

## REVIEW ON MODI HANDWRITTEN CHARACTERS RECOGNITION

**Sahil Das\*1, Krishna Wankhede\*2, Anand Rituraj\*3**

\*1,2,3Student, Dept. Of Computer Science And Engineering, MIT WPU, Pune, India.

DOI : <https://www.doi.org/10.56726/IRJMETS31334>

### ABSTRACT

This review paper is based on MODI handwritten character recognition. Handwritten Characters recognition work is now a challenging task for the last few years because different writing styles for each person are different. Character recognition is nothing but turning human handwritten language into machine language. Many works have been done on other Indian languages until the present moment. Many Indian historical documents are written in many Indian script languages. But not so much work is done for the MODI script. So that it is necessary to preserve such documents for the future, several recognition systems have been created for multiple languages, and research on Indian scripts is ongoing. In recent decades, the different writing patterns of every HOCR have been a severe problem. The world recognized India because of its unity in variety. Due to the country's geographical and cultural diversity, many spoken languages and written scripts were formed and used for regular contact. Several brave efforts were made toward HOCR for different Indian hands. Because of the variety and structure of Indian characters, creating the character recognition system was deemed an active topic of study. This paper is an investigation. Previous research was done on MODI handwritten character recognition.

**Keywords:** Character Recognition, Modi Script, Segmentation, Recognition, Classification.

### I. INTRODUCTION

The brain is where the human central nervous system is located. The brain is complicated since it comprises a vast network of 50–100 billion neurons. A brain tumor is a collection of unneeded and abnormal cells developing inside the brain. Because it occupies space inside the skull and increases intracranial pressure, it is often referred to as an intracranial lesion. Brain tumors can be divided into two main categories: benign and malignant. In contrast to benign tumors, which are non-cancerous and seldom recur, malignant tumors are cancerous and grow quickly, invading the surrounding healthy brain tissue.

According to research in India. India speaks around 780 languages. Out of that, 220 languages are vanishing. MODI Script was an old script that lapsed from proper use in the 19th century before Devanagari was officially accepted. MODI was used for written communication at various phases for around 600 years. MODI became popular and was regularly used to write Marathi, as shown in fig 1. History researcher Chandorkar's MODI script emerged from the Ashoka period's Mouryi (Bramhi) script. According to another research, the Modi letter was created During the reigns of 'Peshwai' (Pune) and 'Chatrapati Shivaji Maharaj,' it was widely employed for writing in the 17th century. Shalgaonkar found that the earliest MODI document is stored in the Bharat Itihas Sanshodhan Mandal (BISM) in Pune between 1429 and 1389 A. D. MODI script, also called 'proto-MODI' or 'Adyakalin' in the 12<sup>th</sup> century, Yadavakalin' as in 13<sup>th</sup> century, Bahamanikalin' as in 14<sup>th</sup>–16<sup>th</sup> centuries, Shivakalin' as in 17<sup>th</sup> century, and Chitnisi' as in 18th century as Peshvekalin' between 1818 and 1952. Fig.1. shows MODI characters with consonants and vowels.



**Fig 1: Basic MODI Characters Set**

Many academics were drawn to handwritten character identification as Pattern Recognition, Image Processing, Natural Language Processing, and Document Analysis are the most promising fields. Today's environment is extremely fast-paced and highly mechanized. Automation has become associated with technology. Everything is because we, as people, want to do our tasks as quickly as possible. As a result, the further we organize, the simpler and quicker our work gets. Digitalization is the next massive thing in today's fast-paced society. Because we live in the computer age, we want all accessible information to be digitized and kept in computers with quicker computational capabilities. However, the challenge with translating facts into the digital world is to teach the computer the real-world facts in question. MODI is an ancient language built in the 12th century and created by Hemadpant. It is used for writing purposes only. These characters are generally cursive, as shown in Fig.1, so it isn't easy to recognize each character written by different writers. Most research now moves towards handwritten character recognition, which is most important for image processing, natural language processing, and pattern recognition.

There are two modes of Handwritten Characters Recognition.

- Online characters recognition Mode
- Offline characters recognition Mode

In online mode, A gadget wherein popularity is achieved when characters are below creation. In offline mode, it may be a framework in which written by hand documents are generated, scanned, put away within the computer, and then recognized. For offline mode, the input picture from the scanner is acquired for the offline character recognition system. Following its acquisition of the scanned digital print, the second stage is preprocessing, as shown in fig.2. It improves the image, making it acceptable for segmentation. Preprocessing procedures include erosion, dilation, opening, closure, and smoothing. The binarization method uses global thresholding to transform a grayscale image into binary. The procedures in the last two phases to generate the preprocessed picture appropriate for segmentation include detecting lines in the binary image using the Sobel approach, dilating the image, and filling the holes.

The primary purpose of this research is to offer a comprehensive review of state of the art in offline handwriting recognition systems for MODI characters. We looked at the research publications on handwritten characters' recognition based on the information collected.

## II. BRAIN TUMORS CLASSIFICATION MODI CHARACTER RECOGNITION SYSTEMS

### A. Machine learning-based Methods for Feature Extraction and Classification

Extraction of features is the initial stage in the data preprocessing method, which separates and minimizes an originating organizing original data into more manageable groups. Consequently, processing will be simplified; many variables are essential to these massive data sets. These variables demand a significant amount of processing power to process. As a result, feature extraction assists in locating the ideal feature from giant data sets by selecting and integrating elements from components, resulting in a reduction in the amount of data.. Feature extraction assists in locating the ideal feature from giant data sets by selecting and integrating elements from components, resulting in a reduction in the amount of data. Table 1. Shows some classification and feature extraction techniques used in machine learning. Classification is the process of classifying a set of data classes; it may be used to organize unstructured data. The procedure begins with forecasting the trajectory of provided data points. Classification predictive analysis emulates the mapping function between categorical input and discrete output variables. The primary purpose is to determine which category the new data will belong to.

Singh, P., & Budhiraja, S. (2011).[21] Introduced zoning, projection histogram, and distance profile feature as extraction methods and KNN, PNN, and SVM methods as classification techniques for Gurumukiscript. With the help of these methods, they got an accuracy rate of KNN 72.54% and SVM 73.02%. Affine and Moment Invariant Moment Invariant [22] is one of two approaches used to feature extracted from handwritten separated data.

Machine learning is a subfield of artificial intelligence inspired by psychology and biology that works to understand such data collection. It may be used to address a wide range of issues. A machine learning formula is presented data examples unique to a given problem and a solution that solves the issue for each case. Once training is finished, the system can achieve higher responses.

Sanjay S. Gharde and Dr. R. J. Ramteke [22] introduce different Machine learning. The support vector machine is utilized as a classifier in machine learning approaches. When doing classification, this support vector machine employs the linear kernel function. Using SVM, three datasets are combined, and this system has a high recognition accuracy for handwritten MODI script data.

Savitri Laxmanrao Chandure and Vandana Inamdar [12] introduced BPM, KNN, and SVM for MODI handwritten character recognition. They analyzed Devanagari and MODI Vowels using an algorithm with the chain code feature extraction method. According to this study, the recognition rate for Devanagari is more than for the MODI vowels.

Omer Aydin [16] proposed two different methods for the recognition of handwritten documents as well as printing document scans a scanner or digital camera with optical character recognition. He used the supervised method is used. One is Machine learning-based OCR, while the other is Microsoft Office Document Management Library. Text classification is done with the Naive Bayes Algorithm. In this, OCR is getting high accuracy for computer-based documents, and accuracy is getting low in handwritten characters. It is around 45%-60%, and using naive Bayes, phase accuracy was 53%.

Juhee Sachdeva and Sonu Mittal [19] suggested machine-learning techniques for handwritten character Devnagari recognition. He works on compound surfaces. Features are applied to the SVM, SMO, M LP, and simple logistic recognition. This classification model achieved a recognition rate: SVM achieved 99.88%, SMO achieved 99.72 %, Simple Logistics achieved 99.04 %, and MLP achieved 97.7%.

Ramteke A. S. and Katkar G.S[3] researched the structural similarity approach, and with that recognition, the rate was satisfactory, as shown in table1. Besekar D.N.[2] used chain code, an image-centroid method of extraction of features, and a Neural network model for classification. S. Joseph and J. George[6].suggested SVM for classification with his model accuracy rate of 99.3%, which is most satisfactory for character recognition. S. Chandure and V. Inamdar[9].also introduced the SVM model for classification, and the result is most satisfactory. As shown in Table.1. they are given below.

**TABLE 1. STUDIES BASED ON MACHINE LEARNING-BASED METHODS FOR CLASSIFICATION AND FEATURE EXTRACTION**

Project Title	Year of publication	Methodology	Result	Future Scope
Recognition of characters in Indian Modi Script. Sanjay S. Gharde and Dr. R. J. Ramteke [22]	2016	support vector machine (SVM)	Accuracy rate Dataset 1 89.71% Dataset 2 90.14% Dataset 3 89.33%	Increasing and adding samples in datasets check the recognition rate for using different kernel functions in the future.
Performance Analysis of Handwritten Devnagari and MODI Character Recognition System. Savitri Laxmanrao Chandure and Vandana Inamdar [12]	2021	BPM, SVM, KNN	Accuracy rate BPM 37.5%, KNN 60%, SVM 65% using chain code as feature extraction	Use different classification algorithms for improving accuracy, such as the neural network method.
Classification of Documents Extracted from Images with Optical Character Recognition Methods. Omer Aydin [16]	2016	Naive Bayes	The accuracy rate was 53%	Taking significant dataset recognition rate needs to improve.
Handwritten Offline	2021	SVM, SMO, M	The accuracy rate	Take a large dataset for

Project Title	Year of publication	Methodology	Result	Future Scope
Devanagari Compound Character Recognition Using Machine Learning. Juhee Sachdeva and Sonu Mittal[19]		LP, simple logistic recognition	was Support Vector Machine 99.88%, SMO 99.72 %, SimpleLogistic 99.04 %, and MuliLayerPercepti on 97.7%.	more accurate rate recognition. And apply more machine learning techniques to it.
Special Approach for Recognition of Handwritten MODI Script's Vowels Besekar D.N.[2]	2012	For feature extraction Chain code, image, and centroid method used, with classification Neural network model used.	Additional features like centroid are used for testing to improve the result for accuracy.	In the future, take many samples and extend the work using all characters.
Recognition of Offline MODI Script. Ramteke A.S., Katkar G.S[3]	2012	A measured structural similarity approach is used for structural handwritten character recognition.	The performance rate was 91 to 97 percent.	Take more text set datasets and compare alphabets that compare handwritten characters using A measured structural similarity approach.
Handwritten Character Recognition of MODI Script using Convolutional Neural Network Based Feature Extraction Method and Support Vector Machine Classifier. S. Joseph and J. George[6].	2020	SVM used for classification	99.3% accuracy rate was reported.	Modi script segmentation.
Handwritten Modi character recognition using transfer learning with discriminant feature analysis. S. Chandure and V. Inamdar[9].	2021	SVM is used to obtain a classifier model.	A satisfactory accuracy rate was obtained.	Improve inter-class misclassification for the respective dataset as well as CNN invariance properties.

**B. Deep learning-based Methods for classification**

Deep learning methods are used for solving complex problems that require finding hidden data for the extensive dataset. It is a deep understanding of intricate relationships between many interdependent variables. Table.2. Shown MODI handwritten character recognition employs deep learning approaches for categorization.

Solley Joseph and Jossy George [6] For the MODI character recognition, As a feature extractor, a CNN autoencoder has been installed is suggested. The feature set size was reduced from 3600 to 300 using CNN autoencoder. Following that, the retrieved characteristics were classified using SVM. Character recognition accuracy was 99.3 percent, which is higher than other recorded accuracy of MODI character recognition accuracy. This work's essential contribution is attaining a high accuracy rate in character recognition of MODI script.

Sidra Anam and Saurabh Gupta [11] looked into 22 MODI scripts with different vowels and consonants with numerals. They developed a recognition system that used otsu's binarization algorithm and Kohonen neural network method got a lower recognition rate for similar characters. Still, the overall recognition rate was 72.6% which is very satisfactory for the handwritten characters dataset—Savitri Chandure & Vandana Inamdar [17] Research on MODI and Devnagari handwritten scripts. Comparers self-created handwritten dataset compares and development is done with the support of supervised transfer learning method on a framework. It utilizes a Deep Convolutional Neural Network (DCNN). They achieved 92.32% and 97.25% recognition rates for MODI and Devanagari scripts. As shown in Table 2. Ritik Dixit, Rishika Kushwah, Samay Pashine[24] CNN classification model was getting more accuracy than other machine learning-based models. Dr. Kirti Mahajan1, Niket Tajne[23] AlexNet is a classification model that achieves a satisfactory recognition rate.

**TABLE 1: DEEP-LEARNING-BASED METHOD FOR CLASSIFICATION.**

Project Title	Year of publication	Methodology	Result	Future Scope
Handwritten Character Recognition of MODI Script using Convolutional Neural Network-Based Feature Extraction Method and Support Vector Machine Classifier. Solley Joseph and Jossy George [6]	2020	CNN autoencoder	The accuracy rate was 99.3%	Take more concentration on Modi segmentation in the future.
An Approach for Recognizing Modi Lipi using Otsu's Binarization Algorithm and Kohonen Neural Network. Sidra Anam and Saurabh Gupta [11]	2011	otsu's binarization algorithm and Kohonen neural network method	72.6% accuracy rate for the handwritten characters dataset	In the future, they will extend the scope of work with different Indian languages script available and improve the recognition rate.
Handwritten MODI Character Recognition Using Transfer Learning with Discriminant Feature Analysis. Savitri Chandure & Vandana Inamdar [17]	2021	DCNN	Modi script accuracy was 92.32%. Devnagari's script accuracy was 97.25%	It will encourage examining the technique using an additional dataset, emphasizing improving inter-class misclassification and the CNN's invariance qualities.

Project Title	Year of publication	Methodology	Result	Future Scope
An Ancient Indian Handwritten Script Character Recognition by Using Deep Learning Algorithm. Dr. Kirti Mahajan1, Niket Tajne[23]	2021	The deep learning model named as AlexNet	The accuracy rate was 89.72%	Improve accuracy rate using different deep learning models.
Handwritten Digit Recognition using Machine and Deep Learning Algorithms. Ritik Dixit, Rishika Kushwah, Samay Pashine[24]	2021	Developed SVM(support vector machine), CNN(Convolutional neural network ), MLP (Multilayer Perceptron)	Compare the three models and conclude that CNN has the best accuracy rate than the other two models for image character recognition.	Use applications in government sectors like medical for the treatment of patients, in national forces for equipment checking in national parties, and fingerprint detection and surveillance system for detection of the suspected person.

### III. DISCUSSION

We observed and analysed some points listed below.

- Need to generate Standard dataset for handwritten MODI script
- Need more research on old Indian handwritten languages
- The handwritten style for the same person may be changed, which affects for accurate recognition of handwritten characters
- In Machine learning-based methods, the accuracy rate is minimum, and machine learning-based methods contain feature extraction methods that may be time-consuming.
- Technology can analyse enormous amounts of data and identify precise trends and patterns humans may miss.
- ML Algorithms were capable of handling many-dimensional and complex information in unexpected environments.
- Data gathering is one of the most challenging aspects of Machine Learning. Furthermore, data collection has a cost. Again, gathering data through surveys may contain many fake and erroneous data. We frequently encounter situations where we discover an unbalance in data, resulting in poor prediction accuracy.
- To solve a Machine Learning challenge, several algorithms may be used. They run models with several algorithms and determine the best accurate method based on the findings.
- Machine learning algorithms are sometimes time-consuming, and sometimes it takes more CPU power to process many datasets.
- In deep learning-based methods, the accuracy rate was satisfactory.
- In deep learning, Properties are quickly inferred and adequately modified to get the required outcome. At that point, it will not be essential to extract characteristics. This reduces the need for time-consuming machine-learning techniques.
- The same Neural Network-based methods apply to any other applications. In deep learning methods, differences in data are learned automatically.

- Mainly parallel computation is performed, so machine performance is much more satisfactory when a large dataset is used. Deep learning architecture is more flexible and will adopt new problems in the future.
- It is not a specific model to aid you in selecting the correct deep-learning resources since it necessitates an understanding of topology, training techniques, and other characteristics. As a result, it is harder to adopt by less competent individuals.

#### IV. CONCLUSION

The overall paper is based on previous studies on MODI handwritten character recognition. According to the survey, we state that there is a need for a lot of work to recognize characters to generate one standard dataset. Future work will be to maximize the accuracy of the published dataset. Researchers are researching this field by applying Deep learning and machine learning methods. Not much research was done for MODI handwritten characters script, so it is now to improve handwritten character recognition research work. But so far, the 100% accuracy rate was not noted, so future research on increasing the accuracy rate needs research. According to observations, MODI character identification is still in its early stages. MODI is an old character. It is not on the list of approved Indian language hands, which adds that it has received less study attention than other Indian writings. Pattern identification in MODI language is problematic for various reasons, such as letters from similar and irregular writing styles in contrast to other scripts. So far, just a few papers have been published, and they have been studied in this paper.

#### V. REFERENCES

- [1] Besekar D.N., 2011, "Recognition of Numerals of MODI Script Using Morphological Approach", International Referred Research Journal, ISSN 0974- 2832, vol. 3, Issue -27, pp. 63-66.
- [2] Besekar D.N., 2012, "Special Approach for Recognition of Handwritten MODI Script's Vowels," International Journal of Computer Applications (IJCA), MAHA-2012, pp. 48-52
- [3] 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.
- [4] 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.
- [5] 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.
- [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] 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.
- [8] P. A. Tamhankar, K. D. Masalkar, and S. R. Kolhe, "A novel approach for character segmentation of offline handwritten Marathi documents written in Modi script," Procedia Computer Science, Third International Conference on Computing and Network Communications (CoCoNet'19), vol. 171, pp. 179-187, 2020.
- [9] S. Chandure and V. Inamdar, "Handwritten Modi character recognition using transfer learning with discriminant feature analysis," IETE Journal of Research, pp. 1-11, 2021.COEP,
- [10] Solley Joseph and Jossy George, "Review article Feature Extraction and Classification Techniques of MODI Script Character Recognition," PERTANIKAJOURNAL ON SCIENCE AND TECHNOLOGY ISSN: 0128-7680.
- [11] 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 (0975 - 8887) Volume 111 - No 2, February 2015.

- [12] Savitri Laxman Rao Chandure and Vandana Inamdar, "Performance Analysis of Handwritten Devnagari and MODI Character Recognition System", 2016 International Conference on Computing, Analytics and Security Trends (CAST) College of Engineering Pune, India. Dec 19-21, 2016.
- [13] Manisha s. Deshmukh, Manoj Patil, and Satish Kolhe, "Offline handwritten Modi numerals recognition using chain code".
- [14] Shruti Sawant, Anita Sharma, Geeta Sharma, Snehal Kulkarni, Talisha Tanna, "Word Transcription of MODI Script to Devanagari Using Deep Neural Network", 2020 3rd International Conference on Communication System, Computing and IT Applications (CSC)
- [15] Solley Joseph and 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.
- [16] Omer AYDIN, "Classification of Documents Extracted from Images with Optical Character Recognition Methods," Anatolian Journal of Computer Sciences, ISSN:2548-1304.
- [17] Chandure & Vandana Inamdar (2021), " Handwritten MODI Character recognition Using Transfer Learning with Discriminant Feature Analysis," IETE Journal of Research.
- [18] A. T. Anju, Binu P. Chacko, and K. P. Mohammad Basheer, "Review of offline handwritten text recognition in south Indian languages", Malaya Journal of Matematik, Vol. 9, No. 1, 751-756, 2021 <https://doi.org/10.26637/MJM0901/0132>.
- [19] Juhee Sachdeva and Sonu Mittal, "Handwritten Offline Devanagari Compound Character Recognition Using Machine Learning," ACI'21: Workshop on Advances in Computational Intelligence ISIC 2021, February 25-27, 2021, Delhi, India.
- [20] Monica Patel, Shital P. Thakkar, "Handwritten Character Recognition in English: A Survey," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 2, February 2015, ISSN (Online) 2278-1021ISSN (Print) 2319-5940.
- [21] Singh, P., & Budhiraja, S. (2011), "Feature extraction and classification techniques in OCR systems for handwritten Gurmukhi Script-a survey," International Journal of Engineering Research and Applications (IJERA), 1(4), 1736-1739.
- [22] Sanjay S. Gharde and Dr. R. J. Ramteke, "Indian Recognition of Characters in Modi Script", 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication, 978-1-5090-0467-6/16/\$31.00 ©2016 IEEE.
- [23] Dr. Kirti Mahajan, Niket Tajne, "An Ancient Indian Handwritten Script Character Recognition by Using Deep Learning Algorithm", 1<sup>st</sup> International Conference on Emerging Scientific Applications in the field of Engineering and Technology", September 2021, ISBN:978-93-91535-10-0, pp 59-67.