Detection and Tracking of Pallets by AGVs

Shengchang Zhang(victorzh),Thomas Young(tomyoung)
Jie Xiang(jxiang), Weijian Han (hanwj)

CS229: Machine Learning
Autumn 2019

Introduction

Automated Guided Vehicles (AGVs) have been used in distribution, fulfillment, and manufacturing for many years to improve operational efficiency and address shortages in labor. In this project, we construct a Faster Region-based Convolutional Neural Network (Faster R-CNN) for performing object detection on the pallets. We then construct a CNN-based classifier and train it to predict whether or not images contain pallets, which will be used to assist the Faster R-CNN model in tracking the vehicle. After verifying the model prediction results, we make improvement by tuning hyperparameters by changing layers for both the Faster R-CNN and the CNN-based classifier to make it more efficient.

Data

We identified and chose the 2D Laser Rangefinder dataset [2], There are a total of 565 scans, 340 of which contains a pallet, while the remaining 225 do not.

Data augmentation For each image, we rotate it by 90, 180, and 270 degrees, and also reflecting each image (including the rotated ones) over the x-axis. Using this data augmentation technique, we were able to increase the size of our data set by a factor of 8. and each image is turned into 8 images as in Figure 2.

In order to train the Faster R-CNN detector, the whole dataset has been divided in two parts: 70% as a training set and 30% as a test set.

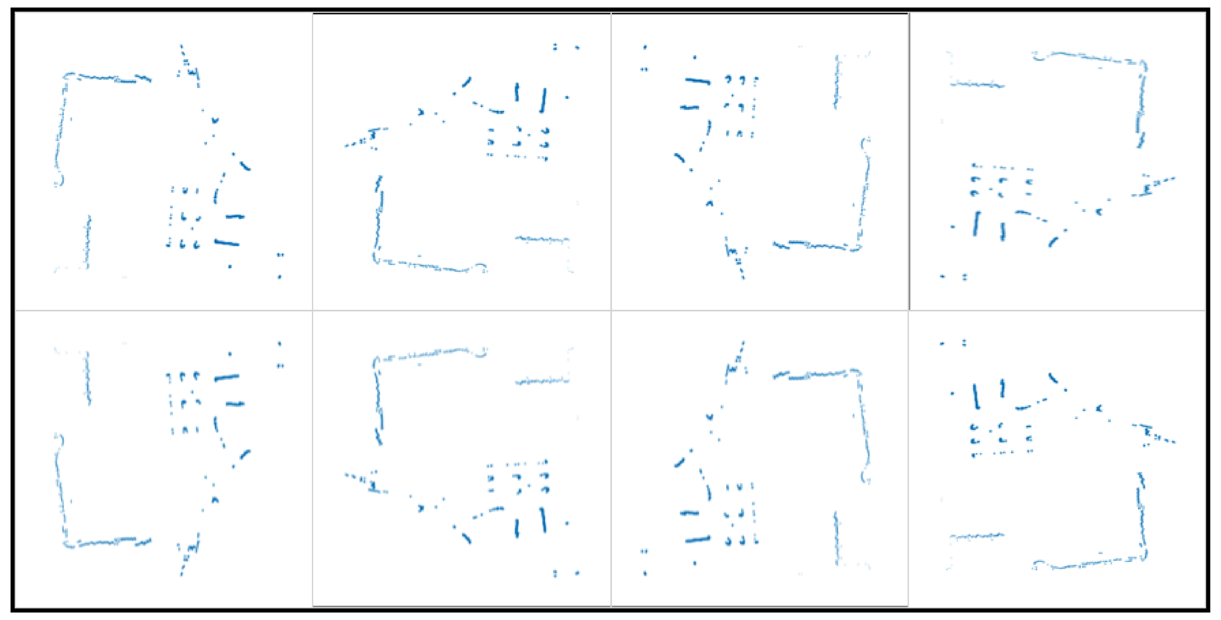

Figure 2: Data Augmentation

Methods

Pallet Detection model is made-up of two steps: a state-of-the-art Faster R-CNN detector which detects the ROIs in each image, and a CNN-based classifier taking as input the previous step detection and discriminating which of them could be a possible pallet candidate.We then constructed CNN, for classifying our input images on whether or not these images contain any pallets. We've improved upon the baseline model and verified the prediction results, re-architected the network and tuned the hyperparameters.

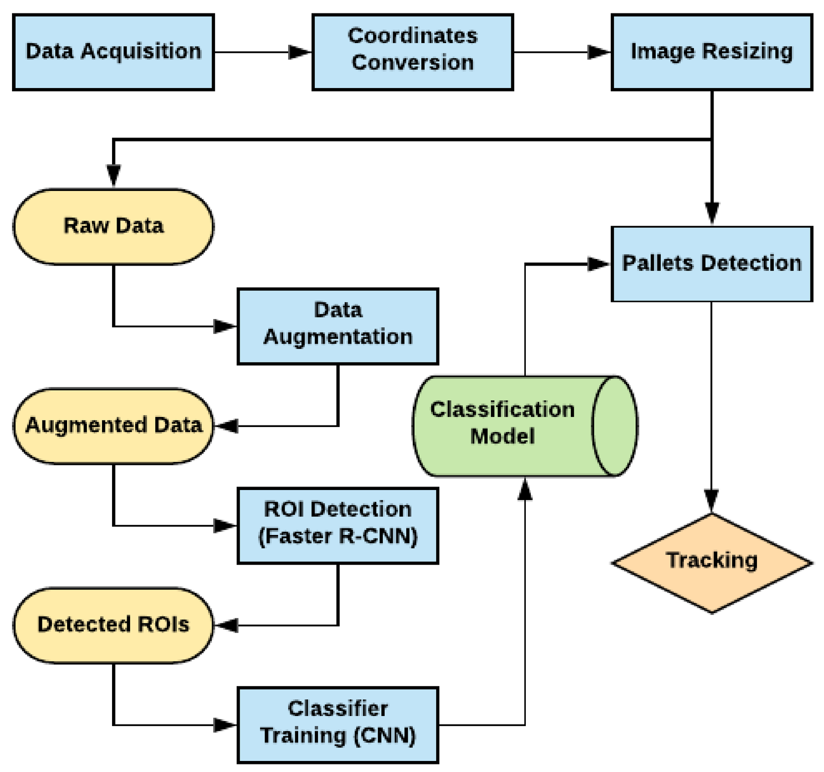

Figure 3: Pallet Detection model

Results

Training and Analysis

We experimented with adding and removing layers for the neural networks and adjusting the filters in each layer, and we also tweaked other hyperparameters such as the learning rate, number of folds for k-fold cross validation, and the number of training epochs. SGD and k-fold cross-validation (with k = 2, 3, 5, 8, 10) are used to train the CNN-based classifier with an initial learning rate $\alpha = 0.001, 0.005, 0.01, 0.03, 0.05, 0.1$, and mini-batch size set to 50, leading to the following data. Considering both performance and computation time, we selected max epochs = 10, folds of cross validation = 3, number of filters = 15, and convolutional layers = 1.

We then compared the performance of our optimized model against the original model by looking at the prediction accuracy, precision, and recall on the test set.Our optimized model reduced the error rate by 25%, kept roughly the same false positive rate, and reduced the false negative rate by 28.5%. Further more, out optimized model architecture and training code reduced the training time by 20%. Figure 1 shows one example of how we selected the best number of filters to use in the first hidden layer of the CNN model. In this example, we selected 15 filters because it offers a good balance of model complexity and performance, and using it gives us a higher accuracy, precision, and recall compared against the baseline model.

Figure 1: Accuracy, Precision, and Recall

Conclusion

- Training our optimized model on the training set, and testing on the test set yielded the following results: accuracy = 0.994, precision = 0.998, and recall = 0.990. Our model not only outperforms the baseline model, which yielded the following: accuracy = 0.992, precision = 0.998, and recall = 0.986, but also took 20% less time to train.

Future Research

We plan to refine the precision of position estimate for pallets with respect to a robot-centred reference frame, as well as integrating orientation estimation. We plan to improve tracking by incorporating data on the algorithm's operating frequency and the maximum speed the vehicle. Furthermore, we intend to explore the possibility of detecting different pallet types or detecting multiple pallets at the same time. Finally, we plan to improve the AGV efficiency by using reinforcement learning so that the AGV can learn the shortest route towards a target.

References & Resource link

[1] Fulvio Mastrogiovannib Stefano Rovettab Renato Zaccariab. Ihab S Mohameda, Alessio Capitanellib. A 2d laser rangefinder scans dataset of standard eur pallets. *ArXiv*, 1805.08564v2, 2019.

[2] Fulvio Mastrogiovanni Stefano Rovetta Renato Zaccaria. Ihab S Mohamed, Alessio Capitanelli. Detection, localisation and tracking of pallets using machine learning techniques and 2d range data. volume 1803.11254v3, 2019.

Code link: `https://github.com/jxiang1984/CS229-ML-Project.git`
Presentation Video link: `https://youtu.be/WO1yopHVNQA`