iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.80587

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Mood Based Recommemdation System Using Deep Learning

B. Sujatha[1], D. Siri[2], A. Mohan Sai[3], B. Navya[4], V. Sai Srinivas[5]
Department of Computer Science & Engineering– Data Science, Anil Neerukonda Institute of Technology and Sciences Visakhapatnam, India

Abstract: Human emotions greatly influence decision-making processes, productivity, and even mental well-being. As technology advances and artificial intelligence emerges, the need for a machine that can identify human emotions and respond to them becomes evident. In this paper, the development of a Mood-based Recommendation System that uses deep learning techniques is proposed. It will be able to identify users' facial emotions and suggest corresponding wellness activities. To do so, the system will use Residual Network and the attention mechanism to ensure precise emotion recognition. In case the user types anything, the system will utilize BERT model to classify user's intent. Based on user's emo- tions and intent, wellness activities such as music, video recommenda- tions, as well as quotes and activities, will be suggested via API integra- tion. Experiments demonstrated satisfactory performance of the system both in terms of real-time emotion recognition and personalized sugges- tions. In a nutshell, the system integrates computer vision, deep learning, NLP, and multimedia services into one emotional wellness application. Keywords: Deep Learning, ResNet, Attention Mechanism, BERT, Emo- tion Recognition, Recommendation System, Computer Vision, NLP.

I. INTRODUCTION

It is important to note that emotions have the power to influence our behaviors, deci- sion-making processes, and psychological well-being. Our emotional state influences how productive we are daily and how we communicate. With the increase in the use of artificial intelligence within our immediate environments, it is becoming imperative for these systems to detect and respond to our emotions accurately. The expressions on a person's face are one of the most intuitive ways of detecting emo- tions. By merely observing a person's face, we are able to determine their basic emo- tional states such as joy, sadness, anger, fear, surprise, disgust, and neutral. The emer- gence of advanced techniques in deep learning and computer vision makes it possible to detect these emotions accurately. Traditional recommendation systems concentrate mainly on the user's interests, browser history, and ratings. However, they do not take into account the user's state of mind in the present time. This is a major drawback when it comes to customizing an experience as much as possible. The implication of the study here would be to create a system that uses mood recog- nition along with the interpretation of a user’s intended context. The system would use a ResNet based framework that incorporates the attention mechanism for the detection of emotions as well as the BERT model to classify the textual intent. Based on all these, the system will recommend items such as songs, videos, quotes, etc.

II. RELATED WORK

However, facial emotion recognition algorithms have evolved considerably com- pared to such approaches as LBP, HOG, and facial landmarks detection since they can- not handle various challenges related to lighting and angles well. In contrast, with CNN-based architectures, one can significantly increase recognition performance by learning the features directly from images.

To be more precise, Pandey et al. state that CNN architectures perform much better than classical techniques, adding that transfer learning may become especially useful here. Moreover, Ranjani et al. pay attention to preprocessing and its integration into real-time applications via OpenCV. In turn, Agrawal et al. emphasize the necessity of balanced datasets and deeper architectures. Shukla et al. even propose combining Haar Cascade with CNNs to detect emotions on the spot. At last, according to Beena Priya et al., pre-trained VGG19 and MobileNet perform much better than less complex archi- tectures.

Nevertheless, while previous studies mostly concentrated on classification accuracy, little attention was paid to the potential integration with recommendation systems. This challenge can be solved thanks to advancements in such areas as ResNet, attention mechanisms, and BERT. The proposed system hopes to address the shortcomings highlighted above through the use of a ResNet attention-based model for emotion detection as well as BERT for intent recognition, to come up with a recommendation system that functions based on users’ mood.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

III. METHODOLOGY

The proposed Mood-Based Recommendation System strives to obtain real-time infor- mation about the mood of users and give personalized recommendations. The recom- mendation system includes facial recognition for emotions, intent analysis, and recom- mendation engine. These components work together to give the user a personalized experience when interacting with the system.

A. Image Acquisition and Preprocessing

Initially, the webcam captures the live video stream. Preprocessing takes place to im- prove accuracy as the images might suffer from noise or bad light conditions. The pre- processing process comprises face detection, cropping to concentrate on that part of the picture, image resizing, and pixel value normalization.

These steps provide the model with preprocessed data, which is necessary for accurate emotion detection.

B. Emotion Detection using ResNet with Attention

The core of the model used in this recommendation system is the ResNet neural net- work with an attention mechanism. In particular, ResNet allows capturing the compli- cated details of a face, and the attention layer focuses on the eyes, eyebrows, and mouth region of the photo. The classification of emotions involves happiness, sadness, anger, fear, surprise, disgust, and neutral state of mind. Real-time performance will be pro- vided with prediction averaging.

C. Intent Classification Based on BERT

If the users type any text, then the system will use BERT in order to understand the needs of the users better. The analysis of the user's intention becomes better since the emotions and text are combined.

It is associated with certain emotions, such as relaxation, motivation, concentration, or comfort, which make the recommendations more relevant.

D. Recommendation Engine

At last, based on the analyzed emotions and intentions, the personalized recommenda- tions are formed. These may include songs from Spotify, videos from YouTube, inspi- rational quotes, or even some yoga exercises such as mudras.

IV. EXPERIMENTAL RESULTS

A. Training and Validation Accuracy

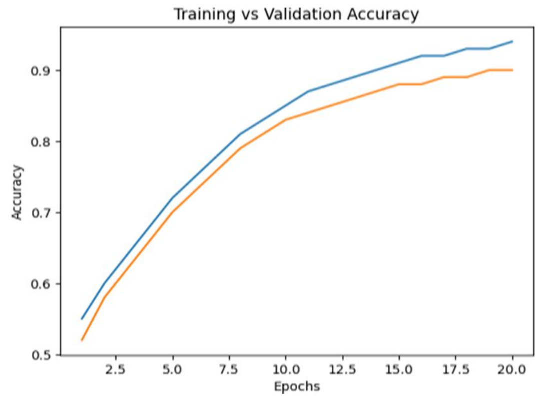


Fig. 1. Training and Validation Accuracy Plot

Fig. 1 displays the train- ing and validation accuracy plot of the ResNet with attention model over several epochs. We observe that training accuracy increases gradually, indicating that the model has effectively learned the features due to residual connections. The validation accuracy follows the training trend, revealing robust generalization and low overfitting tendencies.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

B. Training and Validation Loss

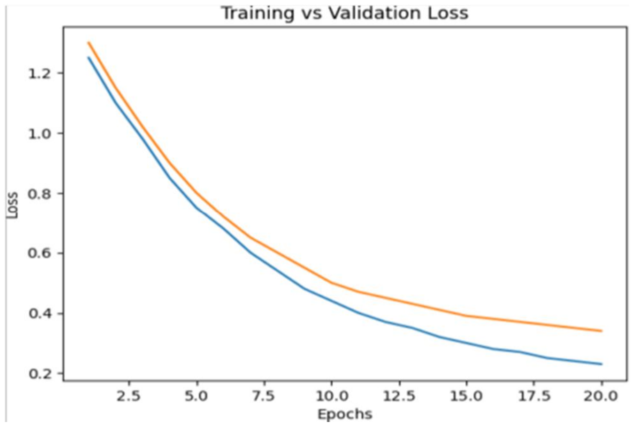


Fig. 2.Training and Validation Loss Plot

Fig. 2 depicts the training and validation loss plots of the ResNet with attention model over several epochs. A steady decline in loss values over several epochs signifies efficient optimization and effective learning in the model. Additionally, there is a negligible gap between training and validation loss, confirming that overfitting is not an issue with the model.

C. Class-wise Emotion Recognition Performance

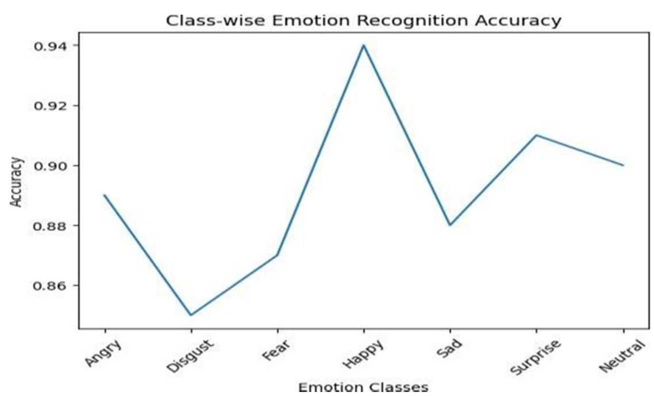


Fig. 3. Class-wise Emotion Recognition Accuracy

This sub-section presents the class-wise emotion recognition accuracy of the model. It demonstrates that the model efficiently recognizes emotions such as Happy and Sur- prise but underperforms slightly when recognizing emotions such as Disgust and Fear.

D. Confusion Matrix Analysis

The diagonal values denote the number of correct classifications, whereas the off-diag-

onal values show the points where the classification was wrong. High performance in all emotional classes is validated by the prevalence of diagonal values. Minor confusion may exist between some emotions that are alike, like Fear and Surprise. This type of confusion is usual in facial emotion recognition tasks.

E. System Overview Interface

This represents the interface of the Mood-Based Recommendation System, which fa- cilitates users to analyze their emotions and get recommended results accordingly. The dashboard incorporates a combination of features, such as emotion recognition, intent understanding, and content recommendation in an easy-to-use way.

F. Real-Time Emotion Detection

Our Mood-Based Recommendation System can perform real-time emotion detection by applying the attention mechanism on a ResNet model. It works through web-camera, capturing the user's face image and focusing on specific regions of interest, such as the eyes and mouth. As an example, our model can recognize the emotion "Happy".

G. Emotion Stabilization and Personalized Recommendations

After performing some stabilization procedures, the ultimate output for the emotion detection process is found to be "HAPPY". Upon this identification process, the system automatically makes recommendations of music, videos, mudras, and inspirational quotes to match the emotion detected.

H. Recommendation System for Music

In order to personalize music according to the detected emotion and intended action, the system relates the two and makes use of APIs for the purpose of acquiring music playlists and tracks. Recommendations are well described with a play button attached to them.

I. Recommendation System for Videos

In this case, personalized video suggestions are made by the system, based on detected emotion and intended actions, using YouTube API. Videos are selected dynamically from YouTube channels.

J. Recommendation System for Mudras

This is concerned with recommending hand gestures to match certain emotions. This module also keeps record of past emotions with corresponding confidence values to help users know themselves better.

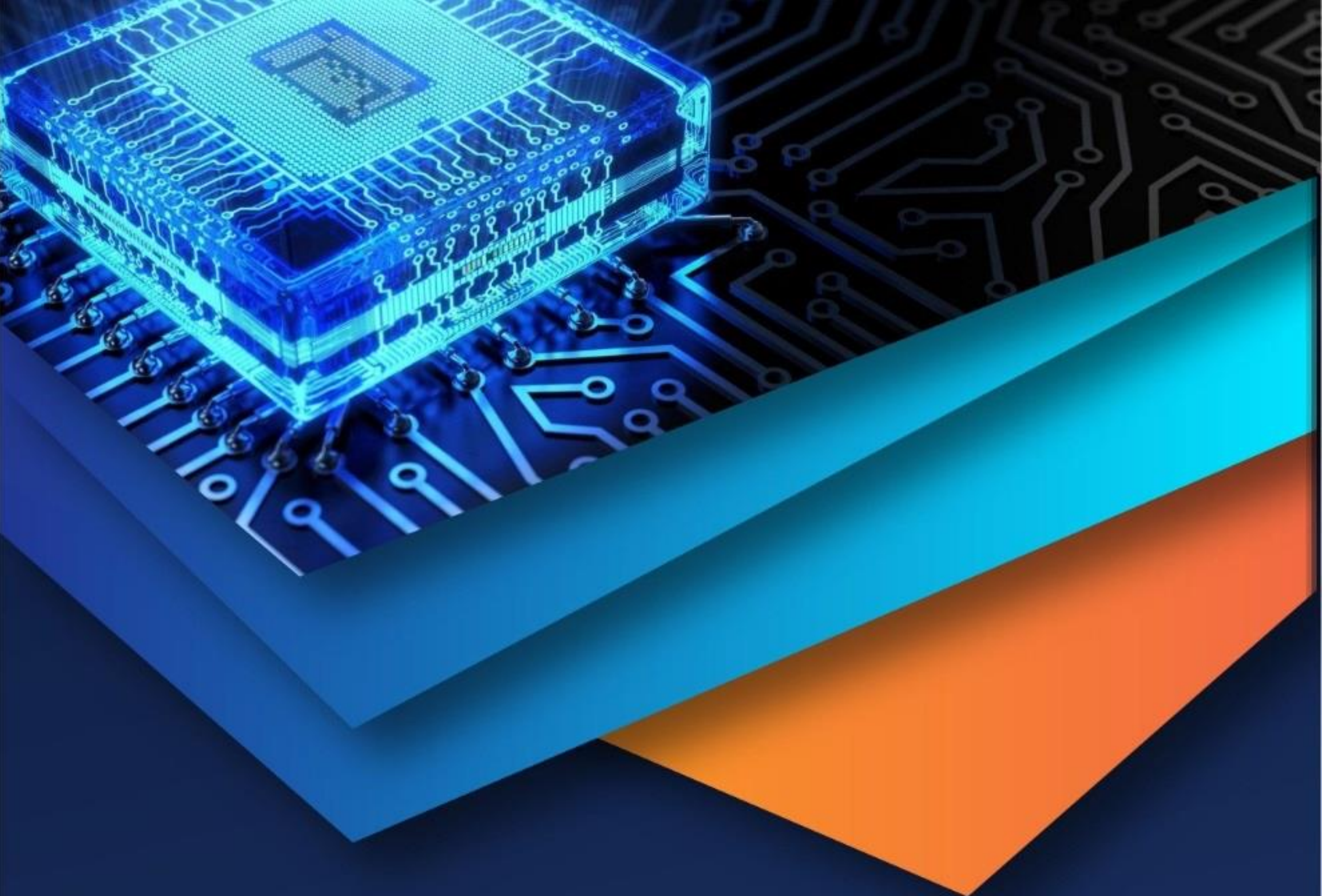
V. CONCLUSION

The core of the mood detection recommendation system is based on recognizing moods through deep learning in combination with facial emotions recognition in real-time. It is implemented using a neural network built with a ResNet architecture, equipped with an attention module. Additionally, it uses the BERT approach for contextualization of emotions detected by the algorithm. Based on the mood, the system will suggest music, videos, motivational quotes, and wellness activities.

The results obtained during its testing have confirmed the high accuracy and reliability
of the algorithm in operation in real-time, and also prove how successfully computer vision and deep learning combined with the API recommendation approach can work together.

REFERENCES

[1] Ranjani, R. M., Nalla, R., M., S., & K, S. G. (2025). "Emotion Recognition using CNN." Proc. 8th Int. Conf. Trends in Electronics and Informatics (ICOEI).
[2] Agrawal, I., V, Y., Kumar, A., Hegde, R., & G, S. D. (2021). "Emotion Recog- nition from Facial Expression using CNN." Proc. IEEE Region 10 Humanitarian Tech- nology Conf.
[3] Shukla, D., Kumari, R., & Bhargavi, A. (2024). "Human Face Detection and Emotion Recognition Using OpenCV through AI." Proc. 12th Int. Conf. Internet of Every-thing, Microwave, Embedded, Communication and Networks (IEMECON).
[4] Beena Priya, R. N., Hanmandlu, M., & Vasikarla, S. (2021). "Emotion Recog- nition Us-ing Deep Learning." Proc. IEEE Applied Imagery Pattern Recognition Work- shop (AIPR).
[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition." Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR).
[6] Vaswani, A., et al. (2017). "Attention Is All You Need." Advances in Neural Information Processing Systems (NeurIPS).
[7] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." Proc. NAACL- HLT.
[8] P-L. Carrier and A. Courville, "FER-2013 Facial Expression Recognition Da- taset", Kaggle, 2013.
[9] Spotify for Developers, "Spotify Web API Documentation" 2024. Available at: https://developer.spotify.com/
[10] Google Developers, "YouTube Data API v3," 2024. [Online]. Available: https://developers.google.com/youtube



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY