

STRESS DETECTOR: A DEEP LEARNING BASED EMOTION RECOGNITION AND LANDMARK DETECTION SYSTEM

ZAEEM NAZIR

Department of CS & IT, University of Narowal, Pakistan. Email: zaeem.nazir@uon.edu.pk

SADIA TARIQ *

Department of Computer Science and IT, The University of Lahore, Pakistan.

*Corresponding Author Email: sadia.tariq.shah@gmail.com

TALHA MAJEED

Department of Computer Science, University of Engineering & Technology Lahore, Pakistan.

Email: icemughal1245@gmail.com

SAIRA TARIQ

Department of CS & IT, University of Narowal, Pakistan. Email: sairatariq892@gmail.com

FARHANA TABASSUM

Department of CS & IT, University of Narowal, Pakistan. Email: thefarhanatabbassum@gmail.com,

USAMA NAWAZ

Department of CS & IT, University of Narowal, Pakistan. Email: usma6126@gmail.com

Abstract

Stress, a prevalent issue in today's fast-paced world, has significant implications on physical and mental health. Recognizing the global impact of stress-related diseases, we embarked on a journey to develop an intelligent stress detection system. Our approach integrates facial expression analysis, landmark detection, and deep learning to provide real-time assessment of stress levels. Leveraging a dataset of facial images and deep learning models, we customized a Convolutional Neural Network (CNN) as an optimal choice for emotion recognition, achieving remarkable results. Incorporating physiological indicators, we devised a stress calculation formula that categorizes stress into four states. Our user-friendly web interface empowers users to monitor their stress trends overtime. This research showcases the effectiveness of learning based assistive technologies for stress management and significant improvement in accuracy of classifier models.

Index Terms: Emotion Recognition, Lips and Eyebrow Landmarks, Stress Calculation, Stress States.

1. INTRODUCTION

In our fast-paced modern society, stress has become a pervasive issue, affecting an ever-growing number of individuals. Stress can cause a person to experience extreme discomfort and distress on a psycho-physiological level, It may result in serious mental health problems including panic episodes or depression. Depending on how stressful events are managed, stress can have both positive and negative effects on a person. Many factors contribute to stress such as biological and genetic history, external environment and personal experiences etc. It is important to detect stress especially negative stress as early as soon as possible because stress can seriously affect people's lives (Gedam & Paul, 2021).

It may be harmful to one's health, As the World Health Organization (WHO) has underlined the worldwide relevance of stress-related disorders, with over 450 million people worldwide grappling with these conditions, ranking mental disorders among the most common causes of illness and disability (World Health Organization, 2001).

Stress can be triggered by a multitude of factors, including workplace stress, financial pressures, interpersonal conflicts, and the loss of loved ones. According to WHO data, in 2019, roughly 15% of working-age persons were diagnosed with a mental condition (World Health Organization, 2001). Uncontrolled stress can be harmful to one's physical health and emotional well-being, resulting in conditions like depressive disorders, anxiety, insomnia, heart disease, and hypertension (World Health Organization, 2001). Given the profound societal repercussions of stress, Effective stress management and detection systems are in high demand.

Recognizing that the signs and causes of stress are unique to each individual, it is imperative to employ diverse approaches for identification and mitigation. Recent technological and artificial intelligence (AI) advancements have opened new doors for innovative stress detection systems that have the potential to significantly enhance our ability to identify and manage stress in real-time.

Our overarching objective is to harness the power of machine learning and AI to develop an intelligent stress detection system capable of reliably assessing stress levels. Our approach integrates physiological cues, landmark detection, and facial expression analysis, drawing inspiration from the work of (Brugman, 2022). By providing individuals under stress with a technical tool to discern their emotional state promptly, we aim to empower them to take proactive steps in stress management.

Our approach is based on the rich fields of facial expression analysis and facial landmark recognition. Human communication relies heavily on facial expressions, and a deeper understanding of these expressions can offer valuable insight into an individual's emotional state (Giannakakis et al., 2017).

In this study, Deep learning models, specifically CNNs, are used, to construct a framework capable of real-time detection of stress-related facial expressions, building on the work of Zenonos et al. (Zenonos et al., 2016). We also include physiological markers, like facial landmarks, to give a thorough evaluation of stress, in line with the conclusions of Ghaderi et al. (Ghaderi et al., 2015).

Our goal with stress detection technology is to help people take control of their emotional health. This technology facilitates proactive stress reduction measures by providing real-time feedback on stress levels, ultimately leading to improved mental and physical health (Nakashima et al., 2016).

To advance beyond previous models and research and create a more precise and reliable stress detection web-app, our aim is to collaboratively use physiological indicators, facial expression analysis and machine learning methods (Liu & Ulrich, 2014). This is especially important in circumstances where excessive stress can negatively impact people's health and wellbeing (Bhattacharyya & Basu, 2018).

2. LITERATURE REVIEW

We seek to examine recent advancements in stress recognition techniques, with a focus on their applicability and accuracy in real-world scenarios.

Su et al. (Su et al., 2020) performed a thorough analysis of deep learning's application and AI to mental health outcome studies. They sub-divided the available studies into four major categories: (1) diagnosis and prognosis, (2) genetic analysis, (3) vocal and visual data based illness detection and (4) estimation of risk of being ill using social media data. The work presented by Gedam et al. (Gedam & Paul, 2021) thoroughly summarizes various stress detection methods/models, surveys and reviews, commercial devices and apps, smart wearable devices and sensors arena for the said task in a convenient manner. These tools are simple to use and generate little error and noise, allowing them to monitor stress levels without getting in the way of the user's daily tasks.

In terms of algorithms, several machine learning techniques were applied to produce categorization models. RF, SVM, KNN, and logistic regression models were the most often utilized classifiers. The most commonly used methods of verifying models used for classification are k-fold cross-validation and leave-one-subject-out. Biological signals that can be reliably assessed as stress generating agents include Physical (respiratory rate, Speech Analysis, skin temperature, pupil size, ocular activity) and physiological (Electroencephalogram, Electrocardiogram, Electro dermal Activity, Electromyogram) testing (Giannakakis et al., 2019). They are mostly sensor based and focus on efficient, robust and consistent stress detection. The following section discusses completely contact-free approaches for analyzing face features using a video camera given by Lim et al. (Lim et al., 2020). When smart wearables come into touch with the skin, physiological data including electro dermal activity and heart-related signals may be discreetly recorded.

Similarly, multimodal affective computing systems promise higher classification accuracy. Next, we analyze five more articles that present innovative approaches similar to ours. Naidu et al. (Naidu et al., 2021) introduced a non-invasive stress detection system that leverages facial landmarks and the AlexNet architecture. Their study underscores the importance of unobtrusive stress detection, addressing limitations of invasive bio-signal-based methods. By applying deep learning techniques to standard camera images, they achieve precise stress level identification, consistent with established psychological theories. Sabry et al. (Sabry et al., 2022) proposed a deep learning-based method for recognizing stress, emphasizing the significance of physiological signals and facial expressions. Their study integrates physiological data with facial features, resulting in high accuracy in stress classification. This work highlights the potential of combining multimodal data sources for enhanced stress detection, aligning with the non-invasive theme presented in (Naidu et al., 2021).

Mayya et al. (Mayya et al., 2015) explored the use of smartphone data, including GPS, call logs, and app usage, to predict stress levels. Their study utilizes machine learning algorithms to analyze behavioral patterns. The findings demonstrate that smartphone data can serve as valuable indicators of stress, offering a non-intrusive and continuous

monitoring solution. This approach complements the themes of non-invasiveness and continuous monitoring presented in (Naidu et al., 2021) and (Rony et al., 2021). Rony et al. (Rony et al., 2021) analyzed real-time stress detection using electro-dermal activity (EDA) and heart rate variability (HRV) using wearable sensors, in particular. Their research emphasizes the potential of wearable devices in monitoring stress levels during daily activities. Integration of physiological data with Machine learning algorithms allow for precise results and real-time stress recognition, echoing the goal of non-invasive and continuous monitoring shared across (Naidu et al., 2021), (Rony et al., 2021), and (Mayya et al., 2015). Atal et al. (Atal & Singh, 2020) studied stress detection methods such as thermal imaging and machine learning. Their study addresses the limitations of conventional thermal imaging methods through a novel algorithm. By analysing thermal images, the authors achieve effective stress inference, presenting an alternative non-invasive approach. This aligns with the theme of non-invasive stress detection found in all the previously mentioned articles. The need of an entirely software system with preferably a web app interface for immediate detection of stress with minimal effort at user end is obvious. Our primary contribution lies in this gap. The remaining section of this article is structured in the following order: Section 3 discusses the proposed model and dataset, whereas Section 4 goes into the experimental setup and evaluation. Section 5 contains the findings, their analysis, and related commentary. Section 6 contains the concluding thoughts and recommendations for future research.

3. PROPOSED METHODOLOGY

The proposed methodology employs a publically available dataset named Fer2013. After rescaling and pre-processing the dataset, the CNN is applied for emotion recognition and landmark detection. Next, landmarks are applied and distance is computed. Finally, stress levels and stated were calculated, displayed on user interface and stored in history. A pictorial summary of the procedure has been shown in figure 1.

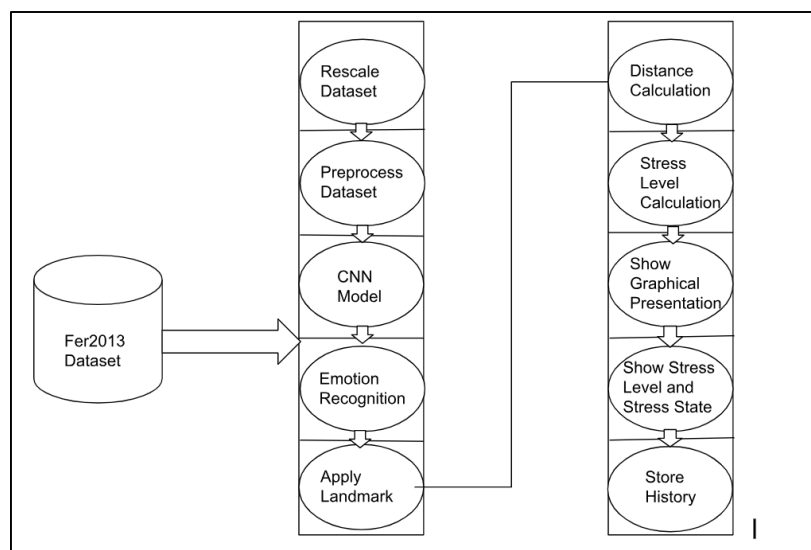


Figure 1: Architecture of our Proposed Model

3.1 Evaluation Corpus

Our data collection process centers on the “FER-2013 dataset”, a publicly available resource. This dataset contains 35,887 facial images in PNG format, each categorized into one of seven distinct emotional expressions: (1) angry (4,953 images), (2) disgusted (547 images), (3) fearful (5,121 images), (4) happy (8,989 images), (5) neutral (6,198 images), (6) sad (6,077 images) and (7) surprised (4,002 images). As a convenience, the dataset is further separated into training and testing sets.

4. EXPERIMENTAL SETUP AND EVALUATION

Next, we present our experimental setup and evaluation of our proposed model on fer213 dataset. Categorically, we explain the process of rescaling and pre-processing, data splitting, model selection and its parameter, brief description of CNN and application of landmarks, distance and stress calculations and the user interfaces.

4.1 Rescaling and Preprocessing:

Preprocessing and rescaling for a machine learning model are typically required for image classification tasks. The editing process is divided into the following steps:

1. 'Rescale=1. /255' is a rescaling factor applied to the pixel values of the image. The pixel numbers are shifted from [0,255] to [0,1]. This is a typical step in neural network preparation since it helps to standardize the input that is being fed into the network.
2. 'target_size= (48, 48)' makes all images 48x48 pixels in size.
3. The value of the parameter 'batch_size=64' instructs the computer that each batch of training data contains 64 pictures.
4. The specification "color_mode: grayscale" instructs the computer to transform the pictures to grayscale, resulting in a single channel rather of three channels as in RGB.
5. The parameter "class_mode='categorical'" indicates that the algorithm will use categorical classification, suggesting that the data has been separated into discrete categories and that it will only examine labels that are unique within the dataset.

4.2 Data Split

To evaluate our stress detection model, Our dataset was partitioned into two separate sections: a practice set (80% of the data, [28,709 images]) used to build and refine the model, enabling pattern recognition for emotional expressions, including stress, and a testing set (20%, [7,177 images]) for assessing its ability to generalize and accurately detect emotion from facial expressions. The parameters of our model have been revealed in figure 2.

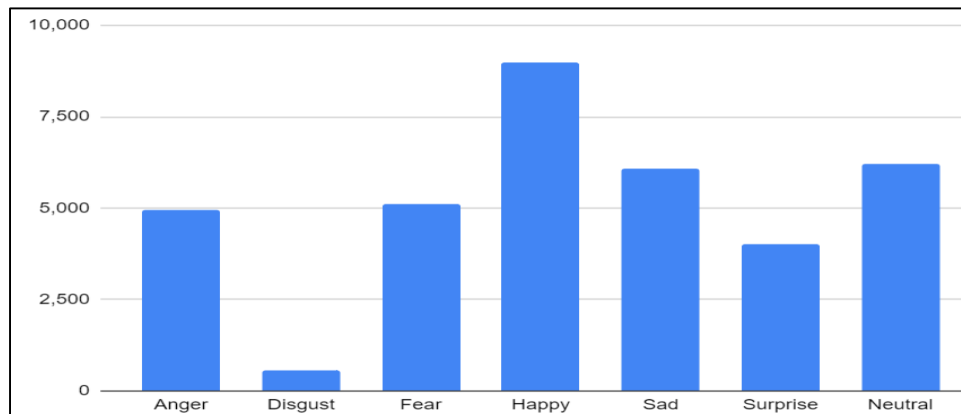


Figure 2: Model Parameters

4.3 Model Selection for Emotion Recognition

In the context of deep learning techniques, a pivotal phase in our research involved the experimentation with various model architectures to ascertain the optimal choice for emotion recognition. We conducted a rigorous evaluation of three distinct models: VGG16, Mini exception, and a Convolutional Neural Network (CNN) that has been custom-designed. The results of training the model on fer2013 dataset have been displayed in table 1.

Table 1: Results of Model Training On Fer2013 Dataset

Parameter	Mini-Exception	VGG	CNN
Training Accuracy	70.00%	89.16%	95.60%
Training Loss	0.81%	0.31%	0.12%
Validation Accuracy	64.50%	61.94%	81.93%
Validation Loss	0.98%	1.25%	0.93%

Each model was meticulously trained and assessed for its performance in recognizing emotional expressions from facial images. After rigorous experimentation, the CNN model emerged as the preferred choice due to its superior accuracy in our specific context. The CNN architecture utilized in Conv2D layers comprised of 32 filters with 3x3 kernels each were strategically chosen to improve computational efficiency in this study. This choice was driven by the model's ability to efficiently capture and recognize patterns associated with emotional expressions, including stress. Hence, the CNN model was selected as the foundation for our stress detection system, demonstrating its aptitude for our research objective.

4.4 Convolutional Neural Network (CNN)

A CNN is a type of network architecture that has been designed for certain activities such as recognizing images and pixel input analysis in the context of deep learning methodologies (P A Riyantoko et al., 2021). To improve computational efficiency, the CNN model used in this study has Conv2D layers made up of 32 filters with 3x3 kernels each. It consists of three components: input data, a filter/kernel, and a feature map. Kernel

maps are used to build an output/feature map given input data, as shown mathematically and visually in Figure 3.

$$(n * n) * (f * f) = (n - f + 1) * (n - f + 1)$$

The input matrix size is denoted by n, the kernel size by f, and the output matrix size is represented by result.

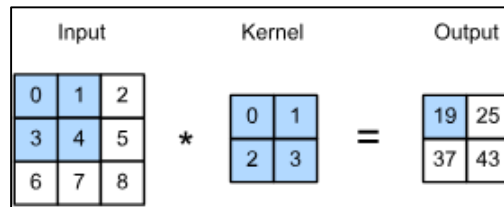


Figure 3: Mapping Example of Conv2D

A 2x2 pool-sized Max-Pooling 2D layer is added to compensate for over-fitting. It is used to lower the dimensions in order to identify images more clearly, as illustrated in Figure 4.

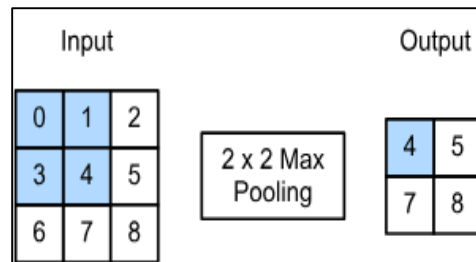


Figure 4: Max Pooling Mapping Example

In addition, a Dropout layer is purposefully added to avoid over-fitting, with a dropout rate of 0.25. The data is transformed into a 1D vector in the following stages by employing a flatten layer. An additional dropout layer with a rate of 0.5 is added to address over-fitting issues. The architecture also comprises two dense layers that are fully coupled for emotion recognition from 48x48 grayscale photos (Riyantoko & Hindrayani, 2021). Figure 5 depicts the internal, fully connected layer of CNN.

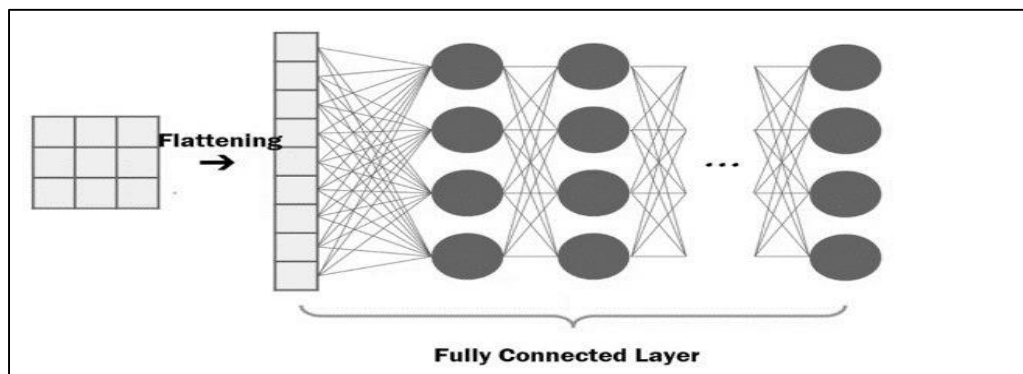


Figure 5: CNN's Fully Connected Layer

For multi-class classification, I added an output layer with softmax activation. After 50 iterations, the validation accuracy is 81.93%, while the training accuracy is 95.9%. Accuracy and loss graphs of our CNN model have been exhibited in figure 6 and figure 7 respectively.

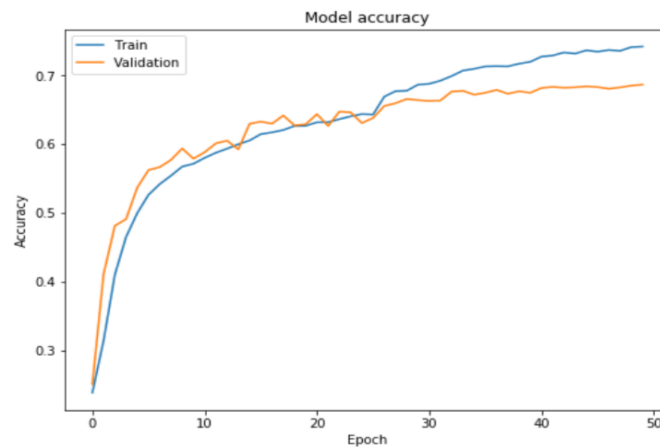


Figure 6: Accuracy Graph for our CNN Model

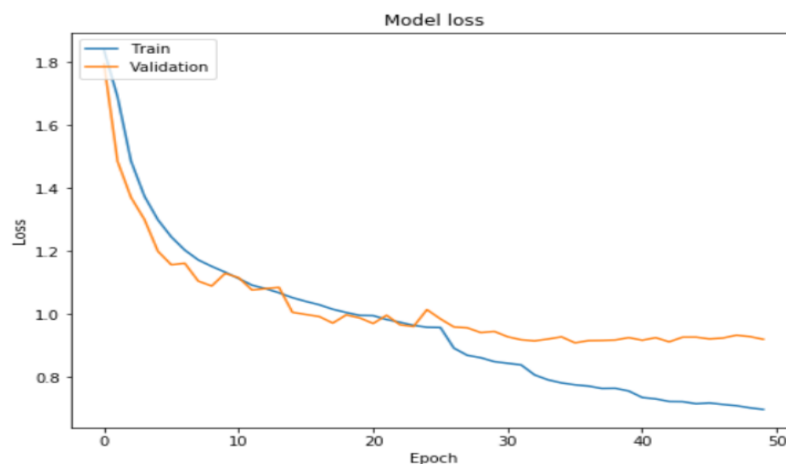


Figure 7: Loss Graph for our CNN Model

4.5 Applying Landmarks

To create a basis for stress detection, we selected three fundamental stress emotions—Sadness, Fear, and Anger—and applied landmark detection techniques to the lips and eyebrows of individuals using the 68-shape-predictor Python library, renowned for its ability to pinpoint 68 distinct facial landmarks. These landmarks, strategically chosen for their sensitivity to emotional cues, include lip corner positions and eyebrow arches, endpoints, and furrows. These landmarks serve as quantitative metrics for detecting stress from individual's face. The formulae used for land mark detection of eyebrows and lips have been shown in figure 8 and figure 9 respectively.


```
def eye_brow_distance(leye, reye):  
    global points  
    distq = dist.euclidean(leye, reye)  
    points.append(int(distq))  
    return distq
```

Figure 8: Procedure for Detecting Landmarks of Eyebrows

```
def lip_distance(lower_lip, upper_lip):  
    global points_lip  
    dist_vert = dist.euclidean(lower_lip, upper_lip)  
    points_lip.append(int(dist_vert))  
    return dist_vert
```

Figure 9: Procedure for Detecting Landmarks of Lips

4.6 Distance Calculation

Facial characteristics are important for determining stress levels. To be more precise, the eyebrows' normal posture is moved and shifted in order to indicate the level of stress (Riyantoko & Hindrayani, 2021). Both the distance between the left and right brows and the gap between the upper and lower lips are measured in order to analyze lip nervousness, biting, or compression linked to stress (Riyantoko & Hindrayani, 2021). After that, an exponential function is used to gauge the stress level and it is normalized between 1 and 100. The formulae used to compute the stress value are given below.

$$\text{Normalized_value_lip} = \frac{\text{abs}(\text{dis_lip} - \text{np.min}(\text{points_lip}))}{\text{abs}(\text{np.max}(\text{points_lip}) - \text{np.min}(\text{points_lip}))} \quad (1)$$

$$\text{Normalized_value_eye} = \frac{\text{abs}(\text{disp} - \text{np.min}(\text{points}))}{\text{abs}(\text{np.max}(\text{points}) - \text{np.min}(\text{points}))} \quad (2)$$

$$\text{Normalized_value} = (\text{normalized_value_eye} + \text{normalized_value_lip}) / 2 \quad (3)$$

This algorithm is used to calculate the victim's stress level. The stress level is subsequently calculated using the normalize value. The way the lips and eyes interact in their normal postures is what determines the "stress value". Its objective is to normalize these variables' deviations from predetermined locations in order to determine stress levels. Exponentiation of the negative average of these normalized variables yields the stress value. It provides a numerical description of a person's emotional condition or level of stress. Lower numbers indicate more stress, whereas higher values indicate lower levels of stress.

4.7 Stress Calculation

Based on the calculated "stress_value," the system categorizes stress as follows: When the "stress_value" argument is set to "None," the stress value that is returned is also set to "None." When a situation's "stress_value" is more than or equal to 0.76, it's classified as "High Stress," and when it's greater than or equal to 0.51, it's classified as "Medium

Stress". Any condition where the "stress_value" is greater than or equal to 0.26 is labeled as "Low Stress". In all other cases, it is called a "Resting State." Table 2 provides a summary of these conditions.

Table 2: Stress States

Stress Level	Stress State
75-100	High
50-75	Medium
25-50	Low
0-25	Relax

Also, a graphical representation of high to medium and low stress levels is presented for better understanding of user as presented in Figures 10, 11, and 12, in that order.

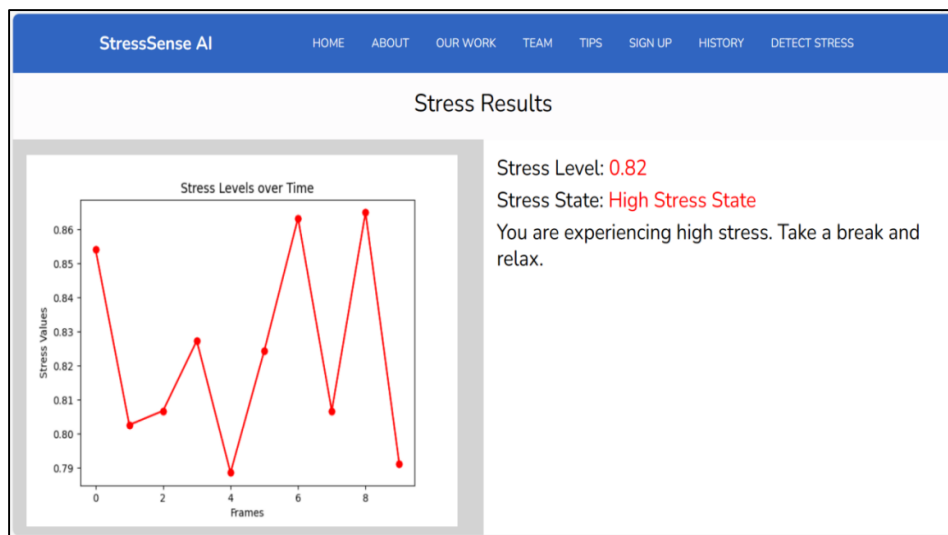


Figure 10: Representation of High Stress State of User



Figure 11: Representation of Medium Stress State of user



Figure 12: Representation of Low Stress State of user

4.8 User Interfaces

Our stress detection system incorporates a user-friendly interface designed using HTML and CSS for an engaging presentation of stress detection results. Python serves as the robust programming language underpinning our system, with the Flask framework facilitating seamless connectivity between the front-end and back-end components.

One standout feature of our system is the capability to store user stress history, enabling individuals to monitor their emotional well-being trends over time. This integration of web technologies and stress history tracking underscores our commitment to providing holistic stress management and mental health support.

5. RESULTS AND DISCUSSION

Our research culminated in the development of an efficient stress detection system. After evaluating three distinct models—Mini-Xception, VGG-16, and a custom CNN—we selected the CNN model as the optimal choice for emotion recognition.

The CNN achieved a remarkable training accuracy of 95.6% and a validation accuracy of 81.93%. A comparison of our model and baselines has been performed through figure 13.

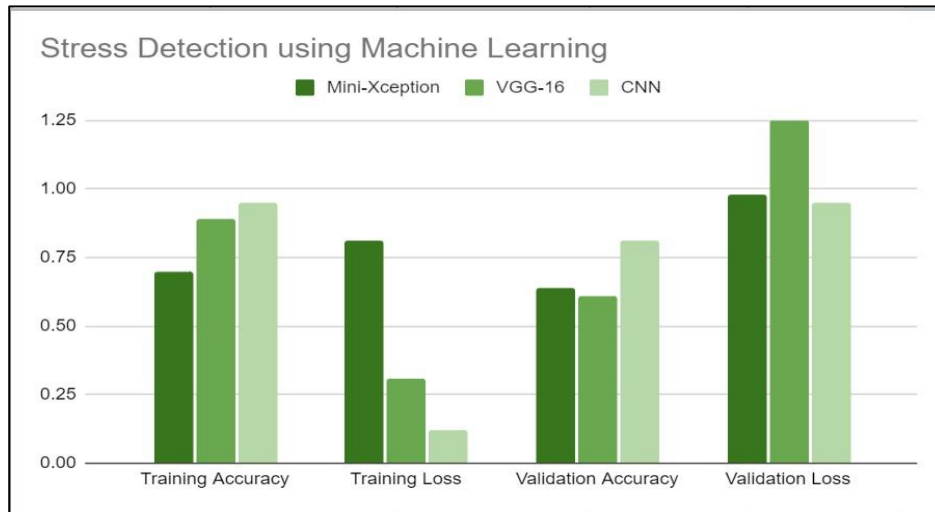


Figure 13: Comparison with Baseline Models

The following results can be derived from the comparison above:

- Training accuracy and validation accuracy are visibly higher than baselines.
- Training and validation loss is considerably lower than baselines.

The system also incorporated landmark detection techniques on lips and eyebrows, which, when combined with a stress calculation formula, accurately categorized stress levels into "High Stress," "Medium Stress," "Low Stress," or "Resting State." Our user-friendly web interface, built with HTML, CSS, Python, and Flask, allows users to access stress detection results and track their emotional well-being trends over time. These findings collectively demonstrate the system's effectiveness in recognizing and categorizing stress levels. A sample result generated by stress detector has been shown in figure 14.

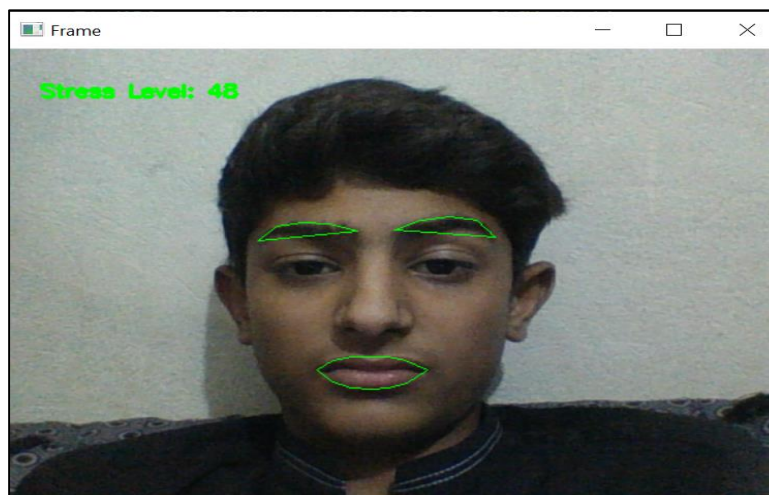


Figure 14: Results Generated by Stress Detector

6. CONCLUSION AND FUTURE RECOMMENDATIONS

In today's world, stress is a common problem that requires early detection to avoid worsening health problems. Algorithms for detection based on deep learning are widely used in many different areas. As a result, we released the Stress Detection online tool, which uses Landmark Detection and emotion recognition to assess a person's degree of stress. Additionally, this gadget keeps track of early detection data so that it can monitor a person's stress level over time. It shows an individual's stress level graphically to improve comprehension.

In the future, it may be possible to create a system that can identify stress using a variety of signs, such as heart rate and audio analysis. Furthermore, by adding new indicators like eye blinking, head tilt, and gaze movement, the existing method might be enhanced. By utilizing these qualities, this technology might help psychologists diagnose stress patients and be used by major corporations to track employee stress levels.

1. Future advancements may steer towards:
2. Compatibility Improved accuracy
3. Multi-modal data integration
4. Audio based Stress Detection
5. Compatibility to all devices

Acknowledgements

The authors would like to thank everyone, just everyone!

References

- 1) Atal, D. K., & Singh, M. (2020). Arrhythmia classification with ECG signals based on the optimization-enabled deep convolutional neural network. *Computer Methods and Programs in Biomedicine*, 196, 105607.
- 2) Brugman, S. (2022). The development of a real-time monitoring system for fatigue detection on truckers [University of Twente].
- 3) Gedam, S., & Paul, S. (2021). A review on mental stress detection using wearable sensors and machine learning techniques. *IEEE Access*, 9, 84045-84066.
- 4) Ghaderi, A., Frounchi, J., & Farnam, A. (2015). Machine learning-based signal processing using physiological signals for stress detection. 2015 22nd Iranian conference on biomedical engineering (ICBME),
- 5) Giannakakis, G., Grigoriadis, D., Giannakaki, K., Simantiraki, O., Roniotis, A., & Tsiknakis, M. (2019). Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing*, 13(1), 440-460.
- 6) Giannakakis, G., Padiaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., Marias, K., & Tsiknakis, M. (2017). Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31, 89-101.
- 7) Lim, J. Z., Mountstephens, J., & Teo, J. (2020). Emotion recognition using eye-tracking: taxonomy, review and current challenges. *Sensors*, 20(8), 2384.

- 8) Liu, D., & Ulrich, M. (2014). Listen to your heart: Stress prediction using consumer heart rate sensors. *Final. Proj. Stanford CS*, 229, 229-233.
- 9) Mayya, S., Jilla, V., Tiwari, V. N., Nayak, M. M., & Narayanan, R. (2015). Continuous monitoring of stress on smartphone using heart rate variability. 2015 IEEE 15th international conference on bioinformatics and bioengineering (BIBE),
- 10) Naidu, P. R., Sagar, S. P., Praveen, K., Kiran, K., & Khalandar, K. (2021). Stress Recognition Using Facial Landmarks and Cnn (Alexnet). *Journal of Physics: Conference Series*,
- 11) Nakashima, Y., Kim, J., Flutura, S., Seiderer, A., & André, E. (2016). Stress recognition in daily work. *Pervasive Computing Paradigms for Mental Health: 5th International Conference, MindCare 2015, Milan, Italy, September 24-25, 2015, Revised Selected Papers 5*,
- 12) Organization, W. H. (2001a). Mental disorders affect one in four people. In.
- 13) Organization, W. H. (2001b). *The World Health Report 2001: Mental health: new understanding, new hope*.
- 14) Riyantoko, P., & Hindrayani, K. (2021). Facial emotion detection using haar-cascade classifier and convolutional neural networks. *Journal of Physics: Conference Series*,
- 15) Rony, R. J., Ahmed, M., Sarcar, S., & Ahmed, N. (2021). An Exploratory Study, Wearable Development and Experiment to Understand Driving Stress in Urban Bangladesh. *Sayan and Ahmed, Nova, An Exploratory Study, Wearable Development and Experiment to Understand Driving Stress in Urban Bangladesh (August 26, 2021)*.
- 16) Sabry, F., Eltaras, T., Labda, W., Alzoubi, K., & Malluhi, Q. (2022). Machine learning for healthcare wearable devices: the big picture. *Journal of Healthcare Engineering*, 2022.
- 17) Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Translational Psychiatry*, 10(1), 116.
- 18) Zenonos, A., Khan, A., Kalogridis, G., Vatsikas, S., Lewis, T., & Sooriyabandara, M. (2016). HealthyOffice: Mood recognition at work using smartphones and wearable sensors. 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops),
- 19) Zhuang, Y., & Xie, R. J. (2021). Mechanoluminescence rebrightening the prospects of stress sensing: a review. *Advanced Materials*, 33(50), 2005925.