Facial Emotion Recognition Using Convolutional Neural Network (CNN)

Mr. Siva Krishna[1], Sathi Asritha Reddy[2], Pabbineedi Ramya Sri[3], Vedullapalli Lakshmi Sandhya[4], Nama Mahesh[5], Guttula Charan Kumar[6]

[1]Assistant Professor, CSE, Department, Vsm College Of Engineering, Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh - 533255
[2,3,4,5,6]Student, CSE, Department, Vsm College Of Engineering, Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh - 533255

ABSTRACT
In the current digital age, the interaction between humans and computers is advancing quickly, with emotion recognition playing an essential role in improving user experience across various applications. Identifying facial emotions can be beneficial in fields such as mental health assessments, analyzing customer feedback, and facilitating human-robot interactions. Human emotions represent a state of mental feelings and arise spontaneously. There is often no distinct link between emotions and facial expressions, and significant variability exists, making facial recognition a complex area of research. Techniques like Histogram of Oriented Gradient (HOG) and Scale Invariant Feature Transform (SIFT) have been explored for pattern recognition. These techniques derive features from images based on manually established algorithms. Recently, Machine Learning (ML) and Neural Networks (NNs) have been applied in the field of emotion recognition. In this report, a Convolutional Neural Network (CNN) is employed to extract features from images for emotion detection. The Python Dlib library is utilized to locate and extract 64 key landmarks on a face. A CNN model is trained using grayscale images from the FER 2013 dataset to categorize expressions into five emotions: happy, sad, neutral, fear, and angry. To enhance accuracy and prevent model overfitting, batch normalization and dropout techniques are implemented. The optimal model parameters are established based on the training outcomes. The results from testing indicate that CNN Model 1 achieves an accuracy of 80% for four emotions (happy, sad, angry, fear) and 72% for all five emotions (happy, sad, angry, neutral, fear), whereas CNN Model 2 records 79% accuracy for four emotions and 72% for five emotions.

INTRODUCTION

Facial Emotion Detection is an emerging field of computer vision that involves detecting and interpreting emotions from facial expressions. Emotions are an essential part of human communication and play a crucial role in social interactions. Emotions such as happiness, sadness, anger, and surprise are expressed through facial expressions that convey a variety of signals, including muscle movements, eyebrow raises, and eye blinks. Facial Emotion Detection systems aim to recognize and interpret these signals and classify the displayed emotion. The ability to recognize and interpret emotions from facial expressions has many potential applications, including improving human-computer interaction, enhancing customer service, and assisting in mental health diagnosis. Facial Emotion Detection is an emerging area in computer vision focused on identifying and interpreting emotions from facial expressions. Emotions play a vital role in human communication and are essential for social interactions. Feelings such as joy, sorrow, rage, and astonishment are expressed through facial gestures that relay a range of signals, including movements of facial muscles, raised eyebrows, and blinking eyes. Systems for Facial Emotion Detection aim to recognize and interpret these signals, classifying the emotions displayed. The capability to detect and interpret emotions from facial expressions has numerous potential uses, including enhancing human-computer interaction, improving customer service, and aiding in mental health assessments. For instance, a Facial Emotion Detection system could analyze customer feedback in real time to gauge their satisfaction level or identify those who might be facing depression or anxiety. Recently, deep learning models, especially Convolutional Neural Networks (CNNs), have demonstrated impressive performance in image classification tasks, including Facial Emotion Detection. CNNs are a subset of neural networks capable of autonomously learning spatial hierarchies of features from images. They have been effectively utilized in various computer vision applications, such as image classification, object detection, and facial recognition. In this project, we intend to create a Facial Emotion Detection system utilizing CNNs. The system will accept an input image of a face and categorize the emotion portrayed in that image. We will train the system using a collection of labeled images of faces along with their corresponding emotions and assess its performance using established metrics. Our aim is to develop a Facial Emotion Detection system that reaches high accuracy in emotion recognition and can be implemented in real-world situations. By harnessing the capabilities of CNNs and advanced computer vision methodologies, we seek to contribute to the expanding body of research in the area of Facial Emotion Detection and its various applications. Security mechanisms, human-computer interaction, and marketing are just a few instances of the myriad practical applications of computer vision, particularly in facial recognition and detection. The past few years have witnessed significant advancements in the precision of face detection and tracking, thanks to methods based on deep learning. There remain several face detection strategies, each with unique advantages and disadvantages. To determine the optimal approach for a specific application, it is essential to compare and evaluate the effectiveness of these techniques.

EXISTING SYSTEM

In the current realm of emotion recognition, various conventional machine learning methods have been utilized. These include:

- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)
- Principal Component Analysis (PCA)
- Linear Discriminant Analysis (LDA)

Although these techniques perform adequately on simple datasets, they tend to falter when dealing with large datasets and intricate facial expressions. Moreover, their efficacy is constrained by the necessity for manual feature extraction. These models depend on engineered features such as shape geometry and landmark detection, which may fail to fully represent the subtle expressions observable in human faces.
Some limitations of the current systems consist of:

- Low accuracy when processing real-world images.
- Inadequate generalization under varying lighting conditions, occlusions, and head orientations.
- Manual feature extraction contributes to longer development times.

Furthermore, current systems lack the robustness required for real-time applications, which further restricts their practical usefulness in fields that demand immediate feedback or rapid processing.

PROBLEM STATEMENT

Facial expressions represent one of the most influential, innate, and universal forms of human communication. The ability to detect emotions through facial expressions is essential for enhancing interaction between humans and computers, improving customer service systems, and aiding in mental health evaluations. Despite significant advancements in computer vision and machine learning, creating a reliable, robust, and real-time facial emotion detection system continues to be a significant challenge.

Traditional methods of facial emotion recognition heavily depend on manually crafted features and shallow classifiers, which frequently struggle to generalize across various datasets. These methods are particularly vulnerable to changes in lighting, positioning, occlusions, and individual differences in emotional expression. Additionally, many current systems have difficulty maintaining reliability in real-world situations where faces may be misaligned or poorly illuminated.

Deep learning techniques, especially Convolutional Neural Networks (CNNs), have demonstrated exceptional performance in image classification and recognition tasks, including detecting facial emotions. However, developing an effective CNN model necessitates large, balanced, and high-quality datasets that are often lacking. Frequently utilized datasets, such as FER2013, tend to be imbalanced and contain noise, resulting in biased models that favor majority emotion categories while performing poorly on minority ones. Moreover, implementing these deep learning models in real-time applications presents difficulties due to the computational demands and resource limitations of edge devices. There is also a shortfall in interpretability within CNN-based models, complicating trust in their predictions in sensitive areas like mental health evaluation.

Consequently, there is a pressing need to create an accurate, efficient, and interpretable facial emotion detection system using CNNs that can function effectively under different real-world conditions, manage imbalanced datasets, and be scalable for implementation on low-resource devices.

PROPOSED SYSTEM

The proposed system aims to develop an efficient and accurate Facial Emotion Detection model using Convolutional Neural Networks (CNNs). The system is designed to classify human facial expressions into distinct emotional categories such as happiness, sadness, anger, surprise, fear, disgust, and neutral. It leverages the powerful feature extraction capabilities of CNNs to recognize subtle patterns in facial expressions and make accurate predictions in real-time.

ADVANTAGES OF PROPOSED SYSTEM

- Exceptional Precision: Delivers enhanced performance relative to conventional methods by utilizing deep learning techniques.
- Instantaneous Processing: Optimized for rapid execution through the use of efficient CNN architectures.
- Expandable: Capable of being broadened to include more emotion categories or integrated with other biometric attributes.
- Cross-Device Compatibility: Built to function across a range of devices, from personal computers to embedded systems.

- Modular Framework: Simple to upgrade, improve, or swap out specific components without impacting the overall system.

IMPLEMENTED ALGORITHMS

a) Convolutional Neural Network (CNN):

A custom Convolutional Neural Network (CNN) is implemented to classify the extracted facial images into one of the seven emotion categories: happy,sad, angry, surprised, fear,disgust,and neutral.

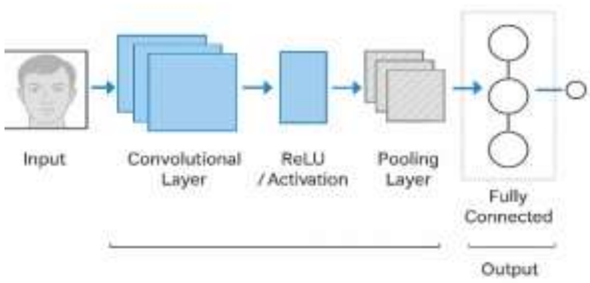

Fig:1 Sample convolutional Neural Network

b) Haar Cascade Classifier (open CV):

To identify faces within an image or video frame, the system employs the Haar Cascade Classifier from the OpenCV library. This method utilizes machine learning, where a cascade function learns from both positive and negative images to recognize objects, specifically human faces.

Advantages:

- Fast and lightweight.
- Effective in detecting frontal faces.
- Suitable for real-time applications.

WORKING OF DEEP LEARNING ALGORITHMS

Convolutional Neural Networks (CNNs) are favored for their effectiveness in tasks related to image recognition. The network autonomously discovers spatial characteristics like edges, shapes, and expressions from data at the pixel level.

- Forward Propagation: Input is passed through convolutional and pooling layers to extract high-level features.
- Backpropagation: The model updates weights using gradient descent based on prediction errors.
- Activation Functions: ReLU is used for hidden layers; Softmax for the output layer.

VIII.SYSTEM ARCHITECTURE

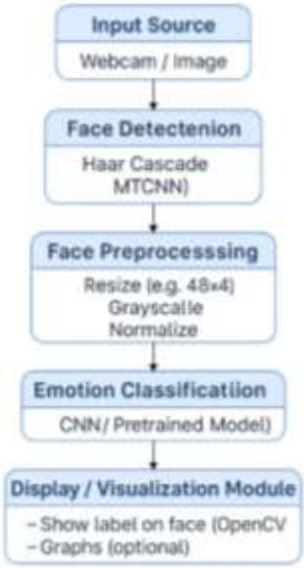

Fig:2 System Architecture Diagram

1. Input Source

Webcam / Image:

The system takes either a live webcam feed or a static image as input to detect and classify facial emotions.

2. Face Detection

These are face detection techniques.

- **Haar Cascade**: A classical method using OpenCV for real-time detection.
- **MTCNN (Multi-task Cascaded Convolutional Networks)**: A more accurate deep learning-based method.

3. Face Preprocessing

This step prepares the detected face for classification:

- Resize: Adjusts the image to a standard size (e.g., 48x48 pixels).
- Grayscale: Converts the image to grayscale (to reduce complexity).
- Normalize: Scales pixel values (usually between 0 and 1) to improve model performance.

4. Emotion Classification

- Uses a **Convolutional Neural Network (CNN)** or a **pretrained model** to predict the emotion.
- Outputs labels like *happy*, *sad*, *angry*, *surprised*, etc.

5. Display / Visualization Module

Shows the result on screen:

Label on face (using OpenCV): Emotion name is displayed on the detected face.

Graphs (optional): Can include charts showing confidence levels or emotion distribution.

IX. RESULT ANALYSIS

Fig:3 Emotion Happy Detection

Fig:4 Emotion Neutral Detection

Fig:5 Emotion Surprise Detection

Fig:6 Emotion Angry Detection

Fig:7 Emotion Fear Detection

Fig:8 Emotion Sad Detection

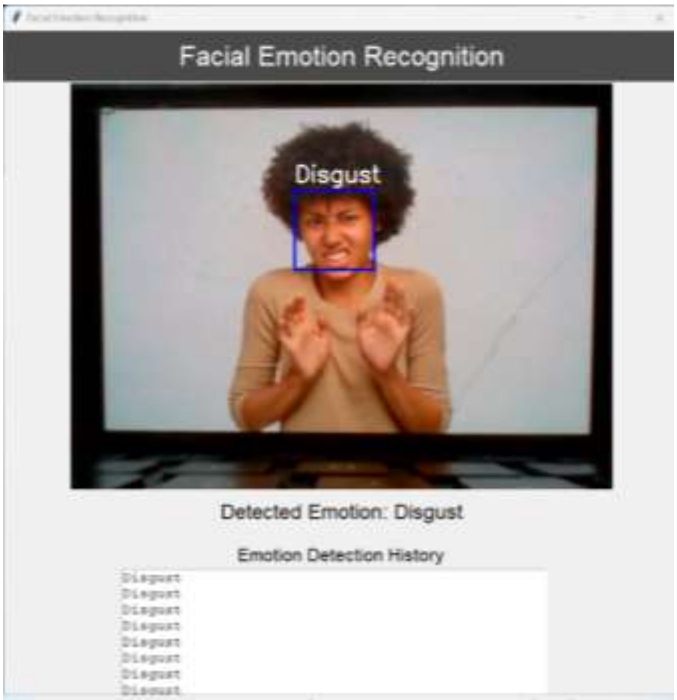

Fig: Emotion Disgust Detection

Fig:8 Multiple Emotions Detection

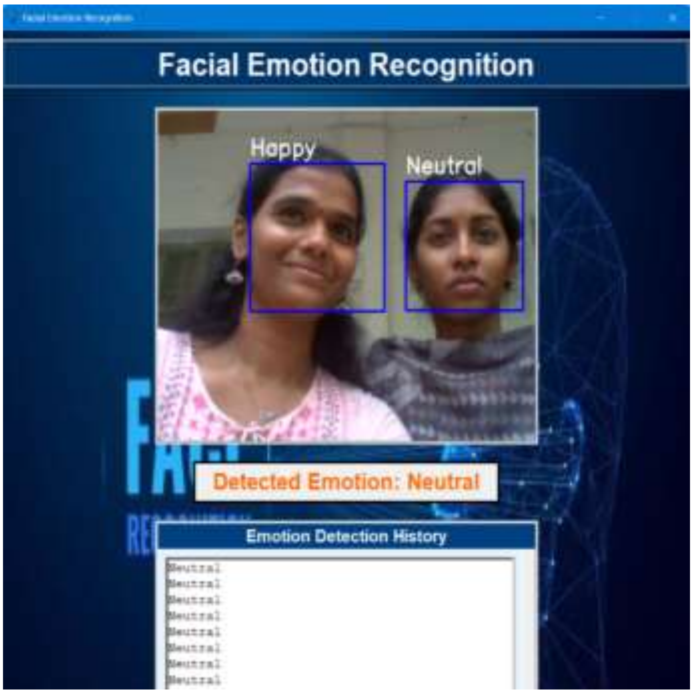

Fig:8 Multiple Emotions Detection

CONCLUSION

In this endeavor, we effectively created a Facial Emotion Recognition (FER) system utilizing Convolutional Neural Networks (CNN). The system reliably identifies human faces from either webcam or image inputs, processes them through resizing, grayscale conversion, and normalization, and categorizes emotions with the help of a trained CNN model. The incorporation of face detection techniques such as Haar Cascade and MTCNN improved the accuracy of face localization. The final results were presented in real-time using OpenCV, rendering the system both interactive and visually engaging. This project showcases the real-world use of deep learning in computer vision, especially in the identification of human emotions—a vital aspect in areas like mental health assessment, human-computer interaction, and intelligent surveillance.

REFERENCES

1. **Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2016)** *AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild* Proceedings of the IEEE Transactions on Affective Computing.
2. **Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012)** *ImageNet Classification with Deep Convolutional Neural Networks* Advances in Neural Information Processing Systems (NeurIPS).
3. **Barsoum, E., Zhang, C., Ferrer, C. C., & Zhang, Z. (2016)** *Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution* Microsoft Research.
4. **Tang, Y. (2013)** *Deep Learning using Support Vector Machines*
5. **Goodfellow, I. J., Erhan, D., Luc Carrier, A., Courville, A., & Bengio, Y. (2013)** *Challenges in Representation Learning: A Report on Three Machine Learning Contests* Springer Lecture Notes in Computer Science.
6. **Zhao, X., Zhang, S., Liang, Y., & Xu, T. (2021)** *Facial Expression Recognition Based on Deep Learning* Journal of Physics: Conference Series, IOP Publishing.

7. **Li, Y., Zeng, J., Shan, S., & Chen, X. (2017)** *Occlusion Aware Facial Expression Recognition Using CNN with Attention Mechanism* IEEE Transactions on Image Processing.
8. **Hasani, B., & Mahoor, M. H. (2017)** *Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks* IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
9. **Szegedy, C., Liu, W., Jia, Y., et al. (2015)** *Going Deeper with Convolutions* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
10. **Khorrami, P., Paine, T. L., & Huang, T. S. (2015)** *Do Deep Neural Networks Learn Facial Action Units When Doing Expression Recognition?* IEEE International Conference on Computer Vision (ICCV