



ISSN: 2350-0328

**International Journal of Advanced Research in Science,
Engineering and Technology**

Vol. 6, Issue 4, April 2019

Survey on Optical Character Recognition (OCR) using Machine Learning techniques

Srinivas BV, Sunil Biradar, Sathyanarayana, Jyothi S, Suprabhat

Assistant Professor, Department of Information Science & Engineering, Atria Institute of
Technology, Bangalore, India

Student, Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India
Student, Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India
Student, Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India
Student, Department of Information Science & Engineering, Atria Institute of Technology, Bangalore, India

ABSTRACT: The present demand of digitization of text and manuscripts requires an immediate solution so that they can be accessed electronically. While in text to speech, there occurs various systems which convert normal language generated text into speech, thus enables the user to identify them. The proposed system aims to study image recognition technology (Optical Character Recognition) with text to speech conversion technology and to develop a cost effective user friendly system. In this system we have tried to make a system by which we can get the text through scanned image and then that text is converted to speech. Simple Arithmetic operation are performed based on the operations mention in the scanned image.

KEYWORDS: Optical character recognition, Convolutional neural network, Text-to-speech, Artificial neural network.

I. INTRODUCTION

After the advent of digital computers, incorporating human functions to computers has been an interesting and exciting research field. For over years, humans have been thinking of machines with the ability to read and interpret printed textual documents, so that they can be automatically converted into an alternate medium or format efficient algorithms have been developed so far so as to enable the machines to recognize characters. Such a system is named as Optical Character Recognition. This is a system developed for deriving character-based files from digitized images of printed or typewritten documents and/or handwritten manuscripts. Digitizing is done by using flatbed scanners or digital cameras. It is thus a process of visual recognition, which converts text documents into editable or searchable text.

II. RELATED SURVEY

An Optical character recognition (OCR) model is been presented that uses neural network (NN) for both the scanned and handwritten characters which has a proved to be an efficient one. This uses various algorithms which are very much unique in their own. Thus this model implements various advanced techniques for the detection of characters and thus study the behaviour of various algorithms. Optical character recognition is a process of detecting, segmenting and recognizing various characters from their respective text or image. This could be very much used for the historical manuscripts or any older documents. The recognition system first generates the document, then digitalizes it and finally stored in the system. Various external factors need to be kept in mind that it doesn't disturb the system. The accuracy of HCR is been pushed up to 90 percent due to enormous changes in the shape, size, scale and orientation. Various researches are also done for the Arabic languages where the reading pattern entirely changes. Thus image processing and recognition plays very crucial role in character recognition. Once a document is fetched using the camera, the software tends to pick the text, follows few processing steps and finally tries to extract the scanned document. The features extracted from the document are classified into high level or low level. The low level extraction includes height, width, curliness and aspect ratio. On the other hand high level extraction includes number of position of loops, straight lines etc. Due to presence of various processing activities, these techniques could be classified into various techniques. The proposed system is been developed on the basis of neural network which learns from training data set and provides fruitful result. This technique uses optical mechanism to recognize character. The used system uses few very strong algorithms to produce efficient output. Firstly the input is fetched to the input layer. Further these inputs are

propagated through the entire network until the last output unit receives. Due to this forward pass the actual output are predicted. The actual network are further subtracted from the initial input and generates error signal which is based on the back propagation method where the errors are propagated back. The weights are further adjusted and the neural network has just learned from the experience. These network further undergoes supervised training which leads to present at the input layer which passes the pattern to the next neuron presented at the hidden layer. Thus the final output is determined by the activation from output layer. This process is repeated various times and once the network has learned the correct set of classifications for a particular set of input, it can be tested on further test inputs.

Visually challenged individuals find various challenges to find a scanned or printed content that leads to difficulties to differentiate between various properties like exactness, arrangement, center, viability and versatility. A smart system is been used to enable visually challenged person to overcome such problem more productively and precisely. Here an individual uses a camera based assistive device to pursue the entire text paragraph. The following system uses a certain camera so as to fetch the input through the user and later is been recognized by the process of digitization. The scanned document is further been processed by the software called OCR which does the rest of the work. The existing system has a PC program which enables the individual or the user to recognize what is been typed on the screen. Similarly when visually impaired person writes something and if the screen reader finds it wrong then he would press delete to eradicate it. The major issue that arises from this system is that it can read only text and not an image. Thus the screen reader helps to identify character which uses a sounding device that many people find it difficult and complex. Thus to rectify and overcome the problems and consequences from the earlier system an OCR is developed using neural OCR in open computer vision. One of the important aim is to read the document and identify the text and later the image is been captured using webcam which later follows image processing. The later is then recognized using assistance of speaker which helps the user to read the content. The entire process starts with the image capturing where an inbuilt camera captures the fetches input. Special care should be taken to ensure the image is clear. This step is followed by pre-processing where different techniques are used for skew detection like skew-detection, linearization and noise removal. This process verifies whether the point of orientation lies to an extent 15 degrees which ensures a skew detected image and used for further processing. [2]

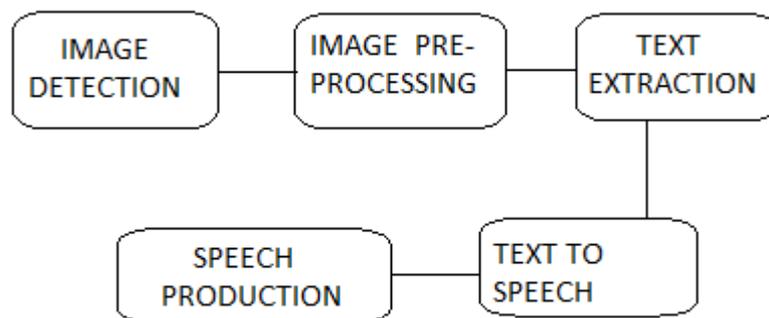


Figure1. System flow process

Once the pre-processing is done the scanned document is fetched to segmentation process. This process results in a breakdown of a scanned image to further simple characters. Further the image is binarized and the spacing is checked. A simple paragraph is first distinguished between various lines which in turn is differentiated into many words and finally to various characters. The feature extraction process of glyph is also done to ensure the consistency. A glyph is a well-defined symbol which has a unique dimension. Thus its recognition plays a very important role in image and text recognition. Finally the last stage includes the feature extraction of the scanned and pre-processed image which uses tesseract OCR engine that differentiates between various characters and finally the visually paired person can overcome their difficulties using text to speech recognition

The following system proposes the recognition of various characters, math expressions that includes various operators. Particularly, there is a need for a portable scanner that would be affordable and easily. This paper gives the complete recognition of characters with the mathematical problems and its solution designed for the visually challenged individuals. The system consists of a webcam interfaced with raspberry pi which accepts a page of handwritten text in scanned document or the math expression written by the user. The proposed system is controlled under Matlab software. In this MATLAB with the help of image processing toolbox Optical Character Recognition is performed with text to speech recognition technique. An image with handwritten text written by user gets converted with each pixel converted with binary "0" and "1" conversion, labelling and segmentation, Synthesizer, the result of math expression which is recognized by the OCR tool is later converted with speech which is heard by external speaker. Controller coding for the

Raspberry pi is done through PYTHON language. The audio output is gained once the result of math expression is detected and then converted to speech. Raspberry pi has the audio port where the output can be heard through the headphone or the speaker. Once we get the result of our math expression, the raspberry pi will take few milli seconds to convert it as a speech. The proposed system is validated with both simulation and experimental verification it achieves the handwritten document is converted with speech for the use of visually impaired people. Optical Character recognition is useful for visually impaired individuals who cannot read Text document, but need to access the content of the Text documents. It is used to get the solution of mathematical problems just by writing the math expression on a paper. This paper is on Methodology of a camera based device that can be used by people to read any handwritten document or Text document. [2].

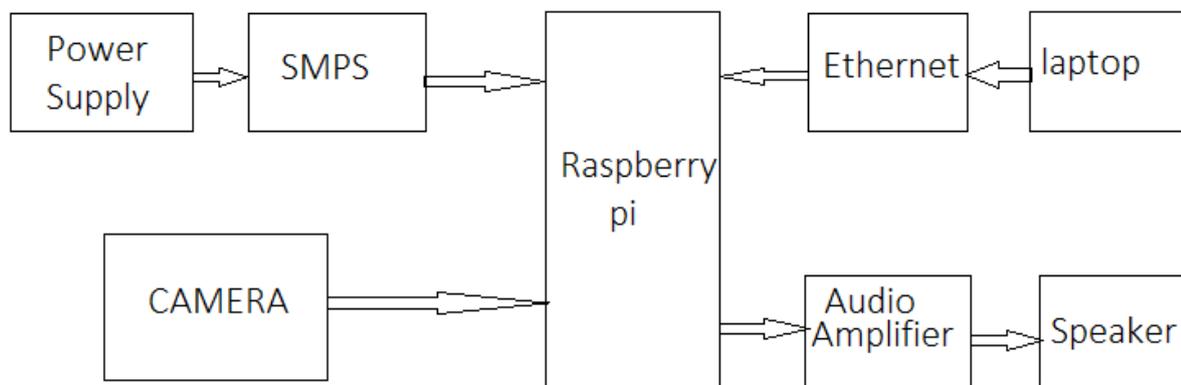


Figure 2- Architecture diagram

The web camera is connected to raspberry pi. The raspberry pi has an operating system called RASPION that process the conversions. The audio output is received from audio jack of the raspberry pi. The converted speech output is then amplified using an audio amplifier. The Internet is connected through the Ethernet port in. The page which is to be detected is placed on a particular base and the camera is focused to capture the image with optimum and ideal conditions. The captured image is processed by the optical character recognition software which is installed in raspberry pi. The captured image is later converted to text by the pre-processing steps. The text is converted into speech by the Text to speech engine. The final output is given to the audio amplifier that is connected to speaker. The speaker can also be replaced by a headphone for convenience or external speaker.

III.CONCLUSION

As described, a quantitative evaluation of character recognition using OCR technology is presented with mathematical equation solver and text to speech method is used for the results. OCR system can be effectively used to speed up the translation of image based documents into structured documents that are currently easy to discover, search and process. The recognition of new font characters by the system is very easy and quick. The result of mathematical equations can be calculated very quickly and these results will be converted to speech easily.

REFERENCES

- [1] G. Takacs, I. Pil'aszy, B. N'emeth, and D. Tikk, "Scalable collaborative filtering approaches for large recommender systems," Journal of machine learning research, vol. 10, no. Mar, pp. 623–656, 2009.
- [2] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," Computer, vol. 42, no. 8, 2009.
- [3] J. McAuley, R. Pandey, and J. Leskovec, "Inferring networks of substitutable and complementary products," in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015, pp. 785–794.
- [4] J. Bobadilla , F. Ortega, A. Hernando and A. Gutiérrez,"Recommender systems survey", Knowledge-based systems, vol. 46, pp.109-132,2013.