Intel® Unnati Industrial Training Project, 2023

Report on:

Road Object Detection with Deep learning

Team TRON
Members:
Noel John Robert (noeljohnrobert.b20cs1147@mbcet.ac.in)
Josh Danny Alex (joshdannyalex.b20cs1131@mbcet.ac.in)
Nubi Fathima N (nubifathiman.b20cs1149@mbcet.ac.in)

Internal Mentor
Mr. Shibu VS
Asst. Prof Dept of CSE(MBCET)
(shibu.vs@mbcet.ac.in)

Mar Baselios College of Engineering and Technology (Autonomous)
Mar Ivanios Vidyanagar
Nalanchira P.O, Thiruvananthapuram,
Kerala, India
Pin: 695015

Aim: To develop a model that could identify objects in roads, specifically Indian roads.

Abstract:

This project presents a robust and efficient approach to road object detection using the YOLOv5 deep learning framework. By leveraging the capabilities of YOLOv5, the project achieves detection and classification of various road objects, including cars, pedestrians, trucks etc. The YOLOv5 architecture enables efficient object detection, and it is trained using a carefully designed process that includes learning from previous knowledge (pre-trained weights). Experimental evaluations confirm the system's accuracy, speed, and robustness, demonstrating its effectiveness in real-world scenarios and potential integration into autonomous vehicles, transportation systems, and traffic surveillance applications. Overall, this project contributes to the field of computer vision and autonomous systems, emphasizing the significance of precise and efficient detection systems in enhancing road safety, traffic management, and transportation efficiency.

Motivation to choose Road Object Image Detection with Deep learning:

Choosing road object image detection using deep learning as a project is motivated by several factors. Firstly, it addresses a real-world problem by improving road safety, traffic management, and autonomous systems. Secondly, the project enables technological innovation, pushing the boundaries of accuracy and robustness in object detection. Thirdly, it facilitates learning, skill development, collaboration, and networking opportunities, leading to academic and career growth in the field of computer vision and artificial intelligence. Additionally, contributing to the open-source community further enhances the project's impact.

In summary, the motivation for choosing this project stems from its ability to address real-world challenges, advance technology, foster learning and collaboration, and contribute to the broader community. By developing road object image detection using deep learning, we can make meaningful contributions to road safety, transportation efficiency, and the field of computer vision, while also gaining valuable skills and networking opportunities.

Approach:

The dataset collected was IDD_Dataset provided by IIIT Hyderabad and it had a size of 22.8 GB, the dataset had two folders named JPEGImages and Annotations. There are also three files train.txt, val.txt and test.txt these contain the path to specific files.

The model that we used was YOLOv5 and it had five different model configurations mainly nano, small, medium, large and extra large. We had to write a program to convert annotations in XML format that was already present in the dataset to Pascal VOC format that was needed for YOLOv5.

Training was initially done with the small model configuration having a batch size of 8 and 50 epochs. Using the weights obtained from this training we re-trained on the same dataset using the medium configuration and using the weights obtained from this, we re-trained the large configuration.

Training was completed and the final precision we obtained was around 78%, we decided to proceed on this as further training did not yield better results.

Validation was done using the already provided data and results are also included.

Random images containing the objects were taken from the internet and testing was done on that and its results are also included.

Detailed instructions of the same are provided in the README.md file in the Github repository.

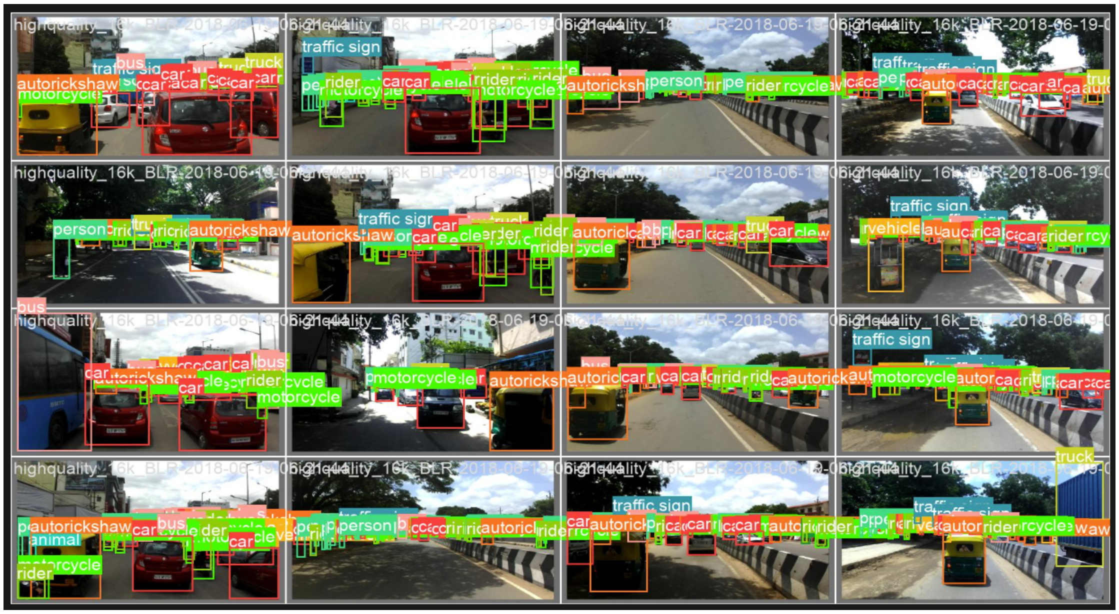
Challenges that we encountered

- The dataset exhibited limitations in its quality, as certain bounding boxes were observed to have dimensions smaller than 3 pixels. Consequently, these small bounding boxes presented challenges for accurate detection by the model.
- A program had to be developed so as to facilitate the conversion of existing XML format annotations in the dataset to the required Pascal VOC format for YOLOv5.
- Despite extensive efforts, we encountered limitations in our ability to surpass the accuracy of our model.

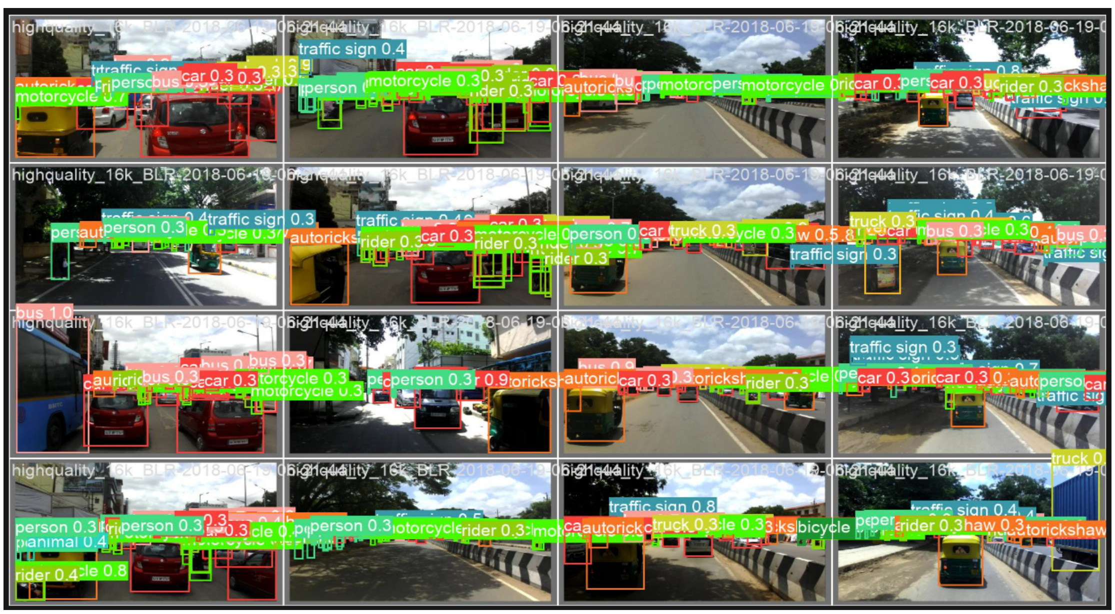
Result

Three rounds of training were done, first with small, then medium and finally with large size model - 50 epochs each and batch size 8. Weights from previous training were used to assist the next model. Results for training are present in models/yolov5/runs/train/<exp_no>.

Given below are few examples taken from the validation set:
The first image was created using annotation present in the dataset.



And the below was obtained after prediction by model.



Here, we see that the model was in fact able to detect most of the objects, but the enormous number of objects within a single frame and their small sizes will really hamper the effectiveness of the model.

Precision-Confidence curve, Precision-Recall curve, and Recall-Confidence curve are also present in the results folder. The results.csv file contains losses, precision, and other metrics for each epoch. The final value of precision was 0.78335 and that of recall was 0.35397.

Random images were also taken from the internet to detect objects, and results for these are present within models/yolov5/runs/detect/<exp_no>.