Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 143 (2018) 528–535

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Conference on Advances in Computing and Communication (ICACC-2018)

BornoNet: Bangla Handwritten Characters Recognition Using Convolutional Neural Network

AKM Shahariar Azad Rabby[a]*, Sadeka Haque[a]†, Md. Sanzidul Islam[a], Sheikh Abujar[a] and Syed Akhter Hossain[a]

[a]*Dept. of Computer Science and Engineering, Daffodil International University, Dhanmondi, Dhaka-1205, Bangladesh*

Abstract

Bangla handwriting recognition is becoming an important issue in several years but it becomes a challenge to get good performance due to the alignment and many of them are similar. A simple, lightweight CNN model has been proposed in this paper for classifying Bangla Handwriting Character, which contains 50 basic Bangla characters (11 vowels and 39 consonants). Experiments have been made on three datasets along with the BanglaLekha-Isolated [1] CMATERdb [2] and the ISI [3] dataset. For character recognition, the proposed BornoNet model gets 98%, 96.81%, 95.71%, and 96.40% validation accuracy respectively for CMATERdb, ISI, BanglaLekha-Isolated dataset and mixed dataset. Also proposed model was trained with one dataset and cross-validated with other two datasets. Proposed model achieved the best accuracy rate so far for BanglaLekha-Isolated, CMATERdb and ISI datasets. The proposed BornoNet model can be found on https://github.com/shahariarrabby/BornoNet

© 2018 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the scientific committee of the 8th International Conference on Advances in Computing and Communication (ICACC-2018).

Keywords: Handwritten Recognition; Pattern Recognition; Document Image Analysis; Machine Learning; Computer Vision.

* AKM Shahariar Azad Rabby.
E-mail address: azad15-5424@diu.edu.bd
* Sadeka Haque
E-mail address: sadeka15-5210@diu.edu.bd

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the scientific committee of the 8th International Conference on Advances in Computing and Communication (ICACC-2018).
10.1016/j.procs.2018.10.426

1. Introduction

In recent years Convolutional neural network (CNN) becomes popular for complex visual recognition in due to its architecture and a lot of research on using deep CNN to recognize handwritten digits, characters, etc. And we will explore the power of deep CNN on the classification of handwritten Bangla characters. Handwritten recognition has importance for its various applications as Optical Character Recognition (OCR). Several scientific types of research have been carried out for character recognition of English, Chinese, Arabic handwritten digits and alphabet. Though automatic recognition of handwriting remains a great challenge: the performance of layout analysis, word/character segmentation, and recognition is still far behind the human recognition capability.

Bangla is the mother language of Bangladesh, apart from it is the official language of Bangladesh, West Bengal of India, Tripura, Assam and Jharkhand, Sierra Leone a West African country. Though Bangla is the 7th most popular language and writing scripts like about 250 million speak in Bangla and 2nd most beautiful language in the world. Considering those all circumstance the technology in different sectors in these regions, Bangla Handwritten recognition plays an important role and should overcome the challenge. However, to compare other language writing script only a few studies are attested on handwritten characters of Bangla scripts there have a sturdy model such as Latin [4], Chines [5], Japanese [6] have achieved a great success on machine learning and deep learning application. Conducting with the handwritten Character or digit recognition always a big challenge in due to its variation of shape, size and complexity of Bangla text such as there have many misleading characters. But several years few works have done for Bangla handwritten digit recognition using CNN and gain some success of classifying those type of recognition to the researcher for its educational and economic interesting applications.

Bangla writing scripts different from English or other languages writing scripts because of it came from Sanskrit script which is completely different and it has alinement and some character are similar to another character as some of them are different from small dot and line. There is 50 basic character (11 vowels and 39 consonants), 10 modifiers, 10 numerals, more than 300 compound characters. Due to those, all fact Bangla scripts make difficult to achieve a good result and better performance with Bangla Handwritten character recognition. If we overcome all these challenges and make a model then we improved many kinds of Handwritten recognition-based application such as Bangla Handwritten character base OCR (Optical Character Recognition), Picture to text to speech, Bangla ID card reading, Number plate reading, vehicle tracking, Post office automation etc. Fig 1 is example of Bangla characters.

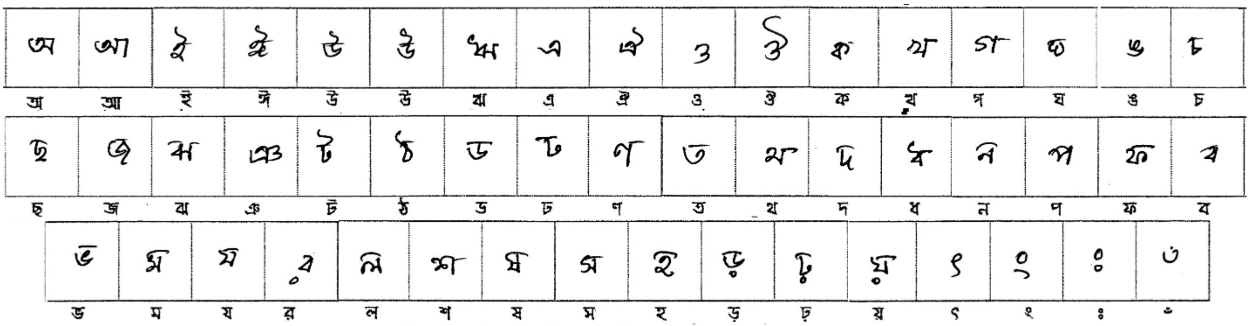


Fig 1: Example of Bangla Characters

2. Literature review

In the area of Bangla character classification, previous works have mainly focused on Bangla digit, which contains 10 digits.

There are a few works are available for Bangla handwritten character recognition. Some literature has been reported on Bangla characters recognition in the past years as “A complete printed Bangla OCR system” [7], “On the development of an optical character recognition (OCR) system for printed Bangla script” [8] but few research on handwritten Bangla numeral recognition that reaches to the desired recognition accuracy. Pal et al. have conducted some exploring works for recognizing handwritten Bangla characters those are “Automatic recognition of unconstrained offline Bangla hand-written numerals” [9], “A system towards Indian postal automation” [10] and

"Touching numeral segmentation using water reservoir concept" [11]. The proposed schemes are mainly based on extracted features from a concept called water reservoir. Apart from there also present several Bangla Handwritten Character Recognition and had achieved pretty good success. Halima Begum et al, "Recognition of Handwritten Bangla Characters using Gabor Filter and Artificial Neural Network" [12] works with own dataset that was collected from 95 volunteers and their proposed model achieved without feature extraction and with feature extraction around 68.9% and 79:4% of recognition rate respectively. "Recognition of Handwritten Bangla Basic Character and Digit Using Convex Hull Basic Feature" [13] accuracy for Bangla character 76.86% and Bangla numeral 99.45%. "Bangla Handwritten Character Recognition using Convolutional Neural Network" [14] achieved 85.36% test accuracy using their own dataset. In "Handwritten Bangla Basic and Compound character recognition using MLP and SVM classifier" [15] handwritten Bangla character recognition with MLP and SVM has been proposed and they achieved around 79.73% and 80.9% of recognition rate, respectively.

3. Proposed methodology

In this paper to build the proposed model "BornoNet" using Convolutional Neural Network (CNN) which has many phases that are following in below.

3.1. Dataset

The proposed method used three datasets BanglaLekha-Isolated, CMATERdb and ISI datasets. These three datasets consist of Bangla characters, Bangla numeral digits and Bangla compound characters.

CMATERdb database has total 15,000 characters images. The image was 32 x 32px noise-free images and the image edge looks blocker. ISI Dataset has total 37,858 images. The image has different pixel size and image edge look smother. BanglaLekha-Isolated has total 98,950 images. The image is inverted and resized and add some padding to be square with preserving the aspect ratio.

By joining all these three datasets we make a mixed dataset of 151580 images. Fig 2 showing an example of the different datasets.

(a) (b) (c)

Fig 2: (a) BanglaLekha (b) ISI (c) CMATERdb

3.2. Dataset preparation

Initially we prepared dataset for building the model. First fixed some incorrect labeling images from the BanglaLekha-Isolated dataset. Delete some incorrect images such as blank images. This dataset already containing inverted image prepossessed with its inverted foreground and background, removal of noise with the median filter, edge thickening.

For CMATERdb and ISI dataset first thresholding, grayscaling and invert all the images.

Secondly for all datasets resized all images to 28 x 28 pixel. Then converted all dataset's 28 x 28-pixel image into a 784 + (1 label) D matrix and store all the image pixel into a CSV file to reduce hard disk read and fast computation. Also, perform a Minmax (1) normalization to reduce the effect of illumination differences. Moreover, the CNN converge faster on (0...1) data than on (0...255)

$$Z_i = \frac{X_i - \text{minmum}(X)}{\text{maximum }(X) - \text{minmum}(X)} \quad (1)$$

Then convert the 10 labels into one hot encoding
(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0) = অ
(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1) = ঁ

Then convert the 784 D into 28 x 28 image matrix.

3.3. Proposed model

The proposed model is a 13-layer convolutional neural network with 2 sub-layers. Proposed model use ADAM [16] optimizer. For first two-layer same padding and ReLU (2) activation used with 32 filters with the 5x5 kernel. Then a max-pooling layer added with a 2x2 followed by 25% dropout layer. All dropout layer used to reduce overfitting.

$$ReLU(X) = MAX(0, X) \quad (2)$$

The output from this two-layer goes as an input of two sublayers. Where both sublayers have same 2 convolutional layers with the same padding, ReLU activation, 64 filters with a 5x5 kernel, followed by another convolutional layer with a 3x3 kernel. Later the output of last 2 sub convolutional layer added together and pass through a Max-Pooling layer with 25% dropout [17] layer.

Then flatten the layer and used a fully connected layer with 1280 hidden node followed by 25% dropout. Final output layer has 50 nodes with SoftMax (3) activation. Fig 3 is showing the architecture of the network.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \; for \; j = 1, \dots k \quad (3)$$

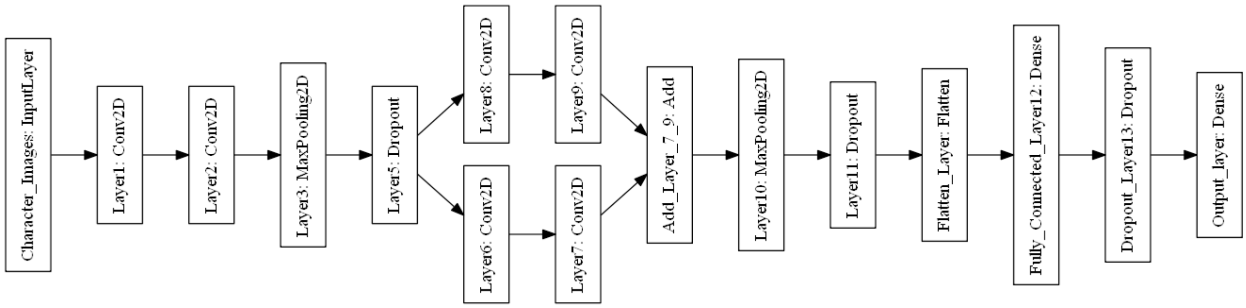


Fig 3: Prosed BornoNet architect.

Table 1. BornoNet model architect summary.

Layer No (type)	Output Shape	Param	Connected to	Layer No (type)	Output Shape	Param	Connected to
(Input Layer)	28, 28, 1	0	-	8 (Conv2D)	14, 14, 64	36928	7
1 (Conv2D)	28, 28, 32	832	Input Layer	Add 6 8 (Add)	14, 14, 64	0	6, 8
2 (Conv2D)	28, 28, 32	25632	1	9 (MaxPooling2D)	7, 7, 64	0	Add_ 6_8
3 (MaxPooling2D)	14, 14, 32	0	2	10 (Dropout)	7, 7, 64	0	9
4 (Dropout)	14, 14, 32	0	3	Flatten (Flatten)	3136	0	10
5 (Conv2D)	14, 14, 64	51264	4	11 (Dense)	1280	4015360	Flatten
7 (Conv2D)	14, 14, 64	51264	4	12 (Dropout)	1280	0	11
6 (Conv2D)	14, 14, 64	36928	5	Output (Dense)	50	64050	12

Total parameters: 4,282,258
Trainable parameters: 4,282,258
Non-trainable parameters: 0

3.4. Optimizer and Learning rate

In Deep Learning and computer vision work, the optimization algorithm can change the result and make it pretty sufficient. The Adam paper says, "...many objective functions are composed of a sum of subfunctions evaluated at different subsamples of data; in this case, optimization can be made more efficient by taking gradient steps w.r.t. individual sub-functions ..." [16]. The Adam optimization algorithm is straightforward to implement, computationally efficient and little memory requirements that's why recently adopting most of the computer vision and natural language processing application. Proposed method used ADAM (4) Optimizer with learning rate 0.001.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt[1]{\hat{v}_t + \varepsilon}} \hat{m}_t \tag{4}$$

When using a neural network to perform classification and prediction task we need to calculate the error rate. A recent study shows that cross entropy function performs better than classification error and mean square error. [18] Proposed method used categorical cross entropy (4) as loss function.

$$\mathrm{L_i} = -\sum_j \mathrm{t_{i,\ j}} \log(\mathrm{p_{i,\ j}}) \tag{5}$$

To make the optimizer converge faster and closer to the global minimum of the loss function, using an automatic Learning Rate reduction method [19]. Learning rate is the step by which walks through the minimum loss. If the learning rate is too low, it will take more time to reach the global minima, and if the learning rate is too high then the training may not be converging or even diverge. To keep the advantage of the fast computation time we set a high learning rate which is automatically decreased by monitoring the validation accuracy.

3.5. Data augmentation

To avoid overfitting, artificially expand the handwritten dataset. This data transformation will create some variance that can occur when someone else writing the digits. For Data augmentation, several methods are chosen:

- Randomly shifting height and width 10% of the images.
- Randomly rotate our training image 10 degrees.
- Randomly 10 % zoom the training image.

3.6. Training the model

The proposed model was trained with different training and validation set with the batch size of 86. During training, automatic learning rate reduction formula mentored the validation accuracy and reduce the learning rate if required.

4. Evaluate the model

The proposed model was applied to the three-different dataset and gain good result on train, test, validation sets.

4.1. Train, Test and Validation sets

For CMATERdb 12,000 images used as train image and 3,000 for the test image. We take 10% of train images for validation purpose. After training model was tested with 3000 images of CMATERdb test dataset also with 37,858 images of ISI dataset and 98,722 images of the BanglaLekha-Isolated dataset.

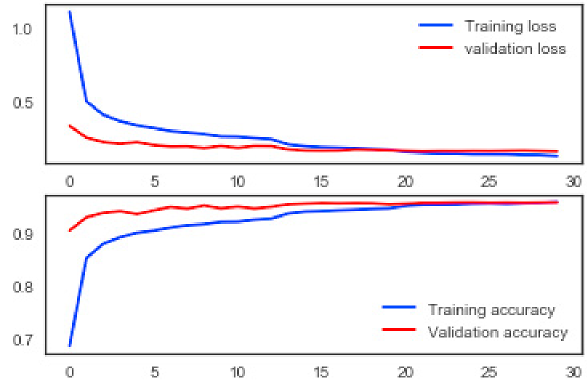
For ISI, the dataset has 37,858 images, where we use 7000 as validation images and 12,859 as test images also used 15,000 images of CMATERdb dataset and 98,722 images of BanglaLekha-Isolated for testing the model.

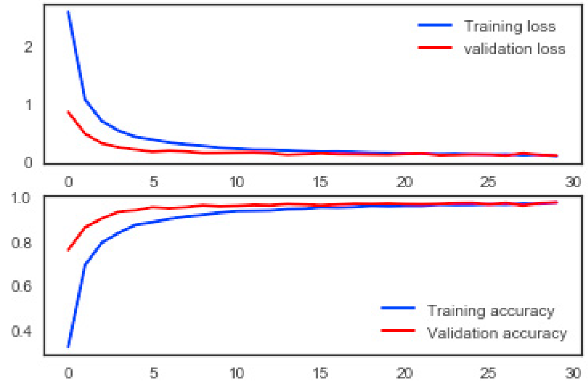
For BanglaLekha-Isolated 98,722 images we took 88,849 images for training and 9,873(100%) image as a validation set. After training model tested with 15,000 images of CMATERdb dataset and 37,858 images of ISI dataset.

For Mixed Datasets 151,580 images we make the training set of 121,263 images by taking 80% of each dataset. 10% used as validation and other 10% used as test set.

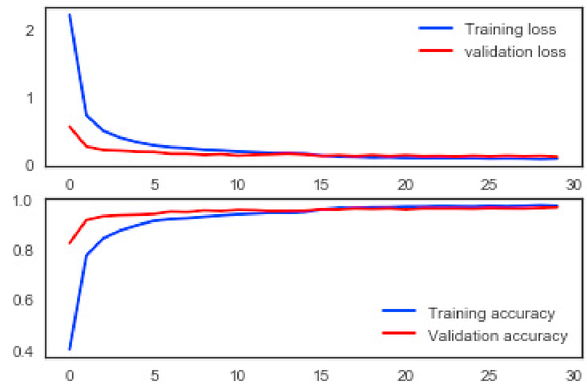
4.2. Model performance

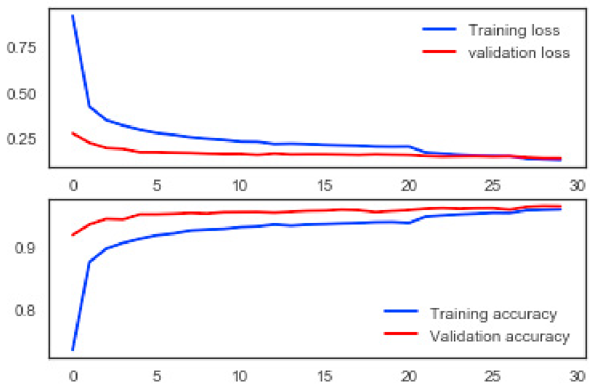
After 30 epoch proposed model gets 98%, 96.81%, 95.71%, and 96.40% validation accuracy respectively for CMATERdb, ISI, BanglaLekha- Isolated dataset and mixed dataset. Also, all of this dataset cross-validate with each other and perform accurately. Fig 4 (a) (b) (c) (d) is showing the accuracy and loss of training and validation set of BanglaLekha-Isolated, CMATERdb, ISI and mixed dataset respectively. Table 2 showing the details accuracy on different datasets.





(a) (b)





(c) (d)

Fig 4: (a) BanglaLekha-Isolated (b) CMATERdb (c) ISI (d) Mixed Database

Table 2. Result comparison in different Dataset.

Dataset Name	Tr. Loss	Val. Loss	Tr. Acc.	Val Acc.	Test Dataset	Test Acc.
BanglaLekha	0.1455	0.1679	95.45%	95.71%	CMATERdb	95.22%
					ISI	93.59%
CMATERdb	0.0475	0.0817	98.44%	98.00%	BanglaLekha	78.29%
					ISI	92.19%
ISI	0.0814	0.1120	97.37%	96.81%	BanglaLekha	84.63%
					CMATERdb	94.94%
Mixed Dataset	0.1298	0.1396	95.81%	96.40%	Mixed Test	96.44%
					BanglaLekha	97.96%
					CMATERdb	99.12%
					ISI	99.18%

4.3. Result analysis

Analyzing the result and confusion, we found that CMATERdb dataset got best validation accuracy but perform poorly with another dataset. That because CMATERdb dataset contains the image that is noise free. That's why when noise image from other datasets came it fails to recognize. Confusion matrix can be found on https://github.com/shahariarrabby/BornoNet/graph.

On the other hand, BanglaLekha-Isolated gets less accuracy but perform very well on other datasets. BanglaLekha-isolated has the noisy image. When a noise-free image from another dataset came it perform very well.

4.4. Result comparison

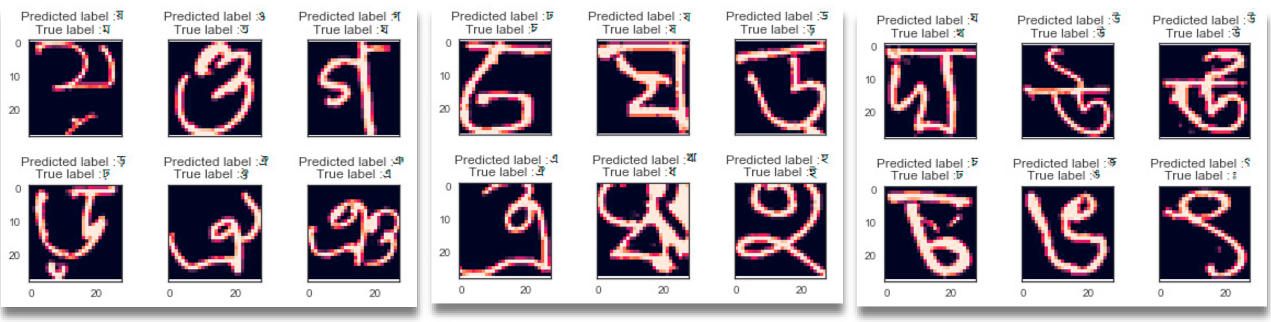
Table 3 showing a comparison between some the previous work. From this table, we found that our proposed BornoNet got so far, the best accuracy rate for all of the three datasets.

Table 3. Result Comparison in different Model.

Work	Accuracy	Work	Accuracy
Recognition of Handwritten Bangla Characters Using Gabor Filter and Artificial Neural Network [12]	79.4%	HMM-Based Online Handwritten Bangla Character Recognition using Dirichlet Distributions [21]	91.85%
Recognition of Bangla handwritten basic characters and digits using convex hull-based feature set [13]	76.86%	Bangla Hand-Written Character Recognition Using Support Vector Machine [22]	93.43%
Handwritten Bangla Character Recognition Using Neural Network [20]	84.00%	Bengali handwritten character recognition using Modified syntactic method [23]	95.00%
Bangla Handwritten Character Recognition using Convolutional Neural Network [14]	85.36%	**BornoNet (Proposed)**	**95.71% (Isolated)** **98% (CMATER)** **96.81% (ISI)**

5. Error observation

Analyzing the error from validation set we found that most of the incorrect classification is caused by error labeling on the dataset. It is clear that model performing great for classifying characters. Some of this mistake can also make by humans. Fig 5 showing top 6 validation error from each dataset.



(a) (b) (c)

Fig 5: Top six Error form (a) BanglaLekha-Isolated (b) CMATERdb (c) ISI

6. Conclusion and Future work

This research work presented a new CNN model which performs better classification accuracy in the different dataset for both train and validation set for lesser epochs and less computation time compared to the other CNN model. CNN's in general costly to train but extremely effective deep learning models. Hence, fast convergence should be viewed as an important front of research. Also, the cross-validation from different distribution's data proposed model achieve a great result that makes it a robust model that improve any other previous model.

Sometimes proposed model confused to understand overwritten character and dataset contained some incorrect labeling images. Also, the model performed poorly if the train on noise-free data. In future work fixing dataset and overcoming the limitation of overwriting Character should fix. Making a benchmark model for Bangla Handwritten all characters that include the numeral, basic characters, Bangla modifier and compound letter.

References

[1] Biswas, Mithun & Islam, Rafiqul & Kumar Shom, Gautam & Shopon, Md & Mohammed, Nabeel & Momen, Sifat & Abedin, Anowarul. (2017). BanglaLekha-Isolated: A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters. Data in Brief. 12. 103-107. 10.1016/j.dib.2017.03.035.

[2] R. Sarkar, N. Das, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, "Cmaterdb1: a database of unconstrained handwritten Bangla and Bangla–English mixed script document image," International Journal on Document Analysis and Recognition (IJDAR), vol. 15, no. 1, pp. 71–83, 2012.

[3] Bhattacharya, Ujjwal & Chaudhuri, Bidyut. (2009). Handwritten Numeral Databases of Indian Scripts and Multistage Recognition of Mixed Numerals. IEEE transactions on pattern analysis and machine intelligence. 31. 444-57. 10.1109/TPAMI.2008.88

[4] E. Kavallieratou, N. Liolios, E. Koutsogeorgos, N. Fakotakis, G. Kokkinakis, "The GRUHD Database of Greek Unconstrained Handwriting", ICDAR, 2001, pp. 561-565.

[5] F. Yin, Q.-F. Wang, X.-Y. Zhang, and C.-L. Liu. ICDAR 2013 Chinese Handwriting Recognition Competition. 2013 12th International Conference on Document Analysis and Recognition (ICDAR), pages 1464–1470, 2013

[6] B. Zhu, X.-D. Zhou, C.-L. Liu, and M. Nakagawa. A robust model for on-line handwritten Japanese text recognition. IJDAR, 13(2):121–131, Jan. 2010.

[7] B. B. Chaudhuri, U. Pal, "A complete printed Bangla OCR system," Pattern Recognition, vol. 31, pp. 531–549, 1998.

[8] U. Pal, "On the development of an optical character recognition (OCR) system for printed Bangla script," Ph.D. Thesis, 1997.

[9] U. Pal, B. B. Chaudhuri, "Automatic recognition of unconstrained offline Bangla hand-written numerals," T. Tan, Y. Shi, W. Gao (Eds.), Advances in Multimodal Interfaces, Lecture Notes in Computer Science, vol. 1948, Springer, Berlin, pp. 371–378, 2000.

[10] K. Roy, S. Vajda, U. Pal, B.B. Chaudhuri, "A system towards Indian postal automation," Proceedings of the Ninth International Workshop on Frontiers in Handwritten Recognition (IWFHR-9), pp. 580–585, October 2004.

[11] U. Pal, A. Belad, Ch. Choisy, "Touching numeral segmentation using water reservoir concept," Pattern Recognition Lett. vol. 24, pp. 261-272, 2003.

[12] Halima Begum et al, Recognition of Handwritten Bangla Characters using Gabor Filter and Artificial Neural Network, International Journal of Computer Technology & Applications, Vol 8(5),618-621, ISSN:2229-6093

[13] Nibaran Das, Sandip Pramanik, "Recognition of Handwritten Bangla Basic Character and Digit Using Convex Hall Basic Feature". 2009 International Conference on Artificial Intelligence and Pattern Recognition(AIPR-09)

[14] Md. M. Rahman, M. A. H. Akhand, S. Islam, P. C. Shill, "Bangla Handwritten Character Recognition using Convolutional Neural Network," I. J. Image, Graphics and Signal Processing, vol. 8, pp. 42-49, July 15, 2015.

[15] N. Das 1, B. Das, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, "Handwritten Bangla Basic and Compound character recognition using MLP and SVM classifier," Journal of Computing, vol. 2, Feb. 2010.

[16] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization.CoRR, abs/1412.6980,2132014..

[17] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting.Journal of Machine Learning Research,20915:1929–1958, 2014.

[18] Katarzyna Janocha and Wojciech Marian Czarnecki. On loss functions for deep neural networks in classification. Arxiv, abs/1702.05659, 2017

[19] Tom Schaul, Sixin Zhang, and Yann LeCun. No more pesky learning rates. In Sanjoy Dasgupta and David McAllester, editors, Proceedings of the 30th International Conference on Machine Learning, volume 28 ofProceedings of Machine Learning Research, pages 343–351, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

[20] Md. Alamgir Badsha et al, Handwritten Bangla Character Recognition Using Neural Network, International Journal of Advanced Research in Computer Science and Software Engineering 2 (11), November- 2012, pp. 307-312

[21] Chandan Biswas et al, HMM Based Online Handwritten Bangla Character Recognition using Dirichlet Distributions, 978-0-7695-4774-9/12 $26.00 © 2012 IEEE DOI 10.1109/ICFHR.2012.213

[22] Riasat Azim et al, Bangla Hand-Written Character Recognition Using Support Vector Machine, International Journal of Engineering Works vol. 3, Issue 6, PP. 36-46, June 2016.

[23] Mohammad Badiul Islam et al, Bengali handwritten character recognition using modified syntactic method, NCCPB-2005 © Independent University, Bangladesh