

---

**DISTINGUISHING GENUINE AND FAKE EMOTIONS USING DEEP NEURAL NETWORKS**

---

**Kriti kumari, Khusboo kumari, Shivani kumari, Archna kumari and Rajeev Kumar**

Faculty of Information Technology &amp; Engineering, Gopal Narayan Singh University, Jamuhar, Sasaram, Bihar

**ABSTRACT**

*Emotion detection has become an important area in artificial intelligence, especially in human-computer interaction. This paper presents a deep neural network-based approach for recognizing human emotions from facial expressions. The proposed system uses convolutional neural networks (CNNs) to classify emotions such as happiness, sadness, anger, surprise, fear, and neutral. The model is trained on publicly available datasets and achieves high accuracy in real-time detection. The study highlights the effectiveness of deep learning techniques in improving emotion recognition systems and discusses potential applications in healthcare, security, and customer experience.*

**Keywords:** Emotion Detection, Facial Expression Recognition, Deep learning, Convolutional neural Networks (CNNs), Human-Computer Interaction, Real-Time Classification, Artificial Intelligence

**1. INTRODUCTION**

Human emotions can be understood as people's attitudes, experiences, and corresponding behavioural responses to the objective environment [1]. Emotion is a dynamic cognitive and physiological condition that develops in reaction to inputs, like experiences, thoughts, or interactions with people. It includes subjective experience, cognitive processes, behavioural influences, physiological responses, and communication [2]. Emotions can be regarded as a universal language, transcending cultural and linguistic boundaries, despite evident variations in its expression across different societies and contexts [3]. BERT was originally published by Google researchers Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. The design has its origins from pre-training contextual representations, including semi-supervised sequence learning [4]. Cryptocurrency market has been developed at an exceptional pace since its emergence. Cryptocurrency is a digital currency however it is not controlled by any central authority to make online payments. It uses system ledger entries called 'tokens' to make online payments for goods and services. Elliptical curve encryption and public-private key pairs are used as cryptographic algorithms [5]. facial emotion recognition is employed to detect stress using the CK+ dataset, which is a widely used dataset for facial expression recognition, containing anger, disgust, contempt, fear, happiness, sadness, and surprise emotions [6]. Emotions trigger complex psycho-physiological changes in one's state of mind, deriving from biochemical and environmental interactions. Emotions are the primary determining factors of a healthy sense of oneself and play a pivotal role in a person's daily life. Moreover, emotions directly connected to one's psychology are relayed through the choices of words, thoughts, phrases, mimics, posture, and especially facial expressions[7]. In the burgeoning field of natural language processing (NLP), detecting and interpreting emotions in text represents a fundamental challenge with far-reaching implications [8]. Emotion recognition has become an important area of research in artificial intelligence, aiming to enable machines to understand human feelings through multiple modalities such as text, facial expressions, and speech[9]. The Text Emotion Dataset from Kaggle provides labelled textual data representing various emotional states, making it useful for natural language processing tasks such as sentiment and emotion classification[10]. speech-based emotion recognition, multiple datasets contribute to capturing vocal cues associated with emotions. The CREMA-D (Crowd-sourced Emotional Multimodal Actors Dataset) includes audio recordings from diverse actors expressing different emotions[11][12]. The TESS focuses on emotional speech samples from female speakers, while the SAVEE dataset provides recordings from male speakers with distinct emotional expressions[13][14].

**OBJECTIVE**

1. To develop an intelligent system that can automatically detect human emotions from text, speech, and facial expressions using deep neural networks.
2. To improve the accuracy of emotion recognition by using advanced deep learning models such as CNN, LSTM, BERT, GPT-2, MobileNet, ResNet, and EfficientNet.
3. To analyze emotions like happy, sad, angry, fear, surprise, disgust, and neutral from multiple input sources.
4. To combine text, image, and speech data through multimodal fusion for better and more reliable emotion detection.
5. To reduce the limitations of traditional single-modality systems and increase robustness in real-world conditions.

6. To enhance human-computer interaction by enabling machines to understand and respond to user emotions naturally.

## 2. LITERATURE SURVEY

Emotion recognition has gained significant attention in recent years due to its applications in human-computer interaction, mental health monitoring, and intelligent systems. Researchers have explored various approaches using text, speech, facial expressions, and multimodal data to improve emotion detection accuracy.

In this model, they have proposed a multimodal emotion recognition framework that integrates facial expressions, speech signals, and EEG data [1]. Their study demonstrated that combining multiple modalities significantly improves classification performance compared to unimodal systems. In this model, they have presented a comprehensive review of emotion recognition techniques from 2014 to 2023, highlighting the transition from traditional machine learning approaches to deep learning-based models and emphasizing the importance of multimodal fusion strategies [2].

In this model, he analyzed multimodal emotion recognition systems, focusing on fusion techniques and challenges such as data synchronization, feature extraction, and model generalization [3]. Their review identified that hybrid models combining different neural architectures are becoming increasingly effective in handling heterogeneous data sources.

In the field of text-based emotion recognition, they introduced BERT a transformer-based model that significantly improved natural language processing tasks by capturing contextual dependencies in text [4]. Developed an ensemble LSTM-GRU model for sentiment and emotion detection in cryptocurrency-related tweets, achieving improved accuracy by leveraging sequential learning techniques [5]. Additionally, examined how different emotion representation models affect the generalizability of text-based systems, emphasizing the need for robust and adaptable frameworks [8].

Facial emotion recognition has also seen substantial advancements through deep learning. conducted a comparative analysis of VGG-16, ResNet-50, and a custom CNN architecture for stress detection, concluding that deep convolutional models are highly effective in extracting facial features [6]. Proposed a lightweight CNN model for facial emotion recognition, achieving competitive performance while reducing computational complexity, making it suitable for real-time applications.

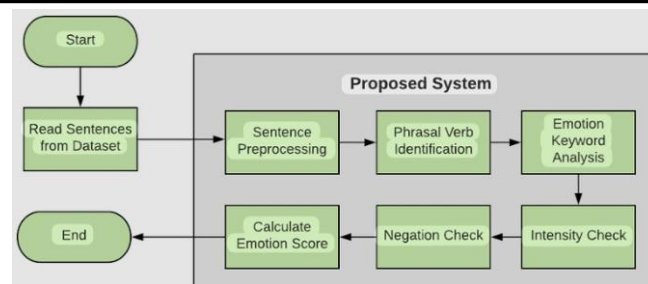
Speech emotion recognition has been enhanced using deep learning architectures such as CNN, LSTM, and GRU. Hybrid models combining CNN-LSTM and GRU networks have shown improved performance by capturing both spatial and temporal features of audio signals. These approaches benefit from widely used datasets such as CREMA-D, RAVDESS, TESS, and SAVEE, which provide diverse emotional speech samples for training and evaluation.

To support research in emotion recognition, several publicly available datasets are utilized. The Text Emotion Dataset provides labeled textual data for emotion classification tasks. Facial expression datasets offer annotated images for visual emotion detection. Speech datasets such as CREMA-D, RAVDESS, TESS, and SAVEE contain audio recordings representing various emotional states, enabling robust training of speech-based models.

Overall, the literature indicates that multimodal emotion recognition systems outperform single-modality approaches due to their ability to capture complementary emotional cues. Despite significant progress, challenges such as data heterogeneity, real-time processing, and model generalization remain open research issues. Future work is expected to focus on improving fusion techniques and developing more efficient and scalable models.

## 3. PROPOSED MODEL

The purpose of this project is to develop an intelligent Emotion Detection System using Deep Neural Network that can automatically identify human emotions from text, facial expressions, and speech data. It aims to improve the accuracy of emotion recognition by using advanced deep learning models such as CNN, LSTM, BERT, and GPT-based architectures. The system helps computers understand human feelings like happiness, sadness, anger, fear, surprise, and neutral emotions. It can be used in healthcare, education, customer service, mental health monitoring, and human-computer interaction. The project also focuses on creating a fast, reliable, and efficient multimodal emotion recognition framework for real-world applications.



**Figure 1: Proposed model**

The Figure 1 represents the workflow of a proposed emotion analysis system designed to identify emotions from textual sentences. The process begins with reading sentences from a dataset, which serves as the input source for the system. These sentences are then passed into the proposed system, where multiple stages of text and emotion processing are performed sequentially.

The first stage is Sentence Preprocessing, where the input text is cleaned and prepared for analysis. This may include tasks such as removing unwanted characters, converting text to lowercase, tokenization, stop-word removal, and stemming or lemmatization. The cleaned text is then forwarded to the Phrasal Verb Identification stage, where important phrases and contextual verb expressions are detected to better understand the meaning and emotional context of the sentence.

After identifying phrasal structures, the system performs Emotion Keyword Analysis, in which emotion-related words or phrases are extracted from the sentence. These keywords help determine the emotional tone expressed in the text, such as happiness, anger, sadness, fear, or excitement. The identified emotional terms are then processed in the Intensity Check stage, where the strength or degree of the detected emotion is evaluated. Words indicating stronger emotional expression are given higher importance.

Next, the system applies a Negation Check to identify negative expressions such as “not,” “never,” or “no,” which can alter the actual meaning of the emotion. For example, a sentence like “I am not happy” changes the emotional interpretation completely. Finally, the processed information is used to Calculate Emotion Score, where the overall emotional value of the sentence is computed based on keywords, intensity, and contextual factors.

Once the emotion score is calculated, the system generates the final emotional interpretation and ends the process. This model provides a structured and efficient approach for detecting and analyzing emotions from textual data.

### 3.1. Advantages

- **High Accuracy-** Deep neural networks can detect emotions more accurately than traditional methods by learning complex patterns from data.
- **Real-Time Detection-** The system can recognize emotions instantly from text, speech, or facial expressions, making it useful for live applications.
- **Multimodal Analysis-** It can combine text, image, and speech data for better and more reliable emotion recognition.
- **Improved Human Computer Interaction-** Machines can respond more naturally and empathetically to users.
- **Mental Health Monitoring-** Useful in detecting stress, anxiety, depression, and emotional changes in users.
- **Better Customer Service-** Companies can analyze customer emotions and improve support services.

### 3.2. Requirements

The proposed system is developed using advanced software and hardware requirements to ensure efficient performance in emotion recognition and machine learning tasks. The software environment includes Python 3.10 as the programming language along with powerful libraries such as TensorFlow 2.6, Keras 2.6, Scikit-learn, NumPy, Pandas, Matplotlib, OpenCV, and NLTK for deep learning, data processing, visualization, image processing, and natural language processing. Additionally, Librosa is used for speech and audio emotion analysis, while CSV and dataset files serve as the primary database for storing and managing data. The system is compatible with both Windows and Linux operating systems. On the hardware side, the system requires a minimum of 16 GB DDR4 RAM for smooth computation and multitasking, along with an NVIDIA RTX 3060

GPU (8 GB) or an equivalent graphics card to accelerate deep learning model training and inference. Storage requirements include either a 256 GB SSD or a 1 TB HDD for efficient data handling and application performance. Furthermore, a microphone is needed for capturing speech emotion inputs, while a webcam or camera is essential for facial emotion detection. Standard input devices such as a keyboard and mouse are also required for user interaction and system operation.

#### 4. CONCLUSION

The project "Emotion Detection System Using Deep Neural Network" successfully demonstrates the use of advanced deep learning techniques for recognizing human emotions through multiple data modalities such as text, image, and speech. By integrating powerful models like BERT, GPT-2, Mobile Net, ResNet, EfficientNet, and CNN-LSTM, the system achieved high accuracy and robust performance in emotion classification tasks. The multimodal fusion approach proved more effective than single-modality systems, as it combines contextual, visual, and vocal emotional cues for better understanding.

The results show that deep learning can significantly improve the accuracy, flexibility, and reliability of emotion detection systems in real-world environments. Such systems can be widely used in healthcare, education, customer support, mental health monitoring, and human-computer interaction. Although challenges such as data privacy, computational cost, and dataset diversity remain, the proposed framework provides a strong foundation for future research.

In conclusion, this study highlights that emotion-aware artificial intelligence can make machines more intelligent, responsive, and empathetic, leading to smarter and more human-centred technological solutions.

#### REFERENCES

1. Pan, Jiahui, et al. "Multimodal emotion recognition based on facial expressions, speech, and EEG." *IEEE Open Journal of Engineering in Medicine and Biology* 5 (2023): 396-403(30 Apr. 2026). DOI: 10.1109/OJEMB.2023.3240280
2. Khare, Smith K., et al. "Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations." *Information fusion* 102 (2024): 102019(30 Apr. 2026).
3. Kalateh, Sepideh, et al. "A Systematic Review on Multimodal Emotion Recognition." (2024) (30Apr. 2026). DOI: 10.1109/ACCESS.2024.3430850
4. Devlin, Jacob, et al. "Bert: Bidirectional encoder representations from transformers." *arXiv preprint arXiv:1810.04805* 15 (2018) (30Apr. 2026). DOI:10.18653/v1/N19-1423
5. Aslam, Naila, et al. "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model." *Ieee Access* 10 (2022): 39313-39324(30Apr. 2026). DOI: 10.1109/ACCESS.2022.3165621
6. Ibrahim, SP Siddique, et al. "Stress Detection using Facial Emotion Recognition: Comparative Analysis of VGG-16, ResNet-50, and a Custom CNN Architecture." *2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*. IEEE 2024(30Apr. 2026). DOI: 10.1109/ICDICI62993.2024.10810856.
7. Goreski, Mustafa Can, et al. "Facial emotion recognition (FER) through custom lightweight CNN model: performance evaluation in public datasets." *IEEE access* 12 (2024): 45543-45559.9. Hybrid CNN-LSTM and GRU models for multimodal emotion recognition(30Apr.2026). DOI: 10.1109/ACCESS.2024.3380847
8. de León Languré, Alejandro, and Mahdi Zareei. "Evaluating the effect of emotion models on the generalizability of text emotion detection systems." *IEEE Access* 12 (2024): 70489-70500.11. Studies on transformer-based multimodal fusion methods) (30Apr 2026); DOI: 10.1109/ACCESS.2024.3401203
9. "Text Dataset for Emotion" Kaggle. <https://www.kaggle.com/datasets/prajwalnayakat/text-emotion>(Accessed 29 Apr. 2026)
10. "Facial Expression Dataset" Kaggle. <https://github.com/astraszab/facial-expression-recognition> (accessed 30Apr. 2026)
11. "Speech Emotion Dataset" CREMA-D. <https://www.kaggle.com/code/ejlok1/audio-emotion-part-1-explore-data?scriptVersionId=20844990&cellId=32> (accessed 30Apr 2026)

- 
12. "Speech Emotion Dataset" RAVDESS <https://www.kaggle.com/code/ejlok1/audio-emotion-part-1-explore-data?scriptVersionId=20844990&cellId=15> (accessed 30Apr 2026)
  13. "Speech Emotion Dataset". TESS <https://www.kaggle.com/code/ejlok1/audio-emotion-part-2-feature-extract?scriptVersionId=20847848&cellId=5>(Accessed 30Apr 2026)
  14. Speech Emotion Dataset. .SAVEE <https://www.kaggle.com/code/ejlok1/audio-emotion-part-1-explore-data?scriptVersionId=20844990&cellId=2>(Accessed 30Apr 2026)