International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Analyzing Different Techniques to Generate Image Datasets for MODI Script Handwritten Character Recognition

[1]Kanchan Varpe, [2]Sachin Sakhare

Submitted: 20/10/2023 **Revised**: 08/12/2023 **Accepted**: 16/12/2023

Abstract: Character recognition of MODI script is a proedure to recognize multifarious handwritten characters using distinct image documents. It is one of the oldest scripts utilized in Marathi language. Across Maharashtra India, Marathi Language is still a major spoken language. Applications like optical character recognition (OCR), handwriting analysis, and text extraction depend heavily on character recognition, which is a key task in computer vision and artificial intelligence (AI). Character recognition systems performance is strongly influenced by the caliber and variety of the training picture collection. This research is conducted to analyze distinct techniques to create the image dataset for MODI script handwritten character recognition. Furthermore, the significance of numerous datasets in handwritten character recognition, emphasizing the need for differences in writing styles of MODI script, typefaces, and backgrounds is examined. We explored the key methods to increase the variety of datasets utilizing multifarious approaches such as data augmentation, AI, and many others for Modi script handwritten character recognition. The intricacy of every data generation approach is also covered in the research, including assurance of data quality as well as preserving dataset, effectively. In this work, a new approach is proposed for the creation of a novel dataset of Handwritten MODI script character recognition. Further, there is proposed a framework rooted in enhanced recurrent neural network (RNN) and convolutional autoencoder-based framework for character recognition of the MODI lipi from handwritten documents. It was observed that our MODI lipi character recognition framework offers high accuracy in image analysis and takes less time in pre-processing the document images.

Keywords: *Data Augmentation, Handwritten Character Recognition, Image Datasets, Image Filtering, MODI Script, MODI Lipi.*

1. Introduction

Handwritten character recognition is becoming a vital component of computer vision and artificial intelligence. It plays a pivotal role in various modern applications, from automated document processing to text recognition in scanned document images containing the MODI script characters. The accuracy and performance of character recognition systems heavily depend on the quality and diversity of the image dataset [1]. There have been explored critical facets of this field, shedding light on the multifaceted methods and strategies used to curate and create image datasets of handwritten MODI script for character recognition. In a world increasingly reliant on digital information, the ability to accurately identify and interpret characters from diverse languages and scripts such as MODI is becoming essential in the modern digitized era. The practical implications span a wide spectrum, from enabling efficient data entry to aiding in the translation of historical documents, and automating data extraction in the forms of enhancing accessibility [2], [3].

Recently, a distinct examination dives into the intricate processes and tools employed by researchers, data scientists, and computer vision experts to collect, preprocess, and organize datasets that accurately reflect real-world scenarios for MODI lipi character recognition. It is through these meticulously crafted datasets that character recognition models learn to understand and differentiate between a vast array of characters, each with its unique nuances, shapes, and styles [4], [5]. The challenges in character recognition are manifold, ranging from variations in handwriting to the complexities of fonts and writing styles in image document text. Consequently, the techniques for generating image datasets must address these challenges, ensuring that recognition systems are robust, versatile, and capable of handling a wide array of scenarios [6]. It delves into the rich landscape of methodologies, both time-tested and innovative, utilized to create high-quality image datasets for MODI script character recognition. It covers aspects such as data augmentation, synthesis, crowdsourcing, and expert annotation, highlighting how these approaches are instrumental in building datasets that can empower MODI character recognition models to achieve high accuracy and adapt to the dynamic nature of written and printed text in image documents [7]. By exploring the myriad techniques employed to generate image datasets for character recognition. This topic not only deepens a key understanding of the underlying technology but also serves as a valuable resource for those engaged in the development of character recognition systems. The

[1]*Research Scholar, Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune 411048, India. Kanchanv2007@gmail.com*
[2]*Professor, Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune 411048, India. sachin.sakhare@viit.ac.in*

advancement of deep learning (DL) methods has the potential to transform and analyze the handwritten MODI lipi document images for effective character recognition.

1.1. Various Techniques for Generating MODI Script Image Datasets for Character Recognition:

Generating image datasets for MODI script character recognition involves employing a variety of techniques to compile a diverse and comprehensive collection of handwritten image documents. These techniques encompass manual data collection, where characters are sourced from scanned documents or handwritten samples. Data augmentation is a key strategy that is utilized for increasing the raw datasets. Computer-generated fonts and synthetic handwriting enable the creation of characters in multiple styles and sizes, while character extraction from text documents allows for real-world character acquisition [8]. Camera images and video frame extraction may capture handwritten characters of MODI script in everyday contexts, and 3D rendering enables the generation of images with different perspectives and lighting conditions. Image-to-image translation models can convert sketches or symbols into characters, and web scraping techniques collect character images [9]. Synthetic data generation through models like generative adversarial network (GAN) and the use of mixed datasets further enrich the dataset, facilitating more robust character recognition model training. Careful consideration of data diversity, balance, and ethical concerns is crucial in the dataset generation process [10], [11]. Various techniques for generating image datasets for character recognition are described as follows:

1.1.1. Manual Data Collection:

A vast number of pictures of the MODI script characters or symbols must usually be gathered and labeled to create an image dataset for character recognition. Although gathering data by hand might take a lot of time, it is necessary to guarantee that the dataset is diversified and of excellent quality [12]. The key methods utilized for the manual data collection can be classified into distinct categories as follows:

a. Photograph or scan physical documents containing characters.

b. Write characters by hand and scan them.

c. Take pictures of signs, billboards, or printed materials with MODI script characters.

1.1.2. Online Databases:

There are multifarious online sources for the creation of raw datasets for MODI script character recognition.

a. Utilize publicly available character datasets, such as MODI-HHDoc and many more

b. Explore websites or repositories offering character datasets for download.

1.1.3. Data Augmentation:

To enhance the performance of the character recognition models and diversify the dataset, data augmentation is an essential stage to increase datasets for character recognition in an automated manner. Using different transformations, data augmentation methods generate new, slightly altered copies of the original Handwritten MODI script images.

a. Apply geometric transformations such as rotation, scaling, and shearing to existing images.

b. Removal of noise or distortions to images to simulate real-world variability.

1.1.4. Computer-Generated Fonts:

A handy and flexible method for creating picture datasets for character recognition is to use computer-generated typefaces. A large range of characters and styles may be created, and the process of gathering data is simpler to manage [9].

a. Use font rendering libraries to generate images of characters in various fonts, styles, and sizes.

b. Customize font colors and backgrounds.

1.1.5. Handwriting Generation:

While training the models it is vital to identify handwritten characters, and one useful method is to generate picture datasets for character identification using handwritten text.

a. Create synthetic handwriting by using handwriting generation models.

b. Modify existing fonts to mimic different handwriting styles.

1.1.6. Character Extraction from Text:

A crucial stage in creating picture datasets for character identification in MODI script image documents is character extraction from text. To train a character recognition model, this method entails separating individual characters or symbols from text pictures.

a. Extract characters from text in various languages and fonts.

b. Randomly select characters from text documents or web sources.

1.1.7. Capture Camera Images:

Using a camera to take MODI script image documents of text samples is one way to capture camera images from text to produce an image dataset for character

identification. This method may help gather data from ancient literature documents or existing datasets [13].

a. Use a camera or smartphone to capture images of characters contained in the MODI script image documents.

b. Collect text images from objects, signs, or labels.

1.1.8. Video Frame Extraction:

To analyze multifarious videos containing the MODI script, one helpful method for creating picture datasets for character identification is to extract video frames from the MODI script image document.

a. Extract frames from videos containing MODI script characters and annotate them.

b. Useful for creating datasets with dynamic or moving characters.

1.1.9. 3D Rendering:

Creating a 3D rendering from text to generate image datasets for character recognition of MODI script is a unique and advanced approach. It involves rendering text in a three-dimensional space and capturing images from various angles and perspectives.

a. Creating the 2D images from different angles for finding the MODI script handwritten document images.

b. This generated dataset can be used in character recognition of the MODI script.

1.1.10. Image-to-Image Translation:

Textual descriptions are translated into pictures that reflect the specified characters in a process known as an image-to-image translation from text to create image datasets for character recognition of the MODI lipi. This is a novel and cutting-edge method that creates pictures of characters using text descriptions [14].

a. Use image-to-image translation models like Cycle GAN, to generate character images from other types of images.

b. Translate distinct symbols into characters.

2. Literature Review

In [6], S. Joseph et al. proposed a MODI script handwritten character recognition model. In this work, the CNN model was utilized for image document feature extraction and the classification of images was done utilizing the SVM classifier. DL-based approaches are mostly selected by researchers for various tasks namely the character recognition and segmentation of the MODI Lipi. Nevertheless, a developed model for MODI script handwritten characters has some limits such as being sensitive to Kernel function selection and computationally complex. D. N. Besekar et al. [15] conducted research on special techniques for the recognition of vowels in handwritten documents. This research was conducted to develop a novel framework that is capable of analyzing the handwritten text of the MODI script. This developed framework utilizes the chain code as well as picture centroid which is aimed to extract the image feature utilizing the double-layered feed-forward neural network for the image's classification.

S. Bhalerao et al. [16], conducted a systematic analysis of advanced approaches to MODI lipi character identification. It is a multipart procedure in which handwritten image documents of MODI lipi are translated into machine-readable code. In the past major research focused on distinct script character recognition such as Devanagari and many others. Due to the complex pattern of MODI lipi, very less investigators are conducted in this field. P. A. Patil [17] investigated novel OCR development for effective recognition of the MOID lipi characters. OCR is the electronic translation of MODI lipi image documents in machine-encoded text. The earlier OCR models support the distinct scripts namely the Devanagari, English, etc. Nevertheless, the existing OCR models built for MODI lipi are less effective in the recognition of the complex characters of MODI lipi.

I. Hussain et al. [18] carried out research for document image analysis (DIA), which focuses on transforming document pictures into machine-readable codes. A vital part of DIA systems is optical character recognition (OCR), which makes it possible to digitize document pictures. Applications for OCR system output may be found in many fields, including speech recognition, sentiment analysis, natural language processing (NLP), and translation services. The research emphasizes how crucial standard datasets are to the creation and assessment of text recognition algorithms. This work also tackles the issue of no standard dataset for handwritten Pashto text, which is a language with limited resources. The authors of the research have developed a dataset known as Pashto Handwritten Text Image-base (PHTI) to close this gap. This contains 4,000 scanned photos overall, equally divided between 200 male and 200 female authors, from PHTI. Further 36,082 text-line images were created by further segmenting these pictures, and each one was annotated or transcribed using UTF-8 codecs. For diverse DL applications, such as text recognition and classification, the PHTI dataset is pragmatic.

M. Ahmed et al. [19] conducted a review of OCR analysis. Character recognition is an automated procedure for reading diverse characters from photographs or documents whether printed or handwritten, and translating them into computer-readable text, such as ASCII and Unicode. Because OCR is affordable and time-

efficient, it is widely used in a variety of industries, including enterprises, post offices, banks, security systems, and robots. This study summarizes the present state of OCR technology research and applications in languages including Chinese, Hindi, Arabic, Russian, English, and MODI script-based documents. It also discusses a review of studies done on Arabic, Devanagari, and English characters, outlining the approaches used and the difficulties encountered in the character recognition of multifarious scripts.

Z. Jia et al. [20] discussed the significance of evaluating text picture quality in the context of scene text recognition (STR) systems. It draws attention to the difficulty in accurately assessing text image quality and presents the idea of text image quality assessment as a solution. The research presents a novel technique termed FSR- text image quality, which is predicated on integrating robustness and texture information at the character level. This approach generates character embedding and character pictures using an attention-based recognizer, then evaluates the robustness and textural clarity of the character embedding using Haralick features and similarity distribution distance. The purpose of a quality score network is to generate a quality score by normalizing these characteristics. The method's promise for dataset analysis and collection direction is shown by the strong discrimination it can achieve for varied-quality text pictures across many datasets.

S. Patil et al. [21] discussed an approach for improving the OCR on the picture along with the mixed character utilizing semantic segmentation. However, it draws attention to the lack of advancement in text recognition systems that can accurately perform OCR on handwritten and printed text at the same time. The unique characteristics of handwritten and machine-printed text give rise to this challenge. To improve the efficiency of conventional OCR engines, the research presents a method for locating, classifying, and cropping boxes in an image that contains relevant text by using semantic segmentation. To precisely find text in photos, the authors analyze well-known OCR engines such as Microsoft Cognitive Services, Google Cloud Vision, as well as AWS Recognition. They also suggest a pixel-by-pixel categorization method. As experimental investigations have shown, this technology attempts to increase the quality and accuracy of OCR findings, especially for mixed-typed text documents.

3. Methodology

Accurate character analysis and recognition of the MODI lipi is one of the crucial jobs because this script's nature is very complex. In past years, many researchers worked to develop the MODI script character recognition models rooted in deep learning (DL). There have been utilized multifarious techniques for recognition of the characters of the MODI script such as Zoning, HU Moments and Zernike moments, vertical projection, transfer learning, etc. Nevertheless, these MODI script character recognition approaches have many limits such as information losses of various character global patterns, less efficiency in accuracy measures, requirement of large datasets for optimal analysis of script characters, intricacy in complex character identification, etc. This research is aimed to analyze different techniques to generate image datasets for MODI script handwritten character recognition. To do so, there is proposed a DL framework based on enhanced RNN, which is found effective in processing the MODI script image documents in less time, and document cleaning is performed for effective recognition of the MODI lipi in generated datasets.

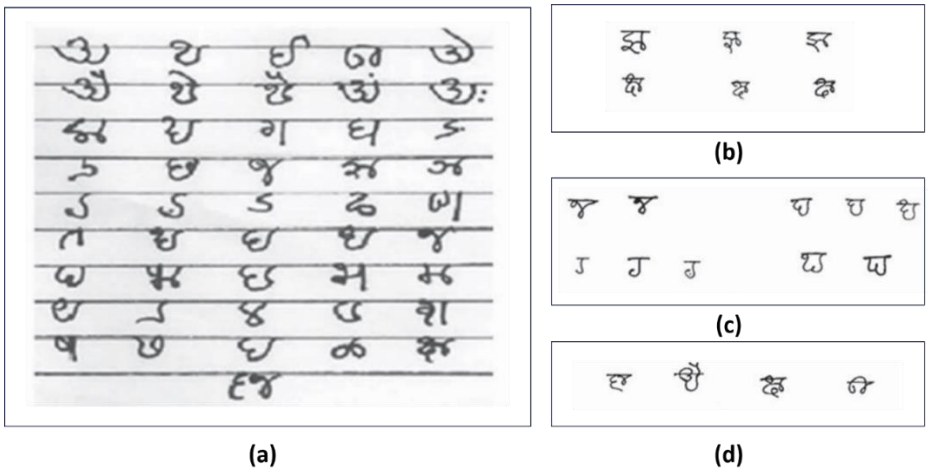

Fig 1: Illustrates the MODI lipi distinct characters (a) represent consonants and vowels set (b) depict characters variants owing to distinct writing patterns (c) characters having identical shapes (d) shows MODI lipi characters having complex shapes [22].

3.1. *Design:*

This research is conducted to formulate a novel dataset for MODI Script handwritten character recognition for accurate analysis of the distinct character image documents written in MODI lipi. Therefore, there is proposed a framework that is used to explore the dataset for character recognition by using enhanced pre-processing and classification of the images. The researcher developed the new MODI lipi dataset and standardized this using enhanced pre-processing and classification algorithms based on a deep learning approach for correct recognition of the distinct intricate characters of the MODI lipi. Figure 2 shows the proposed framework to generate image datasets for MODI lipi and make it standardized for effective character recognition. This proposed model is explained in multiple steps. The first step depicts that there may be chosen various techniques for creating MODI script image datasets from handwritten image documents of MODI lipi and many others to create a raw dataset from scratch. We have taken a total of 595 handwritten documents of the MODI script scanned with a higher resolution of 650 DPI. To make the datasets standardized, enhanced pre-processing techniques are the main building block in the correct identification of the handwritten character of MODI lipi. Hence, in our proposed model, we have applied four different pre-processing techniques to the collected raw dataset. Initially, the image thresholding is done which is an effective and simple way of partitioning an image into a foreground and background. In the subsequent step, an image thinning operation is done for chosen foreground pixels removal through binary pictures. Slant correction is used to improve the recognition accuracy of the image documents of MODI lipi. Further, image segmentation is performed which is a widely used method to divide an image document into various portions or areas, usually depending on the properties of the image's pixels. These pre-processed handwritten MODI script images are split into train-test groups through 10-fold cross-validation to validate the developed framework. An enhanced RNN model is selected to categorize the handwritten MODI lipi image documents divided in the train-test split. Further, this MODI lipi image documents processed dataset obtained in Step 3 goes through three stages. In this technique, noise is removed from the data set and enhanced for high image restoration for correct recognition of MODI lipi characters. The Mean filter was opted for enhancing the overall pixel values of the MODI lipi images. Further, the Median filter is utilized for the removal of the static noise from the image documents, thereby aiding in fast recognition of the MODI lipi characters. The Wiener filter was integrated the eliminate the additive noise and improve the blurred image resolution. Further, a convolutional autoencoder (CAE) has been utilized for the Handwritten image documents feature extraction and final model prediction is done based on the decision algorithm to identify the handwritten character of the MODI script.

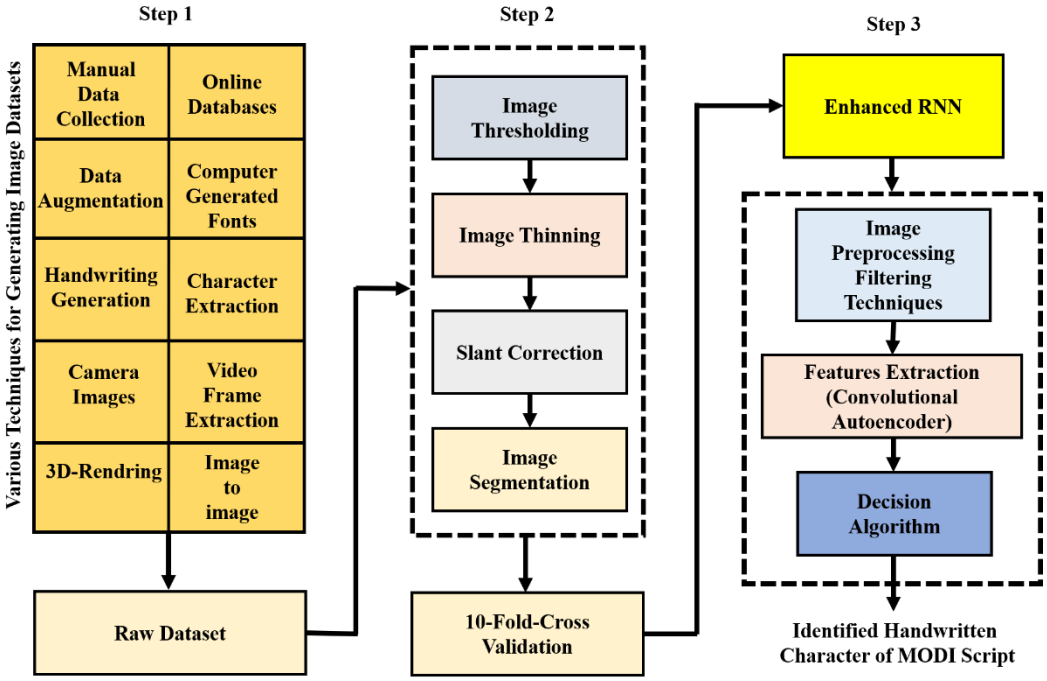

Fig 2: Proposed framework to generate image datasets for character recognition.

3.2. *Instruments:*

The system setup utilized in this research is described in this section. We selected recent computer systems as well as software applications to perform this research. Table 1 shows the system configuration for experiments. For performing this research, the researcher uses Python programming language and Windows 11 operating system with 8GB RAM and Processor (intel core i7, 12th

Generation) for smooth and fast work. This research needs some extra tools and packages with current versions like Labelme (v5.3.1), NumPy (v1.26.1), Pillow (v10.1.0), and OpenCV (v4.7.0).

Table 1: System configuration for experiments.

Sr. No.	System Configuration	Details
1.	RAM	8 GB
2.	Operating System	Windows 11
3.	Processor	12th Generation Intel Core i7
4.	Tools/Packages	**a)** Labelme v5.3.1 **b)** NumPy v1.26.1 **c)** Pillow v10.1.0 **d)** OpenCV v4.7.0
5.	Programming Language	Python v3.10.10

3.3. Dataset Preparation:

The dataset generation for MODI script handwritten character recognition is intricate due to several regions. MODI lipi is one of the oldest scripts, therefore handwritten documents of this script are generally available with degraded quality which makes it complex to understand. Further, the MODI lipi character writing pattern of the individuals is distinct and it takes time to formulate an identical dataset for the analysis of documents. The collection of this script dataset is complicated as the character symbols of this script are larger. In addition to this, the MODI script character analogy is very high, and the multifarious characters of this script may be broken or distorted which leads to high differences in sample collection. Still, there is very little research done on the MODI script handwritten character recognition and hence there are very less datasets available on this script. However, the creation of the MODI script datasets is a time-consuming procedure and requires massive effort, which makes it essential to explore some novel approaches that may aid in developing the new standardized datasets for effective analysis of the MODI script handwritten character recognition for the novel model development. This research explores the optimal approaches utilized for the development of MODI lipi new dataset of handwritten character recognition. The MODI lipi research is conducted in India, Maharashtra by the research institution Bharat Itihas Sanshodha Mandal [23]. In this institution, multifarious history researchers are working on MODI scripts. We have collected 595 handwritten documents of the MODI script and scanned them with a higher resolution of 650 DPI utilizing the standardized scanner application. The selected documents were segmented and classified in an automated manner utilizing the enhanced RNN and autoencoder-based model. All the image documents have been collected in separate folders based on the distinct categories of the images. All the images were saved in JPEG format having the identical dimensions of $256 \times 256 \times 3$. We have collected a sample of 5245 MODI script characters handwritten by the 120 experts of the MODI script. Our newly formulated images dataset contains 36 consonants and 10 vowels based on overall 46 categories of characters. Table 2 illustrates the MODI script dataset details.

Table 2: Illustrates the MODI script dataset details.

S. No.	Category	Categories of characters	Images size	Total Images documents
1	Vowels	10	$256 \times 256 \times 3$	186
2	Consonants	36	$256 \times 256 \times 3$	409
	Overall	46		595

3.4. Data Analysis:

In this research, for MODI script character recognition, a new dataset has been formulated from scratch which involves 595 image documents containing 36 consonant and 10 vowel-classified characters. The image analysis of the handwritten MODI lipi character images is done utilizing multiple filtering methods. These image document filtering techniques involve mean, median,

NLM (Non-Local Means), and Winer filter. These filtering techniques are utilized to perform various operations on the MODI lipi document images such as overall improvement in image resolution, and aid in smooth feature extraction of document images containing the handwritten MODI lipi characters.

3.4.1. Mean Filter:

Mean filtering is a straightforward image processing method for smoothing and reducing noise in images. It is sometimes referred to as an average filter or a box filter. It is a member of a group of linear filters that are used in spatial filtering. The primary concept of a mean filter is to substitute the average color or intensity of a pixel within a specified neighborhood or kernel for the color or intensity of the original pixel. Let, R_{ab} represents a neighborhood sub-image dimension window m × n rooted at the point (a, b). Further, the gray level of image pixel (a, b) in the defined K variable, the filter alters the gray level j(a, b) along with the adjacent pixel account particulars as displayed in equation 1.

$$\mathrm{J(a,b)} = \frac{1}{M\times N}\sum_{u,v\in R_{a,b}} K(\alpha,\beta) \quad (1)$$

3.4.2. Median Filter:

In image processing, a median filter is a distinct nonlinear digital filtering method that lowers noise in a picture while maintaining its edges and crucial features. Salt-and-pepper noise, which consists of errant white and black pixels dispersed across a picture, may be effectively removed with this technique. To get the median, first sort all of the neighborhood window's pixel values numerically, then replace each pixel that is taken into consideration along with the middle pixel value. The mathematical formula of median filtered document picture J(a,b) of the picture K(α, β) is given in equation 2.

$$J(a,b) = median_{(\alpha,\beta)\in R_{ab}}\{K(\alpha,\beta)\} \quad (2)$$

3.4.3. Wiener Filter:

A well-liked and effective linear filter image document processing is the Wiener filter which lowers noise level and improves overall handwritten MODI script image quality. It bears Norbert Wiener's name, who popularized the filter in the 1940s. Estimating the original, noise-free signal or picture from an observed, noisy version is the main objective of the Wiener filter. Equation (3) illustrates the minimized error, whereas Equation (4) represents the minimal wiener filter function in the frequency domain. The Wiener filter instantly eliminates additive noise very effectively and reverses picture pixels in a significant manner.

$$e^2 = E\{J(a,b) - \hat{\jmath}(a,b)\} \quad (3)$$

$$\hat{\jmath}(\alpha,\beta) = \left[\frac{1}{D(\alpha,\beta)}\,\frac{|D(\alpha,\beta)|^2}{|D(\alpha,\beta)|^2 + \frac{P_\eta(\alpha,\beta)}{P_J(\alpha,\beta)}}\right] K(\alpha,\beta) \quad (4)$$

Where, D (α, β) degradation function, $|D(\alpha,\beta)|^2$, D^*(α, β) = D (α, β) D^*(α, β); D^*(α, β), complex conjugate of D (α, β); P_η (α, β), power spectrum of the noise; P_j (α, β) power spectrum of processed undegraded picture.

3.4.4. Non-local means (NLM) Filter:

One of the common image-denoising approaches is the Non-Local Means (NLM) filter for the handwritten MODI script image document pre-processing and effective restoration. It is very effective in reducing Gaussian noise in Handwritten MODI script image documents while keeping fine details and textures. NLM is a non-linear filter that estimates and eliminates noise by considering the similarities between pixel patches within the Handwritten MODI script image documents. The weight given to a pixel in NLM is independent of both spatial as well as intensity distances. The weighted average of each pixel intensities u(Xa) within picture J is the restored intensity v(i, j) of pixel Xa.

$$NL_v(i,j) = \sum_{X_b\in J} W(X_a,X_b)\,u(X_a)$$

Where, W (X_a, X_b) is the weighted average allocated to the variable u(X_a) to restore the pixel X_a.

4. Result and Discussion

MODI script handwritten character recognition procedure attaining popularity, as it is one of the oldest scripts. This research analyzes the distinct techniques to generate a novel image dataset for MODI script handwritten character recognition. Furthermore, a framework is proposed which allows image preprocessing and filtering to identify the handwritten characters of MODI script from the preprocessed image documents containing distinct vowels and characters of MODI lipi. Our proposed framework for handwritten character recognition of MODI lipi image documents is based on the enhanced RNN as well as the convolutional autoencoder approach. The enhanced RNN is a distinct modified architecture employed for the classification of the image of handwritten MODI script character images containing a variety of vowels and consonants of MODI lipi. The RNN is one of the popular categories of neural network (NN). This proposed RNN architecture is utilized for processing of sequential dataset of MODI script and maintains the memory consumption, while taking the previous input data, thereby minimizing the overall time in the handwritten character recognition of

the MODI script while analyzing a large number of the images in the training process. In this research time series analysis and sequential data are the main functions of recurrent neural networks (RNNs). It was employed in MODI script handwritten image processing task when the temporal or spatial dependencies are found in the dataset as it makes it optimally. This proposed framework based on enhanced RNN and convolutional autoencoder performance was measured and compared with other distinct techniques namely LSTM (Long Short-Term Memory), GAN (Generative Adversarial Network), and DBN (Deep-Belief Network). The LSTM is mostly utilized for sequential data analysis, including time series data and jobs involving natural language processing. It is intended to represent and capture sequences of long-range interdependence. However, since these tasks do not inherently require sequential data, LSTMs are not often employed for classic image processing tasks like object recognition or picture categorization. GAN is one of the common approaches utilized in computer vision and document image analysis using data augmentation and many others. The deep learning models known as GANs comprise the generator and discriminator network. Further, the deep learning model called DBN is based upon many layers of latent, and stochastic variables to be utilized in image analysis. DBNs are utilized in certain image processing settings to analyze and classify distinct images.

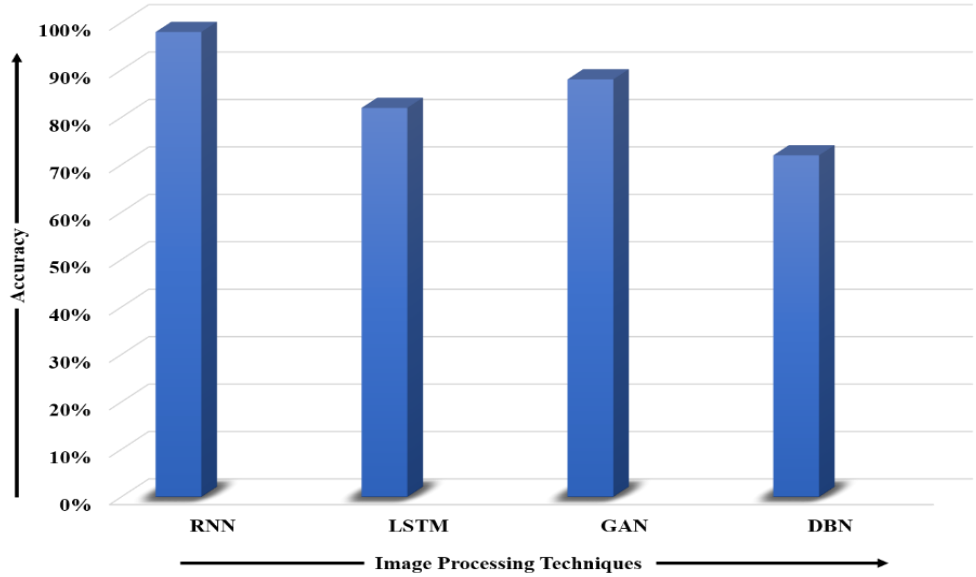

Fig 3: Proposed Framework accuracy for MODI Script Handwritten Character Recognition over other methods.

RNNs are used for image captioning, where they take a sequence of image features and generate a textual description of the image. It can also be used for video analysis, where RNNs process a sequence of video frames to detect and track objects over time. The LSTMs are particularly useful for handling sequences of data with long-term dependencies. In image processing, LSTMs can be used for video frame prediction, anomaly detection in video streams, and generating sequences of images, such as predicting the next frame in a video sequence. Further, GAN is widely used in image analysis, style transfer, as well as super-resolution tasks. The GAN network consists of a generator as well as a discriminator that competes against each other. The GAN may generate realistic images from random noise, which can be used for image synthesis, data augmentation, and even creating deep fake images. Lastly, the DBN is responsible for unsupervised feature learning and dimensionality reduction, making them valuable for image pre-processing and feature extraction. It can also be utilized in the classification of image and object detection tasks, where the learned features are passed to a classifier or detector. Figure 3 shows the proposed framework accuracy for MODI Script handwritten character recognition over other methods. The proposed MODI script handwritten character recognition framework attains the highest accuracy on enhanced RNN and convolutional autoencoder which is observed as 88.00%. However, other techniques namely LTSM, GAN, and DBN obtain accuracy of 82.00%, 88.00%, and 72.00%, respectively. This comparison of accuracy signifies that the proposed model based on enhanced RNN and convolution autoencoder archives the largest accuracy for handwritten character recognition for MODI script image document from the novel proposed datasets containing a distinct class of the MODI lipi characters.

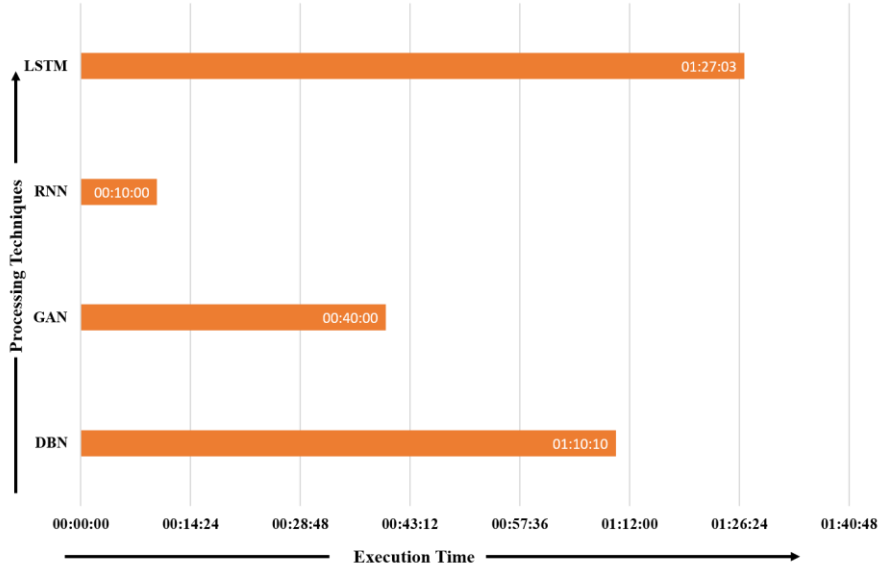

Fig 4: Proposed Framework Execution time analysis for MODI Script Handwritten Character Recognition over other methods.

Figure 4 depicts the proposed Framework execution time analysis for MODI Script handwritten character recognition over other methods. The effectiveness and efficiency of different image processing techniques in terms of execution time are shown and compared via graphical representations. The proposed enhanced RNN-based framework execution time analysis was done to compare our approach with the distinct other techniques namely the LTSM, GAN, and DBN. The proposed framework for handwritten MODI script image analysis for character recognition consumes very little time only 10 minutes over a large image dataset. The LSTM, GAN, and DBN approaches consume time of 01:27:03, 00:40:00, and 01:10:10, respectively. From the above observations, it is clear that our approach is optimal in terms of both time as well as prediction accuracy.

5. Conclusion

Handwritten documents of MODI scripts are very complex to analyze the characters due to the intricacy of the character's size and writing patterns. The previous handwritten image analysis models offer less recognition accuracy in the analysis of the MODI lipi images containing the varied vowels and consonant sets. This research proposed a new dataset creation method and effectively performs pre-processing and filtering of the handwritten MODI script images for accurate prediction of the characters in real-time. Further a novel framework based on enhanced RNN and convolutional autoencoder is presented for effective analysis of the handwritten MODI script characters. It is observed that our approach offers high accuracy in MODI script character recognition and less time consumption in the pre-processing of the MODI script images. In conclusion, this research investigates several methods for creating picture datasets for MODI script character identification and sheds light on the complexity and variety of distinct methods. The techniques used to generate synthetic data, supplement data, and collect it rely much on the particular needs and limitations of the character recognition system under consideration. Although techniques for creating synthetic data are flexible and scalable, they may not always fully represent the richness of real-world data. Character recognition models may be made more robust by adding variations to existing datasets using data augmentation approaches. However, in situations when domain-specific subtleties are important and high-quality ground truth annotations are required, human data gathering is still necessary. Character recognition systems ultimately depend on how well the selected approach is thought out, concerning the goals of the application and the available resources. Researchers and practitioners may help to enhance character recognition technology by making well-informed judgments by knowing the advantages and disadvantages of various dataset creation techniques. It is crucial to modify and improve dataset generation methods as the field develops to satisfy the rising needs of character recognition applications in a world that is becoming more and more digital.

References:

[1] S. Joseph and J. George, “Feature extraction and classification techniques of modi script character recognition,” *Pertanika Journal of Science and Technology*. 2019.

[2] Savitri Jadhav and Vandana Inamdar, “Convolutional Neural Network and Histogram of Oriented Gradient Based Invariant Handwritten

MODI Character Recognition," *Pattern Recognit. Image Anal.*, 2022, doi: 10.1134/S1054661822020109.

[3] S. S. Gharde and R. J. Ramteke, "Recognition of characters in Indian MODI script," in *Proceedings - International Conference on Global Trends in Signal Processing, Information Computing and Communication, ICGTSPICC 2016*, 2017. doi: 10.1109/ICGTSPICC.2016.7955304.

[4] R. K. Maurya and S. R. Maurya, "Recognition of a Medieval Indic-Modi Script using Empirically Determined Heuristics in Hybrid Feature Space," *Int. J. Comput. Sci. Eng.*, 2018, doi: 10.26438/ijcse/v6i2.136142.

[5] K. SadanandA., B. Prashant L., M. Ramesh R., and Y. Pravin L., "Review on Recent Advances in Automatic Handwritten MODI Script Recognition," *Int. J. Comput. Appl.*, 2015, doi: 10.5120/20257-2636.

[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," in *2020 IEEE 5th International Conference on Signal and Image Processing, ICSIP 2020*, 2020. doi: 10.1109/ICSIP49896.2020.9339435.

[7] A. K. Sadanand, L. B. Prashant, R. M. Ramesh, and L. Y. Pravin, "Offline MODI Character Recognition Using Complex Moments," in *Procedia Computer Science*, 2015. doi: 10.1016/j.procs.2015.08.067.

[8] S. Joseph, A. Datta, O. Anto, S. Philip, and J. George, "OCR system framework for modi scripts using data augmentation and convolutional neural network," in *Lecture Notes in Networks and Systems*, 2021. doi: 10.1007/978-981-15-5309-7_21.

[9] R. Shah, M. K. Gupta, and A. Kumar, "Line Level Modi (Heritage script) OCR using Attention based Encoder-Decoder Architecture," in *Proceedings of the IEEE International Conference Image Information Processing*, 2021. doi: 10.1109/ICIIP53038.2021.9702605.

[10] S. Sawant, A. Sharma, G. Suvarna, T. Tanna, and S. Kulkarni, "Word Transcription of MODI Script to Devanagari Using Deep Neural Network," in *2020 3rd International Conference on Communication Systems, Computing and IT Applications, CSCITA 2020 - Proceedings*, 2020. doi: 10.1109/CSCITA47329.2020.9137781.

[11] P. A. Tamhankar, K. D. Masalkar, and S. R. Kolhe, "Character Recognition of Offline Handwritten Marathi Documents Written in MODI Script Using Deep Learning Convolutional Neural Network Model," in *Communications in Computer and Information Science*, 2021. doi: 10.1007/978-981-16-0507-9_40.

[12] A. K. Sadanand, L. B. Prashant, R. M. Ramesh, and L. Y. Pravin, "Impact of zoning on Zernike moments for handwritten MODI character recognition," in *IEEE International Conference on Computer Communication and Control, IC4 2015*, 2016. doi: 10.1109/IC4.2015.7375516.

[13] P. S. Chidrawar and V. Dhamdhere, "A Systematic Review on MODI Script Character Recognition," 2023. doi: 10.1007/978-981-99-3878-0_24.

[14] A. Shirkande and A. Agarwal, "A Review on Various MODI Text Recognition Techniques," *J. Image Process. Artif. Intell.*, 2023, doi: 10.46610/joipai.2023.v09i01.001.

[15] D. N. Besekar, "Special Approach for Recognition of Handwritten MODI Script's Vowels," *Int. J. Comput. Appl.*, 2012.

[16] S. Bhalerao and H. D. Gadade, "A systematic review on advancement in Modi Handwritten Character Recognition," in *Recent Advances in Material, Manufacturing, and Machine Learning*, 2023. doi: 10.1201/9781003370628-44.

[17] P. A. Patil, "Character Recognition System for Modi Script," *Int. J. Comput. Eng. Res.*, 2016.

[18] I. Hussain, R. Ahmad, S. Muhammad, K. Ullah, H. Shah, and A. Namoun, "PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3216881.

[19] M. Ahmed, M. A. Awel, and A. I. Abidi, "REVIEW ON OPTICAL CHARACTER RECOGNITION," *Int. Res. J. Eng. Technol.*, 2019.

[20] Z. Jia, S. Xu, S. Mu, and Y. Tao, "Learning-Based Text Image Quality Assessment with Texture Feature and Embedding Robustness," *Electron.*, 2022, doi: 10.3390/electronics11101611.

[21] S. Patil *et al.*, "Enhancing Optical Character Recognition on Images with Mixed Text Using Semantic Segmentation," *J. Sens. Actuator Networks*, 2022, doi: 10.3390/jsan11040063.

[22] S. Chandure and V. Inamdar, "Handwritten MODI Character Recognition Using Transfer Learning with Discriminant Feature Analysis," *IETE J. Res.*, 2023, doi: 10.1080/03772063.2021.1902867.

[23] A. Institutions, "Bharat Itihas Sanshodhak Mandal." https://map.sahapedia.org/article/Bharat-Itihas Sanshodhak Mandal/2769