IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Project Tourist Translator Using

Hybrid Translation

Pooja Nimhan, Saurabh Pawar, Rushikesh Mane

Student,
Project Guide: Prof. Prateeksha Chouksey.
Computer Department,
Genba Sopanrao Moze College of Engineering, Pune, India.

Abstract: ***When it comes to learning the local language as a visitor in a foreign country, it's not easy. Everything requires an understanding of the local language, from reading signs to being overcharged while buying, booking cabs and hotels for their stay, sightseeing, and communicating with the locals. Nowadays, almost everyone owns a smartphone, which has proven to be the most useful instrument for travelling. We want to create an Android application that can translate text, speech, and textual information written on signboards from one language to another. For signboard translation, users will need to use their smartphone camera. For language translation, we are using a hybrid machine translation approach. Recognition of text from images is done using digital image processing. The goal of this project is to minimize the likelihood of any single component failing. To reduce miscommunication between tourists and locals, a machine translation approach is being used.***

Index Terms **- Tesseract OCR, Image Translation, Hybrid Translation.**

1. Introduction

In this world, 1.35 billion out of 7.8 billion people speak English which is almost less than one-fourth of the population. Due to this reason, tourists find it difficult to communicate with locals, and traveling in areas where English is not spoken might make one feel as if they are trapped inside a mysterious bubble. It becomes harmful if one is not able to read the signboard of the speed limit or no parking, it can cause chaos, thus it becomes a necessity to read the signboard. Because of their incapability to communicate, they are cut off from the rest of the world. To avoid this issue, the software is developed where we can convert text, speech, and signboards to the desired languages using the Yandex Application programming interface.

2. System Flow

In this paper, we propose the following stages for text detection, recognition, and translation algorithms.

2.1 Digital Image Processing

2.1.1 Capturing Image

Tourists can use the camera capture module to capture a single textual picture from a signboard or a natural sight by adjusting the camera capture box on the screen by touching the edges of the capture box. The camera is set to autofocus mode for the duration of the session. After capturing the desired image, the image is given to the Tesseract OCR engine module.

2.1.2 Tesseract OCR Module

This section completes the binarization technique for the taken image, which is then evaluated after the text form. Various phases receive the term or string. In these phases, each letter or character is retrieved from the string, and the requirement to attach the extracted character or disrupt associated characters is assessed. Finally, extracted characters are identified using embedded fuzzy features and training data for a given Unicode character language. After that, each character is combined to form a word in the same order as it was extracted, and the word is compared to the terms in the language dictionary.

2.1.3 Word Matching in the dictionary

Each group of a word sequence is transmitted to the Yandex server's Dictionary software, which will aid in more exact and accurate recognition of the phrase than just presenting the result of a meaningless word. The identified text will now be passed to the Unicode text post-processing module.

2.1.4 Unicode text Processing

After the recognized text is supplied to this module, Unicode characters will be shown, and the user will have the option to convert the identified text into a certain known language by selecting from the drop-down list under the options.

2.2 Hybrid Translation

Hybrid machine translation is a technique that provides multiple machine translation strategies. The major goal of creating a hybrid model is to lessen the likelihood of any single machine translation approach failing. We are combining our system with the Yandex Translation server, which employs the CatBoost algorithm, in the proposed system.
The CatBoost method is a combination of statistical and neural machine translation techniques.

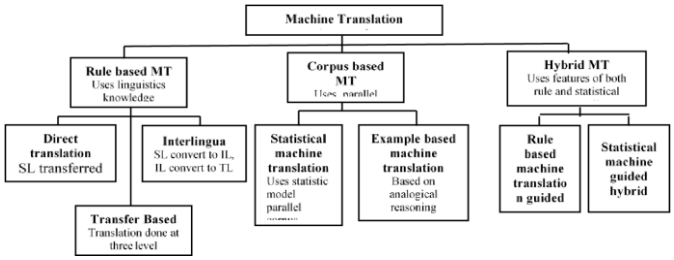

Fig. 1 A diagram showing Hybrid Machine Translation

2.3 Capturing the Signboard Image

The distance between the mobile and the signboard should be approximately 10 meters for optimum image quality. The images were captured with a smartphone camera, and they are frequently blurry and feature complex backgrounds.

2.4 Recognition and text Extraction

Optical Character Recognition is the process of extracting and identifying text from captured images. In this project, we are using the Tesseract OCR engine to extract text from the image you've obtained. Pre-processing, Segmentation, Feature Extraction, Classification, as well as Post Processing, are all steps in the optical character recognition process.

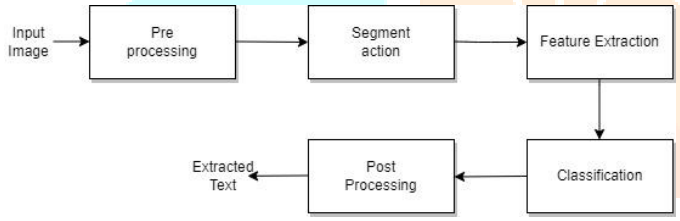

Fig. 2 Optical Character Recognition

2.4.1 Pre-processing Image

Image pre-processing
Image contrast enhancement, noise removal, binarization, and smoothing Transform a grayscale or color image into a binary image, with the backdrop preferably represented by one color and the text by another.

2.4.2 Segmentation of Image

Image segmentation is a typical approach for dividing images into various portions or areas, frequently based on the image pixels, in the processing and analysis of digital images. Color or structure could be used to segment an image. Resemblances between the foreground and background areas in clusters.

2.4.3 Extracting Features

Classification is the process of identifying each character and allocating it to the proper character class. There are two sorts of classification approaches. 1) Methods based on decision theory: The main techniques to decision-theoretical recognition are statistical methods, artificial neural networks, and minimum distance classifiers. 2) Structural Methods: Similarity acts can be expressed using grammatical ideas depending on the structural component's relationship.

2.4.4 Post-processing

Symbol clustering and error handling are performed at this level. The process of gathering symbols into strings is known as a grouping. It is dependent on the location of the symbol in the text. Symbols that have a lot in common are grouped together.

2.3 Text and Voice Translation

For text and voice translation, we used the Yandex translator (CatBoost algorithm), and we used Google speech recognition to recognize words or characters uttered by the user. You must next select the text translator option, select the preferred language in which you want the text to be translated, write text or speak into your microphone, and click the Go button. The app interface will take your data as input and transmit it to the Yandex API Server, where it will be processed, then retrieved by the app interface, and the result displayed in the specified language. This application recognizes over 100 languages from around the world and supports them.

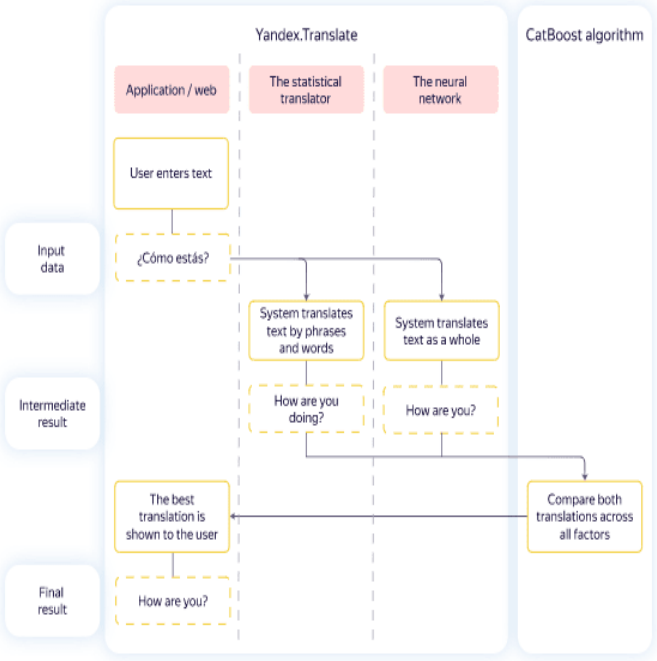

Fig. 3 Yandex Server using Catboost Algorithm for more precise output.

You must next select the text translator option, select the preferred language in which you want the text to be translated, write text or speak into your microphone, and click the Go button. The app interface will take your data as input and transmit it to the Yandex API Server, where it will be processed, then retrieved by the app interface, and the result displayed in the specified language. This application recognizes nearly all available standard symbols and supports over 100 languages around the world.

3. CONCLUSIONS

The proposed technology effectively translated 100+ languages spoken and practiced around the world, and the results were satisfactory. To make better use of technology, we used a hybrid machine translation (CatBoost algorithm) instead of a single machine translation and discovered that hybrid machine translation is superior to single machine translation in terms of accuracy and efficiency. Finally, we can state that this application will be more beneficial in reducing misunderstandings amongst people caused by language differences.

References

[1] N. Mishra and C. Patvardhan, “Atma: Android travel mate application,” International Journal of Computer Applications, vol. 50, no. 16, 2012.
[2] N. Mishra and C. Patvardhan, “Atma: Android travel mate application,” International Journal of Computer Applications, vol. 50, no. 16, 2012.
[3] Thakare, Sahil & Kamble, Ajay & Thengne, Vishal & Kamble, U.R.. (2018). Document Segmentation and Language Translation Using Tesseract-OCR. 148-151. 10.1109/ICIINFS.2018.8721372.
[4] Mobile Camera-Based Text Detection and Translation by Derek Ma, Qiuhau Lin, Tong Zhang.
[5] https://github.com/tesseract-ocr/
[6] https://yandex.com/dev/translate/
[7] A. Canedo-Rodriguez, S. Kim, J. H. Kim, and Y. Blanco-Fernandez, "English to Spanish translation of signboard images from a mobile phone camera," IEEE Southeastcon 2009, Atlanta, GA, 2009, pp. 356-361.DOI: 10.1109/SECON.2009.5174105.