Open Access and Peer Review Journal ISSN 2394-2231 https://ijctjournal.org/

Real-Time Human Facial Emotion Recognition System Using Deep Learning

Prachi Urgunde[1], Deepali Gille[2], Gauri Hushangabadkar[3], Neha Bokad[4], Mohini Shrikhande[5], Prof. Pravin Kaware [6]

[1,2,3,4,5,6]*Department Of Computer Science & Engineering,*

HVPM's College of Engineering & Technology, Amravati, Maharashtra, India

--************************----------------------------------

Abstract:

This Facial Emotion Recognition (FER) is a pivotal element of Human & Computer Interaction. In This paper we present the creation of a instantaneous system that can identify and categorize human expressions into seven basic emotions: 1. anger, 2. disgust, 3. fear, 4. happiness, 5. sadness, 6. surprise, and 7. neutral. By leveraging Deep Learning (DL) architectures & specifically Convolutional Neural Networks (CNNs), we achieve high accuracy in varying environmental conditions. The system works by capturing live video from a webcam, detecting faces in each frame, and then using a trained model to recognize and predict emotions instantly in real time.

Keywords — **Artificial Intelligence (AI), Facial Emotion Recognition, Deep Learning, CNN.**

--************************----------------------------------

I. INTRODUCTION

Human sentiments plays a crucial contribution in communication & social interaction. With the progress in Artificial Intelligence and computer vision, machines are now capable of understanding facial expressions and recognizing human emotions. Facial expression is an important way for humans to express emotions. It can be used to recognize and determine the emotional states of a person [1]. Emotion recognition systems analyse facial features & classify them into categories such as 1. happiness, 2. sadness, 3. anger, 4. surprise, & 5. neutral. The movements of muscles convey the expressions of individual to people who see them. This research focuses on developing an AI-powered human emotion detection system capable of identifying live emotions using deep learning techniques.

In an era dominated by AI, the ability of machines to understand & human affect is crucial. From supporting mental health tracking to enabling personalized marketing and improving gaming experiences, FER systems help connect digital technology with human emotions. The primary challenge lies in the complexity of facial expressions, which vary across different ethnicities, lighting conditions, & head poses.

II. Literature Review

Facial Emotion Recognition (FER) is a useful application of Artificial Intelligence and Computer Vision. It is used to identify human emotions by analysing facial expressions. In the early stages, researchers used traditional machine learning methods. These methods depended on manually designed features like Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG), along with classifiers such as Support Vector Machines (SVM) [2]. However, these techniques did not perform well in real-world situations. With the growth of deep learning, Convolutional Neural Networks (CNNs) have greatly improved the accuracy of emotion detection. Many real-time systems have been developed using tools like OpenCV and TensorFlow [3]. To reduce computational cost while keeping good performance, lightweight CNN models were introduced [4]. Models such as MobileNetV3 have further improved efficiency, especially for devices with limited resources [5]. FER has many applications, including smart environments and systems that can respond to human emotions [6]. Common datasets like FER-2013 are widely used for training and testing these models [7]. In addition, face detection methods like the Viola-Jones algorithm are often used as a preprocessing step [8]. Various studies have explored different techniques for facial expression recognition, highlighting the importance of feature extraction and deep learning models [9].

Even with these improvements, some challenges still remain. Factors like changes in lighting, different face angles, and uneven data distribution can affect the overall performance of the system.

Figure 1: Examples of facial expressions representing different human emotions.

III. Methodology

1. Data Collection

To begin, we require a dataset that contains human faces displaying a variety of emotions. For this, we use dataset FER-2013. This dataset contains many images of people with different facial expressions such as 1. happy, 2.sad, 3. angry, 4. surprise, 5. fear, 6. disgust, 7. neutral etc. The dataset should include people of various ages and genders, along with different lighting conditions, to ensure the model performs effectively in real-world scenarios.

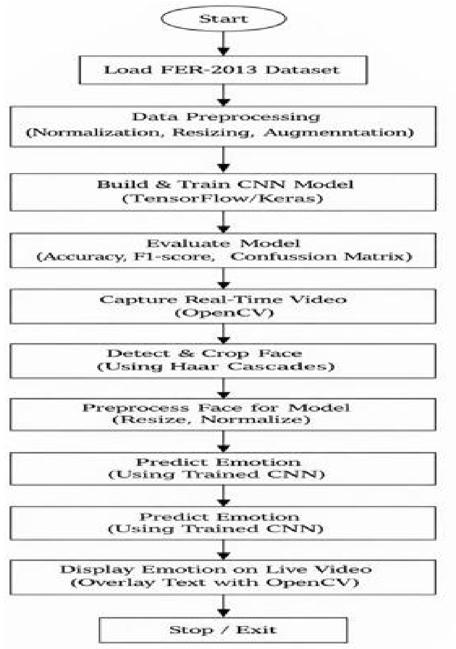

Figure 2: Workflow of the Proposed Emotion Detection Methodology

2. Pre-processing of Data

In this stage, the gathered images are processed to enhance model performance. Face preprocessing proves to be a crucial step for good recognition performance. It helps to remove irrelevant noise and unifies all faces to the same domain. Since we decide to pre-train our deep network model on FER, the detected faces on SFEW are all resized to 48×48 and are transformed to grayscale, which is exactly the same as the FER data [10]. Face detection techniques such as Haar Cascade is applied to extract the facial region. The images are then transformed to Black-and-White in order to reduce algorithmic complexity & dimensions are kept fixed (e.g., 48×48 pixels) for uniformity. The values of pixel are normalized between 0 & 1 for faster convergence. To further improve the model, techniques like expanding, reversing, revolving are applied towards data, which helps increase variation in the dataset and reduces the chances of overfitting.

3. Model Design

The model can assign different importance weights to facial regions, allowing it to focus on visible and informative parts of the face [11]. The system uses a CNN-based architecture consisting of multiple convolutional layers to extract characteristic, accompanied by functions of activation such as ReLU. Pooling layers help decrease the size of the data, which reduces both complexity and computational effort. Dense layers are accustomed to perform classification, & the final layer applies the function name Softmax to assign the I/P image member of the emotion groups.

4. Model Training

Training of model is done on labeled data, with the help of categoric cross-entropy as the loss function & the Adam optimizer to secure productive learning. Training of model is done across several epochs using an appropriate batch size, such as 32 or 64. The data is segregated testing and training sets, to 80:20 ratio, which track the model's performance of model & prevent high variance.

5. Emotion Classification

After training, the model predicts emotions by assigning probability scores to each class using the Softmax function. The emotion having probability higher is chosen for increased prediction output. Classification is a crucial step used to categorize the values into different classes [12].

6. Real-Time Implementation

The trained model is implemented in a real-time setup where it uses input from a webcam. Video frames are captured continuously, & face detection is applied to each frame. The identified face is first preprocessed and is send for prediction corresponding emotion to CNN Model and the output is displayed.

7. Performance Ranking

The metrics such as accuracy, precision, recall, F1-score, & confusion matrix are used for the evaluation of performance of the system and also helps in measuring the effectiveness and reliability.

IV. Emotion Detection using CNN

A Convolutional Neural Network (CNN) helps in analysing visual data and recognizing objects. Convolutional neural networks have emerged as the master algorithm in computer vision in recent years [13]. In this paper, a CNN will be utilized to process an input image containing the face of the user & provide an automatic classification of the user's emotion.

In a CNN model, different layers capture features step by step, learning from simple patterns to more complex representations. [14]The convolutional layer has a set of learnable filters to convolve through the whole input image and produce various specific types of activation feature maps.

The convolutional layer applies learned filters to each input image to detect important areas, such as facial features or other relevant patterns. ReLU, are used in the model to help it grasp and indicate complex relationships within the data.

Subsampling layers reduce the size of the activation maps, which lowers computation time and helps reduce overfitting. The features obtained are then the extracted features are converted into a 1D vector and passed towards the dense layers, where the classification is performed. Convolutional neural networks (CNNs) have become the traditional approach for researchers studying vision and deep learning [15].

In the end, the Softmax layer generates probability scores for each emotion group, & the emotion with the most probability is chosen as output. The model uses grayscale facial images as input and classifies them according to their emotional expressions, such as happy, sad, angry, and neutral.

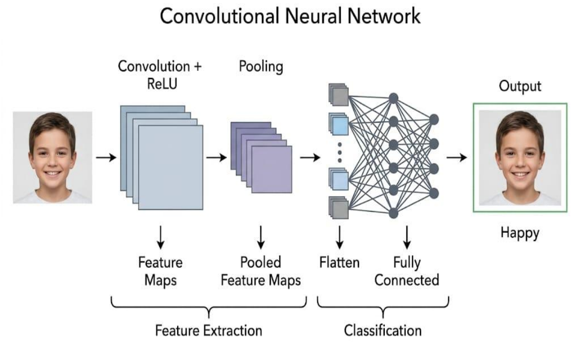

Figure 3: CNN Model for Emotion Detection.

V. System Architecture

This system architecture outlines all of the individual modules & their interactions that will comprise live emotion recognition system. TensorFlow uses dataflow graphs to represent computation, shared state, and the operations that mutate that state[16].Each module in recognition system communicates directly with every other module as the system is made up of sequential components processing input data & generating output data.

These modules include; Input Acquisition, Face Detection, Preprocessing, Emotion Detection, & Result Visualization. The system captures video from a webcam that is continually streaming video frames into the system.

Each frame captured by the webcam is processed by the face detection module, identifying & extracting the facial area. Once the facial area has been identified, the extracted facial area is processed by the preprocessing module to enhance the resolution of the image prior to inputting the input into the CNN used for emotion classification.

Once the facial area has been pre-processed, the image is sent to the emotion detection module where the emotion of the subject is classified. The classified emotion is then viewed on the screen through a user interface.

In addition to displaying the classified emotion on the screen, the system also records the classified emotion for further analysis & viewing. This ensures an easy & continuous data flow among modules allowing for high quality real-time performance.

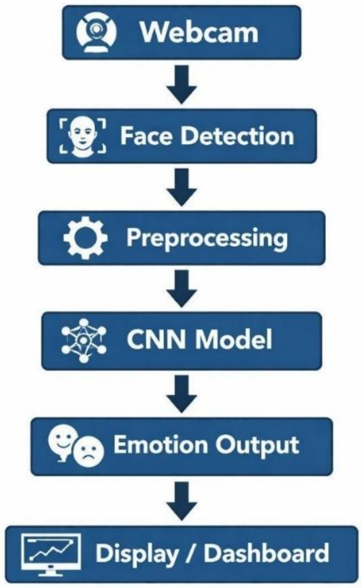

Figure 4: System Architecture & Data Flow Diagram of the Live Human facial emotion Recognition system

VI. Results

The proposed AI-powered human emotion detection approach was evaluated on a facial expression dataset containing seven emotion classes: 1.angry, 2.disgust, 3.fear, 4.happy, 5.neutral, 6.sad, & 7.surprise. The dataset was divided into training & validation sets containing **28,709 images** and **178 images** respectively and also resized to **48 × 48 black-and-white format**, trained with **batch size of 64**.

A Convolutional Neural Network (CNN) model was used for emotion classification. Based on the preserved notebook outputs, the final model achieved a **trained accuracy of 80.42%** & a **testing accuracy of 65.34%**. The gap between training & testing accuracy shows that the model learned meaningful emotion-related patterns, but some degree of overfitting may still be present. Even human-level accuracy in facial emotion classification is approximately 65%, which highlights the complexity of the task [17].

The class-wise execution of the model was analyzed using precision, recall, & F1-score. Among all emotion categories, the **happy** class achieved the best performance with a **precision of 0.84, recall of 0.84, & F1-score of 0.84**. The **surprise** class also performed well with an **F1-score of 0.76**. These results indicate that highly expressive & visually distinct emotions are easier for the CNN model to classify.

On the other h&, the model showed weaker performance on classes such as **fear**, **sad**, & **neutral**. In particular, the **fear** class obtained the lowest **F1-score of 0.44**, suggesting that the facial features of fear may overlap with other emotions, making classification more difficult. Similarly, the **sad** & **neutral** classes produced moderate results, showing that subtle expressions are more challenging for the model to distinguish accurately.

The **disgust** class had very high precision (**0.86**) but lower recall (**0.46**) which means that the model predict disgust correct, but it failed to identify many actual disgust samples. This imbalance may be due to the smaller number of disgust images in the dataset compared to other classes.

Generally, the replica achieved an **precision of approximately 64%**, with a **macro average F1-score of 0.62** & a **weighted average F1-score of 0.63**. These results demonstrate that the proposed model is effective for recognizing common facial emotions, especially those with strong visual characteristics. However, performance can be further improved by using data augmentation, transfer learning, class balancing techniques, & deeper neural network architectures.

Result Table

Emotion	Precision	Recall	F1-Score	Support
Angry	0.54	0.58	0.56	958
Disgust	0.86	0.46	0.60	111
Fear	0.52	0.38	0.44	1024
Happy	0.84	0.84	0.84	1774
Neutral	0.57	0.62	0.59	1233
Sad	0.50	0.53	0.52	1247
Surprise	0.74	0.79	0.76	831

Overall Performance Table

Metric	Value
Training Accuracy	80.42%
Testing Accuracy	65.34%
Overall Accuracy	64%
Macro Average F1-Score	0.62
Weighted Average F1-Score	0.63

VII-Conclusion

In this paper, a live human emotion detection system on artificial intelligence & deep learning has been presented. The project successfully developed an accurate and efficient system for detecting and classifying emotions through facial expressions [18]. The system utilizes computer vision techniques & a Convolutional Neural Network (CNN) to accurately recognize & classify human emotions.

The proposed approach is capable of processing live video I/P & providing live emotion predictions. The results demonstrate that the system performs efficiently under different conditions & is used in human-computer interaction, mental health monitoring, & smart user interfaces. Deep convolutional neural networks have achieved robust performance in facial expression recognition across different environmental conditions [19].

Future improvements may include increasing model accuracy, supporting more emotion categories, & enhancing performance under complex real-world conditions. Future research directions include multimodal information fusion, personalized emotion recognition, and interdisciplinary cooperation [20].

REFERENCES

[1] F. Zhang, T. Zhang and Q. Mao, "Real-Time Facial Expression Recognition Based on Deep Learning," IEEE International Conference on Intelligent Computing and Communication Technologies, 2021.

[2] R. Praveena, V. Bavan Kumar, S. Diwahar, S. Gunaseelan, & S. Suriya Moorthi, "Emotion Detection with Facial Feature Recognition Using CNN & OpenCV," International Journal of Innovative Research in Technology (IJIRT), vol. 11, no. 12, May 2025.

[3] N. Zhou, R. Liang, & W. Shi, "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection," IEEE Access, vol. 9, pp. 1–10, Jan. 2021.

[4] S. Depuru, A. Nandam, P. A. Ramesh, M. Saktivel, K. Amala, & Sivanantham, "Human Emotion Recognition System Using Deep Learning Technique," Journal of Pharmaceutical Negative Results, vol. 13, no. 4, 2022.

[5] X. Liang, J. Liang, T. Yin, & X. Tang, "A Lightweight Method for Face Expression Recognition Based on Improved MobileNetV3," IET Image Processing, vol. 17, no. 6, pp. 1798–1811, Mar. 2023.

[6] K. Sudha, S. P. Ashvitha, & D. Nivethashree, "Smart Study Mirror with Mood Gauge," International Journal of Research Publication & Reviews (IJRPR), vol. 6, no. 11, pp. 5995–5998, Nov. 2025.

[7] I. Goodfellow et al., "Challenges in Representation Learning: A Report on Three Machine Learning Contests," FER-2013 Dataset, 2013.

[8] P. Viola & M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition (CVPR), 2001.

[9] I. M. Revina and W. R. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," Journal of King Saud University – Computer and Information Sciences, vol. 33, no. 6, pp. 619–628, 2021.

[10] Z. Yu and C. Zhang, "Image Based Static Facial Expression Recognition with Multiple Deep Network Learning," Proceedings of the ACM International Conference on Multimodal Interaction, 2015.

[11] Y. Li, J. Zeng, S. Shan and X. Chen, "Occlusion Aware Facial Expression Recognition Using CNN with Attention Mechanism," IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2439-2450, 2019.

[12] S. K. Khare, V. Blanes-Vidal, E. S. Nadimi, and U. R. Acharya, "Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations".

[12] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions,"CVPR, 2017.

[14] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," IEEE Transactions on Affective Computing, vol. 13, no. 3, pp.1195-1215,2022,doi:10.1109/TAFFC.2020.2981446.

[15] A. Mollahosseini, D. Chan, and M. H. Mahoor,"Going Deeper in Facial Expression Recognition using Deep Neural Networks,"
IEEE Winter Conference on Applications of Computer Vision (WACV), 2016.

[16] M. Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning," OSDI, 2016.

[17] O. Arriaga, M. Valdenegro-Toro and P. Plöger, "Real-Time Convolutional Neural Networks for Emotion and Gender Classification," arXiv preprint arXiv:1710.07557, 2017.

[18] Dr. L. Math and N. G., "Facial Emotion Recognition in Real Time Using Deep Learning," International Research Journal of Engineering and Technology (IRJET), vol. 11, no. 07, pp. 805–812, July 2024.

[19] S. Minaee et al., "Facial Expression Recognition Using Deep Learning: A Survey," arXiv preprint arXiv:1902.01019, 2019.

[20] C. Qian, J. A. L. Marques, and A. R. de Alexandria, "Real-time emotion recognition based on facial expressions using artificial intelligence techniques: A review and future directions," Multidisciplinary Reviews, vol. 8, e2025328, Apr. 2025.