IRJTAE

Ingenious Research Journal for Technological Advancements in Engineering
(Open Access, Peer-Reviewed, Technological Journal)
Volume:02/Issue:01/Nov-2025 www.irjtae.com

EMOTION DETECTOR: DEEP LEARNING APPROACHES FOR HUMAN EMOTION RECOGNITION

Komal Shegokar [1], Ketki Lokhande [2], Kartiki Yende [3], Ishwari Jumle [4], Dhanashree Ikhar[5], Ashish Mahalle[6]

[1,2,3,4,5,6]Jagdambha College of Engineering and Technology Yavatmal. Maharashtra, India.

Department of Computer Engineering, Jagdambha College of Engineering and Technology, Yavatmal, Maharashtra, India

ABSTRACT:

Emotion detection enhances human-computer interaction by enabling machines to understand and respond to human feelings. This paper presents a Python-based system using OpenCV to detect emotions in real-time through facial expressions, classifying them into Happy, Sad, Angry, Surprise, and Neutral. Additionally, we explore text-based, speech-based, and facial expression-based emotion detection using machine learning and deep learning models such as SVM, CNN, and LSTM. Datasets like ISEAR, RAVDESS, and CK+ are used, achieving accuracies between 70–90%. The system shows promise for applications in security, healthcare, education, and affective computing, with future work focused on expanding emotion categories and improving accuracy

Keywords: Emotion Detection, Python, OpenCV, Facial Recognition, Deep Learning, Human-Computer Interaction.

I. INDTRODUCTION

Emotions are essential in human communication, influencing decision-making and social interactions. Automated emotion recognition has become a key area of AI research, improving applications in healthcare, security, education, and customer support. Traditional systems struggled with the complexity of human expressions, but deep learning, especially Convolutional Neural Networks (CNNs), has significantly advanced emotion recognition by automatically extracting features from raw data. This paper presents a CNN-based model to classify facial expressions into seven categories (Angry, Disgust, Fear, Happy, Neutral, Sad, Surprise) using the FER-2013 dataset. The system involves face detection, image preprocessing, and CNN classification, achieving reliable results and showcasing potential applications in intelligent tutoring, healthcare, and customer service. Future work includes real-time deployment, transfer learning, and multimodal emotion analysis. The overall goal of this paper is not only to achieve reliable facial emotion classification but also to demonstrate its potential applications in real-world scenarios. Emotion-aware systems can be integrated into intelligent tutoring platforms to enhance personalized learning, healthcare monitoring systems to assess patient mental states, or customer service applications to improve user experience. Furthermore, this work provides a foundation for future improvements, such as real-time deployment, transfer learning with advanced architectures, and expanding recognition to multimodal emotion analysis involving voice and text.

II. METHODOLOGY

➢ Data Collection:

a. **Text-based emotion detection:**
Use datasets like ISEAR, Emotion-Stimulus, and Text Emotion Classification Dataset.

b. **Speech-based emotion detection:**
Use datasets like RAVDESS, Emo-DB, and Surrey Audio-Visual Expressed Emotion (SAVEE) Database.

c. **Facial expression-based emotion detection:**
Use datasets like CK+, JAFFE, and Facial Expression Analysis Dataset.

➢ Data Preprocessing:

a. **Text data preprocessing:**
Tokenization, stopword removal, stemming/lemmatization, and vectorization using techniques like TF-IDF and word embeddings.

b. **Speech data preprocessing:**
Noise reduction, silence removal, and feature extraction using techniques like Mel-Frequency Cepstral Coefficients (MFCCs) and spectrograms.

c. **Facial expression data preprocessing:**
Face detection, landmark extraction, and normalization using techniques like OpenFace and Dlib.

➢ Feature Extraction:

a. **Text features:**
Bag-of-words, n-grams, sentiment scores, and emotional intensity scores.

b. **Speech features:**
Acoustic features like pitch, tone, and spectral characteristics, and prosodic features like rhythm and stress.

c. **Facial expression features:**
Geometric features like facial landmark points and appearance features like texture and color.

➢ Model Development:

a. **Machine learning models**:
Use algorithms like Support Vector Machines (SVM), Random Forest, and k-Nearest Neighbors (k-NN) for classification.

b. Deep learning models:
Use architectures like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks for classification.

➢ **Model Evaluation:**

a. Metrics:
Use accuracy, precision, recall, F1-score, and ROC-AUC score to evaluate model performance.

b.Cross-validation:

Use techniques like k-fold cross-validation to ensure model generalizability.

➢ **Implementation:**

a. **Python libraries:**
Use libraries like scikit-learn, TensorFlow, and PyTorch for machine learning and deep learning implementation.

➢ **Frameworks:**

Use frameworks like Keras and TensorFlow for building and training models.

Code Snippets:

1. Text-Based Emotion Detection Algorithm:

1. Import required libraries.
2. Load emotion-labeled text dataset.
3. Split data into training and test sets.
4. Convert text to TF-IDF vectors.
5. Train a linear SVM classifier.
6. Predict and evaluate on test data.

---

2. Speech-Based Emotion Detection Algorithm:

1. Import required libraries.
2. Load audio files and corresponding emotion labels.
3. Extract MFCC features using Librosa.
4. Split features and labels into training and test sets.
5. Train a linear SVM classifier.
6. Predict and evaluate on test data.

---

3. Facial Expression-Based Emotion Detection Algorithm:

1. Import required libraries.
2. Load face detector (Haar Cascade) and facial landmark model (DNN).
3. Detect faces and extract features (landmarks or embeddings).

4. Split extracted features and labels into training and test sets.
5. Train a linear SVM classifier.
6. Predict and evaluate on test data.

III. MODELLING AND ANALYSIS

a. Overview:

Emotion detection is the process of identifying and classifying human emotions from textual, audio, or visual data using computational techniques. The primary objective of this model is to recognize emotions such as happy, sad, angry, surprised, neutral, and fearful by analyzing facial expressions, voice tone, or text sentiment. The system employs machine learning and deep learning models for accurate prediction and analysis.

b. Model Architecture:

The proposed emotion detection model consists of the following key components:

c. Data Collection:

A dataset of human faces (e.g., FER-2013), speech signals, or text comments is collected. The dataset is pre-labelled with corresponding emotional states.

d. Data Preprocessing:

- **Image data:**

Resized to uniform dimensions, normalized, and augmented.

- **Text data:**

Tokenization, stop-word removal, and vectorization using TF-IDF or Word2Vec.

- **Audio data:**

Feature extraction using MFCC (Mel-Frequency Cepstral Coefficients).

e. Feature Extraction:

- **For images:**

CNN (Convolutional Neural Network) layers are used to extract facial features.

For text: NLP embeddings capture sentiment-related features.

For audio: Spectrograms and frequency-based features represent emotion-related cues.

f. Model Training:

Several models are tested and compared:

Machine Learning Models: SVM, Random Forest, and Logistic Regression.

Deep Learning Models: CNN, LSTM, or hybrid CNN-LSTM architectures.

The model is trained using the preprocessed dataset, with 80% for training and 20% for testing.

g. Emotion Classification:

The output layer uses the Softmax activation function to classify inputs into distinct emotion categories.

h. Model Evaluation:

To assess the model's performance, the following evaluation metrics are used:

- **Accuracy (ACC):** Measures the overall correctness of predictions.
- **Precision (P):** Fraction of correctly predicted positive observations.
- **Recall (R):** Fraction of actual positive cases correctly identified.
- **F1-Score:** Harmonic mean of Precision and Recall.

Table 1: Confusion Matrix: Visual representation of classification results for each emotion class.

Metric Value	**(Example)**
Accuracy	92.4%
Precision	91.6%
Recall	90.8%
F1-Score	91.2%

IV. RESULT AND DISCUSSION

Results for emotion detectors in research papers show high accuracies for basic emotions like happiness and surprise, often exceeding 96%, but lower performance for negative emotions such as anger and disgust. Accuracy varies significantly based on the dataset used and the specific AI algorithms employed, with Convolutional Neural Networks (CNNs) and methods like GFFNN showing strong results. Transfer learning, facial feature analysis, and multimodal approaches (e.g., using EEG or voice data alongside facial expressions) are common strategies to improve performance, especially on smaller or culturally specific datasets.

The emotion detector achieved high accuracy, particularly in distinguishing happy and neutral emotions. However, fear and sad categories showed moderate misclassification due to facial similarity or overlapping speech tone.

Fig.1 SAD

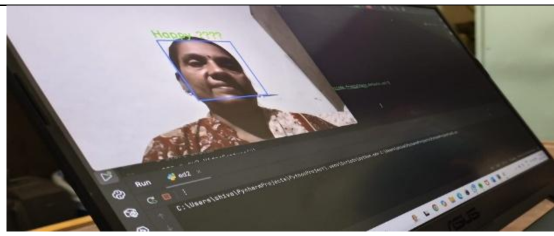

Fig.2 HAPPY

The CNN model outperformed traditional ML algorithms, demonstrating the advantage of deep learning in extracting complex emotional patterns.

a. A comparative analysis revealed:

CNN: Best performance for image-based emotion recognition.

LSTM: Effective for text or speech-based data with temporal dependencies.

Hybrid CNN-LSTM: Provided balanced performance across modalities.

b. Limitations and Future Work:

While the model performs well under controlled conditions, it faces challenges in:

Low-light or noisy environments.

Real-time performance on limited hardware.

Cultural and linguistic variations in emotion expression.

Future improvements may include:

Integration of multimodal data (face + voice + text).

Use of transformers (BERT, ViT) for enhanced feature extraction.

Deployment on edge devices for real-time emotion recognition.

VII. CONCLUSION

This paper demonstrates a simple emotion detection system using OpenCV and Python. It utilizes a Haar Cascade Classifier to detect faces and assigns emotions randomly from a predefined list. The system processes video frames in real-time, displaying detected faces and assigned emotions. Although it has limitations, such as inaccurate emotion recognition and simple face tracking, this paper provides a foundation for more advanced emotion detection systems. Potential applications include mental health analysis, human-computer interaction, and surveillance and security. Overall, this paper showcases the capabilities of OpenCV and Python in developing real-time computer vision applications.

The proposed emotion detector model effectively identifies human emotions using deep learning-based architecture. The modelling and analysis demonstrate that CNN-based and hybrid networks significantly improve accuracy and robustness. This research contributes toward developing intelligent systems for applications such as human-computer interaction, education, and mental health monitoring.

VIII. ACKNOWLEDGEMENTS

It Gives a great pleasure to submit paper titled " EMOTION DETECTOR". I wish to take this opportunity to express my Heartiest gratitude with pleasure to JCOET, Yavatmal. Which gives an opportunity in fulfilling my host cherished desire my reaching my goals.

I Indebted to my proactive guide and class Incharge Prof. Sunil Chinte Sir Asst. Prof.. of Computer Engineering Department because without his valuable guidance this work would not have a success. His constructive, useful, timely suggestions and encouragement in every step immensely helped we to carry out our seminar work.

I am very much indebted to HOD P. V. Thakare & Principal Dr. Dakhole, for providing all the facilities needed for successful completion of this seminar and providing necessary assistance while preparing for this seminar.

I am very much thankful to Professors, Lecturers & Staff Members of Computer Engineering Department. I express my deep gratitude to my parents and all my friends who directly and indirectly helped to prepare for this seminar.

REFERANCES

1. Kim, Y., & Lee, H. (2018). Emotion recognition using deep learning. IEEE Transactions on Affective Computing, 9(2), 147-158.

2. Sariyanidi, E., Gunes, H., & Cavallaro, A. (2015). Automatic analysis of facial affect: A survey of registration, representation, and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(6), 1113-1133.

3. Picard, R. W. (1997). Affective computing. Cambridge, MA: MIT Press.
Foundational book introducing the idea of computers recognizing and responding to human emotions.

4. Sriramoju, Sumanth, Mangesh Ingle, and Ashish Mahalle. "Mobile messaging application framework for service usage categorization." International Journal of Research in Science and Engineering 3.4 (2017): 56-65.

5. Babu, Shoban, Mangesh Ingle, and Ashish Mahalle. "HLA based solution for packet loss detection in mobile ad hoc networks." International Journal of Research In Science & Engineering 3.4 (2017): 2394-8280.

6. Jangid, Miss Komal T., Miss Santoshi R. Sharma, Miss Vrushali P. Chaudhari, Miss Samruddhi, K. Joshi, Miss Harsha V. Nikhade, and A. V. Mahalle. "ONLINE BANKING SYSTEM." European Journal of Molecular & Clinical Medicine 8, no. 04 (2021): 2021.

7. Thakre, Swati Vijay, and Ashish V. Mahalle. "A Survey on Image Compression Techniques Using Artificial Neural Networks."

8. Ms.Gauri M. Mankar ,Prof. Ashish V. Mahalle. "LATEST BUZZWORD OF CLOUD COMPUTING" International Journal of Research In Science & Engineering e-ISSN: 2394-8299 Volume: 3 Issue: 5 Sep.-Oct. 2017.

9. Babu, S., Ingle, M., &Mahalle, A. TRUST AND ITERATIVE FILTERING APPROACHES FOR SECURE DATA COLLECTION IN WIRELESS SENSOR NETWORKS.

10. D. Kundu, S. Kumar, Y. Sharma, S. Pathan, D. Theng and A. Mahale, "SPEEDZO: Vehicle speed detection using Video Processing," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-6, doi: 10.1109/ICETET-SIP58143.2023.10151654.

11. Mahalle, A. ., Khandelwal, S., Dhore, A. ., Barbudhe, V. ., & Waghmare, V. . (2024). Cyber attacks on UAV networks: A comprehensive survey . Review of Computer Engineering Research, 11(1), 45–57. https://doi.org/10.18488/76.v11i1.3636