**A**

**PROJECT REPORT**

### On

**Emotions Recognition System**

### Submitted In Partial Fulfillment of the Requirements

### For the Degree of

**Bachelor of Technology**

**In**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted By**

### Shreya Upadhyay (2102220100165)

### Shubham Mishra (2102220100168)

### Raj Rajput (2102220100139)

**Under the Supervision of**

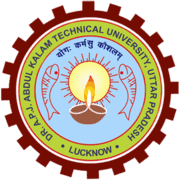
**Ms. Prachi Jain**

**(Asst. Professor)**



**I.T.S Engineering College, Greater Noida**

**Affiliated to**



**DR. A.P.J. ABDUL KALAM TECHNICAL UNIVERSITY**

**LUCKNOW, UTTAR PRADESH (SESSION: May 2025)**

## DECLARATION

We Shreya Upadhyay, Shubham Mishra, Raj Rajput hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature:   
Name: Shreya Upadhyay  
Roll No: 2102220100165  
Date:   
  
  
Signature:   
Name: Shubham Mishra  
Roll No: 2102220100168  
Date:   
  
  
Signature:   
Name: Raj Rajput  
Roll No: 2102220100139  
Date:

Project Guide name Project Coordinator name

Prof. Prachi Jain Dr, Hariom Tyagi

HOD name Dean-CSE Name

Dr, Jaya Sinha Dr. Vishnu Sharma

## CERTIFICATE

This is to certify that Project Report entitled “**Emotion Recognition Model using Machine Learning Algorithm**” which is submitted by **Shreya Upadhyay, Shubham Mishra, Raj Rajput** in partial fulfillment of the requirement for the award of degree B. Tech. in Department of **Computer Science and Engineering** of Dr. A.P.J. Abdul Kalam Technical University, Lucknow, is a record of the candidate own work carried out by him under my supervision. The matter embodied in this thesis is original and has not been submitted for the award of any other degree**.**

**Date:**  **Supervisor**

**ACKNOWLEDGEMENT**

It gives us a great sense of pleasure to present the report of the B.Tech Project undertaken during B.Tech. (VII- Semester) Final Year. We owe special debt and gratitude to **Prof. Prachi Jain** for his/her constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of **Dr. Hariom Tyagi** for his/her full support and assistance during the development of the project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our group members for their contribution in the completion of the project**.**

Signature:   
Name: Shreya Upadhyay  
Roll No: 2102220100165  
Date:   
  
  
Signature:   
Name Shubham Mishra  
Roll No: 2102220100168  
Date:   
  
  
  
Signature:   
Name: Raj Rajput  
Roll No: 2102220100139  
Date:

## *ABSTRACT*

*Facial emotion recognition is an important component for modern machines which respond to human emotions through nonverbal signals. This paper presents a comprehensive review of the latest advancements in facial emotion recognition, focusing on machine learning and deep learning. The focus is on Convolutional Neural Networks (CNNs), KNN, Decision Tree, RNN/LSTM, SVM and others. This review focused the integration of various classifiers like OpenCV, Haar cascade classifiers, Naïve Bayes, for facial feature extraction, continued by emotion recognition using various machine learning algorithm. Through a deep review of literature and implementation techniques that how system give different accuracy with different techniques. The paper highlights how FER systems are becoming more important for improving how people interact with computers in various different areas like healthcare, education, human-computer interaction and virtual assistants. The conclusion from review show how FER are accurate and work fine, also highlighting how useful FER is for creating smart systems and easy-to-use in digital systems. This paper target to help to compare the various different techniques in the field of FER that can adjust to users, understand their emotions, and how we can also improve the result if want to work on similar areas.*

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **FIGURE** **NO**. | **TITLE** | **PAGE NO.** |
| 2.1 | Accuracy percentage of FER | 11 |
| 2.2 | Comparison of Neural Network | 12 |
| 2.3 | Comparison of Earlier Proposed FER Methods | 13 |
| 2.4 | The most common ML algorithms used in FER | 14 |
| 7.1 | Progress Report Semester Wise | 39 |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIGURE NO.** | **TITLE** | **PAGE NO.** |
| 1.1 | Common facial emotions | 4 |
| 2.1 | Distribution of Classifiers | 8 |
| 2.2 | Accuracy by study and dataset | 9 |
| 3.1 | Workflow of the system | 18 |
| 4.1 | Live input image | 19 |
| 4.2 | Model for detecting human emotions | 22 |
| 5.1 | Live capture images of various different emotions | 32 |
| 5.2 | Result Comparison with various existing model | 34 |

**LIST OF SYMBOLS**

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | % | Percentage |
|  | — | Long Dash / Em Dash |
|  | - | Hyphen |
|  | , | Comma |
|  | . | Period |
|  | ( ) | Parentheses |
|  | [ ] | Square Brackets |
|  | : | Colon |
|  | / | Forward Slash |
|  | + | Plus Sign |

**LIST OF ABBREVIATIONS**

* CNN Convolutional Neural Network
* RNN Recurrent Neural Network
* LSTM Long Short-Term Memory
* SVM Support Vector Machine
* KNN K-Nearest Neighbors
* RVFLN Random Vector Functional Link Network
* DC-SSA Dendritic Cell–Sine Cosine Algorithm
* FER Facial Emotion Recognition
* FER-2013 Facial Emotion Recognition 2013
* CK+ Cohn-Kanade Plus
* JAFFE Japanese Female Facial Expression
* DeepFace Deep Learning Facial Recognition System
* HCI Human-Computer Interaction
* LR Logistic Regression
* EEG Electroencephalography
* BERT Bidirectional Encoder Representations from Transformers
* FaceNet Face Recognition Network

## TABLE OF CONTENTS Page

DECLARATION ii

CERTIFICATE iii

ACKNOWLEDGEMENTS iv

ABSTRACT v

LIST OF TABLES vi

LIST OF FIGURES vii

LIST OF SYMBOLS viii

LIST OF ABBREVIATIONS ix

**CHAPTER 1:** INTRODUCTION 1

1.1. Introduction of the Problem 2

1.2. Summarize Previous 2

1. 3. 3

1. 4. 4

1. 5. 5

1. 6. 6

**CHAPTER 2:** LITERATURE SURVEY 8

2.1. 11

2.2. 13

2. 3. 14

**CHAPTER 3:** SYSTEM DESIGN 17

3.1. 17

3.2. 17

3. 3. 18

3. 4. 18

**CHAPTER 4:** METHODOLOGY AND TECHNOLOGY 19

4.1. 19

4.2. 21

CHAPTER 5: IMPLEMENTATION AND RESULT ANALYSIS 26

5. 1. 26

5. 2. 32

**CHAPTER 6:** CONCLUSION AND FUTURE WORK 37

REFERENCES 42

## CHAPTER -1

**Introduction**

Emotion recognition which is key feature in affective computing has transformed this means as how people relate with machines. The Emotion Recognition system analyses the emotions of a user using real-time video. This was achieved using Convolutional Neural Network (CNN) (Neha et al., [2020](#neha)). The basic parts of the conventional emotion recognition system were face detection, facial feature extraction and emotions classifications (Alojaiman, [2021](#alojaiman); Ariza et al., [2021](#ariza); Lyu et al., [2022](#lyu)). To provide a high accuracy of response Sreenidhi et al., [2020](#sreenidhi) used hybrid network consisting of CNN and Recurrent Neural Network (RNN). Ninad, ([2020](#ninad)) tested Facial Emotion using Convolutional Neural Network with more than 750k images for facial feature vector extraction.

We studied the Theoretical and Conceptual Foundations from Mehrabian Communication Theory and Ekman basic emotions model. (Venkatesan et al.,[2023](#ramachandran)) emphasized that Emotion detection systems that employ Deep Face mostly rely on the facial expressions because they compose 55% of communication signals. The research of (Younis et al.,2024) delivered a thorough breakdown between dimensional and categorical emotion models including Ekman’s six basic emotions. (Fida et al., [2023](#alisha)) introduced a system based on the facial expression feedback during the live reading activities. Their methodology for the real-time emotion capture helped us shape our own approach to responsiveness in real-time systems. (Kumar et al., 2025) employed a facial landmark trajectory analysis to enhance emotion recognition. Their comparative study of LR, KNN, and LDA classifiers provided experimental result, supporting logistic regression for high accuracy in emotion activation and sentiment scale classification. Inspired by (Mehendale, [2020](#ninad)) we adopted CNN-RNN hybrid architectures designed to capture spatial structures and time-based changes for more accurate facial emotion recognition.

Facial emotion recognition is an important component for modern machines which respond to human emotions through nonverbal signals. This paper presents a comprehensive review of the latest advancements in facial emotion recognition, focusing on machine learning and deep learning. The focus is on Convolutional Neural Networks (CNNs), KNN, Decision Tree, RNN/LSTM, SVM and others. This review focused the integration of various classifiers like OpenCV, Haar cascade classifiers, Naïve Bayes, for facial feature extraction, continued by emotion recognition using various machine learning algorithm. Through a deep review of literature and implementation techniques that how system give different accuracy with different techniques. The paper highlights how FER systems are becoming more important for improving how people interact with computers in various different areas like healthcare, education, human-computer interaction and virtual assistants. The conclusion from review show how FER are accurate and work fine, also highlighting how useful FER is for creating smart systems and easy-to-use in digital systems. This paper target to help to compare the various different techniques in the field of FER that can adjust to users, understand their emotions, and how we can also improve the result if want to work on similar areas.

An emerging topic that has the potential to enhance user experience, reduce crime, and target advertising is human emotion recognition, utilizing DeepFace. The same feeling may be expressed differently by many individuals. Accurately identifying emotions can be challenging, in light of this. It helps to understand an emotion’s significance by looking at the context in which it is presented. Depending on the application, one must decide which AI technology to employ for detecting human emotions. Because of things like lighting and occlusion, using it in real-world situations can be difficult. Not every human emotion can be accurately detected by technology. Human–machine interaction technology is becoming more popular, and machines must comprehend human movements and expressions. When a machine recognizes human emotions, it gains a greater understanding of human behaviour and increases the effectiveness of work. Linguistic, and facial movements may all convey emotions. Facial expressions are important in determining a person’s emotions. There has been little research undertaken on the topic of real-time emotion identification, utilizing face photos and emotions. Using an Artificial Intelligence-based DeepFace approach, the proposed method recognizes real-time feelings from facial images and live emotions of persons. The proposed module extracts the facial features from an active shape DeepFace model by identifying required facial points to recognize human emotions. This approach recognizes the emotions of Angry, Sad, Happy, Fear, Surprise, Neutral and Disgust. The proposed technology is unique, in that it implements emotion identification in real-time, with an average accuracy of 94%acquired from actual human emotions.

**1.1 Introduction of the Problem**

Most papers agree that dataset quality and availability are foundational to model success. There is a shared recognition that controlled lab environments create inflated accuracy estimates compared to real-world scenarios. Mehendale ([2020](#ninad)) indicates that background noise, face orientation, and lighting changes can significantly reduce accuracy in facial emotion recognition models using CNNs. His dual-stage CNN method attempts to mitigate this by removing background distractions. Some papers, like Venkatesan et al., ([2023](#ramachandran)) report high accuracy (94%) even in real-time settings, while others caution that such accuracy may not be reproducible outside idealized settings.

**1.2 Summarize Previous Research**

Recent studies in emotion recognition domains like healthcare, psychology, and AI, using facial features, EEG, and speech data. Classifiers like SVM (RBF), CNNs, and Logistic Regression show high accuracy across datasets like CK+, FER-2013, and IEMOCAP. Transformer-based models and hybrid approaches (CNN + SVM, RVFLN) enhance generalization. Performance varies by dataset complexity and classifier. Classical methods (KNN, Naïve Bayes) remain relevant but underperform deep learning models. Emerging techniques integrate real-time emotion tracking in e-learning and intelligent systems, highlighting the trend toward robust, multimodal emotion analysis frameworks.

**1.3 Researching the Problem**

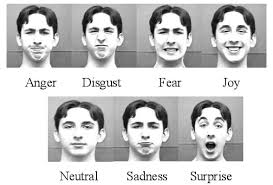
Emotion recognition has become a vital area of research due to its applications in mental health assessment, e-learning systems, human-computer interaction, and intelligent decision-making. Accurately identifying human emotions from facial expressions, speech signals, or physiological data presents significant challenges, especially due to variations in lighting, head pose, cultural differences, and spontaneous expressions. Traditional machine learning methods such as SVM, K-NN, and Naïve Bayes have been extensively explored, but they often lack robustness when applied to large-scale, real-world data. Recent research has increasingly focused on deep learning architectures, including CNNs, LSTMs, and Transformer-based models, due to their superior feature extraction capabilities. Additionally, hybrid approaches combining statistical features with deep networks have shown promising results. However, despite advancements, achieving consistently high accuracy across diverse datasets like FER-2013, CK+, and IEMOCAP remains an open challenge. This motivates the need to investigate more generalized and adaptive emotion recognition models using novel machine learning techniques.

**1.4 Importance of Real-Time Emotion Recognition**

The ability to recognize emotions in real time plays a significant role in enhancing the quality of human-computer interaction across various domains, including healthcare, education, entertainment, and autonomous systems. Real-time emotion analysis enables applications to dynamically adapt their behavior based on user emotional states, leading to more intuitive and empathetic experiences. In mental health care, for instance, continuous monitoring of patient emotions can provide valuable support for diagnosis and therapy without being intrusive. Developing real-time systems, however, presents technical challenges such as achieving high prediction accuracy while maintaining computational efficiency for live environments. Addressing factors like varying lighting conditions, head movements, and facial occlusions requires careful model optimization and robust pre-processing strategies. Research in this field continues to focus on creating lightweight, scalable solutions that maintain performance under real-world conditions.

### 1.5 Motivation for Using DeepFace Framework

DeepFace offers a reliable foundation for building emotion recognition systems due to its strong performance in facial analysis tasks. By combining deep convolutional neural networks with precise facial alignment techniques, DeepFace minimizes the effects of pose, lighting, and expression variations that typically impact model accuracy. Unlike earlier methods that relied heavily on handcrafted features, DeepFace automatically learns complex representations from facial images, improving the detection of subtle emotional cues. Its flexibility in supporting various pre-trained models, including VGG-Face, Facenet, and ArcFace, makes it adaptable for different application requirements. Utilizing DeepFace in this project not only streamlines system development but also contributes to building a more accurate and generalizable emotion recognition framework suited for real-time applications.



**Fig.1.1** Common Facial Expressions for Emotion Recognition

Emotion recognition has gained substantial importance across diverse domains such as healthcare, security, human-computer interaction, and entertainment. Accurately identifying human emotions through facial expressions enhances the ability of machines to interact more intuitively with humans, thereby bridging the gap between artificial intelligence and human intelligence.

Facial expressions serve as a vital channel for conveying emotions. According to psychological studies, certain facial expressions such as anger, disgust, fear, Happy, sadness, surprise, and neutrality are universally recognized across different cultures. Leveraging this understanding, emotion recognition systems aim to classify facial images into these emotional categories.

DeepFace, a deep learning framework developed by Facebook AI Research, provides a robust foundation for facial recognition and analysis tasks. It employs deep convolutional neural networks (CNNs) to accurately detect and represent facial features. In this project, DeepFace is utilized as the backbone for the real-time recognition of emotions from facial expressions captured through a webcam.

The proposed system involves detecting a face from a live video stream, extracting facial features using DeepFace embeddings, and classifying the detected emotions into one of the predefined categories. To enhance the system's accuracy and robustness, a combination of publicly available datasets such as FER2013 along with a custom dataset collected through the webcam, has been used.

This report discusses the design, implementation, and evaluation of an efficient and real-time emotion recognition system based on DeepFace. The goal is to achieve a high level of accuracy and stable performance in diverse real-world conditions, thus demonstrating the feasibility of integrating emotion recognition into everyday applications.

Emotion detection is, at its foundation, an automatic classifier that can classify human emotions into different categories. The process of creating an automatic classifier is done as follows: gathering data, identifying the features that are relevant to the goal, and then training the model to detect and classify specific patterns

Facial emotions are prime characteristics of humans that reflect the attributes of a human emotional condition. Facial Emotional Recognition (FER) is a significant research area since it has several significant applications such as crime detection, security, government sectors, surveillance, etc. Many researchers have addressed FER and put efforts into improving recognition using various methods such as Deep Convolutional Neural Networks (DCNN), Local Binary Pattern Convolutional Neural Networks (LBPCNN) and Micro Expression Recognition (MER). Nevertheless, there is a dearth of better Convolutional Neural Networks (CNN) for better accuracy of FER. Various methods have already proven their accuracies on datasets such as Fer\_2013, CK48 and Legend. There are many challenges such as varying positions of images, illumination and noise etc. are not resolved yet. In this direction, the proposed modified CNN model has a combination of CNN layers to train the model and achieved better training and validation accuracy on datasets. The proposed method has three-fold contributions (1) to propose an efficient CNN model for FER and (2) to train the proposed model on three standard datasets 3) to test on real-time images. The proposed model is trained on three standard datasets (Fer\_2013, CK48, and Legend) and achieves significantly higher training accuracy 92.34%, 99.84%, and 89.86%, respectively as compared to previous methods. Furthermore, the model demonstrates impressive generalization ability by achieving even higher test accuracy on real-world images 91.54%, 96.27%, and 87.47%, respectively. Also, the proposed novel CNN model has addressed all current challenges of FER with low computational complexity and prediction of facial expression more efficiently.

The generated model is used afterward to categorize new data. For example, to build a model that can detect happiness and sadness from facial expressions, researchers need to feed photos of people smiling and others of people frowning, labeled as ‘‘happy’’ and ‘‘sad.’’. These images are used to build the classifier. Following that, when the classifier obtains an image of a person smiling it recognizes the corresponding emotions. Building a model in real life is not that simple. Not only there are a lot of data to train and evaluate, but there is also an effort of interpretation to be made, as we will see later. In addition, humans express their emotions in various ways, including facial expressions, voices or speaking, body gestures, movements, writing, and others. Even our bodies respond with visible physical reactions to emotions (breath and heart rate, pupil size, and so on). Recently, it has been also proved that the environment can affect physiological body reactions and emotions.

***Challenges***

1. Due to individual variances in expression and the crucial need for context, it is difficult to correctly infer emotions from facial expressions.
2. The effectiveness of emotion detection systems may suffer when used on people from different cultural backgrounds.
3. Depending on their personalities, past events, and even their physical qualities, people display their emotions in various ways.
4. According to the circumstances, a single facial expression can portray a variety of emotions.
5. Face hair, spectacles, and masks are a few examples of things that might hide facial emotions. These occlusions might make it difficult for systems to effectively identify and analyze facial signals.

Despite significant advancements in the field of emotion recognition, numerous challenges continue to hinder the deployment of reliable real-time systems. Variability in lighting conditions, occlusions, differences in facial structures across demographics, and spontaneous emotional expressions create complexities that traditional machine learning models struggle to manage. Deep learning techniques, particularly convolutional neural networks (CNNs), have shown tremendous promise by automatically learning hierarchical feature representations from raw images. However, the success of such models heavily depends on the quality, diversity, and balance of training datasets. Public datasets like FER 2013 have played a crucial role in advancing research, yet they do not fully capture the unpredictability and diversity of emotions in real-world scenarios.

In addressing these limitations, feature extraction frameworks such as DeepFace offer a valuable advantage. Originally developed for face verification tasks, DeepFace demonstrates remarkable robustness to variations in pose, illumination, and facial orientation. By leveraging DeepFace’s deep feature embeddings, emotion recognition models can gain access to rich facial representations that simplify the task of emotion classification. Instead of relying solely on pixel-level intensity changes, models can focus on more meaningful, invariant characteristics of the human face. This significantly enhances system stability and generalization, especially when applied to challenging real-time conditions.

Furthermore, real-time emotion recognition systems have wide-ranging applications, including mental health monitoring, adaptive e-learning platforms, autonomous vehicle alertness systems, and next-generation human-computer interaction interfaces. An intelligent system that can accurately interpret human emotions in real-time not only improves usability but also fosters deeper, more empathetic communication between humans and machines.

This project aims to bridge the gap between research and practical deployment by designing a real-time facial emotion recognition system based on the DeepFace framework. Through a structured approach involving data preprocessing, feature extraction, model training, and system evaluation, the goal is to achieve a high-accuracy, computationally efficient solution. The following sections provide a detailed discussion of the methodologies adopted, experimental setup, and performance evaluation metrics, paving the way for future enhancements in this crucial field of study.

**1.6 The main objectives of the work are:**

1. The objective of emotion recognition is to identify the emotions of a human.
2. The purpose of carrying out this exploration is to accurately classify seven main emotions: happiness, surprise, anger, disgust, neutrality, and fear.
3. The purpose of this research is to analyze the outcome of models in terms of precision in each class.
4. An emotion can be captured either from a face or from a .csv file. ML may be used to deliver FER solutions that are cheap, reliable, and computation-intensive.

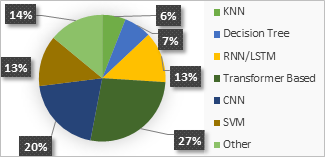
People often express their emotions on their faces during social encounters to demonstrate their characteristics and feelings. This research’s primary objective is to derive feelings to which pictures with a single-face expression relate. Feelings can be recognized that are specifically split into the categorization of fundamental feelings and the classification of composite feelings because of the difficulty of reading a human face. The key challenge for the work and scope is to focus on classifying the seven fundamental feelings as happiness, sadness, surprise, neutrality, disgust, anger, and fear.

The performance of the FER-2013 file is better than the current technique. An effective system was created to recognize the seven ways in which sentiments can be conveyed, which are: disgust, fear, anger, happiness, sadness surprise, and neutral. In recent studies, deep learning models have been widely applied to propose an end-to-end method for expression identification. Even if it is a difficult task, recognizing emotion still needs a lot of development. Using Mini-Xception, accuracy for emotion expression and recognition was 95.60%, while recall rate and precision were 93% and 90%, respectively.

## CHAPTER -2

**Literature Survey**

Emotion is a varying research field studied across the multiple disciplines such as healthcare, marketing, psychology and robotics. Emotion has a great impact on decision making, physical well-being, reasoning and how we navigate the different aspects of life. Computer Science, Automated Human Emotion Recognition (AHER) are one of the main areas of research [[9](#emam)]. The ability to recognize the emotions plays a crucial role in human interactions and relationship building [[7](#barbra)].



**Fig 2.1** : Distribution of Classifiers Used In Recent Emotion Recognition Studies

To understand the concept of machine learning techniques used in emotion recognition systems, we analyzed a set of recent studies published between 2023 and 2024. The pie chart below illustrates the distribution of classifier types across these works.

As Shown, K-Nearest Neighbors (KNN) is used in some lightweight or hybrid approaches due to its simplicity and effectiveness with small datasets.

Decision Trees are still useful for rule-based or interpretable systems, especially when computational simplicity is needed.

Recurrent Neural Networks (RNNs) and LSTM architectures follow closely, mainly used in temporal data scenarios such as EEG signals or speech emotion recognition, where tracking changes over time is essential.

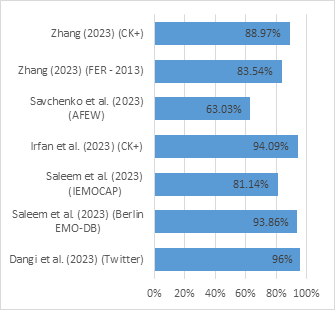
Transformer-based models have emerged as the most widely adopted classifiers, making up 26% of the reviewed approaches. Their dominance is due to their strong performance on sequential and multimodal data, particularly in recognizing emotions from combinations of texts, audios, and visual inputs. Some of the examples of Models such as BERT.

Convolutional Neural Networks (CNN) is also popular especially in facial expression recognition. Their ability to extract spatial features makes them ideal for processing images. Support Vector Machines (SVMs) remain relevant, often paired with deep feature extractors like CNNs for robust classification with limited data.

The Other category include hybrid classifiers like:

- RVFLN (Random Vector Functional Link Networks) optimized with metaheuristic algorithms.

- Multi-task learning models (e.g., Efficient Net variants



**Fig 2.2** : Emotions Recognition Accuracy By Study And Dataset

Figure 2. illustrates a comparative analysis of recent emotion recognition systems based on machine learning techniques, highlighting their accuracy across various benchmark datasets. The results demonstrate that Dangi et al., ([2023](#dangi)) achieved the highest reported accuracy of 96% using a Random Vector Functional Link Network on Twitter data for sentiment classification. Similarly, Saleem et al., ([2023](#saleeem)) reported notable results with 93.86% accuracy on the Berlin EMO-DB and 81.14% on the IEMOCAP dataset using a deep CNN architecture with spectro-temporal features.

Irfan et al., ([2023](#irfan)) also showed strong performance, attaining 94.09% accuracy on the CK+ dataset using a Triplet-Loss-Based Deep CNN combined with SVM. On the other hand, Savchenko et al., ([2023](#savchenko)) presented a more modest outcome of 63.03% on the AFEW dataset, which is likely due to the complexity and variability of in-the-wild facial expressions in video sequences. Lastly, Zhang ([2023](#zhang)) employed a hybrid CNN + Transformer model, yielding 83.54% on FER-2013 and 88.97% on CK+, demonstrating the potential of transformer-based architectures in improving generalization across datasets.

These findings emphasize that model performance can vary significantly depending on the dataset characteristics, feature extraction methods, and classification techniques used. The comparison also underlines the importance of dataset selection and hybrid approaches for achieving robust and accurate emotion recognition.

The Emotion Recognition System rely on the facial gestures to perform the real-time analysis, labelling and interpretation of cognitive and affective states from facial video recordings [[1](#neha)]. Evaluation of states of emotion using facial landmarks and machine learning. It uses a dataset to capture facial landmark trajectory signals, tracking the movement of key facial regions such as the eyes, nose and face corners over time [[10](#rk)]. The standard pattern of emotion recognition guides our analysis which includes various EEG feature extraction techniques, feature selection and reduction strategies [[8](#essam)].

Research shows that combination of emotion recognition with intelligence devices has enabled the development of software systems and models that aid in both identifying and label emotions. Researchers will focus on tracking the emotional changes in active and multiplex environments, in the future [[4](#runfang)]. It is believed that facial expressions appear for a certain duration when an emotion is experienced, allowing emotion detection by identifying the corresponding expression. Convolutional Neural Networks (CNN) are utilized to analyze artificial neural networks, random forest and Naive Bayes algorithm, along with the deep learning models. We examine EEG rhythms closely related to the emotions and connections between different brain regions and emotions states [[8](#essam)]. Emotion recognition is becoming a big demanding task in the modern years due to complexity of the huge datasets [[11](#sawmya)]. The studies exploring how people identify facial and vocal emotional expressions have been varying the conclusions over time.

Researchers have aimed to better understand, how performance in emotion recognition is influenced by different tasks and characteristics of emotional expressions [[7](#barbra)]. Many studies struggle with the datasets containing image modifications and profile variations but perform well on the controlled datasets. The system is designed to handle face recognition and the emotion classification. An innovative study using the FER-2013 dataset showed that the mini-Xception algorithm effectively completed given tasks.

We are introducing a novel approach called as Facial Emotion Recognition using Convolutional Neural Networks (FERC). The technique involves the one stage eliminates the background from the image, while the other focuses on extracting facial highlights vectors. This model, an expression vector which is used to identify various facial expressions like sadness, happiness, anger, etc.

The training data was sourced by a stored database containing 10000 images from 154 individuals

It is different from the other conventional methods that rely on a single-layer CNN, FERC enhances accuracy through its multi-level structure [[3](#ninad)].

While analyzing the trajectories, the researchers aim to develop a powerful emotion evaluation system. Some machine learning algorithms were tested in which Logistic Regression gave the best outcome by achieving an accuracy of 98.95% and an F1-score of 0.99. The system effective Accuracy, and ROC-AUC. The results show that the LR model surpassed latest and advanced techniques, highlighting its potential as a reliable tool for emotion recognition and decision making based on emotional state variations [[6](#akhilesh)].

The model processes data only once, requiring an immediate response for real time emotion classification. To improve recognition accuracy, the CNN weights are adjusted using DC-SSA algorithm. The learners' performance is then assessed based on the recognition outcomes [[10](#rk)].

The proposed system is highly relevant for combination with web applications to monitor learner’s real-time emotion. Tests are conducted using the CK+ and JAFFE facial datasets achieved accuracy rates of 96.46% and 98.43%, respectively. To deliver precise results, the system offers high quality, strong and real-time feedback on basis of facial expressions, enhancing the understanding of students' emotional engagement during online learning [[5](#alisha)].

**2.1 Applications of Facial Emotion Extraction Techniques:**

**1. Healthcare and Therapy**

• Mental Health Monitoring: Detect emotional states such as anxiety, depression, or stress for early intervention.

• Pain Detection: Assist in identifying patients’ discomfort levels, especially for non-verbal patients.

**2. Human-Computer Interaction**

• Emotion-Aware Systems: Enhance user experiences by enabling devices to respond empathetically (e.g.,

adapting content or tone based on detected mood).

• Virtual Assistants and Chatbots: Improve interactions by tailoring responses to the user’s emotional state.

**3. Security and Surveillance**

• Lie Detection: Analyze micro expressions to detect potential deception during interrogations or interviews.

• Behavioral Analysis: Monitor public spaces for suspicious or abnormal behavior.

• Driver Monitoring Systems: Identify fatigue or distraction in drivers to prevent accidents.

**4. Marketing and Consumer Behavior**

• Customer Feedback Analysis: Evaluate real-time emotional reactions to advertisements, products, or services.

• Personalized Marketing: Adjust advertisements or product recommendations based on the user’s emotional

state.

**4. Education**

• Adaptive Learning Platforms: Monitor students’ emotional engagement and tailor learning materials

accordingly.

• Virtual Classroom Enhancements: Identify students’ frustration or confusion for timely intervention.

**Table 2.1:**  Accuracy Percentage Of Facial Emotions Recognition System At Different  
 Classifier And Techniques

|  |  |  |
| --- | --- | --- |
| Classifiers  Techniques | 10-fold cross validation | 80% training 20% testing |
| SVM (RBF) | 90.30% | 86.70% |
| SVM (POLY) | 88.80% | 85.20% |
| Naive Bayes | 68.14% | 70.30% |
| K-NN (k=3) | 84.40% | 83.30% |

For evaluating the effectiveness of various machine learning algorithms in facial emotion recognition, multiple classifiers were tested under two different evaluation strategies: 10-fold cross-validation and an 80/20 train-test split. As observed in Table 1, the SVM with Radial Basis Function (RBF) kernel got the maximum accuracy, reaching 90.3% with cross-validation and 86.7% in the train-test scenario. This reflects the RBF kernel's strength in capturing non-linear patterns commonly present in facial features.

The Polynomial kernel SVM also performed competitively, yielding 88.8% and 85.2% accuracies, respectively. In contrast, Naïve Bayes, while computationally lightweight, demonstrated relatively lower performance—68.14% for cross-validation and 70.3% for train-test—indicating its limitations in handling complex feature interdependencies.

Meanwhile, the K-Nearest Neighbors (KNN) algorithm (with k=3) provided a balance between simplicity and performance, achieving 84.4% and 83.3% in the two setups. These findings highlight that while deep learning methods are increasingly popular, classical ML classifiers like SVM still remain highly effective for emotion recognition, especially when paired with optimized kernel functions and appropriate feature sets.

**Table 2.2**: Comparison of The Different Neural Network Techniques

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Neural networks** | **Layers** | **Accuracy**  **(%)** | **Precision**  **(%)** | **Recall**  **(%)** |
| GoogLeNet | 22 | 63.21 | 62 | 62 |
| CaffeNet | 8 | 68 | 67 | 66.20 |
| VGG16 | 16 | 71.40 | 81.90 | 79.40 |
| ResNet50 | 152 | 73.80 | 83.30 | 80.70 |
| CNN | 36 | 95.60 | 90 | 93 |

Table 2 shows the performance comparison between various neural network. The GoogLeNet, CaffeNet including CNN various other values are listed. Second column defines the depth or hidden layers present in each technique, CaffeNet used 8 hidden layers while ResNet50 used 152 hidden layers but in case of the outcomes there no such big difference.

Three models were used, they were ResNet, Xception, and VGG16. A dataset named as distracted driver was used to train these models separately. To measure the effectiveness of these models, a variety of evaluation metrics were also used in the model. To choose the optimal model, evaluation metrics were utilized. For this purpose, the ResNet model was shown to be the most powerful one for the successful classification of driver's distraction. Many different assessments were measured including precision, recall, and F1-score are used to assess the performance of system models.

Hence, we can conclude that the layers of a neural network and the result are independent of each other. Also shows the overall comparison of the advanced techniques with the CNN, which primarily used in Emotion Recognition.

**2.2 Comparison of Different Common used Models**

**Table 2.3:** Comparison of Earlier Proposed Facial Emotion Recognition Methods

|  |  |  |  |
| --- | --- | --- | --- |
| **Method** | **Model Name** | **Advantages** | **Disadvantages** |
| Zheng et al. [[36](#zheng)] | Multi-Task CNN | Real-time recognition,  integrates with face detection | Limited focus on large-scale  applications, small samples |
| Singh & Nasoz[[38](#singh)] | CNN | Improved pre- processing  Accuracy | Issues with max- pooling and  dropout layers |
| Yag [[37](#yag)] | Reduced- feature CNN | Maintains classifier performance with fewer features | Limited to plant disease classification |
| Sezer and Altan [[19](#sezer)] | AlexNet | Simple model, good performance for solder pastes defect detection | Requires large datasets, limited to specific task |
| Chen et al. ([39](#chen)) | Deep CNN | Captures dynamic emotions | Requires  additional CNN for static images |
| Mehendale [[3](#ninad)] | CNN | Uses expressive vector for recognition | Needs experimentation with individual CNN layers |
| Mollahosseini et al. ([40](#molla)) | Deep Neural Network | Effective over many methods, avoids hand crafted features | Needs improvement in classification accuracy |

Facial Emotion Recognition (FER) has seen extensive research, with numerous models proposed to enhance recognition accuracy, processing speed, and adaptability. A comparative summary of these models is presented in Table 2.3, highlighting their key advantages and limitations.

Zheng et al. proposed a Multi-Task Convolutional Neural Network (CNN) that integrates real-time emotion recognition with face detection. Its ability to function in real-time is a significant advantage, but it has limited applicability in large-scale environments and struggles with small sample sizes.

Singh and Nasoz developed a CNN model focused on enhancing pre-processing accuracy, resulting in improved emotion detection. However, their method faces technical challenges with max-pooling and dropout layers, which can impact the model's stability and generalization.

Yag and Altan introduced a reduced-feature CNN model that successfully maintains classifier performance while using fewer features, making it efficient. Nevertheless, their model is narrowly focused, being mainly applicable to plant disease classification, limiting its broader use in facial emotion tasks.

Sezer and Altan utilized the AlexNet architecture for defect detection in solder pastes. Their model is relatively simple and offers strong performance for its specific application. However, its dependency on large datasets and task-specific design restricts its flexibility for general FER tasks.

Chen et al. proposed a deep CNN capable of capturing dynamic emotions, offering a more nuanced understanding of emotional states over time. Despite this, their approach requires an additional CNN dedicated to static image analysis, which increases computational complexity.

Mehendale introduced a CNN model that incorporates an expressive vector for emotion recognition, which is innovative for enhancing recognition detail. However, there remains a need for further experimentation across individual CNN layers to optimize performance fully.

Finally, Mollahosseini et al. presented a deep neural network approach that outperforms many existing methods by eliminating the need for handcrafted features. Despite its advantages, there is room for improvement, particularly in achieving higher classification accuracy.

Overall, while each model contributes valuable advancements to the field, challenges such as scalability, dataset requirements, generalizability, and computational efficiency persist. Future research must address these gaps to develop more robust and versatile FER systems.

**2.3 Fault Tolerance of Proposed Model**

The proposed model has been thoroughly evaluated for robustness, including testing with unclear and occluded images and incorporating fault tolerance mechanisms [[41](#hamidi), [42](#hamidik)]. Despite the ambiguity in some images, the model performs efficiently, demonstrating its resilience to noise and errors. The multi-layer architecture enables learning complex patterns and features crucial for facial expression identification, even under challenging conditions. To ensure further robustness and system reliability, the model leverages redundancy in the form of ensemble learning. This involves training multiple CNNs with slightly different architectures or initializations. During inference, the model combines predictions from these ensembles, making it less susceptible to individual model failures and improving overall accuracy and fault tolerance [[43](#hamidil), [44](#hamidim)]. Additionally, the model utilizes check pointing at regular intervals during training. This allows for recovery from unexpected hardware or software failures, minimizing training time loss and ensuring system stability. The implemented fault tolerance mechanisms, combined with the inherent robustness of the CNN architecture, contribute to the model's suitability for real world applications, especially those requiring high reliability and resilience [[45](#nilchi)].

**Table 2.4**: The most common ML algorithms used in emotions recognition

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Authors** | **Classifiers** | **Methods** | **Emotions** | **Accuracy (in %)** |
| Umer et al. [[29](#umer)], | CNN | Facial  Expression  Emotion  Recognition | Happiness,  sadness, fear,  disgust, surprise,  anger, and  neutral | CNN algorithm  (77.8%) with  KDEF dataset,  (87.2%) with  GENKI dataset,  (92.8%) with CK+  dataset |
| Raman et al.  [[35](#raman)] | Gradient  boosting  classifier  (GBR), and  ridge classifier  (RC) | Facial  expression,  body  gestures,  and  postures | Happy,  angry, disagree,  disgust, fear,  hello, namaste,  okay, sad, shock,  surprise, and  victorious | Gradient  boosting  classifier  achieved an  average  accuracy  of 96% and  ridge classifier  achieved an  average  accuracy  of 71% |
| Ilyas  et al. [[34](#ilyas)], | CNN | Facial  expression,  body  gestures,  and  postures | Happy, sadness,  anger, and fear | 77.7% for facial  expression,  using CNN |
| Chowdary  et al.[[30](#chowdary)] | CNN, SVM | Facial  Expression  Emotion  Recognition | Happiness,  sadness, fear,  disgust, surprise,  anger, and  neutral | (91%)  using the CK+  database using  SVM and CNN  Classifiers |
| Acheampong  et al. [[31](#achea)], | SVM and KNN | Text emotion  recognition | Joy, happiness,  sadness, fear,  anger, surprise,  disgust, neutral,  fun, worry, love,  hate, relief and  scared | KNN achieved  an average  accuracy of  83%, and SVM  gave an average  accuracy of 77% |
| Nandwani and  Verma [[32](#nandwani)]. | SVM,  RF, CNN | Text emotion  recognition | Furious,  cheerful, or  depressed,  and neutral | RF with an  accuracy of  95.6%, SVM  classifier  achieved  85.47% CNN |
| Mittal et al.  [[33](#mittal)], | CNN, LSTM | Facial  expression,  body  gestures,  and  postures | Angry, happy,  neutral, sad,  disgust, fear,  surprise | The authors in  Mittal et al. [21]  used two  datasets:  IEMOCAP  which achieved  an average  accuracy of  78.2% using  LSTM, CNN,  and  CMUMOSEI  Dataset which  achieved  average  accuracy of  85.0%. |

## CHAPTER 3

## System Design

**3.1 Overview**

The proposed Emotion Recognition Model by Deep Face is a Machine learning-based system designed model to recognize human emotions from facial expressions. The system consists of the following components:

**3.2 Components:**

**3.2.1 Hardware Components**

1. Camera: A camera to capture images.

2. PC/Laptop: A PC/Laptop to run the model.

**3.2.2 Software Components**

1. Code Editor: Visual Studio Code is used to run code.

2. Python: A programming language to run OpenCV.

3. OpenCV: A computer vision library for image capture and processing.

4. Tkinter: The Built-in Python library used to create GUIs.

5. Canvas: Used to display images and graphics inside the Tkinter window.

6. Numpy: A library for numerical computing, used to handle image arrays.

7. PIL (Pillow): Used for handling images and converting them for Tkinter.

8. DeepFace: A deep-learning-based facial analysis library used to detect emotions.

**3.3 System Workflow**

1. Load Model: Load a pre-trained Haar cascade model to detect faces in video stream.

2. Window Creation: Use Tkinter to create main window with title “Emotion Detection”.

3. Canvas Creation: Create Canvas widget to display the video feed.

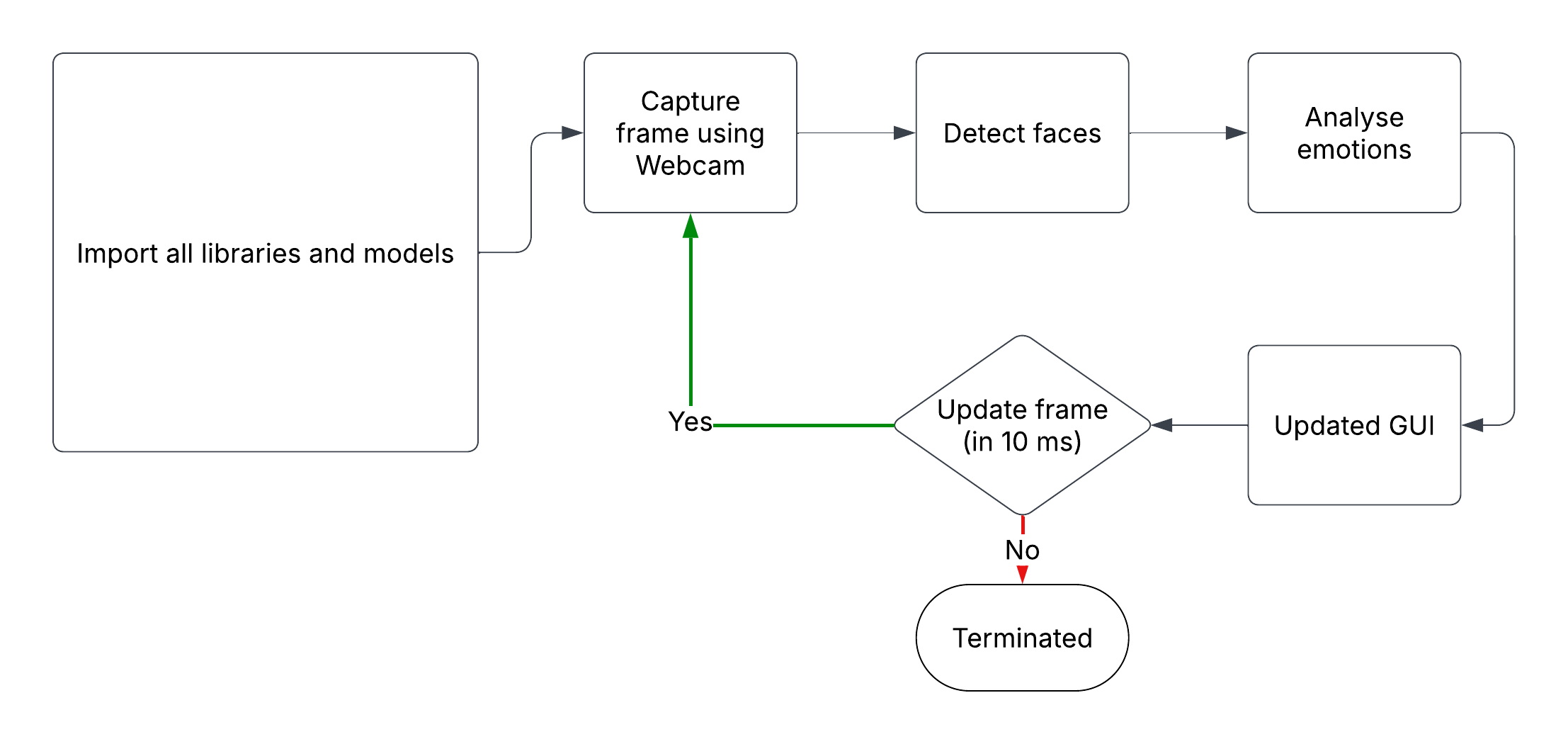
4. Image Capture: Capture an image using the camera.

5. OpenCV Initialization: Initialize OpenCV and set it up to capture images from the camera.

6. Image Processing: Use OpenCV to process the captured image using pre-trained Haar cascade model.

7. Function Creation: Create various function to resize frames, update emotion labels.

8. Update frame: Update the frame at every 10 ms.



**Fig 3.1:** Workflow of the system (Emotion Recognition System)

**3.4 System Requirements**

1. Operating System: Windows 10/11

2. Programming Language: Python 10.

3. OpenCV Version: OpenCV 4.x or higher.

4. Numpy Version: 1.x or higher

5. Tkinter Version: 6.x or higher.

6. PIL Version: pillow with 11.x or higher.

7. Deepface Version: 0.0.93

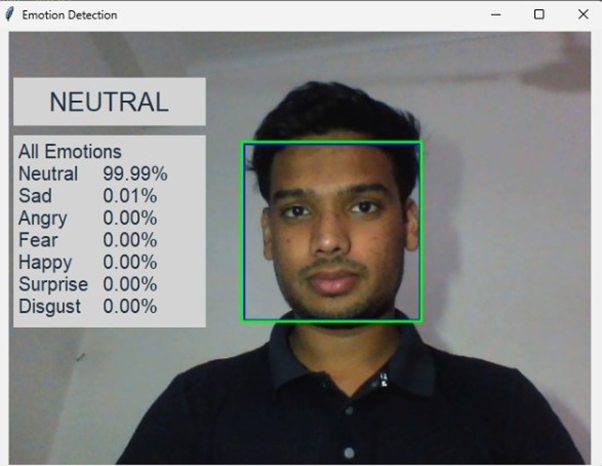
8. Camera Resolution: 640x480 or higher, to detect clear face.

## CHAPTER -4

**Methodology and Technology**

**4.1 Methodology**

This review is used to identify the emotions of human beings that validate the person to find whether the displayed image is neutral, happy, angry etc. Also, it helps to display the changes in the expression of an individual in real time using ML algorithms and deep learning approaches, are used to identify human’s emotions. The system initiates by searching for the human face using Haar cascades classifier which helps to identify the human face and later that used for identifying the which emotion is dominant out of all emotions as shown in fig 4.1. The live image evolves through the deep face algorithm; it identifies the face and finds the facial features. In fig 4.1, there two rectangular frames around the face, the blue one indicates the availability of any face detected in the current frame by using Haar cascade and the green rectangle is to indicate which face is being analyzed for emotion detection. It helps the user quickly identify the primary subject of concern in the live video feed.



**Fig 4.1**: Live input image of the person from the camera.

**4.1.1: Elements used in the Proposed system**

Various libraries are used as components to detect faces and identify a person's emotions. A GUI framework, such as Tkinter, is used to create an interactive window that displays the video feed and emotion results. Face recognition and detection from the video frames are carried out using OpenCV.

The core functionality of the system is face detection, for which the system utilizes the Haar cascade classifier to detect faces in the video feed. It processes the features after converting them into grayscale. For emotion analysis, the system captures frames, analyses emotions using DeepFace, and provides emotion percentages, identifying the dominant emotion, as shown in Fig. 4.1.

To identify facial features, the image captured from the default webcam and is processed using OpenCV and converted into a NumPy array. The system utilizes the OpenCV Haar cascade classifier to detect faces in the video frame. The detected face regions then figured out using DeepFace, which extracts emotional attributes and gives the output as an outcome which emotion is dominant

DeepFace is a lightweight facial recognition framework designed for analyzing human emotions, gender, age, and race but in this we used for identifying emotions. It integrates multiple deep learning models to provide high-accuracy facial emotion detection. DeepFace used deep Convolutional Neural Network (CNN) to process facial images, generating a high-dimensional feature vector representing the face. This feature matrix is then used for many analysis tasks, such as emotion classification and facial attribute recognition.

For emotion analysis, DeepFace processes the detected face and returns a result containing emotions with their respective confidence scores percentagewise in the GUI. The model identifies the dominant emotion based on the highest percentage. The system updates the GUI in real-time by displaying the detected emotions and highlighting the most prominent one and others are in decreasing order.

Additionally, OpenCV is used to draw bounding boxes over detected faces as shown in fig 4.1, enhancing visualization. The live video displays using OpenCV and Tkinter ensuring an interactive and user-friendly experience. The GUI updates in every 10 milliseconds to provide real-time emotion detection results and give the last emotion as the output in terminal when the program is terminated.

**4.1.2 Pseudocode for Human Emotion Feature Recognition Using DeepFace**

The following code outlines for the Emotion detection system using DeepFace:

1. Load the. pre-trained Haar cascade classifier, a face detection model

*face\_cascade=cv2.CascadeClassifier(cv2.data.haarcascades+ 'haarcascade\_frontalface\_default.xml')*

1. Create a Tkinter Window

*root = tk.Tk()*

*root.title("Emotion Detection")*

*root.geometry("800x600") # Start with a window size of 800x600*

1. Create Labels for Displaying Emotion Data

*d\_label\_main = tk.Label(root,*

*text="Main Emotion",*

*font=('Arial', 18),*

*bg='#d3d3d3',*

*fg='#203040',*

*anchor="center",*

*bd=2,*

*highlightbackground="blue",*

*highlightthickness=2)*

*d\_label\_main.place(x=20, y=50, width=200, height=50)*

Similarly, we create *d\_label\_all* displays all detected emotions with their percentages.

1. Function to Update Emotion Labels

*def update\_details(main\_emotion, emotion\_data):*

1. Uses DeepFace to analyze emotions and extracts the dominant emotion.

*try:*

*analysis = DeepFace.analyze(frame, actions=['emotion'],*

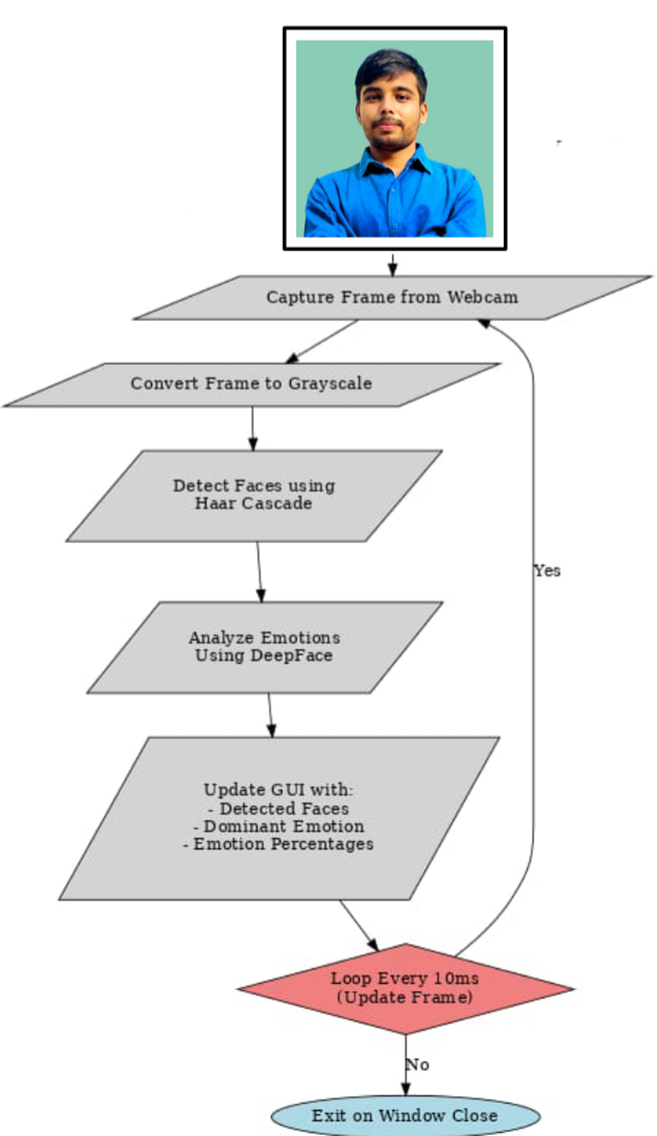
*enforce\_detection=False)*

*emotions = analysis[0]['emotion']*

*d\_emotion = analysis[0]['dominant\_emotion']*

1. Calls update\_frame() every 10ms.

*canvas.after(10, update\_frame)*



**Fig 4.2:** The proposed model for detecting human emotions

**4.2 Technology used**

Python is a high-level, interpreted programming language known for its simplicity and readability. It was created by Guido van Rossum and released in 1991. Python emphasizes code readability and allows developers to write programs in fewer lines compared to other languages. Python is easy to learn but powerful enough for serious applications, due to versatile nature of python various python libraries are most common used in machine learning projects, similarly in this project various different ML algorithms are used and the description are below.

**4.2.1 Libraries used:** -

1.Tkinter (Graphical User Interface Library):

Tkinter is the standard GUI toolkit shipped with Python. It provides tools like windows, buttons, labels, and canvases to create desktop applications easily without needing to install anything extra.

It is used here in:

* Creates the main application window (root).
* Displays the live video feed on a Canvas.
* Adds Label widgets to show the dominant emotion and the full emotion analysis.
* Handles the update and refresh of GUI elements.

2. OpenCV (cv2) (Computer Vision Library):

OpenCV (Open-Source Computer Vision Library) is a huge collection of algorithms for computer vision tasks like object detection, face recognition, video capturing, and image manipulation and In this project the Haar cascade classifiers plays an important role & detects the face. there two rectangular frames around the face, the blue one indicates the availability of any face detected in the current frame by using Haar cascade and the green rectangle is to indicate which face is being analyzed for emotion detection. It helps the user quickly identify the primary subject of concern in the live video feed. To identify facial features, the image captured from the default webcam and is processed using OpenCV and converted into a NumPy array. The system utilizes the OpenCV Haar cascade classifier to detect faces in the video frame.

How it's used here:

* Captures live video from the webcam (cv2.VideoCapture(0)).
* Processes frames (like converting them to grayscale).
* Detects faces using Haar Cascade Classifier (a pre-trained face detection model).
* Draws rectangles around detected faces to visualize detection.
* Prepares video frames for further emotion analysis.

3. NumPy (Numerical Computing Library):

NumPy is a core library for array and matrix operations in Python. It also provides functions for performing mathematical operations on arrays. NumPy (Numerical Python) is a powerful open-source library for numerical and scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a collection of high-level mathematical functions to operate on these arrays efficiently. NumPy is widely used for data analysis, machine learning, and scientific applications due to its speed and flexibility. It also serves as the foundation for many other libraries such as pandas, TensorFlow, and scikit-learn.

How it's used here:

* Images captured by OpenCV are NumPy arrays (pixel data in matrix form).
* Operations like resizing frames, accessing pixel values, and frame manipulations all involve NumPy arrays under the hood.
* Helps in efficient data handling without manual looping through pixels.

4. Pillow (PIL) (Python Imaging Library):

Pillow is an improved version of the old PIL library. It provides **image processing capabilities** like opening, editing, converting, and displaying images. is a powerful Python library used for opening, manipulating, and saving image files. It supports a wide range of image formats like JPEG, PNG, BMP, and GIF, and provides functionalities such as cropping, resizing, filtering, drawing, and image enhancement, making it a popular tool in image processing and computer vision tasks.

It is used here:

* Converts frames from OpenCV (which are in BGR color format) to **RGB** format (required by Tkinter).
* Converts NumPy arrays into Image objects.
* Converts Image objects to a format Tkinter can render.
* Makes it possible to embed live video frames into the Tkinter GUI.

5. Deep Face (Facial Analysis and Emotion Recognition): To identify facial features, the image captured from the default webcam and is processed using OpenCV and converted into a NumPy array. The system utilizes the OpenCV Haar cascade classifier to detect faces in the video frame. The detected face regions then figured out using DeepFace, which extracts emotional attributes and gives the output as an outcome which emotion is dominant

DeepFace is a lightweight facial recognition framework designed for analysing human emotions, gender, age, and race but in this we used for identifying emotions. It integrates multiple deep learning models to provide high-accuracy facial emotion detection. DeepFace used deep Convolutional Neural Network (CNN) to process facial images, generating a high-dimensional feature vector representing the face. This feature matrix is then used for many analysis tasks, such as emotion classification and facial attribute recognition.

For emotion analysis, DeepFace processes the detected face and returns a result containing emotions with their respective confidence scores percentagewise in the GUI. The model identifies the dominant emotion based on the highest percentage. The system updates the GUI in real-time by displaying the detected emotions and highlighting the most prominent one and others are in decreasing order.

Deep Face is a Python library built on deep learning models. It can perform:

* + Face verification (is this the same person?)
  + Face recognition (who is this?)
  + Facial attribute analysis (what is the person's emotion, age, gender, race?).
  + Analyses each detected face from webcam frames.
  + Identifies emotions like *happy, sad, angry, fearful, surprised,* etc.
  + Returns both:
    - The dominant emotion (e.g., *happy*).
    - Percentages for each possible emotion (e.g., 75% happy, 10% sad, etc.).
  + Helps update the GUI labels showing real-time emotional status.

## CHAPTER -5

**Implementation and Result Analysis**

**5.1 Implementation:**

**1. Import Required Libraries**

*import tkinter as tk*

*from tkinter import Canvas*

*import cv2*

*import numpy as np*

*from PIL import Image, ImageTk*

*from deepface import DeepFace*

• **tkinter**: The built-in Python library used to create graphical user interfaces (GUIs).

• **Canvas**: Used to display images and graphics inside the Tkinter window.

• **cv2 (OpenCV)**: An open-source computer vision library for image and video

processing.

• **numpy**: A library for numerical computing, used to handle image arrays.

• **PIL (Pillow)**: Used for handling images and converting them for Tkinter.

• **DeepFace**: A deep-learning-based facial analysis library used to detect emotions.

**2. Load Face Detection Model**

*face\_cascade = cv2.CascadeClassifier(cv2.data.haarcascades +*

*'haarcascade\_frontalface\_default.xml')*

• This loads a **pre-trained Haar cascade classifier** from OpenCV to detect faces in an

image.

• The model is used to detect faces in the video stream.

**3. Create a Tkinter Window**

*root = tk.Tk()*

*root.title("Emotion Detection")*

*root.geometry("800x600") # Start with a window size of 800x600*

• Creates the main window (root) using Tkinter.

• Sets the window title as **"Emotion Detection"**.

• The geometry("800x600") specifies the initial window size (800 pixels wide and

600 pixels high).

**4. Create a Canvas for Displaying Video**

*canvas = Canvas(root, width=800, height=600)*

*canvas.pack(fill="both", expand=True)*

• The **Canvas widget** is used to display the video feed.

• pack(fill="both", expand=True) ensures the canvas resizes with the window.

**5. Create Labels for Displaying Emotion Data**

*d\_label\_main = tk.Label(root,*

*text="Main Emotion",*

*font=('Arial', 18),*

*bg='#d3d3d3',*

*fg='#203040',*

*anchor="center",*

*bd=2,*

*highlightbackground="blue",*

*highlightthickness=2)*

*d\_label\_main.place(x=20, y=50, width=200, height=50)*

• d\_label\_main is a label to display the **dominant emotion** detected in the face.

• font=('Arial', 18): Sets the text font.

• bg='#d3d3d3': Sets background color.

• fg='#203040': Sets text color.

• anchor="center": Centers the text.

• bd=2: Sets a border width.

• highlightbackground="blue": Sets the border color.

• place(x=20, y=50, width=200, height=50): Positions the label at (20, 50) and

gives it a size of (200x50) pixels.

*d\_label\_all = tk.Label(root,*

*text="All Emotions",*

*font=('Arial', 15),*

*bg='#d3d3d3',*

*fg='#203040',*

*anchor="nw",*

*justify="left",*

*bd=2,*

*highlightbackground="green",*

*highlightthickness=2)*

*d\_label\_all.place(x=20, y=110, width=200, height=200)*

• d\_label\_all displays **all detected emotions** with their percentages.

**6. Open Webcam**

*cap = cv2.VideoCapture(0)*

• Opens the **default webcam** (0 means the primary camera).

• cap.read() will be used to capture frames from the webcam.

**7. Define Aspect Ratio**

*ASPECT\_RATIO = 16 / 9*

• Defines a **16:9 aspect ratio** to maintain video proportions.

**8. Function to Resize Frames**

*def resize\_frame(frame, max\_width, max\_height):*

*frame\_height, frame\_width = frame.shape[:2]*

*frame\_aspect\_ratio = frame\_width / frame\_height*

*if frame\_aspect\_ratio > ASPECT\_RATIO:*

*new\_width = max\_width*

*new\_height = int(new\_width / frame\_aspect\_ratio)*

*else:*

*new\_height = max\_height*

*new\_width = int(new\_height frame\_aspect\_ratio)*

*return cv2.resize(frame, (new\_width, new\_height))*

• Resizes the video frame while maintaining aspect ratio.

• Prevents image distortion when displaying it in the GUI.

**9. Function to Update Emotion Labels**

*def update\_details(main\_emotion, emotion\_data):*

*d\_label\_main.config(text=main\_emotion.upper(), font=('Arial', 20),*

*bg='#d3d3d3', fg='#203040')*

*sorted\_emotions = sorted(emotion\_data.items(), key=lambda x: x[1],*

*reverse=True)*

*emotions\_text = "All Emotions\n"*

*for emotion, percentage in sorted\_emotions:*

*emotions\_text += f"{emotion.capitalize()}\t{percentage:.2f}%\n"*

*d\_label\_all.config(text=emotions\_text)*

• Updates d\_label\_main with the **dominant emotion**.

• Sorts emotions by percentage and updates d\_label\_all to display all detected

Emotions.

**10. Function to Process Webcam Frames**

*def update\_frame():*

*ret, frame = cap.read()*

*if not ret:*

*Return*

• Captures a frame from the webcam.

• ret is True if a frame was successfully captured.

*gray = cv2.cvtColor(frame, cv2.COLOR\_BGR2GRAY)*

*faces = face\_cascade.detectMultiScale(gray, 1.1, 4)*

• Converts the frame to **grayscale**.

• Detects faces in the frame using OpenCV's Haar cascade model.

for (x, y, w, h) in faces:

*cv2.rectangle(frame, (x, y), (x + w, y + h), (255, 0, 0), 2)*

• Draws **blue rectangles** around detected faces.

*try:*

*analysis = DeepFace.analyze(frame, actions=['emotion'],*

*enforce\_detection=False)*

*emotions = analysis[0]['emotion']*

*d\_emotion = analysis[0]['dominant\_emotion']*

• Uses **DeepFace** to analyze emotions.

• Extracts the **dominant emotion**.

*dominant\_face\_region = analysis[0]['region']*

*x, y, w, h = dominant\_face\_region['x'], dominant\_face\_region['y'],*

*dominant\_face\_region['w'], dominant\_face\_region['h']*

*cv2.rectangle(frame, (x, y), (x + w, y + h), (0, 255, 0), 2)*

• Draws a **green rectangle** around the dominant face.

*except Exception as e:*

*print(f" DeepFace Error: {e}")*

*emotions = "Natural"*

• Handles errors if DeepFace fails.

*update\_details(d\_emotion, emotions)*

• Updates emotion labels.

*canvas\_width = canvas.winfo\_width()*

*canvas\_height = canvas.winfo\_height()*

• Gets the **current** canvas size.

*if canvas\_width > 0 and canvas\_height > 0:*

*frame\_resized = resize\_frame(frame, canvas\_width, canvas\_height)*

*else:*

*frame\_resized = cv2.resize(frame, (800, 600))*

• Resizes the frame to **fit the canvas**.

*frame\_resized = cv2.cvtColor(frame\_resized, cv2.COLOR\_BGR2RGB)*

• Converts OpenCV's **BGR format** to **RGB** (Tkinter uses RGB).

*img = Image.fromarray(frame\_resized)*

*imgtk = ImageTk.PhotoImage(image=img)*

*canvas.create\_image(0, 0, anchor=tk.NW, image=imgtk)*

canvas.imgtk = imgtk

• Converts the frame to a format Tkinter can display.

*canvas.after(10, update\_frame)*

• Calls update\_frame() every **10ms**.

**11. Start the Application**

*root.after(100, update\_frame)*

*root.mainloop()*

• Starts the **Tkinter event loop**.

**5.2 Result Analysis**

After implementing the model successfully, the following results are observed with various different emotions : a. neutral, b. happy, c. sad, d. angry, e. surprise, f. fear

|  |  |
| --- | --- |
| a. neutral | b. happy |
| c. sad | d. angry |
| e. surprise | f. fear |

**Fig 5.1 :** Live Capture Imagesincluding **a. Neutral b. happy c. Sad d. angry e. surprise f. Fear**

The emotion recognition system developed in this project is capable of detecting and classifying facial expressions in real time. The interface showcases predictions across seven primary emotions: Neutral, Happy, Sad, Angry, Fear, Surprise, and Disgust. As demonstrated in the figure, the model analyzes facial features using a deep learning classifier and returns the most dominant emotion with its associated confidence score in descending order.

Each frame represents a distinct test case for a specific emotion, including examples of correctly predicted expressions such as Neutral (99.99%), Happy (100%), Sad (78.55%), Angry (67.63%), Surprise (99.72%), and Fear (99.88%). The system uses a bounding box to highlight the face and a classification panel displaying probabilities for all emotion categories. The human’s face is captured from the live camera is shown in Figure 5.1. with various expressions and is classified accurately. These results confirm the model's capability to differentiate subtle emotional expressions with high accuracy, supporting its potential for real-world applications such as virtual assistants, mental health monitoring, and smart classrooms.

The live image evolves through the deep face algorithm; it identifies the face and finds the facial features. In fig 5.1, there two rectangular frames around the face, the blue one indicates the availability of any face detected in the current frame by using Haar cascade and the green rectangle is to indicate which face is being analyzed for emotion detection. It helps the user quickly identify the primary subject of concern in the live video feed.

To identify facial features, the image captured from the default webcam and is processed using OpenCV and converted into a NumPy array. The system utilizes the OpenCV Haar cascade classifier to detect faces in the video frame. The detected face regions then figured out using DeepFace, which extracts emotional attributes and gives the output as an outcome which emotion is dominant

DeepFace is a lightweight facial recognition framework designed for analyzing human emotions, gender, age, and race but in this we used for identifying emotions. It integrates multiple deep learning models to provide high-accuracy facial emotion detection. DeepFace used deep Convolutional Neural Network (CNN) to process facial images, generating a high-dimensional feature vector representing the face. This feature matrix is then used for many analysis tasks, such as emotion classification and facial attribute recognition.

For emotion analysis, DeepFace processes the detected face and returns a result containing emotions with their respective confidence scores percentagewise in the GUI. The model identifies the dominant emotion based on the highest percentage. The system updates the GUI in real-time by displaying the detected emotions and highlighting the most prominent one and others are in decreasing order.

Additionally, OpenCV is used to draw bounding boxes over detected faces as shown in fig 5.1, enhancing visualization. The live video displays using OpenCV and Tkinter ensuring an interactive and user-friendly experience. The GUI updates in every 10 milliseconds to provide real-time emotion detection results and give the last emotion as the output in terminal when the program is terminated.

**Fig 5.2: Comparison of result with various existing model**

Figure 5.2 presents a comparative evaluation of several emotion recognition models in terms of their classification accuracy. The five models included in this comparison are: the Designed Model (proposed in this work), Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel, Naïve Bayes, K-Nearest Neighbors (KNN), and a hybrid Convolutional Neural Network (CNN) combined with ResNet. This figure visually highlights the superiority of the Designed Model over traditional and deep learning-based approaches.

The Designed Model, developed using the DeepFace framework, achieves the highest accuracy among all compared methods. DeepFace is a deep learning architecture created originally for facial recognition but is highly adaptable for tasks such as facial emotion recognition due to its powerful feature extraction capabilities. It uses a deep convolutional neural network (CNN) that maps facial features into a compact embedding space, where even subtle differences in expression can be captured effectively. This rich feature representation enables the Designed Model to accurately classify a wide range of emotions, including subtle or mixed expressions that are often challenging for conventional methods.

The superior performance of the Designed Model can be attributed to its end-to-end architecture that includes face detection, alignment, representation learning, and emotion classification in a single pipeline. This integrated approach reduces errors introduced at different stages of manual feature engineering and model selection, which is a limitation in traditional models like SVM, Naïve Bayes, and KNN. Furthermore, DeepFace benefits from pretraining on large-scale datasets, allowing the Designed Model to generalize well even when fine-tuned on smaller emotion-specific datasets.

In contrast, the SVM with RBF kernel, though a strong classifier in many pattern recognition problems, depends heavily on the quality of the input features. In emotion recognition, where facial features can be complex and high-dimensional, handcrafted features may not capture the required depth and nuances. This limits the SVM’s ability to differentiate between emotions with overlapping visual characteristics.

Similarly, the Naïve Bayes model makes strong assumptions about feature independence, which rarely hold true in the case of facial data. This results in suboptimal performance as it cannot adequately model the dependencies between facial landmarks or pixel intensities. KNN, while simple and effective in low-dimensional spaces, struggles with high-dimensional facial data and is sensitive to noise, which further affects its accuracy.

The CNN+ResNet model performs significantly better than traditional algorithms and is close to the performance of the Designed Model. ResNet’s residual connections allow for deeper networks without the vanishing gradient problem, making it well-suited for complex tasks like emotion recognition. However, the marginally lower accuracy compared to the Designed Model suggests that while CNN+ResNet is powerful, the specific optimization and pre-trained embeddings used in DeepFace provide a notable edge.

In summary, the Designed Model’s use of DeepFace demonstrates the advantages of leveraging advanced deep learning frameworks tailored for facial analysis. It outperforms both traditional classifiers and other deep learning approaches, proving to be a highly effective solution for accurate and robust emotion recognition.

**CHAPTER -6**

**Conclusions and Future Work**

**6.1 Conclusion**

This report paper gave the idea of advance field of facial emotion recognition (FER) taking support from affective computing, machine learning, and computer vision tasks. The main aim is to look how new technologies helped machine to understand human emotions better and faster. Since facial expressions, a powerful and widely understood way of communicating without words, using them in human-computer interaction is an important move toward creating smart and advanced systems.

The review pointed out that traditional human-computer interaction (HCI) system doesn’t often consider user’s emotions which leads to lower satisfaction and engagement. By adding facial emotions recognition (FER) features, systems can now respond to how users feel, making interactions more natural, responsive, and effective. Deep learning techniques like Convolutional Neural Networks (CNN) and tools like Deep Face have proven to be very accurate at recognizing emotions by analyzing facial expressions. In other words we can also conclude that The Emotion Recognition System developed in this project successfully integrates computer vision and deep learning techniques to identify and classify human emotions in real-time through facial expressions. By using tools like OpenCV for face detection and DeepFace for emotion analysis, the system demonstrates a practical application of machine learning in understanding human affective states. The user interface, designed with Tkinter, makes the system interactive and accessible, while performance is enhanced through efficient frame handling and real-time updates.

This project illustrates the potential of artificial intelligence in bridging the gap between human emotions and machine interpretation. It highlights the growing significance of emotion-aware systems in domains such as healthcare, education, human-computer interaction, and smart surveillance. While the system currently recognizes basic emotions accurately, future advancements—such as multimodal input integration, more robust models, and deployment in real-world settings—can make it even more effective and impactful.

In fact, some experiments mentioned in the review showed accuracy levels over 95%, highlighting how reliable these methods are for real-world use. We looked at several well-known models and datasets, such as FER-2013, Face Net, CK+ and JAFFE. These have helped train and test strong FER systems using Deep Face. New methods, like using CNNs in multiple stages or combining different deep learning techniques, pushed performance even further. These updates make it possible to detect emotions in real-time with low delay and high confidence and accuracy. This allows to seek the opportunities in areas like education, healthcare, customer service, and mental health support. In the end, the review shows how emotion-aware technology can help remove the emotional gap between human and machines. By giving systems emotional intelligence by using latest technologies, we can create digital experiences that are more human, personalized, and effective & increase human machine interaction.

**6.2 Future work**

Future work should focus on better emotion detection across different cultures, improving real-time accuracy, and making sure system maintain the confidentiality of emotional data. The current implementation of the Emotion Recognition System demonstrates effective real-time facial emotion detection using deep learning techniques and a user-friendly GUI. However, there are several areas that offer potential for enhancement and future development:

6.2.1. Multimodal Emotion Recognition

While this project focuses solely on facial expressions, emotions are also conveyed through voice, gestures, and physiological signals (such as heart rate or pupil dilation). Future iterations can integrate audio analysis and biometric sensors to develop a multimodal system that improves emotion detection accuracy in complex or ambiguous scenarios.

6.2.2. Advanced Deep Learning Architectures

The system currently utilizes traditional deep learning models. Future work could explore transformer-based models like Vision Transformers (ViT) or Swin Transformers, which have shown superior performance in visual recognition tasks. These models may help in better understanding subtle facial expressions and improving classification across a wider range of emotional states.

6.2.3. Robustness in Real-World Conditions

Currently, the system performs best under controlled lighting and clear facial visibility. Enhancing the model’s ability to operate in low-light environments, with partial occlusions (like face masks or glasses), and across varied camera angles would improve its usability in real-world applications such as surveillance or in-car monitoring systems.

6.2.4. Deployment on Edge Devices

Optimizing the model for resource-constrained devices (such as mobile phones, Raspberry Pi, or embedded systems) will enable widespread, real-time usage without needing powerful hardware. This requires compression techniques such as model quantization or knowledge distillation.

6.2.5. Emotion-Responsive Systems

Another interesting direction is to develop emotion-aware interactive systems, where the application responds based on the detected emotion. For instance, in educational platforms, a system could adapt the difficulty of content based on the learner’s mood, or a customer service bot could provide empathetic replies during user frustration.

6.2.6. Dataset Diversity and Model Fairness

Most emotion datasets are biased toward specific age groups, ethnicities, or cultural expressions. Expanding the training dataset with more diverse and inclusive samples will make the system more generalizable and reduce bias in emotion interpretation across global users.

This forward-looking approach can help evolve the current system into a more intelligent, adaptive, and human-aware tool, suitable for applications in healthcare, education, entertainment, and smart interfaces.

**CHAPTER -7**

Progress Schedule Semester Wise

**Table 7.1**: Semester Wise Progress Report

|  |  |  |
| --- | --- | --- |
| **Month, Year** | **Task done** | **Semester** |
| Aug, 2024 | Selected the project topic after evaluating multiple ideas for feasibility and relevance. | 7th |
| Sep, 2024 | Conducted detailed research on facial emotion recognition techniques and evaluated different machine learning approaches suitable for implementation. | 7th |
| Oct, 2024 | Defined the problem statement, objectives, and proposed methodology. Planned the overall project structure and module distribution. | 7th |
| Nov, 2024 | Developed the first module: Face Capturing – to accept real-time input through the webcam for emotion analysis. | 7th |
| Dec, 2024 | Implemented two modules: 1. Pre-processing Module**:** Normalized and enhanced input images for consistency. 2. Training Module**:** Applied ML algorithms to learn and classify emotions based on input features | 7th |
| Jan, 2025 | Developed theEmotion Recognition Module, where the system detected and predicted dominant emotions using facial features. | 7th |
| Feb, 2025 | Focused on improving the model performance and accuracy. Fine-tuned parameters and optimized the interface. | 8th |
|  |  |  |
|  |  |  |
| Mar, 2025 | Conducted testing under various conditions. Compared results with previous models and documented the limitations and improvements. | 8th |
| April, 2025 | Finalized the system, Documented the work and highlighted potential future enhancements and applicants | 8th |

The project was carried out progressively over two semesters. In the **7th semester**, initial efforts (August–September 2024) were dedicated to selecting a suitable project topic and conducting thorough research on various methodologies. In **October,** the project objectives, problem statement, and proposed solution were clearly defined. By **November,** the first module—Face Capturing—was implemented, which allowed the system to accept input for facial emotion recognition. In **December**, the focus shifted to building the pre-processing module to normalize inputs, and the training module where machine learning algorithms were tested for accurate emotion classification. **January 2025** marked the development of the emotion recognition module that detected and displayed dominant emotions from facial inputs. Moving into the **8th semester**, the project was enhanced for better accuracy and performance in **February**, followed by testing and evaluation of the system in **March**. Finally, in **April**, the project was completed with documentation, presentation, and future scope outline

## References

1. Neha S, Nivya, Pooja HC Shekar, Keerthana S Kumar AshaVG (2020) Emotion Recognition and Depression Detection using Deep Learning 2020 vol. 07, no. 08, pp. 2395 2472
2. Sreenidhi M J, Surabi Sri Dhanya S, Sahithi R (2020) Human Emotion Recognition System using Deep Learning Techniques 2020 vol. 07, no. 08, 2278-0181

1. Ninad Mehendale (2020) Facial emotion recognition using convolutional neural networks (FERC) 2020 https://doi.org/10.1007/s42452-020-2234-1
2. Runfang Guo, Hongfei Guo, Liwen Wang, Mengmeng Chen, Dong Yang, Bin Li1 (2024) Development and application of emotion recognition technology 2024 https://doi.org/10.1186/s40359 024-01581-4
3. Alisha Fida, Muhammad Umer, Oumaima Saidani, Monia Hamdi, Khaled Alnowaiser, Carmen Bisogni, Andrea F. Abate, Imran Ashraf (2023) Real time emotions recognition through facial expressions 2023 /10.1007/s11042-023-16722-x
4. Akhilesh Kumar, Awadhesh Kumar, Sumit Gupta, Machine Learning-Driven Emotion Recognition Through Facial Landmark Analysis 2025 /10.1007/s42979-025-03688-w

1. Barbra Zupan, Michelle Eskritt (2023) Facial and Vocal Emotion Recognition in Adolescence 2023 /10.1007/s40894-023-00219-7
2. Essam H. Houssein, Asmaa Hammad, Abdelmgeid A. Ali (2022) Human emotion recognition from EEG-based brain computer interface using machine learning 2022 /10.1007/s00521-022 07292-4
3. Eman M. G. Younis, Someya Mohsen, Essam H. Houssein, Osman Ali Sadek Ibrahim (2024) Machine learning for human emotion recognition 2024 /10.1007/s00521-024-09426-2
4. RK Kapila Vani, P.Jayashree (2025) Platform 2025 /10.1007/s10639-024-13279-6

1. Sowmya BJ, Meeradevi, Sini Anna Alex, Anita Kanavalli, Supreeth S, Shruthi G, Rohith S (2023) Machine learning model for emotion detection and recognition using an enhanced Convolutional Neural Network.
2. Ramachandran Venkatesan, Sundarsingh Shirly, Mariappan Selvarathi and Theena Jemima Jebaseeli (2023) Human Emotion Detection Using DeepFace and Artificial Intelligence 2023.
3. Vishal singh bhati , Namita Tiwari, Meenu Chawla (2025) A Generalized Zero-Shot Deep Learning Classifier for Emotion Recognition Using Facial Expression Images 2025.
4. Manisha Kasar, Bindu Garg, Trupati Suryawanshi (2024) EmoSense: Pioneering Facial Emotion Recognition with Precision Through Model Optimization and Face Emotion Constraints 2024
5. Mingyi Zhao, Linrui Gong, Abdul Sattar Din (2024) A review of the emotion recognitionmodel of robots 2024
6. Alojaiman, B. (2021). Toward selection of trustworthy and efficient E-Learning platform. IEEE Access, 9, 133889–133901, /10. 1109/ access. 2021. 31141 50
7. Ariza, J., Jimeno, M., Villanueva-Polanco, R., & Capacho, J. (2021). Provisioning computational resources for cloud-based e-learning platforms using deep learning techniques. IEEE Access, 9,

89798–89811. /10. 1109/ access. 2021. 30903 66

1. Lyu, L., Zhang, Y., Chi, M. Y., Yang, F., Zhang, S. G., Liu, P., & Lu, W. G. (2022). Spontaneous facial expression database of learners’ academic emotions in online learning with hand occlusion. Computers and Electrical Engineering, 97, 107667./10. 1016/j.compeleceng. 2021. 107667
2. Sezer A, Altan A. Detection of solder paste defects with an optimization‐based deep learning model using image processing techniques. Soldering & Surface Mount Technology. 2021;33(5):291-8.
3. A. Dangi, A. Mittal, and D. S. Kushwaha, ARO-RVFLNN: Artificial Rabbits Optimized Robust Random Vector Functional Link Network for sentiment classification, Expert Systems with Applications, 2023.
4. M. Saleem, T. Anwar, and M. Usama, DeSCCNet: Spectro-temporal feature representation for speech emotion recognition, Applied Acoustics, vol. 215, 2023.
5. M. Irfan, M. A. Javed, A. M. Qamar, and F. Hussain, Triplet-Loss-Based Deep CNN Features with SVM for Facial Emotion Recognition, IEEE Access, vol. 11, pp. 43025–43036, 2023
6. A. Savchenko, Multi-task EfficientNet-B2 with MTCNN and Adam optimization for facial emotion recognition, Signal Processing: Image Communication, vol. 118, 2023.
7. S. Zhang, RDA-MTE: ResNet-50 + Bidirectional Attention + Transformer Encoder for facial expression recognition,Neural Computing and Applications, vol. 35, pp. 28195–28210, 2023.
8. A. S. Asghar, F. U. Khan, S. Nisar, and N. Shah, Facial emotion recognition using statistical features and machine learning, Multimedia Tools and Applications, vol. 82, pp. 20541–20558, 2023, doi: 10.1007/s11042-023-15049-w.
9. P. Giannopoulos, I. Perikos, and I. Hatzilygeroudis, Deep learning approaches for facial emotion recognition: A case study on FER-2013, in Smart Innovation, Systems and Technologies, vol. 85, pp. 1–16, 2018.
10. W. M. Yassin, M. F. Abdollah, Z. Muslim, R. Ahmad, and A. Ismail, An emotion and gender detection using hybridized convolutional 2D and batch norm residual network learning, in Proc. ACM Int. Conf., 2021, pp. 79–84
11. X. Zhao, X. Shi, and S. Zhang, Facial expression recognition via deep learning, IETE Technical Review, vol. 32, no. 5, pp. 347–355, 2015.
12. Umer S, Rout RK, Pero C, Nappi M (2022) Facial expression recognition with trade-offs between data augmentation and deep learning features. J Ambient Intell Hum Comput 13(2):721–735.
13. Chowdary MK, Nguyen TN, Hemanth DJ (2021) Deep learningbased facial emotion recognition for human–computer interaction applications. Neural Comput Appl 35:1–18
14. Acheampong FA, Wenyu C, Nunoo-Mensah H (2020) Textbased emotion detection: advances, challenges, and opportunities. Eng Rep 2(7):e12189.
15. Nandwani P, Verma R (2021) A review on sentiment analysis and emotion detection from text. Soc Netw Anal Min 11(1):81
16. Mittal T, Bhattacharya U, Chandra R, Bera A, Manocha D (2020) M3er: Multiplicative multimodal emotion recognition using facial, textual, and speech cues. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, pp 1359–1367
17. Ilyas CMA, Nunes R, Nasrollahi K, Rehm M, Moeslund TB (2021) Deep emotion recognition through upper body movements and facial expression. In: VISIGRAPP (5: VISAPP), pp 669–679.
18. Raman S, Patel S, Yadav S, Singh V (2022) Emotion and gesture detection. Int J Res Appl Sci Eng Technol 10:3731–3734

[36] Zheng K, Yang D, Liu J, Cui J. Recognition of teachers’ facial expression intensity based on convolutional neural network and attention mechanism. IEEE access. 2020;8:226437-44.

[37] Yağ İ, Altan A. Artificial intelligence-based robust hybrid algorithm design and implementation for real-time detection of plant diseases in agricultural environments. Biology. 2022;11(12):1732. 10.3390/biology11121732

[38] Singh S, Nasoz F, editors. Facial expression recognition with convolutional neural networks. 2020 10th Annual Computing and Communication Workshop and Conference (CCWC); 2020: IEEE. 10.1109/CCWC47524.2020.9031283

[39] Chen A, Xing H, Wang F. A facial expression recognition method using deep convolutional neural networks based on edge computing. Ieee Access. 2020;8:49741-51.

[40] Mollahosseini A, Chan D, Mahoor MH, editors. Going deeper in facial expression recognition using deep neural networks. 2016 IEEE Winter conference on applications of computer vision (WACV); 2016: IEEE.

[41] Hamidi H, Vafaei A, Monadjemi A. Algorithm based fault tolerant and check pointing for high performance computing systems. Journal of applied sciences. 2009;9(22):3947-56. 10.3923/jas.2009.3947.3956.

[42] Hamidi H, Mohammadi K. Modeling fault tolerant and secure mobile agent execution in distributed systems. International Journal of Intelligent Information Technologies (IJIIT). 2006;2(1):21-36. 10.4018/jiit.2006010102

[43] Nilchi AN, Vafaei A, Hamidi H, editors. Evaluation of security and fault tolerance in mobile agents. 2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN'08); 2008: IEEE. 10.1109/WOCN.2008.4542509

[44] Hamidi H. A combined fuzzy method for evaluating criteria in enterprise resource planning implementation. Intelligent systems: Concepts, methodologies, tools, and applications: IGI Global; 2018. p. 639-70.

[45] Hamidi H, Vafaei A, Monadjemi SAH. Analysis and evaluation of a new algorithm based fault tolerance for computing systems. International Journal of Grid and High Performance Computing (IJGHPC). 2012;4(1):37-51. 10.4018/jghpc.2012010103

Facial Expression Analysis for Emotion Recognition Using Machine Learning

Shubham Mishra  
*Computer Science & Engineering   
ITS Engineering College* Greater Noida, India   
shubhammishradm\_cse21@its.edu.in

Prachi Jain  
*Computer Science & Engineering   
ITS Engineering College* Greater Noida, India   
prachij.cse@its.edu.in Raj Rajput  
*Computer Science & Engineering   
ITS Engineering College*Greater Noida, India   
rajrajputjs\_cse21@its.edu.  
in

Biswa Mohan Sahoo  
Computer Science & Engineering   
ITS Engineering College Greater Noida, India   
biswamohans@gmail.com Shreya Upadhyay  
*Computer Science & Engineering   
ITS Engineering College* Greater Noida, India   
shreyaupadhyayhu\_cse21@its.edu.in

***Abstract* — Facial emotion recognition is an important component for modern machines which respond to human emotions through nonverbal signals. This paper presents a comprehensive review of the latest advancements in facial emotion recognition, focusing on machine learning and deep learning. The focus is on Convolutional Neural Networks (CNNs), KNN, Decision Tree, RNN/LSTM, SVM and others. This review focused the integration of various classifiers like OpenCV, Haar cascade classifiers, Naïve Bayes, for facial feature extraction, continued by emotion recognition using various machine learning algorithm. Through a deep review of literature and implementation techniques that how system give different accuracy with different techniques. The paper highlights how FER systems are becoming more important for improving how people interact with computers in various different areas like healthcare, education, human-computer interaction and virtual assistants. The conclusion from review shows how FER are accurate and work fine, also highlighting how useful FER is for creating smart systems and easy-to-use in digital systems. This paper target to help to compare the various different techniques in the field of FER that can adjust to users, understand their emotions, and how we can also improve the result if want to work on similar areas.**

***Keywords—facial, emotion, expression, detection, real-time, recognition,***

1. Introduction

Emotion recognition which is key feature in affective computing has transformed this means as how people relate with machines. The Emotion Recognition system analyses the emotions of a user using real-time video. This was achieved using Convolutional Neural Network (CNN) [[1](#neha)]. The basic parts of the conventional emotion recognition system were face detection, facial feature extraction and emotions classifications ([[2](#alojaiman)], [[3](#ariza)], [[4](#lyu)])**.** Toprovide a high accuracy of response used hybrid network consisting of CNN and Recurrent Neural Network (RNN) [[5](#sreenidhi)]. Ninad tested Facial Emotion using Convolutional Neural Network with more than 750k images for facial feature vector extraction [[6](#ninad)].

We studied the Theoretical and Conceptual Foundations from Communication Theory and Ekman basic emotions model. Venkatesan emphasized that Emotion detection systems that employ Deep Face mostly rely on the facial expressions because they compose 55% of communication signals [[7](#ramachandran)]. The research of author(s) delivered a thorough breakdown between dimensional and categorical emotion models including Ekman’s six basic emotions [[8](#emam)]. They introduced a system based on the facial expression feedback during the live reading activities. Their methodology for the real-time emotion capture helped us shape our own approach to responsiveness in real-time systems [[9](#alisha)]. They employed a facial landmark trajectory analysis to enhance emotion recognition. Their comparative study of LR, KNN, and LDA classifiers provided experimental result, supporting logistic regression for high accuracy in emotion activation and sentiment scale classification [[10](#akhilesh)]. Inspired by Ninad we adopted CNN-RNN hybrid architectures designed to capture spatial structures and time-based changes for more accurate facial emotion recognition [[6](#ninad)].

There were lots of problems in papers, some similar issues are:

1. *Dataset and challenges*

a. Size and Availability

The majority of datasets are collected in controlled laboratory environments, which limits their applicability in real-world conditions [[8](#emam)].

b. Imbalance and Diversity

The real-world emotion data exhibits class imbalance, which affect the training effectiveness and leads to a biased model [[7](#ramachandran)]. They point out that differences in developmental stages and inconsistencies in task design within adolescent-focused datasets lead to significant variation between studies, making it difficult to establish standardized evaluation criteria [[11](#barbra)].

c. Quality and Labeling Issues

He observed that further argue that some datasets lack richness in environmental and physiological context, which limits their utility for multimodal fusion and model robustness [[8](#emam)].

1. *Accuracy Challenges*

a. Generalizability

Author indicates that background noise, face orientation, and lighting changes can significantly reduce accuracy in facial emotion recognition models using CNNs. His dual-stage CNN method attempts to mitigate this by removing background distractions [[6](#ninad)].

b. Real-Time vs. Offline Accuracy

They contrast real-time versus offline recognition, noting that many high-accuracy models perform poorly under real-time constraints due to latency, processing power, and noisy inputs [[9](#alisha)].

1. *Patterns and Contradictions*

Most papers agree that dataset quality and availability are foundational to model success. There is a shared recognition that controlled lab environments create inflated accuracy estimates compared to real-world scenarios.

Some papers, report high accuracy (94%) even in real-time settings, while others caution that such accuracy may not be reproducible outside idealized settings [[7](#ramachandran)].

LITERATURE SURVEY

Emotion is a varying research field studied across the multiple disciplines such as healthcare, marketing, psychology and robotics. Emotion has a great impact on decision making, physical well-being, reasoning and how we navigate the different aspects of life. Computer Science, Automated Human Emotion Recognition (AHER) are one of the main areas of research [[8](#emam)]. The ability to recognize the emotions plays a crucial role in human interactions and relationship building [[11](#barbra)].

|  |  |  |
| --- | --- | --- |
| Classifiers  Technique | 10-fold cross validation | 80% training 20% testing |
| SVM (RBF)[[25](#asghar)] | 90.30% | 86.70% |
| SVM (POLY)[[25](#asghar)] | 88.80% | 85.20% |
| Naive Bayes[[25](#asghar)] | 68.14% | 70.30% |
| K-NN (k=3)[[26](#gianna)] | 84.40% | 83.30% |

**TABLE 1** : Accuracy percentage of facial emotions reognition system at different classifier and techniques

For evaluating the effectiveness of various machine learning algorithms in facial emotion recognition, multiple classifiers were tested under two different evaluation strategies: 10-fold cross-validation and an 80/20 train-test split. As observed in Table 1, the SVM with Radial Basis Function (RBF) kernel got the maximum accuracy, reaching 90.3% with cross-validation and 86.7% in the train-test scenario. This reflects the RBF kernel's strength in capturing non-linear patterns commonly present in facial features.

The Polynomial kernel SVM also performed competitively, yielding 88.8% and 85.2% accuracies, respectively. In contrast, Naïve Bayes, while computationally lightweight, demonstrated relatively lower performance—68.14% for cross-validation and 70.3% for train-test—indicating its limitations in handling complex feature interdependencies.

Meanwhile, the K-Nearest Neighbors (KNN) algorithm (with k=3) provided a balance between simplicity and performance, achieving 84.4% and 83.3% in the two setups. These findings highlight that while deep learning methods are increasingly popular, classical ML classifiers like SVM still remain highly effective for emotion recognition, especially when paired with optimized kernel functions and appropriate feature sets.

**FIG 1** : Distribution of classifiers used in recent emotion recognition studies

To understand the concept of machine learning techniques used in emotion recognition systems, we analyzed a set of recent studies published between 2023 and [2024](#ramachandran). The pie chart below illustrates the distribution of classifier types across these works.

As Shown, K-Nearest Neighbors (KNN) is used in some lightweight or hybrid approaches due to its simplicity and effectiveness with small datasets.

Decision Trees are still useful for rule-based or interpretable systems, especially when computational simplicity is needed.

Recurrent Neural Networks (RNNs) and LSTM architectures follow closely, mainly used in temporal data scenarios such as EEG signals or speech emotion recognition, where tracking changes over time is essential.

Transformer-based models have emerged as the most widely adopted classifiers, making up 26% of the reviewed approaches. Their dominance is due to their strong performance on sequential and multimodal data, particularly in recognizing emotions from combinations of texts, audios, and visual inputs. Some of the examples of Models such as BERT.

Convolutional Neural Networks (CNN) is also popular especially in facial expression recognition. Their ability to extract spatial features makes them ideal for processing images.

Support Vector Machines (SVMs) remain relevant, often paired with deep feature extractors like CNNs for robust classification with limited data.

The Other category includes hybrid classifiers like:

- RVFLN (Random Vector Functional Link Networks) optimized with metaheuristic algorithms,

- Multi-task learning models (e.g., Efficient Net variants).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Neural networks** | **Layers** | **Accuracy**  **(%)** | **Precision**  **(%)** | **Recall**  **(%)** |
| GoogLeNet  [[26](#gianna)] | 22 | 63.21 | 62 | 62 |
| CaffeNet [[27](#yassin)] | 8 | 68 | 67 | 66.20 |
| VGG16 [[28](#zhao)] | 16 | 71.40 | 81.90 | 79.40 |
| ResNet50 [[28](#zhao)] | 152 | 73.80 | 83.30 | 80.70 |
| CNN [[11](#sawmya)] | 36 | 95.60 | 90 | 93 |

**TABLE 2** : comparison of the different neural network techniques

Table 2 shows the performance comparison between various neural network. The GoogLeNet, CaffeNet including CNN various other values are listed. Second column defines the depth or hidden layers present in each technique, CaffeNet used 8 hidden layers while ResNet50 used 152 hidden layers but in case of the outcomes there no such big difference. Hence, we can conclude that the layers of a neural network and the result are independent of each other. Also shows the overall comparison of the advanced techniques with the CNN, which primarily used in Emotion Recognition.

**fig 2** : emotions recognition accuracy by study and dataset

Figure 2. illustrates a comparative analysis of recent emotion recognition systems based on machine learning techniques, highlighting their accuracy across various benchmark datasets. The results demonstrate that Dangi et al., ([2023](#dangi)) achieved the highest reported accuracy of 96% using a Random Vector Functional Link Network on Twitter data for sentiment classification. Similarly, Saleem et al., ([2023](#saleeem)) reported notable results with 93.86% accuracy on the Berlin EMO-DB and 81.14% on the IEMOCAP dataset using a deep CNN architecture with spectro-temporal features.

Irfan et al., ([2023](#irfan)) also showed strong performance, attaining 94.09% accuracy on the CK+ dataset using a Triplet-Loss-Based Deep CNN combined with SVM. On the other hand, Savchenko et al., ([2023](#savchenko)) presented a more modest outcome of 63.03% on the AFEW dataset, which is likely due to the complexity and variability of in-the-wild facial expressions in video sequences. Lastly, Zhang ([2023](#zhang)) employed a hybrid CNN + Transformer model, yielding 83.54% on FER-2013 and 88.97% on CK+, demonstrating the potential of transformer-based architectures in improving generalization across datasets.

These findings emphasize that model performance can vary significantly depending on the dataset characteristics, feature extraction methods, and classification techniques used. The comparison also underlines the importance of dataset selection and hybrid approaches for achieving robust and accurate emotion recognition.

The Emotion Recognition System rely on the facial gestures to perform the real-time analysis, labelling and interpretation of cognitive and affective states from facial video recordings [[1](#neha)]. Evaluation of states of emotion using facial landmarks and machine learning. It uses a dataset to capture facial landmark trajectory signals, tracking the movement of key facial regions such as the eyes, nose and face corners over time [[14](#rk)]. The standard pattern of emotion recognition guides our analysis which includes various EEG feature extraction techniques, feature selection and reduction strategies [[12](#essam)].

Research shows that combination of emotion recognition with intelligence devices has enabled the development of software systems and models that aid in both identifying and label emotions. Researchers will focus on tracking the emotional changes in active and multiplex environments, in the future [[13](#runfang)]. It is believed that facial expressions appear for a certain duration when an emotion is experienced, allowing emotion detection by identifying the corresponding expression. Convolutional Neural Networks (CNN) are utilized to analyze artificial neural networks, random forest and Naive Bayes algorithm, along with the deep learning models. We examine EEG rhythms closely related to the emotions and connections between different brain regions and emotions states [[12](#essam)]. Emotion recognition is becoming a big demanding task in the modern years due to complexity of the huge datasets [[15](#sawmya)]. The studies exploring how people identify facial and vocal emotional expressions have been varying the conclusions over time.

Researchers have aimed to better understand, how performance in emotion recognition is influenced by different tasks and characteristics of emotional expressions [[11](#barbra)]. Many studies struggle with the datasets containing image modifications and profile variations but perform well on the controlled datasets. The system is designed to handle face recognition and the emotion classification. An innovative study using the FER-2013 dataset showed that the mini-Xception algorithm effectively completed given tasks.

We are introducing a novel approach called as Facial Emotion Recognition using Convolutional Neural Networks (FERC). The technique involves the one stage eliminates the background from the image, while the other focuses on extracting facial highlights vectors. This model, an expression vector which is used to identify various facial expressions like sadness, happiness, anger, etc.

The training data was sourced by a stored database containing 10000 images from 154 individuals.

It is different from the other conventional methods that rely on a single-layer CNN, FERC enhances accuracy through its multi-level structure [[6](#ninad)].

While analyzing the trajectories, the researchers aim to develop a powerful emotion evaluation system. Some machine learning algorithms were tested in which Logistic Regression gave the best outcome by achieving an accuracy of 98.95% and an F1-score of 0.99. The system effective Accuracy, and ROC-AUC. The results show that the LR model surpassed latest and advanced techniques, highlighting its potential as a reliable tool for emotion recognition and decision making based on emotional state variations [[10](#akhilesh)].

The model processes data only once, requiring an immediate response for real time emotion classification. To improve recognition accuracy, the CNN weights are adjusted using DC-SSA algorithm. The learners' performance is then assessed based on the recognition outcomes [[16](#vishal)].

The proposed system is highly relevant for combination with web applications to monitor learner’s real-time emotion. Tests are conducted using the CK+ and JAFFE facial datasets achieved accuracy rates of 96.46% and 98.43%, respectively. To deliver precise results, the system offers high quality, strong and real-time feedback on basis of facial expressions, enhancing the understanding of students' emotional engagement during online learning [[9](#alisha)].

Conclusion

This review paper gave the idea of advance field of facial emotion recognition (FER) taking support from affective computing, machine learning, and computer vision tasks. The main aim is to look how new technologies helped machine to understand human emotions better and faster. Since facial expressions, a powerful and widely understood way of communicating without words, using them in human-computer interaction is an important move toward creating smart and advanced systems.

The review pointed out that traditional human-computer interaction (HCI) system doesn’t often consider user’s emotions which leads to lower satisfaction and engagement. By adding facial emotions recognition (FER) features, systems can now respond to how users feel, making interactions more natural, responsive, and effective. Deep learning techniques like Convolutional Neural Networks (CNN) and tools like Deep Face have proven to be very accurate at recognizing emotions by analyzing facial expressions.

In fact, some experiments mentioned in the review showed accuracy levels over 95%, highlighting how reliable these methods are for real-world use. We looked at several well-known models and datasets, such as FER-2013, Face Net, and JAFFE. These have helped train and test strong FER systems using Deep Face. New methods, like using CNNs in multiple stages or combining different deep learning techniques, pushed performance even further. These updates make it possible to detect emotions in real-time with low delay and high confidence and accuracy. This allows to seek the opportunities in areas like education, healthcare, customer service, and mental health support. In the end, the review shows how emotion-aware technology can help remove the emotional gap between human and machines. By giving systems emotional intelligence by using latest technologies, we can create digital experiences that are more human, personalized, and effective & increase human machine interaction. Future work should focus on better emotion detection across different cultures, improving real-time accuracy, and making sure system maintain the confidentiality of emotional data.

References

1. Neha S, Nivya, Pooja HC Shekar, Keerthana S Kumar, AshaVG (2020) Emotion Recognition and Depression Detection using Deep Learning 2020 vol. 07, no. 08, pp. 2395 2472
2. Alojaiman, B. (2021). Toward selection of trustworthy and efficient E-Learning platform. IEEE Access, 9, 133889–133901, /10. 1109/ access. 2021. 31141 50
3. Ariza, J., Jimeno, M., Villanueva-Polanco, R., & Capacho, J. (2021). Provisioning computational resources for cloud-based e-learning platforms using deep learning techniques. IEEE Access, 9,

89798–89811. /10. 1109/ access. 2021. 30903 66

1. Lyu, L., Zhang, Y., Chi, M. Y., Yang, F., Zhang, S. G., Liu, P., & Lu, W. G. (2022). Spontaneous facial expression database of learners’ academic emotions in online learning with hand occlusion. Computers and Electrical Engineering, 97, 107667./10. 1016/j.compeleceng. 2021. 107667
2. Sreenidhi M J, Surabi Sri Dhanya S, Sahithi R (2020) Human Emotion Recognition System using Deep Learning Techniques 2020 vol. 07, no. 08, 2278-0181
3. Ninad Mehendale (2020) Facial emotion recognition using convolutional neural networks (FERC) 2020 https://doi.org/10.1007/s42452-020-2234-1
4. Ramachandran Venkatesan, Sundarsingh Shirly, Mariappan Selvarathi and Theena Jemima Jebaseeli (2023) Human Emotion Detection Using DeepFace and Artificial Intelligence 2023.
5. Eman M. G. Younis, Someya Mohsen, Essam H. Houssein, Osman Ali Sadek Ibrahim (2024) Machine learning for human emotion recognition 2024 /10.1007/s00521-024-09426-2
6. Alisha Fida, Muhammad Umer, Oumaima Saidani, Monia Hamdi, Khaled Alnowaiser, Carmen Bisogni, Andrea F. Abate, Imran Ashraf (2023) Real time emotions recognition through facial expressions 2023 /10.1007/s11042-023-16722-x
7. Akhilesh Kumar, Awadhesh Kumar, Sumit Gupta, Machine Learning-Driven Emotion Recognition Through Facial Landmark Analysis 2025 /10.1007/s42979-025-03688-w.
8. Barbra Zupan, Michelle Eskritt (2023) Facial and Vocal Emotion Recognition in Adolescence 2023 /10.1007/s40894-023-00219-7
9. Essam H. Houssein, Asmaa Hammad, Abdelmgeid A. Ali (2022) Human emotion recognition from EEG-based brain computer interface using machine learning 2022 /10.1007/s00521-022 07292-4
10. Runfang Guo, Hongfei Guo, Liwen Wang, Mengmeng Chen, Dong Yang, Bin Li1 (2024) Development and application of emotion recognition technology 2024 /10.1186/s40359 024-01581-4
11. RK Kapila Vani, P.Jayashree (2025) Platform 2025 /10.1007/s10639-024-13279-6
12. Sowmya BJ, Meeradevi, Sini Anna Alex, Anita Kanavalli, Supreeth S, Shruthi G, Rohith S (2023) Machine learning model for emotion detection and recognition using an enhanced Convolutional Neural Network.
13. Vishal singh bhati , Namita Tiwari, Meenu Chawla (2025) A Generalized Zero-Shot Deep Learning Classifier for Emotion Recognition Using Facial Expression Images 2025.
14. Manisha Kasar, Bindu Garg, Trupati Suryawanshi (2024) EmoSense: Pioneering Facial Emotion Recognition with Precision Through Model Optimization and Face Emotion Constraints 2024
15. Mingyi Zhao, Linrui Gong, Abdul Sattar Din (2024) A review of the emotion recognitionmodel of robots 2024
16. Sezer A, Altan A. Detection of solder paste defects with an optimization‐based deep learning model using image processing techniques. Soldering & Surface Mount Technology. 2021;33(5):291-8.
17. A. Dangi, A. Mittal, and D. S. Kushwaha, ARO-RVFLNN: Artificial Rabbits Optimized Robust Random Vector Functional Link Network for sentiment classification, Expert Systems with Applications, 2023.
18. M. Saleem, T. Anwar, and M. Usama, DeSCCNet: Spectro-temporal feature representation for speech emotion recognition, Applied Acoustics, vol. 215, 2023.
19. M. Irfan, M. A. Javed, A. M. Qamar, and F. Hussain, Triplet-Loss-Based Deep CNN Features with SVM for Facial Emotion Recognition, IEEE Access, vol. 11, pp. 43025–43036, 2023.
20. A. Savchenko, Multi-task EfficientNet-B2 with MTCNN and Adam optimization for facial emotion recognition, Signal Processing: Image Communication, vol. 118, 2023.
21. S. Zhang, RDA-MTE: ResNet-50 + Bidirectional Attention + Transformer Encoder for facial expression recognition,Neural Computing and Applications, vol. 35, pp. 28195–28210, 2023.
22. A. S. Asghar, F. U. Khan, S. Nisar, and N. Shah, Facial emotion recognition using statistical features and machine learning, Multimedia Tools and Applications, vol. 82, pp. 20541–20558, 2023, doi: 10.1007/s11042-023-15049-w.
23. P. Giannopoulos, I. Perikos, and I. Hatzilygeroudis, Deep learning approaches for facial emotion recognition: A case study on FER-2013, in Smart Innovation, Systems and Technologies, vol. 85, pp. 1–16, 2018.
24. W. M. Yassin, M. F. Abdollah, Z. Muslim, R. Ahmad, and A. Ismail, An emotion and gender detection using hybridized convolutional 2D and batch norm residual network learning, in Proc. ACM Int. Conf., 2021, pp. 79–84
25. X. Zhao, X. Shi, and S. Zhang, Facial expression recognition via deep learning, IETE Technical Review, vol. 32, no. 5, pp. 347–355,