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

OPTICAL CHARACTER RECOGNITION AND TRANSLATION

Ms. K. Thamizharasi[1], T. Kirubagaran[2], Z. Mohammed Islauddin[3]

[1]Assistant Professor, Computer Science, Jeppiaar SRR Engineering College, Tamil Nadu, India
[2,3]Student, Computer Science, Jeppiaar SRR Engineering College, Tamil Nadu, India

---

Abstract – *This task referred to as OCR and translation, targets at processing a given image, extracting text, translating it and provides output as speech. In this project, we develop a mobile application to implement the following functionalities and process the given input image to extract text by text mining technique with the help of Tesseract -OCR Engine. The application is pre-processed which includes the training phase where multiple input sets are provided which improves accuracy of text mining and character recognition. The Extracted text is translated to user's preferred language. The Translated output text can be converted into a pdf file and multiple files can be created files can be created and saved which will be helpful for future uses. These output files can also be viewed across different platforms. The process and results on sample inputs show that this technique significantly outperforms existing applications and techniques in terms of features, effectiveness and availability.*

Keywords: Optical Character Recognition (OCR), text mining, text recognition, text translation, PDF converter.

1. INTRODUCTION

Mobile services and applications have become an indispensable part of daily life, helpful in various domains. Nowadays, many mobile services and applications depend mainly on internet connection. This application differs in a way that it won't need internet connection so it's available offline and provide easy access point. This application is trained with multiple sample input sets to ensure fast and high accuracy during text mining and recognition. It used Tesseract engine with AI for character recognition. The primary motive of this application is to reduce the language barrier and to assist users with poor eyesight. This application includes various functionalities and features bundled into a single package. The application can also be improved further with support for languages and functionalities.

2. LITERATURE SURVEY

A literature survey or narrative survey is a type of review article. It's a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic. We present a literature survey in the field of Optical Character Recognition and Translation. Table 1 presents a survey overview. It includes all the sources which we have referred for our project along with its details. We have overcome the limitations of existing sources and papers provided in this survey.

Title	Author Name	Algorithms	Limitations
Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data	Mingchen Feng, Jinchang Ren, Amir Hussain, Xiuxiu Yue and Qiyuoyan Liu	Root Mean Square Error	Data Variety support is limited
Signboard Detection and Text Recognition	Mohammad A.Panhwar, Adeelabro	SVM Classifier	Accuracy of text detection is only70.2%.

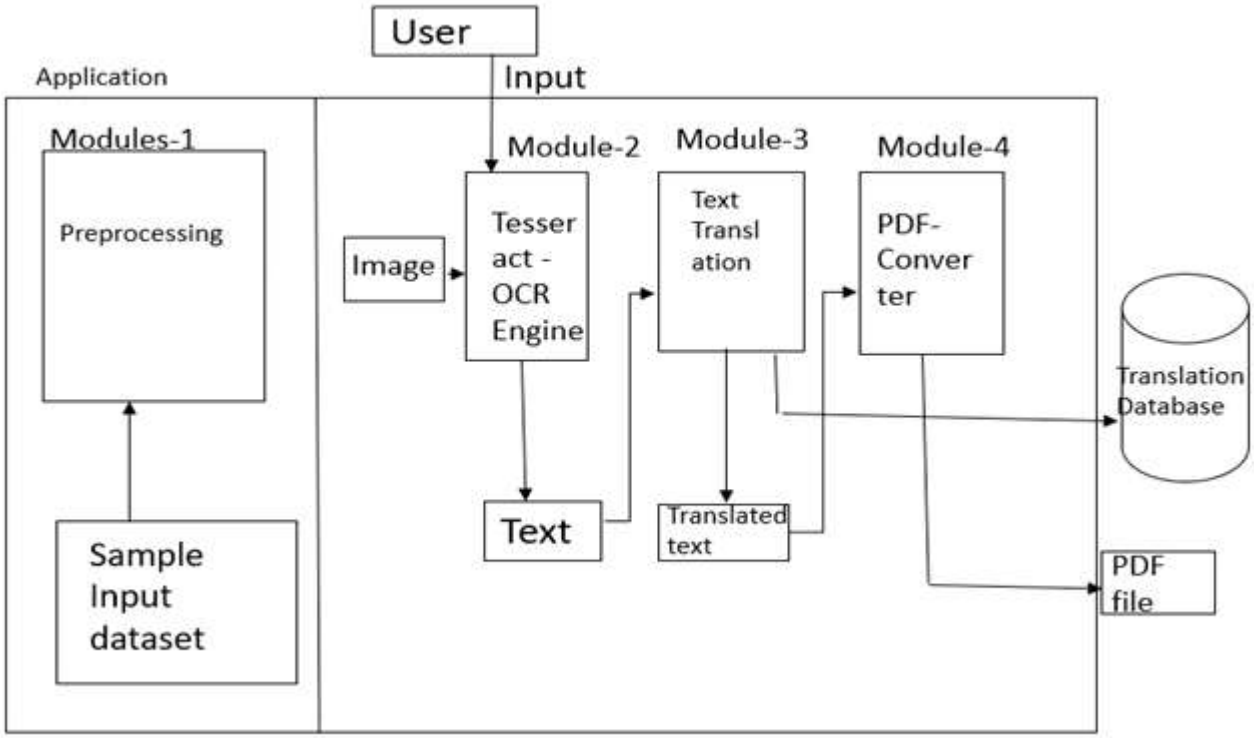
3. CHALLENGES IN EXISTING WORK

Google lens is a popular and commonly used application for image processing. This application requires an active internet connection constantly. U-dictionary is also a similar application which can translate text from one language to another language. This application also requires internet connection.

In many existing systems the converted text can't be converted into PDF and be used in future. The existing works on this field which we have referred and provided in the literature survey all have limitations of their own like limited support for data variety, accuracy of text detection, etc.

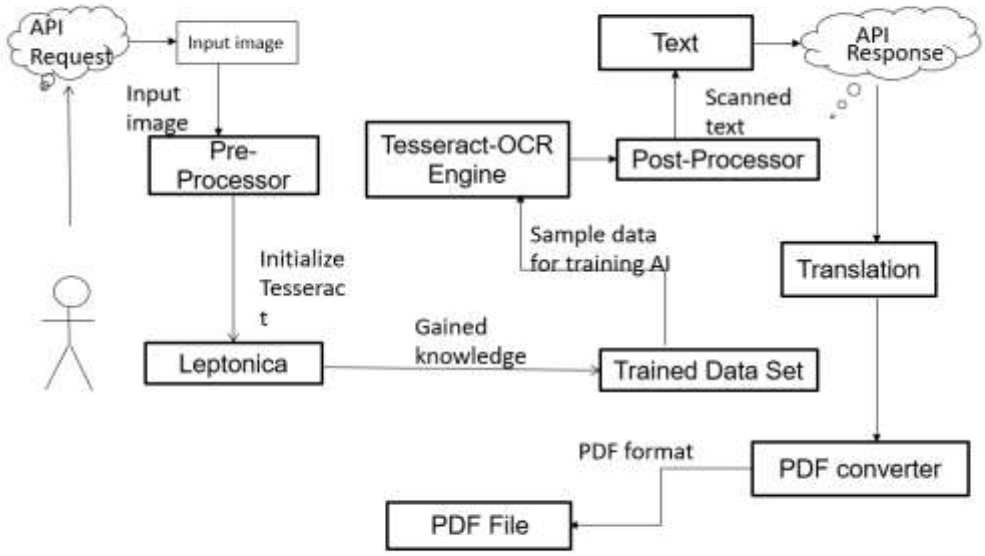
4. ARCHITECTURE DIAGRAM

The architecture diagram explains about the set of concepts that are part of architecture, including the principles, elements and components. It shows all the modules present in the application, the system-user interaction and the overall basic working of the application.



5. DATA-FLOW DIAGRAM

A data-flow diagram (DFD) is a way of representing a flow of a data of a process or a system. The DFD also provides information about the outputs and inputs of each entity and the process itself. The diagram represents specific operations based on the data by a flowchart which includes all the data-flow from the user's input to the output produced by the application. It includes all the modules of the application and the interactions among them in a sequential fashion. It helps greatly in understanding the basic flow of data in the application.



6. PROPOSED MODEL

This mobile application is standalone and does not require internet connection. It uses Tesseract engine with natively provided multiple sample input sets for training the application with AI for Optical Character Recognition to ensure high speed and accuracy. The text can be then translated if necessary to user's preferred language. This application can support multiple languages through updates if needed. It also includes functionality like PDF conversion to convert the translated text to PDF file for future references. Our proposed model helps us to develop an application which can be free of cost and in turn will be very helpful for common people. This application will help illiterate and users with poor eyesight alike and will also reduce language barrier.

7. IMPLEMENTATION

Preprocessing phase is used. This allows the application to train with several sample input sets and improve its AI and machine learning ability. The Tesseract engine is used which uses various techniques like line finding, baseline fitting, fixed pitch detection and chopping, etc. to detect the text from the given input image. Then the text is translated to user's preferred language. The translated text can be converted into PDF file.

8. MODULE DESCRIPTION

We have implemented this application using four modules. These modules include:

- Preprocessing,
- OCR,
- Translation and
- PDF Conversion.

1. PREPROCESSING:

This is the training phase for the application. In this phase, the application is trained with multiple input sets to assist in character recognition. This training is in turn helpful for optical character recognition during runtime.

2. OCR:

This is the phase where the given input image is processed for availability of any text with the help of preprocessing training. Here, all the text present in the given input image is recognized and extracted and it's provided as input to the next module.

3. TRANSLATION:

In this phase the extracted and recognized text is analyzed to determine the language. If the language detected matches the user's preferred language, then it's provided as input to next phase. Else, using Translator API, the text is translated to the user's preferred language and is sent to next module.

4. PDF CONVERSION:

In this phase, the final translated text is converted into a PDF file using this module. Then the PDF file is produced as output and is displayed to the user which can then be saved for future uses and references.

9. CONCLUSION

The mobile application extracts the text from the image with the help of AI and machine learning. The text is translated to the user's preferred language. The translated text can then be saved in PDF format and can be viewed across different platforms for future use. It is helpful for people with poor eyesight to identify text and it will also reduce language barrier. This application can be used in multiple platforms like Windows, Ubuntu, Linux and Android.

10. FUTURE WORK

We have worked on this application to support only translation from one language to another as of now. Our future work involves inclusion for further different languages support for translation and to implement Text-Speech conversion. We also want to make this application work completely offline in the future.

11. REFERENCES

[1] A. J. Hunt and A. W. Black, "Unit selection in a concatenative speech synthesis system using a large speech database," in Proc. IEEE Int.Conf. Acoust, Speech, Signal Process., Atlanta, GA, USA, vol. 1, May 9, 1996, pp. 373–376.

[2] H. A. Murthy et al., "Building unit selection speech synthesis in Indian languages: An initiative by an Indian consortium," in Proc. Co-ordination Standardisation Speech Databases Assessment Technology., Kathmandu, Nepal, Oct. 4, 2010, pp. 1–6.

[3] T. Yoshimura, "Simultaneous modelling of phonetic and prosodic parameters, and characteristic conversion for HMM-based text-to-speech systems (chapter 6)," Ph.D. dissertation, Dept. Elect. Comput. Eng., Nagoya Inst. Technol., Nagoya, Japan, Aug. 2017.

[4] B. Ramani et al., "A common attribute based unified HTS framework for speech synthesis in Indian languages," in Proc. 8th Int. Speech Commun. Assoc. Workshop Speech Synthesis, Barcelona, Catalonia, Aug. 31– Sep. 2, 2013, pp. 291–296.

[5] A. Pradhan, A. Shanmugam, A. Prakash, K. Veezhinathan, and H. Murthy, "A syllable based statistical text to speech system," in Proc. Eur. Signal Process. Conf., Shanghai, China, Sep.9–13, 2013, pp. 1–5.

[6] A. Baby, N. L. Nishanthi, A. L. Thomas, and H. A. Murthy, "A unified parser for Indian languages," in Proc. Text Speech Dialogue, Brno, Czech Republic, Sep.12–16, 2016, pp. 514–521.

[7] A. Baby, J. J. Prakash, R. Vignesh, and H. A. Murthy, "Deep learning techniques in tandem with signal processing cues for phonetic segmentation for text to speech synthesis in Indian languages," in Proc. INTERSPEECH, Stockholm, Sweden, Aug. 20–24, 2017, pp. 3817–3821.