


RESEARCH

Open Access



Automated detection of human mental disorder

Shereen A. Hussein* , Abd El Rahman S. Bayoumi and Ayat M. Soliman

*Correspondence:
sam26@fayoum.edu.eg

Department of Computer
Science, Faculty of Computers
& Artificial Intelligence, Fayoum
University, Fayoum, Egypt

Abstract

The pressures of daily life result in a proliferation of terms such as stress, anxiety, and mood swings. These feelings may be developed to depression and more complicated mental problems. Unfortunately, the mood and emotional changes are difficult to notice and considered a disease that must be treated until late. The late diagnosis appears in suicidal intentions and harmful behaviors. In this work, main human observable facial behaviors are detected and classified by a model that has developed to assess a person's mental health. Haar feature-based cascade is used to extract the features from the detected faces from FER+ dataset. VGG model classifies if the user is normal or abnormal. Then in the case of abnormal, the model predicts if he has depression, anxiety, or other disorder according to the detected facial expression. The required assistance and support can be provided in a timely manner with this prediction. The system has achieved a 95% of overall prediction accuracy.

Keywords: Mental health, Depression, Anxiety, CNN, VGG16, Facial emotions, Haar-cascades

Introduction

Mental health is a state of well-being in which a person understands his/her abilities, can cope with the everyday stresses of life, work productively, and contribute to the community. When the demands on someone exceed their resources and coping abilities, their mental health will be negatively affected, including their emotional, psychological, and social well-being. In low- and middle-income countries, between 76 and 85% of people with mental disorders receive no treatment for their disorder [1]. There are many different mental disorders such as depression, mood swings, stress, anxiety, and other psychoses. Unfortunately, these disorders are frequently undetected and are responsible for various morbidities, either directly or indirectly [2]. Recent studies have given the required attention to monitor a person's mental health using machine learning and artificial intelligence techniques. The unconsciously transmitted behavioral symptoms expressed by head-pose, eye-gaze direction, and facial expressions were used to model structures that could predict mental health conditions [3].

Related work

Most of vision-based approaches for mental health automatic prediction like depression and anxiety relied on models that used behavioral features usually extracted from faces only and did not take personality into account. The latest developments in machine learning field encouraged its use in automatic facial expressions analysis for mental disorders detection.

Jaiswal et al. [4] combined the facial behavior and self-reported personality information to predict depression and anxiety disorders using deep neural networks (DNN). The facial behavior is encoded as a set of behavior primitives, to with facial action units (AUs), head-pose, and eye-gaze direction.

Yoon et al. [5] investigated the recognition accuracy of simple and complex emotions to examine accuracy for depression and anxiety disorders facial expressions recognition. With hypotheses of patients with major depressive disorder (MDD) and anxiety disorders (AnD) would show lower recognition accuracy for complex emotions, and also patients with MDD would show more difficulties in recognizing pleasant emotions compared to unpleasant emotions.

Giannakakis et al. [6] developed a framework for detecting and analyzing stress/anxiety emotional states through video-recorded facial cues. A feature selection procedure was employed to select the most robust features, followed by classification schemes discriminating between stress/anxiety and neutral states.

Venkataraman et al. [7] extracted the features of depression for different frontal faces. Based on the level of depression features, they were classified as depressed or non-depressed using the SVM classifier. Then, the level of depression is estimated through the level of contempt and disgust.

Khairuddin et al. [8] achieved an optimized model for facial emotion recognition with single-network classification accuracy on FER2013 dataset. He adopted the VGG network and constructed various experiments to explore different optimization algorithms and learning rate schedulers.

Dataset

FER+ dataset [9] which was designed to detect depression and anxiety by training the dataset (FER+ 2013) is used in the proposed model. The data comprises 38,537 grayscale images of faces at a resolution of 48×48 pixels. The aim is to categorize each face into one of seven emotion categories, all labeled, depending on the emotion expressed in the facial expression. The FER2013 dataset contains images that vary in viewpoint, lighting, and scale. Emotions in the dataset are “angry,” “disgust,” “fear,” “happiness,” “sad,” “surprise,” and “neutral” as in Fig. 1. The training set consists of 32,114 examples. The public test set consists of 6423 examples.

The extended Cohn–Kanade dataset (CK+) [10] that contains the facial behavior of 123 participants was from 18 to 30 years of age. It includes 327 out 593 video sequences that match one the seven emotions we interested in as in Fig. 1. Participants display different expressions began and ended in a neutral face. Image sequences for frontal views and 30° views were digitized with 8-bit grayscale values.

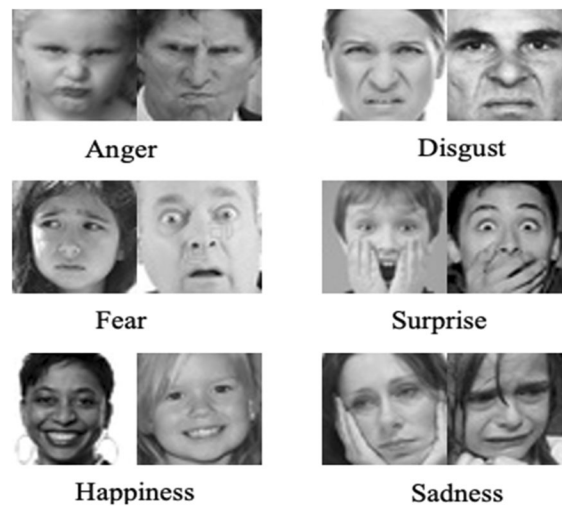


Fig. 1 Samples of different emotions

Only the final frame of each sequence with peak expression is been used in the experiment, which results in 309 images.

Proposed system

The overall architecture of the proposed model with four main phases is presented in Fig. 2.

Frame preprocessing

This phase is used to improve the input image frames by suppressing undesired distortions or enhancing some frame features relevant for further processing and analysis tasks. Fourteen frames per second (14/FPS) [11] are considered from an uploaded video as the landmark of the user in the video will be more clear and more useful for the process of detection. Beginning with grayscale conversions as colored frames not a matter in the detection process. Followed by resizing of frame in 48×48 and rescaling by normalization algorithms. Finally, representing the behavior attributes by histogram.

Face detection

The main objective of face detection process in grayscale image is to determine whether there is any face in an image or not. Haar-cascades algorithm [12] is a machine learning method that involves drilling a classifier from a large number of positive and negative pictures. It is an object detection algorithm that detects faces in images and real-time videos. Haar feature-based cascade is used to extract the features from the detected faces. Each feature is a distinct value obtained by subtracting the sum of pixels in the white rectangle from the total of pixels in the black rectangle, in which it recognizes the faces of various people in various settings. Because of integral pictures, the Haar-like feature of any size may be determined in constant time. The system automatically detects faces using Haar-cascade then segments it and fed it to the model for classification and prediction [7].

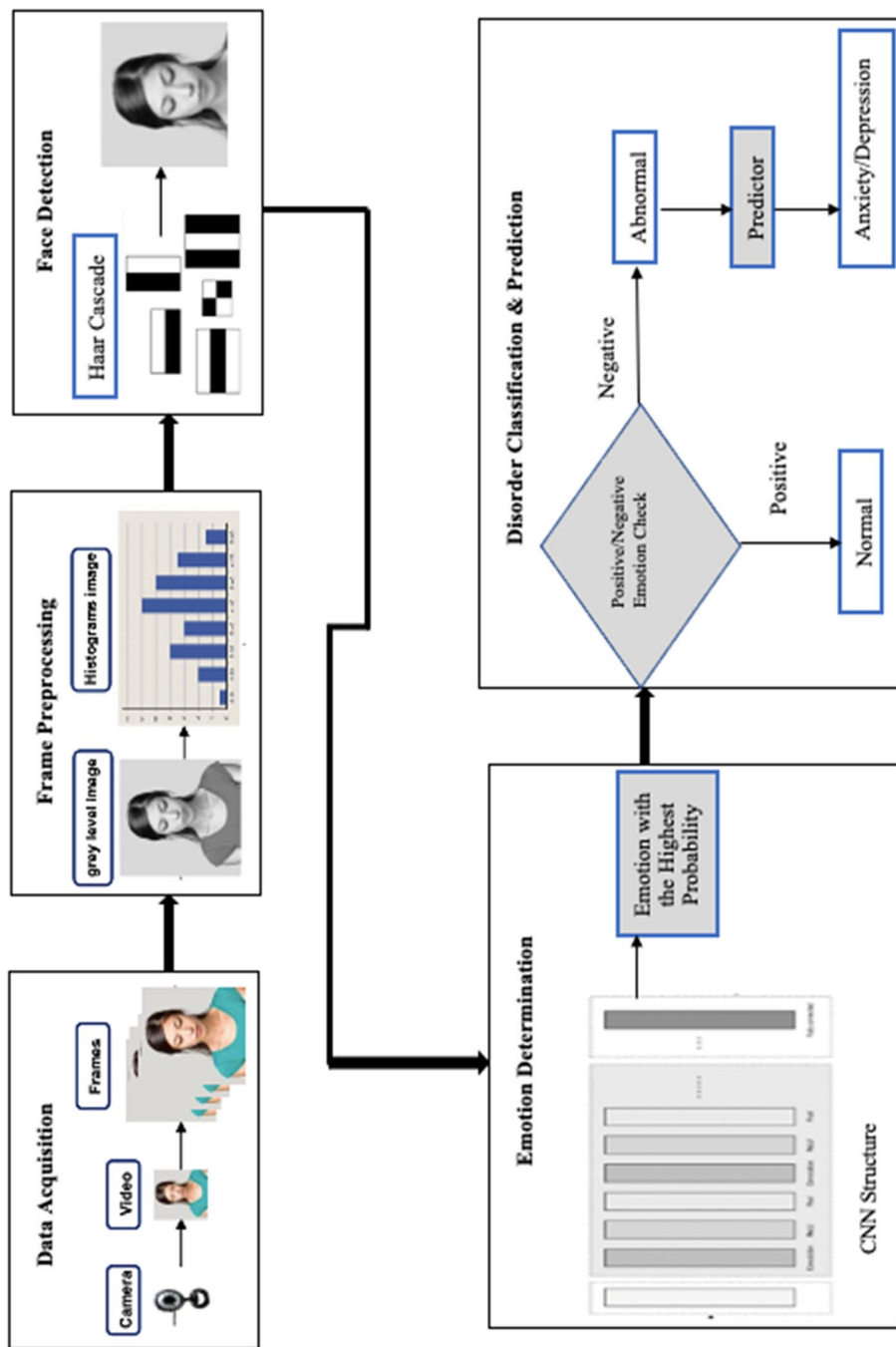


Fig. 2 Proposed model architecture

Emotion determination model

This work used a transfer learning model VGG that concatenates the best weights with deep learning to enhance the classification of emotions. The proposed model utilizes VGG model with a pretrained network with ImageNet without final classification layer. The transfer learning allows training a given dataset using a pretrained CNN model to overcome the difficulty and time consumption problems [13, 14].

The VGG model [15] is a classical convolutional neural network architecture used in large-scale image processing and pattern recognition using six convolution layers, interleaved with max-pooling, batch normalization, and activation function layers as presented in Fig. 3. More specifically, after the input layer, there are two convolution layers with 64 kernels of size 3×3 . The structure repeats but changes in the number of convolution layers and number of kernels. A Conv 1×1 layer with 128 filters with padding-zero and stride-one is added to offer feature map pooling. It decreases the depth from 256 to 128 feature maps with retaining the most salient features to reduce the dimensionality [16]. Then, a flattening layer turns the matrix into one vector. After all the convolution layers, three dense layers are added, each with 128 hidden nodes. Adding these dense layers to the convolutional base, which has its weights frozen while additional dense layers are trained. The final dense layer with soft-max of 7 nodes is to generate the output emotions.

The proposed model compiled with the Adam optimizer [17] is a stochastic gradient descent method with a learning rate equal $1e-7$. The proposed model is fitted at epochs where value equals 70, train batch size equals 64, and validate batch size equals 64 and avoided the overfitting problem. The best-saved weights were loaded after those testing images were loaded. Finally, images were resized to $48 * 48$ and fed the testing images to the model to classify seven classes of the facial emotions.

Disorder classification and prediction

The final phase in the system architecture where the model firstly classifies if the user is *normal* or *abnormal* then in the case of abnormal the model predicts if he has depression, anxiety, or other disorder according to the facial expression that was detected. The CNN algorithm is used to extract emotion from the frame to detect emotion with the highest probability to obtain one emotion from seven emotions [18] as Disgusted, Fearful, Happy, Neutral, Sad, Surprised, and Angry.

The emotions are classified to two classes: positive and negative [19–21]. Positive emotions are considered if the emotion is “Happy” or “Neutral” or “Surprised,” the results became that detect Human is *Normal*. Negative emotions are considered if the emotion is “Fearful” or “Disgusted” or “Sad” or “Angry,” the results became that detect Human is *Abnormal*.

Based on the result of classification, the abnormal case could be an indicator for three disorders which are “Anxiety,” “Depression,” and “other disorders.” The prediction is an anxiety disorder where the detected emotions are combination of fearful and disgusted [22]. Depression disorder is a combination of sad and angry detected emotions. Other disorder if there are another combinations [23].

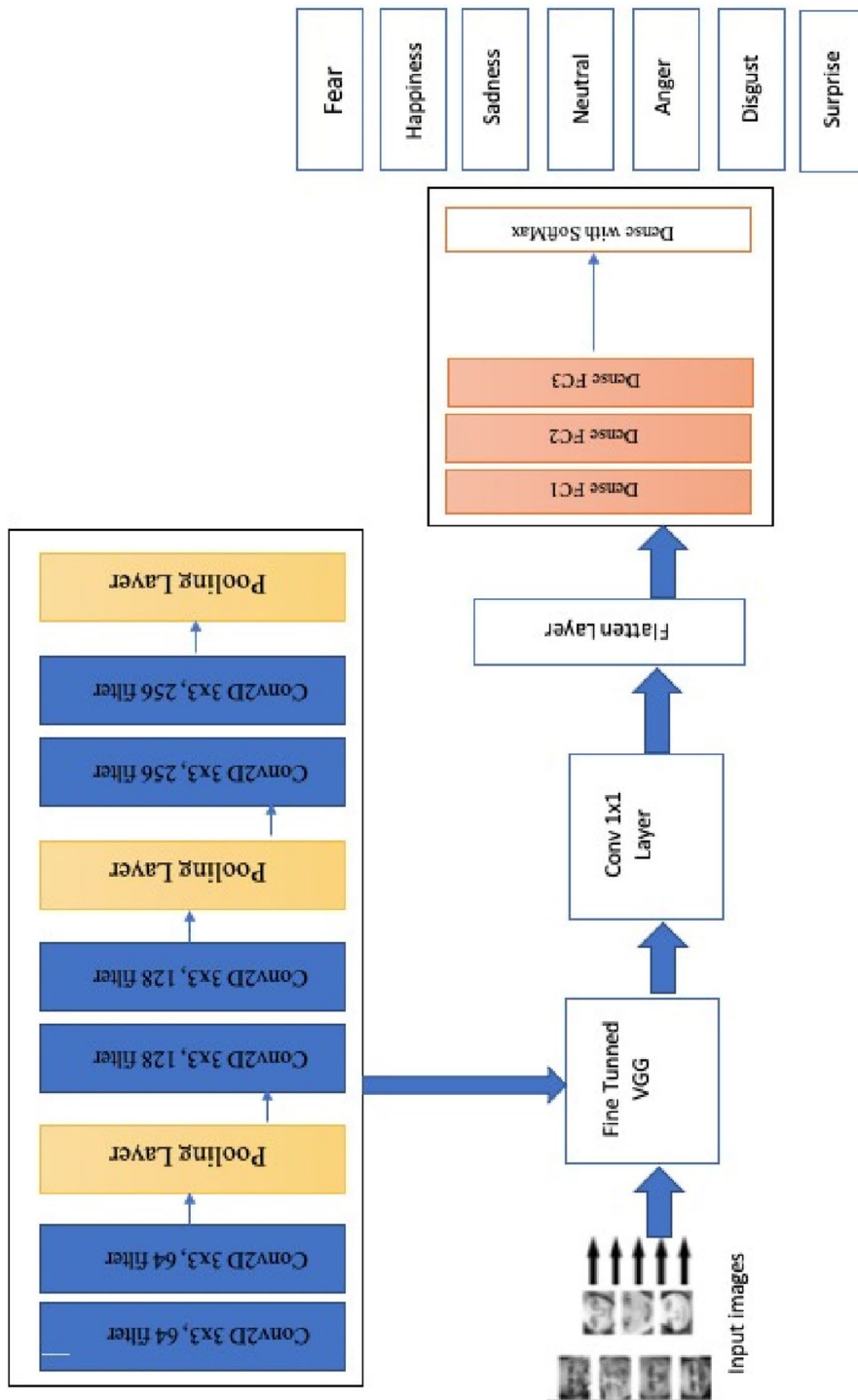


Fig. 3 The structure of the VGG model

Discussion and experimental results

This work demonstrates that the highest accuracy is achieved with FRB+ and CK+ dataset using VGG model of transfer learning. The dataset is divided into a ratio of 8:2. All experiments are implemented using Colab [24] that is provided by Google, the Keras framework [25] that can run on top of TensorFlow, and the Python programming language. All experiments were conducted on a 12 GB NVIDIA Tesla K80 GPU (Graphical Processing Unit) and 12 GB of RAM.

The performance of the proposed model is evaluated by these measures' precision (1), recall (2), F1-score (3), and accuracy (4) in the testing phase equations shown in Tables 1

Table 1 Performance equations summary

Assessments	Equation	Equ. No
Precision (P)	$\frac{TP}{TP+FP}$	(1)
Recall (R)	$\frac{TP}{TP+FN}$	(2)
F1-Score (F)	$2 * \frac{P * R}{P+R}$	(3)
Accuracy (Acc)	$\frac{TP+TN}{TP+TN+FP+FN}$	(4)

In true positive (TP), the model predicts the positive class correctly. In true negative (TN), the model correctly classifies the negative class. In a false positive (FP), the model predicts the positive class incorrectly. In false negative (FN), the model predicts the negative class incorrectly

Table 2 Performance measure values

Emotion	Precision	Recall	F1-Score
Angry	0.98	0.94	0.96
Disgust	0.96	0.99	0.97
Fear	0.96	0.89	0.92
Happy	0.94	1.00	0.97
Sad	0.95	0.95	0.95
Surprise	0.94	0.98	0.96
Neutral	0.98	0.88	0.93
Accuracy			0.95

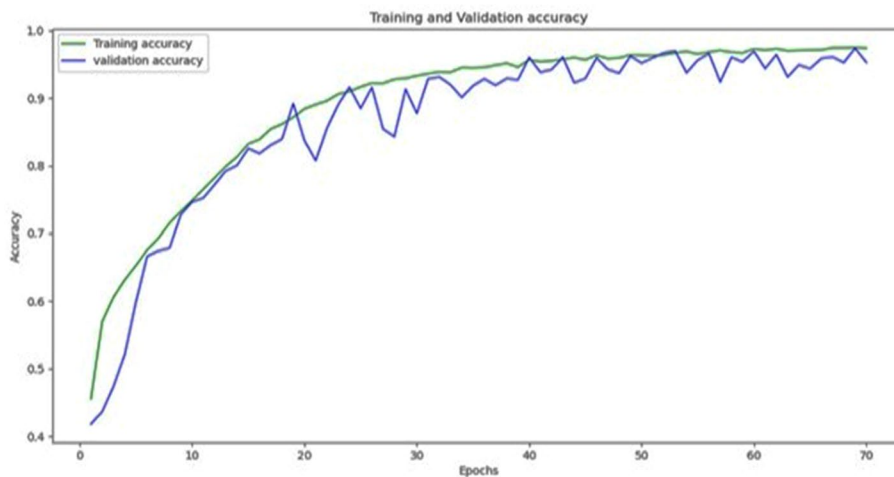


Fig. 4 The accuracy of training and validation

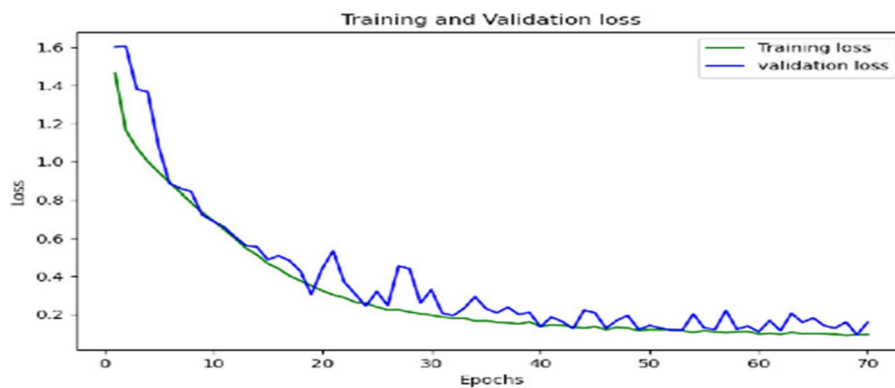


Fig. 5 The loss of training and validation

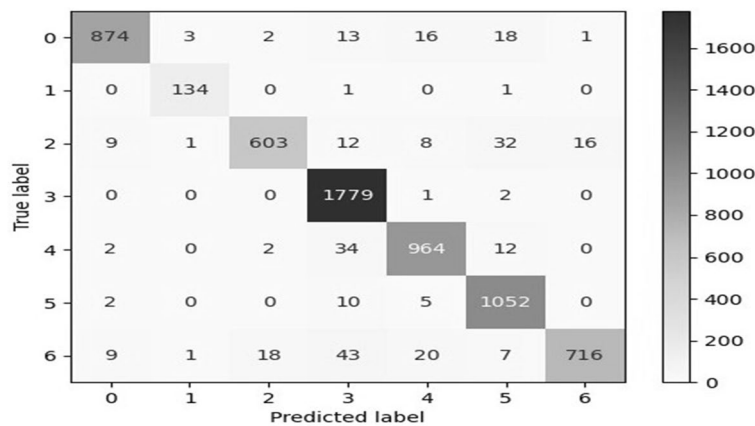


Fig. 6 Confusion matrix for the model

and 2. Moreover, loss, accuracy, validation loss, and validation accuracy are calculated during different epochs in the training phase, as shown in Figs. 4, 5, and 6. Also, the proposed model performance is compared with other models that work on the same dataset as shown in Table 3.

Conclusion

Daily life stress has caused a considerable degree of fear, concern, and anxiety among the entire population their mental health deteriorates that may slip into depression or anxiety. Not everyone has the luxury to get in contact with therapists and psychologists for his mental suffering determination. Many papers had discussed this option and provided some ideas, such as using facial expressions, mobile/computers logs, and eye expressions with gaze tracking, and others. This study has demonstrated a model using facial expression and the FER+ dataset using a VGG network of transfer learning. All hyperparameters have tuned toward an optimized model for facial emotion recognition. Different optimizers and learning rate schedulers are explored and the best testing classification accuracy achieved is 95%, surpassing all network accuracies previously reported that helped us to detect the mental health of the patient if he/she has depression, anxiety, or normal.

Table 3 Classification accuracies summary

Paper name	Dataset	Method	Accuracy rate (%)
Facial emotion recognition: state of the art performance on FER2013 [8]	FER2013	VGG13	73.28
Automatic prediction of depression and anxiety [18]	FER2013	CNN	63
Local learning to improve bag of visual words model for facial expression recognition [26]	JAFFE + (CK+) + MMI	Bag of word	67.484
Deep learning approaches for facial emotion recognition: a case study on FER-2013 [21]	FER2013	GoogLeNet	65.2
Local learning with deep and hand-crafted features for facial expression recognition [27]	FER+	VGG-SVM	66.31
Going deeper in facial expression recognition using deep neural networks [28]	(CK+) + DISFA, + FER	Conv + inception layer	66.4
Learning to amend facial expression representation via de-albino and affinity [29]	RAF-DB + AffectNet + SFEW	ARM(ResNet-18)	71.28
Facial expression recognition using convolutional neural networks: state of the art [30]	FER+	ResNet	72.2
		VGG	72.7
The proposed model	FER++ (CK+)	VGG	95

Acknowledgements

Not applicable.

Author contributions

AS prepared the input image frames and applied the Haar-cascade algorithm for face detection. AB built the VGG-16 model for the emotion determination. SH performed the testing & classification of the mental disorder, and was a major contributor in writing the manuscript. All authors read and approved the final manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The dataset (FER+ 2013) is used in the proposed model. The data comprises 38,537 grayscale images of faces at a resolution of 48 × 48 pixels. The aim is to categorize each face into one of seven emotion categories, all labeled, depending on the emotion expressed in the facial expression. Emotions in the dataset are "angry," "disgust," "fear," "happiness," "sad," "surprise," and "neutral." It can be freely and openly accessed via <https://www.kaggle.com/datasets/msambare/fer2013>.

Declarations**Competing interests**

The authors declare that they have no competing interests.

Received: 21 September 2022 Accepted: 2 February 2023

Published online: 08 February 2023

References

- World Health Organization (2022) Mental disorders. WHO, Geneva
- Government of Canada (2013) Canadian centre for occupational health and safety. Ccohs.ca
- Foley GN, Gentile JP (2010) Nonverbal communication in psychotherapy. *Psychiatry* (Edgmont) 7:38
- Jaiswal S, Song S, Valstar MF (2019) Automatic prediction of depression and anxiety from behaviour and personality attributes. In: 8th International conference on affective computing and intelligent interaction (ACII)
- Yoon S, Kim H, Kim J, Lee S, Lee S (2016) Reading simple and complex facial expressions in patients with major depressive disorder and anxiety disorders. *Psychiatry Clin Neurosci* 70:151–158

6. Giannakakis G, PEDIADITIS M, Manousos D, Kazantzaki E, Chiarugi F, Simos PG, Marias K, Tsiknakis M (2017) Stress and anxiety detection using facial cues from videos. *Biomed Signal Process Control* 31:89–101
7. Venkataraman D, Parameswaran NS (2018) Extraction of facial features for depression detection among students. *Int J Pure Appl Math* 118:455–463
8. Khairuddin Y, Chen ZL (2021) Facial emotion recognition: state of the art performance on FER2013
9. FER-2013 (2013) [www.kaggle.com. https://www.kaggle.com/datasets/msambare/fer2013](https://www.kaggle.com/datasets/msambare/fer2013)
10. CK+. [www.kaggle.com. https://www.kaggle.com/datasets/shawon10/ckplus](https://www.kaggle.com/datasets/shawon10/ckplus)
11. Zarif NE, Montazeri L, Leduc-Primeau F, Sawan M (2021) Mobile-optimized facial expression recognition techniques. *IEEE Access* 9:101172–101185
12. Padilla R, Filho CF, Costa MG (2012) Evaluation of Haar cascade classifiers designed for face detection. *World Acad Sci Eng Technol Int J Comput Electr Autom Control Inf Eng* 64:362–365
13. Barros B et al (2021) Pulmonary covid-19: learning spatiotemporal features combining CNN and LSTM networks for lung ultrasound video classification. *Sensors* 21:5486
14. Sheng S et al (2020) Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. *Appl Energy* 260:114296
15. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: 3rd International conference on learning representations, ICLR 2015—conference track proceedings
16. Christen S et al (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR)
17. Diederik PK, Jimmy B (2015) Adam: a method for stochastic optimization. Cornell University, Ithaca
18. Gopika, Haritha H, Sabira Reshni IK, Hema PM (2021) Automatic prediction of depression and anxiety. *Int J Curr Eng Sci Res IJCESR* 8:109–118
19. Prabhu S, Mittal H, Varagani R, Jha S, Singh S (2022) harnessing emotions for depression detection. *Pattern Anal Appl* 25:537–547
20. Pathak P, Gangwar H, Agarwal A (2021) Detecting negative emotions to counter depression using CNN
21. Giannopoulos P, Perikos I, Hatzilygeroudis I (2018) Deep learning approaches for facial emotion recognition: a case study on FER-2013
22. Cleveland Clinic (2022) Anxiety disorders: types, causes, symptoms and treatments
23. Mayo Clinic (2022) Depression (major depressive disorder)—symptoms and causes
24. Bisong E (2019) Google colab. In: Bisong E (ed) Building machine learning and deep learning models on google cloud platform. Apress, Berkeley
25. Keras Team. Getting started. Keras.io
26. Ionescu RT, Grozea C (2013) Local learning to improve bag of visual words model for facial expression recognition. In: ICML workshop on representation learning, Atlanta, Georgia, USA
27. Georgescu M, Ionescu RT, Popescu MC (2019) Local learning with deep and handcrafted features for facial expression recognition. *IEEE Access* 7:64827–64836
28. Mollahosseini A, Chan D, Mahoor MH (2016) Going deeper in facial expression recognition using deep neural networks. In: IEEE winter conference on applications of computer vision (WACV)
29. Shi J, Zhu S (2021) Learning to amend facial expression representation via de-albino and affinity. ArXiv
30. Pramerdorfer C, Kampel M (2016) Facial expression recognition using convolutional neural networks: state of the art. ArXiv

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
