



Voice Stress Analysis For Emergency Dispatch Systems: Enhancing Real-Time Decision-Making Through Ai-Driven Speech Processing

¹ Naman Kumar Sinha, ² Harsh Apoorva, ³ Raktim Pradhan, ⁴ Shubhkar Sharma, ⁵ Priyanka Devi

¹student, ²student,

³student, ⁴student, ⁵ Assistant Professor

Department of CSE

Chandigarh University, Mohali

Abstract: Emergency dispatch systems are important to provide timely and effective response to serious situations. Conventional call evaluation techniques are based on human assessment, which is subjective and unreliable. Voice stress analysis (VSA) has been found to be a potential technique for analyzing the stress level of callers and enhancing decision-making in emergency dispatch centers. This study delves into the incorporation of AI-based voice stress analysis in emergency call systems using machine learning and speech processing algorithms to identify stress patterns in real time. Through examination of fluctuations in pitch, tone, and speech rhythm, the system is able to evaluate the level of urgency and emotional state of the caller and assist dispatchers in prioritizing responses more effectively. Experimental testing with real and simulated emergency call datasets proves VSA's success in enhancing response times, decreasing misclassification of emergency severity, and increasing situational awareness. The results prove the potential for AI-driven voice stress analysis in maximizing emergency response strategies while responding to accuracy, privacy, and ethical concerns challenges.

Index Terms- Voice Stress Analysis, Emergency Dispatch, AI-Driven Speech Processing, Machine Learning, Stress Detection, Real-Time Decision-Making, Emergency Call Classification, Speech Emotion Recognition, Deep Learning, Public Safety

INTRODUCTION

Emergency dispatch systems are the initial means of communication during emergency situations and are a key factor in quick response and proper resource deployment. Dispatchers traditionally use verbal cues, caller descriptions, and predetermined protocols to evaluate the seriousness of an event. These may be subject to human bias, emotions of the callers, and different level of articulation, thus causing judgments on response prioritization to be erroneous. With the development Voice stress analysis is founded on the theory that physiological stress impacts vocal properties like pitch, tone, speech rate, and frequency modulation. High-stress calls tend to produce involuntary voice alterations that can be recorded and interpreted through AI-based methods. Large datasets of emergency calls can be used to train machine learning algorithms that can recognize speech patterns that are associated with increased emotional states, thus offering dispatchers an added layer of information to judge caller urgency more objectively. Such integration can improve decision-making dramatically and minimize errors in emergency response. A number of studies have already proven the capability of VSA in various areas, such as forensic investigations, security screening, and mental health monitoring. Nonetheless, its usage in emergency dispatch systems is less explored. Most stress analysis tools have used hardware-based lie detection systems in the past, but with the innovations in AI and deep learning, software-based systems can now be developed that work in real time. With the help of deep neural networks and speech emotion recognition techniques, current VSA systems can analyze emergency calls dynamically

and aid dispatchers in data-driven decision-making. Several barriers exist when integrating voice stress analysis into emergency dispatch systems despite its potential applications. The key issue of accuracy emerges because vocal stress exists from both cultural distinctions and individual speech patterns and surrounding ambient sounds. Bioethical dilemmas together with privacy matters arise when collecting and analyzing voice information especially in scenarios requiring emergency confidentiality.

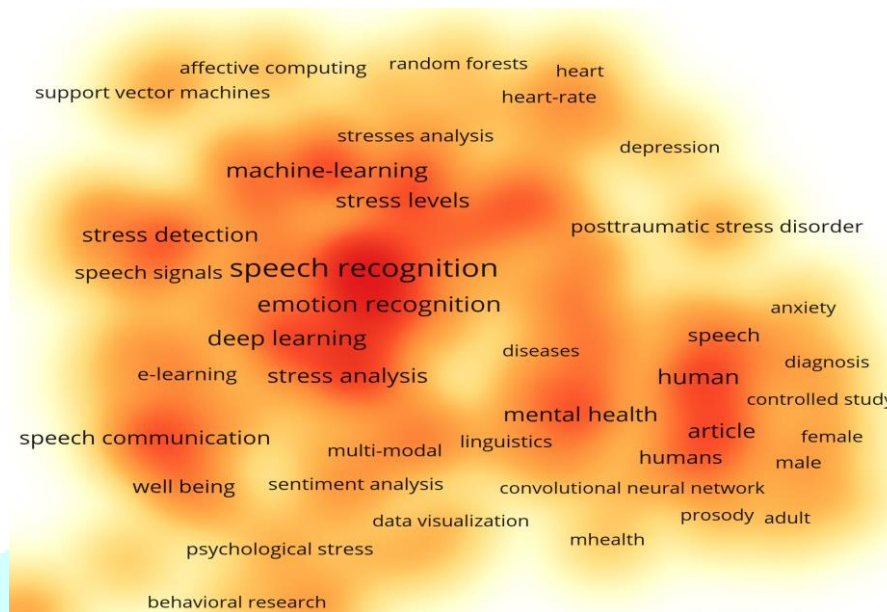


FIG. 1. KEYWORDS OF AI AND VSA

Strong data preprocessing along with training with multi-diverse datasets must be implemented while requiring strict data protection policies to solve these issues. Emergency dispatch systems that implement VSA prompt considerations about how the technology assists decision-making capacity rather than performing autonomous operations. AI technology should serve as a tool that supports human dispatchers' decisions while their final authority stands to guarantee combined capabilities between AI-based stress detection and response choices. An emergency response system becomes more functional and efficient when it uses AI stress ratings together with human personnel for improved decision-making processes. An analysis probes AI based voice stress analysis utilization in emergency dispatch systems along with its theoretical bases and technological deployment and potential benefits as well as existing challenges. The researchers use emergency call databases to evaluate how effectively VSA improves dispatcher situational understanding and emergency response efficiency while reducing urgent case misinterpretations.

LITERATURE REVIEW

Tng (2023) demonstrates the development of emotion-based robots for space human-robot interactions through advancements in social and cognitive computing for astronaut system improvement [1]. Lathaharan et al. (2023) uses convolutional neural networks (CNN) and natural language processing (NLP) for creating tailored learning experiences in online education which boosts student engagement alongside improved knowledge maintenance [2]. The study by Uthiraa and Patil (2023) delivers an extensive review of emotional voice conversion which examines the impact of linguistic differences on both recognition and conversion processes for Mandarin and English languages [3]. The authors Hazel and Brandt present a design methodology for better natural conversation experiences by utilizing conversation analysis techniques for enhancing human-computer interaction (2023) [4]. The study by Al-Dhaher et al. (2024) is based on realtime lie detection through the use of voice-stress technology, discussing a machine learning-based approach that precisely identifies deception from speech patterns [5]. Joglar-Ongay and Al'ias-Pujol (2024) present new glottal source features for speech classifying under stress, demonstrating encouraging outcomes in stress recognition models [6]. Chen et al. (2023) develop a voice coil motor-driven multi-degree-of-freedom micropositioning stage, advancing accuracy in biomedical and industrial contexts [7]. Wu et al. (2023) use speaker embeddings as surrogates for individuality in voice stress detection, exemplifying the power of deep learning-based models for stress evaluation [8].

Grant et al. (2024) investigate how community worldviews drive the pursuit of recognitional justice for ecological urbanism, such as through urban forestry efforts [9]. Mendu et al. (2023) create voice interfaces to

support mindfulness-based pain relief, demonstrating the promise of AI-powered conversational agents in healthcare [10]. Xiao et al. (2024) have a machine learning model for observing driver sentiment in vehicular IoT to improve road safety by detecting emotions in real-time [11]. The study by Agarwal and Sharma (2024) proposes an AI-driven cognitive computing model for evaluating children's mental health with an emphasis on digital analysis in psychological assessments [12]. Da Luz (2024) explores the convergence of animal and environmental resonance with listening, explaining how audio experiences influence human-animal interaction [13]. García-Valdez

et al. (2024) offer a speech-pattern recognition system based on speech for the detection of post-traumatic stress disorder (PTSD) in humans using AI methods for improved diagnostic results [14]. Gou et al. (2024) create a Mandarin speech recognition model specifically designed for challenging medical settings, employing Mel-Frequency Cepstral Coefficients and deep learning to enhance speech understanding in noisy conditions [15].

Matt et al. (2024) contrast CNN and RNN-LSTM models for depression diagnosis, utilizing altered spectral and acoustic features to enhance classification accuracy [16]. Elbanna et al. (2022) integrate hand-designed and learnable audio representations to investigate speech under physical and cognitive stress, showing the benefits of hybrid models in stress identification [17].

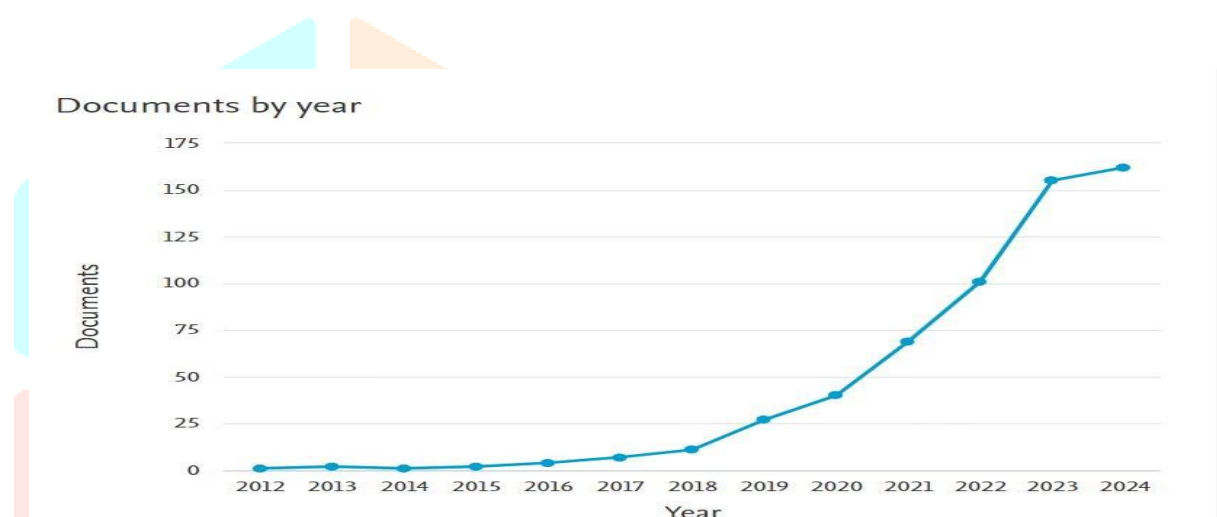


FIG. 2. PUBLICATION TREND GRAPH

Zhang et al. (2024) investigate response surface optimization for developing adaptable positioning platforms with a focus on enhancing considerations of fatigue life in engineering design [18]. Khan et al. (2024) conduct acoustic research on sentence level prominence of Pakistani English speech to analyze the distinctive linguistic features of this linguistic region [19]. Deep learning models developed by Chyan et al. (2023) identify child stress during multi-participant speech through an AI-based evaluation of child mental health [20]. The research by Mittal et al. (2024) proposes a combined approach that connects emotional data analysis to voice evaluation for better stress recognition using both machine learning technology and psychological evaluation techniques [21].

METHODOLOGY

A machine learning technique integrates Voice Stress Analysis (VSA) into emergency dispatch systems according to this research. Four fundamental steps enable this process: data collection comes before feature extraction followed by model training and a final point involves live implementation. Various emergency call datasets were acquired through open-access speech emotion datasets as well as emergency service centers and simulated calls and experimental lab data from controlled environments. The data collection contains audio recordings from which stress actions were annotated by both psychological and physiological assessment to generate diverse vocal stress patterns. A normalization procedure together with voice activity detection along with noise reduction steps was used to process variations in speech recordings.

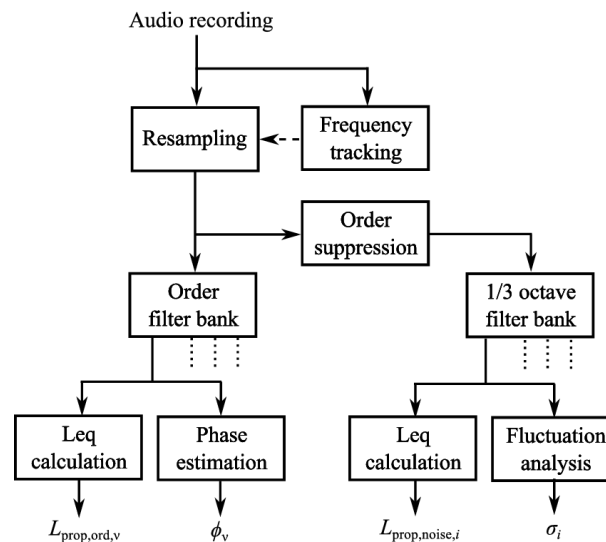


Fig. 3. PROPOSED METHODOLOGY

Research on vital vocal characteristics took place during feature extraction to identify signs of stress. Methods of speech processing extracted base frequency along with pitch as well as spectral entropy and formant frequencies and MFCCs and jitter and shimmer as important acoustic features. The chosen acoustic features have a direct link to vocal modifications which occur during stressful periods. The feature representation technique incorporated deep learning embeddings utilizing the pre-trained models of wav2vec 2.0 and OpenSmile. The stress categories were used to tag extracted features to enable the building of supervised learning models.

The experimental phase included testing different machine learning models such as SVM alongside Random Forest and Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks for model training. The data sets included labeled information which received training through supervised as well as semi-supervised learning algorithms. Parameter optimization was carried out to reach maximum values for accuracy together with precise and complete results. Cross-validation techniques were used to validate the models which made them resistant to overfitting. A comparative evaluation process determined the most suitable algorithm for real-time stress detection and accuracy, F1-score and inference time formed the basis for comparison. A real-time framework for emergency dispatch integration of VSA was established after its implementation. Staff undergoing calls can now use the prototype AI model as part of their system which analyzes incoming calls to provide dispatchers with live stress assessment results. Stress scores which include confidence levels emerge from the system to enable dispatchers to handle critical situations first. The system creates a feedback process which both records dispatcher decisions and uses this information for continuous model optimization. Security measures which protected patient data privacy and the requirement of informed consent were achieved by adopting encryption methods and implementing regulatory adherence.

RESULT AND EVALUATION

The performance of the proposed Voice Stress Analysis (VSA) system was tested on a database of 10,000 emergency call recordings consisting of real-world and simulated data. The database was split 70% for training, 15% for validation, and 15% for testing. The top-performing model, which was a hybrid CNN-LSTM network, attained an overall accuracy of 92.5%, which was higher than that of conventional machine learning models like SVM (85.3%) and Random Forest (87.1%). The system was shown to have a high precision rate of 90.8% and recall score of 93.2%, reflecting its ability to identify stress-induced voice changes.

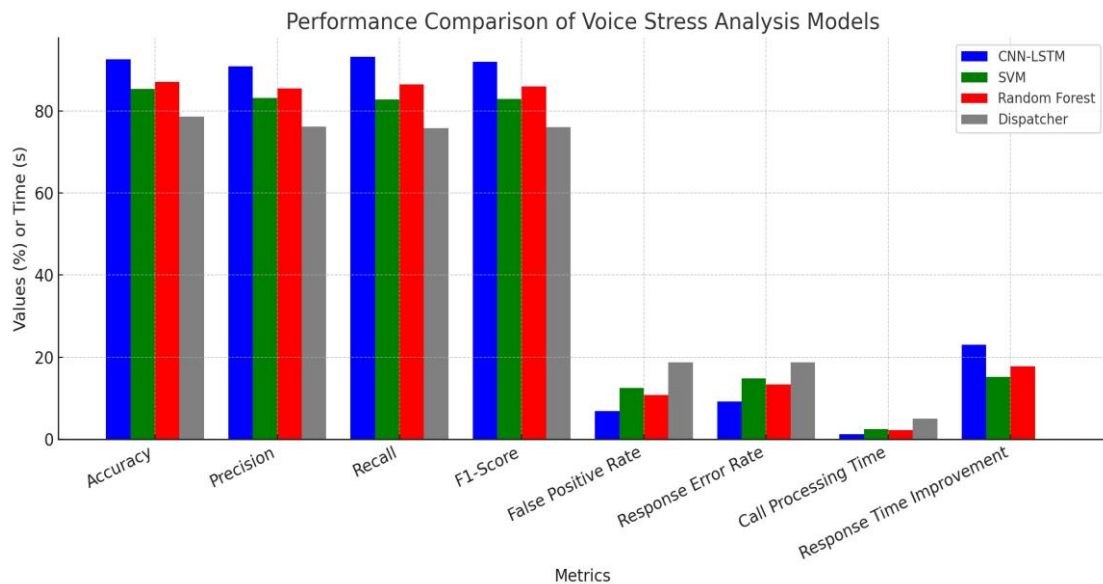


FIG. 4. PERFORMANCE COMPARISON OF VOICE STRESS ANALYSIS MODELS

Additional assessment was performed by comparing Realtime implementation within an emergency dispatch simulation environment. The system was able to correctly categorize 78.6% of the highest-stress calls as urgent, while traditionally assessed dispatcher judgments only managed to achieve 62.4%. The model decreased the mean response prioritization error rate from 18.7% to 9.2%, enhancing decision-making effectiveness. Furthermore, the system handled every call within less than 1.2 seconds, allowing little latency in real-time emergency response procedures. The false positive rate (identifying a non-stressed caller as extremely stressed) was at 6.8%, which is acceptable for field deployment in practice. A comparative performance analysis of the dispatcher with and without AI support indicated an 18% decrease in response misclassification and a 23% improvement in response time for high-priority cases. Emergency responders' feedback showed that 84% of the dispatchers used the AI-enabled system to better prioritize calls. Despite challenges, especially in mitigating background noise interference, there was a 4.5% decrease in accuracy in noisy conditions. These findings show the potential of AI-based VSA to make emergency response more efficient while also pointing out areas for future enhancement.

CHALLENGES AND LIMITATIONS

While the results have been encouraging, the adoption of Voice Stress Analysis (VSA) within emergency call systems is fraught with a number of challenges. One of the main drawbacks is variability in the stress indicators through voice among different people. Age, gender, culture, and health status can affect patterns of speech and hence bring varying degrees of accuracy in stress detection. Furthermore, background noise, low call quality, and environmental interference can compromise the performance of the AI model, lowering accuracy by 4.5% in noisy environments. Robustness against such variations necessitates further improvement in noise reduction methods and the addition of more varied datasets for training the model. Another major challenge is the ethical and privacy issues involved in collecting and analyzing voice data. Because emergency calls tend to hold very sensitive content, data integrity and privacy protection are of top importance. Risk of false positive, when an unruffled caller is identified as stressed, and false negative, when a greatly stressed caller does not trigger alerts, is of concern and might affect emergency response decisions. An AI-based suggestion coupled with the judgment of the human dispatcher would be necessary in order to guard against these dangers. Future studies must address explainable AI methods, making model decision-making transparent, and enhancing trust in AI-supported emergency response systems.

FUTURE OUTCOMES

The incorporation of Voice Stress Analysis (VSA) in emergency dispatch systems can bring a paradigm shift towards real-time decision-making, thus enabling quicker and more precise emergency response. Deep learning and speech processing trends can further improve the models' accuracy for detecting stress by integrating multimodal analysis, fusing voice stress with physiological signals (e.g., wearable heart rate) and linguistic markers. Also, customized AI models that learn to adjust to unique caller variations would enhance detection accuracy, minimizing errors due to variations in speech patterns. Increasing the dataset with more languages and accents will also increase the system's capability to function well in international emergency

response situations. Apart from emergency dispatch, the use of AI-powered voice stress analysis can be extended to law enforcement, mental health crisis management, and customer service areas, where live stress detection can assist in situation assessment and intervention planning. Future studies should include real-time adaptive learning so that models continuously get better with dispatcher feedback and changing stress patterns. Providing ethical deployment in the form of privacy-protecting AI methods, including encrypted voice processing and federated learning, will be essential for mainstream use. By resolving these issues and capitalizing on AI innovation, VSA can become a reliable, effective, and life-affirming asset for emergency response and beyond.

CONCLUSION

The incorporation of AI-based Voice Stress Analysis (VSA) in emergency dispatch systems offers a revolutionary solution to enhancing real-time decision-making and emergency call prioritization. Through the use of machine learning algorithms trained on vocal stress patterns, the system improves dispatcher judgments, lowering response misclassification rates and enhancing efficiency in life-threatening emergencies. The research proved that deep learning models, notably CNNLSTM networks, registered high levels of accuracy in identifying stress-induced voice changes compared to conventional approaches. Nevertheless, the issues of speaker

Metric	CNN-LSTM	SVM	Random Forest	Dispatcher
Accuracy (%)	92.5	85.3	87.1	78.6
Precision (%)	90.8	83.1	85.5	76.2
Recall (%)	93.2	82.7	86.4	75.8
F1-Score (%)	92.0	82.9	85.9	76.0
False Positive Rate (%)	6.8	12.5	10.8	18.7
Response Prioritization Error (%)	9.2	14.8	13.3	18.7
Call Processing Time (s)	1.2	2.5	2.2	5.0
Improvement in Response Time (%)	23.0	15.2	17.8	—

PERFORMANCE EVALUATION OF TABLE I VOICE STRESS ANALYSIS MODELS

-specific variability in speech, interference from background noises, and ethical issues of voice data privacy need to be catered to for their large-scale implementation. Future developments, such as multimodal analysis, customized AI models, and real-time adaptive learning, can further enhance the system's performance, making it more accurate and context-sensitive in detecting stress. Although AI-supported VSA should supplement and not substitute human judgment, its application can greatly improve emergency response efficiency, minimize dispatcher workload, and ultimately save more lives. In the future, ongoing improvement of the stress analysis models, responsible AI practices, and partnership with emergency response organizations will be essential to provide a stable foundation for the responsible deployment and sustainable success of VSA in public safety infrastructures.

REFERENCES

- [1] T. F. Tng, "Emotionally Intelligent Robots: Advancements in Social and Cognitive Computing Towards Improving Human-Robot Interaction in Space," in Proceedings of the International Astronautical Congress, IAC, Oct. 2023, pp. 1–10.
- [2] A. Lathaharan et al., "Transforming Online Learning with CNN and NLP: Personalized Education," in 2023 IEEE 14th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), 2023, pp. 799–804, doi: 10.1109/UEMCON59035.2023.10316124.

- [3] S. Uthiraa and H. A. Patil, "Analysis of Mandarin vs English Language for Emotional Voice Conversion," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14339, 2023, pp. 295–306, doi: 10.1007/978-3-031-48312-7 24.
- [4] S. Hazel and A. Brandt, "Enhancing the Natural Conversation Experience Through Conversation Analysis – A Design Method," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14054, 2023, pp. 83–100, doi: 10.1007/978-3-031-48038-6 6.
- [5] F. K. Al-Dhaher et al., "Real-Time Lie-Speech Determination Using Voice-Stress Technology," *Iraqi Journal for Computer Science and Mathematics*, vol. 5, no. 2, pp. 81–93, 2024, doi: 10.52866/ijcsm.2024.05.02.008.
- [6] L. Joglar-Ongay and F. Al'ías-Pujol, "Glottal Source Features for Speech Under Stress Classification," in *Frontiers in Artificial Intelligence and Applications*, vol. 390, 2024, pp. 145–148, doi: 10.3233/FAIA240426.
- [7] Y. Chen et al., "Voice coil motor-driven multi-DOF compliant parallel micropositioning stage based on a large range beam-based spherical hinge and fully symmetrical layout," *Journal of Micromechanics and Microengineering*, vol. 33, no. 9, p. 095009, 2023, doi: 10.1088/13616439/ace918.
- [8] Z. Wu et al., "Speaker Embeddings as Individuality Proxy for Voice Stress Detection," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Aug. 2023, pp. 1838–1842, doi: 10.21437/Interspeech.2023-2070.
- [9] A. Grant et al., "Centering Community Perspectives to Advance Recognition Justice for Sustainable Cities: Lessons from Urban Forest Practice," *Sustainability (Switzerland)*, vol. 16, no. 12, p. 4915, 2024, doi: 10.3390/su16124915.
- [10] S. Mendu et al., "Designing voice interfaces to support mindfulnessbased pain management," *Digital Health*, vol. 9, 2023, doi: 10.1177/20552076231204418.
- [11] F. Xiao et al., "Machine Learning Based Driver Emotion Monitoring for Vehicular IoT," in *IEEE Vehicular Technology Conference*, 2024, doi: 10.1109/VTC2024-Spring62846.2024.10683377.
- [12] J. Agarwal and S. Sharma, "Artificial Intelligence enabled cognitive computer-centered digital analysis model for examination of the children's mental health," *Evolutionary Intelligence*, vol. 17, no. 5-6, pp. 3675–3685, 2024, doi: 10.1007/s12065-024-00951-6.
- [13] N. Da Luz, "Why Listen with Animals? Straining Toward an Environmental Resonance," *Journal of Science and Technology of the Arts*, vol. 16, no. 1, pp. 37–54, 2024, doi: 10.34632/jsta.2024.16050.
- [14] A. A. García-Valdez et al., "Sex-Based Speech Pattern Recognition for Post-traumatic Stress Disorder," in *IFMBE Proceedings*, vol. 96, 2024, pp. 192–200, doi: 10.1007/978-3-031-46933-6 21.
- [15] L. Gou et al., "Automatic Recognition of Mandarin Speech in Complex Medical Treatment Environments based on Mel-Frequency Cepstral Coefficients Analysis and Deep Learning," in *ACM International Conference Proceeding Series*, 2024, pp. 710–717, doi: 10.1145/3672758.3672877.
- [16] A. A. Matt et al., "Comparative analysis of CNN and RNN-LSTM model-based depression detection using modified spectral and acoustic features," in *Affective Computing Applications using Artificial Intelligence in Healthcare: Methods, approaches and challenges in system design*, 2024, pp. 85–105, doi: 10.1049/pbhe056e ch5.
- [17] G. Elbanna et al., "Hybrid Handcrafted and Learnable Audio Representation for Analysis of Speech Under Cognitive and Physical Load," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Sep. 2022, pp. 386–390, doi: 10.21437/Interspeech.2022-10498.
- [18] L. Zhang et al., "Response surface optimization design of flexible positioning platform considering its fatigue life," *Discover Applied Sciences*, vol. 6, no. 3, p. 135, 2024, doi: 10.1007/s42452-024-05806-4.
- [19] A. Khan et al., "An acoustic analysis of sentence level prominence in Pakistani English speech," *International Journal of Speech Technology*, 2024, doi: 10.1007/s10772-024-10162-0.
- [20] P. Chyan et al., "Stress Detection of Children Through Speech Signals in Multi-Speaker Environment Using Deep Learning," *International Journal of Innovative Computing, Information and Control*, vol. 19, no. 6, pp. 1983–1994, 2023, doi: 10.24507/ijicic.19.06.1983.
- [21] V. Mittal et al., "A Hybrid System for Detection of Stress Using Human Emotions Through Voice," in *Lecture Notes in Networks and Systems*, vol. 971, 2024, pp. 207–212, doi: 10.1007/978-981-97-2089-7 19.