International Journal on Robotics, Automation and Sciences

Real-Time Emotion Detection Using Artificial Intelligence: A Review

Zoobiya Aalam, Saman Aziz, Kai Liang Lew*, and Chia Shyan Lee

***Abstract* – The integration of artificial intelligence (AI) in emotion recognition has significantly transformed human-computer interaction and revolutionized fields such as medicine, education, and entertainment. This paper reviews 30 papers on the detection of emotional signs through various biometric inputs, including electroencephalography (EEG), electrocardiography (ECG), facial expressions, and speech patterns. Despite advancements in AI-driven emotion recognition systems, challenges persist, particularly in data variability, computational inefficiency, and ethical dilemmas associated with privacy, security, and algorithmic bias. Recent innovations in feature extraction techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have enhanced the precision of emotional state recognition across multiple input channels. The transition to edge computing has further enabled real-time processing with low latency, facilitating integration into wearable devices and IoT ecosystems. Multimodal systems, which leverage data sources such as physiological signals, facial expressions, and speech, show great promise but face challenges related to inclusivity and system fragility. To address these issues, the study recommends for robust training datasets, ethical guidelines, and hardware optimizations. Incorporating contextual information and accounting for individual differences can improve recognition accuracy and user trust. However, ethical concerns remain critical, emphasizing the need for strict standards of privacy, security, and equitable access to ensure AI emotion recognition systems are trustworthy and inclusive. Overall, this paper highlights the potential of AI-driven emotion recognition systems while underscoring the importance of continuous research to address technical and ethical challenges, paving the way for broader applications in pattern recognition, cognitive studies, and specialized tools.**

***Keywords*— Emotion Recognition And Detection, Multimodal Systems, Deep Learning, Affective Computing, Ethical Ai**

I. Introduction

Real-time emotion perception and interpretation are fundamental innovations within affective computing technologies that integrate emotional understanding into decision-making processes across various applications such as personalized health services, intelligent tutoring systems, and mental health monitoring. These sophisticated emotion recognition systems interpret diverse sensory data derived from human emotional behaviors including facial expressions and neuroscientific analyses such as brain and physiological responses to transform these behaviors into meaningful insights about users' emotions.

They hold the potential not only to enhance user experience by providing immediate feedback in educational or customer service contexts but also to detect stress indicators or mental health concerns thus offering essential support despite their promise deploying advanced algorithms for emotion recognition presents enduring challenges dealing with

*Corresponding Author email: 1132703002@student.mmu.edu.my, ORCID: 0000-0002-0376-2970
Zoobiya Aalam is with Riphah International University, Islamabad, Pakistan (email: zoobiyaaalam12@gmail.com)
Saman Aziz is with Riphah International University, Islamabad, Pakistan (email: samanaziz955@gmail.com)
Kai Liang Lew is with Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia (e-mail: 1132703002@student.mmu.edu.my).
Chia Shyan Lee is with Curtin University, Perth, Australia (email: cat_lee97@hotmail.com).

International Journal on Robotics, Automation and Sciences (2025) 7, 1:104-110
https://doi.org/10.33093/ijoras.2025.7.1.12
Manuscript received:4 Jan 2025 | Revised: 13 Feb 2025 | Accepted: 27 Feb 2025 | Published: 31 Mar 2025
Published by MMU PRESS. URL: http://journals.mmupress.com/ijoras

unpredictable noisy environments requires adaptability while cross-cultural variations necessitate sensitivity in data interpretation moreover limitations in the scope and scale of available datasets pose significant hurdles exacerbated by ethical considerations regarding user consent and data privacy which complicate system deployment amidst ongoing debates in this area.

In consolidating insights from over 30 peer-reviewed studies, this review aims to provide a holistic overview of the latest advancements, imminent threats, and future outlook within this rapidly evolving field. With a focus on deep learning techniques, multimodal approaches, and ethical considerations, the sources were selected based on their suitability for real-time emotion recognition. Older or fewer comprehensive works were ignored in favor of more recent, peer-reviewed studies with strong empirical backing to produce a comprehensive and up-to-date review.

Recent progress includes hybrid systems that merge deep learning techniques with traditional classifiers demonstrating noteworthy robustness improvements in emotional recognition tasks the integration of multimodal approaches such as combining audio signals with visual cues extends promising avenues for enhancing emotion detection accuracy by leveraging contextual data. Furthermore, the advent of Edge AI solutions represents a significant leap forward by enabling scalable and privacy-preserving implementations these systems process information locally at or near the data source thus reducing latency and enhancing real-time responsiveness in emotionally-aware applications developments in transfer learning techniques have brought additional optimism by demonstrating improved model performance even with constrained training datasets thereby widening applicability across diverse settings.

In designing future emotion recognition technologies an emphasis on ethical considerations remains paramount incorporating transparency and user control mechanisms is essential to fostering trust and broader acceptance of these innovative systems as we continue refining emotional intelligence in human-computer interactions achieving a harmonious blend of creativity with ethical integrity will shape the guiding principles for advancing affective computing technologies.

II. METHODOLOGY

A. *Modalities for Emotion Detection*

Emotion detection systems rely on two primary types of data, including physiological signals and behavioral cues.

Physiological Signals

Electroencephalographic (EEG) recordings, reflecting electrical activity in the brain, constitute one of the superior modalities used for emotion recognition on the grounds of their depiction of emotional states [1]. Basic features such as power spectral density, asymmetry in EEG, and entropy are extracted in a very robust manner concerning arousal and valence [2] [3]. The activities from ECG provide information about the inter-assay heart-rate variability, relating to emotional states like stress, relaxation, and excitement[4]. An enhancement of additional physiological signals, like GSR and respiration, improves classification accuracy when combined with EEG or ECG. This multimodal approach has tied the best of the two sensory systems together to achieve a more comprehensive understanding of emotional response. The recent advent of sophisticated machine learning techniques has made feature extraction and classification operations more convenient, leading to more cheerful emotion detection prototypes in real-time applications. As the research in this domain develops, there is a pressing need for robust emotion recognition systems, which can be applied in a variety of contexts-all addressing the challenges of variability in data and environmental noise [5].

Behavioral Cues

The facial expressions are the most intuitive and the most widely recognized indicators of human emotions. The reason of high identification could be given to the sophisticated convolutional neural network architectures that were being enforced in the identification of "affect-related" subtle movements, thus allowing the identification of emotions and their classification by emotional effectiveness. Given that facial emotion detection is susceptible to external hindrances, including light-dependent issues and cultural circumstances, one should no less examine speech features [6] [7]. Technologies can be expertly used in conducting voice features, such as pitch, tone, and rhythm to categorize the different emotions. These evidence-based systems worked quite well in controlled circumstances but were limited in detecting true data when the analysis was completed in the presence of by noise and language variances. The integration of facial and speech-based emotion recognition can improve overall accuracy through the extra data streams, with further research demonstrating the need to address environmental challenges associated with emotion detection. Investigating the limitations imposed by external factors in future works would perhaps pave the way for more robust emotion-detecting systems. Therefore, improving algorithms and integrating diverse training datasets further will propel the field of emotion detection toward more solid and less error-prone applications in different fields, including mental health monitoring, and user experience optimization [8] [9].

Multimodal Approaches

With an increasing focus in recent research on the importance of integrating multimodal systems that might overcome the limitations of single-modal systems, for example, EEG signals combined with facial expression data or speech demonstrating higher classification accuracy and more resilience to noise [10] [11]. Although multimodal architectures need voluminous computable resources, they are the most up-to-date for real-time emotion recognition [12].

Prospectively, further new multimodal fusion strategies such as attention models and graph-based learning could enhance systems by dynamically neglecting the most informative modalities, yet with the condition of lessening the computational load, while sustaining accuracy in operation. Such advanced strategies use attention models by attending to those inputs that are detected as most relevant to optimize computational resources so that emotion accrual could be impacted positively in varied environments. This is augmented through deep learning, which is destined to improvise the fusion process so that systems become able to adapt themselves to diverse conditions and user contexts. It is rational to hope that carrying out genuine research will provide opportunities for robust and scalable emotion detection applications that can be employed in various crucial sectors like health, education, and marketing support. Table 1 shows the emotion detection modalities with their strength and weaknesses

TABLE 1. Emotion detection modalities with their strength and weaknesses.

Modality	Description	Strengths	Weakness
EEG	Measures brain activity to deduce emotions	High accuracy	Complex setup
ECG	Tracks heart rate variability	Complemen-tary to EEG	Susceptible to noise
Facial Expression	Analyze facial muscle movement	Intuitive and Easy to capture	Affected by lighting
Speech	Detects, Tone, Pitch and rhythm in voice	Language independent features	Sensitive to Background noise

B. *AI Techniques for Emotion Detection*

AI algorithms form the backbone of modern emotion recognition systems, particularly deep learning methods and techniques

Convolutional Neural Networks (CNNs)

CNNs are utilized significantly for image data analysis, particularly facial expression recognition. They ensure a balance between accuracy and program performance load, being exceptionally well-equipped for on-spot measuring applications. The very light version of the MobileNet model is primarily designed for resource-restricted settings where the combination of fast processing speed and high accuracy is preserved [13] [14]. More advanced versions, such as attention-based CNNs, aim at enriching the ability of distinguishing very nuanced facial emotions intending to learn specific critical features from the input data [15]. This way, the proposed method can more brutally pinpoint subtle variations in expression which very likely indicate different emotional states. Besides, using techniques such as data augmentation and transfer learning can go a long way toward greater model robustness and generalization on different datasets. As a result, advancements in CNN architecture would potentially speed up emotion recognition research and inject real-time applications with greater meaning in terms of the end-user experience.

Recurrent Neural Networks (RNNs)

RNN architectures that include LSTM features, especially for EEG signal processing, have a good impact because LSTMs are very good at capturing the inherent temporal dependencies needed to consider evolving emotional states [16]. However, somewhat recently, research has continued to investigate the incorporation of attention in LSTM networks, which helps in pinpointing critical points in the timeline of the signal being analyzed. Not only does this improvement boost the emotion-detection accuracy of the model in structured and "noisy" environments, but also attention mechanisms now allow LSTMs to more efficiently utilize resources by focusing on important data points and ensuring that the information that most matters is attended to and weighed with regard to its relevance in its context. These processes together allow for training neural networks to be more adaptable to the variability found in input sequences, therefore making the neural networks robust against noise and fluctuations. Hence, deploying models that include combinations of LSTM networks and attention mechanisms significantly stimulates further research into emotion recognition, with the prospect of creating systems that are more precise and responsive and fit for adapting to real-world applications

Hybrid Models

One of the successful trends regarding emotion recognition consists of hybridizing CNNs, which capture spatial information with RNNs, which cater for time-variant data. These hybrid systems excel in multimodal methods where varied data types require specialize processing[17]. Furthermore, Transformer-based architectures are being considered as a more potent alternative since they provide excellent means for capturing spatial-temporal dependencies across various data streams, thus overcoming issues with traditional CNN-RNN hybrids. Transformers serve sequential input information via self-attention mechanisms that allow dynamic allocation of importance toward different input elements; this affects the model to increasingly recognize complex patterns and relationships within the data. As research continues, merge these advanced architectures into emotion recognition systems for great enhancement in accuracy and efficiency, therefore creating fast-track implementations in various fields, including healthcare, security, and human-computer interaction. Table 2 shows the AI techniques used in Emotion Detection

TABLE 2. AI Techniques used in Emotion Detection.

Techniques	Descriptions	Applications
CNNs	Used for image based data analysis (e.g., Facial expressions)	Real time emotion recognition

RNNs	Suitable for sequential data (e.g.,EEG Signals)	Capturing temporal dependencies
Hybrid Models	Combines CNNs and RNNs for data processing)	Enhanced accuracy in emotion detection

C. Hardware-Accelerated Systems

With the surge in demand for real-time performance, edge computing along with hardware accelerators (FPGA) is now integral to any emotion detection system. These advancements while maintaining accuracy, and low latency processing lend itself to the practical use of emotion recognition in portable and embedded devices. Data processing can now be pushed to the source at the edge, allowing for reduced response time and improved user experience. On another hand, FPGAs provide customizable hardware platforms for enhanced processing efficiency so that complex algorithms can run smoothly in real-time applications. The combination of these technologies not only fulfills modern emotion detection system requirements but also eases the deployment in several environments of mobile devices and IoT applications, thereby extending the real-time emotional analysis capabilities across industries [18] [19].

D. Dataset Challenges

The datasets are the mainstay of emotion recognition research, yet openly available datasets such as DEAP (for EEG use), FER-2013 (for facial expressions), and RAVDESS (for speech) widely used suffer from class imbalance and lack of cultural diversity. [20]. Many datasets predominantly consider basic emotions and, in this way, sidestep more complex emotional states like ambiguity or neutrality [21]. Future endeavors must prioritize the developmental objectives of diverse and well-annotated datasets to address these gaps. Including mixed emotions and a wider context of cultures would increase the usability of emotion recognition systems for various populations. If they increase the variety of available datasets, researchers can assist model training and validation processes, ultimately leading to increased accuracy and reliability of emotion detection in real-world situations. Addressing the above issues is necessary for this field to advance and for the fair effectiveness of emotion recognition technology in various application [22].

III. Discussion

A. Achievements and Current Trends

Emotion detection systems have made great advances, particularly for multimodal integration. Hybrid systems integrating EEG with facial expressions and speech inputs have reached accuracies above 95% in controlled condition [23]. Further advancements, aided by deep learning-based techniques such as transfer learning and generative models, have augmented the classifiers addressing the more complex dynamics of emotional states. Such methodologies thus improve the feature extraction and generalization capabilities of the model on varied datasets, thereby enhancing the robustness of emotion recognition. In conclusion, improvements in data preprocessing and augmentation techniques ensure better performance of models by tackling issues such as class imbalance and noise. As these technologies develop, their prospective applications in real-world settings become increasingly plausible for mental health monitoring, user experience assessment, and interactive systems which create opportunities for stronger emotional intelligence in many fields [24]. Figure 1 shows the classification accuracy of different modalities

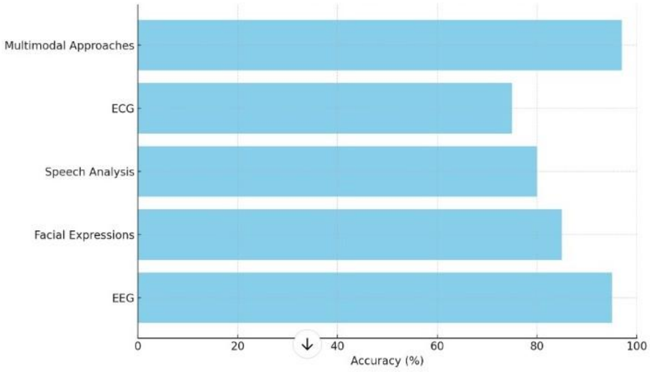

FIGURE 1. Classification accuracy of different modalities.

B. Challenges

Although emotional recognition has made remarkable advancements in science, there are still quite a few roadblocks in the field. Cultural differences interfere with the reliability of such systems, as people from different societies may display very different "joy" and "sad" expressions" and therefore may be prone to classification errors [25]. Computational concerns become more serious, as some real-time systems are an enormous drain on resources in processing multimodal data streams [26]. On the other hand, ethical concerns, particularly surveillance and privacy, deter further acceptance of these technologies [27]. Therefore, the further course should focus on adaptive emotion recognition models that consider cultural specifications, performing optimally by using light-weight architecture to minimize computational overheads, and employing federated learning in favor of user privacy concerning decentralized systems. If these issues are taken care of, we can work toward increasing the efficacy and acceptability of emotion detection technology in many applications and populations.

C. Future Directions

Addressing open research difficulties and broadening the systems' range of applications should be the key goals of future emotion recognition studies. To enhance model generalization and reduce biases, one important topic is the development of inclusive and culturally representative datasets. Additionally, the implementation of emotion detection technologies in portable and real-time applications will be made possible by furthering hardware simplification and optimizing algorithms for resource-constrained contexts. Furthermore, combining methods like federated learning, transfer learning, and data augmentation can improve model generalization and lessen reliance on sizable labelled datasets. These techniques will increase the flexibility of emotion detection systems for practical applications by allowing them to dynamically adjust to various populations and situations. Labeled

In this discipline, ethical considerations are also quite important. To guarantee user trust and fair deployment, precise rules for data collection, storage, and use must be established. Research on emotion detection should also look into how it might be used in human-computer interaction and mental health monitoring, as these fields offer important insights into user experience and emotional health. Future advancements in emotion detection technology can result in more precise, inclusive, and morally sound systems with broad applications in human-computer interaction, healthcare, and education by tackling these issues. Table 3 shows the comparison of emotion recognition methods across different modalities, feature extraction techniques, and ml/dl algorithms.

TABLE 3. Comparison of Emotion Recognition Methods Across Different Modalities, Feature Extraction Techniques, and ML/DL Algorithms.

Reference	Modality Used	Feature Extraction Method	ML/DL Algorithm	Dataset Used	Accuracy (%)	Key Findings
Fang et al. (2019) [1]	ECG	Power Spectral Density	CNN	DEAP	85%	EEG-based recognition is highly accurate for valence and arousal.
Hassouneh et al. (2020) [2]	ECG	Hybrid Model	Deep Neural Network	AMIGOS	80%	EEG-based recognition is highly accurate for valence and arousal.
Turabzadeh et al. (2018) [3]	ECG + EEG	CNN Feature Extraction	CNN	SEED	88%	Multimodal approaches improve performance.
Jaiswal et al. (2019) [4]	Facial Expressions	CNN-based Feature Extraction	Deep Learning Model	FER-2013	92%	Deep learning models outperform traditional classifiers.
Kartal et al. (2018) [5]	Speech	MFCC Feature Extraction	LSTM	RAVDEES	78%	Speech emotion recognition struggles with environmental noise.
Talast et al. (2023) [6]	EEG + Facial	Feature-Level Fusion	Bi-LSTM	IEMOCAP	90%	Fusion models increase robustness against data variability.
Jiang et al. (2020) [7]	Speech + Facial	Fusion Network	Transformer	AFEW	87%	Multimodal learning improves generalizability.

IV. Conclusion

Recently, artificial intelligence systems for detecting emotion in real-time have gained significant momentum primarily due to multimodal approaches and deep learning techniques. A combination of heterogeneous data sources such as facial expressions, vocal signals, and physiological signals gives a much deeper perspective into human emotions. Diversity in the datasets, computational constraints, and ethical issues are challenges that still remain. Indeed, many available databases do not have fair representation across cultures, which can introduce undesired bias in emotion recognition. The processing of multimodal data streams may require a great deal of computational resources, further restricting the use of such systems in real time. By developing inclusive datasets, working on alternatives with lower power consumption like FPGAs, and instituting ethical guidelines, these problems, when resolved, will herald a full realization of emotion detection technologies, leading to the creation of revolutionary applications: in healthcare for diagnosing mental health; in education for the personalization of learning; in public safety for threat identification; and human-computer interaction for a more empathetic user experience. Emotion detection technologies would thus have wide-ranging application in virtually all sectors, which means that with the resolution of these challenges will surely enhance not just how technology interacts with us, but they will also improve societal wisdom.

Acknowledgment

There is no funding agencies supporting the research work.

Author Contributions

Zoobiya Aalam: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation;

Saman Aziz: Project Administration, Writing – Review & Editing;

Kai Liang Lew: Writing – Review & Editing.

Chia Shyan Lee: Writing – Review & Editing

Conflict of Interests

No conflict of interest was disclosed.

Ethics Statements

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org/

References

[1] W. Fang, K. Wang, N. Fahier, Y. Ho and Y. Huang, "Development and Validation of an EEG-Based Real-Time Emotion Recognition System Using Edge AI Computing Platform With Convolutional Neural Network System-on-Chip Design," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 4, pp. 645-657, 2019. DOI: https://doi.org/10.1109/JETCAS.2019.2951232

[2] A. Hassouneh, A. Mutawa and M. Murugappan, "Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods," *Informatics in Medicine Unlocked*, vol. 20, pp. 100372, 2020. DOI: https://doi.org/10.1016/j.imu.2020.100372

[3] S. Turabzadeh, H. Meng, R. Swash, M. Pleva and J. Juhar, "Facial Expression Emotion Detection for Real-Time Embedded Systems," *Technologies*, vol. 6, no. 1, pp. 17, 2018.
DOI: https://doi.org/10.3390/technologies6010017

[4] S. Jaiswal and G.C. Nandi, "Robust real-time emotion detection system using CNN architecture," *Neural Computing and Applications*, vol. 32, no. 15, pp. 11253-11262, 2020. DOI: https://doi.org/10.1007/s00521-019-04564-4

[5] A. Kartali, M. Roglić, M. Barjaktarović, M. Đurić-Jovičić, and M. M. Janković, "Real-time algorithms for facial emotion recognition: A comparison of different approaches," *14th Symposium on Neural Networks and Applications (NEUREL)*, pp. 1-6. 2018.
DOI: https://doi.org/10.1109/NEUREL.2018.8587011

[6] S. Gupta, P. Kumar and R.K. Tekchandani, "Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models," *Multimedia Tools and Applications*, vol. 82, no. 8, pp. 11365-11394, 2023.
DOI: https://doi.org/10.1007/s11042-022-13558-9

[7] M. Oki, S. Akintola, A. Adediran, P. Ikubanni and B.T. Ogunsemi, "Reinforcement bar Corrosion - Causes and Management," *Journal of Physics: Conference Series*, vol. 1378, no. 2, pp. 022095, 2019.
DOI: https://doi.org/10.1088/1742-6596/1378/2/022095

[8] S. Gupta, "Facial emotion recognition in real-time and static images," *2nd International Conference on Inventive Systems and Control (ICISC),* pp. 108, 2018.
DOI: https://doi.org/10.1109/ICISC.2018.8398861

[9] M. Awais, M. Raza, N. Singh, K. Bashir, U. Manzoor, S.U. Islam and J.J.P.C. Rodrigues, "LSTM-Based Emotion Detection Using Physiological Signals: IoT Framework for Healthcare and Distance Learning in COVID-19," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16863-16871, 2021.
DOI: https://doi.org/10.1109/JIOT.2020.3044031

[10] A. R., A. K., and D. S. Jaiswal, "Facial emotion detection using deep learning," *International Conference for Emerging Technology (INCET)*, pp. 1-6, 2020.
DOI: https://doi.org/10.1109/INCET49848.2020.9154121

[11] K. Zhang, Y. Li, J. Wang, E. Cambria and X. Li, "Real-Time Video Emotion Recognition Based on Reinforcement Learning and Domain Knowledge," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1034-1047, 2022.
DOI: https://doi.org/10.1109/TCSVT.2021.3072412

[12] F. M. Talaat, "Real-time facial emotion recognition system among children with autism based on deep learning and IoT," *Neural Computing and Applications,* vol. 35, no. 2, pp. 1817–1830, 2023.
DOI: https://doi.org/10.1007/s00521-023-08372-9

[13] Y. Jiang, W. Li, M. S. Hossain, M. Chen, A. Alelaiwi, and M. Al-Hammadi, "A snapshot research and implementation of multimodal information fusion for data-driven emotion recognition," *Information Fusion*, vol. 53, pp. 55–62, 2020. DOI: https://doi.org/10.1016/j.inffus.2019.06.019

[14] K. Bahreini, W.v.d. Vegt and W. Westera, "A fuzzy logic approach to reliable real-time recognition of facial emotions," *Multimedia Tools and Applications*, vol. 78, no. 14, pp. 18943-18966, 2019.
DOI: https://doi.org/10.1007/s11042-019-7250-z

[15] M.A. Helal, A.I. Chowdhury, A. Islam, E. Ahmed, M.S. Mahmud and S. Hossain, "An Optimization Approach to Improve Classification Performance in Cancer and Diabetes Prediction," *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2019. DOI: https://doi.org/10.1109/ECACE.2019.8679413

[16] M.A. Ozdemir, B. Elagoz, A. Alaybeyoglu, R. Sadighzadeh and A. Akan, "Real Time Emotion Recognition from Facial Expressions Using CNN Architecture," *2019 Medical Technologies Congress (TIPTEKNO)*, pp. 1-4, 2019. DOI: https://doi.org/10.1109/TIPTEKNO.2019.8895215

[17] L. Zahara, P. Musa, E.P. Wibowo, I. Karim and S.B. Musa, "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi," *2020 Fifth International Conference on Informatics and Computing (ICIC)*, pp. 1-9, 2020. DOI: https://doi.org/10.1109/ICIC50835.2020.9288560

[18] E. Pranav, S. Kamal, C.S. Chandran and M. Supriya, "Facial Emotion Recognition Using Deep Convolutional Neural Network," *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020. DOI: https://doi.org/10.1109/ICACCS48705.2020.9074302

[19] A.I. Siam, N.F. Soliman, A.D. Algarni, F.E.A. El-Samie and A. Sedik, "Deploying Machine Learning Techniques for Human Emotion Detection," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-16, 2022.
DOI: https://doi.org/10.1155/2022%2F8032673

[20] K. S. Mustaqeem, "Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network," *Mathematics*, vol. 8, no. 8, 2020.
URL: https://ph01.tci-thaijo.org/index.php/easr/article/view/252462 (accessed: 27 Feb 2025)

[21] P. M. Ghate, B. D. Jadhav, P. N. Kota, S. D. Chavan and P.B. Chopade, "Optimized intelligent speech signal verification system for identifying authorized users," *Engineering & Applied Science Research*, vol. 50, no. 3, pp. 233–245, 2023. DOI: https://doi.org/10.14456/easr.2023.20

[22] S. R. Shahamiri, "An optimized enhanced-multi learner approach towards speaker identification based on single-sound segments," *Multimedia Tools and Applications*, vol. 83, pp. 24541–24562, 2024.
DOI: https://doi.org/10.1007/s11042-023-16507-2

[23] M.J.A. Nahian, T. Ghosh, M.N. Uddin, M.M. Islam, M. Mahmud and M.S. Kaiser, "Towards Artificial Intelligence Driven Emotion Aware Fall Monitoring Framework Suitable for Elderly People with Neurological Disorder," *Lecture Notes in Computer Science*, pp. 275-286, 2020.
DOI: https://doi.org/10.1007/978-3-030-59277-6_25

[24] S. Kwon, M. Khan, "A CNN-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 19, no. 23, 2019.
DOI: https://doi.org/10.1002/int.22505

[25] V. Vielzeuf, C. Kervadec, S. Pateux, A. Lechervy and F. Jurie, "An Occam's Razor View on Learning Audiovisual Emotion Recognition with Small Training Sets," *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, pp. 589-593, 2018.
DOI: https://doi.org/10.1145/3242969.3264980

[26] K. S. Mustaqeem, "Optimal feature selection based speech emotion recognition using two-stream deep convolutional neural network," *International Journal of Intelligent Systems*, vol. 36, no. 2, pp. 708–731, 2021.
DOI: https://doi.org/10.1002/int.22505

[27] P. Mohan, L. Sharma and P. Narayan, "Optimal Path Finding using Iterative SARSA," *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 811-817, 2021. DOI: https://doi.org/10.1109/ICICCS51141.2021.9432202
[28] I. Sergi, T. Montanaro, F.L. Benvenuto and L. Patrono, "A Smart and Secure Logistics System Based on IoT and Cloud Technologies," *Sensors*, vol. 21, no. 6, pp. 2231, 2021. DOI: https://doi.org/10.3390/s21062231
[29] R. Venkatesan, S. Shirly, M. Selvarathi and T.J. Jebaseeli, "Human Emotion Detection Using DeepFace and Artificial Intelligence," *RAiSE-2023*, pp. 37, 2023. DOI: https://doi.org/10.3390/engproc2023059037
[30] J. Chen, K. Yu, F. Wang, Z. Zhou, Y. Bi, S. Zhuang and D. Zhang, "Temporal Convolutional Network-Enhanced Real-Time Implicit Emotion Recognition with an Innovative Wearable fNIRS-EEG Dual-Modal System," *Electronics*, vol. 13, no. 7, pp. 1310, 2024. DOI: https://doi.org/10.3390/electronics13071310
[31] Y. Xu, Y. Lin, X. Zhou and X. Shan, "Utilizing emotion recognition technology to enhance user experience in real-time," *Computing and Artificial Intelligence*, vol. 2, no. 1, pp. 1388, 2024. DOI: https://doi.org/10.59400/cai.v2i1.1388
[32] B. Bakariya, A. Singh, H. Singh, P. Raju, R. Rajpoot and K.K. Mohbey, "Facial emotion recognition and music recommendation system using CNN-based deep learning techniques," *Evolving Systems*, vol. 15, no. 2, pp. 641-658, 2024. DOI: https://doi.org/10.1007/s12530-023-09506-z
[33] C. Suneetha and R. Anitha, "Advancements in Speech-Based Emotion Recognition and PTSD Detection through Machine and Deep Learning Techniques: A Comprehensive Survey," *International Journal of Electronics and Communication Engineering*, vol. 11, no. 5, pp. 220-234, 2024. DOI: https://doi.org/10.14445/23488549/IJECE-V11I5P121
[34] H.B.U. Haq, W. Akram, M.N. Irshad, A. Kosar and M. Abid, "Enhanced Real-Time Facial Expression Recognition Using Deep Learning," *Acadlore Transactions on AI and Machine Learning*, vol. 3, no. 1, pp. 24-35, 2024. DOI: https://doi.org/10.56578/ataiml030103
[35] A. M. Mutawa and A. Hassouneh, "Multimodal Real-Time patient emotion recognition system using facial expressions and brain EEG signals based on Machine learning and Log-Sync methods," *Biomedical Signal Processing and Control*, vol. 91, pp. 105942, 2024. DOI: https://doi.org/10.1016/j.bspc.2023.105942
[36] M.S. Rao, A. Kiran, S.L. Naik, A.P. Kumar, K.K. Swamy and P. Devika, "Facial Emotion Recognition with Convolutional Neural Networks using DL," *2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, pp. 1-5, 2024. DOI: https://doi.org/10.1109/ASSIC60049.2024.10507979