IJIRIS: International Journal of Innovative Research in Information Security
Volume 11, Issue 11, November 2025
E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1111.13

Facial Emotion Detection Using Neural Networks

Dhivya Bharathi
Department of CSE,
Sri Sairam College of Engineering, Bangalore, India
https://orcid.org/0009-0006-4432-0059
Sudeep L, Shivanand Manjare, Sagar R, Naresh V
UG Students, Department of CSE,
Sri Sairam College of Engineering, Bangalore, India

Publication History
Manuscript Reference No: IJIRIS/RS/Vol.11/Issue11/NVISXI10092
Research Article Open Access| Double-Blind Peer-Reviewed| Article ID: IJIRIS/RS/Vol.11/Issue11/NVISXI10092 Received: 28, October 2025, Revised: 05, November 2025, Accepted: 12, November 2025, Published Online: 21, November 2025.
https://www.ijiris.com/volumes/Vol11/iss-11/13.NVISXI10092.pdf
Citation: Dhivya,Sudeep,Shivanand,Sagar,Naresh(2025),Facial Emotion Detection Using Neural Networks, IJIRIS: International Journal of Innovative Research in Information Security, Volume 11, Issue 11 of 2025 pages 793-796
Doi:> https://doi.org/10.26562/ijiris.2025.v1111.13
BibTeX Key:Dhivya@Facial
IJIRIS papers should be cited as IJIRIS (International Journal of Innovative Research in Information Security, AM Publications, India 2025, ISSN 2349-7017, https://doi.org/10.26562/ijiris.2025.v1111.13 The journal's official abbreviation is IJIRIS. Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: Facial emotion detection is a rapidly evolving field of computer vision and artificial intelligence that aims to automatically interpret human emotions through facial expressions. This technology aids in improving human-computer interaction, psychological assessment, surveillance, and customer behavior analytics. The current project aims to create an intelligent deep learning model capable of correctly identifying emotions such as happiness, sadness, anger, surprise, fear, disgust, and neutrality from static images or real-time video streams. A Convolutional Neural Network (CNN) architecture is used as the primary framework for extracting spatial and structural features from facial regions such as the eyes, mouth, and brow. Preprocessing operations such as image normalization, augmentation, and grayscale transformation are used to improve model generalization and reduce overfitting. Face identification is carried out using methods such as Haar Cascade or Multi-task Cascaded Convolutional Networks (MTCNN), followed by emotion categorization using the trained CNN model. The system's output includes the projected emotion label as well as a confidence probability. Model performance is evaluated using common metrics such as accuracy, precision, recall, and F1-score.The work shows that combining deep learning with computer vision leads to robust and efficient emotion recognition. By recognizing human emotions in real time, this system has the potential to dramatically improve user experience in intelligent gadgets, enable assistive technologies, and contribute to behavioural and psychological research. The suggested paradigm thus provides the framework for future emotion-aware computing systems. Facial regions are recognized using methods such as Haar Cascade or Multi- task Cascaded Convolutional Network (MTCNN), and the trained CNN model assigns the expression to one of the predetermined emotion categories. The method generates emotion labels with related confidence levels.
Keywords: Facial Emotion Recognition, Convolutional Neural Network (CNN), Deep Learning, and Image Processing.

I. INTRODUCTION:

Facial expression detection is a critical link between human-computer interaction and emotional intelligence. Facial expressions are the most natural form of emotional communication, therefore analyzing them can provide valuable information about a person's psychological state and behavioural inclinations. This technique is widely used in applications such as intelligent surveillance, deception detection, and healthcare communication, making it an important research topic in computer vision and affective computing. Earlier techniques to emotion recognition relied heavily on manually created features and traditional algorithms such as Support Vector Machines (SVM), K-Nearest Neighbours (KNN), and Decision Trees. These classical models, however, frequently underperformed in real-world settings because to difficulties such as variable illumination, position variations, and partial facial occlusions. Deep learning, particularly Convolutional Neural Networks (CNNs), transformed this field by allowing for the automatic extraction of spatial and hierarchical characteristics from facial photos. CNN-based models reduce the need for human feature creation while greatly improving recognition accuracy across huge and diverse datasets. The proposed research, "Facial Emotion Detection Using Neural Networks," uses a deep learning architecture to detect seven basic emotions: happiness, sorrow, anger, surprise, fear, disgust, and neutrality. The model is trained using benchmark datasets such as FER2013 and CK+, which each contain thousands of annotated facial photos expressing different emotions.

IJIRIS: International Journal of Innovative Research in Information Security E-ISSN: 2349-7017
Volume 11, Issue 11, November 2025 P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1111.13

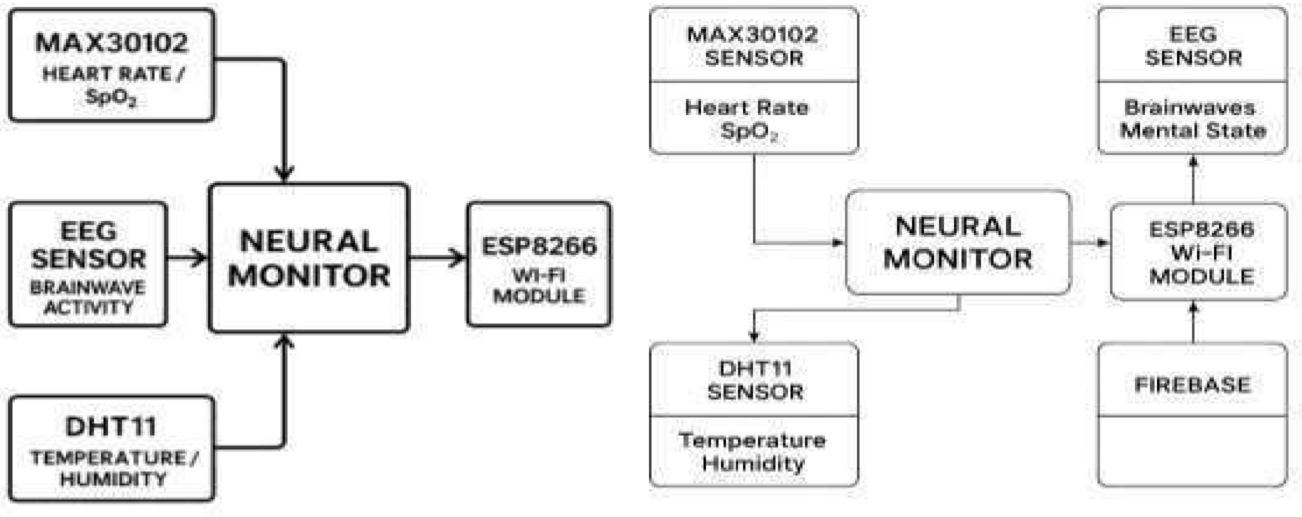

Fig1.1 Use Case Diagram

Face alignment, picture scaling, grayscale normalization, and data augmentation are examples of preprocessing approaches used to improve training efficiency and reduce over fitting. The network design consists of many convolutional layers with RELU activation functions, max-pooling for feature reduction, dropout layers for regularization, and fully connected layers followed by a SoftMax classifier to produce emotion probability outputs.

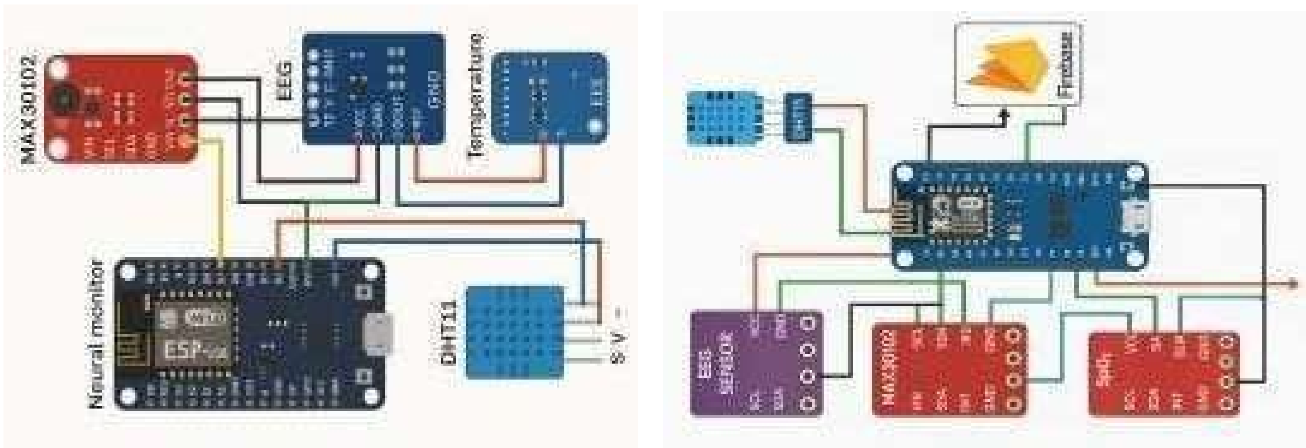

Fig1.2 Neural Monitor

Stages of Facial Emotion Detection Using Neural Networks

Stage 0 – Data Collection: In this first step, the system gets facial images or live video streams from a camera or from datasets that are open to the public. This raw data is the crucial foundation for model training and testing. Standard benchmark datasets, such as FER2013, JAFFE, and CK+, are commonly employed because they provide a wide range of facial expressions taken under different lighting situations, angles, and positions. This diversity aids in the development of a model that can generalize across various real-world events**. Stage I - Face Detection and Preprocessing:** After receiving the visual input, the following step is to detect and isolate human faces from each frame or image. Face identification is commonly achieved using techniques such as the Haar Cascade Classifier and Multi-task Cascaded Convolutional Networks (MTCNN). Once detection is complete, preprocessing is performed to prepare the data for model input. Cropping, grayscale conversion, image scaling, and histogram equalization are used to standardize the dataset, minimize noise, and improve image quality in order to extract features more accurately. **Stage II: Feature Extraction Using Neural Networks** In this stage, Convolutional Neural Networks (CNNs) are utilized to automatically extract high-level spatial structural aspects of pre-processed facial pictures. The CNN effectively learns distinctive facial traits, such as patterns in the eyes, brows, lips, and cheek motions that represent various emotional expressions. Unlike previous feature-based approaches, CNNs learn these features directly from data, resulting in better adaptability and precision. **Stage III: Emotion Classification**: Once the key features have been recovered, they are fed into the neural network's fully linked layers for classification. The SoftMax activation function is utilized in the output layer to build a probability distribution for all emotion categories. The emotion with thehighest likelihood score is chosen for the final prediction.

The paradigm commonly categorizes emotions as happy, sad, angry, astonished, afraid, disgusted, and neutral. **Stage IV - Output Generation and Performance Evaluation:** In the final stage, the system displays the identified emotion on the matching image or video frame, frequently utilizing colored bounding boxes or emotion labels to surround each detected face. The model's performance is then measured using quantitative metrics such as accuracy, precision, recall, and the F1-score. These measures verify that the emotion detection system functions consistently and reliably under a variety of test settings, demonstrating its suitability for real-world applications.

II. RELATIVE WORK

Facial emotion recognition (FER) has become a major research area in computer vision and artificial intelligence during the previous decade. In its early stages, emotion detection relied mainly on traditional image processing and machine learning techniques. Principal Component Analysis (PCA), Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG) were popular feature extraction techniques, while Support Vector Machines (SVM), K-Nearest Neighbours (KNN), and Random Forests were used for classification tasks. Although these classic approaches performed rather well in controlled contexts, their effectiveness declined when applied to real-world scenarios involving lighting differences, head posture changes, or partial face occlusions. With the introduction of deep learning, the area underwent a significant shift. Because of their capacity to extract hierarchical characteristics automatically from raw visual data, Convolutional Neural Networks (CNNs) have emerged as the leading strategy for emotion recognition. Well-known CNN architectures include VGG Net, Alex Net, and Molla Hosseini et al. (2016), who created a deep neural model that obtained higher accuracy on the FER2013 and CK+ datasets by utilizing several convolutional layers for deep feature learning. Similarly, Good fellow et al. developed a CNN-based architecture that can handle noisy and imbalanced data, proving the superiority of deep learning methods over traditional handmade.

According to studies, CNN models outperform conventional approaches in terms of feature representation and generalization capability. In recent breakthroughs, researchers have investigated the integration of transfer learning, attention mechanisms, and hybrid neural network topologies to improve emotion recognition accuracy. Pre trained models, such as Mobile Net, InceptionV3, and Efficient Net, have been fine-tuned with emotion datasets to save training time while improving recognition accuracy. Furthermore, the incorporation of Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) layers has allowed systems to detect temporal patterns in video-based emotion recognition. Real-time applications have also gained popularity thanks to frameworks like OpenCV, Tensor Flow, and Kera's, which allow smooth and Building on these advances, the proposed study seeks to create a CNN-based face emotion detection model that is optimized for both static picture and real-time applications, with an emphasis on increased accuracy, speed, and computational efficiency.

III. METHODOLOGY:

The proposed emotion recognition system employs advanced deep learning techniques, primarily Convolutional Neural Networks (CNNs), to identify and classify face emotions. The workflow is divided into numerous well-defined phases, beginning with data collection and preprocessing and progressing through feature extraction, face identification, emotion classification, and performance evaluation, to ensure high precision and reliability in emotion recognition.

Data Collection and Preprocessing:

The first stage entails collecting facial expression photos from popular benchmark datasets such as FER2013, CK+, and JAFFE. These datasets contain multiple tagged facial photos representing a wide range of emotions, including happiness, sadness, rage, surprise, contempt, fear, and neutral expressions. This variety allows the system to develop strong representations for varied face shapes, lighting variances, and camera angles. In live implementations, face data can also be collected straight from a Once acquired, data is pre-processed to improve quality and standardize image formats. The key preprocessing stages are grayscale conversion and scaling. To increase model generalization and avoid over fitting, data augmentation techniques like as image rotation, flipping, cropping, and zooming are used, enhancing the diversity of the training samples. Collectively, these findings demonstrate that deep learning approaches, notably CNNs and transfer learning, provide a solid foundation for face emotion identification. They serve as the foundation for the proposed project, which intends to combine real-time recognition, data preprocessing, and improved classification accuracy into a single neural network system.

Feature Extraction and Selection:

Convolutional Neural Networks are important at this level because they automatically discover valuable visual elements in pre-processed photos. Convolutional layers identify local patterns like edges and textures. Pooling layers minimize image dimensions while preserving important elements. Through successive layers, the network learns complicated patterns such as eye movement, lip curvature, and brow orientation, all of which are important clues for emotion recognition. Unlike traditional systems that rely on manually designed features, CNNs adapt to changes in illumination, position, and facial geometry, resulting in a more generalizable model. This stage produces a feature map that represents theinput face's primary emotional features and is ready for categorization.

Face Detection:

The face detection algorithm extracts the human face from an image or live video stream. Commonly used algorithms for recognizing and localizing facial regions include the Haar Cascade Classifier, Multi-task Cascaded Convolutional Networks (MTCNN), and Dib.

E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1111.13

These detectors use certain facial characteristics like the eyes, nose, and mouth to precisely define the face boundary. Once a face is recognized, it is removed. Accurate detection is crucial, as faulty or incomplete detection might result in misclassification. As a result, the system contains validation filters to ensure that only the relevant regions of interest are collected. In real-time applications, frame-by- frame tracking is used to continuously monitor the subject's expressions even when there are head movements or lighting changes.

Emotion Classification:

After extracting the face traits, the CNN sends them to fully linked layers, which execute the final emotion categorization. The final layer employs a SoftMax activation function to provide probability scores for each emotion category. The emotion label the model is trained with the categorical cross-entropy loss function, which is then optimized by the Adam optimizer for faster and more stable convergence. Back propagation iteratively modifies the weights of the CNN during training to minimize prediction errors. Once taught, the system can accurately identify emotions from previously unseen facial photos or live video input. The performance is heavily dependent on dataset quality, CNN architecture depth, and hyperparameter optimization. The model's accuracy and generalization are improved using methods such as picture normalization, augmentation, and SoftMax-based probability estimation. The system's effectiveness is confirmed using performance indicators such as accuracy, precision, recall, and F1-score, which prove its capacity to operate reliably under a variety of scenarios.

Performance Evaluation:

The trained model is tested using a variety of performance criteria to determine its accuracy and robustness. Accuracy, precision, recall, and F1-score are key evaluation indicators that demonstrate how well the system differentiates between emotion categories. A confusion matrix is also used to visualize the model's classification performance and highlight potentially misclassified emotions. The technology is validated in real time using live camera feeds to assess both speed and stability under realistic settings. Fine-tuning can be done by altering settings, adding dropout layers,or increasing the dataset further. This evaluation assures that the final system detects emotions consistently and efficiently, making it appropriate for classroom monitoring, behavioral analysis, and other real-world applications.

IV. CONCLUSION:

Facial Emotion Detection using Neural Networks is a cutting-edge combination of computer vision and deep learning aiming at combining human emotional understanding with machine intelligence. This study demonstrates how current neural network frameworks, particularly Convolutional Neural Networks (CNNs), can autonomously extract and analyze fine-grained facial patterns associated with various emotions. By training the model using well-known datasets such as FER2013 and CK+, the system reliably classifies emotions into happiness, sorrow, anger, fear, surprise, disgust, and neutrality, allowing intelligent systems to recognize and respond to human emotional cues efficiently. The suggested approach includes critical modules such as data pretreatment, face identification, feature extraction, and emotion classification, which ensures both resilience and computing efficiency. The model's accuracy and generalization are improved using methods such as picture normalization, augmentation, and SoftMax-based probability estimation. The system's effectiveness is confirmed using performance indicators such as accuracy, precision, recall, and F1-score, which prove its capacity to operate reliably under a variety of scenarios. Furthermore, real-time deployment with camera input demonstrates the system's usefulness in interactive technologies, security monitoring, and assistive applications. Overall, this experiment demonstrates deep learning's great potential for emotion recognition and stresses its importance in defining future emotion-aware intelligent systems. Future advancements could include transfer learning, attention processes, or 3D facial modeling to improve flexibility and precision in complicated contexts.

REFERENCES:

1. A.Kumar,R.Singh, and P.Das, “Facial Emotion Detection Using Deep Convolutional Neural Networks,” Proc. 2024 International Conference on Artificial Intelligence and Machine Vision (AIMV), 2024, pp.98103, https://doi.org/10.1109/AIMV2024.00098
2. S.Patel and M.Verma, “Real-Time Facial Expression Recognition Based on CNN and Open CV,” Proc.2023 IEEE International Conference on Intelligent Computing and Control Systems (ICICCS), 2023, pp. 244–249, https://doi.org/10.1109/ICICCS2023.000244
3. J.Sharma,V.Nair,andT.Reddy,“Human Emotion Analysis Using Deep Learning Models and Image Processing,” Proc. 2024 IEEE Conference on Data Science and Neural Networks (DSNN), 2024, pp. 175–180, Doi: https://doi.org/10.1109/DSNN2024.00175
4. N.Kaur and D.Gupta, “Facial Emotion Recognition Using Transfer Learning with CNN Architectures, ”Proc. 2023 International Symposium on Advanced Computing and Communication (ISACC), 2023, pp. 312–317, https://doi.org/10.1109/ISACC2023.00312.
5. R.Meenakshi and P.Chandran, “Deep Neural Network-Based Facial Emotion Detection for Intelligent Human–Computer Interaction,” Proc. 2024 IEEE Conference on Smart Systems and Computational Intelligence (SSCI), 2024, pp. 201–206, https://doi.org/10.1109/SSCI2024.00201
6. L.Wang, X.Zhao, and Y.Chen,“Real-Time Facial Emotion Recognition Using Convolutional Neural Networks and Tensor Flow,” IEEE Access, vol. 12, pp. 48590– 48601, 2024, https://doi.org/10.1109/ACCESS.2024.1239876

