**Biometrics (COMP 388-002 / 488-002)**
Loyola University Chicago, Fall 2023
Assignment 2: Fingerprint Recognition
Due date: Oct 16, 11:59 PM ET
Total: 10 points

## 1. Introduction

The purpose of this assignment is to train and evaluate the students' abilities to adapt, use, and evaluate the performance of a third-party library of fingerprint recognition, leveraging the fingerprints collected in class. To do so, one zip file is being 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://bit.ly/3LFyw0s and is password protected. To obtain the password, please contact the instructor (Daniel Moreira) through e-mail (dmoreira1@luc.edu). **This dataset must not be shared by the students with anybody and must be deleted right 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 has 30 image files in bitmap format (*.bmp*), each one containing one fingerprint. In addition, the text files *genuine.txt* and *impostor.txt* respectively define 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 illustration sake.

```
91081.bmp,15465.bmp
30767.bmp,78494.bmp
06185.bmp,73226.bmp
(...)
```

**Figure 2**. The first 3 lines of the file *genuine.txt*. Take line 1 for an example: it says image files *91081.bmp* and *15465.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 15 lines (and therefore 15 fingerprint pairs) that the students are expected to process while answering this assignment.

### 1.2. Fingerprint recognition and metrics libraries

The fingerprint recognition library is available as a Google Colab notebook (https://bit.ly/48z6v4q), 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/bdjeeh5k) or the *scitkit-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://bit.ly/48z6v4q) and make your own copy. After downloading and unzipping the dataset file 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.

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 16, 2023, 11:59 PM ET.

## 2. Questions

Considering the content of both *genuine.txt* (with 15 pairs of fingerprint image file paths) and *impostor.txt* (also with 15 pairs of fingerprints), please answer the following questions.

2.1. For each one of the 30 available pairs of fingerprint images (15 from *genuine.txt* and 15 from *impostor.txt*), provide the minutiae-based similarity score, as defined in slide 41 of the presentation available at https://bit.ly/3PD35EZ. To present these scores, generate a single *output.csv* file with 30 data lines; the first 15 data lines must be respective to the 15 lines of *genuine.txt*, while the following 15 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 of the input files used in the first assignment. (4 points)

```
# System output. Line format: label [0: impostor, 1: genuine],
score
1,0.6285714285714286
1,0.6046511627906976
(...)
0,0.1326530612244898
0,0.2732919254658385
(...)
```

**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 the scores you have obtained, what score threshold (a.k.a. operating point) should you use for this system? Please explain your answer and describe how you have obtained 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)