SPARK Program Student Research Project, Summer 2022

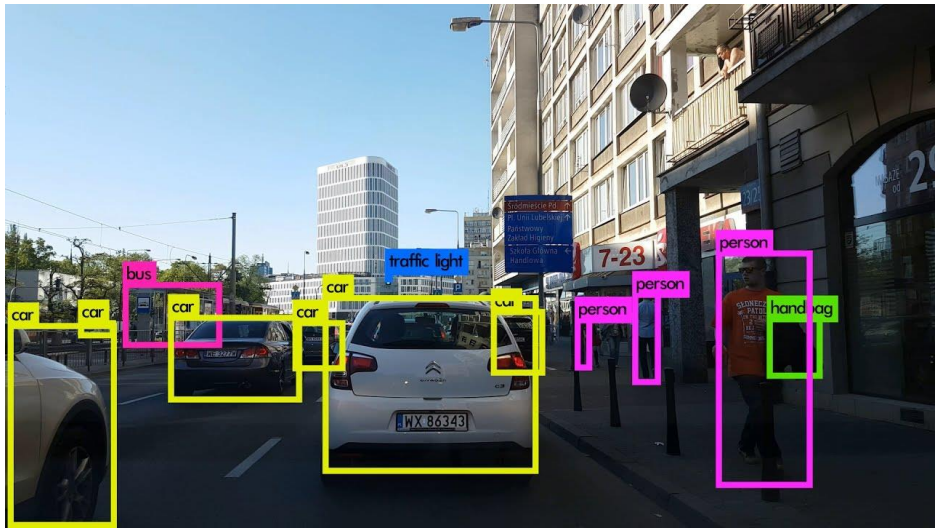
Faster R-CNN Analysis in Autonomous Vehicles

Student: Aadit Juneja
Advisor*:* Xiaoyu Xie

Northwestern | McCORMICK SCHOOL OF ENGINEERING

Motivation and Objective

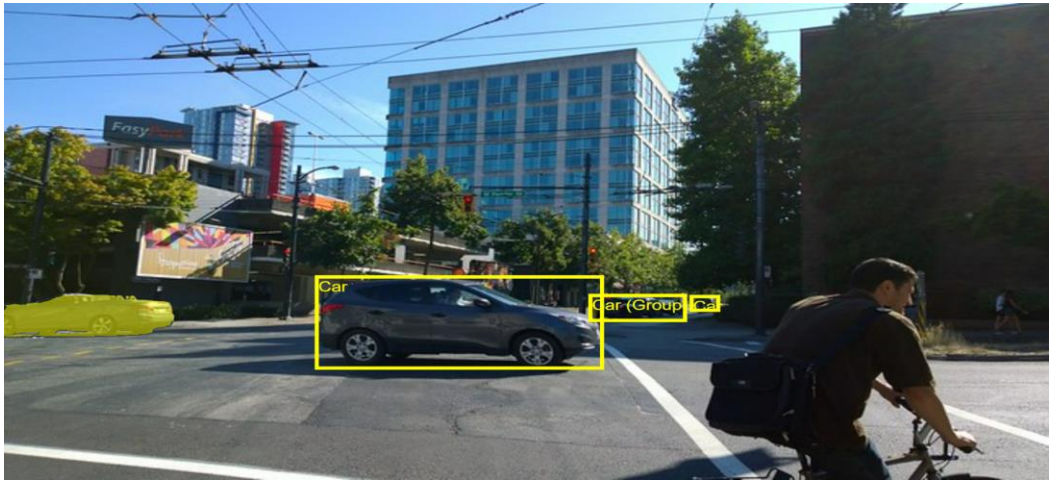
- Autonomous vehicles object detection algorithms have to run quickly in practice to be effective
- Hyperparameters in object detection models like Faster R-CNN must be tuned to minimize average testing time
- Objective: Analyze the relationship between hyperparameters such as learning rate, number of hidden layers, etc. and average testing time of a Faster R-CNN



https://public.roboflow.com/object-detection/self-driving-car

Dataset Information

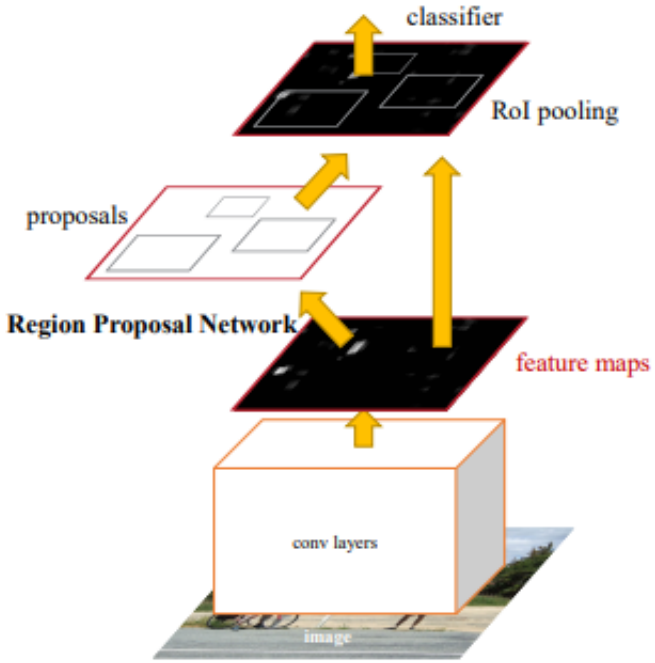
- Google Open Images v6
- Selected 3 classes: car, person, bicycle, 800 images for each class
- Only used the labels of those 3 classes in those images (e.g. building labels ignored)



https://storage.googleapis.com/openimages/web/index.html

Faster R-CNN

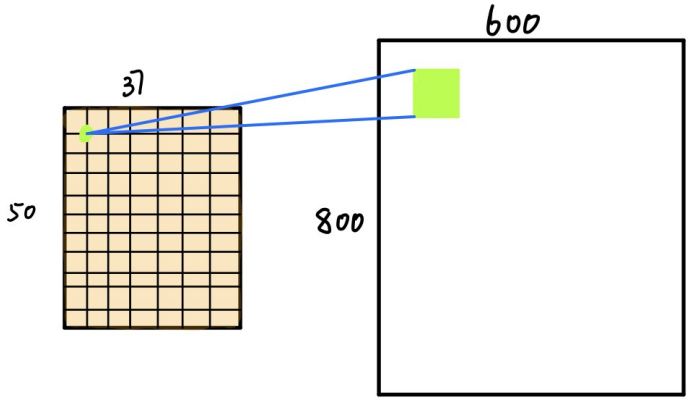
- Faster R-CNN released in 2016 by eliminating the need for fixed search algorithm to generate region proposals
- Built on R-CNN (2013) and Fast R-CNN (2015) models
- One of the most popular object detection algorithms today alongside YOLO and SSD



https://paperswithcode.com/lib/torchvision/faster-r-cnn

Faster R-CNN

- First stage of model is Region Proposal Network (RPN)
- Image initially sent through pretrained model (VGG16) to generate feature map
- Each point on feature map is mapped back to 9 boxes (anchors) through combinations of 3 scales and 3 aspect ratios
- Anchors sent through convolutional layer
- RPN outputs classification (object or background) and regression (bounding box)
- RPN trained using IoU method

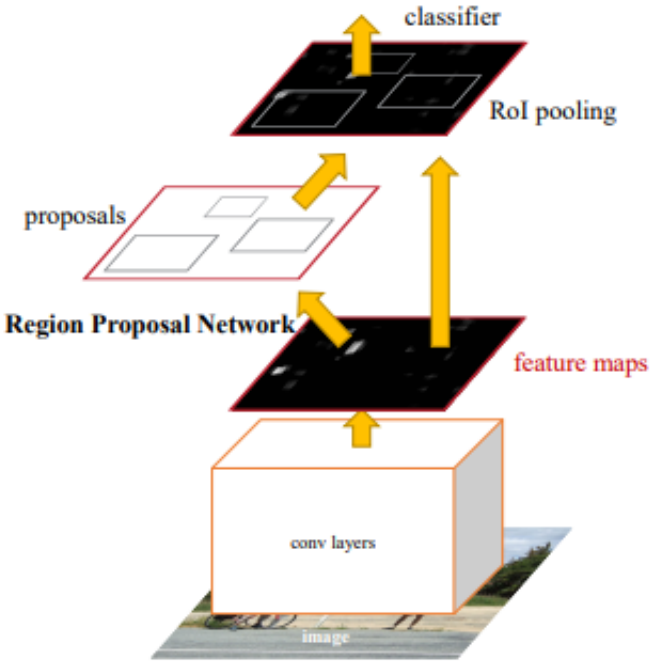


https://towardsdatascience.com/faster-r-cnn-object-detection-implemented-by-keras-for-custom-data-from-googles-open-images-125f62b9141a

https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/

Faster R-CNN

- Anchors classified as objects now called Regions of Interest (ROI)
- ROIs sent through pooling layers
- Flattened and sent through fully connected layers
- Outputs classification (car, person, bicycle) and regression (bounding box)



https://paperswithcode.com/lib/torchvision/faster-r-cnn

https://public.roboflow.com/object-detection/self-driving-car

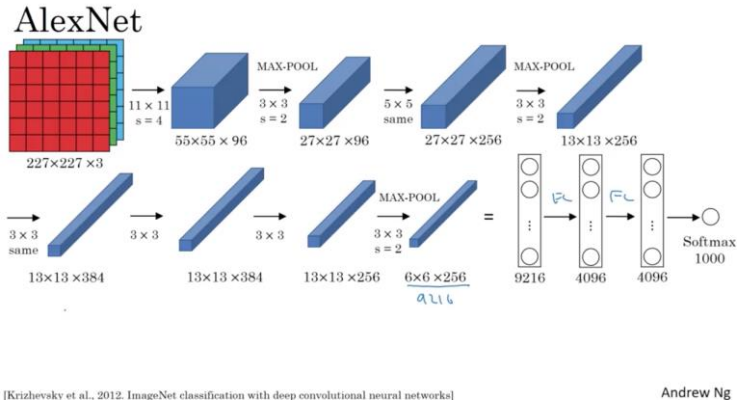
Results

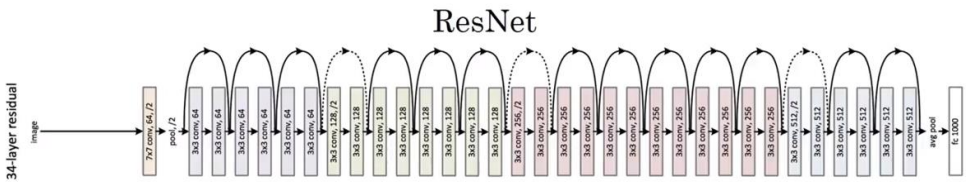
Number of FC Layers	Recall	Average Testing Time (s)
3	72.56%	0.691
5	75.88%	0.691
7	76.91%	0.692
15	78.26%	0.693

Recall: The number of correctly detected objects divided by the total number of objects in the image

Interpretation of Results and Future Work

- Little difference is seen in average testing time, which was likely due to VGG16 having 138 million parameters, while each additional fully connected layer adds about 1000
- 15 layers could be an overfit result, as it is possible that the recall was maximized at a number of FC layers between 7 and 15. Alternatively, recall may still be increasing at 15 layers
- Future exploration can be done by examining the difference in average testing time when changing the base model from VGG16 to models like ResNet 50(~23 million parameters) or LeNet(~61 million parameters)





https://medium.com/@RaghavPrabhu/cnn-architectures-lenet-alexnet-vgg-googlenet-and-resnet-7c81c017b848

Acknowledgements

- Xiaoyu Xie for advising me on this project
- Professor Wing Kam Liu for organizing the MDS experience
- The whole MDS group for such a great summer!

References

Gad, Ahmed Fawzy. “Faster R-CNN Explained for Object Detection Tasks.” *Paperspace Blog*, Paperspace Blog, 9 Apr. 2021, https://blog.paperspace.com/faster-r-cnn-explained-object-detection/.

Ren, Shaoqing, et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” *ArXiv.org*, 6 Jan. 2016, https://arxiv.org/abs/1506.01497.

Xu, Joyce. “Deep Learning for Object Detection: A Comprehensive Review.” *Medium*, Towards Data Science, 29 Sept. 2017, https://towardsdatascience.com/deep-learning-for-object-detection-a-comprehensive-review-73930816d8d9.

Xu, Yinghan. “Faster R-CNN (Object Detection) Implemented by Keras for Custom Data from Google's Open Images...” *Medium*, Towards Data Science, 25 Feb. 2019, https://towardsdatascience.com/faster-r-cnn-object-detection-implemented-by-keras-for-custom-data-from-googles-open-images-125f62b9141a.