

# Emotion Recognition and Content Recommendation System Using Deep Learning

Muppidi Sai Kumari<sup>1</sup>, Manogna Nannapaneni<sup>1</sup>, Pothula Gayathri<sup>2</sup>,  
Reddybattula Jahnvi<sup>3</sup>, Janjanam Jasmitha<sup>4</sup>, Midatani Manoj<sup>5</sup>

<sup>1,2,3,4,5</sup>Department of Information Technology,  
Dhanekula Institute of Engineering & Technology, JNTUK, Vijayawada, Andhra Pradesh, India.

Corresponding Author: Manogna Nannapaneni

DOI: <https://doi.org/10.52403/ijrr.20260427>

## ABSTRACT

Facial Emotion Recognition (FER) has become Emotion Recognition in the Context of Human Computer Interaction and Affective Computing. An important component in intelligent human computer interaction and affective computing systems. An integrated approach to Emotion Recognition is needed. Deep learning-based emotion recognition system is developed for affect detection and classification model for facial images with the goal of enabling secure interaction and providing users with personalized suggestions. The affect detection and classification model for facial images uses a pre-trained MobileNetV2 architecture which is fine-tuned for the classification of seven emotional states: angry, disgust, fear, happy, neutral, sad, and surprise. We apply pre-processing and image augmentation techniques on the input images to improve the model generalization performance and robustness. We then proceed to optimize using the Adam optimizer with cross-entropy loss and label smoothing. At inference time, this model generates emotion predictions along with confidence scores and applies time-aware bias adjustment to improve contextual relevance. Recommendations tailored to the current emotional state of the user are obtained via a secure login-enabled interface based on the

emotions identified. Experimental validation of the proposed approach is also provided. accuracy of 97.89%, achieving high precision and recall across all classes with minimal overfitting. The lightweight architecture and integrated system design enable efficient real-time deployment in user-centric applications.

**Keywords:** Facial Emotion Recognition, Deep Learning, Transfer Learning, MobileNetV2, Affective Computing.

## INTRODUCTION

Facial expressions have become crucial part of nonverbal communication and they help us understand the emotions of humans. Automatic Facial Emotion Recognition (FER) has become a major area of research in the field of computer science due to its applicability in human-computer interaction, affective computing, intelligent tutoring systems, mental health evaluation, etc. in smart surveillance environments. Emotion recognition leads to more natural interaction between humans and the system provided, hence it helps in achieving more user-friendly interface. Deep learning has seen a significant progress in recent years. As a result, the accuracy of Facial Emotion Recognition (FER) models has also seen a steep rise. The most commonly used architecture in the context of FER is the Convolutional Neural Networks (CNNs).

These networks can automatically learn relevant discriminative features from images, hence making the labour-intensive handcrafted feature engineering less significant. Thus, there is a need for a working FER model that strikes a balance between its classification metrics, computation speed, awareness of environmental context and finally, ability to run in real time. The proposed work deals with the design of an integrated deep learning-based facial emotion recognition system. The suggested system is based on a pre-trained MobileNetV2 model which is fine-tuned for the recognition of seven basic emotions namely angry, disgust, fear, happy, neutral, sad and surprise. Its light weight architecture facilitates the swift computation of emotions without compromising the accuracy. The data preprocessing and augmentation strategies are used to increase the robustness and the generalizability of the model. In addition to facial emotion recognition, the proposed system has been designed as a complete user-centric system with features such as: secure login mechanism, confidence score estimation for emotion recognition, time-aware bias correction for contextualizing the current time with respect to the emotional expression of the user and emotion driven recommendations for tailoring the content that is to be displayed to the current emotional state of the user. Thus, the proposed system facilitates more humanized interaction and content customization with respect to the emotional state of the user. Experimental results show that the proposed model achieves good classification performance with a high validation accuracy and relatively stable results for all emotion classes. Finally, the proposed model has the merits of fast real-time execution and can be easily integrated into intelligent interactive systems.

## **LITERATURE REVIEW**

Recent developments in artificial intelligence and deep learning have improved what we see in facial emotion

recognition and emotion driven recommendation systems. We have reported on many studies which looked at different ways to go about this including the use of convolutional neural networks, fuzzy logic systems, transformer-based architectures, and API based solutions for emotion detection and personal content recommendation. Ansari et al. [1] reported they had put together a facial emotion detection and recommendation system which they used Convolutional Neural Networks (CNN) in. They trained the system to put face expressions into multi emotional categories which included happy, sad, angry and neutral. After determining the users' emotion, the system put forth content recommendations that played into the determined emotional state. The architecture contained image preprocessing, feature extraction, and classification stages. While the system did report success in emotion recognition it mainly used CNN based architectures which we did not see to include exploration of lightweight transfer learning models for real time deployment. Aldeshev et al. [2] put forth an emotion aware music recommendation system which they based on fuzzy intelligence. We put forth that they used a Fuzzy Inference System (FIS) for face emotion detection and a Fuzzy Nearest Mean (FNM) classifier for emotion classification of music. Also, they used Russell's circumplex model which represents emotions in terms of valence and arousal. From what we found out from the use of the Radboud Faces Database the results were very promising in terms of accuracy of emotion detection. That said the use of fuzzy rule-based systems did increase computational complexity and also required manual rule design. Mustafa et al. [3] reported they developed a personal stress relief system which puts together facial emotion recognition with machine learning. The system which uses a camera to capture the user's facial image and which in turn uses the Google Vision API for emotion identification. Based on what is detected the system puts forth personal stress relief

content which may be in the form of music, movies or books. They reported an emotion detection accuracy of 92.19%. Although the system did show practical value it did so by use of external cloud-based APIs which in turn means it isn't a full scale trainable deep learning architecture.

Also, Richard et al. [4] put forth a framework they named Vivify which puts together emotion recognition and music recommendation which they did so using transformer-based models. The present system uses the RoBERTa model for text emotion recognition and LSTM networks for music emotion classification. We use Cosine similarity to determine the emotional tone between input text and recommended music. We report that we saw better personalization and user engagement. At the same time the system does mostly text to music emotional matching which is not to be confused with real time face emotion recognition. Also, Gajula [5] did a review of sentiment aware recommendation systems in e-commerce. That work reports on the integration of sentiment analysis into recommendation systems which in turn improves personalization and user satisfaction. His paper reports on a number of deep learning and transformer -based architectures which put forth sentiment information from text to improve recommendation accuracy. While they are very useful in what they do for text-based sentiment analysis these approaches do not put forward facial emotion recognition. Also, which is related to our discussion, Luo [6] reported a deep learning-based media recommendation which puts together user behaviour data and semantic analysis to better improve recommendation accuracy. They use deep learning to model user preference and put out very personal media recommendations. We saw that the results of this to be very improved in terms of recommendation performance and user satisfaction. That said the approach they put forth does mainly focus on behavioural data as opposed to emotion-based recommendations. Vijayalakshmi et al. [7]

reported a study which they designed that put together facial expression recognition with music recommendation using Convolutional Neural Networks. We had them use a webcam to capture facial images which in turn were put through CNN models which in past work they used the FER-2013 data set. The system put out music via the Spotify API based on the determined emotional state. They report an accuracy of 89% in emotion recognition. While the put forth method did very well it did not take into account context or time-based elements in the recommendation process. Based from our analysis of past studies we see they have mainly looked at individual aspects of emotion detection accuracy and separate study algorithms. Also, which is few there are not many systems which present an integrated model for light weight deep learning, real time emotion detection, confidence score calculation, context adaptability and end user centered recommendation systems. In this case our introduced system includes the use of MobileNetV2 architecture from the perspective of transfer learning which in turn we use for a real time emotion recognition and also a recommendation engine. We also include in this model confidence score assessment and what we term time aware bias correction to increase the context relevance of the recommendation and at the same time put forth personalized content based on the identified emotional state of the user.

## **MATERIALS & METHODS**

Our approach in this project can be broadly classified into four stages, namely, dataset preparation and preprocessing, the deep learning part in which models are created using the concept of transfer learning, real time emotion detection, and at the final stage recommending suitable context based on the inferred emotions.

### **Dataset and Preprocessing**

Our system is trained and tested on the Balanced Real-world Affective Faces

Database (RAF-DB) which consists of a large number of facial images, corresponding to 7 basic affective expressions (Angry, Disgust, Fear, Happy, Neutral, Sad and Surprise). The balanced version of the dataset (RAF-DB) is used for this purpose in order to have approximately the same number of samples per class, which prevents class imbalance. This tweak affects how context affects recommendation priorities. The emotional recommendation engine uses this tweak to match each detected emotion to the appropriate content recommendations. The actual recommendation is then computed based on the current detected emotion, confidence level and the time shift, allowing for a very contextual content recommendation.

### Emotion Classification Using Transfer Learning

Transfer learning is implemented using a pretrained MobileNetV2 model. The convolutional base is retained for feature extraction, while the final classification layer is replaced with a fully connected layer corresponding to seven emotion classes. A dropout layer is included before the final layer to improve generalization. The model is trained using the Adam optimizer with cross-entropy loss incorporating label smoothing. The SoftMax activation function generates probability scores for each emotion class, and the class with the highest probability is selected as the predicted emotion.

### Real-Time Emotion Inference

During deployment, facial images are captured in real time using a webcam. Each frame undergoes resizing and normalization before being passed to the trained MobileNetV2 model. The model outputs class probabilities, and the emotion with the highest probability is identified as the detected emotional state. A confidence score is derived from the predicted probability to indicate classification reliability and is displayed alongside the detected emotion.

### Time-Aware Bias Adjustment and Recommendation

To improve contextual relevance, a time-aware bias adjustment mechanism is applied by incorporating temporal information such as time of day. This adjustment modifies recommendation priorities based on contextual conditions. The recommendation engine maps each detected emotion to predefined content suggestions. The final recommendation is generated based on the predicted emotion, confidence score, and temporal adjustment, enabling adaptive and context aware content delivery

### RESULT

We evaluate the performance of the proposed emotion recognition model using the Balanced RAF-DB dataset. The dataset is split into training, validation and testing sets to guarantee the randomness and unbiasedness of the experimental results. The seven basic emotions are classified by using a pre-trained MobileNetV2 model with the transfer learning technique, including angry, disgust, fear, happy, neutral, sad and surprise.

### Overall Performance Metrics

As we can see in Table I, the final performance of the proposed model is presented. The validation set achieved an accuracy of 97% which clearly shows the robustness of the network to unseen data samples. To evaluate the performance and classify the behaviour of the network, the precision, recall and F1 score values are shown in this table. All the values are high and in accordance with each other for every class, confirming the correctness of the network's performance.

Table I. Overall Performance Metrics

Metric	Value
Accuracy	0.97
Precision	0.97
Recall	0.97
F1-Score	0.97

**Class-wise Performance Analysis**

For the class wise analysis, please refer to Table II. As it can be observed, the precision and recall value of almost all the emotions are very high. The emotions with the near to perfect recall values were those

of anger and fear. The slight drop in precision and recall values for the classes neutral and happy is mainly due to the fact that the facial expressions for these emotions are quite similar to each other.

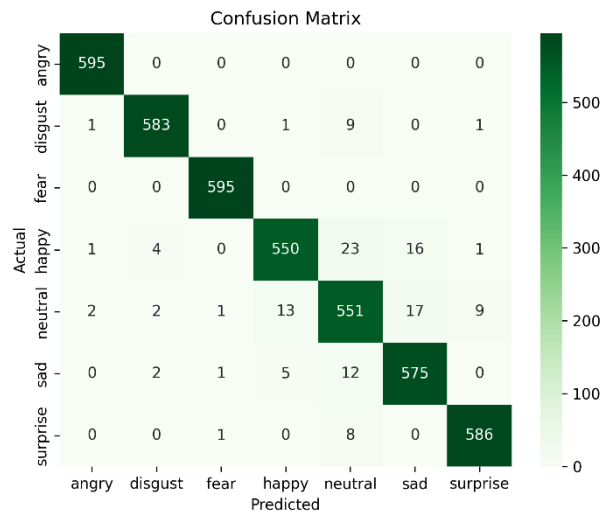
**Table II. Class-wise Precision, Recall, and F1-Score**

Emotion	Precision	Recall	F1-Score
Angry	0.99	1.00	1.00
Disgust	0.99	0.98	0.98
Fear	0.99	1.00	1.00
Happy	0.97	0.92	0.95
Neutral	0.91	0.93	0.92
Sad	0.95	0.97	0.96
Surprise	0.98	0.98	0.98

**Confusion Matrix Analysis**

The confusion matrix depicted in Fig. 1 summarizes our classification performance for all emotion classes. As can be seen from the matrix, the majority of the classifications fall on the diagonal. The

small number of misclassifications occur between neutral and sad and between neutral and happy. These are caused by the similar facial features and are not disturbing, however, the generalization performance is still very good.

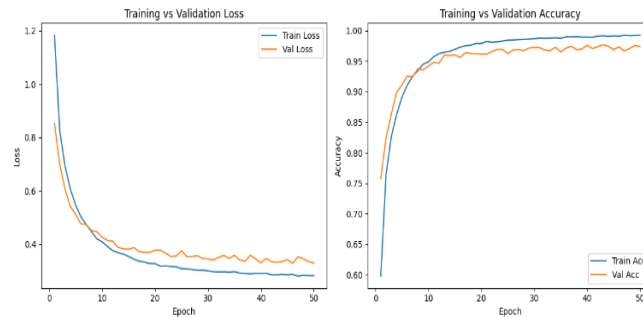


**Fig. 1. Confusion Matrix for Emotion Classification on Balanced RAF-DB.**

**Training and Validation Behavior**

See Fig. 2 for training and validation curves for this model. The accuracy and loss curves of this model are increasing and decreasing, respectively, in the desired fashion. The small difference between training and

validation accuracy indicates that the model generalizes well and does not overfit. The data augmentation, dropout and label smoothing all contribute to well-behaved training.



**Fig. 2. Training and Validation Loss and Accuracy Curves.**

### Real-Time Inference Performance

During the real-time experiment using the webcam, it can be observed that the proposed system can carry out the emotion recognition and output the relevant recommendation and the corresponding confidence in a very short time period. It is clearly observed that the MobileNetV2 is very efficient in terms of the time required to classify the emotion, while still maintaining the required classification rate. The real time emotion recognition and the output of the recommendations with the corresponding confidence level further validate the efficiency, robustness and practicality of the proposed system for real time application.

### DISCUSSION

The experimental results demonstrate the excellence of the MobileNetV2-based transfer learning method in facial expression emotion recognition. High validation accuracy validated the robustness of the features learned by the network to distinguish different emotions. The well distributed datasets also verified the stability of the model performance for all the seven basic emotion classes in the Balanced RAF-DB. The classification result from the confusion matrix shows that the majority of the classifications falls on the diagonal which represents the correct classification given by the model. The misclassifications occur mainly in between adjacent classes, i.e. neutral to sad and happy to neutral. This is due to similar movements of the facial muscles when going from one expression to another and also due to the large variance in the intensity of the expressions within one

class. But still the mis-classification rate is quite low showing that the network was trained with relevant features, and also that the fine-tuning of the pre-trained net was successful. The training and validation curves demonstrate stable convergence with little divergence between training and validation accuracy which implies that overfitting is successfully avoided because of the dropout, data augmentation and label smoothing. In addition, the loss consistently decreases which is also a good sign for the optimization process. Besides evaluating the proposed approach in offline settings, the learned models are implemented and validated in a real time application environment. The resulting system is capable of classifying emotional expressions recognized by the webcam and outputs the classification result along with the corresponding confidence level. Additionally, our time aware bias adjustment is shown to further improve the contextual appropriateness of generated recommendations. In total, our results indicate that a computationally efficient deep learning model, coupled with the contextual adaptation strategy, is capable of realizing a fully deployable emotion aware system.

### CONCLUSION

A deep learning based facial emotion recognition system based on the MobileNetV2 model pre-trained by the transfer learning way is proposed in this paper. Experiments show that the model has the best classification effect on the RAF-DB dataset consisting of seven emotion categories. It shows a good generalization

ability, a stable convergence performance and a smaller overlap among similar facial emotion expressions. The learned model was implemented into an application environment, using the webcam to classify the emotion from the user and output the corresponding confidence, a time aware bias adjustment is also implemented to improve the contextual adaptability of the emotion-driven recommendation generation. Our implementation has proven the applicability of the proposed ultra-lightweight deep learning models for building high-quality and applicable emotion-aware systems. The overall results validate the use of transfer learning in conjunction with a suitable CNN design for multi-class facial emotion recognition and hence for adaptive user interfaces.

#### **Declaration by Authors**

**Acknowledgement:** None

**Source of Funding:** None

**Conflict of Interest:** No conflicts of interest declared.

#### **REFERENCES**

1. Ansari Abu Affan Mustafa Kamal Ah., Sufiyan M. Ahsan Ansari, Dhananjay Dinkar Deore, Afeefa Rafeeqe, Mubashshera Mirza Gufran Beg, and Md. Salman Baig, "Facial Emotion Detection and Recommendation Using Deep Learning," published in conference proceedings in 2025.
2. Azat Aldeshev, Sanzhar Seitbekov, Amandyk Kartbayev, Parasat Tynysbekov, Olzhas Dairov, and Nurzhan Momynkul, "Harmonizing Emotions and Music with Fuzzy Intelligence for Personalized Recommendations," published in the IEEE 5th International Conference on

Smart Information Systems and Technologies (SIST) in 2025.

3. Mohammed M. Mustafa, Ahmed A. Khalifa, and Korhan Cengiz, "Personalized Stress Relief System Using Emotion Recognition and Machine Learning," published in IEEE conference proceedings in 2025.
4. Bettina Shirley Richard, Bhuvana Jayaraman, Ratnavel Rajalakshmi, and Dhivya Chinnappa, "Vivify—Emotion Recognition and Music Recommendation Using Transformers," published in IEEE Access in 2025.
5. Yogesh Gajula, "Sentiment-Aware Recommendation Systems in E-Commerce: A Review from a Natural Language Processing Perspective," published in an academic journal article in 2025.
6. Qi Luo, "Development of Deep Learning-Based Media Content Recommendation System (DL-MCRS) for User Satisfaction," published in Scalable Computing: Practice and Experience in 2025.
7. V. Vijayalakshmi, Pooja Shrivastav, and Gomathi Thiyagarajan, "Facial Expression-Based AI System for Personalized Music Recommendations," published in the Proceedings of the International Conference on Advances and Applications in Artificial Intelligence (ICAAAI) in 2025.

How to cite this article: Muppidi Sai Kumari, Manogna Nannapaneni, Pothula Gayathri, Reddybattula Jahnavi, Janjanam Jasmitha, Midatani Manoj. Emotion recognition and content recommendation system using deep learning. *International Journal of Research and Review*. 2026; 13(4): 278-284. DOI: <https://doi.org/10.52403/ijrr.20260427>

\*\*\*\*\*