

## **Image to Audio Conversion for Impaired People**

**M. Jaswanth, N. Lokesh, K. Lakshmi Narasimha , K. Mounika Sai**

**Mr. N.Gopala Krishna** (*Associate Professor*)

*Department Of Computer Science and Engineering,*

*Tirumala Engineering College, Narasaraopet, Andhra Pradesh*

### **ABSTRACT**

*Blind people and low vision people are known to suffer from many other problems. Some of the problems they face in their daily activities are reading, identifying the objects, and especially they are facing the problem to identify their medicine while they are buying from the medical stores. For visually impaired people would to read a text, Optical Characters Recognition devices exist that allow them to scan the text line by line and either convert it to braille. This is not very convenient. There arises the need for some application to help the visually impaired people and old people to detect the text in order to identify the text and be able to read other written work with the help of audio. Thus, we design a system that helps the visually impaired people to upload the image and convert the detected text into voice message.*

**Keywords:** Digitization, Conversion of hardcopy of text to speech, Synthetic speech signals, Language translation

### **1.1 INTRODUCTION**

The Image to Sound Conversion for Impaired People project aims to help visually impaired individuals and older adults to identify text in images by converting the detected text into an audio message. This system will utilize Optical Character Recognition (OCR) technology to scan the text in the image and then Text-to-Speech (TTS) technology to convert it into an audio message. The project is designed to assist people with low vision or blindness to read and comprehend written materials, particularly medication labels, which are crucial to their daily lives. The project is a valuable tool that will increase accessibility and independence for individuals

### **1.2 AIM OF THE PROJECT**

The Image to Sound Conversion for Impaired People project aims to help visually impaired individuals and older adults to identify text in images by converting the detected text into an audio

message. This system will utilize Optical Character Recognition (OCR) technology to scan the text in the image and then Text-to-Speech (TTS) technology to convert it into an audio message. The project is designed to assist people with low vision or blindness to read and comprehend written materials, particularly medication labels, which are crucial to their daily lives. The project is a valuable tool that will increase accessibility and independence for individuals with visual impairments.

### **1.3 PROJECT DOMAIN**

The project domain for Image to Sound Conversion for Impaired People using TTS and OCR technologies falls under the field of assistive technology for people with visual impairments. Specifically, the project utilizes OCR and TTS technologies to assist visually impaired individuals and older adults with text identification and comprehension, which is an essential aspect of their daily lives. The project is designed to increase accessibility and independence for individuals with visual impairments, and falls under the domain of accessibility and assistive technology.

### **1.4 SCOPE OF THE PROJECT**

There is scope to add more functionality to the present application. One of the addons can include support for languages other than English. Algorithms used to preprocess a natural image works sufficiently in this system. However, there is a scope to improve it further.as well as the OCR techniques. Also, adding support for more formats of image inputs can be provided. Algorithms can be developed to recognize text from low resolution and blur images. This will help users to upload old historic manuscripts and scrolls which have been damaged and extract text from it.

**LITERATURE SURVEY [1]** K. C. SHAHIRA, “Towards Assisting the Visually Impaired: A

Review on

Techniques for Decoding the Visual Data From Chart

Images,” IEEE Access,

Volume 9, 2021

The objective of the task is to develop a system that is both practical and accurate. The schedule is also quite wide because it involves both audio and text translation. This paper presents a method for defining the area of interest (ROI) in which objects can be segregated from cluttered backgrounds that is both efficient and effective. To obtain text information, this ROI extracts the text localization and recognition.

**[2]** Sai Aishwarya Edupuganti; Vijaya Durga Koganti; Cheekati Sri Lakshmi;



Ravuri Naveen Kumar, “Text and Speech Recognition for Visually Impaired People using Google Vision,” 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021

A printed copy of a book is the usual means of reading it. Books came in handy because it was impossible to carry a physical copy of a book everywhere. Both printed and electronic books have an impact on the eyes. Visual impairment is one of humanity’s most significant constraints, particularly in this day and age where information is obtained by reading a large number of books (electronic and paper based).

[3] Asha G. Hagargund, Sharsha Vanria Thota, Mitadru

Bera, Eram Fatima Shaik “Image to speech conversion for visually impaired,” International Research Journal of Engineering and Technology (IRJET), Volume 03, Issue, 2017 The study is based on the Android platform, as well as machine learning. For blind persons, reading text from text pictures and text boards is a difficult effort. In a time when more and more communication takes the shape of written or digital text rather than spoken word, the inability to see is one of mankind’s most severe handicaps. [4] Samruddhi Deshpande; Revati Shriram, “Real time text detection and recognition on hand held objects to assist blind people,” 2016 International Conference on Automatic Control and Dynamic

Optimization Techniques (ICAC-DOT), 2016

According to surveys, many find living in today’s world to be extremely difficult. The writers have created a new device called the Blind Book Reader System to assist them. When blind people are in need, it’s an easy remedy to get them to read books and newspapers.

[5] Prabhakar Manage; Veeresh Ambe; Prayag

Gokhale; Vaishnavi Patil, “An Intelligent Text Reader based on Python,” 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020

In this study, we describe a photograph (I2AD) task for creating audio descriptions from photographs, which can be used to enhance the visual experience of blind and visually impaired individuals.

Proposed System

CNN LSTM is a hybrid deep learning model that combines the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. Here are some advantages of using CNN LSTM for image to sound conversion for impaired people using TTS and OCR technologies:



**Effective Image Processing:** CNNs are effective in image processing tasks such as object recognition and segmentation, making them well-suited for analyzing visual data. By using CNNs as a pre-processing step, CNN LSTM can extract meaningful features from the input images and enhance the accuracy of the final output.

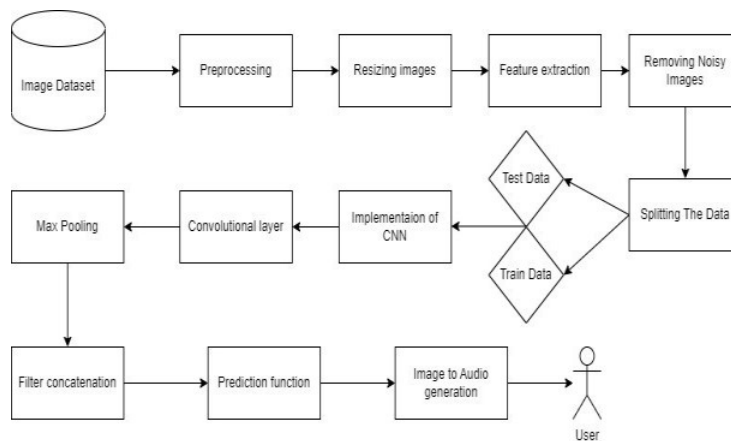
**Memory Capabilities:** LSTM networks are designed to handle sequential data and have memory capabilities, making them well-suited for modeling the temporal nature of speech. By using LSTM networks in conjunction with CNNs, CNN LSTM can learn to recognize and generate more complex patterns in the quality Sound Generation: By using a combination of CNNs and LSTMs, CNN LSTM can generate highquality sounds that accurately represent the visual information in the input images. This can be especially important in the context of image to sound conversion for impaired people, where clear and accurate sounds are crucial for conveying important information.

**Accessibility:** CNN LSTM can be designed to be accessible to a wide range of users, including those with limited language abilities or cognitive impairments. By using visual information to generate sounds, CNN LSTM can make information more accessible and easier to understand for people with different types of impairments. **Scalability:** CNN LSTM can scale to handle large amounts of data and can be trained on large datasets to improve its accuracy. This can be important in the context of image to sound conversion, where the algorithm needs to be able to process large amounts of visual data in real-time.

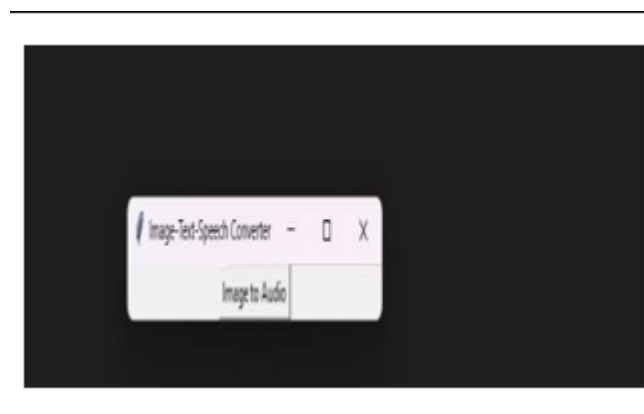
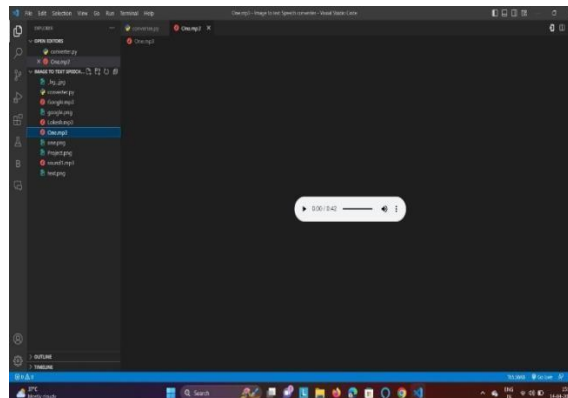
Therefore, CNN LSTM has several advantages for image to sound conversion for impaired people using TTS and OCR technologies. It can effectively process visual data, generate high-quality sounds, be designed for accessibility, and scale to handle large amounts of data.

#### Architecture

The Architecture Diagram for Image to Sound Conversion for Impaired People using TTS and OCR technologies depicts the high-level structure of the system, including the components and their relationships. It shows how the OCR and TTS technologies are integrated into the system and how the user interacts with the system to upload an image and based information.



**OUTPUT SCREENS**





## **CONCLUSION**

In conclusion, the Image to Sound Conversion for Impaired People using TTS and OCR technologies is an important project that aims to assist visually impaired individuals in reading and understanding text. The proposed system utilizes OCR technology to extract text from images and TTS technology to convert the extracted text into an audio message, making it accessible to visually impaired individuals. The system provides an easy-to-use and accessible interface for uploading images and receiving audio output, improving the user experience for visually impaired individuals. The proposed system offers improvements in accuracy, processing time, user interface, feature set, and scalability compared to existing systems. However, there are still opportunities for further research and development to improve the accuracy and efficiency of the OCR and TTS technologies used in the system. Overall, the Image to Sound Conversion for Impaired People using TTS and OCR technologies is an important step towards providing equal opportunities and accessibility for visually impaired individuals in accessing and understanding text-based information.

## **REFERENCES**

- [1]. Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik "Image to speech conversion for visually impaired," International Research Journal of Engineering and Technology (IRJET), Volume 03, Issue, 2017
- [2]. Samruddhi Deshpande; Revati Shriram, "Real time text detection and recognition on hand held objects to assist blind people," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), 2016
- [3]. Prabhakar Manage; Veeresh Ambe; Prayag Gokhale;
- [4]. Vaishnavi Patil, "An Intelligent Text Reader based on
- [5]. Python," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020
- [6]. D.Velmurugan, M.S.Sonam, S.Umamaheswari,
- [7]. S.Parthasarathy, K.R.Arun[2016]. A Smart Reader for Visually Impaired People Using Raspberry PI. International Journal of Engineering Science and Computing IJESC Volume 6 Issue No. 3.
- [8]. K Nirmala Kumari, Meghana Reddy J [2016]. ImageText to Speech Conversion Using



- [9]. OCR Technique in Raspberry Pi. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 5, Issue 5, May 2016.
- [10]. Silvio Ferreira, Celine Thillou, Bernard Gosselin. From Picture to Speech: An Innovative Application for
- [11]. Embedded Environment. Faculté Polyte