Lab A3 Boosting

Tore Andersson, Keshav Padiyar Manuru

March 9, 2021

Q1. Plot how the classification accuracy on training data and test data 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.

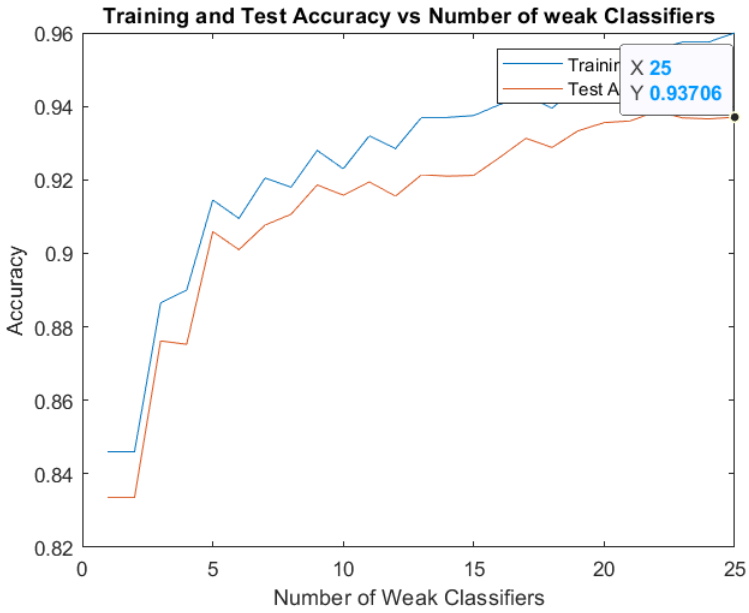

Figure 1: Accuracy vs Number of weak classifiers using 200 Haar-features,training images = 2000, weak classifiers = 25, test accuracy = 93.7%

There are total of 4196 images with faces and 7872 images without faces. In our experiment we considered 2000 images for training, the algorithm combines first 1000 images of each face and non-faces to form a training dataset of 2000 images. Also, the remaining images i.e images from 1001 to total number of images of faces and non-faces are combined to for test dataset i.e 10068. Figure 1 shows the training and test accuracy of classification.

Q2: Q2. 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 25 weak classifiers while training, and considered all 25 of them as final classifier. If we change the number of weak classifiers from the training to testing phase, we would be considering a new model compared to the one used in the training set which is not become a fair comparison.

Q3. What is the accuracy on the training data and test data after applying the optimized strong classifier?
Discuss your choice of hyper-parameters and how they influence the accuracy.

The best combination of the different hyper-parameters were 200 haar-features, 2000 training images and 25 weak-classifiers which achieves a test-accuracy of 93.7% which is shown in figure 1.

Number Of Haar Features	Number Of Training Samples	Number of Weak Classifiers	Accuracy	Overfits	Remarks
64	2000	25	91 / 93.3	No	Not Stable / Stable
128	2000	25	92.2	No	Stable
200	2000	25	93.7	No	Stable
256	2000	25	94	No	Not Stable
256	750	25	93.5	Yes	Not Stable
256	2000	50	95	Yes	Stable

Number Of Training Samples:
Training data with 750 images and testing on 11318 images would be highly disproportionate ∼(1:16). Increasing the number of images in the training set from 750 to 2000 ∼(1:6) improved the test-accuracy and reduced the over fitting of the model on the data.

Number Of Weak Classifiers:
The model gained stability by using around 25 weak classifiers. Doubling the number of weak classifiers improved the test-accuracy however there were a higher training accuracy at over 98% as well which tends to overfit the model.

Number Of Haar-Features:
Each haar-feature is a responsible for detection of a particular feature as the number of haar-feature increases the number of feature detected increases. We identified that 200 haar-features are sufficient to obtain the required accuracy and less computationally heavy.

Q4. 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?

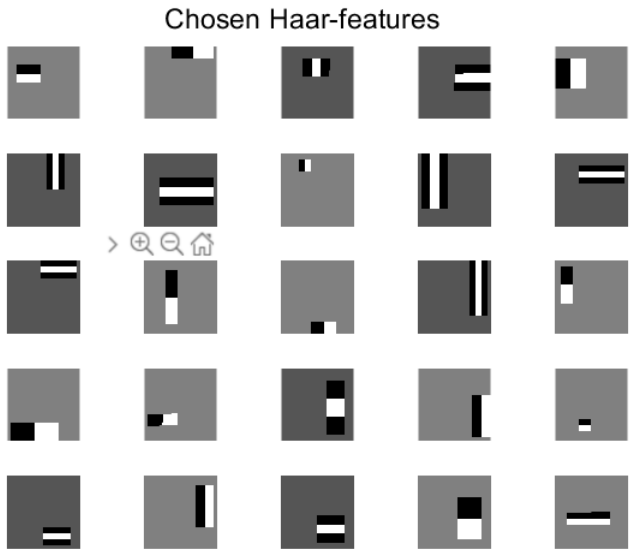

Figure 2: A subset of haar-features used

Figure 2 depicts the haar-features subsets used by the classifier.

Haar-features are the re-scaled rectangular/square shapes with combination of darker and lighter areas. The darker areas consist pixel value = 1 and 0 at lighter region. This structure of haar feature is used to detect features like lines and edges where there is contrast in intensity, this nature of Haar-features can be used to detect faces. For example, if we consider the facial image in black and white channel, the pixel intensity at the eyebrows are darker than at the forehead. A horizontal haar feature would detect that. Another example is detection of lips, the pixel intensity will be darker on lips region and lighter at both above and below lips. So, light-dark-light combination haar-filter can be used in that region.

Q5. Plot some of the misclassified faces and non-faces that seem hard to classify correctly. Why do you think they are difficult to classify?

From figure 3 the missclassified pictures have very dark features and bright features. One reason could be that the haar-features used can't explain the facial features. Where some haar-features could not be differentiate distinct enough facial features from the shading, as there are not clear enough edges for specific features. For example the eye-brows are not dark enough and lacks clear edges compared the rest of the face, this could be of the complexion of the person in

Figure 3: Missclassified faces

the picture or shading.

From figure 4 the reasoning for miss classification could be that the pictures have similar pixel-densities for positioning of the features that could correspond to a facial features.

Q6. Are your results reasonable? Can you think of any way to improve the results?

Our results looks reasonable as we trained over only 2000 images (training on more images was computationally expensive for our systems), if we had trained on more images with higher black and white resolution the haar-features would have better values to discriminate between the intensities in pixels.

In the current implementation to reduce the weight of the outliers in the dataset, if an observation receives a weight higher than 0.9 it's redefined with a weight of 0.8 instead. The way to reduce the weights of outliers could be improved with a more sophisticated methods as "logit boost" or "Gentle boost".
Q7. Can we expect perfect results? Motivate your answer

It's not reasonable to expect a 100% accuracy for classification as the variation in image pixel density is not constant. Also, haar-features can't actually

Figure 4: Missclassified non-faces

identify the facial features itself, it represents the difference in contrast between the edges of facial structures. This means that pictures which share the similar layout of pixel-densities can be classified as a face even if it is not a face.