Grenze International Journal of Engineering and Technology, June Issue

Handwritten MODI Digit Recognition using Convolutional Neural Network

Lekha Gadpade[1], Aruna Chamatkar[2] and Amina Vali[3]
[1]Department of Electronics and Comp. Science, R.T.M.N.U, Nagpur, India
Email: lekha.thawkar@gmail.com
[2]Computer Science and MCA Department, Kamla Nehru Mahavidyalaya, Nagpur, India
[3]Sadabai Raisoni Women's College, Nagpur, India

***Abstract*— The writing system that was extensively used in Maharashtra for administrative purpose during the medieval period was the MODI script. The following features of MODI script are (i) high intra- class variability, (ii) Similarity of strokes (iii) No standardized style of writing. Due to these features, automatic recognition of the handwritten script is difficult. This paper presents a study of Convolutional Neural Network (CNN) based approach for recognition of the handwritten MODI script. A diverse dataset of 101,095 handwritten digit images belonging to ten classes was used for this study. The CNN model was designed to learn the discriminative, spatial features directly from the raw grayscale images without the extraction of manual features. An overall accuracy of 88% with a strong performance class-wise across most digits was achieved. Confusion matrices and misclassified sample visualization was used for detailed analysis, this highlighted the challenges that came up with digits that were visually similar. The proposed approach establishes a strong baseline for MODI digit recognition and contributes toward the preservation and digitization of historical Indic script.**

***Index Terms*— Pattern Recognition, Handwritten Digit Recognition, Convolutional Neural Network.**

I. INTRODUCTION

During the era of the Maratha Empire, MODI script, an ancient Indian script was used for writing Marathi language. The script remained underexplored in the domain of Optical Character Recognition. Recognition of digits of MODI script is difficult due to variations in styles, incomplete strokes, and the similarities in shapes of digits.

Scripts such as Latin, Arabic and Devanagari have been studied extensively for the recognition of handwritten digits on the other hand the research for recognition of MODI script is limited. Handcrafted features such as Histogram of Oriented Gradients (HOG), Zoning and Structural descriptors used by traditional OCR methods, often fail to capture the variations in the handwritten script.

Superior performance in the recognition of visual patterns tasks has been seen with the recent advancement in the field of deep learning, Particularly Convolutional Neural Networks (CNN). Hierarchical representations are automatically learned by CNN, from the data given thereby making them a good choice for scripts such as MODI which are complex.

Grenze ID: 01.GIJET.12.2.406_22

A CNN based handwritten MODI digit recognition system, along with an in-depth performance analysis that uses quantitative evaluation has been proposed in this paper. The ancient Indian scripts have been studied for character recognition by using traditional machine learning as well as deep learning methods. Moderate success has been achieved using a feature based methods like HOG, SIFT and statistical descriptors in a combination with the classifiers like Random Forest and support vector machine (SVM). The drawback of these methods is that they are sensitive to variations in handwriting and require careful feature engineering.
Handwritten character recognition using CNN-based methods have gained prominence in the recent years. This is because of their ability to automatically learn spatial features, this has been proved by the results of the studies on Urdu, Bangla and Devanagari scripts using CNN architecture. The research on digit recognition of MODI, script till date remains scarce thereby highlighting the importance of deep learning – based solutions that are robust.

II. Literature Review

A number of studies exploring different approaches for recognition of MODI script, using both traditional Machine language methods as well as modern deep learning architecture have been carried out by various researchers.
"Reference [1]" presents a review on several traditional machine learning algorithms for MODI Lipi character recognition, with SVM Method showing superior accuracy and overall performance.
Several recognition techniques for character recognition were compared according to "Ref. [2]", with the conclusion that convolutional Neural Network (CNN) gave the best accuracy among all the models that were evaluated by them.
According to "reference [3]", researcher have been reviewed the various approaches for hand-written digit recognition using machine learning and deep learning. These methods have used the benchmark datasets like MNIST, Classical classifiers such as SVM, KNN, Decision trees and reinforcement learning methods along with modern CNN based architectures have been analysed, The study has shown that deep learning model like optimised CNN have outperformed the traditional methods with an accuracy of 99.9% as the authors came to the conclusion that the most suitable approach for digit recognition was CNN based as it felicitated automatic feature extraction and also proved robust where handwriting variations were concern.
Digitisation of ancient palm leaf manuscripts in Tamil by means of an intelligent four stage pipeline that consisted pre-processing followed by segmentation and finally CNN based recognition was proposed by "Ref. [4]".
"Reference [5]" addressed a major research gap in OCR for languages like Kurdish which had low resources for data sets. They introduced a dataset for Kurdish language, Kurdset the first large scale dataset that was publicly available for recognition of handwritten characters. Kurdset contains over 45,000 samples that were collected from diverse writers. Multiple CNN architectures including custom CNN, Inception V3, X-ception and Dense Net 121 were evaluated using transfer learning. The highest accuracy 99.80% was achieved by Dense Net 121. This highlighted the fact that densely connected networks for scripts that had characters which were visually similar was extremely effective. Kurdset was set a benchmark for future Kurdish OCR research.
"Reference [6]" used CNN model along with extensive data augmentation. They achieved an accuracy of 99.78% with on-the-fly augmentation and 99.47% with offline augmentation.
A duel threshold character segmentation method to address the continuous writing characteristics of the script was developed by "Ref. [7]".This method gave a segmentation accuracy of 67%.
According to "reference [8]", MODI image was enhanced using an Ostu based threshold approach that was evaluated through Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE) metrics.
A feature recognition system that was hybrid was developed by "Ref. [9]". This achieved an average accuracy of 91.20% and a maximum accuracy of 99.10%, showing robustness to variations in writing.
"Reference [10]" applied Zone based extraction for MODI numerals that were handwritten. The pixel values were converted into polar coordinates the accuracy obtained was 93.5%.
"Reference [11]" created a hybrid (neural and non-neural) approach to classifying MODI and Devanagari vowels utilizing Chain Code histograms and Intersection algorithms. Classifiers including BN, k-NN, and SVM were used. While MODI recognition rates were lower than Devanagari due to misclassification of similarly shaped vowels, the study emphasized the importance of different methodologies and larger datasets for improving MODI character recognition.
Table 1 shows the representation of feature extraction methods to recognize MODI script by various researchers.

TABLE 1. TABULAR REPRESENTATION OF VARIOUS FEATURE EXTRACTION METHODS TO RECOGNIZE MODI SCRIPT

Sr. No	Author	Research year	Feature Extraction / Other Methodology	Classifier	Dataset	Accuracy in %
1	Varpe K, Sakhare S “Ref. [12]”	2023	CNN	CNN, KNN, Deep Learning	-	99.78
2	Altaf Ramjon, et.al. ”Ref. [13]”	2023	CNN	CNN	-	98.7
3	Vishal Pawar, et.al ”Ref. [14]”	2022	CNN	CNN	-	91.62
4	Joseph S., et.al. “Ref. [15]”	2021	Wavelet Transform-based Character Recognition	Decision Tree Classifier-Daubechies, Haar, Symlet Wavelet	4600 MODI characters(3220) training, 1380(testing)	91.75
5	Joseph S., et.al. “Ref. [6]”	2021	CNN- based Character Recognition	CNN-Argumentation(On-the-Fly, Offline)	-	99.78(on-the-fly), 99.47(Offline).
6	Joseph S., et.al. “Ref. [16]”	2021	CNN Auto encoder-based Character Recognition	KNN,SVM	-	KNN- 99.4, SVM-99.3
7.	Tamhankar P.A.,et.al. “Ref. [17]”	2021	Character Segmentation and Recognition	Vertical Projection Profile (VPP)	-	67
8.	Gadpade L., et.al. “Ref[1] “	2025	Traditional Machine learning Methods	SVM, k-NN, Decision Tree, Random Forest, Naïve Bayes(NB)	10110 MODI Characters	SVM-96.4, k-NN-92.2, Decision Tree-91.3, Random Forest-92.1, Naïve Bayes(NB)-92.1

III. PROPOSED METHODOLOGY

The proposed model is divided into pre-processing, classification algorithm, and comparison performance evaluation as shown in Fig. 1.

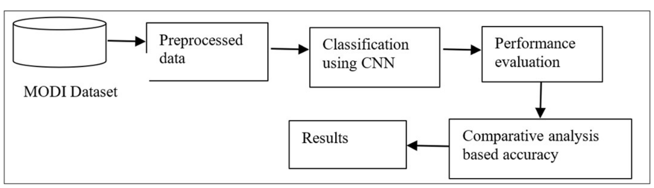

Figure 1. Proposed model of CNN

A. Dataset Description

The dataset is available on public domain by “Ref. [18]”. A dataset consisting of 101,095 images of handwritten MODI digits that was divided into ten classes representing the digits from 0 to 9 as shown in Fig.2, was used in this study. An 80: 20 ratio spilt was used to divide the dataset into training and validation sets respectively, giving 80,876 training images and 20,219 validation images.Fig.3 shows the sample images of class six. In order to ensure uniform input dimension the images were converted to grayscale images and resized to 64x64 pixel. The CNN learned the features directly from the pixel intensities as no handcrafted extraction of features was performed.

Figure 2. MODI Digit Dataset

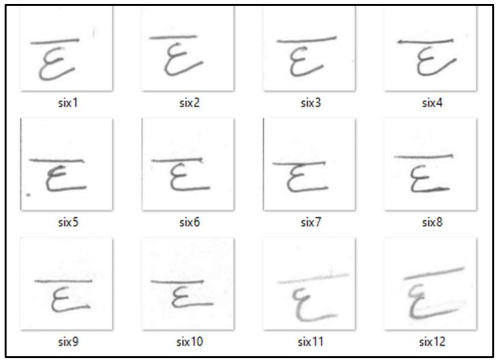

Figure 3. Sample Images for Class Six

B. Classification Algorithm

A three convolutional block CNN architecture with each block including and convolutional layer, normalisation and max- pooling has been proposed in this study. The convolution operation for k^{th} feature map is defined in equation (1).

$$\mathcal{Y}_{i,j}^{(k)} = \sum_m \sum_n x_{i+m+n} \cdot w_{m,n}^{(k)} + b^{(k)} \qquad (1)$$

Where,
x represents the input image,
$^{(k)}$ denotes the convolutional kernel associated with the k^{th} feature map,
$b^{(k)}$ is the bias term, and
$^{(k)}$ is the resulting feature map.
Then the Rectified Linear Unit (ReLU) activation function is applied as shown in equation (2):

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

For Max-pooling reduces spatial dimensions as shown in equation (3) and equation (4): (2)

$$Y = \max(x_1, x_2, \ldots \ldots, x_n) \qquad (3)$$

$$P(\mathcal{Y} = i|x) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_i}} \qquad (4)$$

Where, C is the number of classes and z_i is the logit for class i.

C. Training Strategy

Equation (5) is used for Adam optimiser and categorised the cross entropy loss used to brain the network:

$$L = -\sum_{i=1}^{C} \mathcal{Y}_i log\,(\hat{y}_i) \qquad (5)$$

Where, y_i denotes the true label, and $\hat{y}_i$ represents the predicted probability for the i^{th} class. In order to prevent over lifting a stabilized the training, early stopping and rate reduction techniques were used. *C.*

D. Comparison of Accuracy and Loss Analysis Using CNN

In this study, Accuracy is measure the propagation of correctly classified instances out of the total number of predictions as given in equation (6).

$$Accuracy = \frac{(TruePositive + TrueNegative)}{(TP + TN + FP + FN)} \qquad (6)$$

Where, *TP* -True Positive, correctly predicted positive class instance.
TN -True Negative, correctly predicted negative class instance.
FP -False Positive, incorrectly predicted as positive as positive when it is actually negative.
FN -False Negative, incorrectly predicted as negative when it is actually positive.

E. Performance Evaluation

For performance evaluation in this study, confusion matrix used for analysing how well the classification model is performing for each class. It helps identify which classes are correctly predicted and which classes are being misclassified, provide deeper insight beyond overall accuracy.

IV. EXPERIMENTAL RESULTS

A. Accuracy and Loss Analysis

Rapid conversions of the model was demonstrated by the training and validation curves. The performance of multiclass CNN of MODI digits recognition shown in Fig. 4 and Fig. 5. The accuracies obtained were approximately 98.99% and 88% for training and validation respectively thereby indicating that generalisation was collective a stable optimization behaviour was confirmed by the loss curves that were obtained.

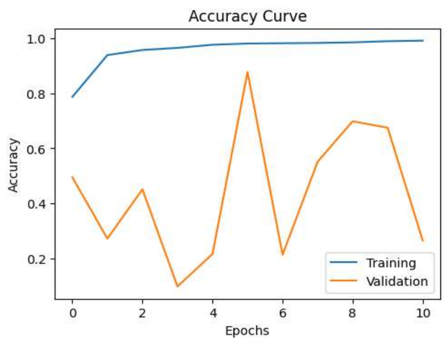

Figure 4. Accuracy of CNN Model

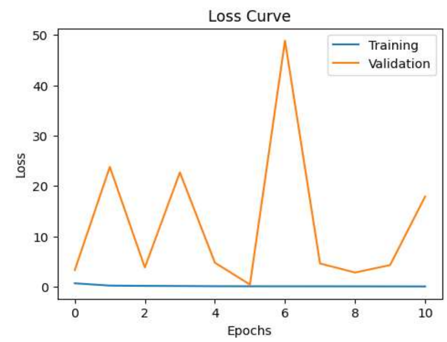

Figure 5. Loss of CNN Model

B. Confusion Matrix Analysis

A strong diagonal dominance that indicated that the classification for most of the digits was appropriate was revealed by the confusion matrix. Class wise prediction distribution that provided insight into systematic misclassification pattern was highlighted by a normalised confusion matrix as shown in Fig. 6 and Fig. 7.

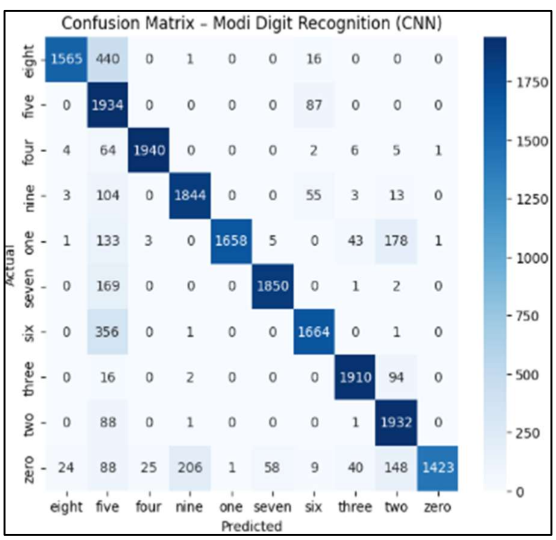

Figure 6. Confusion Matrix –MODI digits using CNN

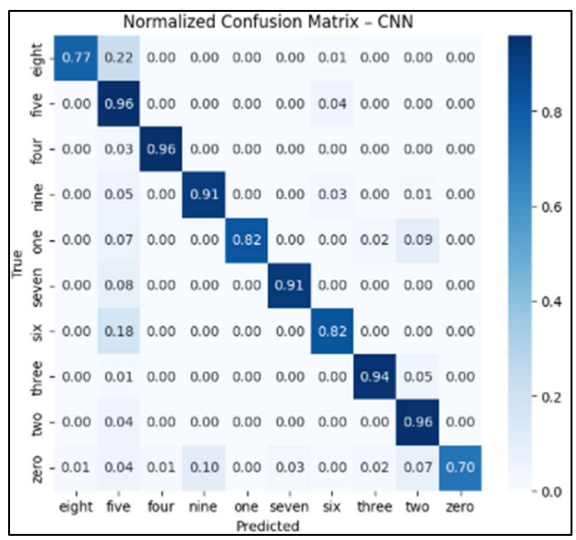

Figure 7. Normalized Confusion Matrix

C. Class wise recognition accuracy

A recognition rate of above 90% for most of the digit was observed by analysing the class wise accuracy. Incomplete circular strokes and variation in writing the zero digit led to a lower accuracy of approximately (70%) for this digit as shown in Fig. 8.

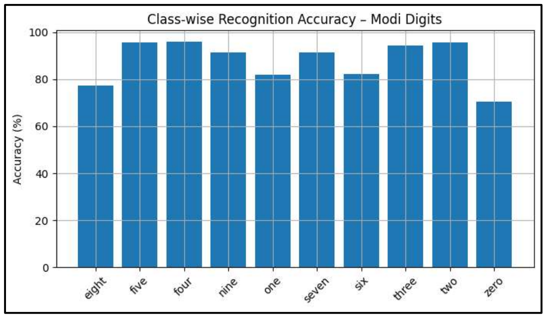

Figure 8. Class-wise recognition accuracy of MODI digits

D. Misclassified sample Analysis

Misclassified sample visualisation showed in Fig. 9 that there are predominant errors between digits that were visually similar. In the observation obtained from this further validated the findings of the conclusion matrix and emphasized the complexity of the MODI digits that were handwritten.

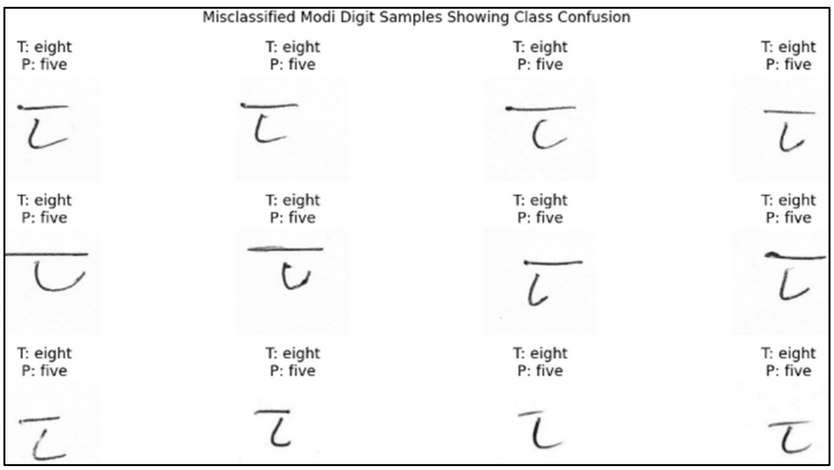

Figure 9. Misclassified MODI digit samples highlighting visual similarity between confusing classes.

VI. DISCUSSION

The effectiveness of CNN based feature learning for the recognition of the MODI digit was confirmed by the experimental results, as an overall high accuracy was achieved. Misclassification between certain pairs of digits points towards the complexity of intrinsic script rather than model failure. The recognition performance is significantly impacted by the absence of standardisation in the handwriting styles and also the historical variability. The noticeable gap between the training and validation accuracy as depicted in Fig. 4 and Fig. 5 is due to small dataset and similar strokes for depicted the digit in handwritten MODI. CNN focuses on feature selection during training hence level of accuracy is high. The validation data may contain diverse or unusual styles which may not be predicted accurately by CNN. The difference in accuracy and loss curve is common in handwriting recognition of ancient scripts.

VII. CONCLUSION AND FUTURE

A CNN based approach for the recognition of handwritten MODI digits presented in this paper. An overall accuracy of 88% was obtained with a large scale dataset. The need for handcrafted features has been eliminated in the method proposed. It also provides a comprehensive performance analysis by means of visual inspection of the misclassified samples and confusion matrix. Suggested future work includes exploring deeper architectures for further improvement in accuracy of recognition. Also the approach to full MODI character sets can be extended and data augmentation incorporated.

REFERENCES

[1] L. Gadpade and A. Chamatkar, "MODI Lipi character recognition using traditional machine learning algorithms," Int. J. Comput. Sci. Trends Technol., ICAMT25, 2025, ISSN: 2347-8578. pp. 199–202,

[2] K. Varpe and S. Sakhare, "Review of character recognition techniques for MODI script," Indian J. Sci. Technol., vol. 16, no. 26, pp. 1935–1946, 2023

[3] P. Patil and B. Kaur, "Handwritten digit recognition using various machine learning algorithms and models," Int. J. Eng. Res. Technol. (IJERT), vol. 9, no. 6, pp. 1234–1238, 2020

[4] Maheswari, S.U., Maheswari, P.U. and Aakaash, G.R.S. "An intelligent character segmentation system coupled with deep learning based recognition for the digitization of ancient Tamil palm leaf manuscripts," (2024). https://doi.org/10.1186/s40494-024-01438-4

[5] S. H. Ali and M. B. Abdulrazzaq, "KurdSet: A Kurdish handwritten characters recognition dataset using convolutional neural network," Heritage Science, 2024, doi: 10.1186/s40494-024-01438-4.

[6] S. Joseph and J. George, "Data augmentation for handwritten MODI character recognition using deep learning," in ICT for Intelligent Systems. Singapore: Springer, 2021, doi: 10.1007/978-981-15-7062-9_51, pp. 515–522,

[7] P. A. Tamhankar, K. D. Masalkar, and S. R. Kolhe, "Offline handwritten MODI document recognition using deep CNN," in Commun. Comput. Inf. Sci.. Singapore: Springer, 2021, pp. 478–487.

[8] B. Solanki and M. Ingle, "Performance evaluation of thresholding techniques on MODI script," in Proc. Int. Conf. Adv. Comput. Telecommun., 2018.

[9] R. K. Maurya and S. R. Maurya, "Recognition of medieval MODI script using hybrid feature space," Int. J. Comput. Sci. Eng., vol. 6, no. 2, pp. 136–142, 2018.

[10] D. N. Besekar "Special Approach for Recognition of Handwritten MODI Script's Vowels", Proceedings published by International Journal of Comp. Applications® (IJCA). (2012); https://research.ijcaonline.org/medha/number1/medha1023.pdf. pg. 48–52.

[11] S. L. Chandure, V. Inamdar. "Performance analysis of handwritten Devnagari and MODI Character Recognition system", International Conference on Computing, Analytics and Security Trends (CAST). (2016); Available from: https://doi.org/10.1109/CAST.2016.7915022. p. 513–516.

[12] K. Varpe, S. Sakhare, "Review of Character Recognition Techniques for MODI Script", Indian Journal of Science and Technology 16(26) (2023); 1935-1946. https://doi.org/ 10.17485/IJST/v16i26.485

[13] A. Ramjon, A. Bwatiramba, S. Venkataraman, "Handwritten Character Recognition Using Deep Learning (Convolutional Neural Network)", Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol.14, No.1, (2023).

[14] V. Pawar, D. Wadkar, S. Kashid, P. Prakare, V. More, A. Babar, "MODI Lipi Handwritten character Recognition using CNN and Data Augmentation Techniques". International Research Journal of Engineering and Technology. (2022); p. 1880–1884. Available from: https://www.irjet.net/archives/V9/i6/IRJET-V9I6337. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[15] S. Joseph, J. George, "Efficient Handwritten Character Recognition of MODI Script Using Wavelet Transform and SVD". Data Science and Security. (2021); Available from: https://doi.org/10.1007/978-981-15-5309-7_24, p. 227–233.

[16] S. Joseph, J. George, "Handwritten Character Recognition of MODI Script using CNN Based Feature Extraction Method and Support Vector Machine Classifier", 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP). (2020: Available from:https://doi.org/10.1109/ICSIP49896.2020.9339435.

[17] P. Tamhankar, K. Masalkar, and S. Kolhe, "A Novel Approach for Character Segmentation of Offline Handwritten Marathi Documents written in MODI Script", Procedia Computer Science. (2020);. Available from: https://doi.org/10.1016/j.procs.2020.04.019, 171:179–187

[18] M. Deshmukh, S. Kolhe, "MODI-HChar: Historical MODI Script Handwritten Character Dataset", (2023). IEEE Dataport. https://dx.doi.org/10.21227/v2kt-rr94