Optical Character Recognition Using Machine Learning

Harshasudhan[1] , Mounika M[2], Pandiselvi T[3], Nithya L[4]
[1, 2, 3] Dept of computer science and Engineering
[4]Assistant professor, Dept of computer science and Engineering
[1, 2, 3, 4] Nehru Institute Of Technology Coimbatore

Abstract- *The device we have proposed aims to help people with visual impairment and also to convert the hardcopy hospital data into Soft Copy. The converted softcopy file will be stored as a backup in the drop box. In this project, we developed a device that converts an image's text to speech. The basic framework is an embedded system that captures an image and extracts only the region of interest (i.e. region of the image that contains text) and converts that text to speech. It is implemented using an Ubuntu and a Web-Camera. The captured image undergoes a series of image pre-processing steps to locate only that part of the image that contains the text and removes the background. Two tools are used to convert the new image (which contains only the text) to speech. They are OCR (Optical Character Recognition)software and TTS (Text-to-Speech)engines. The audio output heard through the Ubuntu's audio jack using speakers or earphones.*

I. INTRODUCTION

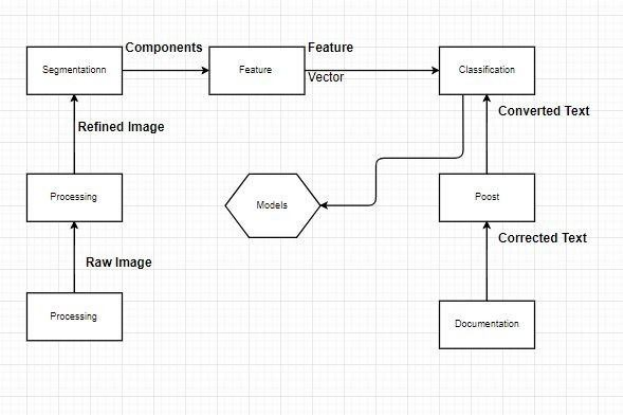
In our planet of 7.4 billion humans, 285 million are visually impaired out of whom 39 million people are completely blind i.e. have no vision at all, and 246 million have mild or severe visual impairment (WHO, 2011). It has been predicted that by the year 2020, these numbers will rise to 75 million blind and 200 million people with visual impairment. As reading is the prime importance in the daily routine) text being present everywhere from newspaper, sign-board etc.) of mankind.

II. OPTICAL CHARACTERRECOGNITION

Optical Character Recognition is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine coded text, whether from a scanned document or a photo of a document, from subtitle text that is superimposed on an Here we use two main processes. In the first stage we acquire the image from the web camera and then we convert it into the text document using Optical Character Recognition (OCR). In the second stage we use natural language processing and digital signal processing for converting the text into speech using Text to Speech synthesizer (TTS). image. The basic junction of this Technology is to automatically capture the printed text present scanned images and convert it into a text searchable document. It has been widely used in scanned or photographed documents, converting them into soft copies where one can edit, search, play its content and easily carry.

The Text to Speech (TTS) Synthesizer is a computer based system that reads the text aloud automatically, whether the test is introduced by a computer input stream or a scanned input submitted to implement by both hardware and software. Speech is based on concatenation of natural speech i.e. The units that are taken from the natural speech altogether are formed as a word or sentence. This technology of speech synthesis (TTS) enables a text which is in digital format to be synthesized into human voice and that is played through an audio system. The objective of the TTS is the automatic conversion of text in to audio without any restrictions algorithms are compared (S K Hese and M R Banwaskar, 2013) and they found to have similar features but Linear Discriminant Analysis (LDA) outperforms Principal Component Analysis (PCA) algorithm when large training sets are involved in recognition. Also, LDA discriminates most of the information present in the image efficiently by computing the intra class and inter class scatter matrices. Using the database which contains normalized face images, the recognition is performed in the vehicle security system through the LDA algorithm. LDA performs the feature extraction of the stored images in the database which are called the training images and the camera acquired face image which is called the test image. The test image is to be compared with the database images and the classifier used in the algorithm decides the image as known or unknown using the Euclidean distance and the threshold value. The Euclidean distance is calculated between the corresponding weights of features and the image which produces minimum distance is best matched with the test image

III. BLOCK DIAGRAM



EXISTING SYSTEMDISADVANTAGES

There is no complex background for the input images which are taken in the previous works, i.e. the test inputs are placed on the plain white sheet. It is easy to convert such images to text without pre-processing, but this kind of work will not be useful for real-time systems.

IV. PROPOSEDMETHOD

To overcome the problems in the existing system we have developed a project for Blind People and for the Hospitals using neural OCR in OpenCV. The proposed system is to assist blind people to read the text from challenging pattern and background for the purpose of reading document. It is used in the Hospitals to store the report for future reference. The main objective of our system is to identify the text in the document. Firstly the object image is captured by using a webcam which is embedded within Ubuntu and it is followed by the image processing. To execute an automated system, which checks a document and reads out its substance into the individual on clicking the button. The vocal is delivered with the assistance of a speaker which would help the individual to readout the content in the scanned document. Our system helps the blind people for the purpose of reading without consuming much space. The Documents that are read will be uploaded in the drop box for the future usage.

V. METHODOLOGY

Here we use two main processes. In the first stage we acquire an image from the web camera and then we convert it into a text document using Optical Character Recognition (OCR). In the second stage we use natural language processing and digital signal processing for converting the text into speech using Text to Speech synthesizer (TTS).

VI. CONCLUSION

This system is a novel implementation of printed documents into soft copy for Hospital report and text to speech conversation for blind people. Our algorithm has processed successfully and reads the text efficiently and then converts into speech. We demonstrated the software implementation with different samples. The device is compact as we require the image through the laptop's web camera which is able to locate anywhere and enable the retrieval of the read documents. Thus we have proposed a multifunctional system that acts as an aid for enhancing the quality of one's life. This project is made on Python platform with its various packages which is programmed and run on Ubuntu 3B+ model. By implementing this system we are performing image to text and then to speech conversion, Google based personal assistant with home automation. It was observed that personal assistants like Amazon Echo are not so efficient and they are unresponsive. These personal assistants are expensive too.

REFERENCES

[1] Breen A.P., "The future role of text to speech synthesis in automated services".
[2] Tanprasert, C.; Koanantakool,T., "Thai OCR: a neural network application"
[3] Lucas S.M., „"High performance OCR with syntactic neural networks".
[4] https://github.com/aquatiko/Image-Text-Speech-Synthesizer-Converter
[5] https://ieeexplore.ieee.org/abstract/document/1296349/
[6] Morgan Kaufmann, 1990, pp. 612-626. [7] http://code.google.com/p/tesseract-ocr.
[7] https://github.com/prabhakar267/ocr-convert-image-to-text
[8] M. Cai, J. Song and M. R. Lyu. A New Approach for Video Text Detection. In Proc. of International Conference On Image Processing, Rochester, New York, USA, pp. 117-120, 2002.
[9] https://ieeexplore.ieee.org/abstract/document/7919526/
[10] http://ijarcsse.com/Before_August_2017/docs/papers/Volume_7/7_July2017/V7I7-0117.pdf
[11] A. K. Jain and B. Yu. Automatic Text Location in Images and Video Frames. In Proc. of International Conference of Pattern Recognition (ICPR), Brisbane, pp. 1497-1499, 1998.
[12] https://www.hackster.io/rgrokett/pitextreader-for-impaired-vision-695331
[13] https://pdfs.semanticscholar.org/99c9/90ff841da4923771c5d2b6692c3a4656d447.pdf