

English Accent Recognition System Using Deep Learning

¹Mr. P. Srinivasa Rao, ²B. Varun, ³D. Sundar Daniel, ⁴C. Akshaya,
⁵B. Shalini

*1Associate Professor, 2,3,4,5Student
1,2,3,4,5Department of Computer Science and Engineering,
1,2,3,4,5ACE Engineering College, Hyderabad, India*

ABSTRACT:

Accent detection is an emerging field that aims to identify a person's pronunciation or accent based on their geographic location. Deep learning techniques, such as Convolutional Neural Networks (CNNs), have been employed to recognize English accents. This software is designed to determine the closest English accent spoken in a particular location. Categorizing languages or accents provides insights into people's backgrounds and the demographic regions they inhabit. This information can be valuable in various domains. In this study, CNN is utilized for accent detection among English speakers. Experimental results demonstrate the effectiveness of deep learning in accent detection and the potential of the algorithms used for accent classification. This process holds significant utility in security domains.

Date of Submission: 13-06-2023

Date of acceptance: 28-06-2023

I. INTRODUCTION:

English is one of the most widely spoken languages globally, with over 70 countries using it as a primary or official language. However, English speakers from different countries exhibit distinct accents. An accent refers to the way words are pronounced in a language. Accent variations can lead to communication difficulties, hindering mutual understanding and causing trust issues among individuals. Studies have shown that people with accents are often less trusted. Additionally, accents and dialects can provide information about a person's origin and ethnicity. In the context of border control, identifying a suspect's origin and ethnicity can aid law enforcement in gathering crucial details about their background and country of origin, facilitating the identification process. Artificial intelligence and machine learning techniques enable automatic accent and dialect detection and classification, serving as a crucial stage in speech recognition systems.

Speech processing is a specialized branch of signal processing that plays a vital role in various commercial applications, including banking, travel, telecommunications, and voice dialing. It encompasses several areas such as language identification, speech enhancement, spoken dialog systems, speaker recognition and verification, speech coding, emotion and attitude recognition, speech segmentation and labeling, speech recognition, prosody, text-to-speech synthesis, and audio-visual signal processing. Speech recognition systems convert spoken language into text format, allowing machines to analyze and understand human speech. Techniques like vector quantization, feature extraction, dynamic time warping, hidden Markov models, Gaussian mixture models, decision-tree-based clustering, and language models are commonly employed in speech recognition systems. This paper focuses on continuous and connected speech, utilizing MFCC (Mel-frequency cepstral coefficients) features extracted from speech corpora. The extracted speech signals are trained using HMM (Hidden Markov Model) models, and the output results are compared with continuous and connected speech.

1.1. EXISTING SYSTEM

Existing systems for English accent recognition range from gradient boosting to MLP (Multi-Layer Perceptron). For instance, gradient boosting is a machine learning technique used in regression and classification tasks. It constructs an ensemble of weak prediction models, often decision trees, to form a prediction model. Gradient-boosted trees, built in a stage-wise manner, generally outperform random forest algorithms. However, existing systems suffer from certain drawbacks, including a lack of accuracy, issues with quality control, and limited data set sizes.

1.2. PROPOSED SYSTEM

The proposed English accent recognition system employs classification algorithms. The following steps are involved in the proposed system:

1. Data pre-processing

2. Feature extraction
3. Training and testing
4. Performance evaluation based on the accuracy metric
5. System optimization

ADVANTAGES

Efficient Information Retrieval: An English accent recognition system facilitates the organization and retrieval of large volumes of English accent audio data. By categorizing documents into different classes or topics, researchers and translation professionals can efficiently search for relevant articles and research papers within their area of interest.

Automated Decision Support: English accent recognition systems can serve as valuable tools in automated decision support systems. By analyzing a user's voice, these systems can identify the location associated with a particular accent, assisting translators in categorizing accents.

II. CONCLUSION

An accent detection and classification system has been proposed for use in deception detection and order management. The proposed system utilizes deep learning techniques, resulting in significantly improved results. Accurate accent classification provides insights into a speaker's origin and heritage, which is particularly valuable in languages like English with diverse regional dialects and widespread usage. Extensive simulation results have demonstrated excellent accuracy, surpassing 90% when employing Deep Transfer Learning. Further research is required to collect and simulate unique datasets for applications such as border control, litigation, healthcare, border security, and domestic lie detection.

ACKNOWLEDGEMENTS

We extend our gratitude to our guide, Mr. P. SRINIVASA RAO, and the project coordinator, Mr. V. Chandra Sekhar Reddy, for their invaluable guidance throughout this research project. We also acknowledge the support of Dr. M. V. Vijaya Saradhi, Head of the Computer Science and Engineering Department at ACE Engineering College, for his unwavering assistance during the completion of this project.

REFERENCES:

- [1]. FluentU, English Language and Culture Blog, [Link]. Accessed on 1/5/2022.
- [2]. Yunus Korkmaz Y., Boyacı A. (2022), "A comprehensive Turkish accent/dialect recognition system using acoustic perceptual formants," *Applied Acoustics*, Volume 193.
- [3]. Faria A. (2005), "Accent Classification for Speech Recognition," *Machine Learning for Multimodal Interaction*, pp. 285–293.
- [4]. Chen, P., Lee, J., Neidert, J. (n.d.), "Foreign accent classification," [Link]. Accessed on 2/5/2022.
- [5]. Radzikowski, K., Wang, L., Yoshie, O. et al. (2021), "Accent modification for speech recognition of non-native speakers using neural style transfer," *J AUDIO SPEECH MUSIC PROC.*, 11. [Link]
- [6]. Chen T., Huang C., Chang C., and Wang J. (2002), "On the use of Gaussian mixture model for speaker variability analysis," the Int. Conf. SLP, Denver, CO.
- [7]. The Polygraph and Lie Detection (2003), The National Academies of Sciences Engineering Medicine, [Link]. Abootalebi V, Moradi MH, Khalilzadeh MA. (2009), "A new approach for EEG feature extraction in P300-based lie detection," *Comput Methods Programs Biomed*, 94 (1), pp. 48-57.