

# Scalable Computing Project 2

Jiaming Deng Christopher Lohse

## 1. Introduction

For this assignment, we need to fetch files from a server, train a model and classify 2000 images at pi. This report mainly describes the problems we solve and summary details of this assignment and the conclusion describes the principle of scalable computing that teacher wants to tell us through this assignment.

## 2. Task1

For the task1, we need to find out a way to fetch a file list and images from the server. Firstly we try to type “cs7ns1.scss.tcd.ie/?shortname=dengji” at chrome but it doesn’t work and return a Forbidden sign. The reason is it’s not in the environment of pi. Next, we try to use “curl -L cs7ns1.scss.tcd.ie/?shortname=dengji -o file.txt” at pi, but it doesn’t work too. Finally our teacher advises us to connect to VPN, set proxy and try again, at this time, we get the file list and then get the images successfully.

## 3. Task2

For the task2, we need to train a model and recognize the images’ symbols and submit the output file in the submitty and get a score.

Firstly, we train a model using the project1’s code and get a very low score. After comparing at the output file and images’ symbols, we recognize that the number of images are not constant, but from 1 to 6. Next, we change some codes at generate file and get some training images with different length of symbols and then train the model again. At this time, the score increases a lot, but at this time teacher posts top3 scores at piazza, we know we have to move on. Then we find that we need to add some elements into the symbols.txt file and expand the dataset, and we get more than 1000 score at this time. Finally, we have to classify the files in the pi and get a problem that pi is 32-bits machine but tensorflow only support 64-bits machine. We change the classify file using tensorflow-lite and run correctly at pi and get 1375/2000 score at submitty.

## 4. Conclusion

The teacher deliberately asks us to run the classification code on a 32-bits pi. We think the intention is to show us that we might have to train a model with high detection accuracy in a very complex environment, but the trained model could be used to apply and solve problems in a very simple environment.