REAL-TIME STRESS DETECTION USING CNN IN DEEP LEARNING

Tandra Debarati Shome

Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Chennai, India

Laxmi Maurya

Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Chennai, India

ABSTRACT

Stress has become a part of everyday life, affecting people of all ages. It creates significant challenges for well-being and productivity. Despite advancements in physiological techniques for stress detection, there are still hurdles in making these solutions real-time, affordable, and accessible to everyone. Psychological stress is closely tied to emotions, and understanding this connection plays a key role in analyzing human behavior, particularly in computational psychology. While deep learning techniques, like Convolutional Neural Networks (CNNs), have shown great promise in detecting facial emotions from images, their potential for identifying mental stress remains underexplored. The system provides a holistic approach to understand and evaluate stress through images and video processing.

Keywords- Stress detection, CNN model, Emotions classes, image processing.

I. INTRODUCTION

In the present era, stress has become an unavoidable element of daily lifestyles, affecting humans throughout all age groups and professions. From students and younger experts to the ones inside the middle-aged adults balancing a couple of obligations, strain has emerged as a prominent issue that affects no longer only man or woman fitness but also the well-being financial system of countries like India. Despite the increasing recognition and physiological improvements in stress recognition, there are still sizeable obstacles to detect pressure in real-time in a very expensive way. Here artificial intelligence comes into picture where technology can be used to detect the stress level through different kind of emotions. In computational psychology, stress and emotions are closely connected, and knowledge this relationship is prime to interpreting human behavior.

This study has made widespread development in using deep learning strategies to detect emotions from facial expressions, however the express use of such strategies to discover psychological stress has no longer yet been absolutely explored. Emotions like anger, disgust, fear, happy, sad, surprise and neutral are regularly clear signs of someone's mental state, and these emotional cues are meditated in facial musculature. By studying these expressions through image processing, insights have been collected into both emotional and stress degrees. This paper proposes a hybrid model that leverages a Convolutional Neural Network (CNN) to detect stress with the aid of reading facial musculature from pictures and video. The CNN model is trained to categories facial expressions into discrete emotion categories consisting of anger, disgust, neutral, fear, sad, happy, and surprise, which are critical signs of stress degrees.

Emotion reputation is accomplished using a real-time image dataset wherein the CNN is trained to identify emotional states based on facial expressions. This hybrid approach no longer simplest detects emotions but also provides a real-time classification of emotions by integrating deep learning techniques with real-time image processing, and also gives suggestions like take a deep breath, have a glass of water to reduces the stress level. This has a look at pursuits to bridge the distance in stress detection research, offering a scalable and powerful solution for real-time stress monitoring.

II. LITERATURE SURVEY

A study on the use of facial signals for stress and anxiety detection was conducted by Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., & Tsiknakis, M. Stress signs were identified by analysing facial features such as head movements, jaw movements, and eye activity. In comparison to neutral or calm states, the study discovered that stressful conditions caused considerable alterations in these traits. The results point to the possibility of non-invasive stress assessment through the use of facial clues [13].

A real-time Convolutional Neural Network (CNN) architecture for emotion and gender recognition is proposed by Singh, J., Singh, A., Singh, K. K., Lal, B., Verma, H., Samudre, N., & Raperia, H. Convolutional, pooling, and fully linked layers make up the CNN architecture, which is intended to extract and categorize characteristics from facial photos. Using the CK+ and AffectNet datasets, the model demonstrated high recognition accuracy for both gender (97.58%) and emotion (91.35%). Future studies would possibly observe integration with realistic programs and switch studying [14].

Sharma, N., and Gedeon, T.'s research look at investigates the region of strain categorization and recognition. It goes over modeling stress stages computationally and using sensors to screen pressure. The observe highlights the ability advantages of stress recognition in a number of contexts and makes guidelines for in addition research, together with combining one-of-a-kind sensors, using clustering procedures, and growing greater complicated stress fashions.[12]

The goal of the studies via Jaiswal, D., Mukhopadhyay, S., & Sharma, V. is to use wearable Electrodermal interest (EDA) sensors to construct a useful resource-efficient strain detection gadget. Deep Neural Networks (DNNs) are used by the machine to classify pressure ranges and examine EDA records. The paintings investigate automation strategies such as Neural structure search (NAS) to optimize the version for actual-time inference on confined devices. The findings display that, whilst significantly decreasing model length and computational complexity, NAS-generated models obtain accuracy degrees comparable to humanly built models, making them appropriate for on-device stress detection. As a way to enhance prediction accuracy, future research will integrate extra physiological alerts and contextual information. It will also investigate hybrid systems that combine NAS with conventional machine learning methods [1].

Using HRV data taken from the WESAD dataset, Benita, D. S., Ebenezer, A. S., Susmitha, L., Subathra, M. S. P., & Priya, S. J.'s research study suggests a stress detection model. The have a look at focuses on a 1D CNN model and employs Random woodland for function choice and hyperparameter tuning to enhance the version's overall performance. The study emphasizes the fee of non-invasive pressure tracking techniques, particularly those who employ PPG sensors. The version's first-rate type accuracy of 98% for stress and the performance of function choice in performance optimization are crucial discoveries. The examine advances stress detection era, which has capability uses in some of industries which includes education, healthcare, and administrative center wellbeing [2].

Hafeez, M. A., & Shakil, S. use EEG facts to discover ability biomarkers of their examine work, which examines the association among student stress ranges and exam performance. Exams with a time constraint had been proven to dramatically boost strain degrees and have a unfavorable impact on overall performance. Theta, alpha, and beta frequency bands have been shown to alter at some stage in an EEG exam, which is connected to pressure. The usage of EEG facts, the researchers categorized strain tiers the usage of LSTM and CNN fashions, with 70.67% and 90.46% accuracy charges, respectively. The look at involves the belief that CNN is a good method for classifying strain and that EEG-based brainwave pictures can be beneficial indicators for pressure identity [3].

Extremely-brief heart charge variability (HRV) evaluation is a unique technique to pressure detection supplied in the article with the aid of Adarsh, V., & Gangadharan, G. R. The approach reduces version length and computational complexity substantially while achieving excessive accuracy and performance thru the aggregate of explainable graph convolutional networks with pruning and quantization techniques. Because of this, it can be implemented on devices with limited resources. The suggested methodology exhibits encouraging outcomes on the WESAD and SWELL datasets, providing a beneficial resolution for practical stress identification uses [4].

An IoT-based system for music therapy and stress detection is proposed in this work. According to Mate, A., Narkhede, P., Kulkarni, P., Lokhande, H., Prasad, J., & Prasad, R., the system applies a machine learning model to determine stress levels and uses wearable sensors to gather physiological data (blood pressure, heart rate). The technique suggests suitable ragas (ancient Indian musical modes) to reduce stress and enhance well-being based on the degree of tension. The study shows how machine learning and the Internet of Things can be used to create individualized stress management programs [5].

In this study, Deulkar, Narvekar, Gandhi, P., Gada, D., & Kamath, S. explore the possibility of utilizing social media data to identify stress. The study evaluates the efficacy of various text representation strategies (TF-IDF and Bag-of-Words) and machine learning algorithms (Logistic Regression, SVM, Random Forest, Naive Bayes) in recognizing stress patterns by examining text data from Reddit and Twitter. Additionally, deep learning methods including RoBERTa, LSTM, BERT, and RNN were investigated. The results indicate that social media data can be a useful tool for stress identification, with deep learning approaches outperforming more conventional machine learning techniques [6].

According to Tasnim, M., Ramos, R. D., Stroulia, E., & Trejo, L. A., this research presents a longitudinal dataset for the purpose of using speech to predict mental health problems (stress, anxiety, and depression). The dataset carries self-stated DASS-21 rankings and speech samples from forty topics that have been amassed over a -month period. According to the look at, there's an instantaneous hyperlink between the seriousness of mental fitness issues and following the pointers for amassing data. Its miles suggested to use a 1D CNN version with VGG-19 features in a system mastering pipeline to are expecting DASS-21 rankings from speech. In relation to forecasting the depth of stress, tension, and depression, the version performs quite nicely [7].

In this look at, physiological indicators recorded by a wristband are utilized by Ramachandra, R. A., Santhosh, J., Dengel, A., & Ishimaru, S. to take a look at stress detection among college students. To create a number stress tiers, participants finished responsibilities of differing complexity. To categorize strain ranges, deep mastering fashions (FCN, ResNet,

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

LSTM) had been utilized. The have a look at observed that, in particular for LSTM, nice-tuning the models with a tiny quantity of unknown records significantly accelerated category accuracy. The effects underscore the capability of deep studying in exactly figuring out stress and the importance of custom designed calibration in enhancing generalization [8].

This study looks into how yoga influences college students' stress ranges and looks into the usage of gadget gaining knowledge of to forecast changes in stress. The DASS-21 questionnaire changed into used to gauge the pressure ranges of a collection of students who underwent a 6-week yoga intervention, in step with Ghosh, S., Tripathi, ok., Garg, A., Singh, D., Prasad, A., Bhavsar, A., & Dutt, V. Based on pre-intervention data, system gaining knowledge of fashions (such as synthetic Neural Networks, Linear Regression, Random forest, and aid Vector Regression) had been evolved to are expecting modifications in stress. The results of the look at showed that the Random Forest model changed into the maximum successful in forecasting changes in strain ranges and that yoga intervention dramatically lowered pressure stages [9].

The authors of this paper, Fontes, L., Machado, P., Vinkemeier, D., Yahaya, S., Bird, J. J., & Ihianle, I. K., suggest a deep getting to know and face video method for far flung pressure detection. The technique uses pyVHR to extract heart rate data from face videos, and it classifies stress using 1D-CNN, GRU, and LSTM models. The study shows the efficacy of the suggested method for remote stress monitoring and achieves 95.83% accuracy on the UBFC-Phys dataset [10].

A hybrid deep learning model is proposed by Tanwar, R., Phukan, O. C., Singh, G., Pal, P. K., & Tiwari, S. for the identification of stress utilizing physiological markers (ECG, EDA). The model efficiently extracts features and ranks pertinent data by combining CNN, LSTM, and an attention mechanism. Surpassing conventional approaches, the model achieves high accuracy (92.70%) and F1-score (90%) on the WESAD dataset. The study emphasizes how crucial attention mechanisms and multimodal fusion are to precise stress identification [11].

III. PROPOSED METHODOLOGY

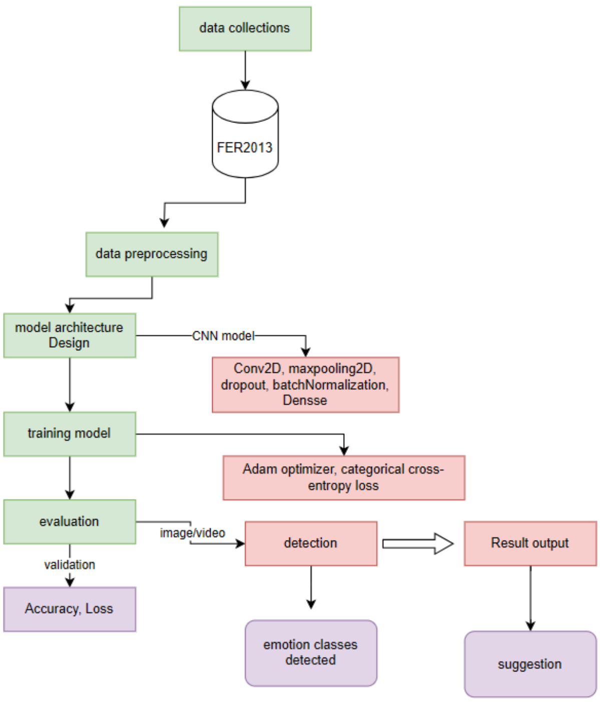


Fig.1 Architecture of the Proposed Method

A. Data collection

The FER2013 (facial expression reputation 2013) dataset is a broadly used dataset for facial emotion detection duties. It was at the beginning added at some stage in the ICML 2013 competition on facial features reputation. The dataset incorporates 35,887 grayscale pics of 48x48 pixels, wherein each image captures a single human face with considered one of seven emotional expressions: angry, Disgust, fear, happy, sad, surprise, and neutral. Those images were accumulated from diverse assets and preprocessed to be in a uniform length, supplying a balanced but various set of facial expressions that may be used for schooling, validation, and trying out in deep gaining knowledge of models. The dataset is split into 3 subsets; Training Set: 28,709 snap shots, used to train gadget getting to know models. Validation Set: three,589 pix, normally used to track model hyperparameters. Test Set: three,589 photographs, for evaluating the version's performance.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

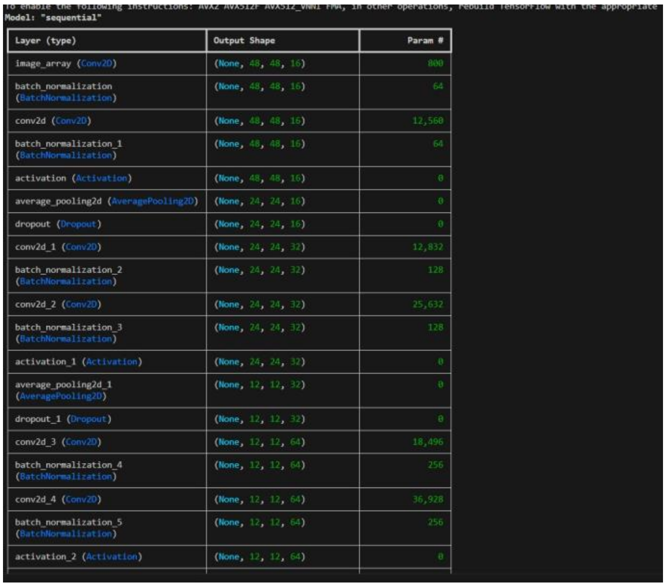
Each image is labeled with an integer similar to one of the seven emotions, making it a supervised getting to know undertaking. The surprisingly low resolution of the images (48x48) demanding situations models to seize facial features and expressions no matter the dearth of best information, making this dataset treasured for both benchmarking and actual-world application research in emotion class obligations. Given its shape, FER2013 has been widely used with deep gaining knowledge of strategies like convolutional neural networks (CNNs) to increase and evaluate fashions for emotion detection. Its large range in facial pics, masking a selection of age businesses, ethnicities, and lights situations, makes it suitable for generalization to actual-international eventualities.

B. Data preprocessing

Preprocessing is necessary to improve model performance in emotion recognition for the FER2013 dataset. Normalization and Rescaling: To guarantee consistent input for model training, the grayscale pixel values (0–255) are scaled to [0,1] by dividing by 255.Data augmentation involves applying methods such as random flips, rotations, and shifts to improve sample variety and minimize overfitting, which facilitates the generalization of the model. Images (48x48) are reshaped to (48x48x1) in order to conform to the CNN architectures' needed input format. Label Encoding: To help the model perform multi-class classification efficiently, integer labels (0–6) are one-hot encoded. Divide the dataset into training, validation, and test sets in order to conduct a fair assessment of performance and adjust the model.

C. Model Training

Convolutional layers are used to extract face characteristics, ReLU activation is used to seize intricate styles, and pooling layers are used to decrease spatial dimensions even as preserving important functions. These CNN model additives are used to educate the FER2013 dataset. A SoftMax layer generates chances for every of the seven emotion classes, at the same time as fully related layers cope with categorization. Accuracy is used as the performance metric and the Adam optimizer collectively with categorical pass-entropy loss are used to assemble the version. Preprocessed photos (48x48x1) are fed into the version in mini-batches (32–128) at some point of 20–50 epochs for training. Validation facts is applied to keep the model from overfitting. strategies for augmenting statistics, such rotating, zooming, and flipping at random, resource in improving generalization. The version is then further adjusted to maximize overall performance after being assessed using accuracy, confusion matrix, and F1 rating. the usage of CNN at the FER2013 dataset, this manner allows for reliable emotion identification.



Model: "sequential"

Layer (type)	Output Shape	Param #
image_array (Conv2D)	(None, 48, 48, 16)	800
batch_normalization (BatchNormalization)	(None, 48, 48, 16)	64
conv2d (Conv2D)	(None, 48, 48, 16)	12,560
batch_normalization_1 (BatchNormalization)	(None, 48, 48, 16)	64
activation (Activation)	(None, 48, 48, 16)	0
average_pooling2d (AveragePooling2D)	(None, 24, 24, 16)	0
dropout (Dropout)	(None, 24, 24, 16)	0
conv2d_1 (Conv2D)	(None, 24, 24, 32)	12,832
batch_normalization_2 (BatchNormalization)	(None, 24, 24, 32)	128
conv2d_2 (Conv2D)	(None, 24, 24, 32)	25,632
batch_normalization_3 (BatchNormalization)	(None, 24, 24, 32)	128
activation_1 (Activation)	(None, 24, 24, 32)	0
average_pooling2d_1 (AveragePooling2D)	(None, 12, 12, 32)	0
dropout_1 (Dropout)	(None, 12, 12, 32)	0
conv2d_3 (Conv2D)	(None, 12, 12, 64)	18,496
batch_normalization_4 (BatchNormalization)	(None, 12, 12, 64)	256
conv2d_4 (Conv2D)	(None, 12, 12, 64)	36,928
batch_normalization_5 (BatchNormalization)	(None, 12, 12, 64)	256
activation_2 (Activation)	(None, 12, 12, 64)	0

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

average_pooling2d_2 (AveragePooling2D)	(None, 6, 6, 64)	0
dropout_2 (Dropout)	(None, 6, 6, 64)	0
conv2d_5 (Conv2D)	(None, 6, 6, 128)	73,856
batch_normalization_6 (BatchNormalization)	(None, 6, 6, 128)	512
conv2d_6 (Conv2D)	(None, 6, 6, 128)	147,584
batch_normalization_7 (BatchNormalization)	(None, 6, 6, 128)	512
activation_3 (Activation)	(None, 6, 6, 128)	0
average_pooling2d_3 (AveragePooling2D)	(None, 3, 3, 128)	0
dropout_3 (Dropout)	(None, 3, 3, 128)	0
conv2d_7 (Conv2D)	(None, 3, 3, 256)	295,168
batch_normalization_8 (BatchNormalization)	(None, 3, 3, 256)	1,024
conv2d_8 (Conv2D)	(None, 3, 3, 7)	16,135
global_average_pooling2d (GlobalAveragePooling2D)	(None, 7)	0
predictions (Activation)	(None, 7)	0

Total params: 642,935 (2.45 MB)
Trainable params: 641,463 (2.45 MB)
Non-trainable params: 1,472 (5.75 KB)

Fig.2 Model Architecture Design

IV. RESULT

Fig.3 Positive emotion detected from image

Fig.4 Negative emotion detected from image

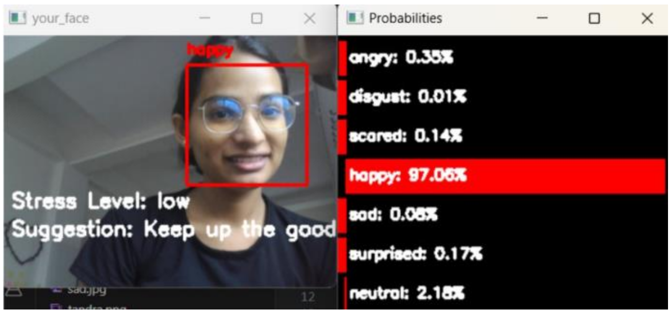


Fig.5 Positive emotion detected from real-time video

Fig 3, shows that device detects the face from the input picture and classifies the emotion as happy. The output indicates the possibilities for each emotion elegance, with the highest probability being 0.4946 for happiness. The system also calculates a corresponding pressure value of 1.748, indicating a low stress stage. The message "All properly! keep smiling!" confirms the detected high-quality emotion and presents a pleasant response.

Fig 4, identifies the emotion as unhappy, with a chance of 0.6914. The model also calculates a stress level of 8.368, suggesting better emotional distress. At the side of detecting the emotion, the model affords counseled moves to reduce strain, including yoga, being attentive to calming music, walking, or drinking water.

In fig 5, The CNN model correctly identifies the emotion from the facial photo in real time. The gadget processes the enter picture and classifies the dominant emotion as glad, with a completely high chance of 99.16%, as seen within the displayed chances. The output panel indicates the breakdown of chances for all other emotion classes, along with angry (0.35%), disgust (0.01%), scared (0.14%), sad (0.08%), surprised (0.17%), and neutral (2.18%), with happy being the maximum outstanding emotion detected. On the right-hand aspect, the real-time emotion detection is visualized via a bounding box around the detected face, with the label "happy" displayed in red above the face, confirming the correct classification by means of the model. Also stress level as low is detected and suggestion is shown as keep up the good.

V. **CONCLUSION**

Our model correctly detects a huge range of human feelings with an excellent accuracy of approximately 95% which shows excessive reliability in distinguishing between effective and terrible emotional states. The version is trained the use of the _mini_XCEPTION structure, combined with OpenCV for face detection. The model accurately classifies emotions such as happiness, sadness, anger, fear, surprise, disgust, and neutrality by thoroughly analyzing facial features. Beyond identifying basic emotions, it also quantifies the individual's stress level based on the detected emotion, offering a comprehensive assessment. As an example, negative emotions like anger, sadness, fear, and disgust notably make a contribution to better stress ranges, even as high-quality feelings together with happiness or neutrality bring about lower or negligible strain stages. Based at the calculated strain stage, the gadget indicates pressure-relieving activities like practicing yoga, meditating, paying attention to tune, going for a stroll, or ingesting water. If the detected emotion is happiness, the gadget provides advantageous reinforcement with a cheerful message like "All proper!" accompanied by a smiling emoji, enhancing the users enjoy and interaction.

In less difficult terms, the model no longer best reliably detects feelings but additionally provides personalized pointers to reduce stress, making it a complete tool for each emotion reputation and nicely-being improvement. This mixture makes it highly appropriate for actual-global programs together with pressure detection and intellectual health support structures.

REFERENCES

1. Jaiswal, D., Mukhopadhyay, S., & Sharma, V. (2024, March). Tinystressnet: On-device stress assessment with wearable sensors on edge devices. In 2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops) (pp. 166-171). IEEE.
2. Benita, D. S., Ebenezer, A. S., Susmitha, L., Subathra, M. S. P., & Priya, S. J. (2024, February). Stress Detection Using CNN on the WESAD Dataset. In 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC) (pp. 308-313). IEEE.
3. Hafeez, M. A., & Shakil, S. (2024). EEG-based stress identification and classification using deep learning. Multimedia Tools and Applications, 83(14), 42703-42719.
4. Adarsh, V., & Gangadharan, G. R. (2024). Mental stress detection from ultra-short heart rate variability using explainable graph convolutional network with network pruning and quantisation. Machine Learning, 1-28.
5. Mate, A., Narkhede, P., Kulkarni, P., Lokhande, H., Prasad, J., & Prasad, R. (2024, April). Musical Therapy For Stress Reduction Using Machine Learning And IoT. In 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon) (pp. 1-5). IEEE.
6. Deulkar, K., Narvekar, M., Gandhi, P., Gada, D., & Kamath, S. (2024, May). Evaluating the Influence of Text Representation techniques on Diverse Machine Learning Algorithms for Stress Detection in Social media users. In 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE) (pp. 1-7). IEEE.
7. Tasnim, M., Ramos, R. D., Stroulia, E., & Trejo, L. A. (2024, April). A Machine-Learning Model for Detecting Depression, Anxiety, and Stress from Speech. In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7085-7089). IEEE.

8. Ramachandra, R. A., Santhosh, J., Dengel, A., & Ishimaru, S. (2024). Enhancing stress detection for students: Exploring the impact of fine-tuning and user-specific data calibration in deep learning. International Journal of Activity and Behavior Computing, 2024(1), 1-21.

9. Ghosh, S., Tripathi, K., Garg, A., Singh, D., Prasad, A., Bhavsar, A., & Dutt, V. (2024, June). Predicting Stress among Students via Psychometric Assessments and Machine Learning. In Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 662-669).

10. Fontes, L., Machado, P., Vinkemeier, D., Yahaya, S., Bird, J. J., & Ihianle, I. K. (2024). Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques. Sensors, 24(4), 1096.

11. Tanwar, R., Phukan, O. C., Singh, G., Pal, P. K., & Tiwari, S. (2024). Attention based hybrid deep learning model for wearable based stress recognition. Engineering Applications of Artificial Intelligence, 127, 107391.

12. 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.

13. 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.

14. Singh, J., Singh, A., Singh, K. K., Lal, B., Verma, H., Samudre, N., & Raperia, H. (2024). Real-Time Convolutional Neural Networks for Emotion and Gender Classification. Procedia Computer Science, 235, 1429-1435.