

## **Individual Contribution [Lizhen Lin 82226790]**

When I first started this project my first task was to read a lot of different academic papers. My goal was to try and find a model that would not take a long time to train, without sacrificing a solid level of accuracy. The two papers I read that were most beneficial to me were Deep Residual Learning for Image Recognition from Kaiming He Xiangyu Zhang, Shaoqing Ren, Jian Sun and Robust Hand Detection and Classification in Vehicles and in the Wild from T. Hoang Ngan Le Kha Gia Quach Chenchen Zhu, Chi Nhan Duong Khoa Luu and Marios Savvides. The main thing I gained from these papers were better understanding the nature of the problem we were working on. Further, these papers helped me think about our problem in a step by step manner.

My next task was to find an appropriate dataset, and preprocess this data. After doing some research online, the most famous one were EgoHands and the dataset from the following link: <http://www.robots.ox.ac.uk/~vgg/data/hands/> Reading through other people's experiences, the models trained with EgoHands usually had better results. Therefore, we decided to use the EgoHands dataset.

After we obtained the data, I began working on setting up the Google Cloud account that we needed to use. I read the corresponding documentation and was responsible for explaining how to add GPU to the instance and also how to transfer data from local sources to the cloud to my teammates.

I next looked for a method to load the matlab files and combine them with the images. I wrote a script that renamed all the files with its directory name prefix to the filename since initially all the file names are just their frame number. This meant among all the folders/videos there were some conflict between all the files. After doing some research I learned from Victordibia's github on how to convert the images and matlab files into csv file format. After this I started to transfer csv files into a tfrecord, by adding all the important labels (filename, image(encoded) xmin[], xmax[], ymin[], ymax[], and classes.) to the array. I did all this so that when we actually used this data in the jupyter notebook it was really easy to decode the related information. Finally, in terms of the training process, I was tasked with helping adjust the parameters with JiaHao Yao based on each one time performance.

Lastly, I helped write our final report (focusing mainly on Section 6) and was also tasked with the final editing of our report. In the end any typos that remain are my fault!

At the beginning of the project, I read introductions of all project topic we can choose and related articles. After discussing with team members and making a decision, I searched and read others' projects outlines and the datasets they used. Egohands dataset, dataset we got from course and VIVA can be used to train our models. In order to check the quality and feasibility of these datasets, I used these datasets and ran a online model to test the accuracy of result. Since pictures we got from course are too vague and the data from VIVA are too complex, we finally choose Egohands dataset. Next, my teams members and I found a way and wrote script to convert images in dataset to csv format which contains box coordinates. In the training stage, Finally, I test the accuracy of our model with my team members. When preparing the presentation powerpoint, I collected the results, datasets, and problems we already had and met. In the final report, I wrote and parts named "Introduction and problem statement", "Dataset", "Software" and "Discussion and Conclusion".