World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Leveraging deep CNNs for accurate facial emotion recognition and analysis

P Chiranjeevi, Nagalaxmi Kalluri *, Sai Saket Gurubhagavatula, Abhishek Kuncham and Mohammed Sami

Department of CSE (Data Science), ACE Engineering College, Hyderabad, Telangana, India.

World Journal of Advanced Research and Reviews, 2025, 25(02), 204-212

Publication history: Received on 23 December 2024; revised on 01 February 2025; accepted on 03 February 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.25.2.0342

Abstract

Facial Emotion Detection is imaginative and prescient and device mastering; it determines human emotions based totally on facial expressions, thereby the distance among emotional intelligence from a human and synthetic intelligence allowing natural human-gadget interactions. Spotting feelings efficiently is a need for scientific, educational, advertising and marketing, leisure, and human useful resource control fields. Enhancements in such structures might be completed on this challenge based on their obstacles and creation to capabilities to cater feelings together with disgust, wonder, and neutrality, supplying emotional insights for the adaptive mastering or intellectual fitness tracking applications. The system enables simultaneous analyses of several humans within the frame as it can efficiently deal with multi-face detection as pertinent to some organization situations-which includes lecture rooms, conferences, or social gatherings-and gives a closing self-belief score to every prediction growing reliability and trustworthiness.

Keywords: Facial Emotion Detection; Deep Learning; Convolutional Neural networks; Computer Vision

1. Introduction

In the modern digital landscape, comprehending human emotions is vital for building genuine connections and improving user experiences. Unfortunately, the current facial affective emotion recognition systems are primarily capable of identifying basic emotions and face challenges when dealing with intricate, real-life situations. This limitation restricts their application in crucial fields such as education, healthcare, marketing, and mental health monitoring, where a deep understanding of emotions is vital.

Our project, facial affective emotion recognition and detection, tackles these challenges by utilizing convolutional neural networks (cnns) to create a real-time system that can accurately identify a wide range of emotions, even for multiple individuals. By extracting intricate facial features, cnns enable precise detection of subtle emotional cues that traditional methods overlook, opening doors for empathetic and interactive digital environments.

Affective Emotion detection has the power to bring about significant changes in various fields. In the medical field, it assists in evaluating patients' emotional well-being to facilitate early intervention and tailor care to their specific needs. In education, it aids educators in recognizing students' challenges or abilities, while in marketing, it fuels campaigns by analyzing emotional reactions. By empowering machines to comprehend and respond empathetically, ferd plays a crucial role in the advancement of empathetic artificial intelligence systems that improve the interaction between humans and machines.

Facial affective emotion recognition systems are crucial in connecting human emotions with technological interaction, improving user experiences in various domains. Emotions play a crucial role in communication, and comprehending them allows for the development of applications in healthcare, education, and customer service to become more empathetic and responsive. This project is driven by the need for intelligent systems that can understand and respond

* Corresponding author: K Nagalaxmi

to emotions, thanks to recent breakthroughs in deep learning and computer vision, enabling real-time and precise emotion recognition. Equipped with advanced features such as multi-face detection and an accuracy meter, the system guarantees precision, scalability, and reliability, even in intricate situations. By considering cultural diversity, environmental differences, and real-time limitations, this project aims to develop technology that prioritizes human needs and facilitates impactful and meaningful interactions.

2. Related Work

The literature on facial affective emotion recognition and detection using convolutional neural networks (cnns) emphasizes the significant impact of deep learning in understanding human emotions. Foundational work by paul ekman,wallace v. Friesen, and phoebe ellsworth in emotion in the human face: guidelines for research and an integration of findings introduces the facial action coding system (facs), which links facial muscle movements, or action units , to specific emotions. This study forms the foundation for emotion analysis, driving progress in affective computing, human-computer interaction, and clinical psychology.

Building upon this foundation, de la torre and cohn offer a comprehensive examination of facial expression analysis methodologies, encompassing geometric modeling, statistical approaches, and machine learning techniques. Their research focuses on showcasing the scalability and efficiency of cnns in automating emotion recognition and classification, especially for real-time applications. The incorporation of deep learning has greatly enhanced the accuracy and flexibility of facial expression recognition systems.

Zhang's research centers around facial expression recognition based on features, conducting sensitivity analyses and experiments using multilayer perceptrons. This research emphasizes the significance of extracting features and utilizing neural networks to enhance recognition accuracy. Zhang suggests methods to improve cnn-based models for emotion detection, focusing on key performance indicators, and validating their effectiveness through experiments. In addition to the contributions by bavkar, sandeep, rangole, and deshmukh, other researchers have developed a geometric framework for recognizing human emotions.

Their work combines theoretical principles with CNN-based feature extraction to achieve high accuracy in classifying emotional states. Additionally, advancements in cnn architectures, including resnet and vgg, pre-trained models, and data augmentation techniques, have addressed technical challenges like lighting variations, occlusions, and cultural diversity, further enhancing the robustness of emotion detection systems.

These studies collectively contribute to a comprehensive understanding of facial emotion recognition using CNNs, providing valuable insights into system design, implementation, and optimization. This study sets the foundation for practical applications in healthcare, education, entertainment, and human-computer interaction, propelling advancements in affective computing with exceptional accuracy and real-time performance.

3. Existing System

Facial affective emotion recognition systems are designed to detect emotions like happiness, sadness, anger, fear, and neutral expressions, utilizing machine learning and deep learning algorithms, with cnns being the most effective due to their feature extraction capabilities. These systems frequently utilize pre-existing models and datasets, such as fer-2013, which offers a vast collection of annotated facial images for classifying basic emotions.

They are extensively used in various industries, including customer service, marketing, and human-computer interaction. Nevertheless, current systems have significant drawbacks. They are limited in their ability to recognize basic emotions and find it challenging to understand more complex ones like surprise or disgust. Furthermore, they only support single-face detection, do not have real-time processing capabilities, and do not provide confidence metrics to ensure reliability. These limitations impede their efficiency in complex, group-based, and intricate emotional situations.

4. Proposed Model

The facial affective emotion recognition and detection (ferd) system tackles the shortcomings of existing models by incorporating cutting-edge features. It will be able to categorize both simple and intricate emotions, such as astonishment and revulsion, while also allowing for the simultaneous analysis of emotions for multiple individuals. This makes the system well-suited for dynamic environments and group scenarios. By utilizing efficient algorithms and real-time processing, ferd guarantees quick response times and instant feedback.

Furthermore, the system will include confidence metrics for each detected emotion, improving reliability and accuracy in various applications. The system will utilize convolutional neural networks (cnns) to extract robust features and classify emotions accurately. Advanced deep learning frameworks, like tensorflow, will be employed for constructing and training models.

5. Methodology

The methodology for the facial emotion recognition and detection (ferd) system centers around gathering a wide range of facial expression data, constructing an intelligent convolutional neural network (cnn), and facilitating real-time emotion detection. The process guarantees easy-to-understand visualizations and ongoing enhancements to make the system flexible and suitable for practical applications.

The machine learning algorithms and techniques selected for this project include libraries such as scikit-learn, TensorFlow, and Keras. Those Libraries imported are:

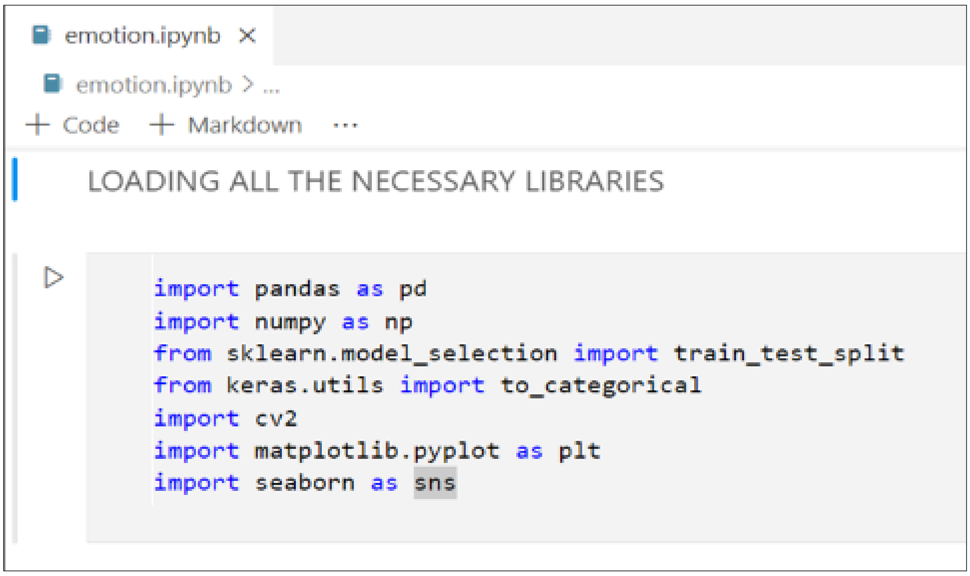

Figure 1 Libraries imported

5.1. Data Collection

Data collection is the foundation for developing a robust facial emotion recognition model. This step ensures the system has access to diverse and representative data.

5.1.1. Dataset Selection

Use datasets like FER-2013.These dataset cover a wide range of emotions (e.g., happiness, sadness, anger, fear, surprise, and disgust, neutral).

5.1.2. Data Preprocessing

Resize images to a uniform size (e.g., 48x48 pixels for FER-2013).Normalize pixel values to a range of 0–1 for better compatibility with the CNN. Perform data augmentation (e.g., flipping, rotation, zooming, noise addition) to enhance the dataset's diversity and prevent model overfitting.

5.2. System Architecture

The system architecture for Facial Emotion Recognition and Detection (FERD) visually represents the structural components and their interactions, enabling real-time emotion detection and classification.

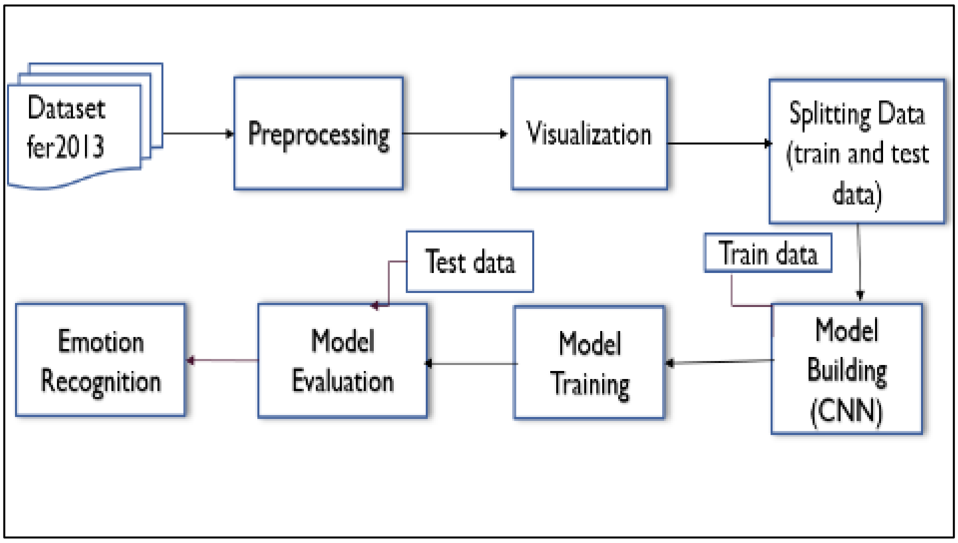

Figure 2 System Architecture

It illustrates how input data flows through various processing layers, resulting in meaningful outputs.

5.2.1 System Architecture Components

- **Input Module:** Captures images or video streams using cameras, ensuring a continuous flow of data for analysis.
- **Face Detection Module:** Detects and localizes faces in the input frames using algorithms like Haar cascades, MTCNN, or OpenCV.
- **Emotion Recognition Module:** Uses a Convolutional Neural Network (CNN) for emotion classification. It extracts facial features and identifies emotions like happiness, sadness, anger, etc.
- **Confidence Scoring Module:** Assigns confidence levels to the detected emotions, providing a measure of prediction reliability.
- **User Interface (UI):** Displays real-time results, including detected faces, their emotions, and confidence scores, in an accessible and user-friendly format.
- **Output:** Provides results in real-time with a visual display of emotions and confidence percentages.

5.3. Model Development

The model is the core component of the system, responsible for learning and predicting emotions.

5.3.1. Define CNN Architecture

Use convolutional layers to extract spatial features from facial images. Add pooling layers (e.g., max pooling) to reduce dimensionality while preserving key features. Include fully connected layers for classification into emotion categories.

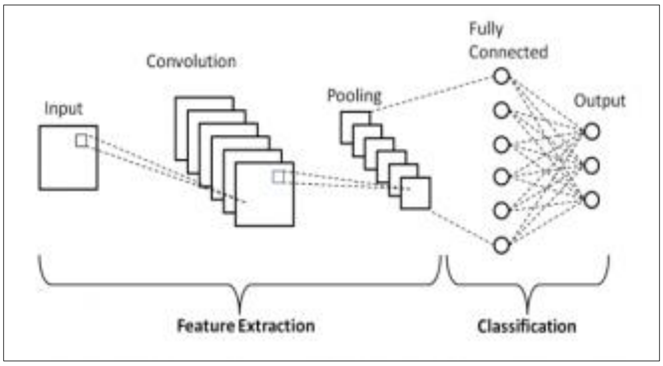

Figure 3 CNN Architecture

5.3.2. Validation

Split the data into training, validation, and test sets. Monitor metrics like accuracy and loss on the validation set to identify and mitigate overfitting. Use techniques like dropout layers to improve generalization.

```
from keras.models import Sequential
from keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout

# Build the CNN model
model = Sequential()

# First Convolutional Layer
model.add(Conv2D(64, (3, 3), activation='relu', input_shape=(48, 48, 1)))
model.add(MaxPooling2D(pool_size=(2, 2)))

# Second Convolutional Layer
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
#Third Convolutional layer
model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))

# Flatten the output
model.add(Flatten())

# Fully connected layer
model.add(Dense(512, activation='relu'))
model.add(Dropout(0.5))  # Dropout for regularization

# Output layer (7 emotions)
model.add(Dense(7, activation='softmax'))
```

Figure 4 Model Building

5.3.3. Training

Train the model using batch training, where the dataset is divided into smaller subsets for efficient learning.

Apply backpropagation to minimize the loss function and update model weights.

MODEL TRAINING

```
# Train the model
history = model.fit(X_train, y_train, epochs=30, batch_size=32)
```

```
Epoch 1/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 167s 186ms/step - accuracy: 0.3867 - loss: 1.5698
Epoch 2/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 208s 231ms/step - accuracy: 0.4794 - loss: 1.3646
Epoch 3/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 160s 178ms/step - accuracy: 0.5126 - loss: 1.2804
Epoch 4/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 169s 188ms/step - accuracy: 0.5428 - loss: 1.2021
Epoch 5/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 138s 153ms/step - accuracy: 0.5649 - loss: 1.1448
Epoch 6/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 131s 146ms/step - accuracy: 0.5916 - loss: 1.0844
Epoch 7/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 138s 154ms/step - accuracy: 0.6099 - loss: 1.0334
Epoch 8/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 129s 144ms/step - accuracy: 0.6295 - loss: 0.9869
Epoch 9/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 130s 144ms/step - accuracy: 0.6519 - loss: 0.9265
Epoch 10/30
898/898 ━━━━━━━━━━━━━━━━━━━━ 129s 144ms/step - accuracy: 0.6656 - loss: 0.8919
```

Figure 5 Model Training

5.3.4. Testing

Evaluate the model on unseen test data to ensure its ability to generalize to new inputs. Use performance metrics such as accuracy, precision, recall to assess effectiveness.

5.4. Real-Time Data Processing

This step ensures the system operates effectively in a real-time environment.

5.4.1. Input Capture

Use webcams or cameras to stream video or capture images in real-time. Integrate software like OpenCV to manage video streams and frame capture.

5.4.2. Face Detection

Employ robust face detection algorithms (e.g., Haar cascades) to identify face regions in each frame. Ensure the face detector works effectively under varying lighting conditions, angles, and occlusions.

5.4.3. Feature Extraction

Use algorithm Facial Landmarks Detection to pinpoint key facial features such as eyes, eyebrows, nose, and mouth. Focus on emotion-relevant regions to reduce noise and improve accuracy.

5.5. Emotion Detection

This step involves predicting emotions based on the extracted features.

5.5.1. Model Prediction

Feed the preprocessed facial data into the trained CNN model. Use softmax activation in the final layer to output probabilities for each emotion category.

5.5.2. Confidence Scoring

Assign confidence scores to predictions, indicating how certain the model is about each emotion. Use thresholds to filter out low-confidence predictions, ensuring reliability.

5.6. Visualization

Delivering results in a user-friendly manner is crucial for system usability.

5.6.1. Real-Time Display

Developed a graphical user interface (GUI) using tool like Stream lit. Displaying emotion labels, confidence scores, and bounding boxes over detected faces in real-time.

6. Results and Discussion

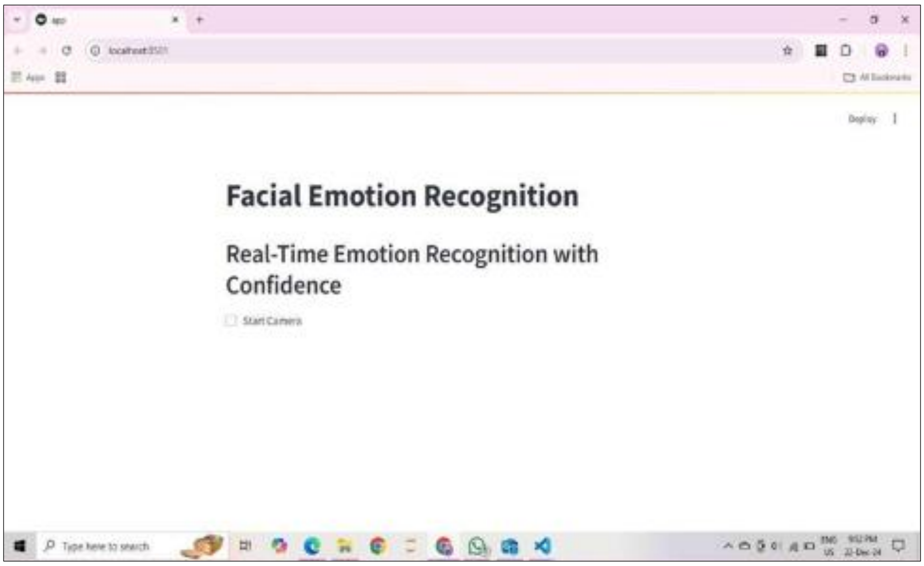

Figure 6 User Interface

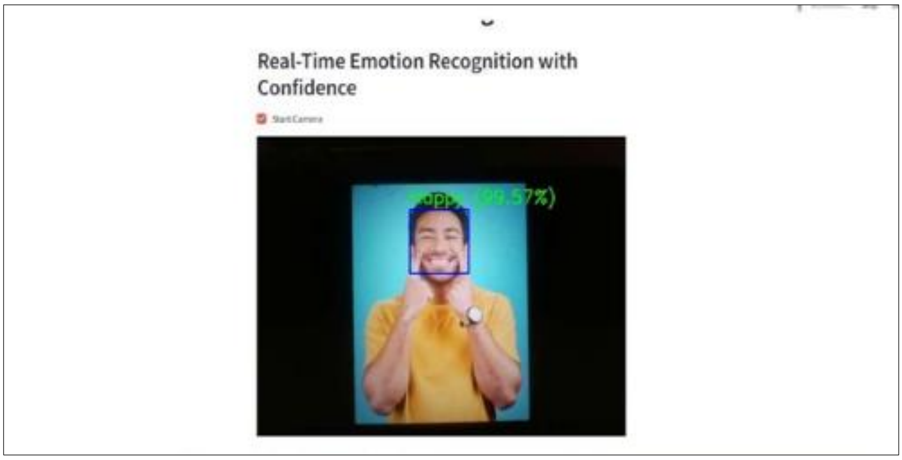

Figure 7 Facial Emotion Recognition

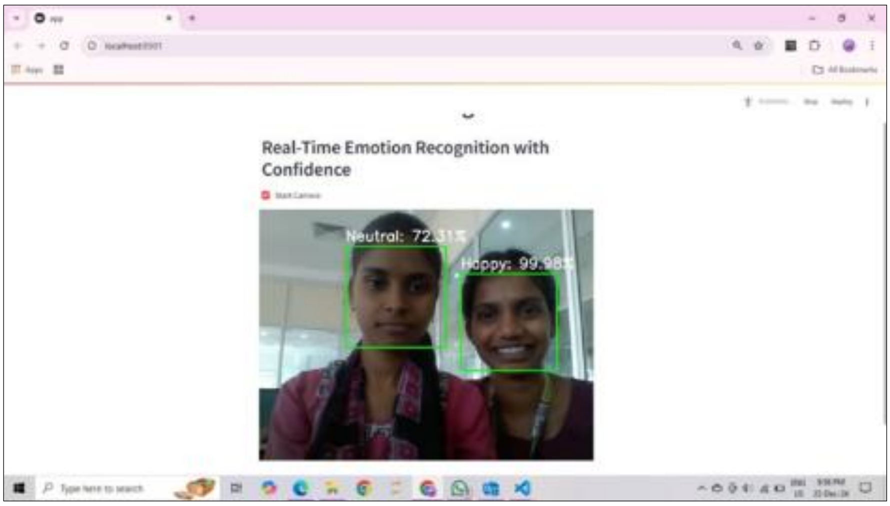

Figure 8 Multi-Facial Emotion Recognition

Figure 9 Multi-Facial Emotion Recognition

7. Conclusion

To summarize, the facial emotion recognition system equipped with multi-face detection and an accuracy meter is an effective tool for real-time analysis of human emotions based on facial expressions. By utilizing sophisticated algorithms for face detection and emotion classification, the system can effectively identify multiple faces in diverse settings and accurately classify emotions with a high level of confidence. By incorporating a dynamic accuracy meter, users gain visibility into the system's reliability, enabling them to evaluate the accuracy of its predictions. This method, backed by an intuitive interface, guarantees seamless interaction and instant feedback, making it suitable for various domains including security, human-computer interaction, and mental health monitoring. The system's flexible design, ability to handle varying workloads, and high performance guarantee its dependability and versatility for diverse applications.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Emotion in the Human Face: Guidelines for Research and an Integration of Findings by Paul Ekman, Wallace V. Friesen, Phoebe Ellsworth

[2] Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and RecognitionEvangelos Sariyanidi, Hatice Gunes, Andrea Cavallaro

[3] "FER-2013 Dataset": https://www.kaggle.com/datasets/nicolejyt/facialexpressionrecognition

[4] Zhang's research centers around facial expression recognition

[5] A. M. Martinez, "Recognition of Facial Expressions: The State of the Art," in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 1998, pp. 1493-1498

[6] Z. M. L. L. D. C. Kim, H. W. S. S. L. K. Cho, "MTCNN: Multi-Task Cascaded Convolutional Networks for Face Detection," in Proceedings of the 2016 IEEE International Conference on Computer Vision (ICCV), 2016.

[7] KUMAR, P. ASHOK, G Satish KUMAR, and SETTI NARESH KUMAR. "Improve the Capacity of Uniform Embedding for Efficient JPEG Steganography Based on DCT."(2015).

[8] Kumar, P. Ashok, B. Vishnu Vardhan, and Pandi Chiranjeevi. "Investigating Context-Aware Sentiment

[9] Classification Using Machine Learning Algorithms." In XVIII International Conference on Data Science and Intelligent Analysis of Information, pp. 13-26. Cham: Springer Nature Switzerland, 2023.

[10] Sunkavalli, Jayaprakash, B. Madhav Rao, M. Trinath Basu, Harish Dutt Sharma, P. Ashok Kumar, and Ketan Anand. “Experimentation Analysis of VQC and QSVM on Sentence Classification in Quantum Paradigm.” In 2024 International Conference on Computing, Sciences and Communications (ICCSC), pp. 1-5. IEEE, 2024.

[11] PANDI, CHIRANJEEVI, THATIKONDA SUPRAJA, P. ASHOK KUMAR, and RALLA SURESH.”A SURVEY: RECOMMENDER SYSTEM FOR TRUSTWORTHY.”;

[12] Kumar, P. Ashok, B. Vishnu Vardhan, and Pandi Chiranjeevi. “Correction to: Investigating Context-Aware Sentiment Classification Using Machine Learning Algorithms”In XVIII International Conference on Data Science and Intelligent Analysis of Information, pp. C1-C1. Cham: Springer Nature Switzerland, 2023.

[13] Facial emotion recognition using deep convolutional neural networks.”Proceedings of the International Conference on Artificial Intelligence and Pattern Recognition (AIPR).

[14] “Facial expression recognition using deep convolutional neural networks.” Journal of Computer Vision and Image Processing, 4(1), 15-28.

Author’s short biography

Dr. P Chiranjeevi Dr. P Chiranjeevi is working as an HOD & Associate Professor in the Department of CSE (DATA SCIENCE) at ACE Engineering College, Hyderabad (India). He had completed Ph.D at JNTUH University at Hyderabad (India). He is in software Industry for more than 1 year. He is in teaching profession for more than 18 years. His main area of interest includes Opinion Mining, Sentiment Analysis and NLP.	
Nagalaxmi Kalluri I am K Nagalaxmi, a B.Tech student in Computer Science and Engineering (Data Science) with a strong interest in Machine Learning and Data Science. As an undergraduate researcher, I am keen on exploring trending technologies and innovative techniques in predictive analytics, and intelligent systems to solve real-world challenges.	
Sai Saket Gurubhagavatula G Sai Saket is currently pursuing a B.Tech in Computer Science and Engineering with a focus on Data Science. He has developed a strong interest in data-driven technologies, particularly in machine learning, artificial intelligence, and statistical analysis. Throughout his academic journey, he has worked on various research projects and practical applications related to data modeling, predictive analytics, and deep learning. Passionate about leveraging data science to solve real-world problems	
Abhishek Kuncham K Abhishek is currently pursuing a B.Tech in Computer Science and Engineering (Data Science). His research interests include Deep Learning, with a focus on leveraging advanced computational techniques for data-driven applications. As an undergraduate researcher, he is passionate about exploring machine learning models to solve real-world challenges, particularly in intelligent automation and pattern recognition.	
Mohammed Sami I am Mohammed Sami, currently pursuing a B.Tech in Computer Science and Engineering with a specialization in Data Science. My academic journey has been driven by a deep interest in computer science, particularly in the field of machine learning. I have gained valuable experience. As an undergraduate, I am passionate about using data science to tackle real-world problems. I look forward to continuing to explore and contribute to this rapidly evolving field.