**HW1: Face Detection Report**

111550107 顏芝佑

# Part I. Implementation (5%):

一張含有 文字, 螢幕擷取畫面, 軟體 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 軟體, 多媒體軟體 的圖片

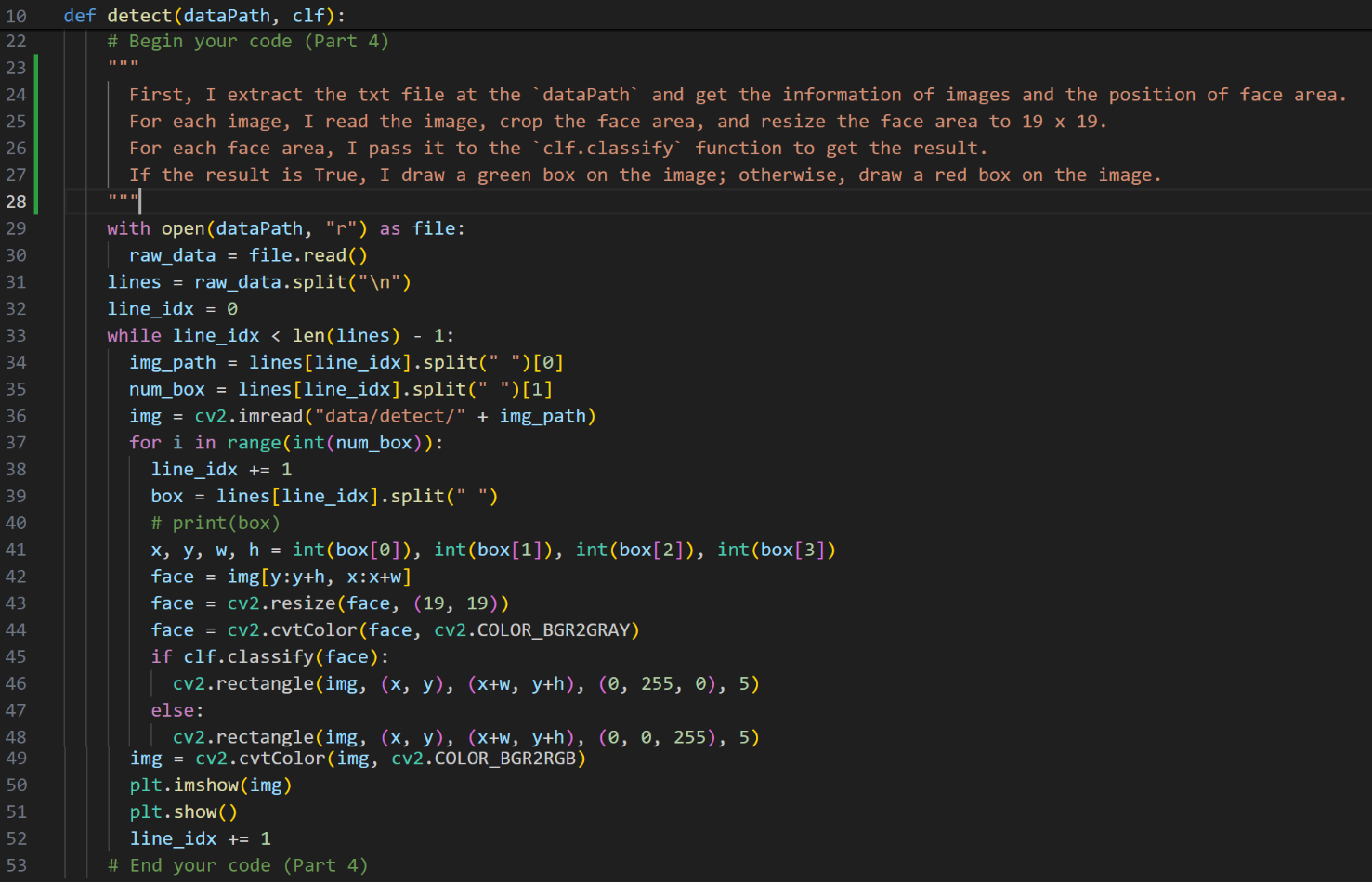
自動產生的描述

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 軟體, 多媒體軟體 的圖片

自動產生的描述



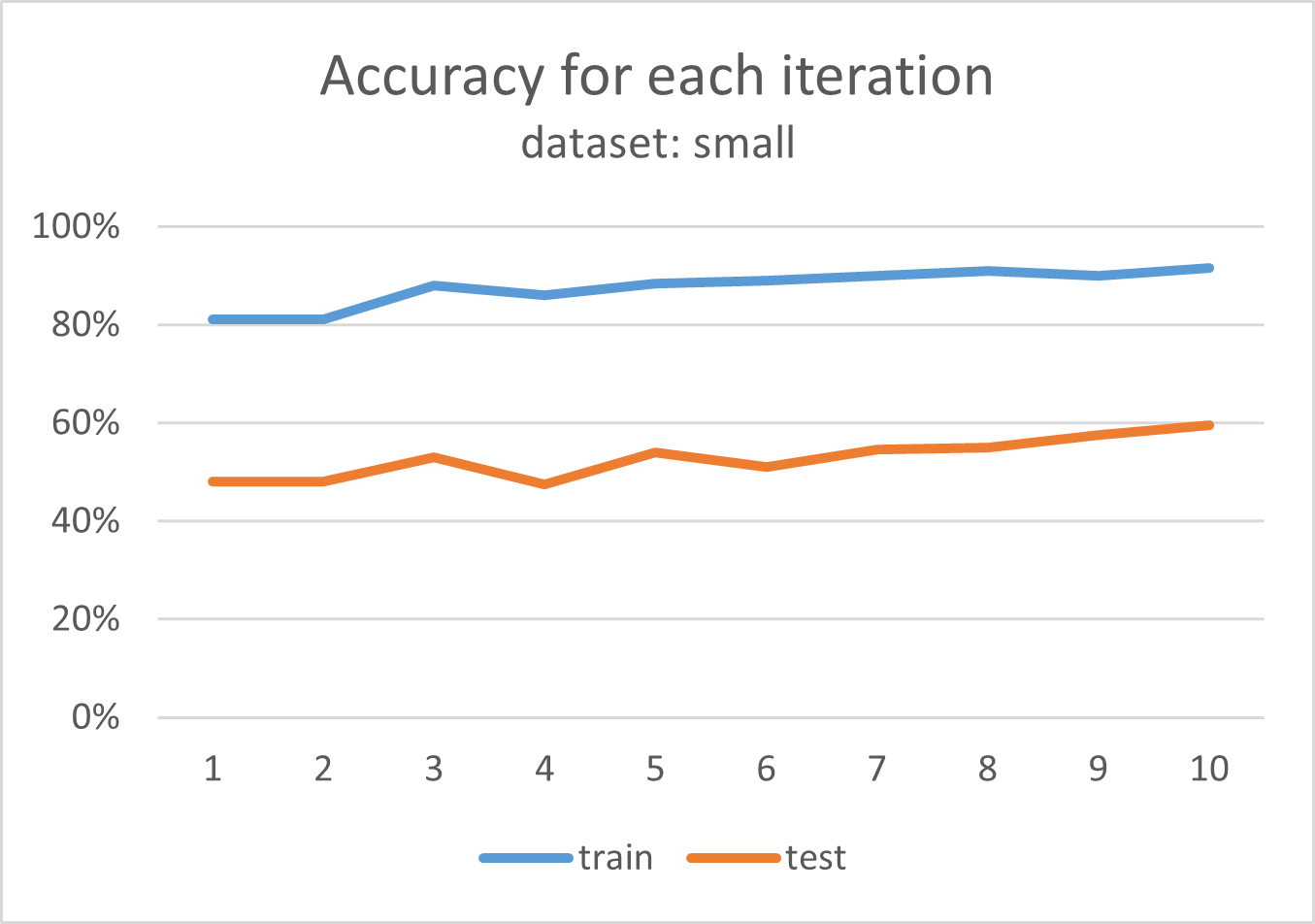
# Part II. Results & Analysis (10%):

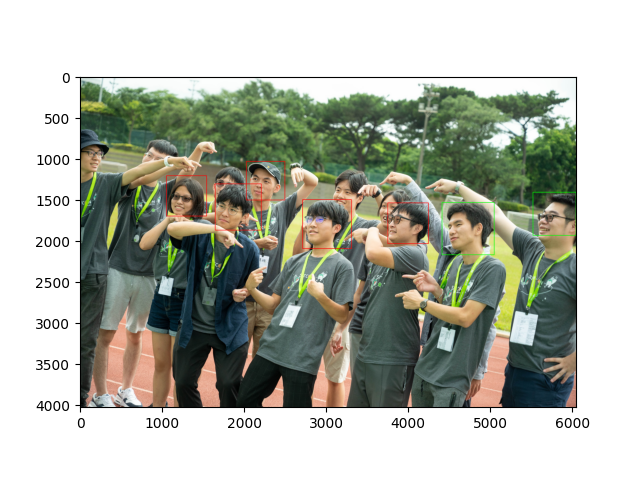
1. Classifier trained by small dataset

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 數字, 字型 的圖片

自動產生的描述 

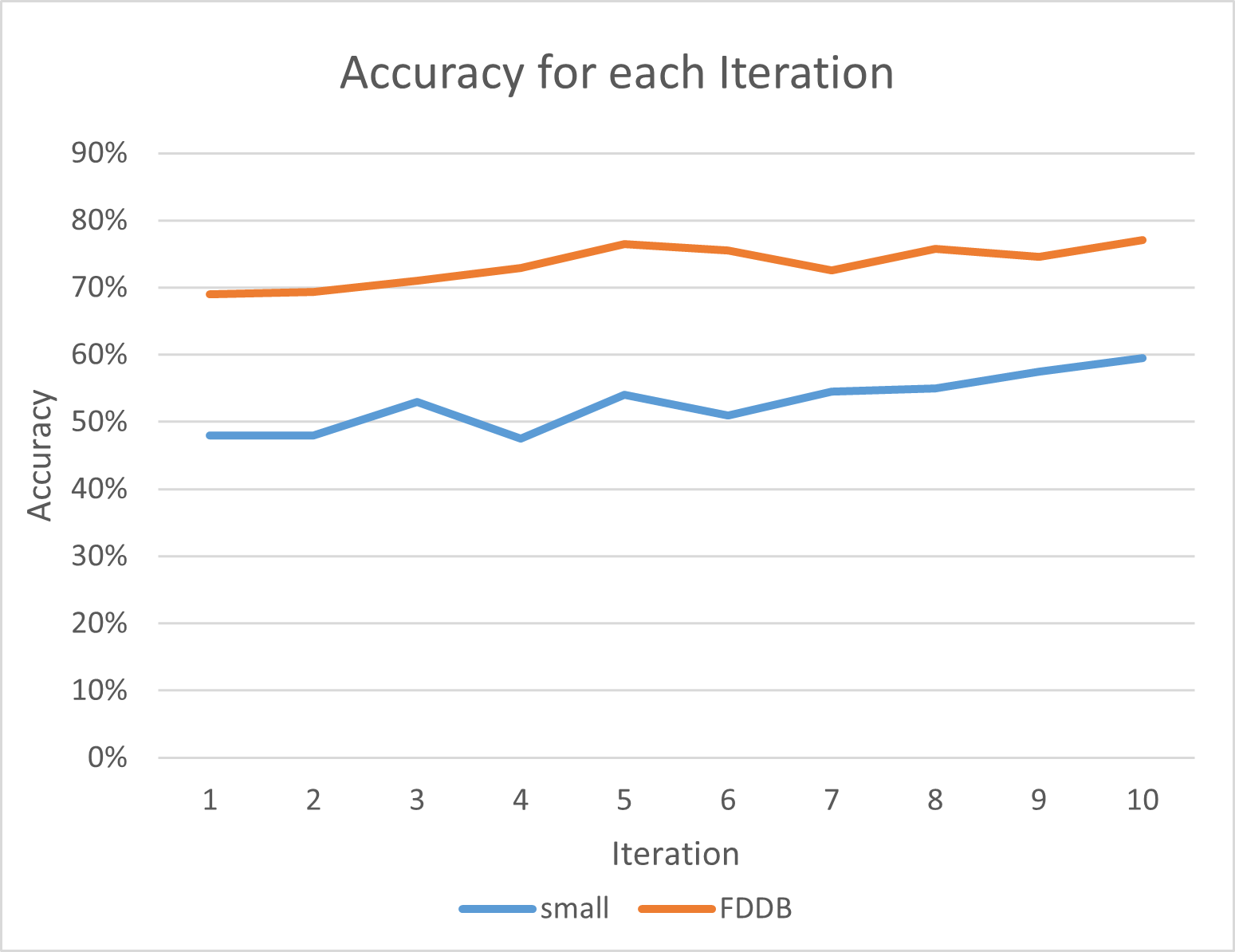
 

The accuracy of the training dataset is around 90%, while the accuracy of the testing dataset is only around 50%. I guess this situation occurs because the model is overfitting due to the insufficient training data.

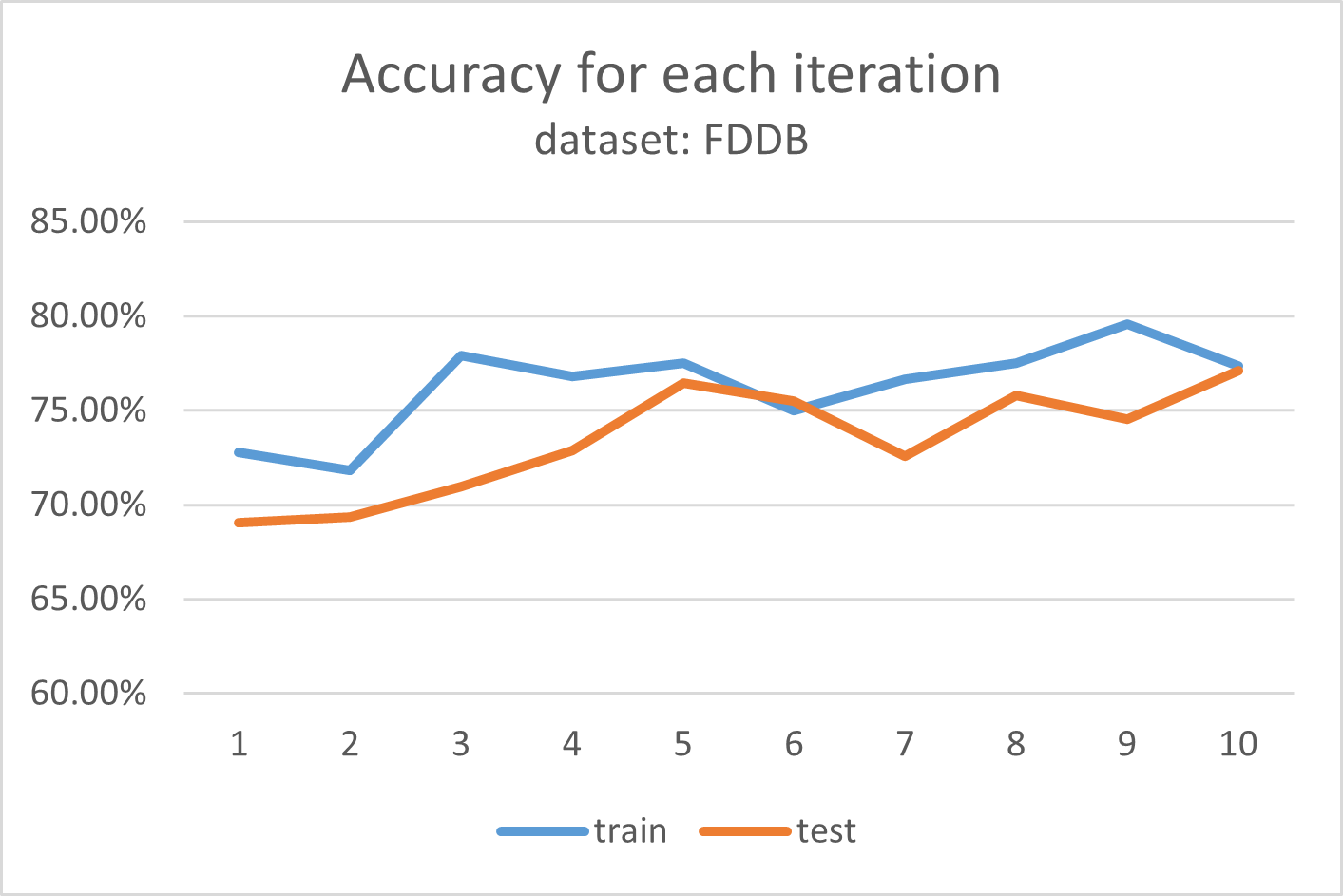
1. Classifier trained by FDDB dataset

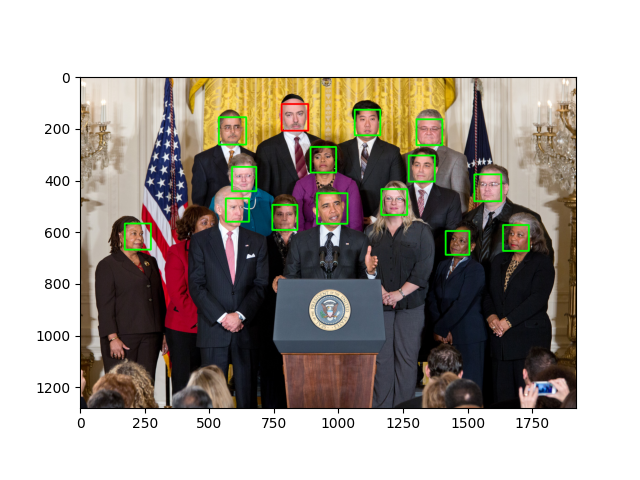
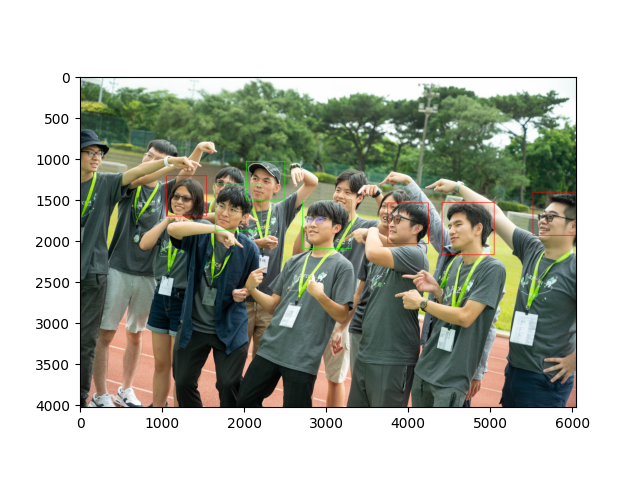
一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述



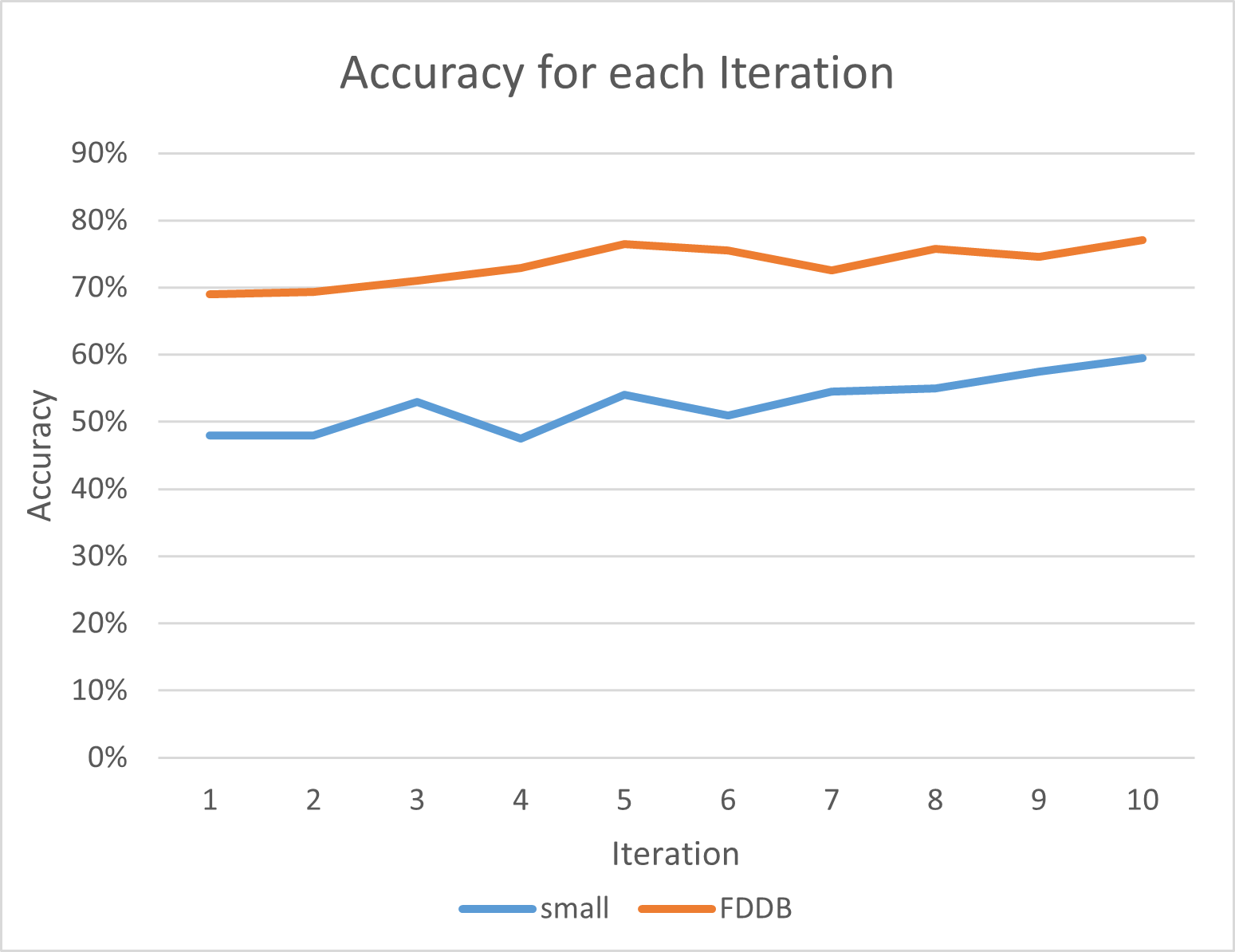
一張含有 文字, 字型, 數字, 螢幕擷取畫面 的圖片

自動產生的描述 

For FDDB dataset, the accuracy of the training dataset and the testing dataset are quite similar, which are around 80%. Also, the result of applying the model on the other images is good, too.

1. Comparison of two classifiers



From this chart, it is obvious that the classifier trained by FDDB dataset performs better than the classifier trained by small dataset. This result is reasonable because the more data we used to train the model, the more accurately the model can perform the detection.

Part III. Answer the questions (15%):

1. Please describe a problem you encountered and how you solved it.

Problem: When I call `plt.draw()` function, I got an error. The error message is as the following: “FigureCanvasAgg is non-interactive, and thus cannot be shown plt.show().”

Solution: I search for the Internet and found this error occurred because the non-GUI environment cannot show a figure. I run my program in WSL, so there is no graphic user interface. To solve this problem, I installed `python3-tk` in the environment.

1. How do you generate “nonface” data by cropping images?

I implemented a `overlap` function to check if two area overlap with each other. Assume the height and width of the image are and respectively. I generate two random variables , and check if the square with being the top-left corner and 19 being the side length overlap with any of the face area. If yes, then randomly choose and check again. Otherwise, we successfully generate a non-face region.

1. What are the limitations of the **Viola-Jones’ algorithm**?
2. It can only deal with binary classification
3. It is sensitive to lighting condition
4. It may not perform well when detecting objects with different viewpoint, such as side faces.
5. It takes lots of time to train the model.
6. It is vulnerable to noise.
7. Based on **Viola-Jones’ algorithm**, how to improve the accuracy except changing the training dataset and parameter T?

We can adjust the threshold parameters applied in the weak classifier to improve the accuracy.

1. Other than **Viola-Jones’ algorithm**, please propose another possible **face detection** method (no matter how good or bad, please come up with an idea). Please discuss the pros and cons of the idea you proposed, compared to the Adaboost algorithm.

The convolutional neural network(CNN) model can also be used to solve face detection problems. CNN is commonly used to deal with image classification and object recognition. Compared to Viola-Jones’ algorithm, there are some pros and cons of CNN model.

Pros

1. CNN is not sensitive to lighting conditions.
2. CNN can handle variations in object size, viewpoint, and orientation.

Cons

1. Training a CNN model is computationally complex, and therefore requires huge amounts of time or computation resources.
2. As a deep learning model, it requires large amounts of annotated data.