ISSN 2454-8065 **International Journal of Applied Theoretical Science and Technology**
Volume 18, Issue 12, pp01-3 December 2023

Speech Transcription Using Raspberry Pi

Dr.A.Rosi,Assoc.Prof
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

Dr.Shaik Noor Mohammad,Assoc.Prof
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

J. Maniharika (225U1A0455)
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

A. Saicharan (225U1A0405)
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

A. Anil (225U1A0410)
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

D. Naresh (225U1A0434)
Department of Electronics & Communication Engineering
AVN Institute of Engineering and Technology (AVNIET)
Hyderabad, India

***Abstract*—**Speech-to-text transcription has become a crucial component in human-computer interaction, assisting applications ranging from accessibility tools to intelligent assistants. This paper presents the design and implementation of a low-cost, portable speech transcription system using Raspberry Pi. By integrating open-source speech recognition frameworks, microphone input, and real-time processing, the system enables efficient conversion of speech into text. The proposed solution is lightweight, scalable, and suitable for rural and academic environments where affordability and reliability are key concerns.

***IndexTerms*—**Speech Recognition, Raspberry Pi, Transcription, IoT, Human-Computer Interaction**.**

I. INTRODUCTION

Speech transcription plays a vital role in bridging the gap between spoken language and digital applications. Traditional transcription systems require powerful computing resources or subscription-based cloud services, limiting accessibility in low-resource areas. The Raspberry Pi, being an affordable and versatile microcomputer, provides an excellent platform to develop low-cost transcription systems.

This paper focuses on developing a Raspberry Pi-based speech transcription system that captures voice input through a microphone and processes it using open-source speech-to-text libraries. Such systems can be used in education, healthcare, and public service applications where quick transcription is required.

LITERATURESURVEY

Several works have explored speech-to-text systems, with a focus on accuracy and language support. Google Speech API and Microsoft Azure offer robust solutions, but they rely heavily on cloud infrastructure. Offline solutions such as CMU Sphinx and Vosk have been integrated into Raspberry Pi with moderate accuracy. Recent research emphasizes the need for lightweight, localized systems that can function without continuous internet connectivity.

Table I presents a comparison of existing transcription systems:

Table I. Comparison of Speech Transcription Approaches

Approach	Platform	Connectivity	Cost	Accuracy
Google Speech API	Cloud-based	Internet	High	High

ISSN 2454-8065 International Journal of Applied Theoretical Science and Technology
Volume 18, Issue 12, pp01-3 December 2023

Microsoft Azure STT	Cloud-based	Internet	High	High
CMU Sphinx	Local	Offline	Low	Medium
Vosk API	Local + Cloud	Optional	Low	Medium
Proposed Raspberry Pi	Raspberry Pi 4	Optional	Very Low	Medium+

PROPOSEDSYSTEM

The system architecture is shown in Fig. 1. A microphone captures speech signals, which are processed by Raspberry Pi. The audio input is converted into text using open-source libraries such as Vosk or CMU Sphinx. The transcription results are displayed on a connected screen or stored in a text file for later use.

Fig. 1. Block Diagram of Speech Transcription System Using Raspberry Pi

The main modules of the proposed system include:

- **Input Module:** Microphone for capturing speech.
- **Processing Unit:** Raspberry Pi 4 with Raspbian OS.
- **Speech-to-Text Engine:**Vosk API or CMU Sphinx for offline recognition.
- **Output Module:** Display/Storage of transcribed text.

The system is designed to function both online (with internet access for cloud APIs) and offline, ensuring versatility in usage.

RESULTS AND DISCUSSION

The prototype was tested with different accents and background noise conditions. Results indicated that the system achieved an average transcription accuracy of 82% in a quiet environment and 70% in a noisy environment. The processing delay was minimal, with an average transcription latency of less than 1.2 seconds.

Compared to cloud-based solutions, the proposed system offered reduced costs and independence from constant internet access. However, accuracy was slightly lower than advanced cloud APIs.

CONCLUSION AND FUTURE SCOPE

This work demonstrates the feasibility of implementing a cost-effective speech transcription system using Raspberry Pi. The system provides satisfactory performance in offline environments and can be deployed in schools, healthcare centers, and public offices to improve accessibility.

Future improvements may include integrating noise cancellation algorithms, supporting multiple languages, and enhancing accuracy through machine learning-based models optimized for Raspberry Pi.

FIG1: RESULT

ACKNOWLEDGMENT

The authors express their gratitude to Mrs. M. Arpana (Coordinator) and Ms. K. Soumya (Internal Guide) for their continuous support and guidance throughout the project. They also extend their heartfelt thanks to the Department of Electronics & Communication Engineering, AVNIET, for providing necessary laboratory facilities and resources. Special thanks are due to fellow classmates and peers who contributed through valuable discussions and feedback during the development of the system. Finally, the authors acknowledge their families for encouragement and moral support throughout the completion of this work.

ISSN 2454-8065 **International Journal of Applied Theoretical Science and Technology**
Volume 18, Issue 12, pp01-3 December 2023

REFERENCES

[1] P. Lamere, "CMU Sphinx: Open source speech recognition toolkit," *CMU Sphinx Documentation*, 2018.
[2] A. Ghoshal, P. Swietojanski, and S. Renals, "Multilingual training of deep neural networks," *ICASSP*, pp. 7319–7323, 2013.
[3] A. Stolcke, "SRILM – An Extensible Language Modeling Toolkit," *Proceedings of ICSLP*, 2002.
[4] Vosk Speech Recognition Toolkit. [Online]. Available: https://alphacephei.com/vosk
[5] T. Schultz and A. Waibel, "Language-independent and language-adaptive acoustic modeling for speech recognition," *Speech Communication*, vol. 35, no. 1–2, pp. 31–51, Aug. 2001.
[6] K. Audhkhasi et al., "Building Speech Recognition Systems from Scratch Using Deep Neural Networks," *IEEE Signal Processing Letters*, vol. 21, no. 12, pp. 1596–1600, Dec. 2014.
[7] Raspberry Pi Foundation, "Raspberry Pi Documentation." [Online]. Available: https://www.raspberrypi.org/documentation/