

Devanagari Character Recognition

Rajat Belgundi

*School of Computer Engineering and Technology,
Dr. Vishwanath Karad MIT WPU
Pune,India
rajatbelgundi@gmail.com*

Divyang Bagla

*School of Computer Engineering and Technology,
Dr. Vishwanath Karad MIT WPU
Pune,India
bagladivyang03@gmail.com*

Kartik Bhutada

*School of Computer Engineering and Technology,
Dr. Vishwanath Karad MIT WPU
Pune,India
kartikbhutada21@gmail.com*

Amey Bhide

*School of Computer Engineering and Technology,
Dr. Vishwanath Karad MIT WPU
Pune,India
ameybhide2406@gmail.com*

Abstract— Character recognition has been a topic studied by quite a few researchers which includes use of neural networks to recognize the individual characters from the images. The correctness of recognition is provided by the neural network as a part of the classification performed by the neural network.

In this paper, a tool(web-application) is proposed, which assists the user to learn how to write devanagari characters. The written character's correctness is validated by the prediction of a Convolutional Neural Networks(CNN) which helps to identify whether the user has written the character correctly or not.

Keywords— Devanagari Characters, Character recognition, CNN, Web Application, activation function,Dropout

I. INTRODUCTION

Literacy is critical to economic development as well as individual and community well-being. Our economy is enhanced when learners have higher literacy levels. Hence, literacy is extremely crucial for all. The people who can read and write find it easy to lead life in this fast paced world.

For the people in India, it is important to know Hindi as it is the national language of India.Hindi language is written in the Devanagari. In order to read and write Hindi one must know how to read and write devanagari characters.

In the current pandemic situation,remote learning has become of immense importance. Our proposed tool will be able to help in improving literacy and by

adhering to the current remote learning demand. The current generation of young kids is inclined more towards electronic devices than pen-paper, hence, this tool will be able to reduce the burden on parents(working) as the small kids can open the application and start learning.

The challenges faced during the pandemic of how to teach has forced us to find new solutions for developing applications for remote learning. Also, the new situation has forced people to think if we can set up a complete environment for remote learning and working in future so as to be prepared for further pandemic-like situations.

Artificial Neural Networks are gaining prominence with its vast set of applications in various domains with applications ranging from character recognition to self-driving cars. There lies a huge potential to use ANN based Convolutional Neural Networks (CNN) in ed-tech based applications to assist students to learn better. As our tool involves teaching how to write characters the model was trained on images which is why CNN was the best way to proceed in developing the tool.

Seeing the potential of CNN in edtech, we have designed our tool as a web based application to help children teach how to write devanagari characters and check whether they have written the character correctly or not. Here we have proposed a basic prototype of our tool. We have used CNN and trained

our model on the dataset containing around 92000 images of devanagari characters of 32x32 size and achieved an accuracy of around 98%.

The paper is structured as follows: In section II, we have reviewed the work done in this field which provides the base for our proposed idea and assists in developing the tool.

Section III consists of the methodology for our proposed tool in which all the steps taken to develop the tool have been explained in detail and with use of necessary figures it is made more clear to the reader.

The results of this work along with a discussion of a few key points regarding the results are included in section IV. The section V consists of future scope which has key points about how this approach can be used further and further additions to this tool are discussed. Section VI gives a conclusion to our work.

II. Literature Review

Researchers have been working for quite some time in this domain.

On conducting a literature review we have the following understandings:

The paper[1] helped us to know about CNN in which the authors have tried to bridge the gap between the mathematical structure and the computational implementation of a convolutional neural network (CNN) using a minimal model (Minimal CNN). The proposed minimal CNN is presented using a layering approach.

For our proposed tool, the dataset was taken from [2] in which the authors have introduced a public image dataset of Devanagari script. The dataset consists of 92 thousand images of 46 different classes of characters of Devanagari script segmented from handwritten documents. Also, the challenges in recognition of Devanagari characters have been explored. The authors have also proposed a CNN architecture focused on Dropout and dataset increment approach to improve the accuracy.

This dataset is publicly available on [3].

We were also motivated by this study [4], which focuses on the categorization of Devanagari Script handwritten numbers. The main contributions of this paper are 1) the development of a dataset for

handwritten numerals that is similar to the MNIST dataset, 2) the analysis of Pattern Recognition tools based on Neural Network and Convolution Neural Network, and 3) a detailed discussion of the results by calculating Precision, Recall, and F-measure values and comparing them to other datasets available online. The collection contains 4,282 Devanagari handwritten numbers. The method used in this study was shown to be more than 95% accurate.

Papers[9][10][11][12][13] discuss devanagari character recognition using Convolutional Neural Network techniques.

Here, in this paper we discuss a CNN based application to help students learn to write devanagari characters.

III. Proposed Methodology

Learning to write a new language requires time, resources and a lot of practice. A lot of manpower too needs to be invested to teach a particular skill. With the pandemic situation, schools have been shut and online assistance is the only way to ensure that students don't miss out on education, still the students miss out on practical aspects of the learning. Artificial Intelligence based applications can overcome this situation where AI will help and monitor practical skill based assessments. Here, we have developed an application which will help children and anyone who wants to learn how to write devanagari characters and monitor them. This involved building a CNN model being trained on a lot of images. The training part for the CNN architecture has been explained in Fig 1.

1. Data Collection

The dataset has been taken from [3] which is publicly available. The dataset consists of 92 thousand images of 46 different classes of characters of Devanagari script segmented from handwritten documents.

2. Data Preparation

The dataset was run through a Python script to create data for training and testing. The training and testing data were then converted to proper dimensions which ensures easy usage of data for model building.

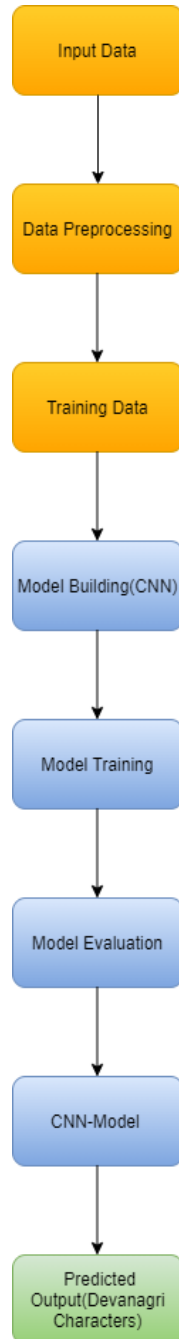


Fig. 1: CNN life cycle

3. Model Building

Our proposed tool uses a Convolutional Neural Network(CNN) to read image data and predict the correctness of the written character.

Before building our model we have to normalize all the pixel values which is done by dividing pixel value by 255.

The normalised data is obtained by using the above mentioned technique on our training and test data.

3.1 Model architecture

The model is built sequentially by adding layers one at a time. The initial layer is the input layer with filters and activation function.

The next layer is a convolutional layer followed by a pooling layer with appropriate parameters. Subsequently, another convolutional layer followed by a pooling layer is added. This is followed by adding Dropout layer (deactivation of 20% neurons).The next layer is a flattening layer which will flatten the data for inputting to the next layer. We then add two dense(fully connected) layers before the final output layer. Finally, we add the output layer with the number of neurons equal to the number of characters in the dataset. The 'adam' optimizer along with 'categorical cross entropy' were used to build the CNN.

The model was trained for 10 epochs and was then saved for further use in the web application.

4. Web Application

The proposed tool displays a character that is to be drawn in the sketchpad and the character drawn in the sketchpad is captured and sent to the model to predict the character. The character drawn and the character displayed is compared and checked whether they match or not. If they match, a correct symbol is displayed, else wrong symbol is displayed. The user can clear the sketchpad or choose to try some other character he/she wishes to try on. (optional) The model uses Django in the backend to process the image and HTML in the frontend.The tool also has a voice feature which delivers the result.

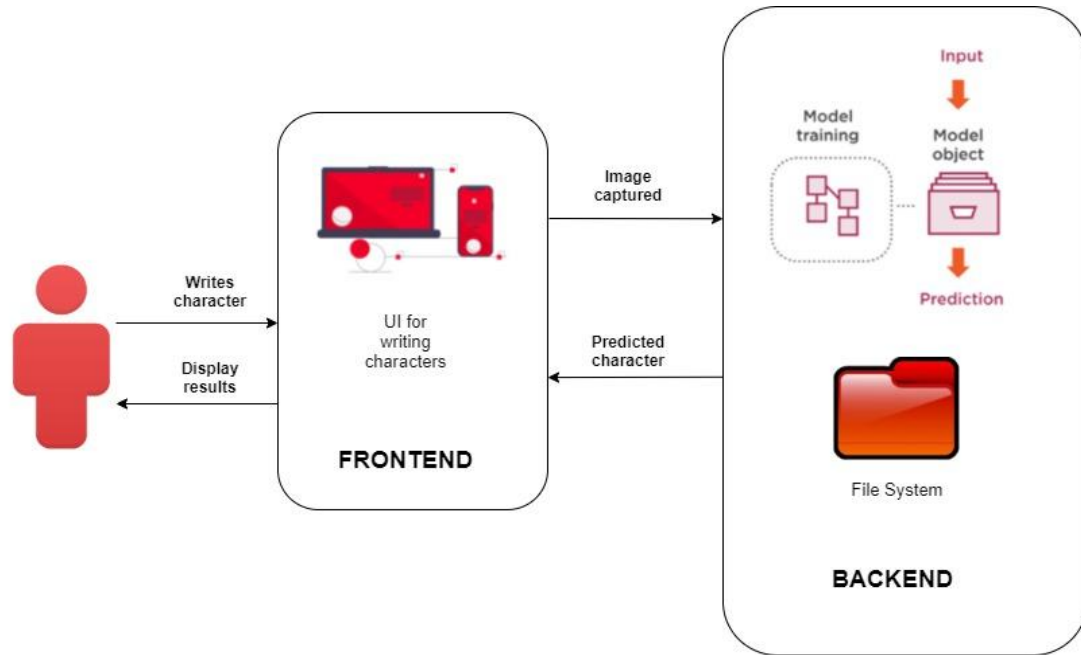


Fig. 2: Overall Architecture

working of the application. With Our model, we have achieved an accuracy of 98% for the dataset which predicts the drawn characters accurately. Accuracy report can be seen in Fig 5.

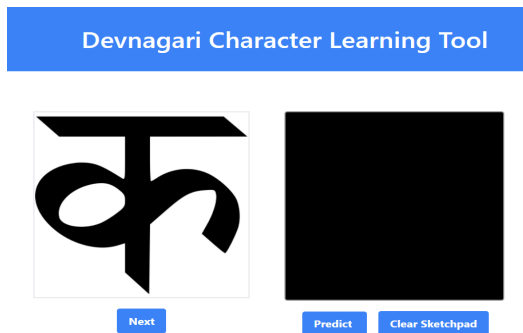
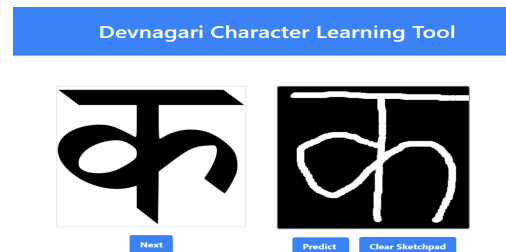


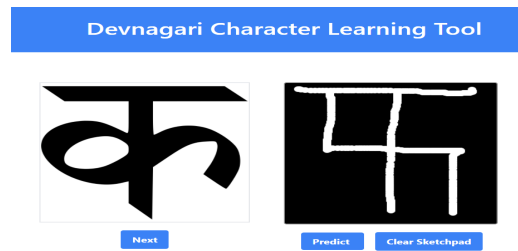
Fig. 3:3: Web application

IV. Results and Discussion

The proposed tool will help students/ children or anyone wishing to learn how to write Hindi characters and help them to practise and write better. With the model trained on handwritten characters it is reliable and hence can be used for daily practice as well. With this, a lot of potential can be seen in Neural networks based architecture and it is yet to be explored with different applications to achieve solutions for various problems. Fig 4. Shows the

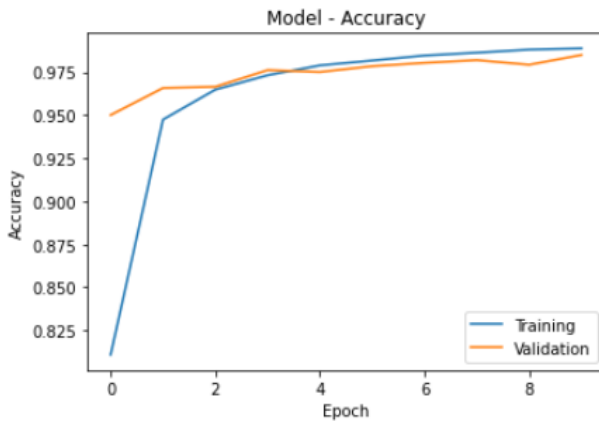


(a) Model predicts written character as correct

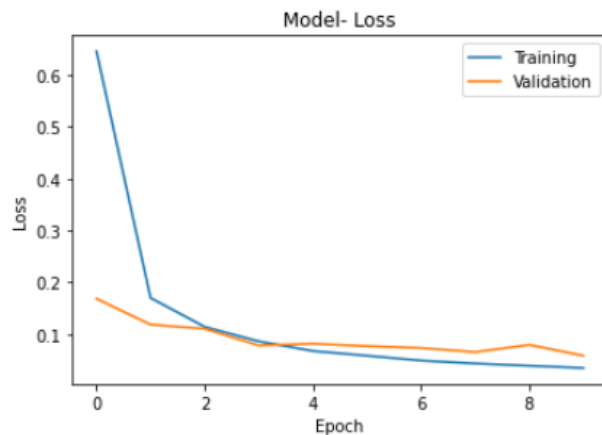


(b) Model predicts written character as wrong

Fig. 4: Working of application



(a) Model Accuracy report



(b) Model Loss report

Fig. 5: Validation & Loss curves

V. Future Scope

Feedback Systems - A feedback system can be introduced which will help the user to know how accurately he/ she is writing a character based on a beauty score predicted by the probability of the predicted character by the model.

Support for multiple languages - Similarly, multiple datasets of other languages' scripts can be trained and an option can provide what language the user wants to learn and practise writing.

A scoring based system can be made - The learning activity can be made more fun by introducing a quiz-like training mode in which the number of correct characters will be scored and shown to the user so as to improve based on his/her score.

A voice based assistant can be introduced to make the application more interactive and insightful.

Here, we have just built a web application, a mobile based application can also be developed to improve the reach of the tool.

Advanced CNNs for character recognition can also be used to make the application more accurate and precise and also to ensure use of the latest technology to improve the application.

VI. Conclusion

So in this paper we have discussed our implementation of a web based application which uses CNN to teach how to write devanagari characters. The model is still not completely accurate and can be improved by training on more data. The applications have a lot of scope in remote learning and AI based teaching assistants. The web application is very suitable for anyone wanting to learn to write Hindi characters

REFERENCES

- [1] Teow, M. Y. (2017, October). Understanding convolutional neural networks using a minimal model for handwritten digit recognition. In *2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 167-172). IEEE.
- [2] Acharya, S., Pant, A. K., & Gyawali, P. K. (2015, December). Deep learning based large scale handwritten Devanagari character recognition. In *2015 9th International conference on software, knowledge, information management and applications (SKIMA)* (pp. 1-6). IEEE.
- [3] Acharya, S., Pant, A. K., & Gyawali, P. K., Devanagari Handwritten Character Dataset Data Set, Available on- <https://archive.ics.uci.edu/ml/datasets/Devanagari+Handwritten+Character+Dataset>
- [4] Prashanth, D. S., Mehta, R. V. K., & Sharma, N. (2020). Classification of Handwritten Devanagari Number—An analysis of Pattern Recognition Tool using Neural Network and CNN. *Procedia Computer Science*, 167, 2445-245.
- [5] Tang, C., Zhu, Q., Wu, W., Huang, W., Hong, C., & Niu, X. (2020). PLANET: improved convolutional neural networks with image enhancement for image classification. *Mathematical Problems in Engineering*, 2020.
- [6] Sushma, L., & Lakshmi, K. P. (2020). An Analysis of Convolution Neural Network for Image Classification using Different Models.

International Journal of Engineering Research & Technology, 9(10).

- [7] Guha, R., Das, N., Kundu, M., Nasipuri, M., & Santosh, K. C. (2020). Devnet: an efficient cnn architecture for handwritten devanagari character recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 34(12), 2052009.
- [8] Bhardwaj, A. (2020). Handwritten Devanagari Character Recognition Using Deep Learning-Convolutional Neural Network (CNN) Model. *PalArch's Journal of Archaeology of Egypt/Egyptology*, 17(6), 7965-7984.
- [9] Kavya, A., Vivek, N., Harika, M., & Nidumolu, V. (2021). Handwritten Devanagari Character Classification Using CNN. In *Inventive Communication and Computational Technologies* (pp. 309-317). Springer, Singapore.
- [10] Sharma, A. K., Adhyaru, D. M., & Zaveri, T. H. (2020). A Survey on Devanagari Character Recognition. In *Smart Systems and IoT: Innovations in Computing* (pp. 429-437). Springer, Singapore.
- [11] Gupta, P., Deshmukh, S., Pandey, S., Tonge, K., Urkunde, V., & Kide, S. (2020, October). Convolutional Neural Network based Handwritten Devanagari Character Recognition. In *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)* (pp. 322-326). IEEE.
- [12] Deshpande, P., Bondre, A., Bhise, V., Ghorpade, H. B., Bhosale, S. S., & Patil, C. G. (2021). Devanagari Character Recognition Using Convolutional Neural Network. *International Journal of Future Generation Communication and Networking*, 14(1), 4189-4198.
- [13] Bhati, G. S., & Garg, A. R. (2020, September). Handwritten Devanagari Character Recognition Using CNN with Transfer Learning. In *Congress on Intelligent Systems* (pp. 269-279). Springer, Singapore.