International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Survey on Intelligent Reader for Visually Impaired People

Anusha D, K U Anjali
Department of Computer Science And Engineering
Sapthagiri college of engineering
Bengaluru India

Revathi D, Yashaswini S
Department Of Computer Science And Engineering
Sapthagiri college of engineering
Bengaluru India

Under the guidance of
Prof. Praveen Kumar. K.V
Department of Computer Science And Engineering
Sapthagiri college of engineering
Bengaluru India

Abstract— OCR (Optical Character Recognition) is a process or technology that recognizes text inside a digital image. It is mostly used to convert transcribed, handwritten, or printed text into editable and reuseable text data. People want faster, more convenient, more dependable tools to meet their daily needs as technology advances. With this philosophy in mind, we went ahead and researched the available tools before creating this WebApp, which gives a smooth experience (no advertisements and easy-to-use) as well as high accuracy. Although OCR technology was designed to recognise printed text, it may also be used to recognise and verify handwritten content. The goal of this project is to enable for the automatic extraction of data.

Keyword →Optical Character Recognition, Tesseract, Handwritten Text Recognition, Image to Text

I. INTRODUCTION

Optical Character Recognition (OCR) is a technology that converts transcribed, handwritten, or printed text documents, such as scanned pages or photographs acquired with any camera or phone, into editable and reuseable text data. [1]. To put it another way, OCR examines a photograph of a text document (thus the name "optical" process) and then recognises the various alphabets, numerals, and other characters. This subprocess is known as character recognition, and it is used to extract characters from an image before converting them to text sentences for further use. This primarily tries to reduce human workload, which it accomplishes because it is convenient and time-saving.

Figure 1 depicts the traditional OCR workflow paradigm, which includes nine processes (excluding the first and end phases) for extracting text from a document. Preprocessing, segmentation, feature extraction, classification, and recognition are the major five processes in this procedure.workload, which it accomplishes because it is convenient and time-saving.
Figure 1 depicts the traditional OCR workflow paradigm, which includes nine processes (excluding the first and end phases) for extracting text from a document. Preprocessing, segmentation, feature extraction, classification, and recognition are the major five processes in this procedure.

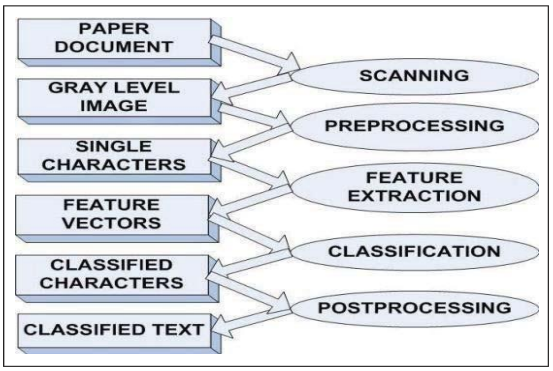

Fig. 1. Model of the traditional OCR method

II. RELATED WORKS

Some of the recent works related to this topic are:

1. Jaewoo Park[1]proposed a multilingual OCR system integrating three neural blocks (a segmenter, a switcher, multiple recognizers for different languages), and the reinforcement learning of the segmenter. Unlike conventional methods for multi-language OCR systems, we aimed to optimize the edit distance of recognition results (i.e., the overall performance of OCR systems) and achieved the goal with the reinforcement learning method. Experimental results show that our method outperforms conventional methods.

2. Ayatullah Faruk Mollah,,Subhadip Basu and Mita Nasipuri,[2]proposed that their study limited to light-weight and computationally efficient techniques Compared to Tesseract, acquired recognition accuracy (92.74%)is good enough. Experiments shows that the

recognition system presented in this paper is computationally efficient which makes it applicable for low computing architectures such as mobile phones, personal digital assistants (PDA) etc

3. Abin M Sabu and Anto Sahaya Das [3] proposed various modern techniques were introduced to remove the noise and to recognise the characters. Each step in optical character recognition is important such that each steps are interrelated to one other and to obtain an accurate result, results at each stage should be observed.

4. S.Gayathri and R.S.Mohana,[4]It is found that fine tuning of hper -parameters like learning rate, activation function helps in enhancing the frutition of CNN architecture. when adam optimizer is used, training and validating accuracy obtained was above 95% which is outplay than existing system results.

5. Khaustov, . Spitsyn and . Maksimova [5], This method is based on structural components extraction and has no requirements to the number of reference images. The proposed approach outperforms its analogs in conditions of a small number of reference images. Another advantage of our method is its high performance as well as low memory consumption.

6. Muiz Ahmed Khan , Pias Paul , Mahmudur Rashid[6], The processed image is then fed to the Tesseract optical character recognition (OCR) engine to extract the text from it The ultrasonic sensors can detect an object within 300 cm by generating a 40 kHz signal and receiving reflected echo from the object in front of it. The distance is calculated based on the pulse count and time-of-flight (TOF).The best image is then fed to a convolutional neural network (CNN),which is trained on big data and runs on a cloud device. The audio feedback in most devices is provided through a headset or a speaker. The audio is either a synthetic voice from the text-to-speech synthesis system or a voice user interface generating a beep sound.

7. I N 1914 Fournier d'Albe [7], Demonstrated a direct translation reading aid for the blind with which printed material was moved horizontally across an image of a vertical row of illuminated dots. The light in each dot was chopped at a different rate. several common properties that these reading aids share are 1) maximum reading rates obtained by the majority of the subjects have been less than 10 correct words per cause of the 10 words per minutelimitation on reading rate has been due to poor resolution.

8. Joselin Villanueva and René Farcy[8], Proposed, An active optical pathfinder using a LED and a photodiode is implemented as an electronic travel aid to improve the mobility of persons who are blind The aim of these devices is to protect users from obstacles that are at knee to head level. The device can be fastened to the cane or held by the user. It gives a vibrating warning when an obstacle is ahead of the white cane or of the hand.

9. B. Ando S. Baglio-Fellow V.Marletta and A.Valastro[9], Mobility and navigation aids exploiting wireless or mobile technologies are available also in the field of context-aware systems Mobility aids should improve life quality and well being of visually impaired people moreover the proposed methodology enhances the visually impaired people's autonomy by avoiding inconvenient environmental interactions. To reinforce the choice of using a cane as the vector for detecting obstacle information, the white cane is the very first aid for visually impaired people. The white cane is able to provide the user with a lot of useful information about the surroundings.

10. Md. Nafiz Hasan Khan1 , Md. Amit Hasan Arovi[10], Use of voice modality is nothing new in the field of HCI but there are few works done to assist visual impaired for text entry and editing. SpeeG2 is a smart text editor that uses speech for writing text. But it uses hand gestures for error correctionIt is a very fast and complete text entry tool. It provides GUI to insert, delete, and edit words Use of text-to-speech (TTS) and speech-to-text (STT) A study investigated the application of speech-to-text to assist learning in a face to face seminar A transcript was generated using speech-to-text recognition. Some participants were able to use the transcript effectively while others were not able to.

11. Charles C Tarpert ,Ching Y Suen ,Toru Wakahara[11], This survey describes the state of the art of on-line hand-writing recognition during a period of renewed activity in the field. It is based on an extensive review of the literature, including journal articles, conference proceedings, and patents. Shape recognition algorithms, preprocessing and post processing techniques, experimental systems, and commercial products are examined.

12. . Jinqiang Bai , Shiguo Lian , Zhaoxiang Liu [12], To help the blind people walk to the destination efficiently and safely in indoor environment, a novel wearable navigation device is presented in this paper. The locating, way-finding, route following and obstacle avoiding modules are the essential components in a navigation system, while it remains a challenging task to consider obstacle avoiding during route following, as the indoor environment is complex, changeable and possibly with dynamic objects. To address this issue, we propose a novel on a collection of individuals and proved to be effective on indoor navigation tasks. The sensors embedded are of low cost, small volume and easy integration, making it possible for the glasses to be widely used as a wearable consumer device.

13. Jinqiang Bai ,Shiguo Lian , Zhaoxiang Liu [13],To overcome the travelling difficulty for the visually impaired group, this paper presents a novel ETA (Electronic Travel Aids)-smart guiding device in the shape of a pair of eyeglasses for giving these people guidance efficiently and safely. For totally blind people, three kinds of auditory cues were developed to inform the direction where they can go ahead. Whereas for weak sighted people, visual enhancement which leverages the AR (Augment Reality) technique and integrates the traversable direction is adopted. The prototype consisting of a pair of display glasses and several low-cost sensors is developed, and its efficiency and accuracy were tested by a number of users. The experimental results show that the smart guiding glasses can effectively improve the user's travelling experience in complicated indoor environment. Thus it serves as a consumer device for helping the visually impaired people to travel safely.

14. Edwige E Pissaloux ,Ramiro Velaquez , Flavien Maingreud [14],This paper proposes a novel framework for a better understanding of human cognitive locomotion and its interaction with the new tactile technologies that assist mobility

and the acquisition of spatial knowledge in the absence of sight [case of visually impaired people (VIP), seniors, etc.]. Unlike the existing mobility models, the proposed framework encompasses four elements: 1)walking with obstacle avoidance, 2) orientation, 3) spatial awareness, and 4) actual physical displacement, and is based on the concept of the tactile gist, a refreshable (dynamic) egocentric tactile representation of the perceived scene. The proposed approach may impact the design of mobility aids for VIP as it reinforces their natural mobility skills and may lead to new mobility strategies of VIP in (un-)known environment.

15. Ujjwal Bhattacharya, B.B Chaudhari[15],This paper primarily concerns the problem of isolated handwritten numeral recognition of major Indian scripts. The principal contributions presented here are 1) pioneering development of two databases for handwritten numerals of the two most popular Indian scripts, 2) a multistage cascaded recognition scheme using wavelet-based multi resolution representations and multilayer perceptron(MLP) classifiers, and 3) application of 2 for the recognition of mixed handwritten numerals of three Indian scripts—Devanagari, Bangla, and English .This scheme has been extended to the situation when the script of a document is not known a priori or the numerals written on a document belong to different scripts. Handwritten numerals in mixed scripts are frequently found in Indian postal mail and tabular form documents.

16. Alex Graves ,Marcus Lickwici ,Santiagio Fernandiz[16], This paper proposes an alternative approach based on a novel type of recurrent neural network, specifically designed for sequence labeling tasks where the data is hard to segment and contains long-range bidirectional interdependencies. In experiments on two large unconstrained handwriting databases, our approach achieves word recognition accuracies of 79.7 percent on online data and 74.1 percent on offline data, significantly outperforming a state-of-the-art HMM-based system. In addition, we demonstrate the network's robustness to lexicon size, measure the individual influence of its hidden layers, and analyze its use of context. Last, we provide an in-depth discussion of the differences between the network and HMMs, suggesting reasons for the network's superior performance.

17. QixiangYe, David Doermann[17], This paper analyzes, compares, and contrasts technical challenges, methods, and the performance of text detection. It summarizes the fundamental problems and enumerates factors that should be 99% on scanned documents [100]. Complex backgrounds, variations of text layout and fonts, and the existence of uneven illumination, low resolution and multilingual content present a much greater challengethan clean, well-formatted documents. Solving these problems requires the application of advanced computer vision and pattern recognition techniques.

18. Huiping Li, David Doermann, and Omid Kia[18], In this work, we present algorithms for detecting and tracking text , Our text tracking scheme consists of two modules: a sum of squared difference (SSD) —based module to find the initial position and a contour-based module to refine the position content has been indexed primarily by manual annotation [1], closed caption [2], or transcribed audio [3], For example, sports scores, product names, scene locations, speaker names, movie credits, program introductions and often appear in the image text and supplement or summarize the visual content, but may not be present in the transcript. Searches can easily be refined if access to this textual content is available.

19. Chucai Yi and Yingli Tian[19],This paper proposes a method of scene text recognition from detected text regions. In text detection, our previously proposed algorithms are applied to obtain text regions from scene image, CAMERA-BASED text information serves as effective tags or clue, Text-based tags are much more applicable than barcode or quick response , While text detection aims to localize text regions in images by filtering out nontext outliers from cluttered background [3], [7], [10], [16], [21], [28], text recognition is to transform image-based text information in the detected regions into readable text codes.

20. Asghar Ali Chandio,MD.Asikuzzaman[20],Text recognition is a challenging problem in computer vision. Different than the optical character recognition (OCR), text recognition is more complex due to variations in text size, colors, fonts, orientations, complex backgrounds, occlusion, illuminations and uneven lighting conditions.The traditional methods developed for text recognition are either based on the isolated character or whole word recognition. The character classifiers with a combination of different feature descriptors such as CNN [26], a histogram of oriented gradient (HOG) with random ferns [25].

21. Yi-Feng Pan, Xinwen Hou, and Cheng-Lin Liu[21],A text region detector is designed to estimate the text existing confidence and scale information in image pyramid, which help segment candidate text components by local binarization.An earlier method proposed by Wu et al. [44] uses a set of Gaussian derivative filters to extract texture features from local image regions. With the corresponding filter responses, all image pixels are assigned to one of three classes ("text", "nontext" and "complex background" then -means clustering and morphological operators are used to group text pixels into text regions.

CONCLUSION AND FUTURE WORK

A. Conclusion:

The proposed OCR WebApp has been experimentally proved to work satisfactorily by consistently providing more accuracy as compared to the existing systems using the Tesseract library. It also provides a features such as real- time OCR and handrwitten text recognition on a single WebApp.

B. Future Work:

- *UI Development*: Better responsive UI with features such as drag-and-drop.
- *Tainting Model*: Model will be trained with more amount of dataset to output higher accuracy.
- *Deploying*: The WebApp will be deployed on a host and will be accessible on a domain to people to use as free and contribute-and-learn.

PROTOCOL	ADVANTAGES	DISADVANTAGES	ROBUST	ACCURACY
MLP	capability to learn non-linear models	non-convex loss function	good	85%
MNIST	no much effort for Data preprocessing	unrealistic image	good	99.8%
MLOCR	high productivity and accuracy	not worth doing for small amount of text	good	95%
PDA	broad internet Connectivity	time consuming	good	86%
AEG	Multidimensional and multitasking	more complicated	weak	----
HMM	strong statistical foundation	unstructured parameters	good	87%
CNN	high accurate in Image Recognition	lots of training data is required	good	95%
RNN	process inputs of any length	gradient vanishing	good	82%
LSTM	can solve problem of Vanishing gradients	over fit of data	good	85.87%
TOF	high accuracy	complex due to calibration requirement	weak	1%
BiLSTM	solves the problem of fixed sequence	expensive	good	90%

REFERENCES

[1] Jaewoo Park; Eunji Lee, Yoonsik Kim, Isaac Kang, Hyung Il Koo and Nam Ik Cho "Multi-Lingual Optical Character Recognition System Using the Reinforcement Learning of Character Segmenter," in IEEE Access (Volume: 8).

[2] Ayatullah Faruk Mollah, , Subhadip Basu and Mita Nasipuri, "Design of an Optical Character Recognition System for Camerabased Handheld Devices,"in IEEE - IJCSI, Vol. 8, Issue 4. P Anjusha and R Karthik, "Off-line Telugu handwrittencharacters recognition using opticalcharacter recognition," in IEEE - 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA).

[3] Anto Sahaya Das, "A Survey on various Optical Character Recognition Techniques," 2018 Conference on Emerging Devices and Smart Systems (ICEDSS).

[4] Gayathri and R.S.Mohana, "Optical Character Recognition in Banking Sectors Using Convolutional Neural Network," I-SMAC 2019, pp. 753 – 756.

[5] Khaustov, and E.I. Maksimova, "Algorithm for Optical Handwritten Characters Recognition Based on Structural Components Extraction," IEEE - IFOST-2016, pp. 355 – 358.

[6] Muiz Ahmed Khan , Pias Paul , Mahmudur Rashid , Student Member, IEEE, Mainul Hossain , Member, IEEE, and Md Atiqur Rahman Ahad , Senior Member, IEEE "An AI-Based Visual Aid With Integrated Reading Assistant for the Completely Blind"

[7] JAMES C. BLISS "A Relatively High-Resolution Reading Aid for the Blind"

[8] Joselin Villanueva, Student Member, IEEE, and René Farcy "Optical Device Indicating a Safe Free Path to Blind People"

[9] B. Ando`, Senior Member, IEEE, S. Baglio, Fellow, IEEE, V. Marletta, Member, IEEE, and A. Valastro "A Haptic Solution to Assist Visually Impaired in Mobility Tasks"

[10] Md. Nafiz Hasan Khan1 , Md. Amit Hasan Arovi2 , Hasan Mahmud3 , Md. Kamrul Hasan4 , and Husne Ara Rubaiyeat5 "Speech based text correction tool for the visually impaired".

[11] C. C. Tappert, C. Y. Suen, and T. Wakahara, ''The state of the art in online handwriting recognition,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 8, pp. 787–808, Aug. 1990, doi: 10.1109/34.57669.

[12] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Virtual-blind-road following-based wearable navigation device for blind people," IEEE Trans. Consum.Electron., vol. 64, no. 1, pp. 136–143, Feb. 2018

[13] J. Bai, S. Lian, Z. Liu, K. Wang, and D. Liu, "Smart guiding glasses for visually impaired people in indoor environment," IEEE Trans. Consum.Electron., vol. 63, no. 3, pp. 258–266, Aug. 2017.

[14] E. E. Pissaloux, R. Velázquez and F. Main greaud, "A New Framework for Cognitive Mobility of Visually Impaired Users in Using Tactile Device," in *IEEE*

Transactions on Human-Machine Systems, vol. 47, no. 6,pp.10401051,Dec.2017.

[15] U. Bhattacharya and B. B. Chaudhuri, ‘‘Handwritten numeral databases of indian scripts and multistage recognition of mixed numerals,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 3, pp. 444–457, Mar. 2009.

[16] A.Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, ‘‘A novel connectionist system for unconstrained hand-writing recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 31,no. 5, pp. 855–868, May 2009.