IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Design and Implementation of a Deep Learning-Based Facial Emotion Recognition System using Convolutional Neural Networks

Ms. Shraddha Rajendra Tayade, Mrs. Vaishali Bagade

[1]*PG student* :shraddhatayade02@gmail.com

Alamuri Ratnamala Institute of Engineering & Technology, Sapgoan, India, 421601

[2]*Assistant Professor* ::vaishali.bagade2@gmail.com

Alamuri Ratnamala Institute of Engineering & Technology, Sapgoan, India, 421601

Email of corresponding Author: vaishali.bagade2@gmail.com

Abstract – *Facial emotion recognition is an important research area in artificial intelligence and computer vision. Human emotions are commonly expressed through facial expressions, voice, and body language. Among these communication channels, facial expressions are considered one of the most reliable indicators of emotional states. Automatic emotion recognition systems analyse facial images and classify emotions using computational techniques.*

Traditional emotion recognition systems relied on handcrafted feature extraction methods and classical machine learning algorithms such as Support Vector Machines and Decision Trees. Although these methods provided moderate accuracy, they often failed to capture complex variations in facial expressions due to changes in lighting conditions, facial orientation, and image quality.

With the advancement of deep learning, Convolutional Neural Networks (CNNs) have shown remarkable performance in image recognition tasks. CNN models automatically extract hierarchical features from images and can identify subtle facial patterns that represent emotional expressions.The proposed system focuses on the design and implementation of a deep learning-based facial emotion recognition system using a CNN-based pre-trained model. The system captures facial images through a webcam, performs preprocessing, extracts meaningful features, and classifies emotions such as happiness, sadness, anger, fear, surprise, disgust, and neutral in real time.The system is evaluated based on its ability to detect emotions accurately under different conditions and demonstrates the effectiveness of deep learning techniques in real-time web-based applications.

Keywords *- Facial Emotion Recognition, Deep Learning, Convolutional Neural Network (CNN), Computer Vision, Real-time Systems*

INTRODUCTION

Facial emotion recognition (FER) is a rapidly growing domain within the fields of artificial intelligence, computer vision, and human-computer interaction. Emotions are fundamental to human communication and social interaction. They provide essential cues about a person's mental state, intentions, and reactions. Automated recognition of these emotions through facial expressions has far-reaching implications across numerous real-world applications, from healthcare and security to entertainment and education. The human face is one of the most expressive channels of emotional communication. Research in psychology, particularly the pioneering work of Paul Ekman and Wallace Friesen, established that there exist six universal basic emotions — happiness, sadness, anger, fear, surprise, and disgust — that are expressed consistently across cultures worldwide.

This cross-cultural consistency makes facial expression analysis a reliable biometric modality for emotion detection. Traditional approaches to facial emotion recognition relied on hand-crafted feature extraction techniques such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), Gabor filters, and Active Appearance Models (AAM). While these methods achieved reasonable results under constrained conditions, they struggled with real-world variability in pose, illumination, occlusion, and individual differences in expression intensity.

The emergence of deep learning, specifically Convolutional Neural Networks (CNNs), has brought transformative improvements to the field of facial emotion recognition. CNNs possess the unique ability to automatically learn hierarchical feature representations from raw image data, eliminating the need for manual feature engineering. This project presents the design, development, and evaluation of a deep learning based facial emotion recognition system using CNNs, trained on benchmark datasets to classify facial expressions into the seven core emotion categories.

LITERATURE REVIEW

The field of facial emotion recognition has a rich history of research spanning several decades. Early approaches relied on geometric-based methods that model the spatial relationship between facial landmarks, while appearance-based methods captured texture and shape information from the entire face or specific regions. Ekman and Friesen (1978) laid the foundational theoretical framework for emotion analysis by establishing the Facial Action Coding System (FACS), which categorises facial movements into Action Units (AUs). FACS-based approaches have been extensively used in automated facial analysis, but their computational demands and sensitivity to facial landmark localisation accuracy remained significant limitations.

With the advent of machine learning, researchers began applying Support Vector Machines (SVM), Bayesian classifiers, and ensemble methods to feature vectors extracted from facial images. Bartlett et al. (2005) demonstrated the effectiveness of SVM classifiers combined with Gabor wavelet features for AU detection. Similarly, Shan et al. (2009) showed that LBP features combined with boosted-LDA achieved strong performance on the CK+ database. These methods improved accuracy but still depended heavily on handcrafted feature extraction techniques.

The deep learning revolution fundamentally changed the facial emotion recognition landscape. Krizhevsky et al. (2012) demonstrated the power of deep Convolutional Neural Networks (CNNs) on large-scale image classification tasks, inspiring researchers to apply similar architectures to facial analysis. Kahou et al. (2013) proposed a CNN-based approach for the EmotiW challenge, showing that deep networks could outperform traditional methods in unconstrained environments. Tang (2013) introduced a modified CNN with a linear SVM classifier and achieved top performance in the FER2013 challenge. Further advancements include the work of Kim et al. (2016), who proposed a hierarchical ensemble of CNN models trained on different facial regions.

Transfer learning emerged as an effective strategy for improving performance, especially due to limited emotion-labelled datasets. Mollahosseini et al. (2016) demonstrated that fine-tuning pre-trained networks such as VGGNet significantly enhances recognition accuracy. Li et al. (2017) addressed annotation noise through a Reliable Crowdsourcing framework. More recently, attention mechanisms have been introduced to focus on important facial regions. Wang et al. (2020) proposed the Self-Cure Network (SCN) to handle noisy labels, while Ma et al. (2021) developed transformer-based approaches such as POSTER for improved feature extraction.

Despite these advancements, several challenges remain. Class imbalance across emotion categories leads to biased models, while occlusion, pose variation, and subtle expressions reduce recognition accuracy. Additionally, cross-dataset generalisation remains a significant issue, as models often perform poorly on unseen data. These challenges highlight the need for more robust and generalisable deep learning-based emotion recognition systems.

METHOLOGY

The proposed system follows a structured pipeline for facial emotion recognition, consisting of face detection, preprocessing, facial alignment, feature extraction, emotion classification, and result visualisation. The system is designed to process both static images and real-time video streams efficiently.

The overall architecture of the system is shown in Fig. 1. The process begins with face detection, where human faces are identified from input images or video frames using deep learning-based detectors. The detected faces are then preprocessed by resizing and normalising illumination conditions to improve input quality. Facial alignment is performed to standardise the orientation of faces, ensuring consistency across different poses.

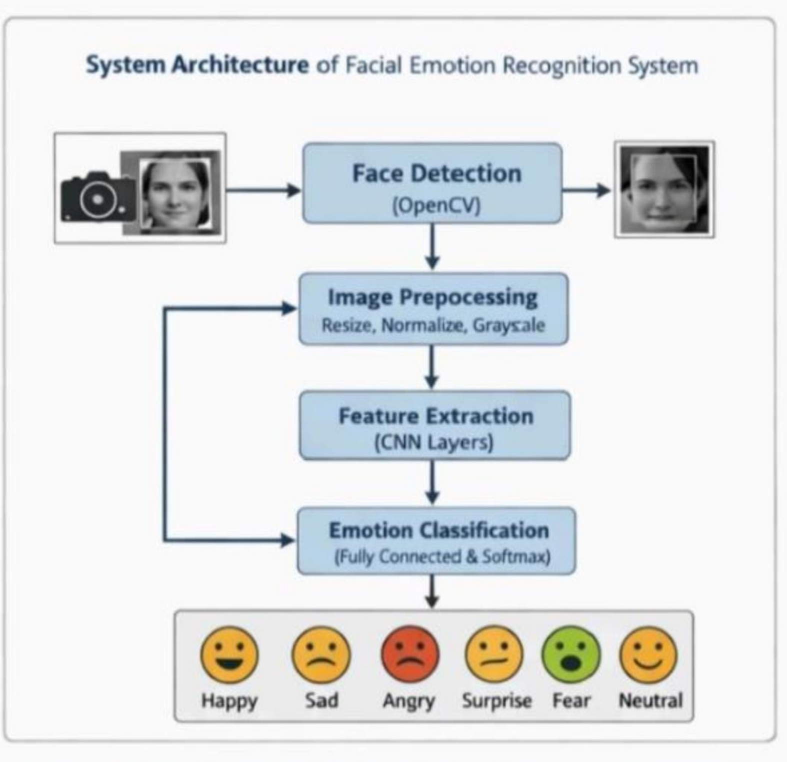

Fig. 1- System Architecture of Facial Emotion Recognition System

After preprocessing, the aligned facial images are passed to the Convolutional Neural Network (CNN) for feature extraction and classification. The CNN automatically learns hierarchical features from facial images, capturing important patterns related to different emotional expressions.

The CNN architecture used for emotion recognition is illustrated in Fig. 2. It consists of multiple convolutional layers that extract spatial features such as edges and textures, followed by batch normalisation and activation functions to enhance learning. Pooling layers are used to reduce spatial dimensions and computational complexity. A dropout layer is included to prevent overfitting, and fully connected layers are used for final classification. The output layer uses a softmax function to classify emotions into categories such as happiness, sadness, anger, fear, surprise, disgust, and neutral.

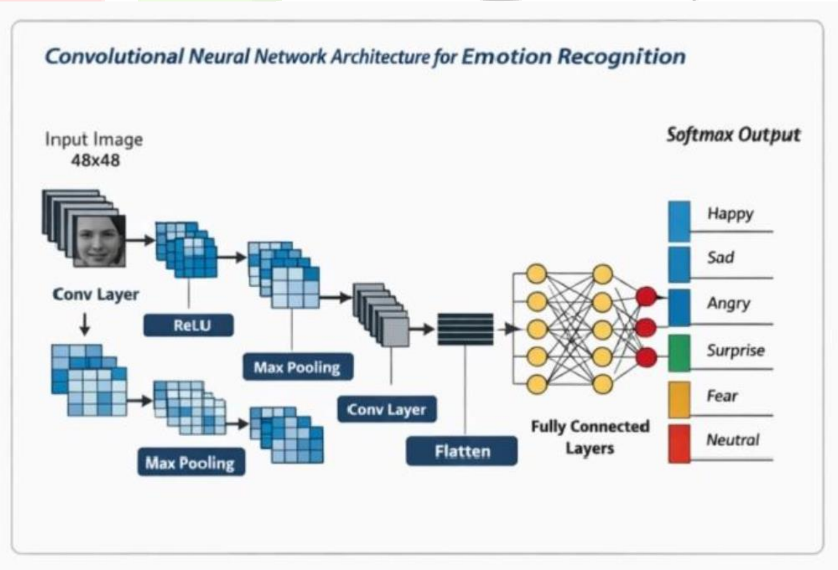

Fig. 2- Convolution Neural Network Architecture for Emotion Recognition

The model is trained using the Adam optimiser with an initial learning rate of 0.0001. Focal loss is employed to address class imbalance by focusing more on difficult samples during training. Data augmentation techniques such as horizontal flipping, rotation, brightness adjustment, and zooming are applied to improve generalisation performance.

For real-time implementation, the system processes video frames continuously and predicts emotions with confidence scores. A smoothing mechanism is applied across frames to reduce fluctuations and improve prediction stability.

DESIGN

The proposed system follows a modular architecture consisting of multiple layers responsible for data processing, feature extraction, classification, and output generation.

The system begins with a data layer that handles dataset loading, splitting, and augmentation. The FER2013 dataset is used, which consists of grayscale facial images categorised into different emotion classes. Data augmentation techniques such as horizontal flipping, rotation, zooming, and brightness adjustment are applied to increase dataset diversity and improve model generalisation.

The model layer implements the CNN architecture using a pre-trained ResNet-50 network. Initially, the pre-trained layers are frozen, and only the classification layers are trained. In later stages, selected layers are fine-tuned to adapt to facial emotion recognition tasks.

The evaluation layer computes performance metrics such as accuracy, precision, recall, and F1-score, along with confusion matrices for each emotion class. Grad-CAM visualisation is used to highlight important facial regions influencing model predictions.

The inference layer performs real-time emotion recognition by processing video frames. Each frame is passed through detection, alignment, and classification stages, and predicted emotions are displayed with confidence scores.

CONCLUSION

This work presents the design and implementation of a deep learning-based facial emotion recognition system using Convolutional Neural Networks. The system demonstrates that transfer learning, combined with data augmentation and optimization techniques, can achieve effective performance in emotion classification tasks.

The model achieves strong performance across benchmark datasets and supports real-time emotion detection with stable predictions. The use of Grad-CAM visualisation improves interpretability and provides insight into the model's decision-making process.

Despite these achievements, challenges such as class imbalance, occlusion, and cross-dataset generalisation remain. Future work will focus on improving minority class detection, enhancing robustness, and exploring advanced architectures for efficient deployment.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to our project guide and faculty members for their continuous support and guidance throughout the development of this project. Their valuable suggestions and encouragement helped us in successfully completing this research work.

We also thank our institution for providing the necessary resources and environment to carry out this project. Finally, we are grateful to all researchers and authors whose work has contributed to the development of facial emotion recognition systems.

REFERENCES

1. D. Pruthviraja, U. Mohan Kumar, S. Parameswaran, V. G. Chowdary and V. Bharadwaj, “Deep Convolutional Neural Network Architecture for Facial Emotion Recognition,” *PeerJ Computer Science*, vol. 10, 2024.
2. V. Ajitha, “CNN-Driven Enhancement in Facial Emotion Recognition Systems,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 2343–2352, 2024.
3. R. Khalid Sahi, “Classifying Feelings Using Facial Expression Recognition,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 3303–3310, 2024.
4. A. Babisha, A. Swaminathan, D. Anuradha, C. Gnanaprakasam and T. Kalaichelvi, “Advancements in Facial Expression Recognition: State-of-the-Art Techniques and Innovations,” *International Journal of Intelligent Systems and Applications in Engineering*, 2024.
5. C. Zhu, “A Review of Facial Expression Recognition Based on Convolutional Neural Networks,” *Applied and Computational Engineering*, 2024.
6. Z. F. Hassan, M. N. Al-Turfi and F. Al-Alawy, “Facial Expression Recognition Enhancement Using Convolutional Neural Network,” *Al-Iraqia Journal for Scientific Engineering Research*, vol. 3, no. 2, 2024.
7. A. Miskow and A. Altahhan, “Emotion Recognition with Facial Attention and Objective Activation Functions,” *arXiv preprint*, 2024.
8. A. Roy, H. Kathania, A. Sharma and A. Dey, “ResEmoteNet: Bridging Accuracy and Loss Reduction in Facial Emotion Recognition,” *arXiv preprint*, 2024.
9. C. Białek, A. Matiolański and M. Grega, “An Efficient Approach to Face Emotion Recognition with Convolutional Neural Networks,” *Electronics (MDPI)*, vol. 12, no. 12, 2023.
10. P. S. Reddi and A. Sri Krishna, “CNN Implementing Transfer Learning for Facial Emotion Recognition,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 4, pp. 35–45, 2023.