

Facial action coding-based facial sub-structures for anxiety emotion classification

Rawinan Praditsangthong, Pattarasinee Bhattarakosol

Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Article Info

Article history:

Received Jun 25, 2022

Revised Nov 15, 2022

Accepted Nov 19, 2022

Keywords:

Decision tree
Emotion classification
Facial features
Stroke patients
Sub-structures

ABSTRACT

Most stroke patients usually have problems in communication and body movement, such as speaking, sitting, walking, and picking up items. Moreover, the number of caregivers is smaller than the number of stroke patients. Nevertheless, these patients need 24-hour caregivers for the patients' safety. Therefore, the objective of this research is to determine patterns of anxiety emotion via facial expressions from the sub-structures on the face, such as the inner brow raiser, brow lower, lid raiser, and lip part. Random samples of 360 facial images from horror-thriller movies based on the internet movie database (IMDb) website were selected. Then, 68 facial landmarks for classifying the emotions were applied to each facial image. The differences in these 68 positions before and after the changed emotions were used as the emotional indicators. Furthermore, these different values are applied to implement a decision tree with all the boundaries of the sub-structures disclosed as a suitable classification model for emotion detection. Consequently, the accuracy when applying this decision tree with other facial images is 83.33%.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Pattarasinee Bhattarakosol

Department of Mathematics and Computer Science, Faculty of Sciences, Chulalongkorn University

254 Phayathai, Pathumwan, Bangkok, Thailand

Email: pattarasinee.b@chula.ac.th

1. INTRODUCTION

A human face is an interesting organ that automatically express emotion, such as happy, angry, disgusted, surprised, sad, and fearful. Moreover, every emotion causes a reaction after facing stimuli. The response of the sentiment will be expressed through facial muscle movement. In the past decade, various studies have investigated facial detection and recognition using the facial action coding system (FACS) based on Paul Ekman's theory [1]-[4]. Furthermore, FACS has been applied in the areas of neuroscience and psychology to detect sentiments of mental illness [3], including autism disorder [5] through facial expressions. Therefore, FACS is relevant to the structure of facial movement. The measurement of facial emotional expression using encoding on the face is composed of many action units (AUs).

Ekman and Friensen [1] developed the FACS from anatomical text of the facial muscles in the expression of six basic emotions, namely happiness, anger, disgust, surprise, sadness, and fear. FACS consists of action units (AUs), the facial action code (FAC) name, and muscular basis. Each AU is a number that represents a basic action of a particular muscle. Moreover, each FAC name is the title of the activity that describes the response on the face related to the AU. Additionally, the muscular basis explains the motions of muscles. Thus, each fundamental emotion can be represented by AUs such as AU6 refers to the muscle under an eyelid when the cheek raises, and AU12 describes the corner of the mouth when smiling, as displayed in Figure 1.

For example, happiness contains AU6 and the AU12, as shown in Figure 1(a), whilst fear comprises AU1, AU2, AU4, AU5, AU7, AU20, and AU25, as presented in Figure 1(b). Surprise consists of AU1, AU2, AU5, AU25, and AU26, as presented in Figure 1(c). Anger refers to AU4, AU5, AU7, and AU23, as show in Figure 1(d). Moreover, sadness and disgust compose three AUs with AU1, AU4, AU15, and AU9, AU15, AU16, as presented in Figure 1(e) and Figure 1(f), respectively.

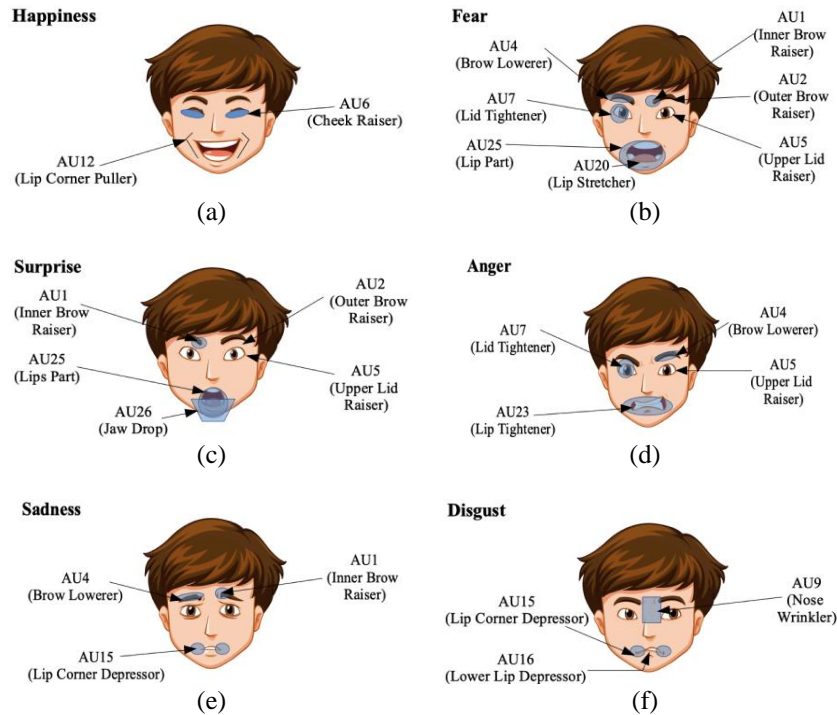


Figure 1. Action units in basic emotions for (a) happiness, (b) fear, (c) surprise, (d) anger, (e) sadness, and (f) disgust

The mechanisms of FACS use the AUs not only for emotion detection, classification, and recognition but also for the machine learning process to improve the classification and recognition model. For example, a decision tree algorithm was applied to recognize emotions on social networking sites using the sentence level classification [6], which was used to detect emotions on the faces from images [7]. The steps to classify an emotion are face detection, feature extraction, and classification [8], [9]. The face detection procedure detects a frontal face from images using an open-source library such as OpenCV, and DLib. Therefore, implementing feature extraction techniques, such as pixel intensities or pixel coordinates, can determine the significant positions on the face [9], [10]. The last step is to categorize emotions from the features using machine learning algorithms, as mentioned above. Based on the techniques and algorithms mentioned preciously, the facial detection, facial recognition, and facial classification of expressions are automatically performed. As a consequence, an automated facial expression system was applied in several fields, such as security, factories, and healthcare, In the security area, the facial expression and recognition system is deployed to authenticate online services using a camera phone [11], [12]. This system guarantees the access of smartphones by encoding a facial image as a completely automated public turing test to tell computers and humans apart (CAPTCHA). Furthermore, a facial expression analysis method has also been applied to the construction industry, in developing robotic scaffolding assembly systems. This robot can identify humans using voice and face recognition, including a response with verbal commands [13], [14]. In addition, facial expression detection can assist an elderly or bedridden persons regarding their safety concerns and improve their quality of life since this detection system can monitor the facial changes of elderly persons or patients via a camera or similar specific equipment. For this reason, a healthcare framework employing facial expressions has been developed as a smart healthcare system [15], [17]. Therefore, the contributions of this research are as:

- Sub-structure features on the face such as inner brow raiser (IBR), brow lower (BL), lid raiser (LR), and lip part (LP).
- An emotional classification model between neutral emotions and fear emotions based on the feature extraction.

The remainder of this paper is as shown in: section 2 presents the proposed method. Then, methods are described in section 3, while the results and discussion are explained in section 4. Finally, the conclusion and acknowledgements appear in section 5 and section 6, respectively.

2. PROPOSED METHOD

2.1. Problem statement and solution

Whenever people cannot control an organ, such as their arms, legs, or head, according to an uncontrollable muscle, these people are called paralyzed. The paralysis usually refers to the inability to move some part of the body's muscles when needed. For example, a stroke patient may be unable to move their arms and legs or speak. Thus, they need intimate care all of the time, which requires many caregivers. Unfortunately, the number of stroke patients is continuously increasing [18], [19]; in contrast, the number of caregivers is not commensurate with the increasing number of patients, including the mental effects and burden on caregivers [20], [21]. Thus, these issues cause a negative impact on the patients' safety. Therefore, using technology to support medical care is a possible solution with a sustainable cost of payment. Consequently, the idea of implementing small cameras to monitor a patient's expressions via the face and then interpreting the patient's critical requirements and feelings, such as breathlessness, chest pain, etc., can be an appropriate solution for all paralyzed patients. Generally, the captured images from the camera are analyzed and classify the patient status, such as normal and abnormal.

Consider Figure 2, in which a set of cameras is implemented over the patient's headboard. The set of cameras consists of the camera stand, as shown in Figure 2(a), and Raspberry Pi devices with a Raspberry Pi camera module, as shown in Figure 2(b). Cameras are responsible for capturing images of the patient's face for every movement from every perspective. All processes are computed in the intranet system, as shown in Figure 2(c). Then, all images will be sent to a computer server, as shown in Figure 2(d), for interpreting the patient's emotion; the alarm will provide an alert when an abnormal situation is detected.

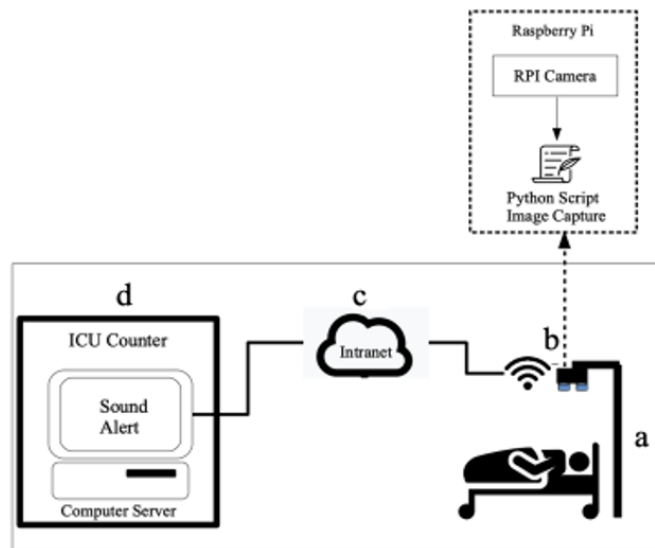


Figure 2. The infrastructure of the smart monitoring system (SMS) for (a) camera stand, (b) Raspberry Pi device, (c) internet system, and (d) computer server

2.2. Proposed method

The structure of the human face consists of a forehead, two eyes, one nose shape, one cheek, one chin, and one mouth [22]. Additionally, these structures also have sub-structures, as shown in Table 1. Once the emotion is altered, the sub-structures are also altered. Thus, this research will focus on the variations of the inner brow raiser (IBR), brow lower (BL), lid raiser (LR), and lip part (LP) under two diverse emotions: neutral and fear. To recognize all sub-structures on the face, this research applied 68 facial landmarks to detect the structural components of the face, as presented in Figure 3. From Figure 3(a), the IBR is calculated by the number pair (20, 27) of the right eyebrow and the number pair (23, 27) of the left eyebrow. The BL is the distance between position number 21 and position number 22, as shown in Figure 3(b).

Consider the LR value. This LR contains 4 distance values that are computed from position number 37 to position number 41, and position number 38 to position number 40 of the right eye, including position number 43 to 47 and position number 44 to position number 46 to the left eye, as displayed in Figure 3(c). Similar to the LR, the LP value, an essential structure for the changing of the lip’s shape, is composed of six positions: 61, 62, 63, 64, 65, and 66. Then, all six positions are paired and the length of each pair is calculated; the order pairs of these positions are (61, 65), (62, 66), and (63, 67), as shown in Figure 3(d). After all of the input parameters and output results of this research are clearly defined, the emotional classification model using a decision tree model can be written as a model with the input parameters of the IBR, the BL, the LR, and the LP; the output of the model can be either neutral or fear.

Table 1. Structure and sub-structure on the face

Structures	Sub-Structures
Forehead	Eyebrow, Ears
Eyes	Eyelid, Eyelash
Nose Shape	Nose
Cheek	Cheek, Chin
Mouth	Inner Lip, Outer Lip

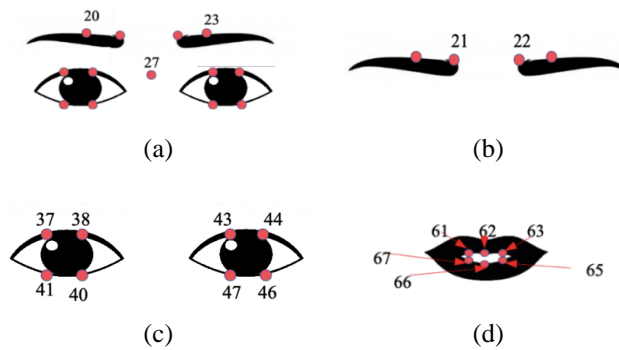


Figure 3. Structure on the face of (a) inner brow raiser, (b) brow lower, (c) lid raiser, and (d) lip part

3. METHOD

The research uses frontal face images for differentiating between a neutral state and a fear state. Based on the assumption that every human, including stroke patients, can express their anxiety via their faces, all images were chosen from horror-thriller movies on the internet movie database (IMDb) website, the well-known ratings online database for movies, and only 24 films were selected within three months because these films have high ratings of over 6.5 out of 10. As a result, the total number of actors/actresses is 60 with the age range of 18-61 years. In addition, six images were captured from each of the individual players; three neutral emotion images from neutral emotion scenes and three fear emotion images from fear emotion scenes. Therefore, 360 images are used in this experiment.

Once all frontal face images are captured, their size must be transformed to the particular 800x800 pixels known as the real scale size of the face images [23]. After re-sizing, 68 facial landmarks were determined over a piece of the facial image. Thereupon, with 68 facial landmarks, 240 out of 360 images were set as the training data while the remainder were set as the testing data. Moreover, based on 240 images of the training data set, 40 actors are separated to comprise groups that are represented as 50.0%, male and female. The range of 18-25 years old is the highest percentage of the age range, which is 35.0%. In the statistical testing, the non-parametric Wilcoxon signed ranks test is deployed to compare the differences between the neutral emotion and the fearful emotion. The Wilcoxon signed ranks test is performed employing a 95% confidence level or the 0.05 significance level.

3.1. The distance average and standard deviation of the inner brow raiser (IBR)

Given the distribution of the data, the values of the inner brow raisers are not normal; therefore, the non-parametric analysis is applied to analyze and conclude the differences between neutral and fear situations. Therefore, when using the Wilcoxon signed ranks test with a 95% confidence level, the statistical values indicated that there is no significance of the means’ differences between the IBRs under the neutral and fear emotions for both males and females, p -value=0.417>0.05 and p -value=0.454>0.005, respectively. Moreover,

the calculation also points out that in the fear state, the mean value of the right IBR of males is significantly different from females, but not the left side. Nevertheless, the mean of the IBRs for both sides in males under the normal mode shows no significant difference from the mean of the IBRs in females.

Considering the differences of means of the IBRs under the neutral state and fear state, the non-parametric Wilcoxon signed ranks test can determine that there is a significant difference between the means of the right IBRs of males when the emotion has been changed from neutral to fear, $p\text{-value}=0.00<0.05$, similar to the left IBRs. Similar to the males, there is also a significant difference between the means of the right IBRs of females, which is the same as the left IBRs, when their temperament was altered from normal to the scared state, $p\text{-value}=0.00<0.05$. The distance averages and standard deviations of the IBRs on all states of both males and females are displayed in Table 2.

Table 2. The distance average and standard deviation of inner brow raiser (IBR) (millimeters)

Emotion	Inner Brow Side	Gender		Total Mean
		Male	Female	
Neutral	Right	37.18 ± 3.58	39.42 ± 3.12	38.30 ± 3.35
	Left	37.45 ± 3.13	38.42 ± 2.58	37.94 ± 2.86
Fear	Right	38.14 ± 3.39	39.32 ± 4.08	38.73 ± 3.74
	Left	39.85 ± 3.64	39.81 ± 2.59	39.83 ± 3.12

3.2. The distance average and standard deviation of the brow lower (BL)

Consider the brow lower (BL), which is the distance between the right end of the left eyebrow to the left end of the other eyebrow, this value was measured and compared in both males and females for differences, as shown in Table 3. With the same reasoning applied in the previous section, running the Wilcoxon signed ranks test indicates that there is no significant difference between the mean values of the BLs of males and females, with a $p\text{-value}=0.514>0.05$. In the neutral situation, the distance average of the BL in males is significantly different from the distance average of the BL in females, with a $p\text{-value}=0.000<0.05$. Furthermore, in the fear situation, there is a significant difference between the mean values of the BL in males and females, with a $p\text{-value}=0.000<0.05$. Furthermore, in both males and females, the mean value of the BL in the normal mode is significantly different from the mean value of the BL in the fear mode, with a $p\text{-value}=0.000<0.05$.

Table 3. The distance average and standard deviation of brow lower (BL) (millimeters)

Emotion	Gender		Total Mean
	Male	Female	
Neutral	25.42 ± 3.04	29.15 ± 2.66	27.29 ± 2.85
Fear	26.27 ± 3.30	27.17 ± 3.32	26.72 ± 3.31

3.3. The distance average and standard deviation of the lid raiser (LR)

This section focuses on the positions of the LRs that can be classified as the inner lid and outer lid for each eye. Based on Figure 4, the outer LRs are the number pairs (37, 41) of the right eye and (44, 46) of the left eye while the inner LRs are the number pairs (38, 40) of the right eye and (43, 47) of the left eye. Since the distribution of the LRs of both eyes from all samples is not normal, the Wilcoxon signed ranks test, which is a non-parametric method, is applied. From computation of the Wilcoxon signed ranks test, the statistical values showed that there is a significant difference between the mean values of the LRs under the neutral state of both eyes, with a $p\text{-value}=0.001<0.05$, which is unlike the fear state, with a $p\text{-value}=0.105>0.05$. In addition, this method also pointed out that the average sizes of the LRs from both eyes of both genders are significantly different when the emotions have been changed from neutral to fear, with a $p\text{-value}=0.000<0.05$. The distance averages and standard deviations of the inner lid raiser and the distance averages and standard deviations of the outer lid raiser are shown in Table 4 and Table 5, respectively.

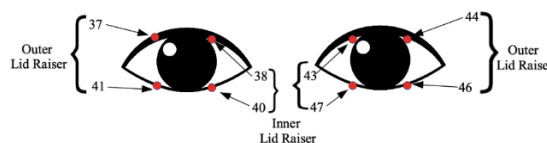


Figure 4. The inner and outer lid raiser (LR)

Table 4. The distance average and standard deviation of the inner lid raiser (inner LR) (millimeters)

Emotion	Inner Brow Side	Gender		Total Mean
		Male	Female	
Neutral	Right	8.87 ± 1.72	9.93 ± 1.85	9.40 ± 1.79
	Left	8.99 ± 1.79	10.08 ± 1.79	9.54 ± 1.79
Fear	Right	9.23 ± 1.58	10.70 ± 2.02	9.97 ± 1.80
	Left	9.32 ± 2.01	10.77 ± 2.01	10.05 ± 2.01

Table 5. The distance average and standard deviation of the outer lid raiser (outer LR) (millimeters)

Emotion	Inner Brow Side	Gender		Total Mean
		Male	Female	
Neutral	Right	8.88 ± 1.87	8.83 ± 1.87	9.45 ± 1.87
	Left	8.83 ± 1.85	10.02 ± 1.88	9.43 ± 1.87
Fear	Right	9.22 ± 2.03	10.72 ± 2.05	9.97 ± 2.04
	Left	9.16 ± 2.06	10.69 ± 2.17	9.93 ± 2.12

3.4. The distance average and standard deviation of the lip part (LP)

This section describes the average and standard deviation of the inner LP where six positions on the lip were defined, as presented in Figure 5. To study the change of the lip’s shape, six points were connected as three ordered pairs and three links were drawn. Consequently, the first link was obtained from the pair of (61, 65), the right top LP to the left bottom LP; the second link was the line between the position 62 and 66, which can be written as an ordered pair of (62, 66), the middle top lip part to the middle bottom LP. Then, the remaining points, 63 and 67, connected the left top LP to the right bottom LP represented as an ordered pair of (63, 67).

Once all the links were created, the lengths of all the links were measured and tested to determine whether they were normally distributed using the Shapiro-Wilk test. As a result, the distance averages, and standard deviations of all three links of LP are displayed in Table 6. Unfortunately, the result of the statistical test indicated that the distribution of the distances was not normal. Therefore, the non-parametric method, namely, the Wilcoxon signed ranks test, was applied using a 95% confidence level. Consequently, the Wilcoxon signed ranks test determined that under the neutral feeling, the means of all lengths were significantly different from those under fear, with the following findings: $p\text{-value}=0.001<0.05$, $p\text{-value}=0.000<0.05$, and $p\text{-value}=0.003<0.05$. Moreover, with all links, there are significant differences of means from the changed neutral emotion to fear emotion of both males and females, with the same $p\text{-value}=0.000<0.05$.

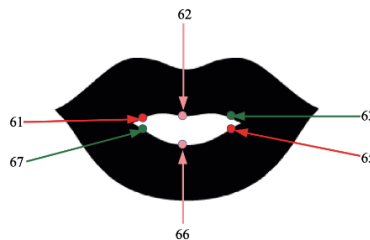


Figure 5. The inner lip part

Table 6. The distance average and standard deviation of all link’s lip part (LP) (millimeters)

Emotion	The First Link		Total Mean
	Gender		
	Male	Female	
Neutral	17.27 ± 2.55	16.88 ± 3.39	17.08 ± 2.97
Fear	18.03 ± 3.73	18.55 ± 4.38	18.29 ± 4.06
The Second Link			
Emotion	Gender		Total Mean
	Male	Female	
Neutral	4.68 ± 3.80	4.71 ± 4.38	4.70 ± 4.09
Fear	6.91 ± 5.59	8.53 ± 6.27	7.72 ± 5.93
The Third Link			
Emotion	Gender		Total Mean
	Male	Female	
Neutral	17.25 ± 2.56	16.80 ± 3.40	17.03 ± 2.98
Fear	17.80 ± 3.66	18.58 ± 4.01	3.84

3.5. Selected parameters

Referring to the statistical results above, it can be concluded that the efficiency of emotion detection can be achieved by five parameters: gender, the inner BR, the BL, the LR, and the LP. Nonetheless, there are small details for each parameter that is of concern, such as defining whether the inner BR and the LR are on the right or the left side of the face, or whether the LR is associated with the outer lid or the inner lid, as well as the inclusion of age, and the links over the LP. Therefore, to obtain a precise classification of emotion expression from a face, various parameters must be applied; these parameters are gender, age, side, position, the inner BR, the BL, the LR, and the LP. Table 7 presents all values of all parameters.

Table 7. The values of all parameters

Independent Variable	Values
Gender	Male, Female
Age-Range	18-24, 25-31, 32-38, 39-45, 46-53, 54-61
Side	Right, Left
Position	Outer, Inner
Link	First, Second, Third
Inner Brow Raiser (IBR)	[37.18, 39.85]
Brow Lower (BL)	[25.42, 29.15]
Lid Raiser (LR)	[8.83, 10.77]
Lip Part (LP)	[4.68, 18.58]
Dependent Variable	Values
State of Emotion	Normal, Fear

Furthermore, Figure 6 shows the relationship between the dependent variable and the independent variable. The first link is the link of correlation between the change of IBR/BL/LR/LP and emotion, including the relationship of the correlation of the interactions among gender, side, and emotion. However, there are direct links of correlations between interactions with gender, or side, or position, or link. Moreover, there are hidden correlations between gender and emotion, side and emotion, the position and emotion, link and emotion, and age and emotion. Thus, every independent variable has an effect on the dependent variable.

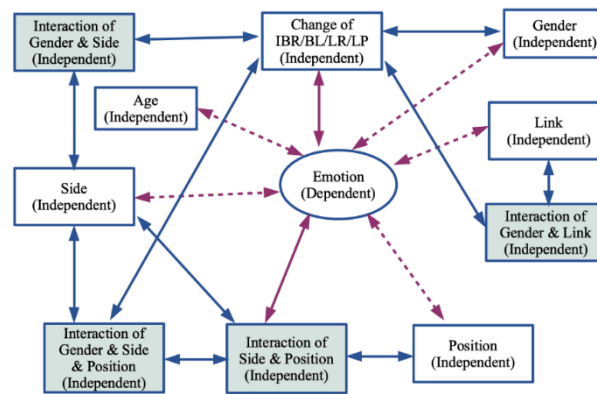


Figure 6. Relationships of all variables

3.6. Creating the model

This research uses Scikit-Learn, a machine learning library, to interpret all parameters via the PyCharm community edition version 2021.2 with Python language. The pattern was modified to a general classification model, as presented in (1).

$$Emotion = M(gender, age, side, position, IBR, BL, LR, LP) \tag{1}$$

From (1), M refers to the model of gender, age, side, position, IBR, BL, LR, and LP; the values of all parameters are defined in Table 7.

The emotion classification model is applied to classify emotion as either neutral or fear using a decision tree algorithm. Thus, 360 images of the data set were imported to PyCharm community edition with Scikit-Learn. Then, this data set was separated into two data sets: the training data set and the testing data set, with a ratio of 70:30. Therefore, 240 images of data comprised the training set that contains the cluster attributes, for which the expected result after the learning process, namely, a label, was obtained. Furthermore, the decision tree algorithm obtained as another outcome of the learning process is displayed in Figure 7.

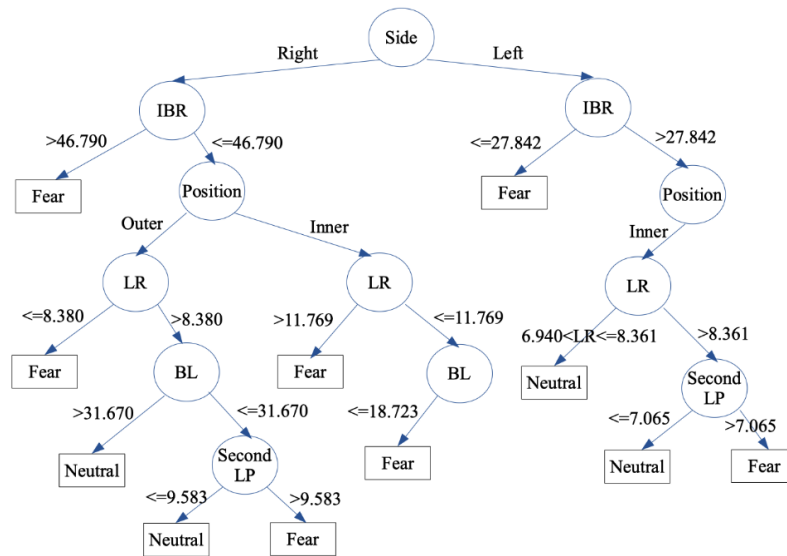


Figure 7. Decision tree algorithm of emotion classification

4. RESULTS AND DISCUSSION

4.1. Results

The face is a significant human organ that automatically shows feelings such as neutral, fear, happy, and sad. Nevertheless, there is only one important emotion that people express for help, which is the fear emotion. Therefore, this research aims to uncover a mechanism to distinguish between the neutral and fear emotions of people via their faces. Since there are various features on a face, namely the IBR and LR of the right and the left, the BL, and the LP, the positions of these organs will slightly change when the emotion has been changed. Consequently, this research intended to determine the values and the positions of all components of the face so that the right emotion can be identified automatically. Nonetheless, a face image will be marked on the basis of 68 standard facial landmarks before execution throughout the research procedures, and all changes are compared with these standard markers.

The step of emotion classification is composed of adjusting the size of the images, detecting 68 facial landmarks, creating a decision tree model, predicting the emotion, and optimizing the model. The first step is to adjust the size of the frontal face image to be 800x800 pixels, which is close to the real size of the human face, followed by the second step, where the position of IBR, LR, BL, and LP are detected based on the 68 facial landmarks using the DLib library. Then, these values are calculated to find the changed distance between the neutral emotion and fear emotion. The calculated values are representatives as features for creating a decision tree model in the emotional classification procedure. Afterward, the obtained decision tree model was evaluated according to the accuracy testing process using 120 images of the testing data set. Finally, the optimization of the classification and the prediction accuracy of the parameters in the decision tree model based on a grid search were performed. This optimized accuracy depends on true positive (TP), true negative (TN), false positive (FP), and false negative (FN) findings. Four values are used to calculate the precision, recall, and accuracy, as shown in (2)-(4).

$$Precision = \frac{TP}{(TP+FP)} \tag{2}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{3}$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (4)$$

From the experimental results in Table 8, TP is 15, FP is 3, FN is 3, and TN is 15. Consequently, the precision value is 83.33%, the recall is 83.33%, and the accuracy of this model is 83.33%. Moreover, the comparison of the performance and classification error from the same data sets with the decision tree model and random forest model is displayed in Table 8.

Table 8. Comparison of performance and classification error

	Performance	Classification Error
Decision Tree	83.33%	16.67%
Random Forest	75.00%	25.00%

In addition, Figure 8 shows the receiver operating characteristic (ROC) curves, where the performance of the decision tree model and the random forest model have been compared using RapidMiner studio version 9.5. According to the graph, the red line belongs to the decision tree while the blue line is the random forest. Thus, it can be interpreted that the decision tree has better performance than the random forest.

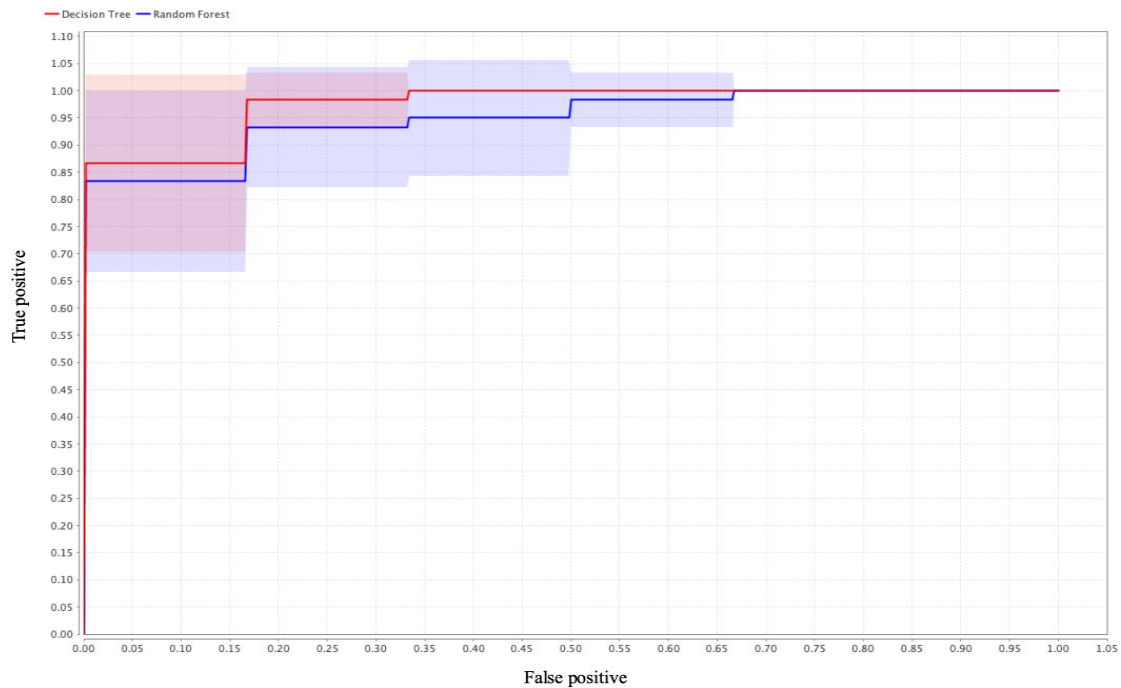


Figure 8. Comparison ROCs

4.2. Discussion

Whenever human emotion receives a meaningful stimulus, an automatic facial expression always occurs immediately and is called a micro-expression [24], [25]. This micro-expression refers to only five basic emotions: happy, angry, fearful, sad, and surprised. According to the emotional expression, the position of the forehead, eyes, nose, shape, cheek, and mouth are usually altered from the normal mood [24], [25]. Thus, these variations are suitable to be deployed in the emotion detection algorithm. Unlike other studies, this study considers the sub-structures on the face as individual components in the detection process, whereas other studies used the main structures for emotional detection. Hence, instead of using the eyes in the same manner as other studies, the eyelid and the eyelash are considered. To detect such small details on the face, specific equipment with a high capability for image detection is needed. Fortunately, there is still some equipment that is similar to the specialized equipment that has the same quality, such as the Raspberry Pi with the Raspberry Pi camera module. This equipment set can be embedded with some specific scripts as required.

One vital application of image processing relates to medical applications, especially applications for life-saving endpoints. Currently, there is a large volume of paralysis patients. These patients can send a request signal via their faces. Thus, using the main structural components or the sub-structure components to determine the feelings of patients may not be considered dissimilar. Nonetheless, in the case of a patient with facial palsy, the use of the main structure items might not allow interpretation of their feelings because of muscle incapability. In contrast, the sub-structures can be applied, and the caregivers can be notified of patients' needs by the detection system. Moreover, the application of the proposed technique does not require a large volume of facial data from a database, and the classification mechanism relies on the proposed decision tree, where the boundary value of changes in each sub-structure component is defined. Consequently, the proposed emotion detection mechanism has various benefits regarding resource optimization and fast classification because the classification does not have the time overhead of searching a database, as is the case for the other existing mechanisms.

5. CONCLUSION

The objective of this research is to investigate patterns of emotional expression on the human face. To achieve this substantial goal, the focus variables are the sub-structures on the face, together with the attributes of gender, age, side, position, and link. Furthermore, the values of these variables are compiled by the Scikit-Learn library via the PyCharm community edition version 2021.2. Therefore, all values are applied to classify the difference in the emotions between the neutral emotion and fear emotion using a decision tree algorithm. As a result of this run, the decision tree classification model with all defined boundary values is obtained. Likewise, using this obtained model, the output is the predicted emotion based on all sub-structure values. The accuracy of the emotion prediction model is 83.33%. Consequently, this mechanism is suitable to be deployed in medical face detection for facial palsy because of the usage of sub-structures. In addition, the required storage to execute this platform is small because this system is a non-database platform. The future work in similar studies will be to explore the correlations of other features on the right side and the left side of the face, such as the nose, cheek, and chin, for precise emotional prediction. Therefore, the capability of improving the quality of lives in various applications, such as transport, education, and medicine, can be achieved.

ACKNOWLEDGEMENTS

Author thanks acknowledgments of the Asahi Glass Foundation, Japan.




REFERENCES

- [1] P. Ekman and W. V. Friesen, "Measuring facial movement," *Environmental Psychology and Nonverbal Behavior*, vol. 1, no. 1, pp. 56–75, 1976, doi: 10.1007/BF01115465.
- [2] C. F. Benitez-Quiroz, R. Srinivasan, and A. M. Martinez, "Discriminant functional learning of color features for the recognition of facial action units and their intensities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2835–2845, Dec. 2019, doi: 10.1109/TPAMI.2018.2868952.
- [3] K. Wolf, "Measuring facial expression of emotion," *Dialogues in Clinical Neuroscience*, vol. 17, no. 4, pp. 457–462, Dec. 2015, doi: 10.31887/DCNS.2015.17.4/kwolf.
- [4] A. Moeini, H. Moeini, A. M. Safai, and K. Faez, "Regression facial attribute classification via simultaneous dictionary learning," *Pattern Recognition*, vol. 62, pp. 99–113, Feb. 2017, doi: 10.1016/j.patcog.2016.08.031.
- [5] J. Manfredonia et al., "Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder," *Journal of Autism and Developmental Disorders*, vol. 49, no. 1, pp. 279–293, Jan. 2019, doi: 10.1007/s10803-018-3757-9.
- [6] A. Dixit, A. K. Pal, S. Temghare, and V. Mapari, "Emotion detection using decision tree technique," *International Journal of Advance Engineering and Research Development*, vol. 4, no. 02, pp. 145–149, 2017, doi: 10.21090/ijaerd.24629.
- [7] F. Z. Salmam, A. Madani, and M. Kissi, "Facial expression recognition using decision trees," in *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGIV)*, Mar. 2016, pp. 125–130, doi: 10.1109/CGIV.2016.33.
- [8] S. Zafeiriou, C. Zhang, and Z. Zhang, "A survey on face detection in the wild: Past, present and future," *Computer Vision and Image Understanding*, vol. 138, pp. 1–24, Sep. 2015, doi: 10.1016/j.cviu.2015.03.015.
- [9] N. Wang, X. Gao, D. Tao, H. Yang, and X. Li, "Facial feature point detection: A comprehensive survey," *Neurocomputing*, vol. 275, pp. 50–65, Jan. 2018, doi: 10.1016/j.neucom.2017.05.013.
- [10] H. Shim, K. H. Cho, K. E. Ko, I. H. Jang, and K. B. Sim, "Multi-tasking deep convolutional network architecture design for extracting nonverbal communicative information from a face," *Cognitive Systems Research*, vol. 52, pp. 658–667, 2018, doi: 10.1016/j.cogsys.2018.08.006.
- [11] E. Uzun, S. P. H. Chung, I. Essa, and W. Lee, "rtCaptcha: A real-time CAPTCHA based liveness detection system," 2018, doi: 10.14722/ndss.2018.23253.
- [12] Y. Guo, Y. Xia, J. Wang, H. Yu, and R.-C. Chen, "Real-time facial affective computing on mobile devices," *Sensors*, vol. 20, no. 3, p. 870, Feb. 2020, doi: 10.3390/s20030870.
- [13] C. Follini, A. L. Cheng, G. Latorre, and L. F. Amores, "Design and development of a novel robotic gripper for automated scaffolding assembly," in *2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)*, Oct. 2018, pp. 1–6, doi: 10.1109/ETCM.2018.8580276.




- [14] S. J. Rosula Reyes, K. M. Depano, A. M. A. Velasco, J. C. T. Kwong, and C. M. Oppus, "Face detection and recognition of the seven emotions via facial expression: integration of machine learning algorithm into the NAO robot." in *2020 5th International Conference on Control and Robotics Engineering (ICCRE)*, Apr. 2020, pp. 25–29, doi: 10.1109/ICCRE49379.2020.9096267.
- [15] G. Muhammad, M. Alsulaiman, S. U. Amin, A. Ghoneim, and M. F. Alhamid, "A facial-expression monitoring system for improved healthcare in smart cities," *IEEE Access*, vol. 5, pp. 10871–10881, 2017, doi: 10.1109/ACCESS.2017.2712788.
- [16] M. Yang, Y. Ma, Z. Liu, H. Cai, X. Hu, and B. Hu, "Undisturbed mental state assessment in the 5G Era: A case study of depression detection based on facial expressions," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 46–53, Jun. 2021, doi: 10.1109/MWC.001.2000394.
- [17] Y. Zhuang *et al.*, "Facial weakness analysis and quantification of static images," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2260–2267, Aug. 2020, doi: 10.1109/JBHI.2020.2964520.
- [18] N. C. Suwanwela, "Stroke epidemiology in Thailand," *Journal of Stroke*, vol. 16, no. 1, p. 1, 2014, doi: 10.5853/jos.2014.16.1.1.
- [19] Y. Béjot *et al.*, "Impact of the ageing population on the burden of stroke: The dijon stroke registry," *Neuroepidemiology*, vol. 52, no. 1–2, pp. 78–85, 2019, doi: 10.1159/000492820.
- [20] K. Jaracz, B. Grabowska-Fudala, K. Górna, J. Jaracz, J. Moczko, and W. Kozubski, "Burden in caregivers of long-term stroke survivors: Prevalence and determinants at 6 months and 5 years after stroke," *Patient Education and Counseling*, vol. 98, no. 8, pp. 1011–1016, Aug. 2015, doi: 10.1016/j.pec.2015.04.008.
- [21] D. Hekmatpou, E. M. Baghban, and L. M. Dehkordi, "The effect of patient care education on burden of care and the quality of life of caregivers of stroke patients," *Journal of Multidisciplinary Healthcare*, vol. 12, pp. 211–217, Mar. 2019, doi: 10.2147/JMDH.S196903.
- [22] T. Von Arx, M. J. Nakashima, and S. Lozanoff, "The face-a musculoskeletal perspective," *Swiss Dental Journal Sso*, vol. 128, pp. 678–688, 2018.
- [23] A. Poston, "Human engineering design data digest," in *US Army Missile Command, Redstone Arsenal, AL*, Department of Defense Human Factors Engineering Technical Advisory Group Washington, 2000, p. 82.
- [24] F. Xu, J. Zhang, and J. Z. Wang, "Microexpression identification and categorization using a facial dynamics map," *IEEE Transactions on Affective Computing*, vol. 8, no. 2, pp. 254–267, Apr. 2017, doi: 10.1109/TAFFC.2016.2518162.
- [25] M. Takalkar, M. Xu, Q. Wu, and Z. Chaczko, "A survey: facial micro-expression recognition," *Multimedia Tools and Applications*, vol. 77, no. 15, pp. 19301–19325, Aug. 2018, doi: 10.1007/s11042-017-5317-2.

BIOGRAPHIES OF AUTHORS



Rawinan Praditsangthong    received the B.Sc. degree in computer science from the Rangsit University, Pathum Thani, Thailand, the M.Sc. degree in Information Technology from the Rangsit University, Pathum Thani, Thailand. Now, she is Ph.D. Candidate in Computer Science and Information Technology, Department of Mathematics and Computer Science, the Chulalongkorn University, Bangkok, Thailand. She is a lecturer in the College of Digital Innovation Technology, the Rangsit University. Her research areas are machine learning, facial recognition, and healthcare emotion recognition. She can be contacted at email: rawinan.p@student.chula.ac.th.



Pattarasinee Bhattarakosol    received the B.Sc. degree in mathematics from the Chulalongkorn, Bangkok, Thailand, the M.Sc. degree in applied statistics from the National Institute of Development Administration, Bangkok, Thailand, and the Ph.D. degree in computer science from the Wollongong University, Australia. She is a lecturer in Department of Mathematics and Computer science, Faculty of Science, Chulalongkorn University, Thailand. Her research areas are computer network and security, software architecture. She can be contacted at email: pattarasinee.b@chula.ac.th.