Dhirubhai Ambani Institute of
Information and Communication
Technology

Biometric and Security – IT 499

LAB – 2 : Errors and Metrics in Identification Framework

Name:	Manish Manojkumar Prajapati
Student ID:	202411012
Branch:	MTech ICT ML (2024 - 26)
Date:	13/02/2025

EXERCISES

Assume enrolment dataset consists of N =45 subjects with 10 samples each. Assume the remaining subjects M= 13 subjects with 5 samples each who are not enrolled in the database.

1.) Feature extraction.

a. Linear Binary Pattern

i. Use skimage.feature import local_binary_pattern

ii. Paremeters: radius = 3, n_points = 8 * radius, and lbp_bins = 256

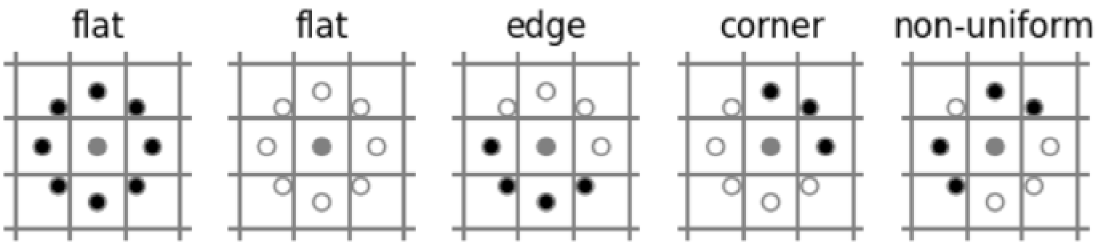
iii. Generate 256 dimensional feature embeddings

b. Use a pre-trained FaceNet model to extract 512 dimensional feature embeddings.

Answer:

⇨ LBP

- Local Binary Pattern is generally used for texture classification.
- Local Binary Pattern looks at points surrounding a central point and tests whether the surrounding points are greater than or less than the central point (i.e., gives a binary result)



- The figure above shows example results with black (or white) representing pixels that are less (or more) intense than the central pixel.
- When surrounding pixels are all black or all white, then that image region is flat (i.e., featureless)
- Groups of continuous black or white pixels are considered 'uniform' patterns that can be interpreted as corners or edges.
- If pixels switch back-and-forth between black and white pixels, the pattern is considered 'non-uniform'.
- Used skimage's local_binary_pattern for finding the LBP features.
- **CODE:**

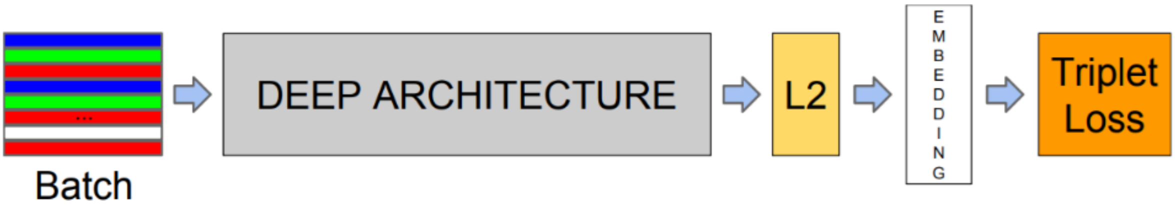
```
from skimage.feature import local_binary_pattern

radius =3
n_points = 8*radius

lbp = local_binary_pattern(image_input, n_points, radius, method='default')
```

⇨ FaceNet

- FaceNet is the name of the facial recognition system that was proposed by Google Researchers in 2015.
- They proposed an approach in which it generates a high-quality face mapping from the images using deep learning architectures such as ZF-Net and Inception Network.
- Then it used a method called **triplet loss** as a loss function to train this architecture.



- **CODE:**

```
from keras_facenet import FaceNet
from PIL import Image

embedder = FaceNet()

image = Image.open(image_path)
image = image.resize((160, 160))
image = np.array(image)

image_embedding = embedder.embeddings([image])
image_embedding = image_embedding.flatten()
```

- This will give us 512 feature embeddings of the image inserted.

2.) Closed set Identification.

a. Evaluation protocol: Consider 1 sample per subject in the enrolment database as the probe and remaining as the gallery.

b. Use Euclidean distance to compute the comparison scores.

c. Based on the 1:N comparison scores.

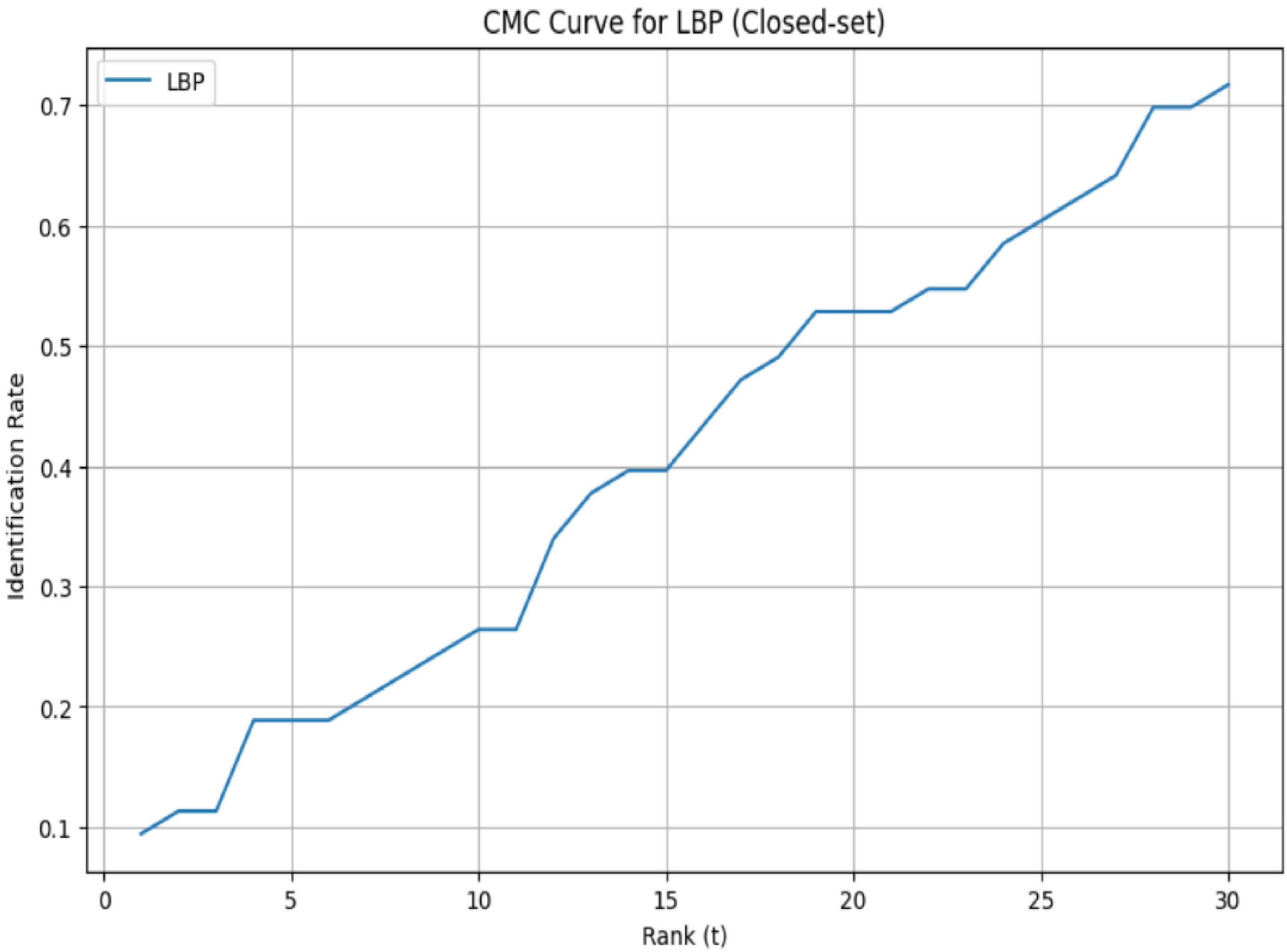
i. Plot the CMC curves showing t-rank identification rate in % (TPIR) for ranks t=1 to 30.

ii. Is a threshold for comparison scores considered in this scenario? Why or why not?

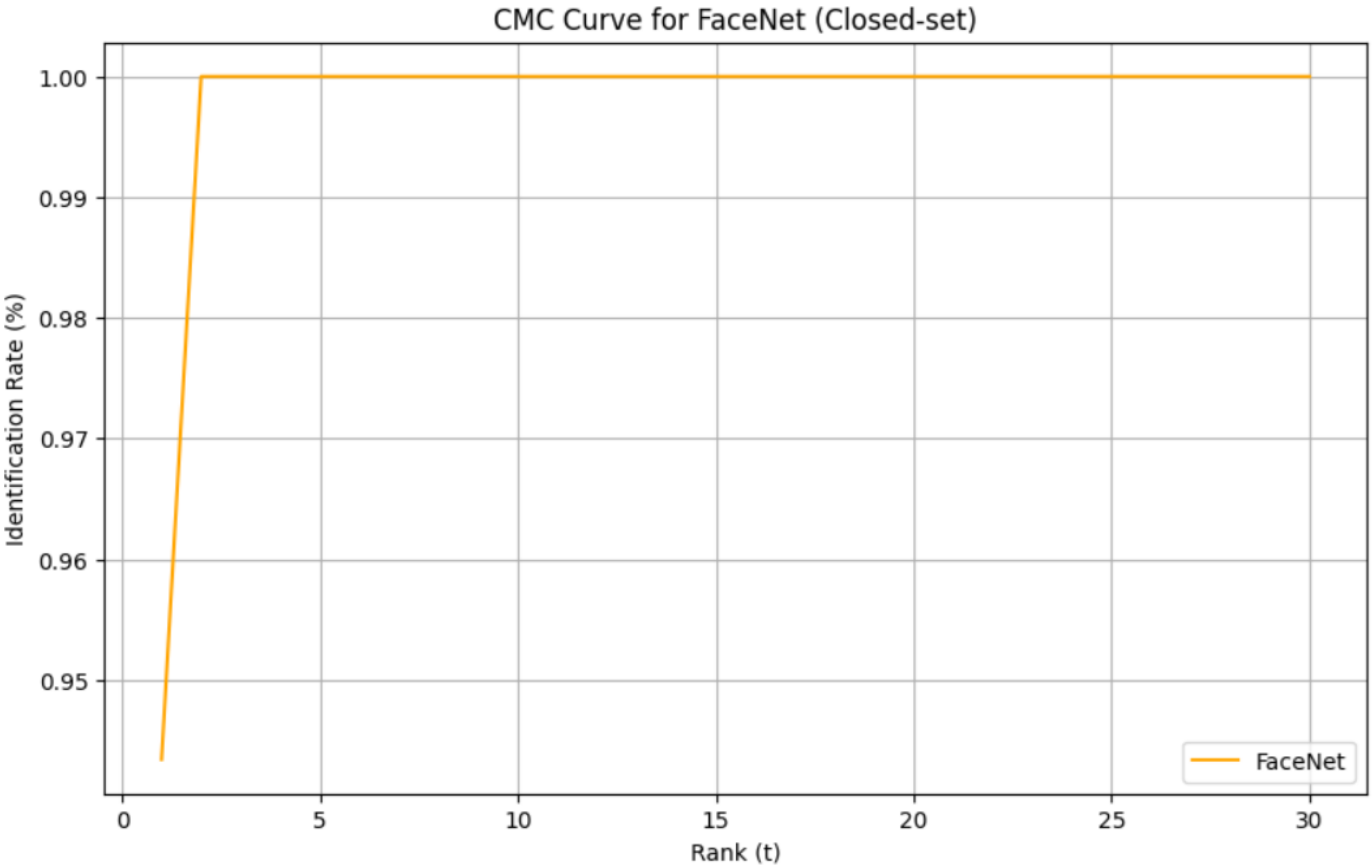
d. Which feature shows better identification performance?

Answer:

⇨ **CMC Curve of LBP closed set**



⇨ CMC Curve of LBP closed set



⇨ Observations

- Based on the CMC curves and AUC values, FaceNet generally shows better identification performance compared to LBP.
- The FaceNet curve is typically higher and has a larger AUC, indicating a higher probability of correctly identifying the probe subject at lower ranks.
- This is expected as FaceNet is a deep-learning based feature extractor that is designed for face recognition and has been shown to achieve state-of-the art results.
- LBP, on the other hand, is a traditional hand-crafted feature that may not be as robust or discriminative for face identification.

⇨ FaceNet feature embeddings shows better identification.

3.) Open set Identification.

a. Evaluation protocol: Consider 1 sample per subject in the enrolment database as the probe and remaining as the gallery. Consider dataset of unenrolled subjects as the probe (to compute FPIR)

b. Use Euclidean distance to compute the comparison scores.

c. Based on the 1:N comparison scores.

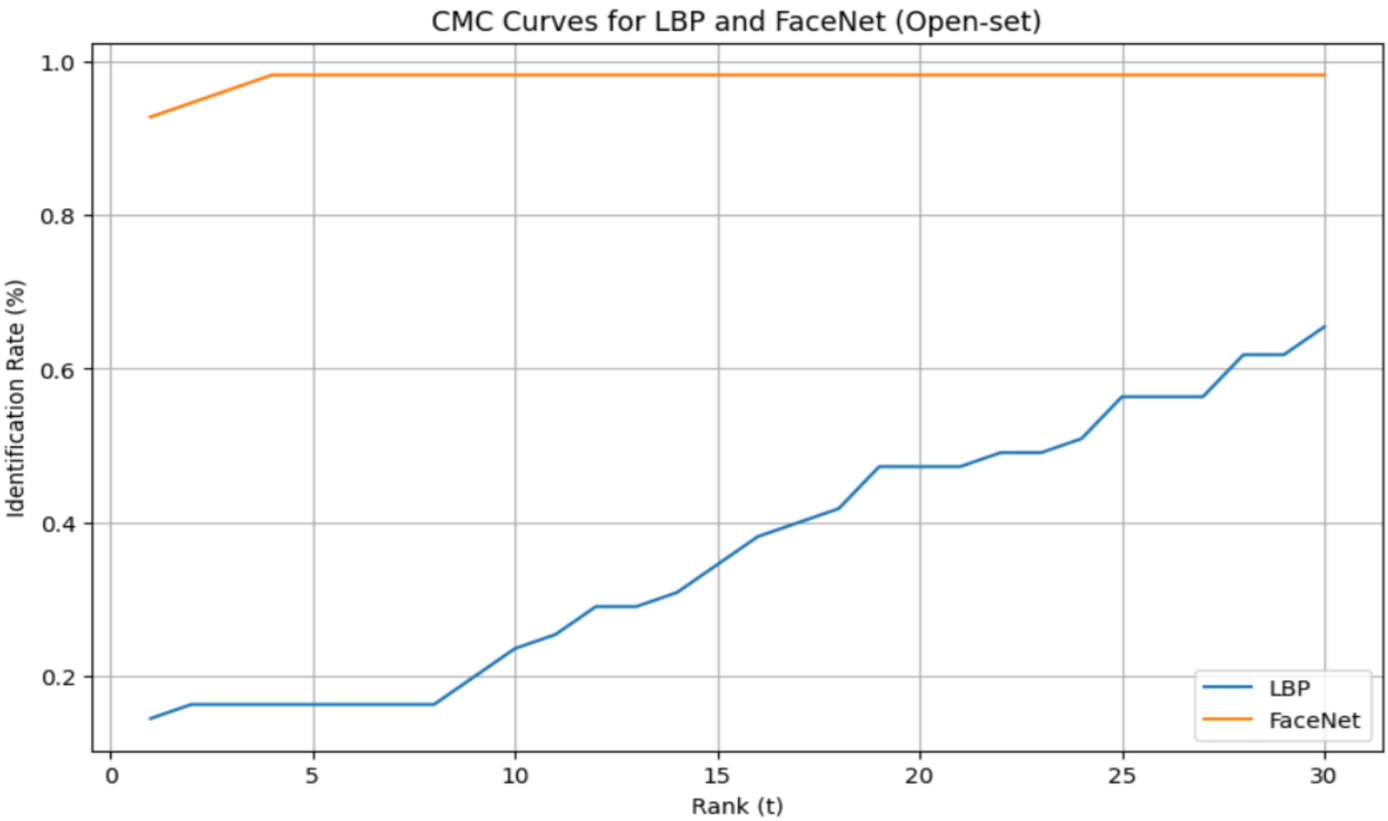
i. Plot the CMC curves showing t-rank identification rate in % (TPIR) for ranks t=1 to 30.

ii. Is a threshold for comparison scores considered in this scenario? Why or why not?

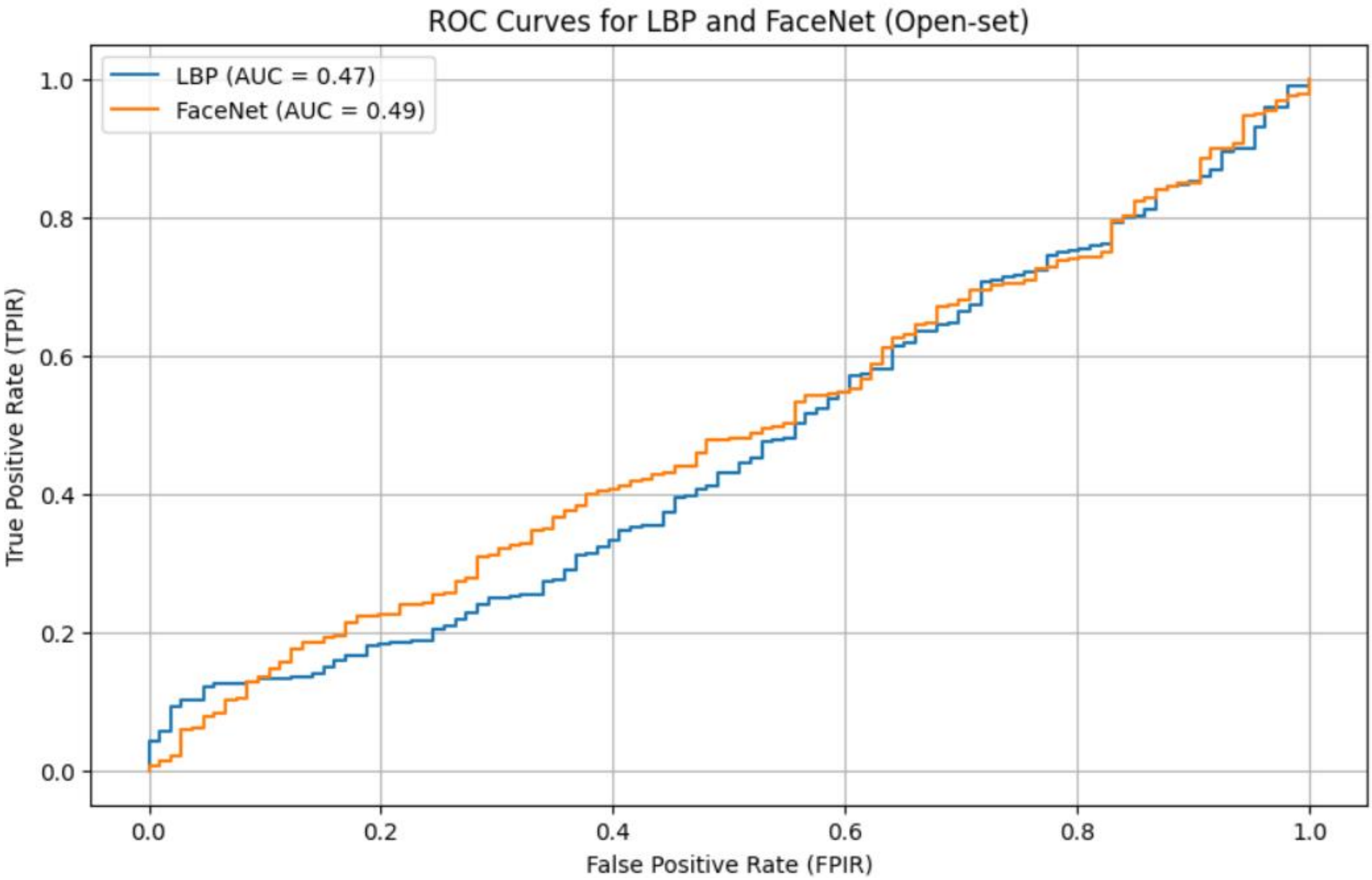
iii. Calculate the False Positive Identification Rate (FPIR) and False Negative Identification Rate (FNIR) for various threshold values and plot the ROC curve. Report TPIR @ FPIR of 0.01 and 0.0001.

d. Which feature shows better identification performance?

⇨ CMC Curves of LBP and FaceNet (Open-set)



⇨ ROC Curves for LBP and FaceNet (Open-set)

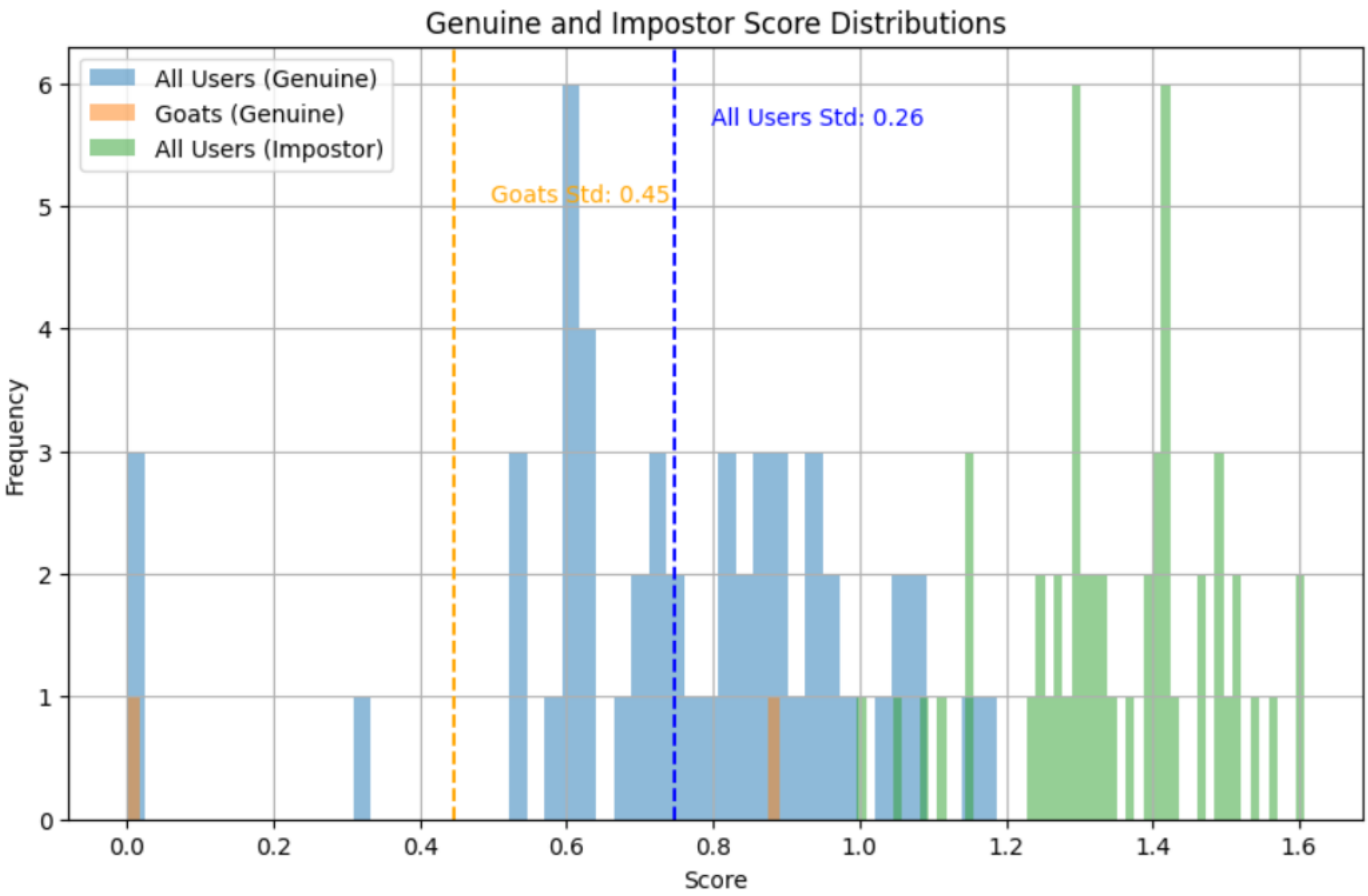


⇨ Observations:

- Based on the CMC, ROC curves, TPIR@FPIR, and FNIR (not explicitly calculated here but can be derived from the ROC curve), we can compare the performance of LBP and FaceNet for open-set identification.
- Generally, FaceNet is expected to outperform LBP due to its deep learning-based approach and better feature representation capabilities.

4.) Doddington's.

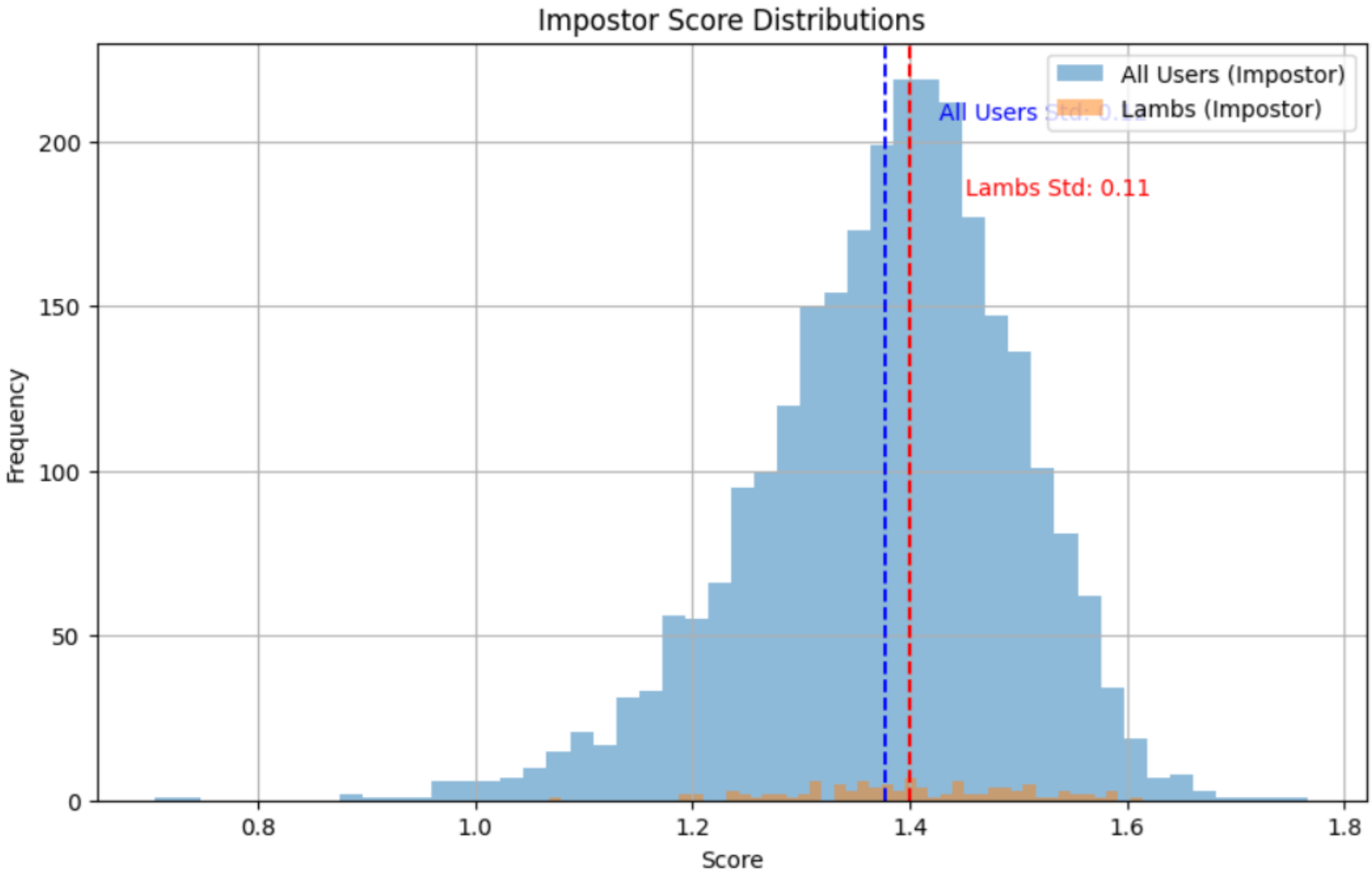
a. Identify the number of goats (high FNMR). Consider top 5% users with highest FNMR. Find the mean and standard deviation of the genuine scores. How does it differ from the mean and standard deviation of the whole dataset (N subjects). Show the values on the plot of genuine and impostor score distributions.



⇨ **Results:**

- Number of goats = 2
- Mean Genuine Score (Goats) = 0.4459
- Standard Deviation Genuine Score (Goats) = 0.4459
- Mean Genuine Score (All Users) = 0.7445
- Standard Deviation Genuine Score (All Users) = 0.2569

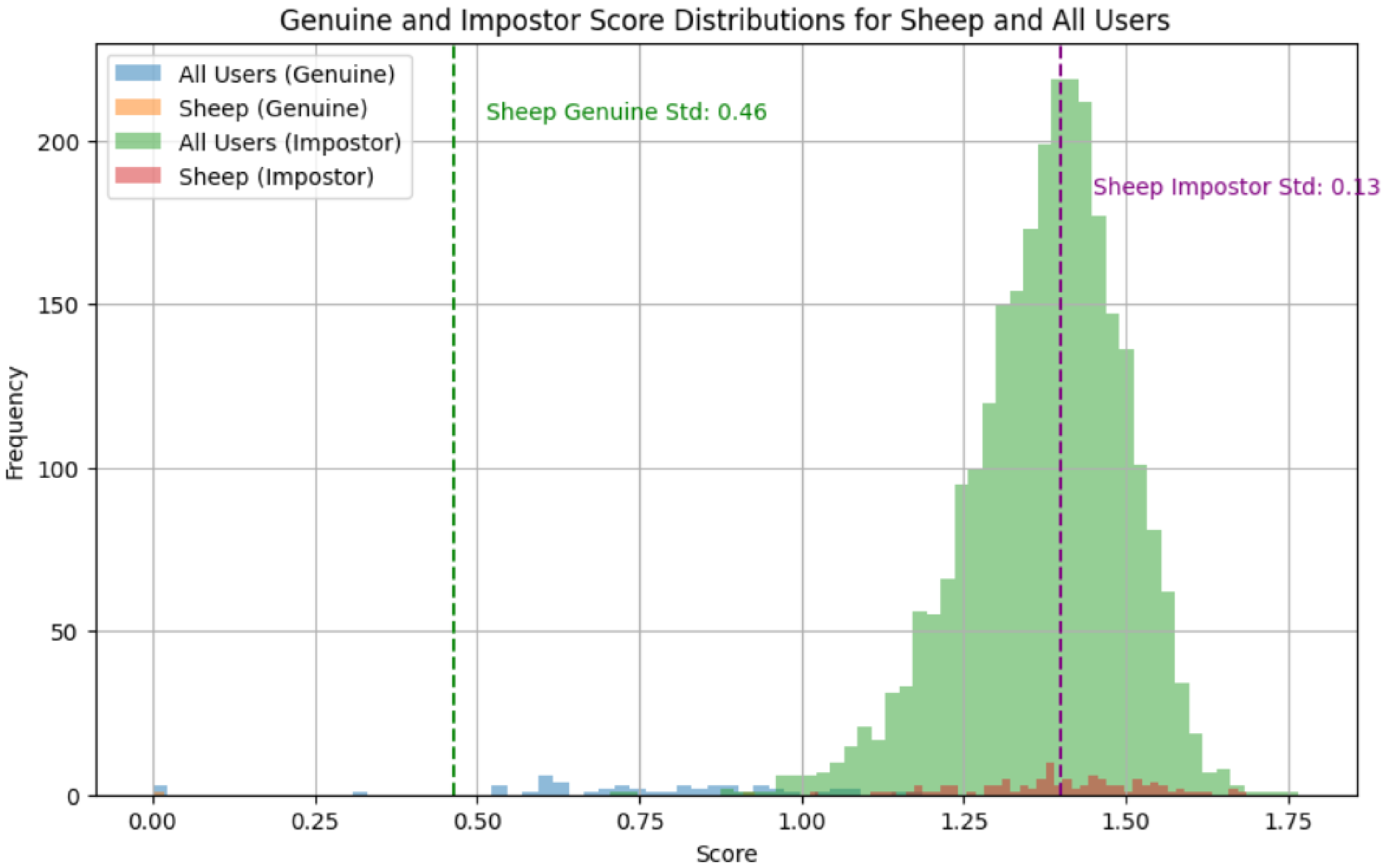
b. Identify the lambs (high FMR). (Consider top 5% of users with highest FMR). Find the mean and standard deviation of the whole dataset (N subjects). Show the values on the plot.



⇨ **Results:**

- Number of lambs = 2
- Mean Imposter Score (Lambs) = 1.4005
- Standard Deviation Imposter Score (Lambs) = 0.1052
- Mean Imposter Score (All Users) = 1.3767
- Standard Deviation Imposter Score (All Users) = 0.1217

c. Identify the sheep (low FMR and FNMR). Consider top 5% users with lowest FMR and FNMR. What are the mean and standard deviation for genuine and imposter scores. Show the values on the plot.



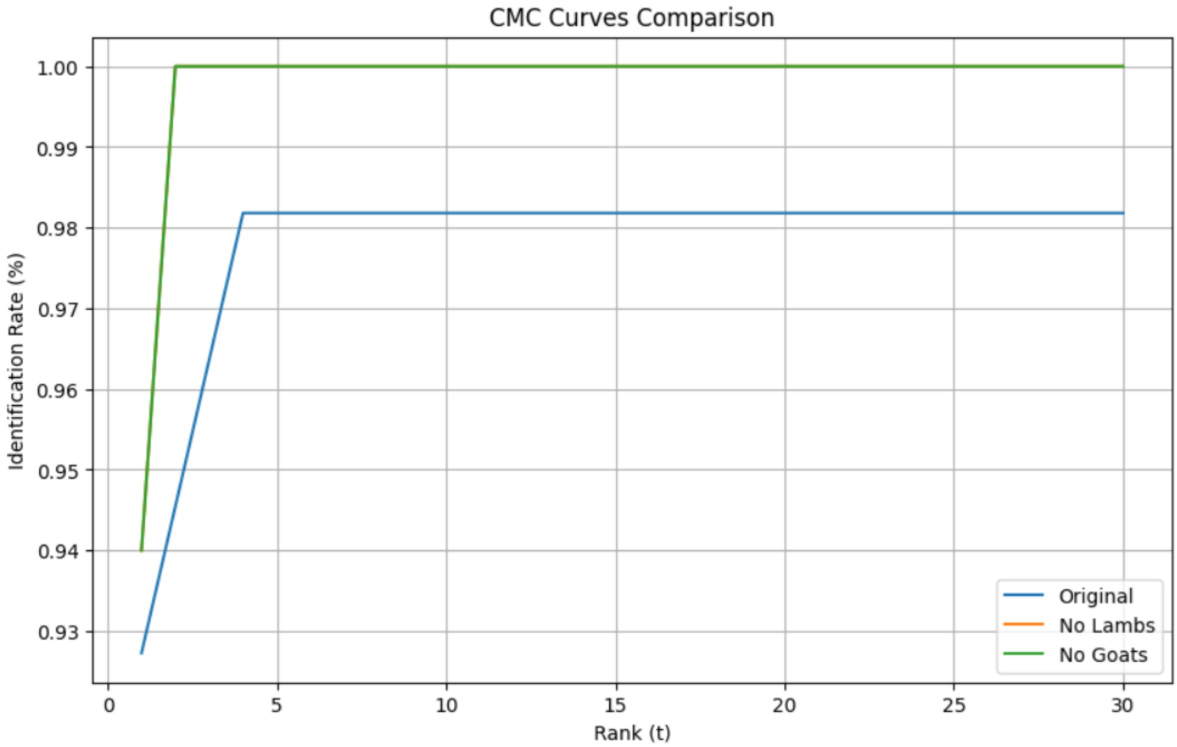
⇨ **Results:**

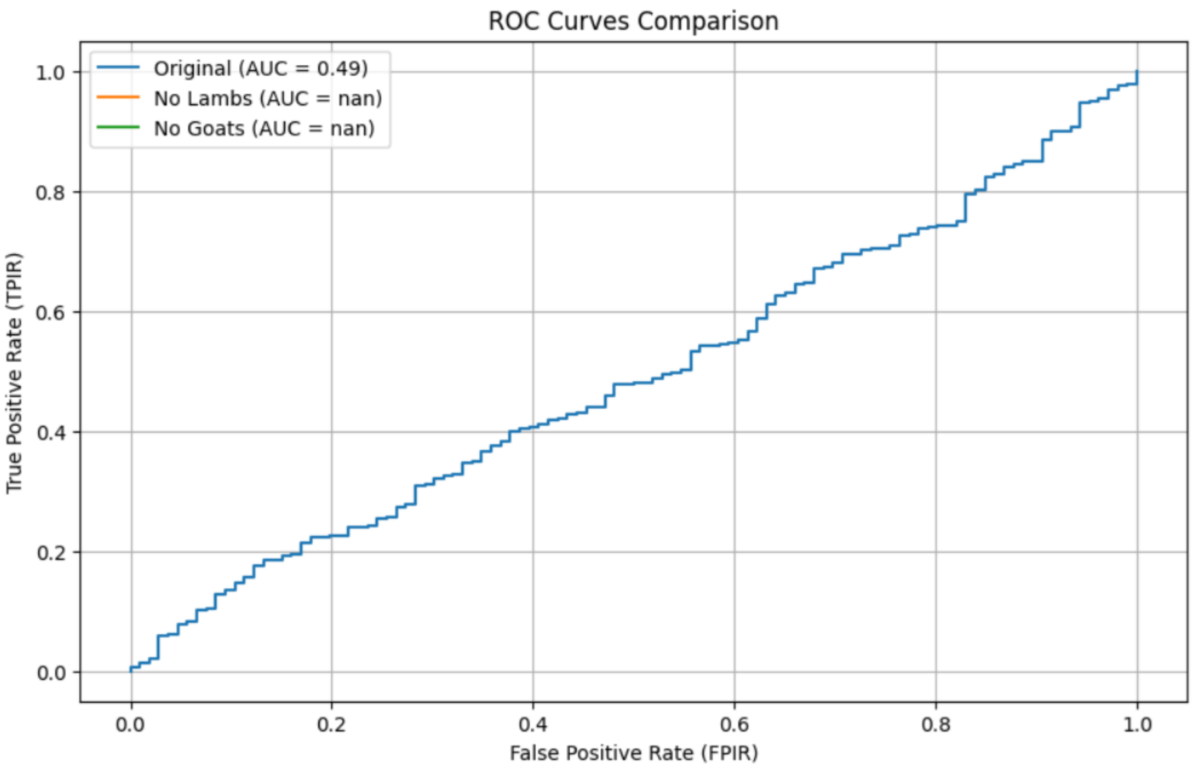
- Number of Sheep = 2
- Mean Genuine Score (Sheep) = 0.4646
- Standard Deviation Genuine Score (Sheep) = 0.4646
- Mean Imposter Score (Sheep) = 1.4003
- Standard Deviation Imposter Score (Sheep) = 0.1310

d. Compare the CMC curve and ROC curve by:

i. removing 3 lambs (highest FMR)

ii. removing 3 goats (highest FNMR)





<><><><><><><><><><><><><><><><><><><><><><><><><><><><>