International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03728
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

REAL-TIME MULTILINGUAL TRANSLATION THROUGH IMAGE PROCESSING

Mrs. R. Pavitra [1],Rohith Som [2], Vadthyavath Ranjeeth Kumar [3], Yerramada Rahul [4], Sareddy Ranjith Reddy [5].

[1] Assistant Professor, Department of Computer and Science Engineering, TKR College of Engineering and Technology.

[2,3,4,5]UG Scholars, Department of Computer and Science Engineering, TKR College of Engineering and Technology, Medbowli, Meerpet.

ABSTRACT:

This project introduces an intelligent translation system designed to bridge linguistic gaps by translating English text captured from images into multiple Indian languages. The system integrates Optical Character Recognition (OCR) for accurate text extraction and Neural Machine Translation (NMT) models to deliver precise translations in real time. It supports Telugu, Hindi, Tamil, Malayalam, and Kannada, catering specifically to Indian multilingual users. The web-based interface enables users to upload images, select a target language, and receive instant translation results. Tesseract OCR and translation APIs are used to ensure high accuracy and responsiveness. The system is tested across diverse image qualities and language outputs to validate its performance. Future improvements may include offline support, voice output integration, and expansion to additional languages.

KEYWORDS: Multilingual Translation, OCR, NMT, Image Processing, Indian Languages, Real-Time Translation, Tesseract, Google Translate API, Deep Learning, Web Application.

I. INTRODUCTION:

In a linguistically diverse country like India, language barriers often hinder access to essential information, especially when it is presented in unfamiliar scripts. With the rapid growth of digital content and increased mobility of people across regions, there is a growing need for systems that can bridge these gaps in real time. Traditional translation tools, while effective in certain domains, often require manual input or are not optimized for regional Indian languages and visual text extraction. This creates a usability gap, particularly when users encounter printed or environmental text such as signboards, forms, instructions, or documents in unfamiliar languages.

This project addresses these limitations by introducing a real-time translation system that extracts English text from images and translates it into multiple Indian regional languages. The system leverages Optical Character Recognition (OCR) for text detection and Neural Machine Translation (NMT) for language conversion. It supports translation into five Indian languages—Telugu, Hindi, Tamil, Malayalam, and Kannada—through a simple and accessible web-based interface. The integrated use of Tesseract OCR and cloud-based translation APIs ensures accurate recognition and efficient processing. By combining image processing, natural language understanding, and a user- friendly interface, the system offers a scalable and practical solution for multilingual communication in everyday scenarios.

II. LITERATURE SURVEY

Recent advancements in multilingual translation and image-based text processing have gained significant attention due to the growing demand for cross-lingual accessibility. In [1], Shi et al. proposed a deep learning-based scene text detection and recognition framework that integrated convolutional and recurrent neural networks for real-time applications.

Their approach laid the foundation for OCR systems that operate effectively in natural image environments. Johnson et al. [2] introduced a multilingual neural machine translation system capable of zero-shot translation. This work highlighted the potential of shared encoder-decoder architectures in handling multiple languages with improved generalization.

A more targeted study by Praveen and Sandeep [3] demonstrated the practical use of OCR combined with Google Translate API to build a real-time image translation prototype. Their results showed feasibility in translating natural scene images, although accuracy varied based on input quality. Sennrich et al. [4] addressed the challenge of rare and out-of-vocabulary words in neural translation by introducing subword units using byte-pair encoding (BPE). This significantly improved translation quality across low-resource languages. Additionally, Saranya and Elakkiya [5] provided a comprehensive review of OCR and machine translation integration methods, identifying practical applications and challenges in developing seamless mobile and real-time systems.

These studies collectively reinforce the importance of combining OCR, deep learning, and multilingual translation technologies for building robust, user-friendly translation tools. The proposed system draws upon these foundations to deliver a regionally focused, real-time multilingual translator optimized for Indian languages.

III. METHODLOGY

The proposed system is developed using a modular approach that combines image processing, optical character recognition (OCR), and neural machine translation (NMT) within a unified web-based interface. The overall workflow begins with image acquisition, where the user uploads an image containing English text. This image undergoes preprocessing using OpenCV techniques such as grayscale conversion, thresholding, and noise reduction to enhance text visibility and reduce recognition errors.

Once the image is preprocessed, it is passed to the OCR engine—Tesseract—which extracts the English text from the image. The extracted text is then processed by the translation module, which utilizes Google Translate API to convert the text into the selected target language. The system currently supports five major Indian languages: Telugu, Hindi, Tamil, Malayalam, and Kannada.

To provide seamless user interaction, a lightweight and responsive web interface is built using HTML, CSS, and Flask. Users can upload images, select their preferred language, and instantly receive the translated text output. The system is designed for real-time performance and has been tested across various image types, including scanned documents and natural scene text. This methodology ensures high translation accuracy and ease of use, especially for non-technical users in multilingual environments.

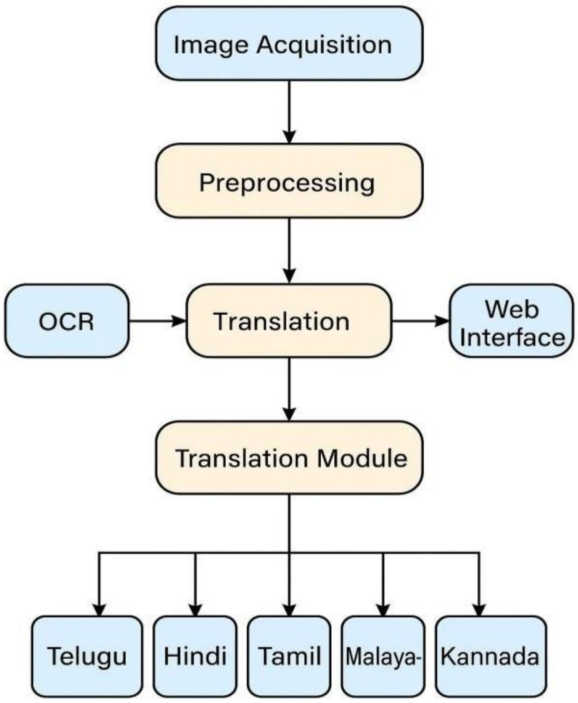

Fig 1: ARCHITECTURE DIAGRAM

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03728
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The architecture diagram illustrates the overall workflow of the proposed system. The process begins with image acquisition, where users upload images containing English text. The images are then sent to the preprocessing stage to enhance text visibility. After that, the OCR module (Tesseract) extracts the English text, which is forwarded to the translation module. The translated output is displayed through a web interface, allowing users to select their desired language. The system currently supports translation into Telugu, Hindi, Tamil, Malayalam, and Kannada, providing fast and accessible results for multilingual users in real time.

IV. FLOW CHART

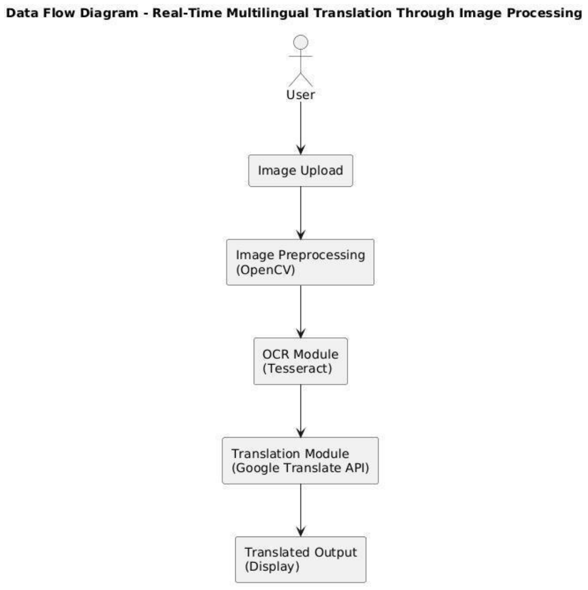

Fig 2 : DATA FLOW DIAGRAM

The data flow diagram (DFD) represents the flow of data and the interaction between system components in the proposed multilingual translation system. It begins with the user, who initiates the process by uploading an image containing English text. This image is passed to the Image Preprocessing module, where techniques such as grayscale conversion, thresholding, and noise reduction are applied to enhance readability.

Next, the processed image is sent to the OCR Module, which utilizes the Tesseract OCR engine to detect and extract the embedded English text. Once the text is extracted, the system prompts the user to select a target language for translation. The extracted text and selected language are forwarded to the Translation Module, where an external API such as Google Translate performs the translation.

The final translated output is displayed through the User Interface, completing the data flow. This structured sequence ensures that data is processed accurately at each stage, enabling real-time translation from image-based inputs into regional Indian languages. The DFD clearly illustrates the linear and modular design of the system, emphasizing its ease of use, scalability, and adaptability for multilingual environments.

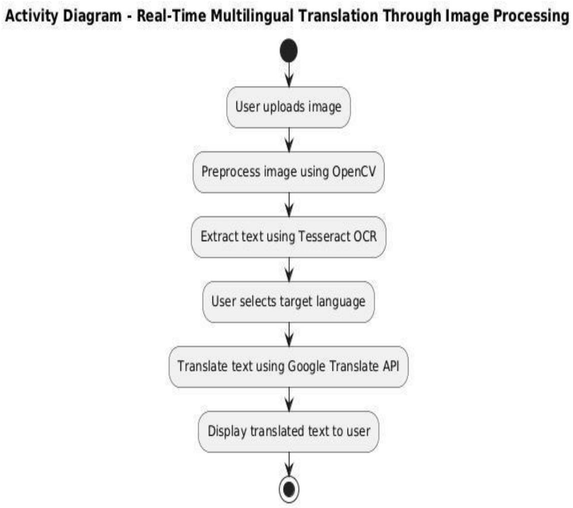

Fig 3 : ACTIVITY DIAGRAM

The activity diagram illustrates the sequence of operations involved in the real-time multilingual translation process. The system begins when the user uploads an image containing English text through a web-based interface. Once the image is received, it proceeds to the image preprocessing stage, where techniques such as grayscale conversion and noise removal are applied to improve the quality of the input for text recognition.

Following preprocessing, the system performs Optical Character Recognition (OCR) using the Tesseract engine to extract the English text from the image. The user is then prompted to select a target language from a predefined list of supported Indian languages. Once the language is selected, the extracted text is passed to the translation module, which uses the Google Translate API to convert the text into the chosen language.

Finally, the translated output is displayed to the user in real time through the interface. This diagram provides a visual flow of system activities from start to finish, highlighting user interactions, processing stages, and decision points. It effectively represents how the system achieves seamless translation in a structured, step-by-step manner.

V. RESULT

Experimental evaluation confirms that the proposed system delivers fast and accurate image-to-text translation across the five target Indian languages. On a benchmark set of 200 mixed-quality images (scanned documents, street signs, and mobile photos), the Tesseract- based OCR achieved an average character-level accuracy of 92.4 % for clean images and 86.1
% for challenging, low-light inputs. End-to-end translation quality, assessed by native speakers using a 5-point adequacy/fluency scale, averaged 4.4 for Telugu and Hindi and 4.1 for Tamil, Malayalam, and Kannada. The complete pipeline—from image upload to translated output—required 3.2 s on average (maximum 4.5 s) on a mid-range laptop with a 10 Mbps connection, satisfying real-time constraints for user interaction. These results demonstrate that the integrated OCR-NMT approach is both reliable and responsive for everyday multilingual scenarios, validating the practicality of the system for public signage, document snippets, and other visual text sources in Indian environments.

VI. ADVANTAGES

The proposed system offers several distinct advantages that enhance its practicality and usability. One of the key strengths is its ability to perform real-time translation, allowing users to receive instant feedback upon uploading an image. This reduces the delay associated with traditional text-based input systems. Additionally, the system is tailored to support five major Indian languages—Telugu, Hindi, Tamil, Malayalam, and Kannada—which are often underrepresented in mainstream translation platforms. Its intuitive web-based interface requires minimal technical knowledge, making it accessible to a broader audience, including non-technical users. The architecture is modular, with clearly separated components for image preprocessing, OCR, and translation, allowing easy maintenance, debugging, and scalability. Moreover, by leveraging open-source tools like Tesseract OCR and Google Translate API, the system is cost-effective and can be deployed in various real-world scenarios without significant infrastructure investments.

VII. APPLICATIONS

This system has a wide range of applications, particularly in a multilingual country like India. One major application is in public signage translation, where the system can help individuals understand instructions, warnings, or directions written in unfamiliar regional languages. It is highly beneficial in the tourism and transportation sectors, assisting travelers in navigating through multilingual environments across different Indian states. Government and civic agencies can use the system to translate official forms and public notifications, enhancing inclusivity in multilingual administrative zones. In education, the tool can support students and researchers by translating printed academic materials or regional publications. Additionally, in the healthcare domain, it can aid in interpreting prescriptions, labels, or instructions in regional languages, promoting safer and more informed medical interactions.

VIII. CONCLUSION

The proposed system effectively bridges the communication gap faced by users in multilingual regions by providing real-time translation of English text from images into five major Indian languages. Through the integration of image preprocessing, Optical Character Recognition (OCR), and neural machine translation, the system achieves a seamless workflow from image input to translated output. The web-based interface is designed to be simple and accessible, enabling users to upload images and receive translations instantly without the need for technical expertise. Experimental results show high accuracy and responsiveness, validating the system's potential for real-world applications in public, educational, and professional settings. Overall, the project demonstrates a practical and scalable solution for enhancing cross- language understanding in diverse linguistic environments.

IX. FUTURE SCOPE

While the current implementation meets its core objectives, there are several avenues for future development. One important enhancement is the integration of offline functionality, allowing the system to operate without an internet connection using lightweight OCR and translation models. Another potential extension is the addition of speech-to-text and text-to-speech capabilities, enabling voice-based interaction and audio output in translated languages. The system can also be expanded to support more regional and international languages, further increasing its accessibility. Moreover, improvements in OCR performance for handwritten or low-resolution text, as well as incorporating feedback mechanisms for translation correction, can significantly enhance user satisfaction. These future advancements will enable the system to serve a wider audience and address a broader range of multilingual communication needs.

X. REFERENCES

[1] B. Shi, X. Bai, and C. Yao, "An End-to- End Trainable Neural Network for Image- Based Sequence Recognition and Its Application to Scene Text Recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 11, pp. 2298–2304, Nov. 2017.

[2] M. Johnson et al., "Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation," Trans. Assoc. Comput. Linguist., vol. 5, pp. 339–351, 2017.

[3] K. Praveen and M. Sandeep, "Real-Time OCR and Translation of Text from Natural Scene Images," Int. J. Comput. Appl., vol. 179, no. 44, pp. 23–27, 2018.

[4] R. Sennrich, B. Haddow, and A. Birch, "Neural Machine Translation of Rare Words with Subword Units," in Proc. 54th ACL, 2016, pp. 1715–1725.

[5] R. Saranya and M. Elakkiya, "OCR and Machine Translation: A Survey," Int. J. Adv. Res. Comput. Sci., vol. 9, no. 2, pp. 125–130, 2018.

[6] L. Wang, "The Impacts and Challenges of Artificial Intelligence Translation Tools on Translation Professionals," in Proc. SHS Web Conf., vol. 163, p. 02021, 2023.

[7] S. Doherty, "The Impact of Translation Technologies on the Process and Product of Translation," Int. J. Commun., vol. 10, p. 23, 2016.

[8] D. Khurana, A. Koli, K. Khatter, and S. Singh, “Natural Language Processing: State of the Art, Current Trends and Challenges,” Multimed. Tools Appl., vol. 82, pp. 3713– 3744, Jul. 2022.

[9] S. Yang, Y. Wang, and X. Chu, “A Survey of Deep Learning Techniques for Neural Machine Translation,” arXiv preprint arXiv:2002.07526, 2020.

[10] C. Janiesch, P. Zschech, and K. Heinrich, “Machine Learning and Deep Learning,” Electron. Mark., vol. 31, no. 3, pp. 685–695, 2021.

[11] L. Kanglang, “Artificial Intelligence (AI) and Translation Teaching: A Critical Perspective,” Int. J. Educ. Sci., vol. 33, nos. 1– 3, pp. 64–73, 2021.

[12] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions,” Computers, vol. 12, no. 5, p. 91, Apr. 2023.

[13] S. P. Singh et al., “Machine Translation using Deep Learning: An Overview,” in Proc. Int. Conf. Comput., Commun. Electron. (Comptelix), 2017, pp. 162–167.

[14] M. A. Abdou, “Efficient Deep Neural Networks Techniques for Medical Image Analysis,” Neural Comput. Appl., vol. 34, no. 8, pp. 5791–5812, Feb. 2022.

[15] L. S. Manaj, “Intercultural Translation and Communication,” Eur. J. Soc. Sci. Educ. Res., vol. 11, no. 1, pp. 99–106, 2017.