Depression Level Analysis Using Face Emotion Recognition Method

Sudarshan Khandelwal [1*], Shridhar Sharma [2], Suyash Agrawal [3], Gayatri Kalshetti [4], Bindu Garg [5], Rachna Jain [6]

[1,2,3,4,5] Department of CSBS, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

[6] Department of Information Technology, Bhagwan Parshuram Institute of Technology, New Delhi-110089

sudarshankhandelwal99@gmail.com, shridharsharma200@gmail.com, agrawalsuyash2001@gmail.com, gayatrikalshetti3@gmail.com, bindugarg80@gmail.com, rachnabtec@gmail.com

Abstract.
Globally, most of the population faces depression or stress for a variety of reasons and at different stages of their lives. Stress in the modern world assists to depression over time, because of the hectic pace of our lives. Artificial intelligence systems can mimic the empirical system of a human person. Comprehensively capable machines and robots are capable of automatically identifying the mental state of a person from their facial expressions and body language. In order to determine and categories the amount of depression, artificial intelligence (AI) and deep learning algorithms are employed to recognize the facial expressions of people in real-time. By using this software analysis to obtain revised depression levels, this model will enable psychiatrists and other medical personnel involved in human psychology to gain a new viewpoint. This model is trained on FER Plus dataset obtained from Kaggle and later CNN model is used to channelize the output with the accuracy of 62.44%. The primary driving force behind this effort is to increase the precision of the following model statement, which has the potential to have an effect on subsequent work and long-term implementation.

Keywords: Convolutional Neural Network, Deep Learning, Deep Neural Network, Emotion Recognition, Face Detection, Haar Cascade

1 INTRODUCTION

A psychiatric condition known as depression affects more than 300 million people globally. A depressed person has anxiety on a daily basis, which negatively impacts their relationships with their family and friends, worsens their health, and in the worst-case scenario, results in suicide. People are under greater strain than ever because of how quickly work and life is moving forward, which raises their risk of developing depression. However, many patients might not get a timely diagnosis because of the extreme imbalance in the doctor-to-patient ratio worldwide. The primary indicators of depression include a lack of appetite and sleep, thinning hair, a loss of interest in once enjoyed hobbies, and a general lack of social life. All of these symptoms may occasionally exacerbate depression. According to [1], about 15% of Indian individuals need active support for one or more mental health issues, and one in every 20 of them is depressed. According to the World Health Organization (WHO), depressive disorders affect more than 350 million people globally of all ages. Depression is one of the most severe but prevalent mental illnesses in the world. (Also known as depressive disorder or clinical depression).

The inability to perform necessary everyday activities for a minimum period of two weeks could be caused by depression's severe impairments. Apart from either a low mood or loss of interest, at least four other symptoms such as issues with concentration, self-image, food, energy or persistent thoughts related to death or suicide must exist consistently for two weeks. Depression is receiving more and more attention from a variety of connected areas due to its dangers and the recent increase in occurrence. Depression may be treated with medicine, psychotherapy, and other therapeutic techniques even if it is a serious condition. The more quickly that therapy can start, the better it is. The detection of depression in its initial phase is imperative in managing the condition and curtailing the societal and financial burdens associated with this disorder.

The primary sources of data utilized in conventional methods of diagnosing depression include patient self-reports during clinic interviews, behaviors noticed by family or friends, and questionnaires like the Patient Health Questionnaire (PHQ-9) and the Beck Depression Inventory. (BDI-II). However, because they all rely on subjective evaluations, the outcomes might vary depending on the situation or the setting. To arrive at a somewhat objective diagnosis, numerous clinical specialists must be involved. Early-stage diagnoses and reassessments for monitoring therapy results are frequently restricted and time-consuming as the number of depressed individuals rises. In light of this, it is envisaged that automatic potential depression risk detection or recognition based on machine learning will enable objective and speedy diagnosis, assuring excellent clinical treatment quality and considerably reducing the probability of harm in actual life.

Depression-related behavior disorder-based indications for depression detection, such as voices, facial expressions, gestures, gaits, and eye movements, are becoming more prevalent under its impact. This study focuses on analyzing facial expressions to detect people who may be at risk for depression. Videos or photos are mostly used in studies over depression based on facial expressions. With video-based technology relying heavily on picture processing through converting motion footage into static shots, these limitations hinder its accuracy and reliability. The recognition performance will be impacted if these aspects are not properly addressed.

Therefore, to identify stress or depression levels in a certain facial image in real-time, this model will be employing Convolutional Neural Network (CNN) and face recognition techniques that might be done via the haar cascade approach. Deep Learning (DL) has been used to extract a representation of depression signals from video and translate them into frames of visuals for a depression diagnosis in order to improve existing medical therapy. Here are the objectives mentioned that are being pursued:

- To detect refined depression level statistics from facial expressions.
- To close the gap created by the unbalanced doctor-patient ratio.
- To improvise the performance and accuracy through large-scale datasets.

1.1 MOTIVATION AND MAIN CONTRIBUTION

According to a meta-analysis of 41 studies, general practitioners and sometimes expert psychiatrists can only identify depression in 47% of the cases. More precise tools are required, with the goal of assisting psychiatrists in their decision-making rather than replacing humans in the diagnosis process. We will provide psychiatrists or psychologists a piece of software to use in order to assess the precise level of depression a patient is likely to be experiencing. This software will provide a number along some suggestion as per scale from 1 to 10 as an output which depicts a simple mental pain health scale, by dividing the scale from:

- Mild Depression – Range 1 to 3
- Moderate Depression – Range 4 to 6
- Severe Depression – Range 7 to 10

Advantages of Proposed Solution:

- Depression detection using computer vision has higher accuracy than average general human practitioners, requiring only 30 to 40 seconds of image data to get processed and detect the depression level.
- Using this software along with their own conventional approaches will give psychiatrists a new perspective on decision-making and a quick, accurate understanding.
- A one-time software investment can produce superior outcomes for the long-term diagnosis of a patient at various stages of depression.

The remaining sections of the article is arranged as follows: Section 2 provides a quick review of similar work in this field; Section 3 about the research methodology and description of dataset used. Section 4 gives us insight about results and decision and Section 5 concludes up the paper with its future scopes.

2 RELATED WORKS

The identification of facial expressions has attracted the attention of certain academics. Fan and Tjahjadi provided a framework for recognizing facial expressions that combines CNN and custom characteristics. They discovered that the neural network could extract texture information from face patches to provide outstanding recognition results, and that the incorporation of CNN had a positive impact on the detection of facial expressions [5].

Reddy et al. proposed an organic fusion of deep learning features with the manual face expression detection technique. In trials, they confirmed the method's applicability in natural settings, which showed the method's efficacy when deep learning and manual production were combined [6]. Liang et al. proposed the old handcrafted facial representations, could only show superficial characteristics. They presented the Patch Attention Layer of embedding handmade features, a novel approach to facial emotion detection that is based on patches of interest, in order to overcome this constraint and learn the local shallow characteristics of each patch on face pictures [7].

Jain et al. (2018) presents the theory in which Convolutional layer and Recurrent Neural Network (RNN) were combined to extract information from face photos, hybrid convolutional-recursive neural network for facial emotion detection [8]. Avots et al. discovered human emotions by analysing audio-visual data. Also, they used the Viola-Jones facial recognition algorithm, CNN (AlexNet), and multiple datasets as test sets to classify the emotions on face photos [9].

Li et al. based on the L1 norm for face recognition, built a two-dimensional principal component analysis network using deep learning, tested it against a database of facial images, and concluded that the network was reliable [10]. Bernhard et al. realizes since emotions significantly affect human decision-making so, they used deep learning to enhance the outcomes of emotion identification. They discovered that RNNs and transfer learning outperformed standard machine learning, which was heavily influenced by applications for emotion computing [11]. Kumar et al. covered the modelling of deviant facial expressions based on computer vision tasks and emotional aberrations. They discovered that deep CNN might be crucial in the training and classification of face expressions, which gave visual surveillance systems a new visual modelling technique [12].

Mishra et al. used CNN to identify various emotions and intensity levels on human faces, laying the groundwork and providing support for further research on computer emotion identification [13]. Björn Schulle et al. offered the first combined open Audio/Visual Emotion and Depression identification Challenge, AVEC 2013. It tackles two sub-challenges: the estimate of a self-reported state of depression, and the detection of the valence and arousal of the emotional dimension in continuous time and value [14]. Young-Shin Lee et al. put forth a model based on artificial intelligence (AI) that can help in the diagnosis of depressive illness. Fast region-based convolutional neural networks (R-CNN), a deep learning technique that understands vector-based information, can be used to create a model that aids in the diagnosis of depressive disorder by analysing changes in the position of the eyes and lips as well as inferring emotions from a collection of photos [15].

Sharifa Alghowinem et al. in order to conduct a binary classification job, examined the effectiveness of eye movement characteristics collected from face films using Active Appearance Models. (Depressed vs. non-depressed). Model discover that eye movement low-level features provided 75% accuracy when utilising statistical measurements with SVM classifiers over the full interview and 70% accuracy when using a hybrid classifier comprising Gaussian Mixture Models and Support Vector Machines [16]. P Ramesh Naidu et al. in order to take advantage of the fact that eye, mouth, and head movements differ from normal situations when a person is under stress, proposed an algorithm that can recognise stress from photographs taken with a camera and a deep neural network that receives facial landmarks as input [17]. Lang He et al. proposed a promising psychological investigation by discovering some variations in speech and facial expression between depressed patients and healthy people. They provide the objective markers for automated depression estimate in the databases [18].

Karen Schepman et al. following a standardised evaluation of diagnostic and mood symptoms, conducted an experiment in which participants supplied accurate data after being given a computerised face emotion detection test. The AVEC2014 dataset was selected so that studies utilising unprocessed visual and audio data could be conducted [19]. Ninad Mehendale put forth emotion recognition and proposed that it is crucial to creating successful human-computer interactions. The input is a picture. Prior to identifying the face, the face is first detected in the picture, from which key traits are then extracted. Extraction of the expression features from the picture is the next stage. The classifier is then given the retrieved characteristics in order to classify the output as expressions [20]. Bhavna Singh Parihar et al. says that one of the most severe mental diseases, depression can present with a variety of symptoms in different people, making it challenging to diagnose. Here, we have created a CNN model that analyses a person's facial traits to determine whether or not they are depressed [21].

Qian Chen et al. propose a sequential fusion approach for facial depression identification in order to concurrently learn face movements and appearance in a single framework, a chained-fusion technique is presented for mining the associated and complimentary depression patterns in multimodal learning. We demonstrate how this type of sequential fusion may offer a probabilistic view of the model complementarity and correlation between two separate data modalities for enhanced depression identification [22]. Asim Jan et al. asserted that DL approaches can be useful in the field of mental health since they recognise the value of gathering thorough information to characterise the various psychiatric diseases. Visual and verbal data are essential for building an effective artificial system for recognising sadness since a system utilising a camera and microphone can quickly collect them [23].

S. Modi et al. put forth that the primary driving force behind the effort is to increase the correctness of the following model statement, which might have an effect on further research. Additionally, a comparison of the used model and the transfer learning model will provide the study effort the necessary innovation [24]. Weitong Guo et al. comes up with a unique method for identifying probable depression risk and was based on two separate deep belief network (DBN) models. While the other model extracts 3D dynamic characteristics from 3D face points gathered by a Kinect, the first model extracts 2D appearance features from facial photos captured by an optical camera. The two models are combined to provide the final decision outcome. Finally, we assess each deep model on our constructed dataset [25].

G. Giannakakis et al. says, video-recorded facial signals can be used to identify and analyse emotional states associated with stress and anxiety. Mainly concentrated on non-voluntary and semi-voluntary facial signals to more accurately evaluate the emotion representation [26]. Nandita Sharma et al. developed a theory where stress is being modelled using Bayesian networks, artificial neural networks, and support vector machines [27]. Amir Hasanbasic et al. used wearable sensors to monitor 10 students in order to gauge their degrees of exam-related stress. Different categorization techniques were utilised as input with characteristics of the ECG and electro dermal activity signals [28]. Andre Teixeira Lopes et al. for the purpose of recognising facial expressions suggested a straightforward approach that combines well-known techniques like Convolutional Network and certain picture pre-processing processes [29].

Octavio Arriaga et al. proposed a real-time vision system that can recognise faces, identify genders, and identify emotions [30]. Ali Mollahosseini et al. suggested a deep neural network design to handle the FER problem across several well-known standard face datasets. Our network has four Inception layers after two convolutional layers, each followed by max pooling. The network, which consists of a single component, classifies recorded face pictures as input into either the six fundamental expressions or the neutral expressions. On seven publicly accessible facial expression databases—MultiPIE, MMI, CK+, DISFA, FERA, SFEW, and FER2013—we conducted extensive studies [31].

In summary, deep learning may considerably increase the recognition efficacy of facial expression, especially the hybrid model, whereas the classic manual approaches are no longer appropriate for the present study on facial expression recognition. Even while facial expression recognition has made considerable progress in the past, there aren't many research that pair it with psychological analysis.

3 RESEARCH METHODOLOGY

Let's understand this segmentation with the help of a flowchart which depicts the flow of tasks processing in the system from input to output. As seen in (Fig. 1) a proper procedure has been showed as how the program will work and what could be the possible outcomes, which are Mild Depression, Moderate Depression and Severe Depression. A CNN model will be learning and getting trained based on the input images in the ML Classification stage.

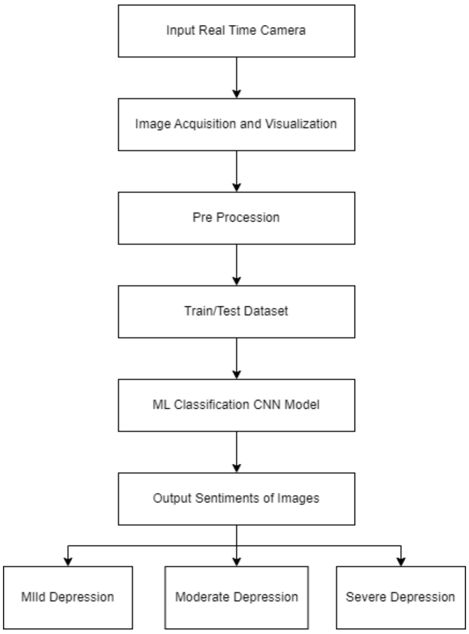

Fig. 1 - Flow Chart for Depression Detection

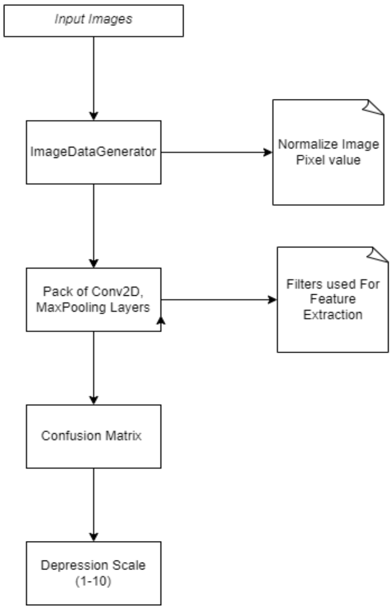

Fig. 2 - System Architecture for Depression Detection

The process of analyzing emotions from a video stream involves several stages. The algorithm first broadcasts a live video feed, after which it gathers a certain number of frames for use. Then, a CNN model that was trained and validated on a big dataset of images representing six to seven different emotions is compared to these frames.

The ImageDataGenerator script is then used to turn all grayscale-textured pictures into datasets, and batches of tensor image data are generated with real-time data augmentation. The training dataset is fitted across 60 epochs with various layers of Conv2D and MaxPooling, with a learning rate of 0.0001 and category cross entropy loss. There are separate training and validation components for the dataset. With a 25% dropout rate, the set of available classes are used to map labels to pictures.

A confusion matrix is created to provide an accuracy score for each image's predicted emotion, and graphs showing the model's accuracy and loss are shown during training and testing. After utilizing the CNN model, the model analyzes frames and identifies emotions in each of them and there after retrieve the emotions for each image in the stack. It has the ability to detect positive and negative emotions and ascertain the degree of these emotions. Subsequently, it can also estimate whether there has been a shift in emotions over time. The percentage level of emotion falls under a certain threshold value, then the model generates results displaying depression levels as mild, moderate, or severe and providing an appropriate suggestion.

3.1 DATASET DESCRIPTION:

The FER-2013 database, often known as FERPlus, has about 35k face expressions based on 7 fundamental expressions. Published in 2013 as part of a Kaggle challenge. This dataset is taken from Kaggle and the major contributor is Mr. Manas Sambare. The web-based photos are gathered, transformed to grey scale, and scaled to (48 × 48). Since this database reports a 68% ±5% human accuracy, theoretically it might be mislabeled. However, this model used it as pre-training data because it is a sizable spontaneous collection of facial expressions. The faces have been automatically registered such that each face roughly fills the same amount of space in each image and is roughly centered.

Based on the emotion shown in the facial expression, it is categorized into one of seven types, (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Neutral, 5=Sad, 6=Surprise). The training set consists of 28,709 examples.

Fig. 3 - Facial Expression in FER-2013

4 RESULT AND DISCUSSION

Here two activation functions are tested on the same dataset to find out the best results. First function is **ReLU**, which is a piecewise linear function that will output the input directly if it is positive, else, it will output zero and other is **SIGMOID** that is a non-linear function which maps any number between 0 and 1, inclusive, to itself. The figure given below is the over-fitted model and accuracy is 72.23% and split is of (80-20), but still it’s giving false positive results. So, this model was discarded and the test-train split is changed.

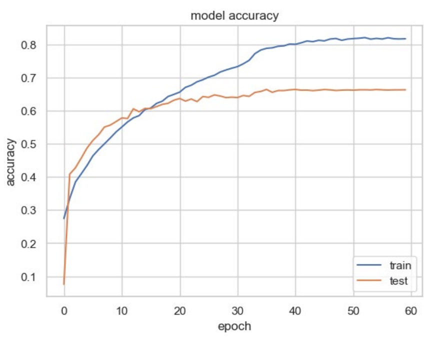

Fig. 4 - Model Accuracy (80-20) using ReLU

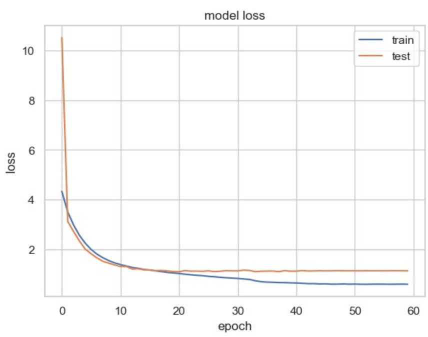

Fig. 5 - Model Loss (80-20) using ReLU

In the graph (fig. 5) model loss is decreased after many epochs retries which makes it less optimized. Feature extraction is stabilized between 10 and 20 x scale values.

Here SIGMOID non-linear function is used to attain a higher accuracy and the splitting ratio of train and test is set to (70-30) which still resulted in unstable test case accuracy as shown in the graph (fig. 6) and this makes it non-reliable.

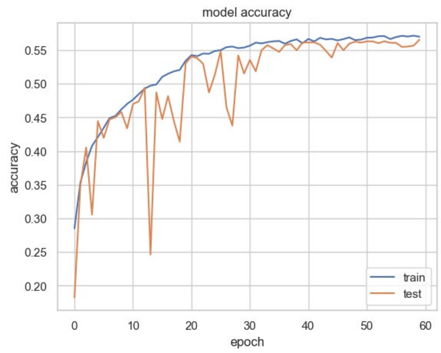

Fig. 6 - Model Accuracy (70-30) using SIGMOID

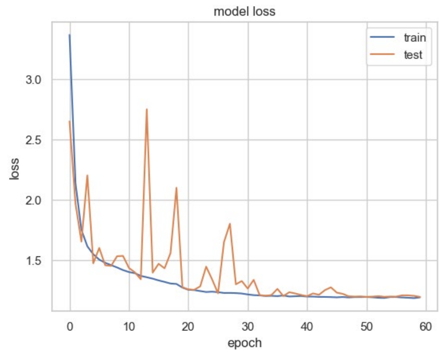

Fig. 7 - Model Loss (70-30) using SIGMOID

But here as shown (fig. 7) that model loss is not getting stabilized so the feature extraction is very unoptimized. So, this model is also not very much reliable either.

Now the same splitting ratio for training and testing of (70-30) is reused, but this time again using the ReLU function on the same dataset. And as the result it does provide us with a much higher accuracy of 83.88% and very less model loss as seen in (fig. 8) and (fig. 9).

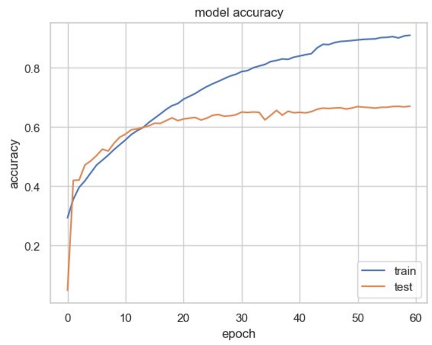

Fig. 8 - Model Accuracy (70-30) using ReLU

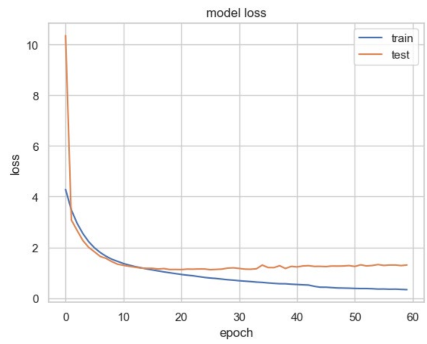

Fig. 9 - Model Loss (70-30) using ReLU

Table. 1 - Accuracy values from all functions used

Function Used	Split Ratio	Final Train Accuracy (%)	Validation Accuracy (%)
ReLU	80-20	72.23	63.66
SIGMOID	70-30	59.85	56.51
ReLU	70-30	83.88	64.03

So, at last, it is decided to go with ReLU function in our model as it is giving the best accuracy for the (70-30) split and with vey less model loss and high feature extractions.

5 CONCLUSION AND FUTURE SCOPE

The research study found that the model was successful in predicting the nature of images, but improvements were needed for future purposes. The suggested AI treatment had potential applications in detecting drowsiness in drivers and monitoring elderly patients' tension and anxiety using emotion detection. The model could also provide audio-based refinement for depression detection and suggest mental activities, a healthy diet, and music to promote stability and prevent depressive thoughts. Further enhancements were necessary to integrate the model into an Android or iOS application for identifying potential depression in friends and family.

REFERENCES

[1] Health & Consumer Protection Directorate General, “Mental health in the EU,” 2008.

[2] M. Marcus, M. T. Yasamy, M. van Ommeren, and D. Chisholm, “Depression, a global public health concern,” pp. 1–8, 2012.

[3] World Health Organization, “Depression and other common mental disorders: global health estimates,” Tech. Rep., 2017.

[4] D. D. Luxton, Artificial intelligence in behavioral and mental health care. Academic Press, 2015.

[5] X. Fan and T. Tjahjadi, “**Fusing dynamic deep learned features and handcrafted features for facial expression recognition**,” Journal of Visual Communication and Image Representation, vol. 65, Article ID 102659, 2019.

[6] G. Viswanatha Reddy, C. Dharma Savarni, and S. Mukherjee, “**Facial expression recognition in the wild, by fusion of deep learnt and hand-crafted features**,” Cognitive Systems Research, vol. 62, pp. 23–34, 2020.

[7] X. Liang, L. Xu, J. Liu et al., “**Patch attention layer of embedding handcrafted features in cnn for facial expression recognition**,” Sensors, vol. 21, no. 3, p. 833, 2021.

[8] N. Jain, S. Kumar, A. Kumar, P. Shamsolmoali, and M. Zareapoor, “**Hybrid deep neural networks for face emotion recognition**,” Pattern Recognition Letters, vol. 115, pp. 101–106, 2018.

[9] E. Avots, T. Sapiński, M. Bachmann, and D. Kamińska, “**Audiovisual emotion recognition in wild**,” Machine Vision and Applications, vol. 30, no. 5, pp. 975–985, 2019.

[10] Y. K. Li, X. J. Wu, and J. Kittler, “**L1-2D2PCANet: a deep learning network for face recognition**,” Journal of Electronic Imaging, vol. 28, no. 02, Article ID 023016, 1 page, 2019.

[11] B. Kratzwald, S. Ilić, M. Kraus, S. Feuerriegel, and H. Prendinger, “**Deep learning for affective computing: text-based emotion recognition in decision support**,” Decision Support Systems, vol. 115, pp. 24–35, 2018.

[12] R. K. Kumar, J. Garain, D. R. Kisku, and G. Sanyal, “**Estimating attention of faces due to its growing level of emotions**,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 2018.

[13] S. Mishra, G. R. B. Prasada, R. K. Kumar, and G. Sanyal, “**Emotion recognition through facial gestures - a deep learning approach**,” in Proceedings of the International Conference on Mining Intelligence and Knowledge Exploration, , Springer, 2017.

[14] Valstar, M., Schuller, B., Smith, K., Eyben, F., Jiang, B., Bilakhia, S., ... & Pantic, M. (2013, October). Avec 2013: the continuous audio/visual emotion and depression recognition challenge. In Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge (pp. 3-10).

[15] Lee, Y. S., & Park, W. H. (2022). Diagnosis of depressive disorder model on facial expression based on fast R-CNN. Diagnostics, 12(2), 317.

[16] Alghowinem, S., Goecke, R., Wagner, M., Parker, G., & Breakspear, M. (2013, September). Eye movement analysis for depression detection. In 2013 IEEE International Conference on Image Processing (pp. 4220-4224). IEEE.

[17] Naidu, P. R., Sagar, S. P., Praveen, K., Kiran, K., & Khalandar, K. (2021, November). Stress Recognition Using Facial Landmarks and Cnn (Alexnet). In Journal of Physics: Conference Series (Vol. 2089, No. 1, p. 012039). IOP Publishing.

[18] He, L., Niu, M., Tiwari, P., Marttinen, P., Su, R., Jiang, J., ... & Dang, W. (2022). Deep learning for depression recognition with audiovisual cues: A review. Information Fusion, 80, 56-86.

[19] Schepman, K., Taylor, E., Collishaw, S., & Fombonne, E. (2012). Face emotion processing in depressed children and adolescents with and without comorbid conduct disorder. Journal of abnormal child psychology, 40, 583-593.

[20] Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). SN Applied Sciences, 2(3), 446.

[21] B. S. Parihar, S. S. Satam, S. S. Satam, and K. Dange, “CNN model for Depression Detection using JAFFE Dataset”, Pices, vol. 4, no. 6, pp. 135-139, Oct. 2020.

[22] Qian Chen, Iti Chaturvedi, Shaoxiong Ji, Erik Cambria, “Sequential fusion of facial appearance and dynamics for depression recognition”, Pattern Recognition Letters, Volume 150, 2021, Pages 115-121, ISSN 0167-8655.

[23] A. Jan, H. Meng, Y. F. B. A. Gaus and F. Zhang, "Artificial Intelligent System for Automatic Depression Level Analysis Through Visual and Vocal Expressions," in IEEE Transactions on Cognitive and Developmental Systems, vol. 10, no. 3, pp. 668-680, Sept. 2018, Doi: 10.1109/TCDS.2017.2721552.

[24] S. Modi, M.H. Bohara Facial Emotion Recognition using Convolution Neural Network 2021 5th International Conference on Intelligent Computing and Control Systems.

[25] Guo, Weitong & Yang, Hongwu & Liu, Zhenyu & Xu, Yaping & Hu, Bin. (2021). Deep Neural Networks for Depression Recognition Based on 2D and 3D Facial Expressions Under Emotional Stimulus Tasks. Frontiers in Neuroscience. 15. 609760. 10.3389/fnins.2021.609760.

[26] Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., ... & Tsiknakis, M. (2017). Stress and anxiety detection using facial cues from videos. Biomedical Signal Processing and Control, 31, 89-101.

[27] Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Computer methods and programs in biomedicine, 108(3), 1287-1301.

[28] Hasanbasic, A., Spahic, M., Bosnjic, D., Mesic, V., & Jahic, O. (2019, March). Recognition of stress levels among students with wearable sensors. In 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-4). IEEE.

[29] Lopes, A. T., De Aguiar, E., & Oliveira-Santos, T. (2015, August). A facial expression recognition system using convolutional networks. In 2015 28th SIBGRAPI conference on graphics, patterns and images (pp. 273-280). IEEE.

[30] Arriaga, O., Valdenegro-Toro, M., & Plöger, P. (2017). Real-time convolutional neural networks for emotion and gender classification. arXiv preprint arXiv:1710.07557.

[31] Mollahosseini, A., Chan, D., & Mahoor, M. H. (2016, March). Going deeper in facial expression recognition using deep neural networks. In 2016 IEEE Winter conference on applications of computer vision (WACV) (pp. 1-10).