Tech Science Press

Suicide Ideation Detection of Covid Patients Using Machine Learning Algorithm

R. Punithavathi[1,*], S. Thenmozhi[2], R. Jothilakshmi[3], V. Ellappan[4] and Islam Md Tahzib Ul[5]

[1]M Kumarasamy College of Engineering, Karur, Tamil Nadu, India
[2]Dayananda Sagar College of Engineering, Bangalore, Karnataka, India
[3]R.M.D Engineering College, Kavaraipettai, Tamil Nadu, India
[4]Adama Science and Technology University, Adama, Ethiopia
[5]Dhaka International University, Dhaka, Bangaladesh
*Corresponding Author: R. Punithavathi. Email: r.punithavathi@gmail.com

Abstract: During Covid pandemic, many individuals are suffering from suicidal ideation in the world. Social distancing and quarantining, affects the patient emotionally. Affective computing is the study of recognizing human feelings and emotions. This technology can be used effectively during pandemic for facial expression recognition which automatically extracts the features from the human face. Monitoring system plays a very important role to detect the patient condition and to recognize the patterns of expression from the safest distance. In this paper, a new method is proposed for emotion recognition and suicide ideation detection in COVID patients. This helps to alert the nurse, when patient emotion is fear, cry or sad. The research presented in this paper has introduced Image Processing technology for emotional analysis of patients using Machine learning algorithm. The proposed Convolution Neural Networks (CNN) architecture with DnCNN preprocessing enhances the performance of recognition. The system can analyze the mood of patients either in real time or in the form of video files from CCTV cameras. The proposed method accuracy is more when compared to other methods. It detects the chances of suicide attempt based on stress level and emotional recognition.

1 Introduction

In past few years, Coronavirus and its variants have affected the world socially and economically. Emotion recognition systems analyse vocal tone, facial expressions using machine learning and deep learning algorithms. The Emotions such as fear, happiness, sadness, angry and surprise can be detected using the recent methods. The Conventional facial emotion recognition (FER) system consists of stages namely: image acquisition, pre-processing, feature extraction, classification or regression. Fig. 1 represents a conventional FER system. The first step is pre-processing where the input image obtained

from several peoples are processed. The pre-processed data will be send to detection unit. The detection will be made on intensity, smoothing and data augmentation (DA).
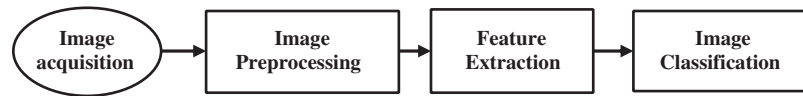

Figure 1: Conventional FER system

In second step, pre-processed features will be extracted by Actions Units (AUs). In final step, features are transferred to the classifier. The classification unit consists of Support Machine Vectors (SVMs) or Convolutional Neural Networks (CNNs). The proposed method detects the patient's emotion and suicide ideation based on the stress and emotional variations on Temperamental Dysregulation. The stress level is categorized based on different classes.

2 Literature Survey

Wang et al. 2018 analyzed the depression patient's videos and facial features are extracted. From the located facial features, the classification is done using a SVM method [1]. With the use of Image processing technology Ketcham et al. 2018 recognized emotions of patients in real time by the use of video files from the CCTV camera [2]. Shojaeilangari et al. 2015 presented an extreme sparse learning method [3] which uses the dictionary and nonlinear classification model. This method gives accurate results even when the input data is affected by noise signals. Compound facial emotion analysis method was presented in literature by the use of iCV-MEFED data set [4]. Yan et al. 2016 described an emotion recognition method from facial expression and sparse kernel reduced-rank regression (SKRRR) fusion technique [5]. Here an openSMILE algorithm is used for extracting features from speech. A scale invariant feature transform for facial analysis is employed. Facial expressions are synthesized by Mumenthaler et al. 2020 [6]. Zhang et al. 2019 developed an algorithm based on deep learning framework named as spatial-temporal recurrent neural network (STRNN) [7] which uses multidirectional recurrent neural network (RNN) for emotion recognition. Chakraborty et al. 2009 made use of external stimulus to excite certain emotions in human [8] such as eye opening, mouth opening, eyebrow length etc., The image was analyzed by segmentation method which divides the image into separate regions of interest. A hybrid cascaded Gaussian mixture model and deep neural network based classifier [9] is presented by Shahin et al. 2019, in which the emotions are contrasted with support vector machine and multilayer perceptron. Tzirakis et al. 2017 described an emotion recognition method with audio and video systems [10]. The convolutional neural network is used for recognizing audio and deep residual network for video recognition. In gaming environment, the player's emotions are detected and recognized with their heart beat. The use of Bidirectional long- and short-term memory (Bi-LSTM) network for heart beat recognition and CNN for facial recognition was done [11]. Further Du et al. 2020 presented a SOM-BP network which combines the heart rate and facial features to detect the player's emotion. Li et al. 2019 [12] presents a maximization algorithm to express compound or mixture of emotions. For this multi-expression analysis, a deep CNN is used. It increases the detection capability. Hossain et al. 2017 described the usage of Bandlet transform to analyze the features in images for facial expression detection [13]. The method is applied on images captured from the camera. For feature selection, Kruskal-Wallis method is used and the dominant bins are sent to Gaussian model for further classification of emotions. A sampling approach method based on human perceptual psychology [14] is proposed by Cruz et al. 2014 which extracts the temporal dynamics of facial emotions. It is tested in Audio/Visual Emotion Challenge event and found that there is considerable increase in accuracy of detection. Wang et al. 2015 developed a tensor independent color space (TICS) [15] for identification of

micro-expressions. The Region of interests is identified, and dynamic texture histograms are estimated for every ROI. It is found that CIELab and CIELuv are useful for identification of micro-expressions. Xia et al. 2017 described a Deep Belief network (DBN) for acoustic emotion recognition [16]. Features are extracted and classified with the use of support vector machine. Tariq et al. 2012 addressed an emotion recognition method [17] based on subject dependent and subject-independent emotion recognition. Emotions recognition are person-specific and person-independent and that can be done either manually or automatically. A light weight neutral emotion classification [18] is presented by Chiranjeevi et al. 2015, using statistical texture model. The neutral appearance is identified for emotion recognition. It decreases the computational complexity. Zhang et al. 2016 developed a emotion recognition method [19] for extracting multiscale features using biorthogonal wavelet. For classification, fuzzy multiclass support vector machine is used. Wang et al. 2010 applied four different methods [20] namely Eigen face approach, fisher face approach, Active Appearance model (AAM), and AAM based LDA method on visible and infrared facial expression database for emotion recognition. The relationship between facial temperature and emotion are studied using statistical methods. A new CNN model is presented by Zhang et al. that to integrate the content information of the deep network [21]. An audio-visual emotion recognition based on machine learning is presented by Seng et al. 2018. For visual path BDPCA and LSLDA [22] are used. Prosodic features and spectral features are used to recognize emotions. FER system based on hierarchical deep learning scheme [23] is developed by Kim et al. 2019. Here the features extracted are fused with the geometric features. This method employs autoencoder technique for emotion detection. The emotion recognition was carried out with the absence of sequence data. Wang et al. presents a method to analyze the multimodal spontaneous facial expression database which contains natural visible and infrared facial expressions (NVIE). The method can recognize emotions [24] with respect to their gender and can recognize the existence of multiple emotions/expressions. It is found by Wang et al. that there is a difference in emotion recognition for various environment. Zhang et al. 2019 described a method on the basis on CNN in which the facial image is first normalized. The edge of the faces is extracted to get implicit features [25] by the use of maximum pooling method. Finally, the image is classified by the use of Softmax classifier. Mistry et al. 2017 uses evolutionary particle swarm optimization (PSO) based feature optimization technique [26] for facial emotion recognition. He also presented a novel method of feature extraction using mGA-embedded PSO algorithm. Finally for recognition of the facial expression, different classifiers are employed. Meléndez et al. 2020 done a study on variation of emotion of adults due to covid-19 [27]. The study focuses on finding the basic emotions of two groups of adults one confined and other unconfined during covid-19. The results show that there is a high variation in emotions with increasing sadness, depression and reduced happiness.

3 Background Methodology

In medical records, it's been identified that the possible propositions for suicide during Covid situation depends on the factors like insufficient integration, stress, frustration, extreme social regulation, enmeshed social regulations, thwarted belongingness, loneliness etc,. Fig. 2 presents various factors inducing suicide ideations during Covid times. Some of the methods used in clinical diagnosis for suicide ideations detection is shown in Fig. 2.

3.1 Pre-Processing

The image quality and amount of noise present determines the face recognition rate. But noise can be removed when pre-processing is carried out before extraction process. During preprocessing the image is converted into gray color to obtain a normalized uniform intensity so that color grading process can be carried out. The Effective color grading processes such as cropping is utilized in face detection. After that

images are filtered with low pass filter. In this step, unwanted regions in image are eliminated. The Target regions are highlighted using Viola–Jones algorithm.
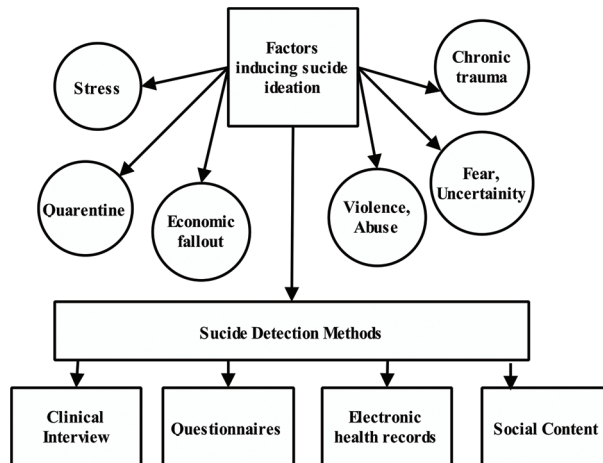

Figure 2: Factors and detection methods for suicide ideation

3.2 Viola and Jones Based Face Detection

The Viola-Jones algorithm scans the sub window to detect the faces in given input image [28]. In few cases, image rescaling of given input with fixed size is done. Scale invariant detectors are used to rescale the images in minimum time. The Rescaling and feature extraction are done using Haar wavelets.

3.3 Feature Extraction and Dimensionality Reduction

3.3.1 Histograms of Oriented Gradients (HOG)

Navneet Dalal and Bill Triggs (2005) proposed a framed histogram based oriented gradient to extract dense features from all regions present in the image [29]. For object detection, the proposed feature extraction does not use segmentation techniques to extract the features. HOG attempts to capture the pattern by extracting data based on gradients. The pattern divides the cells into 8×8 pixels. Each cell consists of gradients to normalize the histogram. This represents 1-D array of histogram known as descriptor.

3.3.2 Dimensionality Reduction

Even though the Data with larger dimension can provide better object description, the recognition speed becomes slow. To speedup the recognition process, the data/dimensionality reduction is done. Dimensionality reduction methods such as PCA or Bag of Words are utilized. In facial recognition application, PCA performs an ortho-normal transformation for converting correlated variables into linearly uncorrelated variables. The Principal components are represented as symmetric covariance matrix with Eigen values. For analysis, ORL image database are utilized. The wavelet transform algorithm has been utilized to reduce noises in the input images.

3.3.3 Principal Component Analysis: PCA

The Principal Component Analysis (PCA) has the capability to reduce the data quantity, but it does not minimize the accuracy [30]. Each sample variance is increased between each covariance. Karkpearson described visual recognition with size reduction for description vector in SIFT and SURF. The Eigen space describes the declaration syntax of input samples with the training samples.

3.4 Classification

In Classification stage, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) classifiers are utilized.

3.4.1 K-Nearest Neighbors (KNN)

Each classifier performs its training with random images. The K-Nearest Neighbors classifiers are utilized in machine learning for predicting the test sample. This Classifier encloses a nonparametric technique to classify unknown objects by identifying its closest neighbors [31]. It is most common classifier utilized in pattern recognition for classification. Based on k value, the KNN classifier evaluation gets varies. The Training samples are collected when it's nearest to test sample. In next stage, average operation will be performed on them.

3.4.2 Multi-Class SVM Problem

To solve multi class's analysis, One-vs-One approach, One-vs-All approach and Weston and Watkins' approaches are utilized. The classes correspond to different facial expressions. There are 10 facial expression classes in ADFES database, 8 expressions in TFEID database, and 7 expressions in MUG, KDEF, WSEFEP, and JAFFE database. The One-vs-All approach performs correlation of one class with other classes. Class-1 is defined to represent the positive objects while the other "not Class_n" is used to represent the negative objects. The 'n' represents the expressions number used in the classifier. At last, the classifier selects an effective class for its relevant tested samples.

4 Proposed Emotion Detection Using Deep Learning

Human psychological stress and emotion are very much interconnected. In computational psychology, the relationship between stress and emotions is used to understand the human behavior. A CNN is trained to detect, recognize and classify facial expressions and discrete emotion categories (Anger, Disgust, Neutral, Fear, Sad, Happy and Surprise). Further, logarithmic regression is applied to evaluate stress as a function of deciphered emotions. We performed experiments on Facial Expression Recognition (FER2013) dataset to evaluate our architecture. The feature extraction no longer required for CNN method. The method works well with images. Deep Learning understands the variables, relations and performs extraction according to its characteristics. Sometimes it is complicated to extract high level abstract features from raw data. So that Convolutional Neural Networks (CNN) is utilized for image detection and classification. This is the main advantage of the existing method, but the preprocessing stage should be enhanced.

4.1 Proposed Deep learning's Model for Facial Emotion Detection

The CNN extracts the features of the image data sets at different levels. The Outputs of convolutional layer are known as feature maps. The Feature map are based on the number of filters it utilizes. Then it will be passed to non-linear gating function. The proposed depression detection is shown in Fig. 3.
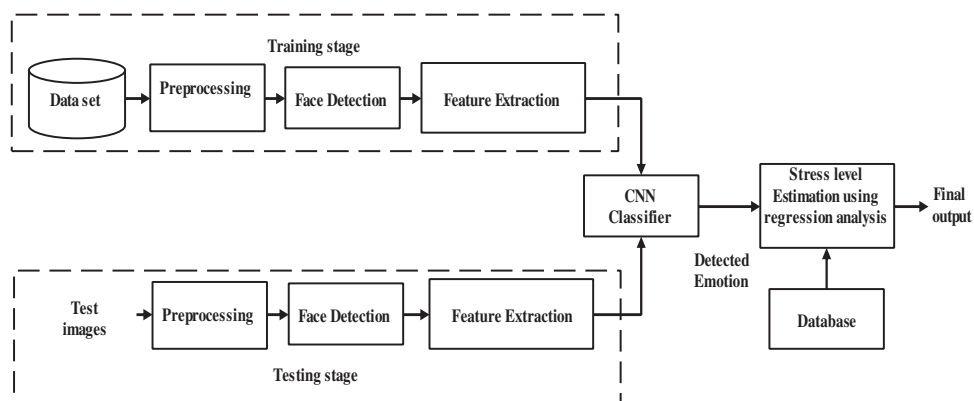

Figure 3: Architectural diagram for the proposed 'Depression Detection' system

Consider xi and yj represents i-th input and j-th output of feature map. Fig. 4 shows the block diagram of the proposed architecture.
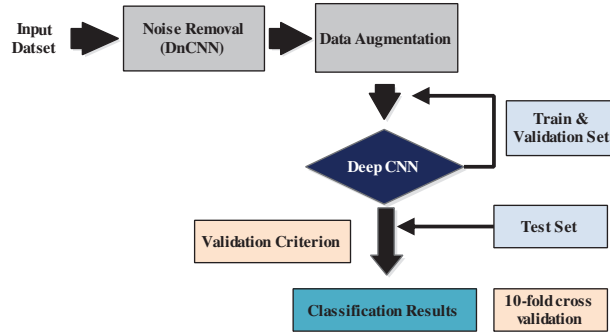

Figure 4: Block diagram of proposed architecture

4.1.1 DnCNN for Preprocessing

The Deep CNN architecture with denoising models based on MLP and CSF are shown in Fig. 5. The original mapping function to predict ECG image is done through residual learning by DCNN. The Convolutional filters dimension and the receptive field of Denoising Convolutional Neural Network (DnCNN) are initialized based on depth value. Proper depth values will optimize the performance and efficiency for larger images. Considering the Gaussian denoising, the receptive field size of DnCNN is initialized to 35 × 35 with the corresponding depth of 17. Based on the depth, the DnCNN has three types of layers as shown in Fig. 5. Batch normalization is added between convolution and ReLU units. The DnCNN works on the residual learning formulation and batch normalization is used to speed up training. This integration of residual learning and batch normalization also increases the denoising performance.
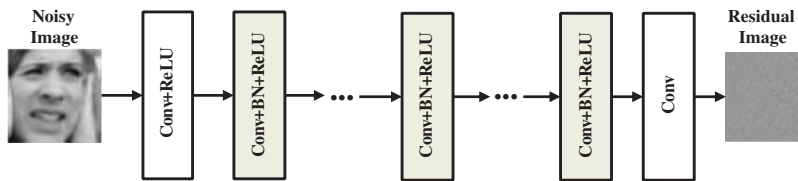

Figure 5: Proposed DnCNN architecture for image Pre- processing

4.1.2 Proposed Deep CNN Structure

This Network has large number of convolutional networks mixed with nonlinear and pooling layers (Fig. 6). If image gets passes on one convolution layer, the first layer output acts as input for the second layer. This process gets repeated for entire layers. After completing convolution operations, the output of the convolution layer will be added with nonlinear layer. The Nonlinear layer performs nonlinear properties on obtained result. If nonlinear properties are not involved, the network fails to detect and analysis the response variable. The Pooling layer follows the nonlinear layer. In this layer, the down sampling operations are performed with respect to image dimensions. Finally, image volume is minimized. After the entire process, the convolutional, nonlinear, and pooling layers will be linked with fully connected layer. The Fully connected layer extracts output data from convolutional networks. Fully connected layer ouput is said to be N dimensional vector. In this architecture N = 10. Where N denotes the number of classes'.
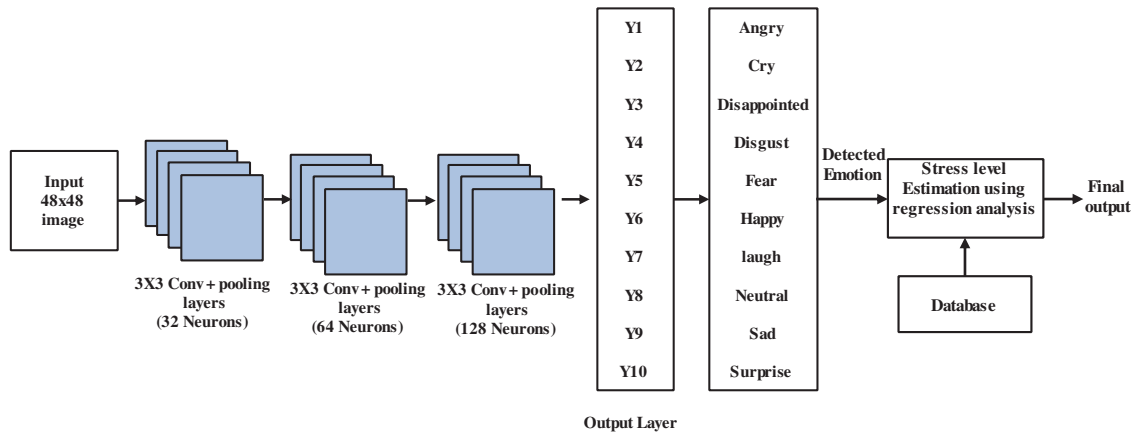
Figure 6: Proposed deep learning architecture for emotion detection and stress level analysis

4.1.3 Convolution Layers

The Hidden layer learns the local patterns in small 2D windows. This Layer detects the visual features of images. It also extracts the property of an image at certain point which is used to recognize horizontal edges. Each output result is sum and product of an original image. The Output will be 4×4 dimensions. The Input image of 6×6 dimensions is applied to 3×3 filter at 4×4 position. The Matrix evaluates the image at upper left and then moves towards right. If row is progressed, another row will be progressed from below. Multiple filters can be utilized to detect multiple features.

4.1.4 Pooling Layers

The Pooling layers combines with neurons to extract the input data properties. An "average pooling" extracts the average value from group of neurons. The "Max pooling" selects the maximum value from each group. In pooling layer when neurons are grouped, entries will be minimized. This Minimization is based on group's parameters dimensions. For example, in Fig. 7, a single size is transformed into a group with lesser dimension. Even though there is image transformation, spatial relationship exists within it.
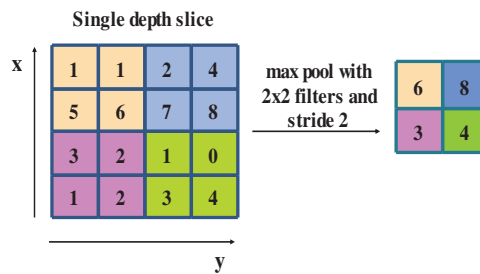

Figure 7: Pooling example

4.1.5 Reducing Overfitting

The Architecture with several layers may lead to overfit the training data. To eliminate those issues, dropout is applied in neural network as represented in Fig. 8. The Dropping unit eliminates the neurons and their weight contacts from the network. In Dropout, fixed probability is utilized but dropping choice of units is random. The Forward and back ward propagation networks use dropout of units, So that it leads to thin network.
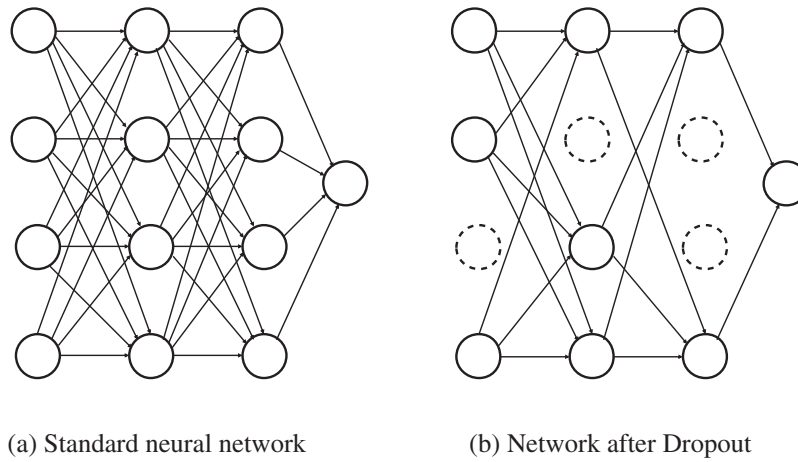
(a) Standard neural network   (b) Network after Dropout

Figure 8: Scheme of dropout performance

The Deep learning is performed on neural networks. In that, image will be passed into classifier. The Classifier alters the image size to 48 × 48. These Images consists of 3 channels for each color (RGB). The Matrix form for each image is 48 × 48 × 3. When number of neurons are larger evaluation time will increase. To avoid those issues ReLU activation function must be utilized.

5 Results and Discussion

The proposed algorithm for covid patient emotion detection is performed and its results are noted. This experiment is carried out in MATLAB software.

5.1 Dataset

The Datasets with different emotions and models are needed to train and test the images. The Data contains 48 × 48 pixel grayscale images [31]. The Database is generated with Google image search API and automatic registration processes are set, so that face gets centred. It maintains same amount of space for each image. Each picture categorized under 7 major categories given in Fig. 9.

Figure 9: FER example

Case 1: Train the Model with Lesser Dataset

In Case1 (Figs. 10,11), FER dataset consists of 7 emotions. Additionally, 3 emotions are included. Angry, cry, disappointed, disgust, fear, happy, laugh, neutral, sad and surprise emotions are evaluated. Before execution, the data has to be organized within a set and labelling has to be finished. Tab. 1 lists the 10 facial expressions and labels. Tab. 2 lists the number of training and validation dataset utilized under each emotion.
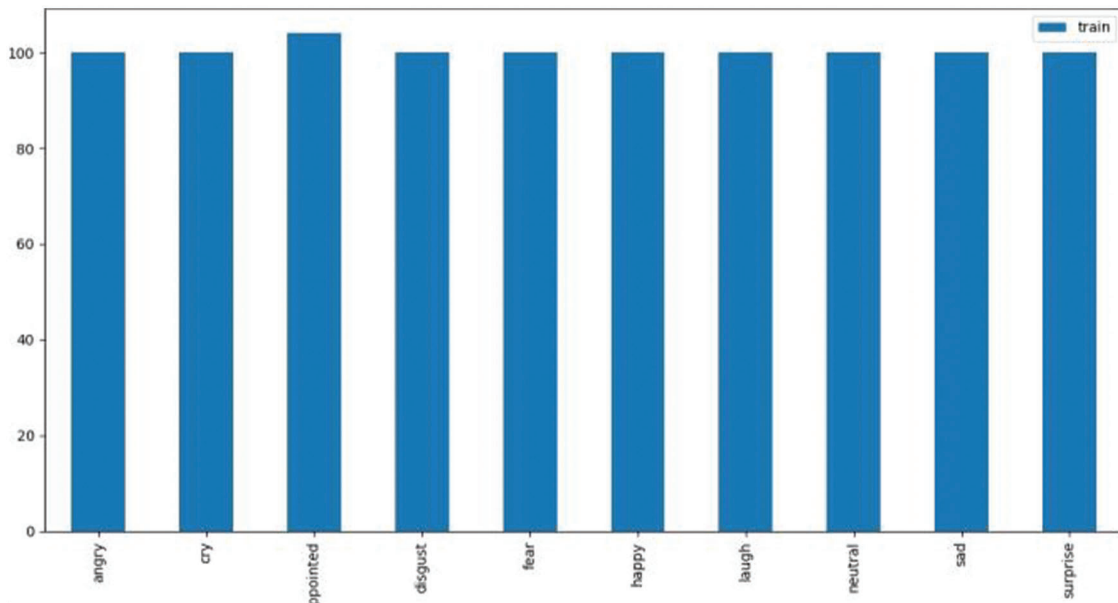

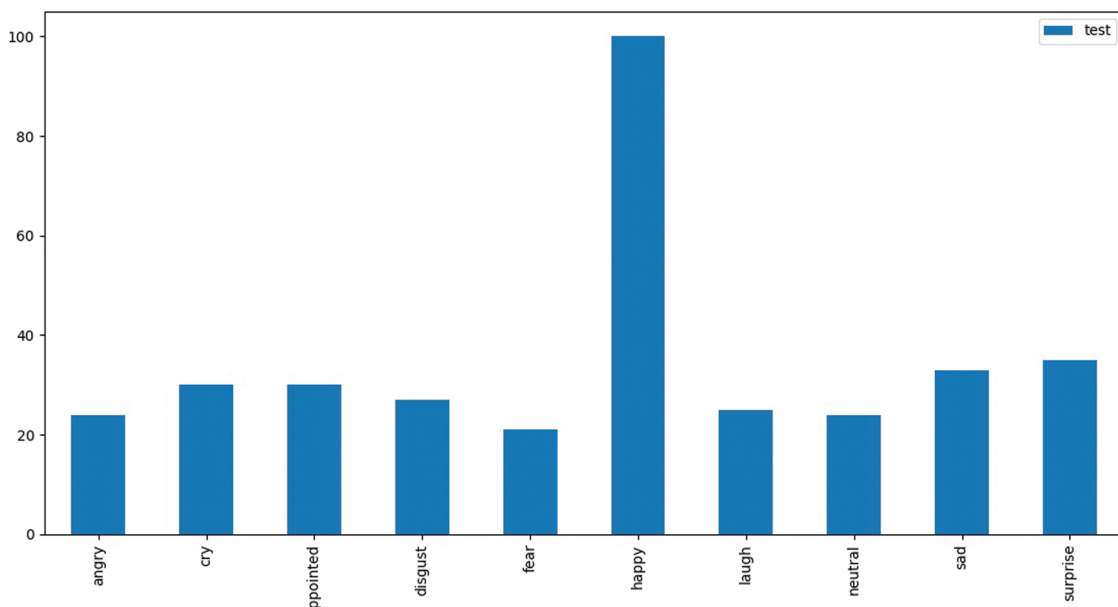**Figure 10:** Number of images in training dataset

Figure 11: Number of images in testing dataset

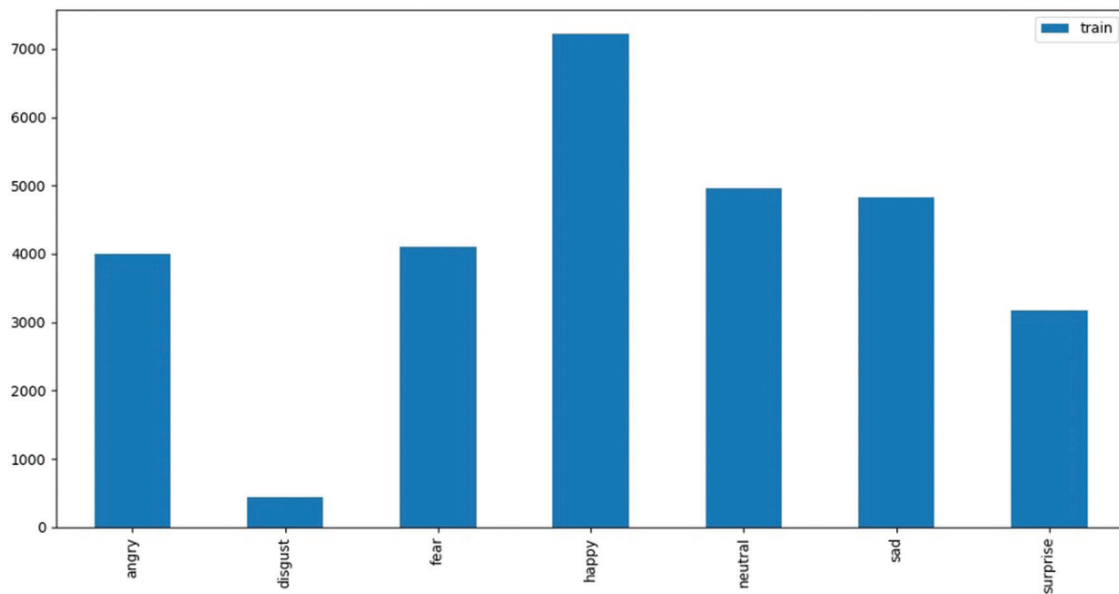**Table 1:** Ten facial expression categories and the corresponded labels

	Angry	Cry	Disappointed	Disgust	Fear	Happy	Laugh	Neutral	Sad	Surprise
Label	0	1	2	3	4	5	6	7	8	9

Table 2: Number of training and validation images used

	Angry	Cry	Disappointed	Disgust	Fear	Happy	Laugh	Neutral	Sad	Surprise
Training	100	100	104	100	100	100	100	100	100	100
Validation	24	30	30	27	21	100	25	24	33	35

Case 2: Train the Model with Larger Dataset

In case 2 (Figs. 12,13), FER dataset has several images categorized into 7 emotions. Tab. 3 lists the number of training and validation Dataset utilized in each emotion.

Figure 12: Number of images in training dataset

5.2 CNN Training Program

The Database must fulfil both quality and quantity (Fig. 10). The Quantity related failures are solved in 2 approaches. 1) Reducing the complexity of failures, 2) Increasing the number of datasets. The Learning rate, batch size, epochs count and drop out parameter are initiated with 1e−4, 20, 60 and 50%. The experiment results show that the small size input images has utilized minimum timing in training process. Figs. 14a and 15a represents the overall training accuracy results of datasets. Initially training accuracy obtained was 20% but finally it was improved to 1. When training reaches it, the accuracy will get stabilized by minor fluctuations. The Minor fluctuation was the new feature found by the neural network at learning. The Case 2 provides a better good accuracy when compared with case1. Figs. 14b and 15b

represents the loss occurred in training phase and these losses are reduced to 0 at final stage. Tab. 5 shows the performance of existing and proposed method.
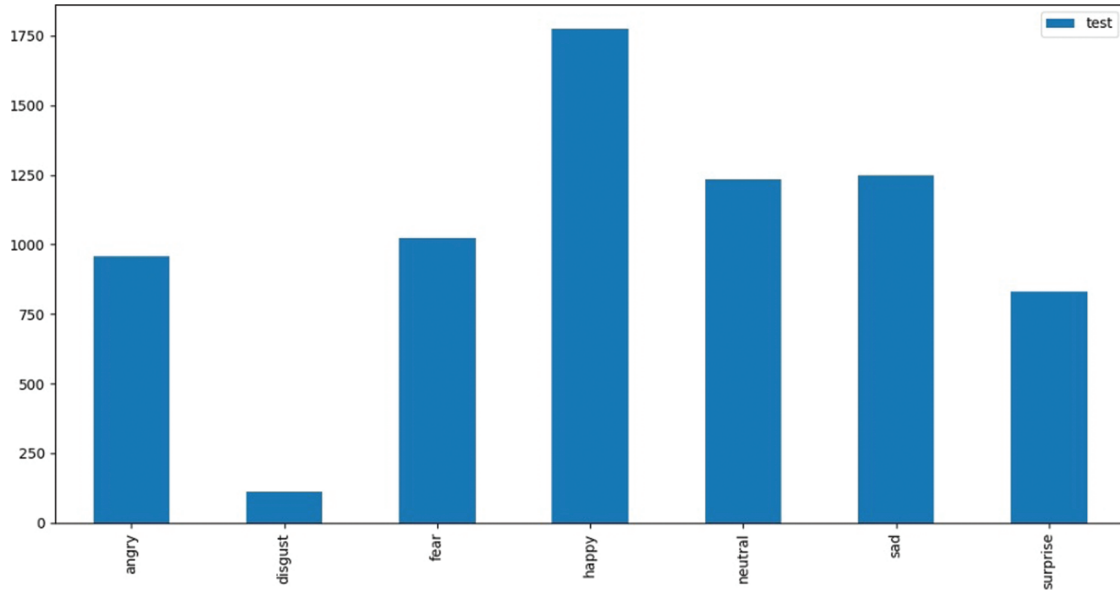

Figure 13: Number of images in test dataset

Table 3: Number of training and validation images used

	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Training	3995	436	4097	7215	4965	4830	3171
Validation	958	111	1024	1774	1233	1247	831

Figure 14: Case 1 (a) overall training accuracy (b) overall training loss of the datasets
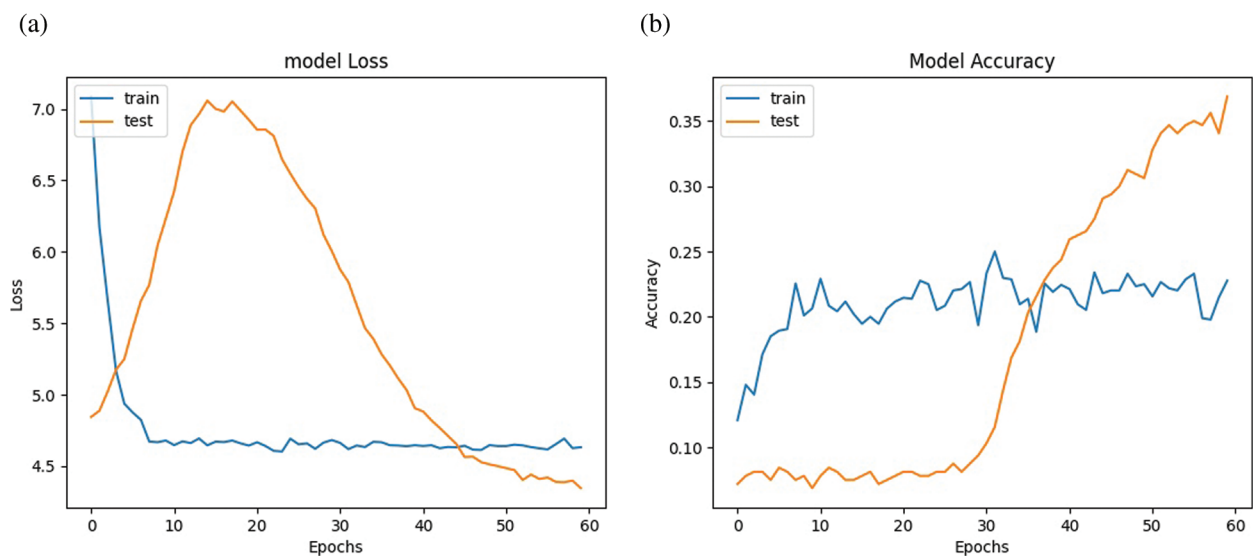
(a)   (b)
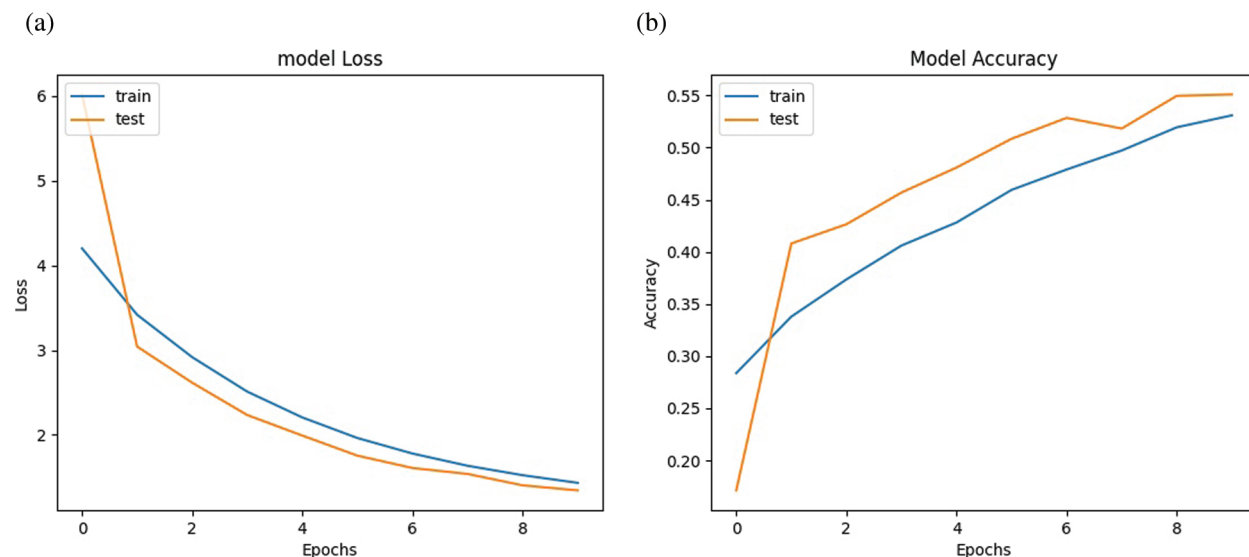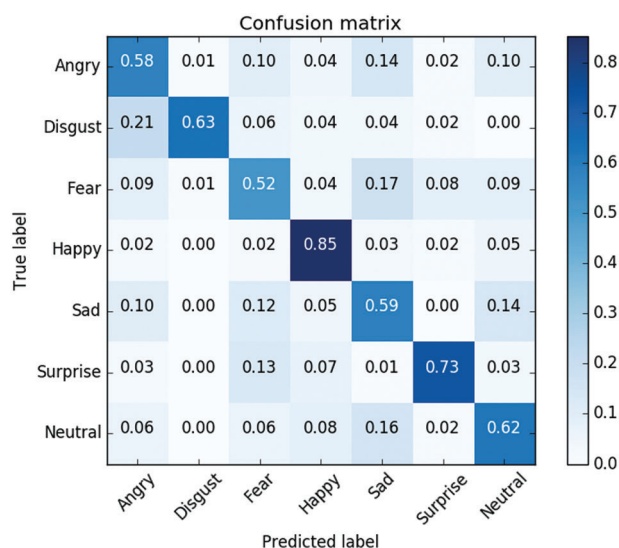

Figure 15: Case 2 (a) overall training accuracy (b) overall training loss of the datasets

5.3 Clinical Outcomes and Discussion

The investigation shows that the covid affected patients' facial emotion processing represents a vulnerability factor for depression. The confusion matrix of emotion detection is shown in Fig. 16. The levels of depression can be low, moderate, and severely dysregulated depressive profile. The Tab. 4 shows a 5 data from the pool to classify between low, moderate, and severely dysregulated depressive profiles. The suicide ideation can be confirmed through the depressive symptoms. These results are consistent with literature. The affective temperaments with stress level above 5 shown a high depression and suicidal behavior. These behaviors can be strongly related to anxious temperament. The clinical history can be verified that the presence of cyclothymic–depressive–irritable–anxious temperamental constellation pattern in the patients. The suicidal ideation follows from a dysphoric–dysregulated depressive profile. It's a challenging task to differentiate the low temperamental and severely dysregulated depressive component. The challenge also lies on the recognition of neutral facial expressions by severe affective temperament dysregulation. Further research is required to accurately classify the neutral facial expressions. Nevertheless, the results of the present study did not cover the interpretation of neutral facial expressions and problems arising due to psychological or social risk factors which also contributes to the development of depressive symptoms and disorders.

Figure 16: Confusion matrix of emotion detection

Table 4: Performance analysis

S. No.	Image	Stress level (0–9)	Emotional recognition	Chances of sucide attempt (%)
1.	Test 1	4	Low Temperamental Dysregulation	35
2.	Test 2	6	Severely Dysregulated Depressive nature	70
3.	Test 3	5	Moderate Temperamental Dysregulation	57
4.	Test 4	7	Severely Dysregulated Depressive and Affective Temperaments nature	75
5.	Test 5	2	Low Temperamental Dysregulation	10

Table 5: Performance analysis with existing methods

S. No.	Preprocessing	Classification	Accuracy (%)
1.	Convnets	LSTM	76
2.	Time and F	SVM	79
3.	Proposed	CNN	82

6 Conclusion

In this paper, the recognition of human feelings using image processing technology has been presented. The COVID patient's emotional disorder which leads to suicide ideation are monitored. The monitoring system based on face detection and emotional analysis are used to identify and solve those disorders in patients and doctors. Here CNN model with DnCNN pre-processing establishes efficient accuracy. This Algorithm is progressed on real time data sets. During covid times large number of individuals is suffering from suicidal ideation in the world due to different issues. The detection through self-expression, emotion release, and personal interaction, will find the suicidal thoughts in them. This work

uses facial expression recognition technique that automatically extracts the features on human face to recognize the patterns of suicide ideation. The proposed emotion recognition system monitors the emotion of COVID patients and alerts the nurse, when patient emotion is fear, cry or sad. This research has introduced Image Processing technology for emotional analysis of patients using Machine learning algorithm. Proposed Convolutional Neural Networks (CNN) architecture with DnCNN pre-processing have enhanced the performance. The system can analyze the mood of patients either in real time or in the form of video files from CCTV cameras.

The Hardware for the alert sending is completed using a Raspberry pi hardware unit. The program is done in python and the embedded unit interacts through the GSM modules –SIM 900 with the mobile phone. Since this paper presents the Image processing methods and results, the hardware unit is to be presented in a future paper. The work can be extended to detect the infection level in the patients along with the emotion detection

References

[1] Q. Wang, H. Yang and Y. Yu, "'Facial expression video analysis for depression detection in chinese patients," *Journal of Visual Communication and Image Representation*, vol. 57, no. 10, pp. 228–233, 2018.

[2] M. Ketcham, M. Piyaneeranart and T. Ganokratanaa, "Emotional detection of patients major depressive disorder in medical diagnosis," in *14th Int. Conf. on Signal-Image Technology & Internet-Based Systems (SITIS)*, Las Palmas de Gran Canaria, Spain, pp. 332–338, 2018.

[3] S. Shojaeilangari, W. Yau, K. Nandakumar, J. Li and E. K. Teoh, "Robust representation and recognition of facial emotions using extreme sparse learning," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2140–2152, 2015.

[4] J. Guo, Z. Lei, J. wan, E. Avots, N. Hajarolasvadi *et al.,* "Dominant and complementary emotion recognition from still images of faces," *IEEE Access*, vol. 6, pp. 26391–26403, 2018.

[5] J. Yan, W. Zheng, Q. Xu, G. Lu, H. Li *et al.,* "Sparse kernel reduced-rank regression for bimodal emotion recognition from facial expression and speech," *IEEE Transactions on Multimedia*, vol. 18, no. 7, pp. 1319–1329, 2016.

[6] C. Mumenthaler, D. Sander and A. S. R. Manstead, "Emotion recognition in simulated social interactions," *IEEE Transactions on Affective Computing*, vol. 11, no. 2, pp. 308–312, 2020.

[7] T. Zhang, W. Zheng, Z. Cui, Y. Zong and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 839–847, 2019.

[8] A. Chakraborty, A. Konar, U. K. Chakraborty and A. Chatterjee, "Emotion recognition from facial expressions and its control using fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 4, pp. 726–743, 2009.

[9] I. Shahin, A. B. Nassif and S. Hamsa, "Emotion recognition using hybrid gaussian mixture model and deep neural network," *IEEE Access*, vol. 7, pp. 26777–26787, 2019.

[10] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller and S. Zafeiriou, "End-to-End multimodal emotion recognition using deep neural networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 8, pp. 1301–1309, 2017.

[11] G. Du, S. Long and H. Yuan, "Non-contact emotion recognition combining heart rate and facial expression for interactive gaming environments," *IEEE Access*, vol. 8, pp. 11896–11906, 2020.

[12] S. Li and W. Deng, "Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 356–370, 2019.

[13] M. S. Hossain and G. Muhammad, "An emotion recognition system for mobile applications," *IEEE Access*, vol. 5, pp. 2281–2287, 2017.

[14] C. Cruz, B. Bhanu and N. S. Thakoor, "Vision and attention theory based sampling for continuous facial emotion recognition," *IEEE Transactions on Affective Computing*, vol. 5, no. 4, pp. 418–431, 2014.

[15] S. Wang, "Micro-expression recognition using color spaces," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6034–6047, 2015.

[16] R. Xia and Y. Liu, "A multi-task learning framework for emotion recognition using 2D continuous space," *IEEE Transactions on Affective Computing*, vol. 8, no. 1, pp. 3–14, 2017.

[17] U. Tariq, K. H. Lin, Z. Li, X. Zhou, Z. Wang *et al.,* "Recognizing emotions from an ensemble of features," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 4, pp. 1017–1026, 2012.

[18] P. Chiranjeevi, V. Gopalakrishnan and P. Moogi, "Neutral face classification using personalized appearance models for fast and robust emotion detection," *IEEE Transactions on Image Processing*, vol. 24, no. 9, pp. 2701–2711, 2015.

[19] Y. Zhang, Z. J. Yang, H. M. Lu, X. X. Zhou, P. Phillips *et al.,* "Facial emotion recognition based on bi-orthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation," *IEEE Access*, vol. 4, pp. 8375–8385, 2016.

[20] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu *et al.,* "A natural visible and infrared facial expression database for expression recognition and emotion inference," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 682–691, 2010.

[21] W. Zhang, X. He and W. Lu, "Exploring discriminative representations for image emotion recognition with CNNs," *IEEE Transactions on Multimedia*, vol. 22, no. 2, pp. 515–523, 2020.

[22] K. P. Seng, L. Ang and C. S. Ooi, "A combined rule-based & machine learning audio-visual emotion recognition approach," *IEEE Transactions on Affective Computing*, vol. 9, no. 1, pp. 3–13, 2018.

[23] J. Kim, B. Kim, P. P. Roy and D. Jeong, "Efficient facial expression recognition algorithm based on hierarchical deep neural network structure," *IEEE Access*, vol. 7, pp. 41273–41285, 2019.

[24] S. Wang, Z. Liu, Z. Wang, G. Wu, P. Shen *et al.,* "Analyses of a multimodal spontaneous facial expression database," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 34–46, 2013.

[25] H. Zhang, A. Jolfaei and M. Alazab, "A face emotion recognition method using convolutional neural network and image edge computing," *IEEE Access*, vol. 7, pp. 159081–159089, 2019.

[26] K. Mistry, L. Zhang, S. C. Neoh, C. P. Lim and B. Fielding, "A micro-GA embedded PSO feature selection approach to intelligent facial emotion recognition," *IEEE Transactions on Cybernetics*, vol. 47, no. 6, pp. 1496–1509, 2017.

[27] J. C. Meléndez, E. Satorres, M. Reyes-Olmedo, I. Delhom, E. Real *et al.,* "Emotion recognition changes in a confinement situation due to COVID-19," *Journal of Environmental Psychology*, vol. 72, pp. 101518, 2020.

[28] W. Lu and M. Yang, "Face detection based on viola-jones algorithm applying composite features," in *Int. Conf. on Robots & Intelligent System (ICRIS)*, Haikou, China, pp. 82–85, 2019.

[29] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 886–893, 2005.

[30] A. L. Ramadhani, P. Musa and E. P. Wibowo, "Human face recognition application using PCA and Eigen face approach," in *Second Int. Conf. on Informatics and Computing (ICIC)*, Jayapura, Indonesia, pp. 1–5, 2017.

[31] https://www.kaggle.com/msambare/fer2013.