IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ASSISTING VISUALLY IMPAIRED PEOPLE

Mr.ARAVINDHRAJA. R
Department of Computer Science
SRM Valliammai Engineering College
Chennai,

Mr.DEEPAK RISHIKESH. R
Department of Computer Science
SRM Valliammai Engineering College
Chennai,

Mr.HARISH KUMAR. P
Department of Computer Science
SRM Valliammai Engineering College
Chennai,

Abstract

Our eyes are the most essential sensory organ. Because most tasks require visual information, visually impaired people are at a disadvantage because crucial information about their surroundings is unavailable. It is now possible to extend the support provided to people with visual impairments with advancements in technology. This project proposes to use Artificial Intelligence, Machine Learning, Image and Text Recognition to assist persons who are blind or visually impaired. The habit of reading is slowly diminishing because of the busy lives of people.The situation is significantly worse for visually impaired people. To deal with this issue, we have developed a text summarisation technique. A camera device is added to the system to scan texts. The summarized output generated is read out to the user using text to speech conversion technique. This project also focuses on helping visually impaired people by detecting obstacles on their way and alerting them.

Keywords: Deep learning, OpenCV, OCR, SequencetoSequence, NLP

1. INTRODUCTION

Individuals are members of the community, and their lives are inextricably linked to the society's established values and customs. Blind individuals are members of society, and their mobility in the environment and in social situations has been constrained. These individuals rely on assistive aids such as guiding dogs, canes, and others to get from one location to another. For these individuals, there are several navigational issues. As a result, there is a need to design a system for visually impaired persons that would assist them in effectively and accurately navigating from one site to another. With the support of the system, these people would be able to live a self-governing existence. Individuals who are blind will get the same confidence as sighted people as a result of this technique.

Therefore, someone with visual impairments needs assistance by providing substitutes for their eye function, namely the visual function. In addition to the usual aids with the sense of touch system, the blind also needs a switch for their sense of sight, so that the ultrasonic sensor can be used by the user. The ultrasonic sensor informs blind persons whether there are any obstructions in their path, allowing the blind person to avoid an object in front of him. Within a 90-cm range, the sensor recognises a solid item. The sensor is both a transmitter and a receiver of waves. The sent wave would collide with the impediment before being picked up by the sensor. They can also read novels, pamphlets, and notices on the go without the assistance of others with text summation and text to speech conversion.

2. MODULES

2.1 Ultrasonic Sensor

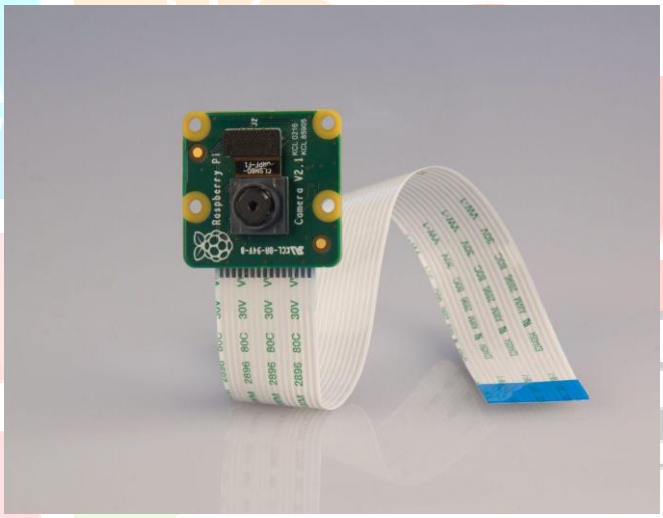
Ultrasonic sensors work by emitting a sound wave that is above the human hearing range. The sensor's transducer functions as a microphone, receiving and transmitting ultrasonic sound. To deliver a pulse and receive the echo, our ultrasonic sensors, like many others, use a single transducer.High-frequency sound waves reverberate off surfaces, creating different echo patterns. It detects the object and alerts the impaired person by creating an alarm sound.

2.2 Raspberry pi

The Raspberry Pi is a low-cost computer with a set of GPIO (general purpose input/output) ports that can be used to operate electronic components for physical computing and to investigate the Internet of Things (IoT). The ultrasonic sensor, beeper are connected to the raspberry pi.

2.3 Camera

High-definition video and still pictures can both be captured with the Camera Module.The Raspberry Pi Board features a CSI (Camera Serial Interface) interface, which we may use to directly connect the PiCamera module.



2.4 OpenCV

OpenCV (Open Source Computer Vision Library) is a free computer vision and machine learning software library.A module for image processing that contains linear and non-linear image filtering, geometric image transformations (resize, affine and perspective warping, general table-based remapping), colour space conversion, histograms, and other features. OpenCV is used to capture an image and convert it into text.

2.5 NLP

Natural language processing (NLP) is a subject of computer science—specifically, a branch of artificial intelligence (AI)—concerning the ability of computers to understand text and spoken words in the same manner that humans can. NLP is used to power computer programmes that translate text from one language to another, respond to spoken commands, and quickly summarize vast amounts of material—even in real time.

2.6 Obstacle Detection using Ultrasonic Sensor

The ultrasonic sensor transmits a sound pulse at high frequency, and then measures the period to obtain the sound echo signal to mirror back. There are 2 circles inside the sensor. One of them transmits the ultrasonic waves and serves as the transmitter. The other each serves as a receiver and collects the repeated sound signal (mostly a small microphone). The sensor is adjusted according to air velocity of the echo. With that measured information, the time difference between sound pulse propagation and detection is determined by calculation of the distance to the target.

2.7 Text to Speech

The portable text to speech converter is meant to assist the visually challenged in hearing any scanned material read aloud.The scanned image may contain text with background pictures which are simply ignored and only the text in the scanned image is extracted by the optical character recognition application. People with visual impairments will benefit greatly from the creation of a text to speech synthesizer, as it will make reading vast amounts of material much easier.

3. PROPOSED SYSTEM

In this work, there are two main parts:

- Optical Character Recognition System for Paper Text
- Text Summarization
- Text to Speech Conversion

3.1 Optical Character Recognition System

There are three sections in this topic, as indicated below:

- Template file Creation
- Creating the Neural Network
- Recognition of characters

3.2 Template file creation

Images of the letters A through Z and the numbers 0 to 9 are collected. Each image is converted into a 5 x 7 character representation in a single vector. These data are recorded as a data file for neural network training.

3.3 Creating the neural network

With 25 hidden neurons, a feedforward neural network is employed to set up for pattern recognition. The network's weights and biases are also initialized when it is created, making it ready for training. The target is set at a value of 0.01 to 0.05. The data file and target file are used to train the Neural Network that has been formed. The neural network must be trained by altering the network's weight and bias until the performance achieves the desired level.

Working

- Webcam captures the image then the image can be read.
- Pre-processing is done in the 2nd step, colour image is converted into grayscale and grayscale is converted into the binary image.
- Characters are extracted and resized.
- Load templates that can be matched.
- Remove the background
- Edge detection is done in the last step of character recognition.

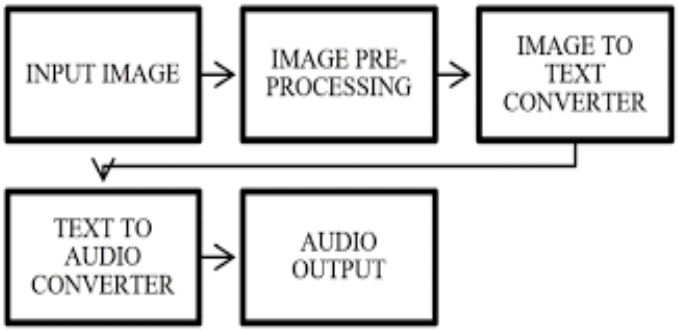


Figure 1.1 Text to speech

3.4 Text Summarizer

The method of reducing a certain set of data to acquire the most significant information is known as automated summary of any text. It is the process of making a brief summary of important facts by writing down the key points in the text. Manual text summarization requires a lot of time, money, and effort. Our text summarization service analyses and summarizes your material using innovative AI-based algorithms. The summary is created by extracting the most important words and phrases from the original text and putting them together. This summarizing approach does not include any rephrasing or the use of synonyms. The words are kept in place and altered slightly to give the sentence structure. It takes the most suitable sentences and creates a faultless summary of the provided content. In our system we have used a sequence-to-sequence model.

Working

1. Importing the Dataset
2. Cleaning the Data
3. Determining the Maximum Permissible Sequence Lengths
4. Selecting Plausible Texts and Summaries
5. Tokenizing the Text
6. Removing Empty Text and Summaries
7. Creating the model
8. Training the model
9. Generating predictions

Tokenization:

Tokenization is the process of breaking down a phrase, sentence, paragraph, or an entire text document into smaller parts, such as individual words or phrases. Tokens are the names given to these smaller units.

Before we can comprehend a natural language, we must first identify the words that make up a string of letters. As a result, tokenization is the most fundamental step in moving forward with NLP (text data). This is significant because the meaning of the text can be easily deduced by studying the words in the text.

For example,

"I have bread" - ['I','have','bread']

Stemming:

Stemming is a natural language processing approach that reduces inflection in words to their root forms, hence assisting in text, word, and document preparation for text normalization.

Inflection is the process through which a word is altered to transmit a variety of grammatical categories like tense, case, voice, aspect, person, number, gender, and mood. As a result, while a word may have several inflected forms, having multiple inflected forms within the same text adds redundancy to the NLP process.

As a result, we use stemming to reduce words to their simplest form or stem, which may or may not be a valid word in the language.

Eg: writing - write, baskets - basket

Word Embedding:

Word embeddings are a sort of word representation that allows words with similar meanings to be represented in the same way.

They are a distributed representation for text that is maybe one of the fundamental breakthroughs for deep learning approaches' excellent performance on difficult natural language processing challenges.

A word embedding is a learnt representation for text in which words with the same meaning are represented similarly. This way of modeling words and documents is one of the fundamental successes of deep learning on difficult natural language processing tasks.

Sequence to Sequence Model

We can use a Seq2Seq model to solve any problem involving sequential data. Our goal is to create a text summarizer using a long series of words (in a text body) as the input and a summary as the output (which is a sequence as well). As a result, this can be modeled as a Many-to-Many Seq2Seq issue. Seq2seq takes a sequence of words (sentences or sentences) as input and produces a sequence of words as output. Encoder and Decoder are the two fundamental parts of a many to many seq2seq architecture.

Encoder

Each token in the input sequence is processed by the encoder. It aims to squeeze all of the information about the input sequence into a fixed-length vector known as the 'context vector.' The encoder provides this vector to the decoder after passing through all of the tokens. The input length of the encoder is Maximum Permissible Sequence Lengths.

Decoder

The encoder delivers the internal states to the decoder after reading the entire input sequence, and this is where the output sequence prediction occurs. The decoder scans the context vector and attempts, token by token, to anticipate the target sequence. An embedding layer is defined in the decoder, followed by an LSTM network.

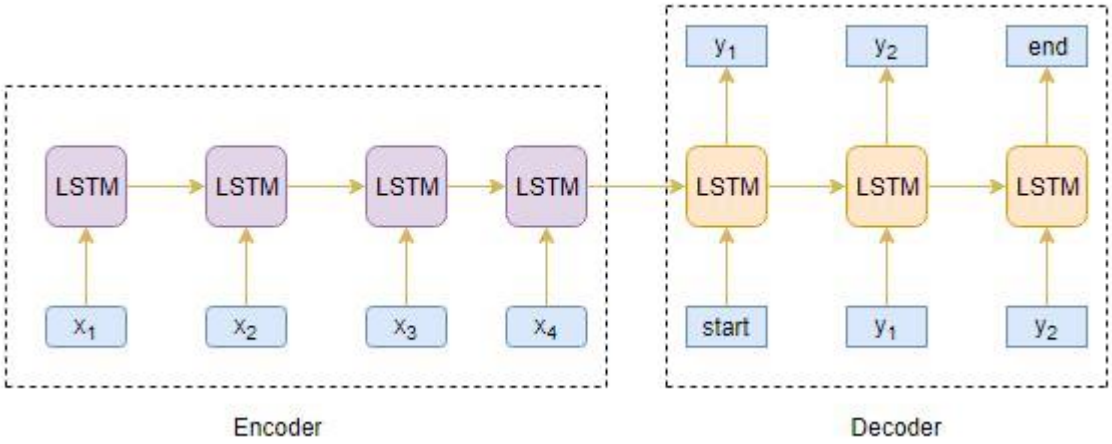


Fig 1.2 Sequence to Sequence architecture

The objectives of our proposed system include,

1. Create a model using ultrasonic sensors to detect objects and alert them.
2. To create a model that summarizes large contents into small paragraphs which contains only short and required information.
3. Program to read texts through images and provide output to the user as audio messages.
4. Build a cost effective solution that benefits visually imapired people.

The primary objective is to provide an integrated system to assist visually impaired people, by designing a product to help them transport and to read novels or get the summary of the novel.

The flow of the system is managed by a voice assistant as depicted in the Control Flow diagram shown in figure 1.2 The essential task is carried out based on the user's demand, whether to read the text or prepare a summary.

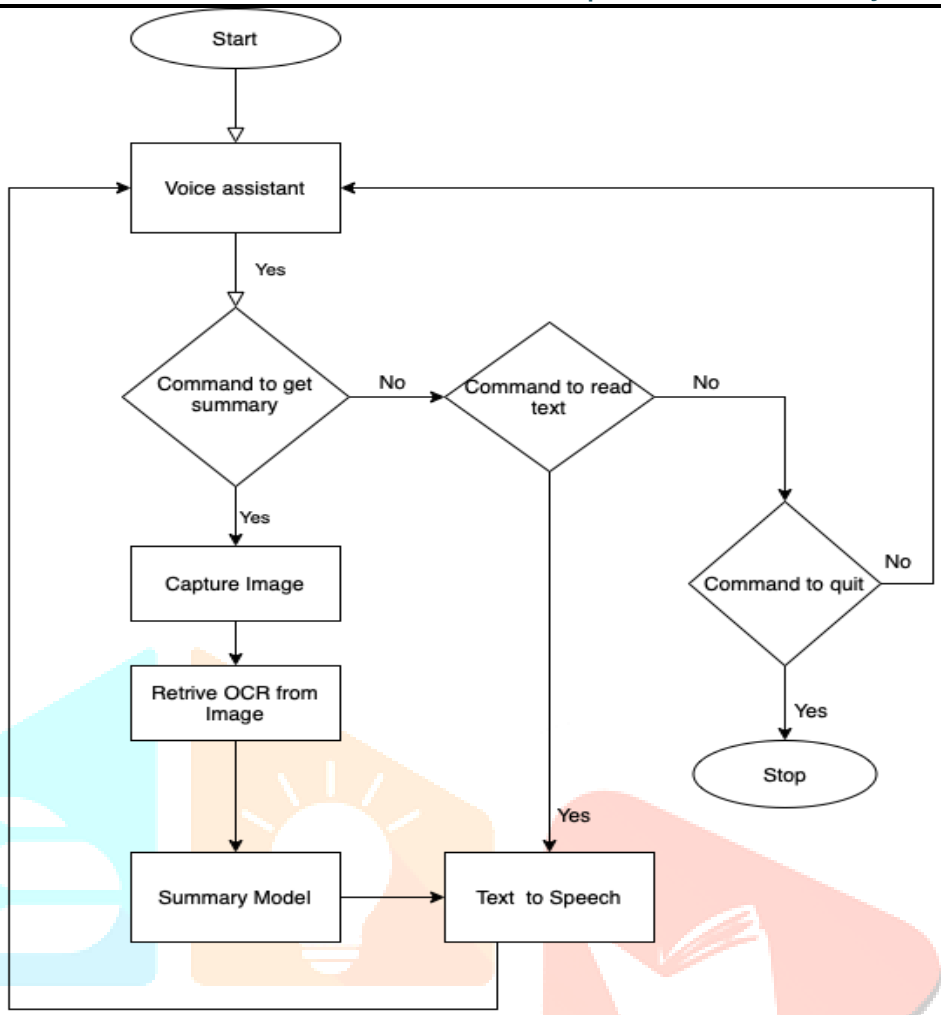


Figure 1.3 Control flow diagram of the Proposed System

4. Future Work

In future we will try to minimize the processing time and provide output more quickly than before. Further, we will try to add the extended modules support and try to reduce the complexity of the program. We can also add many languages and provide support to it. We have planned to expand the size of the dataset, play with the network's hyperparameters, try making it larger, and increase the number of epochs to get more accurate results from the text summarisation model.

5. Conclusion

In this paper, we have created a model design to help visually impaired people to detect objects and alert them using ultrasonic sensor, to create a model that summarizes large contents into small paragraphs which contains only short and required information, program to read texts through images and provide output to the user as audio messages, build a cost effective solution that benefits visually imapired people.

6. References

1.ÁNGEL, HERNÁNDEZ-CASTAÑEDA, RENÉARNULFO,GARCÍA-HERNÁNDEZ, YULIA LEDENEVA, AND CHRISTIAN EDUARDO MILLÁN-HERNÁNDEZ in " Extractive Automatic Text Summarization Based

on Lexical-Semantic Keywords" in IEEE Acess, Digital Object Identifier 10.1109/ACCESS.20202.2980226

2.R. M. Alguliyev, R. M. Aliguliyev, N. R. Isazade, A. Abdi, and N. Idris,''COSUM: Text summarization based on clustering and optimization,'' Expert Syst., vol. 36, no. 1, Feb. 2019, Art. no. e12340.

3.R. A. García-Hernández and Y. Ledeneva, ''Single extractive text summarization based on a genetic algorithm,'' in Proc. Mex. Conf. Pattern Recognit. Berlin, Germany: Springer, 2013, pp. 374–383.

4.K.S.Manikanta, T. S. S. Phani, A.Pravin,"Implementation and Design of Smart Blind Stick for Obstacle Detection and Navigation System",International Journal of Engineering Science and Computing, 2018.

5.Hatem, H. Raheem, A. I. Abdalla and Z. N. A. Al-Rawi, "Design & Implementation of ultrasonic radar system for distance measurements using Arduino," International Journal of Engineering & Technology, vol. 7, no. 4, pp. 3115-3118, 2018

6.M. Afsharizadeh, H. Ebrahimpour-Komleh, and A. Bagheri, ''Query-oriented text summarization using sentence extraction technique,'' in Proc.4th Int. Conf. Web Res. (ICWR), Apr. 2018, pp. 128–132

7.S. Charitha, N. B. Chittaragi, and S. G. Koolagudi, ''Extractive document summarization using a supervised learning approach,'' in Proc. IEEE Distrib. Comput., VLSI, Electr. Circuits Robot. (DISCOVER), Aug. 2018, pp. 1–6.

8.Y. Zhou and Z. Dong, "A vision-based autonomous detection scheme for obstacles on the runway," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 832-838, doi: 10.1109/CAC.2017.8242881.

9.Chandran, P., Aravind, S., Gopinath, J. and Saranya, S.S. (2015) 'Text to speech conversion system using OCR', International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol. 5, No. 1, pp.389–395.

10.Ragavi, K., Radja, P. and Chithra, S. (2016) 'Portable text to speech converter for the visually impaired', in Proceedings of the International Conference on Soft Computing Systems, Springer, pp.751–758

11.ÜstünErcan, Seda & MOHAMMED, Mohammed. (2020). Distance Measurement and Object Detection System Based On Ultrasonic Sensor and XBee. DüzceÜniversitesiBilimveTeknolojiDergisi. 10.29130/dubited.634256. 202

12.J. Majchrzak, M. Michalski and G. Wiczynski, "Distance estimation with a long-range ultrasonic sensor system," IEEE Sensors Journal, vol. 9, no. 7, pp. 767-773, 2019