2023 International Conference on Communication System, Computing and IT Applications (CSCITA) | 978-1-6654-5987-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/CSCITA55725.2023.10104985

Deep Learning Based Marathi Sentence Recognition using Devnagari Character Identification

Rupali Patil
Electronics and Telecommunication Engineering
K.J. Somaiya College of Engineering
Somaiya Vidyavihar University
Mumbai, India
rupalipatil@somaiya.edu

Bhairav Narkhede
Computer Engineering
K.J. Somaiya College of Engineering
Somaiya Vidyavihar University
Mumbai, India
bhairav.n@somaiya.edu

Stuti Gaonkar
Electronics and Telecommunication Engineering
K.J. Somaiya College of Engineering
Somaiya Vidyavihar University
Mumbai, India
stuti.g@somaiya.edu

Tirth Dave
Electronics and Telecommunication Engineering
K.J. Somaiya College of Engineering
Somaiya Vidyavihar University
Mumbai, India
t.dave@somaiya.edu

***Abstract*—There are multiple algorithms available to recognize Marathi Devnagari characters. Most of these methods are limited because of the large variety of character variations due to Kana, Matra, Ukar, Velanti, and Anusvar, which are specific to the Marathi grammar called Barakhadi. There is a need to have a dictionary-based word formulation to achieve full Marathi sentence recognition. In the proposed work, a Marathi sentence is recognized using a combination of full 454 variation detection of Devnagari characters and nearest dictionary word mapping using the k-nearest neighbour (KNN) model to achieve full sentence recognition. This is the first time full 454 (Vyanjan variation as per Barakhadi) character recognition instead of the traditional 58 characters (Vyanjans) has been attempted which leads to sentence recognition. The proposed method could achieve a sentence recognition accuracy of 86.84%, a 454 character classification accuracy was 89.52%, and the execution speed of the proposed system was 1.464 secs per word. For the training of the character recognition network, a separate dataset was created for all Vyanjan variations as per Barakhadi. This novel contribution of the proposed system will surely inspire researchers to explore Devnagari sentence recognition.**

***Index Terms*—Devnagari sentence recognition, Character recognition, Marathi Barakhadi, KNN classification**

I. Introduction

In Marathi sentences, a large number of feature variations and complex grammatical variations are expected. It is a very difficult problem because none of the existing technologies developed were focusing on sentence recognition. The computation complexity of sentence recognition is way more than Marathi character recognition. There are multiple efforts made to recognize the Marathi Devnagari characters, such as Customized CNN as shown by D. T. Mane *et al.* [1]. Only recognition of character may not help in any way as an entire sentence might carry a different meaning. Hence, the logical step is to use the existing Marathi character recognition methodologies and perform sentence recognition. There are 36 Vyanjans and 12 Swars (vowels) which makes up a total of 48 Mulakashar. Hence, it is quite difficult to produce a good character recognition system with 36 X 12 + 12 + 10 (454) different characters. In the dataset, numbers from 0 to 9 in Marathi were also considered. Although the deep learning technologies recently developed can take care of this 454-class classification, there are further special Jodakshar that make it impossible to do a standard optical character recognition.

Marathi is a language spoken by 120 million people and is ranked 10 th in the list of native speakers. The oldest literature in the Marathi language was found in 600 AD. Various dialects of Marathi exist apart from the standard spoken Marathi. There are Koli, Agari, Malvani, and Konkani dialects which make the Marathi language processing a huge challenge for the research community as mentioned in the paper by S. Kayte *et al.* [2]. Marathi is derived from the early forms of the 'Prakrit' language as shown by P. Ajmire *et al.* [3]. There are stone carvings found in Marathi which are dated to approximately 3 century BCE. There are a total of 42 spoken dialects of Marathi. Marathi is typically written with the Balbodh version of the Devnagari script with 36 consonant letters and 16 vowels. The proposed system was designed to recognize Marathi sentences from images of handwritten sentences. The proposed system consists of a Marathi character recognition network and a Marathi dictionary matching network. Initially, the character recognition network segments individual characters from the input image of the handwritten Marathi sentence. Then the network performed 454-class classification on each of the segmented characters. The output of the multi-class classification was then used to generate the digitized version of the recognized characters. The recognized digital characters were fed to the Marathi dictionary matching network. The

978-1-6654-5987-7/23/$31.00 ©2023 IEEE

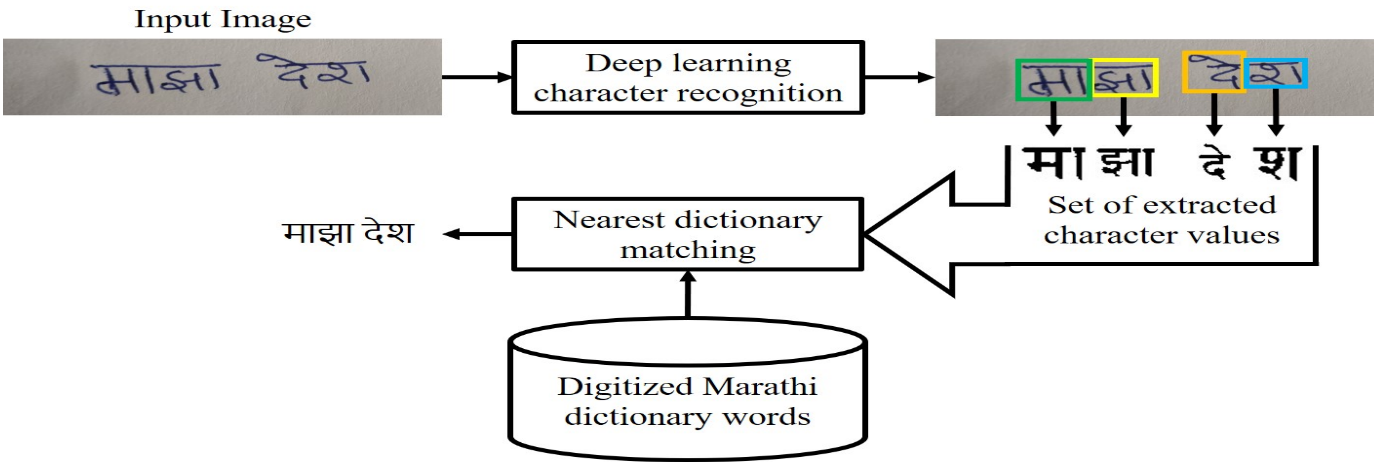


Fig. 1: Flow diagram for sentence recognition using deep learning from handwritten Marathi records. The handwritten Marathi text (sentences) is scanned and given as input to a deep-learning network for single-character extraction. Then the nearest matching network is used for the extraction of digitized sentences using the set of extracted values from the handwritten Marathi sentences and digitized Marathi dictionary words.

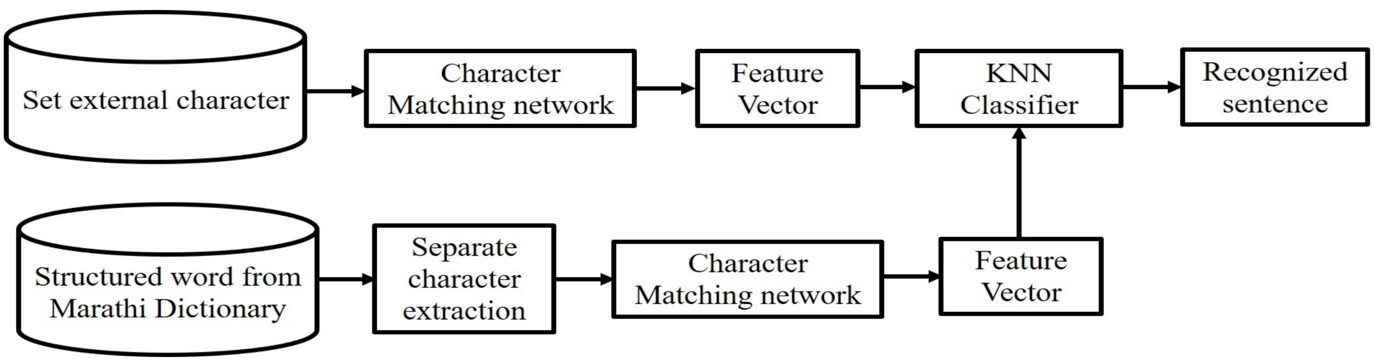
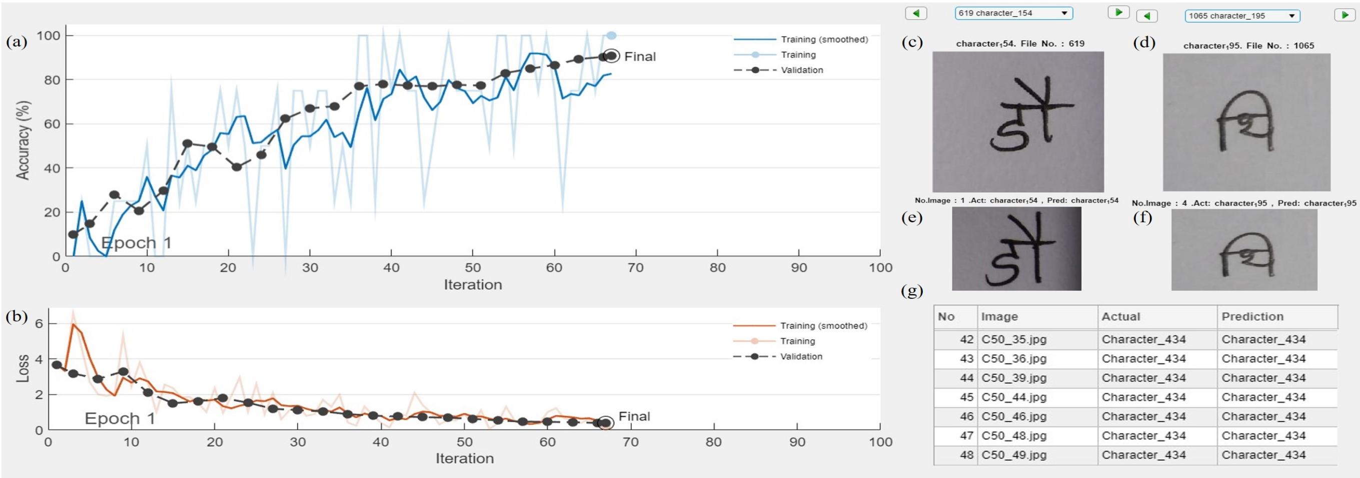


Fig. 2: The flow diagram for extraction of a digitized sentence from a handwritten Marathi sentence. The digitized character extracted from handwritten Marathi sentences is fed to a character-matching network for feature vector extraction. The feature vector is also extracted from the structured word of the Marathi Dictionary after the separation of each character using a character-matching network. The feature vector of characters extracted from handwritten Marathi sentences and the feature vector of structured word of the Marathi Dictionary is used by the KNN classifier for digitized sentence recognition.

matching network found the word which closely matches the input digital characters. The same procedure was applied to each word in the handwritten input image. Finally, the matching network outputs the digitized version of the input handwritten sentence. It is notable to mention that almost 120 languages were derived from the Devnagari script, hence the proposed work done for the Marathi language can be extended to almost 120 languages to create sentence recognition systems for those languages.

II. Literature Review

P. M. Kamble *et al.* [4] have used Rectangle Histogram Oriented Gradient (RHOG) for handwritten Marathi character recognition. They have used 40 different characters with 8000 sample datasets. 10 different handwriting were collected and each character is normalized to 20X20 pixels. They have used Support Vector Machine (SVM) and feed-forward Artificial Neural Network (ANN). They achieved an accuracy of 97.15 %. D. Chikmurge textit *et al.* [5] has developed a SVM and k-nearest neighbour (KNN) based Optical Character Recognition (OCR) system for Marathi handwritten characters. They have used Histogram Oriented Gradient (HOG) method with preprocessing steps such as segmentation of characters. As mentioned by M. Agrawal *et al.* [6], the complex structural properties of the Marathi script create a lot of issues for handwritten optical character recognition (OCR) which is not present in most other scripts. S. P. Ramteke *et al.* [7] has shown the OCR framework for handwritten Marathi document classification and recognition system. They have used morphological operations in the preprocessing and modified the Pihu method for three levels of segmentation; line segmentation, word segmentation, and character segmentation. They achieved high accuracy of 95.14 % for the SVM classifier. Their group also proposed the Adaptive Cuckoo Search (ACS) algorithm



No	Image	Actual	Prediction
42	C50_35.jpg	Character_434	Character_434
43	C50_36.jpg	Character_434	Character_434
44	C50_39.jpg	Character_434	Character_434
45	C50_44.jpg	Character_434	Character_434
46	C50_46.jpg	Character_434	Character_434
47	C50_48.jpg	Character_434	Character_434
48	C50_49.jpg	Character_434	Character_434

Fig. 3: Recognition of 454 Marathi handwritten characters. (a) Graph of training and validation accuracy for character matching network. The accuracy for training and validation keeps increasing with successive iterations. The final training and validation accuracy were 100 % and 89.52 %, respectively. (b) Graph of training and validation loss for character matching network. The loss for training and validation keeps decreasing linearly. The final training and validation loss were 0.3 and 0.6, respectively. (c-d) Input handwritten character images of size 227 X 227 X 3 pixels were used for training. (e-f) Corresponding handwritten character numbers are predicted by the character-matching network. (g) The output of the character-matching network shows the actual and predicted character numbers.

for the optimization procedure after the feature extraction process as mentioned by S. P. Ramteke *et al.* [8]. They have used a projection profile segmentation technique that generates less error. They have used the real-time Marathi character datasets and achieved 99.36 % accuracy, 90 % sensitivity, 91 % precision, 89.51 % recall, 99.67 % specificity, and 89.93 % F-score compared to the existing FireFly Selection (FFS) and Bat Selection (BBS) approach. The research paper by Y. Baek *et al.* [9] provides a method for effectively detecting text areas by analysing each character and the affinity between characters.

D. T. Mane *et al.* [10] has used a stacked ensemble neural network to determine the handwritten Marathi numerals. The stack ensemble method uses a pre-trained-based pipeline to create a multi-head meta-learning classifier. The exact results are concatenated instead of averaged. S. T. Deokate *et al.* [11] used CNN classification for analysis and recognition of the Marathi manuscript. They have used KNN with the CNN approach to recognize the Marathi script. They could achieve a validation accuracy of 99 % with 3600 files in the dataset and a batch size of 128 files per epoch. They carried out the training for 20 epochs.

S. S. M. N. Akhter *et al.* [12] have shown a semantic segmentation of Marathi text using deep learning methods. They have used standard U-Net and residual U-Net (ResU-Net) for classifying Marathi documents. Their dataset was extracted from various Marathi books. The skip connections avoid vanishing gradient problems and give 95 % and 98 % accuracy for U-Net and ResU-Net, respectively. S. Prabhanjan *et al.* [13] has proposed deep learning architecture for the recognition of handwritten Devnagari script. It is very difficult to determine an optimal set of good features to get a good recognition rate for handwritten Devnagari characters. The proposed deep learning architectures learn hierarchies of features using unsupervised stacked Restricted Boltzmann Machines (RBM). Fine-tuning of the network parameters with supervised learning yields an accuracy of 91.81 % when tested on a large set of handwritten numerical, character, vowel modifiers, and compound characters.

Most of these reported literature papers have only shown character extraction and recognition but none of the work has been carried out into a sentence-level extraction.

III. Methodology

Figure 1 shows the overall flow diagram for the Marathi sentence recognition using deep learning character matching network and Marathi sentence dictionary matching network. The YOLOv5 model segmented the input image i.e a handwritten Marathi sentence to extract the individual characters of all the words in the sentence. The YOLOv5 algorithm was trained using 44 images of which 89 individual words were marked in each image. MATLAB function, Image Labeler was used to mark individual words on the ground truth image. The images of the segmented handwritten Marathi characters were fed as input to a 454-class character-matching network. The set of swapping blocks (i.e. sub-image generated by cropping the main image) was compared to the database of 454 characters written in Devnagari. If the threshold crossed 80% then the character was recognized otherwise it was discarded. This sub-image was swapped throughout the input image to extract all the possible characters in an image. For the post-processing, the x-y dimensions of some images were considered and

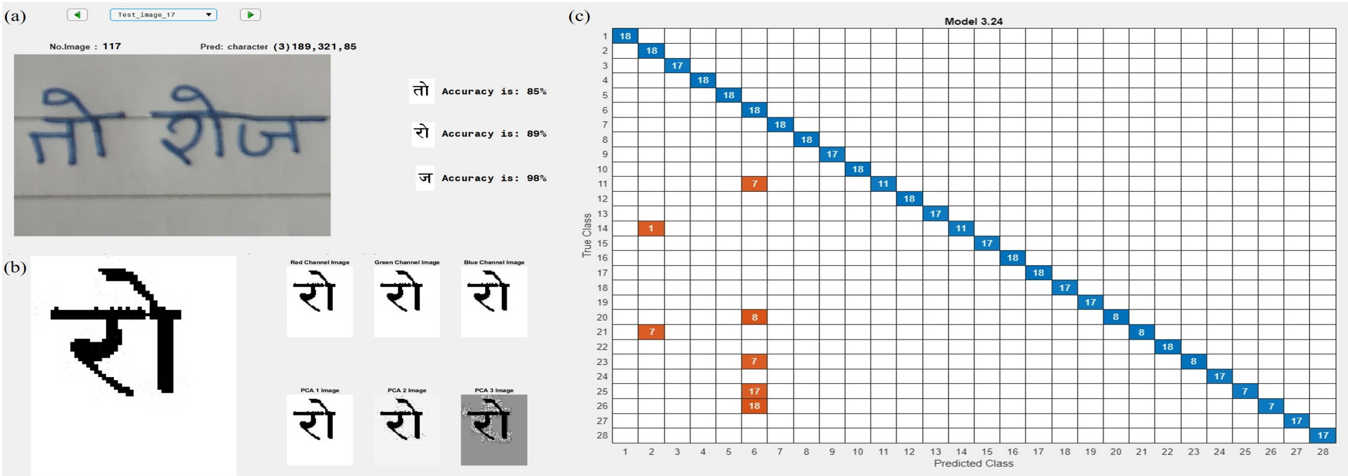


Fig. 4: (a) The image of a handwritten sentence was segmented to extract images of individual handwritten characters. These extracted handwritten character images were given as input to the character-matching network. The network gives the corresponding character numbers of the recognized handwritten characters. An indexed list of all 454 digitized characters was created. Each index was mapped to a digital image of a Marathi character. (b) The PCA components are extracted from digitized characters as well as from structured words from Marathi Dictionary and provided as input to the KNN classifier. (c) Confusion matrix for sentence recognition using KNN classifier for a total of 28 sentences (each sentence has unique characters) dataset used for validation. The accuracy achieved was 86.84 %.

average dimensions with the top 90% of sub-images are filtered. If the noise wasn't removed, most extracted characters would be garbage values. To understand the word from the character flat horizontal line on the top of the word was used. All the pixels belonging to a single horizontal line are grouped as a set of extracted character values. Hence the extracted characters were grouped as words. The extracted characters obtained at the output of the network were mapped to one of the 454 stored digital images. All the digital characters belonging to the recognized characters from the handwritten sentence were grouped to form digital words.

Using principal component analysis (PCA), feature vectors of the grouped digital characters (digital words) were calculated. These feature vectors were given as input to the nearest dictionary-matching network. A list of a total of 165 digital Marathi words was prepared. Out of the total number of words, 86 words were used for training, 29 words were used for testing, and 50 words were used for validation. Using PCA the feature vectors of all these words were also stored in three separate indexed lists. The nearest matching network was created using KNN. The KNN-based model compares the feature vectors of the grouped digital characters with all the feature vectors stored in the indexed list. The index of the feature vector matching the closest to the input feature vectors was stored. The corresponding digital Marathi word of the stored index was the output of the nearest matching network.

The flow diagram for extracting a digitized sentence from a handwritten Marathi sentence is shown in figure 2. A character-matching network was used to generate feature vectors from the digitized characters taken from the handwritten Marathi text. The feature vector was also extracted from the structured words of the Marathi dictionary after each character was separated using a character-matching network. KNN classifier used the feature vector of characters extracted from handwritten Marathi sentences and the feature vector of structured words from the Marathi Dictionary for digitized sentence recognition.

IV. RESULTS AND DISCUSSION

Figure 3 shows the various components of the character-matching network for classifying 454 Marathi handwritten characters. Figure 4 shows the outputs of the nearest dictionary-matching network at various stages. A total of 28 sentences (each sentence includes unique characters) were used for generating the confusion matrix. 86.84 % accuracy was achieved using the KNN-based classifier as shown in figure 5.

Different types of character recognition systems were compared to the proposed system in table I. As compared to the datasets used by the systems reported in the literature, the proposed system used the largest dataset that consisted of 11004 images. The highest accuracy reported in the literature was 99.36 % by Ramteke *et al.* [8], while the accuracy of the proposed system was 86.84 %. None of the reported literature had classified all 454 Marathi characters instead, they had generally classified 58 characters. As the proposed system was classifying all 454 characters its accuracy value was lower than the values reported in the literature. A specificity of 99.92 % and a sensitivity of 96.9 % were achieved using the proposed system. The sensitivity and specificity of the proposed system were greater than all of the systems reported in the literature.

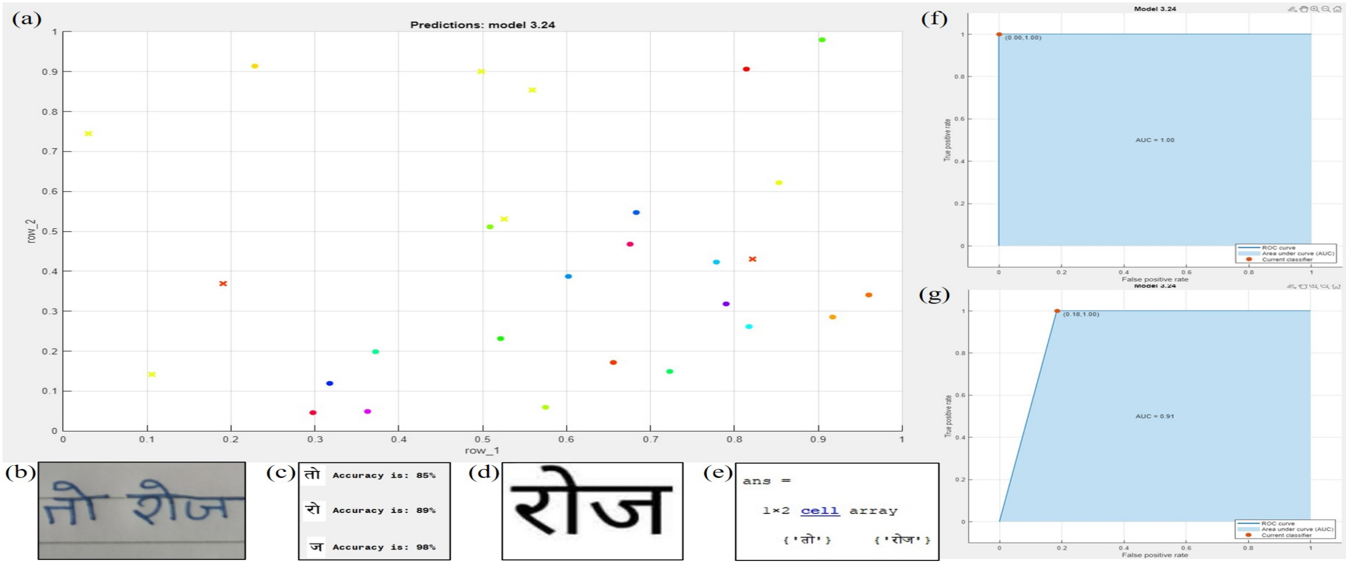


Fig. 5: (a) The scatter plot of principal components for all 28 words in the dictionary. (b) Handwritten input image of a Marathi sentence with a length of two words. (c) The output of the character matching network after digital mapping. (d) Nearest matching image from the dictionary whose PCA components are the nearest matching to the PCA components of the set of characters in the second word of the input image. (e) Final Marathi sentence recognized in terms of string cell array after Marathi dictionary mapping. Two sample ROC curves out of 28 different classes have AUC of 1.00 (f) for class 1, and AUC of 0.91 (g) for class 6.

TABLE I: Comparison of different types of character recognition systems with the proposed system. The proposed system was trained using the 11004 images which were greater than the training dataset of any reported literature. The proposed system has the greatest specificity (99.92 %) and sensitivity (96.9 %) among all the systems reported in the literature.

Method	Number of images	Accuracy (%)	Specificity (%)	Sensitivity (%)
Kamble *et al.* [4]	8000	97.15	-	-
Akhter ResU-Net *et al.* [12]	-	98	-	-
Akhter U-Net *et al.* [12]	-	95	-	-
Ramteke *et al.* [7]	-	95.14	-	-
Deokate *et al.* [11]	3600	99	-	-
Ramkete *et al.* [8]	-	99.36	99.67	90
Prabhanjan *et al.* [13]	-	91.81	-	-
Proposed method	11004	86.84	99.92	96.9

A. Discussion

The proposed system identified characters from handwritten scanned images using YOLO v5. Stronger approaches than ResNet-18 can enhance character detection accuracy. After character detection, matching with 454 classes was difficult, so people can classify into 58 classes and then use a basic Barakhadi classifier. The suggested technology detects 1 word every 1.5 seconds. Given modern processing power, this speed is insufficient, but there was a trade-off between accuracy and speed. Researchers can use the strategies to optimize speed vs. accuracy. There were a lot of issues when using PCA for dictionary matching as there were very closely related words such as ’dhaar’ and ’ghar’, also ’maaz’ and ’bhaaz’, etc as ’m’ and ’bh’, and ’dh’ and ’gh’ are closely related. The major achievement of the proposed system was creating a deep learning classifier that can classify into 454 categories in a single go, that too using a user-defined network, which can be called as DevnagariNet. The proposed system worked with at least 10 images per class and used auxiliary data to generate further images. Students typically write thousands of times the same word when they are learning, so a fast scanning mechanism can help build such a large dataset. Though the proposed system used an open-source dictionary it was difficult to get the Marathi character as a variable as its ASCII does not exist with standard programming languages such as MATLAB. Future directions of the proposed system could be live translators using computer vision and text-to-speech translators for blind Marathi people.

B. Novelty of the proposed method

Literature has mostly focused on recognising Devnagari characters. Since Devnagari sentences usually contain Swaras and Jodakshar, it was not enough to recognize the basic letters. We developed a dataset with 10 images of each 454 Devnagari characters, including all basic Swaras and Jodakshar. Previous investigations only recognized Devnagari characters. Sentence recognition in Devnagari or other scripts has not advanced. One of the first Devnagari sentence recognition attempts is proposed. This challenge needs efficient character recognition and dictionary matching to group characters into words.

V. Conclusion

For recognizing Marathi Devnagari characters, there are several algorithms available. To achieve full-sentence recognition in the proposed study, a Marathi sentence is recognized using a combination of complete 454 variation detection of Devnagari characters and nearest dictionary word mapping using a KNN model. The proposed technique achieved an accuracy of 86.84 % for sentence identification and 89.52 % for 454 character categorization. The proposed system's execution speed is 1.464 seconds per word. A specificity of 99.92 % and sensitivity of 96.9 % are achieved, which is greater than the systems reported in the literature. The proposed system's innovative contribution will undoubtedly stimulate scholars to investigate Devnagari sentence recognition. By digitizing Marathi phrases, a vast amount of printed material can be digitized and made accessible to a wider audience.

References

[1] D. T. Mane and U. V. Kulkarni. Visualizing and understanding customized convolutional neural network for recognition of handwritten marathi numerals. *Procedia computer science*, 132:1123–1137, 2018.

[2] S. Kayte, M. Mundada, and C. Kayte. Di-phone-based concatenative speech synthesis systems for marathi language. *OSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume*, 5:76–81e, 2015.

[3] P. Ajmire and S. E. Warkhede. *Handwritten Marathi character (vowel) recognition*. Lap Lambert Academic Publ, 2012.

[4] P. M. Kamble and R. S. Hegadi. Handwritten marathi character recognition using r-hog feature. *Procedia Computer Science*, 45:266–274, 2015.

[5] D. Chikmurge and R. Shriram. Marathi handwritten character recognition using svm and knn classifier. In *International Conference on Hybrid Intelligent Systems*, pages 319–327. Springer, 2019.

[6] M. Agrawal, B. Chauhan, and T. Agrawal. Machine learning algorithms for handwritten devanagari character recognition: A systematic review. *vol*, 7:1–16, 2022.

[7] S. P. Ramteke, A. A. Gurjar, and D. S. Deshmukh. A novel weighted svm classifier based on sca for handwritten marathi character recognition. *IETE Journal of Research*, pages 1–13, 2019.

[8] S. P. Ramteke, A. A. Gurjar, and D. S. Deshmukh. A streamlined ocr system for handwritten marathi text document classification and recognition using svm-acs algorithm. *International Journal of Intelligent Engineering and Systems*, 11(3):186–195, 2018.

[9] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee. Character region awareness for text detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9365–9374, 2019.

[10] D. T. Mane and Swati V Tapdiya, R. S. Handwritten marathi numeral recognition using stacked ensemble neural network. *International Journal of Information Technology*, pages 1–7, 2021.

[11] S. T. Deokate and N. J. Uke. Cnn classification approach for analysis and recognition of marathi manuscript. *International journal of scientific & technology research*, 2020.

[12] S. S. M. N. Akhter and P. P Rege. Semantic segmentation of printed text from marathi document images using deep learning methods. In *2019 IEEE 16th India Council International Conference (INDICON)*, pages 1–4. IEEE, 2019.

[13] S. Prabhanjan and R. Dinesh. Deep learning approach for devanagari script recognition. *International Journal of Image and Graphics*, 17(03):1750016, 2017.