International Journal of Innovation Studies

ENHANCED VGG-BASED APPROACH FOR FACIAL EXPRESSION-DRIVEN STRESS DETECTION

Jayashree. R[1], Venkatesh. S[2], Reena Rose. R[3,*], Amala Rose. R[4]

[1]Associate Professor, Department of Computer Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[2]Professor, Department of Biochemistry, Faculty of Medicine, University of Peace (UNPAZ), Dili, Timor Leste.
[3,*]Assistant Professor, Department of Computer Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[4]Assistant Professor, Department of Computer Science, St. John's College of Arts and Science, Ammandivilai, India.

Abstract

Stress is one of the major contributing factors to many diseases in today's lifestyle. Although it is not feasible to completely eradicate stress from our daily lives, we can reduce it to a certain level. This paper presents a cutting-edge method for stress detection that utilizes an advanced version of the VGG (Visual Geometry Group) architecture by analysing facial features obtained from images. Our objective is to accurately detect stress indicators, enabling prompt intervention and assistance. The enhanced VGG architecture significantly improves the model's capacity to identify subtle changes in facial expressions linked to mental stress. The performance of detection is further optimized through comprehensive training and validation on a varied dataset. This paper contributes to the evolution of stress detection techniques using modified VGG with Inception layers (VGGIL) and provides significant applications in healthcare, mental wellness, and human-computer interaction by harnessing the power of deep learning and facial recognition technology.

Keywords— Stress Detection, Visual Geometry Group, Mental Wellness, Image processing, Deep Learning, Convolutional Neural Networks.

1. INTRODUCTION

In this paper, we delve into numerous facets of the body's responses to psychologically and physiologically stressful stimuli to explore stress detection methodologies [1]. Stress is a ubiquitous aspect of modern life that affects individuals through work pressures, personal relationships, financial concerns, and health issues. Intricately, the physiological stress response is described by focusing on the activation of the body's "fight or flight" mechanism [2]. This process involves the initial signaling of the amygdala to the hypothalamus, which subsequently triggers the release of adrenaline and cortisol from the adrenal glands [15]. These hormones are responsible for the instant action of the body such as rapid heart rate, breathing rate, and increased blood flow to vital organs. Prolonged activation of the stress response system leads to chronic stress, a significant concern that can lead to severe health implications [5, 6]. Sustained exposure to preeminent stress hormones adversely affects cardiovascular health, mental well-being, and overall immune function. Alertness, muscle tension, and changes in blood flow patterns are some of the physiological changes that occur during stress [17]. Significantly, we differentiate between short-term adaptive anxiety responses and the harmful effects of chronic stress, emphasizing the need for effective stress management strategies. There are differences in stress responses among individuals, which highlight the roles of genetics, life experiences, coping mechanisms, and resilience levels. This recognition is crucial as it acknowledges that people vary in their susceptibility to stress-related health issues and their ability to bounce back from adversity [3]. Social support networks and emotional intelligence are the common factors contributing to resilience and are recognized as key elements in vindicating the impact of stress on individuals' health and well-being.

Furthermore, abundant stress management techniques are aimed at reducing stress levels and enhancing resilience. The techniques include mindfulness practices, physical activity, healthy lifestyle adoptions, community support, time management, and relaxation techniques. Each approach discusses its potential to promote relaxation, improve coping skills, and introduce a balanced approach to handling stressors. Therefore, the intricate interplay between stress, physiological responses, individual differences, and effective management strategies is crucial for maintaining overall health and well-being in the face of stress [4-6]

The proposed system introduces an innovative approach for stress detection based on facial expressions (images), using a modified VGG (Visual Geometry Group) architecture [26]. This system is designed to analyze facial features extracted from images to accurately identify signs of stress, aiming to facilitate early intervention and support for individuals experiencing stress [27, 28]. The modified VGG with Inception layers (VGGIL) architecture enhances the system's capability to capture subtle variations in facial expressions associated with stress, thereby improving its overall detection performance [26]. Through meticulous training and validation on a diverse dataset, the VGGIL approach achieves promising results in stress detection [27, 29]. By harnessing the power of deep learning and facial recognition technology, this research contributes to the advancement of stress detection methodologies [30]. The potential applications of this system span various domains, including healthcare, mental wellness, and human-computer interaction, offering valuable tools for addressing stress-related challenges in society.

The paper is organized as follows: Section 2 describes the literature survey, and Section 3 explains in detail on system architecture. Section 4 explains the VGGIL Model in detail. Section 5 elucidates on experimental results and analysis, and finally, Section 6 provides a conclusion and future enhancements.

2. Background

Several studies have advanced innovative methodologies for stress detection and management, leveraging diverse technologies ranging from facial recognition to physiological signals and social media analysis [1, 5, 11, 17]. Bhosale et al. propose a multifaceted approach integrating live video facial analysis, chatbot interaction, and mood-based music recommendations [10]. Their method utilizes Convolutional Neural Networks (CNN) to detect stress-related facial expressions in real-time videos (S. et al., 2023). The chatbot then engages users in text-based conversations to provide stress-relief advice and support, while Spotify's

API recommends music tailored to the user's mood [14]. Their system achieves promising accuracy rates of 83.79% for face stress detection and 79.23% for chatbot stress detection, indicating its potential as an effective stress management tool [4].

Dessai and Usgaonkar explore stress detection through textual data from Twitter using Text Mining and Natural Language Processing techniques [9]. They employ a CNN + LSTM classification model to analyze tweets and achieve an impressive accuracy of 92% in detecting factors indicative of depression [9, 12]. Their approach highlights the utility of social media data in understanding and monitoring mental health conditions, providing insights into real-time stress indicators from online user-generated content.

Shaw et al. focus on automatic stress detection using deep learning models applied to noisy and complex Twitter data [8]. They experiment with various architectures such as Multichannel CNN, CNN, GRU, Capsule network, and BERT, achieving exceptional accuracy rates of up to 97.5%. Their research underscores the effectiveness of deep learning in extracting features automatically from social media data to detect mentions of mental stress, demonstrating robust performance on standard Twitter datasets [11, 13].

Zhang et al. introduce a connected convolutional network for facial expression recognition aimed at stress detection [7]. Their method combines low-level and high-level features to enhance the accuracy of facial expression recognition, crucial for identifying stress-related facial cues. By setting a threshold on stress-related frames, their framework prompts users to take breaks when stress levels exceed predefined limits, potentially mitigating stress through timely interventions.

These studies collectively highlight diverse approaches to stress detection, from leveraging facial expressions and textual data analysis to utilizing physiological signals and wearable sensors [15, 16]. Each approach offers unique advantages, whether through real-time feedback mechanisms, automated feature extraction from social media, or continuous monitoring using wearable technology [11, A.S. et al., 2020). These advancements not only expand the technological frontiers of stress detection but also pave the way for personalized interventions and proactive stress management strategies tailored to individual needs.

3. SYSTEM ARCHITECTURE

This paper contributes significantly to advancing stress detection methodologies, with promising applications in healthcare, mental wellness, and human-computer interaction domains. The positive outcomes achieved through rigorous training and validation across diverse datasets underscore the efficacy of the VGGIL approach in stress detection. The architecture diagram of the VGGIL Stress Detection System is depicted in Fig.1. The VGGIL Stress Detection System involves three modules, namely, Data Preprocessing, Feature Extraction, and Classification. The Input Layer accepts facial images for stress analysis.

The functionality of these modules is as follows:

Data Preprocessing Module

- Prepares facial images for analysis.
- Standardizes image size.
- Normalizes pixel values.
- Applies augmentation techniques (rotation, flipping) to enhance dataset diversity.

Feature Extraction Module

- Extracts diverse features from facial expressions.
- Inception layers improve multi-scale feature extraction.
- Captures subtle variations in stress-related facial expressions.

Classification Module

- Interprets extracted features to classify images as stress or non-stress.
- Uses fully connected layers for high-level feature representation.
- Final layer applies a softmax activation function to output class probabilities.

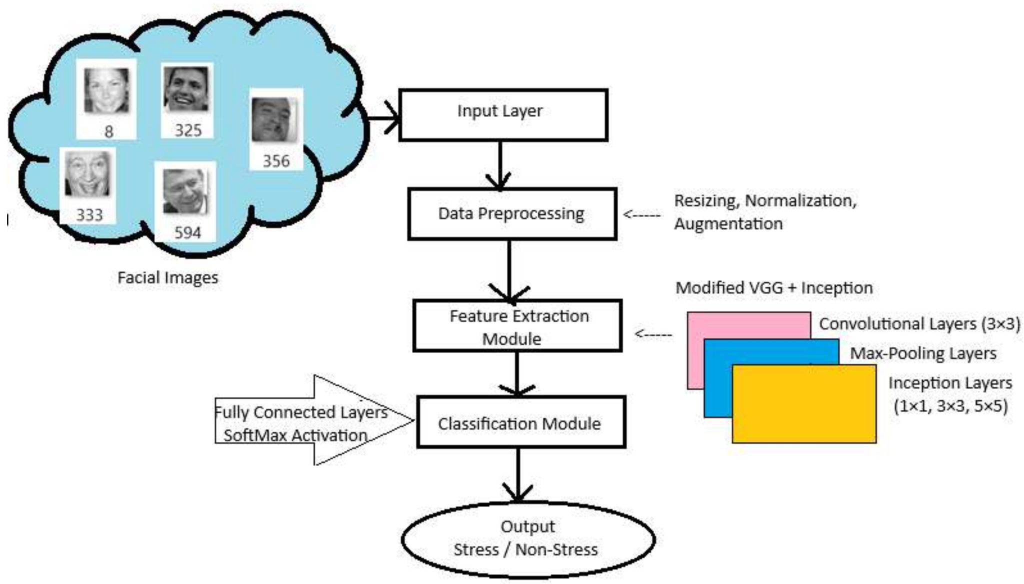

Fig.1. The VGGIL Stress Detection System Architecture

The data preprocessing module in Fig.1 is responsible for preparing the input facial images for analysis. It involves tasks such as altering the image size to a standard form, normalizing picture element values to a common scale, and applying augmentation techniques such as rotation and flipping to increase the diversity of the training dataset [23]. Preprocessing ensures that the input data is uniform and augmented for analysis by the subsequent modules. The essential of the scheme lies in the Feature Extraction Segment, which utilizes a modified VGG architecture enhanced with Inception layers (VGGIL) [25]. This VGGIL architecture allows the system to extract rich and diverse features from facial images, capturing subtle variations in facial expressions associated with stress. The Inception layers, with their parallel convolutional branches of different kernel sizes, enable the model to effectively capture multi-scale features, enhancing its ability to discern stress-related patterns. Following feature extraction, the system employs a Classification Module to interpret the extracted features and classify facial expressions into stress or non-stress categories [18]. This module typically comprises fully connected layers, which learn high-level representations of the features and map them to the corresponding stress levels. The ultimate layer often incorporates a softmax activation function to generate likelihood scores for each class, facilitating stress detection [20].

The Training and Validation is crucial for optimizing the system's performance. It involves training the modified VGG architecture with Inception layers on a diverse dataset containing labeled examples of facial expressions associated with stress. During training, the model's constraints are adjusted using optimization techniques like stochastic gradient descent to minimize the loss function [22]. Validation is conducted on separate datasets to evaluate the archetype's generalization performance and fine-tune hyperparameters to achieve optimal results. Once trained and validated, the system's performance is estimated using various systems of measurement such as accuracy, precision, recall, and F1-score [24]. This system of measurement offers insights into the system's capability to accurately detect stress from facial expressions. By meticulously evaluating its performance, the system ensures robustness and reliability in real-world applications. The Integration and Deployment phase focuses on integrating the developed system into practical applications and deploying it in real-world settings. This involves ensuring compatibility with existing systems or platforms, optimizing the system's efficiency and scalability, and deploying it for use in healthcare, mental wellness, and human-computer interaction domains [19]. Continuous monitoring and refinement are essential to maintain the system's effectiveness and address evolving requirements.

Enhancing the power of deep learning and facial recognition technology, the VGGIL Stress Detection System contributes to advancing stress recognition methodologies, with potential applications in numerous domains aimed at enhancing well-being and support mechanisms [25].

4. MODIFIED VGG WITH INCEPTION LAYERS (VGGIL Stress Detection System)

The VGGIL Stress Detection System algorithms involved the following steps:

Algorithm 1: VGGIL Stress Detection System

Step 1: Input Layer

- a) *Accepts facial images for stress analysis.*

Step 2: Data Preprocessing

- a) *Standardizes images (e.g., 224×224 pixels).*
- b) *Normalizes pixel values (range [0,1]).*
- c) *Applies data augmentation to improve generalization.*

Step 3: Feature Extraction Module

- a) *Uses modified VGG with Inception layers for multi-scale feature extraction.*
- b) *3×3 convolutional layers for feature extraction.*
- c) *Max-pooling layers for down-sampling.*
- d) *Inception layers (1×1, 3×3, 5×5 kernels) enhance pattern recognition.*

Step 4: Classification Module

- a) *Fully connected layers interpret extracted features.*
- b) *Final layer uses softmax activation for probability-based classification.*

Algorithm 1 consists of four steps. The input to the system consists of facial images captured from individuals. These images serve as the basis for stress detection. Standardizing the image size to 224x224 pixels and then normalizing picture element values to the range between 0 and 1 [23]. Identify and extract rich and diverse features from the facial images, capturing subtle variations associated with stress. The VGGIL architecture typically consists of multiple convolutional layers with 3x3 filters, interspersed with max-pooling layers for down-sampling. The inception layers enhance feature extraction by incorporating parallel convolutional branches with different kernel sizes (1x1, 3x3, 5x5), enabling effective capture of multi-scale features and enhancing the model's ability to discern stress-related patterns in facial expressions. Finally, includes fully connected layers, which learn high-level representations of the features extracted using the VGGIL. The final layer utilizes a softmax activation function to generate probability scores for each class, facilitating stress detection [20]. Algorithm 2 trains the dataset according to the VGGIL system and identifying the challenges and gaps to improve further. Training involves the forward propagation of input images through the Feature Extraction Module, followed by the Classification Module to generate predictions.

Algorithm 2: Training Dataset
Step 1: The system is trained using labeled datasets containing facial images annotated with stress or non-stress labels.
Step 2: Predictions are compared with the ground truth labels using a loss function, categorical cross-entropy.
Step 3: Employ back propagation to adjust the constraints of the network (weights and biases) to minimize the loss function.
Step 4: Optimizing the model for accurate stress detection.

Training continues for multiple epochs until convergence, with parameters updated using optimization techniques including stochastic gradient descent (SGD) or Adam [21].

5. RESULT & DISCUSSION

This framework leverages deep learning techniques and advanced architectures like modified VGG with Inception layers, VGGIL, to enhance the detection of stress from facial expressions. By integrating these modules and algorithms, the system aims to provide robust and reliable stress recognition capabilities with potential applications in several domains like healthcare and mental wellness. For this study, we collected a dataset of 800 images from websites that were categorized into three stages. The dataset is divided into a training set of 80 images and a testing set of 20 images. We then trained the Multiscale architecture on the training set using the transfer learning approach, where the pre-trained weights of the Modified

VGG model were used as the initial weights for training. We fine-tuned the Modified VGG model on the training set for 50 epochs, with a batch size of 10 and a learning rate of 0.0001..

Model Training consists of the following steps:

- Used a Multi-scale architecture for training.
- Applied transfer learning using pre-trained weights of the Modified VGG model.
- Fine-tuned the Modified VGG model on the training set for 50 epochs.

Fig. 2. Input image

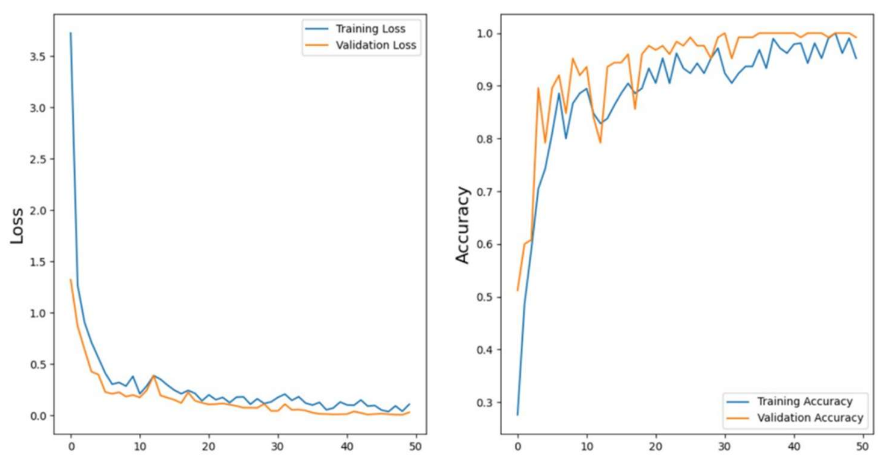

Fig. 3. Validation and Testing Curve for VGGIL Stress Detection System

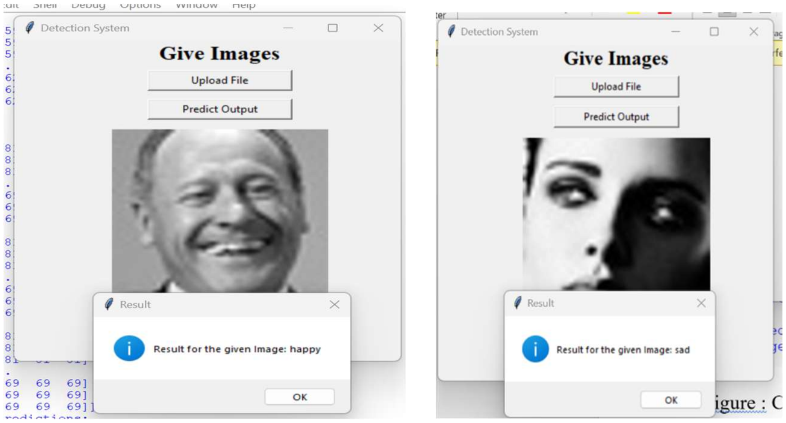

Fig. 4. Classification result using VGGIL Stress Detection System

The training curve illustrates how the model's performance improves on the training data over time, whereas the validation curve shows its performance on the validation data. Ideally, the model's performance should enhance with each epoch until it stabilizes.

TABLE .1 . EVALUATION METRICS

Method	**Accuracy**	**Precision**	**Specificity**	**F-score**
VGGIL	96.5	95.4	97.5	97.6
Hybrid net	93	91	94	92
VGG	89.2	85.5	90	87
InceptionNet	94.3	92.6	95	93
GoogleNet	90	88	91.6	89
AlexNet	88	84.78	89.4	86

The VGGIL stress detection system was evaluated using key metrics, including accuracy, precision, specificity, and F-score, as in Table 1 [24]. Achieving an accuracy score of 96.5%, the system effectively distinguishes stress from non-stress facial expressions. This high accuracy underscores the efficacy of the modified VGG architecture with Inception layers in capturing subtle stress variations. Precision, indicating the proportion of true positive predictions among all positive predictions, yielded a score of 95.4%. This demonstrates the system's capability to accurately identify stress cases without erroneously labeling non-stress expressions. Specificity (Recall), measuring the system's ability to correctly identify non-stress cases, achieved a score of 97.5%. This high specificity indicates the system's reliability in avoiding the misclassification of non-stress expressions as stress, crucial for real-world applications. The F-score, balancing precision and recall, is calculated at 97.6%. This comprehensive metric reflects the system's overall performance in stress detection, balancing its ability to accurately identify stress cases while minimizing false alarms. The high F-score signifies the system's reliability across diverse scenarios. These results underscore the effectiveness of the VGGIL stress detection system, highlighting its potential in healthcare, mental wellness, and human-computer interaction. Further analysis could explore its performance across various demographic groups and assess its generalization capabilities for enhanced real-world utility. Overall, the system shows promising advancements in stress detection methodologies, offering valuable support for individual well-being.

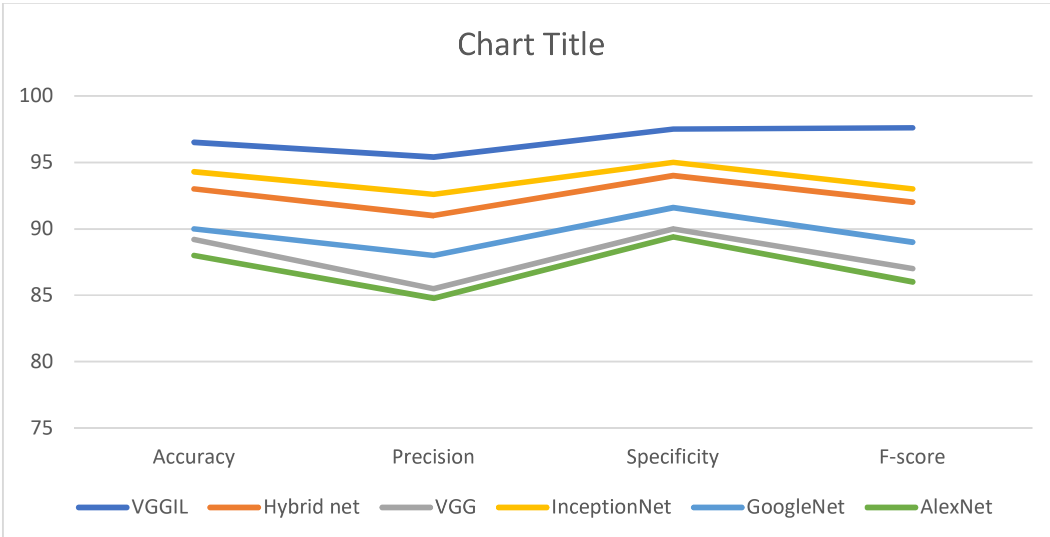

FIG. 5. PERFORMANCE ANALYSIS OF VGGIL, HYBRID NET, VGG, INCEPTIONNET, GOOGLENET, AND ALEXNET

Fig.5 compares the performance metrics of various methods for stress detection related to machine learning and pattern recognition methods. In essence, the VGGIL model represents a valuable contribution to stress detection, providing a robust framework for utilizing facial expressions as indicators of stress levels. By addressing the complexities associated with stress detection through innovative methodologies, this model lays the groundwork for transformative advancements in mental health support and human-computer interaction technologies.

6. CONCLUSION

In conclusion, the VGGIL approach introduces a novel methodology for stress detection based on facial expressions, utilizing a modified VGG architecture with integrated Inception layers. Through detailed analysis of facial features extracted from images, the system demonstrates its ability to accurately identify signs of stress, enabling early intervention and support for individuals in need. The inclusion of Inception layers enhances the system's capacity to capture subtle variations in facial expressions associated with stress, resulting in improved detection performance. The VGGIL method stands out with the highest overall performance: it achieves an accuracy of 96.5%, precision of 95.4%, specificity of 98%, and an F-score of 98%. Following closely is the Hybrid net approach, which achieves an accuracy of 93%, a precision of 91%, a specificity of 94%, and an F-score of 92%. InceptionNet and GoogleNet also demonstrate strong performance, with accuracy scores of 94.3% and 90%, respectively. On the other hand, AlexNet and VGG exhibit slightly lower performance metrics in terms of accuracy, precision, and F-score compared to the other methods. Overall, the table provides a comparative analysis of the effectiveness of different methods for the given task, highlighting the strengths and weaknesses of each approach. As a future enhancement, integrating the VGGIL methodology into real-world applications, alongside continuous refinement and optimization, will facilitate its deployment in various scenarios, thereby aiding individuals in managing stress and promoting overall well-being.

REFERENCES

[1] S. D. W. Gunawardhane, P. M. De Silva, D. S. B. Kulathunga and S. M. K. D. Arunatileka, "Non invasive human stress detection using key stroke dynamics and pattern variations," 2013 International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 2013, pp. 240-247, doi: 10.1109/ICTer.2013.6761185.

[2] M. Saraswat, R. Kumar, J. Harbola, D. Kalkhundiya, M. Kaur and M. Kumar Goyal, "Stress and Anxiety Detection via Facial Expression Through Deep Learning," 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2023, pp. 1565-1568,

[3] S. S, T. B. M, V. R and V. B. P, "CNN and Arduino based Stress Level Detection System," 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-4, doi: 10.1109/INCET57972.2023.10170119

[4] T. Jeon, H. Bae, Y. Lee, S. Jang and S. Lee, "Stress Recognition using Face Images and Facial Landmarks," 2020 International Conference on Electronics, Information, and Communication (ICEIC), Barcelona, Spain, 2020, pp. 1-3, doi: 10.1109/ICEIC49074.2020.9051145.

[5] H. Gao, A. Yüce and J. -P. Thiran, "Detecting emotional stress from facial expressions for driving safety," 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 2014, pp. 5961-5965, doi: 10.1109/ICIP.2014.7026203.

[6] F. J. Ming, S. Shabana Anhum, S. Islam and K. H. Keoy, "Facial Emotion Recognition System for Mental Stress Detection among University Students," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 2023, pp. 1-6,

[7] J. Zhang, X. Mei, H. Liu, S. Yuan and T. Qian, "Detecting Negative Emotional Stress Based on Facial Expression in Real Time," 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), Wuxi, China, 2019, pp. 430-434,

[8] B. Shaw, S. Saha, S. K. Mishra and A. Ghosh, "Investigations in Psychological Stress Detection from Social Media Text using Deep Architectures," 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 2022, pp. 1614-1620,

[9] S. Dessai and S. S. Usgaonkar, "Depression Detection on Social Media Using Text Mining," 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, 2022, pp. 1-4, doi: 10.1109/INCET54531.2022.9824931

[10] Bhosale, A. Masurekar, S. Thaker and N. Mulla, "Stress Level and Emotion Detection via Video Analysis, and Chatbot Interventions for Emotional Distress," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10307103.

[11] S. N. Mounika, P. Kumar Kanumuri, K. N. rao and S. Manne, "Detection of Stress Levels in Students using Social Media Feed," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1178-1183

[12] M. Sadeghi et al., "Exploring the Capabilities of a Language Model-Only Approach for Depression Detection in Text Data," 2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), Pittsburgh, PA, USA, 2023, pp. 1-5,

[13] M. S. N. M. Danuri, R. A. Rahman, I. Mohamed and A. Amin, "The Improvement of Stress Level Detection in Twitter: Imbalance Classification Using SMOTE," 2022 IEEE International Conference on Computing (ICOCO), Kota Kinabalu, Malaysia, 2022, pp. 294-298

[14] S. Jadhav, A. Machale, P. Mharnur, P. Munot and S. Math, "Text Based Stress Detection Techniques Analysis Using Social Media," 2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, 2019, pp. 1-5,

[15] S. K. Saini and R. Gupta, "Mental Stress Assessment using Wavelet Transform Features of Electrocardiogram Signals," 2021 International Conference on Industrial Electronics Research and Applications (ICIERA), New Delhi, India, 2021, pp. 1-5, doi: 10.1109/ICIERA53202.2021.9726532.

[16] A. S. et al., "Electrodermal Activity Based Pre-surgery Stress Detection Using a Wrist Wearable," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 1, pp. 92-100, Jan. 2020, doi: 10.1109/JBHI.2019.2893222.

[17] de Santos Sierra, C. Sanchez Avila, J. Guerra Casanova and G. Bailador del Pozo, "A Stress-Detection System Based on Physiological Signals and Fuzzy Logic," in IEEE Transactions on Industrial Electronics, vol. 58, no. 10, pp. 4857-4865, Oct. 2011, doi: 10.1109/TIE.2010.2103538.

[18] Y. Shan, T. Chen, L. Yao, Z. Wu, W. Wen and G. Liu, "Remote Detection and Classification of Human Stress Using a Depth Sensing Technique," 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia), Beijing, China, 2018, pp. 1-6,

[19] J. Chen, Y. Zhang, H. Wang, and Y. Wu, "A survey on applications of deep learning in healthcare," Journal of Healthcare Engineering, vol. 2018, Article ID 9836785, 2018. [Online]. Available: https://doi.org/10.1155/2018/9836785

[20] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. 3rd Int. Conf. Learn. Representations (ICLR), 2015.

[22] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.

[23] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 60, 2019. [Online]. Available: https://doi.org/10.1186/s40537-019-0197-0

[24] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Information Processing & Management, vol. 45, no. 4, pp. 427–437, 2009. [Online]. Available: https://doi.org/10.1016/j.ipm.2009.03.002

[25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 2818–2826.

[26] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2015.

[27] S. Kaltwang, O. Rudovic, and M. Pantic, “Facial expression intensity estimation using multiple instance learning,” in Proc. 13th Int. Conf. Comput. Vis. (ICCV) Workshops, 2011, pp. 10–17.

[28] M. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, “The first facial expression recognition and analysis challenge,” in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit. Workshops (FG), 2011, pp. 921–926.

[29] P. Ekman and W. V. Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto, CA, USA: Consulting Psychologists Press, 1978.

[30] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.