VEHICLE DETECTION AND COUNTING IN IMAGES

Anurag Hambir Ryan Furtado FNU Himanshu Akash Bhapkar

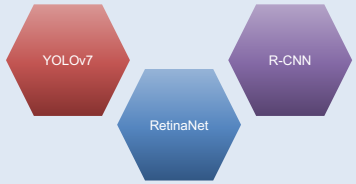
Indiana University – Luddy School of Informatics, Computing and Engineering

Introduction

This project is inspired by the problem of heavy traffic observed on streets and highways which tends to be a major problem faced by a lot of metropolitan cities these days especially during holiday periods or during peak work hours. We aim to count the number of vehicles on such roads so that we can take preventive measures to control the flow of traffic by detecting the presence of traffic in real time and suggesting alternative routes to drivers beforehand.

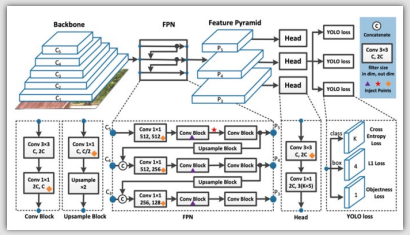
Many researchers have used various models to work on this kind of problem. Researchers have employed PASCAL VOC, Counting CNN, KLT Tracker and Deep Neural Networks.

To solve this problem, we will train models using the following well known architectures and compare their performance in order to understand their strengths and weaknesses.



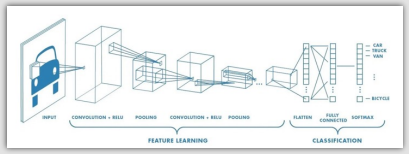
Methodology

YOLO is a single stage object detection technique based on efficient layer aggregation network (ELAN). ELAN controls the shortest and longest gradient path to make the network efficient thereby allowing deeper networks to be trained easily.



RetinaNet is a single stage object detection technique that extracts features at various scales from an input image using a feature pyramid network, and then utilizes a feature fusion module to fuse the features at various scales. It uses a focal loss function which assigns higher weights to harder examples in order to address the class imbalance problem.

Region Based CNN(R-CNN) is a deep learning two stage object detection algorithm. It has categorization and localization responsibilities as two distinct stages. A list of potential object regions is generated during the first step, along with boxes around the regions. Next, Convolutional Neural Networks are applied to each region to produce fixed-length feature vectors. The final step is to transmit feature vectors through an object categorization and Non-Maximal suppression layer.

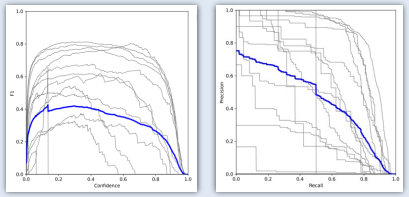


Experiments

- We have used Traffic Image Dataset from Kaggle consisting of around 3000 images of traffic in the country of Bangladesh.
- The dataset contains annotations in the format of text files with bounding boxes.
- We trained **YOLOv7**, **RetinaNet** and **R-CNN** on the Traffic Image Dataset from Kaggle consisting of 2704 training images.
- We also ran different experiments with 40, 70 and100 epochs with varying batch sizes of 8, 12 and 16.
- The test data contains 300 images and calculated Mean Average Precision (MAP) according to different parameters.
- The Mean Average Precision (MAP) is calculated on the correct count of the objects (vehicles) in the test dataset.
- The table below shows the results on each model, with YOLOv7 performing best for our dataset.

Mean Precision Error (MAP)						
Number Of Epochs	40		70		100	
Batch Size	16	32	16	32	16	32
YOLOv7	40.8	38.3	35.3	33.6	32.8	30.5
Retina Net	43.6	42.5	40.4	39.2	36.7	35.4
R-CNN	44.7	42.8	41.5	40.8	38.6	36.3

Here are some of the F-1 and precision-recall curves for the models that we have trained.



Results

With our dataset, YOLOv7 performs better than RetinaNet and R-CNN, since it does a better job of predicting the count of the objects in the images with respect to the ground truth. The results can be seen below.

YOLOv7 Input Images Vs. Predicted Images

RetinaNet Input Images Vs. Predicted Images

R-CNN Input Images Vs. Predicted Images

Conclusion

Based on the experiments on traffic dataset we concluded that Yolov7 is the fastest and the most accurate model out of the three models.

Practitioners can use these models in real time traffic detection and suggest alternative routes to avoid traffic congestion and in case of people, avoid overcrowding.

These models were trained on a very small dataset of around 3000 images and yet they gave amazing results. If trained with more data and on better hardware architecture, they will definitely be faster and provide more accurate results.

Future Work

- These models can be repurposed to be used on videos for efficient real time predictions giving us a better idea of their performance against each other.
- Distinguishing between different types of vehicles as well as their make and model can also be considered for denser prediction tasks.

References

- Road Vehicle Images Dataset: https://www.kaggle.com/datasets/ashfakyeafi/road-vehicle-images-dataset
- P. Gao, J. Tian, Y. Tai, T. Zhao and Q. Gao, "Vehicle Detection with Bottom Enhanced RetinaNet in Aerial Images," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020, pp. 1173-1176, doi: 10.1109/IGARSS39084.2020.9323216.
- P. Gao, J. Tian, Y. Tai, T. Zhao and Q. Gao, "Vehicle Detection with Bottom Enhanced RetinaNet in Aerial Images," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020, pp. 1173-1176, doi: 10.1109/IGARSS39084.2020.9323216.
- M. Manana, C. Tu and P. A. Owolawi, "Preprocessed Faster RCNN for Vehicle Detection," 2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC), Mon Tresor, Mauritius, 2018, pp. 1-4, doi: 10.1109/ICONIC.2018.8601243.
- M. V., V. V. R. and N. A., "A Deep Learning RCNN Approach for Vehicle Recognition in Traffic Surveillance System," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0157-0160, doi: 10.1109/ICCSP.2019.8698018.
- Song, H., Liang, H., Li, H. et al. Vision-based vehicle detection and counting system using deep learning in highway scenes. Eur. Transp. Res. Rev. 11, 51 (2019). https://doi.org/10.1186/s12544-019-0390-4
- R. Kejriwal, R. H J, A. Arora and Mohana, "Vehicle Detection and Counting using Deep Learning basedYOLO and Deep SORT Algorithm for Urban Traffic Management System," 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, 2022, pp. 1-6, doi: 10.1109/ICEEICT53079.2022.9768653.