IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Multitrans: An Indian Language Translator

[1]D Likitha Raju, [2]M Vaishnavi, [3]Nandigam Sravitha, [4]Varshini B S, [5]Sneha Girish

[-4]Students, [5]Proffessor

[1-5] Department of Artificial Intelligence and Machine Learning

[1-5] KS Institute of Technology, Bangalore, India

Abstract: The Multilingual Translator has the ability to regulate a range of input formats, consisting of speech, images, documents and text. By offering precise and effective translations between several Indian languages, the objective is to remove linguistic obstacles and promote international contact. To accomplish its goals, the project makes use of already existing machine translation technology and APIs. With text translation, users can enter text in a specific language and get an output in the language of their choice. When translating images, text is first extracted from the images using optical character recognition (OCR), then the translated text is displayed below the original image. Users can upload documents in supported formats (such as txt) in .txt form for translation using document translation. The system processes the document, extracts text, translates it, and presents the translated text. Audio translation allows users to speak in one language, and the system converts the speech to text, translates it, and outputs both the translated text and synthesized speech.

Index Terms **-** Text, Image, Audio, Document, Streamlit, Google Translator APIs, Optical Character Recognition (OCR).

I. INTRODUCTION

A multilingual translation tool that supports text, image, document, and voice inputs is essential for overcoming language barriers and enabling seamless global communication. This project aims to develop an advanced system using neural machine translation and quality training data to ensure accurate, context-aware translations across selected language pairs. Key features include automatic language detection, translation memory, glossary management, and post-editing support. The tool integrates four main components: text translation, image translation using OCR, document translation with formatting retention, and voice translation with real-time speech synthesis. Leveraging technologies like NLP, computer vision, and speech recognition, the system prioritizes usability, accessibility, and ethical considerations such as bias reduction and data privacy. Designed for a wide range of users—from individuals seeking quick translations to organizations needing professional accuracy—it supports applications in education, healthcare, international business, and travel, promoting cross-cultural collaboration and inclusivity.

II.LITERATURE REVIEW

[1] "Converting Images to Multilingual Text for Literacy Education"

The goal of this [1] project is to create a visual literacy program that can translate text embedded in images into several languages. For precise text extraction and recognition in a variety of scenarios, it makes use of optical character recognition, or OCR. By using a smartphone to take pictures, the program allows for instant text translation, improving language and literacy instruction, and acting as a visitor's aide.

[2] "Image-Based Multilingual Speech and Text Recognition and Translation"

The goal of this study [2] is to automate language translation to break down obstacles to communication between nations and regions. The program combines image-based text translation, speech synthesis,

translation, and voice recognition. It gives audio output for translations and translates written or spoken input into a user-specified language. Furthermore, it can identify text from stored or captured images, translates it, and displays the results, facilitating seamless multilingual communication.

[3] "A Survey of Deep Learning Methods for Document Understanding and OCR"

With an emphasis on English-language texts, this survey [3] examines cutting-edge methods for autonomous document interpretation in domains such as technology, finance, and law. It showcases deep learning developments in computer vision and natural language processing, combining approaches to direct further study in this area.

[4] "Using OCR Tesseract and Deep Learning to Transcribing and Translating Bilingual Text."
In order to improve bilingual text transcription and translation, this study [4] investigates the combination of deep learning and OCR Tesseract. Bilingual datasets are cleaned and arranged for model training using OCR for text recognition and Deep Learning for translation. The results demonstrate how well deep learning preserves semantic meaning, leading to improvements in language localization, document digitization, and intercultural communication.

[5] "Real-Time Sequence to Sequence Approach for Text and Speech Translation"
This study [5] describes the creation of a real-time application that uses better translation techniques to facilitate smooth cross-linguistic conversation. Language barriers are effectively broken by the program, which offers chat, audio, and video translations by augmenting conventional convolutional networks with a sequence-to-sequence technique.

[6] "Translation from Speech to Text Facilitating Multilingualism"
This study [6] improves on the current Google Speech Recognition model by creating a multilingual speech recognition model that translates spoken input into text in the user's preferred language. It uses the concepts of natural language processing to bridge communication gaps, allowing even illiterate people to interact with computers in their native tongue.

[7]"An Integrated Model for Linguistic Conversion from Text to Text, Image to Text, and Audio to Text Using Machine Learning Approach"
With an emphasis on Indian languages, this survey [7] offers a machine learning-based model for text, image, and audio-to-text conversion. It uses cutting-edge methods for precise transcription and translation, tested on big datasets, and supports applications such as accessibility, language learning, and cross-language communication to successfully close linguistic gaps.

[8]"An analytical investigation into multilingual machine translation"
In order to overcome communication impediments, this study [8] highlights developments in multilingual machine translation (MT) systems that translate several language pairs. With an emphasis on issues like corpus size and the shift from monolingual to multilingual MT systems for international communication, it explores AI-driven NLP applications.

[9] “OCR and Speech Recognition System Using Machine Learning”
This paper [9] discusses the development of a system that integrates optical character recognition (OCR) and speech recognition using machine learning techniques. It focuses on converting text and speech inputs into digital formats, which can be further utilized for translation and other applications.

III. SOFTWARE REQUIREMENTS

• Streamlit: A lightweight and interactive web application framework designed for Python. It enables easy deployment of data-driven applications, making it ideal for showcasing the multilingual translator's functionalities.

• Speech Recognition: A library for performing speech recognition with support for various APIs. It is used to convert audio input into text, enabling the audio-to-audio translation module.

• gTTS (Google Text-to-Speech): A Python library that generates speech from text using Google's Text-to-Speech API. It facilitates the audio output for translated text in the audio-to-audio module.

• Googletrans/DeepTranslator: Libraries for leveraging Google Translate API, offering robust and context-aware translations between multiple languages for text, images, documents, and audio.
• EasyOCR: An Optical Character Recognition library that extracts text from images. It supports multiple languages, making it a core component of the image-to-text extraction translator.

• OpenCV: A popular computer vision library used for image preprocessing tasks. It ensures the quality of images fed into the OCR module for improved text recognition.

• Pillow(PIL): A Python Imaging Library used to handle and manipulate image files. It is essential for processing and displaying images uploaded by users.

• Matplotlib: A visualization library used for image annotation and displaying results in the image text extraction module.

• tempfile: A standard Python library to create temporary files and directories. It manages temporary audio and image files efficiently during the application's runtime.

• Base64: A Python module for encoding images into base64 format. This is used to embed images as backgrounds within the Streamlit web application.

• NumPy: A fundamental library for numerical computations in Python. It supports efficient array manipulation for image data processing

IV. DESIGN

The multilingual translator project is designed with a modular and user-friendly architecture to ensure seamless integration of diverse functionalities. The key aspects of the design include:

1.System Architecture:

The system follows a modular design, where each component (text-to-text translator, image-to-text extraction translator, document translator, and audio-to-audio translator) operates independently but integrates cohesively into the overall framework.

2.User Interface Design:

- A web-based interface built using Streamlit provides an intuitive platform for users to interact with the system. Users can upload inputs (text, images, documents, or audio), specify language preferences, and receive results effortlessly.
- The design emphasizes accessibility, making the application easy to use for non-technical users.

3. Backend Components:

- The backend handles core functionalities such as Optical Character Recognition (OCR), translation, and speech processing using efficient Python libraries. Each module is optimized to perform specific tasks accurately and in real-time.

4. Output Presentation:
- Outputs are displayed in a clean and organized manner, ensuring that users can easily interpret results for text, images, documents, and audio.

5. Scalability and Maintainability:
- The system is designed to be scalable, allowing for the addition of more languages or advanced features in the future without major modifications.

V. METHODOLOGY

The methodology outlines the step-by-step approach used to implement and integrate various components of the multilingual translator:

1.Input Handling:
- Text-to-Text Translation: Directly process user-provided text for translation.
- Image-to-Text Extraction: Uploaded images are pre-processed using OpenCV and passed through Easy OCR to extract text accurately.
- Document Translation: Uploaded text files are read and processed for translation while preserving formatting.
- Audio-to-Audio Translation: Speech Recognition captures audio input, which is then transcribed into text for translation and synthesized back into speech.

2. Language Translation:

- Translation is performed using Googletrans or DeepTranslator, ensuring accurate and context aware translations across various language pairs.

3. Output Generation:
- Text Output: Translated text is displayed for text, image, and document inputs.
- Audio Output: Translated text is converted into audio using gTTS, enabling real-time playback for audio-to-audio translation.

4.Performance Optimization:
- Techniques like image preprocessing and efficient API calls are employed to ensure low latency and high accuracy.
- Temporary files and memory management are handled using the tempfile library to optimize performance.

5.Integration of Technologies:
- OCR and Image Processing: EasyOCR, supported by OpenCV, extracts text from images.
- Speech Recognition and Synthesis: Speech Recognition captures and processes audio, while gTTS generates speech from translated text.
- Translation APIs: Googletrans/DeepTranslator APIs are utilized to handle multilingual translation needs.

6.Testing and Validation:
- Each component is tested independently and then integrated into the system for end-to-end validation. User feedback is considered to refine the application and ensure reliability.

VI. IMPLEMENTATION

1) TEXT-TEXT TRANSLATOR

2) IMAGE-TEXT TRANSLATOR

3) DOCUMENT TRANSLATOR

4) AUDIO-AUDIO TRANSLATOR

VII. RESULT

1) Text Translator

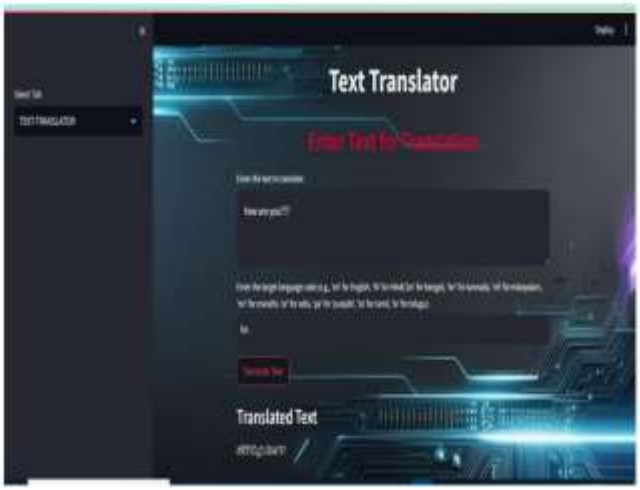

Fig 1

2) Image Translator

Fig 2

Fig 2.1

3) Document Translation

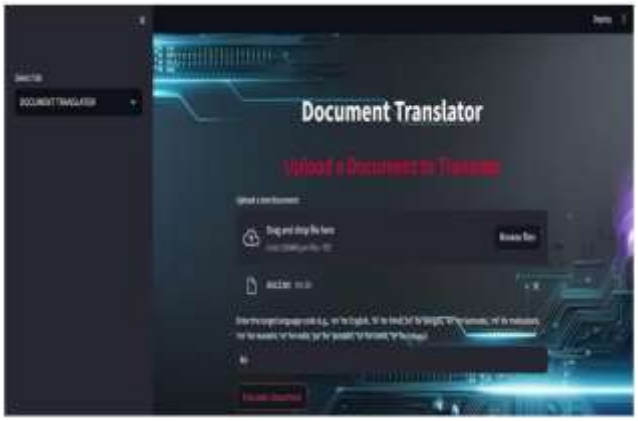

Fig 3

Fig 3.1

4) Audio Translation

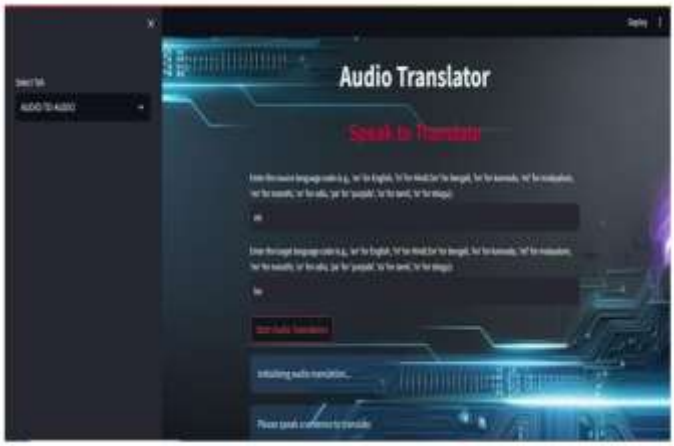

Fig 4

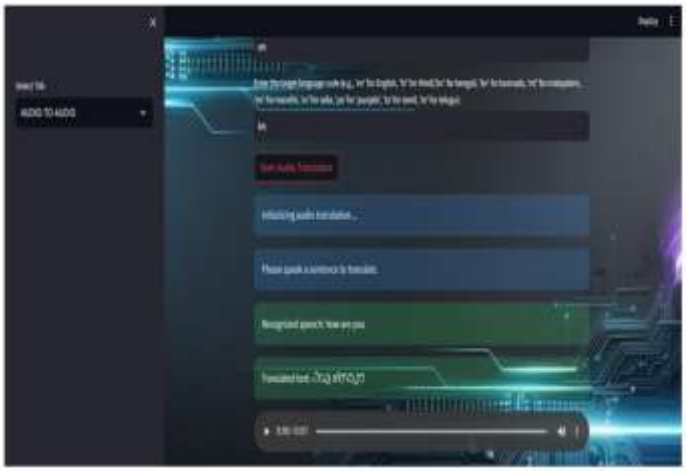

Fig 4.1

VIII. CONCLUSION

One effective technique for overcoming language barriers and advancing accessibility, diversity, and intercultural communication is the Multilingual Translator Project. Through the integration of text, image, document, and voice translation features into a single platform, the project offers a flexible and intuitive way to address a range of linguistic requirements. It is a useful tool for both individuals and businesses because of its uses in a wide range of fields, such as governance, commerce, healthcare, education, and travel.

The project not only helps society but also promotes environmental sustainability by lowering the need for physical travel, reducing paper waste, and promoting digital transformation. All things considered, the Multilingual Translator Project serves as an example of how cutting-edge technology may promote constructive change and create a more sustainable and interconnected society.

IX. FUTURE SCOPE

Enabling offline functionality and increasing language support, notably for indigenous languages, will increase accessibility. Advanced natural language processing methods will improve context-aware translations, and real-time multimodal translation will improve live meetings. Speech synthesis will be improved with the addition of customizable voice styles. Accuracy and speed will be optimized by NMT and other machine learning models. Multi-user collaboration tools will facilitate real-time group translations, and integration with applications such as Zoom and WhatsApp will ensure seamless accessibility. These improvements will make the translator more efficient, flexible and easy to use.

X. REFERENCES

[1] Image to Multilingual Text Conversion for Literacy Education
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8614240&isnumber=8614025

[2] Multilingual Speech and Text Recognition and Translation Using Image
http://dx.doi.org/10.17577/IJERTV5IS040053

[3] A Survey of Deep Learning Approaches for OCR and Document Understanding
https://doi.org/10.48550/arXiv.2011.13534

[4] Transcribing and Translating Bilingual Text Using OCR Tesseract and Deep Learning
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10540714&isnumber=10540606

[5] Real-Time Text & Speech Translation Using Sequence-to-Sequence Approach
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9544509&isnumber=9544092

[6] Speech to Text Translation Enabling Multilingualism
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9298280&isnumber=9298025

[7] An Integrated Model for Text-to-Text, Image-to-Text, and Audio-to-Text Linguistic Conversion Using Machine Learning Approach-
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10112123&isnumber=10111935

[8] Deep Learning for Scene Text Detection and Recognition's
https://doi.org/10.48550/arXiv.1811.04256

[9] OCR and Speech Recognition System Using Machine Learning
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9697030&isnumber=9696444