University of Southampton
Electronics and Computer Science

A report submitted for COMP6211 Biometrics Coursework

By Sarunas Iljeitis
28825004
si1g16@soton.ac.uk

Biometrics system based on human body shape recognition

May 25, 2020

Abstract

Various biometrics recognition systems have become more prominent in the current day society. Aside from traditional methods such as fingerprint, eye retina and face recognition techniques less direct approaches are being investigated. This report describes a straightforward method of person recognition by using certain body parts measurements. The system deals with both front and side images and compares them against two separate front and side data sets giving it capabilities to identify a person from just one picture. Initially fourteen measurements are taken from each image following along key points of human skeleton, although following experimentation results only five measurements were used in the end. This approach yields a best case Correct Classification Rate of 81.81% and an Equal Error Rate of 31.82%. The report also covers covers approach evaluation and possible alternatives as well as future improvements.

Contents

Intro

Computer based human identification technologies are continuously being integrated to our daily lives. In 2017 Apple with the release of Iphone X introduced Face ID feature that allows user to unlock their devices using front camera image of their face[2] and a more conventional method of unlocking smartphones with a fingerprint has been available on majority of mainstream devices since around 2013[5]. Since biometric identification techniques have been so widely used in consumer technologies, it is no surprise that various governments around the globe have integrated more advanced systems on a much wider scale. China, for example, is known to be extensively using computer vision techniques to track certain ethnic minorities living inside it's borders. The project is said to have been scaled to effectively track more than 11 million inhabitants[4]. Identification techniques are widely known to have their drawbacks as majority of them require accurate measurements that often are not possible and their results can be affected by certain clothing and or wearable accessories such as sun glasses. Therefore, depending on their application biometric identification systems vary in complexity and accuracy.

Biometrics systems have two main modes of operation: verification and identification and are described by a number of performance metrics. In the verification mode the user claims to be a certain individual and presents a it's biometric data to be measured such as his finger. The system then only needs to do a one-to-one measurement comparing the provided data with the one stored in it's data set. If the difference between two measurement is below the threshold, the users identity is confirmed giving a binary outcome. The identification mode, in contrast, works by comparing the users biometric data against all the instances in the data set and seeing if any of the entries are similar enough to pass the threshold. The threshold in both cases is necessary to prevent False Acceptance (FA) cases where a user is falsely identified as the claimed person or as a person in the data set. It also has to be balanced against False Rejection (FR) cases that happen when a legitimate user is not identified as part of the set. If expressed as a percentage of the overall test cases, these metrics are called False Acceptance Rate (FAR) and False Rejection Rate (FRR) respectively. The point where the FAR and the FRR are equal, is called the Equal Error Rate (EER) and, generally, the lower it is the better the algorithm. Depending on the circumstances the FAR and FRR can have varying importance as they directly effect robustness and system convenience respectively and have to balanced depending on the circumstances.

Method

The approach used for this system is largely simplistic but is an excellent example that it is possible to achieve moderate results with limited data. The system is essentially 2 Python3 scripts that can be are run separately and is divided into

2 stages: preparation and classification. In the preparation stage the training data set is processed: each subject measurements are taken and stored in the subject data set. In the classification stage the same measurements are taken from the entry images and a KNN classifier (with K = 1) is used to determine best match for the entry. If the match distance is less than the threshold, the systems treats it as a suitable match and if the distance is greater than the threshold, the system treats as if no match was found and an entry does not identify as any of the data set subjects.

The preparation stage consists of these actions as illustrated in Figure 1. First step is to use a pre-trained human body recognition model called "Deeplabv3-ResNet101"[1] from PyTorch to extract the body mask and the mask is applied to remove the background. The image is saved to the disk to be reused on new runs and reduce runtime. Then the person contour is calculated using OpenCV findContour() method. The image is then cropped to 800px width and 1400px height by horizontally centering the image (same space left in both sides of the contour is calculated by $(800 - contourthickestpart)/2$. The contour lowest part is placed at the bottom of the image of the image. Next stage is to use another pre-trained PyTorch model KeypointRCNN[3] which returns 17 body parts points displayed in Figure 1 and 2 3rd picture. Using these points, it is identified whether it is a side or front image using the hips width. These points are then used to device 14 metrics initially planned to be used for the subject identification. These metrics are: height, head width (and height of the thickest part), necks thinnest part width (and it's height), shoulders, hips, knees and ankles widths and heights, also angle between middle of the shoulders and the nose. The measurements together with image name and orientation are then saved to the disk as a YAML file.

Figure 1: Illustration of front image preparation stages: (from the left) initial image, human shape mask, human body parts labelled, points which are measured and then stored

The classification stage follows these steps. Input image goes through the preparation stages, the output is saved as a YAML file to be reused later on. The training data is loaded from the YAML file, divided into front and side views and standardized using the Standard Score (ZScore) scaler. The input image data is also scaled together with one of the sets depending on the image orientation. Finally, a KNN classifier with K = 1 loaded with either front or side data set is used to determine the closest match. The classifier also receives a threshold and the calculated distance between closest matches is compared against it to determine if the match is accepted.

After some experimentation it was concluded that due to subjects clothing irregularities and difference in posture, most of the metrics cannot be used and do not help to classify images, therefore only 5 metrics where left in the final version: height, head width and widest part height, neck thinnest part width and thinnest part height. Surprisingly, the combination of these metrics proved to be the most reliable and consistent.

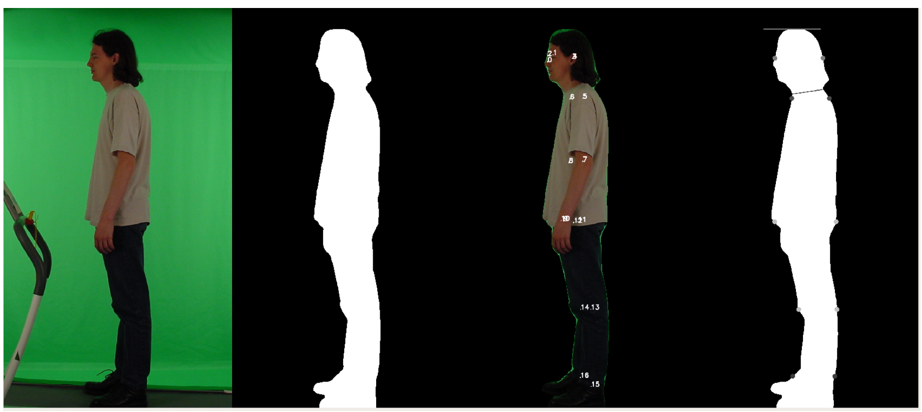

Figure 2: Illustration of side image preparation stages: (from the left) initial image, human shape mask, human body parts labelled, points which are measured and then stored

Results

The system is capable of working with just one person image either it being a front or a side picture. The Figure 3 shows a Bar chart of maximum Correct Classification Rate (CCR) plotted against the number of features used for classification. Surprisingly, the best rate is achieved when using just 5 or 6 features. Therefore, the maximum achieved CCR value is 81.81%. The fall in CCR value when more measurements are included can be attributed to those measurements being inaccurate in general, such as hip width which can be affected by the clothing or distance between ankles which changes depending on person feet position. Both front and side classifiers correctly classify 9 out of 11 people, therefore also having 81.81% CCR rates if used separately. Which is also surprising as the system is built in such a way that side and front views can have different measurements used for calculations but for both views it was the determined that the same 5 measurements: height, head width and widest part height, neck thinnest part width and thinnest part height work equally well for both.

It was also observed, but not implemented due to limited use and time constraints, that if front and side images were combined and used to identify a single person (two images per person classification), the CCR can rise to 90.90% (10 out of 11). This can be achieved by doing separate KNN classifications as before for both front and side images and if two KNN classifiers come up with different results, then the classifier which nearest neighbour distance is smaller would be trusted over the other. As this method would not make a significant improvement and would not affect EER value, the correctness these statements can be verified by manually reviewing distances of both classifiers and mapping front and side test images.

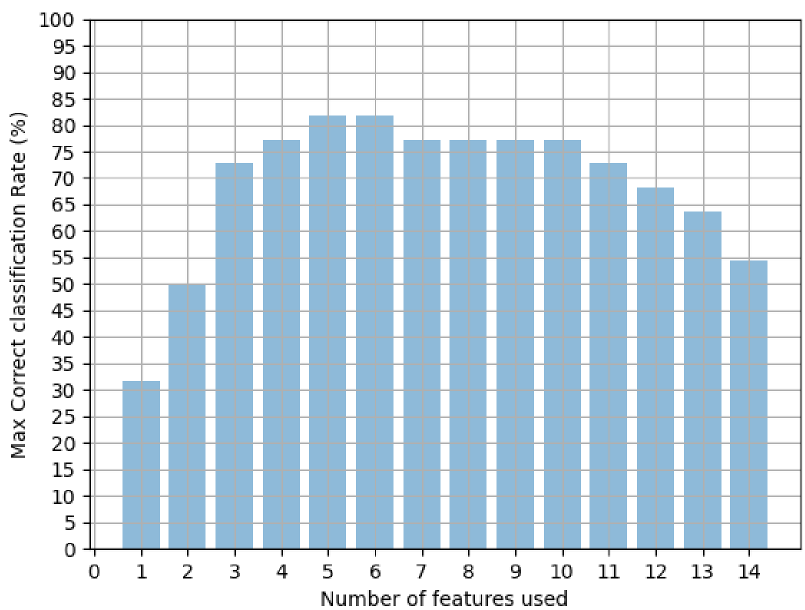

Figure 3: Best Correct Classification Rate plotted against the number of features used

The False Rejection Rate (FRR) was calculated by

$len(tests\_cases) - authenticate(test\_cases, database\_entries))/len(test\_cases)$ where $authenticate()$ is a binary function returning number of test subjects identified under the threshold and $test\_cases \subseteq database\_entries$. The given data set allows 22 individual tests.

The False Acceptance Rate (FAR) was calculated by

$authenticate(test\_cases, database\_entries)/len(test\_cases)$ where $test\_cases \cap database\_entries = \emptyset$.

The Figure 4 displays the FAR and FRR graphs. From these measurements the EER is computed to be 31.82%. The FRR and FAR graphs both seem to have a steep rise/fall immediately after the threshold value introduction and is significantly affected by minimal threshold changes. It seem to indicate that the system is highly dependant on precise biometric measurements and less than ideal conditions can increase it's EER significantly. This is expected to be due to the low number of feature measurements. Additionally the graph also shows the algorithm accuracy by dividing the number of True Positive and True Negative outcomes by the number of all tests taken.

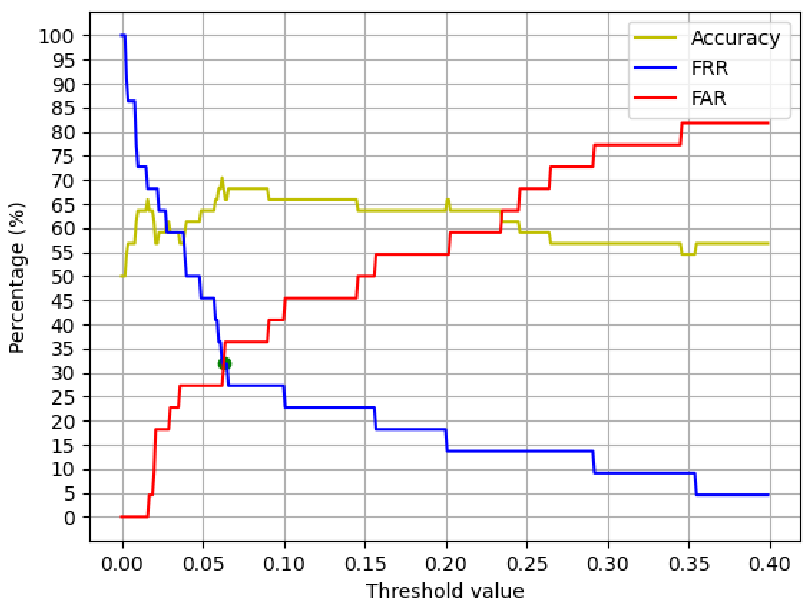

Figure 4: Equal Error Rate graph with addition of Accuracy $(TP + TN)/ALL\_TESTS$

Discussion

These results show that few measurements can be used to classify individuals quite effectively. On the other hand, due to the low number of measurements, such classification is system has high False Acceptance and False Rejection Rates and a low Accuracy. The problems can be addressed by searching for and using more robust body shape measurements. As it has been observed, measurements below the shoulder line are easily affected by clothing and posture, therefore it is likely that continuous evaluation of the neck and the head shape may provide better results.

As the dataset provided is fairly limited, it is complicated to account for measurements precision and as only one picture of each of the individuals is provided in the training dataset, image based measurements (such as EigenFace) are not possible and would be highly susceptible to minimum changes. The employed simplistic model based approach with the use of pre-trained ML models used for feature extraction, is not affected by the environment such as lighting, surroundings or image quality. Therefore, it may have potential to work in uncontrolled surroundings such as outside of the building. The few keypoint measurements along with employment KNN algorithm allows quick comparison as prototyping as most of the data: measurements, image foreground mask, human pose estimation points; can be cached and reused.

Further improvements could be achieved by using a larger dataset, which would contain more diverse data. With the current dataset improvements may be possible by using the segmentation and use measurements of individual body parts such as legs, torso or face (which would have the most distinct features) and comparing their distinct features, as well as measuring and comparing MU moments of the person head, neck and shoulders area as their measurements will have the least variance. It may also be possible to compare the head silhouette and comparing the coordinate sequence using Dynamic Time Warping algorithm[**?**].

Conclusions

The current system setup can be used to identify people whose data is already in the system but provides poor performance when non-dataset entries are provided. The system uses model based approach to compare five most distinct measurements of the image to uniquely identify the person. It's error rate could be further reduced by increasing the number of distinct points measured.

Appendices

Submission Files

- preparations.py - takes in the training set and export measurements as features_training.yaml

- main.py - takes in test data set and features_training.yaml file and produces either err or ccr graphs (change called function in main()) Also will cache test set measurements in features_test.yaml, thus lines 338 and 284 may need to be uncommeted to change use new measurements instead of features_test.yaml from before.

- features_test.yaml & features_training.yaml - cached measurements to be used by main.py for quick testing

- Images and graphs used in the report.

References

[1] Pytorch. https://pytorch.org/hub/pytorch_vision_deeplabv3_resnet101/.

[2] About face id advanced technology. https://support.apple.com/en-us/HT208108, Feb 2020.

[3] Francisco Massa. Pytorch. https://pytorch.org/blog/torchvision03/.

[4] Paul Mozur. One month, 500,000 face scans: How china is using a.i. to profile a minority. https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html, Apr 2019.

[5] Young-Hoo, Jeon, Seong-Yun, Jong-Hyuk, Lee, and Mun-Kyu. Security analysis and improvement of fingerprint authentication for smartphones. https://www.hindawi.com/journals/misy/2016/8973828/, Mar 2016.