iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** IX **Month of publication:** September 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.55673

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue IX Sep 2023- Available at www.ijraset.com*

# Alpha-Numeric Character Recognition for Blinds Using Deep-Learning Techniques

Somanna B M[1], K R Sumana[2]
[1]*PG Student,* [2]*Faculty, The National Institute of Engineering, Mysuru, Visveswaraya Technological University, Belgaum, Karnataka*

***Abstract:*** ***In this study, we are introducing an alphanumeric character recognition system using Deep Learning concepts. This system will classify Alphabets and Numeric characters with high accuracy. Identification and classification of characters would be essential for users with who have problem reading the characters in a document or any other text-based media. We utilize a dataset containing of letters A to Z with a total of 3,89,765 images which is approximately 15000 images per Alphabet. We classify the characters based on the shape of each character and using Deep Learning concepts like CNN, RNN and Random Forest architecture to train out the dataset and develop a system which produces the highest accuracy. The success rate in identifying the characters makes it a very easy and an efficient tool.***
***Keywords: Alphanumeric Character, CNN, RNN, Random Forest, Deep Learning, Recognition.***

## I. INTRODUCTION

Visually impaired or blind people have difficulty in reading documents or texts in situations where reading texts in newspapers ,important documents, letters can be hard in non-ideal conditions where low light or far sightless can hinder one's ability to read or recognize certain texts. Numerous strategies have been used to resolve the problem in helping visually impaired personnel with an optimized a collection of implements that can help lead a comfortable life. With the advancement in Deep Learning techniques and the algorithms several strategies can now be implemented for processing the images. Some notable Deep Learning techniques such as CNN and RNN have been implemented to further address the concern. Developing the Alphanumeric Character-Recognition system for the visually-impaired using Deep-Learning Techniques stems from the fundamental principle of inclusivity and accessibility. Visual impairment affects millions of people globally, limiting their ability to access printed text and written information independently. Among the critical challenges faced by persons with visual impairments is recognizing alphanumeric characters, such as letters and numbers, in everyday life. This task is essential for reading signs, labels, documents, and various forms of printed text, which is necessary for independence, education, employment, and social engagement. Consequently, there is a pressing need for a dedicated solution that leverages the power of Deep Learning Techniques to recognize alphanumeric characters accurately and efficiently, providing a practical and user-friendly tool for the visually impaired.

## II. LITERATURE SURVEY

A remedy for the issue for visually impaired to help with their difficulties in reading texts in non-ideal conditions that may include changing lighting conditions, the different font styles used in the document and the direction of the characters. The proposed system [1], developed using the MATLAB platform, serves as an affordable substitute, making it accessible to a broader user base. Its core functionalities rely on two key components: Optical Character Recognition (OCR) and Text-to-Speech technology. Furthermore, the authors have ingeniously integrated a prediction algorithm into the system. This algorithm enhances the reading experience by predicting the subsequent character in a word, thus reducing errors, and ensuring a seamless conversion into text. This error reduction is crucial for minimizing disruptions during the subsequent Text-to-Speech conversion. It transforms the recognized characters into spoken words, allowing visually impaired individuals to engage with the document's content through auditory means. This ensures an undisturbed and user-friendly reading experience, enabling users to comprehend the text without reliance on visual. The Author used a prediction algorithm to help predict the following character in a word to provide Error less text that may use by the Text to Speech tool for easy and undisturbed conversion.

A system that uses the images and extract details contained in the images that includes recognition of texts and objects that would aid in understand the surroundings. his system leverages cutting-edge deep learning techniques [2] to extract meaningful details from images, aiding in their understanding of the surrounding environment. At the core of this system is the utilization of Convolutional-Neural-Networks and Long Short-Term Memory (LSTM) networks for extraction of relevant features and text-to-speech conversion.

©IJRASET: All Rights are Reserved 

These advanced neural network architectures, including GoogleNet. AlexNet, ResNet, and VGG16, play a crucial part in recognizing and interpreting textual information within images. By harnessing the power of these neural networks, the system can accurately identify and extract text, enabling further processing. this system on a Raspberry Pi, making it a portable and cost-effective solution. The most significant achievement of the referred system lies in its accuracy. This high level of accuracy ensures that the information provided to visually impaired users is reliable and valuable, reducing the likelihood of errors that could hinder their understanding.

This paper focuses on this critical age group and leveraging technology, the "Third Eye" system [3] effectively bridges a gap in infrastructure for the blind children. It empowers them to engage with the world of alphanumeric characters, fostering independence and literacy skills from an early age. One of the key features of this system is its integration of Google API, which plays a pivotal role in converting recognized texts into speech. This functionality ensures that visually impaired users receive audible feedback about the characters they've written, contributing to a more interactive and engaging writing experience. Furthermore, this system aligns with the broader goals of inclusivity and accessibility, ensuring fair opportunities for all youngster's opportunities to explore and express themselves through writing, regardless of their visual abilities.

This paper presents an innovative Android application [4] designed specifically to address the requirements of people with vision impairments. The primary focus of this system is to visually impaired users to recognize textual content within various media, including printed and scanned documents. At the core of this application lies a method known as COCO, which stands for Common-Objects-in-Context. Moreover, it is cost-effective, which means that it can be easily adopted by a broad user base, making it an inclusive solution for the community of someone who is blind. The application operates in conjunction with Google Voice Assistant, responding to audio commands submitted by the client. This voice-controlled functionality enhances its usability and ensures that it can be navigated without reliance on visual or tactile interfaces. While it has a limitation in object recognition distance, its compactness, affordability, and overall usability make it a promising solution to enhance the quality of life for those with vision impairments.

## III. PROPOSED WORK

The proposed system is an Alphanumeric Character Recognition solution designed primarily for visually impaired individuals. It leverages advanced deep-learning techniques to accurately identify and interpret alphanumeric characters from various sources such as printed text, handwritten notes, and digital documents. The system aims to enhance the independence and accessibility of visually impaired users by providing real-time character recognition using cameras. The proposed system addresses the social concern of inclusivity by making character recognition technology accessible and affordable to the visually impaired community. It promotes independence, education, and increased access to information, ultimately increasing the consumers' quality of life. Additionally, the system adheres to ethical guidelines, prioritizing user privacy and data security, and continuously seeks user feedback for enhancements.

## IV. METHODOLOGY

In our work, we leverage Python's robust ecosystem, including key libraries like PyTorch, NumPy, Pandas, Scikit-Learn for data preprocessing and neural network construction. TensorFlow enhances computation with GPU acceleration. Our methodology focuses on deep-learning techniques, training Convolutional Neural Networks (CNNs) Recurrent Neural Networks (RNNs) and Random-Forest Model on labelled data for recognizing intricate patterns in captured texts. The system initiates by capturing video through a strategically positioned webcam, aimed at capturing texts. Following a successful identification of Alphanumeric characters, the system provides the result in a speech-based format.

### *A. Video Capture Unit*

The video capture device records video frames in real-time by positioning a camera in front of the user to capture the text. The video is periodically sampled, and the sampled frames are transmitted to the Alphanumeric Character detection system. These frames are made up of RGB images, and in low-light conditions, we apply adjustments to enhance image quality and lower noise. Prioritizing noise reduction in the image before performing contrast enhancement increases the accuracy of our method. This sequence is crucial as it ensures noise is removed before enhancing the image through contrast adjustments. By doing so it significantly enhances our system's accuracy by eliminating heavy noise, preventing texture blurring, and averting over-enhancement in the processed image. To enhance character detection, we convert the image to grayscale. The image is then transferred to the Alphanumeric character recognition system.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

*B. User Module*

The User Module in the proposed system for the visually impaired plays a pivotal role in ensuring a seamless and user-centric experience. This module is specifically designed to cater to the needs and capabilities of visually impaired users, focusing on usability, accessibility, and user empowerment. Real-time character recognition is made possible by the system's usage of the available camera, which gives users access to a live video stream from the camera on their device. This function guarantees that users can instantly capture and interpret text. The system uses Text- to-Speech tool to convert the recognized text to a speech format that helps visually impaired users to listen to recognized texts. The system must have a feedback mechanism that notifies the user if texts cannot be recognized using a timer to notify the user once every 5 to 7 seconds and is as shown in figure 1.

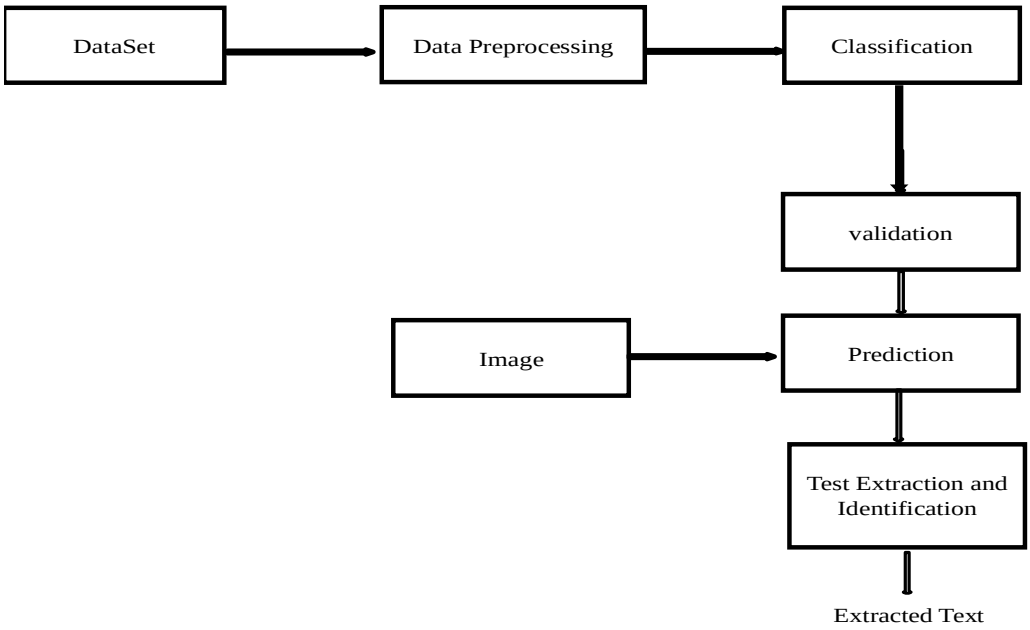



Fig 1 : System Architecture Of The Proposed System

## V. RESULT ANALYSIS

The following analysis of our work gives the self-explanatory for the data values that we collected dataset values.

*A. Accuracy and Loss of the CNN Model*

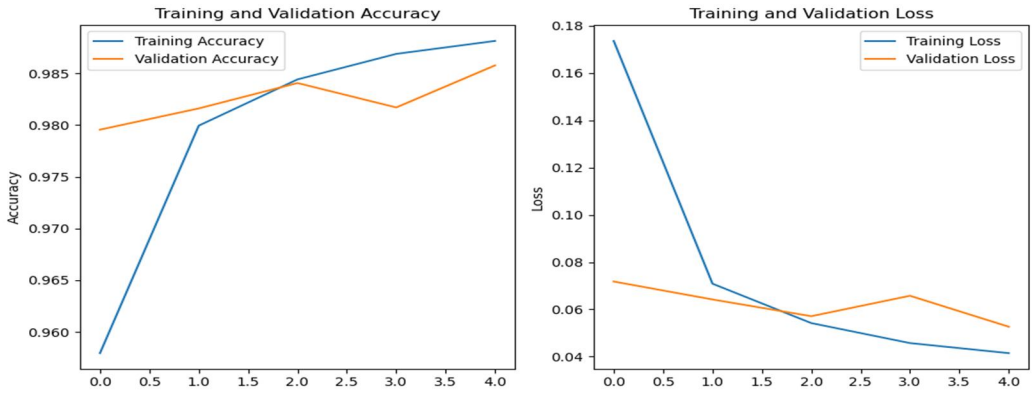



Fig 2: Accuracy and loss of the CNN model

### *B. Accuracy and Loss of the RNN Model*

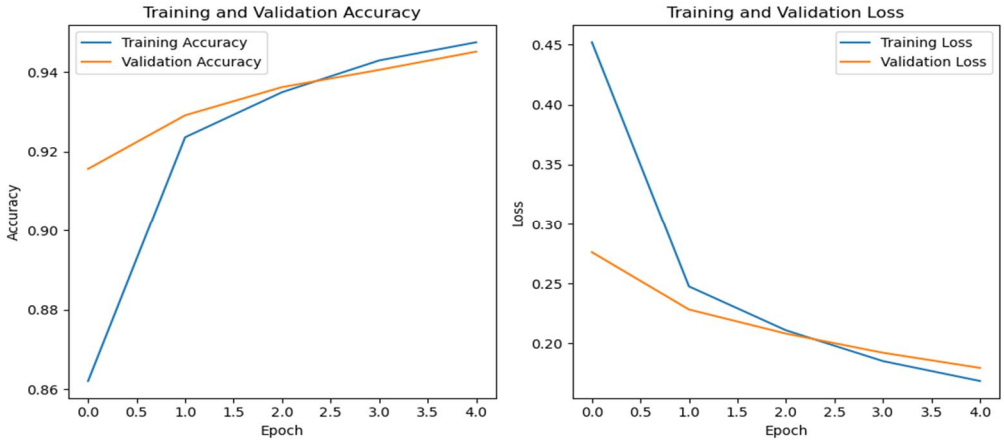



Fig 3: Accuracy and loss of the RNN model

### *C. Accuracy and Loss of the Random Forest Model*

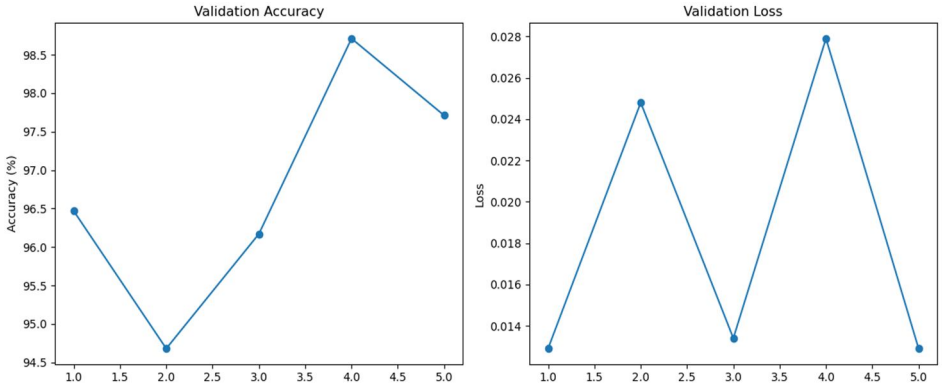



Fig 4: Accuracy and loss of the Random Forest model

The Project is implemented with CNN, RNN, FRCNN. These algorithms have been he subject of a comparison study. As graph indicates, CNN is having more accuracy i.e., 98.58 when compared to RNN and Random Forest.

Table I: Accuracy Score Of Different Models

| Algorithms | Accuracy(%) |
|---|---|
| CNN | 98.58 |
| RNN | 94.51 |
| Random Forest | 96.74 |

The below figure Fig 6 shows the system recognising documents from smart phone and predicting results.

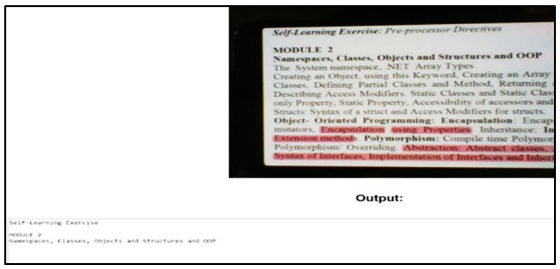



Fig 5: Recognizing text document from a smart phone.

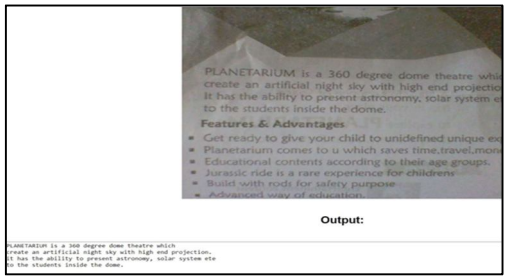



Fig 6: Recognizing text from a leaflet.

©IJRASET: All Rights are Reserved 

The below figure 7 shows the system recognising text from a leaflet and predicting results successfully.

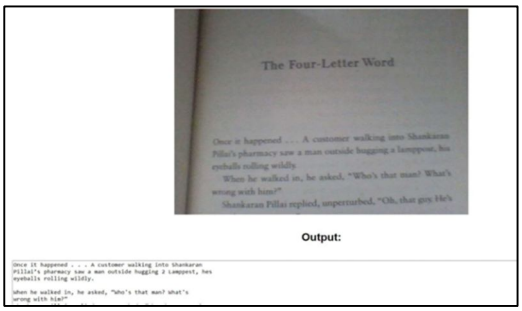



Fig 7: Recognizing text from a Novel.

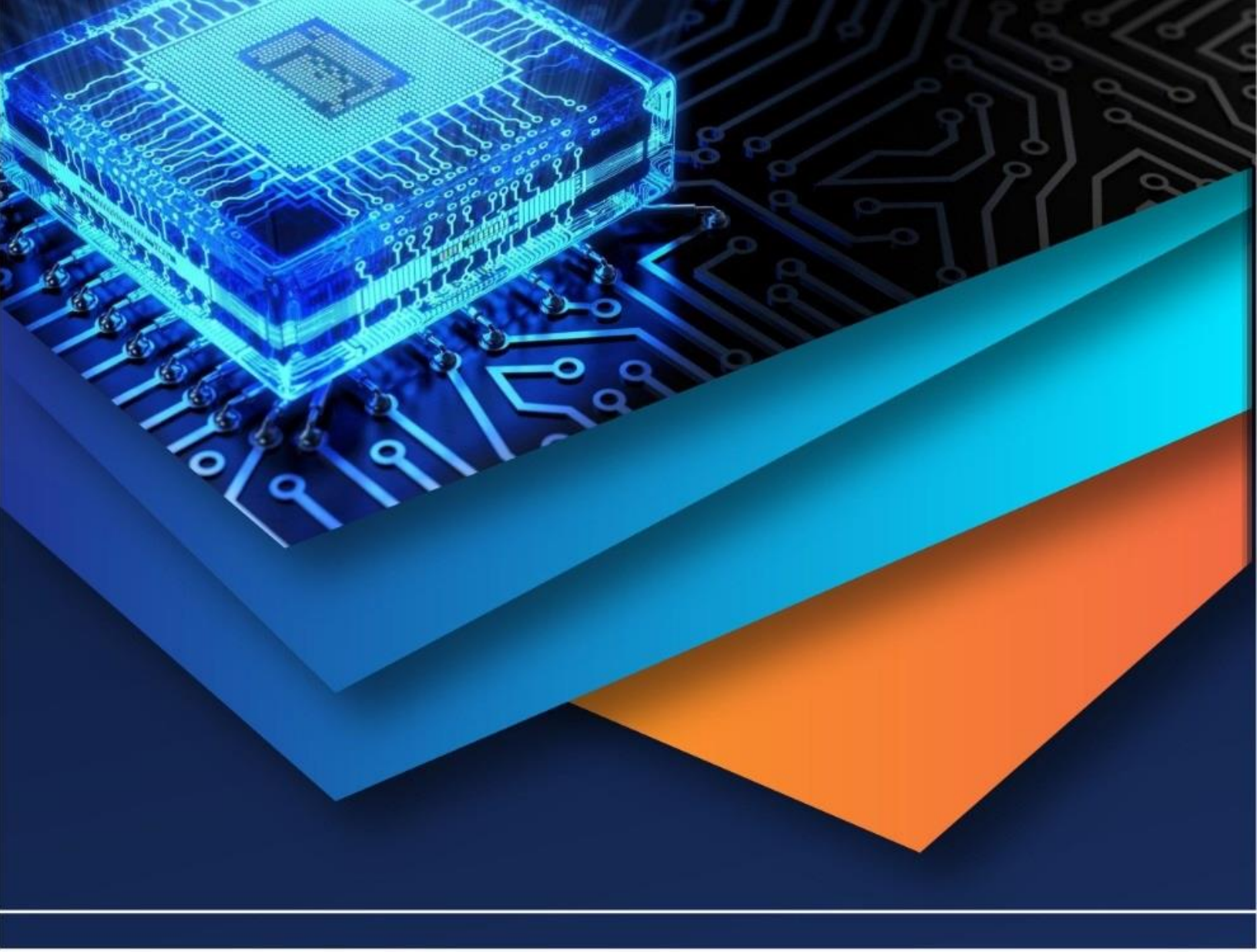
The above figure 8 shows the system recognising text from a Novel and predicting results successfully.

## VI. CONCLUSION

In summary, the system incorporates live video optical-character-recognition (OCR) using deep-learning techniques which offers a valuable solution for enhancing accessibility and engaging with the people who are visually impaired. By harnessing the power of modern computer vision and natural language processing technologies, this system enables real-time recognition of alphanumeric characters from a live video stream. The system's effectiveness has been enhanced for efficiency, ensuring rapid character recognition and minimal processing delay. It provides reliable results through the combining of robust deep learning models and Tesseract OCR, bolstered by the Python programming language. Maintenance of the system is facilitated by its modular design, allowing for straightforward updates and improvements in response to evolving user requirements and technological advancements. The user interface is designed for usability, offering a simple and intuitive experience, while voice feedback enhances the system's accessibility. Overall, this system represents a commendable effort to address the needs of visually impaired individuals by providing them with an innovative tool for real-time character recognition from live video sources, thereby fostering greater independence and inclusivity in an increasingly digital world.

## REFERENCES

[1] “An Easy to Use Versatile Reader for Blind People and Partially Blind People”, Pooja Patel, Nabila Shaikh
[2] “Deep Learning Reader for Visually Impaired”, Jothi Ganesan , Ahmad Taher Azar , Shrooq Alsenan, Nashwa Ahmad Kamal , Basit Qureshi and Aboul Ella Hassanien.
[3] “Third Eye – A Writing Aid for Visually Impaired People”, M.S Divya Rani, T K Padma Gayathri, Sreelakshmi Tadigotla, Syed Abrar Ahmed.
[4] “Blinds Personal Assistant Application for Android” , Dr. Usha B A , Prof. Sangeetha K N , Mr. Vikram C M.
[5] “Third Eye Smart Aid for Visually Impaired”, Sahana V ,Shashidhar ,Bindu S. N, Chandana A. N. ,Nishrutha C. G



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)