Traffic Sign Detection System Through Faster-RCNN Architecture

Jahdiel Alvarez, University of Puerto Rico-Mayaguez
Carnegie Mellon University

Introduction

- ✓ Federal institutions maintain an inventory of the traffic signs around a city. Human inspection is a costly process which can be automated.
- ✓ **The goal is to develop a framework to correctly detect and classify multiple classes of traffic signs.**

Faster R-CNN Detection

- ✓ We used the Faster R-CNN architecture [1] as our traffic sign detection model.

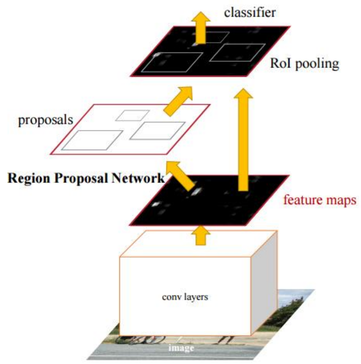

Fig. 1: Faster R-CNN Detection Model

US Traffic Sign Datasets

- ✓ We trained on the LISA-TS Extension Dataset and NavLab's Stop Sign Dataset [2].

Fig. 2: LISA-TS Extension Dataset Sample

Fig. 3: NavLab's Stop Sign Dataset Sample

US Signs Are Different

- ✓ Not all countries follow the Vienna Convention on Road Signs and Signals, such as the US. Therefore signage in this countries may vary drastically from the Vienna Convention standard.

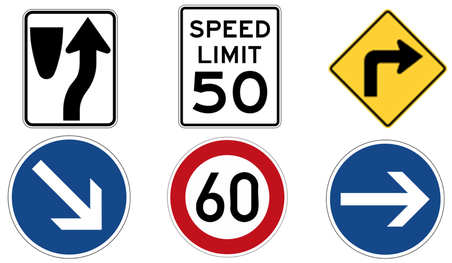

Fig. 4: Same signs in US (top) and Vienna Convention (bottom) contexts.

(left) Keep right, (middle) Speed limit, (right) Turn right

Results

- ✓ Our model was able to obtain a **mean average precision** (**mAP**) of **80.10%** with **IoU of 0.5** on our testing set, which was 20% of the LISA-TS Extension labeled's images.

Fig. 4: Predicted detections of our Faster R-CNN detector

- ✓ The model was trained on **80%** of the **LISA-TS Extension** (LTE). Our testing sets were the other **20%** of the **LTE** set, **LISA-TS Original Set (5839 img)** and **our NavLab Stop Sign Dataset *(2454 img)***.

	AP@.40IOU	**AP**	AP Stop	AP Warning	AP No turn	AP Speed limit
LISA-TS Extension	0.812	**0.801**	**0.866**	0.752	0.816	0.771
LISA-TS Original	0.924	**0.919**	0.874	**0.943**	-	0.940
NavLab Stop Sign	**0.954**	0.887	**0.887**	-	-	-

Fig. 5: Mean average precision results on both LISA-TS Sets and NavLab Stop Signs.

Model Description

- ✓ **Resnet-50**, with PASCAL VOC weights
- ✓ Resize image to **960x960px**, instead of 600x600px
- ✓ 4 Classes: **Stop, Warnings, No Turn, Speed Limit**
- ✓ **Fine-tuned** the network for **53 epochs**.
- ✓ LISA-TS Training Set: **2937 images** Testing Set: **734**
- ✓ Image Size: **1280x960px** Traffic sign sizes: **23 – 222 px**
- ✓ Post-processing: ROI Position Filtering

	AP	AP Stop	AP Warning	AP No turn	AP Speed limit
LISA-TS Ext.	**0.830**	0.857	0.752	**0.919**	0.792

Fig. 6: Mean average precision on LISA-TS Ext. w/ ROI Position Filtering

- ✓ The model was also integrated to a **Traffic Sign Object Tracker** framework as the detector.

Conclusion

- ✓ Our project provides the first public results for a deep learning method based on the 4 LISA-TS Extension super classes.
- ✓ Aggregated Channel Features (ACF) methods currently outperform deep learning methods. This might change by implementing our future works.

Future Work

- ✓ Utilize a smaller network, in order to minimize the receptive field, given traffic signs are very small. Doing more training and obtaining more data are part of our future intentions.

Thanks to my advisor Christoph Mertz.

References

[1] S. Ren, K. He, R. Girshick, and J Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. NIPS, 2015
[2] S. Varadharajan, S. Jose, K. Sharma, L. Wander, and C. Mertz. IEEE Winter Conference on Application of Computer Vision, March 2014.