

DAFTAR PUSTAKA

- [1] Z. Almutairi and H. Elgibreen, "A Review of Modern Audio Deepfake Detection Methods: Challenges and Future Directions," *Algorithms*, vol. 15, no. 5, p. 155, May 2022, doi: 10.3390/a15050155.
- [2] H. Prayoga and H. Tuasikal, "Penyebaran Konten Deepfake Sebagai Tindak Pidana: Analisis Kritis Terhadap Penegakan Hukum Dan Perlindungan Publik Di Indonesia," *Abdurrauf Law Sharia*, vol. 2, no. 1, pp. 22–38, May 2025, doi: 10.70742/arlash.v2i1.194.
- [3] U. Zbezhkhovska and O. Khapilin, "Exploring Challenges and Future Paths in Deepfake Audio Detection".
- [4] A. Hery, O. Joseph, O. Femi, and H. Luz, "Audio deepfakes: threats to voice assistants and voice-activated systems".
- [5] A. Martínez-Serrano, C. Montero-Ramírez, and C. Peláez-Moreno, "Authenticity at Risk: Key Factors in the Generation and Detection of Audio Deepfakes," *Appl. Sci.*, vol. 15, no. 2, p. 558, Jan. 2025, doi: 10.3390/app15020558.
- [6] J. Yi, C. Wang, J. Tao, X. Zhang, C. Y. Zhang, and Y. Zhao, "Audio Deepfake Detection: A Survey," Aug. 29, 2023, *arXiv*: arXiv:2308.14970. doi: 10.48550/arXiv.2308.14970.
- [7] N. M. Müller, P. Czempin, F. Dieckmann, A. Froghyar, and K. Böttinger, "Does Audio Deepfake Detection Generalize?," Aug. 27, 2024, *arXiv*: arXiv:2203.16263. doi: 10.48550/arXiv.2203.16263.
- [8] K. Shahzad, S. Farhan, Yasin-Ul-Haq, R. Sana, and S. Pathan, "Enhancing Voice Spoofing Detection: A Hybrid Approach With VGGish-LSTM Model for Improved Security in Automatic Speaker Verification Systems," *IEEE Access*, vol. 13, pp. 40682–40702, 2025, doi: 10.1109/ACCESS.2025.3544639.
- [9] Y. Sims, A. Mendes, and S. Chalup, "Enhanced Embeddings in Zero-Shot Learning for Environmental Audio," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece: IEEE, Jun. 2023, pp. 1–5. doi: 10.1109/ICASSP49357.2023.10096134.
- [10] O. A. Shaaban, R. Yildirim, and A. A. Alguttar, "Audio Deepfake Approaches," *IEEE Access*, vol. 11, pp. 132652–132682, 2023, doi: 10.1109/ACCESS.2023.3333866.
- [11] C. Li, "Conversation Understanding and Realistic Artificial Crash Data Generation With Deep Learning," University of Missouri-Columbia, 2023.
- [12] S. Chakravarty, R. R. Khandelwal, and K. M. Dhote, "Feature Extraction Techniques for Deep Learning based Speech Classification," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India: IEEE, Jul. 2023, pp. 1–6. doi: 10.1109/ICCCNT56998.2023.10307237.
- [13] H. Widyagustin, H. Kusuma, and T. A. Sardjono, "Deep Learning-Based Classification of Lung Sound Using YAMNet for Identifying Lung Diseases," in *2024 2nd International Symposium on Information Technology and Digital*

- Innovation (ISITDI)*, Bukittinggi, Indonesia: IEEE, Jul. 2024, pp. 131–135. doi: 10.1109/ISITDI62380.2024.10796009.
- [14] R. K. Bhukya, A. Raj, and D. N. Raja, “Audio Deepfakes: Feature Extraction and Model Evaluation for Detection,” in *2024 5th International Conference for Emerging Technology (INCET)*, Belgaum, India: IEEE, May 2024, pp. 1–6. doi: 10.1109/INCET61516.2024.10593405.
 - [15] Z. Khanjani, G. Watson, and V. P. Janeja, “Audio deepfakes: A survey,” *Front. Big Data*, vol. 5, p. 1001063, Jan. 2023, doi: 10.3389/fdata.2022.1001063.
 - [16] L. H. Singh, P. Charanarur, and N. K. Chaudhary, “Advancements in detecting Deepfakes: AI algorithms and future prospects – a review,” *Discov. Internet Things*, vol. 5, no. 1, p. 53, May 2025, doi: 10.1007/s43926-025-00154-0.
 - [17] B. Zhang, H. Cui, V. Nguyen, and M. Whitty, “Audio Deepfake Detection: What Has Been Achieved and What Lies Ahead,” *Sensors*, vol. 25, no. 7, p. 1989, Mar. 2025, doi: 10.3390/s25071989.
 - [18] C. Constantinescu and R. Brad, “An Overview on Sound Features in Time and Frequency Domain,” *Int. J. Adv. Stat. ITC Econ. Life Sci.*, vol. 13, no. 1, pp. 45–58, Dec. 2023, doi: 10.2478/ijasitels-2023-0006.
 - [19] A. Mukhamediya, S. Fazli, and A. Zollanvari, “On the Effect of Log-Mel Spectrogram Parameter Tuning for Deep Learning-Based Speech Emotion Recognition,” *IEEE Access*, vol. 11, pp. 61950–61957, 2023, doi: 10.1109/ACCESS.2023.3287093.
 - [20] K. L. Ong, C. P. Lee, H. S. Lim, K. M. Lim, and A. Alqahtani, “Mel-MViTv2: Enhanced Speech Emotion Recognition With Mel Spectrogram and Improved Multiscale Vision Transformers,” *IEEE Access*, vol. 11, pp. 108571–108579, 2023, doi: 10.1109/ACCESS.2023.3321122.
 - [21] K. Sharifani and M. Amini, “Machine Learning and Deep Learning: A Review of Methods and Applications,” vol. 10, no. 07, 2023.
 - [22] J. Schneider and M. Vlachos, “A Survey of Deep Learning: From Activations to Transformers,” Feb. 10, 2024, *arXiv*: arXiv:2302.00722. doi: 10.48550/arXiv.2302.00722.
 - [23] P. Purwono, A. Ma’arif, W. Rahmانيar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. U. Haq, “Understanding of Convolutional Neural Network (CNN): A Review,” *Int. J. Robot. Control Syst.*, vol. 2, no. 4, pp. 739–748, Jan. 2023, doi: 10.31763/ijrcs.v2i4.888.
 - [24] G. Bompem and D. Pandluri, “Batch Normalization Based Convolutional Neural Network for Segmentation and Classification of Brain Tumor MRI Images,” *Int. J. Intell. Eng. Syst.*, vol. 17, no. 2, pp. 39–49, Apr. 2024, doi: 10.22266/ijies2024.0430.04.
 - [25] M. Helal, “Enhancing Deep Learning Models through Tensorization: A Comprehensive Survey and Framework”.
 - [26] I. Salehin and D.-K. Kang, “A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain,” *Electronics*, vol. 12, no. 14, p. 3106, Jul. 2023, doi: 10.3390/electronics12143106.
 - [27] M. Krichen, “Convolutional Neural Networks: A Survey,” *Computers*, vol. 12, no. 8, p. 151, Jul. 2023, doi: 10.3390/computers12080151.

- [28] A. S. Musallam, A. S. Sherif, and M. K. Hussein, "A New Convolutional Neural Network Architecture for Automatic Detection of Brain Tumors in Magnetic Resonance Imaging Images," *IEEE Access*, vol. 10, pp. 2775–2782, 2022, doi: 10.1109/ACCESS.2022.3140289.
- [29] O. Elharrouss *et al.*, "Task-based Loss Functions in Computer Vision: A Comprehensive Review," Sep. 09, 2025, *arXiv*: arXiv:2504.04242. doi: 10.48550/arXiv.2504.04242.
- [30] X. Liu, H. Qi, S. Jia, Y. Guo, and Y. Liu, "Recent Advances in Optimization Methods for Machine Learning: A Systematic Review," *Mathematics*, vol. 13, no. 13, p. 2210, Jul. 2025, doi: 10.3390/math13132210.
- [31] A. De Carlo, D. Ronchi, M. Piastra, E. M. Tosca, and P. Magni, "Predicting ADMET Properties from Molecule SMILE: A Bottom-Up Approach Using Attention-Based Graph Neural Networks," *Pharmaceutics*, vol. 16, no. 6, p. 776, Jun. 2024, doi: 10.3390/pharmaceutics16060776.
- [32] L. Yang, "Theoretical Analysis of Adam Optimizer in the Presence of Gradient Skewness," *Int. J. Appl. Sci.*, vol. 7, no. 2, p. p27, Oct. 2024, doi: 10.30560/ijas.v7n2p27.
- [33] Z. Liu *et al.*, "Optimizing depthwise separable convolution on DCU," *CCF Trans. High Perform. Comput.*, vol. 6, no. 6, pp. 646–664, Dec. 2024, doi: 10.1007/s42514-024-00200-3.
- [34] F. Daghero, A. Burrello, M. Poncino, E. Macii, and D. J. Pagliari, "Accelerating Depthwise Separable Convolutions on Ultra-Low-Power Devices," Jun. 18, 2024, *arXiv*: arXiv:2406.12478. doi: 10.48550/arXiv.2406.12478.
- [35] K. Altwlkany, S. Delalic, E. Selmanovic, A. Alihodzic, and I. Lovric, "A Recurrent Neural Network Approach to the Answering Machine Detection Problem," in *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*, May 2024, pp. 85–90. doi: 10.1109/MIPRO60963.2024.10569812.
- [36] A. Huerta Herraiz, O. Ayo-Martin, J. J Rieta, and R. Alcaraz, "Assessing ECG Signal Quality Using a Pre-trained Audio Network," presented at the 2024 Computing in Cardiology Conference, Dec. 2024. doi: 10.22489/CinC.2024.449.
- [37] S. Lachenani, H. Kheddar, and M. Ouldzmirli, "Improving Pretrained YAMNet for Enhanced Speech Command Detection via Transfer Learning," in *2024 International Conference on Telecommunications and Intelligent Systems (ICTIS)*, Dec. 2024, pp. 1–6. doi: 10.1109/ICTIS62692.2024.10894266.
- [38] A. Ameh Joseph, M. Abdullahi, S. B. Junaidu, H. Hassan Ibrahim, and H. Chiroma, "Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer)," *Intell. Syst. Appl.*, vol. 14, p. 200066, May 2022, doi: 10.1016/j.iswa.2022.200066.
- [39] L. D. Yulianto and S. Desmana, "Binary Classification of Academic Outcomes Using Ensemble Learning and Neural Networks: A Case Study on OULAD," vol. 15, no. 01, 2025.

- [40] M. Fahmy Amin, "Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial," *J. Eng. Res.*, vol. 6, no. 5, pp. 0–0, Dec. 2022, doi: 10.21608/erjeng.2022.274526.
- [41] Jude Chukwura Obi, "A comparative study of several classification metrics and their performances on data," *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 1, pp. 308–314, Feb. 2023, doi: 10.30574/wjaets.2023.8.1.0054.
- [42] C. A. N. Fadlilah, S. Surorejo, and Z. Arif, "Klasifikasi Jenis Buah Apel Menggunakan Algoritma Convolutional Neural Network," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 5201–5208, Sep. 2025, doi: 10.31004/riggs.v4i3.2759.
- [43] H. Dani, P. Bhople, H. Waghmare, K. Munginwar, and A. Patil, "Review on Frameworks Used for Deployment of Machine Learning Model," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 2, pp. 211–215, Feb. 2022, doi: 10.22214/ijraset.2022.40222.
- [44] S. Liao *et al.*, "Fish-Speech: Leveraging Large Language Models for Advanced Multilingual Text-to-Speech Synthesis," Nov. 09, 2024, *arXiv*: arXiv:2411.01156. doi: 10.48550/arXiv.2411.01156.
- [45] tensorflow, "models/research/audioset/yamnet at master · tensorflow/models," GitHub. Accessed: Nov. 01, 2025. [Online]. Available: <https://github.com/tensorflow/models/tree/master/research/audioset/yamnet>
- [46] J. Sadaiyandi, P. Arumugam, A. K. Sangaiah, and C. Zhang, "Stratified Sampling-Based Deep Learning Approach to Increase Prediction Accuracy of Unbalanced Dataset," *Electronics*, vol. 12, no. 21, p. 4423, Oct. 2023, doi: 10.3390/electronics12214423.