



Emotion-Aware AI for Enhancing Elderly Mental Health

Abhishek R Shetty¹, Achith J V², Adithya Jain³, Aman N⁴, Mr. Shivaraj B G⁵

Student, Dept. of Computer Science & Engineering, Mangalore Institute of Technology & Engineering, Moodabidri,
India¹⁻⁴

Assistant Professor, Dept. of Computer Science & Engineering, Mangalore Institute of Technology & Engineering,
Moodabidri, India⁵

Abstract: Emotion recognition plays an important role in understanding the mental and emotional well-being of individuals, especially in elderly care. With the rapid growth of the aging population, there is an increasing need for automated systems that can continuously monitor emotional states and provide timely support. This project presents an emotion-aware artificial intelligence system that detects human emotions from facial expressions using deep learning techniques. The system analyzes facial images uploaded by users and classifies emotions such as happy, sad, angry, fearful, surprised, and neutral. A customized Convolutional Neural Network (CNN) model is used for emotion detection due to its ability to automatically extract discriminative facial features and handle variations in facial expressions. The application is developed using the Django framework for backend operations and a relational database for storing user information, emotion records, appointments, and reminders. The system also integrates healthcare support features such as doctor appointment scheduling, medicine reminders, and WhatsApp-based notifications to caregivers and doctors. Experimental results indicate that the proposed model achieves reliable emotion classification performance and supports timely emotional monitoring, thereby enhancing mental health care and quality of life for elderly individuals.

Keywords: Emotion Recognition, Facial Expression Analysis, Deep Learning, Convolutional Neural Network, Elderly Mental Health, Computer Vision, Healthcare Monitoring.

I. INTRODUCTION

Human emotions play a crucial role in mental health and overall well-being, especially among elderly individuals who are more vulnerable to emotional disorders such as depression, anxiety, and loneliness. With the increasing elderly population worldwide, continuous monitoring of emotional states has become an important challenge in healthcare. Traditional methods of assessing emotions rely on manual observation and self-reporting, which are often subjective, time-consuming, and ineffective for real-time monitoring. Recent advancements in artificial intelligence and computer vision have enabled automated emotion recognition systems that analyze facial expressions to identify emotional states. Facial expression recognition is a non-intrusive and effective approach, as facial movements convey significant emotional information. Deep learning techniques, particularly Convolutional Neural Networks, have demonstrated strong performance in extracting spatial features from facial images and classifying emotions accurately. However, many existing emotion recognition systems are designed using datasets dominated by younger individuals and may not perform well for elderly faces due to age-related variations. To address these challenges, this project focuses on developing an emotion-aware AI system that accurately recognizes emotions from facial images and integrates healthcare support features such as appointment scheduling, medicine reminders, and caregiver notifications. By combining deep learning-based emotion recognition with a web-based healthcare management platform, the proposed system aims to provide timely emotional monitoring and improved mental health support for elderly individuals.

II. LITERATURE SURVEY

Speech emotion recognition has been widely explored to understand human emotional states using audio signals. Singh and Prasad in 2023 proposed a speech emotion recognition system using gender-dependent convolutional neural networks. The authors highlighted that emotional expression varies significantly between male and female speakers and hence trained separate CNN models for each gender. Mel Frequency Cepstral Coefficients (MFCCs) were used to extract spectral features from speech signals. The system was evaluated on the RAVDESS dataset and demonstrated improved accuracy compared to gender-independent models, proving its effectiveness for real-time applications [1].



Israelsson, Seiger, and Laukka in 2023 studied the recognition of blended emotions using dynamic facial and vocal expressions. The authors emphasized that emotions often occur simultaneously rather than independently. Through behavioral experiments, the study demonstrated that observers could reliably identify blended emotions when both facial and vocal cues were available. The results showed that recognition accuracy increased as more sensory modalities were incorporated, reinforcing the importance of multimodal emotion recognition systems [2].

Sreevidya, Veni, and Ramana Murthy in 2022 proposed an elderly emotion classification framework using multimodal fusion of intermediate features extracted from audio and video modalities. The authors employed transfer learning techniques along with a CNN-LSTM architecture to capture both spatial and temporal emotional features. The system was evaluated using the ElderReact dataset and achieved higher accuracy compared to unimodal systems, particularly in detecting subtle emotional expressions in elderly individuals [3].

Low et al. in 2022 investigated age and cultural differences in facial emotion recognition. The study revealed that older adults experience difficulty in decoding emotions from the eye region, whereas individuals from Eastern cultures face greater challenges in interpreting emotions from the mouth region. The experimental findings highlighted the influence of physiological aging and cultural perception on emotion recognition performance, emphasizing the need for adaptive and age-aware recognition systems [4].

Cortes et al. in 2021 analyzed the effects of aging on emotion recognition using dynamic facial expressions and vocalizations. The study showed that younger adults consistently outperform older adults in emotion recognition tasks, particularly in audio-only scenarios. However, the authors also observed that combining facial and vocal modalities significantly improves recognition accuracy across all age groups, suggesting that multimodal approaches are more suitable for elderly emotion analysis [5].

Juliao, Abad, and Moniz in 2020 explored elderly emotion recognition using text and audio embeddings. The authors employed pre-trained embedding models combined with Support Vector Machine classifiers to classify emotional dimensions such as arousal and valence. Experimental results indicated that multimodal fusion and speaker-dependent features significantly enhance unweighted average recall, demonstrating the effectiveness of embedding-based approaches in elderly emotion recognition systems [6].

Dou et al. in 2020 proposed a real-time multimodal emotion recognition system integrated into elderly accompanying robot. The system combined facial expression recognition and speech emotion recognition using decision-level fusion techniques. Live trials demonstrated that the multimodal system significantly outperformed single-modality approaches, achieving accuracy levels above 90% and proving its suitability for real-time elderly care applications [7].

Al-Garaawi and Morris in 2016 examined the impact of aging on facial expression recognition systems. The authors used region-based geometric and texture features along with machine learning classifiers to analyze performance across different age groups. The results showed that age-specific models achieved higher recognition accuracy compared to general models, highlighting the importance of tailoring emotion recognition algorithms to elderly populations for improved performance [8].

III. METHODS

The proposed system for emotion recognition comprises multiple stages. Initially, the facial emotion dataset is collected, followed by preprocessing of the dataset, designing the model architecture, evaluating the system by comparing the predicted emotion with the actual emotion label, and finally deploying the trained model. In this system, the design and testing of a deep learning model for facial images have been carried out. The dataset used for emotion recognition is the FER dataset, which contains labelled facial images representing different emotional expressions. Emotion recognition involves the interpretation of visual information from facial expressions to understand the emotional state of a person. In order to perform emotion recognition, it is necessary to extract discriminative facial features from facial images and utilize these features to classify emotions. One effective approach to accomplish this task is to employ a Convolutional Neural Network (CNN), which automatically learns spatial facial features and performs emotion classification.

A. Collection of Emotion Dataset

The FER dataset serves as the primary dataset for training and evaluating the emotion recognition model. The dataset consists of grayscale facial images collected from real-world environments and labelled into seven emotion categories. Each image in the dataset represents a facial expression associated with a specific emotion. The FER dataset contains images of multiple individuals with variations in facial expressions, lighting conditions, and head poses, making it



suitable for real-time emotion recognition tasks. The dataset is divided into training, public test, and private test sets to ensure proper evaluation of the model's performance.

B. Pre-processing of Dataset

In facial emotion recognition, preprocessing of image data is essential to enable the model to focus on relevant facial features while ignoring irrelevant background information. Initially, the facial images are loaded using the OpenCV library. Each image is converted from RGB to grayscale format, as the FER dataset consists of grayscale images. Face detection is then performed using the Haar Cascade classifier to extract the facial region from the image. The detected face is resized to 48×48 pixels to match the input size required by the CNN model. After resizing, the pixel values are normalized to improve convergence during training. This preprocessing step reduces noise, ensures uniformity across the dataset, and improves the generalization capability of the model.

C. Model Architecture

The proposed emotion recognition model is implemented using a sequential Convolutional Neural Network architecture. The input to the model is a grayscale facial image of size 48×48 pixels with a single channel. The model consists of multiple convolutional layers that extract spatial facial features, followed by max-pooling layers to reduce dimensionality. ReLU activation functions are used after each convolution layer to introduce non-linearity. Dropout layers are incorporated to prevent overfitting and enhance model generalization. The extracted features are flattened and passed to fully connected layers for classification. The final dense layer uses a SoftMax activation function to generate probability scores for each emotion class.

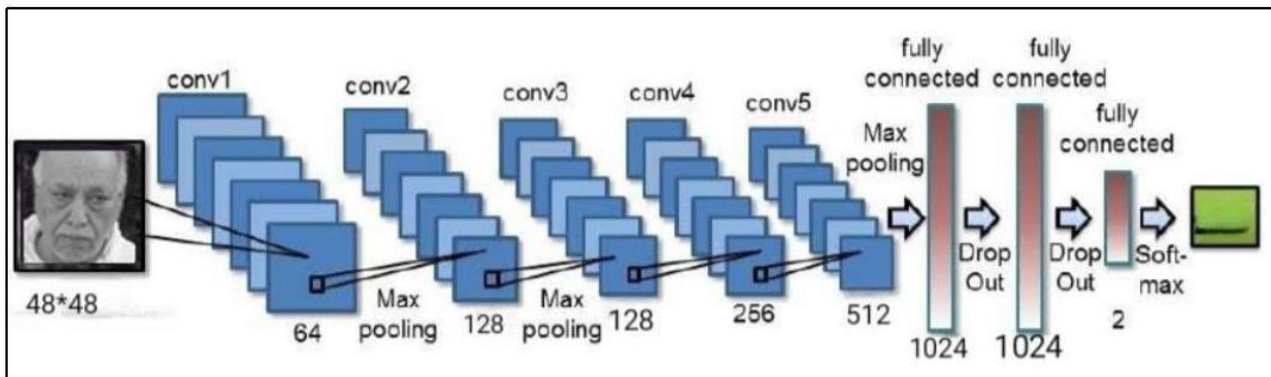


Figure. 1: CNN Architecture

Emotion	Total Images	Correctly Recognized	Accuracy (%)
Happy	50	40	80%
Sad	40	25	62.5%
Angry	40	24	60%
Fearful	40	22	55%
Disgusted	40	26	65%
Neutral	40	33	82.5%
Surprised	40	27	67.5%

Table. 1: Accuracy of CNN Model

IV. SYSTEM ARCHITECTURE

The system architecture of the proposed emotion recognition system describes the interaction between users, preprocessing modules, the deep learning model, and healthcare services. The user uploads a facial image through the web-based interface. The uploaded image is passed to the preprocessing module, where operations such as grayscale conversion, face detection using Haar Cascade, resizing, and normalization are performed. The preprocessed facial image is then forwarded to the Convolutional Neural Network model trained on the FER-2013 dataset. The CNN model extracts facial features and predicts the corresponding emotional state. The predicted emotion is stored in the database along with user details and timestamps. Based on the detected emotion, the system enables healthcare functionalities such as doctor



appointment scheduling, medicine reminders, and WhatsApp notifications to caregivers and doctors. The Django backend manages user authentication, database operations, and communication between all system components, ensuring smooth and secure operation of the entire system.

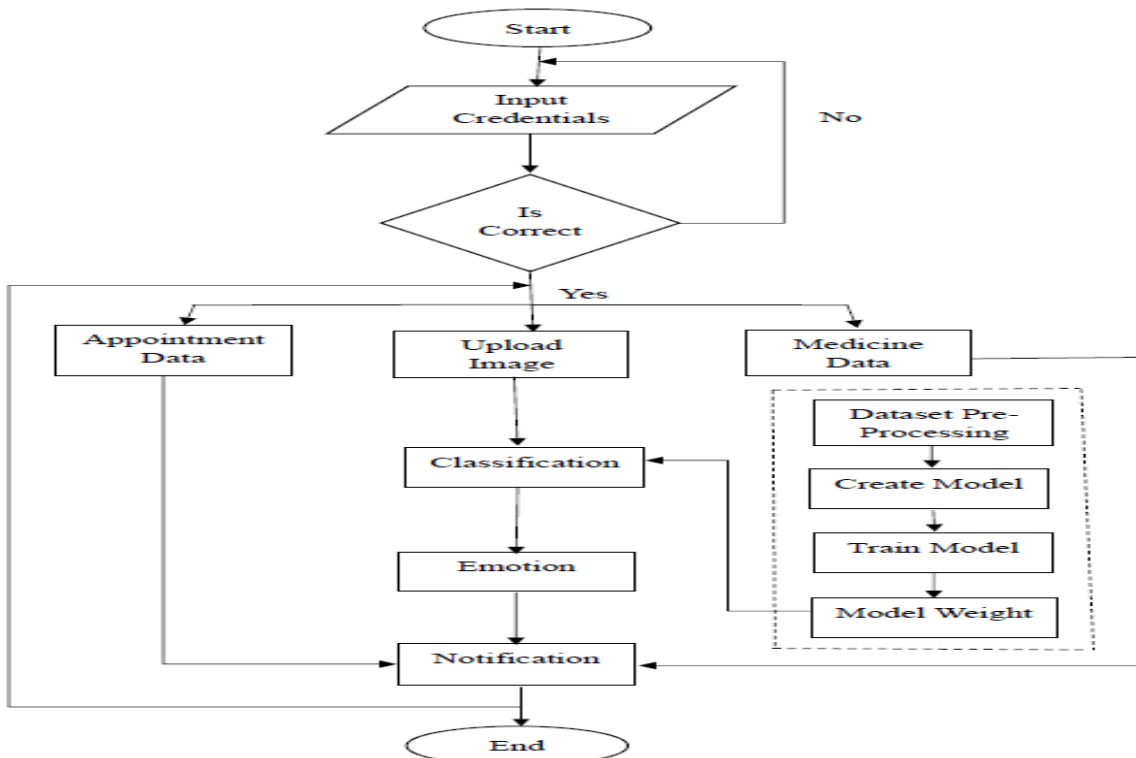


Figure. 2: System Architecture

V. RESULTS AND DISCUSSION

The performance of the proposed emotion recognition system was evaluated based on its ability to correctly classify facial expressions into different emotional categories. The model demonstrated reliable emotion prediction across multiple classes such as happy, sad, angry, fearful, surprised, and neutral. During experimentation, the preprocessing techniques helped improve recognition accuracy by reducing noise and focusing on relevant facial features. The convolutional neural network effectively learned discriminative facial patterns and showed stable performance on unseen test images. The system was also tested within the web application environment, where real-time emotion detection was successfully achieved from user-uploaded images. The predicted emotions were stored correctly in the database and were reflected accurately in the user and doctor dashboards. The overall results indicate that the proposed system performs efficiently and consistently, making it suitable for real-world emotion monitoring and healthcare-related applications.

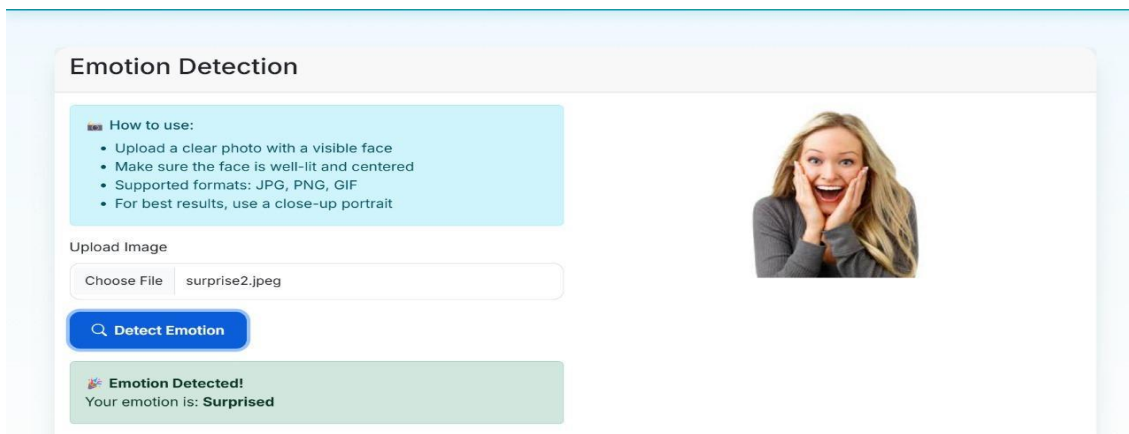


Figure. 3: Surprised Emotion Result

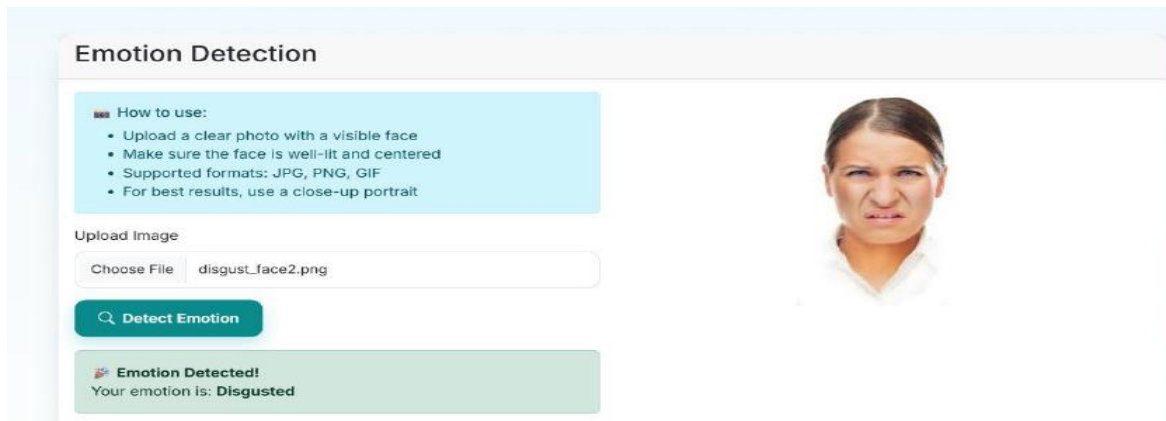


Figure. 4: Disgusted Emotion Result

VI. CONCLUSION

The proposed emotion recognition system successfully demonstrates the effectiveness of deep learning techniques in identifying human emotions from facial expressions. By utilizing a customized Convolutional Neural Network, the system is able to accurately classify multiple emotional states such as happiness, sadness, anger, fear, surprise, and neutrality. The applied preprocessing techniques, including face detection, grayscale conversion, resizing, and normalization, significantly enhance the robustness and reliability of emotion prediction. The integration of the emotion recognition module with a web-based application enables efficient management of user data, emotion records, and healthcare-related functionalities. Furthermore, features such as appointment scheduling, medicine reminders, and WhatsApp-based notifications contribute to timely emotional support and proactive care. Overall, the system provides a scalable and practical solution for continuous emotional monitoring and has strong potential for improving mental health support in real-world healthcare environments.

REFERENCES

- [1]. V. Singh and S. Prasad, "Speech Emotion Recognition System Using Gender Dependent Convolution Neural Network," 2023 IEEE Conference on Computational Intelligence and Communication Technology (CICCT), 2023.
- [2]. A. Israelsson, A. Seiger, and P. Laukka, "Blended Emotions Can Be Accurately Recognized from Dynamic Facial and Vocal Expressions," *Frontiers in Psychology*, vol. XX, 2023, pp. XX-XX.
- [3]. P. Sreevidya, S. Veni, and O. V. R. Murthy, "Elder Emotion Classification Through Multimodal Fusion of Intermediate Layers and Cross-Modal Transfer Learning," *Procedia Computer Science*, vol. XX, 2022.
- [4]. A. C. Y. Low, V. Y. S. Oh, E. M. W. Tong, D. Scarf, and T. Ruffman, "Older Adults Have Difficulty Decoding Emotions from the Eyes, Whereas Easterners Have Difficulty Decoding Emotion from the Mouth," *Scientific Reports*, vol. 12, 2022.
- [5]. D. S. Cortes, C. Tornberg, T. Bänziger, H. A. Elfenbein, H. Fischer, and P. Laukka, "Effects of Aging on Emotion Recognition from Dynamic Multimodal Expressions and Vocalizations," *Scientific Reports*, vol. 11, 2021.
- [6]. M. Juliao, A. Abad, and H. Moniz, "Exploring Text and Audio Embeddings for Multi-Dimension Elderly Emotion Recognition," *Proceedings of Interspeech 2020*, 2020.
- [7]. S. Dou, Z. Feng, X. Yang, and J. Tian, "Real-Time Multimodal Emotion Recognition System Based on Elderly Accompanying Robot," *Journal of Physics: Conference Series*, vol. 1453, 2020.
- [8]. N. Al-Garaawi and T. Morris, "Study on Aging Effect on Facial Expression Recognition," *International Conference on Emerging Security Technologies (EST)*, 2016.