

Translator based Smart Text and Voice Pattern Analyser for Physically Challenged

Srinivasa K

Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumakuru, India
Email: srinivasa@sit.ac.in

Abstract—In the era of human computer interaction, it is important to break the language barrier between physically challenged people. To overcome the language obstacles, it is necessary to analyze the different ways in which people can communicate with each other, like through text and voice. Hence, in this work, a tool is developed to remove the language barrier between disabled people by developing features like text-to-text translation which provides correct possible outputs for the inaccurate/invalid input which might be given by physically challenged people. Unlike the existing tools like Google Trans, the tool considers inaccurate input and gives the possible accurate outputs to help physically challenged people. The tool also considers file of multiple language text and convert the content into desired language, which is currently not performed by existing tools. The tool also extracts the text from an image and as an add on feature this work extracts information/details of patients from medical diagnostics reports. These unique features that have been implemented in the work stand out to be ultimately helpful for people with language diversity and especially to physically challenged ones.

Index Terms—NLP, Multilingual, and Machine Translation.

I. INTRODUCTION

In today's interconnected world, the ability to communicate across language barriers has become increasingly crucial. In this regard, language translation and text processing technologies [1] have witnessed significant advancements in recent years. Even though many tools are developed to overcome the language barriers, the existing tools are not feasible to be used by physically challenged ones. To address this issue, this paper explores the objectives and advancements in several key areas of language translation and text processing, including text-to-text translation [2] which provides all the possible outputs for the given inaccurate input for physically challenged people, extraction of text from an image [10] and translating input text file having content in multiple languages to a target language. Existing systems can translate text, handwriting and speech from one language to another language, given the source language [3]. However, the existing systems expect from the user to provide the proper input to get the accurate input. Hence, this work considers an invalid or wrong input and provides the possible accurate outcome which finds very helpful for physically challenged people as the disabled people are unable to use all the keystrokes to provide the accurate input. The proposed system also includes a feature where a text with multiple languages is given as input and produces output in a user specified target language. The existing tools [11] allow users to take a photograph of a document, signboard etc., and recognize the text from the image using OCR technology and gives the translation.

The enhancement of this feature is achieved in the work by extraction of patient details from a medical diagnostics report for faster information collection of patients in the form of text. The primary contribution of this paper includes introducing a tool to overcome language barrier for people with disability to translate an inaccurate text in a source language to an accurate text in a target language, extraction of text from an image for extracting information from medical reports and translating input text file having content in multiple languages to a target language with minimum key strokes.

II. RELATED WORK

Here, the study of related works is emphasized on the implementation of systems that do translation to remove the communication barrier. Some existing works such as [1,2] have proposed text to text translation methods where advanced deep learning algorithms are used to provide accurate translations of text from one language to another. They rely on large neural networks that are trained on what amounts of multilingual data to learn how to translate text from one language to another.

These services use optical character recognition (OCR) technology to identify and extract the text from the image. [1,2] have proposed text to text translation methods where, advanced deep learning algorithms are used to provide accurate translations of text from one language to another. They rely on large neural networks that are trained on what amounts of multilingual data to learn how to translate text from one language to another.

These services use optical character recognition (OCR) technology to identify and extract the text from the image. Similarly,[3] is a neural machine translation system that has been trained on a massive multilingual corpus of text data. While it is primarily designed for text-to-text translation, it may also be able to extract text from images using OCR technology.

A recent work [6] proposed a framework for developing natural user interfaces. It allows developers to create cross-platform applications that can be used with touchscreens, and it includes built-in support for gesture recognition, multi-touch, and other features that are commonly used in natural user interfaces. Similarly, [5] is the system that works by analyzing the content of the input files and comparing it to a large database of known language patterns.

For text files, Detectsy uses natural language processing (NLP) techniques to analyze the text and identify the language. NLP involves analyzing the structure, grammar, and vocabulary of a text to determine its language. For images, Detectsy uses optical character recognition (OCR) to extract any text from the image and then uses NLP to determine the language of the extracted text.

A recent work [8] is a book that teaches developers how to create graphical user interfaces (GUIs) using the wx-Python toolkit. wxPython is a Python implementation of the cross-platform wxWidgets C++ toolkit, which provides a set of tools and libraries for creating GUI applications that can run on multiple operating systems, including Windows, macOS, and Linux. Works such as [9,10,11] emphasize text extraction and recognition from images. The process [9] involves using machine learning techniques to analyze an image and identify any text present in it, before converting that text into a machine-readable format.

This can be useful in a variety of contexts, such as extracting data from documents or identifying text in photographs. Keras-OCR is a Python library that provides a pipeline for text recognition from images using deep learning models[10].Some of the best ways to extract text from images without Tesseract in Python such as PyOCR, OCRopus, Kraken, EasyOCR, Google Cloud Vision AP, OpenCV have been included.

III. PROBLEM DEFINITION AND OBJECTIVES

As India is a multilingual country, language barrier is one of the main problems that need to be addressed. Even though the problem is addressed to a large extent, the existing tools are not feasible to be used by the disabled people as they can't use the complete keystrokes to provide the accurate input.

Hence, we design and implement a tool that addresses language obstacles that occur during communication with minimum keystrokes and high accuracy. To address the problem described above, the following objectives are set:

- To Translate the text from one language to another, which provides all the possible outputs for the given inaccurate input for physically challenged people.
- To Extract text from any image and translate into any other language which is used to extract information from medical reports.
- To identify and translate an input text file that may have the content written in multiple languages into an user chosen language.

IV. PROPOSED METHOD

The overview of the proposed system is as shown in the figure 1. The design methodology of each module shown in figure 1 is described in the further sub sections.

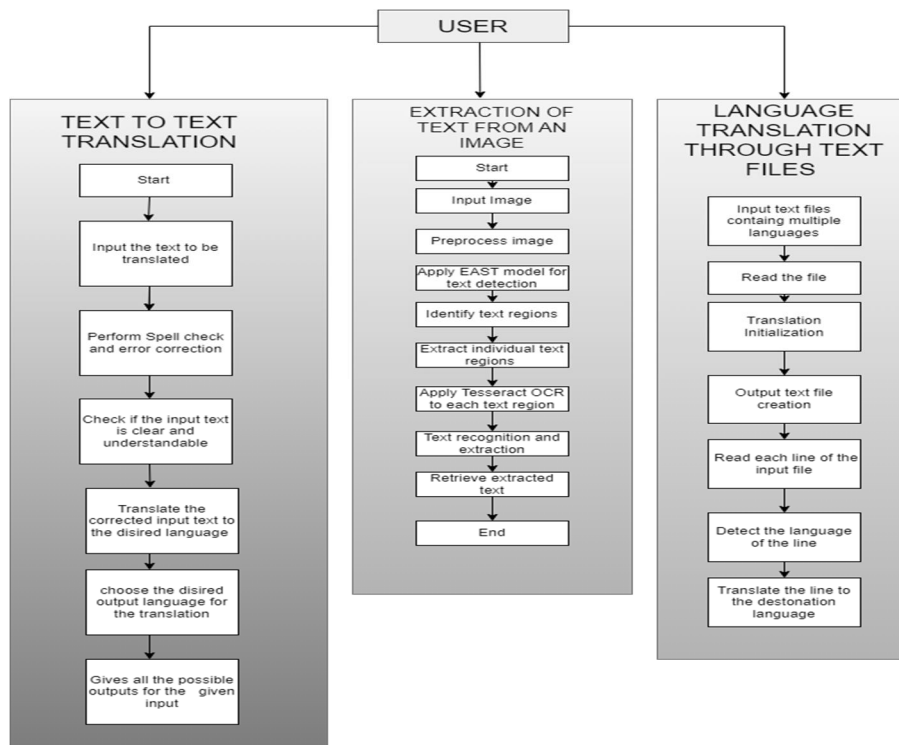


Fig. 1 Overview of the Proposed System

A. Text to Text Translation

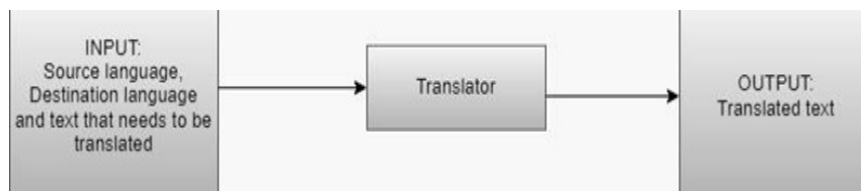


Fig. 2 Working of Text-to-Text Translation

The Fig. 2 represents the working of text to text translation, in which the user needs to provide source language, destination language and the text that needs to be translated as input to the system. The text entered inaccurately will be preprocessed using basic NLP techniques like tokenization, stemming, spell checker, Synonym identification using wordnet and so on to identify the accurate form of the entered text. The libraries used for this model are NLTK and GTRANS which are free and include limitless libraries.

B. Translation Through Text File with Multi-Linugal Texts

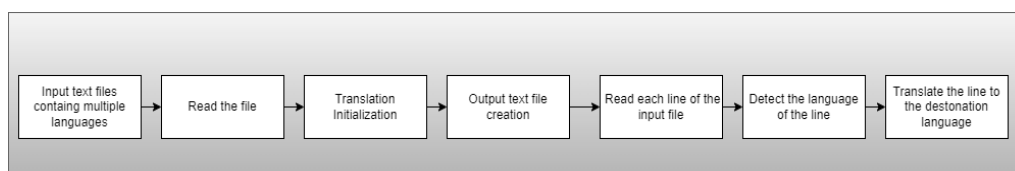


Fig. 3 Flow Diagram of Implementation of language identification and translation through text file

The Fig. 3 represents the implementation of language identification and translation through text file wherein user gives input in the form of text file that contains multiple language texts and obtains the output as translated text

file in user chosen destination language. For each of the sentences in multiple languages, the language is detected by the language detection module and then the sentences will be translated to the desired target language specified by the user.

C. Extraction of Text from an Image

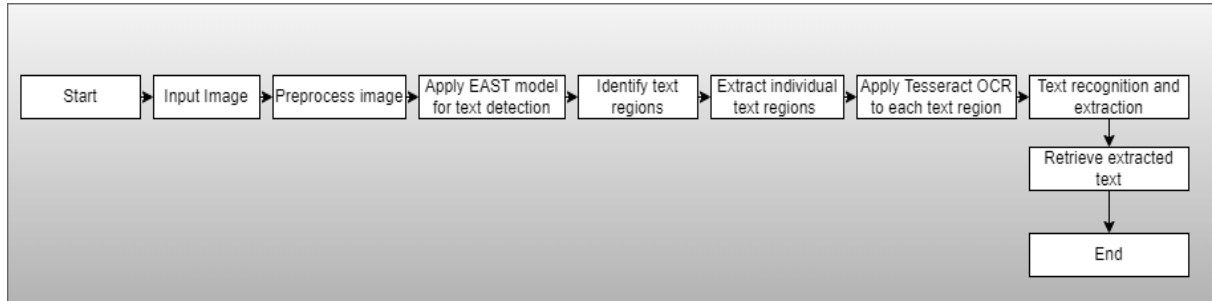


Fig. 4 Flow Diagram of Implementation of Extraction of text from an image

The Fig. 4 represents the implementation of Extraction of text from an image where the user can upload image or the medical diagnostic report. EAST model is used here for the text detection and Tesseract OCR is used for the extraction of text.

V. RESULT AND COMPARISON STUDY

The proposed system is developed in the form of a web application. Fig. [5] represents the home page through which users can navigate. Sample outputs for the Translate text with different languages, Translate text file and Extraction of text from Image are displayed in Fig. [6], [7] and [8] respectively.

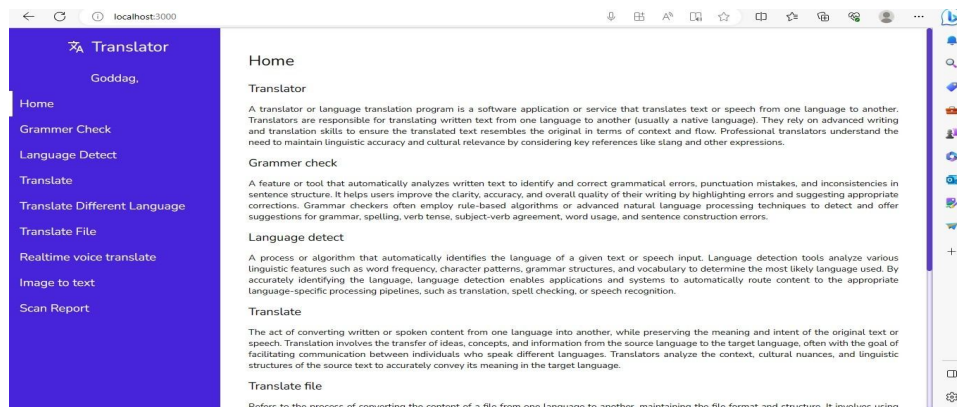


Fig. 5 Snapshot of Home Page

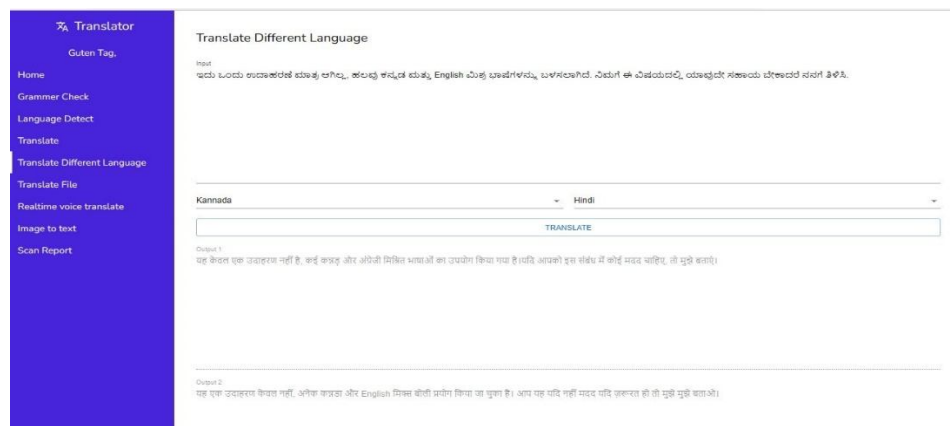


Fig. 6 Text to Text Translator

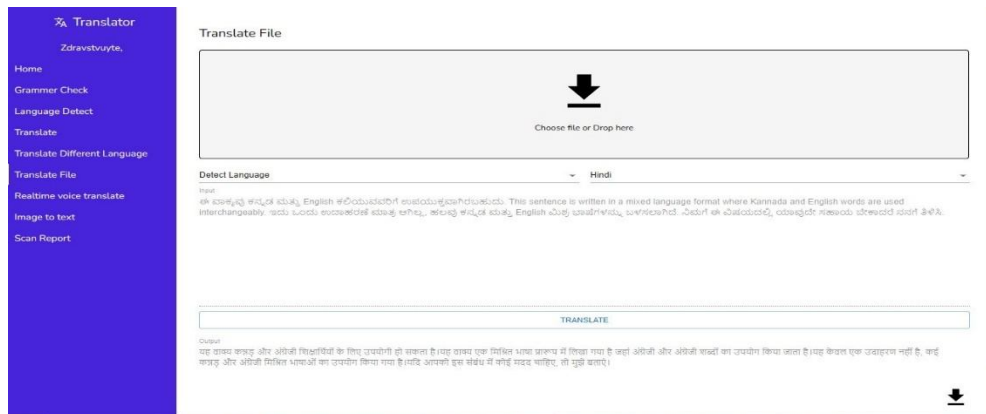


Fig. 7 Text File Translator

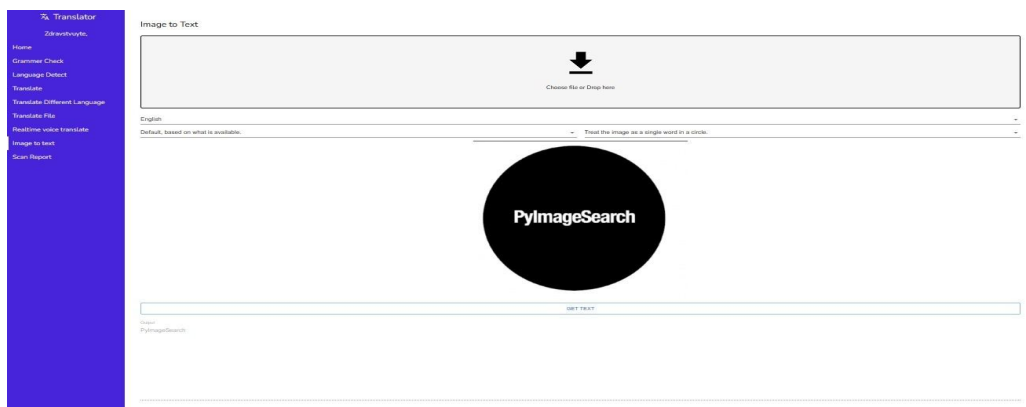


Fig. 8 Extraction of Text from Image

The evaluation metrics for the translation of input text file with multiple language content will be the number of lines versus the percentage of accuracy achieved during the translation. The bar graph in Fig.9 displays the accuracy range accomplished with respect to the number of lines for the text file translation.

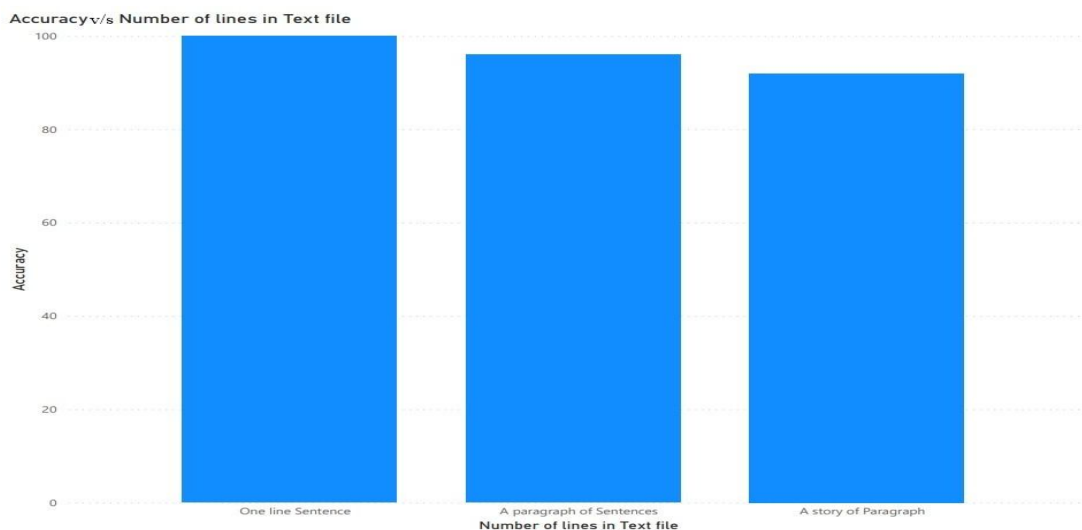


Fig. 9 Graph of Accuracy v/s Number of Lines in Text File

Accuracy for the Extraction of Text from an image is purely dependent on the Clarity or the percentage of Noise present in the image provided. Fig.10 is the graphical representation of how accuracy of the Extraction of text module is directly proportional to the clarity of the Image.

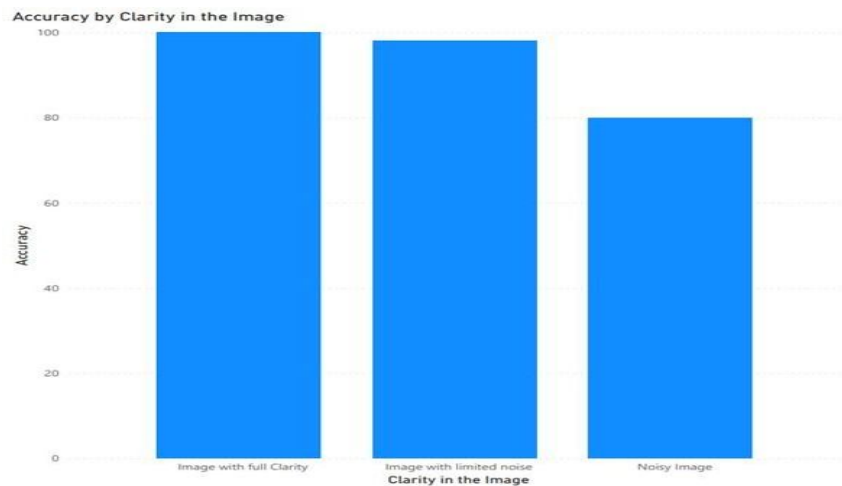


Fig. 10 Graph of Accuracy v/s Clarity in the image

VI. CONCLUSION

The system demonstrated broad language coverage, supporting a wide range of languages for translation. It successfully handled various language pairs. This expansive coverage enhances the system's usability and makes it accessible physically challenged people. The translation system achieved a high level of accuracy in translating text across multiple language pairs. It consistently produced translations that confront to the original meaning of the input text. Future systems can be created to open up new channels for inclusive communication by combining translation capabilities with other modalities like picture recognition or sign language processing. This would make it easier for those who are physically disabled to translate text that is available in images or videos and to translate between spoken and sign languages. In order to offer more handiness, user can specify the source and destination language through voice and the output of the translation of text could be in voice. Users of future translation systems might be able to specify their own language preferences, prioritizing particular languages or setting translation rules for content that is multilingual. User's language choices, dialect preferences, or translation standards can be specified, resulting in more accurate and personalised translations for their particular needs.

REFERENCES

- [1] Takakusagi, Yosuke, Takahiro Oike, Katsuyuki Shirai, Hiro Sato, Kio Kano, Satoshi Shima, Keisuke Tsuchida et al. "Validation of the reliability of machine translation for a medical article from Japanese to English using deep translator." *Cureus* 13, no. 9 (2021).
- [2] Junczys-Dowmunt, Marcin. "Microsoft translator at WMT 2019: Towards large-scale document-level neural machine translation." *arXiv preprint arXiv:1907.06170* (2019).
- [3] Fan, Angela. "Introducing the first AI model that translates 100 languages without relying on English." (2020).
- [4] Millstein, Frank. *Natural language processing with python: natural language processing using NLTK*. Frank Millstein, 2020.
- [5] Foote, Jonathan. "An overview of audio information retrieval." *Multimedia systems* 7, no. 1 (1999): 2-10.
- [6] Fernandez, Ramon A. Suarez, Jose Luis Sanchez-Lopez, Carlos Sampedro, Hriday Bavle, Martin Molina, and Pascual Campoy. "Natural user interfaces for human-drone multi-modal interaction." In *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1013-1022. IEEE, 2016.
- [7] Hassen Ghalila, Ahmed Ammar, Vasudevan Lakshminarayanan. *GUI Programming with Python and Qt Vol. 31(12)*, 2055-2076.
- [8] Rappin, Noel. *wxPython in Action*. 2006.
- [9] Misra, C., P. K. Swain, and J. K. Mantri. "Text extraction and recognition from image using neural network." *International Journal of Computer Applications* 40, no. 2 (2012): 13-19.
- [10] Alrasheed, Nouf, Shivika Prasanna, Ryan Rowland, Praveen Rao, Viviana Grieco, and Martin Wasserman. "Evaluation of Deep Learning Techniques for Content Extraction in Spanish Colonial Notary Records." In *Proceedings of the 3rd Workshop on Structuring and Understanding of Multimedia Heritage Contents*, pp. 23-30. 2021.
- [11] Misra, C., P. K. Swain, and J. K. Mantri. "Text extraction and recognition from image using neural network." *International Journal of Computer Applications* 40, no. 2 (2012): 13-19.