Scopus

Documents

Arifin, F.[a] , Nasuha, A.[a] , Priambodo, A.S.[a] , Winursito, A.[a] , Gunawan, T.S.[b]

Advanced Multimodal Emotion Recognition for Javanese Language Using Deep Learning
(2024) *Indonesian Journal of Electrical Engineering and Informatics*, 12 (3), pp. 503-515.

[a] Department of Electronics and Informatics Engineering of Education, Universitas Negeri Yogyakarta, Indonesia
[b] Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia

Abstract
This research develops a robust emotion recognition system for the Javanese language using multimodal audio and video datasets, addressing the limited advancements in emotion recognition specific to this language. Three models were explored to enhance emotional feature extraction: the Spectrogram-Image Model (Model 1), which converts audio inputs into spectrogram images and integrates them with facial images for emotion labeling; the Convolutional-MFCC Model (Model 2), which leverages convolutional techniques for image processing and Mel-frequency cepstral coefficients for audio; and the Multimodal Feature-Extraction Model (Model 3), which independently processes video and audio features before integrating them for emotion recognition. Comparative analysis shows that the Multimodal Feature-Extraction Model achieves the highest accuracy of 93%, surpassing the Convolutional-MFCC Model at 85% and the Spectrogram-Image Model at 71%. These findings demonstrate that effective multimodal integration, mainly through separate feature extraction, significantly enhances emotion recognition accuracy. This research improves communication systems and offers deeper insights into Javanese emotional expressions, with potential applications in human-computer interaction, healthcare, and cultural studies. Additionally, it contributes to the advancement of sophisticated emotion recognition technologies. © 2024 Institute of Advanced Engineering and Science. All rights reserved.

Author Keywords
Audio-Visual Integration; Cultural Emotion Analysis; Emotion Detection Models; Human-Computer Interaction; Javanese Emotion Recognition; Multimodal Deep Learning

References

- Ahmed, N., Al Aghbari, Z., Girija, S.
  A systematic survey on multimodal emotion recognition using learning algorithms
  (2023) *Intelligent Systems with Applications*, 17, p. 200171.

- Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E., Mihalcea, R.
  (2018) *MELD: A multimodal multi-party dataset for emotion recognition in conversations*, arXiv preprint arXiv:1810.02508

- Ashraf, A., Gunawan, T. S., Arifin, F., Kartiwi, M., Sophian, A., Habaebi, M. H.
  On the Audio-Visual Emotion Recognition using Convolutional Neural Networks and Extreme Learning Machine
  (2022) *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, 10 (3), pp. 684-697.

- Ashraf, A., Gunawan, T. S., Arifin, F., Kartiwi, M., Sophian, A., Habaebi, M. H.
  Enhanced Emotion Recognition in Videos: A Convolutional Neural Network Strategy for Human Facial Expression Detection and Classification
  (2023) *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, 11 (1), pp. 286-299.

- Wani, T. M., Gunawan, T. S., Qadri, S. A. A., Kartiwi, M., Ambikairajah, E.
  A comprehensive review of speech emotion recognition systems
  (2021) *IEEE Access*, 9, pp. 47795-47814.

- Jamiluddin, A. S., Udja, S. K., Safithri, R.
  Meaning and Message of Communication Behaviour of Javanese Ethnic Traders to Prospective Buyers

(2022) *International Conference on Halal, Policy, Culture and Sustainability Issues*, 4 (1), p. 19.

- Wijonarko, P., Zahra, A.
  Spoken language identification on 4 Indonesian local languages using deep learning
  (2022) *Bulletin of Electrical Engineering and Informatics*, 11 (6), pp. 3288-3293.

- Sulistyo, E. T.
  Emotional Intelligence And Balanced Personality In Javanese Cultural Understanding
  (2021) *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18 (4), pp. 3344-3359.

- Kumala, S. A.
  Analysis of Language Attitude and Language Preservation in Javanese Language.: A Case Study of Javanese Speaker in Madiun, East Java
  (2021) *e-LinguaTera*, 1 (1), pp. 11-19.

- Kresna, A. A.
  The Epistemology of Rasa as a Basic Foundation of the Javanese Psychology
  (2023) *East Asian Journal of Multidisciplinary Research*, 2 (8), pp. 3209-3222.

- Yunanto, T. A. R.
  Happiness in the Javanese context: Exploring the role of emotion regulation and resilience
  (2023) *Humanitas: Indonesian Psychological Journal*, pp. 149-158.

- Khurana, Y., Gupta, S., Sathyaraj, R., Raja, S.
  RobinNet: A Multimodal Speech Emotion Recognition System With Speaker Recognition for Social Interactions
  (2022) *IEEE Transactions on Computational Social Systems*,

- Wen, G., Ye, S., Li, H., Wen, P., Zhang, Y.
  Multimodal and Multitask Learning with Additive Angular Penalty Focus Loss for Speech Emotion Recognition
  (2023) *International Journal of Intelligent Systems*, 2023 (1), p. 3662839.

- Patamia, R. A., Santos, P. E., Acheampong, K. N., Ekong, F., Sarpong, K., Kun, S.
  Multimodal Speech Emotion Recognition Using Modality-Specific Self-Supervised Frameworks
  (2023) *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 4134-4141.
  IEEE

- Dong, G.-N., Pun, C.-M., Zhang, Z.
  Temporal relation inference network for multimodal speech emotion recognition
  (2022) *IEEE Transactions on Circuits and Systems for Video Technology*, 32 (9), pp. 6472-6485.

- Arifin, F., Priambodo, A. S., Nasuha, A., Winursito, A., Gunawan, T. S.
  Development of Javanese Speech Emotion Database (Java-SED)
  (2022) *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, 10 (3), pp. 584-591.

- Ahmed, T., Begum, I., Mia, M. S., Tasnim, W.
  Multimodal Speech Emotion Recognition Using Deep Learning and the Impact of Data Balancing
  (2023) *2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI)*, pp. 1-6.
  IEEE

- Lian, H., Lu, C., Li, S., Zhao, Y., Tang, C., Zong, Y.
  A survey of deep learning-based multimodal emotion recognition: Speech, text, and face
  (2023) *Entropy*, 25 (10), p. 1440.

- Abdullah, S. M. S. A., Ameen, S. Y. A., Sadeeq, M. A., Zeebaree, S.
  Multimodal emotion recognition using deep learning
  (2021) *Journal of Applied Science and Technology Trends*, 2, pp. 73-79.
  01

- Nugroho, K., Noersasongko, E., Santoso, H. A.
  Javanese gender speech recognition using deep learning and singular value decomposition
  (2019) *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, pp. 251-254.
  IEEE

Correspondence Address
Arifin F.; Department of Electronic and Informatics Engineering, Indonesia; email: fatchul@uny.ac.id