



Publisher homepage: www.universepg.com, ISSN: 2663-7804 (Online) & 2663-7790 (Print)

<https://doi.org/10.34104/ajeit.022.027031>

Australian Journal of Engineering and Innovative Technology

Journal homepage: www.universepg.com/journal/ajeit

Australian Journal of
**Engineering and
Innovative Technology**



Bangla Handwritten Characters Recognition Using Convolutional Neural Network

Md. Anwar Hossain^{1*}, Mirza A. F. M. Rashidul Hasan², A. F. M. Zainul Abadin¹, and Nafiul Fatta¹

¹Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna-6600, Bangladesh; and ²Department of Information and Communication Engineering, University of Rajshahi, Bangladesh.

*Correspondence: manwar.ice@gmail.com (Md. Anwar Hossain, Associate Professor, Department of Information and Communication Engineering, Pabna-6600, Bangladesh).

ABSTRACT

In the Bangla language, there are 50 complex-shaped characters and working with this huge amount of characters with an appropriate set of features is a tough problem to recognize handwritten characters. Moreover, ambiguity and precision errors are common in handwritten words. Furthermore, among a large number of complex-shaped letters, some are quite similar in shape, making handwritten Bangla characters difficult to recognize. In this work, we proposed a convolutional neural network-based approach for recognizing the handwritten Bangla alphabet. In character recognition, the convolutional neural network (CNN) outperforms most of the other models. However, to guarantee a satisfactory performance, CNNs usually need a great number of samples. Bangla handwriting recognition has been a hot topic for several years, but due to the similarity of many Bangla characters, it's difficult to achieve good results. By training and testing on Bangla character datasets, the model gets a 90.22% validation accuracy for Bangalekha isolated dataset and 93.22% validation accuracy for the Ekush dataset.

Keywords: Deep learning, Image recognition, Convolutional networks and Handwritten character recognition.

INTRODUCTION:

The main objective of this research is to recognize individual Bangla handwritten characters. Which can be extended further by providing appropriate methods and can be used in different scenarios like transforming human written documents to digital Unicode documents, extracting information from national identity cards, driving licenses, bank cheques, and many more. Convolutional Neural Network (CNN) (Albawi *et al.*, 2017), is one of the best methods to perform classification on image data. CNN mimics the visual cortex of the human brain. The visual layers of the human brain can detect complex features from the image to recognize an image. The same principle is

UniversePG | www.universepg.com

applied in CNN. In the CNN layer, a different type of filter is applied over the image to extract feature images from which prediction is performed using a neural network.

Previous work

Previous research in the field of Bangla character categorization has primarily focused on the Bangla digit, which has ten digits. There are a few works available for handwritten character recognition in Bangla. Other people worked with Bangla's hand-written character recognition but all of them worked with 50 letters and 10 numerals separately. Rahmen *et al.* (2015) proposed a model which achieved 85.96%

accuracy for 50 letters. Purkayastha et al. (2017) also proposed a model which achieved higher accuracy 89.01% for 50 letters (Chowdhury et al., 2019) proposed a model which achieved 91.13% accuracy for 50 letters and 98.42% accuracy for 10 numerical.

Apart from there also present several Bangla Hand-written Character Recognition and had achieved pretty success. Halima Begum et al. (2017) worked with their dataset that was collected from 95 volunteers and their proposed model was achieved without feature extraction and with feature extraction around 68.9% and 79.4% of recognition rate respectively (Das et al., 2009) accuracy for Bangla character 76.86% and Bangla numeral 99.45%. (Rahman et al., 2015; Rahman et al., 2022) achieved 85.36% test accuracy using their dataset. In (Das et al., 2010) handwritten Bangla character recognition with MLP and SVM has been proposed and they achieved around 79.73% and 80.9% of recognition rate, respectively.

METHODOLOGY:

Architecture

The proposed CNN model has 10 layers and all of them are connected sequentially. The first layer is an input layer that defines Input image size and the number of color channels. In our model, its static value is (32x32x1). Then two convolutional layers are connected back to back. One of them has 32 filters and another one has 64 filters. Each of them has the same kernel size 3x3 activation function RELU. After that comes a special layer is known as Max-Pooling, which shrinks the size of the image to half by picking up the maximum value. Then comes again the two convolutional layers which mimic the previous two layers' definitions. After that comes a Flatten layer, which transforms 2-dimensional data to 1-dimensional data. The next two dense layers and a dropout layer are connected in a sandwich manner, having a dropout layer in the middle. The first dense layer act as a hidden layer in which 1-dimensional data is mapped. Then dropout layer randomly deletes some weights based on their threshold value. The Last dense layer acts as model output. It has the same number of nodes as the number of classes that need to be classified. It also has an activation function as softmax. So, our model has 637,724 parameters and all of them are trainable. Fig. 1 shows the architecture of the hand-

written character recognition model. Table 1 shows the internal parameters we are using for the model.

Table 1: Internal parameters for our Model.

Layer(type)	Output Shape	Param
Conv2D	(None, 28, 28, 32)	832
Conv2D	(None, 24, 24, 64)	51264
MaxPooling2D	(None, 12, 12, 64)	0
Conv2D	(None, 10, 10, 32)	18464
Conv2D	(None, 8, 8, 64)	18496
Flatten	(None, 4096)	0
Dense	(None, 128)	524416
Dropout	(None, 128)	0
Dense	(None, 128)	16512
Dropout	(None, 128)	0
Dense	(None, 60)	7740

Total params: 637,724

Trainable params: 637,724

Non-trainable params: 0

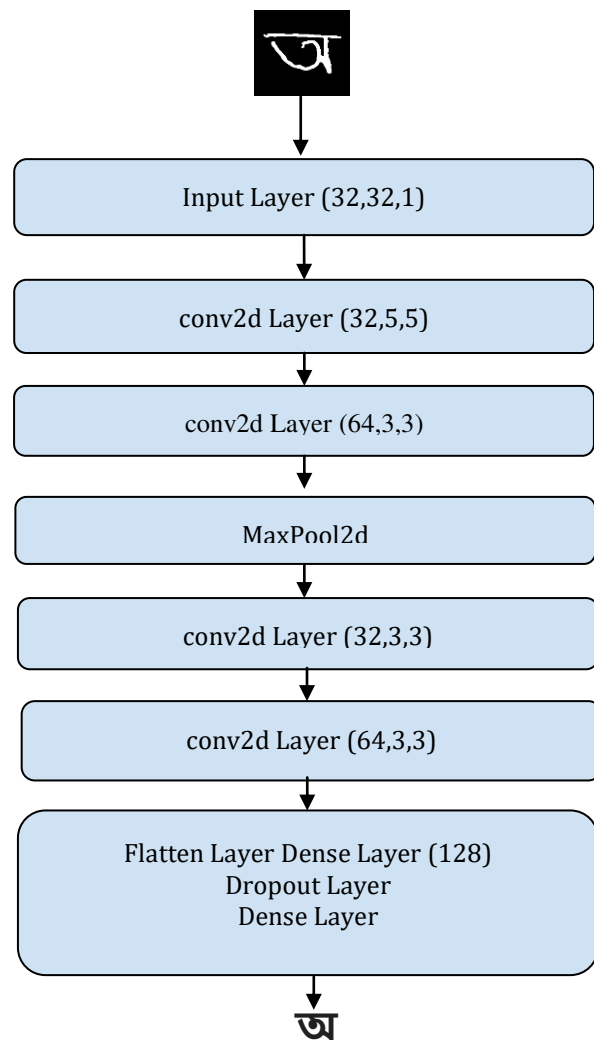


Fig. 1: Architecture of the Handwritten Character Recognition Model.

Compiling

The Model is then compiled with an ADAM optimizer having a learning rate of 0.0001 and loss function as sparse_categorical_crossentropy. Model is trained in Google Colab GPU notebook.

Graphical User Interface

The graphical interface of this trained model is developed using web technologies. For server-side workload, Python Flask (Aslam & Mohammed, 2015). Module is used, which loads the pre-trained model and provides API to the frontend. For the frontend, Javascript React framework is used, it's communicating with the backend Flask server through the API and provides a Graphical User Interface to the user. Fig. 2 shows the output using the graphical user interface.

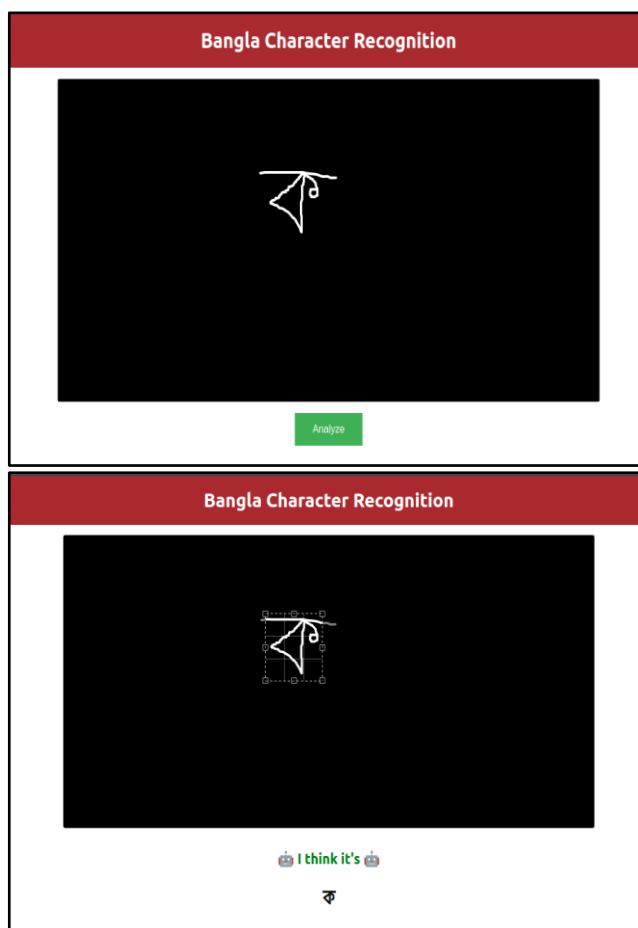


Fig. 2: Graphical User Interface built with ReactJS.

Datasets

Dataset preprocessing

In this work, we have created a combined dataset with 50 letters and 10 numerals totaling 60 alphabets and trained our model on it. Multiple datasets are used to

train the model. we have used TensorFlow 'image_dataset_from_directory' API to create training and validation datasets. This API is capable of generating an image label from its directory name. To do so we have to rename the directory name "0" to "59", where "0" to "9" represent "০" to "৯" and "10" to "59" represents "০" to "৯". All of the images in the dataset are resized to 28x28 images using bilinear interpolation. Images are read in single-channel and normalized from 0-255 to 0-1 and batched together.

- Bangla Lekha-isolated dataset (Mithun et al., 2017): It's a dataset of 84 classes that contain Bengali numbers, vowels, consonants, and compound characters. each class contains 2000 (approx) images. But we are only interested in the first 60 classes, the rest of them are deleted. So the number of total images is 120000 (approx), where 80% (96000) is used for the training set and 20% (24000) is reserved for the validation set.
- Ekushdataset (Rabby et al., 2019): This dataset contains 60 classes of the image in two categories male and female. We merged them into a single folder. Now each class contains 3000 (approx) images. So total number of images 180000 (approx), where 80% (144000) used in the training set and 20% (36000) reserved for the validation set.

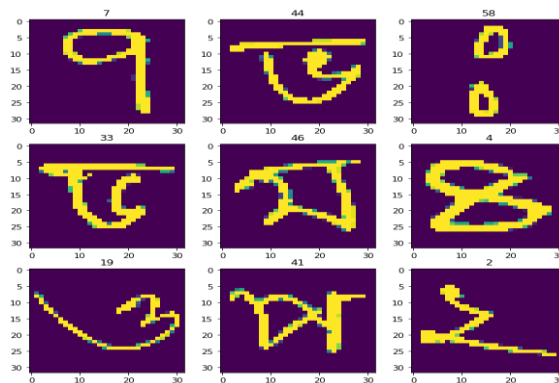


Fig. 3: Visual representation of a chunk of the dataset.

Accuracy and Performance

We have trained our model using both datasets and can achieve 90.22% accuracy for the Bangla Lekha-isolated dataset and 93.40% for the Ekush dataset.

Table 2: Performance comparison of our model.

Dataset	Classes	Accuracy
Bangla Lekha-isolated	60	90.22%
Ekush	60	93.40%

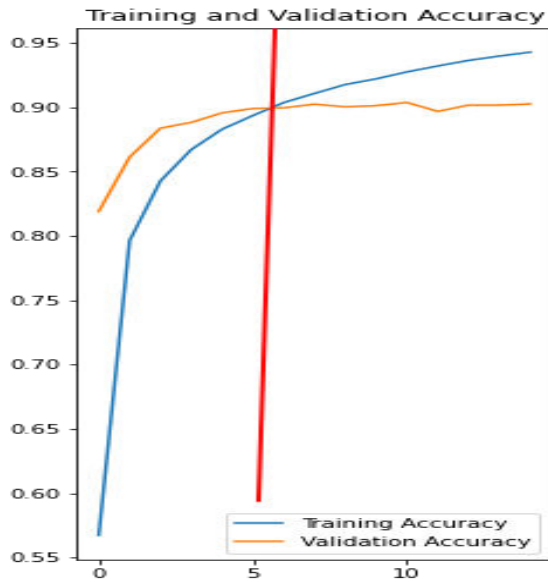


Fig. 4 (A): Training and Validation Accuracy for the Bangla Lekha-isolated Dataset.

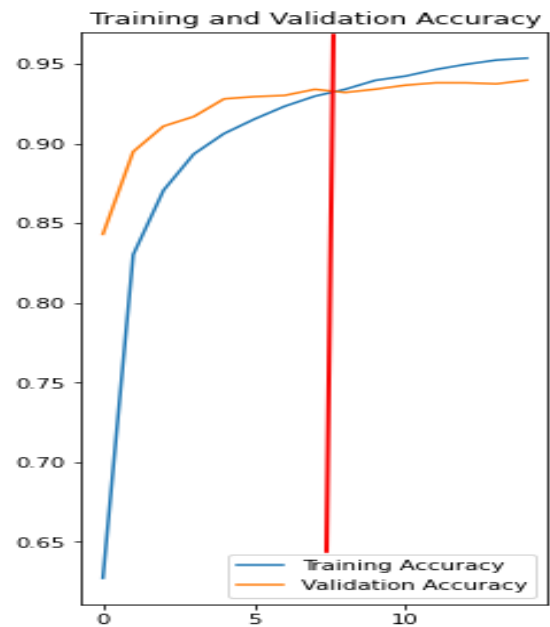


Fig. 5 (A): Training and Validation Accuracy for the Ekush Dataset.

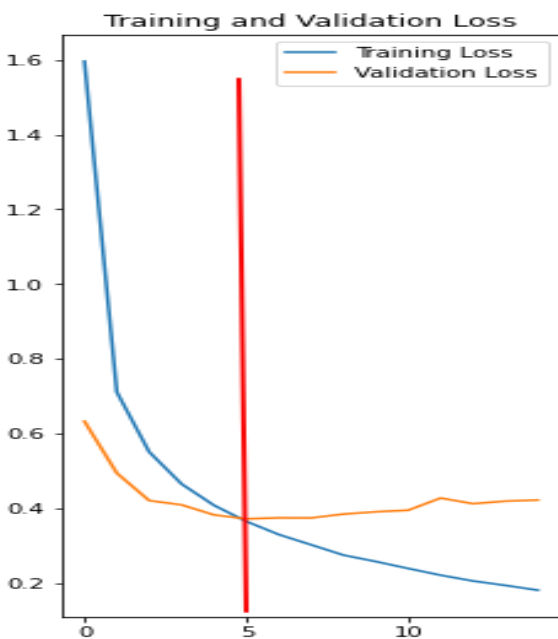


Fig. 4 (B): Training and Validation Loss for the Bangla Lekha-isolated Dataset.

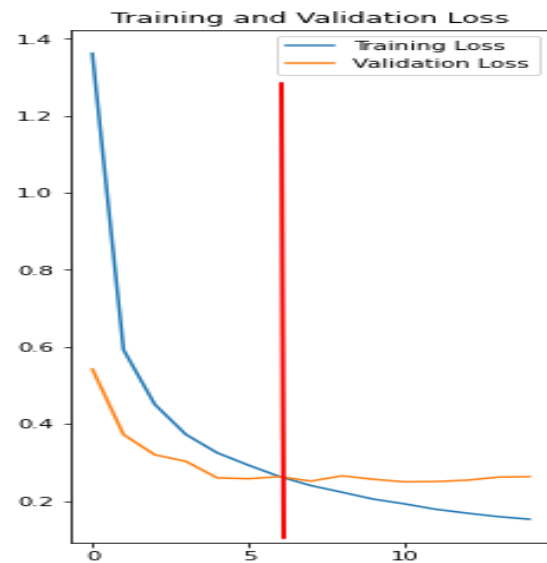


Fig. 5 (B): Training and Validation Loss for the Ekush Dataset.

From the performance graph, **Fig. 4**, we can see that our model gets its highest validation accuracy in 8 epochs after that our model is overfitted. Redline in our graph indicates the best fit for our model.

From the performance graph, **Fig. 5**, we can see that our model gets its highest validation accuracy in 9 epochs after that our model is overfitted. Redline in our graph indicates the best fit for our model.

CONCLUSION:

This research finds out the accuracy of output from our proposed model and is comprised of another model. We can find the accuracy using the noisy dataset also find accuracy using the noiseless dataset. The performance graph shows that our model gets its highest validation accuracy 90.22% for the Bangla Lekha-isolated dataset and 93.40% for the Ekush dataset.

ACKNOWLEDGEMENT:

We are grateful to Pabna University of Science and Technology (PUST) for the support to the research.

CONFLICTS OF INTEREST:

The authors state that there is no potential conflict of interest in publishing this research article.

REFERENCES:

- 1) Albawi, Saad and Mohammed, Tareq Abed and Al-Zawi, Saad. (2017). *International Conference on Engineering and Technology (ICET)*, Understanding of a convolutional neural network. <https://doi.org/10.1109/ICENGTECHNOL.2017.8308186>
- 2) Aslam, F.A., & Mohammed, P.S. (2015). Efficient Way Of Web Development Using Python And Flask. *International Journal of Advanced Research in Computer Science*, **6**, 54-57. <https://doi.org/10.26483/ijarcs.v6i2.2434>
- 3) Begum, H., & Islam, M.M. (2017). Recognition of Handwritten Bangla Characters using Gabor Filter and Artificial Neural Network. *International Journal of Computer Technology & Applications*, **8**(5), 618-621, ISSN:2229-6093
- 4) Chowdhury *et al.* (2019). Bangla Hand written Character Recognition using Convolutional Neural Network with Data Augmentation. <https://doi.org/10.1109/ICIEV.2019.8858545>
- 5) Mithun Biswas *et al.* (2017). BanglaLekhaIsolated: A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters, *Data in Brief*, **12**, 2017, Pages 103-107, ISSN 2352-3409, <https://doi.org/10.1016/j.dib.2017.03.035>
- 6) Nibar Das, Sandip Pramanik., (2009). Recognition of Handwritten Bangla Basic Character and Digit Using Convex Hull Basic Feature. *International Conference on Artificial Intelligence and Pattern Recognition (AIPR-09)*. <https://arxiv.org/abs/1410.0478v1>
- 7) N. Das *et al.* (2010). Handwritten Bangla Basic and Compound character recognition using MLP and SVM classifier. *Journal of Computing*, **2**, <https://arxiv.org/pdf/1002.4040>
- 8) Purkaystha,B., Datta T. and IslamM. S. (2017). Bengali handwritten character recognition using deep convolutional neural network. *20th International Conference of Computer and Information Technology (ICCIT)*, 2017, pp. 1-5. <https://doi.org/10.1109/ICCITECHN.2017.8281853>
- 9) Rabby *et al.* (2019). Ekush: A Multipurposeand Multitype Comprehensive Database for Online Off-line Bangla Handwritten Characters. <https://www.springerprofessional.de/en/ekush-a-multipurpose-and-multitype-comprehensive-database-for-on/16936016>
- 10) Rahman A, Islam MM, Moon MJ, Tasnim T, Siddique N, Shahiduzzaman M, and Ahmed S. (2022). A qualitative survey on deep learning based deep fake video creation and detection method. *Aust. J. Eng. Innov. Technol.*, **4**(1), 13-26. <https://doi.org/10.34104/ajeit.022.013026>
- 11) Rahman *et al.* (2015). Bangla Handwritten Character Recognitionusing Convolutional Neural Network. *International Journal of Image, Graphics and Signal Processing (IJIGSP)*,**7**, 42-49. <https://doi.org/10.5815/ijigsp.2015.08.05>

Citation: Hossain MA, Hasan MAFMR, Abadin AFMZ, and Fatta N. (2022). Bangla handwritten characters recognition using convolutional neural network. *Aust. J. Eng. Innov. Technol.*, **4**(2), 27-31.

<https://doi.org/10.34104/ajeit.022.027031> 