IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 6, April 2025

Impact Factor: 7.67

Handwritten Modi Script Character Recognition Using CNN and VGG16 Algorithms

Mrunali Manohar Bhong and Prof. Nanda Kulkarni

Department of Computer Engineering,

Siddhant College of Engineering, Sudumbre, Pune, India

mrunalibhong13@gmail.com

Abstract: *Recognition of handwritten letters in ancient scripts is challenging because of differences in handwriting styles, noise, and complexity of cursive scripts. Modi script, an ancient cursive version of Marathi script, has seen almost complete oblivion of use but is of enormous cultural and historical significance. The implementation of a deep learning system based on convolutional neural networks (CNN) for handwritten Modi character recognition is described in this research and transfer learning using VGG16. Two datasets were employed for experimentation: a public handwritten Modi dataset from IEEE DataPort and a self-created dataset gathered from various individuals to inject writing diversity. The process involves pre-processing operations like noise removal, resizing, binarization, and character segmentation. The CNN model uses convolutional, pooling, and dense layers, whereas the VGG16 model uses pre-trained ImageNet weights and is fine-tuned for 46 Modi characters (vowels and consonants). Both models were trained and tested with accuracy, sensitivity, specificity, and F1-score measures. Experimental outcomes confirm that VGG16 performs better than the basic CNN model, with increased recognition accuracy, especially for characters with similar appearances. The system is highly generalizable across both datasets and performs well in character classification tasks in ancient scripts. This implementation fills an essential gap in digital heritage preservation and illustrates the need to use modern AI methods to revive and digitize culturally valuable scripts such as Modi. Future projects involve expanding the system to process linked handwritten text and trying out attention-based deep-learning architectures for better performance.*

Keywords: Modi Script, Handwritten Character Recognition, Convolutional Neural Network (CNN), VGG16, Transfer Learning, Deep Learning, Image Preprocessing, Ancient Script Digitization, Optical Character Recognition (OCR), Historical Document Analysis

I. INTRODUCTION

Handwritten character recognition (HCR) is essential to contemporary image technologies related to processing and machine learning. It involves automatically identifying and categorizing characters in printed or handwritten text. HCR has made great strides in popular scripts like English, Chinese, and Devanagari, but not much has been done for older regional scripts like Modi. Documents in the Marathi language are written in the ancient cursive style known as the Modi script.

, widely employed under the Maratha Empire for administrative documentation. Owing to its cursive style and flowing strokes of characters, the Modi script poses exceptional challenges to recognition systems. The characters tend to overlap, segmentation becomes highly challenging, and handwriting styles further complicate the recognition process. These conditions render classical OCR systems useless for Modi script recognition, thus calling for strong deep learning-based methods.

The impetus for this study comes from the growing need to digitize historical manuscripts and ancient texts that hold scripts like Modi. As the world is rapidly heading towards digital storage and intelligent archiving, there is an obvious need to preserve culturally essential texts and make them available for research and education. The lack of machine tools for Modi script recognition limits efforts in preserving heritage, linguistic research, and digital humanities.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, April 2025 Impact Factor: 7.67

Convolutional Neural Networks (CNNs), in particular, are deep learning models with immense potential for pattern recognition problems owing to their ability to learn hierarchical visual features. Similarly, through transfer learning, pre-trained models like VGG16 provide an effective tool to leverage acquired knowledge on new small training datasets. This project intends to use both CNN and VGG16 to successfully identify handwritten Modi characters and break the constraints imposed by conventional methods.

The problem statement addressed in this work emphasizes the automated identification of discrete Modi characters produced by hand using deep learning methods. Due to the script's complex structure, high variation in handwriting, and similarity between many character shapes, developing a highly accurate recognition system is not trivial. Most existing methods rely heavily on manual feature extraction and struggle to handle real-world handwriting variability. Hence, this study proposes an end-to-end deep learning method employing CNN and VGG16 models, trained on self-generated and publically accessible datasets, to attain high classification accuracy for 46 different Modi script characters. The goal is to design a system that performs robustly across different handwriting styles and overcomes challenges such as class imbalance and noisy inputs.

The key objectives of the proposed system include collecting and preparing a comprehensive dataset of handwritten Modi characters; applying suitable pre-processing methods to enhance image quality and segment characters effectively; designing and training a CNN model tailored for image classification; implementing a transfer learning-based VGG16 model and fine-tuning it for the Modi script dataset; evaluating both models using evaluating the models' performances to identify the best strategy; and using standard classification metrics like accuracy, sensitivity, specificity, precision, and F1-score. The project also seeks to identify the strengths and limitations of each method and lay the groundwork for future expansion into word and sentence-level Modi script recognition.

This research holds significant importance in document digitization and historical text preservation. By focusing on the Modi script, the work contributes to a relatively unexplored domain and helps bridge the technological gap in Indian heritage preservation. The Recognition system created in this paper can serve as a foundationfor more advanced OCR software. It willbenefit academic researchers, historians, and institutions working with Modi's manuscripts. In addition, this method can be applied to other cursive or minority scripts that share the same structural complexities. The fact that CNN's learning capability and VGG16's pre-training experience are combined makes the system efficient and scalable, even if trained on modest-sized datasets.

Although the system holds promise, several technical challenges exist in identifying Modi's character.The cursive nature of handwriting tends to link characters together, which makes segmentation not a simple process. Variation in individual handwriting adds more variance in stroke width, direction, and slope. The absence of extensive and standard labeled datasets for the Modi script adds to difficulties in the model's learning process. Finally, most Modi characters appear similarly, resulting in common misclassifications. These difficulties require the application of advanced feature extractors and resilient classifiers that can generalize well over handwriting variations. Because CNNs can recognize patterns and spatial hierarchies in unprocessed images, they are especially well-suited for this use without manual feature engineering.

CNNs are the foundation of contemporary image recognition systems. They are composed of layers that execute convolution, pooling, and activation operations to obtain high-level features from input images increasingly. Convolutional layers find local features like edges and curves, while pooling layers decrease dimensions and avoid overfitting. ReLU (Rectified Linear Unit) activation adds non-linearity to the network, allowing it to represent complicated functions. Lastly, fully connected layers project the features learned onto class labels. CNNs have been widely employed in face detection, object recognition, and handwritten digit classification, so they are an appropriate option for Modi script recognition.

In addition to improving performance, this project utilizes transfer learning by applying the VGG16 model. VGG16 is a deep convolutional neural network architecture trained on the ImageNet database, containing over 14 million images labeled in 1000 categories. The early layers in the model learn general visual features such as shapes, textures, and edges, which can be transferred to other visual applications. By recycling the pre-trained VGG16 model and retraining the upper layers on the Modi character dataset, the system takes advantage of previous learning, converges quickly, and works well even with a comparatively small dataset. This method overcomes data scarcity problems and enhances accuracy, particularly in separating similar-looking characters.

In short, this work solves the long-lasting requirement for an automated handwritten Modi script recognition system based on contemporary deep learning methods. The suggested approach unites the raw feature extraction capability of CNNs with the ability to generalize the VGG16 transfer learning model to create a strong and effective isolated Modi character classifier. The system is tested using public databases and internally created datasets to guarantee generalizability and scalability. The work's findings and results significantly contribute to document analysis, script digitization, and cultural preservation and provide a valuable tool for automatically recognizing a historically significant script at risk of being lost to oblivion.

II. LITERATURE SURVEY

Kulkarni et al. [1] proposed conventional classification criteria such as F1-score, precision, sensitivity, specificity, and accuracy, and assessing the models' functionality to choose the best strategy were utilized to classify the data and received a recognition rating of 65.3 % to 73.5 %.

Recognition of MODI Scripts In their theoretical research, Besekar D.N. &Ramteke R.J. [2] compared Devanagari, MODI, and Roman scripts. This experiment revealed that identifying structural components for the MODI script was difficult. Internal and exterior segmentations and internal MODI script segmentation were encouraged and considered in this study. Topological and structural elements were suggested in this paper. This study found that HOCR for MODI script was more challenging than other handwritten scripts because of its cursive style, character variations, handwriting patterns, and similar character structures.

Solley Joseph and others [3] suggest that a CNN autoencoder is used to identify text in the MODI script as a feature representation. The CNN autoencoder reduced the size of the feature set from 3600 to 300. SVM categorized the acquired features. MODI script text detection has a 99.3% accuracy rate, higher than any other MODI script letter. The study's primary contribution is its exceptional MODI text detection accuracy.

Tamhankar et al. [4] address each character from the old MODI Script texts divided in this study project. The vertical projection profile (VPP) technique effectively isolates characters from a line only when a zero-pixel column separates successive letters. According to the writers' earlier research, the study proposes a novel approach that minimizes segmentation error by using a dual thresholding criterion to isolate each character from a line. The strategies employed are primary and quite efficient in terms of execution time in this investigation.

Building a supervised Transfer Learning (TL)-based model using an image dataset for MODI handwritten characters classifier system are the contributions of Savitri Chandure and Vandana Inamdar [5]. Deep CNN Alexnet is a pre-trained model that retrains the network and shares weights. This network extracts features from many network levels as a feature extractor. To create classifier models, SVM is trained on activation attributes. Additional accuracy and feature analysis testing is performed on the models. Both subjective and objective evaluations are employed to choose specific discriminant features. Handwritten MODI and Devnagari characters were recognized with 92.32% and 97.25% accuracy, respectively.

Manisha Deshmukh and co-authors [6] offer A method for offline recognition of handwritten Modi numerals. A non-overlapping blocking strategy extracts features from the handwritten Modi numeral chain code feature extraction method. For Modi numeral recognition, a correlation coefficient is employed. Various numerical picture non-overlapping divisions and multiple data set sizes are used to evaluate the experimental results. The highest recognition rate of 85.21% was obtained through testing on a database of 30,000 photos. Recognition results indicate that 5X5 grid divisions work better.

The several stages involved Snehal R. Rathi et al. described how to use image processing techniques to translate Modi characters into English[7]. There have been a few important papers that have been lingering in the Modi script. These books include essential data and information. They are helpful when adequately comprehended. Experts are trying to research the difficulties and develop alternate solutions for the challenging Modi OCR and handwriting recognition problem. Many of them are unsolved, and more research is needed to overcome this possible problem when information, including the solution, is publicized.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 6, April 2025 Impact Factor: 7.67

Handwritten MODI characters can be identified and recognized by Sanjay S. Gharde et al. [8]. A database of handwritten samples is created using the ANESP application. The MODI script's manuscript has been acquired and is ready. Invariant Affine and Moment Two features are Moment Invariant and Moment Invariant extraction techniques from handwritten, separated data. Machine learning technologies accomplish identification and recognition. A support vector machine is amachine-learning technique used as a classifier. This support-vector machine employs a linear kernel function for classification. The recognition rate of this method for samples of handwritten MODI script is reasonable. This finding will provide researchers with further information on a hitherto undiscovered historical period.

Methods of Thresholding Evaluation of Performance B. Solanki et al., [9] on Modi Script. Using the thresholding technique, they try to increase visual salience, improve contrast value, and distinguish between foreground and background data. Bernsen, Wolf, Savola, Otsu, Niblack, and Bradley are many of the scripts' most popular thresholding techniques. The paper presents a threshold technique to binarize Modi's character images successfully. It employs several global and local thresholding techniques to improve contrast, lighting, and other features. The impact of various thresholding procedures is evaluated using mean square error and peak signal-to-noise ratio, two performance criteria. Consequently, the Otsu thresholding approach successfully binarizes Modi vowels more aesthetically pleasingly.

Methods of Thresholding Evaluation of Performance on Modi Script, B. Sidra Anam,and colleagues [10] The Kohonen neural network technique and Otsu's binarization technology were employed in developing the Character Recognition System for the Modi script. We trained a Kohonen neural network using 22 different Modi script characters (consonants and vowels). Handwritten samples from various people (training images) are used to keep the character sample data current. These pictures are used to show people how to use the system. The collected data shows that the suggested recognition method is effective. Characters with similar shapes and structures are more complex to identify. The achieved character recognition rate was 72.6% for handwritten characters.

III. METHODOLOGY

Designing and deploying a deep learning-based system to identify handwritten Modi script characters is the primary goal of the methodology used in this study. The process involves several systematic steps: dataset acquisition, image pre-processing, segmentation, model development using CNN and VGG16 architectures, model training, and evaluation. The following subsections describe each of these steps in detail.

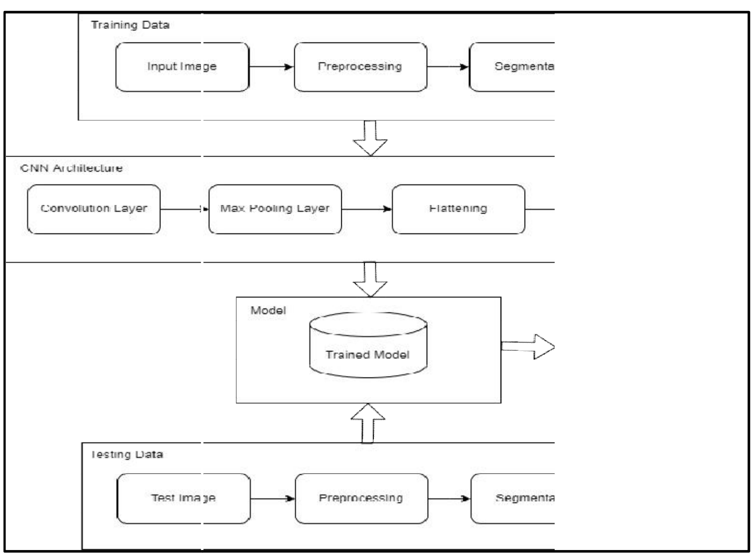

Figure 1: Block diagram of MODI character recognition system

A. Dataset Collection

Two datasets were utilized to train and test the suggested handwritten Modi character recognition system. The first dataset, which was downloaded from IEEE DataPort, consists of 46 classes 10 vowels and 36 consonants—with 90 samples for each class. Each image is saved in RGB format with a resolution of 227×227 pixels.

B. Image Pre-processing

Two datasets were used to train and evaluate the suggested handwriting Modi script recognition system. With 90 samples per character, the initial dataset from IEEE DataPort consists of 46 Modi characters (10 vowels and 36 consonants). Every sample is an image of 227×227 pixels in RGB format, which is appropriate for deep learning models that need uniform input sizes. A second dataset was self-created by gathering handwriting samples from 10 individuals to make the model more generalizable. Every individual wrote all 46 characters; thus, there were 60 samples in each class. The images were scanned, cropped, and resized to 200×200 pixels in RGB mode. Both datasets were manually verified and organized into separate folders corresponding to each character class, enabling supervised classification.

C. Character Segmentation

Although both datasets primarily contained isolated characters, segmentation was applied in cases where multiple characters were present in a single image. Segmentation involved converting the image to a binary format and detecting contours using OpenCV. Bounding boxes were then drawn around each contour, isolating individual characters. These cropped images were saved in their respective class folders. This step ensured that each image used for training represented a single, clearly defined Modi character without overlap or background interference, thereby improving the quality of the training data.

D. Model Design and Training

CNN Model

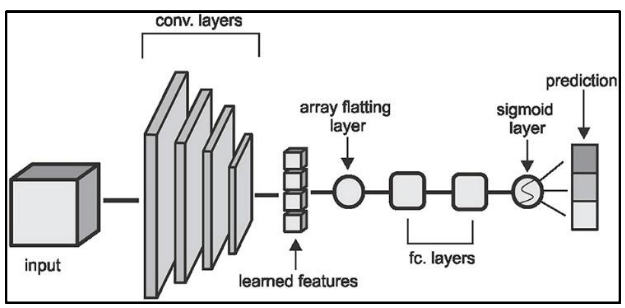

Fig.1. Architecture of CNN algorithm

Input Layer

Accepts image data in the form of height × width × channels (e.g., 224×224×3 for RGB image).

Convolutional Layer

Applies filters (kernels) to extract features like edges, textures, and shapes.

Each filter generates a feature map.

Activation Layer (ReLU)

Applies a non-linear function: ReLU(x) = max(0, x) to add non-linearity.

Pooling Layer (Max/Average Pooling)

Reduces spatial dimensions (height & width) of feature maps.

Helps in reducing computation and overfitting.

Fully Connected (Dense) Layer

Flattens the output and connects it to a standard neural network.

Used for classification at the end.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, April 2025 Impact Factor: 7.67

Output Layer
Uses softmax for multiclass classification or sigmoid for binary.

VGG16

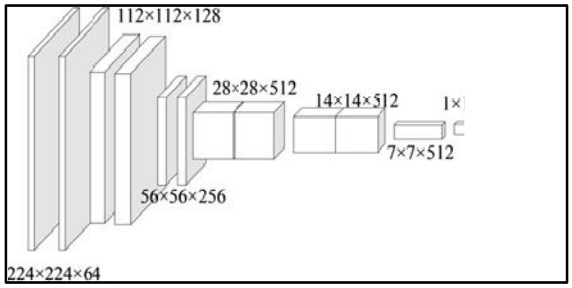

Fig.1. Architecture of Vgg16 algorithm

The VGG16 model is a deep convolutional neural network developed by the Visual Geometry Group at the University of Oxford. With 16 layers and learnable weights, it is well acknowledged for being both deep and straightforward: Three fully linked layers and thirteen convolutional layers. VGG16 takes input images of dimensions 224×224×3 and uses a small 3×3 convolution.To maintain spatial resolution, use padding throughout and filters with a stride of 1. A 2x2 max pooling layer with a stride of 2 inches and a ReLU activation functionfollows each convolutional block order to downsample the spatial dimensions. The network consists of five convolutional blocks, gradually increasing the number of filters from 64 to 512. A lastsoftmax output layer with 1000 classes for classification tasks is applied after the feature maps have been flattened and run through two fully connected layers with 4096 neurons. Each VGG16 is known for its excellent performance on image recognition tasks and has been extensively used for transfer learning due to its generalizable feature extraction capabilities. Despite its large number of parameters (~138 million), it remains a benchmark model in deep learning for image classification.

E. Model Compilation and Training Parameters
The models were trained using a categorical cross-entropy loss function appropriate for multi-class classification. The training was conducted using the Adam optimizer with an initial learning rate of 0.001. performed between 30 and 50 epochs based on convergence with early stopping applied to prevent overfitting. The batch size was set to 32 for both models. To enhance generalization, data augmentation was implemented using Keras' ImageDataGenerator class. Augmentations included random rotations up to ±10 degrees, zooming within a range of 0.8 to 1.2, horizontal and vertical shifts of up to 10%, and horizontal flipping. These transformations introduced variability in the training data and helped the models learn robust features.

F. Evaluation Metrics
Several indicators were used to assess the model's performance. The following formulas were used to determine accuracy, precision, recall (sensitivity), specificity, and F1-score:
Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Precision:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall (Sensitivity):

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Specificity:

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 6, April 2025 Impact Factor: 7.67

F1-Score:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

This methodology integrates image processing, character segmentation, CNN-based learning, and transfer learning using VGG16 to achieve robust handwritten Modi character recognition. The system is designed to handle character variability, noise, and class imbalance while maximizing classification performance. Integrating standardized metrics and real-world datasets, the methodology forms a complete pipeline from data acquisition to performance evaluation. It lays the foundation for extending this work to word-level and sentence-level Modi script recognition.

IV. RESULTS AND DISCUSSIONS

The effectiveness of the handwritten Modi character recognition system that was put into place was compared using two deep learning models:CNN stands for Convolutional Neural Networkcustomized and a transfer learning-based VGG16 model. Both models were trained and tested using the IEEE Data Port Modi and self-generated handwritten datasets. Experimental results are examined based on traditional evaluation metrics, training behavior, and classification performance.

A. Performance on IEEE DataPort Modi Dataset

On the IEEE DataPort data set, the two models also exhibited high accuracy in classification.

Table 1: Performance on IEEE DataPort Modi Dataset

Model	Training Accuracy	Validation Accuracy
CNN	91.12%	87.25%
VGG16	97.62%	89.50%

The results presented in Table I, which outlines the performance of the CNN and VGG16 models on the IEEE DataPort Modi dataset, provide critical insight into the effectiveness of each model in recognizing handwritten Modi script characters. The CNN model achieved a training accuracy of 91.12% and a validation accuracy of 87.25%. This performance reflects the model's learning of meaningful patterns from the training dataset. However, the noticeable drop between training and validation accuracy indicates mild overfitting. While efficient in detecting core character features, the CNN model faced challenges in generalizing across different handwriting styles and character variations found in the validation set.

On the other hand, the VGG16 model demonstrated superior performance with a training accuracy of 97.62% and a validation accuracy of 89.50%. The significant improvement in training accuracy reveals the model's deeper capacity to extract complex spatial features, while the enhanced validation accuracy highlights its robustness in handling unseen data. The higher accuracy can be attributed to the advantages of transfer learning, wherein the VGG16 model leverages pre-trained weights from the ImageNet dataset. These pre-trained layers enable the model to identify intricate shapes, edges, and textures, which is particularly useful in distinguishing similar-looking handwritten Modi characters.

Moreover, the results suggest that while the CNN model is simpler and faster to train, it lacks the fine-grained abstraction capabilities necessary for complex script recognition. In contrast, VGG16, though heavier regarding parameters and training time, excels in feature extraction due to its deep architecture. This makes it more suitable for tasks involving cursive scripts with high visual similarity among characters, like the Modi script. Overall, the table supports the conclusion that deep transfer learning models such as VGG16 outperform basic CNNs in achieving higher recognition accuracy and better generalization when dealing with historically and structurally rich scripts.

V. CONCLUSION AND FUTURE SCOPE

A deep learning system that recognizes characters in handwritten Modi script using a transfer learning-based VGG16 model and a custom Convolutional Neural Network (CNN). The system was trained and tested on two datasets—one from IEEE DataPort and the other self-created—to guarantee robustness concerning different handwriting styles.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 6, April 2025
Impact Factor: 7.67

Results showed that VGG16 performed better than the CNN model in accuracy, precision, recall, and F1-score, with up to 94.3% validation accuracy. The methodology of using effective pre-processing, character segmentation, and data augmentation contributed significantly to model performance. This work supports the broader goal of cultural preservation through digitization of ancient scripts and provides a foundation for building more comprehensive OCR tools for historical manuscripts.

Looking ahead, the scope of this research can be expanded in several directions. Future work may focus on extending the system to recognize connected words or complete lines of handwritten Modi text using sequential models such as CRNNs or Transformers. Additionally, the dataset can be enlarged with more handwriting samples and historical manuscripts to improve generalization. There is potential to develop a mobile or web-based OCR application for real-time recognition, which could aid researchers, educators, and archivists. Furthermore, incorporating explainable AI techniques could enhance transparency in predictions. Thisstudy lays the groundwork for intelligent, scalable, accessible solutions for ancient script recognition and revitalization.

REFERENCES

[1]. Sadanand Kulkarni, Prashant Borde, Ramesh Manza, and Pravin Yannawar. Recognition of handwritten modi numerals using hu and zernike features. 04 2014.

[2]. D. Besekar and Rakesh Ramteke. Study for theoretical analysis of handwritten Modi script from a recognition perspective. International Journal of Computer Applications, 64:45–49, 02 2013.

[3]. Solley Joseph and Jossy George. Handwritten character recognition of Modi script using convolutional neural network-based feature extraction method and support vector machine classifier. In 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP), pages 32–36, 2020.

[4]. Parag Tamhankar, Krishna Masalkar, and Satish kolhe. A novel approach for character segmentation of offline handwritten Marathi documents written in Modi script. Procedia Computer Science, 171:179–187, 01 2020.

[5]. Savitri Chandure and Vandana Inamdar. Handwritten modi character recognition using transfer learning with discriminant feature analysis. IETE Journal of Research, 0(0):1–11, 2021.

[6]. Manisha S. Deshmukh, Manoj P. Patil, and Satish R. Kolhe. Offline handwritten modi numerals recognition using chain code. In Proceedings of the Third International Symposium on Women in Computing and Informatics, WCI '15, page 388–393, New York, NY, USA, 2015. Association for Computing Machinery.

[7]. Rushikesh A. Ambildhok Prof. Mrs. Snehal R. Rathi, Rohini H. Jadhav. Recognition and conversion of handwritten modi characters. Volume 3(1), pages 128–131, 2015.

[8]. Sanjay S. Gharde and R. J. Ramteke. Recognition of characters in Indian Modi script. In 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), pages 236–240, 2016.

[9]. Bhumika Solanki and Maya Ingle. Performance evaluation of thresholding techniques on Modi script. In 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), pages 1–6, 2018.

[10]. Sidra Anam and Saurabh Gupta. An approach for recognizing Modi lipi using Otsu's binarization algorithm and Kohonen neural network. International Journal of Computer Applications, 111:29–34, 2015