

BITS-F464: Machine Learning

Instructions to Class Project 01

February 17, 2021

1 The Database

The database could be downloaded from the course webpage¹. Please note you are not suppose to distribute it and delete it after the submission of your report. The same would be removed from the website on Feb 20, 2021 at 5PM. Make sure that you download it in this duration.

The database has file with name such as 10_P1.S1.4 .jpg it signifies as below.

- Subject identity is 10. So you take it as the data provided by the subject number 10
- Our app was taking two images in one go. First one from distance and the second one from near. P1 signifies the image taken from distance and P2, the second image capture when phone is brought near to the forehead. (So the image 10_P1.S1.5 .jpg is taken from far)
- S1 signifies session one. You know our app was taking images in two sitting. First it takes five pair of images (10 images) and then asks for 12 hour interval. Images taken after 12 hour session are having S2 in their name. So the image 10_P1.S1.5 .jpg was taken in first session.
- In one session app have taken 5 pair of images. The value 4 signifies that this was 4th pair of image.

Based on the above explanation all 20 images provided by subject id say 31 are named as follows

31_P1.S1.1 .jpg, 31_P2.S1.1 .jpg, 31_P1.S1.2 .jpg, 31_P2.S1.2 .jpg, 31_P1.S1.3 .jpg, 31_P2.S1.3 .jpg, 31_P1.S1.4 .jpg, 31_P2.S1.4 .jpg, 31_P1.S1.5 .jpg, 31_P2.S1.5 .jpg, 31_P1.S2.1 .jpg, 31_P2.S2.1 .jpg, 31_P1.S2.2 .jpg, 31_P2.S2.2 .jpg, 31_P1.S2.3 .jpg, 31_P2.S2.3 .jpg, 31_P1.S2.4 .jpg, 31_P2.S2.4 .jpg, 31_P1.S2.5 .jpg, 31_P2.S2.5 .jpg

There course be all the challenges in the database such as missing images and variations in the data acquisition.

2 What need to be done

As you know a code is provided on course webpage². We expect following experiments to be done. However, the list is neither comprehensive.

¹Link: <http://www.ktiwari.in/ml/wp-courseProject.html> by Feb 20, 2021

²Link: <http://www.ktiwari.in/ml/wp-courseProject.html>

1. Do $n \times n$ matching using popular feature extractor
2. Prepare genuin/imposter histogram
3. Explain how scores are obtained
4. Plot ROC
5. Report EER and the corresponding threshold.
6. Report CRR
7. Do some ablation study and modify parameters

Explain what are your observations and conclusions.

2.1 Definitions

Standard performance parameter are,

- **Correct Recognition Rate (CRR):** It is defined as the number of actual matches that are obtained at rank one recognition.

$$CRR = \frac{\text{Number of matches correctly recognized}}{\text{Total number of matches}}$$

- **False Acceptance Rate (FAR) and False Rejection Rate (FRR):** FAR refers to the likelihood of the biometric system to incorrectly accept an unauthorized user as an authorized one *i.e.* the rate of false acceptance over the number of imposter attempts. It is defines as:

$$FAR = \frac{\text{Number of incorrect matches recognized}}{\text{Total number of matches}}$$

FRR on the other hand, is the likelihood of biometric system to incorrectly reject an authorized user by considering him to be an unauthorized user, which is defines as the rate of false rejection over the number of genuine attempts by the user.

$$FRR = \frac{\text{Number of correct matches not recognized}}{\text{Total number of matches}}$$

- **Equal Error Rate (EER):** It is the point at which the False Acceptance Rate (FAR) and False Rejection Rate (FRR) are equal. It gives a threshold to evaluate the recognition performance of a system. Also, a system with lower EER is considered better.
- **Accuracy:** is maximum value of $(100 - (FRR + FAR)/2)$ across all thresholds.
- **ROC Curve:** An ROC curve is obtained by plotting FRR vs FAR, by varying the decision threshold. The area under the curve gives the error rate of the system. A system having less area under the curve is better at classification.

3 Submission

A single .zip file on Nalanda containing .pdf and code. Report need not to be very large, talk about the experimental findings and results only. Result-score file need not to be submitted. Keep it with you. Will be asked if needed. Deadline is Feb 23, 2021

4 Evaluation

Would be done based on the results obtained, number of experiments done, insights gained.

5 Help

Use already shared whatsapp group to ask for any doubt.