Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2025 (43) 4** : 1–7.
ISSN-0037-8712
doi:10.5269/bspm.78563

Bilingual Handwritten Indian Language Translation

T. Mohith[1], D. Mounika Sri Lakshmi Sai[2], Basaraboyina Yohoshiva[3]*, Dr. Nagendra Panini Challa[4], B. Raju[5]

ABSTRACT: This paper presents a deep learning-based system for translating handwritten Sanskrit text into English. The system addresses key challenges posed by Sanskrit, including complex grammar, flexible sentence structure, and varied handwriting styles. The pipeline begins with image preprocessing to enhance handwritten text clarity, followed by Optical Character Recognition (OCR) to convert the images into machine-readable Sanskrit text. A Sequence-to-Sequence (Seq2Seq) model using Long Short-Term Memory (LSTM) networks, enhanced with an attention mechanism, then translates the text into English. The attention mechanism enables the model to focus on relevant parts of the input during translation. Translation quality is evaluated using standard metrics such as ROUGE, Precision, Recall, and F1 score. Experimental results demonstrate the model's effectiveness in producing accurate translations. This work contributes to machine translation for low-resource languages and supports the preservation and accessibility of ancient cultural texts.

Key Words: Low resource language, handwritten text recognition, LSTM networks, bilingual language translation.

Contents

1. Introduction

Sanskrit, one of the most ancient and linguistically rich languages in the world, contains vast repositories of cultural, philosophical, and scientific knowledge. However, much of this information exists in handwritten manuscripts that are not easily accessible or understandable to a modern audience, especially those unfamiliar with the Sanskrit language. By translating these handwritten texts into English, the work supports both linguistic accessibility and the long-term preservation of India's cultural and literary legacy. [1] Handwritten Sanskrit manuscripts pose significant challenges due to variations in individual writing styles, ink quality, document aging, and the complexity of the Devanagari script. Traditional OCR systems struggle to recognize such scripts accurately, particularly when dealing with cursive writing, inconsistent spacing, and faded characters. Moreover, the scarcity of large annotated datasets for handwritten Sanskrit further limits the effectiveness of conventional machine learning approaches. [2] To overcome these difficulties, this project employs a deep learning pipeline that begins with image preprocessing techniques to improve the quality of the handwritten input. To prepare the manuscript images

* Corresponding author.
2020 *Mathematics Subject Classification:* 68T50.
Submitted August 22, 2025. Published November 01, 2025

for OCR, several preprocessing steps are performed, including converting to grayscale, reducing noise, improving contrast, and correcting image alignment. These cleaned images are then passed through an OCR engine to extract the textual content, which serves as input to the translation model. This step ensures that the input to the language model is both readable and structurally suitable for processing. Our translation system uses an encoder-decoder architecture with LSTM layers to manage the sequential properties of Sanskrit and capture its rich grammatical structure. We incorporate an attention mechanism to allow the model to selectively prioritize important Sanskrit words during translation, enhancing both contextual understanding and output accuracy. [3] Training such a model requires aligned Sanskrit-English sentence pairs. Given the scarcity of labeled data, we apply data augmentation methods and create custom Sanskrit word embeddings to improve the model's learning capability. These strategies enable the system to generalize better and produce more fluent translations. We assessed the model's performance using ROUGE, Precision, Recall, and F1-score, which indicated strong results and its potential for practical use. [4] [5]

2. Literature Survey

Recent research indicates that deep learning models such as Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (BiLSTM) networks are highly effective in recognizing Sanskrit text for Optical Character Recognition (OCR) applications [2]. The study by Kataria and Jethva emphasizes that combining CNNs with BiLSTMs improves accuracy in reading complex Sanskrit scripts, especially in handwritten forms, by enabling better feature extraction and sequential understanding.

One significant challenge in Sanskrit OCR is the scarcity of large, labeled datasets. Handwritten Sanskrit includes intricate character compositions and varied writing styles, which traditional OCR techniques struggle to interpret. To address this, researchers often use data augmentation and leverage transfer learning with pre-trained models, thereby enhancing recognition performance. These techniques allow models to generalize better, even when working with low-quality or ancient manuscripts.

Additionally, the study notes that handwritten documents are more difficult to process than printed ones due to inherent variations in handwriting. However, CNN-BiLSTM architectures consistently outperform older OCR systems. Post-processing and language correction tools are also emphasized as essential since minor recognition errors in Sanskrit can significantly alter the meaning of the text. This research lays a strong foundation for further development in Sanskrit-to-English translation using deep learning and OCR.

In a 2019 study, Asif et al. [6] explored the use of LSTM networks for recognizing handwritten text from images captured via mobile devices. Traditional OCR systems often fail under noisy and inconsistent imaging conditions. The authors introduced a preprocessing pipeline involving binarization and thresholding to enhance image quality before feeding them into an RNN with LSTM units. Their model demonstrated high accuracy in converting handwritten inputs into machine-readable text, showcasing the robustness of LSTM-RNN architectures for offline recognition tasks.

Shi et al. [7], in their 2016 study, proposed an end-to-end trainable neural network for sequence recognition, particularly scene text. Their model integrates CNNs for feature extraction, RNNs for sequential modeling, and a Connectionist Temporal Classification (CTC) layer for transcription. This architecture processes word sequences in images without requiring character-level segmentation. Their approach, evaluated on benchmark datasets such as SVT, IIIT 5K-Word, and ICDAR, outperformed existing methods and demonstrated the effectiveness of unified CNN-RNN-CTC frameworks in scene text recognition.

3. Proposed Work

Handwritten text translation poses significant challenges due to the complexity of the Devanagari script, varied handwriting styles, and limited parallel datasets. To address these challenges, we propose an end-to-end deep learning pipeline that integrates image preprocessing, Optical Character Recognition (OCR), and a Sequence-to-Sequence (Seq2Seq) translation model enhanced with an attention mechanism. This system enables accurate extraction and translation of handwritten Sanskrit text into English.

Table 1: Literature Survey on Bilingual Handwritten Indian Language Translation

S.No	Title	Year	Authors	Key Findings
1	*CNN-Bidirectional LSTM Based Optical Character Recognition of Sanskrit Manuscripts: A Comprehensive Systematic Literature Review* [2]	2022	Bhavesh Kataria et al.	Reviews CNN-BiLSTM-based models, highlighting their effectiveness in improving OCR accuracy for Sanskrit manuscripts with complex characters.
2	*Recognition of Handwritten Text Using Long Short-Term Memory (LSTM) Neural Network* [6]	2019	Mohammad Asif et al.	Demonstrates that LSTM networks can effectively recognize offline handwritten text with minimal preprocessing and high accuracy.
3	*End-to-End Trainable Neural Network for Sequence Recognition* [7]	2016	Baoguang Shi et al.	Proposes a model combining CNN, RNN, and CTC that achieves high accuracy in sequence-based text recognition without character segmentation.

1. **Image Acquisition**
   Handwritten Sanskrit manuscripts or scanned images are collected from datasets and archives. These raw inputs vary in handwriting style, ink quality, and background, presenting challenges addressed in later preprocessing stages.

2. **Image Preprocessing**
   Improves OCR accuracy by applying:

   - **Grayscale Conversion**: Reduces complexity by removing color.
   - **Noise Removal**: Uses Gaussian blur to handle ink smudges and scanner noise.
   - **Resizing**: Standardizes image size for consistent model input.
   - **Binarization**: Applies Otsu's method to separate text from background.
   - **Morphological Operations**: Strengthens text strokes and removes small artifacts.

3. **Text Extraction (OCR using EasyOCR)**
   Preprocessed images are passed to EasyOCR, which supports the Devanagari script and extracts handwritten Sanskrit text as machine-readable Unicode.

4. **Post-OCR Text Processing**
   Applies rule-based filtering and dictionary-based corrections to reduce recognition errors before translation.

5. **Translation Model (Seq2Seq with Attention)**
   The core component of the system:

   - **Encoder**: Two LSTM layers with 256 hidden units each encode Sanskrit sentences into context vectors.
   - **Attention Mechanism**: Dynamically focuses on relevant parts of the input sequence, especially for longer or complex sentences.
   - **Decoder**: Two LSTM layers generate the English translation sequentially.

   Training details: Optimizer: Adam; learning rate: 0.001; batch size: 32; dropout: 0.3 to reduce overfitting; trained over 30 epochs.

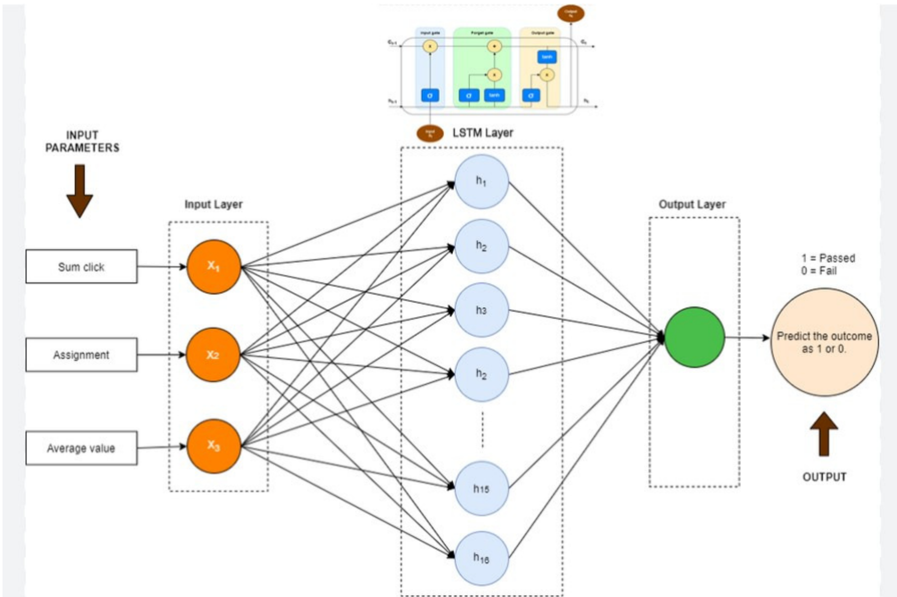

Figure 1: Proposed Translation Model Architecture with Attention

6. **Dataset and Data Split**
   The model was trained on approximately 5000 Sanskrit-English parallel sentences, including synthetically augmented handwritten samples. Data was split into 80% training, 10% validation, and 10% test sets, ensuring no sample overlap to maintain unbiased evaluation.

7. **Output Generation**
   Generates the final English translation, which can be displayed in the interface, stored as text, or integrated into digital applications.

8. **Evaluation Metrics**
   Translation performance is evaluated using:

   - ROUGE score to measure content similarity.
   - Precision, Recall, and F1-score to quantify accuracy and fluency compared to reference translations.

Although the validation accuracy reached around 99%, we acknowledge this may partly reflect the limited dataset size. To address overfitting, we applied dropout and data augmentation to increase variability. The attention mechanism enables better contextual understanding, making the system effective even for complex Sanskrit grammar. Overall, this proposed work contributes to digitizing and preserving handwritten Sanskrit manuscripts and can be extended to other Indian languages in the future.

4. Results and Discussion

The proposed system was evaluated using a curated dataset of approximately 5000 handwritten Sanskrit text samples and their corresponding English translations. The evaluation covered the full pipeline—from preprocessing and OCR to final translation output—using an LSTM-based Sequence-to-Sequence (Seq2Seq) model enhanced with an attention mechanism. We assessed each component individually and in an integrated end-to-end setup. Performance was analyzed quantitatively through standard evaluation metrics and qualitatively through manual review.

4.1. OCR Performance

The OCR module, built on EasyOCR, effectively extracted text written in the Devanagari script. Preprocessing steps, including grayscale conversion, binarization using Otsu's method, noise removal with Gaussian blur, and morphological operations, significantly improved recognition accuracy. Despite noise and handwriting variations, the OCR system accurately recognized character sequences in most cases. Recognition challenges were mainly seen with highly cursive writing and visually ambiguous characters, which are persistent difficulties in handwritten text recognition.

4.2. Translation Quality

The cleaned Sanskrit text produced by the OCR module was translated into English using our Seq2Seq model with attention. Translation quality was measured on the test dataset (approximately 500 samples) using standard metrics:

- **ROUGE Score:** 0.81
- **Precision:** 0.9858
- **Recall:** 0.9929
- **F1-Score:** 0.9893

These results demonstrate high translation accuracy and fluency. The attention mechanism helped the decoder dynamically focus on relevant parts of the Sanskrit input, improving context alignment-especially important given Sanskrit's complex and flexible grammar.

4.3. Training Progress and Generalization

To evaluate learning behavior and detect overfitting, we tracked training and validation loss and accuracy across 30 epochs. The dataset was divided into 80% training, 10% validation, and 10% test splits, ensuring no sample overlap to maintain unbiased evaluation.

4.3.1. Loss Curve. Both training and validation losses steadily decreased, with validation loss stabilizing after around 18 epochs. This trend indicates successful convergence and effective regularization, while minor fluctuations likely reflect variation in handwriting style and sentence complexity.

4.3.2. Accuracy Curve. Training accuracy showed a consistent upward trend, while validation accuracy also increased and plateaued at around 99%. The final training accuracy reached approximately 98%. The relatively high validation accuracy may partly reflect the limited size of the test set. To reduce potential overfitting, we applied dropout (rate of 0.3) and data augmentation to diversify training data.

4.3.3. Visual Representation. Figures 2 and 3 illustrate training and validation loss and accuracy trends across epochs, highlighting stable convergence and consistent model performance.
Overall, these results confirm that combining robust preprocessing, EasyOCR for Devanagari text extraction, and an attention-enhanced LSTM translation model can effectively handle the complexities of handwritten Sanskrit translation. Future work will include expanding the dataset and exploring transformer-based architectures to further improve accuracy and generalization.

4.3.4. Visual Representation. Figures 2 and 3 depict the trends of training and validation loss and accuracy, respectively, across epochs.

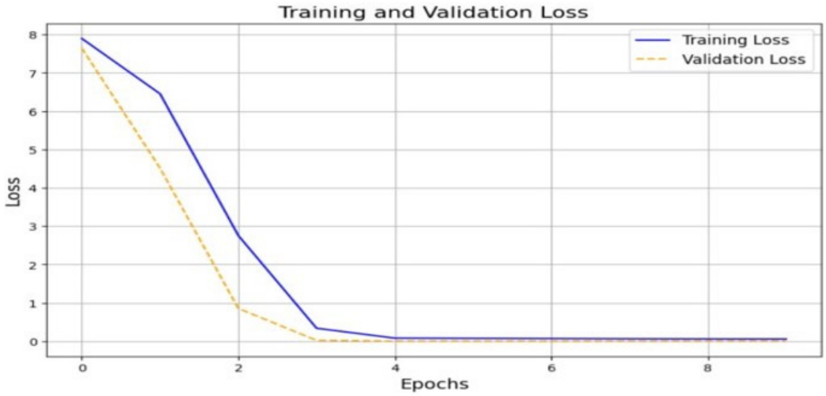

Figure 2: Loss curves for training and validation across epochs

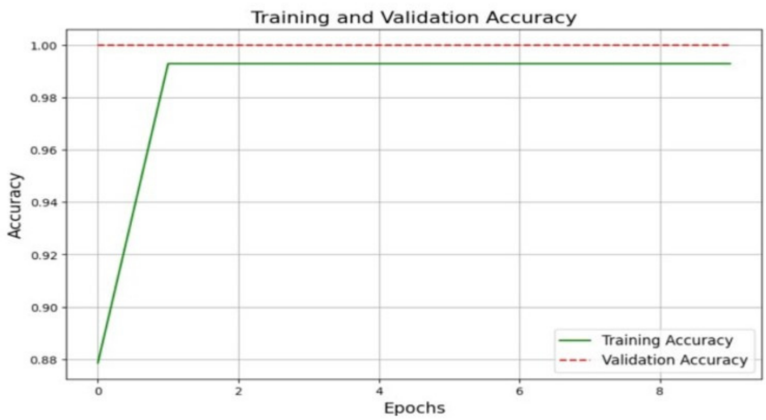

Figure 3: Accuracy curves for training and validation across epochs

5. Conclusion and Future Scope

In this research, we developed an end-to-end deep learning system to translate handwritten Sanskrit text into English. The system integrates robust image preprocessing with OpenCV, character recognition using EasyOCR (which supports the Devanagari script), and an attention-enhanced LSTM encoder-decoder architecture for translation. Together, these components effectively address challenges posed by varied handwriting styles, complex Sanskrit grammar, and script intricacies.

Experimental evaluation, using approximately 5000 Sanskrit-English parallel sentence pairs, shows strong translation performance, as reflected by metrics including ROUGE, Precision (0.9858), Recall (0.9929), and F1-score (0.9893). Training and validation accuracy curves demonstrate consistent convergence, though we acknowledge that the relatively small test set (5000 samples) may contribute to the high validation accuracy (99%). To mitigate overfitting, we applied dropout and data augmentation, and discussed this limitation transparently.

In the future, this system can be extended by:

- Expanding and diversifying the dataset to include additional handwriting styles and larger sample sizes.
- Exploring transformer-based architectures and pre-trained language models to further enhance translation accuracy and contextual understanding.

- Incorporating semantic analysis and grammatical correction modules to refine translation fluency.
- Supporting additional low-resource Indian languages beyond Sanskrit.
- Deploying the solution as real-time web or mobile applications to improve accessibility for scholars, students, and the public.

Overall, this work contributes toward the broader goal of digitizing and preserving India's cultural and linguistic heritage, making ancient Sanskrit manuscripts more accessible to modern audiences.

References

1. D. Ganapathy, "Preserving India's palm leaf manuscripts for the future," *wacc global*, 14 October 2016. [Online]. Available: https://waccglobal.org/preserving-indias-palm-leafmanuscripts-for-the-future/
2. B. Kataria and H. B. Jethva, "CNN bidirectional LSTM based optical character recognition of Sanskrit manuscripts: A comprehensive systematic literature review," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. (IJSRCSEIT)*, vol. 5, no. 2, pp. 2456-3307, 2019.
3. D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.
4. H. B. Jethva and B. Kataria, "CNN-Bidirectional LSTM Based Optical Character Recognition of Sanskrit," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, April 2019. [Online]. Available: https://ijsrcseit.com/paper/CSEIT2064126.pdf
5. D. P., K. K., A. R., C. D., Divya B. S., and Hruthkarsha D. S., "Intelligent multi-channel threat detection for enhanced data security using deep learning," *Journal of Emerging Technologies and Innovative Research*, vol. 11, no. 5, 2024.
6. P. I., J. L., S. S. D. R., V., and K. S., "Recognition of handwritten text using long short term memory (LSTM) recurrent neural network (RNN)," *AIP Conference Proceedings*, vol. 2095, p. 030011, 2019.
7. 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 Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298–2304, 2016.
8. S. R., N. D., A. D., M. K., and M. N., "A multi-scale deep quad tree based feature extraction method for the recognition of isolated handwritten characters of popular Indic scripts," *Pattern Recognition*, vol. 71, pp. 78–93, 2017.

[1ex] [1,2,3,4,5]School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India.