INTEL UNNATI INDUSTRIAL TRAINING
SUMMER 2023

ROAD OBJECT DETECTION WITH DEEP LEARNING

TEAM NAME : THE BRAINIACS

AUTHORS :

1. Tadakala Likitha
2. Kakarla Deepika

DATE OF SUBMISSION : 15TH July 2023

ABSTRACT

Road object detection with deep learning is a technique used to identify and classify objects that appear on the road, such as vehicles, pedestrians, and traffic signs, using deep learning algorithms. This approach has become a key component in various applications, including autonomous driving, advanced driver-assistance systems (ADAS), and traffic analysis. Deep learning models for road object detection typically use convolutional neural networks (CNNs) to extract features from images and make predictions. The model is trained on large datasets to learn the complex patterns and characteristics of different objects. The accuracy and performance of on-road object detection with deep learning depend on various factors, including the quality and diversity of the training data, the chosen model architecture, etc.

INTRODUCTION

Road safety is essential in today's world, and along with it autonomous driving is one of the most anticipated technologies of the 21st century. Autonomous driving attempts to navigate without human intervention ensuring safety. It includes major challenges for computer vision and machine learning. The process of road object detection with deep learning involves data collection, preprocessing, model selection, training, evaluation, optimization, and deployment. Deep learning models, particularly convolutional neural networks (CNN), are used to extract features from road images and make predictions about the presence and location of objects. By analyzing road scenes in real time, road object detection systems can provide valuable information for various applications. The advancements in deep learning and computer vision continue to improve the accuracy and performance of road object detection systems, allowing for more reliable and precise detection of objects on the road.

MOTIVATION

Our cities and roads are very unpredictable dynamic environments where multiple aspects such as pedestrians, animals, streets, or other vehicles coexist together. In this way, it is needed to provide Autonomous Vehicles with robust perception systems to correctly understand the environment, and therefore be able to interpret what is happening in the surroundings to act in consequence. Regarding software development, in the last years, significant advances have been made remarkably in the fields of Deep Learning and specifically Convolutional Neural Networks (CNNs). CNNs have supposed an important breakthrough, and its results are the solutions to problems such as object detection and classification or semantic segmentation and understanding. This context motivates us to participate in the research and development of new and robust Convolutional Neural Network-based algorithms that will perform a key role in the perception systems of autonomous vehicles, substituting standard computer vision approaches.

PRIOR WORK

We haved researched on YOLOv5 which is an object detection model that belongs to the YOLO (You Only Look Once) family of algorithms. It is an evolution of the YOLOv4 model and introduces several improvements and optimizations. YOLOv5 aims to provide a good balance between accuracy and inference speed for real-time object detection tasks.

The few criteria that we discovered are

Architecture: YOLOv5 follows a one-stage object detection approach, where it predicts bounding boxes and class probabilities directly from the input image. It employs a backbone network (typically a convolutional neural network, such as CSPDarknet53) followed by detection heads at different scales to predict objects of varying sizes.

Model Variant: YOLOv5 comes in different model sizes, such as YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, where each variant has a different number of layers and parameters. Smaller models are faster but may sacrifice some accuracy, while larger models provide better accuracy but with increased computational cost.

Backbone and Neck: YOLOv5 utilizes a modified backbone network, typically CSPDarknet53 or EfficientNet, to extract features from the input image. It also incorporates a neck module (e.g., PANet or FPN) to fuse features at different scales and improve detection performance for objects of various sizes.

Training Strategy: YOLOv5 employs a combination of techniques for training, including label smoothing, focal loss, and data augmentation (e.g., random scaling, flipping, and rotation). It also supports transfer learning, where pre-trained models on large-scale datasets like COCO or ImageNet can be fine-tuned on domain-specific datasets.

Inference Speed: YOLOv5 is designed for real-time or near real-time object detection. It achieves fast inference speeds due to its efficient architecture and optimized implementation.

Open-Source Implementation: YOLOv5 is available as an open-source project, with the code and pre-trained models publicly accessible on GitHub. This allows researchers and developers to use, modify, and adapt YOLOv5 for their specific applications.

APPROACH

The approach towards Road Object Detection with Deep Learning includes various steps that are essential for the model development.They are:

- Data Collection: Collect a dataset of road images or videos that include the objects you want to detect, such as vehicles, pedestrians, and road signs. Ensure the dataset covers a diverse range of road scenarios and lighting conditions.

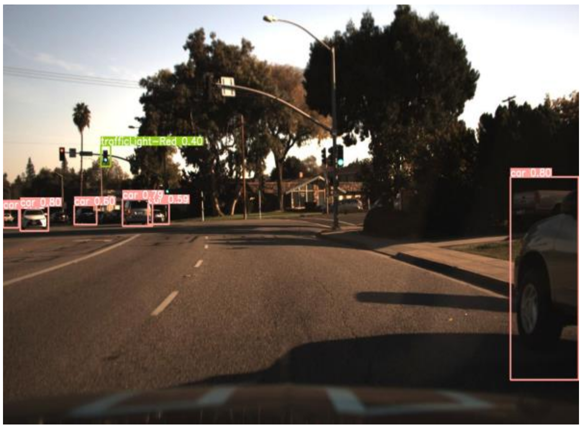
  Data set used : https://roboflow.com/

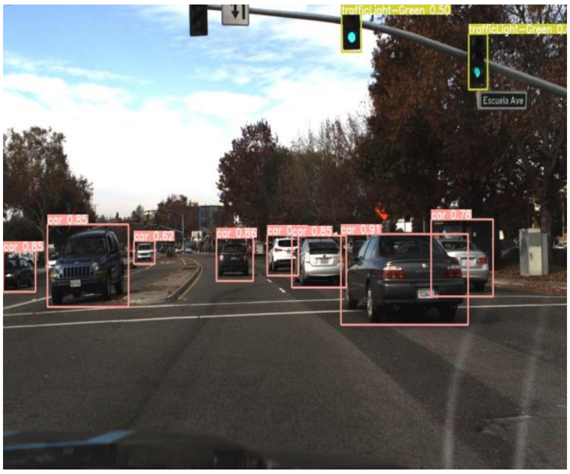
- Data Preprocessing: Resize and normalize the images to ensure they are compatible with YOLOv5's input requirements. Split the dataset into training and validation sets, maintaining a balanced distribution of object classes in both sets.
- Model Configuration: Configure the YOLOv5 model architecture according to your specific requirements. Decide on the model size (e.g., YOLOv5s, YOLOv5m, YOLOv5l, or YOLOv5x) based on the trade-off between accuracy and inference speed. Adjust other parameters, such as the number of detection layers and the anchor box sizes, to suit the characteristics of your dataset.
- Model Training: Train the YOLOv5 model using the preprocessed dataset. Utilize the training data to optimize the model's parameters for accurate object detection. You can use frameworks like PyTorch and the YOLOv5 implementation available on GitHub for training. Monitor the training process, adjust hyperparameters if necessary, and ensure convergence.

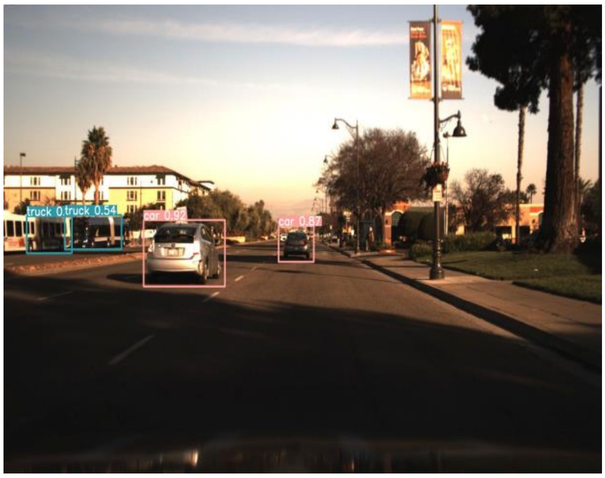
  Github repository for yolov5: https://github.com/ultralytics/yolov5

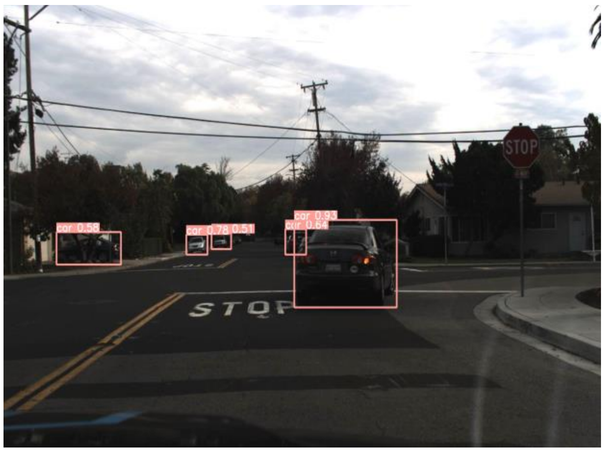
- Model Evaluation: Evaluate the trained model's performance on the validation set.Analyze the results to identify any shortcomings and areas for improvement.

RESULTS









REFERENCES

- Efficient Road Object Detection Using Deep Learning and Transfer Learning" by Y. Zhang, Z. Chen, and H.
- Deep Convolutional Neural Networks for Road Object Detection" by M. D. Zeiler and R. Fergus.
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection.

LINK TO SOLUTION