Towards Real-Time Facial Emotion-Based Stress Detection Using CNN and Haar Cascade in AI Systems

Archana Balkrishna Yadav [*]

[*]**Corresponding Author:** Archana Balkrishna Yadav

Received: 28-09-2024 Revised: 11-10-2024 Accepted: 30-10-2024

ABSTRACT

Understanding human conduct requires the ability to recognise facial emotions, which has applications in everything from human-computer interaction to psychological wellness monitoring. This research provides a new approach to stress detection using Convolutional Neural Networks (or CNNs) and HaarCascade classifiers. The suggested method uses a CNN to recognise facial expressions and Haar Cascade algorithm for face detection. The methodology begins with preliminary processing the input photos, followed by face detection and extraction of facial regions. Those parts are then fed into the CNN model, which classifies emotions. The system has been trained and tested on publicly available datasets, with encouraging results in stress detection accuracy. This method, which detects stress through facial expressions, has potential uses in stress management, mental health evaluation, and personalised therapies.

Face expressions have an important part in transmitting emotions, especially stress, which is a common problem in today's fast-paced world. This research provides a novel approach for detecting stress by analysing facial expressions with Convolutional Neural Networks(CNNs)and Haar Cascade classifiers. The proposed system enhances the precision and effectiveness of stress detection by combining the benefits of both approaches.

The methodology begins by preprocessing the input photos to improve their quality and normalise them for subsequent analysis. Haar Cascade classifiers are then used to detect faces in the images, ensuring precise identification of facial regions even under different lighting conditions and orientations. The discovered faces are removed and resized to produce homogeneous inputs for further processing.

Keywords— CNN, Haar Cascade, AI

I. INTRODUCTION

In modern fast-paced and demanding environment, stress has become a daily occurrence, impacting millions of individuals globally. Chronic stress has a detrimental impact on a person's physical health, work efficiency, and quality of life in addition to their mental and emotional health. Early detection and management of stress are critical for leading a healthy lifestyle and avoiding consequences. Traditional techniques of stress measurement frequently rely on self-reporting, which is subjective and prone to bias. However, new potential for objective stress detection has been made possible by recent technological advancements, particularly in the fields of computer vision and machine learning. One interesting strategy is to analyse facial expressions, which are strong markers of an individual's emotional condition, including stress.

Facial emotion detection has received a lot of attention in both academic research and commercial applications because of its non- intrusiveness and high accuracy. Analysis of facial expressions can reveal underlying emotions such as happiness, sadness, anger, fear, and, most crucially, stress. In a number of situations, such as stress management, human-computer interaction, and mental health monitoring, this information can be quite helpful.

In this regard, this research gives a thorough investigation into the creation of a faceemotion-based stress detection system employing Convolutional Neural Networks (also known as CNNs) and Haar Cascade classifiers. CNNs have performed exceptionally well in applications involving picture categorisation, especially facial recognition. They can learn and extract complex information from facial photos, allowing them to accurately discern between various emotional states.

Haar Cascade classifiers, on the other hand, offer reliable and efficient face detection capabilities. They use a machine learning approach to recognise facial characteristics including the eyes, nose, and mouth, allowing for precise localization of facial regions inside images. This is critical for identifying the areas of interest (i.e., the face) prior to further research.

The combination of CNNs and Haar Cascade classifiers provides a synergistic method to stress detection, taking advantage of the capabilities of both systems. By integrating CNN's capacity to recognise complicated patterns in facial expressions with Haar Cascade's precise face identification, the suggested approach seeks to deliver accurate and reliable stress assessments.

II. LITERATURE SURVEY

[1]The system is intended to identify an individual's stress level based on facial analysis and emotion identification. It's a straight forward application that uses the computer's or smartphone's front camera and doesn't require any additional external gear. It possesses been created with a greater emphasis on students and the younger generation and a lesser emphasison adults due to the younger generation's propensity for excessive use of smart devices.

[2]It is possible to identify emotional stress by examining several face features. In order to distinguish between a neutral and a stressed or anxious state, the automated identification of facial Action Units (AU) as quantitative markers is the main focus of this research. Therefore, a model for the automatic recognition of facial action units is suggested, trained on the UNBC and BOSPHORUS datasets - two publicly available annotated facial datasets.AAM land marks' geometric (non-rigid deformations of their 3D shape) and morphological (Histograms of Oriented Gradients)features were retrieved. Each AU's intensity was regressed using the Support Vector Regression (SVR) technique. Each dataset's matching models were fused together to create a single composite model.

[3]The idea behind the project is to measure and identify people's stress levels by observing their facial expressions, especially how the lips and eyebrows move. It analyzes high- dimensional features taken out of face photos taken with a regular camera by utilizing machine learning and deep learning techniques. Determining if people are under high or low stress and forecasting stress levels and stress intensity are the key goals. The goal of the project is to create a simple, non-invasive way to detect and manage stress by tracking lip and brow movements in real-time video

[4]The contraction and relaxation of muscles beneath the skin of the face is referred to as a facial expression. Facial expressions are a kind of nonverbal communication. Without the use of words, the human face may convey a vast array of emotions. Happiness, sadness, neutrality, surprise, wrath, fear, and contempt are a few universal emotions. Users' emotional well-being can be tracked and their physiological and mental health can be screened for with the usage of emotion detection. Stress, anxiety, and depression at work can all have a negative impact on productivity and absence. Our goal is to operate in real-time, identifying emotions from photos taken by a live webcam.

III. PROPOSED SYSTEM

Although these current technologies offer useful stress management tools, the project's suggested facial-based stress detection systems may be more specific and accurate. To deal with this matter considering the strong correlation between emotional states and facial expressions, facial expression analysis utilizing Haar-Cascade and CNN provides a non- invasive and potentially more intuitive method of stress detection. The system uses deep learning algorithms to study and interpret face expressions in real-time, with the goal of detecting facial landmarks. These algorithms can precisely recognize minute indicators linked to stress since they have been trained ona wide range of datasets.

Aspects like eye and face movements are taken into account to improve the accuracy of stress detection. The suggested system may offer more sophisticated insights into stress levels and facilitate more focused stress management therapies by concentrating on facial clues. The suggested approach combines image processing methods with deep learning algorithms to detect stress in real time by analyzing facial expressions. Exercises, yoga positions, and individualized interventions are examples of remediation techniques that successfully reduce stress.

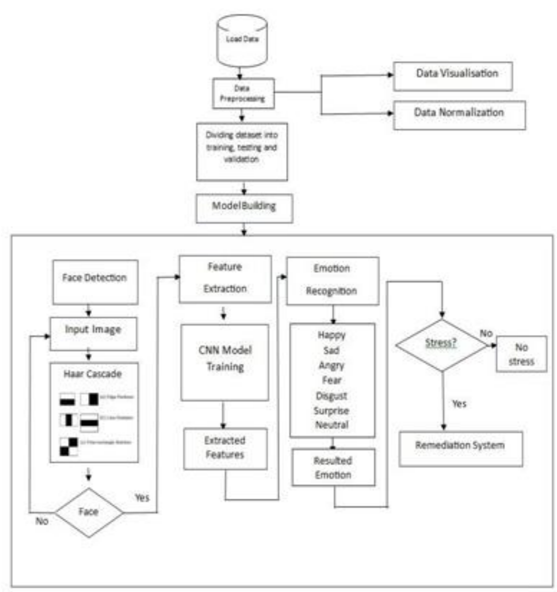

Figure 1: Architecture of the proposed system

IV. DATASET

FER2013

One of the most widely used datasets for tasks involving facial expression recognition is Fer2013. The "Facial Expression Recognition Challenge" was

Peer Reviewed & Refereed Journal
Volume-14, Issue-5 (October 2024)
https://doi.org/10.5281/zenodo.14064731

developed for the 2013 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR). The dataset consists of 35,887 images of faces cropped to 48x48 pixels, with grayscale intensity values. One of seven feeling categories—disgust, anger, fear, happiness, sorrow, surprise, or neutral—is assigned to each image.

Three subsets of the dataset were created: 28,709 photos for the training set, 3,589 images for the public test set, and 3,589 images for the private test set. The training set is commonly used for training machine learning models, while the public test set is used for validation during model development, and the private test set is used for final evaluation to assess the generalization performance of the model.

Figure 2: Categories of Emotions

V. METHODOLOGY

1. Convolutional Neural Networks (CNNs)

Convolutional neural networks, or CNNs, are strong deep learning algorithms that are frequently used in applications involving the processing of images and videos. Its ability to maintain spatial relationships makes it very successful at obtaining pertinent information from visual data. Large labelled datasets, in which each image is assigned to a particular class or category, are used to train CNNs. The CNN gains the ability to recognize patterns and characteristics in the images through training. For image categorisation, CNNs often employ a specific layer design. This architecture's layers are as follows:

The picture being input data is received by the input layer.

Convolutional Layers: These layers serve as the main building blocks of CNNs. They apply convolutions to the input image using learnable filters or kernels.

Activation Function: Following each convolutional layer, an activation function is applied element-wise to the convolution output.

Pooling Layers: These layers oversee preserving important information while shrinking the spatial dimensions of the characteristics that have been retrieved.

These components collectively contribute to the effective feature extraction and classification capabilities of CNNs in image analysis tasks.

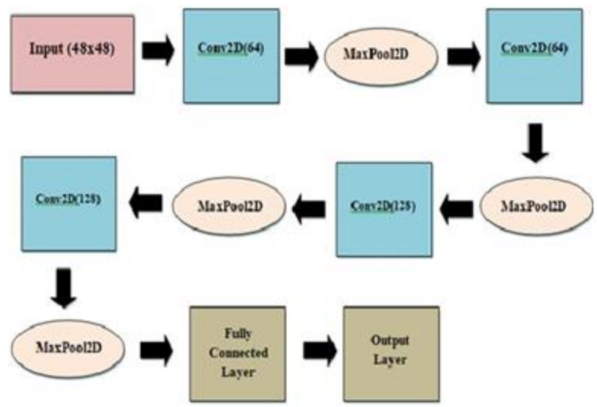

Figure 3: CNN Workflow

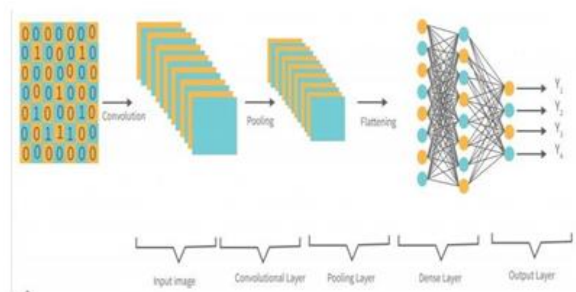

Figure 4: Architecture of CNN

2. HaarCascade Algorithm

Haarcascading is computationally efficient, especially during the detection phase. The cascade structure allows for the quick rejection of regions of an image that are unlikely to contain the object, reducing the computational burden compared to exhaustive search methods.

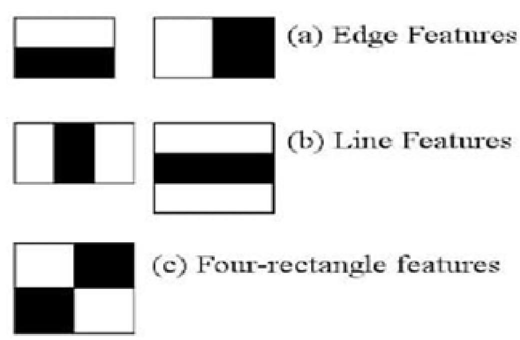

Figure 5: HaarCascade Algorithm

Haarcascade technique is used to detect faces in each image or each frame in the window. Faces in an image can be found with this object detection algorithm. Features for edge or line detection are used by the method.

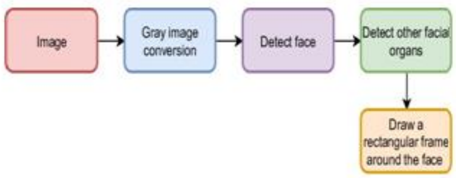

Figure 6: Working of Haar-Cascade

3. Training Procedure

The training procedure for the proposed facial-based stress detection system involves several steps:

Data Collection: Compile a wide range of face expressions linked to various stress levels. Images or videos showing people reacting in different ways to stress-inducing events with different facial expressions should be included in this dataset.

Preprocessing: Preprocess the data to enhance its quality and standardise it for training. This could entail operations like alignment, augmentation, and normalization to guarantee uniformity throughout the collection. Detection of Facial Landmarks: Employ deep learning methods, including Convolutional Neural Networks (CNNs), to precisely identify facial landmarks. These landmarks function as frames of reference for reading expressions on the face.

Feature Extraction: This involves identifying characteristics from the expressions on the face, such as the eye movements, facial movements, and other underlying signs of stress. To train the stress detection model, these features are necessary.

Model Training: Utilizing the tagged and extracted features from the varied dataset, train the stress detection model. To understand the intricate patterns and connections between stress levels and facial expressions, deep learning methods like CNNs are used.

Evaluation: Utilizing validation data, appraise the trained model's efficacy in identifying stress levels from facial expressions in terms of accuracy, specificity, and resilience.

Integration: Integrate the trained model in the proposed system and use image processing methods to combine it with facial expression analysis for real-time stress detection.

Remediation Strategies: Depending on the levels of stress that have been identified, create remediation techniques (e.g., yoga positions, exercises) to successfully reduce stress.

4. DataPreparation

A varied dataset of facial expressions connected to different stress levels must be put together in order to provide the data for the proposed facial-based stress detection system. Preprocessing, such as alignment and normalization, is applied to this collection in order to improve quality and standards. Then, to reliably identify facial landmarks—which are essential for further analysis—deep learning techniques like Convolutional Neural Networks (CNNs) and HaarCascade are used. In order to identify tiny signs linked to stress, features including eye and facial movements are retrieved. Data augmentation techniques are also employed to increase dataset variety and model durability. The method attempts to create deep learning models that can reliably identify stress levels from facial expressions in real- time, enabling more focused stress management interventions, by carefully preparing the data.

VI. RESULTS AND ANALYSIS

The goal of the proposed facial-based stress detection system is to use the specificity and accuracy of facial expression analysis to overcome the limitations of current stress management techniques. The system takes advantage of the close connection between emotional states and facial expressions to provide a non-invasive and user-friendly method of stress detection through the use of Convolutional Neural Networks (CNNs) and Haar-Cascade. Real-time facial expression analysis is accomplished by deep learning algorithms that have been trained on a variety of datasets to precisely recognize minute indicators of stress, such as changes in the eyes and face.

As a result of its emphasis on facial expressions, the proposed system can offer more sophisticated insights into stress levels, allowing for more focused stress management strategies. Real-time stress detection through facial expression analysis is made possible by the system's integration of deep learning algorithms and image processing techniques. To further successfully reduce stress, remediation techniques including yoga positions, exercises, and tailored treatments are included.

Figure 7

Peer Reviewed & Refereed Journal
Volume-14, Issue-5 (October 2024)
https://doi.org/10.5281/zenodo.14064731

Figure 8: .Home Page

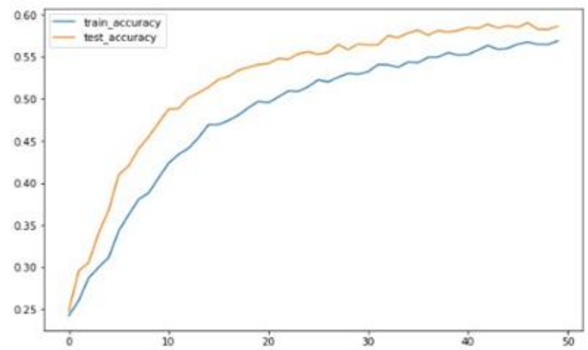

Figure 9: Plot for Training accuracy and Testaccuracy

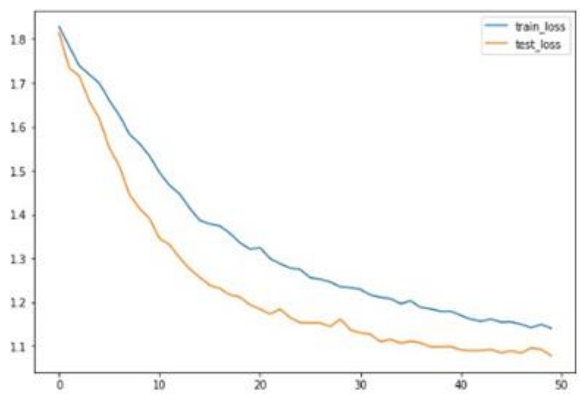

Figure 10: Plot of Training loss and TestLoss

	precision	recall	f1-score	support
0	0.46	0.54	0.50	958
1	0.66	0.21	0.32	111
2	0.48	0.23	0.31	1024
3	0.77	0.86	0.81	1774
4	0.53	0.56	0.54	1233
5	0.44	0.51	0.47	1247
6	0.74	0.71	0.72	831
accuracy			0.59	7178
macro avg	0.58	0.52	0.52	7178
weighted avg	0.58	0.59	0.57	7178

Figure 11: Table of metric measures

Figure 12: Face Detection

Figure 13: Level-1

Figure 14: Level-2

Figure 15: Level-3

VII. CONCLUSION

In conclusion, by examining facial expressions, the suggested facial-based stress detection system provides a non-invasive, real-time method of accurately detecting stress levels. It recognizes minor signs linked to stress by utilizing deep learning algorithms and image processing techniques, allowing for customized interventions for stress management. With rapid and accurate stress evaluations, this approach has the potential to improve mental health awareness and individual well-being.

REFERENCES

[1] Ekman P & Friesen WV. (1978). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.

[2] B. Martinez, M. F. Valstar, B. Jiang & M. Pantic. (2019). Automatic analysis of facial actions: A survey. *IEEE Transactions on Affective Computing*, *10*(3), pp. 325-347.

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

[4] F.Bevilacqua, H.Engstrom & P.Backlund. (2018). Automated analysis of facial cues from videos as a potential method for differentiating stress and boredom of players in games. *International Journal of Computer Games Technology, 1*, 8734540.

[5] Daudelin-Peltier, C., Forget, H., Blais, C., Deschênes, A. & Fiset, D. (2017). The effect of acute social stress on the recognition of facial expression of emotions. *Scientific Reports*, *7*(1), 1036. https://doi.org/10.1038/s41598-017-01053-3.

[6] P.Ekman & W.Friesen. (1978). *Facial action coding system (FACS): Manual*.

[7] Viola, P. & Jones, M.J. (2004). Robust real-time face detection. *International Journal of Computer Vision, 57*, 137–154.

[8] I.Matthews & S.Baker. (2004). Active appearance models revisited. *International Journal of Computer Vision, 60*(2), 135-164.

[9] D.Watson. (2013). Contouring: a guide to the analysis and display of spatial data. *Elsevier, 10*.

[10] T.Baltrusaitis, M.Mahmoud & P.Robinson. (2015). Cross-data set learning and person- specific normalization for automatic action unit detection, in Automatic Face and Gesture Recognition(FG). *11th IEEE International Conference and Workshops 6*. IEEE.

[11] T.F.Cootes & C.J.Taylor. (1992). Active shape models smart snakes, in BMVC92. *Springer*, pp.266–275.

[12] P.Carcagn, I,M.DelCoco, M.Leo & C.Distante. (2015). Facial expression recognition and histograms of oriented gradients: a comprehensive study. *Springer Plus, 4*(1), pp.645.

[13] J.Smola & B.Scholkopf. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222.

[14] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon & I. Matthews. (2011). Painful data: The unbc-mcmaster shoulder pain expression archive database. In: *Automatic Face Gesture Recognition and Workshops*, pp. 57–64.

[15] Savran,N., Alyuz,H., Dibeklio˘glu, O. C¸ eliktutan, B. Gokberk, B. Sankur & L. Akarun. (2008). Bosphorus database for 3d face analysis. In: *European Workshop on Biometrics and Identity Management. Springer*, pp. 47–56.

[16] Ding & H. Peng. (2005). Minimum redundancy feature selection from microarray gene expression data, *Journal of Bioinformatics and Computational Biology, 3*(2), 185–205.

[17] G.Gulgezen, Z.Cataltepe & L.Yu. (2009). Stable and accurate feature selection. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 455–468.

[18] Y. Andreu, F. Chiarugi, S. Colantonio, G. Giannakakis, D. Giorgi, P. Henriquez, E. Kazantzaki, D. Manousos, K. Marias, B. J. Ma tuszewski, M. A. Pascali, M. Pediaditis, G. Raccichini & M.Tsiknakis. (2016). Wizemirror-a smart, multisensory cardio-metabolic risk monitoring system. *Computer Vision and Image Understanding, 148*, 3–22.

[19] G. Giannakakis, D. Manousos, V. Chaniotakis & M. Tsiknakis. (2018). Evaluation of head pose features for stress detection and classification. In: *IEEE EMBS International Conference on Biomedical & Health Informatics*, pp. 406–409.

[20] K. Anis, H. Zakia, D. Mohamed & C. Jeffrey. (2018). Detecting depression severity by interpretable representations of motion dynamics. *13TH IEEE International Conference on Automatic Face & Gesture Recognition (FG2018)*, pp. 739–745.