IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⫻ Peer-reviewed & Refereed journal ⫻ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14728

An Efficient OCR System For Visually Impaired

Arya Chandran V[1], Shalom David[2]

Student, Department of Computer Science, Christ Nagar College Maranalloor[1]

Assistant Professor, Department of Computer Science, Christ Nagar College Maranalloor[2]

Abstract: The major problem faced by visually impaired people is that they are unable to do text recognition on their own, which forces them to depend on others for their day to day activities such as reading newspapers, letters sent through post, referring books etc. The aim of this project is to develop an Optical Character Recognition (OCR) system integrated into an academic system which is designed to improve accessibility for visually impaired students, teachers, and administrators. This system uses advanced OCR technology, combined with image processing, Text-to-Speech Conversion (TTS) and machine learning algorithms, to convert printed text into a digital format. The key objective is to ensure that visually impaired users can access educational, administrative, and communicative resources seamlessly and can be recognized text into spoken words. For students, the OCR system allows them to access a variety of academic resources, such as textbooks, class notes, assignments, schedules, and exam papers, in an accessible digital format. The system can instantly extract text from documents and convert it into a format that is compatible with Text-to-Speech software, enabling students to interact with materials independently. Furthermore, students can benefit from multi-language support, ensuring they can access information in their preferred language. For teachers, the OCR system provides the ability to convert printed or handwritten teaching materials and notices into accessible formats for their visually impaired students. Teachers can scan documents and instantly create digital copies that are compatible with assistive technologies. This facilitates smoother distribution of teaching resources and more inclusive classroom engagement. Teachers can also use the system to modify documents to meet specific needs of students, enhancing personalized learning.

Keywords: OCR, TTS, Visual Impairment, Accessibility, Assistive Technology

I. INTRODUCTION

In the current learning environment, accessibility and inclusivity are crucial in providing equal opportunities to all students, instructors, and administrators. The integration of an Optical Character Recognition (OCR) system in an academic system which is a giant step towards providing more accessibility, particularly for the visually impaired. The innovative integration makes printed and electronic documents, notices, and learning material convertible to machine-readable text, which can be processed by Text-to-Speech (TTS) technology or screen readers. With OCR integration, visually impaired users can read course material, administrative reports, exam papers, and official notices without manual transcription or external assistance. The system allows students to perform academic activities independently, allows instructors to plan coursework more effectively, and allows administrators to manage official documents more effectively. Furthermore, the OCR-enhanced system ensures institutional compliance with accessibility legislation, offering an inclusive academic environment where all can easily interact with digital and printed material. This report offers an overview of the functionality, benefits, and impact of OCR technology integration into an academic system, highlighting how it enhances usability, accessibility, and efficiency in the academic environment. The system is also integrated in harmony with existing learning management systems (LMS), screen readers, and assistive devices, ensuring accessibility features are accessible on multiple platforms. With enhanced artificial intelligence (AI) and machine learning (ML), OCR technology becomes superior, enhancing text recognition accuracy on various fonts, handwriting styles, and languages. This not only enhances user experience but also institutional compliance with global accessibility standards, offering a more inclusive academic environment where every individual, regardless of visual ability, can engage in academic and administrative activities to the fullest.

II. SIMILAR WORKS

➢ The paper titled **"Smart Translation for Visually Impaired People"** [1] is designed to develop a smartphone-based TTS system by using OCR that can help support visually impaired individuals. The proposed system enables the user to take pictures of the printed text that is then read and converted to audio output to make written material more accessible. It is applied to provide an affordable, efficient, and accessible solution for the visually impaired in reading printed text through speech output. OCR, or Optical Character Recognition, pulls text from images. Google Text-to-Speech (GTTS) for converting the extracted text into natural-sounding speech.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14728

Convolutional Neural Networks for pattern recognition in text processing. It employs the technologies of Support Vector Machines (SVMs) for character comparison and recognition. This assists visually impaired people to gain real-time independence and independence with printed materials. Its features include lowering the cost of assistive devices, increasing accuracy in text recognition, and improving naturalness in synthesized speech.

- The paper on **"Intelligent Reader for Visually Impaired Person"** [2] deals with developing an assistive system that would help a visually impaired individual read printed text. The system usually employs OCR technology in the form of scanning and identification of text from documents, books, and other printed material. Then, the recognized text is translated into speech through TTS technology to enable listening on the part of the user. The "intelligent" part will involve the potential to learn with time, hence maybe including elements like context sensing, language translation, or other complex text handling. It uses OCR, Deep learning models and speech synthesis algorithms. The device is meant to be user-friendly, efficient, and easy for visually impaired individuals to use freely to read and interact with print independently.

- The paper **"Text/Object Recognition for Visually Impaired People"** [3] focuses on developing a system that helps visually impaired individuals recognize text and objects using Optical Character Recognition (OCR) and deep learning-based object detection techniques. The system is built using a Node MCU microcontroller, controlling peripherals such as a camera and headset, allowing users to capture images of text, which are then processed and converted into speech. The device, mounted on spectacles, captures images under proper lighting conditions, extracts text using OCR, and converts it into an audio format using Google Text-to-Speech (gTTS) and Python programming. The system also has an object detection model (YOLO - You Only Look Once) to identify objects in the environment, aiding users in recognizing everyday items. For the hearing-impaired users, a soundbite hearing system has been incorporated, which sends sound signals through the teeth via vibration technology and thus does away with the conventional hearing aids. It is being utilized to further improve access for visually and hearing-impaired persons by incorporating computer vision, deep learning, and assistive technologies for live audio feedback.

- The paper **"Image Processing Based on Optical Character Recognition with Text to Speech for Visually Impaired"**[4] is an efficient, cost-effective system where the visually impaired can hear what is printed, rather than read it. It is integrated with OCR and TTS technology by a webcam, where images of printed text are captured, and then the content extracted and transformed into speech output via computer audio or headphones. It is basically an image processing module for extracting the text and a voice processing module for the speech synthesis of the text. It uses Google Tesseract OCR for text recognition and Google Text-to-Speech (gTTS) for audio conversion, thus achieving high accuracy and natural-sounding speech. This approach removes the need for external hardware, thus enhancing accessibility, portability, and affordability, allowing visually impaired users to be more independent in reading newspapers, books, and other printed materials.

- The paper **"Vulnerability Analysis of Transformer-based Optical Character Recognition to Adversarial Attacks"** [5] is concerned with the assessment of the robustness of Transformer-based Optical Character Recognition (OCR) systems against adversarial attacks. With Transformer models gaining popularity in OCR due to their superior performance in recognizing complex text patterns, this study examines how vulnerable these models are to adversarial perturbations, or small, carefully crafted modifications to input images that can mislead the OCR system into incorrect text recognition. The paper delves into the various adversarial attack techniques, evaluates their effect on OCR accuracy, and presents possible defense mechanisms to improve the security and reliability of Transformer-based OCR systems. It uses FGSM, C&W, DeepFool for adversarial attacks. The results have brought to the fore critical vulnerabilities, emphasizing the need for enhanced robustness in real-world applications such as document processing, security systems, and automated text extraction.

- In this paper **"Virtual Assistant and Navigation for Visually Impaired using Deep Neural Network and Image Processing,"** [6] deep neural networks will be used as part of image processing techniques so as to explain what is visible and detect in its surroundings so as to sense possible obstacles or the presence of particular landmarks with all important visual signals. It uses integrated wearable smart glasses for hands-free operation and combines object detection, text recognition and audio feedback for navigation and reading. The processed information will be converted to audio feedback and is given as real-time navigation. The virtual assistant is therefore intended to be used in helping visually impaired persons navigate through space by providing guidance, alerts, and context relevant information for maximizing their independence and safety while traversing public space. It uses Raspberry Pi board for processing, Tesseract OCR for text recognition and YOLOv5 for object detection and localization. The focus of the paper is on developing a strong and intelligent solution through cutting-edge AI and computer vision for enhancing mobility and accessibility.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14728

- The paper **"An Efficient Text Recognition System from Complex Color Image for Helping the Visually Impaired Persons"** [7] deals with the development of a robust text recognition system to extract and process text from complex color images to support visually impaired people. The proposed system encompasses advanced Optical Character Recognition (OCR) techniques in conjunction with image enhancement methods to enhance the readability of the text at noisy, cluttered, or poor-contrast backgrounds. The system increases the accuracy in identifying and extracting text from real-world images such as signboards, product labels, and documents by using machine learning algorithms, color normalization, and deep learning-based text detection. The extracted text is then converted into speech using Text-to-Speech (TTS) technology, thereby allowing visually impaired users to access printed information more independently. This paper emphasizes high-accuracy text recognition in dynamic environments and seeks to deliver the possibility of a more user-friendly and efficient assistive tool for visually impaired users.

- The paper **"A Novel Reading Technique for Visually Impaired Person Using Enhanced Optical Character Recognition Method"** [8] focuses on developing an advanced Optical Character Recognition (OCR) system to assist visually impaired individuals in reading printed text. The study proposes an enhanced OCR technique that improves text recognition accuracy by integrating image preprocessing, noise reduction, and deep learning-based character recognition. The system captures text using a camera, processes the image to extract and refine the text, and then converts it into speech using Text-to-Speech (TTS) technology. This system, with the help of machine learning algorithms and optimized OCR methods, improves the accessibility of reading for visually impaired people, thus providing a more accurate, efficient, and user-friendly solution. The paper discusses how this technology can be a great tool for independence and accessibility in making printed materials accessible to people with visual impairments.

- The paper on **"Autonomous Assistance System for Visually Impaired using Tesseract OCR & gTTS"** [9] focuses on the development of an autonomous assistance system that would assist visually impaired individuals by converting the written text to speech. It consist of open-source implementation using Tesseract OCR, which supports multiple languages. It can process images from scanned images and live captures. Tesseract OCR will be used in detecting and extracting text from printed documents or images. This text will then be converted into speech using gTTS (Google Text-to-Speech), hence providing an accessible solution to read and understand the printed content. This system is aimed at improving independence and accessibility for visually impaired users by providing real-time text recognition and auditory feedback, making it easier for them to interact with printed materials.

- The paper **"A Study on Face, Expression, Object, and Character Recognition to Implement an Aid for the Visually Impaired"** [10] explores the integration of computer vision and machine learning techniques to develop an assistive system for visually impaired individuals. The study focuses on recognizing faces, facial expressions, objects, and text using advanced image processing and deep learning algorithms. The system proposed uses Optical Character Recognition (OCR) for extracting text, object detection models to identify surrounding items, and facial recognition techniques for detecting and interpreting emotions, in order to assist visually impaired people to be independent. The information recognized is transformed into audio output using TTS technology so that the users can navigate around their environment effectively. The paper points out the potential of using a combination of multi-modal recognition techniques in order to design a comprehensive and intelligent assistive tool, enhancing accessibility and quality of life for the visually impaired.

- The paper entitled **"Optical Character Recognition on Live Footages for Visually Impaired People"** [11] focuses on the development of a real-time text recognition system in processing live video footage to enable visually impaired persons to read textual information from their surroundings. By incorporating Optical Character Recognition (OCR) with live video processing, the system continuously captures frames from a camera, detects and extracts text from dynamic scenes such as signboards, labels, documents, digital screens, and converts them into speech using Text-to-Speech (TTS) technology. The research addressed challenges including motion blur, varying lighting conditions, and complex backgrounds when improving text detection accuracy in real-world situations. This is a proposed system that is set to increase independence and accessibility of visually impaired people by giving audio feedback on text content in the environment. Such users are in a better position to engage their surroundings in an effective way.

- The paper **"OCR Using Visually Impaired People Shopping Trolley Technology"** [12] is focused on developing a smart shopping trolley system that integrates Optical Character Recognition (OCR) to assist visually impaired individuals in identifying and selecting products independently while shopping. The system utilizes a camera-mounted trolley to capture product labels, price tags, and other relevant text, which is then processed using OCR to extract and recognize the information. The recognized text is converted into speech output using Text-to-Speech (TTS) technology, allowing users to hear product details in real-time.

Furthermore, the system may include barcode or QR code scanning, object detection, and navigation assistance to enhance the overall shopping experience. This technology aims to empower visually impaired shoppers by providing them with real-time audio feedback to make their shopping experience more accessible, efficient, and independent.

➢ The paper **"Smart Application for the Visually-Impaired"** [13] develops a mobile or digital assistive application that will promote accessibility and independence for visually impaired individuals. This application integrates OCR for text detection, TTS for audio conversion, and object recognition using machine learning to aid users in understanding their surroundings. It implements robust computer vision algorithms like SIFT and FAST techniques. It could include voice commands, real-time object identification, currency recognition, navigation assistance, and smart notifications to make life easier to carry out. Through the use of computer vision, artificial intelligence, and mobile technology, the smart application would look to be an easy-to-use and efficient tool for a visually impaired user to interact with his or her environment in a better way and lead to more independence and accessibility.

➢ The paper on **"Text Reader for Visually Impaired Persons"** [14] focuses on creating a system to help visually impaired people read written text aloud. This solution is a mobile application, which makes it easy to access, portable, and user-friendly with audio output. It employs several technologies, including Optical Character Recognition (OCR), which scans and recognizes printed text from books, documents, or any written material. After the text is identified, the system processes it and converts it into speech using text-to-speech (TTS) software. The goal is to develop a simple, accessible tool that helps visually impaired users read and understand printed information, improving their quality of life and independence.

➢ The paper, **"Raspberry Pi-Based Intelligent Reader for Visually Impaired Persons,"** [15] proposes the design of an affordable and efficient assistive technology using a Raspberry Pi platform to help the visually impaired in accessing printed text. The system integrates a camera to capture text from physical documents, which are then processed through Optical Character Recognition (OCR) software. The extracted text is converted to speech using the Text-to-Speech (TTS) system, and content is listened through by the user. The objective of the designed system is through the use of the Raspberry Pi to provide accessibility at a minimum cost, transportability, and customization for independence in reading for visually impaired persons.

➢ This paper, titled **"A Scene Perception System for Visually Impaired Based on Object Detection and Classification Using CNN"** [16] focuses on creating a system to help visually impaired people understand their environment and be aware of what's around them. The system uses CNN for object detection and classification. It can identify and categorize objects in the scene in real-time. Input comes from cameras or sensors, which detect items like people, obstacles, or landmarks. The system then communicates these detected objects to the user through auditory feedback. This helps them comprehend their surroundings. It also supports complex scripts like Devanagari and can process information offline. The main goal is to provide a smart, real-time perception system that improves the mobility and safety of visually impaired individuals by assisting with navigation and interaction with their environment.

➢ The paper on **"OCR-Based Facilitator with Emotion Detection for the Visually Impaired"**[17] focuses on developing an assistive system that helps visually impaired individuals access printed text and recognize emotional cues in interactions. The system uses Optical Character Recognition (OCR) to extract text from physical documents, which is then converted into speech via Text-to-Speech (TTS) technology, allowing users to hear the content. The other feature added to it is emotion detection, probably using facial recognition or voice tones analysis, which is used to interpret the emotional context in any type of communication whether written or spoken. This is done to provide accessibility to text, but most importantly, it offers insight to users about social interactions, making them better understand both content and emotional tone in their surroundings. The overall goal is to create a more holistic, interactive system that promotes greater independence and social awareness for visually impaired users.

➢ The paper on **"Text Recognition System for Visually Impaired Using Portable Camera"** [18] focuses on a portable and practical access to printed text for the visually impaired. A small handheld camera captures images of text on documents, books, or signs and processes these through Optical Character Recognition (OCR) technology for extraction of text. The recognized text is then synthesized as speech with Text-to-Speech (TTS) technology to let the user listen to it. This system is supposed to offer an accessible, portable device that enables users with a severe visual impairment to be independent as much as possible while they can read texts while traveling from place to place. It uses computer vision, NLP and machine learning techniques also scene text detection algorithms.

➢ The paper on **"Smart Reader for Visually Impaired"** [19] deals with the development of a system intended to aid visually impaired readers in reading printed text. A "Smart Reader" usually includes OCR technology, which scans and captures text from paper documents, books, or any other form of printed text. The extracted text is further processed into speech through TTS technology, so the user can listen to it. It is likely to be made portably friendly and user-friendly. Indeed, this system could give independence to a visually impaired user by providing him with the possibility of accessing written information efficiently. It aims to create a faithful and accessible tool that should bring improvement in the quality of life and increase the level of autonomy for visually impaired users.

➢ The paper on **"A Tesseract-Based Optical Character Recognition for a Text-to-Braille Code Conversion"** [20] discusses a system that helps visually impaired individuals by converting printed text into Braille using Tesseract OCR. The system captures an image of the text with a camera or scanner. Then, Tesseract OCR recognizes the text from the image. After scanning, the text is converted into Braille code. This code can be shown on a refreshable Braille display or printed. This system improves access for people with vision impairments by providing a clear, reliable, and effective way to read in Braille. As a result, it boosts their independence and involvement with written material.

➢ The paper on **"Smart Reader for Visually Impaired People Based on Optical Character Recognition"** [21] focuses on developing an assistive technology system designed to help visually impaired individuals read printed text. The system uses Optical Character Recognition (OCR) to scan and recognize text from physical documents, books, or other printed materials. Once the text is detected, Text-to-Speech (TTS) technology converts it into speech so that the user can hear the content. The system is probably portable, user-friendly, and aims at improving the accessibility of printed material for visually impaired users in order to enable them to browse through and interact with their surroundings effectively. The aim is to raise the daily life standards of the visually impaired who will now be easily equipped with a practical means to read text clearly.

➢ The paper, **"Assistance For Visually Impaired Using FingerTip Text Reader Using Machine Learning,"** [22] focused on developing a novel solution that could help a visually impaired individual read printed text. The system uses a reader based on fingertip, which the user has to place the fingers on the text, and the machine learning algorithms process the scanned information. The system employs machine learning techniques to recognize the text from the document and produce speech from this input, so that the user can listen to what is said and gain access to the information in that material. The aim is for this tool to be very simple, intuitive, and effective so that visually impaired users can attain this quality and improve their quality of life. It uses CNN, RNN, computer vision and NLP techniques.

➢ The paper **"A Dual-Purpose Refreshable Braille Display Based on Real-Time Object Detection and Optical Character Recognition"** [23] aims at developing an advanced assistive device for visually impaired persons. The system focuses on real-time object detection and OCR, with the goal of scanning and interpreting text; detected text is converted into Braille and displayed on a refreshable Braille display for tactile reading. The dual-purpose feature involves the ability of the system to identify objects and report this information in real time to the users, thereby further enhancing the interaction process between the user and their environment. This paper seeks to enhance the accessibility process by offering an adaptable and efficient tool capable of both textual and environmental awareness for visually impaired users.

➢ The paper on **"AI-Based Reading System for Blind using OCR"** [24] deals with designing an assistive system that helps the blind and partially sighted people read print. The main approach is an application of artificial intelligence and OCR in the text extraction process of a variety of documents, ranging from books and labels to road signs. Once the text is recognized, it is converted into speech using Text-to-Speech (TTS) technology, allowing blind users to listen to the content. The goal is to develop a reliable, intelligent system that improves accessibility and empowers blind individuals to independently interact with written materials in their daily lives.

➢ The paper on **"EyeMath: Increasing Accessibility of Mathematics to Visually Impaired Readers"** [25] is focused on the development of a system that will make mathematical content more accessible to visually impaired individuals. EyeMath is designed to address the problems blind or visually impaired readers encounter when trying to interact with complex mathematical formulas, symbols, and equations. It probably uses technologies including OCR and TTS to recognize and read mathematical expressions aloud in a meaningful and usable form for the visually impaired. EyeMath intends to make educational materials more accessible so that visually-impaired people can cope with mathematics on their own, well, independently and effectively.

➢ The paper, **"Cloud-Based Text Extraction Using Google Cloud Vision for Visually Impaired Applications,"** [26] develops a cloud-powered solution to aid visually impaired users by extracting text from images or documents. It uses Google Cloud Vision, which is a strong tool for image recognition, in detecting and extracting text from pictures of printed material, such as books, signs, or documents. Once the text is extracted, it can be converted into speech using Text-to-Speech technology, so the visually impaired users can hear it. The system is cloud-based, which makes it possible to process large volumes of data and access the service from different devices, providing a scalable and convenient solution that improves accessibility and independence for visually impaired users.

➢ The paper on **"Efficient Portable Camera-Based Text to Speech Converter for Blind Person"** [27] focuses on the development of a portable and user-friendly system that enables the blind to access printed text. The system utilizes a portable camera, such as a smartphone or a dedicated handheld device, to capture images of text from documents, books, or signs. The captured images are then processed using Optical Character Recognition (OCR) technology to extract the text. After recognizing the text, the system converts it into speech with Text-to-Speech (TTS) technology, through which blind people can listen to the content. The purpose is to provide an efficient, on-the-go solution that improves the independence and accessibility of blind people, enabling them to read printed material anywhere and anytime.

➢ The paper **"Low-Cost Braille Printer Prototype Design With OCR Technology"** [28] discusses creating an affordable and effective Braille printing solution for visually impaired people. The system uses Optical Character Recognition (OCR) technology to scan and recognize printed text from documents or books. After recognizing the text, the system converts it into Braille code. A low-cost Braille printer prototype prints Braille characters on paper, enabling visually impaired users to read the content by touch. It aims to offer a budget-friendly tool that improves access to Braille materials, helping visually impaired individuals read written information independently.

➢ The paper **"Document Image Classification: Towards Assisting Visually Impaired"** [29] aimed to create a system that classifies document images to help visually impaired people read and organize printed materials. The system was designed to use image classification to automatically identify a document and its content. This makes it easier for visually impaired users to access different types of documents. The system can classify the content and then convert text to speech using TTS technology, allowing users to listen to the document's contents. It incorporates Tesseract OCR, CNN, LSTM, RCNN, and Graph-Based Models. Essentially, the goal is to help visually impaired individuals navigate various documents independently and efficiently.

➢ This paper, titled **"Project Bhashitha – Mobile Based Optical Character Recognition and Text-to-Speech System,"** [30] aims at the development of a mobile-based application that serves as an OCR-TTS mobile assistant for vision-impaired persons. The scanned text from papers, books, or other literature using the cell phone camera shall be read using TTS and shall be voiced over. The text is read, and its content is listened to by users after it is recognized. It aims to offer a solution for the visually impaired, accessible, portable, and user-friendly that enhances their interaction with written material and increases independence.

III. IMPLEMENTATION

Broadly, the whole process is broken into three portions. Section 1 is used to brief regarding the image input, Section 2 use to brief of the major work and operations processed by the system to acknowledge the text appearing in the form of an image and get the textual section appearing in it, and finally, Section 3 describes a technique used here to convert input textually and then processed to receive its speech output.

A. Image Acquisition

The input to the OCR system is a scanned document of the image. The input image should be in a specific format, such as .jpeg, .png, or .bmp. This can be obtained using a scanner, digital camera, or any other digital input device. The image processing library focuses on real-time computer vision and applies to many areas, like 2D and 3D feature toolkits, facial and gesture recognition, human-computer interaction, mobile robotics, object detection, and more. The image processing is done using the OpenCV library.

B. Pre-processing

In this module, the input given is the image captured by the scanned document and the output will be the Pre-processed Image. The raw data is consequently put through a series of preliminary processing steps in order to be used in the descriptive stages of character analysis. Gray - scale conversion, Noise Reduction, Binarization, Edge Detection are the preprocessing steps in the proposed method. For binarization, Otsu's thresholding technique is used, edge detection is performed by Canny Edge Detection, and noise removal can be achieved through Gaussian blur or Median filtering. The goal of pre-processing is to create data that are easy for the OCR systems to work correctly.

- **Grayscale Conversion:** Convert the scanned image or document into a grayscale format, which simplifies further analysis by removing color information.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14728

- **Noise Reduction:** Apply methods like Gaussian blur or Median filtering to remove noise from the scanned image.
- **Binarization:** Use techniques such as Otsu's thresholding to convert the image to black and white, making the text more distinct from the background.
- **Segmentation:** Segmentation is the most important step in pre-processing. It lets the recognizer to pick out features from each character. In the case of more complicated handwriting, the segmentation problem is much more difficult because letters are likely to be joined together, overlapped or skewed. Segmentation is carried out to extract the single character, single word and single line of text from the input document.

In the case of isolated characters or digits, the segmentation process is less complicated. In the segmentation-based OCR system, the most important step initially is to identify the individual elements, suitable for the final recognition process (OCR). Such an approach is strongly dependent on the individual character extraction from the text. The proposed method uses character segmentation which extracts characters from words. Character segmentation is a difficult stage of OCR systems since it isolates useful areas for processing. This stage splits the images into classifiable units called characters.

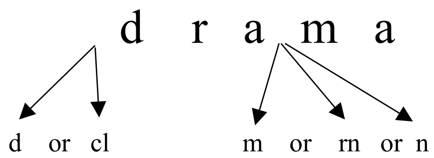

Fig 1. Character Segmentation

- **Edge Detection:** Use Canny Edge Detection to detect edges in the image, making it easier for the OCR to distinguish the characters.

C. Feature Extraction

Character presents at the feature extraction stage. It is represented as a feature vector that identifies it. The main goal of feature extraction is to gather a set of features that boosts the recognition rate while using the fewest elements. To produce, most instances of the same symbol, come up with similar features. Feature extraction methodology examines the input document image and chooses a collection of features that uniquely identify and classify the character. The feature extraction stage gives us the feature vector used for classification. To do the classification we must have a data set to compare with many feature vectors. Tesseract OCR is a widely used open-source OCR engine that can recognize printed text from images.

D. Classification and Recognition

Classification is a decision-making step in the OCR system that makes use of the features extracted from the one previous stage in the process, which requires a classifier to compare the feature vector of input and the feature vector of the data set. CNN algorithms can be used to classify characters from segmented images.

E. Generating Text

We process the images and convert them into text. OCR involves detecting the text on images and translate the images into encoded text that the computer can easily understand. Here we are using LSTM for Handwriting Recognition which process sequential data like handwriting and convert it into text and gTTS (Google Text-to-Speech), a python library interface which Google Translate's text-to-speech API.

There are many APIs to convert text to speech in Python. gTTS is a simple tool that turns entered text into audio, and you can save it as an MP3 file. The gTTS API supports several languages, including English, Hindi, Tamil, French, and German. It is possible to speak as fast or slow. Till the latest version update, it has not been possible to change the voice of generated audio.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14728

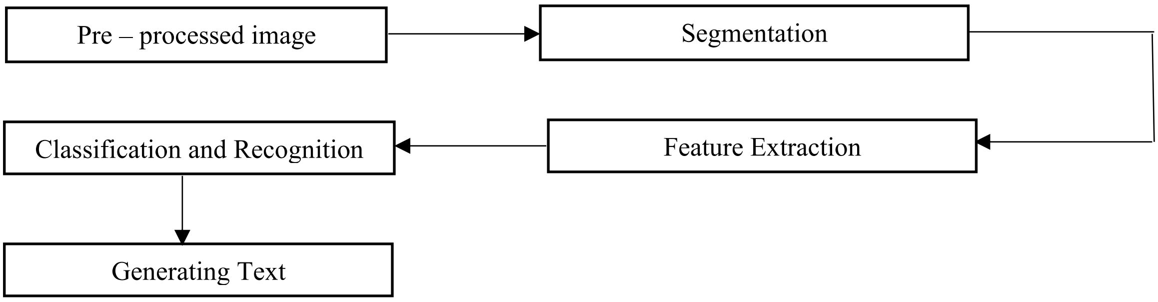

Fig 2. Stages of Implementation

IV. COMPARISON AND RESULTS

Ref. no	Advantages	Disadvantages	Methodology
Vijay Kumar MS et.al [1]	User-friendly smartphone integration, affordable solution	Limited to mobile devices, relies on consistent image capture	Smartphone based OCR with CNN and SVM for processing
A.K. Kavitha et.al [2]	Enhances accessibility and independence for visually impaired individuals.	High cost of advanced devices. Limited performance in low-light or high-glare conditions.	Deep learning models, Machine learning, cloud and edge computing
R. Kumaran et.al [3]	Low-cost hardware, assistive for visually impaired	Limited hardware capabilities, needs good lighting	OCR with Node MCU and PyTesseract, YOLO for object detection
Vijayanarayanan A. et.al [4]	Portable and affordable, integrates TTS	Dependent on Raspberry Pi hardware	Google TTS with Webcam, OpenCV, LSTM for Recognition
Lucas Beerens et.al [5]	Highlights security vulnerabilities, regulatory insights	Highly vulnerable to untargeted adversarial attacks	FGSM, C&W, DeepFool for adversarial attacks using TrOCR
Palakolanu Harsha Vardhan Reddy et.al [6]	Integrated wearable smart glasses for hands-free operation. Combines object detection, text recognition and audio feedback for navigation and reading.	Relatively high computational demands due to deep learning algorithms. Dependence on hardware like Raspberry Pi and smart glasses increases cost.	Raspberry Pi board for processing, Tesseract OCR for text recognition, YOLOv5 for object detection and localization.
Ahmed Ben Atitallah et.al [7]	Handles poor-quality images effectively, low complexity	Limited natural scene detection, needs advanced setups	Text recognition using Gamma Correction Method (GCM) and AGCM

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14728

R. Rajan et.al [8]	High accuracy and precision, structured document focus	Resource-intensive, focused on structured documents	Enhanced OCR (EOCR) leveraging Tesseract and deep learning
Tushar Khete et.al [9]	Open-source implementation using Tesseract OCR, which supports multiple languages. Can process images from scanned images and live captures.	Heavily dependent on the quality of input images (e.g., noise, blurriness, resolution). Requires preprocessing steps like binarization and noise removal for optimal results.	Tesseract OCR engine for text detection and recognition and Google Text-to-Speech (gTTS) for audio output
Shruthi K et.al [10]	Combines multiple assistive features, real-time processing	Struggles with small objects or distant text in low light	Tesseract OCR, YOLO, Viola-Jones for face detection and Object detection
Shalini K V et.al [11]	Provides real-time feedback on objects, emotions, and text	Limited robustness in dynamic environments	Tensor Flow, Pyttsx3, GTTS and OCR integrated with Tensor Flow Object Detection API
Sahithya. V et.al [12]	Enables independent shopping for visually impaired users	Limited recognition for handwritten text, requires specific environments	Python-based OCR with Tesseract, MODI for text recognition.
Olumide Olayinka Obe et.al [13]	Implements robust computer vision algorithms like SIFT and FAST.	Performance issues with blurred, transparent, or fancy fonts.	OpenCV, Hessian Matrix using SIFT and FAST
Mohd Nadhir Ab Wahab et.al [14]	Mobile application-based solution, making it accessible and portable. User-friendly design with audio output.	Accuracy depends on preprocessing (e.g., image quality and text type). Language limitation.	OCR Framework and Text-to-Speech (TTS) Framework
Prof. Vaibhav V et.al [15]	Uses open-source tools and affordable hardware components like Raspberry Pi and its camera module.	Struggles with poor lighting, reflective surfaces, and blurred images.	Raspberry Pi camera captures an image of the text and Tesseract OCR
Lalita Moharkar et.al [16]	High recognition accuracy. Supports complex scripts and Offline processing.	Requires large training datasets. Limited real-world testing.	CNN, Max-pooling, fully connected layers.
Srividhya et.al [17]	Empowers visually challenged individuals to access text and perceive emotions of people around them.	Emotion detection may falter in cases of partial or obstructed faces. High computational demand for real-time OCR and emotion detection.	OCR Algorithm, Emotion Detection Algorithm, Text-to-Speech Algorithm
Gauri Vaidya et.al [18]	Provides real-time text recognition and conversion into speech	Performance may degrade under poor	OCR, TTS, Scene text detection algorithms

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14728

	for visually impaired users.	lighting or high-glare conditions.	
Krishnaraj Rao N S et.al [19]	Enhances accessibility for printed and digital content.	Handwritten or overly stylized fonts can reduce recognition accuracy.	OCR, TTS, Object detection and localization
Robert G. de Luna [20]	Provides a seamless pipeline from text recognition to Braille conversion.	Limited by the accuracy of OCR, especially for noisy or complex images.	OCR Engine, Text Preprocessing and Braille Mapping
Muhammad Farid Zamir et.al [21]	Enhances accessibility by converting textual content into audible or tactile formats.	Performance is sensitive to poor lighting or cluttered backgrounds in captured images.	OCR, TTS, Region of interest detection
Kowshik .S et.al [22]	Enables real-time text recognition for the visually impaired and provides tactile and auditory feedback for text interpretation	Requires optimal lighting conditions for accurate text recognition and struggles with complex layouts, stylized fonts, or handwritten text.	OCR with deep learning enhancement, Tesseract, CRNN
Naimul Hassan K M et.al [23]	Dual functionality and cost-effective compared to existing refreshable Braille displays. Supports multiple languages.	Limited to a small dataset and performance depends on lighting and image quality. The OCR and Braille refresh rate could be slow for real-time usage.	Pre-trained AlexNet CNN, Tesseract OCR, Softmax, ReLU
Abhishek Mathur et.al [24]	Allows visually impaired users to "read" printed materials, such as books, menus, or signboards, independently.	Accuracy decreases under low-light conditions or with motion blur during image capture.	Image Processing, Text Detection and Segmentation, OCR algorithms
Kanlayanee Dumkasem et.al [25]	Makes complex mathematical content accessible via audio or tactile feedback	Challenges with recognizing handwritten or highly stylized equations.	OCR for Text and Symbols, Math Expression Parsing, TTS
Vaithiyanathan D et.al [26]	Leverages Google Cloud Vision's advanced OCR capabilities for high precision. Automatic updates and improvements from Google's AI models.	Dependent on a stable internet connection for accessing cloud services. Limited functionality in offline mode unless cached.	OCR, TTS, NLU, Image processing
Trupti Shah et.al [27]	Portable and lightweight design, suitable for on-the-go use.	Sensitive to environmental factors like poor lighting or complex text backgrounds.	OCR, TTS and machine learning

DOI: 10.17148/IJARCCE.2025.14728

Israel Durán Encinas et.al [28]	Affordable alternative to commercial Braille printers.	Limited performance with noisy or unclear input images.	OCR, Braille mapping, text processing
Shahira K C et.al [29]	Allows visually impaired individuals to access and comprehend documents that they would otherwise be unable to read independently.	OCR and image description accuracy can be lower for difficult layouts, handwritten text, or complex graphics. Processing documents in real-time can leads to delays.	Tesseract OCR, CNN, LSTM, RCNN, Graph-Based Models
De Zoysa D.S.S et.al [30]	The system is implemented as a mobile application, making it portable and accessible to users	Performance may be affected by variable lighting, text complexity, and environmental noise.	OCR, Text-to-Speech Conversion, Post-Processing.

V. CONCLUSION

In conclusion the OCR system offers a comprehensive solution for improving accessibility within an academic system, enabling visually impaired students, teachers, and administrators to interact with digital content through text-to-speech and Braille-compatible formats. By leveraging advanced CNN and LSTM techniques for character and handwriting recognition, the system effectively addresses the limitations of traditional OCR methods, providing accurate and reliable recognition even for handwritten text. However, despite these advancements, the existing systems still face some challenges. Traditional OCR systems often struggle with handwritten text, especially when dealing with diverse handwriting styles, leading to inaccurate character recognition. They may also falter when working with low-quality images or documents containing complex layouts, such as tables or multi-column formats, which traditional systems are not designed to handle. Additionally, current OCR solutions can be computationally expensive, requiring significant processing power and storage, especially when dealing with large volumes of data.

Furthermore, while text-to-speech capabilities have improved, they may not always offer the flexibility and personalization needed for all users, especially for those with unique voice preferences or varying levels of visual impairment. Finally, OCR systems can face difficulties with languages that have complex grammar or characters, which may not be adequately supported by current models. The proposed system, by addressing these issues with modern techniques, offers a more efficient and accessible solution, but it still faces the challenge of continually improving accuracy, processing efficiency, and adaptability to diverse needs within the educational environment. Through continuous innovation, research, and a strong focus on accessibility, the proposed OCR system strives to revolutionize the educational experience for visually impaired students, teachers and administrators. With ongoing improvements and adaptability, this system aspires to create a more inclusive and equitable learning landscape for all.

REFERENCES

[1]. Vijay Kumar MS, Renuka, Sanchitha S Mourya, Sanjana J, Amith S. (2024). Smart Translation For Visually Impaired People. *International Research Journal of Modernization in Engineering Technology and Science. vol.6.* https://www.irjmets.com/uploadfiles/paper//issue_5_may_2024/ 55582/final/fin_irjmets1715229437.pdf

[2]. Kavitha A.K, Hariharaeshwar K, Karthick S, Subramaniam S. (2024). Intelligent Reader For Visually Impaired Person. *International Journal of Creative Research Thoughts (IJCRT). vol.12.* https://www.ijcrt.org/papers/IJCRT24A4470.pdf

[3]. Kumaran .R, Brindha .M, Harunyaa .P, Archana .S (2023). Text/Object Recognition For Visually Impaired People. *International Journal of Creative Research Thoughts (IJCRT). vol.11.* https://ijcrt.org/papers/IJCRT2304264.pdf

[4]. Vijayanarayanan A, Savithiri R, Lekha P, Abbirami R S. (2023). Image Processing Based on Optical Character Recognition with Text to Speech for Visually Impaired. *Journal of Science, Computing and Engineering Research (JSCER).vol.6.* https://doi.org/10.46379/jscer.2023.060414

[5]. Lucas Beerens, Desmond J. Higham. (2023). Vulnerability Analysis of Transformer-based Optical Character Recognition to Adversarial Attacks. https://doi.org/10.48550/arXiv.2311.17128

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14728

[6]. Palakolanu Harsha Vardhan Reddy, Bhumireddy Kashyap, Bhukhya Charan Naik, Prashant Pal, Shashank Kumar Singh, Saurabh Bansod, Yogesh Kumar. (2023). Virtual Assistant and Navigation for Visually Impaired using Deep Neural Network and Image Processing. *Research Square.* https://doi.org/10.21203/rs.3.rs-2867156/v1

[7]. Ahmed Ben Atitallah, Mohamed Amin Ben Atitallah, Yahia Said, Mohammed Albekairi, Anis Boudabous, Turki M. Alanazi, Khaled Kaaniche ,Mohamed Atri. (2023). An Efficient Text Recognition System from Complex Color Image for Helping the Visually Impaired Persons. *Tech Science Press vol. 46.* pp. 702-717, DOI: 10.32604/csse.2023.035871

[8]. Rajan.R, Geetha Devasena M.S. (2022). A Novel Reading Technique for Visually Impaired Person Using Enhanced Optical Character Recognition Method. *Neuro Quantology.vol.20,* pp. 3348-3365, doi: 10.14704/nq.2022.20.9.NQ44386

[9]. Tushar Khete, Aditya Bakshi. (2022). Autonomous Assistance System for Visually Impaired using Tesseract OCR & gTTS. *4th International Conference on Intelligent Circuits and Systems*. doi:10.1088/1742-6596/2327/1/012065

[10]. Shruthi K, Sneha Datta H, Sukruthi A, Vaishnavi Srikanth, Dr. Shivakumaraswamy P.M, Gouri Katageri . (2021).A study on face, expression, object and character recognition to implement an aid for the visually impaired. *2nd International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS 2021).* http://dx.doi.org/10.2139/ssrn.3882625

[11]. Shalini K V, Sai Srujana P H, Shalini B R, Shalini V, V Chandana Chowdary.(2021). Optical Character Recognition On live Footages for Visually Impaired. *International Research Journal of ModernizationinEngineeringTechnologyandScience.vol.3*.https://www.irjmets.com/uploadfiles/paper/volume3/issue_6_ june_2021/12428/1628083496.pdf

[12]. Sahithya. V, Sowmiya. S, Vinothini. P, Ohm Sankar .S. (2021). Ocr Using Visually Impaired People Shopping Trolley Technology.vol.9 .DOI 10.17148/IJIREEICE.2021.9410

[13]. Olumide Olayinka Obe, Akinwonmi A. E. (2021). Smart Application For The Visually – Impaired. *International Journal of Advanced Research in Computer Science. vol. 12.* http://dx.doi.org/10.26483/ijarcs.v12i2.6690

[14]. Mohd Nadhir Ab Wahab, Ahmad Sufril Azlan Mohamed, Abdul Syafiq Abdull Sukor, Ong Chia Teng. (2021). Text Reader for Visually Impaired Person. *5th International Conference on Electronic Design (ICED).* doi:10.1088/1742-6596/1755/1/012055

[15]. Prof. Vaibhav V. Mainkar, Tejashree U. Bagayatkar, Siddhesh K. Shetye, Hrushikesh R. Tamhankar, Rahul G. Jadhav, Rahul S. Tendolkar. (2020). Raspberry pi based Intelligent Reader for Visually Impaired Persons. *Second International Conference on Innovative Mechanisms for Industry Applications IEEE*. DOI: 10.1109/ICIMIA48430.2020.9074948

[16]. Lalita Moharkar, Sudhanshu Varun, Apurva Patil, Abhishek Pal. (2020). A scene perception system for visually impaired based on object detection and classification using CNN. *ITM Web of Conferences*. doi: https://doi.org/10.1051/itmconf/20203203039

[17]. Srividhya, Ravichandran. (2020). Ocr Based Facilitator with Emotion Detection for the Visually Challenged. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH. vol.9*. DOI: 10.36106/ijsr

[18]. Gauri Vaidya, Ketki Vaidya, Kishor Bhosale. (2020). Text Recognition System for Visually Impaired using Portable Camera. *IEEE International Conference on Convergence to Digital World – Quo Vadis (ICCDW 2020).* DOI: 10.1109/ICCDW45521.2020.9318706

[19]. Krishnaraj Rao N S, Sandeep Kumar S, Santhosh S, Sudesh Rao, Krishna Kaushik P. (2020). Smart Reader for Visually Impaired. *Fourth International Conference on I-SMAC IEEE.* DOI: 10.1109/I-SMAC49090.2020.9243598

[20]. Robert G. de Luna. (2020). A Tesseract-based Optical Character Recognition for a Text-to-Braille Code Conversion. *International Journal on Advanced Science Engineering Information Technology, vol.10*, no.1. https://ijaseit.insightsociety.org/index.php/ijaseit/article/view/9956/pdf_1311

[21]. Muhammad Farid Zamir, Khan Bahadar Khan, Shafquat Ahmmad Khan, Eid Rehman. (2020). Smart Reader for Visually Impaired People Based on Optical Character Recognition. *Springer Nature Singapore Pte Ltd. 2020.* https://doi.org/10.1007/978-981-15-5232-8_8

[22]. Kowshik .S, Gautam V.R, Dr.Suganthi .K. (2019). Assistance For Visually Impaired Using FingerTip Text Reader Using Machine Learning. *11th International Conference on Advanced Computing (ICoAC).* DOI: 10.1109/ICoAC48765.2019.246808

[23]. Naimul Hassan K M, Subrata Kumar Biswas, Md Shakil Anwar, Md Shakhrul Iman Siam, Celia Shahnaz. (2019). A Dual-Purpose Refreshable Braille Display Based on Real Time Object Detection and Optical Character Recognition. *IEEE International Conference on Signal processing, Information, Communication &Systems.* doi:10.1109/SPICSCON48833.2019.906510

[24]. Abhishek Mathur, Akshada Pathare, Prerna Sharma, Sujata Oak. (2019). AI based Reading System for Blind using OCR. *Third International Conference on Electronics Communication and Aerospace Technology IEEE.* DOI: 10.1109/ICECA.2019.8822226

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14728

[25]. Kanlayanee Dumkasem, Padchaya Srisingchai, Prapaporn Rattanatamrong. (2019). EyeMath: Increasing Accessibility of Mathematics to Visually Impaired Readers. *2019 23rd International Computer Science and Engineering Conference (ICSEC)*. DOI: 10.1109/ICSEC47112.2019.8974682

[26]. Vaithiyanathan D, Manigandan Muniraj. (2019). Cloud based Text extraction using Google Cloud Vison for Visually Impaired applications. *2019 11th International Conference on Advanced Computing (ICoAC).* DOI: 10.1109/ICoAC48765.2019.246822

[27]. Trupti Shah, Sangeeta Parshionikar. (2019). Efficient Portable Camera Based Text to Speech Converter for Blind Person. *International Conference on Intelligent Sustainable Systems IEEE*. DOI: 10.1109/ISS1.2019.8907995

[28]. Israel Durán Encinas, Andrés Sandoval Bringas. J, Ivan De Casso Verdugo. A, Carlos Sandoval Bringas, Jonathan Giovanni Soto Muñoz. (2019). Low-cost Braille Printer Prototype Design With OCR Technology. *2019 International Conference on Inclusive Technologies and Education (CONTIE).* DOI: 10.1109/CONTIE49246.2019.00047

[29]. Shahira K C, Lijiya A. (2019). Document Image Classification: Towards Assisting Visually Impaired. *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON).* DOI: 10.1109/TENCON.2019.8929594

[30]. De Zoysa D.S.S, Sampath J.M, De Seram E.M.P, Dissanayake D.M.I.D, Wijerathna .L, Thelijjagoda .S. (2018). Project Bhashitha – Mobile Based Optical Character Recognition and Text-to-Speech System. *IEEE 13th International Conference on Computer Science & Education.* DOI: 10.1109/ICCSE.2018.8468858

[31]. Sunitha S. Nair, P. Ranjit Jeba Thangaiah, (2024). A Novel Approach: CNN-RNN and Bi-GRU for Handwritten Character Recognition. *SSRG International Journal of Electronics and Communication Engineering,* vol. 11, no. 1, pp. 1-14, 2024. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V11I1P101

[32]. Nair, S.S. ., & Thangaiah, P. R. J.. (2024). Deciphering excellence: Comparative study of Deep Learning Models in Handwriting Recognition. *International Journal of Intelligent Systems and Applications in Engineering, 12*(16s), 602–614. Retrieved from https://www.ijisae.org/index.php/IJISAE/article/view/4898