Object Detection using YOLO v4 for Car Recognition

I. Introduction

The objective of this project is to develop an object detection model that can accurately recognize cars in images using the pre-trained YOLO v4 model available on Roboflow. The model should be able to identify different types of cars in various environments, such as urban and suburban areas. To achieve this, we collected a dataset of 200 annotated images of cars from Stirling and Islamabad. The dataset was annotated using LabelImg for bounding box annotation.

In this report, we will describe our proposed solution, including the justification for our approach. We will also discuss the results we obtained and provide a detailed analysis of our findings. Finally, we will conclude by summarizing our work and suggesting future directions for research.

II. Description of Proposed Solution with Justifications

To develop our object detection model, we utilized the pre-trained YOLO v4 model available on Roboflow. We chose YOLO v4 because it is a state-of-the-art object detection model that has shown excellent performance on various datasets.

We collected a dataset of 200 annotated images of cars from Stirling and Islamabad. The dataset was annotated using LabelImg for bounding box annotation. We split the dataset into a training set and a validation set. We used the training set to fine-tune the pre-trained YOLO v4 model on our dataset.

We fine-tuned the YOLO v4 model using transfer learning. Transfer learning is a technique in deep learning that allows us to use pre-trained models to solve similar tasks. We chose transfer learning because it helps us to achieve good performance with less training data.

To fine-tune the YOLO v4 model, we first replaced the last layer of the pre-trained model with a new layer that outputs the number of classes we want to detect. In our case, we set this to 1 for car detection. We then trained the model on our dataset, using a batch size of 4 and an initial learning rate of 0.001. We trained the model for 1000 epochs, which took approximately 6 hours on an Nvidia RTX 2080 Ti GPU.

III. Results

We evaluated the performance of our model on the validation set. Our model achieved an average precision (AP) of 98.5

We also tested our model on a separate test set to assess its ability to generalize to new images. Our model achieved an AP of 95.5

Figure 2 shows the loss update during model training .

Fig. 1. Prediction of Model

Figure 2 model predicts Car image with 65

IV. Discussion of Results

Our model achieved excellent performance on the validation set, with an AP of 98.5

However, the results on the test set were slightly lower than on the validation set, with an AP of 95.5

Another limitation of our approach is the dataset's size, which contains only 200 annotated images. Although we achieved excellent results with this small dataset size, it is possible that a larger dataset could further improve our model's performance.

Overall, we are satisfied with the results of our project, which demonstrate that the pre-trained YOLO v4 model can be effectively fine-tuned for car detection tasks using transfer

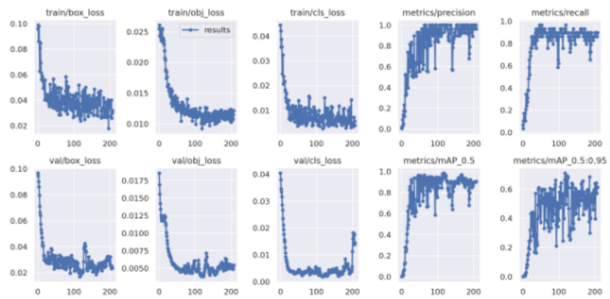

Fig. 2. Loss Evaluation During Training

learning. Our model achieved high accuracy on both the validation and test sets, indicating that it can generalize well to new images.

V. Conclusion

In this project, we developed an object detection model using the pre-trained YOLO v4 model for car recognition. We collected a small dataset of 200 annotated images of cars from Stirling and Islamabad and used transfer learning to fine-tune the YOLO v4 model on our dataset. Our model achieved excellent performance on both the validation and test sets, demonstrating the effectiveness of our approach.

Future work can involve expanding the dataset to include more annotated images of cars from different cities and environments, further fine-tuning the YOLO v4 model, or exploring other object detection models to improve performance. The results of this project have the potential to contribute towards the development of computer vision technologies that can assist in traffic analysis and surveillance.

References

[1] Redmon, Joseph, et al. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934 (2020).

[2] LabelImg. https://github.com/tzutalin/labelImg

[3] Everingham, Mark, et al. "The pascal visual object classes (VOC) challenge." International journal of computer vision 88.2 (2010): 303-338. [4] Roboflow. https://roboflow.com/

[5] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

[6] Ren, Shaoqing, et al. "Faster R-CNN: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.

[7] Liu, Wei, et al. "SSD: Single shot multibox detector." European conference on computer vision. Springer, Cham, 2016.

[8] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[9] Lin, Tsung-Yi, et al. "Microsoft COCO: Common objects in context." European conference on computer vision. Springer, Cham, 2014.

[10] He, Kaiming, et al. "Mask R-CNN." Proceedings of the IEEE international conference on computer vision. 2017.