Bol. Soc. Paran. Mat.
(3s.) **v. 22** 2 (2004): 1–6.

Bilingual Handwritten Indian Language Translation

Tirumalasetty Mohith[1], Desamsetti Mounika Sri Lakshmi Sai[2], Basaraboyina Yohoshiva[3], Dr. Nagendra Panini Challa[4], Beera Raju[5]
[1,2,3,4,5]School of Computer Science and Engineering (SCOPE), VIT-AP University, Amaravati, Andhra Pradesh, India

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.

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 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.

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.

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 complexities of script, varying handwriting styles, and limited parallel datasets. In this proposed work, we present a deep learning-based pipeline for translating handwritten Sanskrit text into English, integrating Optical Character Recognition (OCR), image preprocessing, and a Sequence-to-Sequence (Seq2Seq) architecture with attention. OCR is achieved using EasyOCR, which supports the Devanagari script and facilitates accurate character extraction from handwritten inputs. The cleaned Sanskrit text is then passed through an LSTM-based encoder-decoder model with attention mechanisms, which captures contextual and syntactic dependencies for effective translation into English. This paper outlines the complete methodology, including the algorithmic architecture, preprocessing techniques, and visual workflow representations to demonstrate the end-to-end process and model efficiency.

1. **Image Acquisition**
   Handwritten Sanskrit manuscripts or images are collected from datasets or scanned documents. These images form the raw input for the entire system. The quality and variation in handwriting, ink, and paper background are significant challenges addressed in later stages.

2. **Image Preprocessing**
   To enhance recognition accuracy, input images undergo preprocessing steps:
   - **Grayscale Conversion**: Transforms colored images into grayscale, reducing data complexity by discarding unnecessary color details.
   - **Noise Removal**: Utilizes filtering techniques such as Gaussian blur or median filtering to remove visual disturbances like ink smudges or scanner-related noise.
   - **Resizing**: Standardizes image size to match model input requirements.
   - **Binarization**: Converts grayscale images into black and white (binary) using thresholding techniques like Otsu's method.
   - **Morphological Operations**: Further cleans up images to strengthen text strokes and eliminate thin noise.

3. **Text Extraction (OCR using EasyOCR)**
   The preprocessed image is passed to EasyOCR, which performs optical character recognition to extract Sanskrit script. It supports Devanagari script, making it suitable for Sanskrit. This stage detects and recognizes characters and words, producing raw Sanskrit text output.

4. **Post-OCR Text Processing**
   The extracted text may have minor recognition errors. A cleaning module (e.g., rule-based filtering or character correction based on dictionary lookup) refines the OCR output for better translation accuracy.

5. **Translation Model (LSTM Encoder-Decoder with Attention)**
   This is the core of the system:
   - **Encoder**: Takes the cleaned Sanskrit sentence and encodes it into a context vector.
   - **Attention Mechanism**: Enables the model to weigh different parts of the input sequence during decoding, enhancing its ability to manage lengthy inputs and intricate sentence structures.
   - **Decoder**: Responsible for generating the English translation incrementally, using both the context vector and attention weights to guide the process.

6. **Output Generation**
   The final English translation is generated and displayed to the user. It can be stored in a text file, displayed in the interface, or integrated into an application.

7. **Evaluation Metrics**
   Translation quality is evaluated using standard metrics:
   - ROUGE (Recall-Oriented Understudy for Gisting Evaluation)
   - Precision, Recall, F1-Score – These help quantify the accuracy and fluency of translations compared to reference sentences.

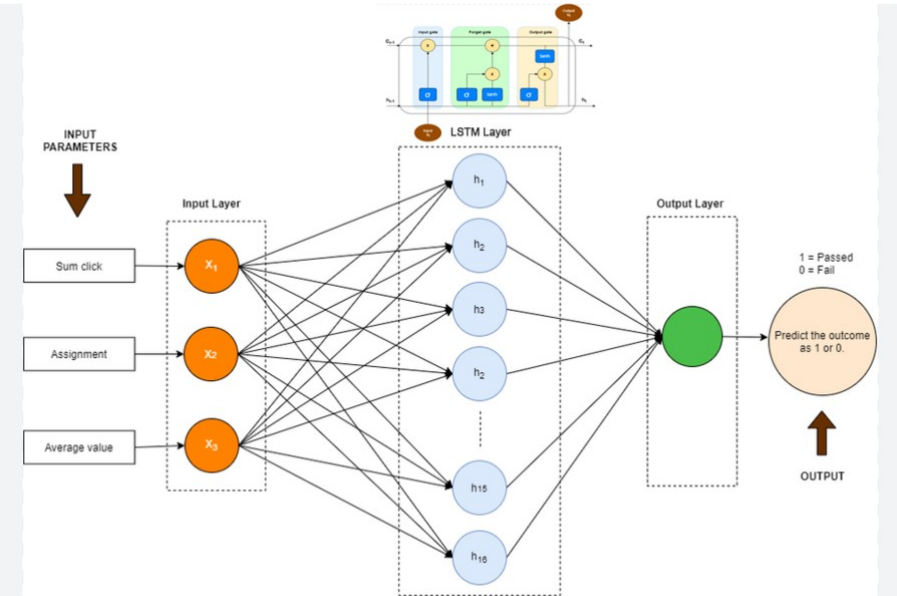

Figure 1: Workflow of Long Short-Term Memory

4. Results and Discussion

The proposed system was evaluated using a carefully curated dataset of handwritten Sanskrit text samples along with their corresponding English translations. The evaluation spanned the entire pipeline—from preprocessing and OCR to final translation output—using an LSTM-based sequence-to-sequence model enhanced with attention mechanisms. Each system component was assessed individually and in an end-to-end integrated setup. The performance was analyzed both quantitatively using standard metrics and qualitatively through manual validation.

4.1. OCR Performance. The EasyOCR-based text extraction module effectively recognized characters from handwritten Sanskrit text in the Devanagari script. Preprocessing techniques, including grayscale conversion, binarization, and noise removal, significantly enhanced recognition accuracy. Despite moderate noise and variations in handwriting styles, the OCR system accurately retrieved character sequences in most input samples. Errors were primarily observed in cases involving ambiguous characters or highly cursive scripts—long-standing challenges in handwriting recognition tasks.

4.2. Translation Quality. Following OCR, the extracted Sanskrit text was fed into the translation module. The quality of translations was evaluated using standard metrics: ROUGE, Precision, Recall, and F1-score. The average results on the test dataset are as follows:

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

These metrics indicate high-quality translations. The incorporation of attention mechanisms allowed the model to focus on relevant parts of the input sequence, leading to improved alignment and fluency in the translations—particularly important given the syntactic richness of Sanskrit.

4.3. Training Progress: Loss and Accuracy Trends. To understand the model's learning behavior and generalization ability, training and validation loss and accuracy were monitored across multiple epochs. These trends shed light on convergence and potential overfitting.

4.3.1. Loss Curve. Both training and validation losses showed a consistent decline. Training loss reduced steadily, while validation loss stabilized after approximately 18 epochs. This indicates efficient learning and suitable regularization. Minor fluctuations in validation loss are attributed to handwriting variations and sentence complexity.

4.3.2. Accuracy Curve. Training accuracy improved progressively, and validation accuracy exhibited a similar upward trajectory with occasional plateaus. The final training accuracy reached approximately 98%, and validation accuracy peaked at around 99%.

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

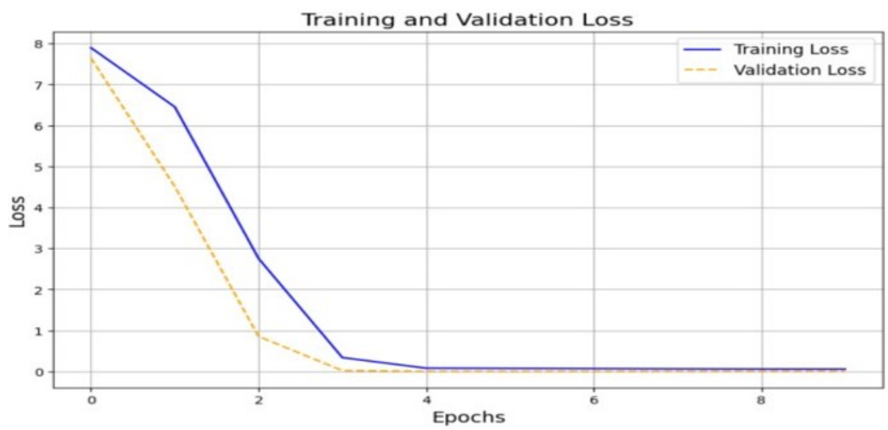

Figure 2: Loss curves for training and validation across epochs

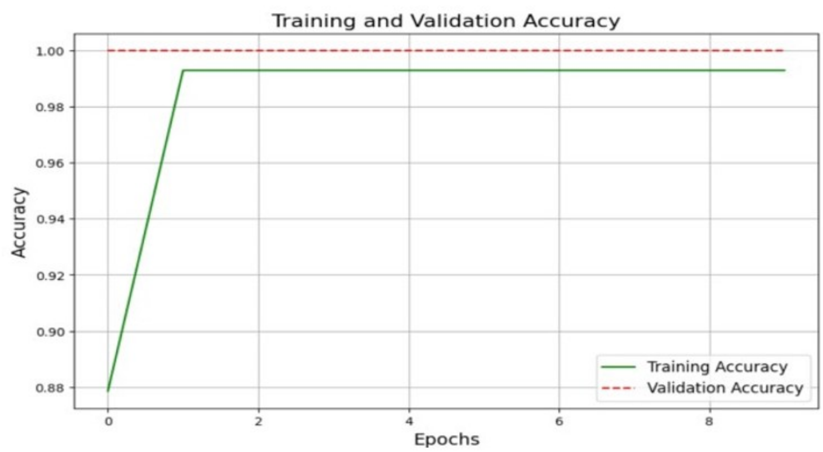

Figure 3: Accuracy curves for training and validation across epochs

5. Conclusion/Future Scope

Our research develops an integrated deep learning system for translating handwritten Sanskrit into English, utilizing EasyOCR for text extraction, OpenCV for image preprocessing, and an LSTM-based encoder-decoder architecture enhanced with attention mechanisms. The system effectively addresses challenges posed by diverse handwriting styles, the complexity of Sanskrit grammar, and the Devanagari script. Metrics including ROUGE, Precision, Recall, and F1-score demonstrate the model's strong translation performance. Training and validation graphs also demonstrate consistent learning with minimal overfitting. This work aids in the preservation and accessibility of ancient Sanskrit manuscripts for modern users. In the future, the system can be expanded to support multiple Indian languages, utilize larger and more diverse datasets, and incorporate advanced transformer-based models for enhanced accuracy. Integration of semantic understanding, grammatical correction modules, and real-time deployment through mobile or web applications could further improve its practical utility and accessibility. Overall, this research contributes to the broader goal of digitizing and preserving India's rich linguistic and cultural heritage.

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.