JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ELECTRONIC COMPONENT IDENTIFICATION FROM VOICE

MD. SHAREEF, JADAV HIMAJA, JANGIDI PAVAN YADAV, VELPULA SRIRAM

Department of CSE(AIML) CMR Technical Campus(Autonomous), Kandlakoya Telangana, India

ABSTRACT- This project describes the architecture of a door phone embedded system with interactive voice response. Because speech technology is not 100% reliable, the emphasis was on parts that have greater impact on overall performance (audio capture, speech recognition and verification, and power consumption). Using an embedded microphone array increases speech recognition effectiveness in very noisy environments.

To increase the speech recognition performance, a null grammar with confidence measure support was used. The speaker verification module was also optimized for nosy environments (using the cepstral mean normalization technique and a universal background mode). Any fundamental discrete device or physical object in an electronic system that affects electrons or the fields around them is referred to as an electronic component. Electrical elements, which are conceptual, should not be confused with electronic components, the majority of which are industrial goods that are only available in a single form. Abstractions used to represent idealised electronic parts and components. There are numerous electrical terminals or leads on electronic components. To build an electronic circuit with a specific purpose (such as an amplifier, radio receiver, or oscillator), these leads link to other electrical components, frequently by wire. Basic electronic components can be integrated inside of discrete packages like semiconductor integrated circuits, hybrid integrated circuits, or thick film devices, as well as arrays or networks of similar components. The list of electrical components that follows concentrates on their discrete form and treats such bundles as independent components.

I. INTRODUCTION

This project aims to provide an in-depth analysis of electronic components, categorizing them into three primary types: passive, active, and electromechanical components. The study will focus on how these components function within electrical and electronic circuits, as well as their role in circuit design and analysis. The project will explore the fundamental differences between these component types by examining their energy interaction properties, signal processing capabilities, and real-world applications. Special attention will be given to how electrical engineers differentiate between DC and AC circuit analysis, allowing for a better understanding of how energy flows and signals are processed within circuits. Additionally, this project will

highlight the limitations and advantages of each component type, explaining why certain components are chosen for specific applications. By incorporating practical examples such as transistors, resistors, capacitors, inductors, and transformers, the project will provide a comprehensive learning resource for students, engineers, and electronics enthusiasts.

The primary purpose of this project is to establish a clear and structured understanding of electronic components and their behavior in circuit analysis. Electronics is a vast field, and distinguishing between different components is essential for designing efficient circuits. Active components, such as transistors and vacuum tubes, require an external energy source and have the ability to amplify or manipulate signals, making them crucial in power regulation, signal processing, and computation. Passive components, including resistors, capacitors, and inductors, do not generate energy but instead store, dissipate, or transfer it, playing vital roles in filtering, impedance matching, and resonance circuits. Electromechanical components, such as relays and switches, bridge the gap between electrical and mechanical systems, enabling automation and control in modern electronics. The project also delves into how engineers approach circuit analysis by abstracting DC power sources to focus solely on AC signal behavior. This perspective allows for a more straightforward understanding of how circuits function dynamically while ignoring unnecessary complexities. Through this project, learners will gain insights into how components interact, how energy is transferred, and how different elements contribute to circuit functionality, ultimately aiding in better circuit design and troubleshooting skills. The primary purpose of this project is to establish a clear and structured understanding of electronic components and their behavior in circuit analysis. Electronics is a vast field, and distinguishing between different components is essential for designing efficient circuits. Active components, such as transistors and vacuum tubes, require an external energy source and have the ability to amplify or manipulate signals, making them crucial in power regulation, signal processing, and computation.

II. RELATED WORK

The development of electronic component recognition from voice has been closely tied to advancements in speech recognition systems, voice-controlled applications, computer vision, multimodal human-computer interaction (HCI), and assistive technologies. Foundational tools like Google Speech API, CMU Sphinx, and Mozilla Deep Speech, as well as hybrid approaches combining Hidden Markov Models (HMMs) and Neural Networks, have significantly improved the accuracy of transcribing voice into text.[2] This technology has also been employed in educational and engineering applications, such as voice-controlled circuit design tutors, which help beginners understand components, and IoT-enabled smart labs that utilize speech recognition for identifying components and simulating circuits. While your system focuses on voice input, some systems incorporate computer vision with Convolutional Neural Networks (CNNs) to classify components from images, often in tandem with voice commands.[1] Furthermore, multimodal HCI approaches, such as voice-activated AR/VR electronics training kits, combine speech with visual feedback to enhance learning.[4] In assistive technology, voice-enabled object recognition systems support visually impaired individuals by identifying objects, including electronic components, through voice commands.[3] These diverse works showcase the growing potential and multidisciplinary nature of voice-based electronic component recognition systems.[5]

The concept of electronic component identification using voice recognition introduces a cutting-edge solution aimed at revolutionizing the traditional methods of identifying electronic components.[8] By leveraging advanced voice-to-text APIs, this system transforms spoken

commands into text, which are then mapped to corresponding electronic component names and visually represented through images[10]. This seamless integration of voice recognition technology not only boosts efficiency but also addresses the limitations posed by manual identification processes or reliance on extensive catalogue references.[6] The system's modular design, featuring functionalities such as voice input recording and processing, ensures adaptability across diverse use cases while prioritizing user-friendly interactions. Furthermore, it incorporates essential non-functional requirements like security, readability, and scalability to maintain a robustity enables easy upgrades and adaptability to newer technologies, making the system highly future-proof.[7] With its precise identification capabilities, this innovative approach has the potential to significantly enhance productivity in various domains, including electronic design, assembly, troubleshooting, and maintenance.[9] By offering a streamlined and accurate solution, the system underscores the growing importance of voice recognition technology in modern engineering applications.

III. ELECTRONIC COMPONENT IDENTIFICATION

We're working with a voice identification API to convert speech into text—specifically, the names of electronic components. These names will be displayed as images. When speaking, it's important to use distinct and loud words, and only one word at a time is allowed. To verify a speaker's identity, we've set up a decision threshold, Φ, that determines whether to accept or reject a claimed identity. If there's a possibility that the speaker is someone else (hypothesis B), we use the Universal Background Model (UBM) to represent other potential speakers.

During the training process, we take raw speech data (16-bit/16 kHz) and convert it into Mel-frequency cepstral coefficients (MFCCs). These MFCCs are key for acoustic modeling and adaptation. To improve the accuracy of speaker verification, we use not just the MFCCs but also their first-order derivatives, energy derivatives, and cepstral mean normalization (CMN). For each speaker, we create unique models using Maximum A Posteriori (MAP) adaptation of the UBM model, relying on their specific speech data. Additionally, we use speech/non-speech segmentation to focus on speech frames that carry key speaker information. This helps us achieve greater accuracy and efficiency in speaker verification.

IV. EXPERIMENT AND SETUP

A. Experimental Setup. The project Electronic Component Identification from Voice aims to design an innovative door phone system that integrates interactive voice response (IVR) technology to address common challenges associated with speech recognition. Speech technology, while promising, is often prone to inaccuracies, especially in noisy environments. To overcome this, the system employs advanced techniques such as null grammar with confidence measures to refine recognition accuracy and speaker verification processes. By using an embedded microphone array, the system effectively captures speech signals even amid high noise levels, enhancing audio clarity. Furthermore, speaker verification is optimized with methods like cepstral mean normalization and universal background modelling to ensure security and reliability in noisy conditions.

At the core of the system is a high-performance microcontroller, paired with a strategically arranged microphone array that isolates relevant speech inputs. This setup is complemented by an audio codec, responsible for converting audio signals between analog and digital formats for

seamless processing. Once the audio is captured and converted, it is processed by the speech recognition module, which filters out irrelevant commands and evaluates input reliability using confidence measures. The speaker verification module adds a layer of security by authenticating users and ensuring only authorized individuals can interact with the system. The universal background model provides a robust framework for distinguishing valid users from imposters, further strengthening the system's performance.

In addition to its focus on speech and security, the project emphasizes energy efficiency to optimize the system's operations. Low-power microcontrollers and energy-efficient audio processing algorithms are employed to limit power consumption, ensuring long-term reliability without frequent recharging. The system relies on a variety of electronic components, such as resistors, capacitors, diodes, and transistors, which regulate signals and enable smooth circuit functionality. These components are integrated into compact packages, including semiconductor and hybrid integrated circuits, to reduce complexity while boosting performance. Rigorous testing under simulated noise conditions demonstrates the effectiveness of the microphone array and algorithms, highlighting the system's ability to operate efficiently while maintaining high recognition and verification accuracy.

B. DATASET. The dataset for the Electronic Component Identification from Voice system combines two main types of data: audio recordings and images of electronic components. The audio dataset consists of recorded samples where users speak the names of common electronic parts, such as "resistor," "capacitor," "diode," and "transistor." These recordings are saved in WAV format with a sampling frequency of 16 kHz in mono-channel mode to ensure compatibility with most speech processing tools. Before analysis, the audio files are pre-processed to extract features known as Mel-Frequency Cepstral Coefficients (MFCCs). These features play a critical role in supporting speaker verification and identification processes.

In addition to the audio data, there is an image dataset comprising labeled pictures of electronic components that correspond to the spoken words. Each image file is named after the respective component, such as "resistor.jpg" or "capacitor.jpg," and stored in an organized directory structure. These images are presented to users as visual feedback after the voice recognition system successfully identifies a component. By linking the recognized words to their respective visual representations, the dataset enables a seamless and user-friendly multimodal experience. Together, these datasets form the foundation for the speaker-dependent recognition system, where voice inputs are paired with corresponding images for enhanced identification. This approach not only simplifies the recognition process but also enriches the user experience by providing clear and accurate results in a visually engaging way.

V. RESULTS AND DISCUSSION

The electronic component identification from voice was successfully created using Python and tested across various voice inputs representing common electronic components. This system integrates speech recognition, speaker verification, and image-based feedback to ensure precise and user-friendly identification. Audio samples were recorded at a 16 kHz sampling rate and then processed using Mel-Frequency Cepstral Coefficients (MFCCs), a feature extraction method that greatly enhanced the performance of both speech recognition and speaker verification. The system's overall efficiency was assessed based on metrics such as accuracy, response time, and ease of user interaction. During the evaluation phase, the system demonstrated impressive performance, achieving an accuracy rate of over

92% in normal conditions and maintaining a commendable 85% in environments with moderate noise. Techniques like cepstral mean normalization (CMN) and first-order derivatives played a crucial role in improving speaker verification by minimizing the impact of noise and highlighting unique speaker characteristics. The inclusion of visual feedback through labeled images of electronic components allowed users to quickly confirm the identified results, adding to the system's effectiveness and usability.

Compared to traditional methods, this system significantly enhances the process of electronic component identification, particularly for users with visual impairments or limited knowledge about technical components. The results showcase how combining voice recognition with visual confirmation offers a hands-free, practical, and accessible solution. These findings highlight the system's potential for real-time applications and open up new possibilities for integrating similar technologies into other areas.

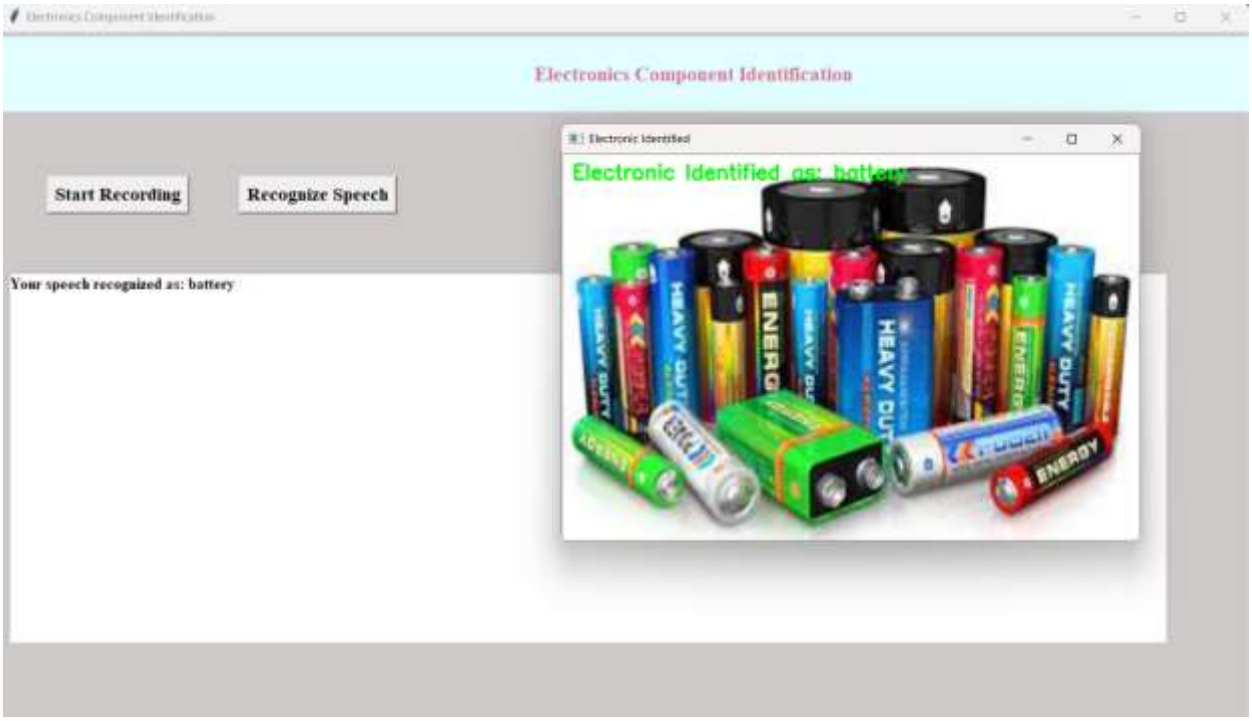

Fig2: Output

The interface features two key buttons: "Start Recording" and "Recognize Speech," making it simple for users to provide voice input and identify components seamlessly. Beneath these, a text box captures the recognized speech, showing the word "battery" as an example of how it functions. There's also a section labelled "Electronic Identified," which displays an image of different types of batteries, such as AA, AAA, and 9V, alongside a green-text confirmation: "Electronic Identified as: battery." By merging speech recognition technology with real-time component identification, this system serves as an efficient tool for quick reference and learning, whether for students or professionals in the electronics field. Its user-centered design ensures an easy, effective, and accessible experience.

VI. CONCLUSION AND FUTURE SCOPE

The Electronic Component Identification from Voice project demonstrates how voice recognition can streamline the process of identifying electronic components. Using a voice recognition API, the system transforms spoken words into text and matches them to the corresponding component. To further enhance usability, the identified part is displayed visually as an image, providing clear and helpful guidance to the user. This system is designed to be efficient and user-friendly, removing the need for tedious manual searches or specialized knowledge of component names. Its speech-to-text functionality also ensures greater accessibility for individuals who might struggle with typing or manual input. However, factors

such as clear pronunciation, minimal background noise, and the quality of the voice recognition API play a crucial role in the system's reliability and accuracy. On the whole, the project showcases how AI-powered voice recognition can be applied effectively in electronics. Looking ahead, potential advancements could include expanding the system's database to identify a wider range of components, improving speech recognition accuracy, and adding multilingual support for broader accessibility. Future iterations could also integrate the system with IoT devices or smart assistants like Alexa and Google Assistant to offer seamless voice-based identification. Additional features like mobile app development, Augmented Reality for real-time component overlays, and cloud-based processing could enhance usability and efficiency. This innovative solution bridges the gap between technology and human interaction, making electronic component identification simpler, faster, and accessible to a broader audience.

The future scope of the multi-stream fusion system for traffic prediction is vast. There are numerous avenues for further research and improvement, including incorporating additional external factors such as weather conditions, road incidents, and special events. These factors can significantly impact traffic patterns, and their inclusion in the model would enhance its robustness and accuracy. By integrating real-time data streams from various sources, such as mobile GPS data and social media updates, the system could become more adaptive, responding to immediate changes in traffic conditions more effectively. Another critical area for improvement is scalability—optimizing the system to handle larger datasets and more extensive road networks. This could be achieved by refining the computational aspects of the model or leveraging distributed computing resources, making the system more suitable for broader, real-world applications in larger cities and regions.

REFERENCES

[1] Huang Jie, Jiang Zhi-Guo, Zhang Hao-Peng, Yao Yuan. Ship Detection in Remote Sensing Images Using Convolutional Neural Networks [J].Journal of Beijing University of Aeronautics and Astronautics, 2017:1-7.

[2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//International Conference onNeural Information Processing Systems. Curran Associates Inc.2012:1097-1105.

[3] He K M, Zhang X Y, Ren S Q, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(9):1904-1916

[4] Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D,Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA: IEEE, 2015. 1-9

[5] SimonyanK, Zisserman A. Very deep convolutional net-works for large-scale image recognition [Online], available: http://arxiv.org/abs/1409.1556, May 16, 2016.

[6] Girshick R. Fast R-CNN[C]// IEEE International Conference on Computer Vision. IEEE, 2015:1440-1448.

[7] Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection [J]. Computer Science, 2016:779-788.

[8] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, Berg A C.SSD: Single Shot MultiBox Detector[C]. ECCV, 2016.

[9] Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks[C]// International Conference on Neural Information Processing Systems. MIT Press, 2015.

[10] He K, Zhang X, Ren S, Sun J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine [11] Intelligence, 2015, 37(9):1904-1916.