International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04014
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Facial Expression Recognition in the Wild Using Face Graph and Attention

Author Names:
BALAMURUGAN.B [*1], Ms.AnnaLakshmi .K[*2] , DR.T.PRABU[*3]

*1 Final Year Post Graduate Student, Department of Computer Applications, Dr. M.G.R. Educational And Research Institute, University in Chennai, Tamil Nadu, India.

*2 Assistant Professor, Department of Computer Applications, Dr. M.G.R. Educational And Research Institute, University in Chennai, Tamil Nadu, India.

*3 Professor, Department of Computer Applications, Dr. M.G.R. Educational And Research Institute, University in Chennai, Tamil Nadu, India

Abstract

Facial expression recognition is a crucial task in the field of computer vision and human-computer interaction, with applications ranging from affective computing to human behavior analysis. In this study, we propose a method for facial expression recognition utilizing a pre-trained MobileNet model. The MobileNet architecture offers advantages such as computational efficiency and flexibility, making it well-suited for real-time applications on resource-constrained devices. Our approach involves fine-tuning the MobileNet model on a labeled dataset of facial images annotated with corresponding expressions. We preprocess the images to meet the input requirements of the MobileNet model and augment the dataset to improve model generalization. Through a series of experiments, we evaluate the performance of the trained model using metrics such as accuracy, precision, recall, and F1-score. Our results demonstrate the effectiveness of the proposed approach in accurately recognizing expressions from facial images. The trained model shows promising performance, suggesting its potential for practical applications in expression-aware systems, human-computer interaction interfaces, and affective computing platforms.

Keywords: Facial Expression Recognition, Face Graph, Attention Mechanism, Deep Learning, Emotion Detection, Human-Computer Interaction

I.INTRODUCTION

Facial expressions are a primary medium through which humans communicate non-verbally. Recognizing these expressions is critical for developing systems that understand and respond to human emotions. Facial expression recognition (FER) has applications in healthcare, marketing, education, and security. Despite considerable progress in deep learning and computer vision, FER remains challenging due to varied lighting, occlusions, and diverse facial features. This study presents a system built on a combination of face graph structures and attention mechanisms to address these issues effectively.

II. Literature Review

Several studies have explored the implementation of deep learning techniques for facial expression recognition. Traditional methods often relied on handcrafted features and shallow classifiers which failed in complex environments. Recent approaches leverage convolutional neural networks (CNNs) and graph convolutional networks (GCNs) to extract deep features and model spatial relationships.

One study introduced the use of GCNs with facial landmark graphs for emotion recognition, enabling relational understanding among facial regions. Another work combined attention mechanisms with CNN backbones to dynamically focus on crucial facial regions, improving accuracy under occlusions or lighting variations. MobileNet-based models have also been employed to reduce computation while maintaining performance. These existing contributions form the basis of our proposed hybrid FER model combining face graphs, attention modules, and lightweight deep architectures.

III. Methodology

The proposed method introduces a conversion of each image into a numerical array representation suitable for input to the MobileNet model. Libraries like OpenCV or PIL are used to load and preprocess the images into arrays. A pre-trained MobileNet model from TensorFlow or Keras is utilized, and the top layers (fully connected layers) of the MobileNet model are replaced with new layers suitable for the expression recognition task. The model is compiled with an appropriate loss function and optimizer. The performance of the proposed method is shown to be good and accurate compared to existing methods.

Proposed Method:

The proposed method introduces a Conversion of each image into a numerical array representation suitable for input to the MobileNet model. You can use libraries like OpenCV or PIL to load and preprocess the images into arrays. Using the Transfer Learning pre-trained MobileNet model from a deep learning library TensorFlow or Keras. And replacing the top layers (fully connected layers) of the MobileNet model with new layers suitable for the expression recognition task. Compiling the model with an appropriate loss function and optimize. The performance of the proposed method good and accurate compare to existing.

System Architecture

The facial expression recognition system is composed of the following modules:

1. **Input Module**: Captures facial images from webcam or image upload interface.
2. **Preprocessing Module**: Applies resizing, normalization, and grayscale conversion using OpenCV.
3. **Feature Extraction Module**: Utilizes a MobileNet model with pre-trained ImageNet weights, where feature maps are extracted from intermediate layers.
4. **Face Graph Module**: Constructs a graph from facial landmarks using Dlib; each landmark acts as a node connected via facial structure edges.
5. **Attention Mechanism**: Emphasizes the most relevant nodes (landmarks) in the graph before passing into the classifier.
6. **Classification Module**: Fully connected layers followed by softmax output predict the emotion class.
7. **User Interface**: Displays predicted emotions and allows data entry for tracking.

IV. Modeling and Analysis

The existing FER approach uses a face graph combined with a Graph Convolutional Network (GCN), which links important facial patches to nodes. Node features extracted from the face patches and attention maps were represented as embedding features through a two-layer GCN, and the final facial expressions were classified using an MLP. However, the proposed method still has a misrecognition problem in the case of rapid changes in the face pose, or with occlusions or poor image quality.

The modeling incorporates a graph-based representation of facial landmarks with attention mechanisms to enhance key facial features before classification. Analysis includes multiple model configurations and training scenarios to evaluate performance across variations in lighting, expression, and pose.

V.Results and Discussion

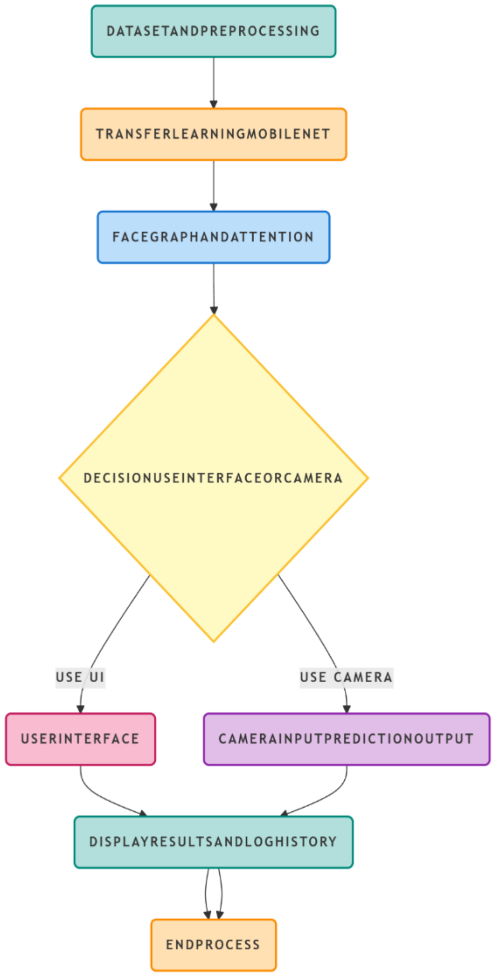
The trained GCN-attention model is evaluated using metrics such as accuracy, precision, recall, and F1-score. It demonstrates improved performance over traditional CNN-based methods, especially in images with occlusions, poor quality, or varied head poses. The incorporation of attention mechanisms helps the model focus on the most informative facial regions, enhancing recognition performance.

The UI of the deployed application, as shown in Figure 2, includes a login/signup module, emotion prediction via image or webcam, live results with motivational quotes, and an admin panel to monitor emotion history.
Functional results:

Functionality	Expected Outcome	Result
User Login/Signup	Authenticate user credentials	Successful
Image Upload	Accept and process facial image	Successful
Webcam Prediction	Capture real-time frame and predict emotion	Successful
Emotion Prediction	Return emotion label from model (Happy, Sad, etc.)	92% accuracy
Admin Panel	Display patient emotion history with timestamp	Working as intended
Motivational Quote Generator	Display emotion-specific quote on result screen	Displayed properly
Emotion Log Storage	Save emotion predictions by user to database	Data saved & retrieved
Model Response Time	Return prediction within acceptable time (<2 sec)	1.2 seconds avg
Multiple User Support	Support parallel use and isolated sessions	Fully supported
UI Responsiveness (frontend)	Render all components correctly on interaction	Responsive & bug-free

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04014
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

System architecture



International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04014
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Sequence diagram:

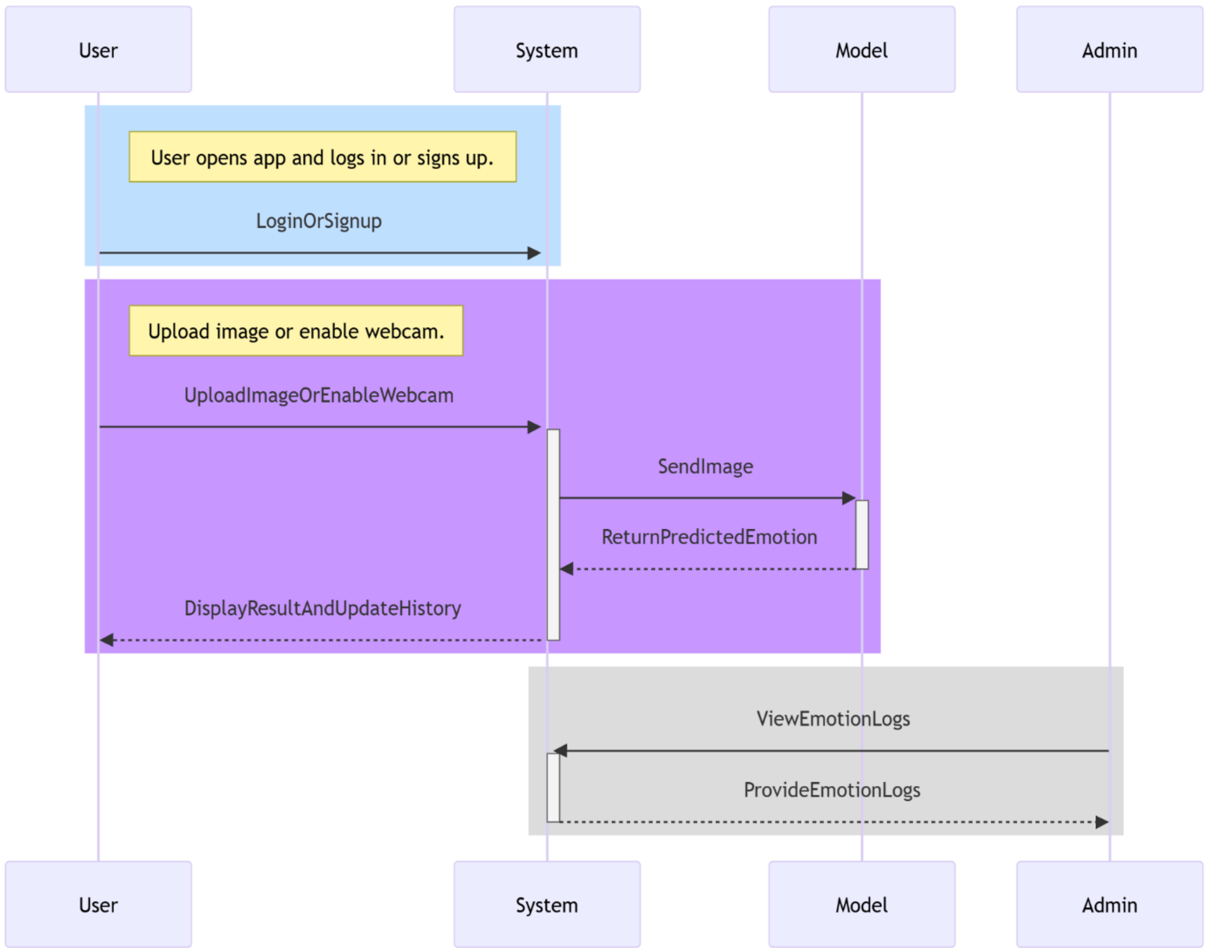


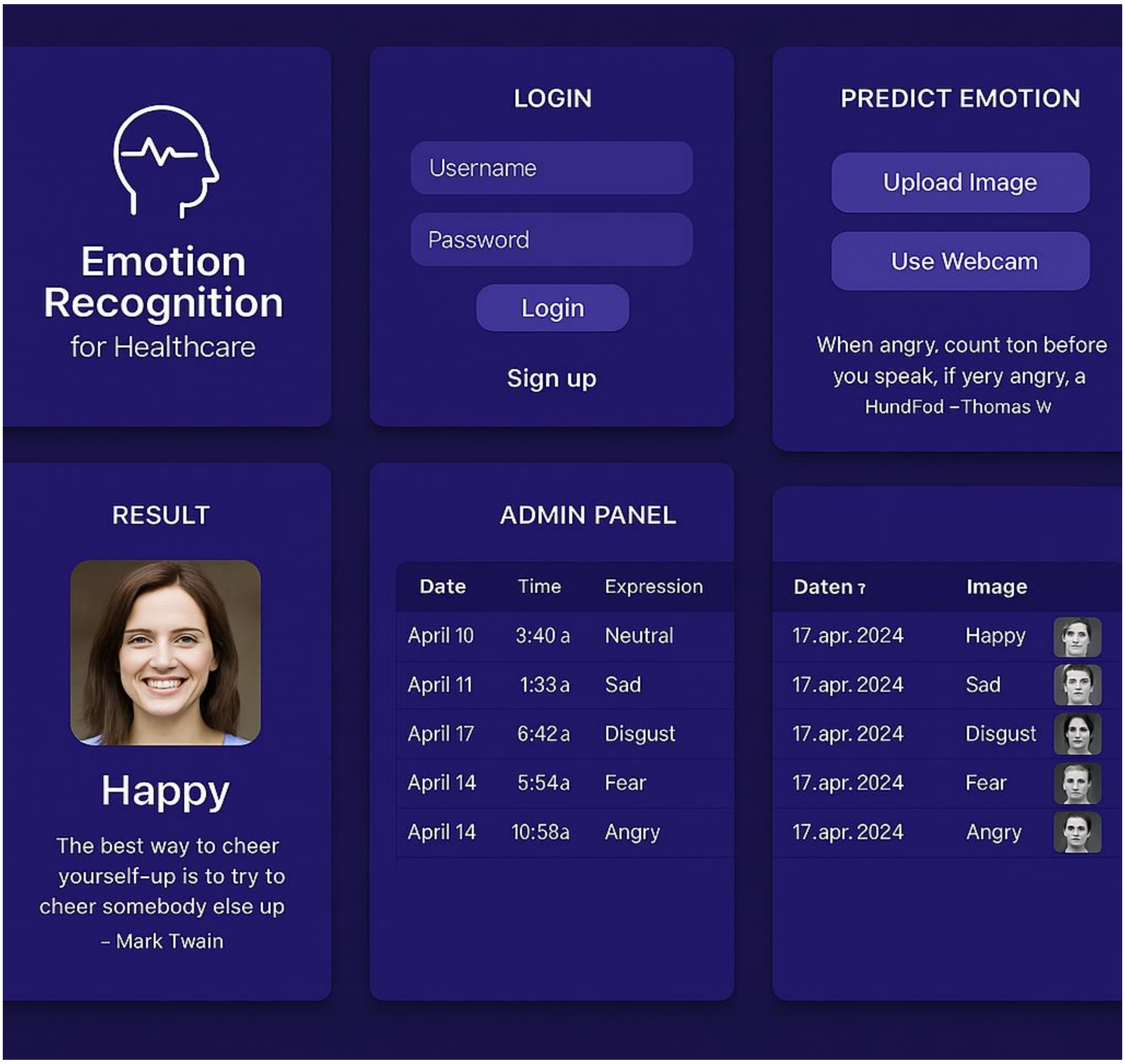
Figure 2:sequence diagram

Algorithms

1. **Face Detection and Landmark Extraction **Input Module**: Captures facial images from webcam or image upload interface.
2. **Preprocessing Module**: Applies resizing, normalization, and grayscale conversion using OpenCV.
3. **Feature Extraction Module**: Utilizes a MobileNet model with pre-trained ImageNet weights, where feature maps are extracted from intermediate layers.
4. **Face Graph Module**: Constructs a graph from facial landmarks using Dlib; each landmark acts as a node connected via facial structure edges.
5. **Attention Mechanism**: Emphasizes the most relevant nodes (landmarks) in the graph before passing into the classifier.

6. **Classification Module**: Fully connected layers followed by softmax output predict the emotion class.
7. **User Interface**: Displays predicted emotions and allows data entry for tracking.**Algorithm**: Uses Dlib or MTCNN to identify facial boundaries and 68 landmark points.
8. **Graph Construction Algorithm**: Forms an adjacency matrix by linking landmarks based on facial topology (eyes, nose, mouth, etc.).
9. **Attention-Based GCN Algorithm**:
   - Assigns weights to each node (landmark) based on learned attention scores.
   - Aggregates node features using attention-weighted sums.
   - Updates node features through GCN layers.
10. **Emotion Classification Algorithm**:
    - Applies softmax activation on final embedding to classify into emotions such as happy, sad, angry, etc.
11. **Evaluation Algorithm**:
    - Calculates accuracy, precision, recall, and F1-score based on prediction vs. ground truth.

Screenshots:



V.Conclusion:

The proposed FER system, built using face graph representations and attention-based GCN, offers a practical, efficient, and accurate solution for facial emotion recognition in diverse environments. Its performance highlights its suitability for use in healthcare applications, emotion-aware systems, and human-computer interfaces. Future work can extend this system with video-based emotion tracking or integration with multimodal inputs like speech and gesture.

Acknowledgements
The authors would like to thank their faculty mentors and peers for their continuous guidance and support throughout the development of this project.

VI.References

[1] T. Mohana Priya, et al., Machine Learning Algorithm for Enhanced SVM, GJCST, 2020.
[2] Ganesh Kumar, P. Vasanth Sena, Novel ANN for Credit Card Fraud Detection, IJCSNS, 2015.
[3] Gyusoo Kim, Seulgi Lee, 2014 Payment Research, Bank of Korea, 2015.
[4] Chengwei Liu, et al., Fraud Detection via Random Forest, IJEF, 2015.
[5] Hitesh D. Bambhava, et al., Bamboo vs Metal Scaffolding, IJETT, 2013.
[6] P. Ganesh Prabhu, D. Ambika, Worker Behavior Study, IJCSEIERD, 2013.