Daftar Pustaka

[1] F. Karray, M. Alemzadeh, J. A. Saleh, and M. N. Arab, "Human-Computer Interaction: Overview on State of the Art," *Int. J. Smart Sens. Intell. Syst.*, vol. 1, no. 1, pp. 137–159, 2008, doi: 10.21307/ijssis-2017-283.

[2] D. A. Reynolds, "An overview of automatic speaker recognition technology," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2002, vol. 4, pp. IV-4072-IV–4075, doi: 10.1109/icassp.2002.5745552.

[3] M. Hamidi, H. Satori, N. Laaidi, and K. Satori, "Conception of Speaker Recognition Methods: A Review," in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology, IRASET 2020*, 2020, pp. 1–6, doi: 10.1109/IRASET48871.2020.9092118.

[4] Z. Özcan and T. Kayıkçıoğlu, "Evaluating MFCC-based speaker identification systems with data envelopment analysis," *Expert Syst. Appl.*, vol. 168, no. December 2020, p. 114448, 2021, doi: 10.1016/j.eswa.2020.114448.

[5] S. Singh and E. G. Rajan, "Application Of Different Filters In Mel Frequency Cepstral Coefficients Feature Extraction And Fuzzy Vector Quantization Approach In Speaker Recognition," *Int. J. Eng. Res. Technol.*, vol. 2, no. 6, pp. 3171–3182, 2013.

[6] V. Tiwari, "MFCC and its applications in speaker recognition," *Int. J. Emerg. Technol.*, vol. 1, no. 1, pp. 19–22, 2010.

[7] I. Stefanus, R. S. J. Sarwono, and M. I. Mandasari, "GMM based automatic speaker verification system development for forensics in Bahasa Indonesia," in *Proceedings of the 2017 5th International Conference on Instrumentation, Control, and Automation, ICA 2017*, 2017, pp. 56–61, doi: 10.1109/ICA.2017.8068413.

[8] J. Martinez, H. Perez, E. Escamilla, and M. M. Suzuki, "Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques," in *CONIELECOMP 2012 - 22nd International Conference on Electronics Communications and Computing*, 2012, pp. 248–251, doi: 10.1109/CONIELECOMP.2012.6189918.

[9] T. Kinnunen and H. Li, "An overview of text-independent speaker recognition: From features to supervectors," *Speech Commun.*, vol. 52, no. 1, pp. 12–40, 2010, doi: 10.1016/j.specom.2009.08.009.

[10] N. Chauhan, T. Isshiki, and D. Li, "Speaker Recognition Using LPC, MFCC, ZCR Features with ANN and SVM Classifier for Large Input Database," in *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, 2019, pp. 130–133, doi: 10.1109/ccoms.2019.8821751.

[11] T. Kinnunen, V. Hautamäki, and P. Fränti, "On the fusion of dissimilarity-based classifiers for speaker identification," in *EUROSPEECH 2003 - 8th European Conference on Speech Communication and Technology*, 2003, pp. 2641–2644.

[12] H. Li *et al.*, "The I4U system in NIST 2008 speaker recognition evaluation," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 4201–4204, 2009, doi: 10.1109/ICASSP.2009.4960555.

[13] L. Chowdhury, H. Zunair, and N. Mohammed, "Robust deep speaker recognition: Learning latent representation with joint angular margin loss," *Appl. Sci.*, vol. 10, no. 21, pp. 1–17, 2020, doi: 10.3390/app10217522.

[14] X. Zhao, Y. Wang, and D. Wang, "ROBUST SPEAKER IDENTIFICATION IN NOISY AND REVERBERANT CONDITIONS Center for Cognitive and Brain Sciences , The Ohio State University , Columbus , OH , USA," pp. 4025–4029, 2014.

[15] S. B. Davis and P. Mermelstein, "Comparison of Parametric Representations for," *Trans. Acoust. Speech, Signal Process.*, vol. 28, no. 4, pp. 357–366, 1980.

[16] Y. Bennani, F. F. Soulie, and P. Gallinari, "A connectionist approach for automatic speaker identification.," 1990.

[17] R. Brunelli and D. Falavigna, "Person Identification Using Multiple Cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 10, pp. 955–966, 1995, doi: 10.1109/34.464560.

[18] J. Gonzalez-Rodriguez, S. Gruz-Llanas, and J. Ortega-Garcia, "Biometric identification through speaker verification over telephone lines.," in *IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology (Cat. No. 99CH36303)*, 1999, pp. 238–242.

[19] S. S. Smith, J. Volkmann, and E. B. Newman, "A Scale for the Measurement of the Psychological Magnitude Pitch," *J. Acoust. Soc. Am.*, vol. 8, no. 3, pp. 185–190, 1937, doi: 10.1121/1.1901999.

[20] N. P. Jawarkar, "Speaker Identification in Noisy Environment," *Int. J. Curr. Eng. Sci. Res.*, vol. 4, no. 7, pp. 37–43, 2017.

[21] Y. A. Alotaibi and A. Hussain, "Comparative analysis of arabic vowels using formants and an automatic speech recognition system," *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 3, no. 2, pp. 11–22, 2010.

[22] A. A. Haydarov and D. Mansurova, "THE ROLE OF INTONATION IN SPEECH AND ARTISTIC SPEECH," *Int. J. Educ. Soc. Sci. Humanit.*, vol. 11, no. 1, pp. 673–676, 2023.

[23] L. L. Koenig and S. Fuchs, "Vowel formants in normal and loud speech," *J. Speech, Lang. Hear. Res.*, vol. 62, no. 5, pp. 1278–1295, 2019, doi: 10.1044/2018_JSLHR-S-18-0043.

[24] L. Nanni, A. Rigo, A. Lumini, and S. Brahnam, "Spectrogram classification using dissimilarity space," *Appl. Sci.*, vol. 10, no. 12, pp. 1–17, 2020, doi: 10.3390/APP10124176.

[25] S. Gupta, M. S. Fahad, and A. Deepak, "Pitch-synchronous single frequency filtering spectrogram for speech emotion recognition," *Multimed. Tools Appl.*, vol. 79, no. 31–32, pp. 23347–23365, 2020, doi: 10.1007/s11042-020-09068-1.

[26] B. Yegnanarayana, A. Joseph, and V. Pannala, "Enhancing formant information in spectrographic display of speech," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2020-Octob, pp. 165–169, 2020, doi: 10.21437/Interspeech.2020-2653.

[27] S. Zhao, Y. Yang, I. Cohen, and L. Zhang, "Speech Emotion Recognition Using Auditory Spectrogram and Cepstral Features," *Eur. Signal Process. Conf.*, vol. 2021-Augus, no. 6177012290, pp. 136–140, 2021, doi: 10.23919/EUSIPCO54536.2021.9616144.

[28] V. Gupta, S. Juyal, and Y. C. Hu, "Understanding human emotions through speech spectrograms using deep neural network," *J. Supercomput.*, vol. 78, no. 5, pp. 6944–6973, 2022, doi: 10.1007/s11227-021-04124-5.

[29] M. Lech, M. Stolar, R. Bolia, and M. Skinner, "Amplitude-frequency analysis of emotional speech using transfer learning and classification of spectrogram images," *Adv. Sci. Technol. Eng. Syst.*, vol. 3, no. 4, pp. 363–371, 2018, doi: 10.25046/aj030437.

[30] Y. Jia *et al.*, "Speaker recognition based on characteristic spectrograms and an improved self-organizing feature map neural network," *Complex Intell. Syst.*, vol. 7, no. 4, pp. 1749–1757, 2021, doi: 10.1007/s40747-020-00172-1.

[31] U. Ayvaz, H. Gürüler, F. Khan, N. Ahmed, T. Whangbo, and A. A. Bobomirzaevich, "Automatic Speaker Recognition Using Mel-Frequency Cepstral Coefficients Through Machine Learning," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 5511–5521, 2022, doi: 10.32604/cmc.2022.023278.

[32] F. Amelia and D. Gunawan, "DWT-MFCC Method for Speaker Recognition System with Noise," in *2019 7th International Conference on Smart Computing and Communications, ICSCC 2019*, 2019, pp. 1–5, doi: 10.1109/ICSCC.2019.8843660.

[33] K. A. Al-karawi and D. Y. Mohammed, "Improving short utterance speaker verification by combining MFCC and Entrocy in Noisy conditions," *Multimed. Tools Appl.*, vol. 80, no. 14, pp. 22231–22249, 2021, doi: 10.1007/s11042-021-10767-6.

[34] K. P. Bharath and M. Rajesh Kumar, "ELM speaker identification for limited dataset using multitaper based MFCC and PNCC features with fusion score," *Multimed. Tools Appl.*,

vol. 79, no. 39, pp. 28859–28883, 2020, doi: 10.1007/s11042-020-09353-z.

[35] K. J. Devi, N. H. Singh, and K. Thongam, "Automatic Speaker Recognition from Speech Signals Using Self Organizing Feature Map and Hybrid Neural Network," *Microprocess. Microsyst.*, vol. 79, no. July, p. 103264, 2020, doi: 10.1016/j.micpro.2020.103264.

[36] H. B.Kekre, V. Kulkarni, P. Gaikar, and N. Gupta, "Speaker Identification using Spectrograms of Varying Frame Sizes," *Int. J. Comput. Appl.*, vol. 50, no. 20, pp. 27–33, 2012, doi: 10.5120/7921-1228.

[37] Y. Lukic, C. Vogt, O. Durr, and T. Stadelmann, "Speaker identification and clustering using convolutional neural networks," *IEEE Int. Work. Mach. Learn. Signal Process. MLSP*, vol. 2016-Novem, 2016, doi: 10.1109/MLSP.2016.7738816.

[38] S. Yadav and A. Rai, "Learning discriminative features for speaker identification and verification," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2018-Septe, no. September 2018, pp. 2237–2241, 2018, doi: 10.21437/Interspeech.2018-1015.

[39] P. Li, Y. Li, D. Luo, and H. Luo, "Speaker identification using FrFT-based spectrogram and RBF neural network," in *Chinese Control Conference, CCC*, 2015, vol. 2015-Septe, no. 1, pp. 3674–3679, doi: 10.1109/ChiCC.2015.7260207.

[40] S. S. Tippannavar, R. Shashidhar, H. R. Sathvik, S. Varun, G. V. Punith, and H. G. Nikshep, "Text Independent Speaker Recognition and Classification using KNN Algorithm," *Proc. 5th Int. Conf. Contemp. Comput. Informatics, IC3I 2022*, pp. 1199–1205, 2022, doi: 10.1109/IC3I56241.2022.10072615.

[41] A. Ashar, M. S. Bhatti, and U. Mushtaq, "Speaker Identification Using a Hybrid CNN-MFCC Approach," 2020, doi: 10.1109/ICETST49965.2020.9080730.

[42] J. P. Campbell, "Speaker Recognition : A Tutorial," *Proc. IEEE*, vol. 85, no. 9, pp. 1437–1462, 1997.

[43] K. P. I. U. Gorai and T. A. Jv, "A GAUSSIAN MIXTURE MODEL-BASED SPEAKER RECOGNITION SYSTEM," *Asian J. Pharm. Clin. Res.*, vol. 10, pp. 140–142, 2017, doi: 10.22159/ajpcr.2017.v10s1.19596.

[44] L. Ferrer, M. McLaren, and N. Brümmer, "A speaker verification backend with robust performance across conditions," *Comput. Speech Lang.*, vol. 71, no. May 2021, p. 101258, 2022, doi: 10.1016/j.csl.2021.101258.

[45] S. Singh, "Forensic and automatic speaker recognition system," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 2804–2811, 2018, doi: 10.11591/ijece.v8i5.pp.2804-2811.

[46] M. Faundez-Zanuy and E. Monte-Moreno, "State-of-the-art in speaker recognition," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 5, pp. 7–12, 2005, doi: 10.1109/MAES.2005.1432568.

[47] A. Revathi, C. Ravichandran, P. Saisiddarth, and G. S. R. Prasad, "Isolated Command Recognition Using MFCC and Clustering Algorithm," *SN Comput. Sci.*, vol. 1, no. 2, pp. 1–7, 2020, doi: 10.1007/s42979-020-0093-x.

[48] Himani Chauhan, S. Samal, and A. Ghoshal, "Voice recognition," *Int. J. Comput. Sci. Mob. Comput.*, vol. 4, no. 4, pp. 296–301, 2015, doi: 10.1201/9781315366760-6.

[49] S. M. Qaisar, "Isolated speech recognition and its transformation in visual signs," *J. Electr. Eng. Technol.*, vol. 14, no. 2, pp. 955–964, 2019, doi: 10.1007/s42835-018-00071-z.

[50] M. M. Hasan, A. M. Nasr, and S. Sultana, "An approach to voice conversion using feature statistical mapping," *Appl. Acoust.*, vol. 66, no. 5, pp. 513–532, 2005, doi: 10.1016/j.apacoust.2004.09.005.

[51] M. Hariharan, L. S. Chee, O. C. Ai, and S. Yaacob, "Classification of speech dysfluencies using LPC based parameterization techniques," *J. Med. Syst.*, vol. 36, no. 3, pp. 1821–1830, 2012, doi: 10.1007/s10916-010-9641-6.

[52] V. Z. Këpuska and H. A. Elharati, "Robust Speech Recognition System Using Conventional and Hybrid Features of MFCC, LPCC, PLP, RASTA-PLP and Hidden

Markov Model Classifier in Noisy Conditions," *J. Comput. Commun.*, vol. 03, no. 06, pp. 1–9, 2015, doi: 10.4236/jcc.2015.36001.

[53] R. Bharti and P. Bansal, "Real Time Speaker Recognition System using MFCC and Vector Quantization Technique," *Int. J. Comput. Appl.*, vol. 117, no. 1, pp. 25–31, 2015, doi: 10.5120/20520-2361.

[54] S.B.Dhonde and S.M.Jagade, "Mel-Frequency Cepstral Coefficients for Speaker Recognition: A Review," *Int. J. Adv. Eng. Res. Dev.*, vol. 2, no. 5, pp. 1115–1119, 2015.

[55] T. Kinnunen, "Designing a speaker-discriminative adaptive filter bank for speaker recognition," in *7th International Conference on Spoken Language Processing, ICSLP 2002*, 2002, pp. 2325–2328.

[56] H. M. S. Naing, "PERCEPTUAL-BASED FEATURE EXTRACTION FOR ROBUST AUTOMATIC SPEECH RECOGNITION," Universitas Gadjah Mada, 2020.

[57] Matlab, "Speaker Identification Using Pitch and MFCC." .

[58] R. Hidayat and A. Winursito, "A Modified MFCC for Improved Wavelet-Based Denoising on Robust Speech Recognition," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 1, pp. 12–21, 2021, doi: 10.22266/ijies2021.0228.02.

[59] I. K. Timotius, "Computational Methods for Gait Analysis in Rodents," 2020. .

[60] R. Lyons, "Windowing functions improve FFT results," *Test Meas. World*, vol. 18, no. 7, pp. 37–44, 1998.

[61] S. Roberts, "Lecture 7 - The Discrete Fourier Transforms." https://www.robots.ox.ac.uk/~sjrob/Teaching/SP/l7.pdf (accessed Nov. 15, 2021).

[62] P. Heckbert, "Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm," 1995. http://www.cs.cmu.edu/afs/andrew/scs/cs/15-463/2001/pub/www/notes/fourier/fourier.pdf (accessed Nov. 16, 2021).

[63] K. S. Rao and K. E. Manjunath, "MFCC Features," in *Speech Recognition Using Articulatory and Excitation Source Features*, Springer, 2017, pp. 85–88.

[64] J. W. Picone, "Signal Modeling Techniques in Speech Recognition," *Proc. IEEE*, vol. 81, no. 9, pp. 1215–1247, 1993, doi: 10.1109/5.237532.

[65] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of MFCC.," *J. Comput. Sci. Technol.*, vol. 16, no. 6, pp. 582–589, 2001.

[66] T. Ganchev, N. Fakotakis, and G. Kokkinakis, "Comparative evaluation of various MFCC implementations on the speaker verification task Comparative Evaluation of Various MFCC Implementations on the Speaker Verification Task," 2005, no. January.

[67] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, no. January, pp. 90–93, 1974.

[68] "Discrete Cosine Transform," *Wikipedia*. .

[69] Y. Afrillia, "Performance Measurement Of Mel Frequency Ceptral Coefficient (MFCC) Method In Learning System Of Al- Qur ' an Based In Nagham Pattern Recognition Performance Measurement Of Frequency Ceptral Coefficient (MFCC) Method In Learning System Of AlQur ' anBa," 2017.

[70] W. Han, C. F. Chan, C. S. Choy, and K. P. Pun, "An efficient MFCC extraction method in speech recognition," *Proc. - IEEE Int. Symp. Circuits Syst.*, pp. 145–148, 2006, doi: 10.1109/iscas.2006.1692543.

[71] H. Meng, T. Yan, F. Yuan, and H. Wei, "Speech Emotion Recognition from 3D Log-Mel Spectrograms with Deep Learning Network," *IEEE Access*, vol. 7, pp. 125868–125881, 2019, doi: 10.1109/ACCESS.2019.2938007.

[72] Y. Zeng, H. Mao, D. Peng, and Z. Yi, "Spectrogram based multi-task audio classification," *Multimed. Tools Appl.*, vol. 78, no. 3, pp. 3705–3722, 2019, doi: 10.1007/s11042-017-5539-3.

[73] S. M. Dol, "Use of Classification Technique in Educational Data Mining," *2021 Int. Conf.*

Nascent Technol. Eng. ICNET 2021 - Proc., no. Icnte, pp. 1–7, 2021, doi: 10.1109/ICNTE51185.2021.9487739.

[74] Y. C. A. P. Reddy, S. P. P. Sagar, R. P. Kalyan, and N. S. Charan, "Classification of Hotel Reviews using Machine Learning Techniques," *8th Int. Conf. Smart Struct. Syst. ICSSS 2022*, pp. 1–5, 2022, doi: 10.1109/ICSSS54381.2022.9782215.

[75] Y. Gao and F. Gao, "Edited AdaBoost by weighted kNN," *Neurocomputing*, vol. 73, no. 16–18, pp. 3079–3088, 2010, doi: 10.1016/j.neucom.2010.06.024.

[76] V. B. S. Prasath *et al.*, "Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review," pp. 1–39, 2017, doi: 10.1089/big.2018.0175.

[77] H. Bhavsar and A. Ganatra, "A Comparative Study of Training Algorithms for Supervised Machine Learning," *Int. J. Soft Comput. Eng.*, vol. 2, no. 4, pp. 74–81, 2012.

[78] J. Walters-Williams and Y. Li, "Comparative Study of Distance Functions for Nearest Neighbors," in *Advanced Techniques in Computing Sciences and Software Engineering*, 2010, pp. 79–84, doi: 10.1007/978-90-481-3660-5.

[79] M. Mohibullah, M. Z. Hossain, and M. Hasan, "Comparison of Euclidean Distance Function and Manhattan Distance Function Using K-Mediods," *Int. J. Comput. Sci. Inf. Secur.*, vol. 13, no. 10, pp. 61–71, 2015.

[80] J. Wong, "K-Nearest Neighbors Algorithm (KNN)," 2020. https://towardsdatascience.com/k-nearest-neighbors-algorithm-d4a8bb1926a3 (accessed Mar. 20, 2022).

[81] R. Beulnan and R. Kalaba, "On adaptive control processes.," *IRE Trans. Autom. Control*, vol. 4, no. 2, pp. 1–9, 1959.

[82] M. W. Wong, "Haar Wavelets," in *Discrete Fourier Analysis*, 2011, pp. 67–78.

[83] D. Johnson, "Signal-to-noise ratio," *Scholarpedia*, vol. 1, no. 12. p. 2088, 2006, doi: 10.4249/scholarpedia.2088.

[84] M. D, "Noise Level Decibels Chart – A Comprehensive Guide," *Electronics Hubs*, 2022. https://www.electronicshub.org/noise-level-decibels-chart/.

[85] A. Maurya, D. Kumar, and R. K. Agarwal, "Speaker Recognition for Hindi Speech Signal using MFCC-GMM Approach," *Procedia Comput. Sci.*, vol. 125, pp. 880–887, 2018, doi: 10.1016/j.procs.2017.12.112.

[86] Aniruddha Bhandari, "Confusion Matrix for Machine Learning," *Analytics Vidhya*, 2020. https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/ (accessed Mar. 20, 2022).