Daftar Pustaka

[1]   J. Wilkins, P. Seetharaman, A. Wahl, and B. Pardo, "Vocalset: A singing voice dataset," *Proc. 19th Int. Soc. Music Inf. Retr. Conf. ISMIR 2018*, pp. 468–474, 2018.

[2]   W. Wang, C. Liao, Q. Cheng, and P. Wang, "CONVOLUTIONAL RECURRENT NEURAL NETWORKS FOR MUSIC CLASSIFICATION," *ACM Int. Conf. Proceeding Ser.*, pp. 1–6, 2017, doi: 10.1145/3371425.3371430.

[3]   S. Kum and J. Nam, "Joint detection and classification of singing voice melody using convolutional recurrent neural networks," *Appl. Sci.*, vol. 9, no. 7, 2019, doi: 10.3390/app9071324.

[4]   A. Pahwa and G. Aggarwal, "Speech Feature Extraction for Gender Recognition," *Int. J. Image, Graph. Signal Process.*, vol. 8, no. 9, pp. 17–25, 2016, doi: 10.5815/ijigsp.2016.09.03.

[5]   P. Zwan, P. Szczuko, and A. Czy, "Neural Networks and Rough Sets," pp. 793–802, 2007.

[6]   I. Wijayanto and R. Dwifebrianti, "Jenis Tipe Jangkauan Suara Pada Pria Dan Wanita Menggunakan Metoda Mel-Frequency Cepstral Coefficient," *Konfrensi Nas. Sist. dan Inform.*, no. October 2013, pp. 2–10, 2013.

[7]   K. Choi, G. Fazekas, M. Sandler, and K. Cho, "Transfer learning for music classification and regression tasks," *Proc. 18th Int. Soc. Music Inf. Retr. Conf. ISMIR 2017*, pp. 141–149, 2017.

[8]   M. Pishgar, F. Karim, S. Majumdar, and H. Darabi, "Pathological Voice Classification Using Mel-Cepstrum Vectors and Support Vector Machine," *Proc. - 2018 IEEE Int. Conf. Big Data, Big Data 2018*, pp. 5267–5271, 2019, doi: 10.1109/BigData.2018.8622208.

[9]   A. Sonawane, M. U. Inamdar, and K. B. Bhangale, "Sound based human emotion recognition using MFCC & multiple SVM," *IEEE Int. Conf. Information, Commun. Instrum. Control. ICICIC 2017*, vol. 2018-Janua, no. August 2017, pp. 1–4, 2018, doi: 10.1109/ICOMICON.2017.8279046.

[10]   T. Van Pham, N. T. N. Quang, and T. M. Thanh, "Deep learning approach for singer voice classification of Vietnamese popular music," *ACM Int. Conf. Proceeding Ser.*, pp. 255–260, 2019, doi: 10.1145/3368926.3369700.

[11]   M. Deng, T. Meng, J. Cao, S. Wang, J. Zhang, and H. Fan, "Heart sound classification based on improved MFCC features and convolutional recurrent neural networks," *Neural Networks*, vol. 130, pp. 22–32, Oct. 2020, doi: 10.1016/j.neunet.2020.06.015.

[12]   A. Elbir, H. O. Ilhan, G. Serbes, and N. Aydin, "Short Time Fourier Transform based music genre classification," *2018 Electr. Electron. Comput. Sci. Biomed. Eng. Meet. EBBT 2018*, pp. 1–4, 2018, doi: 10.1109/EBBT.2018.8391437.

[13]   T. Kobayashi, Y. Suzuki, and A. Kubota, "Audio feature extraction based on sub-band signal correlations for music genre classification," *Proc. - 2018 IEEE Int. Symp. Multimedia, ISM 2018*, no. 3, pp. 180–181, 2019, doi: 10.1109/ISM.2018.00-15.

[14]   D. Nagajyothi and P. Siddaiah, "Speech recognition using convolutional neural networks," *Int. J. Eng. Technol.*, vol. 7, no. 4.6 Special Issue 6, pp. 133–137, 2018, doi: 10.14419/ijet.v7i4.6.20449.

[15]   D. S. Park *et al.*, "Specaugment: A simple data augmentation method for automatic speech recognition," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2019-Septe, pp. 2613–2617, 2019, doi: 10.21437/Interspeech.2019-2680.

[16]   E. Kharitonov *et al.*, "Data Augmenting Contrastive Learning of Speech Representations in the Time Domain," *arXiv*, pp. 1–6, 2020.

[17]   L. Muda, M. Begam, and I. Elamvazuthi, "Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques," vol. 2, no. 3, pp. 138–143, 2010, [Online]. Available: http://arxiv.org/abs/1003.4083.

[18]   R. Hasan, M. Jamil, G. Rabbani, and S. Rahman, "Speaker Identification Using Mel Frequency Cepstral Coefficients," *3rd Int. Conf. Electr. Comput. Eng. ICECE 2004*, no. December, pp. 28–30, 2004.

[19]   E. Zawadzka-Gosk, K. Wołk, and W. Czarnowski, "Deep Learning in State-of-the-Art Image Classification Exceeding 99% Accuracy," in *Advances in Intelligent Systems and Computing*, 2019, vol. 930, doi: 10.1007/978-3-030-16181-1_89.

[20]   J. T. Huang, J. Li, and Y. Gong, "An analysis of convolutional neural networks for speech recognition," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2015, vol. 2015-August, doi: 10.1109/ICASSP.2015.7178920.

[21]   M. Kubanek, J. Bobulski, and J. Kulawik, "A Method of Speech Coding for Speech Recognition Using a Convolutional Neural Network," *Symmetry (Basel).*, vol. 11, no. 9, p. 1185, 2019, doi: 10.3390/sym11091185.

[22]   H. K. Vydana and A. K. Vuppala, "Investigative study of various activation functions for speech

recognition," *2017 23rd Natl. Conf. Commun. NCC 2017*, pp. 6–10, 2017, doi: 10.1109/NCC.2017.8077043.

[23]   C. E. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv*, pp. 1–20, 2018.

[24]   M. Huzaifah, "Comparison of Time-Frequency Representations for Environmental Sound Classification using Convolutional Neural Networks," *arXiv*, pp. 1–5, 2017.