IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

Image to Text Viewer (iTV)

Atul Thorat[1], Deepak Singh[1], Rohit Kadam[1], Prof. Uttara Gogate[1]

Shivajirao S. Jondhale College of Engineering, Dombivli, Maharashtra[1]

Abstract: Image To Text Viewer is a mobile application which converts a textual image captured by the android devices into editable and searchable text document. It uses smart phones of android platform. This paper introduces the functionality of Optical Character Recognition, Tesseract Engine and Android studio. iTV system provides many features like editing raw data, Memory utilization and Quick translation. iTV is more efficient and easier alternative to scanning a document using a scanner as the image captured using OCR is of exactly the same quality like its scanned copy. The advantage of using OCR is that we can get scanned copy of any document image with the help of a simple mobile phone camera without using a bulky scanner. The aim is to develop a user friendly application that is used to convert the image captured by the android devices into editable and searchable text document anywhere and anytime.

Keywords: Optical Character Recognition System (OCR), Tesseract Engine, Camera Captured Document Images, Android smart phone.

I. INTRODUCTION

Earlier scanners were the only working OCR application available in the market. The main disadvantage of scanners was that it was not portable and it takes a lot of time to capture an image.

But with today's devices having better processing speeds, larger internal memory and an excellent back camera, researchers have dared to think of running OCR applications on devices such as android phones for having real time imaging results. Applications such as Cam Scanner and Google translate are the prime examples of Optical character Recognition application. It also showcases the fact that this OCR technology can be put to use in a wide array of streams and hence it is a very important concept which requires more attention towards research [1].

iTV application is fully based on OCR technology. iTV is an android application that is used to convert the image captured by the android devices into editable and searchable text document.

Purpose of the system is to develop an Android based application that are capable of localizing the textual regions in the input image and convert it into text .The work will mainly focused on printed text without skew. This application will widely used in banking, legal industries, education, financial and government agencies.

II. LITERATURE SURVEY

Getting inspiration from well-known iphone applications, Derek Ma, Qiuhau Lin, Tong Zhang developed a text translation application which is based on android operating system, which was able to recognized the character that was been captured through camera, translate the text in Chinese language and return the result back on to the screen of the phone. Their algorithm has a correct recognition rate is greater than 85% on character level. This application was developed to help the tourist so they can easily navigate even in a foreign environment [2].

There are several methods that can be used for extracting text from images such as image documents, scene images etc [3]. Texts that are present in an image contain several useful and important data or information. Extraction of text from an image has been used in variety of application such as mobile robot navigation, text scanner, vehicle license plate detection, etc. In similar applications they employ discrete wavelet transform (DWT) to extract the text from an image. The document retrieving, object identification, image that will be passing as an input can be a color image or it can be grayscale image. If the image is a color image, then preprocessing is done on an image. In order to extract the text edges from an image, Sobel edge detector is used on each part of image. The resultant edge obtained from the process is used form an edge map. Morphological operations are applied over edge map. Then thresholding is applied to improve the performance of the edge map.

A. Types of recognition engines:

1. Optical Character Recognition: turns images of printed characters into machine-readable characters.
2. Intelligent Character Recognition: It reads images of hand-printed characters (not cursive) and converts them into machine-readable characters.
3. Magnetic Ink Character Recognition**:** It is a specialized character recognition technology that adopted by the U.S banking industry to facilitate check processing.
4. Optical Mark Recognition**:** OMR technology detects the existence of a mark but not its shape.

III.OPTICAL CHARACTER RECOGNITION

Optical Character Recognition (OCR) is a technology that we can use to convert different types of PDF files, documents, or images captured by a digital camera or mobile camera into searchable and editable data.
Images captured by a digital camera differ from scanned documents or only PDFs. They often have defects such as

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

distortion at the edges and dimmed light that making it difficult for most OCR applications, to recognize the text. OCR allows for automatically recognizing characters through an optical mechanism.
It is capable of recognizing both handwritten and printed text. Its performance can be judged based on the quality of the documents and the camera being used to capture the raw image [4].
OCR works in Android mobile operating system by combining Google's open-source OCR engine, Tesseract and the text recognition OCR engine.

IV. TESSERACT

Tesseract is an open source engine for optical character recognition. Tesseract was in the top three OCR engines in terms of accuracy of character in 1995. It is available for Linux, Windows, Mac OS and Android platform. Tesseract engine is one of the most accurate open source OCR engines currently available [5]. In late 2005, HP released Tesseract for open source and now available at http://code.google.com/p/tesseract-ocr [6].

Table .1 Participating Organizations

Organization	Version Name	Version No.	Platform	Version Type
Caere Corp. Los Gatos, California	Caere OCR	138.1	Sun SPARCstation	pre-release
Electronic Document Technology Pte. Ltd. Singapore	EDT ImageReader	3.0	PC DOS	commercial release
Hewlett Packard Laboratories Bristol, England	HP Labs OCR	7.0	HP workstation	research prototype
International Neural Machines Inc. Waterloo, Ontario	INM NeuroTalker	2.52	PC DOS	beta release
Ligature Ltd. Jerusalem, Israel	Ligature CharacterEyes Pro	2.6	PC Windows	beta release
MAXSOFT-OCRON, Inc. Fremont, California	MAXSOFT-OCRON Recore	3.2	PC Windows	beta release
Recognita Corp. Budapest, Hungary	Recognita OCR	3.0	PC Windows	beta release

Table 1 shows all the participating organization in "The Fourth Annual Test of OCR Accuracy, 1995".

Table 2 Character Accuracy for 300 dpi Binary

	Errors	% Accuracy	Failures
Caere OCR	4,459	98.61	none
EDT ImageReader	13,162	95.88	1 / 0.30
HP Labs OCR	5,959	98.14	none
INM NeuroTalker	---	< 90.00	none
Ligature CharacterEyes Pro	---	---	1 / 1.07
MAXSOFT-OCRON Recore	8,377	97.38	none
Recognita OCR	11,280	96.47	none
XIS OCR Engine	5,473	98.29	none

Character accuracy has been derived by using this formula

$$\frac{\#character - \#errors}{\#character}$$

We are using tesseract engine because it is most accurate open source OCR engine currently available.

A. Architecture of Tesseract:
The first step is a connected component analysis. In this step outlines of the components are stored. This was a computationally very costly design decision at that time, but had a significant advantage: by inspection of the nesting of outlines and the number of different child and grandchild outlines, it is simple to detect and recognize inverse text as easily as black-on-white text. Tesseract engine was probably the first OCR engine able to handle white-on-black text so paltry. At this stage, outlines are gathered together, purely by nesting, into Blobs.

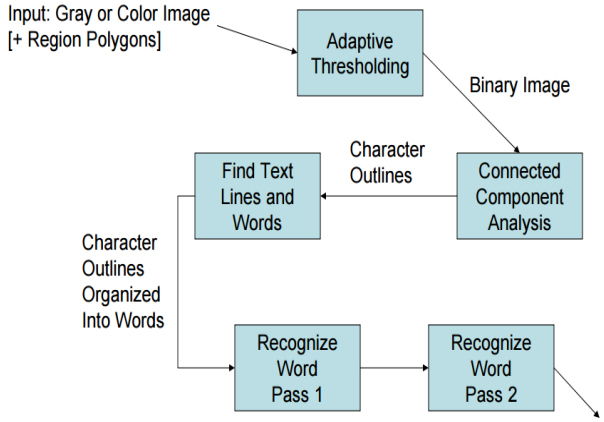

Fig. 1 Architecture of Tesseract

Blobs are organized into text lines, and the lines and regions are analyzed for fixed pitch or proportional text. This text lines are broken into words, according to the kind of character spacing. Fixed pitch texts are chopped immediately by character cells. Proportional text is broken into words using definite spaces and fuzzy spaces.

Recognition then proceeds, which has a two-pass process. In the first pass, it recognizes each word in turn. Each word that is satisfactory is passed to an adaptive classifier as training data and then gets a chance to more accurately recognize characters in text lower down the page. Since the adaptive classifier may have learned something useful to contribution near the top of the page, a second pass is run over the page, in which words that were not recognized that are recognized again. A final phase resolves fuzzy spaces, and checks alternative hypotheses for the x-height to locate small cap text.

V. IMAGE TO TEXT VIEWER (iTV)

Image to Text Viewer is a mobile application that uses smart phones of android platform**.** iTV is OCR based android application that are used to convert the image captured by the android devices into editable and searchable text document [7].

DOI 10.17148/IJARCCE.2016.52121

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

iTV application uses the OCR and Tesseract Engine for character recognition. iTV is more efficient and easier alternative to scanning a document using a scanner as the image captured using OCR is of exactly the same quality like its scanned copy, the only difference being that OCR is done with the help of a simple mobile phone camera whereas scanning is done using a bulky scanner

VI. HOW iTV WORKS

iTV app first asks for an image as input that is to be converted into editable text. This input image can be taken from gallery of owner phone or it can be captured from the camera.

After selecting image iTV app send this input image to crop image function where you can crop the image for particular text that you want to extract. This image croping help in modifying image by extracting extra image components. After applying crop we have to save the image for further work. This saved image go for extracting text by applying OCR (tesseract engine) then all the procedure of character recognition is done by the tesseract OCR engine [4]. In this procedure image binarization is done. Image binarisation is the process that is separates text from the background, producing a black-and-white image that is much smaller in size than the original image.

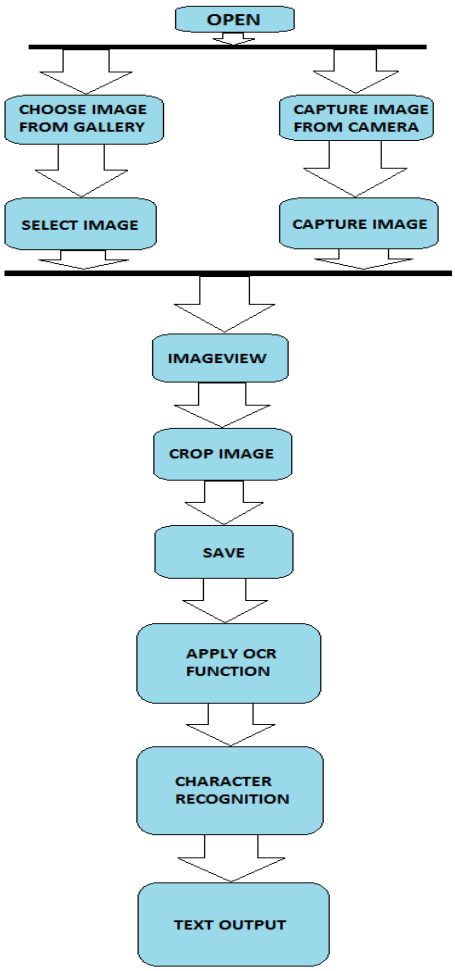

Fig. 2 working of iTV

Additional skew correction and document orientation detection can be applied. Document Analysis is a set of algorithms that analyses the image and it detects letters, joins the letters into words, then into sentences, then into lines of text, and finally, into paragraph. Additionally, the reading area is cleaned and noise removed. Then the detected blocks on the image that are recognized using the special language and pattern definitions, that is include in system [7].

Recognition results are the set of characters with coordinates united in lines. Each character has the level of confidence which shows how recognition engine was sure in final character choice. After that it gives output in editable text format which can be copied from iTV app for further use.

VII. METHODOLOGY

Step 1: User will press the button to select the picture from gallery or captured from camera. Image will remain focused properly on the textual region of the picture and ignoring other details of the scene.

Step 2: After selecting image, User can enhance image. In this process user can crop image, rotate image. After that this image is converted into Bi-colored image i.e. Black & White image.

Step 3: Text recognition will be done by clicking the "OCR!" button. Recognized text will be displayed in the text bar.

All the three components of the application with different functionalities are in a same layout. User will use different components sequentially while using the software like, "**select image**" then "**OCR**" then "**Use text**". While navigating through the mentioned functionalities in application, if a user goes against the intended behavior then application would generate an error message and informing user about the error.

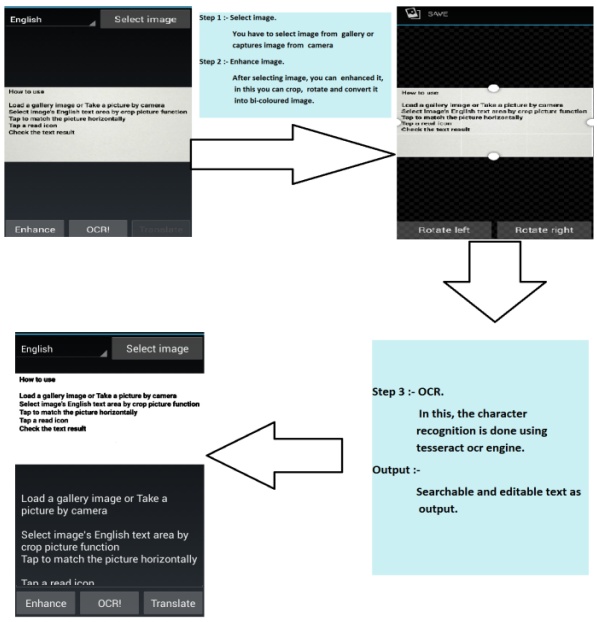

Fig. 3 Steps involving in iTV result

DOI 10.17148/IJARCCE.2016.52121

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

VIII. RESULTS

The following fig.4 shows the results. We take image as an input which contains the text "Shivajirao S. Jondhale College of Engineering." Our system, first converts it into a binarized image or Bi-coloured image that is black and white image as shown in the fig.4. Then this Bi-coloured image processed to OCR engine for character recognition. The recognition of text is done by Tesseract OCR engine. The result is the output text "Shivajirao S. Jondhale College of Engineering." In the document/text form. The user has full control over the result.

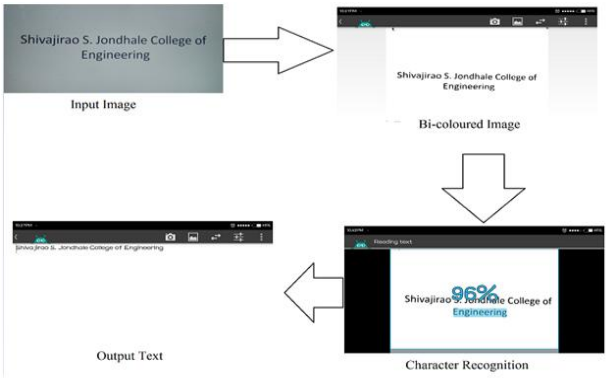

Fig.4 Final Result

A. ADVANTAGES:-

1. Android is an open source & free. iTV mobile application is significantly of lower cost.
2. iTV application can saves time and memory storage.
3. Time and efforts in the documents typing can be avoided and we can edit text for further use.

B. LIMITATION:-

The output depends on quality of input image and the quality of the input image depends on the camera quality. If camera quality is poor then output will be affected. Thus it is necessary to have a good quality camera.

C. FUTURE SCOPE:-

1. Images with different languages can be processed and translated.
2. We can use iTV application as a language converter.
3. We can use iTV application for handwritten data also.

D. APPLICATION:-

1. It can be used in business document as for data entry.
2. It can be use in extraction of data from printed document.
3. It can be used for extracting contact info from any type of business card.
4. It can be used in making e-book from the real and printed books.

IX. CONCLUSION

We presented the Android Application of "iTV (Image to Text Viewer" for converting images into the editable text. The recognition results can be processed and exported. The user of the application has full control over the iTV results. The advantage of the system is that it is easily portable and help in translating the text in various languages.

REFERENCES

[1] "Optical Character Recognition on Handheld Devices", International Journal of Computer Applications (0975 – 8887), Volume 115 – No. 22, April 2015, pg10-13.
[2] Derek Ma, Qiu hau Lin, Tong Zhang, "Mobile Based Text Detection and Translation".
[3] Image Segmentation for Text Extraction by Neha Gupta, V .K. Banga. 2nd International Conference on Electrical, Electronics and Civil Engineering (ICEECE'2012) Singapore, April 28-29, 2012.
[4] Implementing Optical Character Recognition on the Android Operating System for Business Cards By Sonia Bhaskar, Nicholas Lavassar, Scott Green EE 368 Digital Image Processing.
[5] R. Smith. "An overview of the Tesseract OCR Engine." Proc 9th Int. Conf. on Document Analysis and Recognition, IEEE, Curitiba, Brazil, Sep 2007, pp629-633.
[6] The Tesseract open source OCR engine, http://code.google.com/p/tesseract-ocr.
[7] Ray Smith Google Inc. theraysmith@gmail.com.

BIOGRAPHY

Mr. Atul S. Thorat is a student of Computer Engineering in Shivajirao S. Jondhale College of Engineering.

Mr. Deepak L. Singh is a student of Computer Engineering in Shivajirao S. Jondhale College of Engineering.

Mr. Rohit R. Kadam is a student of Computer Engineering in Shivajirao S. Jondhale College of Engineering.

DOI 10.17148/IJARCCE.2016.52121