Grenze International Journal of Engineering and Technology, June Issue

Automated Text Recognition and Translation: An Integrated Web Application Approach

Saggurthi Ramesh[1], Kurakula Yagna Sri[2], Shaik Mubashir Ahamad[3] and Maria Vara Pushpak Nelaturi[4]
[1]Assistant Professor, Department of Artificial Intelligence and Data Science, Lakireddy Balireddy College of Engineering (Autonomous), Mylavaram, Andhra Pradesh, India
Email: rameshsaggurthi1245@gmail.com
[2–4]Students, Department of Artificial Intelligence and Data Science, Lakireddy Balireddy College of Engineering (Autonomous), Mylavaram, Andhra Pradesh, India
Email: yagnasri2828@gmail.com, shaikmubashirahamad@gmail.com, pushpaknelaturi@gmail.com

***Abstract*— The paper proposes the conceptualization and development of a web-based application that uses Optical Character Recognition (OCR) and machine translation to achieve multilingual text extraction and translation. The system allows users to upload images containing text, automatically extract the text using OCR, and translate it into the desired language in real time. By leveraging modern web technologies and cloud-based translation APIs, the platform provides a unified and accessible solution for overcoming language barriers in digital documents. Experimental results demonstrate that the system effectively extracts and translates text across various image formats and languages with high accuracy. The modular design enables easy integration of additional languages and OCR engines, making the platform scalable and adaptable. This work highlights the potential of OCR and translation technologies in improving accessibility, facilitating cross-lingual communication, and enhancing information processing in a digitally connected world.**

***Index Terms*— Optical Character Recognition (OCR), Machine Translation, Multilingual Text Extraction, Digital Accessibility, Cross-Lingual Communication, Image-to-Text, Automated Translation, Information Processing.**

I. INTRODUCTION

The rapid growth of digital content has created a strong need for tools that can extract and understand information across multiple languages. However, much of this content exists in images and scanned documents, making direct text processing difficult. Optical Character Recognition (OCR) and machine translation address this challenge by converting visual text into machine-readable and translatable form.

Despite these advancements, existing solutions often work as separate tools, leading to inefficient workflows. This work proposes an integrated web-based system that combines OCR, translation, and summarization into a single platform. It allows users to upload images or PDFs, automatically extract text, translate it, and generate summaries.

By using advanced preprocessing, multi-engine OCR, and neural machine translation, the system improves accuracy and usability. It enhances accessibility, reduces manual effort, and supports effective cross-lingual communication for both individual and organizational use.

Grenze ID: 01.GIJET.12.2.516

II. LITERATURE SURVEY

The previous work on automated text recognition and translation in web-based settings can be put in perspective with the wider discoveries of online service, e-commerce and digital interaction studies. Online-systems perceived security and risk have been found to have a strong impact on adoption and further usage and this directly applies when users place sensitive documents in an OCR translation system [1]. Gender and age are demographic factors that influence user reactions to persuasive and assistive elements in online interfaces [2], and online feedback and review systems impact intention and trust on digital services [3]. The Theory of Planned Behavior explains how attitude, subjective norm, and perceived behavioral control affect user adoption of automated systems [4]. Trust formation in e-commerce markets highlights the influence of cultural background and perceived platform trustworthiness [5]. Variations in perception of online reviews across user groups emphasize the need for flexible explanations and error reporting in OCR translation systems [6]. Web-based quality factors such as performance, navigation, and visual design significantly influence user satisfaction and can be applied to OCR translation interfaces [7].

Perceived value, trust, and enhanced interaction design also play a crucial role in OCR translation system adoption. Behavioral intention is strongly influenced by perceived trust and value, emphasizing the importance of communi- cating accuracy, privacy, and efficiency benefits [8]. Advanced visualization technologies such as augmented reality improve user interaction and can inspire better presentation of OCR outputs [9]. Demographic differences influence perceptions of convenience, control, and personalization, requiring adaptable system design [10]. Cross-cultural studies show that presentation format and contextual information affect perceived usefulness [11], while cultural differences in credibility perception highlight the importance of transparency and structure [12]. Consistent interface design improves long-term engagement and satisfaction [13]. Statistical modeling techniques provide reliable methods for evaluating system performance and user satisfaction [14, 15].

The credibility of online information significantly affects user trust in automated outputs. Reliable sources enhance trust and influence behavior, similar to how confidence indicators can improve trust in OCR and translation systems [16]. Perceived value influences continued usage of online platforms, highlighting the importance of speed, accuracy, and usability [17]. Feedback systems play an essential role in improving system performance by allowing users to correct errors [18]. The availability of information and reviews accelerates decision-making, similar to how integrated translation systems enable faster comprehension [19]. Conceptual models of customer value emphasize functional, emotional, and social benefits, which align with improved accessibility and reduced cognitive effort in OCR systems [20]. Social influence and credibility factors further enhance trust in automated systems [21, 22]. The influence of short-form content platforms shows how structured presentation of translated information can impact user decisions [23].

TABLE I: LITERATURE SURVEY TABLE

Title of Paper (Year)	Methodology	Dataset Used	Research Gaps Identified	Proposed Handling Strategy
TexTAR (Kumar, 2025)	Context-aware Trans- former using 2D RoPE	MMTAD	Neglect of textual attributes and spatial lay- out	Incorporate layout-aware embeddings to jointly model text style and spatial structure
OCR-MT-Sum (Acharekar, 2025)	Tesseract OCR + MarianMT + LexRank pipeline	OCR-MT-Sum	Lack of coherence across OCR, MT, and summarization	Use OCR-free multilingual transformers for end-to-end translation and summarization
Cross-border- Logis-MT (Lin, 2025)	Domain adaptation with X-LogisKG multitask model	X-LogisKG	Weak domain-specific multilingual generalization	Apply cross-domain fine-tuning with multilingual contextual embeddings
DocOwl2 (Ye, 2024)	High-resolution OCR free DocCompressor	mPLUG-DocOwl2	High computational load; loss of fine details	Optimize with adaptive token compression and lightweight attention layers
GeoLayoutLM (Luo et al., 2023)	Geometric pre-training with relation modeling	GeoLayoutLM	Limited structural relationship extraction	Employ graph-based spatial reasoning for explicit layout-relational learning
ERPA (Abdellaif et al., 2024)	OCR + LLM integration with dynamic lay- out handling	ERPA Benchmark	Poor performance on unstructured layouts	Build adaptive OCR-LLM pipeline with layout-aware prompt adaptation

Studies on specific user groups highlight the importance of usability, aesthetics, and localization in digital systems [24, 25, 26]. Cross-cultural differences in trust mechanisms indicate that interface design must adapt to regional preferences [27]. Research on online platforms shows that trust in system mechanisms significantly

affects user adoption [28]. Statistical validation methods support evaluation of usability and system effectiveness [29]. Eye- tracking studies demonstrate that layout design plays a critical role in how users process and validate information, emphasizing the importance of intuitive interface design in OCR translation systems [30].

III. Proposed System

The proposed system introduces an end-to-end intelligent web platform designed to automate text extraction, recognition, translation, and summarization from images and PDF documents by combining multi-stage OCR pipelines with neural machine translation and adaptive preprocessing.
The workflow begins with image normalization where the uploaded input I is converted into a luminance-based grayscale G using the weighted mapping:

$$G = 0.299R + 0.587G + 0.114B \qquad (1)$$

This converts RGB images into grayscale for consistent intensity representation. To further enhance clarity, Gaussian smoothing is applied:

$$G'(x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} G(x+i, y+j) \cdot h(i,j) \qquad (2)$$

This removes noise while preserving important text structures. The image is then binarized using Otsu's method:

$$T = \arg\max_{\tau} \sigma_b^2(\tau) \qquad (3)$$

This selects the optimal threshold to separate text from background. Text regions are extracted using connected component analysis:

$$C_i = \{(x,y) \mid L(x,y) = i\} \qquad (4)$$

This identifies individual text components in the image. Feature extraction is performed using HOG:

$$H(\theta) = \sum_{(x,y)\in R} m(x,y) \cdot \delta(\theta(x,y) - \theta) \qquad (5)$$

This captures edge and shape features for accurate recognition.
The OCR output T_{rec} is passed into a translation model. The encoder computes hidden states as:

$$h_t = f(h_{t-1}, x_t) \qquad (6)$$

This encodes input sequences into contextual representations. The decoder generates translated tokens:

$$y_t = g(s_{t-1}, c_t) \qquad (7)$$

This produces output text step-by-step.
An attention mechanism improves translation accuracy:

$$\alpha'_{t,i} \quad \frac{\exp(e_{t,i})}{\sum_{j=1}^{n} \exp(e_{t,j})} \qquad (8)$$

This assigns importance to relevant words during translation. The final translated output is constructed as:

$$S_{trans} = \text{Concat}(y_1, y_2, \ldots, y_m) \qquad (9)$$

This combines all generated tokens into the final sentence. Translation quality is evaluated using BLEU score:

$$BLEU = \exp\left(\min\left(1 - \frac{r}{c}, 0\right) + \frac{1}{N}\sum_{n=1}^{N} \log p_n\right) \qquad (10)$$

This measures similarity between generated and reference translations.

IV. Architecture Overview

The system architecture illustrates the integration of preprocessing, OCR, translation, PDF handling, and summa- rization modules within a Flask-based backend. The application provides an interactive user interface and supports downloadable outputs in PDF format.

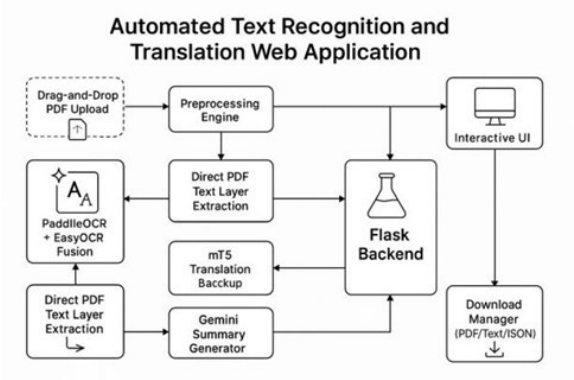

Figure 1: Model Architecture

V. Results and Discussion

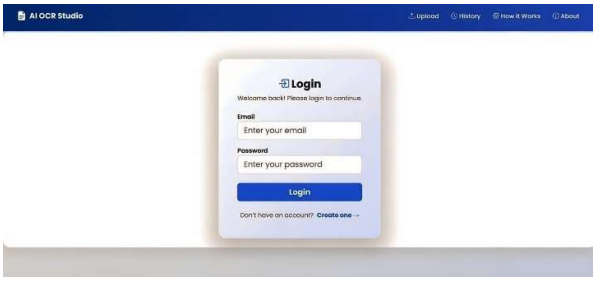

Figure 2: AI OCR Studio Login Interface

The login page provides a simple and focused interface for user authentication. It contains centrally aligned email and password input fields along with a clearly visible blue login button.

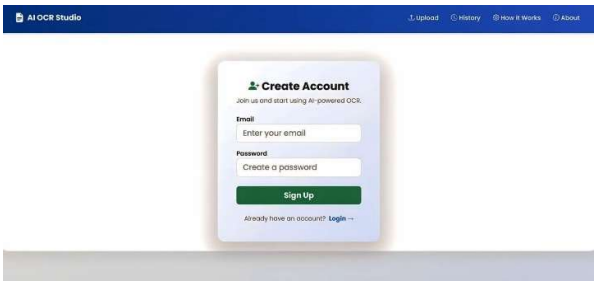

Figure 3: AI OCR Studio Create Account Interface

There is a very simple sign up form on the page which aims to register new users with only basic information. A green “Sign Up” button is used to create a distinct call-to-action.

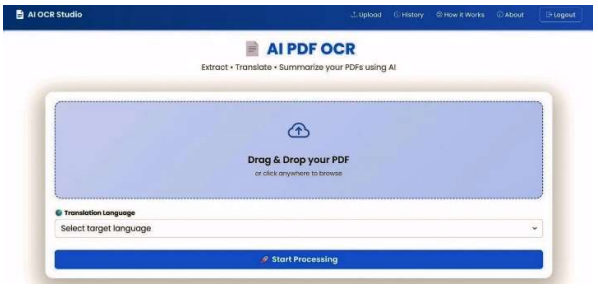

Figure 4: AI PDF OCR Processing Interface

The home dashboard is the primary point of interaction in the application. A drag-and-drop PDF upload feature allows users to easily upload documents for processing.

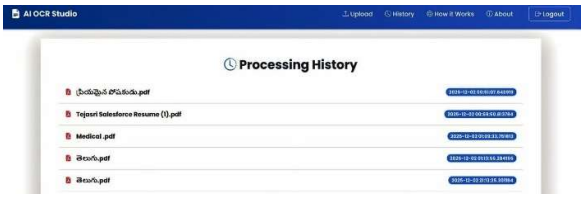

Figure 5: AI OCR Studio Processing History Interface

The history page acts as a repository for previously processed documents. Results are displayed in a structured table along with timestamps.

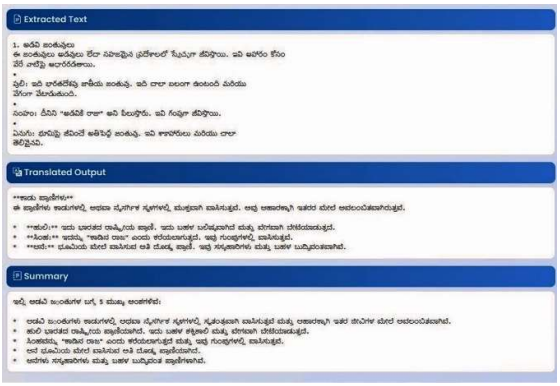

Figure 6: Extracted Text, Translated Output, and Summary Display Interface

The processing page displays the results obtained after document analysis. Extracted text, translated content, and summarized output are shown in separate sections.

VI. Conclusion

The proposed system integrates OCR, neural machine translation, and summarization into a unified web-based platform for processing multilingual text from images and PDF documents. Through effective preprocessing and multi-engine OCR, the system ensures reliable text extraction, while neural translation improves contextual accuracy and fluency. The inclusion of summarization further enhances usability by enabling quick understanding of large documents.

Its modular architecture and real-time processing make the system efficient, scalable, and easy to use for both individual and organizational needs. By combining multiple functionalities into a single platform, it reduces dependency on separate tools and streamlines workflows. Overall, the system improves accessibility, minimizes manual effort, and provides a practical solution for cross-lingual information processing in real-world scenarios.

Future Scope

Future work includes expanding language support, improving recognition of handwritten and low-quality text, and enabling scalable processing through distributed computing. Enhancing security with stronger encryption and access control is also important. Integration with mobile platforms and APIs will further improve accessibility and usability, while continuous development will ensure better performance and reliability.

References

[1] I. Aboobucker, “Consumers’ perceived security risks in online shopping: A survey study in Sri Lanka,” SSRN Electronic Journal, 2019.

[2] I. Adaji, K. Oyibo, and J. Vassileva, “The effect of gender and age on the factors that influence healthy shopping habits in e-commerce,” in Proc. 26th Conference on User Modeling, Adaptation and Personalization, pp. 251–255, 2018.

[3] A. R. Aditya and Y. Alversia, “The influence of online review on consumers’ purchase intention,” GATR Journal of Management and Marketing Review, vol. 4, no. 3, pp. 194–201, 2019.
[4] I. Ajzen, “The theory of planned behavior,” Organizational Behavior and Human Decision Processes, vol. 50, no. 2, pp. 179–211, 1991.
[5] N. T. M. Anh and P. T. T. Huong, “Exploring the Vietnamese consumer’s trust in e-commerce,” VNU Journal of Science: Economics and Business, vol. 35, no. 5E, 2019.
[6] S. Bae and T. Lee, “Gender differences in consumers’ perception of online consumer reviews,” Electronic Commerce Research, vol. 11, no. 2, pp. 201–214, 2011.
[7] B. Bai, R. Law, and I. Wen, “The impact of website quality on customer satisfaction and purchase intentions,” International Journal of Hospitality Management, vol. 27, no. 3, pp. 391–402, 2008.
[8] S. D. Baidoun and M. Z. Salem, “The moderating role of perceived trust and perceived value on online shopping behavioral intention,” Competitiveness Review, vol. 34, no. 1, pp. 125–143, 2024.
[9] F. Baytar, T. Chung, and E. Shin, “Trying on clothes in augmented reality when online shopping,” Journal of Fashion Marketing and Management, vol. 24, no. 4, pp. 667–683, 2020.
[10] S. A. Bhat and M. A. Darzi, “Exploring the influence of consumer demographics on online purchase benefits,” FIIB Business Review, vol. 8, no. 4, pp. 303–316, 2019.
[11] B. Biswas, P. Sengupta, and B. Ganguly, “Exploring the determinants of perceived helpfulness of online reviews,” Electronic Markets, vol. 32, no. 3, pp. 1083–1102, 2022.
[12] B. M. Brand and R. Reith, “Cultural differences in the perception of credible online reviews,” Decision Support Systems, vol. 154, Article 113710, 2022.
[13] H. T. L. Bui, “Fast fashion store image, customer satisfaction and loyalty,” International Journal of Systems and Service-Oriented Engineering, vol. 9, no. 1, pp. 1–19, 2019.
[14] B. M. Byrne, Structural Equation Modeling with AMOS, 3rd ed., Routledge, 2016.
[15] M. K. Cain, Z. Zhang, and K. H. Yuan, “Univariate and multivariate skewness and kurtosis for measuring nonnormality,” Behavior Research Methods, vol. 49, no. 5, pp. 1716–1735, 2017.
[16] U. Chakraborty and S. Bhat, “The effects of credible online reviews on brand equity,” Journal of Promotion Management, vol. 24, no. 1, pp. 57–82, 2018.
[17] H. H. Chang and H. Wang, “The moderating effect of customer perceived value on online shopping behaviour,” Online Information Review, vol. 35, no. 3, pp. 333–359, 2011.
[18] C. Changchit and T. Klaus, “Determinants and impact of online reviews on product satisfaction,” Journal of Internet Commerce, vol. 19, no. 1, pp. 82–102, 2020.
[19] C. D. Chen and E. C. S. Ku, “Diversified online review websites and impulsive buying,” Journal of Internet Commerce, vol. 20, no. 1, pp. 113–135, 2021.
[20] Z. Chen and A. J. Dubinsky, “A conceptual model of perceived customer value in e-commerce,” Psychology & Marketing, vol. 20, no. 4, pp. 323–347, 2003.
[21] Y. H. Cheng and H. Y. Ho, “Social influence’s impact on reader perceptions,” Journal of Business Research, vol. 68, no. 4, pp. 883–887, 2015.
[22] G. Craciun and K. Moore, “Credibility of negative online product reviews,” Computers in Human Behavior, vol. 97, pp. 104–115, 2019.
[23] L. X. Cu et al., “Online review and customer purchase: Evidence from TikTok,” VNU Journal of Economics and Business, vol. 3, no. 5, 2023.
[24] N. H. Cuong et al., “Factors affecting online fashion purchase decisions,” Khon Kaen University, 2022.
[25] T. Do, T. Nguyen, and C. Nguyen, “Online shopping in an emerging market,” Journal of Economics and Management Sciences, vol. 2, no. 2, 2019.
[26] T. T. T. Doan, “Factors affecting online purchase intention,” Management Science Letters, 2020.
[27] B. Dong and M. Li, “Trust in online reviews: A cross-cultural study,” in Human–Computer Interaction, Springer, 2017.
[28] N. Duy et al., “Online reviews on e-commerce platforms in Vietnam,” International Journal of Informatics and Communication Technology, vol. 14, no. 1, pp. 195–206, 2025.
[29] C. D. Dziuban and E. C. Shirkey, “When is a correlation matrix appropriate for factor analysis?,” Psychological Bulletin, vol. 81, no. 6, pp. 358–361, 1974.
[30] T. Chen et al., “The impact of online reviews on purchasing decisions,” Frontiers in Psychology, vol. 13, Article 865702, 2022.