Bangla Handwritten Digit Recognition

Md. Sabik Alam Rahat
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2019-1-60-256@std.ewubd.edu

Sourav Saha
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2018-3-60-052@std.ewubd.edu

Md. Abid Hasan Rafe
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2018-3-60-108@std.ewubd.edu

Naeem Mizan
Department of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2018-3-60-053@std.ewubd.edu

I. Introduction

Now-a-days, handwritten digit recognition has drawn a lot of attention from researchers due to its many possible applications, including the recognition of national ID numbers, automatic license plates for vehicles, postal codes on envelopes, amounts on bank checks, and many more. It enables a computer to recognize and group digits in sample images into 10 categories, ranging from 0 to 9. Despite being one of the most widely spoken languages in the world, research on the recognition of Bangla numbers is not well focused [1]. Furthermore, as a result of the governments' insistence on digitization, the significance of Bangla document analysis has grown more than ever. The recognition process is made more difficult and complex by a number of elements, such as changes in the way that individual numbers are written, similarities in the shapes of different characters, overlaps, and connections between nearby characters, among others.

II. Related Works

Researchers tried a variety of techniques to improve the recognition of handwritten Bangla numbers in order to get over these challenges. Convolutional neural network (CNN)-based solutions, which use deep learning to learn, extract, and classify features automatically, have recently demonstrated noteworthy accuracy in image recognition, classification, annotation, and other disciplines [1,6,9,10,12,13,15,16,17].

Without mentioning response time or recognition reliability, Pal et al. [2–5] used the water reservoir concept to automate the Indian postal system and achieved accuracy rates of 94.13 percent and 93 percent for handwritten Bangla and English numerals, respectively.

To categorize the handwritten Bangla numerals with lower epochs, Sharif et al. [6] suggested a hybrid model incorporating deep CNN and HOG characteristics. They achieved 99.02 percent accuracy for the ISI dataset [7] and 99.17 percent accuracy for the CAMTERDB dataset, both of which contain 6000 photos (600 images per digit) of freely written Bangla digits. Of these, 4000 and 2000 images, respectively, were employed for training and testing [8]. However, more research is required to integrate this hybrid model with numerous customized elements.

For handwritten Bangla digit identification, Shopon et al. [9] combined deep CNN with an unsupervised pre-trained auto encoder. They were able to pre-train larger datasets for better results but only achieved an accuracy of 99.50 percent for CMATERDB.

When applied to the CAMTERDB BanglaLekha-Isolated [11] dataset, the lightweight CNN model by Rabby et al. [10] achieved validation accuracy of 99.74 and 98.93%, respectively. However, it was unable to completely clear up confusions caused by excessive text and incorrect tagging of photos.

Alom et al. [12] used the CMATERDB dataset and CNN with Gabor feature to get an accuracy of 98.78 percent. They also used various deep learning models to work on handwritten Bangla letter identification and demonstrated that DenseNet had the highest recognition accuracy, with 98.31% for the alphabet, 99.13% for numerals, and 98.18% for special characters [13]. They also offered a thorough analysis of many cutting-edge deep models using various learning methodologies. To categorize the ImageNet dataset, they employed the AlexNet and GoogleLeNet models [14].

Shaha and Shaha [15] applied a new deep CNN model to a dataset of isolated Bangla handwritten characters and achieved an accuracy of 97.21 percent. Classifying basic and compound handwritten Bangla letters as well as numerals received particular attention. Recently, the NumtaDB dataset [21], which includes more than 85000 example photos, has been frequently used as a benchmark dataset. As opposed to MNIST and CMATERDB datasets, this dataset is difficult because it lacks preprocessed data.

Using the NumtaDB dataset, Shawon et al. [16]'s suggested deep CNN model achieved testing accuracy of 92.72 percent. However, their method performs poorly in instances involving translation and rotation.

By freezing intermediary layers with fewer epochs and parameters, Zunair et al. [17] were able to categorize photos of various Bangla numerals with an accuracy of 97.09 percent on the NumtaDB dataset. This strategy to transfer learning was unorthodox.

III. METHODOLOGY

We have collected images of each digit from different sources like BengaliAi, Ekush, NumtaDB etc. Then we have merged all this data into a new custom dataset. This custom dataset used for our project. We have used 80% data as training data and 20% for test data. First, we have visualized the data set to check weather we have to make it balanced or not. Then we have generated the model based on this training data. We have implemented two algorithms to classify digit. One is CNN and another one is ANN. We have used the same dataset for those algorithms. We also shown the confusion matrix and classification report for each algorithm. After trained the model we have generated model for each algorithm. Lastly, we have tested our own handwriting for testing the trained model.

GitHub Repo Link:

https://github.com/sabikrahat/bangla_handwriting_digit_classify

IV. RESULT ANALYSIS

CNN:

We have taken around 600 images of each digit. After processing the data, the visualization is:

Fig: Visualization of the Dataset

Data distribution of the dataset:

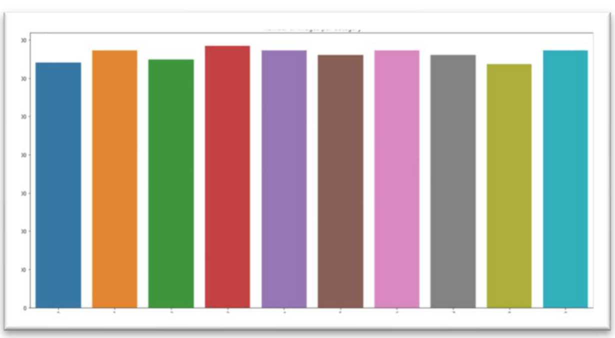

Fig: Data Distribution of the Dataset

Loss & Accuracy Plot:

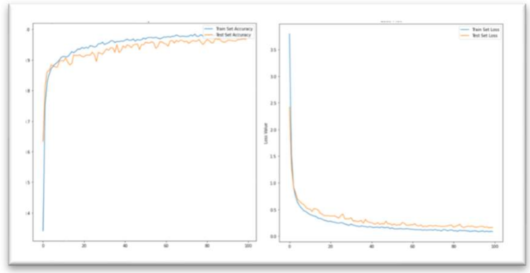

Fig: Loss & Accuracy Plot

Confusion Matrix:

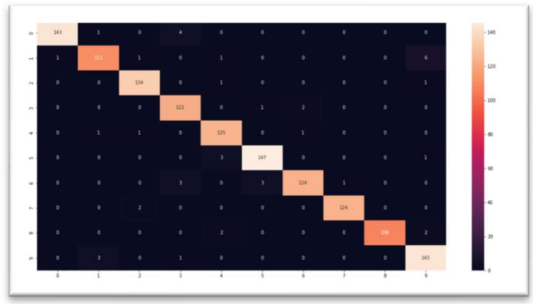

Fig: Confusion Matrix

Classification Report:

	precision	recall	f1-score	support
0	0.99	0.97	0.98	148
1	0.96	0.93	0.94	120
2	0.97	0.99	0.98	136
3	0.94	0.98	0.96	125
4	0.95	0.98	0.96	128
5	0.97	0.97	0.97	151
6	0.98	0.95	0.96	131
7	0.99	0.98	0.99	126
8	1.00	0.96	0.98	112
9	0.93	0.97	0.95	147
accuracy			0.97	1324
macro avg	0.97	0.97	0.97	1324
weighted avg	0.97	0.97	0.97	1324

Fig: Classification Report

Testing the model:

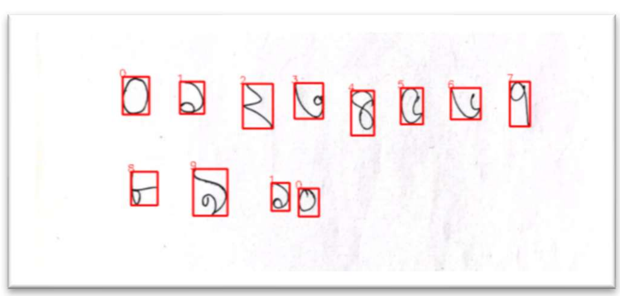

Fig: Testing the model

ANN:

Classification Report:

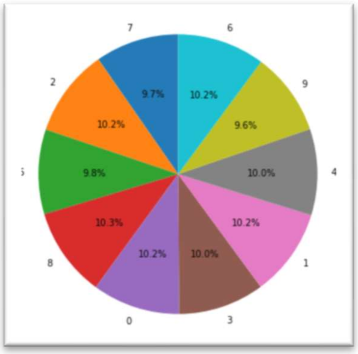

Fig: Classification Report

Process the Data:

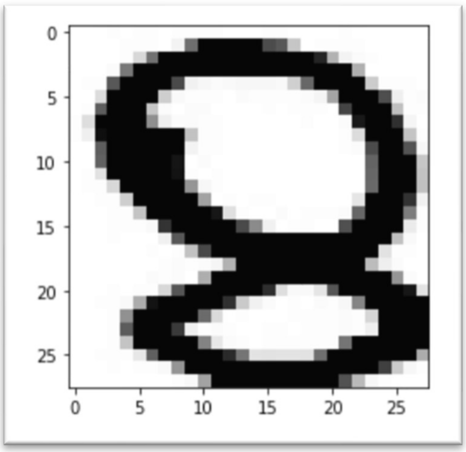

Fig: Process the data.

Loss & Accuracy Plot:

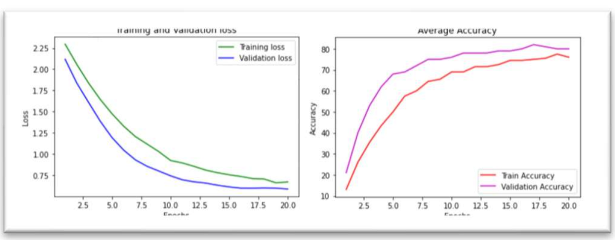

Fig: Loss & Accuracy Plot

Confusion Matrix:

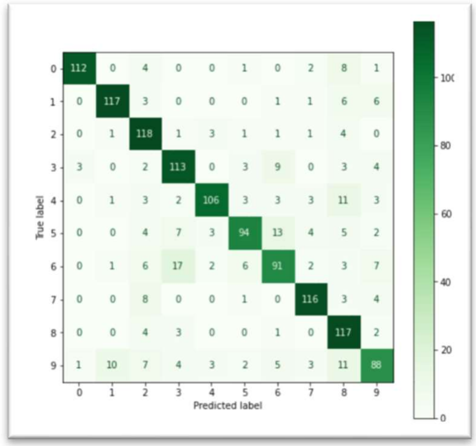

Fig: Confusion Matrix

V. Conclusion and Future Works

Based on these two algorithms results, we have seen that the model of CNN has given the best output. The classification report of the CNN algorithm is much better than the ANN algorithm. There are some logical expansions to this work that could reinforce and expand the conclusions. We first want to integrate the binarization technique presented in this study with additional techniques. We also want to make our dataset more heavy in order to increase the accuracy level.

VI. Data Set Description

Due to the advent of artificial intelligence, natural language processing (NLP) research is currently advancing quickly (AI). Optical character recognition is one of the main areas of NLP (OCR). Digit classification is a useful foundation for creating an OCR in Bengali. Researchers can utilize our extensive dataset (85,000+) of Bengali digits, known as NumtaDB, to assess their system.

Six datasets that were collected from various sources and at various dates were combined to create the dataset. Each of them was extensively examined under the same assessment standard, though, to ensure that at least one human person with no prior knowledge could read all of the digits. On https://bengali.ai/datasets, descriptions of these datasets' collection process, image segmentation and extraction techniques, and image formats are provided.

The sources are categorized as "a" through "f." Depending on the data source, different subsets of the training and testing sets exist (training-a, testing-a, etc.). To ensure that the handwriting from the same subject or contributor is not included in both training and testing sets, all datasets have been divided into training and testing sets. Dataset-f lacked any associated contributor metadata, therefore all of it was put to the testing set (testing-f). The Unweighted Average Accuracy is chosen as the competition's measure (UAA). The competition's starter codes are accessible at https://github.com/BengaliAI.

References

[1] *M. M. Rahman, M. A. H. Akhand, S. Islam and P. C. Shill (2015). Bangla Handwritten Character Recognition using Convolutional Neural Network. I. J. Image, Graphics and Signal Processing, 8, 42-49.*

[2] U. Pal and B. B. Chaudhuri (2004). Indian script character recognition: a survey. Pattern Recognition, 37(9), 1887–1899.

[3] U. Pal and B. B. Chaudhuri (2000). Automatic recognition of unconstrained offline Bangla hand-written numerals. In: T. Tan, Y. Shi, W. Gao (eds) Advances in Multimodal Interfaces, ICMI 2000. Lecture Notes in Computer Science, 1948, 371–378.

[4] K. Roy, S. Vajda, U. Pal and B. B. Chaudhuri (2004). A system towards Indian postal automation. 9th International Workshop on Frontiers in Handwritten Recognition (IWFHR-9), Kokubunji, Tokyo, Japan, 580–585.

[5] U. Pal, A. Belaid and C. Choisy (2003). Touching numeral segmentation using water reservoir concept. Pattern Recognition Letters, 24(1-3), 261-272.

[6] S. M. A. Sharif, N. Mohammed, N. Mansoor and S. Momen (2016). A hybrid deep model with HOG features for Bangla handwritten numeral classification. 9th International Conference on Electrical and Computer Engineering (ICECE), Dhaka, Bangladesh, 463-466.

[7] U. Bhattacharya. Off-Line Handwritten Bangla Numeral Database. Available online: http://www.isical.ac.in/~ujjwal

[8] N. Das, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri and D. K. Basu (2012). A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application. Applied Soft Computing, 12(5), 1592-1606.

[9] M. Shopon, N. Mohammed and M. A. Abedin (2016). Bangla handwritten digit recognition using auto encoder and deep convolutional neural network. International Workshop on Computational Intelligence (IWCI), Dhaka, Bangladesh, 64-68.

[10] A. S. A. Rabby, S. Abujar, S. Haque and S. A. Hossain (2019). Bangla Handwritten Digit Recognition Using Convolutional Neural Network. In: A. Abraham, P. Dutta, J. Mandal, A. Bhattacharya, S. Dutta (eds) Emerging Technologies in Data Mining and Information Security. Advances in Intelligent Systems and Computing, 755, 111-122.

[11] M. Biswas, R. Islam, G. K. Shom, M. Shopon, N. Mohammed, S. Momen and A. Abedin (2017). BanglaLekha-Isolated: A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters. Data in Brief, 12, 103-107.

[12] M. Z. Alom, P. Sidike, T. M. Taha and V. K. Asari (2017). Handwritten Bangla Digit Recognition Using Deep Learning. Available online: http://arxiv.org/abs/1705.02680

[13] M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha and V. K. Asari (2018). Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks. Computational Intelligence and Neuroscience, 2018, Article ID 6747098, 13 pages.

[14] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M.S. Nasrin, B. C. V. Esesn, A. A. S. Awwal and V. K. Asari (2018). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. Available online: http://arxiv.org/abs/1803.01164

[15] S. Saha and N. Saha (2018). A Lightning fast approach to classify Bangla Handwritten Characters and Numerals using newly structured Deep Neural Network. Procedia Computer Science, 132, 1760-1770.

[16] A. Shawon, M. J. Rahman, F. Mahmud and M. M. A. Zaman (2018). Bangla Handwritten Digit Recognition Using Deep CNN for Large and Unbiased Dataset. International Conference on Bangla Speech and Language Processing (ICBSLP). Sylhet, Bangladesh, 1-6.

[17] H. Zunair, N. Mohammed and S. Momen (2018). Unconventional Wisdom: A New Transfer Learning Approach Applied to Bengali Numeral Classification. International Conference on Bangla Speech and Language Processing (ICBSLP), Sylhet, Bangladesh, 1-6.