International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Number/Text Translate from Image

Kaushik Chavan[1], Roshan Bhoir[2], Nikhil Gawade[3], Surekha Khot[4]

[1,2,3,4]*Dept. of Information Technology, ACPCE Kharghar, Mumbai University, Maharashtra, India*

Abstract-*This project is about a Number/Text Translation system using ML based platform such as Firebase ML kit. As hardware integrated in smartphones can operate more task than traditional phones, the smartphones are no longer just a communication device but also considered as a powerful computing device which able to capture images, record videos, surf the internet and performs other operations. The design of this translation system base on android Text Recognition API. It will scan the text or number through the camera of particular android device within certain distance. Once the text is extracted from image it will proceeds to next stage of translation.*

Key Words - Text detection, Text Recognition, Classification, Segmentation

1. INTRODUCTION

Real time world contains different significant messages, labels and useful information but most of them are written in different official languages which depends on the host country. Sometimes a signboard or any other notice or label could contain an important message, information or even danger. If the message is unreachable to person with different language background, it might cause important information to be missed out. Besides that, it is inconvenient for a traveler to carry on their tasks in a foreign country if they don't understand the language used in that country. They need to carry a pocket dictionary or use online translation service in order to understand the message [1].

In 2014, the World Health Organization (WHO) estimated the numbers that there are 285 million people to be visually impaired, total of 246 million with visual impairment and 39 million people are blind. It is also reported that 90% of the world's visually impaired people are from low income group [2]. One of the major problem is visually impaired people they are not capable in accessing translated text. Although there are numbers of assistive technology meant for visually impaired, most of these special devices are not convenient because it require custom modifications and some are too expensive[3]. Almost 70% of visually impaired people are unemployed and most of them are unable to use assistive technology due to its cost. This causes visually impaired users missed the opportunity to access important text that is present in the world to carry out day to day task efficiently [4]. To overcome these issues, this paper proposes to develop an Android application which capture text based image which carries important messages from real world and translate them and finally pronounce it[5].

This paper is formatted as follows: - Section 2 represents methodologies used for text detection, Section 3 is result. in section 4 details about the application. Future scope is presented in section 5 and concluded in section 6

2. METHODOLOGY

The process of detecting and recognizing text is divided into two stages text detection and recognition. Text detection finds the text area from input image, whereas recognition converts the obtained text into characters and words.

With the help of various method and libraries provided by firebase ml kit we can perform various operations such as detection, classification, segmentation and recognition [6].

As android evolve throughout all this years, the device also evolve dramatically. With even budget android device nowadays featuring with superb camera lenses. This makes the task of detection of text in image even easier than before.

The Text Recognition process segments text into blocks, lines, and words. Roughly speaking:

- a Block is a contiguous collection of text lines, such as a paragraph or column,
- a Line is a contiguous collection of words on the same vertical axis, and
- a Word is a contiguous collection of alphanumeric characters on the same vertical axis.

Although there are some Input image guidelines according to the official documentation of firebase kit .

For ML Kit to accurately recognize and identify text, given input images must carry the text that is represented by pixel data. Usually, for Latin text, each character should be at least 16x16 pixels. For Chinese, Japanese, and Korean text (only supported by the cloud-based APIs), each character should be minimum 24x24 pixels.

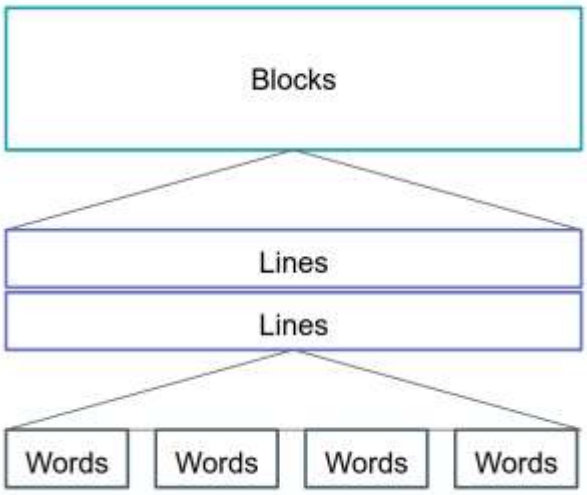
For all languages, there is generally no accuracy advantage for characters larger than 24x24 pixels[7].

So, for example, image having a 640x480 pixels might work well for scanning a business card that has full width of the image. To scan a document printed on letter-sized paper, a 720x1280 pixel image might be required [7].

Poor image focus can hurt text recognition accuracy. If you aren't getting acceptable results, try asking the user to recapture the image [7].

If one recognizes text in a real-time application, might also want to consider the overall dimensions of the input images. Smaller dimension images can be processed more faster to reduce latency, captured images with lower resolutions or dimensions (keeping in mind the above accuracy requirements) and ensure that the text occupies as much of the image as possible [8].

The following image highlights examples of each of these in descending order.



Recognize text in image

To recognize text in an image, create a FirebaseVisionImage object from either a Bitmap file, media.Image, ByteBuffer, byte array, or a file available on the device. Then, pass the Firebase Vision Image object to the Firebase Vision Text Recognizer's process Image (process image is a predefined available with libraries of firebase ml kit) method. This process is further aided by various method like degreesToFirebaseRotation(),FirebaseVisionImageMetadata. Builder(),FirebaseVisionImage.fromBitmap()-this particular method is then used stored image object and finally processImage() method [7].

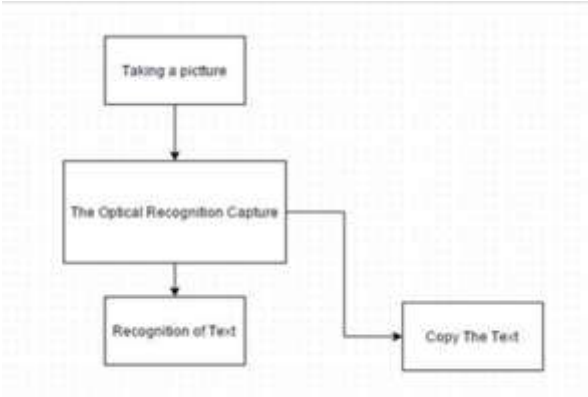
Extract text from blocks of recognized text

Once the text recognition operation performed, a FirebaseVisionText object will be paased on to the success listener. A FirebaseVisionText object holds recognized text from the given image and zero or more TextBlock objects.Each TextBlock represents a rectangular block that contains text, which contains zero or more Line objects. Each Line object carries zero or more Element objects, which represent words and word-like entities (dates, numbers, and so on).

Identify the language

ML Kit identify text from 103 different languages in their native scripts. In addition, romanized text can be recognized for Arabic, Bulgarian, Chinese, Greek, Hindi, Japanese, and Russian [9].For this whole process firebase ml kit provide predefined method such as getLanguageIdentification(). If this method result is success, then it will pass to the BCP-47 code of success listener [9].

Following image describes the flow of the system:-



3. RESULT

4. APPLICATION

Various applications of text detection and recognition from images have been developed in past few years with advancements in image processing techniques. Developments of various embedded and integrated systems in the field of computer vision and machine learning gives further rise to the increase in applications of text detection and recognition.

Learning a new language is always a challenge. One can generally pick it up faster if live in the country where it's spoken. Regardless of who are traveling on holiday or are on a business trip, knowing a few words of the country's language makes it so much easier to communicate with the locals. One doesn't have to pack any heavy dictionaries just take out your smartphone or tablet and you have everything you need.

Text detection and recognition is used in industries for reading labels, numbers, documents etc. It is used to retrieve video captions as well as specific text contents from web pages that contain significant text content.

It is used for street boards reading purpose in case of unmanned vehicles. Text detection and recognition has very important application in form of assisting blind or visually impaired people for reading, learning, making their daily life easy [10]. It is also used in automatic signature reading on cheques. Automatic scanning of various documents is another application of text recognition.

5. FUTURE SCOPE

For future work, an assistant app will be developed which will include speech out feature beneficial for visually impaired users with language-based learning disabilities, such as dyslexia. One of the major problem is visually impaired people they are not capable in accessing translated text. Although there are numbers of assistive technology meant for visually impaired, most of these special devices are not convenient because it require custom modifications and some are too expensive. To overcome these issues, this paper proposes to develop an Android application which capture text based image which carries important messages from real world and translate them and finally pronounce it.

6. CONCLUSION:

This paper reviews an android application "TextTranslator" developed by studying the previous developments to capture text, translate. It will display the translated text on the device screen. Final application has been tested and results obtained are promising which demonstrate that this study successfully addressed the problems discussed. This study proves that the application is mainly convenient for travelers to reduce the language barrier during their visit to another country which uses different language to portray information and access printed/displayed text which may carry significant messages. This application can be further improved and enhanced. It is developed as a prototype to obtain users' feedback and in future it can be upgraded with better translator services or even by multi supported text to speech engine. By doing so, it can significantly improve the performance of the system to ensure a better quality application.

7. REFERENCES

[1] Sathiapriya Ramiah ; Tan Yu Liong Asia Pacific University of Technology & Innovation,''Detecting text based image with optical character recognition for English translation and speech using Android''.

[2] Pascolini D, Mariotti SPM. Global estimates of visual impairment: 2010. British Journal Ophthalmology.

[3] Shailendra Kumar, Gareema Sharma,''Web challenges faced by blind and vision impaired users in libraries of Delhi An Indian scenario''.

[4] Akbar S. Shaik,Gahangir Hossain, "Design, development and performance evaluation of reconfigured mobile Android phone for people who are blind or visually impaired".

[5] Sathiapriya Ramiah , Tan Yu Liong , Manoj Jayabalan ,"Detecting text based image with optical character recognition for English translation and speech using Android".

[6] Chaitanya R. Kulkarni, Ashwini B. Barbadeka, "Text Detection and Recognition".

[7] https://firebase.google.com/docs/ml-kit/android/recognize-text,"Recognize Text in Images with ML Kit on Android".

[8] https://firebase.google.com/docs/ml-kit/android/recognize-text,"Recognize text in images".

[9] https://firebase.google.com/docs/ml-kit/ios/identify-languages ,"Identify the language of text with ML Kit on iOS".

[10] Hanen Jabnoun, Benzarti Faouzi, Hamid Amiri "A New Method for Text Detection and Recognition in Indoor Scene for Assisting Blind People".