



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH


IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 10, Issue 2, February 2022**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 7.542**

 9940 572 462

 6381 907 438

 [ijircce@gmail.com](mailto:ijircce@gmail.com)

 [www.ijircce.com](http://www.ijircce.com)

# Neuro-OCR Based Assistive System for Text Detection with Voice Output

Arun A<sup>1</sup> Akshaya A<sup>2</sup>, Deepa S<sup>3</sup>, Kousik S<sup>4</sup>, Dharnish M<sup>5</sup>

Assistant Professor, Department of Electronics and Communication Engineering, Knowledge Institute of Technology,  
Salem, Tamilnadu, India<sup>1</sup>

UG Student, Department of Electronics and Communication Engineering, Knowledge Institute of Technology, Salem,  
Tamilnadu, India<sup>2,3,4,5</sup>

**ABSTRACT:** Reading text from scene, images and text boards is an exigent task for visually challenged persons. This task has been proposed to be carried out with the help of image processing. Since a long period of time, image processing has helped a lot in the field of object recognition and still an emerging area of research. The proposed system reads the text encountered in images and text boards with the aim to provide support to the visually challenged persons. Text detection and recognition in natural scene can give valuable information for many applications. In this work, an approach has been attempted to extract and recognize text from scene images and convert that recognized text into speech. This task can definitely be an empowering force in a visually challenged person's life and can be supportive in relieving them of their frustration of not being able to read whatever they want, thus enhancing the quality of their lives. The image is processed to produce the output. The output is the information in image and is provided in the form of audio to the blind user itself. The image taken is processed to help blind persons for reading texts from labels, product packaging, papers, books, etc and also the objects and currency in the image are detected.

## I. INTRODUCTION

Every year, the number of visually challenged persons is increasing due to eye diseases diabetes, traffic accidents and other causes. Therefore applications that provide support to the visually challenged persons have become an important tool. Recent developments in computer vision, digital cameras, and computers make it possible to assist these persons by developing camera-based products that merge computer vision technology with other existing beneficial products such as optical

character recognition (OCR) systems. When a visually challenged person is walking around, it is important to get text information which is present in the scene/text boards. Reading is obviously necessary in today's society

Printed text is all over in the form of reports, receipts, bank documents, restaurant menu cards, classroom handouts, product packages, instructions on medicine bottles, etc. As an important form of communication, text is widely used in our daily life. For example, different sign boards, directions, shop names etc contain textual and/or symbolic information that is perceived by a human being to facilitate knowledge of environment and perhaps also help in his navigation.

## II. LITERATURE SURVEY

### 1. Text string detection from natural scenes by structure based partition and grouping.

In 2011, Chucai Yi and YingLi Tian explore a new framework to detect text strings with arbitrary orientations in complex natural scene images. The proposed framework of text string detection consists of two steps: 1) Image partition to find text character candidates based on local gradient features and color uniformity of character components. 2) Character candidate grouping to detect text strings based on joint structural features of text characters in each text string such as character size differences, distances between neighboring characters, and character alignment.

By assuming that a text string has at least three characters, it proposes two algorithms of text string detection: 1) adjacent character grouping method, and 2) text line grouping method. Their proposed methods detect text strings both in horizontal and non horizontal orientation which outperform the state-of-the-art results on the public Robust Reading Dataset which contains text only in horizontal orientation

### 2. Text Detection and Recognition with Speech Output for Visually Challenged Person.

In 2013, Reading text from scene, images and text boards is an exigent task for visually challenged persons. This task has been proposed to be carried out with the help of image processing. Since a long period of time, image processing

has helped a lot in the field of object recognition and still an emerging area of research. The proposed system reads the text encountered in images and text boards with the aim to provide support to the visually challenged persons.

**3. Text Detection and Recognition with Speech Output in Mobile Application for Assistance to Visually Challenged Person.**

In 2009, now a day’s use of mobile is broadly increased every person possesses a mobile phone, in which lots of application run. Using Android mobile phones we can help the visually challenged people by providing easy reading of text boards and printed text information in the form of audio. Reading text from printed text images and text boards is a challenging task for visually challenged persons. The proposed system extracts and recognizes text from scene images and converts that recognized text into speech.

**4. Image to Speech Conversion for Visually Impaired (International Journal of Latest Research in Engineering and Technology).**

In 2017, Asha G. Hagargund et al carried out a work and they concluded that the basic framework is an embedded system that captures an image, extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using a Raspberry Pi and a Raspberry Pi camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition) software and TTS (Text-to-Speech) engines. The audio output is heard through the raspberry pi’s audio jack using speakers or earphones. search in Engineering and Technology).

**5. Scene text detection using graph model built upon maximally stable external region**

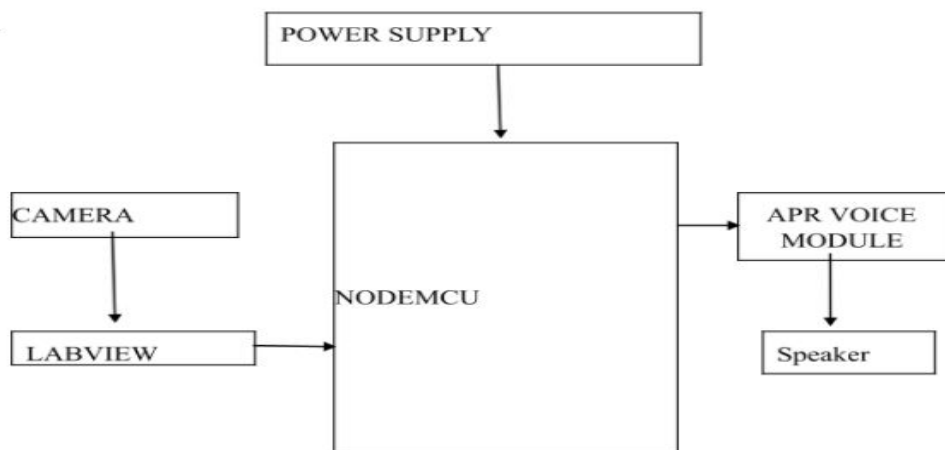
In 2013, Cunzhaoh Shi, Chunheng Wang, Baihua Xiao, Yang Zhang and Song Gao proposed a novel scene text detection approach using graph model built upon Maximally Stable External Regions (MSERs) to incorporate various information sources into one framework. Concretely, after detecting MSERs in the original image, an irregular graph whose nodes are MSERs, is constructed to label MSERs as text regions or non-text ones. Carefully designed features contribute to the unary potential to assess the individual penalties for labeling a MSER node as text or non- text, and color and geometric features are used to define the pairwise potential to punish the likely discontinuities. By 5 minimizing the cost function via graph cut algorithm, different information carried by the cost function could be optimally balanced to get the final MSERs labeling result. The proposed method is naturally context-relevant and scale-insensitive

**III. METHODOLOGY**

In this section, present a general review of previous work on text detection and recognition. There exists some research works for helping visually challenged people with text to speech technology. A number of portable reading assistants system have

been designed particularly for visually challenged persons. P. Blenkhorn, D.G. Evans implemented a computer-based system that allows blind users to read, create and edit one type of schematic diagram, specifically data flow diagrams used in software engineering, is presented, together with the mapping from the original diagram to a suitable generic, tactile diagram. They built an adaptive newspaper reading out system that is sorting headlines in order of user's priority.

**BLOCK DIAGRAM:**



**Proposed System in Text To Speech Output**

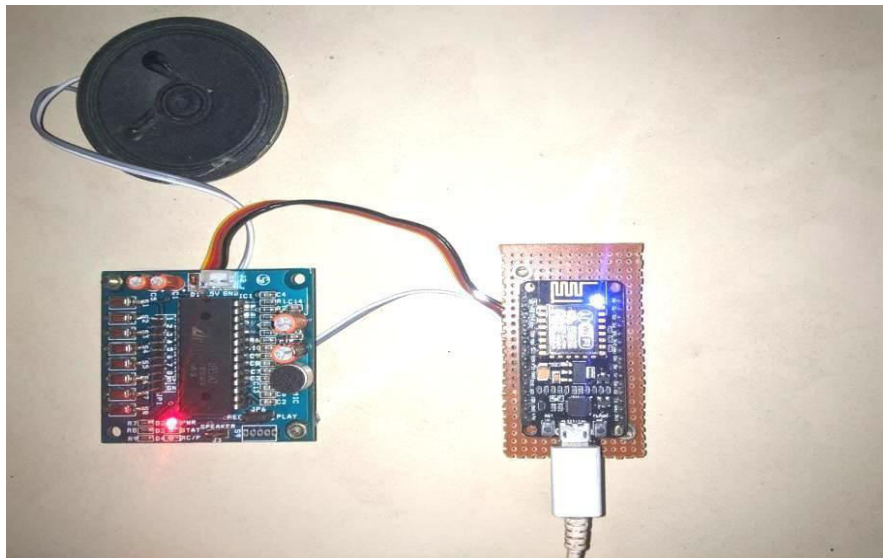
#### IV. WORKING PRINCIPLE

In this system, digitization of the textual information is proposed with the help of camera. It is also possible to use a scanner for immediate processing of information. The gathered information and digitized content is transferred to the OCR for extraction of textual information from the image that can be a photo or paper texts. This conversion happens only after the complete set of files are received by the system. The OCR uses grayscale version of the images for further analysis. Based on the intensity of the areas in the image, character identification is done. The lighter regions are considered as the background and the darker areas as text. This is further processed to identify and categorize the alphanumeric

#### V. RESULT

When the scanner scans the printed text and these texts are stored in an image file. In the first step of result the output of IMAQ Read file reads an image file, this image file will be displayed by using IMAQ windDraw (Display an image in an image window) at the output of IMAQ Read file. The segmentation first three character of word "optical" is shown in image window. After segmentation each character correlate with stored character templates, and recognition of printed text is done by this OCR software system.

#### Connection:



#### VI. CONCLUSION

Today's world is moving towards digitization. In recent years, digital cameras and camcoders are increasingly popular and existing systems have shown potential as other imaging devices. The researchers working in document analysis and recognition have changed their orientation and instead of working with conventional scanner captured document images. Reviewed some text detection and text recognition methods. The proposed system ensures to read printed text on handheld objects for assisting blind persons. In order to solve the common aiming problem for blind users, a motion based method to detect the object of interest is projected, while the blind user simply shakes the object for a couple of seconds. In review following interesting application were found such as newspaper reading out system, signage recognition and conversion to text, automatic electronic pen, camera based text reading system, travelling assistant system of bus detection and recognition

#### VII. FUTURE WORK

The future development will be an obstacle detection process, using haar cascade classifier for recognizing the obstacle from the object. A novel camera based computer vision technology to automatically recognize currency to assist visually impaired people will be enhanced.



#### REFERENCES

1. Mithe R, Indalkar S, Divekar N. Optical character recognition. International Journal of Recent Technology and Engineering.-2013 March.
2. Shah H, Shah A. Optical character recognition . International Conference on Signals, Systems and Automation.-2009.
3. Sim Liew Fong, Abdelrahman Osman Elfaki, Md Gapar bin Md Johar & Kevin Loo Tow Aik, Mobile Language Translator, 5th Malaysian Conference in So ware Engineering (Misses)- 2011.
4. Monk S. Raspberry pi cook Text localization and extraction in images using mathematical morphology and OCR Techniques-2013.
5. Kasarla PK, Kuamarswamy E. Implementation of text- to-speech for real time embedded system using Raspberry Pi processor. International Journal and Magazine of Engineering Technology Management and Research-2015 July.
6. Kumar GS, Krishna MNVLM. Low cost speech recognition system running on Raspberry Pi to support Automation applications. International Journal of Engineering Trends and Technology-2015.
7. Bhargava A, Nath KV, Sachdeva P, Samel M. Reading assis-tant for visually impaired. International Journal of current Engineering and Technology-2015.
8. Gomes LCT, Nagle EJ, Chiquito JG. Text-to-speech conversion system for Brazilian Portuguese using a formant-based synthesis technique. LPSDECOM FEEC- Unicamp.
9. Afritha Farhath K, Amruthavarshini R, Harshitha S, Saranya A, Velumadhavarao R, Development of shopping assistant using extraction of text images for visually impaired. International Conference on Advanced Computing-2014.
10. Kamesh DBK, Nazma SK, Sastry JKR, Camera based text to speech conversion, obstacle and currency detection for blind persons. Indian Journal of Science and Technology-2016 Aug. 43
11. R. Ani, Effy Maria, J.Jameema Joyce, V. Sakkaravarthy, M. A. Raja, smart specs:voice assisted text reading system for visually impaired persons using TTS method. International Conference on Innovations in Green Energy and Healthcare Technologies-2017.
12. D Vaithyanathan, Manigandan, Muniraj Cloud based Text extraction using Google Cloud Vison for Visually Impaired applications. International Conference on Advanced Computing-2019.



**INNO**  **SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 7.542**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details