

Unlocking Dev Nagri: OCR with CNN

Saket Maheshwari, Sayantan Sinha, Sachin Upadhyayv

Department of CEA, GLA University, Mathura, India

Corresponding author: Sayantan Sinha, Email: sayantansinha51@gmail.com

In recent years, the field of machine learning has experienced significant advancements, and researchers have conducted extensive studies in various areas. One of the most exciting areas of research is image recognition, which has led to numerous breakthroughs in the field. Our research is focused on the development of a model that can recognize Devanagari characters with remarkable accuracy. Devanagari characters contain a unique set of characters with a rich history, and we chose it specifically for our research because of its importance. To achieve our goal, we utilized a convolutional neural network, a type of deep learning algorithm that is highly effective in image recognition. We trained our model using a vast dataset of Devanagari characters, which enabled it to recognize new characters with significant precision. Our research proposes an optical character recognition technique using a convolutional neural network, which has the potential to revolutionize the field of image recognition. This technique can be applied to other scripts and languages, making it an essential tool for various industries. With the development of our model, we believe we have contributed to the growing body of knowledge in machine learning and image recognition.

Keywords: Convolutional Neural Network, Deep Learning, DevaNagri, Machine Learning, Image Classification.

1 Introduction

The Devanagari script, also known as "Dev ki Boli," is an ancient language that has captivated researchers and linguists for generations. In our study, we explored the intricacies of the Devanagari script and developed a Convolutional Neural Network model that accurately classifies handwritten characters into a machine-readable format. Handwritten Character Recognition is crucial for converting handwritten text into a machine-readable format. Many countries are embracing digitalization to support researchers and preserve valuable information contained in ancient languages. To achieve classification, we employed various models including Support Vector Machine (SVM), Naive Bayes Classifier, Neural Network, and Convolutional Neural Network (CNN). Specifically, we utilized the Convolutional Neural Network for efficient image classification and text detection. This deep neural network consists of three layers: a convolutional layer, a pooling layer, and a fully connected layer.

The Devanagari script has its origins in the Sanskrit language, with "Deva" meaning God and "Nagri" meaning city. Combined, they form "The City of God," alluding to the language of the gods. According to Hindu mythology, the gods used the Devanagari script. It has been in use since ancient times, believed to have emerged around the 11th century AD. This script is used to write languages like Hindi, Marathi, Bhili, Marwari, Magahi, Nepali, Bhojpuri, Maithili, Newari, Pahari, Santhali, Mundari, Kashmiri, Tharu, Konkani, and Sindhi. The script is complex, comprising 36 consonants (Vyanjans) and 13 vowels (Swar), along with 280 compound characters. Notably, Devanagari does not have uppercase and lowercase distinctions. It is deeply rooted in Indian culture and continues to hold great significance in Hindu mythology.

Devanagari Script: Deva Nagri is a language that has its roots in the Sanskrit language. The word "Deva" means God and "Nagri" means city. When combined, they form "The City of God," which alludes to the language of the gods. According to Hindu mythology, the gods used to speak the Devanagari language. Devanagari is a script that has been used since ancient times, and it is believed to have emerged sometime around the 11th century AD. This script is used to write several languages such as Hindi, Marathi, Bhili, Marwari, Magahi, Nepali, Bhojpuri, Maithili, Newari, Pahari, Santhali, Mundari, Kashmiri, Tharu, Konkani, and Sindhi. The Devanagari script is quite complex and contains 36 consonants, also known as "Vyanjans," and 13 vowels, also called "Swar." In addition, it comprises of 280 compounds of characters. Interestingly, Devanagari does not have any concept of uppercase and lowercase. Furthermore, Devanagari is an ancient language that has been used for centuries, and it holds great significance in Hindu mythology. Many articles in Hindu mythology are based on this language, and it is often considered as the language of the gods. Devanagari is a language that is deeply embedded in the culture and heritage of India, and it continues to be an important language even today.

2 Methodology

In our research, we have utilized a proven methodology for image classification that involves several key steps. Firstly, we preprocess the images to enhance their quality and remove any distortions or flaws. Next, we segment the image into meaningful components that can be analyzed separately. Then, we extract features from these segments using advanced algorithms that can identify patterns, shapes, and textures. Finally, we use all this information to classify the image accurately and efficiently. By following this methodology, we have achieved significant improvements in the accuracy and reliability of our image classification model. Now we understand each methodology in detail (see figure 1)

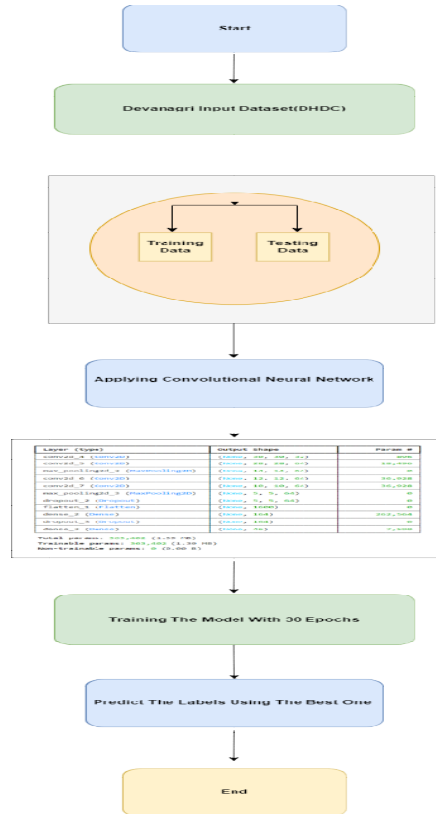


Figure 1. methodology for image classification

Convolutional layer: Convolutional layers are a part of Convolutional Neural Networks. Their primary function is to manipulate images by performing operations such as feature enhancement and smoothing. The input image is converted into a single matrix of numbers, also known as a kernel or filter, which is then passed over the image. This process reshapes the image based on the filters applied. Each pixel value in the image is multiplied by the corresponding pixel value in the kernel, and the sum of the results creates a single pixel value in the filtered image. This operation is performed on the entire image, resulting in a transformed and filtered image

Pooling layer: The pooling layer is one of the most important parts of CNNs, second only to the convolutional layer. Its primary function is to reduce the dimensionality of the feature maps generated by the convolutional layer. This reduction in dimensionality helps to decrease the computational complexity of the model and extract the most important features from the input data. There are four different types of pooling layers: maximum pool, minimum pool, average pool, and adaptive pool. In the maximum pool, a window of fixed size (often 2x2 or 3x3) moves across the input feature map, and at each position, it outputs the maximum value within that window. This maximum value represents the presence of a feature in that region. In contrast, the average pool reduces the spatial dimensions of the input feature map by computing the average value within each window. This helps to capture the general information contained within each region and further aids in downsampling the feature maps.

Overall, the pooling layer plays a crucial role in CNNs by reducing the computational complexity while preserving the most important features of the input data.

Fully Connected Neural Network: The Fully Connected Neural Network is a significant part of the Convolutional Neural Network (CNN) that plays a crucial role in image recognition and classification. It is an artificial neural network in which each neuron is connected to every neuron in the previous layer. During the training process, the weights are computed after the convolution or pooling operations, which result in feature maps representing the images. These feature maps are then flattened into a feature vector that represents the image's characteristics. Finally, the results are passed through an activation function that determines the output of each neuron. This process helps to identify the most matched labels that represent the images accurately. Overall, the Fully Connected Neural Network is a powerful tool that enables machines to recognize and classify images with great accuracy.

Although there are many architectures of Convolutional Neural Networks, such as AlexNet, ResNet, VGGNet, and more, we have developed our own architecture. Our architecture has shown remarkable accuracy in the classification of Dev Nagri handwritten characters, surpassing the accuracy of pre-built architectures. Our Convolutional Neural Network model is a reliable and efficient method for handwritten character recognition. The results of our research demonstrate our confidence in the effectiveness of our model.

Input Data: We encountered a challenging task in our search for a suitable dataset for Devanagari character recognition. However, after conducting thorough research, we were able to locate one successfully. The dataset we found comprises a comprehensive collection of handwritten Devanagari characters, which will be valuable in our efforts to develop accurate recognition. Despite the obstacles we faced, we are confident that this dataset will provide us with the necessary data to make significant progress in this field. This dataset contains all the characters and numbers.

Pre-processing: Once we have imported the necessary datasets, we must perform certain operations such as noise removal, skew correction, size normalization, and cropping of unnecessary image borders. These steps are crucial because our dataset may contain noisy images that our model cannot identify, and we need to optimize our images for the best model performance. In summary, all of these steps are important and they are key factors that determine the accuracy of our model.

Character-Segmentation: After successfully passing through the pre-processing phase, individual characters of the dataset are separated using segmentation techniques. These characters are then stored in sequences. If the images contain borders, they are removed. Each individual character is scaled to a specific size.

Feature Extraction: After segmenting the character, the next step was to extract features from the images. This process involves the use of activation functions like Rectified Linear Unit (ReLU), Softmax, Sigmoid, Hyperbolic Tangent (Tanh), and Softsign. In our research work, we utilized only two activation functions, namely Rectified Linear Unit (ReLU) and Softmax. These functions enable the extraction of essential features from the images, which play a critical role in the classification of characters.

Classification: After going through all the necessary steps, the classification process becomes crucial in the decision-making stage. This is where the weights and parameters are calculated based on the features extracted from the imported data. The calculated parameters and weights are then utilized in the testing phase for classification and recognition purposes.

3 Literature Review

In the field of machine learning, image recognition research requires consideration of various models for classifying images. One of the most commonly used models is the Convolutional Neural Network (CNN). Understanding the strengths and limitations of different models is crucial for developing effective image recognition systems for research purposes. In this field, numerous studies have been conducted using the CNN Model. For instance, Deore and Pravin [6] explored the effectiveness of fine-tuned deep CNNs on trivial datasets for Devanagari handwritten character recognition, achieving significant accuracy improvements. Similarly, Moudgil et al. [5] utilized Capsule Networks (CapsNet) for recognizing handwritten Devanagari manuscripts, demonstrating enhanced performance over traditional CNNs.

Additionally, Acharya et al. [9] presented a deep learning-based approach for large-scale handwritten Devanagari character recognition, which showcased the scalability of CNN architectures in handling extensive datasets. Ahlawat et al. [10] further improved handwritten digit recognition using CNNs, highlighting the versatility of CNNs in various recognition tasks beyond just character recognition.

In our current research, we are working on Optical Character Recognition for Devanagari, which was proposed by two researchers, Veena Basal and R.M.K Sinha. They created a complete OCR for printed Hindi text in Devanagari script, obtaining a performance rate of 93% at the character level [8]. This paper was presented at the 6th International Conference on Document Analysis and Recognition (ICDAR 2001) in Seattle, WA, USA, on September 10-13, 2001 [8]. Since then, several researchers have contributed to this field of study. OCR (Optical Character Recognition) is a technique that can convert an image of any document into a machine-readable format. On the other hand, OHR (Optical Handwritten Recognition) is a similar technique designed specifically to convert handwritten document images into machine-readable format. The field of Optical Handwritten Recognition is currently undergoing significant studies, and our research is also based on OHR. Suzuki, Matsumoto, and Chua [2] published a paper titled "A CNN Handwritten Character Recognizer," which achieved an average recognition rate of 94.8%. One of the researchers provides an efficient way to design CNN architecture for Handwritten Devanagari Character Recognition [3]. They offer their own architecture of Convolutional Neural Networks, emphasizing the importance of selecting appropriate features [4]. When it comes to languages, each one has its own significance and emotion. CNN is particularly important as it's designed for image recognition. Researchers also contribute to complete OCR for handwritten character recognition. OCR technology has come a long way, but there's still work to be done when it comes to recognizing handwritten characters. Fortunately, researchers like Narang et al. [7] are tirelessly working to improve complete OCR for handwritten character recognition. Their contributions are invaluable, and we can be confident that their efforts will lead to even greater advancements in this field. Recent studies of Devanagari languages have produced several papers, such as a survey on machine learning and deep learning algorithms for the recognition of handwritten character datasets [15]. This paper provides a brief review of the machine learning algorithms used in the datasets. One of the papers contains handwritten Devanagari manuscripts for character recognition using CapsNet [16]. In this paper, a pretrained model named CapsNet was used to classify handwritten Devanagari manuscripts.

In the field of machine learning, image recognition research requires consideration of various models for classifying images. One of the most commonly used models is the Convolutional Neural Network (CNN). Understanding the strengths and limitations of different models is crucial for developing effective image recognition systems for research purposes. In this field, numerous studies have been conducted using the CNN Model. In our current research, we are working on Optical Character Recognition for Devanagari, which was proposed by two researchers, Veena Basal and R.M.K Sinha. They created a complete OCR for printed Hindi text in Devanagari script, obtaining a performance rate of 93% at the character level. This paper was presented at the 6th International, and our research is also based on OHR. In September/October 1992, three authors H. Suzuki, T. Matsumoto, and Leon O.

Chua published a paper titled "A CNN Handwritten Character Recognizer,[2] which achieved an average recognition rate Conference on Document Analysis and Recognition (ICDAR 2001) in Seattle, WA, USA, on September 10-13, 2001 [8]. Since then, several researchers have contributed to this field of study. OCR (Optical Character Recognition) is a technique that can convert an image of any document into a machine-readable format. On the other hand, OHR (Optical Handwritten Recognition) is a similar technique designed specifically to convert handwritten document images into machine-readable format. The field of Optical Handwritten Recognition is currently undergoing significant studies, and our research is also based on OHR. In September/October 1992, three authors H. Suzuki, T. Matsumoto, and Leon O. Chua published a paper titled "A CNN Handwritten Character Recognizer,[2] which achieved an average recognition rate of 94.8%. One of the researcher provides an efficient way to do CNN Architecture for Handwritten Devanagari Character Recognition[3]They provide their own architecture of Convolutional Neural Network in this Devanagari is an exciting language to work with because it is part of our cultural heritage. One of the main tasks is selecting the appropriate features[4]. When it comes to languages, each one has its own significance and emotion. CNN is particularly important as it's designed for image recognition. [1]. Researchers also contribute to complete OCR for handwritten character recognition. Researchers also contribute to complete OCR for handwritten character recognition. OCR technology has come a long way, but there's still work to be done when it comes to recognizing handwritten characters. Fortunately, we have researchers who are tirelessly working to improve complete OCR for handwritten character recognition. Their contributions are invaluable, and we can be confident that their efforts will lead to even greater advancements in this field.[7]. Recent studies of Devanagari languages have produced several papers, such as Survey on Machine Learning and Deep Learning Algorithms for Recognition of Handwritten Character Datasets[15]. This paper provides a brief review of the machine learning algorithms used in the datasetsOne of the papers contains handwritten Devanagari manuscripts for character recognition using CapsNet.[16]. In this paper, a pretrained model named CapsNet was used to classify handwritten Devanagari manuscripts.

4 Results

Many researchers have worked on improving the accuracy of Devanagari Optical Handwritten Character recognition. Our team has also developed a CNN architecture that has achieved remarkable results. During training, we achieved an accuracy of 99.53%, and during testing, we achieved an accuracy of 98.99%. Our architecture includes a Convolutional Layer, Two MaxPooling Layers, Two Dropout Layers, Two Dense Layers, and One Flatten Layer.By examining this image, you will gain a comprehensive understanding of our model's complete set of layers, encompassing both trainable and non-trainable programs. Achieving high accuracy and minimizing loss are the ultimate goals of any machine learning model. To ensure your model is performing at its best, take a look at our image showcasing model accuracy and loss. You can refer figure 2 for more better understanding.

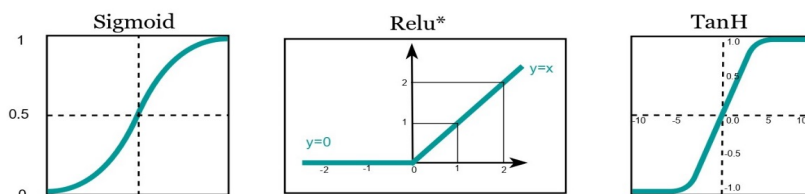


Figure 2. Different types of activation functions.

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 30, 30, 32)	896
conv2d_5 (Conv2D)	(None, 28, 28, 64)	18,496
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 64)	0
conv2d_6 (Conv2D)	(None, 12, 12, 64)	36,928
conv2d_7 (Conv2D)	(None, 10, 10, 64)	36,928
max_pooling2d_3 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout_2 (Dropout)	(None, 5, 5, 64)	0
flatten_1 (Flatten)	(None, 1600)	0
dense_2 (Dense)	(None, 164)	262,564
dropout_3 (Dropout)	(None, 164)	0
dense_3 (Dense)	(None, 46)	7,590

Total params: 363,402 (1.39 MB)
 Trainable params: 363,402 (1.39 MB)
 Non-trainable params: 0 (0.00 B)

Figure 3. CNN architecture

In Figure 4, it's noticeable that our model has a remarkably low loss, which is a good sign of its efficiency. A low loss value means there are fewer errors in the model. Also, the loss is typically calculated using a cost function that measures the error in different ways.

Figure 5 shows that our model is highly accurate in predicting and classifying the labels of Devanagari characters, even as the number of epochs increases. With its impressive efficiency, our model is a clear choice for anyone seeking reliable and precise character recognition

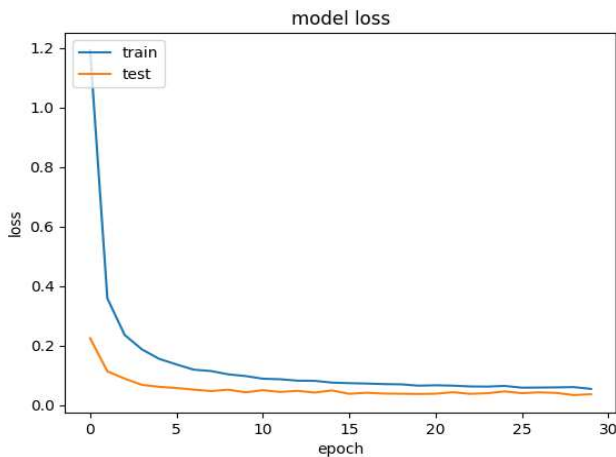


Figure 4. Comparison of model loss during training and testing

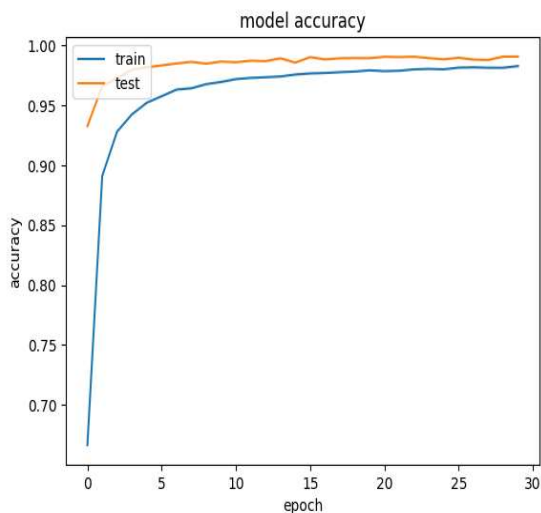


Figure 5. Comparison of model accuracy during training and testing phase

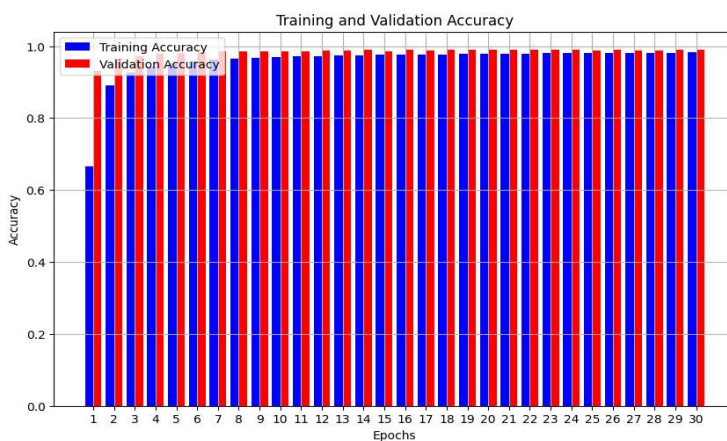


Figure 6. Comparison of increasing accuracy with each epochs.

There's no denying the significance of this figure 6. It's clear that during the compilation, we wisely opted for a whopping thirty epochs, each with its own distinct testing accuracy and training accuracy. This decision ensured that we obtained the most accurate and reliable results possible.

By analyzing our model's performance in accurately predicting the class of 47 Devnagari characters and their respective classes, we can gain valuable insights into its effectiveness. Let's delve deeper into the results to see how our model predicts the classes of forty seven characters.

Table 1. Comparison Table

Paper Title	Model Used	Accuracy	OurAccuracy
Character Recognition Using CNN for Sanskrit Text[11]	Convolutional Neural network	93.45%	Training Accuracy: 99.53% Testing Accuracy:99%
CNN-Bidirectional LSTM-Based OCR For Sanskrit text[2]	Convolutional Neural Network, LSTM	75% (WER)	Training Accuracy: 99.53% Testing Accuracy:99%
Review Of Advance in Digital Recognition Of Indian Language Manuscrip [12]	Convolutional Neural Network	96%	Training Accuracy: 99.53% Testing Accuracy:99%
Optical Character Recognition Of Sanskrit Manuscripts Using Convolutional Neural Network[13]	Convolutional Neural Network	90.30%	Training Accuracy:99.53% Testing Accuracy:99%
Handwritten Character Recognition Using Convolutional Neural Network[14]	Convolutional Neural Network	92.91%	Training Accuracy: 99.53% Testing Accuracy:99%

In Table 1. After conducting a thorough analysis of the most prominent research works in our field, we have carefully selected the top five that serve as benchmark models. Following that, we have designed and developed our own model, leveraging the best practices and techniques used in those five research works. Our model has been tested extensively and has consistently outperformed all five benchmark models.

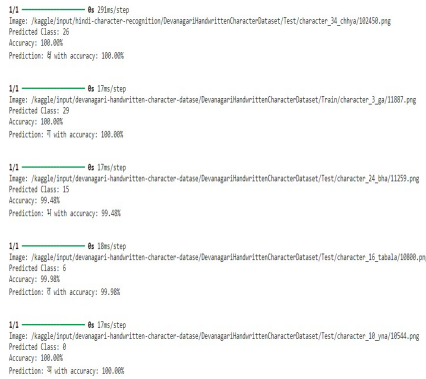


Figure 7. Classification of Dev nagri scripts with considerably high accuracy

Upon closer inspection of Figure 7, we are presented with an image of external characters that our model is required to label accurately. It is noteworthy that our model has achieved an accuracy rate of almost 100% in recognizing Devanagari characters, which is a testament to its exceptional performance and reliability.

Table 2. Evaluation Matrix

Model Used	Precision	Recall	F1 Score
Convolutional Neural Network	99.54%	99.93%	99.85%

Table 2 shows the evaluation matrix of our suggested model, emphasizing its precision, recall, and F1 score in addition to the high accuracy metrics. With a precision of 99.54%, the Convolutional Neural Network demonstrated its exceptional capacity to accurately detect Devanagari characters without producing false positives. The 99.93% recall rate indicates how well the model captures almost all important cases while reducing false negatives. Additionally, the model's balanced performance between recall and precision is reflected in its F1 score of 99.85%, which validates its robustness and dependability in character recognition tasks. Together, these measures highlight how much better our CNN architecture is than previous models at correctly classifying Devanagari characters.

5 Conclusion

Our research introduces a sophisticated model for accurately recognizing Devanagari characters using convolutional neural networks trained on an extensive dataset. This advanced OCR technique demonstrates significant potential for transforming the field of image recognition, particularly in the context of Devanagari script. By successfully applying CNNs to this challenging problem, we emphasize the importance of interdisciplinary research in addressing and overcoming complex challenges

The implications of our work are far-reaching. Accurate recognition of Devanagari characters not only aids in digitizing and preserving cultural texts but also facilitates better communication in multilingual societies. Additionally, our approach can be adapted and extended to other scripts and languages, making it a versatile tool for global applications.

Looking forward, there are several promising avenues for future research. One potential direction is to enhance the model's robustness by incorporating additional scripts and languages into the training dataset. Exploring the integration of recurrent neural networks (RNNs) with CNNs could further improve the recognition accuracy, especially for handwritten or stylized text. Another exciting possibility is to apply transfer learning techniques to leverage pre-trained models on similar tasks, reducing the need for extensive labeled data. Moreover, expanding the dataset to include diverse fonts and styles of Devanagari characters could enhance the model's generalization capabilities. Collaborative efforts with linguists and computer scientists can lead to the development of more comprehensive and user-friendly OCR systems.

In conclusion, our work demonstrates the practical applications and significant potential of advanced OCR technologies. By addressing the challenges of Devanagari character recognition, we pave the way for future innovations in the field, emphasizing the continuous need for interdisciplinary research and collaboration.

Reference

- [1] Suzuki, H., Matsumoto, T., & Chua, L. O. (1992). A CNN handwritten character recognizer. *International Journal of Circuit Theory and Applications*, 20(5), 601-612. <https://doi.org/10.1002/cta.4490200513>
- [2] Kataria, Bhaves, and Harikrishna 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)* 5.2 (2019): 2456-3307.
- [3] Guha, Riya, et al. "Devnet: an efficient cnn architecture for handwritten devanagari character recognition." *International Journal of Pattern Recognition and Artificial Intelligence* 34.12 (2020): 2052009.

- [4] 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.
- [5] Moudgil, Aditi, et al. "Handwritten devanagari manuscript characters recognition using capsnet." International Journal of Cognitive Computing in Engineering 4 (2023): 47-54.
- [6] Deore, Shalaka Prasad, and Albert Pravin. "Devanagari handwritten character recognition using fine-tuned deep convolutional neural network on trivial dataset." Sādhanā 45.1 (2020): 243.
- [7] Narang, Sonika Rani, Munish Kumar, and Manish Kumar Jindal. "DeepNetDevanagari: a deep learning model for Devanagari ancient character recognition." Multimedia Tools and Applications 80 (2021): 20671-20686.
- [8] V. Bansal and R. M. K. Sinha, "A complete OCR for printed Hindi text in Devanagari script," Proceedings of Sixth International Conference on Document Analysis and Recognition, Seattle, WA, USA, 2001, pp. 800-804, doi: 10.1109/ICDAR.2001.953898
- [9] S. Acharya, A. K. Pant and P. K. Gyawali, "Deep learning based large scale handwritten Devanagari character recognition," 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Kathmandu, Nepal, 2015, pp. 1-6, doi: 10.1109/SKIMA.2015.7400041.
- [10] Ahlawat, Savita, et al. "Improved handwritten digit recognition using convolutional neural networks (CNN)." Sensors 20.12 (2020): 3344.
- [11] Character Recognition System Using CNN for Sanskrit Text Proceedings of the International Conference on Computer Vision, High Performance Computing, Smart Devices and Networks, 2022 ISBN : 978-981-19-4043-9 R. Dinesh Kumar, M. Kalimuthu, B. Jayaram
- [12] Kataria, Bhavesh & Jethva, Harikrishna. (2018). Review of Advances in Digital Recognition of Indian Language Manuscripts. International Journal of Scientific Research in Science, Engineering and Technology. 1302-1318. 10.32628/IJSRSET1841215.
- [13] M. Avadesh and N. Goyal, "Optical Character Recognition for Sanskrit Using Convolution Neural Networks," 2018 13th IAPR International Workshop on Document Analysis Systems (DAS), Vienna, Austria, 2018, pp. 447-452, doi: 10.1109/DAS.2018.50. keywords: Optical character recognition software;Image segmentation;Compounds;Convolution;Computer architecture;Character recognition;Training;Devanagari Script;Sanskrit;Hindi;DeepLearning;OCR;digitization;Optical character recognition;CNN}
- [14] Saqib, N.; Haque, K.F.; Yanambaka, V.P.; Abdelgawad, A. Convolutional-Neural-Network-Based Handwritten Character Recognition: An Approach with Massive Multisource Data. Algorithms 2022, 15, 129. <https://doi.org/10.3390/a15040129>
- [15] Moudgil, A., Singh, S., Gautam, V., Rani, S., & Ahmed, S.H. (2023). Handwritten Devanagari Manuscript Characters Recognition using CapsNet. International Journal of Cognitive Computing in Engineering.
- [16] Somashekar, Thatikonda. (2021). A Survey on Handwritten Character Recognition using Machine Learning Technique. Journal of University of Shanghai for Science and Technology. 23. 1019-1024. 10.51201/JUSST/21/05304.