

References

- [1] Lawrence R. Rabiner, Biing-Hwang Juang, and B. Yegnanarayana. *Fundamentals of Speech Recognition*. Pearson Education, Dorling Kindersley (India) Pvt. Ltd., New Delhi, India., 2 edition, 2010.
- [2] J. E. Hawkins. Human ear. Encyclopedia Britannica, March 2023.
- [3] NIDCD. How do we hear? The National Institute on Deafness and Other Communication Disorders (NIDCD), March 2023.
- [4] Fatma Gongor and Onder Tutsoy. Social companion robots: A speech-based emotion analysis algorithm for quarantined patients. *4th International Congress on Engineering Sciences and Multidisciplinary Approaches*, page 5, 2022.
- [5] B. H. Juang and Lawrence R. Rabiner. *Automatic Speech Recognition A Brief History of the Technology*, 2005.
- [6] Daniel Jurafsky and James H Martin. *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, 2 edition, 2008.
- [7] Raviraj Joshi and Venkateshan Kannan. Attention based end to end speech recognition for voice search in hindi and english. In *Proceedings of the 13th Annual Meeting of the Forum for Information Retrieval Evaluation*, pages 107–113, 2021.
- [8] Rahul Jampala, Devisri Santosh Kola, Adithya Nagendra Gummadi, Meghana Bhavanam, and Ithaya Rani Pannerselvam. The evolution of voice assistants: From text-to-speech to conversational ai. In *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, pages 1332–1338. IEEE, 2024.
- [9] Hemakumar, G and Punitha, P. Speech Recognition Technology: A Survey on Indian Languages. *International Journal of Information Science and Intelligent System*, 2(4):1–38, 2013.

- [10] Aditya Vashistha, Pooja Sethi, and Richard Anderson. Respeak: A voice-based, crowd-powered speech transcription system. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 1855–1866, 2017.
- [11] Joost Van Doremalen, Lou Boves, Jozef Colpaert, Catia Cucchiarini, and Helmer Strik. Evaluating automatic speech recognition-based language learning systems: A case study. *Computer Assisted Language Learning*, 29(4):833–851, 2016.
- [12] Mike Wald. Learning through multimedia: Speech recognition enhancing accessibility and interaction. *Journal of Educational Multimedia and Hypermedia*, 17(2):215–233, 2008.
- [13] Bhagath Parabattina, Phool Chandra, Vanshali Sharma, and Pradip K Das. Voice-controlled assistance for robot navigation using android-based mobile devices. In *2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 21–25. IEEE, 2021.
- [14] Yogesh Kumar. A comprehensive analysis of speech recognition systems in health-care: Current research challenges and future prospects. *SN Computer Science*, 5(1):137, 2024.
- [15] Bilva Teja, Nischitha P Chinnari, Prerana Hadadi, Surabhi Shanbhogh, and VR Badri Prasad. A novel approach in the automatic generation of regional language subtitles for videos in english. In *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, pages 1–6. IEEE, 2023.
- [16] Parabattina Bhagath, Savinay Parihar, and Pradip K Das. Speech recognition for indian spoken languages towards automated home appliances. In *2021 2nd International Conference for Emerging Technology (INCET)*, pages 1–5. IEEE, 2021.
- [17] Mrinmoy Bhattacharjee, Shikha Baghel, and SR Mahadeva Prasanna. Driver speech detection in real driving scenario. In *International Conference on Speech and Computer*, pages 189–199. Springer, 2023.
- [18] Amrita S Tulshan and Sudhir Namdeorao Dhage. Survey on virtual assistant: Google assistant, siri, cortana, alexa. In *Advances in Signal Processing and Intelligent Recognition Systems: 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers 4*, pages 190–201. Springer, 2019.
- [19] G. Hemantkumar and P. Punitha. Speech Recognition Technology: A Survey on Indian Languages. *International Journal of Information Science and Intelligent System*, 2(4):1–38, 2013.

- [20] David M. Eberhard, Gary F. Simons, and Charles D. Fennig, editors. *Ethnologue: Languages of the World*. SIL International, Dallas, Texas, twenty-fifth edition, 2022.
- [21] George Cardona and Danesh Jain. *The Indo-Aryan Languages*. Routledge, Abingdon, Oxon, United Kingdom, 1 edition, 2007.
- [22] Cini Kurian. A Survey on Speech Recognition in Indian Languages. *International Journal of Computer Science and Information Technologies*, 5(5):6169–6175, 2014.
- [23] Jinal H Tailor and Dipti B Shah. Review on Speech Recognition System for Indian Languages. *International Journal of Computer Applications*, 119(2):15–18, 2015.
- [24] Amitoj Singh, Virender Kadyan, Munish Kumar, and Nancy Bassan. ASRoIL: a comprehensive survey for automatic speech recognition of Indian languages. *Artificial Intelligence Review*, 53(5):3673–3704, 2020.
- [25] Rachana Bharat Parikh and Hiren Joshi. Gujarati Speech Recognition A Review. *Test Engineering and Management*, 83(4):549–553, 2020.
- [26] Kewal D. Malde, Bhavik B. Vachhani, Maulik C. Madhavi, Nirav H. Chhayani, and Hemant A. Patil. Development of speech corpora in Gujarati and Marathi for phonetic transcription. In *International Conference Oriental COCODA Held Jointly with 2013 Conference on Asian Spoken Language Research and Evaluation, O-COCODA/CASLRE 2013*, pages 1–6, 2013.
- [27] Maulik C. Madhavi, Shubham Sharma, and Hemant A. Patil. Development of language resources for speech application in Gujarati and Marathi. In *International Conference on Asian Language Processing (IALP)*, pages 115–118, 2014.
- [28] Vishal P. Tank and S. K. Hadia. Creation of speech corpus for emotion analysis in Gujarati language and its evaluation by various speech parameters. *International Journal of Electrical and Computer Engineering*, 10(5):4752–4758, 2020.
- [29] H. B. Chauhan and B. A. Tanawala. Comparative Study of MFCC And LPC Algorithms for Gujarati Isolated Word Recognition. *International Journal of Innovative Research in Computer and Communication Engineering*, 3(2):822–826, 2015.
- [30] Jigarkumar Patel and Arun Nandurbarkar. Development and Implementation of Algorithm for Speaker recognition for Gujarati Language. *International Research Journal of Engineering and Technology*, 2(2):444–448, 2015.
- [31] Patel Pravin and Harikrishna Jethva. Neural Network Based Gujarati Language Speech Recognition. *International Journal of Computer Science and Management Research*, 2(5):2623–2627, 2013.

- [32] Vijayendra A. Desai and Vishvjit K. Thakar. Neural Network Based Gujarati Speech Recognition for Dataset Collected by in-ear Microphone. *Procedia Computer Science*, 93(2016):668–675, 2016.
- [33] Anshu Chittora and Hemant A. Patil. Classification of phonemes using modulation spectrogram based features for Gujarati language. In *International Conference on Asian Language Processing (IALP)*, pages 46–49, 2014.
- [34] Jinal H. Tailor and Dipti B. Shah. Speech Recognition System Architecture for Gujarati Language. *International Journal of Computer Applications*, 138(12):28–31, 2016.
- [35] Jinal Tailor and Dipti Shah. HMM-Based Lightweight Speech Recognition System for Gujarati Language. *Lecture Notes in Networks and Systems*, 10:451–461, 2018.
- [36] Sanjay Valaki and Harikrishna Jethva. A hybrid HMM/ANN approach for automatic Gujarati speech recognition. In *International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pages 1–5, 2017.
- [37] Himanshu N Patel and P V Virparia. A Small Vocabulary Speech Recognition for Gujarati. *International Journal of Advanced Research in Computer Science*, 2(1):208–210, 2011.
- [38] Shubham Sharma, Maulik C. Madhavi, and Hemant A. Patil. Development of vocal tract length normalized phonetic engine for Gujarati and Marathi languages. In *17th Oriental Chapter of the International Committee for the Co-ordination and Standardization of Speech Databases and Assessment Techniques (COCOSDA)*, pages 1–6, 2014.
- [39] Vijayendra Desai and Vishvjit Thakar. Word boundary detection for Gujarati speech recognition using in-ear microphone. *1st India International Conference on Information Processing (IICIP)*, pages 1–6, 2016.
- [40] Neil Patel, Sheetal Agarwal, Nitendra Rajput, Amit Nanavati, Paresh Dave, and Tapan S. Parikh. A comparative study of speech and dialed input voice interfaces in rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 51–54. Association for Computing Machinery, 2009.
- [41] Jigisha K Patel, Pritesh N Patel, and Paresh V Virparia. Gujarati Language Speech Recognition System for Identifying Smartphone Operation Commands. *National Journal of System and Information Technology*, 8(2):79–88, 2015.

- [42] Shubham Toshniwal, Tara N. Sainath, Ron J. Weiss, Bo Li, Pedro Moreno, Eugene Weinstein, and Kanishka Rao. Multilingual Speech Recognition with a Single End-to-End Model. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4904–4908, 2018.
- [43] Anjali Kannan, Arindrima Datta, Tara N. Sainath, Eugene Weinstein, Bhuvana Ramabhadran, Yonghui Wu, Ankur Bapna, Zhifeng Chen, and Seungji Lee. Large-scale multilingual speech recognition with a streaming end-to-end model. In *Proc. Interspeech 2019*, pages 2130–2134, 2019.
- [44] Anuj Diwan and Preethi Jyothi. Reduce and Reconstruct: Improving Low-resource End-to-end ASR Via Reconstruction Using Reduced Vocabularies. *ArXiv*, 2020.
- [45] Brij Mohan Lal Srivastava, Sunayana Sitaram, Rupesh Kumar Mehta, Krishna Doss Mohan, Pallavi Matani, Sandeepkumar Satpal, Kalika Bali, Radhakrishnan Srikanth, and Niranjana Nayak. Low Resource Automatic Speech Recognition Challenge for Indian Languages. In *Proc. 6th Intl. Workshop on Spoken Language Technologies for Under-Resourced Languages*, pages 11–14, 2018.
- [46] Jayadev Billa. ISI ASR System for the Low Resource Speech Recognition Challenge for Indian Languages. In *Proc. Interspeech 2018*, pages 3207–3211, 2018.
- [47] Noor Fathima, Tanvina Patel, C. Mahima, and Anuroop Iyengar. TDNN-based multilingual speech recognition system for low resource Indian languages. In *Proc. Interspeech 2018*, pages 3197–3201, 2018.
- [48] Hardik B. Sailor and Thomas Hain. Multilingual speech recognition using language-specific phoneme recognition as auxiliary task for indian languages. In *Proc. Interspeech 2020*, pages 4756–4760, 2020.
- [49] Mohit Dua, Bhavesh Bhagat, Shelza Dua, and Nidhi Chakravarty. A review on gujarati language based automatic speech recognition (asr) systems. *International Journal of Speech Technology*, 27(1):133–156, 2024.
- [50] Mohit Dua and Akanksha. Gujarati language automatic speech recognition using integrated feature extraction and hybrid acoustic model. In *Proceedings of Fourth International Conference on Communication, Computing and Electronics Systems: ICCCES 2022*, pages 45–54. Springer, 2023.
- [51] Yash Sharma, Basil Abraham, and Preethi Jyothi. Gujarati-english code-switching speech recognition using ensemble prediction of spoken language. *arXiv preprint arXiv:2403.08011*, 2024.

- [52] Meera M Shah and Hiren R Kavathiya. Exploring speech corpus for voice recognition in gujarati: An in-depth study. *Educational Administration: Theory and Practice*, 30(6 (S)):168–175, 2024.
- [53] Bhavesh Bhagat and Mohit Dua. Improved spell corrector algorithm and deep-speech2 model for enhancing end-to-end gujarati language asr performance. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 7:100441, 2024.
- [54] Urmila Shrawankar and Vilas M. Thakare. Techniques for feature extraction in speech recognition system : A comparative study. *CoRR*, abs/1305.1145, 2013.
- [55] Z. Tufekci and J. N. Gowdy. Feature extraction using discrete wavelet transform for speech recognition. In *Conference Proceedings - IEEE SOUTHEASTCON*, pages 116–123, 2000.
- [56] Steven Davis and Paul Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, 28(4):357–366, 1980.
- [57] Zekeriya Tufekci, John Gowdy, Sabri Gurbuz, and Eric Patterson. Applied mel-frequency discrete wavelet coefficients and parallel model compensation for noise-robust speech recognition. *Speech Communication*, 48(10):1294–1307, 2006.
- [58] Chun-Liu Liu. A Tutorial of the Wavelet Transform. Technical report, 2010.
- [59] Paolo Prandoni and Martin Vetterli. *Signal processing for communications*, volume 1. EPFL Press, 2008.
- [60] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.
- [61] Toni Giorgino. Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of Statistical Software*, 31(7):124, 2009.
- [62] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.
- [63] A Maas, A Hannun, and A Ng. Rectifier nonlinearities improve neural network acoustic models. In *ICML Workshop on Deep Learning for Audio, Speech and Language Processing*, 2013.
- [64] Diederik P. Kingma and Jimmy Lei Ba. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*, pages 1–15, 2015.

- [65] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. The MIT Press, Cambridge, Massachusetts, USA, 1 edition, 2016.
- [66] David S. Broomhead and David Lowe. Multivariable functional interpolation and adaptive networks. *Complex Syst.*, 2, 1988.
- [67] J. Park and I. W. Sandberg. Universal Approximation Using Radial-Basis-Function Networks. *Neural Computation*, 3(2):246–257, 1991.
- [68] Rudolf Kruse, Christian Borgelt, Frank Klawonn, Christian Moewes, Matthias Steinbrecher, and Pascal Held. *Radial Basis Function Networks*, pages 83–103. Springer London, 2013.
- [69] Yue Wu, Hui Wang, Biaobiao Zhang, and K.-L. Du. Using Radial Basis Function Networks for Function Approximation and Classification. *ISRN Applied Mathematics*, pages 1–34, 2012.
- [70] Yariv Ephraim. *Hidden Markov Models*, pages 704–708. Springer US, Boston, MA, 2013.
- [71] Lawrence R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [72] L. R. Rabiner and B. H. Juang. An Introduction to Hidden Markov Models. *IEEE ASAP Magazine*, pages 4–16, jan 1986.
- [73] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- [74] Dipl.-Inf Frank Krüger. *Activity , Context , and Plan Recognition with Computational Causal Behaviour Models*. PhD thesis, University of Rostock, 2016.
- [75] Purnima Pandit and Shardav Bhatt. Automatic Speech Recognition of Gujarati digits using Dynamic Time Warping. *International Journal of Engineering and Innovative Technology*, 3(12):69–73, 2014.
- [76] Purnima Pandit, Shardav Bhatt, and Priyank Makwana. Automatic Speech Recognition of Gujarati Digits using Artificial Neural Network. In *Proceedings of 19th Annual Cum 4th International Conference of GAMS On Advances in Mathematical Modelling to Real World Problems*, pages 141–146. Excellent Publishers, 2014.
- [77] Purnima Pandit and Shardav Bhatt. Automatic Speech Recognition of Gujarati digits using Radial Basis Function Network. In *International Conference on Futuristic*

- Trends in Engineering, Science, Pharmacy and Management*, pages 216–226. A D Publication, 2016.
- [78] Purnima Pandit and Shardav Bhatt. Automatic Speech Recognition of Gujarati digits using Wavelet Coefficients. *Journal of The Maharaja Sayajirao University of Baroda*, 52(1):101–110, 2017.
- [79] K. G. Sheela and S. N. Deepa. Selection of number of hidden neurons in neural networks in renewable energy systems. *Journal of Scientific and Industrial Research*, 73(10):686–688, 2014.
- [80] Katsunari Shibata and Yusuke Ikeda. Effect of number of hidden neurons on learning in large-scale layered neural networks. In *ICCAS-SICE 2009 - ICROS-SICE International Joint Conference 2009, Proceedings*, number January 2009, pages 5008–5013, 2009.
- [81] Purnima Pandit and Shardav Bhatt. Automatic speech recognition of Gujarati digits using wavelet coefficients in machine learning algorithms. *International Journal of Innovative Computing and Applications (IJICA)*, 14(4):191–200, 2023.
- [82] Purnima Pandit and Shardav Bhatt. Automatic speech recognition of Gujarati digits using Wavelet-based features in Gaussian Mixture Models and Hidden Markov Models. *Global Review of Business and Technology (GRBT)*, 2(2):1–10, 2022.
- [83] Purnima Pandit, Priyank Makwana, and Shardav Bhatt. Automatic Speech Recognition of Continuous Speech Signal of Gujarati Language Using Machine Learning. In *AISC*, pages 147–159. Springer, Singapore, 2021.
- [84] Yunxin Zhao, Jian Xue, and Xin Chen. Ensemble learning approaches in speech recognition. In *Speech and audio processing for coding, enhancement and recognition*, pages 113–152. Springer, 2014.
- [85] Jonathan Le Roux, Shinji Watanabe, and John R Hershey. Ensemble learning for speech enhancement. In *2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 1–4. IEEE, 2013.
- [86] Xibin Dong, Zhiwen Yu, Wenming Cao, Yifan Shi, and Qianli Ma. A survey on ensemble learning. *Frontiers of Computer Science*, 14:241–258, 2020.
- [87] Omer Sagi and Lior Rokach. Ensemble learning: A survey. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 8(4):e1249, 2018.