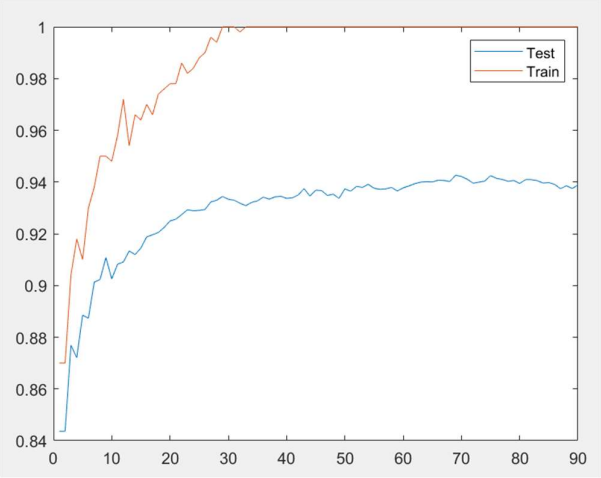
TBMI26 – Computer Assignment Reports Boosting

Deadline – March 14 2021

Authors: Alexander Bois & Daniel Bissessar

In order to pass the assignment you will need to answer the following questions and upload the document to LISAM. Please upload the document in PDF format. **You will also need to upload all code in .m-file format**. We will correct the reports continuously so feel free to send them as soon as possible. If you meet the deadline you will have the lab part of the course reported in LADOK together with the exam. If not, you'll get the lab part reported during the re-exam period.

1. **Plot how the classification accuracy on <u>training data and test data</u> depend on the number of weak classifiers (in the same plot). Be sure to include the number of training data (non-faces + faces), test-data (non-faces + faces), and the number of Haar-Features.**



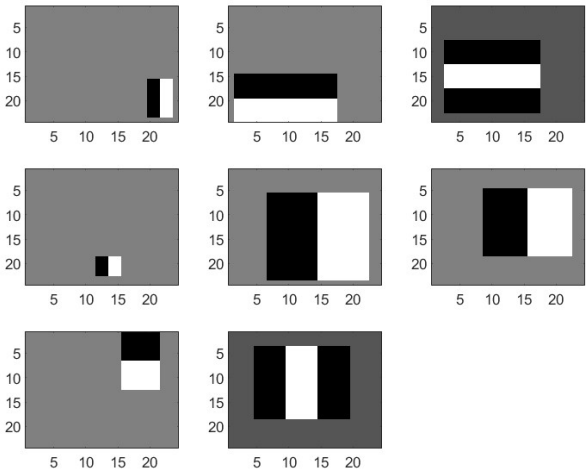
Training data: 500
Test data: ~12000
Haar-Features: 250

2. **How many weak classifiers did you use when training? How many of them did you use for the final strong classifier? Motivate your choices.**
   We used 90 for training, but only 69 for the final strong classifier. We choose a large amount of weak classifiers to train on to be able to choose the best amount from the Test accuracy seen in the graph in the previous question. We then chose 69 weak classifiers since that was the amount that produced best test accuracy.

3. **What is the accuracy on the training data and test data after applying the optimized strong classifier? Discuss your choice of hyperparameters and how they influence the accuracies.**
   Test acc: 94.3%

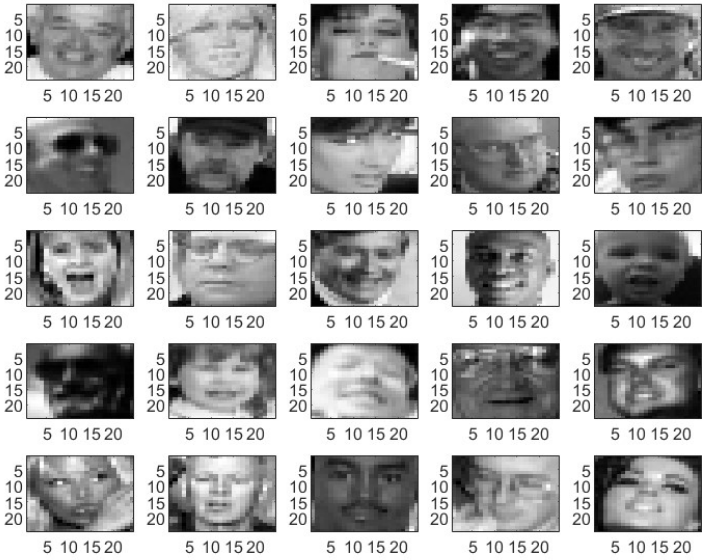
Train acc: 100%

4. **Plot the Haar-features selected by your classifier (one for each weak classifier). If you have many weak classifiers, select some representative subset. Can you think of why they would be useful for classifying faces?**

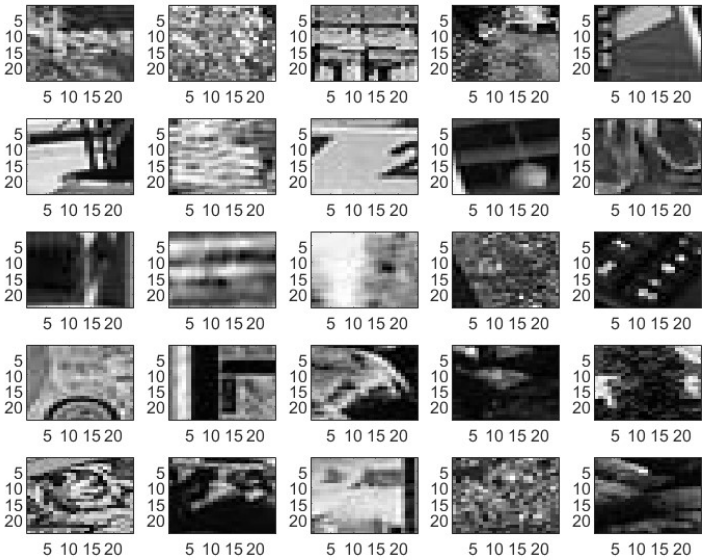


They could capture parts of a face and together figure out if it is a face or not, but it is hard to tell from this subset.

5. **Plot some of the misclassified faces and non-faces that seem hard to classify correctly. Why do you think they are difficult to classify?**



For faces the picture is either very dark or very bright. There is a lot of glasses and mustaches. We think these has had influence on the result.



CHANGED: There is a either a lot of the same/similar color or quite complicated patterns, which we believe to be the reason our model is having a hard time distinguish these pictures from faces.

6. **Are your results reasonable? Can you think of any way to improve the results?**
   Yes, we think the result is quite reasonable. One way of increasing the results would be to use more training data to train the classifier.

7. **Can we expect perfect results? Motivate your answer.**
   We think we can achieve really good result, almost perfect, with enough weak classifiers and training data. A thought is to maybe design some Haar features instead of just having random features, this could maybe improve the result. This would also introduce bias into the model that might not be desirable.