

# OCR-WRITTEN TEXT TO AUDIO CONVERTER

Mr.Arjun Pratap<sup>1</sup>

Mr.Kunal Wavhule<sup>2</sup>

Mr.Viraj Patil<sup>3</sup>

Dr. Vaibhav Narawade<sup>4</sup>

<sup>1</sup>Author Designation, Department of Computer Engineering, RamraoAdik Institute of Technology, Maharashtra, India

<sup>2</sup>Author Designation, Department of Computer Engineering, RamraoAdik Institute of Technology, Maharashtra, India

<sup>3</sup>Author Designation, Department of Computer Engineering, RamraoAdik Institute of Technology, Maharashtra, India

<sup>4</sup>Author Designation, Department of Computer Engineering, RamraoAdik Institute of Technology, Maharashtra, India

## ABSTRACT

Converting text to audio using an optical character reader (OCR) with the help of state-of-the-art technology has solved the main problem of convenient communication. We know that listening is more interactive than reading a book. It is good that the system captures images and detects and recognizes text. Converting text to audio format file. OCR converts input images into editable text, allowing you to turn paper documents into editable and searchable documents. This can help reduce the physical space required to store document and significantly improve the work flow associated with those documents. We built the model using TensorFlow for text recognition after that converting TensorFlow Model to TensorFlow Lite model for using android/mobile application. We integrated the Text-to-Speech (TTS) services for extracted text to audio.

**KEYWORD:-**OCR, TTS, TensorFlow, TensorFlow Lite

## 1. INTRODUCTION

In the communication field This is the most common way of the most common text of the digitalized format used in the Internet signature, editing, signing, editing, and displaying the Internet, and the machine process that is more compact and used in the machine for machine mechanisms. (Extraction), Cognitive Calculation, Text Production and Major Data of Electronic Format. OCR is a study in artificial intelligence, awareness pattern and computer vision[1]. The initial version of the OCR must be trained in each character and the image of each character and operated in one form at the same time. High-level recognition accuracy support for most fonts Advanced systems that can generate different digital image file format inputs through support. Recognition of the optical symbol is an active area that helps the computer system receive an image and to develop the ability to automatically process data from the image [1]. There is a tremendous demand today for using computers to store information on computer disks from data that can be used in printed or handwritten documents to reuse this information[2].When reused, this information is too difficult to reador query text from these image files [3]. Therefore, there is a need for a way to automatically extract and store information from image files, especially text. Successful automation required identifying and addressing several key challenges. Some of the recent issues have to do with character font characteristics, such as images and paper documents. These issues can sometimes cause your computer system to be incorrectly recognized [4].

## 2. LITERATURE SURVEY

A Survey of the research done for Optical Character Recognition and the currently existing system give the following results. Many countries are contribute to the structure and infrastructure of various sectors in order to implement and support the digital process. The road to success, they say, is paved with digitization. Transparency and effectiveness are improved by digitization [1].

OCR is a capable method for integrating analogue life information into a cyberspace world. This technology has long been used to create digital libraries, recognize natural scenes, and understand handwritten office forms. Using OCR technology, documents scanned or captured by the camera are editable electronic versions that can be readily edited, retrieved, duplicated, and communicated [2].

Handwriting recognition is a very necessary method in modern in modern society. Before properly implemented this technique, we relied on to hand – write it, which can lead to errors. Efficient storage and access to physical data is a complex task. Maintaining proper data composition requires manual labor. The latest technology helps to store data in systems, which makes it much easier to store, manage and access data. It also provides highly secure data. An example of handwriting recognition software is Google Lens [3].

Tesseract is an open-source optical character recognition engine that is extensively considered as one of the correct currently available. To recognize text in an image, Tesseract first converts the input image to a binary image by performing an adaptive threshold. It performs an analysis of the connected components to preserve the outlines of each component: Lines of text are split into words based on letter spacing. Text recognition is performed in two steps to increase accuracy. Follow the same sequence of steps for to recognize both printed and handwritten text [12].

## 3. PROPOSED WORK

The proposed system aims at detecting the marks of a mark sheet and recognizes its characters in an efficient and cost-effective way[6].

The methodology followed in extracting the marks from the mark sheet using OCR is stated in the following steps.

1. Take a printed character as a input/image as an input to our system
2. Find the skeleton of the character/image
3. Segmentation and Feature Extraction
4. Comparison between Features/Feature Matching
5. Print the character/Image in the editable format and give the output
6. Output the Text

The considered system comprises the following steps:

First, import the printed binary image of the symbol as input.

Remove the pixel information from this image and save it to the appropriate memory.

After successfully completed, when the successful completion of the skeletal access, find the various functions or geometric formats that are available in this skeleton based on the information on pixel information.

## 4. METHODOLOGY

We're going to make an Android app that extracts text from a picture. As input, the system accepts a white, black-and-white, or color background image. The primary purpose of this research analysis is to use this computer function through an Android application. The Tesseract OCR engine is used to achieve this visual feature on an Android phone.

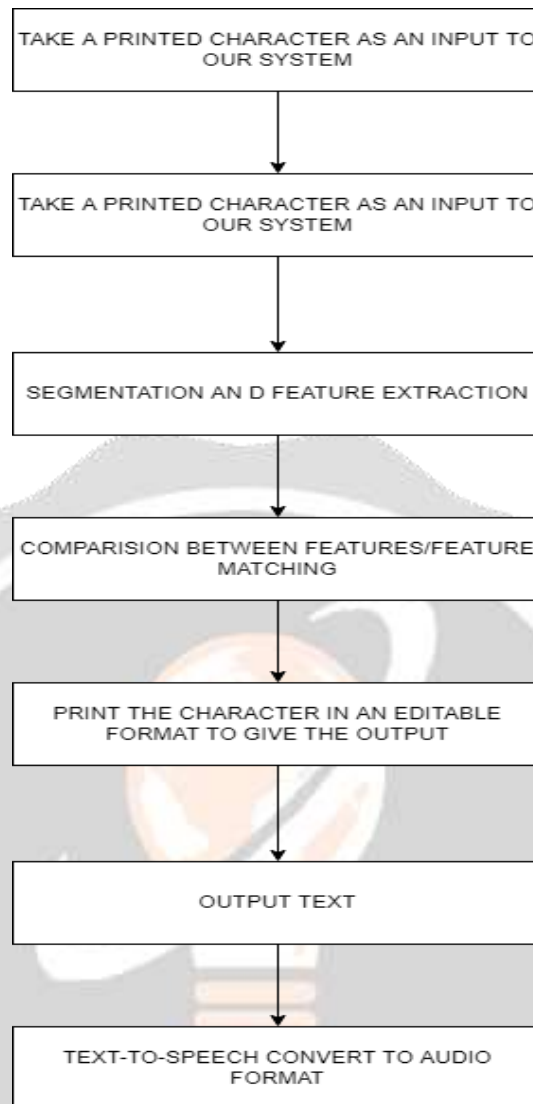


Fig 4.1 Flow Diagram.

This is a text extraction flowchart where the whole process is performed in five steps: image input, image data preprocessing, text extraction upon text detection, and finally text recognition. You can reduce this step by using an API like Tesseract OCR. Because the API has its own process. After onward extracted text can be converted with the help of TTS services into audio.

#### **Android:**

Android is a touchscreen android operating system built on versions of the Linux and other open-source software, built most importantly for smartphones and tablets. Android was created by the Open Handset Alliance, a group of developers with commercial backing from Google. The first Android smartphone, available in September 2008, was introduced in 2007.

As a result, Android is a mobile operating system that was created with the intention of being used on mobile devices.

Mobile phone features and applications The operating system is responsible for everything you see on your device's screen. The Android operating system is distinguished by a series of version numbers, each of which represents a substantial improvement in features, functionality, and stability, and is frequently codenamed. So, if you hear someone say Android versions like Marshmallow, Nougat, Oreo, Pie, or the most recent version of Android 10, you should know what they're talking about.

It's simply the name of the Android version that you can put on your device.

**TESSARECT**

The ability to rapidly search for content is very important for the \sin offices that need to handle massive scans or large document flows. You can now also use paper copy and paste programed instead of revising to fix issue. The Tesseract OCR is quick and dependable. This implies that the structure of the text is retained, saving time. The advantages of tesseract OCR really work when combined with other tools like scanning and file shrinking. If you do not have to spend time

manual workforce, your workflow is improved and faster and more efficient.

These are the Tesseract OCR engine tesseract and the command line program tesseract. Tesseract 4 adds an all-new Neural Network (LSTM) based OCR engine focused on string recognition while still supporting the OCR engine of Tesseract 3 which works with character pattern recognition. Tesseract 3 compatibility comes with mishandling of OCR engine inheritance mode . You also need trained data files to support inheriting engines, such as files from the tesseract data repository.

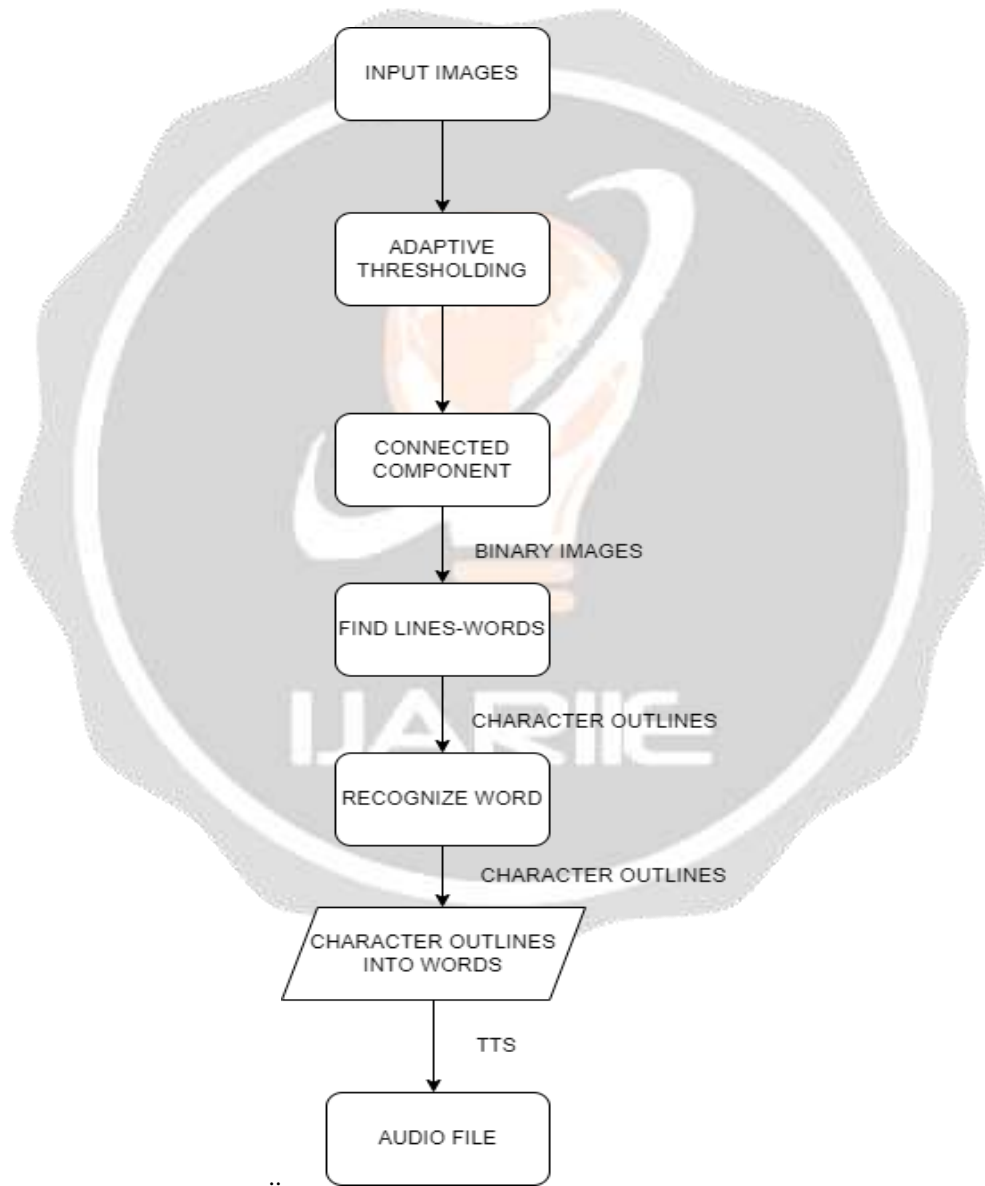


Fig 4.2: Phases of Audio Conversion

**4.1 Adaptive Thresholding**

In its most basic form, adaptive thresholding takes a grayscale or colour image as input and generates a binary image

that represents the segmentation.. We need to compute a threshold for every pixels in the image It is set as the background value if the pixel value is less than the edge, else it is set as the foreground value.

#### **4.2. Connected Component**

Two pixels must be adjacent and the grey level must represent a similarity requirement in order for them to be connected. In a binary image with values of 0 and 1, for example, if two pixels are adjacent but require the same value, it is said to be concatenated.

#### **4.3 Find Lines-Words**

The process begins with BLOB filtering and line design. A simple percentile level of filter removes dropkick and removes the vertical character, and the page layout survey has already been provided with text area of, nearly consistent text size. The average growth is accessing text in this area, so it is safe to filter less than 4, medium heights, perhaps less than the shares of less than 4, medium heights, perhaps, no matter, powdered labeling and noise sharing. The filtered drops are likely to be appropriate for models without overlapping, parallel, parallel slope lines. Positioning and blob x-coordinate handling simplify assignment of blob to a single text string while tracking skew across the entire page and greatly increasing the risk of assigning the wrong text string when skewed. Reduce. After the refined blobs are allocated to the lines, minimum median squares are used to evaluate the baseline, so filtered blobs fit the supportable line. The final step in the road creation process is to horizontally join the blobs that overlap more than half, putting the diacritics together with the correct stems and correctly concatenating parts of some broken characters.

#### **4.4 Recognize word**

Choosing how to break words into characters is an important part of character recognition engines' popularity. When looking for a line, the initial segmentation result is considered first. The remaining word recognition phases are only applicable to text with non-fixedsteps. Symbols that are linked together should be cut. Symbols that are broken can be linked together.

#### **4.5 Character Outlines into Words**

In this Phase Word converted into the machine-encoded format. After onward the character can store monetize, modified.

#### **4.6 TTS**

The process of translating text into human speech is known as text-to-speech (TTS) or speech synthesis. When it comes to generating interactive voice response (IVR) solutions and other speech applications, TTS is a popular choice among developers and corporate users. This is due to the fact that it reduces production time by eliminating the requirement to capture human speech audio files. Each message must be recorded as a human voice in the recorded file, however TTS hints can be dynamically created from raw text.

## **5. EXPERIMENTAL RESULTS**

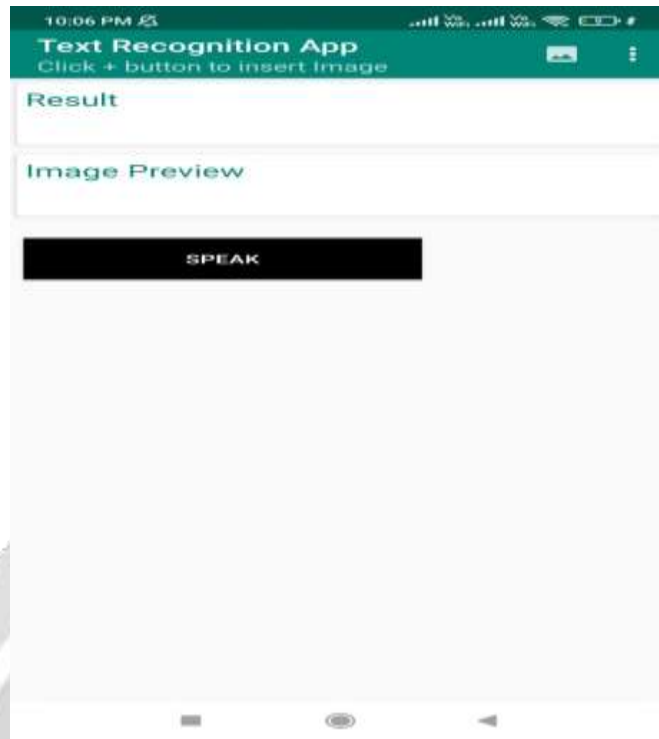


Fig 1: Initial Stage

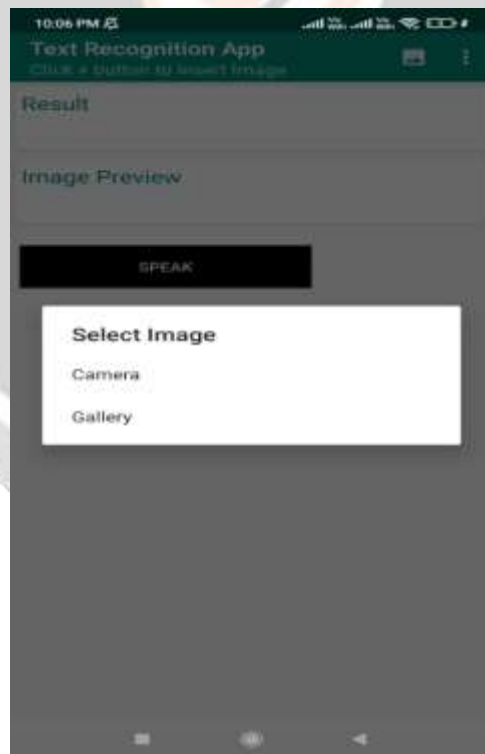


Fig 2: Design of the android Application.

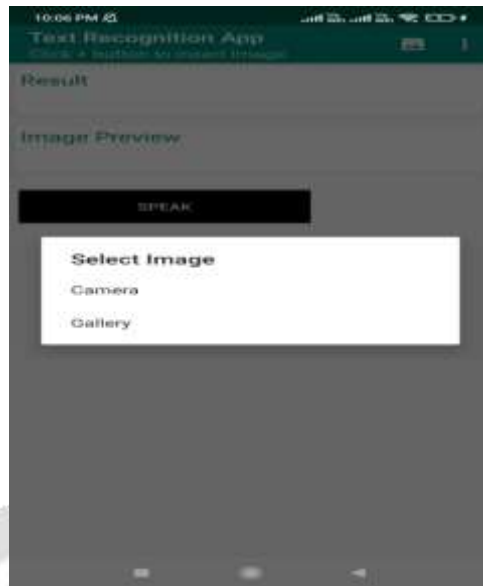


Fig 3: Selection of Option

In the Application user can be scan the image via camera or it can get the image from the Gallery.

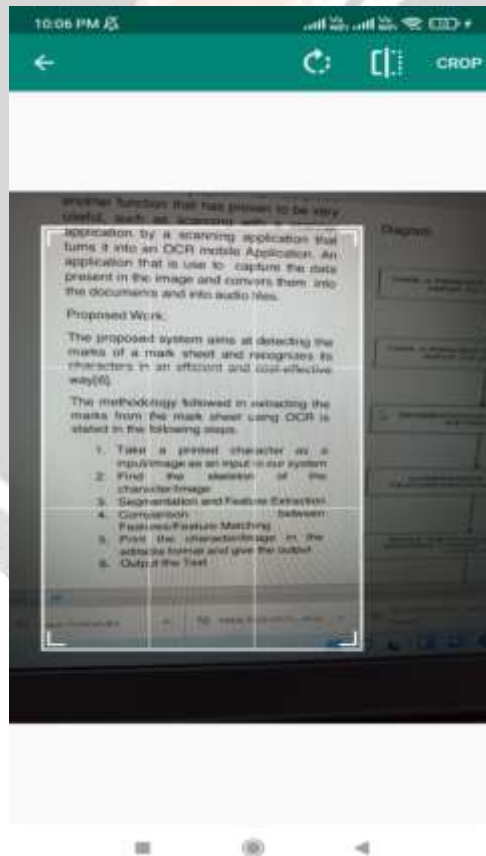


Fig 4 : Input Capturing

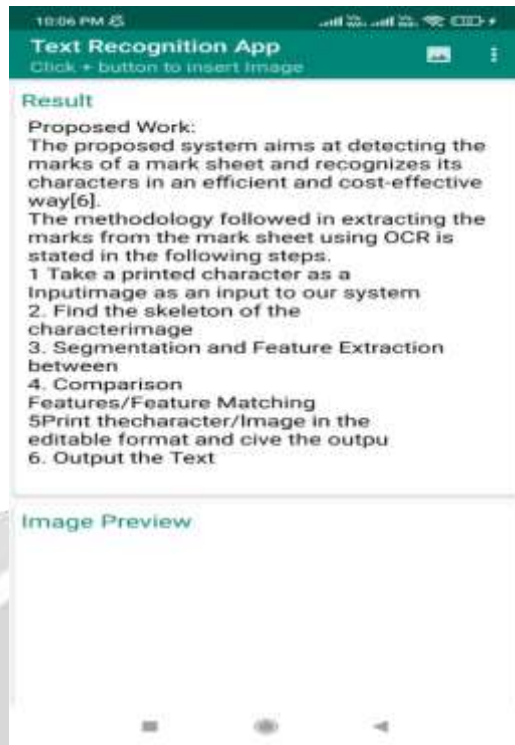


Fig 5: Final Output

Scan the image via camera and its be extracted to text form image and its get the machine editable text. Accuracy of the text be 99.05%

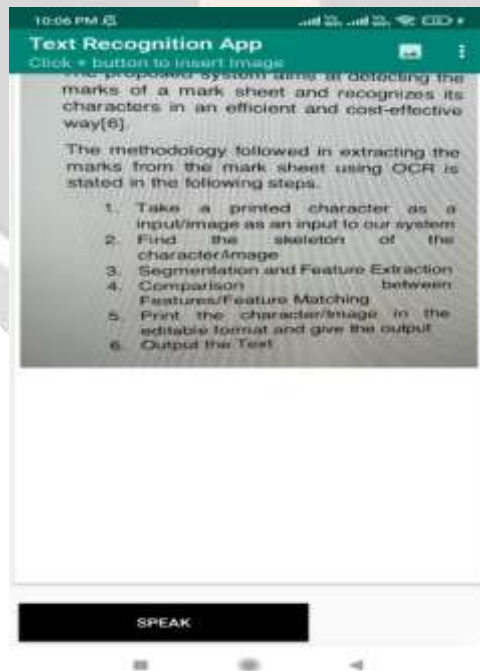


Fig 6 : Captured Content

Here the preview of the image the scanned the image and extracted text converted in the audio file.



## 6. RESULT DISCUSSION

The above the technique of the extraction of the text can be converted into audio its be accuracy of the extraction is more compare to another method its be very fast to implemented and use to the android application it can be used. The sound quality of the using TTS its be good.

## 7. Conclusion

This project is to propose an image conversion approach using optical character and text recognition as a user-developed application, and it is an attempt to propose an effective and efficient image conversion approach that can be applied in real time.

This approach can read text from documents, web pages, or e-books, and even generate synthesized speech through computer speakers or phone speakers or phone speakers. The developed software sets all the rules of a single alphabet and pronunciation methodology used in grammars and dictionaries.

This project is an attempt to propose an approach to OCR speech to text conversion using “speech technology” technology. The developed application is easy to use, economical, and can be applied in real time.

## 8. FUTURE WORK

THE FOLLOWING AREA CAN BE EXPLORE.

Appendixes, if needed, appear before the acknowledgment.

1. Extracted text from the .pdf , .HTML and .HTM etc format its be converted into audio formats.
2. Develop the Module in other languages just like Hindi and other regional language its be now limited in English.

Develop the module the TTS module in Other Language

## REFERENCES

1. Sandeep Dwarkanath Pande, Sandeep Dwarkanath Pande , Pramod Pandurang Jadhav c, Rahul Joshi d, Amol Dattatray Sawant e, Vaibhav Muddebhalkar f , Suresh Rathod d, Madhuri Navnath Gurav b, Soumitra Das, "Digitization of handwritten Devanagari text using CNN transfer learning – A better customer service support", *Neuroscience Informatics* 2 (2022) 100016.
2. ADITYA PALI , ABHIJIT MUSTAFI, " Vartani Spellcheck – Automatic Context-Sensitive Spelling Correction of OCR-generated Hindi Text Using BERT and Levenshtein Distance", *Computer Science and Engineering Birla Institute of Technology, Mesra Ranchi, India.*
3. Mariona Coll Ardanuy<sup>1 c</sup> Kasra Hosseini<sup>1 d</sup> Barbara McGillivray<sup>1,4 e</sup> Giovanni Colavizza, " Assessing the Impact of OCR Quality on Downstream NLP Tasks", *The Alan Turing Institute, London, United Kingdom* <sup>2</sup>University of Amsterdam, Amsterdam, The Netherlands
4. Aditya Rajbongshi<sup>1</sup> , Md. Ibadul Islam<sup>2</sup> , Md. Mahbubur Rahman<sup>4</sup> , Anup Majumder<sup>5</sup> , Dr. Md. Ezharul Islam, " Bangla Optical Character Recognition and Text-to-Speech Conversion using Raspberry Pi", *(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 6, 2020.*
5. Prasanta Pratim Bairagi, " Optical Character Recognition for Hindi", *International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 05 | May-2018*
6. Ravneet Kaur, Assistant Professor, " TEXT RECOGNITION APPLICATIONS FOR MOBILE DEVICES", *Volume 9, No.4, April 2018 Journal of Global Research in Computer Science.*
7. K.Karthick, K.B.Ravindrakumar, R.Francis, S.IIankannan, " Steps Involved in Text Recognition and Recent Research in OCR; A Study", *International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1, May 2019.*

8. YUE YIN1 , (Student Member, IEEE), WEI ZHANG 2,3, (Member, IEEE), SHENG HONG1 , (Student Member, IEEE), JIE YANG1 , (Member, IEEE), JIAN XIONG 1 , (Member, IEEE), AND GUAN GUI 1 , (Senior Member, IEEE),” Deep Learning-Aided OCR Techniques for Chinese Uppercase Characters in the Application of Internet of Things”, Received March 2, 2019, accepted April 1, 2019, date of publication April 4, 2019, date of current version April 18, 2019
9. JishnuMukhoti, Sukanya Dutta, and Ram Sarkar,” Handwritten Digit Classification in Bangla and Hindi Using Deep Learning”, APPLIED ARTIFICIAL INTELLIGENCE 2020, VOL. 34, NO. 14, 1074–1099.
10. Anuradha1 Ms. Nidhi Sengar2 Student1 Assistant Professor,” HANDWRITTEN TEXT RECOGNITION USING TENSORFLOW”, Vol 11, Issue 4 , April/2020 ISSN NO:0377-9254
11. JAMSHED MEMON 1 , MAIRA SAMI 2 , RIZWAN AHMED KHAN 3 , AND MUEEN UDDIN,” Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR), Received June 24, 2020, accepted July 16, 2020, date of publication July 28, 2020, date of current version August 14, 2020.

