International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

Handwritten Vedic Sanskrit Text Recognition Using Deep Learning and Convolutional Neural Networks

Ashi Maheshwari
School of Computer Science and Applicaions
Nirwan University Jaipur, India
ashi.maheshwari@nirwanuniversity.ac.in

Reema Ajmera
School of Computer Science and Applications
Nirwan University Jaipur, India
reema.ajmera@nirwanuniversity.ac.in

Dinesh Kumar Dharamdasani
Department of Computer Science
University Of The People,
Pasadena, California, United Ststes
Dineshkumar.dharamdasani@uopeople.edu

Abstract—Recognizing Vedic Sanskrit text is essential for accessing classical Indo-Aryan language, predominantly utilized in the Vedas. Currently, there is limited awareness about the Vedas, making this field a highly demanding and challenging area in pattern recognition. To accelerate progress in optical character recognition (OCR), deep learning methods are indispensable. This article presents a novel approach to Vedic Sanskrit text recognition, incorporating deep convolutional architectures with their respective interpretations. We introduce three modified 4-fold CNN architectures and the AlexNet model. Our system comprises a handwritten dataset containing 140 distinct Vedic Sanskrit words, with approximately 500 images per word, totaling around 70,000 images. The dataset is partitioned for training and testing in an 80:20 ratio. Training is conducted using 20% of the samples, and the resulting model is applied to the deep convolutional network with varied sets of neurons in their hidden layers. Our proposed method demonstrates robust support for accurate Vedic Sanskrit word classification. The recognition rate achieved in our research is 97.42%, with an average recognition time of 0.3640 milliseconds, surpassing existing CNN-based approaches.

Keywords— text detection, text recognition, natural image, machine learning, text localization, text enhancement.

I. Introduction

Preserving ancient texts is crucial for understanding and appreciating the rich cultural heritage of humanity. Among these texts, Vedic Sanskrit manuscripts hold a special place, offering insights into ancient Indian philosophy, religion, and linguistics. However, many of these manuscripts exist only in handwritten form, making their preservation and digitization challenging tasks. Handwritten Vedic Sanskrit text recognition using deep learning presents a promising solution to this problem, leveraging modern machine learning techniques to automatically transcribe handwritten manuscripts into digital format.

The study of Vedic Sanskrit is not merely an academic pursuit but a cultural imperative. Rooted in the ancient Indian civilization, Vedic Sanskrit texts constitute the oldest layer of Sanskrit literature and are foundational to Hinduism. They encompass a vast array of religious hymns, rituals, philosophical treatises, and linguistic principles, providing invaluable insights into the worldview and intellectual achievements of ancient Indian society. However, many of these texts have been transmitted orally for centuries before being committed to writing, resulting in numerous handwritten manuscripts scattered across libraries, archives, and private collections worldwide.

The digitization of handwritten Vedic Sanskrit manuscripts presents several challenges due to the script's complexity and variability. Vedic Sanskrit is written in the Devanagari script, characterized by intricate ligatures, diacritical marks, and stylistic variations. Moreover, the manuscripts themselves may exhibit fading, smudging, or other forms of degradation, further complicating the task of accurate transcription. Traditional optical character recognition (OCR) methods often struggle to decipher handwritten text, especially in the context of ancient scripts with limited training data.

Deep learning, a subfield of machine learning inspired by the structure and function of the human brain, has emerged as a powerful tool for image recognition tasks, including handwritten text recognition. By leveraging large annotated datasets and sophisticated neural network architectures, deep learning models can learn to extract meaningful features from raw images and make predictions with remarkable accuracy. In the context of handwritten Vedic Sanskrit text recognition, deep learning offers a promising avenue for

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

automating the transcription process and preserving these invaluable cultural artifacts for future generations.

This research aims to develop and evaluate deep learning models specifically tailored to the task of recognizing handwritten Vedic Sanskrit characters and symbols. The proposed approach involves collecting a diverse dataset of handwritten manuscripts, preprocessing the data to enhance its quality and suitability for training, selecting appropriate deep learning architectures, and optimizing model parameters through iterative experimentation. The trained models will be evaluated using standard metrics such as accuracy, precision, recall, and F1-score, providing insights into their performance and potential areas for improvement.

By successfully recognizing handwritten Vedic Sanskrit text using deep learning, this research has the potential to revolutionize the digitization and preservation of ancient manuscripts, making them more accessible to scholars, students, and enthusiasts worldwide. Beyond its immediate applications in cultural heritage preservation, this work contributes to the broader field of deep learning-based image recognition and sets a precedent for leveraging cutting-edge technology to unlock the secrets of our shared human history.

II. REVIEW OF LITERATURE

1) **Historical Context of Vedic Sanskrit**: To appreciate the significance of Vedic Sanskrit text recognition, it is essential to understand the historical and cultural context of the Vedas. Scholars such as Max Müller and F. Max Müller have conducted extensive research on Vedic literature, elucidating its linguistic features, religious significance, and historical background. Their works serve as foundational texts for understanding the linguistic characteristics of Vedic Sanskrit and its importance in ancient Indian civilization.

2) **Handwriting Recognition and Optical Character Recognition (OCR):** The field of handwriting recognition and OCR has witnessed significant advancements in recent decades. Early approaches relied on rule-based algorithms and feature extraction techniques to recognize handwritten characters. However, with the emergence of machine learning and deep learning, researchers have achieved remarkable progress in automatic handwriting recognition. Studies by LeCun et al. (1998) on convolutional neural networks (CNNs) paved the way for modern deep learning-based OCR systems, demonstrating their effectiveness in handling complex patterns and variations in handwriting.

3) **Deep Learning for Handwritten Text Recognition**: Deep learning has revolutionized the field of handwritten text recognition, enabling the development of highly accurate and robust OCR systems. Researchers such as Graves et al. (2009) have explored the use of recurrent neural networks (RNNs) and connectionist temporal classification (CTC) for sequence recognition tasks, achieving state-of-the-art performance on benchmark datasets. More recently, the application of attention mechanisms and transformer-based models has further improved the accuracy and efficiency of handwritten text recognition systems.

4) **Vedic Sanskrit Text Recognition:** Despite the extensive research in OCR and handwritten text recognition, limited attention has been given to the specific challenges posed by Vedic Sanskrit manuscripts. Few studies have addressed the complexities of Vedic Sanskrit script, including ligatures, diacritics, and stylistic variations. Notable contributions include the work of Sharma et al. (2018), who proposed a deep learning-based approach for Vedic Sanskrit text recognition, focusing on character segmentation and classification. However, there remains a need for further research to develop robust and scalable solutions for recognizing handwritten Vedic Sanskrit text.

5) **Proposed Methodologies and Approaches**: Recent studies have proposed novel methodologies and approaches for Vedic Sanskrit text recognition using deep learning techniques. For example, Sharma et al. (2020) introduced a modified CNN architecture tailored to the complexities of Vedic Sanskrit script, achieving promising results on a dataset of handwritten manuscripts. Similarly, Gupta et al. (2021) proposed a hybrid approach combining CNNs with recurrent neural networks for improved sequence recognition in Vedic Sanskrit text.

III. METHODOLOGIES

In optical character recognition (OCR), the input image undergoes a series of preprocessing and classification stages. Our research harnesses the power of deep convolutional neural network (CNN) architecture due to its effectiveness in feature extraction and image classification tasks. CNN operates through layers, extracting nuanced features via matrix operations on the input image matrix and kernel matrix, thereby generating feature maps for subsequent layers. We proposed three distinct modified Deep CNN 4-Folds Architectures, each tailored by adjusting various parameters, and evaluated their performance in terms of model accuracy, recognition accuracy, and recognition time, particularly for Vedic Sanskrit recognition. Given the scarcity of attempts in Vedic Sanskrit text recognition, our research aims to fill this gap and improve optical character recognition (OCR) accuracy for ancient scripts like Sanskrit, Marathi, Pali, etc.

Dataset Creation

We curated a dataset comprising 50 classes and 140 Vedic Sanskrit words, totaling 70,000 images, categorized as clear, blurred, and poorly written words. Each class had approximately 500 images, sourced from both right-handed and left-handed writers. We maintained an 80:20 split ratio for training and testing purposes. Data collection involved leveraging various sources, including YouTube, Kaggle, and the 'LaghuSiddhantaKaumudi' book authored by Kaushal Kishor Pandey.

Preprocessing

To enhance data quality and readability, we conducted preprocessing on input images. This involved converting images to grayscale, then to binary form, and resizing them

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

to 64*64 dimensions. These preprocessing steps facilitated clearer recognition outcomes and streamlined subsequent processing stages.

Feature Extraction & Classification

The core of our approach lies in feature extraction and classification using CNN layers. Initial layers, such as input layers, encode the input image in pixel format. Subsequent layers, particularly the deep convolutional layers, perform matrix operations to generate feature maps, crucial for identifying intricate patterns within the image. We implemented batch normalization and ReLU activation to optimize processing speed and mitigate overfitting. Additionally, max pooling was utilized for dimension reduction, followed by fully connected layers for higher-level reasoning.

Model Optimization

We experimented with different optimization techniques, including Stochastic Gradient Descent (SGD) and Adam optimizers, analyzing their impact on model performance. Our findings indicated that the Adam optimizer yielded superior results, leading us to select Model 2 for further validation.

Learning Rate Analysis:

Furthermore, we conducted an analysis of learning rates, exploring values ranging from 0.01 to 0.025. Through experimentation, we determined that a learning rate of 0.015 exhibited optimal performance for our models.

Model Selection for Real-Time Testing

After rigorous investigation and validation, Model 2 emerged as the most suitable candidate, boasting acceptable validation accuracy. With a validation accuracy of 98.88%, we deemed Model 2 appropriate for real-time testing of our system.

In conclusion, our methodological framework encompasses meticulous dataset creation, preprocessing, and model optimization, culminating in the selection of an optimal CNN architecture for Vedic Sanskrit text recognition. This comprehensive approach ensures robust performance and sets the stage for practical implementation in real-world scenarios.

IV. Flowchart for Methodology

1) Dataset Creation:
- Gather Vedic Sanskrit word samples from various sources (e.g., books, online resources).
- Categorize samples into classes based on clarity (clear, blurred, poorly written).
- Organize dataset with appropriate labels for training and testing.

2) Preprocessing:
- Convert input images to grayscale.
- Apply binary conversion to enhance clarity.
- Resize images to a standardized dimension (e.g., 64x64 pixels).

3) Feature Extraction & Classification:
- Implement deep convolutional neural network (CNN) architecture.
- Define initial layers for input image encoding.
- Configure deep convolutional layers for feature extraction.
- Integrate batch normalization and ReLU activation for optimization.
- Utilize max pooling for dimension reduction.
- Incorporate fully connected layers for higher-level reasoning.

4) Model Optimization:
- Experiment with different optimization techniques (e.g., SGD, Adam).
- Evaluate optimizer performance on model accuracy and speed.
- Select the most effective optimizer for further refinement.

5) Learning Rate Analysis:
- Analyze the impact of various learning rates on model performance.
- Determine optimal learning rate for improved training.

6) Model Selection for Real-Time Testing:
- Validate models using a separate validation dataset.
- Select the model with the highest validation accuracy for real-time testing.
- Ensure the chosen model meets the desired criteria for accuracy and efficiency.

V. Experimental Results

Our experimentation focused on evaluating the performance of our proposed system, which encompasses three 4Fold-CNN models alongside AlexNet, for the task of Vedic Sanskrit text recognition. Given the absence of existing solutions tailored specifically for Vedic Sanskrit, we opted for a comparative analysis with Marathi text recognition. Additionally, we subjected Model 2 (4F-CNN-M2) to testing using the Marathi DHDC dataset to gauge its performance against existing methodologies utilizing the same dataset.

A. Performance Evaluation

In Table 1, we present the recognition accuracy and processing time of each model:

Model	**4F-CNN-M1**	**4F-CNN-M2**	**4F-CNN-M3**
Accuracy (%)	98	98.8	92.6
Time (ms)	0.3621	0.3640	0.6342
Optimizer	Adam	Adam	Adam

Table 1: Model-wise Recognition Accuracy & Time

Model 4F-CNN-M2 demonstrated superior recognition accuracy compared to other models, achieving an accuracy of 98.8%.

Comparative Analysis

Table 2 provides a comparative overview of our proposed algorithm's performance against existing approaches.

Reference Work	**Proposed Algorithm**	**Dataset**	**Recognition Accuracy (%)**
Bhardwaj et al (2022)	Deep Learning Model	DHDC	98.13
Manocha et al (2021)	Deep Learning CNN	DHDC	92
Bhist et al (2020)	Deep CNN	DHDC	98.93
Deore et al (2020)	Fine-tuned Deep CNN	DHDC	96.55
Acharya et al (2015)	Deep CNN	DHDC	98.26
Our Model (2022)	4-Fold Deep CNN Model	DHCD	98.94
	Own Dataset		97.42

TABLE 2: COMPARATIVE STUDY ON DEVANAGARI TEXT RECOGNITION

Our proposed 4-Fold Deep CNN Model demonstrated superior recognition accuracy compared to existing approaches, achieving 98.94% on the DHCD dataset.

Category-wise Performance

Table 3 showcases the performance of our system across various text categories:

Category	**Recognition Accuracy (%)**	**Average Recognition Time (ms)**
Confusing Text	82	0.2654
Blur Images	65	0.4062
Badly Written Images	78	0.3280

TABLE 3: COMPARATIVE ANALYSIS - TEXT CATEGORIES

Our system exhibited satisfactory performance across different text categories, albeit with varying recognition accuracies and processing times.

Model Comparison

Table 4 compares the performance of 4F-CNN-M2 and AlexNet:

Vedic Text Recognition	**4F-CNN-M2**	**AlexNet**
Recognition Accuracy	98.8	53.75
Average Recognition Time	0.364	0.406

TABLE 4: COMPARATIVE ANALYSIS - 4F-CNN-M2 &ALEXNET

4F-CNN-M2 outperformed AlexNet in both recognition accuracy and average recognition time.

VI. CONCLUSION

Our extensive experimentation and analysis underscore the effectiveness and potential of our proposed system for Vedic Sanskrit text recognition. Across various metrics and comparative studies, several key findings emerge, highlighting the robustness and significance of our approach.

A. Superior Performance of 4F-CNN-M2

Among the models evaluated, 4F-CNN-M2 stands out as the top performer, boasting a remarkable recognition accuracy of 98.8%. This superior performance underscores the efficacy of our tailored CNN architecture in accurately recognizing Vedic Sanskrit text. The meticulous design and optimization of this model have resulted in enhanced accuracy and efficiency, making it a promising solution for practical applications.

B. Validation Against Existing Approaches:

Our comparative analysis against existing methodologies in the field of Devanagari text recognition further validates the strength of our proposed algorithm. With recognition accuracies surpassing 98% on standard datasets, our system outperforms several contemporary approaches, demonstrating its superiority and reliability. This validation not only reinforces the efficacy of our approach but also positions it as a leading contender in the domain of ancient text recognition.

C. Addressing Textual Challenges

One notable aspect of our experimentation is the comprehensive evaluation across different text categories, including confusing text, blur images, and badly written images. Despite the inherent challenges posed by these variations, our system exhibits satisfactory performance, showcasing its adaptability and robustness. By effectively addressing these textual challenges, our system proves its versatility and applicability across diverse real-world scenarios.

D. Model Comparison and Efficiency

Our comparative analysis between 4F-CNN-M2 and AlexNet highlights the efficiency and superiority of our tailored CNN architecture. With significantly higher recognition accuracy and lower average recognition time, 4F-CNN-M2 outperforms AlexNet, reaffirming the effectiveness of our approach. This comparison not only emphasizes the importance of model selection but also underscores the advantages offered by our optimized architecture.

REFERENCES

[1] Dessai, Brijeshwar, and Amit Patil. "A deep learning approach for optical character recognition of handwritten Devanagari script." 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT). Vol. 1. IEEE, 2019.

[2] Hazra, Tapan Kumar, Rajdeep Sarkar, and Ankit Kumar. "Handwritten English character recognition using logistic regression and neural network." Int J Sci Res (IJSR) 5.6 (2016): 750-754.

[3] Deoras, Anoop, et al. "Variational approximation of long-span language models for LVCSR." 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2011.

[4] Shevgoor, Sowmya Kamath. "A Morphological Approach for Measuring Pair-Wise Semantic Similarity of Sanskrit Sentences." Natural Language Processing and Information Systems: 22nd International Conference on Applications of Natural Language to Information Systems, NLDB 2017, Liège, Belgium, June 21-23, 2017, Proceedings. Vol. 10260. Springer, 2017.

[5] Saluja, Rohit. Robust Multilingual OCR: from Ancient Indic Texts to Modern Indian Street Signs. Diss. Monash University, 2020.

[6] Yash Gurav, Priyanka Bhagat, “Devanagari Handwritten Character Recognition using Convolutional Neural Networks”, 2020 International Conference on Electrical, Communication and Computer Engineering(ICECCE).

[7] Rohit Saluja, Devaraj Adiga,” A Framework for Document Specific Error Detection and Corrections”, in Indic OCR2017 14th IAPR International Conference on Document Analysis and Recognition.

[8] Anupama Thakur, Amrit Kaur”, Devanagari Handwritten Character Recognition using Convolutional Neural Networks”,2019 International Journal of Scientific & Technology Research.

[9] MeduriAvadesh and Navneet Goya, “Optical Character Recognition for Sanskrit Using Convolution Neural Networks”, 2018 13th IAPR International Workshop on Document Analysis Systems.

[10] Rohit Saluja et al. "A framework for document specific error detection and corrections in indicocr." 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). Vol. 4. IEEE, 2017.

[11] Raj V. Amrutha, R. L. Jyothi, and A. Anilkumar. "Grantha script recognition from ancient palm leaves using histogram of orientation shape context." 2017 International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2017.

[12] Pandey, Bipul, et al. "Multilingual speaker recognition using ANFIS." 2010 2nd International Conference on Signal Processing Systems. Vol. 3. IEEE, 2010.

[13] Bhardwaj, Anuj. "An Accurate Deep-Learning Model for Handwritten Devanagari Character Recognition."

[14] Manocha, Shilpa Kaur, and Piyush Tewari. "Devanagari Handwritten Character Recognition using CNN as Feature Extractor." 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON). IEEE, 2021.

[15] Bisht, Mamta, and Richa Gupta. "Multiclass recognition of offline handwritten Devanagari characters using CNN." International Journal of Mathematical, Engineering and Management Sciences 5.6 (2020): 1429.

[16] Deore, Shalaka Prasad, and Albert Pravin. "Devanagari handwritten character recognition using fine-tuned deep convolutional neural network on trivial dataset." Sādhanā 45.1 (2020): 1-13.

[17] Acharya, Shailesh, Ashok Kumar Pant, and Prashnna Kumar Gyawali. "Deep learning based large scale handwritten Devanagari character recognition." 2015 9th International conference on software, knowledge, information management andapplications (SKIMA). IEEE, 2015.