The Ciência & Engenharia - Science & Engineering Journal
ISSN: 0103-944X
Volume 11 Issue 1, 2023
pp: 2986 - 2994

CNN Based Approach For Modi Script Character Recognition

***Ramiza Kureshi**
Dept. of Electronics and Telecommunication-Communication Network
Dr. D.Y. Patil Institute of Technology, Pimpri, Pune-18, India
*ramizakureshi@gmail.com
Pramodkumar H. Kulkarni
Dept. of Electronics and Telecommunication-Communication Network
Dr. D.Y. Patil Institute of Technology, Pimpri, Pune-18, India
pramod75kulkarni@gmail.com

Abstract—Character recognition is simple for humans, but developing software that is able to do so is a significant challenge. MODI is a language that has been neglected for a long time. We want to construct a character recognition system to read printed and handwritten MODI characters by scanning the papers, and our plan is to use this method. There is not much of a difference in the level of recognition accuracy achieved regardless of the quality of these articles. The established MODI Script Character Recognizer System (MSCR), which makes use of convolutional neural networks, presents challenges for the purpose of keeping a realistic character recognition system in place (CNN). The performance of this approach is evaluated by utilizing the dataset for Modi character recognition provided by IEEE Dataports. The CNN approach is utilized in order to give labels to each of the different Modi characters.
Keywords—Character recognition, Modi script, segmentation, recognition, classification

Introduction

The "MODI" alphabet is quite contemporary when compared to other classical Indian languages. Numerous theories have been put out regarding the origin of the MODI script, a cursive writing system used in the language "Marathi," which is the main language spoken in the state of Maharashtra in western India. There is a theory that "Hemadpant" or "Hemadri" invented MODI in the 12th century. Documents are by far the most common kind of knowledge that can be found in human societies. As a consequence of this, the study of pattern recognition and image processing should place significant emphasis on document picture analysis. The speed of life in today's society is frenetic, and it is dominated by mechanical processes. This is all due to the fact that people like to perform tasks in the quickest and most efficient manner possible. As a direct consequence of this, our work will be simpler and completed in less time as we continue to implement more automation. In our fast-paced world, digitalization is the next big thing that's going to happen. Because we are now living in the computer era, we want all of the information that is easily accessible to be converted into digital format and stored on machines that have improved processing speeds. However, moving knowledge from the physical world into the digital realm is challenging because we have to teach the computer, particularly about actual facts in the physical world.

The "paperless workplace" is gaining importance today, leading to increased online communication and document storage. To speed up additions, searches, and changes and increase the longevity of such data, documents and files previously preserved on paper are now being converted into digital format. Therefore, the need for software that gathers, recognizes and extracts data from physical documents for later retrieval is very significant. Document processing requires text processing using an optical character recognition system (OCR).

MODI characters will be pre-processed, segmented, and recognized in this project. There are many issues in segmenting and identifying the MODI character from a handwritten script. As a result, this technique effectively trains the CNN algorithm to recognize the pattern using the manually typed MODI characters as input.

The state-of-the-art methodologies for the analysis of MODI character identification are described in part II of this study, and Section III explains the CNN-based MODI character recognition system. Section IV discusses the findings from the qualitative and quantitative analyses, and Section V concludes the proposed work.

Literature Survey

A chain coding and centroid-based image recognition model for vowels in MODI script was introduced by Kulkarni et al. [1]. In this study, the noise was reduced using the median filter; binarization was achieved using the global threshold, and size normalization and boundary breaks were avoided using flood fills. The data were classified using two feed-forward neural networks and SVM. This approach achieved 65.3% to 73.5% recognition accuracy.

Utilizing structural similarities, Ramteke A. S. and Katkar G. S. [2] were able to determine the identities of MODI characters. Image quality measures were utilised in conjunction with a structural similarity approach to do the evaluation of the pictures' overall quality. The recognition rate ranged from 91% to 97% thanks to the classification methods of measured structural similarity (SSIM), key-value network (KNN), and backpropagation neural network (NN).

In order to recognise offline handwritten numbers, Besekar D.N. and Ramteke R.J. [3] utilised an approach that was based on zones. In this investigation, the median filter was utilised to lessen the amount of background noise; the global threshold was utilised to accomplish binarization; boundaries were avoided by making use of flood fill; and size normalisation was completed. The feature set was constructed from data pertaining to four 15x15 grid cells, including their polar coordinates, variance, theta angle, and Rh distance. Using a variance table, the authors of this study were able to attain a recognition rate of 93.5%.

Theoretical assessment of the Modified Isolation Protocol The script recognition was done by D.N. Besekar and R.J. Ramteke [4]. This research compared and contrasted the Roman, Devanagari, and MODI scripts with regard to their underlying structural properties. It was found after some research that it was challenging to extract the MODI script. In the course of this research, both internal and exterior segmentations were dissected, and several recommendations were offered for each. MODI scripts frequently have internal segmentation of their own. In this research effort, both structural and topological issues were investigated. According to the findings of this study, HOCR for MODI script proved challenging because to the cursive aspect of the script, the differences in characters, the handwriting patterns, and the comparable letter structures.

A supervised Transfer Learning based classification algorithm was presented by Savitri Chandure et al. [5] in order to recognise MODI handwritten characters. The features are extracted with the assistance of a deep convolutional neural network, and the classification of those features is handled by an SVM classifier. These models' feature analysis and precision are examined in more detail. Subjective and objective assessments are used to select deep discriminant characteristics. The proposed system had a handwritten character recognition accuracy of 92.32%.

The use of a CNN autoencoder that was constructed as a feature representation and presented by Solley Joseph et al.[6] is recommended for text detection in the MODI script. The CNN autoencoder was responsible for the reduction in the size of the feature set from 3,600 to 300. The SVM was utilised in order to classify the obtained characteristics. The accuracy rate for detecting MODI script text is 99.3%, more significant than that for detecting MODI script letters. Having extremely high MODI text detection accuracy is the essential contribution of this study.

Manisha Deshmukh and her colleagues invented a method for reading offline handwritten Modi numerals [7]. To extract the distinctive features of handwritten numbers, a non-overlapping blocking methodology is used with the Modi numeral chain code feature extraction method. Recognizing Modi numerals requires a correlation coefficient. Several dataset sizes and numerical image non-overlapping divisions are employed to evaluate the experimental results. During testing on a database consisting of 30,000 photographs, a recognition rate of up to 85.21% was attained at its highest potential. According to the findings of the recognition research, the most effective grid divisions are those that are 5x5.

In their performance assessment of thresholding algorithms on Modi Script, B. Solanki et al. [8] disclose their findings. They want to boost the image's prominence and contrast value, and they may do so using the thresholding method. They also wish to distinguish between background information and foreground information. Among the numerous diverse scripts, the Bernsen, Wolf, Sauvola, Otsu, Niblack, and Bradley thresholding methods are among the most well-known. In order to effectively binarize images of Modi characters, a threshold approach is provided in this study. It uses various global and local thresholding techniques to offer advantages, including enhanced illumination and increased contrast. As performance metrics, the mean square error and peak signal-to-noise ratio are utilised in order to assess and compare the results obtained from using a variety of thresholding processes. The Otsu thresholding technique is responsible for the more agreeable binarization of MODI vowels as a consequence of this.

Using image processing techniques, Snehal R. Rathi et al. [9] explained the numerous steps necessary in translating Modi characters into English. Numerous significant works written in the "Modi" language are still unpublished. These papers include essential facts and information. OCR and handwriting recognition are significant tasks, and experts are working hard to identify the issues and create workarounds. Numerous issues are still open; thus, more study is required to lessen the severity of potential difficulties whenever remedies are available.

Handwritten MODI character recognition and identification is demonstrated by Sanjay S. Gharde and colleagues [10]. In order to construct a database of handwritten examples, the programme ANESP is utilised. It has been determined that the pre-processed MODI script document may be obtained. The Present and the Unchangeable Affine Two methods, known as Moment Invariant and Moment Invariant, may be used to extract characteristics from handwriting samples that have been previously separated. The application of machine learning technology helps procedures like identification and recognition. One of the machine learning techniques that may be applied as a classifier is the support vector machine. The categorization is carried out via a linear kernel function in this support-vector machine. For MODI script examples, this system has the greatest recognition rate. This research will allow researchers to understand more about a period of history that was previously unknown. According to the literature review, the handwritten dataset of MODI characters and digits is not readily available.

Proposed System

The block diagram of the proposed system is shown in Fig. 1. The block diagram visually divides the training and testing phases.

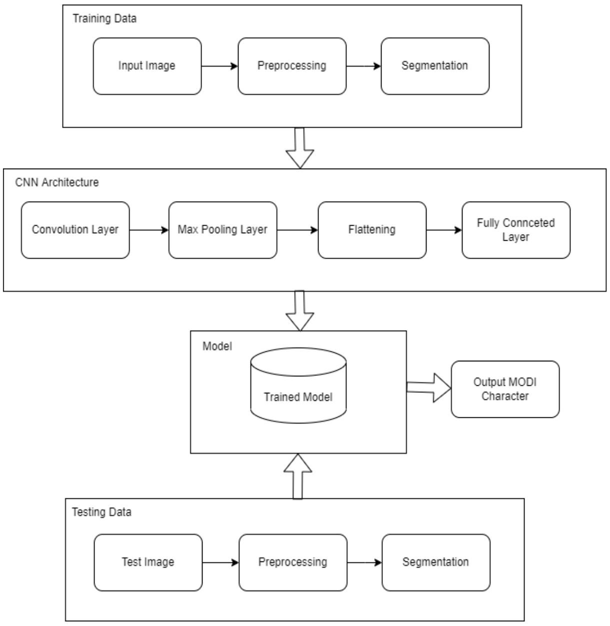


Fig. 1. Block diagram of MODI character recognition system

A. MODI character Dataset

The proposed system uses handwritten MODI characters. Multiple authors write the numeric datasets and the handwritten MODI characters in various writing styles. Because various writers employ diverse writing styles and varying lighting situations when taking photographs, the method is more accurate. Consequently, the dataset is produced using a variety of types and lighting circumstances. There are training and validation sets for the full dataset. 20% of the dataset is used for testing, while the remaining 80% is used for training. An IEEE Datport collection of handwritten MODI characters is used in the proposed method. This dataset contains 46 MODI characters, ten vowels, and 36 consonants. For every character, there are 90 samples. A PNG file with the image's dimensions of 227x227x3 is used. Ninety different people submitted four thousand one hundred forty-character samples. The handwritten example of the Modi character dataset is shown in Fig. 2.

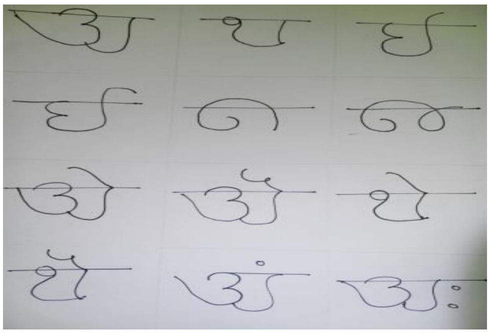


Fig. 2. Handwritten sample of the A4 sheet

B. Pre-processing

Every image in the dataset must be resized to the same size because the collected images may be of varying sizes. This system uses 256x256 downsampling of the images. Gaussian noise, in addition to other noises like salt and pepper, also affects the images. The median filter eliminates the salt-and-pepper noise.

C. Segmentation

The letters on the acquired picture are separated from the backdrop with the use of a tool called thresholding. Through the use of the thresholding process, the picture is transformed into binary. The value 100 is selected to be used as the threshold for this procedure. Clipping occurs, and the character or the region of interest that was selected is then saved for subsequent processing.

D. Training and Testing

a. CNN

CNN is a great example of an outstanding algorithm for classification procedures. There are dense, flattening, pooling, and convolutional layers present in this feed-forward neural network. The filter and the kernels are used in the processing of the picture. Before beginning the training process, it is necessary to have a solid understanding of the CNN fundamentals, which are illustrated in Figure 3. CNNs are a specific type of neural network that are extremely effective at identifying and categorising pictures. One kind of these networks is known as a feed-forward neural network, sometimes abbreviated as CNN.

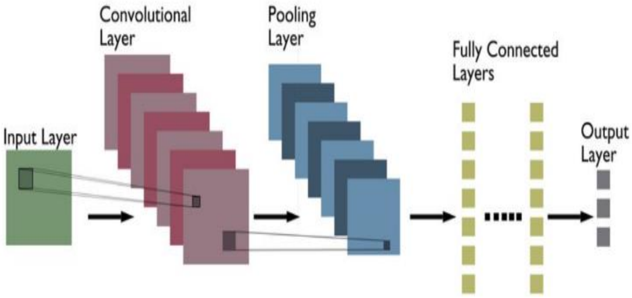


Fig. 3. Block Diagram of CNN algorithm

- Convolutional Layer

The most crucial component of a convolutional neural network is the convolutional layer (CL). CL's main objective is to extract properties from the data that it is given. In convolution, information is learned from the input image by focusing on a very small kernel of that image. This helps to maintain the spatial link between pixels. A network of learning neurons is utilised in order to obscure the original image.

$$G[m,n] = (f * h)[m,n] = \sum_j \sum_k h[j,k] f[m-j, n-k] \qquad (1)$$

G[m,n] is the convolution result, f is the input image, and h is the feature detector kernel.

- ReLU Layer

In non-linear processes, rectified linear units, often known as ReLUs, are utilised. All of the negative feature map values are being manually reset to zero, one pixel at a time. In order to appreciate how the ReLU functions operate, we will assume that the input to the neuron is x and that the rectifier is already provided.

$$f(x) = \max(0, x) \qquad (2)$$

- Pooling Layer

The pooling layer simplifies each activation map while carefully preserving the data that is most important to the analysis. Using the photographs that have been supplied, a number of rectangles that do not overlap each other are made. In order to downsample each zone, such as the average or the maximum, a method that is not linear is applied. This layer, which is typically located in the middle of two CLs, not only prevents distortion and translation but also speeds up the convergence and generalisation processes.

- Flattening Layer

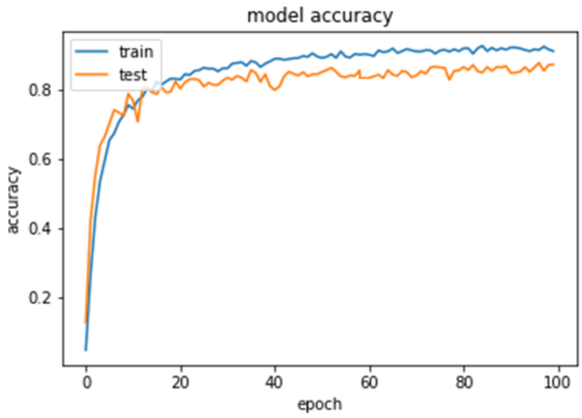
Using CNN, high-resolution data is effectively resolved into representations of things. To "understand" the findings and ultimately offer a classification result, it is possible to think of the fully connected layer as adding a conventional classifier to the network's information-rich output. The CNN dimension's output must be flattened to link this fully connected layer to the network.

- Fully Connected Layer

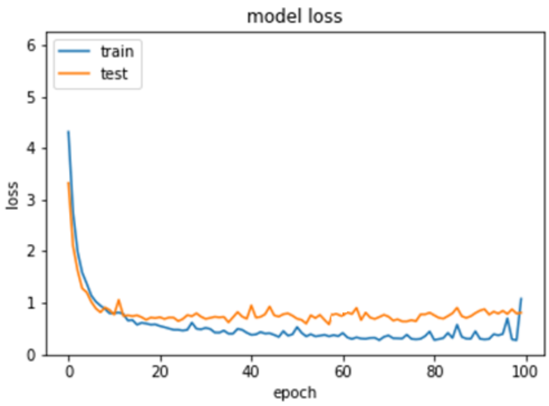
Using these attributes, the FCL divides the input image into several groups depending on the training dataset. The final pooling layer, often known as the FCL, is responsible for providing the Softmax activation function classifier with features. The sum of the probability produced by the FCL is 1. They were utilising Softmax, whose activation function guarantees that this will occur. Every real-valued score may be simplified into a vector of summable values ranging from zero to one with the use of the Softmax function.

Results

In this study, a convolutional neural network method for MODI character identification is presented. The dataset of handwritten MODI characters is taken from IEEE DataPORT. This dataset has 46 MODI characters in it. The training comprises 80% of the dataset, and testing comprises 20%. Fig. 4 displays the accuracy and loss of the proposed system during training and validation.



(a)



(b)

Fig. 4. Training progress graph of CNN algorithm in terms of (a) Accuracy, (b) Loss

The CNN algorithm is utilised in the process of testing and training the MODI character. On the Google Colab platform, the data training is carried out with the help of the CNN algorithm. When presenting the outcomes of the technique that was suggested, both qualitative and quantitative analysis were utilised. The accuracy and loss of the suggested system, which makes use of the CNN algorithm, are depicted in Figures 4 and 5, respectively.

The performance of the system is tabulated in Table I

Table: Quantitative Analysis

Algorithm	***Training***		***Validation***		***Execution Time (Sec)***
	Accuracy(%)	***Loss (%)***	***Accuracy (%)***	***Loss (%)***	
CNN	91.12	10.773	87.25	8.074	3697.23

The training accuracy and loss for the CNN algorithm are 91.12% and 1.0773, respectively. Validation accuracy was 87.25 percent with a validation loss of 0.8074. 3697 seconds were spent on the training execution. The qualitative study of the suggested system is shown in Fig. 6.

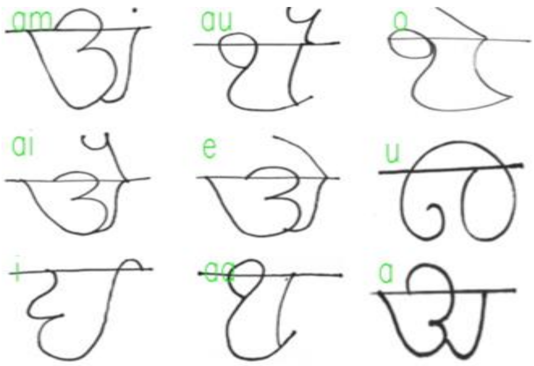


Fig.6. Qualitative analysis of the proposed system

According to qualitative analysis, the CNN algorithm can precisely and accurately recognize the MODI character.

Conclusion

This technique proposes an algorithm for the recognition of MODI characters based on a convolutional neural network. The handwritten MODI character and the numeric datasets are both authored by many writers, each of whom has a unique writing style. The dataset is divided into training, which accounts for 80%, and testing, which accounts for 20%. The CNN method is used to train and test the MODI character inside this system, and accuracy and loss criteria are used to evaluate performance. According to the technique that was presented, CNN was able to attain a training accuracy of 91.12% while suffering a training loss of 10.773%. Validation was accurate 87.25% of the time, with a validation loss of 8.074%.

In the future, it will be able to hyper-tune the settings of the several transfer learning algorithms to increase the system's accuracy. The study has demonstrated the availability of a constrained standard dataset for MODI characters and integers. As a result, a MODI character and number dataset that is sizable in size has to be constructed.

References

[1] Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2015a), "Review on recent advances in automatic handwritten MODI script recognition," International Journal of Computer Applications, 115(19), 975-8887.

[2] Ramteke A.S., Katkar G.S., 2012, "Recognition of Offline MODI Script," International Journal of Research in Engineering, IT and Social

[3] Besekar D.N., Ramteke R.J., 2012, "Feature Extraction Algorithm for Handwritten Numerals Recognition of MODI Script using Zoning-based Approach," International Journal of Systems, Algorithms & Applications, Volume 2, Issue ICRASE12, ISSN 2277 2677, pp. 1-4.

[4] Besekar D.N., Ramteke R.J, 2013, "Study for Theoretical Analysis of Handwritten MODI Script – A Recognition Perspective," International Journal of Computer Applications, vol. 64, no. 3, ISSN 0975-8887, pp. 45-49.

[5] S. Chandure and V. Inamdar, "Handwritten Modi character recognition using transfer learning with discriminant feature analysis," IETE Journal of Research, pp. 1–11, 2021.

[6] S. Joseph and J. George, "Handwritten Character Recognition of MODI Script using Convolutional Neural Network Based Feature Extraction Method and Support Vector Machine

Classifier," 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP), 2020, pp. 32-36.

[7] Manisha s. Deshmukh, Manoj Patil, and Satish kolhe," Offline handwritten Modi numerals recognition using chain code." WCI 15.

[8] 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.

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

[10] 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, 20