IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15548

AFFECTIVE COMPUTING TECHNIQUES FOR HUMAN–MACHINE INTERACTION

Rabiya Fathima [1], C S Swetha [2]

[1] Department of MCA, BIT, Bangalore, India

[2] Assistant Professor, Department of MCA, BIT, Bangalore, India

Abstract: Affective computing is an emerging domain of artificial intelligence that focuses on enabling machines to recognize, interpret, and respond to human emotions. With the rapid advancement of intelligent systems such as chatbots, virtual assistants, and service robots, the ability to understand human emotions has become essential for improving interaction quality and user satisfaction. This paper presents a comprehensive survey and comparative analysis of various emotion recognition techniques, including facial expression recognition, speech emotion recognition, and text-based sentiment analysis. Deep learning models such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks are widely used for emotion detection tasks. Furthermore, multimodal approaches that combine multiple data sources are analyzed for their effectiveness in reducing ambiguity and improving accuracy. The study also discusses key challenges such as real-time processing, data bias, and privacy concerns. Finally, the paper highlights future research directions for developing efficient and human-centric emotion-aware systems.

Keywords- Affective Computing, Emotion Recognition, Deep Learning, CNN, LSTM, Multimodal Systems, Human–Machine Interaction

I. INTRODUCTION

Artificial Intelligence (AI) has significantly transformed modern computing by enabling machines to perform tasks that traditionally required human intelligence. Applications such as virtual assistants, chatbots, and recommendation systems have become an integral part of daily life. However, most existing AI systems primarily focus on logical reasoning and data processing, while lacking the ability to understand human emotions.

Human communication is inherently emotional and involves expressions through facial cues, speech patterns, gestures, and textual interactions. The inability of machines to interpret these emotional signals results in limited interaction quality and reduced user satisfaction. This gap has led to the development of **affective computing**, a field introduced to incorporate emotional intelligence into computing systems.

Affective computing aims to design systems that can recognize, interpret, and respond to human emotions in a natural and meaningful way. By integrating emotional awareness, these systems can enhance user experience, improve decision-making, and enable more personalized interactions. Applications of affective computing are widely seen in healthcare, education, automotive systems, and human–robot interaction.

This paper presents a detailed survey of various emotion recognition techniques used in affective computing and provides a comparative analysis of their performance, advantages, and limitations.

II. RELATED WORK

Several research studies have been conducted in the field of affective computing, focusing on different emotion recognition techniques.

Paper [1]: Li and Deng (2020) conducted a comprehensive survey on deep facial expression recognition techniques. Their work highlights the effectiveness of Convolutional Neural Networks (CNN) in extracting spatial features from facial images and achieving high accuracy in emotion classification.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15548

Paper [2]: Akçay and Oğuz (2019) explored speech emotion recognition using deep learning models such as Long Short-Term Memory (LSTM) networks. Their study demonstrates that LSTM effectively captures temporal dependencies in speech signals, making it suitable for emotion detection.

Paper [3]: Poria et al. (2017) proposed a multimodal emotion recognition framework that combines facial expressions, speech signals, and textual data. Their approach improves accuracy by integrating multiple modalities and reducing ambiguity in emotion detection.

Paper [4]: Zadeh et al. (2017) introduced the CMU-MOSEI dataset and developed a multimodal sentiment analysis model using deep learning techniques. Their work emphasizes the importance of real-world datasets in improving emotion recognition performance.

Paper [5]: Picard (1997) introduced the concept of affective computing and laid the foundation for emotion-aware systems by emphasizing the role of emotional intelligence in computing.

Paper [6]: Calvo and D'Mello (2010) studied emotion recognition in human–computer interaction and highlighted its importance in improving user engagement and system effectiveness.

These studies indicate that while significant progress has been made, challenges such as accuracy, real-time processing, and data limitations still exist.

III. METHODOLOGY AND DATASETS

A. Datasets Used In Affective Computing

Emotion recognition systems rely on well-annotated datasets for training and evaluation. Several benchmark datasets are widely used in affective computing research.

- **FER2013**: Facial emotion dataset with labeled expressions (happy, sad, angry, etc.).
- **AffectNet**: Large-scale dataset with real-world facial images and continuous emotion labels.
- **IEMOCAP**: Multimodal dataset containing speech, video, and transcripts for emotion analysis.
- **CMU-MOSEI**: Multimodal dataset with video, audio, and text for sentiment and emotion recognition.

These datasets enable the development of robust models and help evaluate performance across different modalities.

B. Facial Emotion Recognition

Facial emotion recognition involves analyzing facial expressions to identify emotional states such as happiness, sadness, anger, and surprise. Convolutional Neural Networks (CNN) are widely used due to their ability to extract spatial features from images.

- Input: Facial images
- Model: CNN
- Output: Emotion labels

C. Speech Emotion Recognition

Speech emotion recognition focuses on detecting emotions from voice signals. Features such as pitch, tone, and frequency are analyzed using Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models.

- Input: Speech signals
- Model: LSTM
- Output: Emotional state

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁞ Peer-reviewed & Refereed journal ⁞ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.15548

D. Text-Based Sentiment Analysis

Text-based sentiment analysis uses Natural Language Processing (NLP) techniques to detect emotions from textual data such as messages, reviews, and social media content.

- Input: Text data
- Model: NLP / Transformer
- Output: Sentiment classification

E. Multimodal Emotion Recognition

Multimodal approaches combine multiple data sources such as facial expressions, speech, and text to improve emotion recognition accuracy and reduce ambiguity.

- Input: Image + Speech + Text
- Model: Multimodal Deep Learning
- Output: Accurate emotion prediction

F. Comparative Analysis

Technique	Input	Model Used	Advantage
CNN	Image	Deep learning	High accuracy in facial recognition
LSTM	Speech	Recurrent Neural Network (RNN)	Captures temporal patterns in speech
NLP	Text	Transformer Models	Effective for sentiment analysis
Multimodal	Image + Speech + Text	Fusion Models / Deep Learning	Highest accuracy and reduces ambiguity

IV. SYSTEM ARCHITECTURE FOR EMOTION RECOGNITION

A typical affective computing system consists of multiple stages:

1. **Data Acquisition** – Capture facial images, audio signals, or text input
2. **Preprocessing** – Noise removal, normalization, feature extraction
3. **Feature Extraction** – Identify important patterns (facial landmarks, MFCC for audio, embeddings for text)
4. **Model Processing** – Use CNN, LSTM, or transformer models
5. **Classification** – Predict emotion labels (happy, sad, angry, etc.)
6. **Response Generation** – System reacts based on detected emotion

This pipeline enables machines to process emotional data efficiently and respond intelligently.

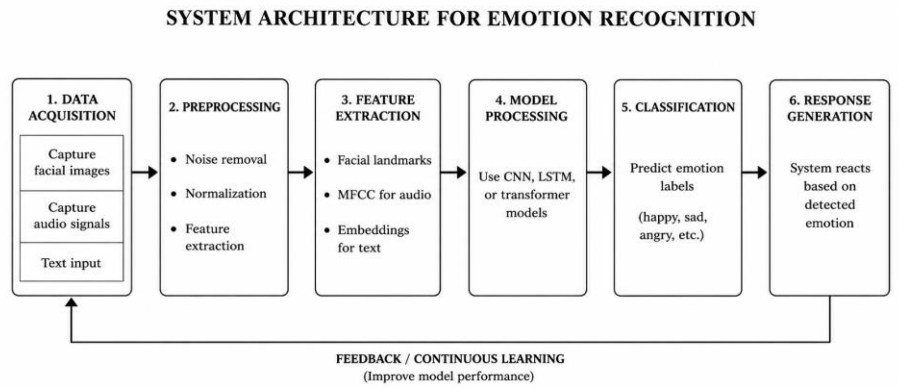

Fig 1. System Architecture Diagram

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15548

V. APPLICATIONS AND CHALLENGES

A. Applications Of Affective Computing

- **Healthcare**: Detect depression, stress monitoring, patient emotion analysis
- **Education**: Intelligent tutoring systems adapting to student emotions
- **Automotive**: Driver fatigue and stress detection
- **Customer Service**: Emotion-aware chatbots and sentiment analysis
- **Human–Robot Interaction**: Service robots responding to human emotions
- **Gaming & VR**: Adaptive gaming experiences based on player emotions

B. Challenges In Affective Computing

Despite advancements, several challenges exist:

- **Ambiguity in emotions** – Same expression can mean different emotions
- **Dataset bias** – Limited diversity affects accuracy
- **Real-time processing** – High computational cost
- **Privacy concerns** – Emotional data is sensitive
- **Cross-cultural differences** – Emotions vary across cultures

Addressing these challenges is crucial for building reliable systems.

VI. ETHICAL CONSIDERATIONS AND PERFORMANCE METRICS

A. Ethical Considerations

The use of affective computing raises ethical concerns:

- Privacy of emotional data
- Misuse in surveillance systems
- Bias in emotion recognition
- Need for transparent AI systems

Ethical AI design is essential for responsible deployment.

B. Performance Metrics

Emotion recognition models are evaluated using:

- Accuracy
- Precision & Recall
- F1-Score
- Confusion Matrix

These metrics help measure model effectiveness and reliability.

VII. RESULTS AND DISCUSSION

The analysis of different emotion recognition techniques reveals that deep learning models significantly improve the performance of affective computing systems. CNN-based approaches are highly effective for image-based emotion detection, while LSTM models perform well in speech-based emotion recognition tasks.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 5, May 2026
DOI: 10.17148/IJARCCE.2026.15548

Text-based sentiment analysis provides useful insights but may suffer from ambiguity due to limited contextual information. In contrast, multimodal approaches that combine multiple data sources provide more accurate and reliable results.

However, several challenges remain in the field:

- High computational cost for real-time processing
- Bias in datasets affecting model performance
- Privacy concerns in collecting emotional data
- Limited availability of diverse datasets

Comparative analysis shows that:

- CNN performs best for image-based emotion detection
- LSTM captures temporal dependencies in speech
- Transformer models improve contextual understanding
- Multimodal approaches outperform single-modal systems

  However, real-world deployment requires balancing accuracy, speed, and privacy.

VIII. CONCLUSION

Affective computing plays a crucial role in enhancing human–machine interaction by enabling systems to understand and respond to human emotions. Various techniques such as CNN, LSTM, and NLP have been widely used for emotion recognition, each offering specific advantages. Among these, multimodal approaches provide the highest accuracy due to their ability to integrate multiple data sources and reduce ambiguity. Despite existing challenges, the field continues to evolve and holds significant potential for future research and real-world applications.

IX. FUTURE WORK

Although the study provides insights into affective computing techniques, several improvements can enhance system performance and real-world applicability:

• **Real-time emotion recognition** for faster and efficient processing

• **Improved multimodal techniques** to increase accuracy and reduce ambiguity

• **Integration with IoT and wearable devices** for better emotion detection

• **Development of diverse datasets** to reduce bias and improve generalization

• **Privacy-preserving frameworks** for secure handling of emotional data

• **Explainable AI models** for better transparency and trust

• **Applications in healthcare and education** for personalized systems

REFERENCES

[1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.
[2] S. Li and W. Deng, “Deep facial expression recognition: A survey,” *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1195–1215, 2020.
[3] M. B. Akçay and K. Oğuz, “Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers,” *Speech Communication*, vol. 116, pp. 56–76, 2020.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 5, May 2026

DOI: 10.17148/IJARCCE.2026.15548

[4] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, “A review of affective computing: From unimodal analysis to multimodal fusion,” *Information Fusion*, vol. 37, pp. 98–125, 2017.

[5] A. Zadeh et al., “Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph,” *Proceedings of ACL*, 2017.

[6] R. A. Calvo and S. K. D’Mello, “Affect detection: An interdisciplinary review of models, methods, and their applications,” *IEEE Transactions on Affective Computing*, vol. 1, no. 1, pp. 18–37, 2010.

[7] P. Ekman, “Facial expression and emotion,” *American Psychologist*, vol. 48, no. 4, pp. 384–392, 1993.

[8] F. Schuller et al., “Speech emotion recognition: Two decades in a nutshell,” *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 1–20, 2021.

[9] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, 2015.

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet classification with deep convolutional neural networks,” *NeurIPS*, 2012.

[11] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[12] A. Vaswani et al., “Attention is all you need,” *NeurIPS*, 2017.

[13] M. Soleymani, D. Garcia, B. Jou, B. Schuller, S. Chang, and M. Pantic, “A survey of multimodal sentiment analysis,” *Image and Vision Computing*, vol. 65, pp. 3–14, 2017.

[14] G. McKeown, M. Valstar, R. Cowie, M. Pantic, and M. Schroder, “The SEMAINE database,” *IEEE Transactions on Affective Computing*, 2012.

[15] C. Busso et al., “IEMOCAP: Interactive emotional dyadic motion capture database,” *Language Resources and Evaluation*, 2008.

[16] A. Mollahosseini, D. Chan, and M. Mahoor, “AffectNet: A database for facial expression recognition,” *IEEE Transactions on Affective Computing*, 2017.

[17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[18] T. Baltrušaitis, C. Ahuja, and L. Morency, “Multimodal machine learning: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[19] E. Cambria, “Affective computing and sentiment analysis,” *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.

[20] J. Russell, “A circumplex model of affect,” *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

