Biometrics (COMP 388-002 / 488-002)
Loyola University Chicago, Fall 2025
Assignment 2: Fingerprint Recognition
Due date: Oct 17, 11:59 PM ET
Total: 10 points

1. Introduction

The purpose of this assignment is to train and evaluate students' abilities to adapt, use, and assess the performance of a third-party library for fingerprint recognition, leveraging the fingerprints collected in class. To do so, a zip file is provided, which contains the dataset of collected fingerprints, properly organized into genuine and impostor pairs, as well as the fingerprint recognition library.

1.1. Fingerprint dataset

The fingerprint dataset is available at https://tinyurl.com/mr5c5ajf and is password-protected. To obtain the password, please contact the instructor (Daniel Moreira) through email (dmoreira1@luc.edu). **This dataset must not be shared with anyone by the students and must be deleted immediately before the end of the course. Anyone breaking these rules will automatically fail the course.** Please refer to Prof. Moreira if you have any questions.

The fingerprint dataset consists of 79 image files in bitmap format (*.bmp*), each containing a single fingerprint. In addition, the text files *genuine.txt* and *impostor.txt* define, respectively, the genuine and impostor pairs of fingerprint images that the students are expected to process during this assignment. The contents of these files follow the same format: two fingerprint image file paths separated by a comma in each line. Figure 2 depicts the content of *genuine.txt* for illustrative purposes.

```
1251.bmp,5184.bmp
7104.bmp,8805.bmp
4223.bmp,1134.bmp
(...)
```

Figure 2. The first three lines of the file *genuine.txt*. Take line 1 as an example: it states that image files *1251.bmp* and *5184.bmp* depict the same fingerprint collected on different occasions (i.e., a genuine pair of fingerprints). File *impostor.txt* follows the same format but, in its case, the line-wise pairs refer to different fingerprints (i.e., impostor pairs of fingerprints). Each file contains 61 lines (and therefore 61 fingerprint pairs) that the students must process while answering this assignment.

1.2. Fingerprint recognition and metrics libraries

The fingerprint recognition library is available as a Google Colab notebook (https://tinyurl.com/3xet5mwx), which was explained in class. In addition to this implementation, you will also need either the first assignment's notebook (metrics' implementation, available at https://tinyurl.com/53knr8y5) or the *scikit-learn* (https://tinyurl.com/nsyu2k9b) library to compute the proper metrics and analyze the performance of the system.

1.3. Assignment directions

To complete this assignment, access the Google Colab notebook (https://tinyurl.com/3xet5mwx) and make your own copy. After downloading and unzipping the dataset files from Sakai to your local computer, upload them to your Google Colab notebook copy. Lastly, follow the instructions and answer the questions presented in Sec. 2. **Plan for a 1-hour software runtime to process all the fingerprint pairs of this exercise.**

There is no formal template for providing your answers. You may use the editor you like. The following option should work fine:

- A single PDF file or Word document containing all your answers and generated figures.

Please submit your file through the respective open assignment in Sakai by October 17, 2025, 11:59 PM ET.

2. Questions

Considering the content of both *genuine.txt* (with 61 pairs of fingerprint image file paths) and *impostor.txt* (also with 61 pairs of fingerprints), please answer the following questions.

2.1. For each one of the 122 available pairs of fingerprint images (61 from *genuine.txt* and 61 from *impostor.txt*), provide the minutiae-based similarity score, as defined in slide 39 of the presentation available at https://tinyurl.com/ydwyz2jx. To present these scores, generate a single *output.csv* file with 122 data lines; the first 61 data lines must be respective to the 61 lines of *genuine.txt*, while the following 61 data lines must be respective to *impostor.txt*. Lines with comments must start with "#". The format of this file is explained in Figure 3 through an example, and it follows the same format as the input files used in the first assignment. (4 points)

```
# System output. Line format: label [0: impostor, 1: genuine],
score
1,0.56
1,0.4367816091954023
(...)
0,0.18018018018018017
0,0.14689265536723164
(...)
```

Figure 3. Expected content for *output.csv*. The scores and number of lines presented here are for the sake of illustration.

2.2. Based on your obtained scores, what score threshold (a.k.a. operating point) should you use for this system? Please explain your answer and describe how you have computed this threshold. (1.5 points)

2.3. Plot and provide a graph with the distribution of the scores obtained by the system. What is the system's d-prime value? (1.5 points)

2.4. Plot and provide a graph with the ROC curve and AUC of the system. Is this system working better than chance? Please explain your answer. (1.5 points)

2.5. In your opinion, would this solution be robust to fake fingerprints such as silicon fingers? Please justify your answer. (1.5 points)