Rajarambapu Institute of Technology, Rajaramnagar

Department of Information Technology

A. Y. 2023-24

Capstone Project Synopsis

Domain Area	Deep Learning
Title of the project	Deep Learning approach for Stress, Anxiety, Depression detection by Real Time Facial Expressions using Image Processing.
Project Guide Name	Prof. P. N. Jadhav
Team Leader's Name	Renuka Gour
Group Number	04

Members

Sr. No.	Roll No.	Name	Email	Phone	Sign
1	2110038	Shraddha Nandkumar Irale	2110038@ritindia.edu	8446850955	
2	2110041	Renuka Sonulal Gour	2110041@ritindia.edu	9322152539	
3	2110057	Sonali Sudam Suryawanshi	2110057@ritindia.edu	7249153670	
4	2110065	Rutuja Vilas Kulkarni	2110065@ritindia.edu	8080252251	

Prof. P. N. Jadhav	Prof. P. N. Jadhav	Dr. A. C. Adamuthe
Sign of Project Guide	**Sign of Project In-charge**	**Sign of HOD, IT**

1. Introduction

Mental health remains a global concern, with stress, anxiety, and depression affecting millions of lives. As technology continues to advance, there is a growing opportunity to harness the power of deep learning and real-time image processing for proactive mental health monitoring. This project endeavors to develop a cutting-edge system that leverages facial expression analysis to detect emotional distress in real-time, offering a non-intrusive and accessible solution for early identification. Facial expressions are potent indicators of one's emotional state, and recent advancements in deep learning have facilitated the creation of highly accurate models for recognizing nuanced expressions. By deploying Deep Learning Algorithm and image processing techniques, this project seeks to capture subtle facial cues associated with stress, anxiety, and depression. The aim is to create a system capable of swift, reliable, and continuous emotional assessment, fostering timely intervention.

The significance of this endeavor lies in the potential to revolutionize mental health care by providing an objective and automated tool for early detection. The integration of real-time image processing ensures the adaptability of the system to dynamic environments, making it suitable for diverse contexts where prompt intervention is essential. Moreover, this project envisions a future expansion by incorporating voice-based recognition alongside facial expression analysis. By combining these two modalities, the system can offer a more comprehensive understanding of an individual's emotional well-being. This introduction sets the stage for a transformative exploration into technology-assisted mental health monitoring, highlighting the project's commitment to addressing a pressing global challenge through innovative and multidimensional approaches.

2. Objectives

1. To determine Depression, Anxiety, Stress levels by analyzing Real Time facial expressions using image processing.
2. To build a low-cost system that provides high accuracy and low prediction time that can be used by health practitioners for assessing and monitoring the severity of Depression, Stress, Anxiety.

3. To evaluate the severity of the Stress, Anxiety, Depression and use them for a wide range of applications.

4. Act as a preventive measure for people who are likely to be get into mental health related issues

3. Literature Survey

A. Phani Sridhar, R. Jahnavi Pramodhani, S. Padmini Priya, Ch Kanoj Kumar (2023) [1] has published a paper on Human Stress Detection using Deep Learning The examination of facial expressions of an employee is handled using a deep learning algorithm named Convolution Neural Network(CNN) and even based on image processing using open-source as OpenCV. A camera is used to collect the face photos, which are then processed to extract information and the convolution neural network, which identifies stressful reactions and applies predictions to the stressed person. The objective of the project is to provide a system that analyze and predict the stress existence in a person by using deep learning algorithm techniques. The paper does not provide a specific accuracy figure for the algorithm implemented. However, it mentions that the project accurately identifies stress and records the user's feelings.

Adel Aref Ali Al-zanam, Omer Hussein Abdou Elsayed Hussein Alhomery , Choo Peng Tan(2023) [2] proposed a Mental Health State Classification Using Facial Emotion Recognition and Detection This paper made approach for determining mental health states by aggregating facial emotion scores over seven days. This innovative technique allows for a comprehensive and dynamic representation of an individual's emotional landscape, facilitating a more accurate classification of mental health status. The method they use comprises several interconnected components, including preprocessing, feature extraction using Principal Component Analysis (PCA) and VGGNet, and classification using Support Vector Machines (SVM) and Multilayer Perceptron (MLP).the dataset they use is FER2013 dataset. The accuracy of the algorithm implemented in the paper varies based on the different models and methods used. The best model, which combined VGGNet feature extraction with SVM classification, achieved an accuracy of 66%.

Sabina Umirzakova, Shabir Ahmad , Sevara Mardieva, Shakhnoza Muksimova, Taeg Keun Whangbo (2023) [3] proposed a Deep learning-driven diagnosis: A multi-task approach for segmenting stroke and Bell's palsy**.**This paper presented a deep-learning-based model for the early detection of facial paralysis disorders, such as Bell's palsy and stroke. can analyze a patient's facial asymmetry and parse six categories, namely, skin, eyebrow, eye, mouth, and stroke. They have investigated network by combining the convolution module and enhanced stroke category (ESC) blocks.The authors in this paper used two datasets a public dataset for Bell's palsy patients and a private Gil Hospital dataset for stroke patients and healthy individuals. The accuracy of the algorithm implemented in the paper is not explicitly mentioned. However, the proposed method demonstrated has improvement of 9.69% on the Gil Medical Center database and achieved a mean IoU of 6.28% and 5.91% of mean AP higher than other state-of-the-art models trained using the same parameters. Additionally, the proposed model achieved a mean F1 score of 98.2% in the area of facial paralysis. This is a significant advancement as early and accurate diagnosis is critical in the management of facial paralysis disorders.

Vandana, Nikhil Marriwala, Deepti Chaudhary **(**2023) has published A hybrid model for depression detection using deep learning [4] the proposed work has designed three models first is textual CNN model, second is audio CNN model, third is hybrid LSTM(long Short-Term model) and Bi-LSTM model. Hybrid model is a combination of audio and text modalities; therefore, it is named as a hybrid structure. It automatically detects depression using deep learning. DAIC-WOZ Depression Database is used for Automatic Depression Detection system. The results show that audio CNN model gives more accurate results in comparison to text CNN model, LSTM model lost the past data whenever new information add in the model. Loss is minimum in all models which make accuracy of models high. it can easily predict early symptoms of depression with an accuracy of 98% and loss of 0.1%, whereas text CNN give accuracy of 92% and loss of 0.2%. Bi-LSTM model and audio CNN has better values as compared to LSTM model and textual CNN model. therefore, text CNN gives accuracy over 92% audio CNN about 98% hybrid LSTM gives 80% whereas hybrid Bi-LSTM 88% this study proves that CNN model provides more accuracy than LSTM model.

Jing Zhang, Hang Yin, Jiayu Zhang, Gang Yang, Jing Qin and Ling He(2022) [5] published Real-time mental stress detection using multimodality expressions with a deep learning framework**.** The objective of this research paper is to propose a real-time deep learning framework for the detection of acute stress by fusing ECG, voice, and facial expressions. The study aims to address the growing concern of mental stress in modern society and to demonstrate the effectiveness of using objective indicators and multimodality fusion for stress detection. the multidimensional features of each modality were extracted through the deep learning framework. The fully connected layers in the framework obtained the information about the stress state, and the framework fused them into a global matrix for stress detection. The real-time deep learning framework was developed by combining ResNet50 and I3D with the temporal attention module. The datasets presented in this article are not readily available because they contain identifiable information such as recognizable faces. The accuracy of the algorithm implemented in this paper varies based on the modality used for stress detection. The method achieves 74.1%, 79.2%, and 83.0% detection accuracy using ECG, facial expressions, and voice, respectively. When fusing the information from all modalities, the accuracy increases to 85.1%. These results demonstrate the effectiveness of the proposed multimodality fusion method for stress detection.

Rehmat Ullah, Hassan Hayat, Afsah Abid Siddiqui, Uzma Abid Siddiqui, Jebran Khan, Farman Ullah, Shoaib Hassan, Laiq Hasan, Waleed Albattah , Muhammad Islam, and Ghulam Mohammad Karami (2022) [6] presents A Real-Time Framework for Human Face Detection and Recognition in CCTV Images. The objective of the research paper is To develop a machine learning and deep learning- based real-time framework for detecting and recognizing human faces in closed-circuit television (CCTV) images. Two feature extraction algorithms, where used one is principal component analysis (PCA) and other convolutional neural network (CNN).the paper compares the performance of the algorithms using K-nearest neighbor (KNN), decision tree, random forest, and CNN.they have used dataset of 40K images with different environ mental conditions, clutter backgrounds, and occlusion the research found was recognized faces with a minimum computing time and an accuracy of more than 90%.the ultimate scope for development is enhancement of this system by making it a complete security system.the system used to recognize a single face from the image the next step is to recognize multiple faces in a live-streaming video.

Di Lu and Limin Yan Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor Journal of Sensors, vol. 2021 [7] The detection result of the OpenCV method has a large deviation, and the angular offset is slightly larger, resulting in an undetectable result. Ordinary photoelectric sensors have only a single light sensor element. In many applications, multiple such sensors are often required to detect the various characteristics of the components, and the vision sensor can capture an image containing millions of pixels, to be able to condition the inspection components shown in detail, which can prevent missed inspections and improve the accuracy of the inspection, which is especially necessary for the inspection of electronic components. The information collected by the image sensor greatly simplifies the bandwidth of the filter in the subsequent image signal processor and enhances the signal-to-noise ratio, ensuring the quality of imaging within the available dynamic range.

Edeh Michael Onyema, Piyush Kumar Shukla, Surjeet Dalal, Mayuri Neeraj Mathur, Mohammed Zakariah, Basant Tiwari, Enhancement of Patient Facial Recognition through Deep Learning Algorithm: ConvNet, Journal of Healthcare Engineering, vol. 2021[8] There are different data augmentations applied to dataset in order to increase the samples in the dataset. Along with data augmentation, they have designed the convolutional neural network model by introducing a residual block in the existing ConvNet model. They have chosen different sets of parameters for optimizing the model and to improve the learning of the model. The purpose of the data augmentation is to resolve the issue of limited samples in the dataset to some extent along with increasing the diversity of the data. Commonly used data augmentations are flip, rotation, scale, crop, translation, and Gaussian noise along with some advanced data augmentation techniques.

Russell Li and Zhandong Liu Stress detection using deep neural networks **(**From The International Conference on Intelligent Biology and Medicine (ICIBM) 2020 August 2020) [9] We developed two deep neural networks: a deep 1D convolutional neural network and a deep multilayer perceptron neural network. The networks analyzed physiological signals measured from chest-worn and wrist-worn sensors to perform the two tasks of binary stress detection and 3-class emotion classification. The performance of the two deep neural networks were evaluated and compared with that of traditional machine learning algorithms

used in previous research. The results indicate that the two deep neural networks performed significantly better for both tasks than the traditional machine learning algorithms. We demonstrated the potential of deep neural networks for developing robust, continuous, and noninvasive methods for stress detection and emotion classification, with the end goal of improving the quality of life.

Pramod Bobade, Vani M Stress Detection with Machine Learning and Deep Learning using Multimodal Physiological Data Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020) [10] The proposed research work has understood the structure and format of the publicly available WESAD dataset, cleaned and transformed data to a set eligible to construct machine learning and deep learning classification methods, explored and constructed various classification models and compared them. WESAD dataset contains data from multiple physiological modalities like three-axis acceleration (ACC), respiration (RESP), electrodermal activity (EDA), electrocardiogram(ECG), body temperature (TEMP), electromyogram (EMG) and blood volume pulse (BVP) which is not available in other datasets, which makes this work suitable for the detection of stress in human being. This model has achieved the accuracy of 84.32% and 95.21% on a three-class and a binary classification problems.

4. Problem Statement

Deep Learning approach for Stress, Anxiety, Depression detection by Real Time Facial Expressions using Image Processing.

5. Problem Description

Mental health disorders, including stress, anxiety, and depression, are widespread and can significantly impact an individual's daily life. Timely detection and intervention are crucial for effective management. The fast-paced nature of modern life, coupled with various socio-economic factors and lifestyle changes, has contributed to a significant increase in the prevalence of these mental health disorders. The impact of stress, anxiety, and depression extends beyond individual well-being, affecting relationships, work productivity, and overall societal health. There is an urgent need for comprehensive and accessible solutions to address these challenges. Doctors use a combination of methods to assess stress, anxiety, and depression in individuals. The process typically involves a thorough evaluation of the

patient's medical history, a physical examination, Laboratory Tests and various psychological assessments, so we need to design a solution which saves time of doctor to analysis patient for that a deep learning system is used for the detection of stress, depression, and anxiety using face recognition techniques.

6. System Architecture

The proposed work involves developing a real-time Stress, Anxiety and Depression detection system using a Deep Learning approach and predictor.

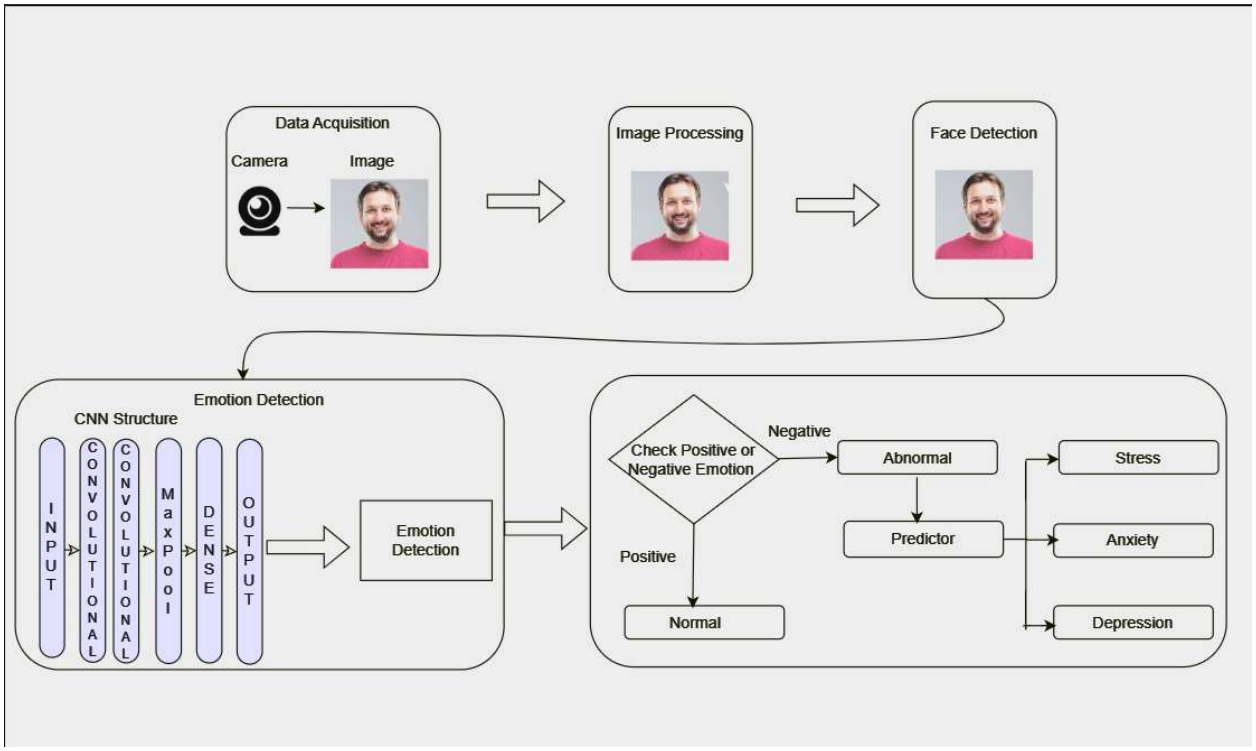

Fig. 1 System Architecture of Mental Health Disorder Detection

Fig. 1 Shows that the system begins by capturing an image through the camera, followed by image processing to detect the face. Subsequently, a Deep Learning approach is applied to recognize emotions. If the detected emotions are positive, it indicates that the person is in a normal emotional state. Conversely, if negative emotions are identified, a predictor is used to determine the Stress, Anxiety, and level of Depression as the output.

7. Modules

7.1 User Interface

Camera

7.2 Input Data

Add image – Capture through camera.

7.3 Output Data

Emotion Detection by facial expression:

1. Predict Stressed or not Stressed.
2. Predict Anxiety Level
3. Predict Depression Level

8. System requirements with justification

- Software requirements:
  - Python
  - Deep Learning
  - JAVA
  - OpenCV
  - VS Code
  - Android Studio
- Hardware requirements
  - Internet connection
  - Mobile device must capture the image.
  - RAM: At Least 6GB
  - Camera with high resolution for better performance

9. Project Outcome

1. A low-cost and low prediction time system that accurately predicts depression, anxiety, and stress levels from facial expressions.

2.Accurate and more precise system which is used in considerations for user interpretability, ethical implications, and collaboration with mental health experts.

3. To aim at publishing a research paper.

4. To participate in various competitions or hackathon.

10.References

[1] A. Phani Sridhar, R. Jahnavi Pramodhani, S. Padmini Priya, Ch Kanoj Kumar (2023), "**Human Stress Detection using Deep Learning"** international journal of progressive research in engineering management and science (ijprems).

[2] Adel Aref Ali Al-zanam, Omer Hussein Abdou Elsayed Hussein Alhomery , Choo Peng Tan Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, Bukit Beruang, Melaka, Malaysia (2023), "**Mental Health State Classification Using Facial Emotion Recognition and Detection"** presented at International Conference on Computational Intelligence and Software Engineering**.**

[3] Sabina Umirzakova, Shabir Ahmad , Sevara Mardieva, Shakhnoza Muksimova, Taeg Keun Whangbo (2023), "**Deep learning-driven diagnosis: A multi-task approach for segmenting stroke and Bell's palsy"** Electronics and Mobile Communication Conference (IEMCON).

[4] Vandana, Nikhil Marriwala, Deepti Chaudhary (2023), "**A hybrid model for depression detection using deep learning"** presented at the 3rd international Conference on Affective Computing and Intelligent Interaction Workshops.

[5] Jing Zhang, Hang Yin, Jiayu Zhang, Gang Yang, Jing Qin and Ling He (2022), "**Real-time mental stress detection using multimodality expressions with a deep learning framework"** IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) in Calgary.

[6] Rehmat Ullah ,Hassan Hayat, Afsah Abid Siddiqui, Uzma Abid Siddiqui, Jebran Khan, Farman Ullah, Shoaib Hassan, Laiq Hasan, Waleed Albattah , Muhammad Islam, and Ghulam Mohammad Karami (2022), "**A Real-Time Framework for Human Face**

Detection and Recognition in CCTV Images" IEEE International Joint Conference on Neural Networks (IJCNN 2008) as part of the IEEE World Congress on Computational Intelligence, which took place in Hong Kong, China.

[7] D. Lu and L. Yan, 2021, "**Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor**" Journal of Sensors, vol. 2021, Article ID 4796768, 16 pages.

[8] Edeh Michael Onyema, Piyush Kumar Shukla, Surjeet Dalal, Mayuri Neeraj Mathur, Mohammed Zakariah, Basant Tiwari (2021), "**Enhancement of Patient Facial Recognition through Deep Learning Algorithm: ConvNet"** Journal of Healthcare Engineering, vol. 2021

[9] Russell Li and Zhandong Liu (2020)**, "Stress detection using deep neural networks"** The International Conference on Intelligent Biology and Medicine (ICIBM) .

[10] Pramod Bobade, Vani M (2020)**,** "**Stress Detection with Machine Learning and Deep Learning using MultiModal Physiological"** Data Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA)

Prof. P. N. Jadhav	Prof. P. N. Jadhav	Dr. A. C. Adamuthe
Sign of Project Guide	**Sign of Project In-charge**	**Sign of HOD, IT**