IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134126

Multi Input Translation Between Indian Languages Using Firebase Machine Learning Kit

Sayali Patil[1], Vaishnavi Patil[2], Jayashree Jadhav[3], Arpit Naik[4], Ashish Bhole[5]

[1,2,3,4] UG Engineering Student, [5] Associate Professor

Department of Information Technology, SSBT's College of Engineering and Technology, KBC North Maharashtra University, Jalgaon, Maharashtra, India, 425001

Abstract: The goal is to create a user-friendly mobile app facilitating accurate translation between English and Hindi, encouraging extensive use of Hindi in official contexts. Bridging language gaps in official matters is crucial in our interconnected world, yet current translation tools often lack efficiency, leading to misunderstandings. Our solution entails developing a robust translation software with multi-input capabilities, leveraging state-of-the-art natural language processing techniques. The app will enable text, voice, and image inputs, aiming to simplify language learning and stress-free communication. Ultimately, our project seeks to break down language barriers, promote linguistic inclusivity, and enhance effective communication by empowering users to communicate effectively in their native language while embracing cross-lingual communication.

Keywords: Multi-input Language Translation, Optical Character Recognition(OCR), Firebase ML Kit, User-friendly Interface, Automatic Language Detection.

I. INTRODUCTION

In a globalized world, effective communication across language barriers is paramount. According to the project, the Multimodal Hindi-English Translation System, is designed to offer a versatile translation experience by accommodating text, image, and voice inputs. Leveraging advanced technologies in natural language processing, computer vision, and speech recognition, the system aims to facilitate seamless and accurate communication between Hindi and English speakers. The innovation addresses the diverse needs of users, providing a comprehensive solution for cross-cultural and multilingual interactions [6].

The language translation application integrates Firebase ML Kit for robust text extraction from images, catering to the increasing need for Image-based Text Extraction. Through sophisticated OCR technology, the app adeptly detects, recognizes, and extracts text from images, ensuring accuracy and reliability [5]. The extracted text, undergoes seamless translation into a wide array of languages facilitated by Firebase Translate API, renowned for its accuracy and versatility [7]. By offering multi-language translation capabilities, the application becomes an indispensable tool for travelers, empowering them to effortlessly overcome language barriers and access pertinent information wherever they go [6].

II. LITERATURE SURVEY

In [1], The development of Android-based language translator applications has gained significant attention due to the widespread use of smartphones. Researchers have focused on creating user-friendly and efficient applications that cater to real-time translation needs. Base papers in this domain often explore techniques like neural machine translation (NMT) and deep learning to enhance translation accuracy. Additionally, studies delve into user experience aspects, such as the integration of voice recognition and natural language processing (NLP) for seamless communication. Some papers also discuss the challenges of handling multiple languages and optimizing resource usage on mobile devices.

In [2], The concept of a multi-lingual translation application involves addressing the complexities of translating between multiple languages seamlessly. Base papers in this area frequently emphasize the use of advanced algorithms, including transformer models, to handle diverse language pairs effectively. Researchers focus on optimizing translation quality and speed while considering the linguistic nuances of various languages. Additionally, the papers discuss the

DOI: 10.17148/IJARCCE.2024.134126

integration of contextual information and domain-specific knowledge for improved accuracy. User-centric design and accessibility across different platforms are also common themes in these papers.

In [3], Document segmentation and language translation using Tesseract-OCR (Optical Character Recognition) is a niche area that combines image processing and translation technologies. Base papers often delve into the challenges of accurately segmenting text from images and then employing machine translation techniques. Tesseract-OCR, being an open-source OCR engine, is frequently explored for its capabilities in extracting text from images. Researchers focus on enhancing the accuracy of the OCR process and integrating it seamlessly with language translation models. Papers also discuss applications in real-world scenarios, such as translating text from scanned documents or images captured by mobile devices.

In [4], While the OCR post-correction paper emphasizes unsupervised techniques for improved accuracy, literature on Android translation applications underscores multidisciplinary approaches. It includes the integration of neural network architectures like CNNs and RNNs, along with speech recognition and OCR technologies for diverse input processing. User experience aspects, such as interface design and real-time translation capabilities, are also focal points. Overall, the literature highlights the importance of integrating machine learning and user-centric design principles to create efficient translation solutions across text, voice, and image inputs on Android platforms.

III. PROBLEM STATEMENT

The challenge is to develop a robust language translator application capable of processing multiple input modalities, including text, images, and voice, to facilitate accurate and contextually relevant translations between Hindi and English. The application needs to navigate the intricacies of linguistic diversity and user preferences, allowing for seamless communication in various scenarios, such as written text, visual content, or spoken interactions. The complexity lies in integrating natural language processing algorithms capable of understanding and translating diverse forms of input while maintaining accuracy and user friendliness

IV. SYSTEM ARCHITECTURE

System architecture refers to the overall design and organization of a complex system, which includes hardware, software, and various components interact with each other to achieve specific goals. It involves defining the system's structure, components, and interfaces, as well as the relationships between them. A good system architecture ensures the system is reliable, scalable, and maintainable, while meeting the functional and non-functional requirements. In Figure 1 the architecture of the system is presented.

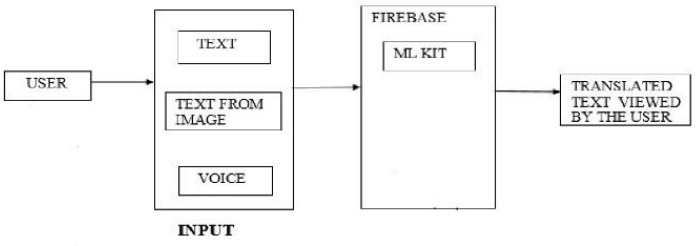

Figure 1 – System Architecture

The Language Translator application features a comprehensive system architecture encompassing the client layer, where users interact with the front end and state management, facilitating input in text, image, and audio formats. The internet layer manages communication between the client and the server, enabling the seamless exchange of translation requests and results. On the server side, the application logic, divided into front-end and back-end components, includes API integration with external language models and translation APIs. The backend server executes core business logic, interacting with an authentication middleware, API layer, and a database server storing language models, translation rules, and user preferences. The architecture ensures secure, responsive, and personalized translations across diverse user inputs.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134126

V. METHODOLOGY

A] Text Extraction (OCR)

A mobile application has been developed using Firebase ML Kit to extract text from images [9,6]. It means the app can recognize and understand text found in pictures taken by users, whether it's text in a scene or document images. It can identify and convert it into digital text that can be used, helping users who may struggle with understanding different languages or need information from images quickly.

To recognize text in an image, utilize Firebase ML Kit's FirebaseVisionImage object, created from various sources like Bitmap files or byte arrays [8]. Then, pass object to the Firebase Vision Text Recognizer's processImage method for text extraction, aided by helper methods for image handling and metadata management.

B] Speech Recognition

Speech recognition for language translation utilizes Firebase ML Kit. The process involves converting acoustic speech into words through feature extraction, building an acoustic models database, using a dictionary and language model, and employing the Viterbi algorithm for decoding. It enables the recognition of spoken utterances and facilitates the reverse operation of translating them into text sequences, leveraging Firebase ML Kit for efficient recognition [10].

C] Translation

The methodology for translating text using Firebase ML Kit involves utilizing Firebase's translation services to process text inputs. The text to be translated is passed to Firebase ML Kit's translation module, which employs advanced algorithms to accurately translate the text into the desired language. The process leverages Firebase ML Kit's powerful language processing capabilities to provide seamless and accurate translations, enhancing user communication and understanding across different languages [3,5].

VI. IMPLEMENTATION

The overall functionality of the Language Translation Application Between Hindi and English along with four different languages, can be divided into 3 modules, which are detailed below:

1. Input Module: The Input Module serves as the entry point for user interactions within the language translation application. It facilitates multiple input methods, including text entry, image selection/upload, and voice input functionalities, ensuring versatility and accessibility for users. The module meticulously validates and preprocesses user input to maintain data integrity and enhance translation accuracy. Through intuitive user interface elements, such as input fields and selection buttons, users can seamlessly navigate between different input modes, optimizing their interaction experience.

2. Translation Module: At the core of the language translation application lies the Translation Module, which leverages Firebase ML Kit to provide seamless translation between Hindi and English languages. The module serve as the backbone for language processing, orchestrating translation requests originating from various input sources, including text, images, and voice inputs. Ensures accurate and efficient language translation while preserving linguistic nuances and contextual relevance. The Translation Module optimizes resource utilization and response time, delivering timely translations to the Output Display Module for presentation to the user.

3. Output Display Module: The Output Display Module is responsible for presenting translated text to users in a visually appealing and comprehensible format within the application's user interface. Orchestrates the formatting and presentation of translated content Implementation based on the input source and user preferences, ensuring readability and clarity. The module dynamically adjusts text layout and styling to accommodate varying input lengths and translation outputs, enhancing user engagement and comprehension.

4. Integration Module: The Integration Module acts as the central nervous system of the language translation application, facilitating seamless communication and coordination among different modules. Serve as the bridge that connects the Input, Translation, and Output Display modules, ensuring cohesive functionality and optimal performance. Through

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134126

efficient integration mechanisms, Module orchestrates data flow and interaction, optimizing resource utilization and enhancing system efficiency

VII. RESULTS

The language translation application, featuring text, image, and voice input capabilities in Hindi, English, and four other languages, leverages Firebase ML Kit for exceptional accuracy and efficiency. Evaluation highlights successful alignment with objectives, robust functionality, and a user-friendly interface, enabling seamless communication. Real-time translation, superior performance metrics, and iterative improvements underscore its significance in bridging linguistic barriers and facilitating cross-cultural communication.

Text to Text Translation between Languages is shown in Table 1.

Table 1 Text to Text Translation between Languages

Language	English	Hindi	Marathi	Telugu	Tamil	Gujarati
English		✓	✓	✓	✓	✓
Hindi	✓		✓	✓	✓	✓
Marathi	✓	✓		✓	✓	✓
Telugu	✓	✓	✓		✓	✓
Tamil	✓	✓	✓	✓		✓
Gujarati	✓	✓	✓	✓	✓	

Image to Text Translation between Languages is shown in Table 2.

Table 2 - Image to Text Translation between Languages

Language	English	Hindi	Marathi	Telugu	Tamil	Gujarati
English		✓	✓	✓	✓	✓
Hindi	✓		✓	✓	✓	✓
Marathi	✓	✓		✓	✓	✓
Telugu	✓	✓	✓		✓	✓
Tamil	✓	✓	✓	✓		✓
Gujarati	✓	✓	✓	✓	✓	

Voice to Text Translation between Languages is shown in Table 3.

Table 3 - Voice to Text Translation between Languages

Language	English	Hindi	Marathi	Telugu	Tamil	Gujarati
English		✓	✓	✓	✓	✓
Hindi	✓		✓	✓	✓	✓
Marathi	✓	✓		✓	✓	✓
Telugu	✓	✓	✓		✓	✓
Tamil	✓	✓	✓	✓		✓
Gujarati	✓	✓	✓	✓	✓	

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134126

VII. CONCLUSION

In conclusion, The study emphasizes the crucial role of language diversity in text translation applications. The integration of languages like English, Hindi, Marathi, Telugu, Tamil, and Gujarati highlights the need for robust development frameworks. Our project, focusing on Hindi and English translation, has significantly improved communication through multi input methods. Leveraging advanced technology ensures accurate translations, fostering better understanding between users.

The language translation app will refine its interface, incorporate real-time collaboration, and enhance translations based on user feedback to ensure accuracy and relevance amid evolving languages.

VIII. REFERENCES

[1]. Ogundokun, Roseline Oluwaseun, Joseph Bamidele Awotunde, Sanjay Misra, Tobe Segun-Owolabi, Emmanuel Abidemi Adeniyi, and Vivek Jaglan. ”An android based language translator application.” In Journal of Physics: Conference Series, vol. 1767, no. 1, p. 012032. IOP Publishing, 2021.
[2]. Li, Yang, and Takayuki Fujimoto. ”A Concept of Multi-lingual Translation Application.” In 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI), pp. 929-931. IEEE, 2018.
[3]. Thakare, Sahil, Ajay Kamble, Vishal Thengne, and U. R. Kamble. ”Document segmentation and language translation using tesseract-ocr.” In 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), pp. 148-151. IEEE, 2018
[4]. Multi-Input Attention for Unsupervised OCR Correction](https://aclanthology.org/P18-1220) (Dong & Smith, ACL 2018)
[5]. Somani, I., Chaudhary, D., Srivastava, D. K., & Shekhawat, D. (2022). Translateit: Android-Based Mobile Application For Multilingual Translation. In Ict Analysis And Applications: Proceedings Of Ict4sd 2022 (pp. 241-248). Singapore: Springer Nature Singapore.
[6]. Bagal, Vipul, Kiran Gaykar, and Ms Purnima Ahirao. "Image based Text Translation using Firebase ML Kit." Grenze International Journal of Engineering & Technology (GIJET) 8.2 (2022).
[7]. Neoh, Desmond Kim Seng, Eric Chun Meng Lock, Kem Ho Lew, Hong Yip Tang, Dong Ling Tong, Tun Tai Tan, and Kwan Lee Tseu. "PictoText: Text Recognition Using Firebase Machine Learning Kit." In 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), pp. 1-5. IEEE, 2021.
[8]. Chavan, Kaushik, Roshan Bhoir, Nikhil Gawade, and Surekha Khot. "Number/Text Translate from Image." International Research Journal of Engineering and Technology (IRJET) 1889 (1886).
[9]. Vasudev, T., K. N. Chaithanya, R. Madhushree, B. P. Apoorva, and N. Arpitha. "Translation of Recognized Text from Images."
[10]. Reddy, B. Raghavendhar, and Erukala Mahender. "Speech to text conversion using android platform." International Journal of Engineering Research and Applications (IJERA) 3, no. 1 (2013): 253-258.

