RIT

Rochester Institute of Technology

CSEC 472 Authentication and Security Models
BS in Cybersecurity
Department of Cybersecurity

Lab 4 Biometrics

Learning Outcome: Students will refine their understanding of applied biometrics by performing 1) fingerprint image processing and 2) feature extraction for authentication decisions.

Lab Set Up:

1) For this lab, you can use your computer or a VM.

2) You can use the programming language of your choice to perform the lab; you have some freedom to choose the tools you want to work with, but you'll probably need to import one or two image processing/machine learning libraries, as in Python's:

Numpy, Scikit-image (skimage), Scikit-learn, PIL/Pillow, OpenCV-Python, SimpleCV, Mahotas, SimpleITK, pgmagick, pycairo, Scipy, Theano, TensorFlow, Keras, PyTorch, Pandas, Matplotlib, etc.

3) Import the NIST 8-Bit Gray Scale Images of Fingerprint Image Groups (FIGS) database from this torrent.

https://academictorrents.com/details/d7e67e86f0f936773f217dbbb9c149c4d98748c6

Publication **DOI:** 10.18434/T4RP4K

Abstract (excerpt):
"The NIST database of fingerprint images contains 2000 8-bit gray scale fingerprint image pairs. Each image is 512-by-512 pixels with 32 rows of white space at the bottom and classified using one of the five following classes: A=Arch, L=Left Loop, R=Right Loop, T=Tented Arch, W=Whorl. The database is evenly distributed over the five classifications with 400 fingerprint pairs from each class… Each of the fingerprint pairs is two completely different rollings of the same fingerprint… Suitable for automated fingerprint classification research, the database can be used for algorithm development, system training, and testing.

The database is a valuable tool for evaluating fingerprint systems on a statistical sample of fingerprints evenly distributed over the five major classifications. System Requirements: software to open zip files

and extract database and software to view PNG formatted images. The size of the zip file is 7895 MB, and the extracted database is 807 MB. For more information on Special Database 4, please contact: Standard Reference Data Program National Institute of Standards and Technology 100 Bureau Dr., Stop 8500 Gaithersburg, MD 20899-8500 (301) 975-2200(VOICE) / (301) 975-4553 (FAX) Contact Us The scientific contact for this database is: Patricia Flanagan Information Access Division National Institute of Standards and Technology 100 Bureau Drive, Stop 8940 Gaithersburg, MD 20899-8940 (301) 975-4965 flanagan@nist.gov"

Figure A.1: Fingerprint file **f0001_01.pct** from *NIST Special Database 4*.

NOTE: Most images used in authentication systems are converted to grey scale. This transforms the image into a matrix (size = n * m, where n and m are the pixel size of the image). This image would become an n * m * c matrix if it were in color, where c represents specific color markers (like RGB).

SIDE NOTE: The NIST dataset can be found through Google's new database search function (https://datasetsearch.research.google.com/). It may be a helpful source of information in your group projects.

Instructions:

EACH GROUP will develop a features extraction system that makes accept/reject decisions. Given a fingerprint, they will design a unified system and compare it to what they built individually.

1). Download and examine the database. It contains 400 image pairs of the five main fingerprint features (arches, loops, etc). The features are encoded on the .txt file that corresponds with the .png, dispersed randomly throughout the dataset. The database is organized as image pairs = $\{f_i, s_i\}$

You'll want to break the overall database into two sets:

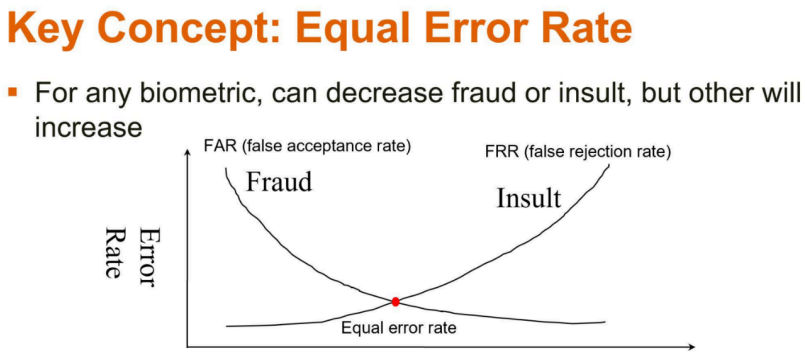
a. TRAIN: the first 1500 image pairs (f0001-f1499 & s0001-s1499)
b. TEST: the last 500 image pairs (f1501-f2000 & s1501-s2000)

Process the TRAIN set to extract features (identify minutia). **Each group will use three different methods**. This will inform the lab experiment.

To do this, you'll calculate minutia for each reference image (f_i) and its corresponding subject image s_i. You'll generally compare distances between these images. **Try three different image matching or ML techniques to do this**.

For each method, document your max, min, and average false reject and false accept rates and calculate your equal error rate.

From lecture slides:



2). Build a table documenting which methods you used and the max, min, and average for both false reject and false accept rates, as well as the equal error rate (error crossover).

Worth 25% of your lab grade: The table should have *three* rows: three different methods. The table should have eight columns: method, FRR avg, FRR min, FRR max, FAR avg, FAR min, FAR max, EER.

3). *Analyze* the table and re-design the feature extraction to reduce the average EER across the three methods. Your new system should apply *all three methods and return the majority decision*; this will create a hybrid system. DON'T TEST THE SYSTEM YET! You'll do that in Step 4.

Worth 25%: Document your hypothesis – what do you think (as a group) will happen with your hybrid system? Will the ERR go up, down, or not change compared to the other systems?

4). Run the hybrid system

Worth 25%: Add a row to your methods table documenting the FRR avg, min, max, FAR avg, min, max, and ERR for your group's hybrid system.

5). Interpret & analyze your findings.

Worth 25%: Write a lab report and answer these questions: What happened? Was your hypothesis correct? What surprised you most? If you were to repeat the experiment, what would you do differently? What questions do you still have?

Deliverables:

1. Lab report. The table & writeup (upload to mycourses)

2. Source code. Please put all team members' code in one repo for this project & submit the link with your lab report to mycourses-> assignments-> lab4

Grading:

25% Table with each method's result (3 rows, 8 columns)
25% Group's hypothesis of what will happen with hybrid system
25% Table with hybrid results (4 rows:3 individual methods+ a hybrid one, 8 columns)
25% Write up your interpretation/analysis

Lab 4 Biometrics

Group Number:

Workload Distribution:

Lab Number: _____ Group Number: _____

Members	Tasks	Total hours spent	Percentage of total effort
Student 1			
Student 2			
Student 3			
Student 4			
Student 5			