

# SpecAssist: Smart Glasses with Real-time speech recognition and transcription for the hearing impaired

Noel Jacob

*Dept. of Computer Science and Engineering  
St. Joseph's College of Engineering and Technology  
Palai, Kottayam, Kerala*

Sathwik P Nair

*Dept. of Computer Science and Engineering  
St. Joseph's College of Engineering and Technology  
Palai, Kottayam, Kerala*

Rajat Mathew

*Dept. of Computer Science and Engineering  
St. Joseph's College of Engineering and Technology  
Palai, Kottayam, Kerala*

Prof. Divya Sunny

*Dept. of Computer Science and Engineering  
St. Joseph's College of Engineering and Technology  
Palai, Kottayam, Kerala*

**Abstract**—SpecAssist is a wearable device, (smart glasses) that enable real-time translation and transcription of speech. The device consists of a head-mounted display that shows the translated text, a microphone to capture the audio, and a translation engine to perform the conversion. It connects to the internet to access the translation engine and can be controlled/configured using a mobile app. SpecAssist allow users to communicate with people who speak different languages without the need for a separate translation device or service. It offers a compact and portable solution for real-time language translation and can improve accessibility for people who are deaf or hard of hearing. SpecAssist harnesses the power of Machine Learning, Automatic Speech Recognition (ASR), Natural Language Processing, and IoT to provide a very low-latency solution for the hearing impaired.

**Index Terms**—Smart Glasses, Automatic Speech Recognition, Machine learning

## I. INTRODUCTION

Hearing loss is a pervasive issue that affects millions of people worldwide, and it can have a significant impact on their quality of life. Communication is a fundamental aspect of human interaction, and those with hearing loss often struggle to participate fully in conversations.

Many proposed solutions to this problem require the deaf or hard of hearing person to divert their attention away from the speaker, which can result in missed nonverbal cues and expressions. However, recent advancements in Automatic Speech Recognition (ASR) technology have opened up new possibilities for assistive wearable devices that can help bridge the communication gap. This paper explores the use of modern neural network models in an assistive wearable device designed specifically for people with hearing disabilities. By leveraging the power of Automatic Speech Recognition, this technology has the potential to revolutionize the way people

with hearing loss interact with the world around them, enabling them to participate more fully in conversations and improve their overall quality of life.

## II. OBJECTIVE AND SCOPE

The objective of SpecAssist is to provide a portable and easy-to-use assistive device with real-time speech recognition and transcription which can provide numerous benefits for individuals who are hearing impaired. Some of these benefits include:

**Improved communication:** By providing a visual representation of spoken language, these smart glasses can improve communication for individuals who are hearing impaired. This can be particularly useful in social situations, where it can be difficult to follow a conversation without being able to hear what is being said.

**Enhanced access to information:** These smart glasses can also improve access to information for individuals who are hearing impaired. For example, they can provide real-time transcriptions of lectures, meetings, and other educational or professional settings, allowing users to stay up to date with the latest information.

**Increased independence:** By providing a visual representation of spoken language, these smart glasses can help individuals who are hearing impaired become more independent and self-sufficient. They can participate more fully in social and recreational activities, such as live performances or sporting events, without needing to rely on a sign language interpreter or other support.

**Improved language learning:** These smart glasses can also be used to assist with language learning by providing real-time translations of spoken language. This can be particularly

helpful for individuals who are learning a new language and need to understand spoken language in order to progress.

SpecAssist is a solution for the aforementioned issues and more. Our ultimate goal is to offer a portable and affordable device, which is crucial in the fast-paced field of smart glasses and augmented reality. As this technology continues to advance, it is expected that these glasses will become increasingly widespread and beneficial for a diverse array of applications.

### III. LITERATURE REVIEW

#### A. INTRODUCTION

In this literature review, a total of 20 papers were studied to understand the current state of the technology and its potential applications. The papers discussed various models of smart glasses available in the market, such as Google Glass, Epson Moverio, and Vuzix Blade. One of the main applications of smart glasses is in the field of Augmented Reality, where they can display virtual images, text, and videos superimposed over the real world. This can be used for entertainment, training, or to enhance the user's experience of the physical world. For example, smart glasses can be used by tourists to learn more about the places they are visiting, or by construction workers to view instructions and diagrams while working on a project. Smart glasses can also be used for VR, which involves fully immersing the user in a computer-generated environment. This can be used for gaming, training, or other immersive experiences.

In addition to these consumer applications, smart glasses have the potential to revolutionize various industries, such as healthcare, retail, and manufacturing. For example, healthcare professionals can use smart glasses to access patient information and perform procedures more efficiently, and retail employees can use them to assist customers in real-time.[4] Despite the many potential benefits of smart glasses, there are also some challenges and limitations to consider. One of the main challenges is the cost and availability of the technology, as many models are still expensive and not widely available. In addition, there are concerns about privacy and security, as the devices can potentially record and transmit sensitive information.

Smart glasses are a promising and rapidly developing technology that has the potential to transform a wide range of industries and applications. However, further research and development is needed to address the challenges and limitations of the technology and to make it more widely available and accessible to the general public.

#### B. APPLICATIONS

There were many real-life and existing applications discussed in the papers. Some of these are :

Augmented reality (AR): Smart glasses can display AR content such as virtual images, text, and videos superimposed over

the real world. This can be used for training, entertainment, or to enhance the user's experience of the physical world.

Virtual reality (VR): Smart glasses can also be used to create immersive VR experiences, in which the user is fully immersed in a computer-generated environment.

Industrial and military training: Smart glasses can be used to provide hands-free training and guidance in industrial and military settings.

Medicine and healthcare: Smart glasses can be used by health-care professionals to access patient information and perform procedures more efficiently.

Retail and customer service: Smart glasses can be used by retail employees to access product information and assist customers in real-time.

Education and training: Smart glasses can be used in education and training settings to provide interactive and immersive experiences.

Gaming: Smart glasses can be used for gaming and other immersive entertainment experiences.

Navigation and transportation: Smart glasses can provide turn-by-turn navigation instructions and information about the surrounding environment, making them useful for transportation and logistics applications.

#### C. CHALLENGES

Building a speech transcribing smart glass is a complex task that involves a range of technical and practical challenges. Some of the main challenges faced in building such a device include:

Accurate speech recognition: This involves developing algorithms that can accurately transcribe spoken words into written text, while also handling different accents, dialects, and languages.

Microphone design: The microphone of a smart glass needs to be able to capture and transmit spoken words accurately and clearly, even in noisy environments. This requires careful design and engineering to ensure that the components are of high quality and perform well.

Display technology: The display of a smart glass needs to be bright and clear, with a wide viewing angle, in order to display information effectively. This requires selecting the appropriate display technology and optimizing the design of the display

module.

**Power and battery life:** Another challenge is ensuring that the smart glass has sufficient power and battery life to operate continuously throughout the day. This requires designing energy-efficient hardware and software, as well as implementing power-saving features.

**Comfort and ergonomics:** Smart glasses need to be comfortable and ergonomic to wear for extended periods of time, which can be a challenge due to the size and weight of the components.

**Cost and availability:** Finally, the cost and availability of the technology is a challenge that must be overcome in order to make speech transcribing smart glasses widely available.

#### IV. FUTURE SCOPE

Smart glasses are a rapidly evolving technology with a wide range of potential applications. Some possible future developments and scopes for smart glasses include:

**Improved design and comfort:** Smart glasses are currently limited by their size and weight, which can make them uncomfortable to wear for extended periods of time. In the future, it is likely that the design of smart glasses will become more compact and lightweight, making them more comfortable to wear.

**Greater accuracy and precision:** Smart glasses have the potential to provide precise and accurate information, but this is currently limited by the accuracy of the sensors and other components. In the future, it is likely that smart glasses will become more precise and accurate, enabling them to be used for even more applications.

**Wider adoption and usage:** Smart glasses are currently used by a relatively small number of people, but in the future it is likely that their adoption will become widespread. This could be due to improvements in technology, as well as increased awareness and acceptance of the devices.

**Increased functionality and capabilities:** As the technology behind smart glasses continues to improve, it is likely that they will become more capable and offer a wider range of features and functionality. This could include the ability to run more complex applications and access a wider range of data and information.

**Greater integration with other technologies:** Smart glasses have the potential to be integrated with a wide range of other technologies, such as Artificial Intelligence (AI), Internet of Things (IoT), and robotics. This could enable them to perform

more complex tasks and interact with the environment in even more sophisticated ways.

#### V. PROPOSED METHOD

The proposed system, SpecAssist is a pair of smart glasses that has 3 modules working together to provide a seamless experience. The solution uses speech recognition and transcription that allows the user to see the transcribed conversation on the user's field of vision. The proposed system would be equipped with a microphone to capture audio from the user's surroundings. This audio would be transmitted to a pre-trained model that is responsible for performing speech recognition and transcription. This model needs to be trained on a large dataset of audio and corresponding transcriptions in order to accurately recognize and transcribe the user's speech. The transcribed text would be displayed on the smart glasses' display, or it could be transmitted to another device, such as a smartphone or laptop, for display. The smart glasses could also be equipped with additional sensors and software to enable additional features, such as voice commands. This is just one possible approach to implementing a system for smart glasses with real-time speech recognition and transcription. The 3 modules present in the device are: Smartglasses, Android Application, and a Backend.

##### A. Smartglasses module

The smartglasses module is a pair of glasses worn by the user. This module contains all the hardware components of this project. It comprises an ESP-32 microcontroller board, Microphone module, 3.7v Li-Po battery, and a Transparent OLED display. This module performs the vital functionalities of the device such as receiving voice input, converting it into digital format, transmitting it to a mobile device, and displaying the text data in the lens of the glasses.

##### B. Mobile Application Module

The mobile application serves as a middleman for the smartglasses module for it to interact with the backend. The main purpose of this application is to facilitate communication with the cloud. The application also contains configurable features and parameters including connection to the wearable device, language, power and display features like font size and brightness.

##### C. Backend Module

The data sent over by the mobile device is received by the backend, fed to the pre-trained NLP model which converts it

into text data which is then sent back to the mobile device. FastAPI, which is a popular python based framework, will be used for building the backend.

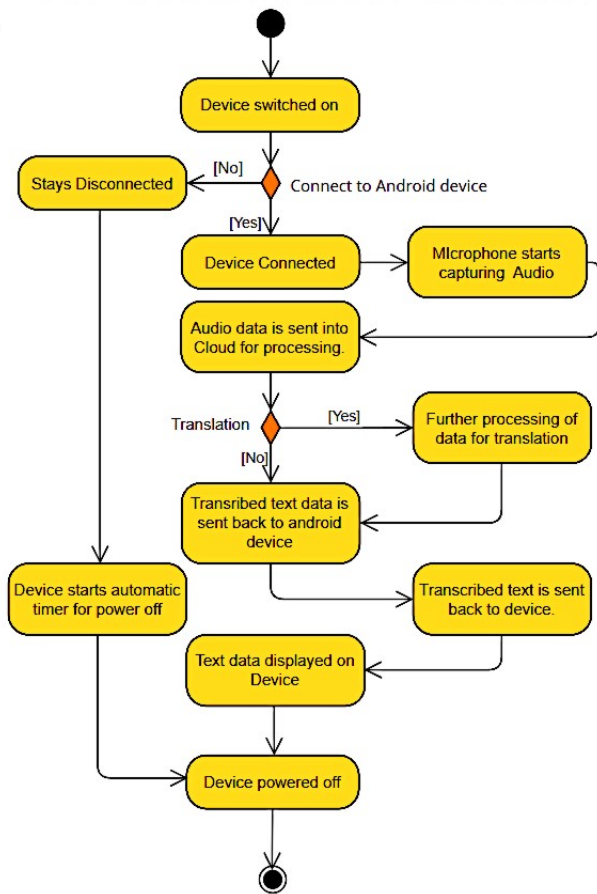


Fig. 1. Activity diagram of proposed system

## REFERENCES

- [1] Lee, Lik-Hang Hui, Pan., *Interaction Methods for Smart Glasses: A Survey*, IEEE Access
- [2] H. AlSaid, L. AlKhatib, A. AlOraidh, S. AlHaidar and A. Bashar, *Deep Learning Assisted Smart Glasses as Educational Aid for Visually Challenged Students*. 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), 2019
- [3] A. Aiordăchioae, O. -A. Schipor and R. -D. Vatavu, *An Inventory of Voice Input Commands for Users with Visual Impairments and Assistive Smartglasses Applications*, 2020 International Conference on Development and Application Systems (DAS), 2020
- [4] O. -A. Schipor and A. Aiordăchioae, *Engineering Details of a Smartglasses Application for Users with Visual Impairments*, 2020 International Conference on Development and Application Systems (DAS), 2020,
- [5] E. Machado, I. Carrillo, D. Saldana, F. Chen and L. Chen, *An Assistive Augmented Reality-based Smartglasses Solution for Individuals with Autism Spectrum Disorder*, 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing
- [6] A. Pražák, J. V. Psutka, J. Psutka and Z. Loose, *Towards live subtitling of TV commentary*, 2013 International Conference on Signal Processing and Multimedia Applications (SIGMAP), 2013

- [7] M. H. Moattar, M. M. Homayounpour and N. Khademi Kalantari, *A new approach for robust realtime Voice Activity Detection using spectral pattern*, 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, 2010
- [8] K. Tada, K. Kutsuzawa, D. Owaki and M. Hayashibe, *Quantifying Motor and Cognitive Function of the Upper Limb Using Mixed Reality Smartglasses*, 2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), 2022,
- [9] S. K. A. X. K, D. Davis and N. Jayapandian, *Internet of Things and Cloud Computing Involvement Microsoft Azure Platform*, 2022 International Conference on Edge Computing and Applications (ICECAA), 2022
- [10] C. -H. Lai and Y. -S. Hwang, *The voice controlled Internet of Things system*, 2018 7th International Symposium on Next Generation Electronics (ISNE), 2018
- [11] B. Sudharsan et al., *OTA-TinyML: Over the Air Deployment of TinyML Models and Execution on IoT Devices* in IEEE Internet Computing
- [12] M. Shafique, T. Theocharides, V. J. Reddy and B. Murmann, *TinyML: Current Progress, Research Challenges, and Future Roadmap*, 2021 58th ACM/IEEE Design Automation Conference (DAC), 2021
- [13] G. Uddin, *Security and Machine Learning Adoption in IoT: A Preliminary Study of IoT Developer Discussions*, 2021 IEEE/ACM 3rd International Workshop on Software Engineering Research and Practices for the IoT (SERP4IoT), 2021
- [14] J. Jiménez, A. M. Iglesias, J. F. López, J. Hernández and B. Ruiz, *Tablet PC and Head Mounted Display for live closed captioning in education*, 2011 IEEE International Conference on Consumer Electronics (ICCE), 2011
- [15] Y. Kim, S. Han, S. Choi and B. Jung, *File-Based Closed Captioning System without Captioning Delay*, SMPTE 2015 Annual Technical Conference and Exhibition, 2015
- [16] Si-Hun Sung and Woo-Sung Chun, *Knowledge-based numeric open caption recognition for live sportscast*, 2002 International Conference on Pattern Recognition, 2002
- [17] A. R. Biswas and R. Giaffreda, *IoT and cloud convergence: Opportunities and challenges*, 2014 IEEE World Forum on Internet of Things (WF-IoT), 2014
- [18] S. Naveen and M. R. Kounte, *Key Technologies and challenges in IoT Edge Computing*, 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2019
- [19] H. -J. Jeong, H. -J. Lee and S. -M. Moon, *Work-in-progress: cloud-based machine learning for IoT devices with better privacy*, 2017 International Conference on Embedded Software (EMSOFT), 2017
- [20] F. Samie, L. Bauer and J. Henkel, *From Cloud Down to Things: An Overview of Machine Learning in Internet of Things*, in IEEE Internet of Things Journal, 2019