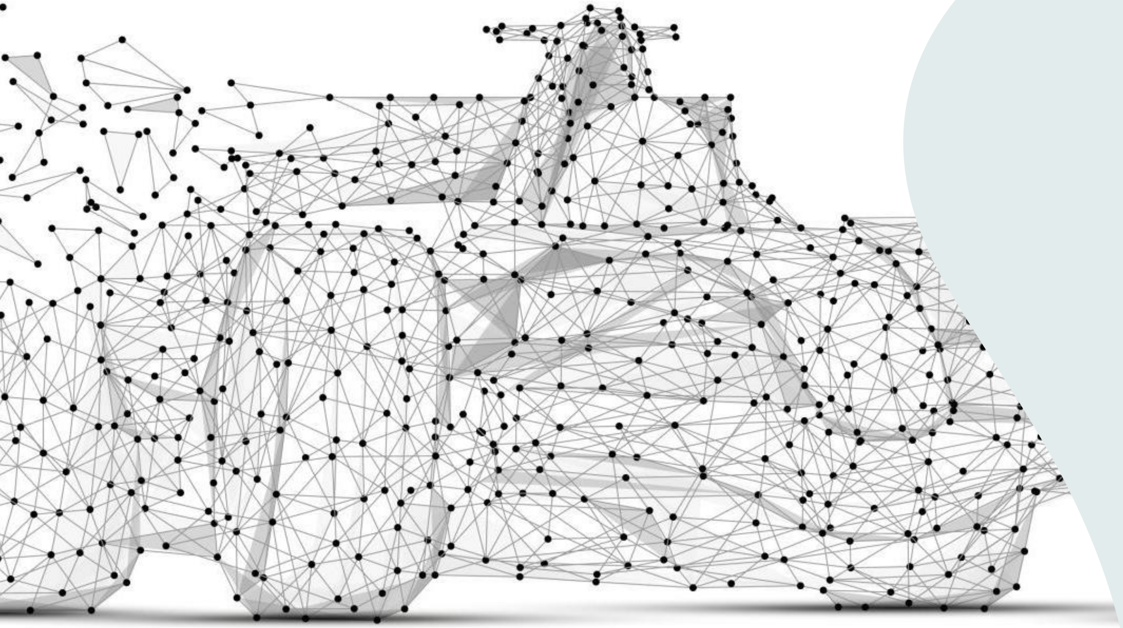
SECJ 3563 Computational Intelligence

Group: F4NT4STIC

Ruhul Quddus Tamim
Md Yusuf Bin Forkan
Md Rasheduzzaman
Shafi Ahmed
Syafiq Ibnu Ramadhan

Application for vehicle detection



Introduction

Vehicle detection-based computer vision is an essential part of perception in autonomous driving.

Designing a visual intelligence system that replaces substitute perception is a tricky task. As a critical part of intelligent perception systems, vehicle detection is one of the classic scientific issues. Thus, current perception-based autopilot algorithms are primarily solving the problem of assisting human drivers to ensure traffic safety.

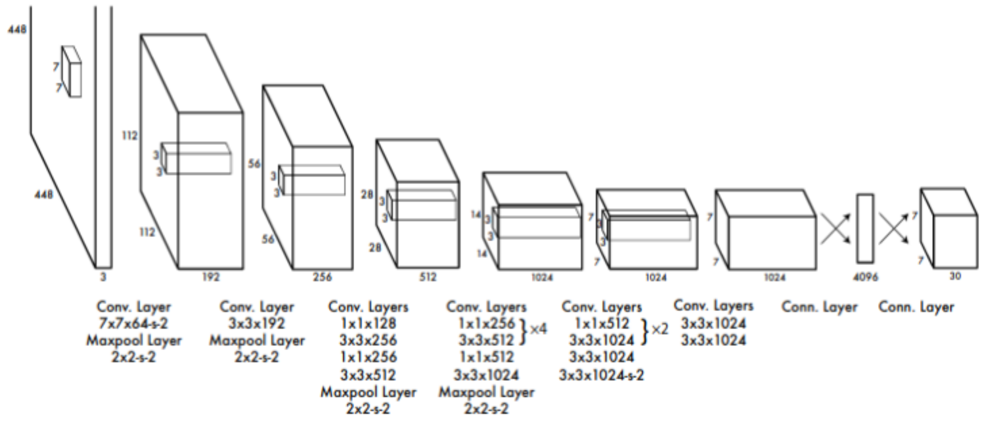
Unified Detection

This presentation talks about the method leverages labeled detection images to learn to precisely localize objects while it uses classification images.

We unify the separate components of object detection into a single neural network. Our network uses features from the entire image to predict each bounding box. It also predicts all bounding boxes across all classes for an image simultaneously.

Network Design

Our network architecture is inspired by the GoogLeNet model for image classification. Our network has 24 convolutional layers followed by 2 fully connected layers.

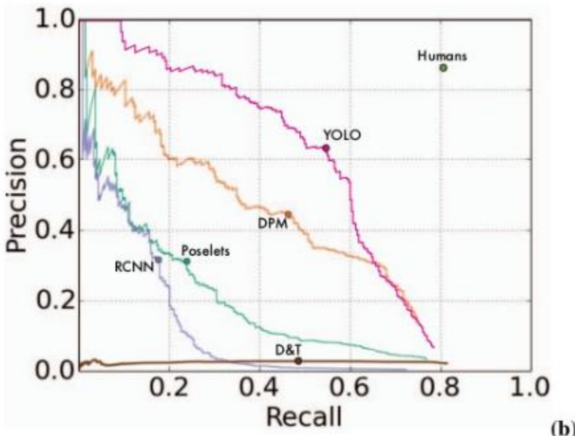


Training

Training is expensive. We pretrain our convolutional layers on the ImageNet 1000-class competition data-set. For pretraining we use the first 20 convolutional layers followed by an average-pooling layer and a fully connected layer.

Comparison to Other Real-Time Systems

YOLO is a fast, accurate object detector, making it ideal for computer vision applications. YOLO is connected to a webcam and verify that it maintains real-time performance, including the time to fetch images from the camera and display the detections. The resulting system is interactive and engaging. While YOLO processes images individually, when attached to a webcam it functions like a tracking system, detecting objects as they move around and change in appearance.



(a) Picasso Dataset precision-recall curves.

	VOC 2007 AP	Picasso AP	Picasso Best F_1	People-Art AP
YOLO	**59.2**	**53.3**	**0.590**	**45**
R-CNN	54.2	10.4	0.226	26
DPM	43.2	37.8	0.458	32
Poselets [2]	36.5	17.8	0.271	
D&T [4]	-	1.9	0.051	

(b) Quantitative results on the VOC 2007, Picasso, and People-Art Datasets. The Picasso Dataset evaluates on both AP and best F_1 score.

Figure 5: Generalization results on Picasso and People-Art datasets.

Conclusion

Thus, we introduce YOLO, a unified model for object detection. Our model is simple to construct and can be trained directly on full images. Unlike classifier-based approaches, YOLO is trained on a loss function that directly corresponds to detection performance and the entire model is trained jointly.

Thank you!