So what Patrick has done with low level features is nice but as humans, we also recognise high level features such as gender.

Two approaches are taken. Once involves using a CNN that trains on the 1200 faces with a filter size of 3 over 100 epochs using MATLAB’s functions. The other is to use eigen faces as the features to input into a state vector machine.

So we should all know what a CNN is by now but what are eigen faces? They are the eigen vectors of the covariance matrix of the face dataset. In English, it captures the variation in a collection of face images in a holistic manner. This is done by performing principle component analysis on a large set of image faces.

As a result, this allows you to represent any human face as a combination of these eigen faces. So my face for instance may be made up of the average face + 10% of face 1, + 30% of face 2, -10% of face 3 etc. Here are the weights of Face 1 over a set of 20 and 1200 eigen faces.

To get the faces from movie posters, MATLAB’s implementation of the viola jones algorithm was used.

So let’s put these weights and Euclidian distances into the SVM shall we?

So unfortunately, the eigen faces aren’t very good as it only has an accuracy of a coin toss. The CNN has a respectable accuracy of 75% but is still rather low. These bad accuracies translate into very funny results.

The issue with the Viola Jones algorithm is that it uses a frontal face classifier and assumes the faces are facing forward. But in movie posters this is not always the case and some faces are not detected. This is because the faces in movie posters are heavily stylised with various lighting and artistic effects.

Furthermore, to use eigen faces correctly several conditions need to be met. The lighting, scale and translation of the faces need to be controlled. But the image training set is random images taken of Wikipedia of various shares, sizes and with objects in the background. So due to the lack of control of the faces in the image dataset, the prediction results are poor.

The CNN results are equally as bad since the faces in movie posters are rather different to the one in the training set. Training it with faces from posters would help improve accuracy.

So, using the number of males and females detected as a feature in Patrick’s module actualy lowers the accuracy down from 79.3% to 75.3%. This is due to the gender data being inaccurate and basically providing redundant information to the classifier, so you shouldn’t expect any improvements.