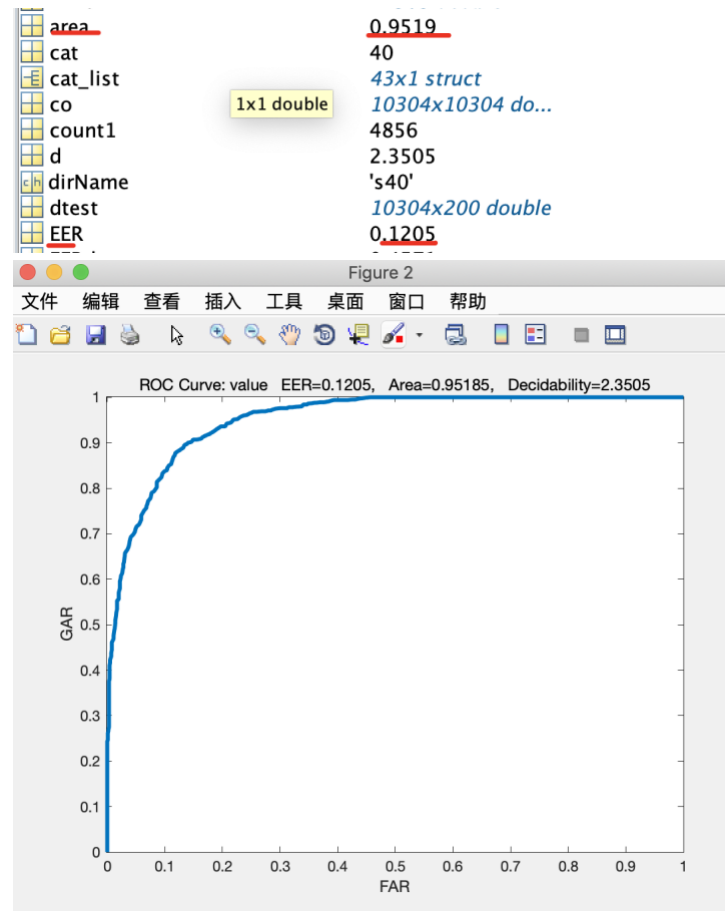


Justification on the difference in the performance of Mode 1 and Mode 2

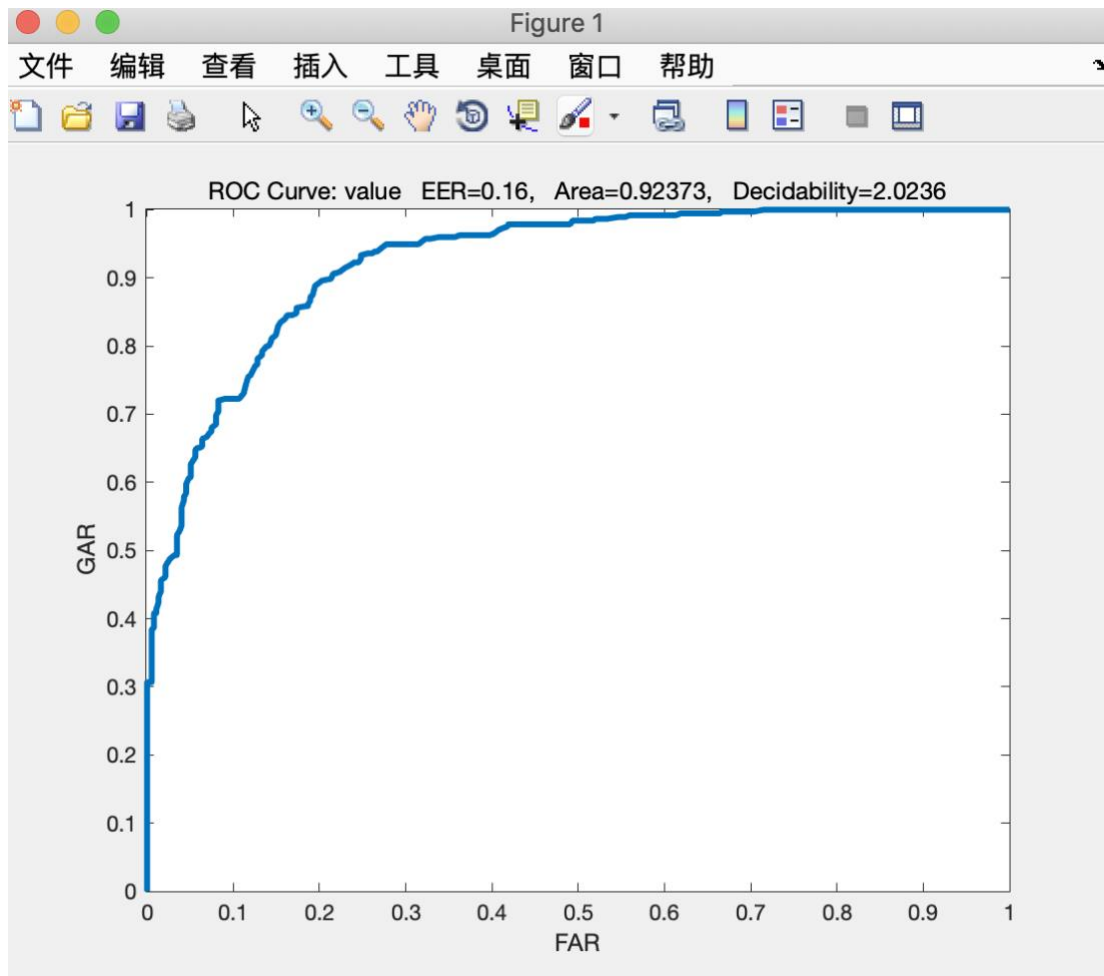
We obtain two models ROC curves by ground truth and scores(genuine and imposter)

For model1:



For model2:

area	0.9237
cat	40
cat_list	43x1 struct
co	10304x10304 do...
count1	4965
d	2.0236
dirName	's40'
dtest	10304x150 double
EER	0.1600



According to the area in above images , we can find out that the performance of model 1 is better than the model 2 . The Area we call it Area Under Curve. It describes the performance for our classifier.

- $AUC = 1$, perfect
- $0.5 < AUC < 1$, good
- $0 < AUC < 0.5$, bad

But for the EER(equals error rate), model2 is higher than model1.

Why model1 is better than model2?

I think for model1 we extract eigen features (eigen vectors and eigen values) from all datasets. We aggregated all subjects' features to find PCA subspace by finding largest variance. But for model2 ,we only find PCA subspace by a subsets(top 20), the PCA subspace we obtained can't represent by all subjects' feature. So when we project our test dataset into PCA subspace , model1 contains all the subjects from training and test , but for model2 only a few features from training subjects.