Unveiling Stress through Facial Expressions: A Literature Review on Detection Methods

Dr.Sinciya P.O
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
posinciya@amaljyothi.ac.in

Dona Maria Joseph
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
donamariajoseph2024@cs.ajce.in

Aaron Varughese Bino
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
aaronvarughese2024@cs.ajce.in

Aathira Krishna
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
aathirakrishna2024@cs.ajce.in

Anamin Fathima Anish
Dept. of Computer Science
Amal Jyothi College of Engineering
Kottayam, India
anaminfathimaanish2024@cs.ajce.in

***Abstract*—Stress detection is crucial in various fields, including healthcare, human-computer interaction, and automotive safety. This paper presents a comprehensive comparison study of three emotion detection modules: facial expression analysis, eyeblink count, and eyebrow movements. The aim is to assess their effectiveness in detecting stress accurately. Each models is evaluated based on its ability to discern stress levels in real-time scenarios. By analyzing the data collected from different research papers related to stress-inducing stimuli, we provide insights into the strengths and limitations of each model. Additionally, we propose a novel framework that integrates these modules to enhance stress detection accuracy. The results indicate promising performance, with the integrated framework demonstrating superior stress detection capabilities compared to individual modules. This research contributes to advancing stress detection methodologies, paving the way for more reliable and efficient stress management systems.**

***Index Terms*—Convolutional Neural Network, Eye aspect ratio**

I. Introduction

Stress is a common health concern that has been seen to be a contributing factor to many other health complications. With stress being such a common problem in modern society, it is necessary to monitor and determine an individual's stress level to understand what activities and environments lead to stress and how to avoid such conditions. As technologyis advancing and the use of automation is becoming more frequent in everyday human life, it would be progressive to develop techniques to measure or monitor stress effectively and efficiently. With stress affecting an individual's thought process, decision making, and reaction time, it is a factor that can affect safety and effectiveness in many everyday tasks for a wide range of professionals. An automated method of measuring stress opens up many possibilities for detecting what environments and conditions cause stress and how to avoid such stressful situations. Stress detection can be applied to monitoring the stress level of a driver, student, teacher, oran IT professional. Stress detection can help individuals to take care of their health, thereby contributing more efficientlyto their work. It is also used in many military applications todetermine the stress level of soldiers in training exercises or inthe field to determine high-stress conditions and environments. One of the best ways to detect stress is stress detection based on facial features. It is a remote method that detectsstress without attaching anything to the subject. However,the detection process must be in real-time. This methodis derived from humans who always show their emotions through their faces, sometimes consciously or unconsciously. Emotion is an expression of a person's psychological state and quality of conscious experience. Emotions always happen to aperson when certain events trigger them. This method is morerelated to psychology, where there have

DOI: 10.5281/zenodo.12549071

been many researchstudies that relate emotion to stress. Usually, emotions canbe measured using subjective methods based on experience and objective methods using physiological signals. Althoughemotions may not be identical to stress, they can be a sign of what happens to a person who feels stressed.

II. PROBLEM STATEMENT

Human stress is defined as an expression of an adaptive reaction to a stimulus that can place physical, physiological, and psychological strain on a person. If these requirements outweigh the resources a person has, this can result in a psychological or physiological dysfunction such as blurred vision or an increase in heart rate. Consistent exposure or high levels of stress can be harmful. The harmful effects of long-term human stress have been identified as a significant health problem by the medical community. It has been linked as a contributing factor to a range of diseases and disorders, including hypertension, cardiovascular disease, obesity, and psychological disorders.

Furthermore, prioritizing the identification and response to short-term stressors, which induce immediate physiological and psychological reactions, becomes imperative. Establishing dependable techniques for real-time detection of acute stress empowers individuals to adopt proactive measures in managing stress levels and mitigating its adverse effects on their well-being. Early recognition of stress facilitates timely interventions, including relaxation techniques or stress management strategies, thereby averting the aggravation of stress-related symptoms and potential health repercussions.

III. LITERATURE SURVEY

In [1] stress levels are calculated based on the eye blink rate of the user. The system monitors the frequency of eye blinks to assess the user's stress levels during digital device usage. The system establishes a baseline blink rate for each user and monitors for deviations. A decrease in blinks suggests higher stress. This real-time feedback allows users to adjust screen time and take breaks, potentially reducing eye strain and stress. The methodology involved collecting 60-second video sequences via webcam to capture facial expressions and eye behavior. A recurrent and convolutional neural network (RCNN)then analyzed the images, extracting features linked to eyefatigue and stress. These features trained a classification modelwith an accuracy of 85% on training data and 80% on a separate testing set. This demonstrates the model's effective-ness in detecting digital eye strain in real-time. Overall, the study contributes to the development of advanced eye strain detection systems using deep learning to monitor user behavior and provide timely alerts to prevent eye strain and stress

caused by prolonged digital device use. The integration of advanced computer vision and machine learning techniques is an efficient method to address the challenges associated with eye-blink detection and eye-state classification, particularly in the context of aiding communication for individuals with severe disabilities [2]. The system's methodology involves a well-modeled pipeline based onhigh-performance models, auxiliary models, and recent Deep Learning (DL) methods, allowing for real-time performance using generic hardware in both the acquisition and processing phases. The system is designed to perform both eye-state classification and eye-blink detection tasks. First, the system detects the user's face and then identifies landmark points on the face. These points are used for face alignment and for extracting coordinates of both the eye patch and the eye. A CNN model is then used to classify these extractions into eye images or non-eye images. The final step involves eye-state classification, for which two models were developed. The first model, based on a CNN, classifies the correct eye patches intoopen or closed eyes, while the second model uses an SVMto sort the extracted eye coordinates. The proposed system also includes auxiliary models such as a rotation compensator, a quality evaluator of the extracted regions of interest (ROIs), and a moving average filter to enhance the system's performance in real-time conditions. The Youtube Eye-state Classification (YEC) and the Autonomus Blink Detection (ABD) datasets were used for training and testing the eye- state classification and eye-blink detection tasks, respectively. Additionally, the system addresses the problem of missing the detection of rotated faces by incorporating a rotation compensation step.

The Eye Aspect Ratio Mapping (EARM) method aids in identifying eye fatigue by accurately detecting spontaneous blinks [3]. This method uses face images to classify blinks with high accuracy at a low cost. It also establishes a strong correlation between the median Spontaneous Blink Rate (SBR) and the time it takes for individuals to become aware of their eye fatigue. The EARM method provides a means to objec-tively estimate eye fatigue sensitivity, potentially allowing for the measurement of visual fatigue sensitivity and the onset of ocular fatigue by tracking spontaneous blinks. This approach could assist in maintaining eye health by providing users with appropriate breaks and awareness of eye fatigue during Visual Display Terminal (VDT) work.

The methodology involves conducting an experiment to es-timate eye fatigue sensitivity by detecting spontaneous blinks with high accuracy. The study used a new dataset and proposed the Eye Aspect Ratio Mapping (EARM) method for blink detection. The experiment involved four subjects in their 20s performing a Visual Display Terminal (VDT) task, specifically a Sudoku puzzle, under controlled indoor conditions. The study also introduced the concept of Eye Fatigue Level (EFL) as an objective quantification index of ocular fatigue. The proposed blink detection method demonstrated high accuracy and a strong correlation between the median Spontaneous Blink Rate (SBR) and the time between objective estimationof eye fatigue and the subject's awareness of eye fatigue.

Electroencephalogram (EEG) signals, speech signals, and audio-visual data can be used for early detection of human mental stress [4]. The system involves the use of various datasets, such as the Database for Emotion Analysis using Physiological Signals (DEAP) for stress detection using EEG, and the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) for recognition of stress using speech signal and audio-visual data. For stress detection using EEG signals, the system utilizes the EEG power spectrum. The system also involves the use of machine learning frameworks for EEG signal analysis of stressed participants. Additionally, the system aims to improve the accuracy of stress recognition from speech signals by implementing deep learning using Long Short-Term Memory (LSTM) structures.

In the case of stress detection using audio-visual data, the system focuses on accurately recognizing the emotional state of individuals. Emotion recognition is performed to determine the stress level, and the system utilizes features such as pitch, energy, and Mel-Frequency Cepstral Coefficients (MFCC) for audio data, and facial expressions and muscle vectors for visual data. The proposed system also aims to use deep learning techniques to improve the results of emotion state recognition from audio-visual data, which will be further used to detect stress.

In [5], system utilizes various parameters like displacement of the eyebrow from its mean position, variations in the position of the eyebrow, and the analysis of facial features. Thesystem operates on real-time, non-intrusive videos that are divided into equal-length sections, from which image frames areextracted and analyzed. By calculating the displacement of theeyebrow coordinates within each section, stress detection decisions are made based on consecutive intervals indicating stress. Additionally, the system employs deep learning algorithms, specifically leveraging the Python framework called Theano, to train models and analyze predictive features. Experimental results demonstrate the system's efficacy in detecting stress levels through the analysis of individual facial expressions.

The architecture of the stress detection system involvesthe use of a camera to capture near-frontal views of individuals, typically working in front of computers. Captured videos undergo segmentation into equal-length sections, with subsequent extraction and analysis of image frames. Image processing techniques are then applied to determine the dis- placement of the eyebrow from its mean position, serving as a key parameter for stress detection based on facial expressions. Moreover, the system integrates modules for image preprocessing, stress detection, and deep learning, where the latteris utilized to train models and predict stress levels based on analyzed facial expressions.

The stress recognition model employs a deep neural network that utilizes facial landmarks for more effective stress detection from face images obtained with a general camera. Haar cascade algorithm is applied for face detection, encompassing steps like converting the input image to grayscale, resizing, and capturing multiple training photos. The system undergoes supervised learning during training and testing phases, covering various stages including image acquisition, face detection, image preprocessing, feature extraction, and classification. Capable of discerning six basic expressions—angry, disgust, normal, sad, surprised, and happy—the facial expression recognition system is equipped to handle diverse emotional cues.

The study involves the use of Convolutional Neural Net- work (CNN) for classification and training, Haar cascade algorithm for face detection, and Local Binary Pattern His- togram (LBPH) for face detection and feature extraction. Additionally, the system uses facial landmarks to detect stress more effectively, leveraging the changes in eye, mouth, and head movements when a person is stressed. The documentalso discusses the use of algorithms such as Support Vector Machines and Local Fisher Discriminant Analysis for facial expression recognition.

In [7] the paper presents a methodology that involves the analysis of human parameters and the application of machine learning algorithms to detect stress in individuals. The method- ology includes the use of various physiological parameters such as age, sex, chest pain, blood pressure, and others, as wellas machine learning models such as Random Forest Classifier, Decision Tree, Naive Bayes, Support Vector Machine, and K-Nearest Neighbor for stress detection.

Various techniques for stress detection are explored, en- compassing the utilization of machine learning algorithms such as data preprocessing, Naive Bayes classification, and Support Vector Machine (SVM) implementation via WEKA The methodology entails statistical scrutiny of the dataset, simulation outcomes, and a comparative analysis of established techniques vis-à-vis the proposed approach. There's a particular emphasis on machine learning algorithms for emotion and stress detection, along with a discussion on deep learning methodologies and real-time EEG signals for prospective research avenues. The stress detection model attains an accuracy rate of 93.2% through the analysis of human parameters like age, sex, chest pain, and blood pressure. Machine learning models such as Random Forest Classifier, Decision Tree, Naive Bayes, SVM, and K-Nearest Neighbor contribute to this achievement. Evaluation of the model's efficacy encompasses comparison between actual and predicted results, complemented by a confusion matrix, offering a comprehensive assessment of its performance in stress detection relative to established methodologies.

IV. METHODOLOGY

A. FACIAL FEATURE EXTRACTION TECHNIQUES

There are several models from facial feature extraction:

1) **Deep Learning-based Feature Extraction (e.g., CNNs):** Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable performance in facial feature extraction tasks. CNNs can automatically learn hierarchical representations of facial features directly from raw pixel data, eliminating the need for handcrafted feature descriptors. They excel in capturing intricate patterns and variations in facial structures and textures,

making them suitable for a wide range of tasks such as face recognition, emotion detection, and facial expression analysis. CNNs can adapt to diverse datasets and handle complex variations in lighting, pose, and facial expressions. However, deep learning approaches typically require large amounts of labeled data for training and considerable computational resources.

2) **Facial Landmark Detection:** Facial landmark detection techniques aim to localize key points on the face, such as the corners of the eyes, nose, and mouth. Methods like Active Appearance Models (AAM), Constrained Local Models (CLM), and deep learning-based approaches can accurately identify facial landmarks, providing valuable information for subsequent feature extraction tasks. Landmark-based techniques are robust to variations in facial expression, pose, and lighting conditions, making them suitable for tasks like face alignment, facial motion tracking, and facial expression analysis.
3) **Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG):** Texture descriptors such as LBP and HOG are effective for capturing texture information in facial images. LBP encodes local texture patterns based on pixel intensities, while HOG computes gradients and orientation histograms to represent textural features. These techniques are useful for tasks like facial expression analysis, as they can capture subtle variations in facial texture and appearance. However, they may not capture high-level semantic features as effectively as deep learning-based approaches.
4) **Eigenfaces and Principal Component Analysis (PCA):** Eigenfaces is a classic technique for facial feature extraction based on PCA. It involves projecting face images onto a lower-dimensional subspace spanned by principal components, allowing for dimensionality reduction and feature extraction. While PCA-based techniques are computationally efficient and straightforward to implement, they may struggle with complex variationsin facial appearance and expression compared to more advanced deep learning methods.

B. MACHINE LEARNING ALGORITHMS FOR STRESS DETECTION

Different machine learning algorithms that are used for stress detection using facial feature extraction include:

1) **Convolutional Neural Networks (CNNs):** CNNs have gained popularity for facial expression analysisdue to their ability to automatically learn hierarchical representations of visual data. They excel at capturing spatial dependencies in images, making them well- suited for tasks like facial feature extraction and emotion recognition. CNN-based approaches often involve training a neural network on labeled facial expression datasets to classify facial expressions into different emotional states, including stress. CNNs have shown promising results in stress detection tasks by analyzing facial microexpressions, subtle changes in facial expressions that occur rapidly and involuntarily, which can indicate underlying stress.
2) **Long Short-Term Memory Networks (LSTMs):** LSTMs are a type of recurrent neural network (RNN) designed to model sequential data. In the context ofstress detection from facial expressions, LSTMs can be used to analyze temporal dependencies in facial expression sequences over time. By capturing the temporal dynamics of facial expressions, LSTMs can effectively identify patterns associated with stress onset and progression. This approach is particularly useful when analyzing video sequences of facial expressions captured in real-time.
3) **Support Vector Machines (SVMs):** SVMs are a classic machine learning algorithm that can be used for binary classification tasks, making them suitable for stress detection where the goal is to classify individuals as stressed or not stressed based on facial expressions. SVMs work well with high-dimensional data and can handle non-linear decision boundaries effectively. While they may not capture temporal dependencies as effectively as RNN-based approaches, SVMs are known for their robust performance and generalization capabilities.
4) **Random Forest:** Random Forest is an ensemblelearning method that combines multiple decision trees to make predictions. It can be used for both classification and regression tasks, making it suitablefor stress detection tasks where the goal is to predict stress levels based on facial expressions. Random Forest is robust to overfitting and works well with high-dimensional data, making it a versatile choice for stress detection applications.
5) **Deep Belief Networks (DBNs):** DBNs are a type of deep learning architecture composed of multiple layers of stochastic, latent variables. They have shown effectiveness in modeling complex, high-dimensional datalike facial expressions. DBNs can learn hierarchical representations of facial features, capturing both low-level and high-level features relevant to stress detection. While less commonly used compared to CNNs and LSTMs, DBNs have shown promise in facial expression analysis tasks.

V. DISCUSSION

In this comparative study of machine learning models for stress detection using facial emotion analysis, Convolutional Neural Networks (CNNs) emerge as a standout choice. The effectiveness of CNNs lies in their ability to automatically learn hierarchical representations of visual data, making them adept at capturing spatial dependencies in images, including facial expressions. Here, we delve into why CNNs outshine other mod~~els in stress detection tasks, highlighting the limitati~~ons of alternative approaches.

CNNs offer unparalleled capabilities in discerning intricate facial expressions, particularly microexpressions, which are rapid and involuntary manifestations of underlying stress. Their hierarchical feature extraction mechanism allows them to capture both subtle and overt cues indicative of stress, providing a comprehensive understanding of emotional states. Unlike traditional machine learning algorithms, CNNs inherently adapt to diverse facial characteristics and expressions, enhancing their robustness across varied datasets and individuals.Contrastingly, models like Long Short-Term Memory Net- works (LSTMs) primarily focus on temporal dependencies within facial expression sequences. While LSTMs excel in analyzing the dynamic evolution of emotions over time, they may overlook crucial spatial cues embedded in facial features. This limitation becomes apparent when dealing with static images or scenarios where immediate stress indicators arepivotal for timely intervention.

Support Vector Machines (SVMs) and Random Forest, while capable of handling high-dimensional data, lack the nuanced feature extraction prowess inherent in CNNs. SVMs rely on defining optimal decision boundaries, potentially over-looking intricate facial patterns crucial for stress detection. Similarly, while Random Forest excels in ensemble learning and mitigating overfitting, its reliance on decision trees may not fully capture the complex relationships between facialfeatures and stress states.

Deep Belief Networks (DBNs), albeit promising in modeling high-dimensional data, are less prevalent in facial emotion analysis compared to CNNs. Their hierarchical architecture enables them to learn intricate representations of facial features, akin to CNNs. However, the relatively limited adoption of DBNs in stress detection tasks may stem from their computational complexity and the robust performance already demonstrated by CNNs. In comparative studies, CNNs consistently demonstrate superior accuracy rates, ranging from 70% to 90%, underscoringtheir efficacy in stress detection from facial expressions. This superiority can be attributed to their inherent ability to learn hierarchical representations of visual data, capturing both spatial and temporal dependencies crucial for discerning stressindicators. As such, CNNs emerge as the optimal choice for stress detection tasks, offering unmatched performance and adaptability across diverse datasets and real-world scenarios.

TABLE I
ACCURACY OF DIFFERENT MACHINE LEARNING MODELS ON FACIAL FEATURES

Machine Learning Model	**Accuracy (%)**
Convolutional Neural Networks (CNNs)	70-90
Long Short-Term Memory Networks (LSTMs)	60-80
Support Vector Machines (SVMs)	60-80
Random Forest	60-80
Deep Belief Networks (DBNs)	60-80

VI. CONCLUSION

In conclusion, Convolutional Neural Networks (CNNs) stand out for stress detection from facial expressions. They excel in learning hierarchical representations of visual data, capturing spatial dependencies effectively. While models like LSTMs, SVMs, Random Forests, and DBNs have merits, theyoften lack nuanced feature extraction compared to CNNs. CNNs consistently demonstrate superior accuracy rates, rang- ing from 70% to 90%, due to their ability to discern intricate facial expressions and capture both spatial and temporal dependencies crucial for stress detection. Thus, CNNs emergeas the optimal choice for stress detection, offering unmatched performance across diverse datasets and scenarios.

REFERENCES

[1] Sharma, Deepti, Archana B. Saxena, and Deepshikha Aggarwal. ”Dig-ital Eye Strain Detection By Applying Deep Learning Technique To Categories Of Images.” Journal of Survey in Fisheries Sciences 10.4S (2023): 2086-2090.

[2] Paulo Augusto de Lima Medeiros, Gabriel Vin´ıcius Souza da Silva, Felipe Ricardo dos Santos Fernandes, Ignacio Sánchez-Gendriz, Hertz Wilton Castro Lins, Daniele Montenegro da Silva Barros, Danilo Alves Pinto Nagem, Ricardo Alexsandro de Medeiros Valentim,”Efficient machine learning approach for volunteer eye-blink detection in real-timeusing webcam”,Expert Systems with Applications,Volume 188,2022,116073,ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2021.116073.

[3] Akihiro Kuwahara, Kazu Nishikawa, Rin Hirakawa, Hideaki Kawano, Yoshihisa Nakatoh,”Eye fatigue estimation using blink detection based on Eye Aspect Ratio Map-ping(EARM)”,Cognitive Robotics,Volume 2,2022,Pages 50-59,ISSN 2667-2413,https://doi.org/10.1016/j.cogr.2022.01.003.

[4] Gupta, Megha and Vaikole, Shubhangi, Recognition of Human Mental Stress Using Machine Learning Paradigms (April8, 2020). Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST) 2020, Available at SSRN: https://ssrn.com/abstract=3571754 or http://dx.doi.org/10.2139/ssrn.3571754

[5] Nisha Raichur , Nidhi Lonakadi, Priyanka Mural.”Detection of Stress Using Image Processing and Machine Learning Techniques”

[6] Mohammad Failzal , Nikhil V , Nikhil V , Prajwal C , Aruna Rao B P”Recognition of Stress using Face Image and Facial Landmarks”

[7] Sudha, V., et al. ”Stress detection based on human parameters using machine learning algorithms.” (2022)