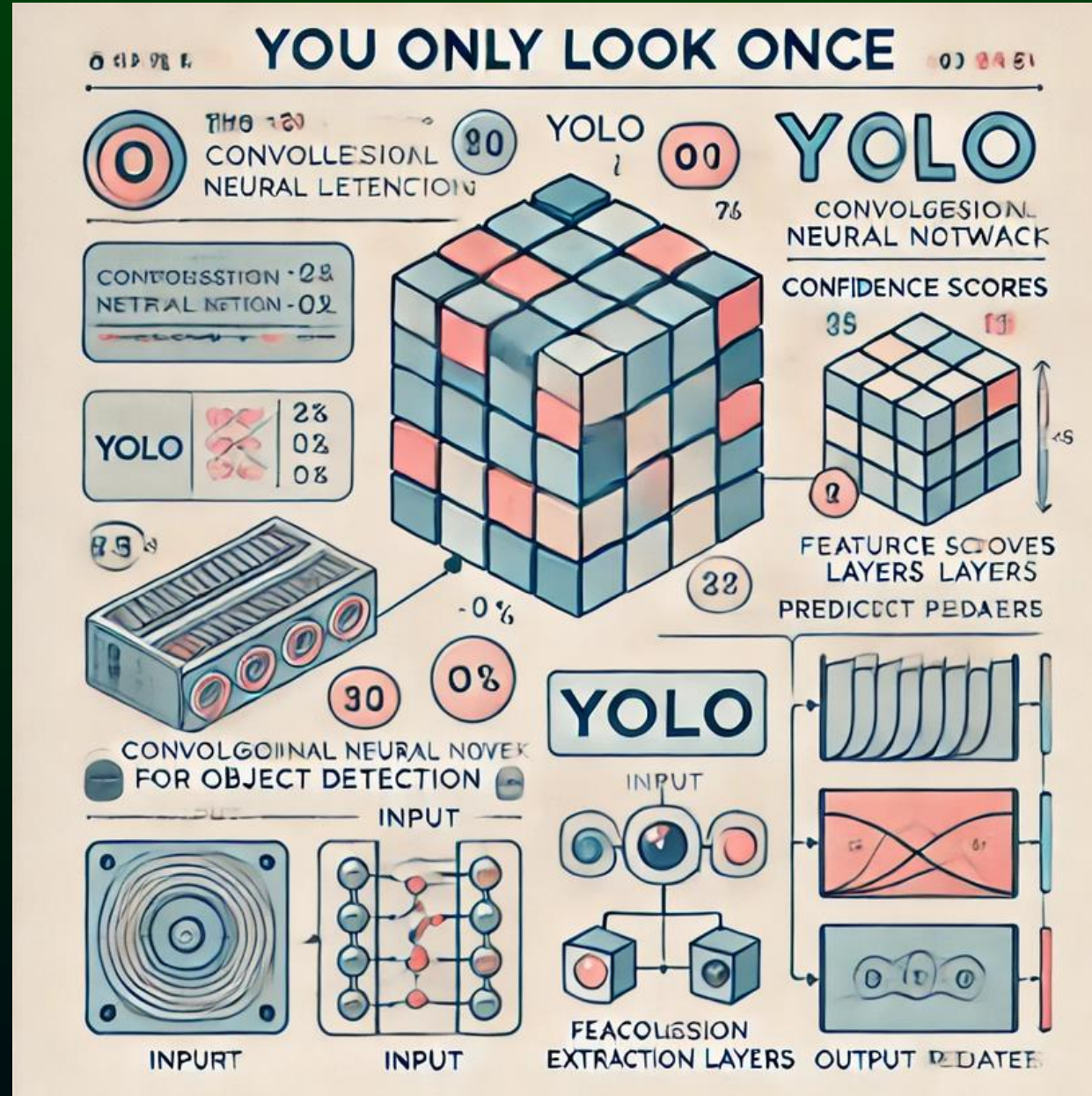
Đánh giá: Đánh giá hiệu suất qua các chỉ số như mAP, độ chính xác, độ nhạy.

04 Triển khai Hệ Thống

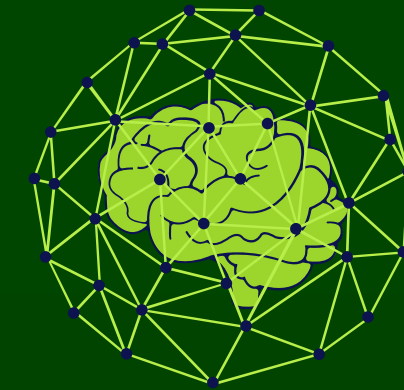
Tích hợp mô hình vào ứng dụng để sử dụng. Tiền xử lý ảnh đầu vào phù hợp với mô hình. Dự đoán bounding box và nhãn đối tượng. Hiển thị kết quả trên ảnh gốc.

MÔ HÌNH YOLO

YOLO (You Only Look Once) là một trong những mô hình hàng đầu cho bài toán nhận dạng và phát hiện đối tượng. Nó tích hợp nhiều kỹ thuật hiệu quả để tăng tốc độ và độ chính xác. Dưới đây là các kỹ thuật chính thường sử dụng trong mô hình YOLO, bao gồm tư tưởng kỹ thuật, hình minh họa, cùng ưu nhược điểm.



MÔ HÌNH YOLO



1. Grid-based Detection

- Tư tưởng kỹ thuật:
 - Chia ảnh thành lưới $S \times SS \times SS$, mỗi ô lưới dự đoán bounding boxes và xác suất đối tượng.
- Ưu điểm:
 - Xử lý nhanh, chỉ cần một lần qua mạng.
 - Kết hợp phát hiện và phân loại.
- Nhược điểm:
 - Hạn chế với đối tượng nhỏ ở giữa ô lưới.

2. Anchor Boxes

- Tư tưởng kỹ thuật:
 - Sử dụng các hộp neo (anchor boxes) có kích thước và tỷ lệ đa dạng.
- Ưu điểm:
 - Phát hiện đối tượng kích thước khác nhau.
 - Xử lý tốt đối tượng chồng chéo.
- Nhược điểm:
 - Tăng phức tạp khi số lượng anchor boxes lớn.

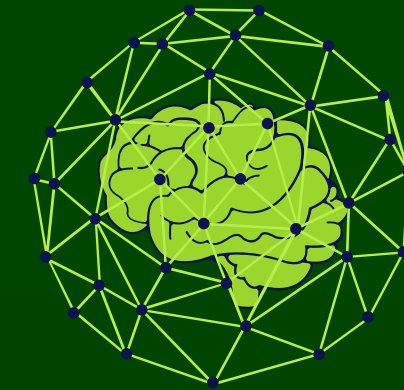
3. Non-Maximum Suppression (NMS)

- Tư tưởng kỹ thuật:
 - Loại bỏ các bounding boxes không tối ưu, giữ lại hộp có độ tin cậy cao nhất.
- Ưu điểm:
 - Loại bỏ trùng lặp, đảm bảo một đối tượng chỉ được phát hiện một lần.
- Nhược điểm:
 - Phụ thuộc ngưỡng confidence và IoU.

4. Feature Pyramid Network (FPN)

- Tư tưởng kỹ thuật:
 - Kết hợp thông tin từ các tầng đặc trưng để phát hiện đối tượng kích thước khác nhau.
- Ưu điểm:
 - Hỗ trợ phát hiện đối tượng nhỏ và lớn.
- Nhược điểm:
 - Tăng độ phức tạp và thời gian xử lý.

MÔ HÌNH YOLO



5. Objectness Score

- Tư tưởng kỹ thuật:
 - Dự đoán objectness score cho bounding boxes để xác định vùng chứa đối tượng.
- Ưu điểm:
 - Giảm dự đoán sai.
 - Đơn giản hóa quá trình lọc bounding boxes.
- Nhược điểm:
 - Phụ thuộc vào ngưỡng score.

6. Intersection over Union (IoU)

- Tư tưởng kỹ thuật:
 - Đo lường mức độ khớp giữa bounding box dự đoán và thực tế.
- Ưu điểm:
 - Đánh giá chính xác độ khớp.
 - Hỗ trợ NMS hiệu quả.
- Nhược điểm:
 - Kém hiệu quả cho đối tượng nhỏ hoặc chồng chéo.

7. Data Augmentation

- Tư tưởng kỹ thuật:
 - Tăng cường dữ liệu bằng các biến đổi như xoay, cắt, hoặc thay đổi màu sắc.
- Ưu điểm:
 - Tăng tính tổng quát của mô hình.
 - Giảm nguy cơ overfitting.
- Nhược điểm:
 - Tăng cường không hợp lý có thể làm giảm hiệu quả.

8. Loss Function

- Tư tưởng kỹ thuật: Kết hợp 3 thành phần:
 - Sai lệch vị trí (Localization Loss).
 - Sai lệch độ tin cậy (Confidence Loss).
 - Sai lệch nhãn lớp (Class Prediction Loss).
- Ưu điểm:
 - Tối ưu hóa đa mục tiêu.
 - Giảm mất mát cho negative samples.
- Nhược điểm:
 - Cần cân bằng trọng số giữa các thành phần.

CÁC BƯỚC THỰC HIỆN BÀI TOÁN VỚI MÔ HÌNH YOLO

1. THU THẬP DỮ LIỆU

- Bộ dữ liệu sử dụng:
 - COCO: Bộ dữ liệu quy mô lớn cho phát hiện, phân đoạn, và chú thích đối tượng, hỗ trợ nghiên cứu và đánh giá mô hình thị giác máy tính.
 - PASCAL VOC: Bộ dữ liệu tiêu chuẩn với 20 lớp đối tượng phổ biến (người, động vật, vật thể hàng ngày).
 - Bộ dữ liệu tùy chỉnh: Dùng khi cần nhận diện các đối tượng đặc biệt, yêu cầu thu thập ảnh và gắn nhãn thủ công.
- Gắn nhãn dữ liệu:
 - Bounding Boxes: Xác định vị trí và phân loại các đối tượng trong ảnh.
 - Công cụ gắn nhãn phổ biến:
 - LabelImg: Công cụ mã nguồn mở, dễ sử dụng.
 - VGG Image Annotator (VIA): Gắn nhãn trực tuyến, tiện lợi.

2. TIỀN XỬ LÝ DỮ LIỆU

- Chuyển đổi dữ liệu:
 - Thay đổi kích thước ảnh: Điều chỉnh kích thước ảnh về chuẩn (416x416 hoặc 608x608) để phù hợp với yêu cầu YOLO.
 - Chuyển đổi nhãn: Biến thông tin đối tượng thành định dạng YOLO (gồm lớp đối tượng và tọa độ bounding box theo tỷ lệ ảnh).
- Tăng cường dữ liệu (Data Augmentation):
 - Kỹ thuật phổ biến:
 - Lật ảnh ngang/dọc.
 - Xoay ảnh một góc ngẫu nhiên.
 - Thay đổi độ sáng và độ tương phản.
 - Thêm nhiễu ngẫu nhiên.
 - Thư viện hỗ trợ: OpenCV và Albumentations.



3. XÂY DỰNG MÔ HÌNH VÀ HUẤN LUYỆN

- Hàm mất và cấu hình YOLO:
 - Framework phổ biến:
 - Darknet: Framework YOLO gốc, nhẹ và nhanh.
 - Ultralytics YOLOv5: Phiên bản dễ sử dụng, nhiều tính năng.
- Cấu hình tham số:
 - Số lớp đối tượng (classes).
 - Số anchor boxes.
 - Kích thước ảnh đầu vào.
- Huấn luyện mô hình:
 - Hàm mất mát: Kết hợp giữa:
 - Mất mát vị trí bounding box.
 - Mất mát xác suất đối tượng.
 - Mất mát phân loại.
 - Optimizer: Thường dùng Adam hoặc SGD.
 - Quy trình: Truyền dữ liệu → Tính độ mất mát → Cập nhật trọng số bằng backpropagation.
- Lưu mô hình: Sau huấn luyện, lưu mô hình và trọng số để sử dụng trong các ứng dụng nhận diện đối tượng.



4. XÂY DỰNG MÔ HÌNH VÀ HUẤN LUYỆN

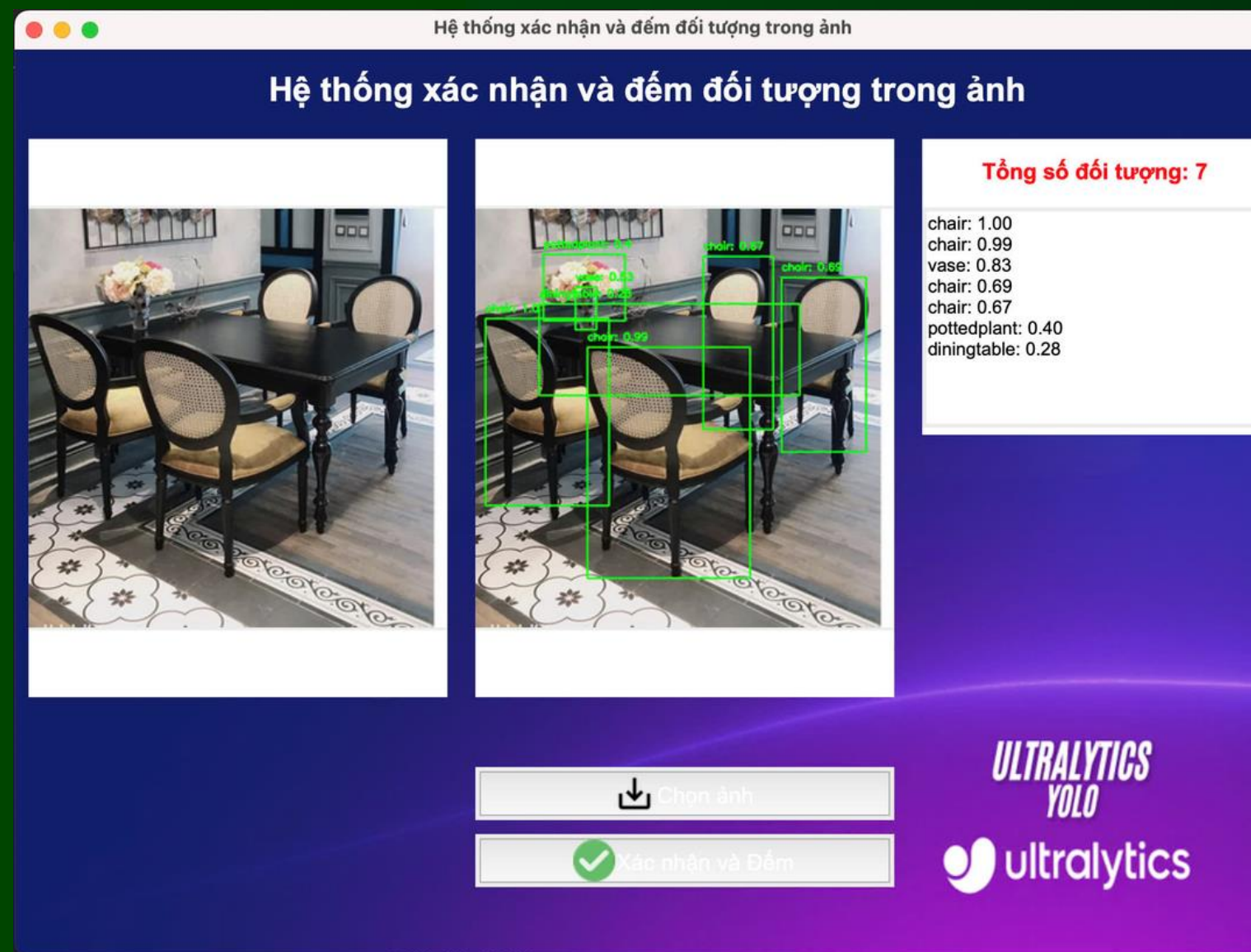
- Tải mô hình đã huấn luyện:
 - Nạp tệp trọng số (weights file) của mô hình YOLO để sử dụng trong dự đoán.
- Dự đoán đối tượng:
 - Nhận đầu vào là ảnh và phân tích toàn bộ ảnh.
 - Đầu ra gồm các bounding boxes với nhãn và xác suất tương ứng.
- Hậu xử lý bằng Non-Maximum Suppression (NMS):
 - Loại bỏ các bounding boxes trùng lặp.
 - Giữ lại bounding box chính xác nhất cho mỗi đối tượng.
- Hiển thị kết quả:
 - Vẽ bounding boxes lên ảnh và hiển thị.
 - Có thể lưu lại ảnh đã được nhận diện.



THỬ NGHIỆM VÀ DEMO SẢN PHẨM

input: là 1 bức ảnh đồ vật

output: là Nhận diện đối tượng là đồ vật, gồm 7 đối tượng, 4 cái ghế, 1 lọ hoa, 1 bó hoa, 1 cái bàn với độ chính xác tương đối đạt



THỬ NGHIỆM VÀ DEMO SẢN PHẨM

input: là 1 bức ảnh động vật

output: Nhận diện đối tượng là chó, gồm 5 con, với độ chính xác gần như tuyệt đối

Hệ thống xác nhận và đếm đối tượng trong ảnh



Tổng số đối tượng: 5

- dog: 1.00
- dog: 0.98
- dog: 0.95
- dog: 0.93
- dog: 0.71

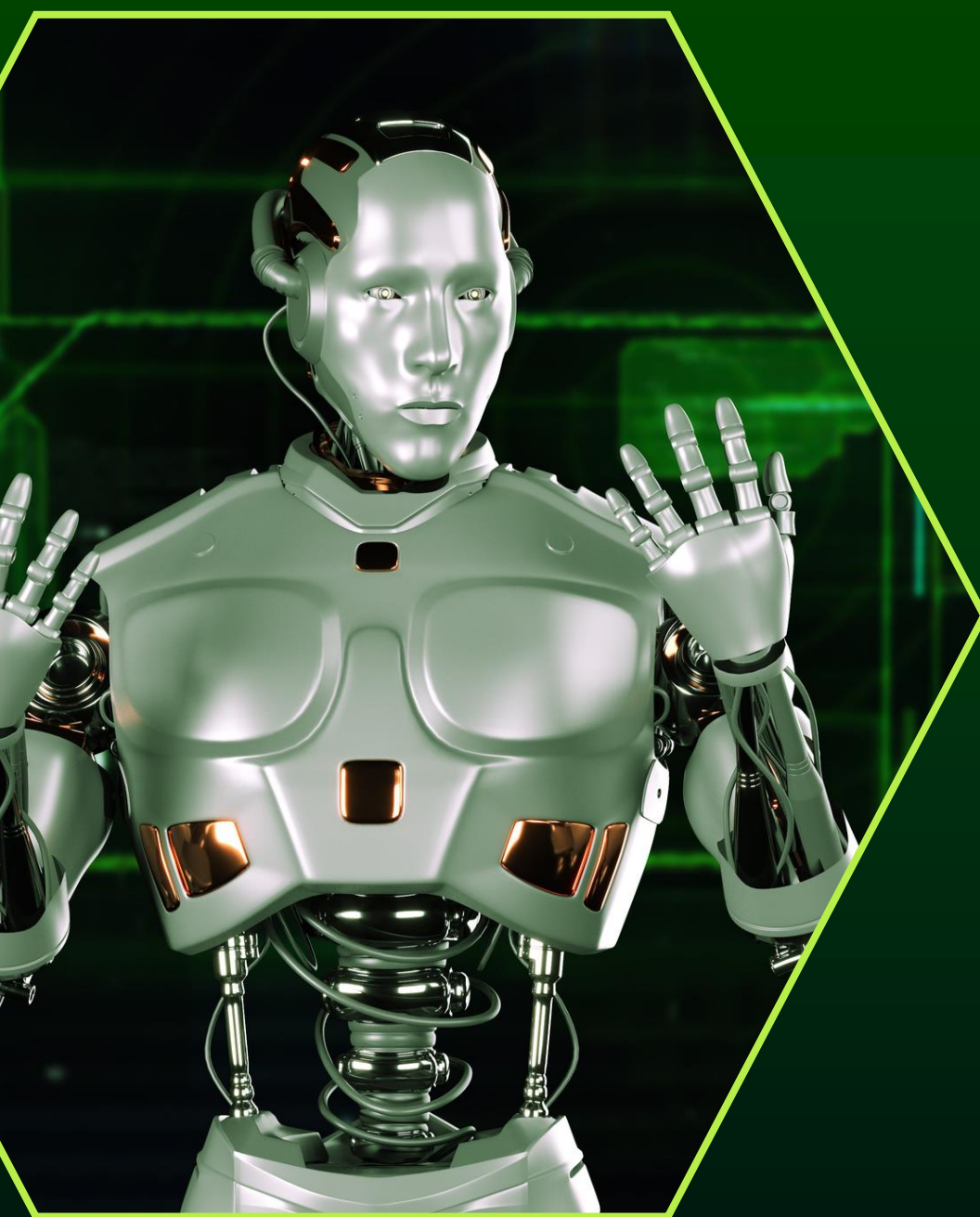
Chọn ảnh

Xác nhận và Đếm

ULTRALYTICS
YOLO

ultralytics





**PHẦN TRÌNH BÀY CỦA NHÓM EM ĐẾN
ĐÂY KẾT THÚC, CẢM ƠN CÔ VÀ CÁC
BẠN ĐÃ LẮNG NGHE !**

