

Article

The Effects of Facial Expressions on Face Biometric System's Reliability

Hind A. Alrubaish *  and Rachid Zagrouba

College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University,
P.O. Box 1982, Dammam 31441, Saudi Arabia; rmzagrouba@iau.edu.sa

* Correspondence: haalrubaish@iau.edu.sa

Received: 9 September 2020; Accepted: 13 October 2020; Published: 17 October 2020



Abstract: The human mood has a temporary effect on the face shape due to the movement of its muscles. Happiness, sadness, fear, anger, and other emotional conditions may affect the face biometric system's reliability. Most of the current studies on facial expressions are concerned about the accuracy of classifying the subjects based on their expressions. This study investigated the effect of facial expressions on the reliability of a face biometric system to find out which facial expression puts the biometric system at greater risk. Moreover, it identified a set of facial features that have the lowest facial deformation caused by facial expressions to be generalized during the recognition process, regardless of which facial expression is presented. In order to achieve the goal of this study, an analysis of 22 facial features between the normal face and six universal facial expressions is obtained. The results show that the face biometric systems are affected by facial expressions where the disgust expression achieved the most dissimilar score, while the sad expression achieved the lowest dissimilar score. Additionally, the study identified the five and top ten facial features that have the lowest facial deformations on the face shape in all facial expressions. Besides that, the relativity score showed less variances between the sample using the top facial features. The obtained results of this study minimized the false rejection rate in the face biometric system and subsequently the ability to raise the system's acceptance threshold to maximize the intrusion detection rate without affecting the user convenience.

Keywords: authentication; face biometric; facial expressions; human moods; false rejection

1. Introduction

Authentication is the mainline in the war to verify a user's identity and reject an illegitimate user from accessing their resources. Three types of authentication can distinguish any person among the population; one approach concerns the user's knowledge—such as a password, the second approach concerns what the user has—such as a national ID card, while the third approach is to define the user themselves using their humanistic traits—"biometrics". This type of authentication is considered the most robust compared to the other approaches as these features cannot be forgotten, shared, or stolen. Biometric authentication is the procedure of recognizing the users through their physiological and behavioral traits, such as fingerprints, iris, gait, keystrokes, face. Although facial biometrics (FBs) is one of the most potent biometric technologies, it is a challenging process. The face recognition process is more complicated than with other biometrics, such as fingerprint and iris identification, since the human face can be viewed from various angles with different poses.

Different factors can affect system reliability, such as illumination, occlusion, aging, facial surgery, and facial expressions. Facial expressions (FE) are means of expressing human feelings and reactions, which include many interconnecting elements of facial muscles movements [1]. Those expressions

result in facial feature shape changes [2]; if the user shows a different expression than the one stored in the database such as a neutral face, this will lead to a different matching result.

In biometric systems, two samples from the same person may give different matching scores due to different causes, such as: FE, lighting effect, and imaging conditions. These causes give the following errors [3–5]: type I error, where the system prevents the authorized person from accessing the resources as they cannot be identified; and type II error, where it gives unauthorized access to the system by identifying an unauthorized user as the authorized one. Those types of error are evaluated using the false rejection rate (FRR)—the measurement of the possibility that the system is willing to deny a genuine user, and the false acceptance rate (FAR)—the measurement of the possibility that the system is willing to accept an illegitimate user [5].

We cannot ignore FRR and the FAR when we need to assess the performance of the FB system's security. Both rates affect the system's security level and user convenience. Moreover, both have an impact on each other. As the FAR goes down, the FRR goes up and vice versa. The FAR is concerned about the system security while the FRR is concerned about user convenience; if we raise the system security, user convenience will be less. As a result, we have one of these two options: a more secure system which is less user-friendly, or a more user-friendly system which is less secure. Most of the entities prioritize user convenience over security. There is no “magic bullet” or “one size fits all” solution. Nevertheless, we can find something that can balance these two issues.

With the rapid increase of cybersecurity crimes day by day, FBs will become vital for authentication in everyday life. Many studies have been done, but even after continuous research, a truly robust and worthwhile outcome has not been achieved yet. Therefore, this study aims to analyze the FB system's reliability under the influence of different FEs to identify a set of facial features with the lowest deformations caused by FEs. These features can then be used during the recognition process, regardless of what expression is presented, to maintain the biometric system performance. The result of this analysis will help in minimizing the FRR in order to raise the acceptance threshold without affecting user convenience.

This paper presents a brief about the FE and the FB performance evaluation in Section 2, while Section 3 reviews the latest studies on the related fields. Section 4 explains the work's methodology; then, Section 5 discusses the results of the work. The findings of this study are listed in Section 6, where Section 7 concludes the work.

2. Background

In contrary to human recognition, automatic recognition is a challenging process. Moreover, face biometrics are more difficult than other biometrics (such as fingerprint and iris) due to the fact that the human face can be viewed from various angles with different expressions. Furthermore, different factors can affect the recognition process which can be summarized into *extrinsic*, such as pose variation and illumination, or *intrinsic*, such as aging and facial expression [6,7]. Facial expressions (FE) are a non-verbal communication role between peoples which can be included in a wide range of applications, e.g., human behavior, customer relationship management, social robots, and expression recognition [8]. Facial expressions of emotion can be categorized into: anger, disgust, fear, happiness, neutrality, sadness, and surprise. Expressions can change the face shape temporally because of the deformation of the face's muscles [9]. A facial biometric system's reliability can be affected by the subject's facial expressions; happiness and sadness and other facial emotions may lead to varying levels of facial identification accuracy and as a consequence have an effect on the system's reliability [10]. There are six basic emotions (BEs) that have been identified [11]: happiness, sadness, surprise, anger, disgust, fear.

The performance of a face biometric system can be evaluated by identifying FRR and FAR errors. To avoid the ambiguity caused by systems that allow multiple attempts or multiple templates and single attempts or single templates, there are two types of performance evaluation: decision error rate and matching error rate [5]:

- False rejection rate (FRR): The measurements of the possibility that the system is willing to deny an authorized transaction. This rate is calculated by the following:

$$FRR_{(\mu)} = \frac{\text{Number of false Rejection Attempts on the System}}{\text{Total Number of Authentication Attempts}} \quad (1)$$

- False acceptance rate (FAR): The measurement of the possibility that the system is willing to accept the unauthorized transactions. This rate is calculated by the following equation:

$$FAR_{(\mu)} = \frac{\text{Number of false Successful Attempts to Access the System}}{\text{Total Number of Authentication Attempts}} \quad (2)$$

3. Related Works

To the best of the author's knowledge, no study has investigated the effect of the facial expression on a face biometric system to find out which facial features have the most impact on the facial deformations in order to improve the system performance. Most of the related papers on facial expression were concerned about the accuracy of recognition and the classifying of samples based on their modes. However, some of the works have been done in different related areas.

3.1. Features Extraction and Facial Landmarks in Face Biometric

Facial landmarks (FL) and the extracted features are further aspects needed in the field of FB. Özseven and Düenci [12] compared FB's performance using distances and slopes between FL with statistical and classification methods. They used the BioID dataset that consists of 1521 pictures for 23 subjects. Their results showed that the best accuracy achieved by distances and slopes then by distances then by slopes. They used FGNet annotation that has 20 points, as shown in Figure 1. They have found that the following landmark points—2,3,4,5,6,7—were very influenced by FE, where they considered only the other 14 points in their analysis, as shown in Figure 2.



Figure 1. FGNet annotation [12].



Figure 2. FGNet annotation [12].

Amato et al. [13] compared between 5-points features and 68-points features, as shown in Figures 3 and 4. They conducted their experiments on videos taken in a real scenario by surveillance

cameras. They used dlib library and the FL detectors to implement the approach represented by [14], which returns an array of 68-points in the form of (x, y) coordinated. The results on the Wiled dataset showed that the 68 points had high mean average precision.

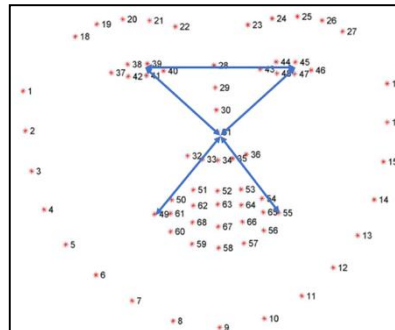


Figure 3. 5-point features [13].

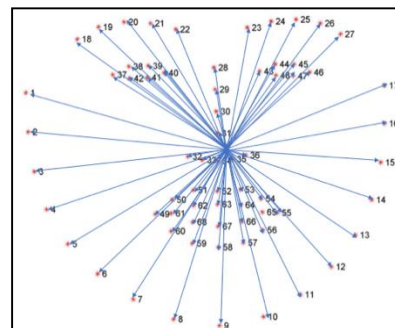


Figure 4. 68-points features [13].

Banerjee [15] measured the distance between the following FL, as shown in Figure 5.

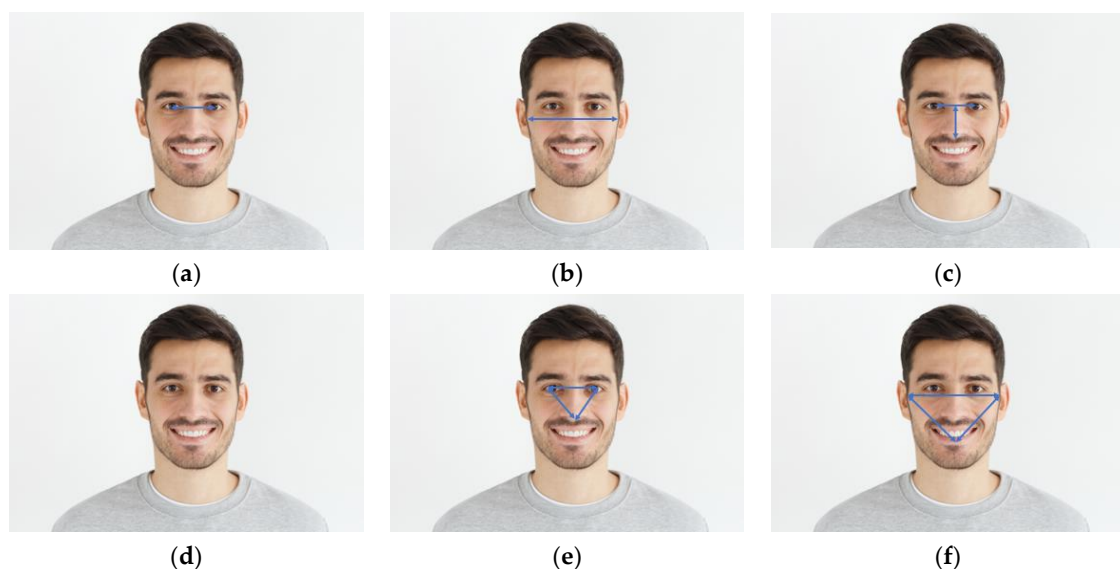


Figure 5. Selected features by Banerjee [15]. (a) distance between eyes, (b) distance between ears, (c) distance between the nose and forehead, (d) width of the leap, in addition to the following angles where the sum is 180°; (e) angles between eyes and nose, (f) angles between ears and mouth. The following were used to measure the distance between the face objects: Euclidian distance (EU), city block metric, Minkowski distance, Chebyshev distance and cosine distance.

Sabri et al. [16] develop a set of algorithms in a 3D face module where the captured face was segmented to obtain FLs, nose tip, mouth corners, left, and right eye corner, as shown in Figure 6. The algorithm computed two triangles; the first one between the eyes center and mouth center, where the second one between the eyes center and nose tip. Additionally, they measured the distance between the left and right eye corners and mouth corners, as shown in Figure 7.



Figure 6. Vector distances in [16].

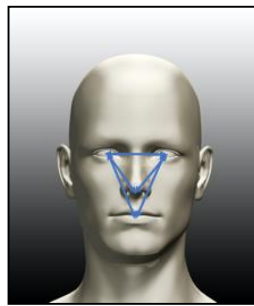


Figure 7. Triangles in [16].

Meanwhile, Napieralski et al. [17] used the Viola–Jones algorithm to detect three facial objects: eyes, nose, and mouth, where the midpoint for each region was calculated, as shown in Figure 8. They used the EU to measure the distance between both eyes, between lips and nose, nose width, lips height, and width.

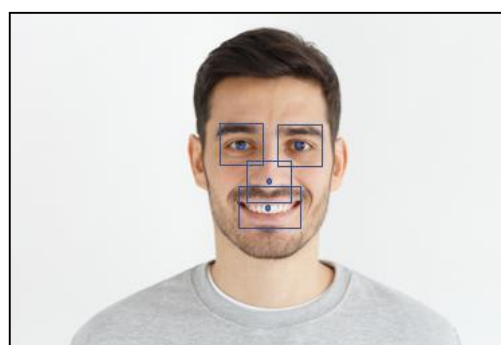


Figure 8. Selected features in [17].

Benedict and Kumar [18] designed geometric shaped facial extraction for face recognition to identify the subjects by finding the center and the corners of the eye using eye detection and eye localization. Then, 11 fiducial points have been derived from the given face: three points on the eye, the lateral extremes, the nose tip, midpoint of the lip, and two points on lateral extremes of lips as shown in Figure 9.



Figure 9. Selected Features in [18].

A study of Gurnani et al. [19] showed that the salient regions, eyes, nose, and mouth were the dominant features that help to classify the facial soft biometric; age, gender, and FE. According to Barroso et al. [20], they found out that the expression recognition performance using the whole face outperforms using some regions.

3.2. Facial Expression Recognition Applications

Some studies tried to improve the facial recognition procedures, such as; Teng et al. in [21] who proposed a 3D Convolutional Neural Networks (CNN) based architecture for FE recognition in videos. Mangala and Prajwala [22] used Eigenfaces and principle component analysis for the same purpose. Meanwhile, Ivanovsky et al. [23] used CNN to detect smiles and FE. Sun et al. [24] proposed FE recognition framework on discovering the region of interest to train the effective face specific of CNN. Yang et al. [25] utilized the facial action unit to recognize the expressions. Liang ji et al. [26] proposed deep learning enhanced gender conditional random forest for expressions in an uncontrolled environment to address the gender influence. Jeong et al. [27] proposed deep joint spatiotemporal features for facial expression recognition based on the deep appearance and geometric neural networks. Mehta et al. [28] recognized emotions based on its intensities while Jala and Tariq [8] aimed to get beyond classification and recognition known FE to cluster unknown facial behaviors.

In terms of the FE recognition applications, the deficits of FE in Huntington's diseases have been studied by Yitzhak et al. in [29] to improve the FE recognition using the predicting the severity of their motor symptoms. Mattavellit [30] studied it in Parkinson's diseases. Flynn et al. [31] assessed the effectiveness of automated emotion recognition in adults and children for the benefits of different applications, such as identification of children's emotions before clinical investigations.

In other aspects, FE can be used as a way of authentication; Delina et al. [32] tried to address the vulnerability of a single biometric authentication model by proposing the subject's physiological and behavioral traits' face. Their approach was to identify users by fusing the face shape and the FE to prove theirs legitimately. Additionally, Ming et al. [33] used FE as liveness detection in addition to the face verification.

3.3. The Effect of Facial Expression on Face Biometric Reliability

A few papers analyzed the performance of the biometric system under the effect of the subject's mode and his expressions, such as Pavael and Lordanescu [34], who analyzed the recognition performance by eyewitness, where their results indicated that happy and sad expressions influenced significantly the process of facial identity. Dalapicola et al. [35] took into their consideration the periocular trait and investigated the effect of FE on this region where it has been found that recognition using CNN was sensitive to the region deformation caused by FE. The experimental study has been done on the extended Cohn-Kanda (CK+) dataset that contains image sequences of 123 subjects where each subject has several samples between (1–11) and the number of frames varied from (4–71). Each sample represents an FE. Azimi [36] investigated whether the emotional faces have a statistically

significant effect on FB's matching score. The experiment was done using python dlib face recognition and Verilook on the Jaffe dataset that involved ten female users with seven different modes: neutrality, happiness, sadness, anger, disgust, fear, surprise. His results showed the following: (1) by comparing the neutral faces and FE, the average genuine similarity has been degraded; (2) sadness and disgust expressions are the most dissimilar expressions among the other expressions; (3) the best class to be verified with was the normal face, as there was no facial deformation and muscle movement; (4) For users who enrolled with happy, angry, surprised, disgusted, sad, and fearful expressions, the best expressions to verify with are fearful, sad, fearful, sad, fearful expressions, respectively; (5) the lowest matching score was achieved when users who provided happy, angry, surprised, disgusted, sad, fearful faces during the enrollment, identified themselves with disgusted, surprised, angry, happy, angry faces respectively during the verification. Another study investigated the effect of FE on FB systems by Márquez-Olivera et al. [37], where they analyzed their FB system under the influence of FE. It has been concluded that failures occurred when the subjects expressed surprise as it has maximum facial deformations while the sadness and anger expressions express high deformation on eye regions. On the other hand, the system performed better when the subject expresses happy expressions. Moreover, they also tried to overcome the effect of FE in the FB system by recognizing the people under their expressions; they proposed a hybrid model of Alpha-Beta Associative memories with correlation Matrix and K-Nearest Neighbors. Although the best face recognition accuracy under the influence of FE was 90% achieved by anger expressions, Khorsheed and Yurtkan [38] claimed that the Local Binary Pattern features form a strong base for face recognition under the influence of FE.

Different aspects have been studied by Azimi and Pacut [1] to investigate whether the effect of FE on the FB system was gender-dependent. Their results on the Stirling dataset showed that 13 females' faces showed more intense FE than ten males' faces using python face recognition and Verilook neurotechnology. This means that the similarity score of neutral faces vs. all FE for male subjects was better than female subjects; therefore, the influence of FE on FB system was gender-dependent.

This paper aims to study the impact of the FE on FB systems due to the lack of such studies in the field, as we can notice from the previous studies. For instance, in [15,16] investigated the utilization of the selected features and landmarks for face recognition purposes only. Although the accuracy was the highest when both slopes and distances were used in [12], this study will use distances only as it analyzes which muscles and facial features are affected by FE, not for recognition purposes [21–33], evaluated the performance of FE classifications. While utilizing the periocular as a biometric trait in [33] has its failures when the face presents posture changes, occlusions, closed eyes, and other changes, in the FB, the recognition process can use other features than the one that exposes failure. Meanwhile, in [35], the study used only ten females' subjects without males. This work aims to fill some gaps within the field, as illustrated in the next sections.

4. Methodology

Humans may express different expressions during daily life, where a robust FB system's performance should not be affected by those expressions and modes. The objective is to analyze the FB system's performance under the influence of different FE to identify facial features that have the lowest deformations caused by FE to be used only during the recognition regardless of what expression is presented. This study aims to achieve the goal by answering these questions; (1) Is the effect of the FE on the FB system significant? (2) Which FE has the best results? (3) Which FE has the worst results? (4) What is the impact of each FE on the similarity score? (5) Which facial features have the lowest facial deformation that can be generalized during the recognition and cannot be affected significantly by the expressed emotion? (6) What is the FRR performance under the influence of FE?

To answer those questions, we used the IMPA-FACES3D dataset [39] to obtain the distances and position for 22 facial features. After that, we determined the relativity shift score (RSS) of different facial features and total similarity score (SS) between the neutral face and six universal expressions for

each subject. Based on the analysis of gained data, we identified a set of facial features with the lowest facial deformations that score higher SS. This section illustrates the methodology in detail.

4.1. Dataset Description

IMPA-FACES3D dataset [39] includes acquisitions of 38 male and female subjects with 12 distinct poses. This study uses neutral mode and the six universal expressions: neutrality, joy, sadness, surprise, anger, disgust, fear. Figure 10 shows an example of those expressions for subject # one. This set is composed of 22 males and 16 females with ages between 20 and 50 years (we used only 36 subjects as there are two subjects are missing (22 male and 14 female)).



Figure 10. Subject 1 in IMPA-FACES3D [37] shows the following expressions: (a) neutral, (b) happy, (c) sadness, (d) surprise, (e) anger, (f) disgust, (g) fear [40].

To achieve the objectives, we have developed an in-house python script built upon OpenCV and dlib that works, as explained in the next sections.

4.2. Face Detection and Acquisition

After uploading the two faces, it will be converted into a greyscale image with a single layer of 8-bit pixels (value ranges between 0–255). The grayscale image was faded into dlib to identify 68 key points FL. FL are key points of the detected face's shape to make up the facial features. Facial features that could be compared with other facial features were made using the distance between and FL. This study uses 68 points templates for FLs, as shown in Figure 11 and explained in Table 1.

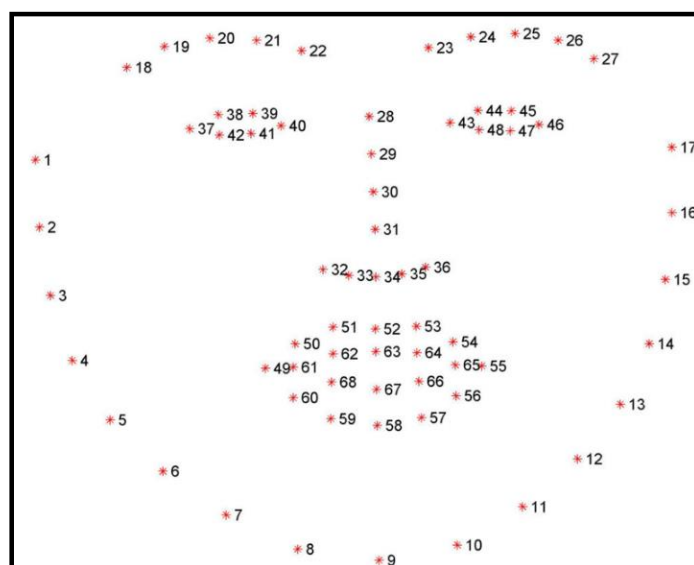


Figure 11. Template image for face's landmark detection using 68-points for a frontal view.

Table 1. The points range for each face feature in 68-points face's landmarks.

Facial Features	Points Range
Chin	1–17
Right Eyebrow	18–22
Left Eyebrow	23–27
Nose	28–36
Left eye	37–42
Right eye	43–48
Mouth	49–68

4.3. Preprocessing

After that, we use static (silent) features to adjust the size of the uploaded faces according to the standard size using points 1 and 17 in Figure 11. Moreover, to align the faces, we kept the angle of the line joining the midpoint between two eyes to zero degree as shown in Figure 12, where the blue line should be aligned with the red line.

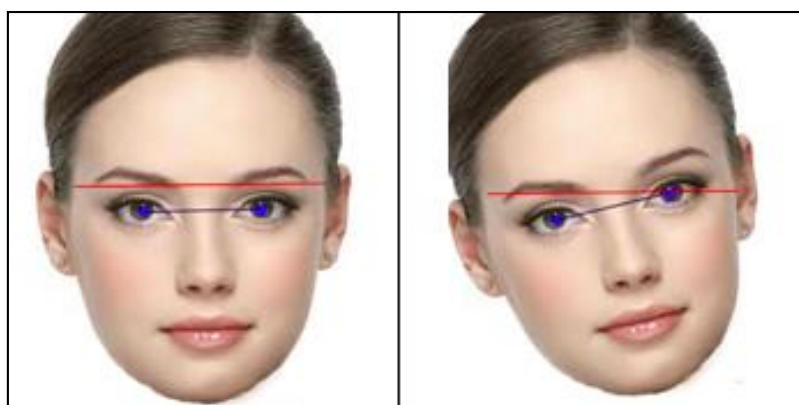


Figure 12. Face alignment.

4.4. Features Extraction and Verification

We have identified 22 facial features to be analyzed as shown in Table 2 and Figure 13. Table 2 explains the facial features and the corresponding points in 68 landmark points, while Figure 13 illustrates the features on the subject's face.

After that, a comparison between neutral mode and other expressions for the same subject is conducted to obtain 22 facial features, RSS and SS. Neutral mode and other expressions were compared to each other to explore the effect of FE on the genuine score. Assuming the provided template in the enrollment session is the neutral mode—as it is the most common mode in our daily routine—the comparison conducted for the subjects were as follows: neutral mode vs. happy expression, neutral mode vs. sad expression, neutral mode vs. surprise expression, neutral mode vs. anger expression, neutral mode vs. disgust expression, neutral mode vs. fear expression. In each comparison, the 68 FL points were obtained for each image (expression) to create a list of face's organs as in Table 1 to help in obtaining the facial features showed in Table 2.

Table 2. Facial features and the corresponding points in 68 landmarks template.

#	Facial Features	Points
1	Left eye width	Distance between 37 and 40
2	Right eye width	Distance between 43 and 46
3	Left eye position	The coordinates of the left eye middle
4	Right eye position	The coordinates of the right eye middle
5	Mouth width	Distance between 49 and 55
6	Mouth position	Coordinates of the intersect of the distance between 49 and 55 and the distance between 52 and 58
7	Nose width	Distance between 32 and 36
8	Nose position	Coordinates of intersect of the distance between 31 and 34 and the distance between 32 and 36
9	Chin width	Distance between 7 and 11
10	Chin position	Coordinates of Point 9
11	Forehead width	Distance between 1 and 17
12	Forehead position	Middle point of the upper line joining 20 and 25
13	Distance between eyes	Distance between middle eyes points
14	Distance between left eye and nose	Distance between 37 and 34
15	Distance between right eye and nose	Distance between 46 and 34
16	Distance between left eye and mouth	Distance between 37 and 49
17	Distance between right eye and mouth	Distance between 46 and 55
18	Distance between left eye and eyebrow	Distance between eye middle point and 20
19	Distance between right eye and eyebrow	Distance between eye middle point and 25
20	Distance between nose and forehead	Coordinates of intersect of the distance between 34 and the middle point of the upper line joining 20 and 25
21	Distance between left ear and mouth	Distance between 3 and 49
22	Distance between right ear and mouth	Distance between 15 and 55

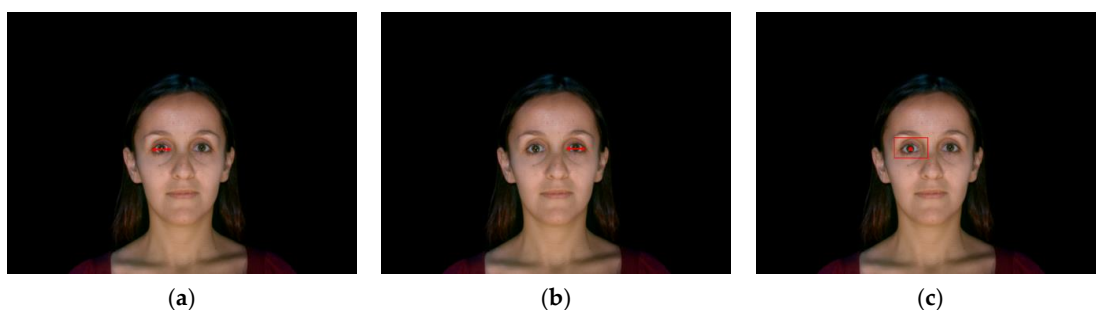


Figure 13. Cont.



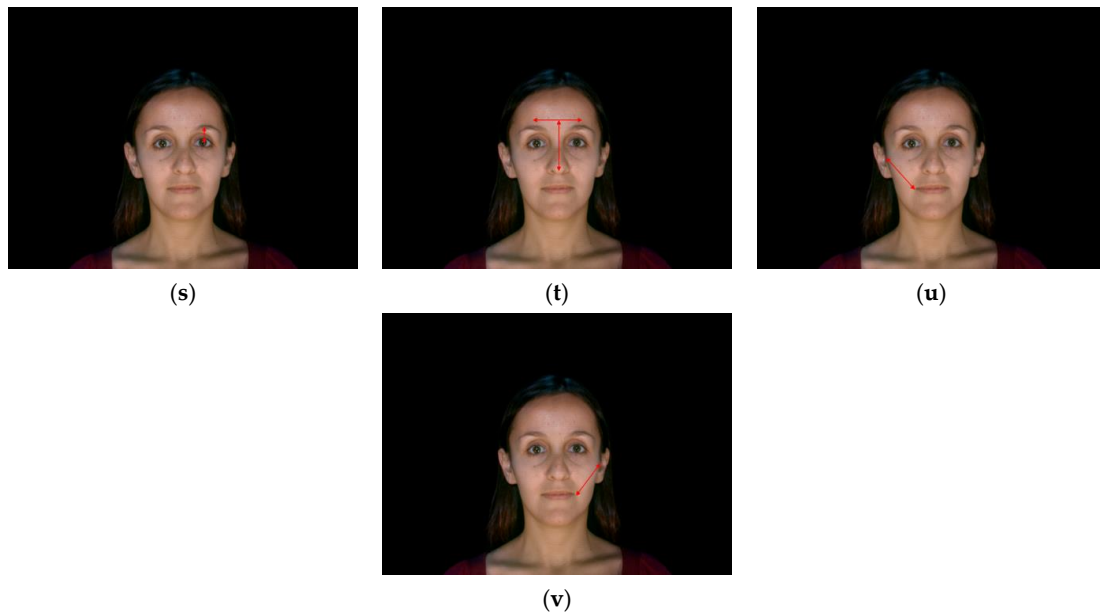


Figure 13. Illustration of the 22 facial features: (a) left eye width; (b) right eye width; (c) left eye position; (d) right eye position; (e) mouth width; (f) mouth position; (g) nose width; (h) nose position; (i) chin width; (j) chin position; (k) forehead width; (l) forehead position; (m) distance between eyes; (n) distance between left eye and nose; (o) distance between right eye and nose; (p) distance between left eye and mouth; (q) distance between right eye and mouth; (r) distance between left eye and eyebrow; (s) distance between right eye and eyebrow; (t) distance between nose and forehead; (u) distance between left ear and mouth; (v) distance between right ear and mouth.

In order to obtain the 22 facial features as described in Table 2 and Figure 13, we used Euclidian distance in Equation (3) to measure the straight-line distance between two FLs where x_1 and x_2 are the coordinate of the first landmark and $y_1 - y_2$ are the coordinates of the second landmark.

$$EU = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3)$$

Up to this point, the values for each facial feature in the neutral mode and the expression in each comparison have been recorded. After that we considered the known measurement error rate [41] and defined it in Equation (4) as follows:

$$ER = \left| \frac{(Accepted\ Value - Experimental\ Value)}{Accepted\ Value} \right| \quad (4)$$

Furthermore, we adapted it to our problem, and called it the “relativity shift score” (RSS) where the value of EU of the expression’s feature in Equation (5) is corresponding to the experimental value in Equation (4); while the EU of the normal mode’s feature corresponding in Equation (5) is corresponding to the accepted value in Equation (4) and defined in Equation (5) as:

$$RSS = \left| \frac{(EU_Feature_Expression - EU_Feature_Normal)}{EU_Feature_Normal} \right| \quad (5)$$

This will measure how two faces are relative to each other in terms of particle facial features for each comparison in a range of (0,1), where 0 means the features were identical (facial feature stayed unchanged, and the expression did not change the feature) while 1 means the features were unidentical, hence lower value means higher similarity.

Next, we summed up all the relativities and divided them by their number to introduce the “similarity score” (SS) measure in Equation (6). Additionally, the results would be in a range of (0,1) where 0 means that (all features) stayed unchanged. We added (1-) to Equation (6) to be “1” as the best case, meaning that two faces are 100% “similar”, i.e., the faces were the same in terms of their all facial features. Thus, to make similarity considering the best value of “1”, we introduce the following formula:

$$SS = \left| \frac{1 - \text{Sum (Relativity for All Features)}}{\text{Number of Features}} \right| \quad (6)$$

After obtaining all the relativities and similarities for 36 subjects in 6 FE, we conducted a statistical analysis to calculate the means $\pm$ SD of RSS and SS for 36 subjects for each expression and all expressions. Based on the results, we ranked the facial features that scored the best RSS, and accordingly, we selected the best five features, best ten features, and the worst ten features, and we compared them in terms of their SS for each FE and all expressions.

Finally, we evaluated the performance using FRR Equation (1) at three acceptance thresholds: 99%, 95%, 90%.

5. Results and Discussion

The below sections described and discussed the achieved results between neutral face and six FE as follows.

5.1. Happy Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face’s features have the best and worst score in terms of the relativity shift score for the happy expression compared with neutral mode.

After comparing the neutral mode with a happy expression for 36 subjects, 22 facial features have been obtained for the neutral mode and the happy expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_Neutral_Happy = \left| \frac{(EU_Feature_Happy - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (7)$$

After that, we ranked the happy facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 3 and Figure 14 showed that the top five features in terms of the RSS were: right eye position, chin position, mouth position, nose position, and left eye position. While of the best ten features, the next best five were: forehead position, forehead width, distance between eyes, chin width, and left eye width. Additionally, we identified the worst ten features as follows: mouth width, distance between left ear and mouth, distance between right ear and mouth, nose width, distance between left eye and mouth, distance between right eye and mouth, distance between right eye and eyebrow, distance between left eye and eyebrow, distance between nose and forehead, and distance between left eye and nose.

Table 3. The relativity shift score (RSS) means between happy expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.0533	10	0.0350
Right eye width	36	0.0592	11	0.0389
Left eye position	36	0.0286	5	0.0233
Right eye position	36	0.0215	1	0.0182
Mouth width	36	0.2447	22	0.1123
Mouth position	36	0.0248	3	0.0166
Nose width	36	0.1483	19	0.0731

Table 3. Cont.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Nose position	36	0.0271	4	0.0186
Chin width	36	0.0416	9	0.0363
Chin position	36	0.0215	2	0.0155
Forehead width	36	0.0360	7	0.0294
Forehead position	36	0.0300	6	0.0263
Distance between eyes	36	0.0387	8	0.0285
Distance between left eye and nose	36	0.0826	13	0.0716
Distance between right eye and nose	36	0.0730	12	0.0651
Distance between left eye and mouth	36	0.1331	18	0.0737
Distance between right eye and mouth	36	0.1293	17	0.0771
Distance between left eye and eyebrow	36	0.0931	15	0.1045
Distance between right eye and eyebrow	36	0.0946	16	0.0981
Distance between nose and forehead	36	0.0855	14	0.0617
Distance between left ear and mouth	36	0.1866	21	0.0816
Distance between right ear and mouth	36	0.1608	20	0.0924
Average		0.0825		

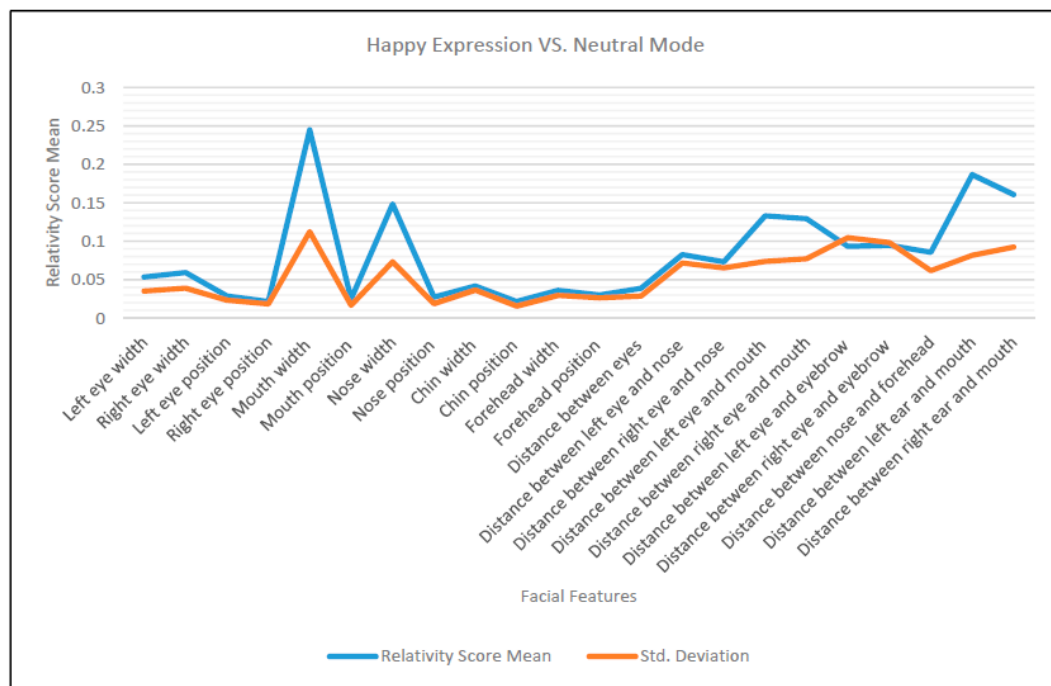


Figure 14. The means of RSS between happy expression and neutral mode of facial features for 36 subjects.

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

$$SS_Happy = \left| \frac{1 - \text{Sum (Relativity score between neutral and happy expression for the selected facial features)}}{\text{number of selected features}} \right| \quad (8)$$

The overall SS for all facial features was 91.7539%, while the SS for the top five was 97.5262 %, the SS for the top ten was 96.7665%, and finally, the SS for the worst ten was 86.4146%.

The SS has been increased by 5.77%, 5.01% after we selected the top five and top ten features, respectively. Additionally, it is noticed from Table 3 and Figure 14 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample are tending to be very close to the mean, while the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

5.2. Sad Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face's features have the best and worst score in terms of the relativity shift score for the sad expression compared with neutral mode.

After comparing the neutral mode with the sad expression for 36 subjects, the values of 22 facial features have been obtained for the neutral mode and the sad expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_Neutral_Sad = \left| \frac{(EU_Feature_Sad - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (9)$$

Then, we ranked the sad facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 4 and Figure 15 show that the top five features in terms of the RSS were: right eye position, chin position, forehead width, forehead position, and mouth position. While of the best ten features, the next best five were: left eye position, nose position, distance between eyes, distance between left eye and mouth, and chin width. Additionally, we identified the worst ten features as follows: right eye width, distance between left eye and nose, mouth width, distance between right eye and nose, distance between nose and forehead, distance between left ear and mouth, distance between right ear and mouth, distance between left eye and eyebrow, distance between right eye and eyebrow, and nose width.

Table 4. The RSS means between sad expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.0568	12	0.0500
Right eye width	36	0.0599	13	0.0595
Left eye position	36	0.0311	6	0.0194
Right eye position	36	0.0264	1	0.0162
Mouth width	36	0.0710	15	0.0730
Mouth position	36	0.0296	5	0.0187
Nose width	36	0.1226	22	0.0875
Nose position	36	0.0327	7	0.0192
Chin width	36	0.0520	10	0.0363
Chin position	36	0.0268	2	0.0178
Forehead width	36	0.0272	3	0.0273
Forehead position	36	0.0284	4	0.0221
Distance between eyes	36	0.0368	8	0.0329
Distance between left eye and nose	36	0.0641	14	0.0556
Distance between right eye and nose	36	0.0816	16	0.0720
Distance between left eye and mouth	36	0.0511	9	0.0416
Distance between right eye and mouth	36	0.0542	11	0.0466
Distance between left eye and eyebrow	36	0.1065	20	0.1052
Distance between right eye and eyebrow	36	0.1094	21	0.1048
Distance between nose and forehead	36	0.0829	17	0.0772
Distance between left ear and mouth	36	0.0832	18	0.0701
Distance between right ear and mouth	36	0.0970	19	0.0833
		0.0605		

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

$$SS_{Sad} = \left| \frac{1 - \text{Sum (Relativity between neutral and sad for the selected facial features)}}{\text{number of selected features}} \right| \quad (10)$$

The SS for all facial features was; 93.9482% while the SS for the top five was 97.2315%; as for the SS for the top ten was 96.4777%. Finally, the SS for the worst ten was 91.2179%.

The SS has been increased by 3.2833%, 2.5295% after we selected only the top five and top ten features. Additionally, it is noticed from Table 4 and Figure 15 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample are tending to be close to the mean, while the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

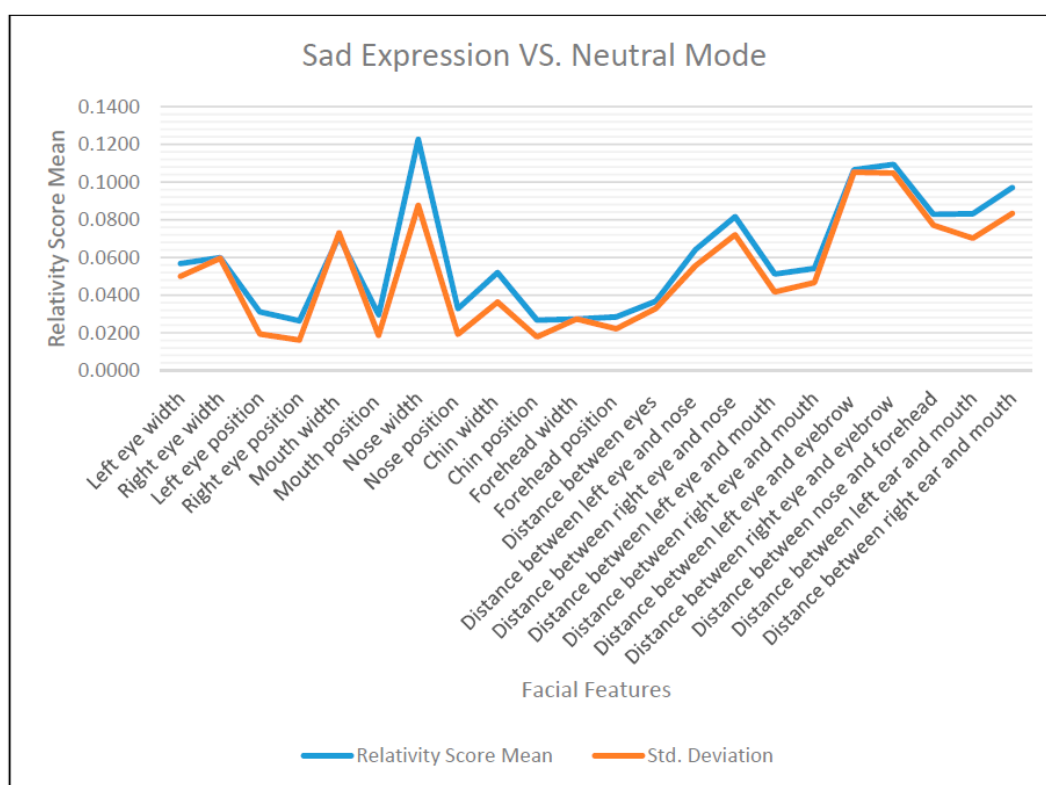


Figure 15. The means of RSS between sad expression and neutral mode of facial features for 36 subjects.

5.3. Surprise Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face's features have the best and worst score in term of the relativity shift score for the surprise expression in comparison with neutral mode.

After comparing the neutral mode with surprise expression for 36 subjects, the values of 22 facial features have been obtained for the neutral mode and the surprise expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_{Neutral_Surprise} = \left| \frac{(EU_Feature_Surprise - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (11)$$

Then, we ranked the surprise facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 5 and Figure 16 show that the top five features in terms of the RSS were: right eye position, left eye position, mouth position, forehead position, and nose position. While of the best ten features, the next best five were: chin position, forehead width, distance between eyes, chin width, and left eye width. Additionally, we identified the worst ten features as follows: distance between left eye and mouth, distance between right eye and nose, distance between right eye and mouth, mouth width, distance between nose and forehead, distance between right ear and mouth, distance between left ear and mouth, nose width, distance between right eye and eyebrow, and distance between left eye and eyebrow.

Table 5. The RSS means between surprise expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.0563	10	0.0446
Right eye width	36	0.0639	11	0.0644
Left eye position	36	0.0276	2	0.0312
Right eye position	36	0.0265	1	0.0264
Mouth width	36	0.0858	16	0.0685
Mouth position	36	0.0288	3	0.0215
Nose width	36	0.1423	20	0.1153
Nose position	36	0.0302	5	0.0309
Chin width	36	0.0452	9	0.0364
Chin position	36	0.0328	6	0.0233
Forehead width	36	0.0357	7	0.0263
Forehead position	36	0.0294	4	0.0338
Distance between eyes	36	0.0414	8	0.0304
Distance between left eye and nose	36	0.0643	12	0.0553
Distance between right eye and nose	36	0.0801	14	0.0745
Distance between left eye and mouth	36	0.0693	13	0.0582
Distance between right eye and mouth	36	0.0814	15	0.0602
Distance between left eye and eyebrow	36	0.1927	22	0.1212
Distance between right eye and eyebrow	36	0.1716	21	0.1173
Distance between nose and forehead	36	0.0931	17	0.0662
Distance between left ear and mouth	36	0.1120	19	0.0941
Distance between right ear and mouth	36	0.1033	18	0.0844
Average		0.0734		

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

$$SS_{Surprise} = \left| \frac{1 - \text{Sum (Relativity between neutral and surprise for the selected facial features)}}{\text{number of selected features}} \right| \quad (12)$$

The SS for all facial features was 92.6649%, while the SS for the top five was 97.1492%; the SS for the top ten was 96.4605%, and finally, the SS for the worst ten was 88.6843%.

The SS has been increased by 4.4843%, 3.7956% after we selected only the top five and top ten features. Additionally, it is noticed from Table 5 and Figure 16 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample is tending to be close to the mean while the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

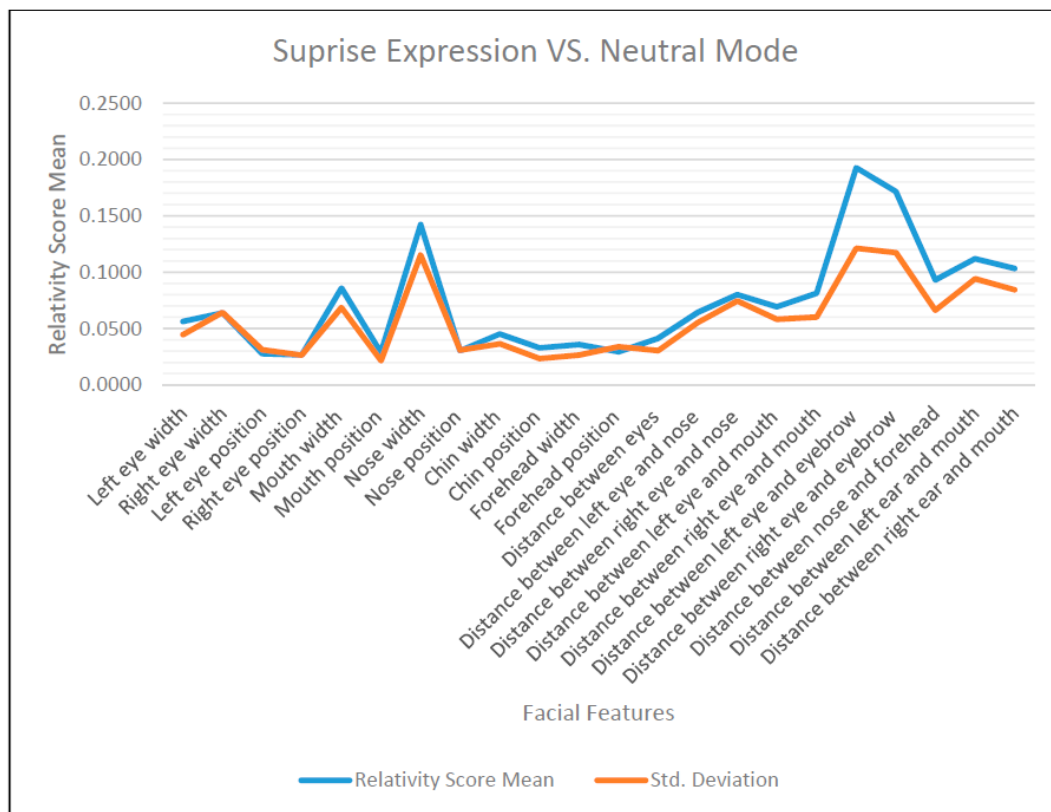


Figure 16. The means of RSS between surprise expression and neutral mode of facial features for 36 subjects.

5.4. Anger Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face's features have the best and worst score in term of the relativity shift score for the anger expression in comparison with neutral mode.

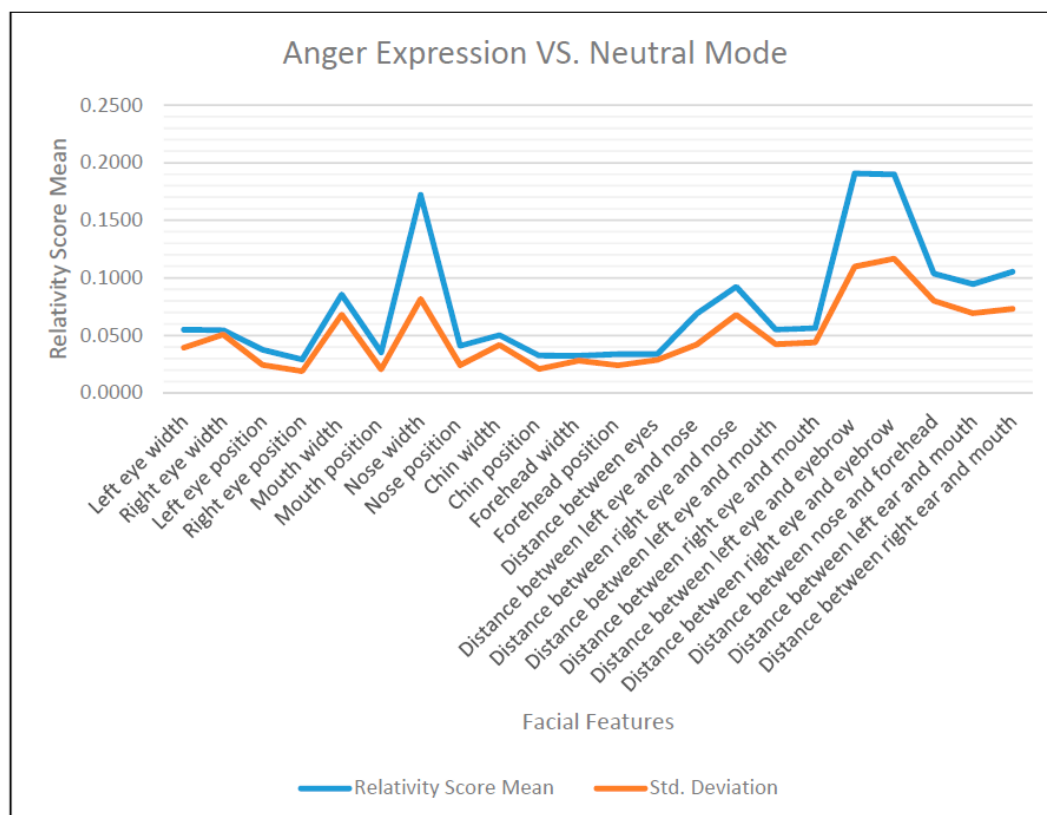
After comparing the neutral mode with anger expression for 36 subjects, the values of 22 facial features have been obtained for the neutral mode and the anger expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_Neutral_Anger = \left| \frac{(EU_Feature_Anger - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (13)$$

Then, we ranked the surprise facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 6 and Figure 17 show that the top five features in terms of the RSS were: right eye position, forehead width, chin position, forehead position, and distance between eyes. While of the best ten features the next best five were: mouth position, left eye position, nose position, chin width, and right eye width. Additionally, we identified the worst ten features as follows: distance between right eye and mouth, distance between left eye and nose, mouth width, distance between right eye and nose, distance between left ear and mouth, distance between nose and forehead, distance between right ear and mouth, nose width, distance between right eye and eyebrow, and distance between left eye and eyebrow.

Table 6. The RSS means between anger expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.0550	11	0.0394
Right eye width	36	0.0546	10	0.0508
Left eye position	36	0.0375	7	0.0243
Right eye position	36	0.0292	1	0.0189
Mouth width	36	0.0856	15	0.0681
Mouth position	36	0.0354	6	0.0205
Nose width	36	0.1722	20	0.0817
Nose position	36	0.0410	8	0.0242
Chin width	36	0.0502	9	0.0417
Chin position	36	0.0326	3	0.0209
Forehead width	36	0.0324	2	0.0282
Forehead position	36	0.0338	4	0.0240
Distance between eyes	36	0.0338	5	0.0288
Distance between left eye and nose	36	0.0691	14	0.0422
Distance between right eye and nose	36	0.0924	16	0.0679
Distance between left eye and mouth	36	0.0551	12	0.0424
Distance between right eye and mouth	36	0.0563	13	0.0439
Distance between left eye and eyebrow	36	0.1906	22	0.1098
Distance between right eye and eyebrow	36	0.1899	21	0.1167
Distance between nose and forehead	36	0.1038	18	0.0801
Distance between left ear and mouth	36	0.0945	17	0.0693
Distance between right ear and mouth	36	0.1053	19	0.0731
Average		0.0750		

**Figure 17.** The means of RSS between anger expression and neutral mode of facial features for 36 subjects.

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

$$SS_Anger = \left| \frac{1 - \text{Sum (Relativity between neutral and anger for the selected facial features)}}{\text{number of selected features}} \right| \quad (14)$$

The SS for all facial features was 92.5003%, while the SS for the top five was 96.7651%; the SS for the top ten was 96.1960%, and finally, the SS for the worst ten was 88.4048%.

The SS has been increased by 4.2648%, 3.6957% after we selected only the top five features and top ten features. Additionally, it is noticed from Table 6 and Figure 17 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample are tending to be close to the mean, while the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

5.5. Disgust Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face's features have the best and worst score in term of the relativity shift score for the disgust expression in comparison with neutral mode.

After comparing the neutral mode with disgust expression for 36 subjects, the values of 22 facial features have been obtained for the neutral mode and the disgust expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_Neutral_Disgust = \left| \frac{(EU_Feature_disgust - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (15)$$

Then, we ranked the disgust facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 7 and Figure 18 show that the top five features in terms of the relativity score were: chin position, right eye position, mouth position, forehead width, and left eye position. While of the best ten features, the next best five were: nose position, forehead position, distance between eyes, chin width, and left eye width. Additionally, we identified the worst ten features as follows: distance between left eye and nose, distance between right eye and mouth, distance between left eye and mouth, distance between nose and forehead, mouth width, distance between right ear and mouth, distance between left eye and eyebrow, distance between right eye and eyebrow, and nose width.

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

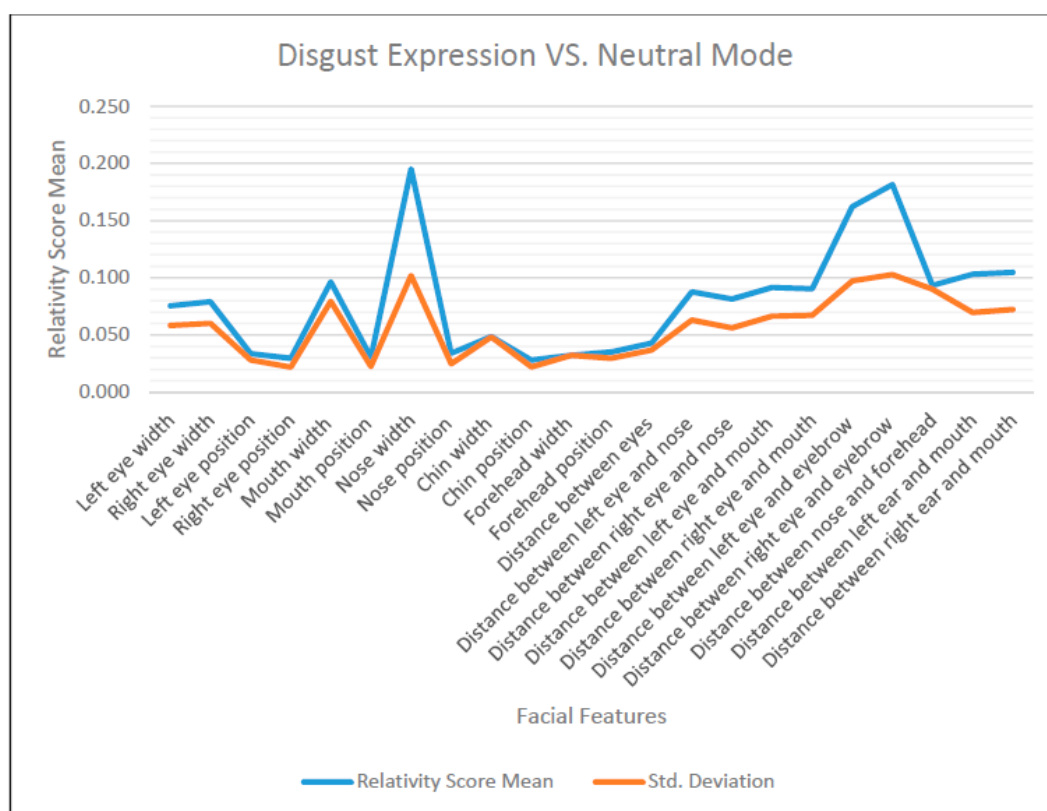
$$SS_Disgust = \left| \frac{1 - \text{Sum (Relativity between neutral and Disgust for the selected facial features)}}{\text{number of selected features}} \right| \quad (16)$$

The SS for all facial features was 92.01283%, while the SS for the top five was 96.9166%; the SS for the top ten was 96.0945%, and finally, the SS for the worst ten was 87.9381%.

The SS has been increased by 4.9037%, 3.0816% after we selected only the top five and top ten features. Additionally, it is noticed from Table 7 and Figure 18 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample is tending to be close to the meanwhile the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

Table 7. The RSS means between disgust expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.076	10	0.0584
Right eye width	36	0.079	11	0.0602
Left eye position	36	0.034	5	0.0282
Right eye position	36	0.030	2	0.0218
Mouth width	36	0.096	17	0.0794
Mouth position	36	0.030	3	0.0225
Nose width	36	0.195	22	0.1019
Nose position	36	0.034	6	0.0248
Chin width	36	0.049	9	0.0482
Chin position	36	0.028	1	0.0221
Forehead width	36	0.032	4	0.0323
Forehead position	36	0.035	7	0.0296
Distance between eyes	36	0.043	8	0.0369
Distance between left eye and nose	36	0.088	13	0.0630
Distance between right eye and nose	36	0.081	12	0.0562
Distance between left eye and mouth	36	0.092	15	0.0665
Distance between right eye and mouth	36	0.090	14	0.0672
Distance between left eye and eyebrow	36	0.162	20	0.0973
Distance between right eye and eyebrow	36	0.182	21	0.1029
Distance between nose and forehead	36	0.093	16	0.0901
Distance between left ear and mouth	36	0.103	18	0.0697
Distance between right ear and mouth	36	0.105	19	0.0723
Average		0.080		

**Figure 18.** The means of RSS between disgust expression and neutral mode of facial features for 36 subjects.

5.6. Fear Expression vs. Neutral Mode

The following analysis shows how, and which set of measured face's features have the best and worst score in term of the relativity shift score for the fear expression in comparison with neutral mode.

After comparing the neutral mode with the fear expression for 36 subjects, the values of 22 facial features have been obtained for the neutral mode and the fear expression. Then we applied Equation (5) to get the results of RSS for each facial feature as follows:

$$RSS_Neutral_Fear = \left| \frac{(EU_Feature_fear - EU_Feature_Neutral)}{EU_Feature_Neutral} \right| \quad (17)$$

Then, we ranked the fear facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results in Table 8 and Figure 19 show that the top five features in terms of the relativity score were: right eye position, chin position, mouth position, nose position, and left eye position. While of the best ten features, the next best five were: forehead position, forehead width, distance between eyes, chin width, and distance between left eye and nose. Additionally, we identified the worst ten features as follows: distance between right eye and mouth, right eye width, mouth width, distance between right eye and nose, distance between right ear and mouth, distance between nose and forehead, distance between left ear and mouth, nose width, distance between right eye and eyebrow, and distance between left eye and eyebrow.

Table 8. The RSS means between fear expression and neutral mode of 22 facial features for 36 subjects.

Facial Features	N	RSS Mean	Rank	Std. Deviation
Left eye width	36	0.0604	11	0.0393
Right eye width	36	0.0742	14	0.0494
Left eye position	36	0.0321	5	0.0258
Right eye position	36	0.0293	1	0.0210
Mouth width	36	0.0837	15	0.0658
Mouth position	36	0.0301	3	0.0185
Nose width	36	0.1053	20	0.0707
Nose position	36	0.0318	4	0.0227
Chin width	36	0.0585	9	0.0312
Chin position	36	0.0295	2	0.0188
Forehead width	36	0.0389	7	0.0318
Forehead position	36	0.0352	6	0.0287
Distance between eyes	36	0.0446	8	0.0411
Distance between left eye and nose	36	0.0590	10	0.0504
Distance between right eye and nose	36	0.0865	16	0.0607
Distance between left eye and mouth	36	0.0660	12	0.0439
Distance between right eye and mouth	36	0.0667	13	0.0471
Distance between left eye and eyebrow	36	0.1285	22	0.1092
Distance between right eye and eyebrow	36	0.1250	21	0.1151
Distance between nose and forehead	36	0.0931	18	0.0618
Distance between left ear and mouth	36	0.0933	19	0.0687
Distance between right ear and mouth	36	0.0928	17	0.0729
Average		0.0666		

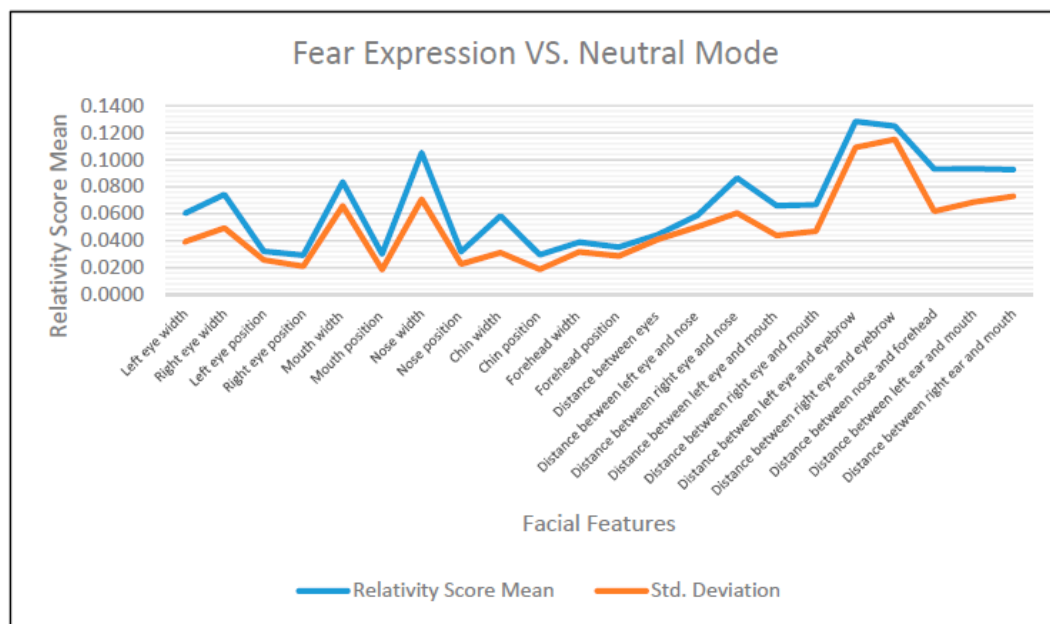


Figure 19. The means RSS between fear expression and neutral mode of 22 facial features for 36 subjects.

Next, we applied Equation (6) to obtain the SS for the 22 facial features, top five features, top ten features, and the worst ten features as follows:

$$SS_{Fear} = \left| \frac{1 - \text{Sum (Relativity between neutral and fear for the selected facial features)}}{\text{number of selected features}} \right| \quad (18)$$

The SS for all facial features was 93.3441%, while the SS for the top five was 96.9428%; the SS for the top ten was 96.1102%, and finally, the SS for the worst ten was 90.5102%.

The SS has been increased by 3.5987%, 2.7661% after we selected only the top five and to ten features. Additionally, it is noticed from Table 8 and Figure 19 that the top five and top ten features have the least standard deviations values which mean that the RSS for the sample are tending to be close to the mean, while the RSS for the worst ten features indicates that they have deviated from the mean as they have a higher standard deviation.

As a summary, Table 9 and Figure 20 show the SS for all six expressions in the following situations; SS for all 22 facial features, top 10, top 5, worst ten. It could be observed that the top five facial features have achieved the best SS.

Table 9. The SS for all six-expression using all features, top five, top ten, worst ten.

Similarity	Happy	Sad	Surprise	Anger	Disgust	Fear
Similarity for all 22 features	91.754%	93.948%	92.665%	92.500%	92.013%	93.344%
Similarity for top five features	97.526%	97.232%	97.149%	96.765%	96.917%	96.943%
Similarity for top ten features	96.767%	96.578%	96.461%	96.196%	96.094%	96.110%
Similarity for worst ten features	86.415%	91.218%	88.684%	88.405%	87.938%	90.510%

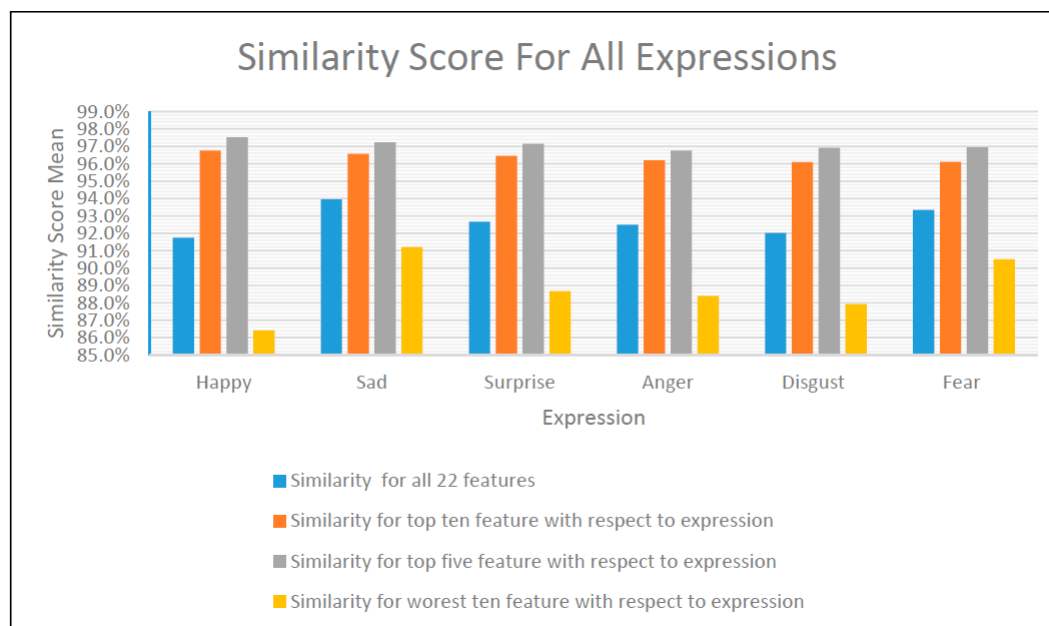


Figure 20. The similarity score (SS) for all six-expression using all features, top five, top ten, worst ten.

5.7. All Expressions vs. Neutral Mode

After obtaining results of the RSS and SS between neutral mode and each expression for 36 subjects, the results showed that the following FEs; happy, sadness, surprise, anger, disgust, and fear showed a lower SS compared to neutral mode. The highest dissimilar FE was achieved by the disgust expression of 92.01%, which was also agreed with results of Azimi [36] and Márquez-Olivera [37]. While the lowest one achieved by sad expression 93.94%, which is contrary to what has been reported in [37] where it was the happy expression. The following expressions: fear, surprise, anger, happy achieved 93.34%, 92.66%, 92.50%, 92.40% respectively, as shown in Table 10 and Figure 21.

Table 10. The mean of the SS for all expressions.

Mode	N	Mean
Happy	36	92.4066%
Sadness	36	93.9482%
Surprise	36	92.6649%
Anger	36	92.5002%
Disgust	36	92.0127%
Fear	36	93.3440%
Average		92.8128%

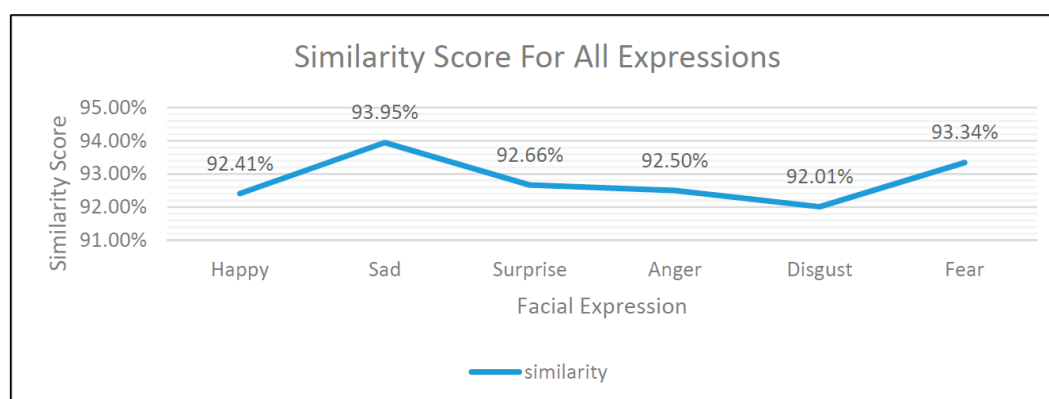


Figure 21. SS means plot of the 6 FE of 36 subjects in comparison to the neutral mode.

Hence, we assumed that there are facial features that cause the changes between neutral's SS and FE's SS and lead to a lower score. To find out which features in order to improve the overall SS, we calculated the mean of the RSS for each facial feature with respect to all expressions for 36 subjects. We then ranked the facial features based on the RSS, where the lowest value is the best in terms of the similarity. The results showed that the top five features in terms of the RSS were: right eye position, chin position, mouth position, left eye position, and forehead position. While of the best ten features, the next best five were: nose position, forehead width, distance between eyes, chin width, and left eye width. Finally, the worst ten features were: distance between left eye and mouth, distance between right eye and mouth, distance between right eye and nose; distance between nose and forehead, distance between right ear and mouth, mouth width, distance between left ear and mouth, distance between right eye and eyebrow, distance between left eye and eyebrow, and nose width. This result is shown clearly in Figure 22 and Table 11.

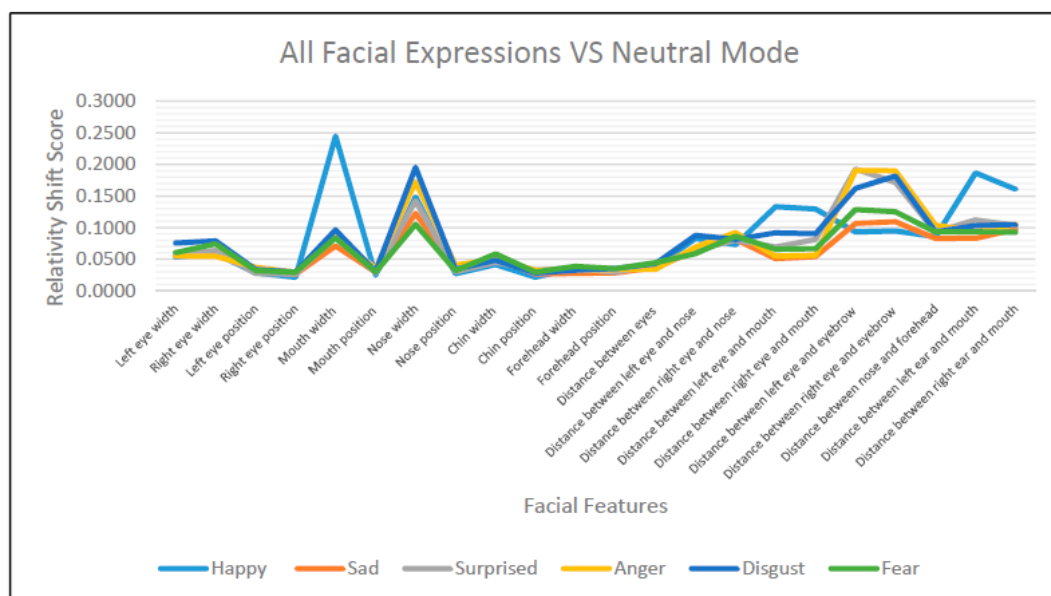


Figure 22. The mean of RSS of 22 facial features for 36 subjects on all expressions.

Table 11. The mean of RSS of 22 facial features for 36 subjects on all expressions.

Features	Happy	Sad	Surprised	Anger	Disgust	Fear	Mean	Rank
Left eye width	0.0533	0.0568	0.0563	0.0550	0.0756	0.0604	0.0595	10
Right eye width	0.0592	0.0599	0.0639	0.0546	0.0790	0.0742	0.0651	11
Left eye position	0.0286	0.0311	0.0276	0.0375	0.0336	0.0321	0.0318	4
Right eye position	0.0215	0.0264	0.0265	0.0292	0.0296	0.0293	0.0271	1
Mouth width	0.2447	0.0710	0.0858	0.0856	0.0963	0.0837	0.1112	18
Mouth position	0.0248	0.0296	0.0288	0.0354	0.0305	0.0301	0.0299	3
Nose width	0.1483	0.1226	0.1423	0.1722	0.1953	0.1053	0.1477	22
Nose position	0.0271	0.0327	0.0302	0.0410	0.0340	0.0318	0.0328	6
Chin width	0.0416	0.0520	0.0452	0.0502	0.0486	0.0585	0.0493	9
Chin position	0.0215	0.0268	0.0328	0.0326	0.0281	0.0295	0.0286	2
Forehead width	0.0360	0.0272	0.0357	0.0324	0.0324	0.0389	0.0338	7
Forehead position	0.0300	0.0284	0.0294	0.0338	0.0352	0.0352	0.0320	5
Distance between eyes	0.0387	0.0368	0.0414	0.0338	0.0429	0.0446	0.0397	8
Distance between left eye and nose	0.0826	0.0641	0.0643	0.0691	0.0877	0.0590	0.0711	12
Distance between right eye and nose	0.0730	0.0816	0.0801	0.0924	0.0814	0.0865	0.0825	15
Distance between left eye and mouth	0.1331	0.0511	0.0693	0.0551	0.0916	0.0660	0.0777	13
Distance between right eye and mouth	0.1293	0.0542	0.0814	0.0563	0.0904	0.0667	0.0797	14
Distance between left eye and eyebrow	0.0931	0.1065	0.1927	0.1906	0.1621	0.1285	0.1456	21
Distance between right eye and eyebrow	0.0946	0.1094	0.1716	0.1899	0.1815	0.1250	0.1453	20
Distance between nose and forehead	0.0855	0.0829	0.0931	0.1038	0.0933	0.0931	0.0919	16
Distance between left ear and mouth	0.1866	0.0832	0.1120	0.0945	0.1031	0.0933	0.1121	19
Distance between right ear and mouth	0.1608	0.0970	0.1033	0.1053	0.1049	0.0928	0.1107	17

Additionally, Figure 22 shows the means for every facial feature in all six expressions; it indicates that there is a clear pattern in the lower values of the RSS where all the facial features have the same pattern. For example, all features have low right eye position's RSS while there is inconsistency in the mouth width feature's RSS as it is high. Next, we applied Equation (6) to obtain the SS for the top five features, top ten features, and the worst ten features with respect to all expressions, as follows: based on the rank in Table 9.

$$SS = \left| \frac{1 - \text{Sum (Relativity between neutral and expression for the selected facial features)}}{\text{number of selected features}} \right| \quad (19)$$

The purpose is to find out a set of facial features that suitable for all expressions where it will achieve a high SS score that does not matter what FE is presented.

Table 12 and Figure 23 show the SS with respect to all expressions.

Table 12. The SS with respect to all expressions using all 22 features, top five, top ten, worst ten.

Similarity	Happy	Sad	Surprise	Anger	Disgust	Fear	Average
Similarity for all 22 features	91.754%	93.948%	92.665%	92.500%	92.013%	93.344%	92.704%
Similarity for top five features	97.469%	97.154%	97.097%	96.631%	96.860%	96.876%	97.014%
Similarity for top ten features	96.767%	96.521%	96.461%	96.192%	96.094%	96.096%	96.355%
Similarity for worst ten features	86.510%	91.405%	88.684%	88.545%	88.001%	90.593%	88.956%



Figure 23. The SS with respect to all expressions using all 22 features, top five, top ten, worst ten.

Table 13 and Figures 24–26 show the SS for all 22 facial features, top five, top ten, worst ten with respect to each expression and top five, top ten, worst ten with respect to all expressions.

Table 13. SS for all 22 facial features, top ten, top 5, worst ten with respect to each expression and with respect to all expressions.

Similarity	Happy	Sad	Surprise	Anger	Disgust	Fear
Similarity for all 22 features	91.754%	93.948%	92.665%	92.500%	92.013%	93.344%
Similarity for top five feature with respect to expression	97.526%	97.232%	97.149%	96.765%	96.917%	96.943%
Similarity for top five feature with respect all expression	97.469%	97.154%	97.097%	96.631%	96.860%	96.876%
Similarity for top ten feature with respect to expression	96.767%	96.578%	96.461%	96.196%	96.094%	96.110%
Similarity for top ten feature with respect all expression	96.767%	96.521%	96.461%	96.192%	96.094%	96.096%
Similarity for worst ten feature with respect to expression	86.415%	91.218%	88.684%	88.405%	87.938%	90.510%
Similarity for worst ten feature with respect all expression	86.510%	91.405%	88.684%	88.545%	88.001%	90.593%

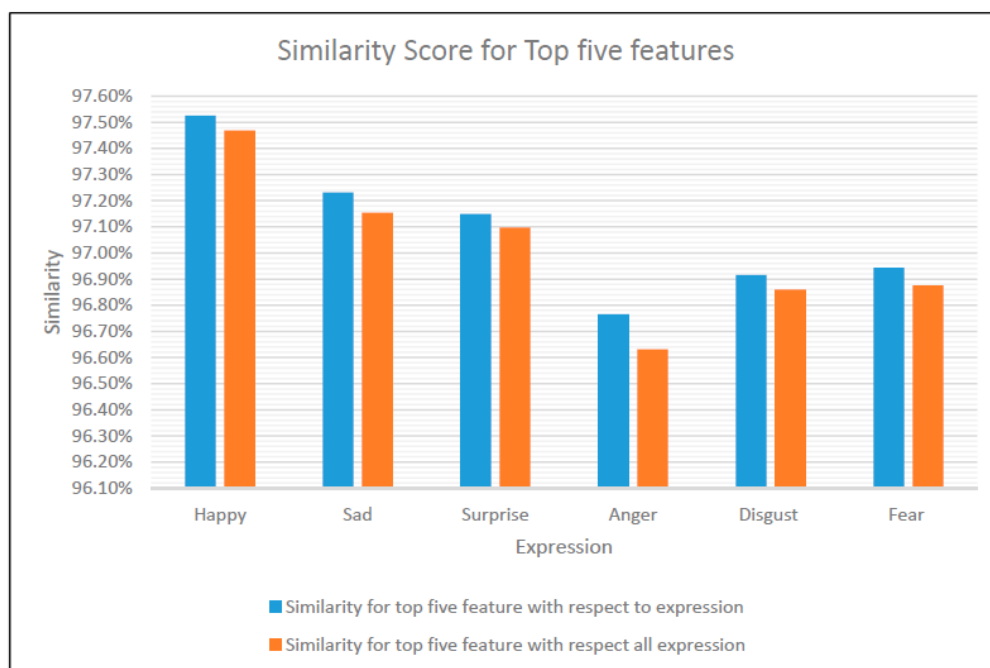


Figure 24. SS for top five features with respect to each expression vs. all expressions.

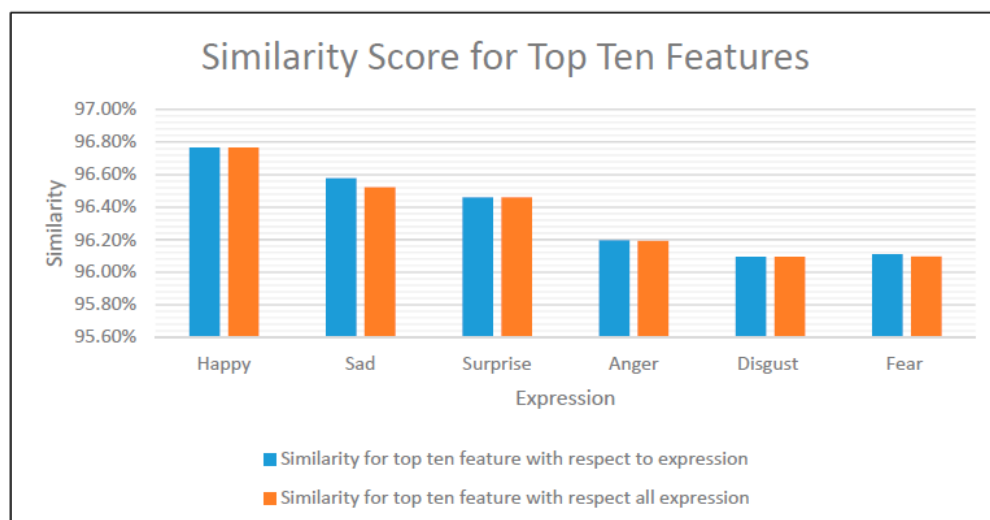


Figure 25. SS for top ten features with respect to each expression vs. all expressions.

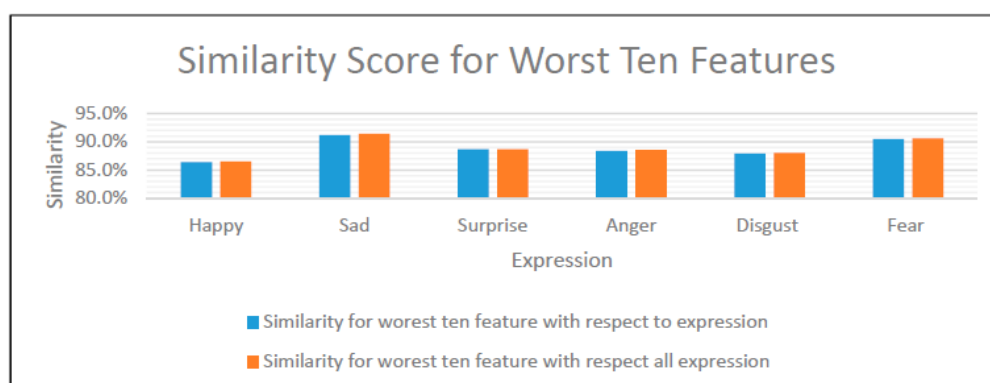


Figure 26. SS for worst ten features with respect to each expression vs. all expressions.

5.8. Face Biometric System Performance

To validate our methodology, we have applied Equation (1) to determine the FRR at three acceptance thresholds; 99%, 95%, and 90%, and compare those rates using all facial features, top five, and top ten features. Considering that we have 216 instances (36 subjects * Six FE comparisons), the results as shown in Table 14 are; At 99% acceptance threshold; out of 216 instances, 216 have been rejected using all facial features, while 214 rejections using top ten features, and 190 rejection using top features. At a 95% acceptance threshold, out of 216 instances, 171 have been rejected using all features, 41 rejections using the top ten features, and 34 rejections using the top five features. At 90% acceptance threshold, out of 216 instances, 30 rejections using all features, one rejection using the top ten features, and no rejection using the top five features.

Table 14. False rejection rate (FRR) at different acceptance threshold.

Similarity Acceptance Threshold	# of Rejection Using All Facial Features (22)	# of Rejection Using Top Ten Features	# of Rejection Using Top Five Features
>99.00%	216 (100%)	214 (99.0741%)	190 (87.9630%)
>95.00%	171 (79.1667%)	41 (18.9815%)	34 (15.7407%)
>90.00%	30 (13.8889%)	1 (0.4630%)	0 (0.0000%)

We can notice that the number of rejections of a genuine user has been decreased using the top five and ten features as follows: The rejection rate at 99% threshold has been decreased by 0.92% between (all) and (top ten) features and 12.03% between (all) and (top five) features. While, it decreased at a 95% threshold by 60.18% and 63.42% between (all) and (top ten), (top five) features, respectively. Finally, at a 90% threshold, it decreased to 13.42% and 13.88% between (all) and (top ten), (top five) features, respectively.

5.9. Top Five vs. Top Ten Facial Features

It could be perceived that the SS for the top five features is very close to the SS of top ten features, as shown in Figure 27. For example, the SS of anger expression in the top ten was 96.192%, while in the top five was 96.631%. This observation gave us the chance to use either of these two choices based on face biometrics' needs and restrictions. However, from a computational cost perspective, there is no difference between five and ten features. As a result, it is recommended to use the top ten features from security respective.

By comparing the neutral mode with six FEs, the average genuine SS has been degraded. This means that there is an effect of FE on the FB system's reliability. Additionally, it could be observed that the results of SS with respect to each expression are similar or very close to being similar to SS with respect to all expressions in each expression—meaning that the facial features that cause the muscle movements during the expressions are the same in all expressions as the same for each expression

Moreover; from all the previous results and observations, it has been proven that the top five facial features are: right eye position, chin position, mouth position, left eye position, and forehead position. The top ten facial features (right eye position, chin position, mouth position, left eye position, forehead position, nose position, forehead width, distance between eyes, chin width, and left eye width) are suitable for all FE in FB and can be generalized during the recognition process since it will provide a higher similarity score no matter what the presented expression is.

Finally, by evaluating the performance using FRR, the results showed that using the top five facial features leads to getting more people correctly accepted and less falsely rejected within the face biometric-based authentication system.

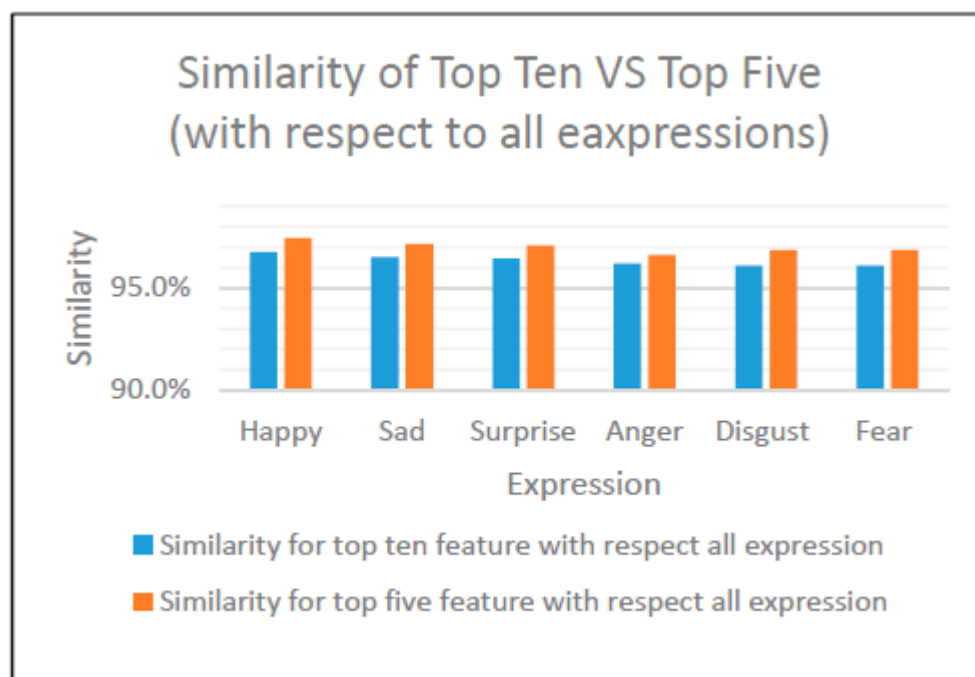


Figure 27. Top ten vs. top five features.

6. Findings

The results of this study found that: (1) The following FEs have impacts on the FB system reliability: happy, sad, surprised, anger, disgust, fear. (2) The sad expression achieved the best SS, 93.94%. (3) The disgust expression achieved the worst SS, 92.01%. (4) Out of the 22 facial features, the following top features have the best RSS as they have the lowest facial deformations: right eye position, chin position, mouth position, left eye position, and forehead position. While the top ten features were: right eye position, chin position, mouth position, left eye position, forehead position, nose position, forehead width, distance between eyes, chin width, and left eye width. Meanwhile, the worst ten features with the highest facial deformations were: distance between left eye and mouth, distance between right eye and mouth, distance between right eye and nose, distance between nose and forehead, distance between right ear and mouth, mouth width, distance between left ear and mouth, distance between right eye and eyebrow, distance between left eye and eyebrow, and nose width. (5) Furthermore, the mean of the RSS showed less variances between the sample using the top facial features. (6) Additionally, it has been found that the performance of the top five and the top ten features were very similar to each other. (7) Finally, the top features can be generalized during the recognition process regardless of what expression is presented during the verification.

By these findings, the FRR has been minimized, and the recognition acceptance threshold raised up to the possible highest without worrying about user convenience. As a result, the intrusion detection will be improved.

7. Conclusions

This paper investigates the effect of facial expressions on the face biometric system's reliability. Happy, sad, surprised, anger, disgust, and fear facial expressions have an impact on the accuracy and may cause false rejection for a genuine user. The statistical analysis of the obtained facial features between the neutral face and the six expressions identified a set of facial features that have the lowest facial deformations. The top features that have been identified in this study can be utilized in a part-based feature representation that removes some parts (regions) from the face and exploits the regions of interest so that will they not affect the recognition accuracy [42]. By the findings of this study, the false rejection rate has been minimized, as the false rejection instances caused by facial expressions

have been minimized. Thus, the matching threshold can be raised without worrying that it will affect the user convenience.

The results of this paper can be utilized in other aspects where the artificial intelligence can be used to preserve the security of the user's identity and its data by authorizing the user using his emotions as behavioral traits where the intensity of his emotion will be used as a verification. By this utilization, the impact of facial expression impact is eliminated. Another area of improvement is to conduct the analysis in an uncontrolled environment where different factors beside the facial expressions are present.

Author Contributions: Conceptualization, H.A.A.; methodology, H.A.A.; software, H.A.A.; validation, H.A.A. formal analysis, H.A.A.; investigation, H.A.A.; resources, H.A.A. and R.Z.; data curation, H.A.A.; writing—original draft preparation, H.A.A.; writing—review and editing, H.A.A. and R.Z.; supervision, R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Data Availability: Previously reported [IMPA-FACES3D] dataset was used to support this study and its available at. [<http://app.visgraf.impa.br/database/faces/>]. This dataset is cited at relevant places within the text as reference [23].

Limitations: Those results were subject to facial expression factor only and did not take into considerations other factors that may causes the false rejection such as illumination, camera perspective, shadowing and subject's pose. The analysis should be run in real recognition scenarios.

References

1. Azimi, M.; Pacut, A. The effect of gender-specific facial expressions on face recognition system's reliability. In Proceedings of the 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), Cluj-Napoca, Romania, 24–26 May 2018; pp. 1–4.
2. Malhotra, J.; Raina, N. Biometric face recognition and issues. In Proceedings of the 2015 International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 11–13 March 2015; pp. 1239–1241.
3. Dasgupta, D.; Roy, A.; Nag, A. *Advances in User Authentication*; Springer: Berlin/Heidelberg, Germany, 2017.
4. Prabhakar, S.; Jain, A.; Ross, A. An Introduction to Biometric Recognition. *IEEE Trans. Circuits Syst. Video Technol.* **2004**, *14*, 4–20.
5. Mansfield, A.J.; Wayman, J.L. *Best Practices in Testing and Reporting Performance of Biometric Devices*; NPL Report CMSC 14/02; National Physical Laboratory: London, UK, August 2002.
6. Anwarul, S.; Dahiya, S. A comprehensive review on face recognition methods and factors affecting facial recognition accuracy. In *Proceedings of ICRIC 2019*; Springer: Berlin, Germany, 2020; Volume 597, pp. 495–514.
7. Gu, L.; Kanade, T.F. Face Acquisition ► Face Device Face Aging Face Alignment. In *Encyclopedia of Biometrics*; Springer Science & Business Media: Berlin, Germany, 2009.
8. Jalal, A.; Tariq, U. Semi-supervised clustering of unknown expressions. *Pattern Recognit. Lett.* **2019**, *120*, 46–53. [[CrossRef](#)]
9. Crowley, T. The Expression of the Emotions in Man and Animals. *Philos. Stud.* **1957**, *7*, 237. [[CrossRef](#)]
10. Jafri, R.; Arabnia, H.R. A Survey of Face Recognition Techniques. *J. Inf. Process. Syst.* **2009**, *5*, 41–68. [[CrossRef](#)]
11. Ekman, P. An argument for basic emotions.pdf. *Psychol. Rev.* **1992**, *99*, 550–553. [[CrossRef](#)] [[PubMed](#)]
12. Özseven, T.; Düğenci, M. Face recognition by distance and slope between facial landmarks. In Proceedings of the IDAP 2017-International Artificial Intelligence and Data Processing Symposium, Malatya, Turkey, 16–17 September 2017; pp. 3–6.
13. Amato, G.; Falchi, F.; Gennaro, C.; Vairo, C. A Comparison of Face Verification with Facial Landmarks and Deep Features. In Proceedings of the 10th International Conference on Advances in Multimedia, Athens, Greece, 22–26 April 2018; pp. 1–6.
14. Kazemi, V.; Sullivan, J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1867–1874.

15. Banerjee, I. Establishing User Authentication Using Face Geometry. *Int. J. Comput. Appl.* **1975**, *8887*, 2014. [[CrossRef](#)]
16. Napieralski, J.A.; Pastuszka, M.M. 3D Face Geometry Analysis for Biometric Identification; In Proceedings of the 21st International Conference Mixed Design of Integrated Circuits and Systems, Lublin, Poland, 19–21 June 2014.
17. Sabri, N.; Henry, J.; Ibrahim, Z.; Ghazali, N.; Mangshor, N.N.; Johari, N.F.; Ibrahim, S. A Comparison of Face Detection Classifier using Facial Geometry Distance Measure. In Proceedings of the 2018 9th IEEE Control and System Graduate Research Colloquium, Shah Alam, Malaysia, 3 August 2018; pp. 116–120.
18. Benedict, S.R.; Kumar, J.S. Geometric shaped facial feature extraction for face recognition. In Proceedings of the 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, India, 24 October 2016; pp. 275–278.
19. Gurnani, A.; Shah, K.; Gajjar, V.; Mavani, V.; Khandhediya, Y. SAF-BAGE: Salient Approach for Facial Soft-Biometric Classification-Age, Gender, and Facial Expression. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 839–847.
20. Barroso, E.; Santos, G.; Proenca, H. Facial expressions: Discriminability of facial regions and relationship to biometrics recognition. In Proceedings of the 2013 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM), Singapore, 16–19 April 2013; pp. 77–80.
21. Teng, J. Facial Expression Recognition with Identity and Spatial-temporal Integrated Learning. In Proceedings of the 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Cambridge, UK, 3–6 September 2019; pp. 100–104.
22. Divya, M.B.S.; Prajwala, N.B. Facial Expression Recognition by Calculating Euclidian Distance for Eigen Faces Using PCA. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSPP), Chennai, India, 3–5 April 2018; pp. 244–248.
23. Ivanovsky, L.; Khryashchev, V.; Lebedev, A.; Kosterin, I. Facial expression recognition algorithm based on deep convolution neural network. In Proceedings of the 2017 21st Conference of Open Innovations Association (FRUCT), Helsinki, Finland, 6–10 November 2017; pp. 141–147.
24. Sun, X.; Xia, P.; Zhang, L.; Shao, L. A ROI-guided deep architecture for robust facial expressions recognition. *Inf. Sci. (N. Y.)* **2020**, *522*, 35–48. [[CrossRef](#)]
25. Yang, J.; Zhang, F.; Chen, B.; Khan, S.U. Facial Expression Recognition Based on Facial Action Unit. In Proceedings of the 2019 Tenth International Green and Sustainable Computing Conference (IGSC), Alexandria, VA, USA, 21–24 October 2019; pp. 1–6.
26. Zhong, L.; Liao, H.; Xu, B.; Lu, S.; Wang, J. Tied gender condition for facial expression recognition with deep random forest. *J. Electron. Imaging* **2020**, *29*, 023019. [[CrossRef](#)]
27. Jeong, D.; Kim, B.G.; Dong, S.Y. Deep Joint Spatio-Temporal Network (DJSTN) for Efficient Facial Expression Recognition. *Sensors* **2020**, *20*, 1936. [[CrossRef](#)] [[PubMed](#)]
28. Mehta, D.; Siddiqui, F.H.M.; Javaid, A.Y. Recognition of emotion intensities using machine learning algorithms: A comparative study. *Sensors* **2019**, *19*, 1897. [[CrossRef](#)] [[PubMed](#)]
29. Yitzhak, N.; Gurevich, T.; Inbar, N.; Lecker, M.; Atias, D.; Avramovich, H.; Aviezer, H. Recognition of emotion from subtle and non-stereotypical dynamic facial expressions in Huntington’s disease. *Cortex* **2020**, *126*, 343–354. [[CrossRef](#)] [[PubMed](#)]
30. Mattavelli, G.; Barvas, E.; Longo, C.; Zappini, F.; Ottaviani, D.; Malaguti, M.C.; Pellegrini, M.; Papagno, C. Facial expressions recognition and discrimination in Parkinson’s disease. *J. Neuropsychol.* **2020**. [[CrossRef](#)] [[PubMed](#)]
31. Flynn, M.; Effraimidis, D.; Angelopoulou, A.; Kapetanios, E.; Williams, D.; Hemanth, J.; Towell, T. Assessing the Effectiveness of Automated Emotion Recognition in Adults and Children for Clinical Investigation. *Front. Hum. Neurosci.* **2020**, *14*, 70. [[CrossRef](#)] [[PubMed](#)]
32. Yin, M.D.B.; Mukhlas, A.A.; Wan, R.Z.; Chik, A.; Othman, T.; Omar, S. A proposed approach for biometric-based authentication using of face and facial expression recognition. In Proceedings of the 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS), Singapore, 28–30 December 2018; pp. 28–33.

33. Ming, Z.; Chazalon, J.; Luqman, M.M.; Visani, M.; Burie, J.-C. FaceLiveNet: End-to-End Networks Combining Face Verification with Interactive Facial Expression-Based Liveness Detection. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 3507–3512.
34. Pavel, F.A.; Iordănescu, E. The influence of facial expressions on recognition performance in facial identity. *Procedia-Soc. Behav. Sci.* **2012**, *33*, 548–552. [CrossRef]
35. Coelho, R.; Dalapicola, R.; Queiroga, C.T.V.; Ferraz, T.T.; Borges, J.T.N.; Saito, H.; Gonzaga, A. Impact of facial expressions on the accuracy of a CNN performing periocular recognition. In Proceedings of the 2019 8th Brazilian Conference on Intelligent Systems (BRACIS), Salvador, Brazil, 15–18 October 2019; pp. 401–406.
36. Azimi, M. Effects of Facial Mood Expressions on Face Biometric Recognition System's Reliability. In Proceedings of the 2018 1st International Conference on Advanced Research in Engineering Sciences (ARES), Dubai, UAE, 1 June 2018; pp. 1–5.
37. Márquez-Olivera, A.G.M.; Juárez-Gracia, V.; Hernández-Herrera, A.; Argüelles-Cruz, J.; López-Yáñez, I. System for face recognition under different facial expressions using a new associative hybrid model $\alpha\beta$ -KNN for people with visual impairment or prosopagnosia. *Sensors* **2019**, *19*, 578. [CrossRef] [PubMed]
38. Khorsheed, J.A.; Yurtkan, K. Analysis of Local Binary Patterns for face recognition under varying facial expressions. In Proceedings of the 2016 24th Signal Processing and Communication Application Conference, SIU 2016-Proceedings, Zonguldak, Turkey, 16–19 May 2016; pp. 2085–2088.
39. Schorr, B.; Schorr, B.S. *Banco de Dados de Faces 3D: IMPA-FACE3D*; IMPA-RJ: Rio DE Janeiro State, Brazil, 2010.
40. Pictures-FacesDB|VISGRAF. Available online: <http://app.visgraf.impa.br/database/faces/pictures/> (accessed on 13 June 2020).
41. Abramowitz, I.A.; Stegun, M. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (Partially Mathcad-enabled)*; U.S. Department of Commerce: Washington, DC, USA, 1972.
42. Li, S.; Deng, W. Deep Facial Expression Recognition: A Survey. *IEEE Trans. Affect. Comput.* **2020**. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).