18th International Conference on Development and Application Systems (DAS)
21-23 May 2026, Suceava, Romania, ISBN 979-8-3315-8387-3

Real-time speech emotion recognition using acoustic features and a probabilistic heuristic classifier

Olesea BOROZAN, Silvia MUNTEANU, Nicu DRUMEA, Andrei CLIMA

https://doi.org/10.1109/DAS69882.2026.11553338

Abstract

This paper presents a real-time speech emotion recognition system based on acoustic feature extraction and a probabilistic-heuristic classifier, intended for intelligent monitoring and the management of exceptional situations. The proposed model formalizes the process of voice signal analysis by segmenting the signal into short frames, extracting a vector of relevant acoustic features, and combining probabilistic inference with heuristic rules to determine the emotional class. The system was implemented on a resource-constrained embedded platform using the ESP32 microcontroller and the INMP441 digital microphone, enabling near-real-time voice signal acquisition and processing, as well as communication via Wi-Fi and Bluetooth. Experimental validation included 50 tests based on short utterances spoken under different emotional states. The system achieved recognition rates of 87% for anger, 92% for fear, 48% for sadness, 94% for the neutral emotional state, and 75% for panic. These results confirm the functionality and feasibility of the proposed solution, as well as its potential for integration into intelligent systems for the early detection of relevant affective states in critical contexts.

Keywords: *affective computing; human-computer interaction*

References:

1. J. Liu, Z. Liu, L. Wang, L. Guo and J. Dang, "Speech Emotion Recognition with Local-Global Aware Deep Representation Learning," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 7174–7178, DOI: 10.1109/ICASSP40776.2020.9053192.
Google Scholar

2. A. Nediyanchath, P. Paramasivam and P. Yenigalla, "Multi-Head Attention for Speech Emotion Recognition with Auxiliary Learning of Gender Recognition," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 7179–7183, DOI:

10.1109/ICASSP40776.2020.9054073.
Google Scholar

3. F. Andayani, B. T. Lau, M. K. T. Tee and C. Chua, “Hybrid LSTM-Transformer Model for Emotion Recognition from Speech Audio Files,” *IEEE Access*, vol. 10, pp. 36018–36027, 2022, DOI: 10.1109/ACCESS.2022.3163856.
Google Scholar

4. E. Morais, R. Hoory, W. Zhu, I. Gat, M. Damasceno and H. Aronowitz, “Speech Emotion Recognition Using Self-Supervised Features,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, 2022, pp. 6922–6926, DOI: 10.1109/ICASSP43922.2022.9747870.
Google Scholar

5. A. A. Abdelhamid, E. S. M. El-Kenawy, B. Alotaibi, G. M. Amer, M. Y. Abdelkader, A. Ibrahim and M. M. Eid, “Robust Speech Emotion Recognition Using CNN+LSTM Based on Stochastic Fractal Search Optimization Algorithm,” *IEEE Access*, vol. 10, pp. 49265–49284, 2022, DOI: 10.1109/ACCESS.2022.3172954.
Google Scholar

6. B. T. Atmaja and A. Sasou, “Effects of Data Augmentations on Speech Emotion Recognition,” *Sensors*, vol. 22, no. 16, Art. no. 5941, 2022, DOI: 10.3390/s22165941.
CrossRef Google Scholar

7. A. Hashem, M. Arif and M. Alghamdi, “Speech Emotion Recognition Approaches: A Systematic Review,” *Speech Communication*, vol. 154, Art. no. 102974, 2023, DOI: 10.1016/j.specom.2023.102974.
CrossRef Google Scholar

8. D. Thiripurasundari, K. Bhangale, V. Aashritha, S. Mondreti and M. Kothandaraman, “Speech Emotion Recognition for Human-Computer Interaction,” *International Journal of Speech Technology*, vol. 27, pp. 817–830, 2024, DOI: 10.1007/s10772-024-10138-0.
CrossRef Google Scholar

9. Y. Wang, J. Huang, Z. Zhao, H. Lan and X. Zhang, “Speech Emotion Recognition Using Multi-Scale Global–Local Representation Learning with Feature Pyramid Network,” *Applied Sciences*, vol. 14, no. 24, Art. no. 11494, 2024, DOI: 10.3390/app142411494.
CrossRef Google Scholar

10. X. Qi, Q. Song, G. Chen, P. Zhang and Y. Fu, “Acoustic Feature Excitation-and-Aggregation Network Based on Multi-Task Learning for Speech Emotion Recognition,” *Electronics*, vol. 14, no. 5, Art. no. 844, 2025, DOI: 10.3390/electronics14050844.
CrossRef Google Scholar

11. D. O’Shaughnessy, “Review of Automatic Estimation of Emotions in Speech,” *Applied Sciences*, vol. 15, no. 10, Art. no. 5731, 2025, DOI: 10.3390/app15105731.
CrossRef Google Scholar

12. C. Barhoumi and Y. BenAyed, “Transformer Encoder and Data Augmentation for Real-Time Speech Emotion Recognition,” *Multimedia Tools and Applications*, vol. 85, Art. no. 226, 2026, DOI: 10.1007/s11042-026-21395-3.

CrossRef Google Scholar

13. S. M. George and P. Muhamed Ilyas, "A review on speech emotion recognition: A survey, recent advances, challenges, and the influence of noise," *Neurocomputing*, vol. 568, pp. 1–23, 2024, DOI: 10.1016/j.neucom.2023.127015.
CrossRef Google Scholar

14. S. R. Gudivaka, P. Polipogu, R. Kodali, et al. "Speech emotion recognition in adults and children: a comprehensive review of traditional features and raw waveform models," *Int J Speech Technol* 29, 21 (2026), DOI: 10.1007/s10772-025-10229-6.
CrossRef Google Scholar

15. X. Qi, Y. Zhang, L. Wang, and J. Chen, "Acoustic Feature Excitation-and-Aggregation Network for Speech Emotion Recognition," *Electronics*, vol. 14, no. 5, 2025, 844, DOI: 10.3390/electronics14050844.
CrossRef Google Scholar

16. S.-W. Byun and S.-P. Lee, "A study on a speech emotion recognition system with effective acoustic features using deep learning algorithms," *Applied Sciences*, vol. 11, no. 4, 2021, 1890, DOI: 10.3390/app11041890.
CrossRef Google Scholar

17. S. Mishra, P. Warule, and S. Deb, "Fixed frequency range empirical wavelet transform based acoustic and entropy features for speech emotion recognition," *Speech Communication*, vol. 166, 2025, 103148, DOI: 10.1016/j.specom.2024.103148.
CrossRef Google Scholar