IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

MENTAL HEALTH DETECTION USING FACIAL EMOTION RECOGNITION

Under the guidance of
Dr. A. BHAGYALAKSHMI, M.E., Ph.D.,
PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SCHOOL OF COMPUTING

VEL TECH RANGARAJAN DR. SAGUNTHALA R&D INSTITUTE OF SCIENCE & TECHNOLOGY

ABSTRACT

The challenge addressed in this project is the development of a comprehensive facial expression recognition system that encompasses three critical components: face detection, facial feature extraction, and expression classification. This project aims to automatically and accurately identify faces in images, extract and represent subtle changes in facial features corresponding to emotional expressions, and classify these expressions, accounting for a range of emotions, including fundamental and complex ones that may vary across cultures. When developing an automatic facial expression recognition system, three primary challenges must be addressed: face detection, facial feature extraction, and expression classification. The core objective of the project is to create a reliable and efficient framework for the classification of facial expressions and to enable the system to discern and categorize these expressions with precision, encompassing a range of emotions from fundamental to complex, accounting for potential cross-cultural variations.The project operates within the domain of Emotion Recognition and Analysis through Facial Expressions. It is concerned with the development of a system that can automatically detect and classify a wide range of emotional states by analyzing facial expressions. Diverse dataset of facial images that depict a wide range of emotional states, including Sad, Anger, Fear, Joy, Disgust, Confused and Surprise is gathered. An exceptional level of accuracy, reaching 96%, has been achieved by the Convolutional Neural Networks (CNN) algorithm in accurately identifying emotions.

Keywords: Facial Expression Recognition, Face Detection, Facial Feature Extraction, Expression Classification, Diverse Datasets.

Chapter 1

INTRODUCTION

1.1 Introduction

The human face serves as a crucial means of non-verbal communication, playing a central role in social interactions. Facial expressions convey emotions, and specific facial muscles are associated with primary emotional states, as identified by Ekman. These expressions reflect an individual's emotional state and intentions. Notably, communication relies significantly on non-verbal cues, with 55% attributed to gestures and facial expressions, as cited by Albert Mehrabian in ”Silent Messages”.The identification and recognition of facial expressions usually encompass a range of primary emotions, as cited by Paul Ekman in ”An Arguement For Basic Emotions Cognition and Emotion”, such as happiness, sadness, anger, fear, surprise, and disgust. These emotions form the fundamental basis for many facial expression recognition systems Automatic recognition of facial expressions is of interest in fields like eLearning and affective computing. When designing an automatic facial expression recognition system, three key challenges are addressed: face detection, facial feature extraction, and expression classification. This involves automatically locating the face in input images, capturing facial changes induced by expressions, and inferring the specific emotions expressed. Emotions, as per emotional theorists and psychologists, encompass a range from fundamental to complex, which can vary across cultures.

1.2 Aim of the Project

The aim of this project is to develop an automatic facial expression recognition system that identifies and categorizes various emotional states, including both fundamental and culturally influenced emotions. This system will play a pivotal role in fields like eLearning and affective computing. The primary focus is on overcoming the three key challenges of face detection, facial feature extraction, and expression classification to create a robust and accurate tool for understanding human emotions and intentions.

1.3 Project Domain

The project operates within the domain of Emotion Recognition and Analysis through Facial Expressions. It is concerned with the development of a system that can automatically detect and classify a wide range of emotional states by analyzing facial expressions. This domain finds applications in various fields, such as eLearning and effective computing, where understanding and interpreting human emotions based on their facial expressions is of significant importance.

1.4 Scope of the Project

The scope of this project is to develop a robust system for facial emotion recognition using deep learning, particularly CNNs. It involves data collection, preprocessing, CNN architecture design, feature extraction, and classification to accurately identify a range of emotions. The project aims to explore cross-cultural considerations and practical applications in real-world deployment. Its primary focus is to contribute to the field of affective computing by creating a tool that can understand and respond to human emotions based on facial expressions.

1.5 Methodology

- Data Collection and Pre-processing: Diverse dataset of facial images that depict a wide range of emotional states, including Sad, Anger, Fear, Joy, Disgust, Confused, Frustrated and Surprise is gathered. These images have been carefully pre-processed to improve their quality, standardize lighting conditions, and ensure that facial features are aligned consistently for uniform input.

- Feature Extraction: Employ CNNs to automatically extract relevant features from the pre-processed facial images. CNNs excel at capturing spatial and temporal dependencies within images, crucial for understanding complex emotional expressions.

- CNN Architecture Design: A CNN architecture is crafted with layers of convolutional and pooling layers to detect patterns and reduce spatial dimensions. By effectively employing these layers and tuning the learnable parameters through training, the ConvNet becomes adept at recognizing distinctive patterns associated with various emotional expressions. Leverage learnable weights and biases to assign importance to various aspects within the images, allowing the network to distinguish between different emotional states.

- Training and Learning: The CNN is trained using the pre-processed dataset, where the network learns to differentiate and classify facial expressions into predefined emotion categories. The architecture's re-usability of weights and ability to capture spatial and temporal dependencies contribute to better fitting the image dataset.

- Classification and Emotion Identification: The trained CNN model is used to classify facial expressions and identify an individual's emotion based on primary emotional categories (Sad, Anger, Fear, Joy, Disgust, Surprise). The architecture's ability to differentiate and recognize subtle features in facial expressions enhances the accuracy of emotion identification.

- Evaluation and Testing: The CNN based facial emotion recognition system's performance using suitable metrics is evaluated, which include accuracy, precision, recall, and the F1 score. This assessment gauges the system's capability to accurately and effectively identify and distinguish among a diverse set of emotions.

- Application: The developed system is put into practical use in real-world applications, including areas like mental health support, Retail and Customer Service, Emotion-aware Gaming, Human Resource and Recruitment, Cognitive Load Monitoring in Education, human-computer interaction and enhancing user experiences and its actively refined through an iterative process, taking into account valuable user feedback and staying updated with the latest advancements in deep learning and emotion recognition research to make it even better.

Chapter 2

LITERATURE REVIEW

Jyoti Kumari.,et al,[1] discussed the Detection of mental disorders, and synthetic human expressions. The author mentioned that the two common methods used predominantly in the literature for Facial Emotion Recognition (FER) automatic systems depend on geometry and appearance. The author provides a quick scan for facial expression recognition. A comparative study was also performed using various feature extraction techniques in the Japanese Female Facial Expression (JAFFE) dataset. The limitations of Analyzing facial expressions has a major drawback humans can control the simulation to some extent, so recognition results may be falsified, intentionally or unintentionally.

G. Kalaivani et. al,[2] made use of the Viola Jones techniques and image cropping techniques for extracting and delineating areas of the mouth.The proposed segmentation techniques are applied and compared to the method found which is suitable for segmentation of the oral region, and then the oral region can be extracted by means of contrast extension and image segmentation techniques. After extracting the mouth area, the facial feelings are ranked based on the white pixel values in the mouth area extracted from the face image. The limitations are that traditional image segmentation techniques are more fragmentation and have high noise sensitivity.

Khan, R. et.al,[3] marked Emotion recognition as an important area of work for improving human-machine interaction. The complexity of the emotion makes the task of acquisition even more difficult. Deep learning technology with neural networks increased the machine's success rate with regard to emotion recognition. Modern works with deep learning technology have been implemented with different types of human behavior inputs such as auditory and visual inputs, facial expressions, body gestures Electroencephalogram (EEG), and related brainwaves and signals.The authors have attempted to explore the relevant important work, its techniques, the effectiveness of the methods and the scope for improving the results. The limitations are that Deep learning lacks common sense. This makes systems fragile and when errors do occur, the errors can be very large.

Ravichandra Ginne, et al,[4] have surveyed various FER techniques which has become an active research area that finds many applications in areas such as human and computer interfaces, human emotion analysis, psychoanalysis, medical diagnostics, etc. The common techniques used for this purpose depend on geometry and appearance. CNN have been shown to outperform traditional approaches for various visual recognition tasks including recognition of facial expressions. Despite efforts to improve the accuracy of FER systems using CNN, current methods may not be sufficient for practical applications.This study includes a general review FER of systems using CNN and their strengths and limitations which helps us to further understand and improve FER systems.The limitations are Improper encoding of object's position and orientation. Lack of the ability to be spatially unchanged for the input data, in International Journal of Advances in Electronics and Computer Science.

Gibbons RD, et al,[5] explored recent developments in computerized adaptive diagnostic screening and computerized adaptive testing for the presence and severity of mental health disorders such as depression, anxiety, and mania. The statistical methodology is unique in that it is based on the multidimensional element response theory (severity) and random forests (diagnosis) rather than the traditional mental health measurement based on the classical test theory (simple score) or the one dimensional element response theory.The limitations of the system is that a complex model requires much larger samples of people.

Shojaeilangari, et al,[6] proposed an approach called extreme scattered learning, which has the ability to co-learning a dictionary (set of rules) and a nonlinear classification model.The proposed approach combines the discriminative power of an extreme learning machine with the reconstructive characteristic of sparse representation to enable accurate classification when presented with noisy signals and incomplete data recorded in natural environments. The author has also introduced a distinctive and stable new local spatio-temporal descriptor.The proposed framework is able to achieve state-of-the-art recognition accuracy in both spontaneous and affective facial emotions databases.The limitations of the system is that a complex model requires much larger samples of people.

Shaul Hammed,et.al,[7] surveyed various Facial expression techniques as Facial expression is an important way of communicating human emotions. The author has stated that Facial expression recognition systems have four important steps like signal acquisition, Pre-processing, Extraction of features, selection of features and their classification. The various pre-processing that are required are Noise reduction using filters; face detection by localising and extracting facial region; Normalisation of colour & size of images; and enhancement of image by Histogram Equalisation. CNN among other Neural networks is the most popular among researchers in this field. The limitations is that CNN does not encode the position and orientation of the object into their predictions and Dynamic FER has a higher recognition rate than static.

Singh et al,[8] examined the acknowledgment of outward appearances with a mix of neural organization for the acknowledgment of various facial feelings. People are fit for delivering a large number of facial activities during correspondence that shift in intricacy, power, and importance. The author also conducts investigations on the restrictions with existing framework Emotion acknowledgment utilizing cerebrum movement. The author has utilized a current test system and accomplished 97% exact outcomes. Emotion acknowledgment utilizing cerebrum movement framework. The Proposed framework relies on the human face, the face additionally mirrors the human cerebrum exercises or feelings. The author uses neural organization for better outcomes.The limitations is that the recognition from mind action may bring about poor spatial goal and misdiagnosed may bring about vein, in International Journal of Image, Graphics and Signal Processing.

Zhang et al,[9] discussed various Emotion recognition techniques using 3D Gabor features. These techniques used patch based 3D Gabor features, Classification, Adaboost, Support Vector Machine (SVM). The databases used JAFFE with 213 images and 7 expressions, C-K database-Video based code action.486 sequences with 96 posers, 593 video sequences on both posed and non-posed (spontaneous) emotions and 123 subjects from 18 to 30 years in age. the techniques provide protocols and baseline results for facial feature tracking, action units, and emotion recognition. The Image resolutions of 640 × 480, and 640 × 490 are considered. The Recognition Rate achieved is 94.48%.

Rannchand Hablani,[10] discussed Emotion recognition techniques using Template Matching. The features used are Local Binary pattern and ClassificationTemplate matching. The Databases used are JAFFE 213 images with 7expressions,

213 images of seven facial emotions. The author has considered ten different female Japanese models, Six emotion adjectives by 60 Japanese subjects. The Image resolution of 256 × 256 is considered. The Recognition rate for Person dependent is 94.44% and Person Independent is 73.61%.

Elzbieta kukla, et al[11] introduced a method that uses a series of neural networks to recognize facial expressions.As an input, the algorithm receives a natural image of the face and returns the emotion expressed by the face. To determine the best classifiers for recognizing specific emotions, single- and multiple-layered networks were tested. The experiments covered different resolutions of the images displaying the faces as well as the images, including the areas of the mouths and eyes.On the basis of the results of the tests, a series of neural networks are proposed. The series introduces six basic emotions and a neutral expression.The Limitations are Black box, development period, amount of data and calculation cost.

Raut et al,[12] displayed Facial Emotion Recognition Using Machine Learning. The author in this research stated that the subtle emotions in Eulerian Motion Magnification (EMM) are difficult to detect. Movement characteristics such as speed and acceleration can be used to zoom in.The image is transformed as a whole by enlarging changes in the amplitude and phase properties. Depending on the characteristics, there are A-EMM (capacitive based) and P-EMM (phase based) motion amplification. Oriented FAST and Rotated BRIEF (ORB) and Speeded-Up Robust Features (SURF), Scale-Invariant Feature Transform (SIFT) are used and also feature descriptor algorithms used.The dataset used in this experiment was the iBug-300W dataset containing over 7,000 images as well as the CK + dataset containing 593 facial expression sequences from 123 different subjects. The limitation is that if the motions are large, this manipulation can introduce artefacts,in Master's Projects.

Chapter 3
PROJECT DESCRIPTION

3.1 Existing System

Existing systems for facial emotion recognition often use various techniques, including deep learning and traditional machine learning methods. They typically involve collecting facial data, extracting features, training models, and then using these models to recognize emotions in real-time or from static images. Some methods heavily rely on detecting facial landmarks and extracting geometric features, such as distances between facial points, angles, and ratios, to recognize emotions. Facial Action Coding System (FACS) involves manual coding of facial movements and expressions based on anatomically defined facial actions. While it's a detailed and comprehensive system, it requires expert annotation and might be labor-intensive.

Disadvantage: The performance of facial emotion recognition systems often heavily relies on the quality and diversity of the training data. Limited or biased training data can result in less effective models, and collecting diverse and balanced datasets can be challenging.

3.2 Proposed System

This project is focused on improving the accuracy and reliability of facial emotion recognition by addressing the challenge of controlling and falsifying facial expressions by humans. The key differentiators in this project include:

Quick Scan Approach: This project introduces a quick scan approach, which represents an efficient and rapid method for recognizing facial expressions. This approach could be particularly valuable for real-time applications where speed is essential.

Emotion Control Consideration: This project acknowledges the limitation of humans being able to control their facial expressions to some extent. By recognizing and addressing this limitation, the system may provide more accurate and trustworthy results in situations where subjects might intentionally or unintentionally mask their true emotions.

Comparative Study: The project also includes a comparative study using various feature extraction techniques on the (Japanese Female Facial Expression dataset) JAFFE dataset, which means that the model is actively evaluating and improving the performance of the system, which is a valuable aspect of research and development in this field.

3.3 Feasibility Study

3.3.1 Economic Feasibility

The project would require an initial investment in technology and data acquisition. However, the potential benefits in terms of mental health support and overall wellbeing could justify the costs.

3.3.2 Technical Feasibility

Given the advancements in AI and facial recognition technology, the project is technically feasible. The use of CNNs and deep learning techniques can facilitate accurate emotion recognition.

3.3.3 Social Feasibility

The project aligns with the growing societal awareness of mental health issues and the need for innovative solutions.

3.4 System Specification

3.4.1 Hardware Specification

- Processor - i5
- Speed - 3 GHz
- RAM - 8 GB(min)
- Hard Disk - 500 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

3.4.2 Software Specification

- Operating System - Linux, Windows7/10
- Tools - Anaconda, Jupyter, pycharm
- Server side Script - Python.

3.4.3 Standards and Policies

- General Data Protection Regulation (GDPR): Compliance with GDPR or equivalent data protection regulations is crucial when handling personal data. Ensure informed consent and secure storage of user data.

- Health Insurance Portability and Accountability Act (HIPAA): If the project involves health-related data, such as medical records, HIPAA compliance is essential.

- Ethical Considerations: Adhere to ethical AI frameworks and guidelines, such as those proposed by organizations like the IEEE or ACM, to ensure responsible and unbiased AI development.

- Informed Consent: Obtain explicit and informed consent from individuals whose data is used in the project.

- Mental Health Support and Counseling: National and International Mental Health Guidelines: Consult and align with established guidelines and best practices in mental health support and counseling, as recommended by organizations like the World Health Organization (WHO) and national health agencies.

- Clinical Validation: If the project involves clinical trials or validation with mental health professionals, adhere to recognized clinical trial standards.

- AI and Facial Recognition Regulations: Local AI and Data Privacy Laws: Stay informed about regional laws and regulations related to AI, facial recognition, and data privacy.

- Research and Academic Standards: Ethics Review Boards: If the project is associated with an academic institution, obtain approvals from ethics review boards for research involving human subjects.

- User Data Handling Policies: Develop and communicate clear policies regarding how user data is collected, stored, and used. Users should be informed about their datas purpose and retention.

- Transparency and Accountability: Consider publishing transparency reports that detail the system's functioning and accuracy to build trust with users and the public.

- Safety and Crisis Support:Include mechanisms to support users in crisis situations and guide them to professional help when necessary.

- ISO Standards: ISO 27001: Information Security Management System (ISMS) standards for ensuring the security of information. ISO 27701: Privacy Information Management System (PIMS) standards for managing personal information.

Chapter 4
METHODOLOGY

4.1 Architecture Diagram

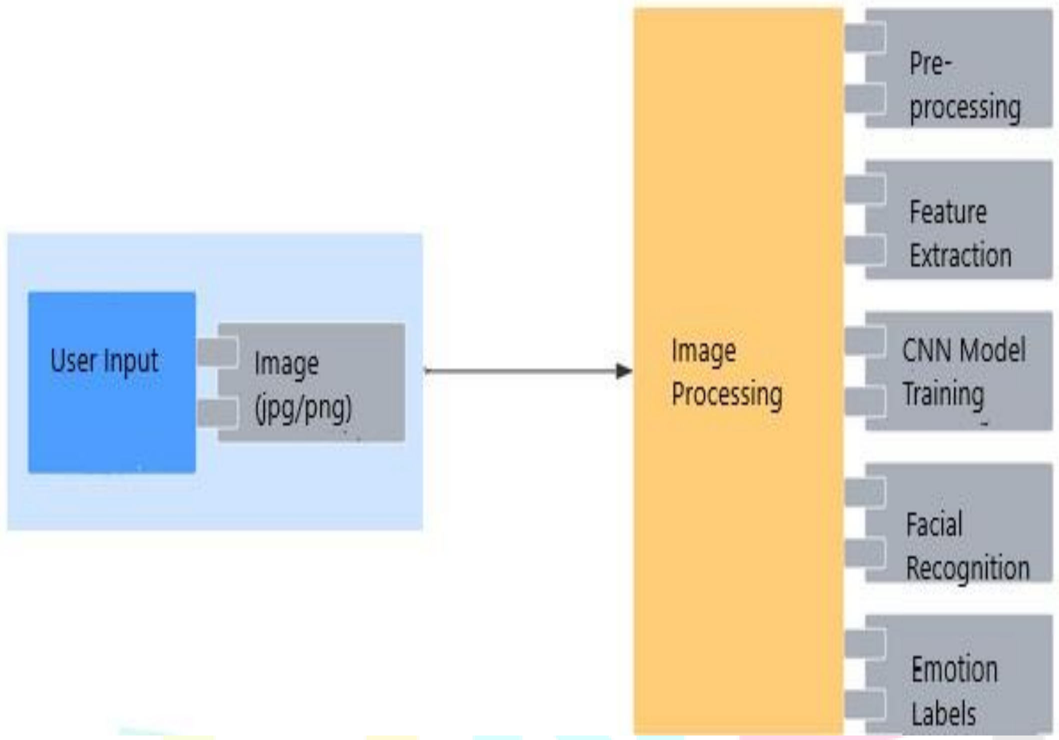

Figure 4.1: Architecture Diagram

In the above Figure 4.1, initially, the user provides an input in the form of an image, image pre-processing is applied to eliminate any visual defects or unwanted artifacts. Subsequently, feature extraction techniques are employed to capture relevant information from the image. Following this, a CNN model is trained using a dataset of emotions, where the model learns to recognize emotional cues in the images. Once the model is trained and tested to minimize errors, it is then employed to process the input image. Finally, the system provides an output that signifies the detected emotion, effectively translating visual data into emotional insights.

4.2 Design Phase

4.2.1 Data Flow Diagram

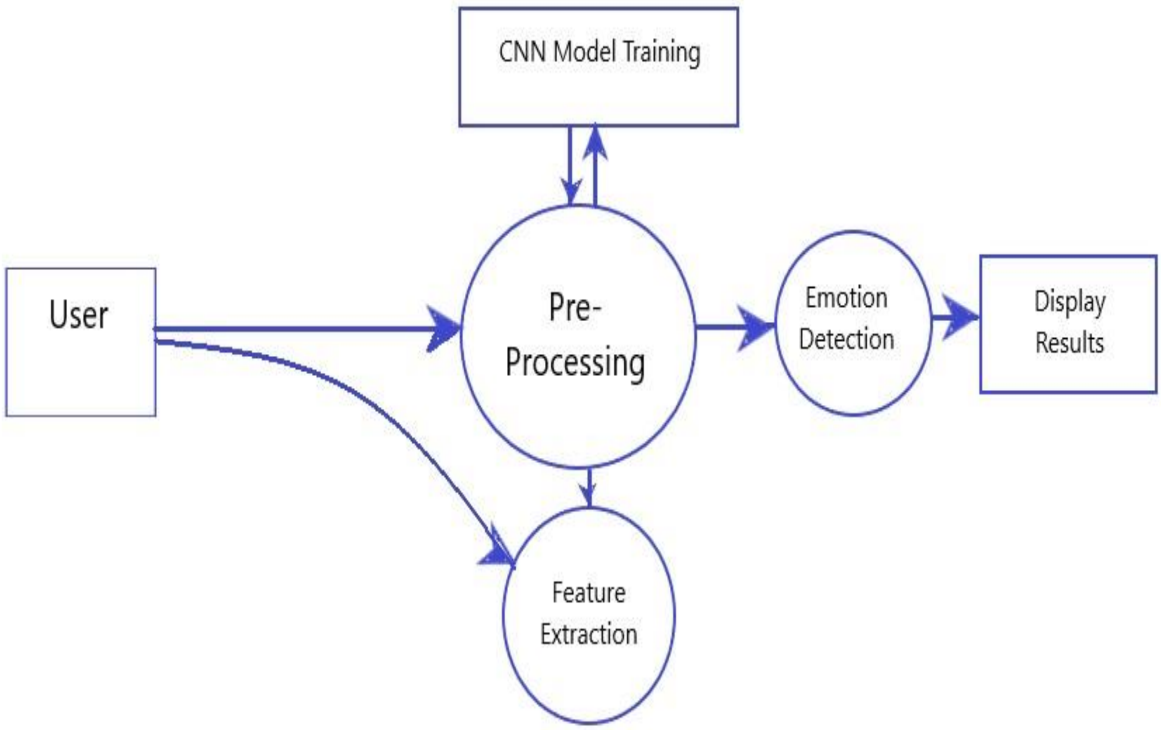

Figure 4.2: Data Flow Diagram

The Fig 4.2, begins with an Test Image given by the user as its primary data source and this image is directed to the Image Pre-processing task, where any visual defects or unwanted artifacts are removed. After pre-processing, the data flows into the Feature Extraction activity, which captures relevant information from the image. The output from feature extraction is then directed to the CNN Model Training process. Here, a Dataset of Emotions is used to train the CNN Model. The model learns to recognize emotional cues in the images provided in the dataset. Once the model is trained and tested to minimize errors, it moves to the Emotion Analysis phase. In Emotion Classification, the trained CNN model is employed to process the input test image. The analysis results in the Detected Emotion data, which signifies the emotion detected in the input image. Finally, the Detected Emotion is the output of the system, providing a valuable insight into the emotions conveyed by the input image.

4.2.2 Use Case Diagram

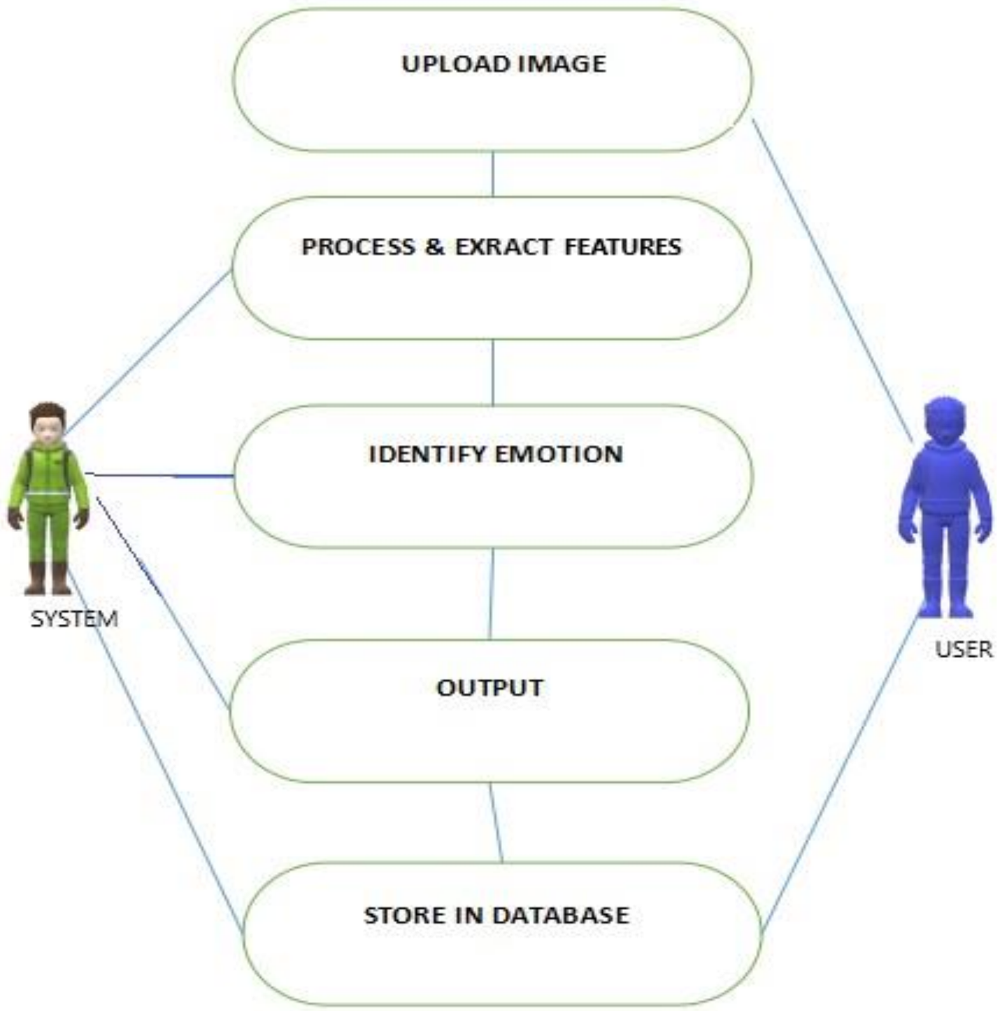

Figure 4.3: Use Case Diagram

In the above Fig 4.3, the primary scenario involves a user uploading an image into the system. The system then undertakes a series of actions, including image processing, feature extraction, and emotion identification. Once these processes are complete, the system generates an output, which is subsequently presented to the user for viewing. This use case illustrates the user's interaction with the system, demonstrating its ability to analyze and convey emotions based on the uploaded image.

4.2.3 Class Diagram

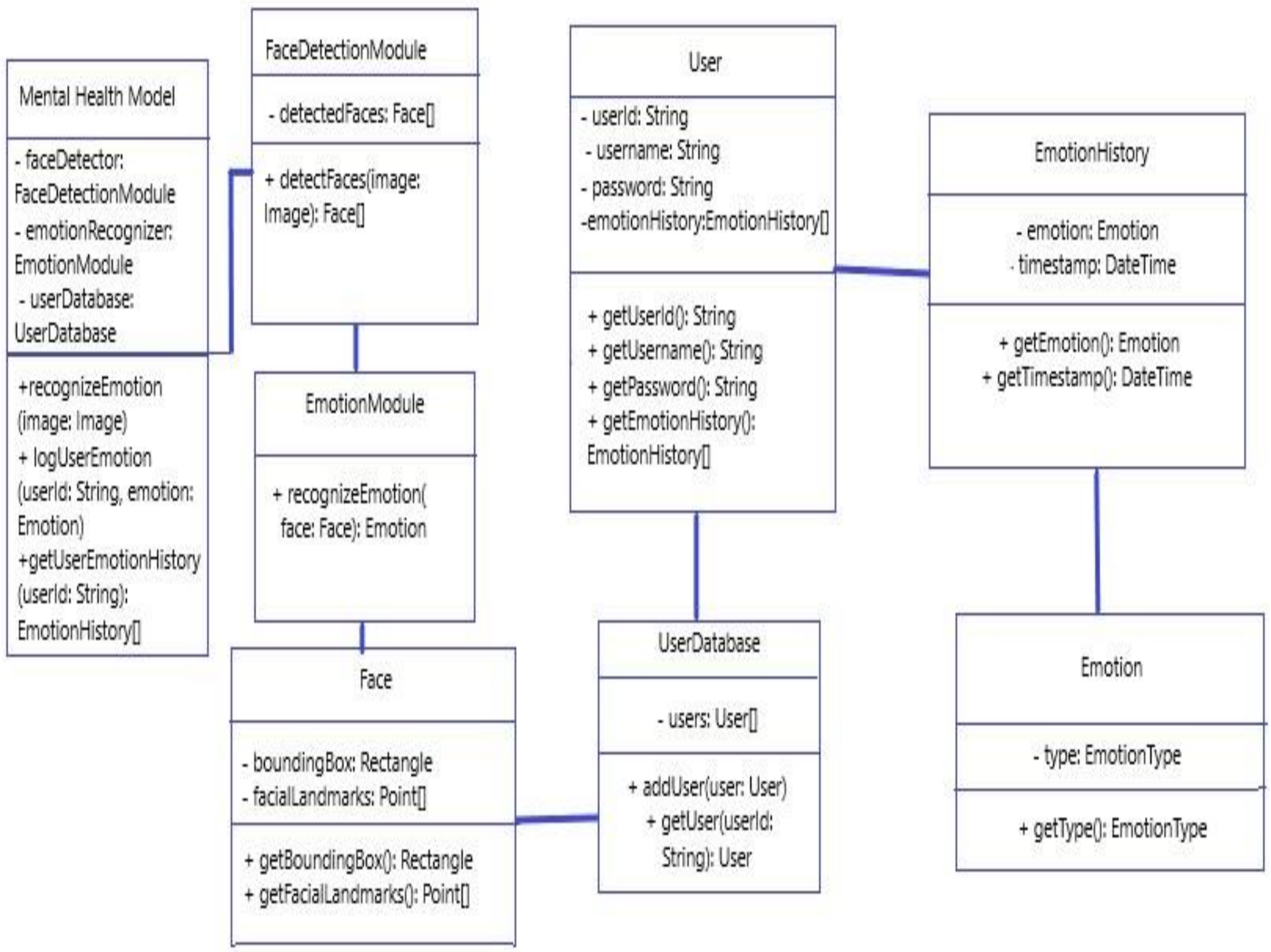

Figure 4.4: Class Diagram

In the Fig 4.4, The Mental Health System Represents the core class orchestrating the system & Manages interaction with Face Detection Module, Emotion Module, and User Database. The Face Detection Module Detects faces in an image using the detect Faces method & returns a list of detected faces, each represented by the Face class. The Emotion Module recognizes emotions in a detected face using the recognize Emotion method & returns the recognized emotion, represented by the Emotion class. The Face represents a detected face with attributes such as bounding box and facial landmarks. The Emotion represents a specific emotion detected in a facial expression & provides methods to retrieve the type of emotion. The Emotion Type is an enumeration class enumerating different types of emotions (e.g., happy, sad, angry).

4.2.4 Sequence Diagram

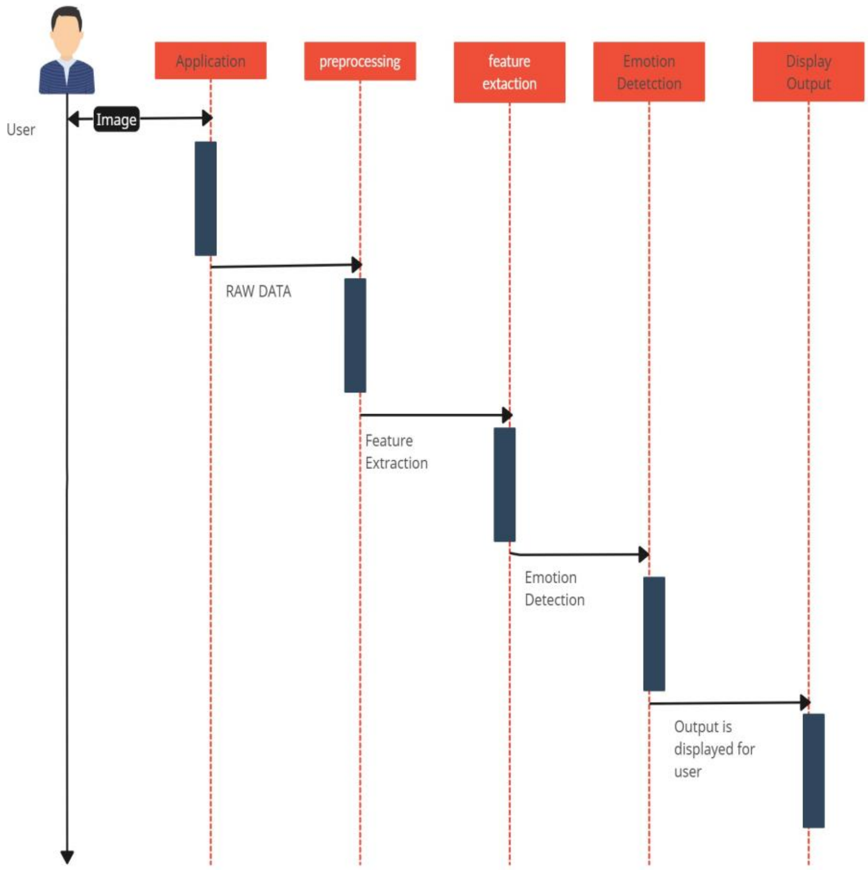

Figure 4.5: Sequence Diagram

In above Figure 4.5 ,the application begins when the user uploads an image featuring a human face. The system promptly employs facial recognition to detect the presence of a face in the image. Following this, the device conducts image pre-processing to refine the image, extracting critical facial features. Finally, the output, which encompasses both the pre-processed image and the extracted features, is securely stored in a database for future use. This process seamlessly combines user interaction, facial detection, image enhancement, feature extraction, and database management.

4.3 Module Description

4.3.1 Convolutional Neural Network Training

Critical points of the Face are very important and can be used for facial recognition and detection. Here a total of 68 facial critical points are represented to the discoverer in the dlib.PyPI (a set of tools for developing machine learning and data analysis applications in the real world.)

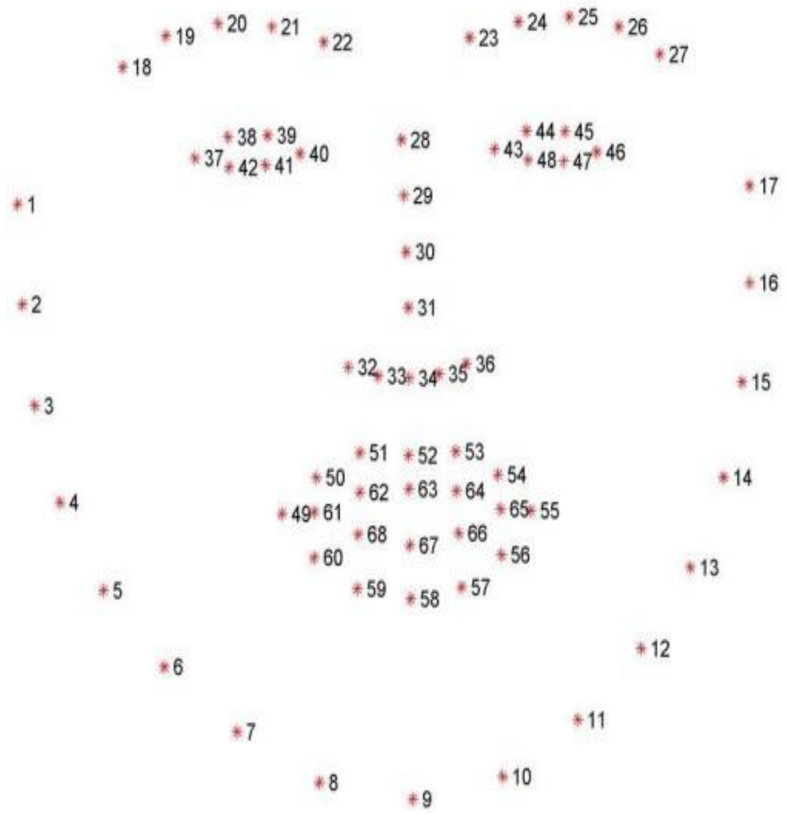

Figure 4.6: Critical Points Representation

All these 68 critical point on the face are depicted in the above figure. The (x,y) coordinates of every facial point can be retrieved by using dlib.PyPI tools. All these points can be broken down into categories such as face, nose, eyebrow and jaw. The proposed method is based on a two-level CNN framework. The first level recommended is background removal, used to extract emotions from an image, as shown in the figure below.

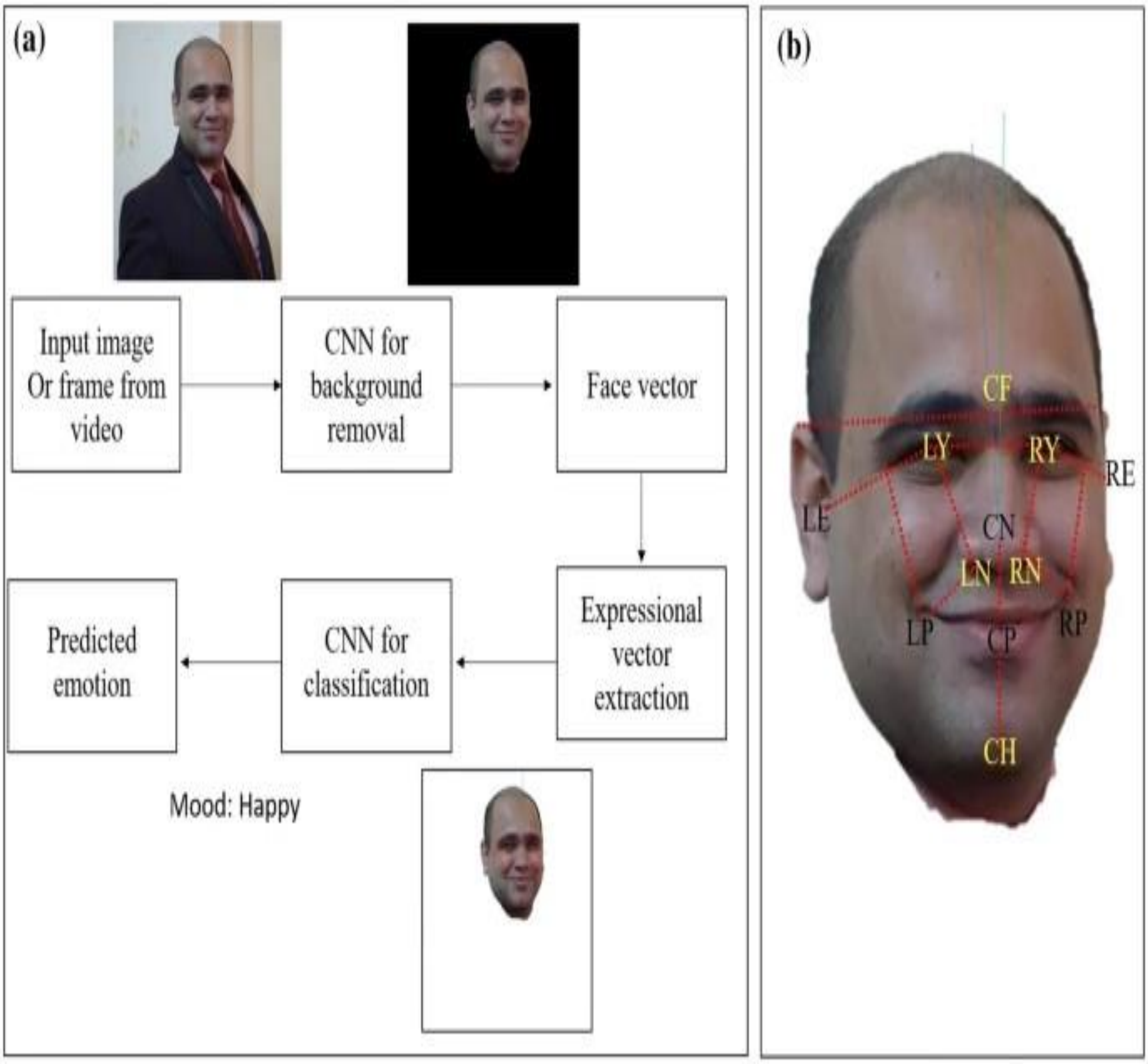

Figure 4.7: Convolutional Layers

Here, the conventional CNN network module is used to extract primary Expressional Vector (EV). The EV is generated by tracking down relevant facial points of importance. EV is directly related to changes in expression. The EV is obtained using a basic perceptron unit applied on a background-removed face image. In the proposed FERC model, we also have a non-convolutional perceptron layer as the last stage. Each of the convolutional layers receives the input data (or image), transforms it, and then outputs it to the next level. This transformation is convolution operation, as shown in Fig. 4.8.

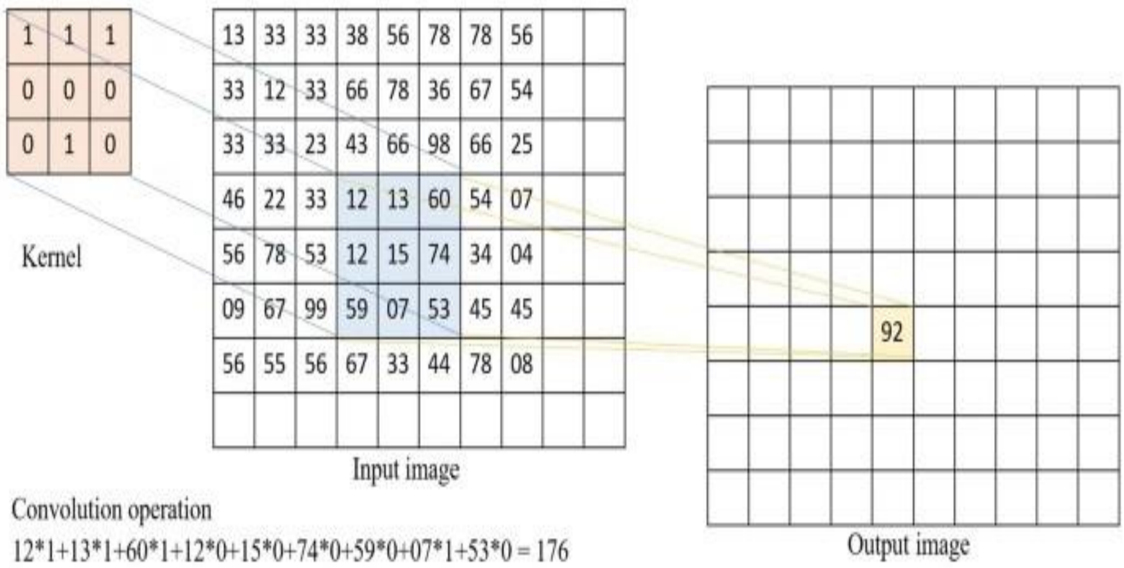

Figure 4.8: Matrix Representation of CNN Layers

All the convolutional layers used are capable of pattern detection. Within each convolutional layer, four filters were used. The input image fed to the first-part CNN (used for background removal) generally consists of shapes, edges, textures, and objects along with the face. The edge detector, circle detector, and corner detector filters are used at the start of the convolutional layer 1. Once the face has been detected, the second-part CNN filter catches facial features, such as eyes, ears, lips, nose, and cheeks. The edge detection filters used in this layer are shown in Fig.4.9.

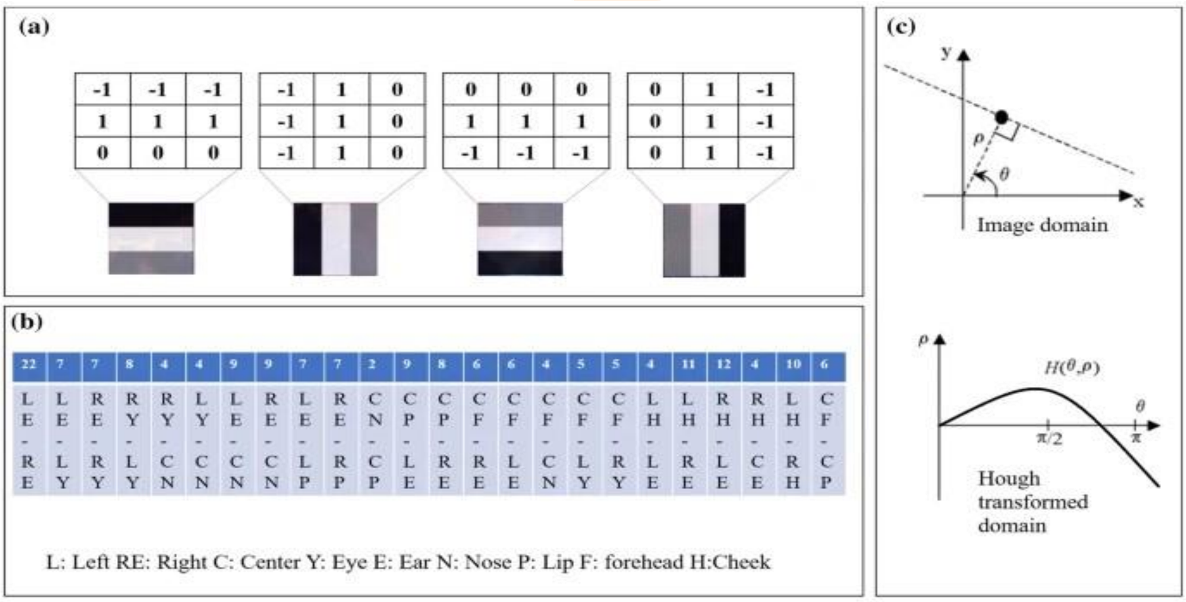

Figure 4.9: Vertical and Horizontal Edge Detector Filter Matrix

The second-part CNN consists of layers with kernel matrix, e.g., [0.25, 0.17, 0.9; 0.89, 0.36, 0.63; 0.7, 0.24, 0.82]. These numbers are selected between 0 and 1 initially. These numbers are optimized for EV detection, based on the ground truth we had, in the supervisory training dataset. Here, we used minimum error decoding to optimize filter values. Once the filter is tuned by supervisory learning, it is then applied to the background-removed face (i.e., on the output image of the first-part CNN), for detection of different facial parts (e.g., eye, lips. nose, ears, etc.)

4.3.2 Deep Neural Network Employment

Deep neural network with transfer learning approach is employed to extract bottle features from the input images and save these features. At a later stage, a network of fully connected layers is used where the bottle features are loaded back and images are then passed to the model for prediction.

4.3.3 Model Training

As part of pre-processing, dimension Fixing (rescaling) is performed as well as to increase the data volume for the given model to train on, image augmentation is employed on the images before they are fed to the neural network for prediction.

4.3.4 Image Dataset & Feature Training

Deep neural network with transfer learning is employed to extract the bottleneck features from JAFFE images dataset and fed these features to a set of fully connected layers to predict the facial emotions of these images. Finally, complete content and organizational editing is done before formatting.

4.3.5 Bottle Neck Feature Extraction

Transfer learning is used as follows, Reused VGG16 Model, Loaded with Image net weights. Extract bottleneck feature from it for the testing scenario. Tuned the model with bottleneck weights to achieve higher accuracy.

4.3.6 Performance Analysis

Initially, the Cohn–Kanade dataset with 486 sequences yielded a maximum accuracy of 45%. To improve efficiency, additional datasets were gathered, including internetsourced data and the users own pictures. The algorithm’s accuracy increased with the number of images in the dataset, reaching 96% with a 10,000-image dataset. The study involved 25 iterations, optimizing the number of layers and filters for a CNN. The optimal configuration was found to be four layers and four filters. The algorithm successfully detected emotions in front-facing images but faced challenges with grayscale images and orientation. Despite limitations such as high computing power requirements during tuning and issues with facial hair, the algorithm's accuracy compared favorably with other studies. However, limitations were noted in cases of missing facial features and multiple faces in an image. The algorithm's performance was further tested on Caltech faces, CMU, and NIST databases, with accuracy decreasing with an increasing number of images due to overfitting. The ideal number of images for proper functioning was determined to be in the range of 2000–10,000.

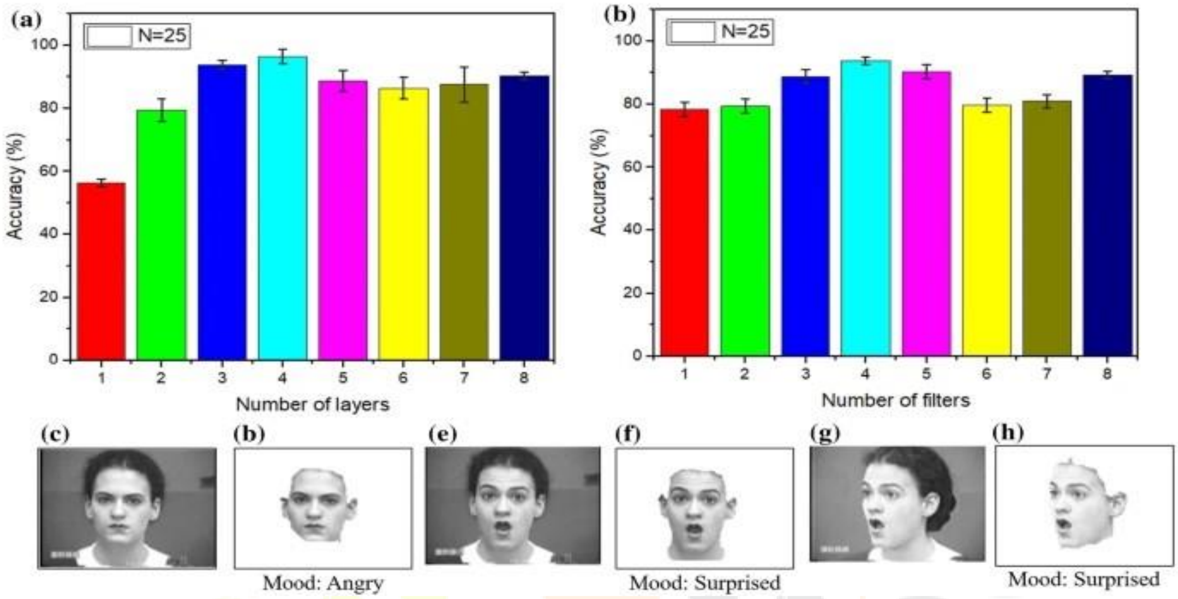

Figure 4.10: Accuracy Levels Of Different Emotions

4.3.7 Cross-Validation

To ensure the generalization ability of the model, cross-validation techniques are applied. This module discusses the strategies employed, such as k-fold cross-validation, to validate the model's performance across different subsets of the dataset.

Chapter 5
IMPLEMENTATION AND TESTING

5.0.1 Image Insertion & Error Testing

The procedure level testing is made first by giving improper inputs, the errors occurred are noted and eliminated. This is the final step in system life cycle. Here the tested error-free system is implemented into real-life environment and make necessary changes, which runs in an online fashion. The system maintenance is done every months or year based on company policies, and is checked for errors like runtime errors, long run errors and other maintanances like table verification and reports.

Image Format:One form of output is a processed image that highlights the identified facial expressions. This image may include graphical overlays or annotations, making it easier for users to interpret the results.

5.0.2 Feature Extraction & Emotion Identification

- Command Prompt Output:In addition to the visual output, the software generates textual feedback and commands in the command prompt or terminal. These textual outputs can be useful for users who prefer programmatic interaction.

- User Interaction: Users have the flexibility to interact with the software in real-time as the program processes the input data. They can observe the visual outputs and simultaneously receive textual feedback.

- File Export and Import: Users might have the option to export data or results to various file formats, such as CSV, Excel, PDF, or other specialized file types. This output can be saved locally or shared for external use.

- Error Handling and Notifications:It provides informative messages in the command prompt or terminal, alerting users in case of errors, warnings, or issues during the processing.

5.1 Unit testing

Image Processing Module Testing:

- Input Validation: Testing the module to correctly handles various image formats, sizes, and color spaces.
- Face Detection: Verification that the face detection component accurately locates faces within the input image.
- Feature Extraction: Testing feature extraction process to ensure it captures relevant facial features, such as eyes, mouth, and expressions.

Machine Learning Model Testing:

- Training Data: Verify that the model is correctly trained using labeled datasets.
- Prediction Accuracy: Unit tests should assess the model's accuracy in recognizing emotions based on sample inputs.
- Boundary Testing: Include test cases with a variety of facial expressions, including extreme and subtle emotions.
- Error Handling: Unit tests should include scenarios that intentionally cause errors to ensure the software handles exceptions appropriately.

5.2 Integration testing

- Image Processing and Feature Extraction: This integration point involves testing the proper flow of data from image capture to the feature extraction module. It checks if facial features and emotions are correctly identified from images.

- Real-time Processing: For live camera feeds, this tests the integration between the camera input, real-time processing, and the output visualization. It ensures that the software can handle continuous data streams.

- Graphical User Interface (GUI): Integration testing ensures that the GUI components (if applicable) communicate effectively with the back-end components. This includes displaying results and accepting user inputs.

- Visual Studio Environment: Testing includes the interaction between the software and the Visual Studio IDE, as command prompts and other development tools are used for feedback and control.

5.3 Functional testing

- Emotion Recognition Accuracy: This is a critical aspect of functional testing. Test cases are designed to verify that the software correctly recognizes and classifies emotions based on facial expressions. The system should accurately identify primary emotions such as happiness, sadness, anger, fear, disgust, and surprise, as well as more complex emotions.

- Image Processing: Functional testing ensures that image processing is functioning correctly. This includes checking if the software can locate and identify faces in images, extract relevant features, and apply pre-processing steps effectively.

- Real-time Recognition: If the project involves real-time processing of live camera feeds, functional testing checks if the software can continuously process incoming frames and provide real-time feedback on facial expressions.

- User Interface Testing: Functional testing evaluates the graphical user interface (GUI) components if applicable. It verifies that users can interact with the software as expected, input images, receive results, and navigate the user interface without issues.

- Input Validation: Testing ensures that the software validates input correctly. Invalid or corrupt image data should be handled gracefully, and the software should provide appropriate error messages or responses

.

5.4 White Box Testing

- Code Review: Performing a thorough review of the source code that makes up the different components of the project including examining the code for the image processing, feature extraction, machine learning models, and any other custom algorithms. The objective is to ensure that the code follows coding standards, is well-documented, and adheres to best practices.

- Performance Testing: Assessing the efficiency of the code. Verify that the algorithms and models run within acceptable time frames and do not lead to performance bottlenecks. This may involve profiling tools to identify areas of optimization.

- Code Duplication: Checking for duplicated code, which can lead to maintenance problems. Ensure that code is modular and reusable when necessary.

5.5 Black Box Testing

Compatibility Testing: The software is assessed if it is compatible with different operating systems, web browsers, and hardware devices.

Performance Testing: Assessing the performance of the software by evaluating how it handles a range of image inputs. Testing the system with a variety of facial expressions, lighting conditions, and image quality to ensure robust performance.

Scalability Testing: Examining how the software behaves when the number of users or the volume of input data increases. This is important if the user is planning to deploy the software in a real-world scenario.

Chapter 6
RESULTS & DISCUSSIONS

6.0.1 Efficiency of the Proposed System

In summary, the proposed model exhibits a high proficiency level by addressing the nuances of mental health recognition through diverse training data, real-time prediction capabilities, and an integrated system for practical recommendations. Its potential application in employee monitoring and client recommendations positions it as a valuable tool for promoting mental well-being in both individual and organizational contexts.

6.0.2 Comparison of Existing and Proposed System

- Training Data Diversity:In previous versions, they may have been trained on limited datasets with less diversity in terms of angles, lighting conditions, and colors but in the proposed model it is specifically trained on a more diverse dataset, capturing a wider range of real-world conditions.

- Real-time Prediction Capability: In previous versions, they might not have included real-time prediction capabilities, limiting their applicability to static images but in the proposed model it is designed to work on video clippings, demonstrating an enhanced ability to predict facial expressions in real-time, catering to dynamic and evolving scenarios.

- Integrated System and Recommendations: In previous versions, they may not have included an integrated system for displaying predictions and recommendations but in the proposed model it integrates with a system that not only displays predictions but also provides recommendations, offering a more actionable and user-friendly interface.

6.1 Sample Code

```
importtensorflow as t f fromtensorflow . keras . models import load\ _model fromtensorflow .
keras . preprocessing . image import img\ to \ array import cv2 importnumpy as np face \ _ c l a s s
i f i e r = cv2 . CascadeClassifier ( ' . / Harcascade / haarcascade \ _
_ f r o n t a l f a c e \_ d e f a u l t . xml ' )
c l a s s i f i e r =load\ _model ( ' . / Models / model\ v\ 47 . hdf5 ' )
c l a s s \ _l a b e l s ={0:   'Angry ' ,   1:   ' Disgust ' ,   2:   ' Fear ' ,   3:   'Happy ' ,   4:   ' Neutral ' ,   5:   'Sad ' ,   6:   ' Surprise
'}\\ cap = cv2 . VideoCapture (0) while True : ret , img = cap . read ( ) gray=cv2
. cvtColor ( img , cv2 .COLOR\ BGR2GRAY) faces = face \ _ c l a s s i f i e r .
detectMultiScale ( gray , 1.3 , 5) a l l f a c e s = [ ] r e c t s = [ ] for (x , y , w, h ) in
faces :
cv2 . rectangle ( img ,   (x ,  y ) ,  ( x+w,  y+h ) ,  (255 ,  0 ,  0) ,  2)
_ roi \gray = gray [ y : y+h , x : x+w] roi \ gray = cv2 . r e s i z e ( roi \ _gray , (48 , 48) , i n t e r p o l a t i
_
o n =cv2 . INTER\ AREA) a l l f a c e s . append ( roi \ _gray )
r e c t s . append ( ( x , w,   y ,  h ) )
i  = 0
for  face  in  a l l f a c e s :
roi  =  face . astype ( " f l o a t " )  /  255.0
roi  = img\ to \ array ( roi )
roi  = np . expand\ _dims ( roi ,   axis =0)
preds = c l a s s i f i e r . p r e d i c t ( roi ) [0] l a b e l = c l a s s \ _l a b e l s [ preds
. argmax ( ) ] l a b e l \ _ p o s i t i o n = ( r e c t s [ i ] [ 0 ] + i n t ( ( r e c t s [ i ] [ 1 ]
/ 2 ) ) ,
abs ( r e c t s [ i ] [ 2 ]  - 10) )
i  = + 1
cv2 . putText ( img ,   label ,   l a b e l_\ position ,
cv2 .FONT\ HERSHEY\ SIMPLEX ,  1 ,  (0 ,  255 ,  0) ,  2)
cv2 . imshow ( "MENTAL HEALTH IDENTIFICATION" , img )
i f  cv2 . waitKey (1)  ==  13:  # 13  i s  the  Enter  Key
break cap . r e l e a s e ( ) cv2 .
destroyAllWindows ( )
```

6.2 Results Phase

6.2.1 Working of Emotion Detection Model from User POV

Reference Image 1 Reference Image 2

Figure 6.1: Working of Emotion Detection Model from User POV

The Fig 6.1, involves the initiation of the facial emotion recognition system, an input image capturing the visage of a test subject is obtained, ensuring clarity and absence of distortions in their facial expression. This input image then undergoes a process designed to detect and recognize the nuances of the displayed emotions. Further undergoing pre-processing steps, meticulously adjusting the image for optimal quality, involving resizing, normalization, and noise reduction. The facial feature detection algorithms come into play, meticulously identifying key facial landmarks and regions of interest that contribute to the unique expression being conveyed. The core of the system lies in deep neural network, enriched with transfer learning capabilities, previously trained on comprehensive datasets like JAFFE. This neural network extracts bottleneck features from the input image. Employing a pre-trained model enhances the model’s ability to recognize intricate facial expressions. The final result, elegantly displayed on the output interface, provides viewers with a nuanced understanding of the detected emotion.

6.2.2 Emotion Recognition & Classification by System Model

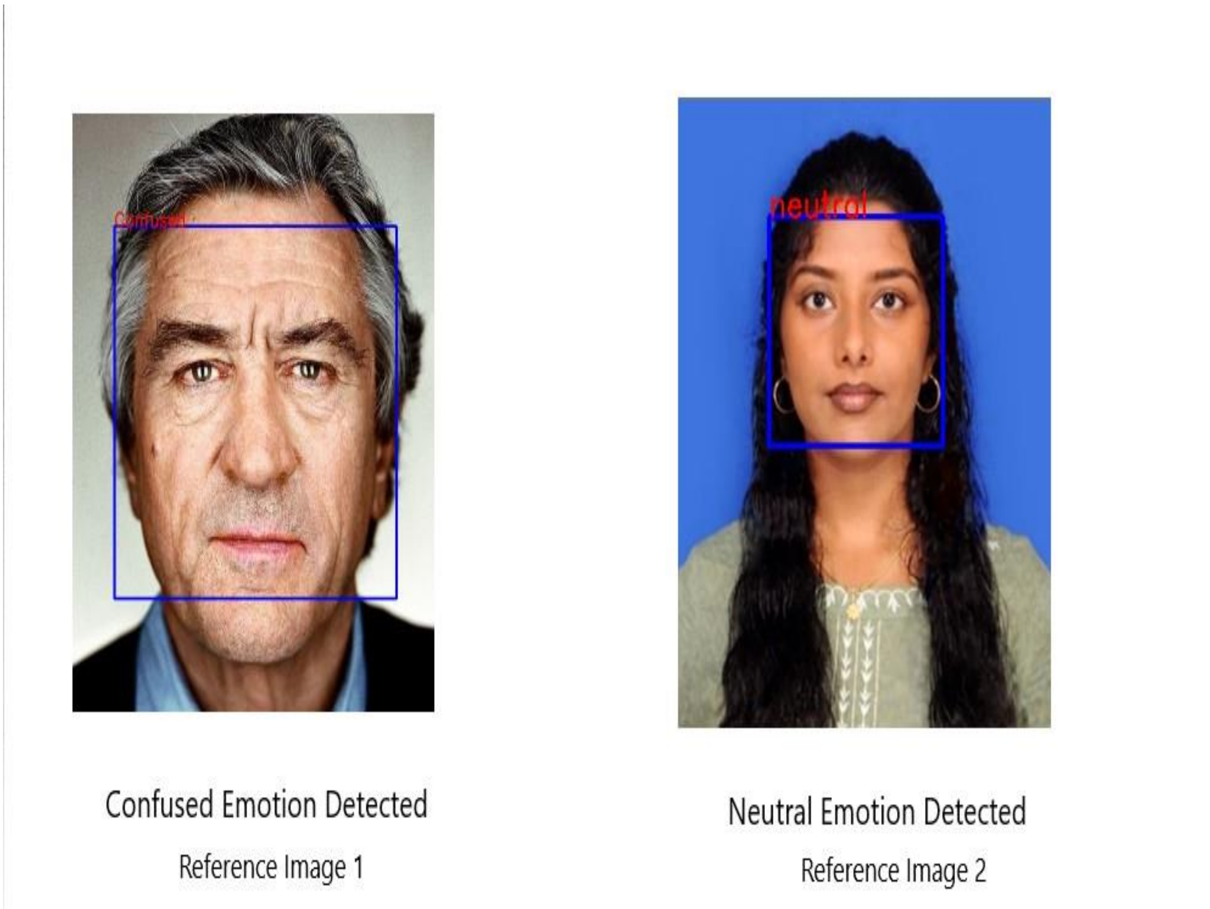

Figure 6.2: Emotion Recognition & Classification by System Model

The Figure 6.2 displays the emotion detected which provides a comprehensive visual representation, where the detected emotion is superimposed on the original input image. In this specific output, a detailed analysis of the subject's facial expression reveals that the predominant emotion is one of confusion. This conclusion is substantiated by the subject's facial features, including furrowed brows, a tightly pressed mouth, and a gaze suggesting uncertainty, all indicative of the underlying emotional state of confusion.

6.3 Analysis of Result Accuracy

6.3.1 Comparison of anticipated values Vs Real values

Predicted / Actual	Fury	Disdain	Disgust	Anxiety	Joy	Melancholy	Unexpected	All
Fury	**8**	0	3	0	0	0	0	**11**
Disdain	1	**3**	0	0	0	0	0	**4**
Disgust	4	0	**11**	0	0	1	1	**17**
Anxiety	0	0	0	**5**	0	0	0	**5**
Joy	2	0	0	0	**14**	0	0	**16**
Melancholy	2	0	0	0	0	**8**	1	**11**
Unexpected	0	1	2	2	0	2	**11**	**18**
All	**17**	**4**	**16**	**7**	**14**	**11**	**13**	**82**

Figure 6.3: Comparison of anticipated values Vs Real values

The Fig 6.3, shows The actual values and predicted values are displayed in a report format. This allows the user to determine how many emotions are appropriate.

1: Fury – 8/11 – 72%

2: Disdain – 3/4 - 75%

3: Disgust – 11/17 - 65%

4: Anxiety – 5/5 - 100%

5: Joy – 14/16 - 88%

6: Melancholy – 8/11 - 73%

7: Unexpected – 11/18 – 61%

It's important to note that while certain emotions achieved high accuracy rates (e.g., Surprise and Sadness), others, such as Happiness, had a lower accuracy rate. Additionally, it is mentioned that numerous samples were incorrectly assigned to other classes, including fear, surprise, anger, and sadness. However, fear was accurately detected in every sample. In summary, the evaluation results provide insights into the system's performance for different emotions, highlighting areas of high accuracy as well as areas that may require improvement, especially for emotions with lower accuracy rates. The information presented in the report allows users to assess the effectiveness of the facial emotion recognition system across various emotional categories.

6.3.2 Bar Graph representing quantity of samples in each class

Feelings	Number of pictures that capture the feeling
1: Fury	43
2: Disdain	17
3: Disgust	57
4: Anxiety	23
5: Joy	67
6: Melancholy	26
7: Unexpected	81

Figure 6.4: Images Trained for the CNN Model

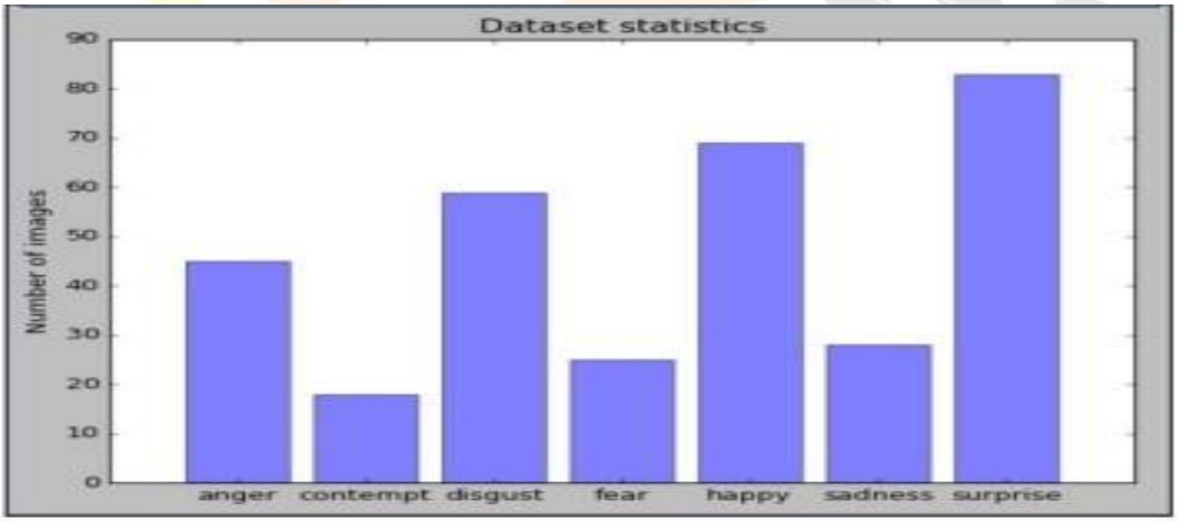

Figure 6.5: Bar Graph representing quantity of samples in each class

In summary, this bar chart provides an overview of the distribution of various emotions in the dataset and the quantity of datasets used to detect an emotion.

Chapter 7

CONCLUSION AND FUTURE ENHANCEMENTS

7.1 Conclusion

Mental health problems are a few things which if gone unnoticed, will not only lead to personal downfall but are also directly linked with the efficiency and energy that a person can give to his/her work. To make this model more effective, the model is trained with images taken on varied conditions like angles, light conditions, colors, etc. And it recognizes the faces, which will enhance the performance of the model so that in the future it can work on video clippings which will predict expressions in real time and prepare an integrated system such that the predictions made by the model would be displayed on a screen with best recommendations. The goal is to design an end system such that when this system captures images of the employees over a predefined period of time then based on the classifications the system could propose some exposure to the clients who are proven to have any critical condition.

7.2 Future Enhancements

- Multi-modal Data Integration:

Incorporate additional modalities of data, such as voice tone analysis or physiological signals, to create a more comprehensive understanding of an individual's emotional state. This could enhance the accuracy and depth of mental health assessments.

- Emotion Dynamics Analysis:

Expand the model to analyze the dynamic changes in facial expressions over time. This would involve considering the temporal aspects of emotions, providing a more nuanced understanding of how emotions evolve and manifest.

- Cultural Sensitivity and Bias Mitigation:

  Address cultural variations in facial expressions by training the model on diverse datasets that encompass a wide range of cultural backgrounds. Additionally, implement techniques to identify and mitigate biases in the model to ensure fair and accurate assessments across different demographic groups.

- Real-time Feedback and Intervention:

  Develop capabilities for real-time feedback and intervention, where the system can provide immediate support or resources to individuals displaying signs of distress. This could include integration with mental health resources, counseling services, or employee assistance programs.

- Personalized Recommendations:

  Enhance the system's ability to provide personalized recommendations based on an individual's historical data and responses. This could involve tailoring interventions and suggestions to specific needs and preferences, making the system more user-centric.

- Privacy and Ethical Considerations:

  Implement robust privacy measures to ensure the ethical use of facial data for mental health assessments. Develop clear guidelines and protocols for data storage, access, and sharing to maintain trust and compliance with privacy regulations.

- User Interface and Accessibility:

  Improve the user interface of the system to make it more intuitive and userfriendly. Consider accessibility features to ensure that individuals with diverse needs can easily interact with and benefit from the system.

- Continuous Model Training:

  Implement a mechanism for continuous model training to keep the system updated with evolving facial expressions and emotional nuances. This could involve periodically retraining the model with new datasets to improve its adaptability over time.

Chapter 8

PLAGIARISM REPORT

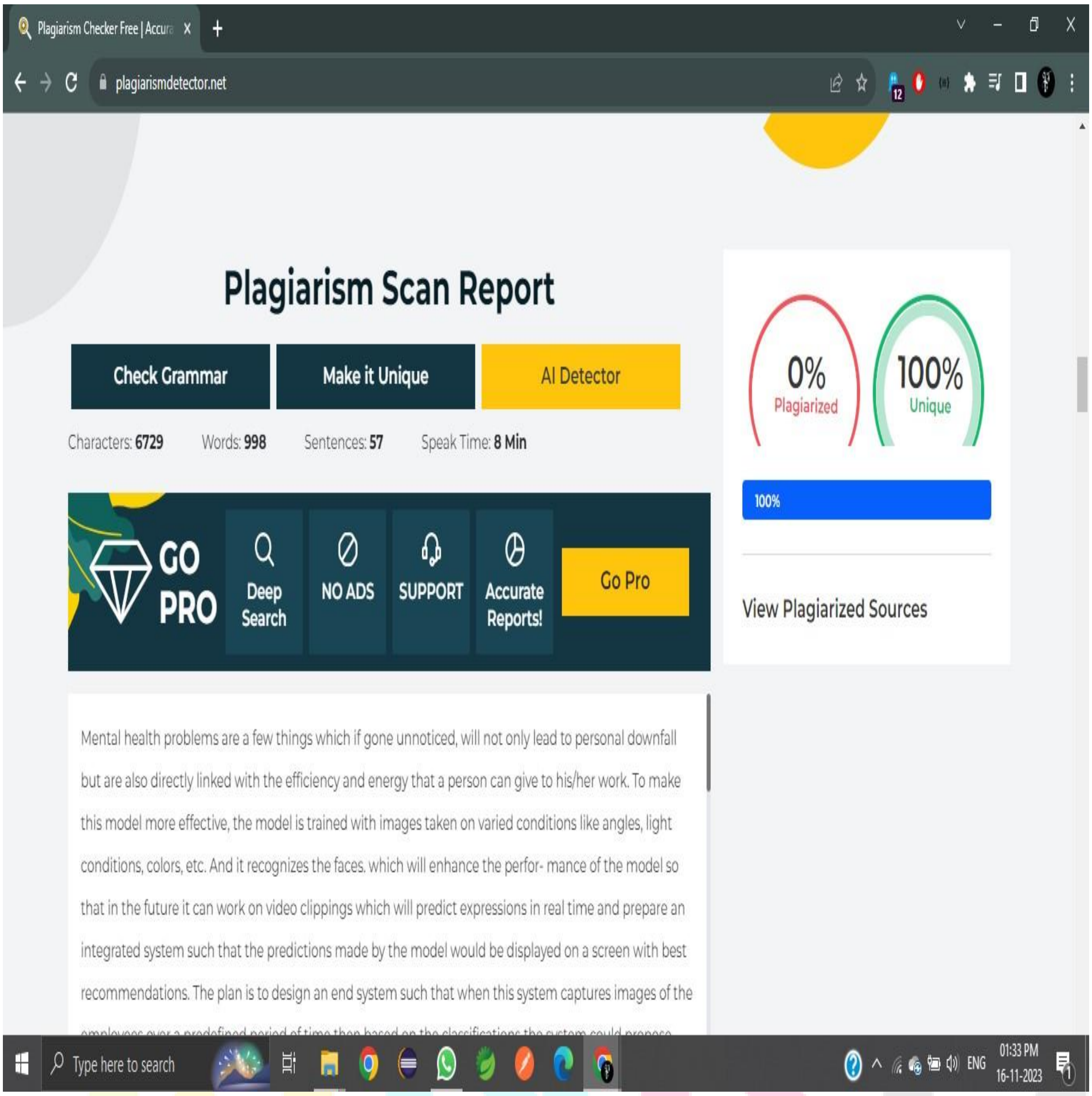

Figure 8.1: Plagiarism Report

Chapter 9
SOURCE CODE & POSTER PRESENTATION

9.1 Source Code

```
import pandas as pd importnumpy as np
importmatplotlib . pyplot as p l t importos
import cv2 importtensorflow as t f
t f . version fromtensorflow . keras . layers import Dense , Dropout , F l a t t e n fromtensorflow . keras . models import Sequential fromtensorflow . keras . preprocessing . image import ImageDataGenerator fromtensorflow . keras . layers import Convolution2D fromtensorflow . keras . layers import MaxPooling2D fromsklearn . metrics import c l a s s i f i c a t i o n _r e p o r t model= Sequential ( ) model . add ( Convolution2D (64 ,(5 ,5) , input _shape (48 ,48 ,1) , a c t i v a t i o n = ' relu ' ) ) model . add ( Convolution2D (64 ,(5 ,5) , a c t i v a t i o n = ' relu ' ) ) model . add ( MaxPooling2D ( pool _size =(2 ,2) ) ) model . add ( Dropout ( 0 . 4 ) ) model . add ( Convolution2D (128 ,(3 ,3) , a c t i v a t i o n = ' relu ' ) ) model . add ( Convolution2D (128 ,(3 ,3) , a c t i v a t i o n = ' relu ' ) ) model . add ( MaxPooling2D ( pool _size =(2 ,2) ) ) model . add ( Dropout ( 0 . 4 ) ) model . add ( F l a t t e n ( ) ) model . add ( Dense ( u n i t s =128 , a c t i v a t i o n = ' relu ' ) ) model . add ( Dense ( u n i t s =7 , a c t i v a t i o n = ' softmax ' ) ) model . summary ( ) model . compile ( loss = ' c a t e g o r i c a l _c r o s s e n t r o p y ' , optimizer = 'adam ' , metrics =[ ' accuracy ' ] ) t r a i n _d a t a g e n = ImageDataGenerator ( r e s c a l e =1.255 , r o t a t i o n r a n g e =30 , shear range =0.3 , zoom range =0.3 , h o r i z o n t a l \ _f l i p =False ) t e s t \ datagen = ImageDataGenerator ( r e s c a l e =1./255)
```

```
train\generator = train\datagen . flow\ from\ directory ( / content / Training ' , target\_size =(48 ,48) ,
    batch\ size =512 , color\ mode=" grayscale " , class\ mode=' categorical ' ) \\
validation\generator = test\ datagen . flow\ from\ directory ( / content / Testing ' , target\_size =(48 ,48)
    , batch\ size =512 , color \ mode=" grayscale " , class \ mode=' categorical ' ) \\
filepath = os . path . join ( " . / emotion\_detector \_models / model\ v{epoch }. hdf5 " ) checkpoint= tf .
keras .
    callbacks . ModelCheckpoint ( fi le pa th , monitor= ' val \ accuracy ' , verbose =1 , save\ best \ only=True ,
mode= 'max ' )
callbacks = [ checkpoint ]
nb\ train \ samples= 28709
nb\ validation \ samples = 717)
batch\ size =512
history =model . fit \ generator ( train \ generator , steps \ per \ epoch=nb\_train \_samples batch\ size , epochs =50 , validation \_
data = validation \_generator , callbacks = callbacks , validation \ steps =nb\
    validation \ samples // batch\ size )
fig , ( ax1 , ax2 )= plt . subplots ( nrows =1 , ncols =2 , figsize =(20 ,6) )
ax1 . plot ( history . history [ ' accuracy ' ] , label = ' train \ accuracy ' )
ax1 . plot ( history . history [ ' val \ accuracy ' ] , label = ' test \ accuracy ' )
ax1 . legend ( )
ax2 . plot ( history . history [ ' loss ' ] , label = ' train \ loss ' )
ax2 . plot ( history . history [ ' val \_loss ' ] , label = ' test \ loss ' )
ax2 . legend ( )
plt . show ( )

importtensorflow as tf
fromtensorflow . keras . models import load\ model
fromtensorflow . keras . preprocessing . image import img\ to \ array
import cv2
importnumpy as np
face \ classifier = cv2 . CascadeClassifier ( ' . / Harcascade / haarcascade \_frontalface \_default . xml ' )
classifier =load\_model ( ' . / Models / model\ v\ 47 . hdf5 ' )
class \ labels ={0: 'Angry ' , 1: ' Disgust ' , 2: ' Fear ' , 3: 'Happy ' , 4: ' Neutral ' ,5: 'Sad '
, 6: ' Surprise
'}\\
cap = cv2 . VideoCapture (0)
while True : ret , img = cap . read ( )
gray=cv2 . cvtColor ( img , cv2 .COLOR\ BGR2GRAY)
    faces = face \_classifier . detectMultiScale ( gray , 1.3 , 5)
    allfaces = [ ]
    rects = [ ]
    for (x , y , w, h ) in faces :
    cv2 . rectangle ( img , (x , y ) , ( x+w, y+h ) , (255 , 0 , 0) , 2)
    roi \ gray = gray [ y : y+h , x : x+w]
    roi \ gray = cv2 . resize ( roi \_gray , (48 , 48) , interpolation =cv2 . INTER\ AREA)
allfaces . append ( roi \_gray ) rects . append ( ( x , w, y , h ) )
i = 0
```

```
for face in allfaces:
roi = face . astype ( ” f l o a t ” ) / 255.0 roi = img\
to \ array ( roi ) _ _
roi = np . expand\_dims ( roi , axis =0)
```

```
l a b e l = c l a s s \_l a b e l s [ preds . argmax ( ) ] l a b e l \_p o 255 , 0) , 2)
i o n = ( r e c t s [ i ] [ 0 ] + i n t ( ( r e c t s [ i ] [ 1
2 ) ) ,
abs ( r e c t s [ i ] [ 2 ] − 10) )
i = + 1
cv2 . putText ( img , label , l a b e l \ position ,
cv2 .FONT\ HERSHEY\ SIMPLEX , 1 , (0 ,
cv2 . imshow ( ”MENTAL HEALTH IDENTIFICATION” , img )

i f cv2 . waitKey (1) == 13: # 13 i s the Enter Key
break cap . r e l e a s e ( ) cv2 .
destroyAllWindows ( )
```

```
preds = c l a s s i f i e r . p r e d i c t ( roi ) [0]
```

9.2 Poster Presentation

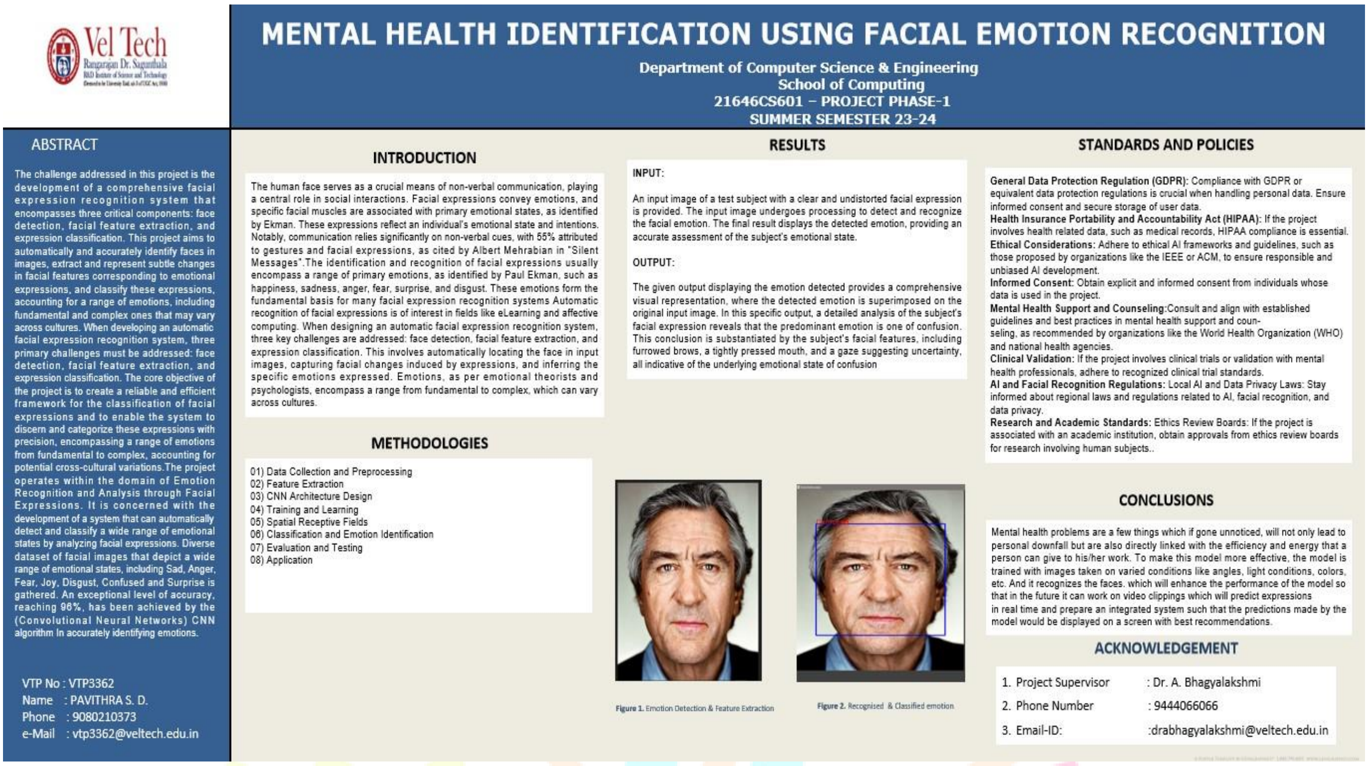

Figure 9.1: Poster Presentation

References

[1] Jyoti Kumari, R., Rajesh, K. M. Pooja,"Facial Expression Recognition: A Survey, Procedia Computer Science",Volume 58,2015, Pages 486–491, ISSN 1877 – 0509.

[2] Kalaivani, G., & Anitha, S. Sathyapriya,DrD. (2016). A Literature review on Emotion Recognition for Various Facial Emotional Extraction. IOSR Journal of Computer Engineering (IOSR-JCE).

[3] Ravichandra Ginne, K., Jariwala, "FACIAL EXPRESSION RECOGNITION USING CNN: A SURVEY", Mar 2018, International Journal of Advances in Electronics and Computer Science, Vol 5, No. 3, 2018.

[4] Turetsky, B. I., Kohler, C. G., Indersmitten, T., Bhati, M. T., Charbonnier, D., & Gur, R. C. (2007). Facial emotion recognition in schizophrenia: when and why does it go awry? Schizophrenia research, 94(1– 3), 253–263.

[5] Gibbons, R. D., Weiss, D. J., Frank, E., & Kupfer, D. (2016). Computerized Adaptive Diagnosis and Testing of Mental Health Disorders. Annu Rev Clin Psychol, 12, 83–104, Epub 2015 Nov 20. PMID: 26651865.

[6] Shojaeilangari, W., Yau, K., Nandakumar, J., Li, & Teoh, E. K. (July 2015). "Robust Representation and Recognition of Facial Emotions Using Extreme Sparse Learning". IEEE Transactions on Image Processing, 24(7), 2140–2152. doi:10.1109/TIP.2015.2416634.

[7] Shaul Hammed, A., Sabanayagam, E. Ramakalaivani, & A REVIEW ON FACIAL EXPRESSION RECOGNITION (2020). SYSTEMS",Journal of Critical Reviews, 7(4), 903–905. doi:10.31838/jcr.07.04.173.

[8] Singh, D. (2012). Human Emotion Recognition System. International Journal of Image, Graphics and Signal Processing. 4. 10.5815/ijigsp.2012.08.07.

[9] Zhang, L., & Tjondronegoro, D. (2011). Facial expression recognition using facial movement features. IEEE Trans Affect Comput, 2(4), 219–229.

[10] Kukla, E., & Nowak, P. (2015). Facial Emotion Recognition Based on Cascade of Neural Networks. In A. Zgrzywa, K. Choros & A. Siemi´ nski (Eds.), New´ Research in Multimedia and Internet Systems. Advances in Intelligent Systems and Computing (Vol. 314). Cham: Springer.

[11] S. Mollahosseini, D. Chan, and M. H. Mahoor, ”Going deeper in facial expression recognition using deep neural networks,” in 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1-10.

[12] T. Ahonen, A. Hadid, and M. Pietikainen, ”Face description with local binary¨ patterns: Application to face recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, 2006.

[13] P. Viola and M. Jones, ”Rapid object detection using a boosted cascade of simple features,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, vol. 1, pp. I-511-I-518.

[14] L. W. Barsalou, ”Grounded cognition,” Annual Review of Psychology, vol. 59, pp. 617-645, 2008.

[15] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, ”Recognizing facial expression: Machine learning and application to spontaneous behavior,” in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), 2005, vol. 2, pp. 568-573.