Biometrics (CSE 40537 / 60537)
University of Notre Dame, Spring 2022
Assignment 2: Fingerprint Recognition and Performance Assessment
Due date: March 4, 11:59 PM ET
Total: 10 points

1. Introduction

The purpose of this assignment is to train and evaluate the students' capabilities to adapt, use, and evaluate the performance of a third-party library of fingerprint recognition, leveraging the fingerprints collected in class. To do so, two zip files are being provided, one containing the third-party library (which was presented in class), and the other containing the dataset of collected fingerprints, properly organized into genuine and impostor pairs.

1.1. Fingerprint recognition library

The fingerprint recognition library is available at https://bit.ly/3gZKpyq. To use it and complete this assignment, you will need a computer with *Python 3* (https://www.python.org/downloads/), plus the *numpy* (https://numpy.org/), *scikit-image* (https://scikit-image.org/), and *OpenCV* (https://opencv.org/) libraries. A quick install of these libraries can be done through *PyPI* (https://pypi.org/), see Figure 1:

```
user@host:~$ pip3 install numpy opencv-contrib-python scikit-image
```

Figure 1. Command line leveraging *PyPI* to install the needed libraries.

A good starting point to solve this assignment is the *Python* program implemented as *main.py*, within the fingerprint recognition library. It currently works by providing the file paths of two fingerprint images and the extension to multiple pairs of images should be straightforward to the students (and it is indeed expected from them).

1.2. Fingerprint dataset

The fingerprint dataset is available at https://bit.ly/3LIVm5z and is password protected. To obtain the password, please contact the instructor (Daniel Moreira) through either Slack https://bit.ly/3H8kajL, *@dmoreira*) or e-mail (daniel.moreira@nd.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 296 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.

```
851475.bmp,025150.bmp
851475.bmp,430501.bmp
851475.bmp,634673.bmp
(...)
```

Figure 2. First 3 lines of file *genuine.txt*. Take line 1 for an example: it says image files *851475.bmp* and *025150.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 577 lines (and therefore 577 fingerprint pairs) that the students are expected to process while answering this assignment.

1.3. Assignment directions

After downloading and unzipping the contents of the fingerprint recognition library and fingerprint dataset, follow the instructions and answer the questions presented in Section 2.

There is no formal template for providing your answers. You may use the editor you like. The following files are expected:

- A single PDF file or Word document containing your answers and generated figures.
- A single text file named *output.csv*, whose content is described in Figure 3.

Please share your answers through Slack (https://bit.ly/3H8kajL, @*dmoreira*) or send them to daniel.moreira@nd.edu by March 4, 2022, 11:59 PM ET.

2. Questions

Considering the content of both *genuine.txt* (with 577 pairs of fingerprint image file paths) and *impostor.txt* (also with 577 pairs of fingerprints), please answer the following questions.

2.1. For each one of the 1154 available pairs of fingerprint images (577 from *genuine.txt* and 577 from *impostor.txt*), provide the minutiae-based similarity score, as defined in slide 40 of the presentation available at https://bit.ly/3gVJ5N7. To present these scores, generate a single *output.csv* file with 1154 data lines; the first 577 data lines must be respective to the 577 lines of *genuine.txt*, while the following 577 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 mere 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 point)

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 point)

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 point)

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