

Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods

Aya Hassouneh^{a,*}, A.M. Mutawa^a, M. Murugappan^b

^a Computer Engineering Department, College of Engineering and Petroleum, Kuwait University, Kuwait

^b Department of Electronics and Communication Engineering, Kuwait College of Science and Technology, Doha, Kuwait

ARTICLE INFO

Keywords:

Face emotion recognition

Virtual markers

LSTM

EEG emotion Detection

ABSTRACT

Real-time emotion recognition has been an active field of research over the past several decades. This work aims to classify physically disabled people (deaf, dumb, and bedridden) and Autism children's emotional expressions based on facial landmarks and electroencephalograph (EEG) signals using a convolutional neural network (CNN) and long short-term memory (LSTM) classifiers by developing an algorithm for real-time emotion recognition using virtual markers through an optical flow algorithm that works effectively in uneven lightning and subject head rotation (up to 25°), different backgrounds, and various skin tones. Six facial emotions (happiness, sadness, anger, fear, disgust, and surprise) are collected using ten virtual markers. Fifty-five undergraduate students (35 male and 25 female) with a mean age of 22.9 years voluntarily participated in the experiment for facial emotion recognition. Nineteen undergraduate students volunteered to collect EEG signals. Initially, Haar-like features are used for facial and eye detection. Later, virtual markers are placed on defined locations on the subject's face based on a facial action coding system using the mathematical model approach, and the markers are tracked using the Lucas-Kande optical flow algorithm. The distance between the center of the subject's face and each marker position is used as a feature for facial expression classification. This distance feature is statistically validated using a one-way analysis of variance with a significance level of $p < 0.01$. Additionally, the fourteen signals collected from the EEG signal reader (EPOC+) channels are used as features for emotional classification using EEG signals. Finally, the features are cross-validated using fivefold cross-validation and given to the LSTM and CNN classifiers. We achieved a maximum recognition rate of 99.81% using CNN for emotion detection using facial landmarks. However, the maximum recognition rate achieved using the LSTM classifier is 87.25% for emotion detection using EEG signals.

1. Introduction

One of the important ways humans display emotions is through facial expressions. Facial expression recognition is one of the most powerful, natural and immediate means for human beings to communicate their emotions and intentions. Humans can be in some circumstances restricted from showing their emotions, such as hospitalized patients, or due to deficiencies; hence, better recognition of other human emotions will lead to effective communication. Automatic human emotion recognition has received much attention recently with the introduction of IOT and smart environments at hospitals, smart homes and smart cities. Intelligent personal assistants (IPAs), such as Siri,

Alexia, Cortana and others, use natural language processing to communicate with humans, but when augmented with emotions, it increases the level of effective communication and human-level intelligence.

Due to the fast advancement of artificial intelligence (AI) and machine learning, its application is actively being used in many domains including spam detection, in which a spam classifier is utilized to rearrange email according to some specific standards and to move unwanted and unsolicited email to spam folder [1]. As well as significantly being used in data mining it is used for market analysis to support the large amount of data being produced every day and to detect fraud probability through the customer's fraud insurance [2]. For example, enhanced

* Corresponding author.

E-mail address: aia.hassouneh@gmail.com (A. Hassouneh).

<https://doi.org/10.1016/j.imu.2020.100372>

Received 23 March 2020; Received in revised form 8 June 2020; Accepted 9 June 2020

Available online 12 June 2020

2352-9148/© 2020 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Fraud Miner uses the clustering-based data mining method ‘Lingo’ in order to identify frequent patterns [3]. In addition, it is used for machine learning driven advances in medical space, such as revenue cycle management (i.e. payments) and understanding patient health through focusing on clinical data-rich environment [4,5].

Moreover, machine learning algorithms have played a significant role in pattern recognition and pattern classification problems, especially in facial expression recognition and electroencephalography (EEG), over the past several decades [6–8]. It has come of age and has revolutionized several fields in computing and beyond, including human computer interaction (HCI) [9,10]. Human computer interaction has become a part of our daily life. Additionally, emotion-related declaration is an important aspect in human interaction and communication due to its effective cost, reliable recognition and shorter computational time, among other advantages [10]. In other words, it may be a potential nonverbal communication media for creating diverse situations that can illustrate superior interaction with humans by close collaboration with human-human communication [10–13]. Facial expression analysis is an interesting and challenging problem and impacts important applications in many areas, such as human–computer interactions and medical applications. Several works are based on facial landmarks to extract some features to help in emotion detection [16]. presents a potential approach that uses 68 facial landmarks to detect three kinds of emotions in real time; negative, blank, and positive using one camera. Their proposed system can detect emotions using both 79 new features and 26 geometrical features (10 eccentricity features, 3 linear features and 13 differential features) from Ref. [17] with the average accuracy of 70.65%. Palestira et al. [18] computes 32 geometric facial (linear, eccentricity, polygonal and slope) features based on 20 facial landmarks for automatic facial expression recognition. Although considerable progress has been made, recognizing facial expressions with high accuracy has been performed in real-time systems for several applications, such as behavioral analysis, machine vision and video gaming. Thus, the expressions of humans can easily be “understood” by recent HMI systems [10,12–14]. Among the many methods studied in the literature, studies that made use of still images and emotion were perceived by measuring the dimensions of lips and eyes [10,11,15]. Biosensors, such as electromyograms (EMGs) and electroencephalograms (EEGs), have been used to perceive facial muscle changes and to conceive brain activities.

There have been efforts to develop multimodal information using facial expressions [11]. Facial recognition suffers limitations such as light intensity, face position, and background changes. EEG also suffers limitations such as noise, artifacts, placement of wired/wireless sensors, speech-noise and interferences, gesture-light intensity, and background changes. Therefore, combining both signals adds a new layer of the complexity of the environment, which meets the requirements of both modalities. Indeed, the computational complexity has been increasing exponentially in the case of multimodal systems. The multimodal strategy has a better performance in terms of the emotion recognition rate compared to the single modal strategy [19]. However, combining facial expressions with other modalities, such as speech signals, gestures and biosignals, will not efficiently detect emotions in dumb, deaf and paralyzed patients when developing intelligent human machine interface (HMI) assistive systems due to the above-mentioned issues. The existence of such HMI can be convenient for those who are totally disabled physically as well as patients with special needs, when seeking help. To date, most of the existing methods work offline, and will not be useful for real-time applications [10,11]. Hence, the present work is mainly focused on developing a multimodal intelligent HMI system that works in a real-time environment. It will recognize emotions using facial expressions and EEG based on machine learning and deep neural network methods. It aims to be a more reliable real-time emotion recognition system compared to systems in earlier works. To identify emotional facial expressions in real-time, the changes in facial reactions are measured. The facial action coding system (FACS) is a human

observer-based system designed by Ekman and Friesen [15], which is used to detect subtle changes in facial features and controls facial models by manipulating single actions, which are called action units (AUs). These AUs are considered a reference for identifying ten virtual markers for detecting six basic emotional facial expressions: sadness, anger, happiness, fear, disgust and surprise. Accordingly, if these facial muscle responses are studied and used as an instruction to pinpoint the expressions, the emotional state of humans can be perceived in real time [11]. Hence, the main objective of this study is to recognize six basic emotional expressions of the subjects using facial virtual markers and EEG signals. Therefore, the proposed system has less computational complexity (execution time, memory) and works in real-time applications. As it is using the markers based approach in which ten virtual markers are placed on the face rather than using the image pixels based approach that requires a longer computational time to detect the face [20], this will simplify the system design as well as reduce the computational time and system memory. As in real-time systems, face recognition is implemented on either by using the pixels of the image, or Haar-like features that efficiently use the AdaBoost cascade classifier, to detect any object in a given image or a video sequence involving human faces [21,22]. These Haar-like features based face detection can process 384×288 pixels of a face image in approximately 0.067 s [23]. They have used an AdaBoost cascaded classifier for a lesser computational time [20]. Most of the works in the literature focus on developing an offline emotion recognition system. In addition, this study aims to develop an algorithm for real-time emotion recognition using virtual markers through optical flow algorithm that works effectively in uneven lighting and subject head rotation (up to 25°), different backgrounds, and various skin tones. The distance feature computed from face virtual markers and the fourteen signals recorded from an EPOC + device are used to classify the emotional expressions using a convolutional neural network (CNN) and long short-term memory (LSTM) classifier. This will lead to achieve the aim of the system to help physically disabled people (deaf, dumb, and bedridden), in addition to its benefit for Autism children to recognize the feelings of others. Moreover, it can drive business outcomes and judge the emotional responses of the audience. Rather than helping to maximize learning, it has a good benefit in personalized e-learning. Three performance measures, namely, the mean emotional recognition rate, specificity, and sensitivity, are used to evaluate the performance of the classifiers.

2. Methods and materials

Fig. 1 shows the structure of the proposed system in this study. As illustrated in Fig. 1, we used two approaches to detect the subject's emotion: emotion detection using facial landmarks and emotion detection using EEG signals.

2.1. Emotion detection using facial landmarks

2.1.1. Facial landmarks database

Regarding emotion detection using facial landmarks, data collection occurred on the same day after informed consent was obtained from each subject volunteered in this study. Thus, two facial expression databases were developed: one with a total of 30 subjects (15 male, 15 female) for automated marker placement, and another with a total of 55 subjects (25 male, 30 female) for testing and validating the proposed system. All subjects were chosen from a mean age domain of 22.9 years. They were mixed males and females and healthy undergraduate university students with a history free from any cognitive, muscular, facial or emotional disorders. The subjects were requested to sit in front of a computer that had a built-in camera, and express six different emotional expressions (happiness, anger, fear, sadness, surprise, and disgust) in a video sequence for the purpose of data collection. They were requested to express particular emotion in a controlled environment (room temperature: 26° , lighting intensity: 50 lux; the distance between the

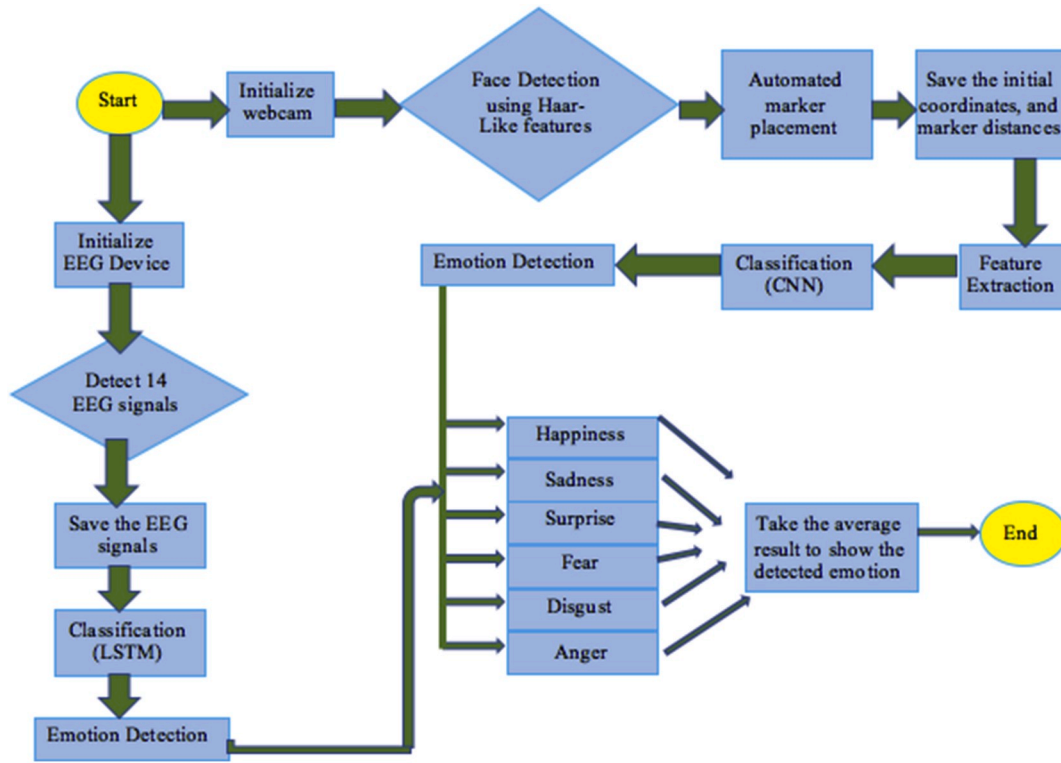


Fig. 1. System structure.

camera and subject: 0.95 m).

Facial expression for each emotion lasted 60 s and was performed once by each subject. Therefore, each subject required 6 min to express the emotions, as illustrated in Fig. 2. However, each subject took approximately 20 min as a total time, including instructions and baseline state. The total input virtual distances collected from each subject on all emotions is 2284 on average, saved in a format of CSV file, and the total virtual distance collected from all subjects on all emotions is 125,620.

2.1.2. Face feature extraction

In this study, an HD camera was used to capture the subjects' faces and create a grayscale image. This simplified the facial image process in

facial expression recognition. Then, by using a grayscale image, the subject's eyes are detected, and ten virtual markers (action units) are placed on the subject's face at defined locations using a mathematical model, as shown in Fig. 3.

The Lucas-Kande optical flow algorithm is used to transfer each virtual marker position to track its position during the subjects' emotional expression. The ten features are derived as the distance between each marker and the point as depicted in Fig. 4. In the current study, all the distance data were calculated using the Pythagorean theorem [24]. Then, they are stored in CSV format during the data acquisition process for further processing.

In Fig. 5, in the right mouth column, line m1 is the hypotenuse of a right triangle, wherein the line parallel to the x-axis is dx [the difference

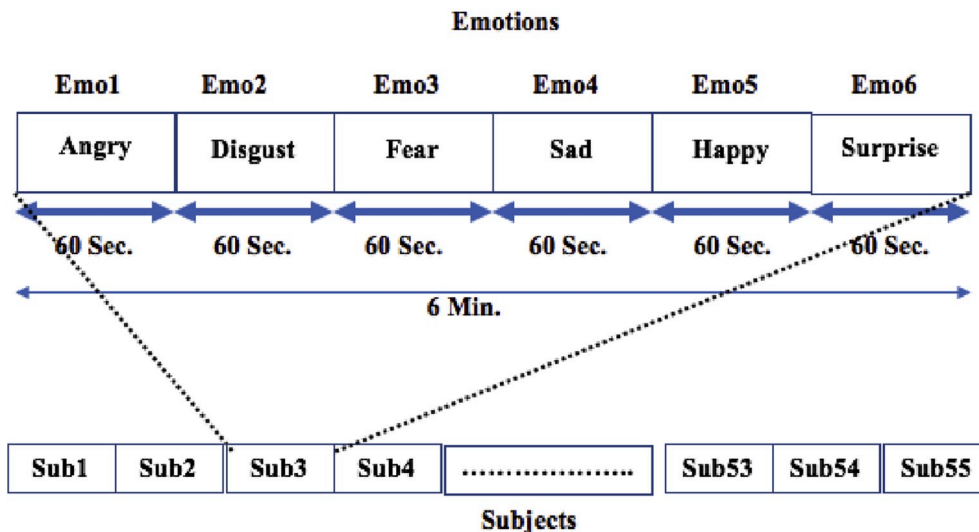


Fig. 2. Data acquisition protocol for facial emotion detection.



Fig. 3. Ten virtual locations using a mathematical model.

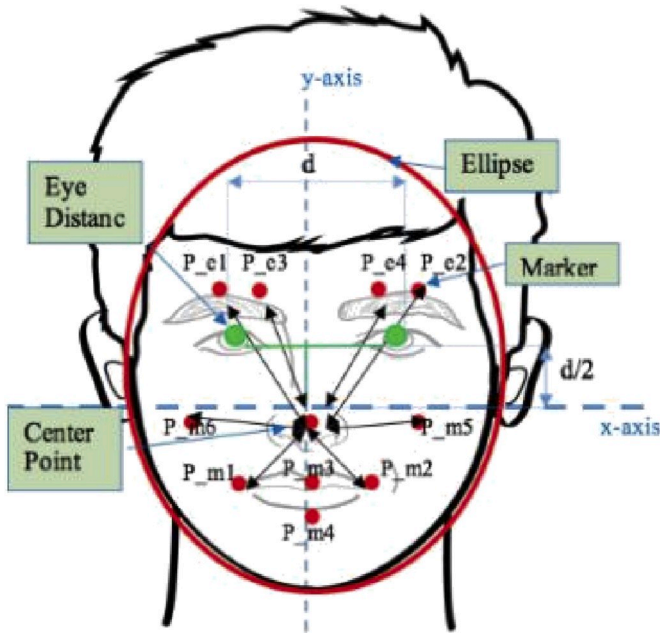


Fig. 4. Ten distances between each marker point and the center point.

between coordinates of p_{m1} ($x_{p_{m1}}$) and the point (x_c), and the line parallel to the axis is dy [the difference between y -coordinates of p_{m1} ($y_{p_{m1}}$) and the point (y_c)]. Thus, the formula for the computation of distance is given in Equation (1):

$$Distance(m1) = \sqrt{(X_c - X_p - m1)^2 + (Y_c - Y_p - m1)^2} \quad (1)$$

2.1.3. Facial landmarks classification

A convolutional neural network was used in our system to obtain improved facial emotion detection as it is applied to other computer fields such as face recognition [25] and object detection [26]. In

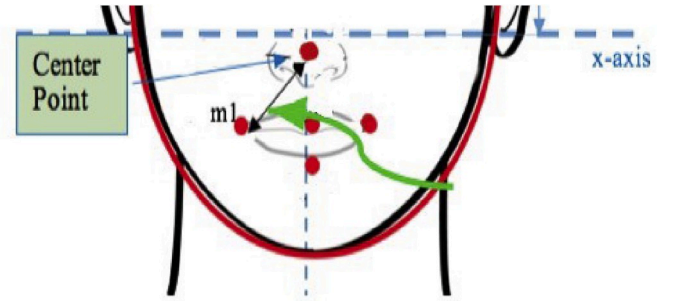


Fig. 5. The center point (C) and p_{m1} coordinates at the distance of $m1$.

addition, predictions are based on information given at a particular time [27].

Fig. 6 shows the network structure that is used for emotion detection using facial landmarks. This network takes an input image and attempts to predict the output emotion. It has eight stages, including convolutions, pooling and fully connected layers with rectified linear unit (ReLU) operations, which preserve good quality while making convergence much faster [28]. The number of filters was 32, 64, and 128 with a filter size of 5×5 for the convolutional layers, and the number of output nodes in the fully connected layer was 6 with the “Adam” optimizer and a dropout rate of 0.3.

2.2. Emotion detection using EEG signals

2.2.1. EEG database

Regarding emotion detection using EEG signals, data collection took place on the same day after written consent was obtained from each subject volunteered in this study.

For the purpose of EEG data collection, a video was created from six main clips to obtain the best reaction of the brain in terms of electrical activities. In the beginning, video clips were created using emotional images from the International Affective Picture System [29] and music [30]. Those clips were supposed to elicit the six emotions to be

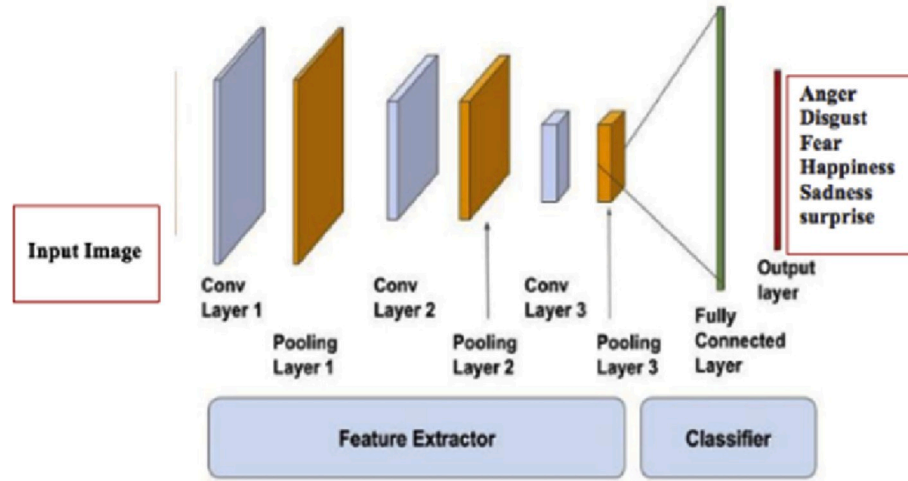


Fig. 6. Facial landmarks system structure.

predicted. After that, the video clips were tested on a small test sample to confirm the elicited emotions by prompting the volunteers to vote for the emotion they felt while watching the clips. Accordingly, other video clips were selected based on certain criteria, one of which must have a minimum of one million views on YouTube. Each clip was 2 min long and referred to one emotion. Therefore, the total length of the video, including the gaps between the clips, was 13 min and 10-s. After the video clips were determined, all the subjects were asked to watch the video and classify the emotional categories following their proto-think and explain the effect that they consciously believe they caused. In this way, EEG raw data of the 14 channels were collected at a sampling rate of 128 Hz using the Emotiv EPOC device [31]. Thus, an EEG signal database was developed with a sum of 19 subjects (7 males and 12 females) for EEG signal investigation with a total of 1,700,000 records of signal data. All subjects were chosen from a mean age domain of 22.9 years. Collecting data was conducted in a laboratory environment and required approximately 1 h to perform. All subjects were healthy undergraduate university students with a history free from any cognitive, muscular, facial, or emotional disorders.

EEG recording was performed using the EPOC + interface. It is connected wirelessly to the Emotiv EPOC device and records EEG signals that are coming from the brain while the subject is watching the video. As each emotion's video clip lasted 2 min with a 10 s gap between the videos and was performed once by each subject, each subject required 13 min and 10 s to record the EEG signals, as illustrated in Fig. 7. However, each subject took approximately 1 h as a total time, including instructions and a baseline state. Then, the same interface was used to

export the recorded signals into CSV files and feed them to a Python system to filter using an infinite impulse response (IIR) filter before classification and training [32].

2.2.2. EEG signal preprocessing

Due to the artifacts that come from different sources during EEG signal recording, such as eye blinks, eye movements, muscle movements, respiration, and sweat, a process is needed to remove this unwanted information from the EEG signals that can cause significant distortion. First, to reduce the artifact effects, the EEG signal voltage was taken in an amplitude range of $\pm 85 \mu\text{V}$. Second, a sixth-order Butterworth infinite impulse response (IIR) filter utilized in Ref. [32] with a cut-off frequency of [1–49] Hz was used to remove noise from the EEG signals.

2.2.3. EEG signal classification

A long short-term memory (LSTM) network model was used for training the affective model and obtaining improved EEG signal emotion detection. LSTM is a special type of artificial recurrent neural network (RNN) architecture. It is used in the field of deep learning and well-suited to time series data to classify, process, and make predictions. It can process entire sequences as analog data in this study.

Nineteen participants were asked to watch the same video that contained six different parts, to recognize the emotion elicited from these videos. Then, the collected data (EEG raw data) were fed to the proposed model, as shown in Fig. 8. The proposed model achieved the highest accuracy among three models (conventional, fully connected

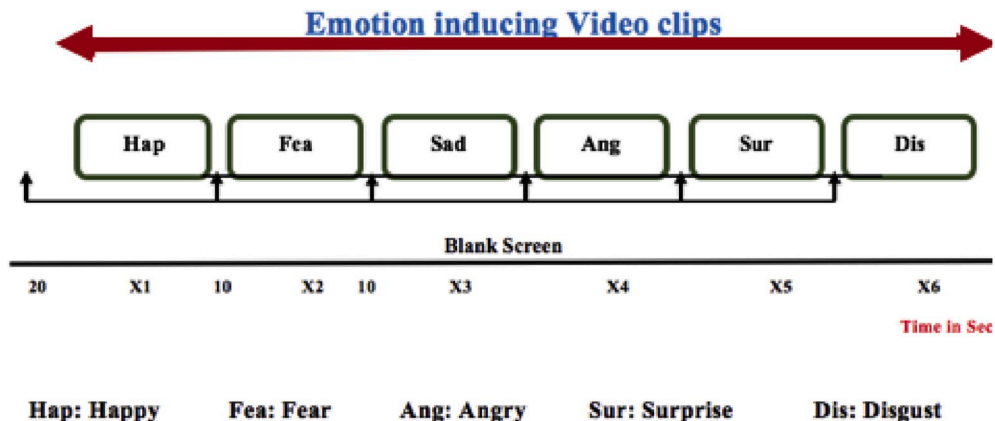


Fig. 7. Data acquisition protocol for EEG emotion detection.

layers, grid search). Each part of the video was segmented into 20 segments with a 6-s window. Thus, each participant had an array of EEG data records with their labels. The array of labels represents the six emotions 'happy, fear, anger, sad, surprise and disgust' of each video performed by each participant. Each participant produced approximately 118,000 records of EEG raw data from 14 EEG channels for the six videos.

Fig. 8 shows the proposed deep learning neural network model that achieves the highest accuracy. It consists of a fully connected embedding layer and three LSTM layers with 128, 64, and 32 neurons, two dropout layers, and a dense layer. The LSTM and dropout layers are used to learn features from raw EEG signals. However, the dropout layer is used to reduce the overfitting by preventing too many units from 'coadapting'. Finally, the dense layer that uses the Softmax activation function is used for classification.

The model is trained on 70% of the EEG records using 3-fold cross-validation and tested on 30% of them. One hundred epochs are used for each cross-validation iteration. The Adam optimizer is used in the training process with a 0.001 learning rate.

Additionally, the grid search method and fivefold cross-validation strategies are used to tune the dimension of features, the threshold value, and the learning rate.

3. Data validation

System validation aims to evaluate the accuracy after development. It was tested by collecting data especially for the testing.

3.1. Facial landmarks database validation

For emotion detection using facial landmarks, two facial expression databases were developed, one with a total of 30 subjects (15 male, 15

female) for automated marker placement, and another one with a total of 55 subjects (25 male, 30 female) for testing and validating of the proposed system. They were requested to sit in front of a computer that has a built in Camera and express six different emotional expressions (happiness, anger, fear, sadness, surprise, and disgust) in a video sequence for the purpose of data collection. They were requested to express particular emotion in a controlled environment (room temperature: 26°, lighting intensity: 50 lux; the distance between the camera and subject: 0.95 m). Thus, the proposed method achieves 99.81% as a highest accuracy.

3.2. EEG signals database validation

The facial landmarks data for the 19 subjects are given into the three-fold cross-validation method to split the set of features into training and testing sets. Thus, the proposed method achieves the higher emotional detection rate of 87.25%.

4. Experimental results and discussion

For emotion detection using facial landmarks, data were collected from 55 subjects for testing; their ages were between 20 and 25. They are undergraduate students (25 males, 30 females). Then, the accuracy was found at 100 epochs for the collected data in both cases of normalized and not normalized data, as illustrated in Table (1).

The collected facial landmark data were normalized using Equation (2) to make all data values between [0, 1].

$$Z_i = (X_i - \text{Min}(X)) / (\text{Max}(X) - \text{Min}(x)) \quad (2)$$

where Z_i is the i th normalized data.

Both models were run till 300 epochs to check the shape of the curve and obtain the highest accuracy for emotion detection. As shown in Fig. 9, emotion detection using facial landmarks has a direct relationship till 300 epochs. Therefore, the system can recognize emotions in 99.81% of facial landmarks using the proposed model at 300 Epochs. Whereas, the proposed model of EEG signals has a direct relationship between accuracy and number of epochs till 100, and the curve stops increasing after 100, as shown in Fig. 10. As a result, the highest emotion recognition accuracy of EEG signals is 87.25% at 100 Epochs.

Additionally, the sensitivity, specificity, F-score, and accuracy are calculated for each class by using the following calculated confusion matrix (Fig. 11) for the six classes for emotion detection using facial landmarks. Each class is used against all classes in order to find those performance factors related to it.

The results in Table (2) are calculated using the following equations:

$$\text{precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{specificity} = \text{TN} / (\text{FP} + \text{TN})$$

$$\text{F-score} = 2 * \text{TP} / (2 * \text{TP} + \text{FP} + \text{FN})$$

Where; TP: True Positive, FP: False Positive, TN: True Negative, FN: False Negative.

Table (2) shows the performance factors for each emotion; 0:Anger, 1:Disgust, 2:Fear, 3:Sad, 4:Smile and 5:Surprise. It shows the perfect precision and the proportion of actual positives and negatives of

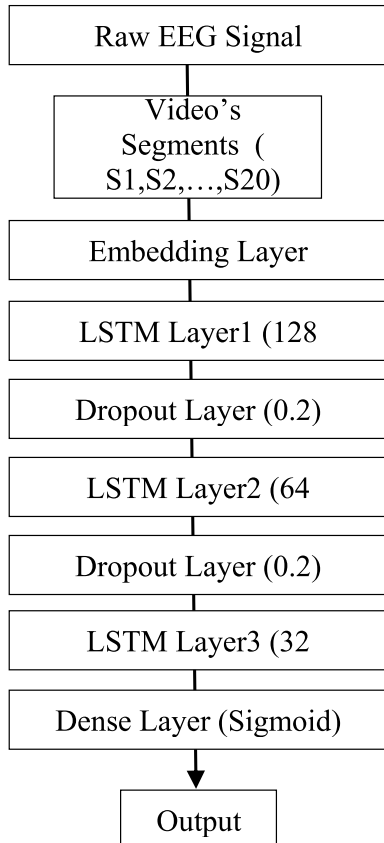


Fig. 8. NN model to handle and train EEG signals.

Table 1
Facial landmark accuracies.

Data Sets	Avg. Accuracy
Training Data from 30 Subjects (Not Normalized)	89.04%
Training Data from 30 Subjects (Normalized)	96.93%
Testing Data from 55 Subjects (Normalized)	93.02%

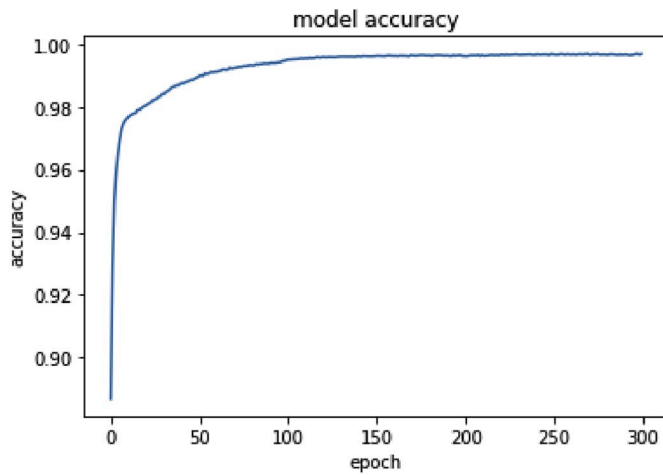


Fig. 9. Accuracy vs. #Epochs (Facial Landmarks).

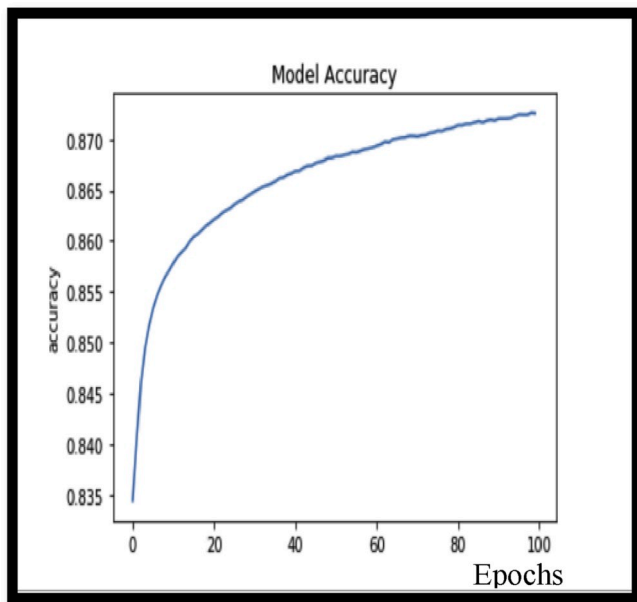


Fig. 10. Accuracy vs. #Epochs (EEG).

emotions that are accurately identified. Thus, the collected data can be used as a benchmark for other researchers. Our system is based on just 10 virtual distances for facial landmarks and 14 EEG raw data for the EEG signals. In other words, it increased the performance of the system as the number of training facial landmark features is fewer than in previous work, as shown in Table (2). Table (3) shows that [33–37] authors use more than 10 virtual markers and those markers are manually placed on the face to detect emotions. They used the CK + database [38], and the maximum accuracy they achieved was 93.2 [37]. Whereas, our system places exactly 10 virtual markers automatically to detect emotions. It is trained using our own created database, and the maximum accuracy was 99.81%. This means that the system detects emotions using facial landmarks with high accuracy and without too many features.

Some of the major aspects that can affect the results found by the system are the user's accuracy in his/her feelings during the watched video, especially in the detection of EEG signals. However, in the detection using facial landmarks, it was more accurate since it depends on external physical features. Additionally, the external environment had an effect on the accuracy.

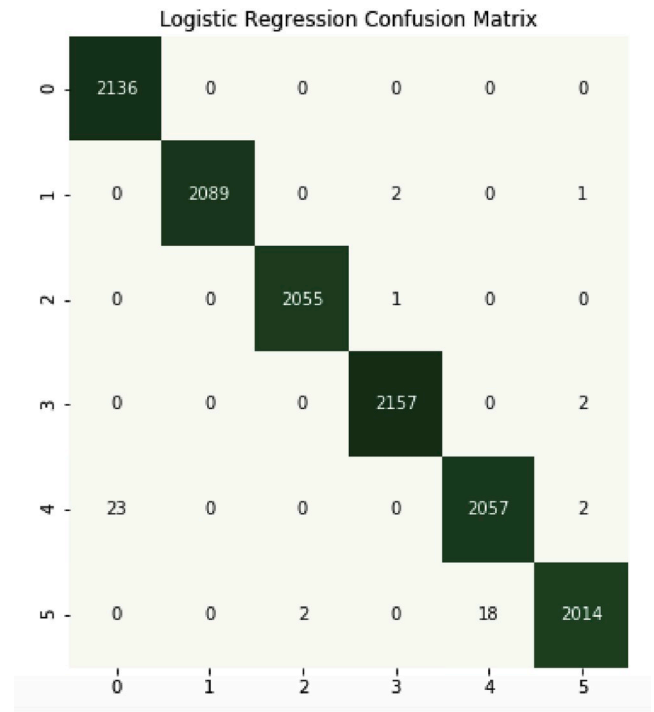


Fig. 11. Confusion matrix.

Table 2

Facial landmarks performance factors.

	0	1	2	3	4	5
Precision (%)	98.9	100.0	99.9	99.9	99.1	99.8
Sensitivity (%)	100	99.8	99.9	99.9	98.8	99.0
Specificity (%)	99.8	100	99.9	99.9	99.8	99.9
F-score (%)	99.5	99.9	99.9	99.9	98.9	99.4
Acc (%)	99.8	100	99.9	99.9	99.1	99.8

Many exhaustive tests were conducted to minimize the errors in obtaining the desired accuracy. Tests were performed to obtain a reliable classifier, normalization method, and an optimal number of markers with the best-proposed features to reduce the errors. The highest recognition rate obtained in this research was 99.81% for emotion detection using facial landmarks with the 10 virtual markers. However, the highest accuracy for emotion detection using EEG signals was 87.25%, which can be improved by collecting more data from more subjects and finding techniques to extract more features from the EEG signals.

5. Conclusion and future work

An algorithm for real-time emotion recognition using virtual markers through an optical flow algorithm has been developed to create a real-time emotion recognition system with less computational complexity (execution time, memory) using facial expressions and EEG signals. This algorithm works effectively in uneven lightning and subject head rotation (up to 25°), different backgrounds, and various skin tones. The system aims to help physically disabled people (deaf, dumb, and bedridden), in addition to its benefit for Autism children to recognize the feelings of others. Moreover, it can drive business outcomes and judge the emotional responses of the audience. Rather than helping to maximize learning, it has a good benefit in personalized e-learning. The system can recognize six emotions in real time for facial landmarks and in offline settings for EEG raw data. Users of the system need to wear an EPOC₊ headset and are faced in front of a camera to record the EEG raw

Table 3

Accuracy of earlier and present work on emotion detection using facial landmarks.

\Authors	[33]	[34]	[35]	Related Work	[36]	[37]	Our Work
#Virtual Markers	22	77	68	22	68		10
Marker Placement		Manual			Auto		
Database	CK+[38]		6 different ready DBs	Own DB	CK+[38]		Own Database
Classifier	SVM		DNN	Neural Network	FURIA		CNN
Acc. (%)	86.0	85.8	93.2	84.27	83.2		99.8

data wirelessly and collect the ten virtual landmarks placed on the subject's face.

The results show that the system can recognize emotion in 99.81% of facial landmarks and 87.25% of EEG signals. The cases that were used to collect the data were performed at Kuwait University. It was difficult to collect data from many subjects due to the student's schedules and timings. Thus, only a few subjects were available for collecting the data.

For future work, the system's precision and accuracy can be improved by collecting more data from more subjects. Additionally, techniques can be used to extract more features from EEG signals [14]. In addition to improving the system techniques, putting the subjects in real situations to express the exact feelings can help to improve the system's accuracy for the EEG.

Ethical statement

Each named author assures that there are no financial or any personal relationships (potential competing interests, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding) with other people or organizations that could inappropriately influence their work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The first two authors would like to thank the continuous support of the graduate school of Kuwait University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.imu.2020.100372>.

References

- [1] Dada EG, Bassi JS, Chiroma H, Abdulhamid SM, Adetunmbi AO, Ajibuwa OE. Machine learning for email spam filtering: review, approaches and open research problems. *Heliyon* 2019;5(6):e01802. <https://doi.org/10.1016/j.heliyon.2019.e01802>.
- [2] Xie M. Development of artificial intelligence and effects on financial system. *J Phys Conf* 2019;1187:032084. <https://doi.org/10.1088/1742-6596/1187/3/032084>.
- [3] Hegazy O, Soliman OS, Salam MA. A machine learning model for stock market prediction. *Int J Comput Sci Telecommun* 2014;4(12):16–23.
- [4] Beckmann JS, Lew D. Reconciling evidence-based medicine and precision medicine in the era of big data: challenges and opportunities. *Genome Med* 2016;8(1):134–9.
- [5] Weber GM, Mandl KD, Kohane JS. Finding the missing link for big biomedical data. *Jama* 2014;311(24):2479–80.
- [6] Loconsole C, Chiaradia D, Bevilacqua V, Frisoli A. Real-time emotion recognition: an improved hybrid approach for classification performance. *Intelligent Computing Theory* 2014:320–31.
- [7] Huang X, Kortelainen J, Zhao G, Li X, Moilanen A, Seppänen T, Pietikäinen M. Multi-modal emotion analysis from facial expressions and electroencephalogram. *Comput Vis Image Understand* 2016;147:114–24. <https://doi.org/10.1016/j.cviu.2015.09.015>.
- [8] Raheel A, Majid M, Anwar SM. Facial expression recognition based on electroencephalography. In: 2019 2nd international conference on computing, mathematics and engineering technologies (iCoMET), Sukkur, Pakistan; 2019. p. 1–5.
- [9] Vassilis S, Herrmann Jürgen. Where do machine learning and human-computer interaction meet?. 1997.
- [10] Keltner D, Ekman P. In: Lewis M, Haviland Jones JM, editors. *Facial expression of emotion, hand book of emotions*. New York: Guilford Press; 2000. p. 236–49.
- [11] Ekman P. Darwin and facial expression: a century of research in review. *United State Of America: Academic Press* Ishk; 2006. p. 1973.
- [12] Ekman P, Friesen WV. Constants across cultures in the face and emotion. *J Pers Soc Psychol* 1971;17(2):124.
- [13] Ekman P. Darwin and facial expression: a century of research in review. *United State Of America: Academic Press* Ishk; 2006. p. 1973.
- [14] Ekman P, Friesen WV, Ancoli S. Facial signs of emotional experience. *J Pers Soc Psychol* 1980;39:1123–34.
- [15] Ekman P, Friesen WV, Ancoli S. Facial signs of emotional experience. *J Pers Soc Psychol* 1980;39:1123–34.
- [16] Nguyen BT, Trinh MH, Phan TV, Nguyen HD. An efficient real-time emotion detection using camera and facial landmarks. In: 2017 seventh international conference on information science and technology (ICIST); 2017. <https://doi.org/10.1109/icist.2017.7926765>.
- [17] Loconsole C, Miranda CR, Augusto G, Frisoli A, Orvalho V. Real-time emotion recognition novel method for geometrical facial features extraction. *Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP)* 2014:378–85.
- [18] Palestra Giuseppe, Pettinichio Adriana, Del Coco Marco, Carcagn Pierluigi, Leo Marco, Distanto Cosimo. Improved performance in facial expression recognition using 32 geometric features. In: *Proceedings of the 18th international conference on image analysis and processing. ICIAP*; 2015. p. 518–28.
- [19] Zhang J, Yin Z, Cheng P, Nichele S. Emotion recognition using multi-modal data and machine learning techniques: a tutorial and review. *Information fusion*. 2020.
- [20] Wilson PI, Fernandez J. Facial feature detection using Haar classifiers. *J. Comput. Small Coll.*, ročnik 21, č. 2006;4. s. 127–133, ISSN 1937-4771.
- [21] Zhao G, Pietikäinen M. Dynamic Texture Recognition Using Volume Local Binary Patterns. In: Vidal R, Heyden A, Ma Y, editors. *Dynamical Vision. WDV 2006, WDV 2005. Lecture Notes in Computer Science*, vol. 4358. Berlin, Heidelberg: Springer; 2007.
- [22] Das PK, Behera HS, Pradhan SK, Tripathy HK, Jena PK. A modified real time A* algorithm and its performance analysis for improved path planning of mobile robot. In: *Computational intelligence in data mining, springer India*, vol. 2; 2015. p. 221–34.
- [23] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. 2001 [20] aeed, A. and Al-Hamadi, A. (2015). Boosted human head pose estimation using kinect camera. In *International Conference on Image Processing (ICIP)*, pages 1752–1756. IEEE.
- [24] Sally Judith D, Paul Sally. "Chapter 3: Pythagorean triples". *Roots to research: a vertical development of mathematical problems. American Mathematical Society Bookstore*; 2007. p. 63. ISBN 0821844032.
- [25] Sun Y, Chen Y, Wang X, Tang X. Deep learning face representation by joint identification-verification. In: *Proc. Adv. Neural inf. Process. Syst.*; 2014. p. 1988–96.
- [26] Ouyang W, Wang X, Zeng X, Qiu S, Luo P, Tian Y, Li H, Yang S, Wang Z, Loy C-C, Tang X, Dong C, Loy CC, He K. Deepid-net: deformable deep convolutional neural networks for object detection. In: *In proc. IEEE conf. Comput. Vis. Pattern recogn.*; 2015. p. 2403–12.
- [27] Tang X. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell* 2016;38(2):295–307. <https://doi.org/10.1109/tpami.2015.2439281>.
- [28] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. In: *Proc. Adv. Neural Inf. Process. Syst.* 2012: 1097–105.
- [29] Lang PJ, Bradley MM, Cuthbert BN. *International affective picture system (IAPS): affective ratings of pictures and instruction manual*. Gainesville, FL: University of Florida; 2008. Technical Report A-8.
- [30] Bhattacharya J, Lindsen JP. Music for a brighter world: brightness judgment bias by musical emotion. *PLoS One* 2016;11(2):e0148959. <https://doi.org/10.1371/journal.pone.0148959>.
- [31] <https://www.emotiv.com/>.
- [32] Sangaiah AK, Arumugam M, Bian G-B. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. *Artif Intell Med* 2019; 101788. <https://doi.org/10.1016/j.artmed.2019.101788>.
- [33] Michel P, El Kaliouby R. Real time facial expression recognition in video using support vector machines. In: *Proc. 5th int. Conf. On multimodal interfaces*; 2003. p. 258–64. Vancouver.

- [34] Suk M, Prabhakaran B. Real-time mobile facial expression recognition system—a case study. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*; columbus, OH, USA; 24–27 June 2014. p. 132–7.
- [35] Jeong M, Ko BC. Driver's facial expression recognition in real-time for safe driving. *Sensors* 2018;18:4270.
- [36] Magdin Martin, Prikler F. Real time facial expression recognition using webcam and SDK affectiva. *International Journal of Interactive Multimedia and Artificial Intelligence* 2017;1. <https://doi.org/10.9781/ijimai.2017.11.002>. InPress.
- [37] Bahreini K, van der Vegt W, Westera W. A fuzzy logic approach to reliable real-time recognition of facial emotions. *Multimed Tool Appl* 2019:1–24.
- [38] Lucey P, et al. The Extended Cohn-Kanade Dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: *2010 IEEE computer society conference on computer vision and pattern recognition – workshops*. IEEE; 2010. p. 94–101.