International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-12 Issue-1, October 2022

Implementing a Hybrid Deep Learning Approach to Achieve Classic Handwritten Alphanumeric MODI Recognition

Maitreyi Ekbote, Aishwary Jadhav, Dayanand Ambawade

Abstract**: *MODI, synonymous with the Devanagari script, is an ancient script from the 17th century used by the Maratha empire as a symbol of culture and power to propagate Marathi. Due to a decline in its usage, absence of quality script database and an unavailability of good literature, identification and translation of MODI script is demanding. The present work deals with a novel study on the recognition of MODI characters and numerals by using Convolutional Neural Network (CNN) architecture. By using a traditional machine learning classifier, classification is performed, and then through a comparative analysis of Random Forest and XGBoost, the study achieves recognition accuracy of 92% for characters and 93.3% for numerals.***

Keywords**: *CNN, Handwritten Character Recognition, MODI Script, Random Forest***

I. INTRODUCTION

The MODI script is one of the obsolete languages from ancient India. MODI emerged as an administrative writing script during the 16th century, before the inception of the Marathas. The Marathas adopted it in the 17th century as a symbol of culture and power, and it continued to be regularly used until the mid-20th century for social, economic, and political purposes. Most papers and letters written since Chhatrapati Shivaji Maharaj's era are written in MODI scripts, therefore it plays a crucial role in historical research. An account of all these historic documents is noted to be preserved with temples and libraries even today [1]. The origin of the MODI script is still a question - several researchers and historians have claimed MODI to have originated in the 12th century in the Devagiri region, during the Ashoka empire in the form of Brahmi script or during the 17th century Shivaji period [2]. By the 20th century, Marathi was written in 2 scripts - Balbodh and MODI. However, due to the need for maintaining uniformity and conformity across all records, it was decided to discontinue the MODI style and maintain the Balbodh way of the Devanagari script to pen Marathi, resulting in a decline in the usage and popularity of the script.

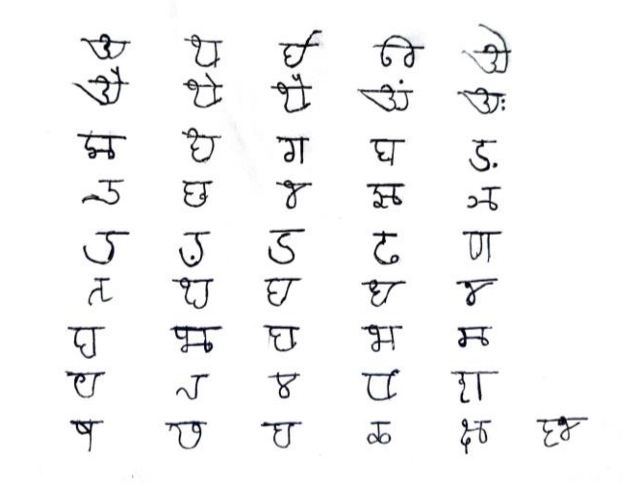

Fig. 1. MODI Characters

The case of the MODI script makes HWCR (Handwritten Character Recognition) even more challenging due to the significant similarity of the characters' shapes and the lack of a word-stopping symbol. Another issue imposed by the MODI lipi is segmentation. This process is demanding because of several issues, including text deterioration, structural variations in different handwritten manuscripts, non-cursive nature of characters and use of strokes among others [3].

Fig. 2. MODI Numbers

Since the advent of complex machine learning models, there have been numerous studies on the recognition of regional characters in recent years. Most of this research deals with the use of KNN (K-Nearest Neighbors), CNN (Convolutional Neural Network), ANN (Artificial Neural Network), SVM (Support Vector Machine), etc.

Manuscript received on 26 September 2022 | Revised Manuscript received on 29 September 2022 | Manuscript Accepted on 15 October 2022 | Manuscript published on 30 October 2022.
* Correspondence Author (s)
Maitreyi Ekbote*, Student, Department of Electronics and Telecommunication, S.P.I.T., Mumbai (Maharashtra), India. Email: maitreyi.ekbote@spit.ac.in
Aishwary Jadhav, Student, Department of Electronics and Telecommunication, S.P.I.T., Mumbai (Maharashtra), India. Email: aishwary.jadhav@spit.ac.in
Dayanand Ambawade, Associate Professor, Department of Electronics and Telecommunication, S.P.I.T., Mumbai (Maharashtra), India. Email: dd_ambawade@spit.ac.in

However, there are few studies that deal with the identification of the characters of a script as outdated as MODI - few studies have used CNN and SVM, while others have used the centroid based recognition model, zone-based approach, and structural similarities. Therefore, there must be a need for a holistic analysis of MODI alphanumeric inputs by using a combination of algorithms. The study focuses on implementing a HWCR model on a dataset of alphabets and numerals of the MODI script using Convolutional Neural Network with an additional machine learning classifier. This classifier helps improve the model performance - this is further fine-tuned by using different classifiers for a comparative analysis. Through this approach, the paper contributes to the extant literature by presenting a hybrid deep learning model (CNN + Random Forest or XGBoost), performing recognition on alphanumeric characters, and achieving it with optimum system complexity. The organization of the paper is as follows: Section II details the present work on MODI HWCR, Section III outlines the proposed research methodology, and Section IV elaborates on the results from the study which is followed by the conclusion.

II. LITERATURE SURVEY

Extant literature focuses predominantly on HWCR research on various common regional languages such as Hindi, Marathi, Urdu, Bengali, Gujarati, etc. However, there has been little research on the MODI script on character recognition. With the advent of complex systems and algorithms, MODI script started garnering attention among researchers in recent years. Some of these studies have been outlined below. Pawar et al. [4] conducted character recognition of handwritten MODI characters using CNN and achieved an accuracy of 91.62%. Similarly, for the recognition of handwritten MODI script, Joseph et al. [5] implemented a method called ACNN (Augmented CNN) which makes use of both data augmentation and CNN. This method achieved a very high accuracy of 99.78%. Tamhankar et al. [6] also employed a CNN model for the recognition of 31 MODI characters and realised 64% accuracy. The Joseph & George [7] approach, which uses SVM classifier with CNN as a feature extractor, performs with a high accuracy of 99.3%. Chandure & Inamdar [8] also used an SVM-based classifier, and developed a Transfer Learning-based system for the recognition of both Devanagari and MODI characters. Using Deep CNN AlexNet for feature extraction, the recognition accuracy achieved was 92.32% for MODI characters and 97.25% for Devanagari characters. Similarly, AlexNet was also used by Mahajan & Tajne [9] for recognition of MODI characters as well as numbers. The accuracy generated by this model is 89.72%. A technique that employs CNN for MODI script character feature extraction and classification was implemented by Sawant et al. [10]. These recognized MODI characters are mapped to Devanagari characters. Acharya et al. [11] used a Deep CNN model for the recognition of handwritten Devanagari characters as well as numbers and were able to achieve an accuracy of 98.47%. Patil [12] defined the Affine Moment Invariants method and calculated moments of each character. Classification has been performed using Fuzzy Logic. [13] proposed a hybrid approach for extraction of feature by combining Moment Invariant and Affine Moment Invariant for feature extraction. This system reached a recognition accuracy of 89.72% by applying SVM as a classification method. A chain code histogram-based method was put forward by Chandure & Inamdar [14] in combination with Intersection Junction feature extraction methods. The accuracy achieved when using BPNN (Back Propagation Neural Network), KNN, and SVM for training, testing, and character classification was 37.5%, 60% and 65%, respectively. In [15], a structured similarity technique is utilised to determine the likeness between handwritten and conventional MODI script characters. For the recognition of MODI script, a feature extraction method utilising backpropagation neural networks and structure similarity was designed, and an accuracy of 93.5% was realised. The current study presents an implementation of a HWCR model on a dataset of alphabets and numerals of the MODI script using a type of Convolutional Neural Network for feature extraction and a traditional machine learning classifier. The model is fine-tuned with the help of the classifier for classification purposes. The paper contributes to the extant literature by using a hybrid deep learning model to perform recognition on alphanumeric characters and achieving it with optimum system complexity.

III. METHODOLOGY

The paper proposes a Convolutional Neural Network based character recognition model to facilitate effective identification of MODI alphabets and numerals due to the ability of the CNN to be widely applied in image recognition scenarios. The model is enhanced by using a classifier - Random Forest or XGBoost, and the performance of each is compared with the aid of performance parameters.

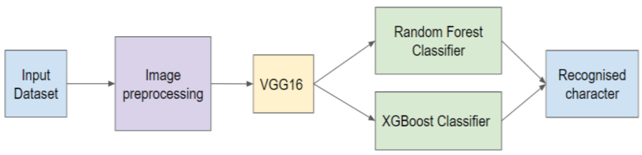

Fig. 3. Block Diagram of Proposed System

A. Image Preprocessing

Data should be simple, accurate, consistent, and sufficient, before employing it for any purpose, but this is not always the case. To get the intended outcomes, the data must be preprocessed, that is, cleaned and processed to the proper format. It can, thus, enhance the quality, suppress unwanted distortions, or extract useful information from the images.To preprocess image data, many techniques such as image resizing, grayscale conversion, normalization, data augmentation, morphology, smoothing or blurring, are utilised. The techniques used can vary depending on the application.

Retrieval Number: 100.1/ijeat.A38461012122
DOI: 10.35940/ijeat.A3846.1012122
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-12 Issue-1, October 2022

B. VGG16

A feed-forward neural network called a convolutional network analyses visual images by processing data in a grid-like architecture. It is also referred to as a ConvNet. To find and categorise objects in an image, the Convolutional Neural Network is employed. The object detection and classification algorithm VGG16 is a type of CNN.

VGG stands for Visual Geometry Group and the 16 refers to 16 weighted layers, which are learnable parameters layers. In VGG16, there are 13 Convolutional layers, 5 Max Pooling layers, and 3 Dense layers. Throughout the whole architecture, the convolution and max pool layers are uniformly ordered. A deep network and small convolutional filters may perform accurate image identification, as demonstrated by the well-known CNN architecture VGG. Hence, in this study, VGG16 is being used. The process of removing useful elements from an image begins with the convolution layer. Multiple filters work together to perform the convolution action in this layer. Each image can be thought of as a matrix of pixel values. In this layer, we try to manoeuvre the filters, strides, and padding functionality of the feature detectors to improve the model performance. We also implement 2 convolution layers together to improve the hierarchical reception of visual images. Upon convolution, the visuals are converted into numerical values, allowing the model to interpret graphical outcomes.

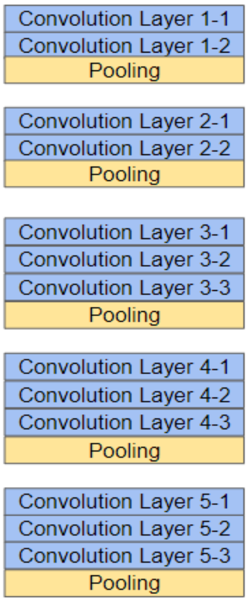

Fig. 4. Architecture of VGG16 used in the study

Pooling layers, also known as down sampling, reduces the number of parameters in the input. It also performs dimensionality reduction. The pooling operation sweeps a filter across the entire input similarly to the convolutional layer, with the exception that this filter lacks weights. Instead, the kernel populates the output array by applying an aggregation function to the values in the receptive field [16].

Dense Layers were not used in the study. Instead, the model is used for feature extraction, and then combined with traditional machine learning classifiers for classification.

C. Random Forest Classifier

The Random Forest classifier generates a set of distinct decision trees from a randomly chosen subset of the training data. To choose the ultimate prediction, it compiles the votes from various decision trees. The individual trees each spit forth a class prediction. The class with the highest votes becomes the prediction made by our model.

Ensemble combines several models to make forecasts. An example of Bagging Ensemble technique is Random Forest. It develops a separate training subset with replacement. The outcome is decided on majority voting.

D. XGBoost Classifier

XGBoost stands for Extreme Gradient Boosting. Trees are built sequentially in this algorithm. Each independent variable is given a weight before being fed into the decision tree that forecasts outcomes. The incorrect weights are updated in the following iteration. These distinct classifiers or predictors are then combined to produce a robust and accurate model. XGBoost employs trees that differ slightly from conventional decision trees. They go by the name CART (Classification and Regression Trees) and include real-value scores indicating whether a certain instance belongs to a particular group. The choice can be made once the tree has reached its maximum depth by classifying the scores using a predetermined threshold. XGBoost is an example of Boosting Ensemble, which tries to construct sequential models with the aim of maximising accuracy in the final model.

IV. RESULTS AND DISCUSSION

The study makes use of a handwritten MODI script dataset consisting of 46 distinct characters – 36 consonants and 10 vowels. Additionally, it also consists of 10 numerals, making the total size of the dataset to nearly 5000 samples. Since the dataset for MODI script is not available, it was created by writing down each character and subsequently scanning it. The dataset was split by 80%-20% pattern into training and test sets and subsequently the 2 classifiers were applied to evaluate model performance.

From Table I, the train and the test accuracy of Random Forest and XGBoost classifiers can be observed.

Table- I: Train v/s Test Accuracy

Accuracy of Classifier	Train (80%)	Test (20%)
Random Forest	0.96	0.9203
XGBoost	0.92	0.843

For analysing the model performance as well as comparing the classifiers, an array of performance metrics was employed: Accuracy, Precision, Recall, F1 score, AUC-ROC (Area Under the Curve - Receiver Operating Characteristic Curve) Score, Cohen's Kappa Score, MCC (Matthew's Correlation Coefficient) and Log Loss. The VGG16 optimised recognition model helped improve the accuracy for both models when compared with native CNN where the accuracy achieved was less than 90%.

In order to determine how frequently two raters who are evaluating the same quantity are in agreement, Cohen's Kappa Score, a statistical measurement, is utilised. An average level of agreement amongst all raters is indicated by a Kappa Score that is close to 1.0.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$k = \frac{p_o - p_e}{1 - p_e} \qquad (1)$$

where, p_o is the probability of agreement and p_e is the probability of random agreement.

MCC calculates the discrepancies between real and anticipated values. A high number, close to 1, indicates that both classes are correctly predicted by MCC, even if one class is disproportionately under or over represented.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \qquad (2)$$

MCC considers all these four values of the confusion matrix – TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative).

According to Table II, although Log Loss is higher for Random Forest it outperforms XGBoost for the characters dataset for most of the performance metrics.

Table- II: Performance Parameters of MODI Characters Dataset

Performance Parameters	Random Forest	XGBoost
Accuracy	0.9203	0.843
Precision	0.9277	0.850
Recall	0.9203	0.843
F1 Score	0.9175	0.836
AUC – ROC Score	0.9934	0.9932
Cohen's Kappa Score	0.9185	0.8395
MCC	0.9188	0.84
Log Loss	1.1439	0.612

Similarly, comparative analysis across the aforementioned performance metrics for the numbers dataset is performed, where we observe Random Forest to be better than XGBoost across all metrics.

Table- III: Performance Parameters of MODI Numbers Dataset

Performance Parameters	Random Forest	XGBoost
Accuracy	0.933	0.833
Precision	0.95	0.833
Recall	0.933	0.833
F1 Score	0.9314	0.825
AUC – ROC Score	0.9951	0.979
Cohen's Kappa Score	0.9259	0.8148
MCC	0.9282	0.8168
Log Loss	0.7825	0.5894

In the study, in an effort to understand the relation between estimators and model accuracy, we iterated the model over an array of estimator values ranging from 10 to 100. As outlined in Figure 5, we observe that as the number of estimators increase, the accuracy of the model also increases.

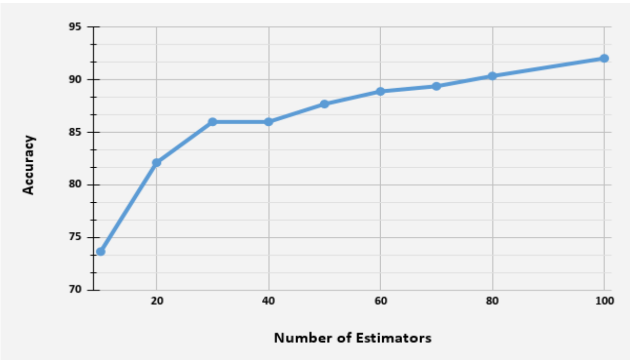

Fig. 5. Number of Estimators v/s Accuracy

Results of this study corroborate the efficiency of Random Forest over XGBoost: it is relatively easier to model and even performs correctly despite the presence of noise or missing data. Moreover, it does not easily lead to overfitting of the model. Further, the results of our study also help in concluding that our model effectively deciphers text of a complex script such as MODI with an accuracy of up to 92% for characters and 93.3% for numerals.

V. CONCLUSION

The present study deals with developing a model to achieve MODI script recognition of both numerals and characters. Present research in this space deals with the application of numerous techniques such as ANN, CNN, KNN, Complex moments and SVM to achieve HWCR of the MODI characters. However, few studies deal with HWCR on both numerals and characters using alternative approaches such as Random Forest and XGBoost which the current work aims to fulfill. Convolution Neural Networks help to effectively carry out traditional character recognition, therefore, we use a VGG16 optimised CNN model which is further modified by classifiers such as Random Forest and XGBoost individually to improve recognition accuracy over the traditional CNN. We achieve a model accuracy of 92% and 93.3% for character and numbers datasets respectively. The current study achieves handwritten alphanumeric character recognition using a hybrid deep learning model by merging the native CNN approach modified by VGG16 with transfer learning to 2 relatively less researched about classifiers in this space, particularly for MODI. The results obtained support the use of these classifiers to facilitate character recognition and singles out the Random Forest to be more efficient than the XGBoost classifier.

REFERENCES

1. Ramteke A.S., Katkar G.S., 2012, "Recognition of Offline MODI Script", International Journal of Research in Engineering, IT and Social Science (IJREISS), Volume 2, Issue 11, ISSN 2250-0588, pp.102-109.
2. Kulkarni Sadanand, A., Borde Prashant, L., Manza Ramesh, R., & Yannawar Pravin, L., "Review on recent advances in automatic handwritten MODI script recognition", International Journal of Computer Applications, 975, 8887.
3. Besekar D.N., Ramteke R.J, "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. [CrossRef]
4. Vishal Pawar, Deepak Wadkar, Shivkumar Kashid, Premraj Prakare, Vinayak More, Prof. S.A. Babar, "MODI Lipi Handwritten character Recognition using CNN and Data Augmentation Techniques", International Research Journal of Engineering and Technology, 2022.
5. Joseph, S., Datta, A., Anto, O., Philip, S., \& George, J. (2021). OCR system framework for MODI scripts using data augmentation and convolutional neural network. In Data Science and Security (pp. 201-209). Springer, Singapore. [CrossRef]
6. Tamhankar, P.A., Masalkar, K.D., Kolhe, S.R., "Character Recognition of Offline Handwritten Marathi Documents Written in MODI Script Using Deep Learning Convolutional Neural Network Model", In: Santosh, K.C., Gawali, B. (eds) Recent Trends in Image Processing and Pattern Recognition, RTIP2R 2020. Communications in Computer and Information Science, vol 1380. Springer, Singapore. [CrossRef]

Retrieval Number: 100.1/ijeat.A38461012122
DOI: 10.35940/ijeat.A3846.1012122
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

7. Solley Joseph, Jossy 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, 2020. [CrossRef]
8. Savitri Chandure, Vandana Inamdar, "Handwritten MODI Character Recognition Using Transfer Learning with Discriminant Feature Analysis", IETE Journal of Research, 2021. [CrossRef]
9. Dr. Kirti Mahajan, Niket Tajne, "An Ancient Indian Handwritten Script Character Recognition by Using Deep Learning Algorithm", 1st International Conference on Emerging Scientific Applications in the Field of Engineering and Technology, 2021.
10. Shruti Sawant, Talisha Tanna, Ankita Sharma, Snehal Kulkarni, Geeta Suvarna, "Word Transcription of MODI Script to Devanagari Using Deep Neural Network", 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), 2020. [CrossRef]
11. Shailesh Acharya, Ashok Kumar Pant, Prashnna Kumar Gyawali, "Deep Learning Based Large Scale Handwritten Devanagari Character Recognition", 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), 2015. [CrossRef]
12. Patil, P. A. (2016). Character recognition system for Modi script. Int J Comput Eng Res, 6(9), 102.
13. S. S. Gharde and R. J. Ramteke, "Recognition of characters in Indian MODI script," 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), Jalgaon, 2016, pp. 236-240. [CrossRef]
14. S. L. Chandure and V. Inamdar, "Performance analysis of handwritten Devnagari and MODI Character Recognition system," 2016 International Conference on Computing, Analytics and Security Trends (CAST), Pune, 2016, pp. 513-516. [CrossRef]
15. A.S. Ramteke, G.S. Katkar, Recognition of Off-line Modi Script: A Structure Similarity Approach, Proceedings- Int. J. ICT Manag. 1 (2013) 12–15.
16. Analytics Vidhya, "How to Approach CNN Architecture from Scratch?" 2022,[Online].Available:https://www.analyticsvidhya.com/blog/2022/05/how-to-approach-cnn-architecture-from-scratch/ [Accessed: Sep. 18, 2022].

AUTHORS PROFILE

Maitreyi Ekbote is a student of the Electronics and Telecommunication Department of Bhartiya Vidya Bhavan's Sardar Patel Institute of Technology, Mumbai. Her research interests include Deep Learning, Machine Learning, Data Analytics, Data Science and Cyber Security. She has a publication on Deep Learning Neural Networks presented at an international conference.

Aishwary Jadhav is a student of the Electronics and Telecommunication Department of Bhartiya Vidya Bhavan's Sardar Patel Institute of Technology, Mumbai. His research interests include Deep Learning, Machine Learning, Data Analytics and Data Science He has 3 research papers published in international conferences in the domain of Machine Learning, Finance and Digital Signal Processing.

Dayanand Ambawade is an Associate Professor at Bhartiya Vidya Bhavan's Sardar Patel Institute of Technology, Mumbai. He holds a M. Tech degree from IIT Bombay and has a teaching experience of more than 20 years. He is an active researcher with over 30 paper publications and 3 books. He has served as the Chairman of Forum of Electronics and Telecommunication Students (FETs), was the Secretary of National Conference on Biomedical Engineering (NCBME) and an Expert Member in the Staff Selection Commission (SSC) of the Government of India.

Retrieval Number: 100.1/ijeat.A38461012122
DOI: 10.35940/ijeat.A3846.1012122
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)