Exploring the Feasibility of using Wrinkles as Biometrics

Miloni Mittal
Birla Institute of Technology and Science, Pilani (BITS Pilani)
f20170243@pilani.bits-pilani.ac.in

I. Introduction

Face recognition is one of the most important research areas in the field of computer vision. These methods depend on typical features such eyes, nose, mouth and more. More recently, micro-features such as wrinkles and freckles are being explored for the purpose of facial recognition. This has been made possible due to the increase in accessibility to higher resolution cameras. In the recent times, the use of mask has increased leading to an obstruction in full face recognition. This also demands for an alternate method to recognise faces, hence reinforcing the need to explore wrinkles as a biometric measure.

For a feature to be used as a biometric measure successfully, it should have the following characteristics:

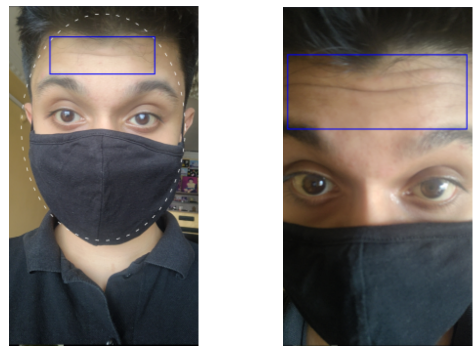
1) **Universality**: A biometric feature should be such that it is present in all people.
2) **Uniqueness**: A biometric feature should be unique to a person. Otherwise, it would lead to confusion and chaos.
3) **Permanence**: A biometric feature should not change over time. If it does, it would lead to discrepancies in identification.
4) **Collectability**: A biometric feature must be easily collectible by any authority. The device to collect it must be easily available.
5) **Acceptability**: People should be willing to provide the particular data.
6) **Circumvention**: It should be extremely difficult or impossible to forge the feature.

Here, an attempt has been made to explore the potential of using forehead wrinkles as a biometric.

II. Data Collection

The data was collected for about 320 subjects. For each of the subject, 20 pictures with combinations of face picture and forehead closeup picture (as shown in Fig 1) were taken. The subjects were instructed to wear a mask and lift their eyebrows to accentuate their forehead wrinkles. The collection was split into two sessions:

1) Session 1- This consisted of 5 sub-sessions where the subject clicked 2 pictures in each sub-session; one face picture and one forehead close-up.
2) Session 2- A repeat of Session 1 12 hours later. Two sessions were done so as to explore the results when data is collected over a varied time period.



(a) Face (b) Forehead Closeup

Fig. 1: Images collected in dataset

III. Models Used

A. SIFT (Scale-Invariant Feature Transform)

The SIFT technique was proposed by David G Lowe in [1]. It is a feature detection algorithm which is scale and rotation invariant. This means that this model can detect similarity between two images even if they are not of the same size or taken at different angles. SIFT operates on finding the keypoints or the local features in an image.

B. ORB (Oriented FAST and Rotated BRIEF)

The ORB techniques was proposed in [2]. It is an amalgamation of FAST (Features from Accelerated Segment Test) key point detector and BRIEF descriptor (with some modifications) [3] and considered to be more robust as compared to SIFT and SURF.

Both these models are available on OpenCV [4].

IV. Algorithm

For the purpose of this experiment, the data of 30 subjects was used. Each image was compared to the other images and a matching score was obtained as shown in Algorithm 1.

V. Performance Metrics

A. CRR (Correct Recognition Rate)

CRR is the number of actual matches that are obtained at rank one recognition.

Algorithm 1: Image pair score computation

```
Result: dataframe df consisting of image pair scores
initialization;
for All images in dataset do
    for All other images in dataset do
        read images;
        find match using openCV model;
        calculate score;
        save in df;
    end
end
```

B. Genuine/Imposter Histogram

The genuine/imposter histogram is the plot between frequency of score and score for genuine matches and imposter matches. It helps us to obtain a rough idea of the distribution of the scores in the two classes.

C. True Positive Rate vs False Positive Rate Curve

As the name suggests, a graph between True Positive Rate (TP) vs False Positive Rate (FP) was obtained for personal understanding of the models.

$$TP\_Rate = \frac{TP}{TP + FN}$$

$$FP\_Rate = \frac{FP}{FP + TN}$$

D. ROC Curve (Receiver Operating Characteristic Curve)

The ROC curve is a graph showing the performance of a model at various thresholds. It is a plot of False Rejection Rate (FRR) vs False Acceptance Rate (FAR). The point at which FRR = FAR is called the equal error rate point. This can be obtained by finding an intersection between the graph and a line with slope 45° passing through the origin.

$$FAR = \frac{FP}{FP + TN}$$

$$FRR = \frac{FN}{FN + TP}$$

E. FAR/FRR vs Threshold Graph

It is the plot between False Acceptance Rate, False Rejection Rate against various thresholds. It helps us to understand the trend between the two errors with respect to the thresholds used. The intersection point gives us the threshold at Equal Error Rate.

F. Accuracy

Accuracy is maximum value of (100 - $\frac{FAR+FRR}{2}$) across all thresholds.

VI. Experimentation and Results

The quantitative results and the graphs obtained have been presented in Table I and II respectively.

A. Steps for the Results

- First, the CRR was calculated.
- Then the genuine/imposter histogram was obtained. Relative scaling was used for the frequency axis (y-axis) so as to obtain easily readable graphs. This was important due to the presence of more imposter matches as compared to the genuine matches. Moreover, data points with score 0.0 and 1.0 were removed for the purpose of graph legibility. Two approximately Gaussian curves can be observed.
- A true positive rate vs false positive rate graph was obtained to understand the relation between the two metrics based on thresholds.
- The ROC Curve was obtained by plotting FRR vs FAR. A y=x line was plotted, to understand the characteristics of FRR and FAR at Equal Error Rate.
- FAR/FRR vs Threshold graph was plotted to verify the value of threshold at EER.
- Next, a dataframe with the thresholds and its respective performance metrics was built. FAR and FRR columns were checked for approximately equal values and the threshold was finalised on that basis. The value of FAR, FRR and threshold were cross verified by visual inspection of the ROC curve and FAR/FRR vs Threshold graphs.

B. Results and Observations

The resultant values of CRR, EER, Threshold at EER and Accuracy have been presented in Table I. Since many of the models did not have an EER point (run 1, 2, 3), the threshold for the point having least difference between FAR and FRR have been recorded. The FAR and FRR values have been presented in (FAR | FRR) form. The models for which EER occurs (run 4, 5, 6), have approximately equal FAR and FRR. The resultant graphs of the 6 experiments have been shown in Table II. We see that in models where there is an excess of score= 0.0 (run 1, 2, 3) in the case of genuine matches, the False Rejection Rate (FRR) shoots up to a high value hence preventing the 45° line from passing through the ROC curve. Additionally, the FAR/FRR vs Threshold curves also do not intersect at any point in such cases. In case we have a good model (run 4, 5, 6), we see that the 45° line from the origin passes through the ROC curve and the FAR/FRR vs Threshold curves also intersect, indicating an Equal Error Rate point.

VII. Ideas for betterment

There could have been discrepancies in collecting the data. This could be due to various reasons such as a photo being cut off at essential points, varying hairstyle across the sessions, presence of a background and so on. One method to streamline a process of uniform data collection could be the usage of special machines to collect the data in controlled environment. This idea is inspired from how fingerprints are collected (in a specific machine with controlled lighting) for biometric verification at offices. Though SIFT and ORB are robust techniques and capable of handling such scenarios, an attempt could be made towards mirroring the techniques used in fingerprint verification.

VIII. My Takeaways and Conclusion

Through this project, the scope of using forehead wrinkles as a biometric measure was explored. By varying the hyper-parameters for the model and the model itself, the similarity scores were obtained. I was able to understand:

- The methodology behind building a Machine Learning Model for various purposes.
- Usage of OpenCV library.
- The concept behind SIFT and ORB models.
- How data is collected for Machine Learning projects

Given the current scenario, with the increase of use of masks, I believe that wrinkles can definitely be explored for facial recognition and biometric purposes.

Acknowledgment

I would like to thank Kamlesh Tiwari Sir for giving me the opportunity to learn through a project which has real-world application and has helped in understanding Machine Learning concepts better. I also want to thank Geetika Arora Ma'am and Stuti Chug Ma'am for guiding me through the project and lab work.

References

[1] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[2] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *2011 International conference on computer vision*, pages 2564–2571. Ieee, 2011.

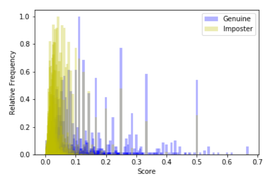
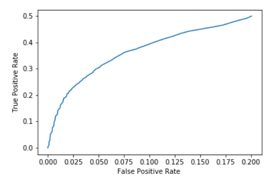
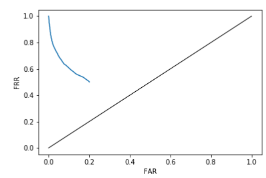
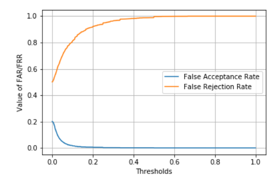
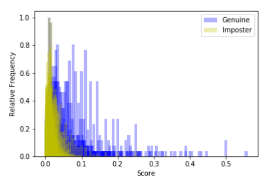
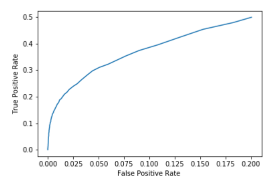
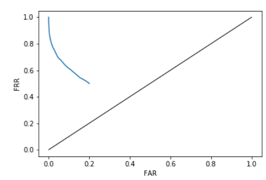
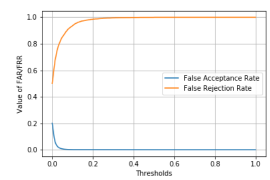
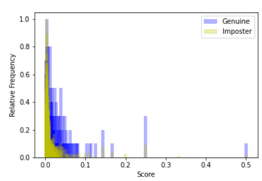
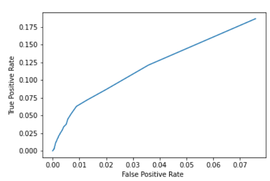
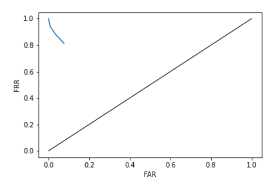
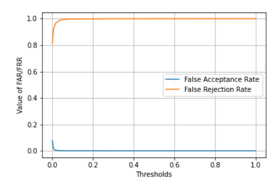
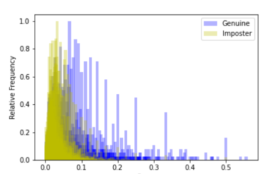
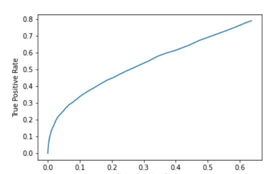
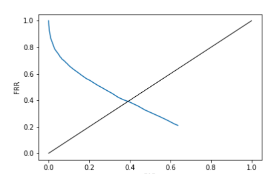
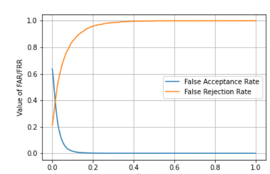
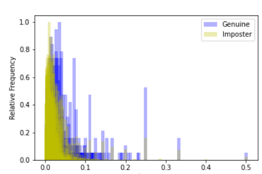
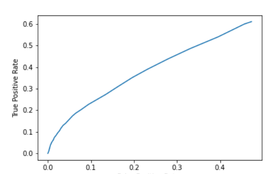
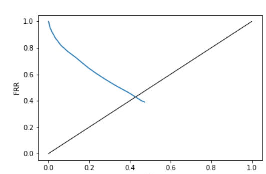
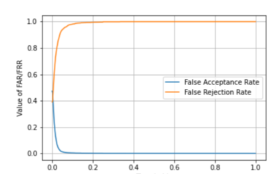
[3] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *European conference on computer vision*, pages 778–792. Springer, 2010.

[4] OpenCV. opencv.org.

TABLE I: Quantitative results of the different models

Run	Model Description	CRR	EER	Th @ EER	Accuracy %
1	SIFT- Used number_keypoints=min(len(kp1),len(kp2)) and m.distance <0.7*n.distance	0.394	0.200 | 0.501	0.001	65.15
2	SIFT- Used number_keypoints=max(len(kp1),len(kp2)) and m.distance <0.7*n.distance	0.744	0.200 | 0.500	0.001	65.05
3	ORB- Used number_keypoints=max(len(kp1),len(kp2)) and m.distance <0.7*n.distance	0.410	0.075 | 0.813	0.001	55.54
4	SIFT- Used number_keypoints=max(len(kp1),len(kp2)) and m.distance <0.8*n.distance	0.620	0.397 | 0.387	0.015	62.52
5	ORB- Used number_keypoints=max(len(kp1),len(kp2)) and m.distance <0.8*n.distance	0.402	0.456 | 0.400	0.003	57.92
6	SIFT- Used number_keypoints=max(len(kp1),len(kp2)) and m.distance <0.85*n.distance	0.557	0.407 | 0.418	0.039	60.31

TABLE II: Graphs obtained with the different models

Run No.	Genuine/Imposter Histogram	TP vs FP	ROC Curve	FAR/FRR vs Th
1				
2				
3				
4				
5				
6