

Bangla Handwritten Character Recognition

September 2021

1 About us

Submitted by: Group: EWUSM2021CSE475012

Name: Zainab Azad ID: 2017-1-60-148

Name: Md. Mazharul Islam ID: 2017-1-60-093

Name: Md. Fahim Sabab Siyam ID: 2017-1-60-142

Submitted to: Md Mahmudur Rahman Department of Computer Science and Engineering

2 Introduction

Bangla handwriting character recognition has a variety of academic and commercial areas, however due to the arrangement and familiarity of many of the characters, it is difficult to accomplish consistent results. Another difficulty in handwritten character detection is dealing with the large variety of handwriting styles used by various researchers. [1]. To identify characters and interpret them, existing methods are applied. Convolutional Neural Network (CNN) has recently been introduced to be successful at handwritten character recognition [2]. CNN models are proposed in this study for classifying Bangla handwriting characters. In this study, we study HBCR using state-of-the-art convolutional neural network (CNN) models [3]. We'll be using EfficientNet and LeNet models, that will be analyzed using a Bangla alphabet dataset and the results will be presented. The results will then be compared based on the performance and accuracy rate of these models, and we will be able to identify which model performs best for HBCR. Our goal is to use a program and an existing dataset to find a higher functioning algorithm to detect Bangla handwriting character recognition. HBCR has a wide range of applications, including Bangla optical character recognition, national ID number recognition, automatic license plate recognition for automobile and parking structure management, post automation of processes, and online banking, among many others [1]. Some publications that are relevant to your topic. As a result, it is important that HBCR applications be processed at a high rate. Our main goal in this study is to find out how to improve our performance in recognizing Bangla handwriting characters, as well

as to show how a convolutional neural network can improve accuracy for Bangla handwritten characters.

3 Methodology

In this project, we use two types of convolutional neural network methods. Which are: 1. LeNet 2. EfficientNet LeNet: LeNet is a basic convolutional neural network that is frequently known to as LeNet-5. LeNet-5 was one of the first convolutional neural networks, and it assisted deep learning progress. Convolutional neural networks are a set of feed-forward neural network with artificial neurons which can respond to a section of the surrounding cells in the coverage range and can do large-scale image processing [3]. LeNet, as an early convolutional neural network, possesses the basic units of convolutional neural networks, such as the convolutional layer, pooling layer, and complete connection layer, providing the basis for future convolutional neural network development [12]. The LeNet-5 denotes CNN's emergence and identifies its main components. However, it was not widely used at the time because to a shortage of hardware, mainly GPU, and other algorithms, such as SVM, can receive equal or even better returns than LeNet [4]. Since AlexNet's success in 2012, CNN has become the optimal option for computer vision applications, and several different types of CNN, like the R-CNN series, have been established [5]. CNN models now differ dramatically from LeNet, but they are all based on LeNet. Features:

- Convolution, pooling, and nonlinear activation functions should be included in every convolutional layer.
- Convolutional neural networks have been used to extract spatial features (Convolution was called receptive fields originally)
- Subsampling average pooling layer
- Tanh activation function
- As the final classifier, MLP was used.
- To lower the complexity of computing, it use sparse connections between layers.

EfficientNet: EfficientNet is a convolutional neural network structure and scaling method that uses a compound coefficient to scale all depth/width/resolution dimensions uniformly [6]. The EfficientNet scaling method consistently increases network breadth, depth, and resolution with a set of preset scaling coefficients, unlike conventional practice, which increases these factors individually. EfficientNet use a compound coefficient to scale network width, depth, and resolution evenly and continuously [7]. The compound scaling method is described by the concept that if the input image is larger, the network demands more layers to extend the receptive field and more channels to capture more fine-grained patterns on the larger image.

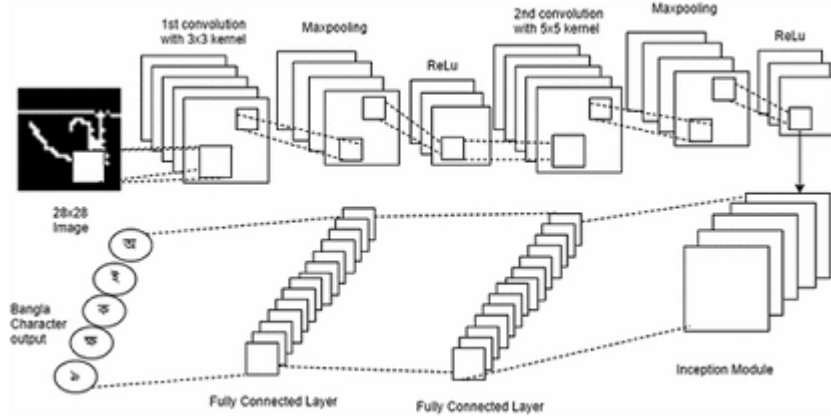


Figure 1: Figure of LeNet method



Figure 2: Samples of Bangla handwritten characters

4 Implementation

4.1 Data Collection

The text recognition algorithm is trained on the Bangla character level with the help of a dataset called ‘Cmaterdb’ that was found in the Google Code Archive [8]. We divided this dataset into smaller chunks to make the task go faster. There are 15001 handwritten character images in our dataset. It is made up of a collection of basic and compound letters collected from the Bangla language [2]. The dataset in Figure 1 has various labels for each character and group. It has a large number of frequently used Bangla basic characters and is shaped in a complex way.

4.2 Data Processing

Duplicate data in the training and test sets can cause the results to be biased. As a result, we proposed to reduce all data overlap between training and testing. For each photograph, we used a constant image size of 64 x 64 pixels. We employed the python imaging library pillow package to resize the photos, which made it effortless to do so.

4.3 Model Development

To create our application, we used the jupyter notebook, which is based on the Python programming language. Many python packages were utilized, such tensorflow, keras, pillow, numpy, and others. We used tensorflow to train our model on two separate methods for processing and working with handwritten picture detection and recognition. We used the pillow package for image resizing and processing, and numpy for exact mathematical functions and calculations. The versions we used to develop for this project are python 3.9.7, tensorflow 2.4.0. We did rescaling, resizing 12000 images to 50 different classes and 3000 to another 50 different classes. We had 100 steps per Epochs and 15 Epochs and 30 validation steps to train our model for both of the algorithms. After training the model, the weights are saved to .h5 format.

4.4 Results

We have used two different algorithms to test our model for Bangla handwriting character recognition which is EffectiveNet and LeNet.

After using EffectiveNet our result is

loss: 0.2410 - accuracy: 0.9266 - val_loss : 0.2231 - val_accuracy : 0.9365

After using LeNet our result is,

loss: 1.4851 - accuracy: 0.5694 - val_loss : 0.7116 - val_accuracy : 0.7844

Where our value is losing less in EffectiveNet and all the other results are quite similar. We decided to use LeNet algorithm to use to run the project and check the success ratio. We tested the model several times and our average success rate is 75

5 Conclusions

We attempted to recognize Bangla handwritten characters from a large number of input devices or devices and built a system that allows us to perform activities at the same time and efficiently [1]. Within the sense of imitating human intelligence to a small amount, artificial intelligence offers several benefits in classification and prediction. The first step into the broad area of Artificial Intelligence and Image Processing is handwritten character recognition [1]. With the least amount of planning, a convolutional neural network (CNN) may identify visual patterns straight from image patches [2]. This study that looked into

CNN structure without any feature selection for identification of Bangla handwritten patterns [2]. We placed this method to the test on a huge handwritten character dataset and compare these results to existing Bangla methods. On the basis of test set accuracy, the proposed method is competitive with existing methods, however the proposed scheme looks to be efficient in terms of size and computation. The given data indicates that CNN training could be improved, high performance. CNN, which again is modeled on the human visual cortex, can recognize visual patterns directly from pixel images with minimal pre-processing.

5.1 Challenges

1. For Bangla handwritten pattern recognition, CNN architecture is examined without any classification.
2. Although the deep CNN architecture is efficient in recognition of handwritten characters in Bangla, the dataset is large [2].
3. There's really yet no established method for correctly identifying all basic and compound characters [9].
4. A full Bangla file does not work in the system. Separate characters must be written and processed.

5.2 Limitations

1. For the classification of Bangla handwriting patterns, the CNN structure is examined without using any features.
2. Even though the deep CNN architecture is effective in recognition of handwritten characters in Bangla, the dataset is big. We also didn't have a powerful graphics unit to train a large dataset, so we had to split it up into smaller chunks.
3. Despite using two alternative algorithms, we were unable to find a method that correctly recognized all of the basic and complex characters.
4. A entire Bangla document does not function in the system. On this project, just characters are being used.

5.3 Future Scope

There are tremendous scopes of future extensions of this work. Some of the scopes are listed out below:

1. To obtain majority-based decisions, multiple CNN channels (CNN ensemble) may be used. The ensemble's projected error is always smaller than an individual predictor's moral judgments.

2. In the deep CNN model utilized in this research, a dropout layer could be introduced. Dropout is a regularization method that prevents complex co-adaptations on training data, avoiding overfitting in neural networks.
3. It's possible that inception modules will be introduced. The inception layer's goal is to cover a larger area while having a fine resolution for minute details on the photos. The theory is that a sequence of Gabor filters of various sizes will be able to handle a variety of object scales. With the added added bonus of being able to learn all of the filters on the conception layer. Using more layers and data is the most straightforward technique to improve deep learning performance. According to a study, adding Inception modules improves the accuracy. Nine Inception modules are used by Google Net [10].
4. By feeding the output of two previous convolutional layers AND bypassing the input to the next layers, residual network layers can be formed. The idea behind the residual network is to use blocks that reroute the input and build on the previous layer's concept. The concept is that during learning, the following layer will pick up on the previous layer's concepts as well as the preceding layer's data. This would be preferable to simply memorizing a notion without having a reference to refer to [5] .
5. The performance of the proposed CNN could be analyzed for Bangla compound characters and digits[2].
6. We can get a soft version of this data by employing OCR so that we can save it and use it when we need it or for an editable version. Day by day, it will grow into a large dataset, with the accuracy rate possible for all Bangla handwritten characters [11].

6 References

1. M. Z. Alom, P. Sidike, M. Hasan, T. M. Taha, and V. K. Asari, "Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks," *Comput. Intell. Neurosci.*, vol. 2018, 2018, doi: 10.1155/2018/6747098.
2. M. M. Rahman, M. A. H. Akhand, S. Islam, P. Chandra Shill, and M. M. Hafizur Rahman, "Bangla Handwritten Character Recognition using Convolutional Neural Network," *Int. J. Image, Graph. Signal Process.*, vol. 7, no. 8, pp. 42–49, Jul. 2015, doi: 10.5815/ijigsp.2015.08.05.
3. B. B. Chaudhuri U. Pal. A complete printed Bangla OCR system. *Pattern recognition*. 1988.31.5:531-549p.
4. S. Afshar, G. Cohen and J. Tapson, A. Schaik. EMNIST: An extension of MNIST to handwritten letters. *arXiv:1702.05373[cs.CV]*, 2017.

5. D. Kingma and J. Ba, Adam: A method of stochastic optimization. arXiv:1412.6980[cs.LG], 2015.
6. R. Gonzalez and R.E. Woods. Digital Image Processing. Third Edition, 162-163p.
7. S. Afshar, G. Cohen and J. Tapson, A. Schaik. EMNIST: An extension of MNIST to handwritten letters. arXiv:1702.05373[cs.CV], 2017.
8. “cmaterdb.” <https://storage.googleapis.com/google-code-archive-downloads/v2/code.google.com/cmaterdb/CMATERdb3.1.3.3.7z> (accessed May 20, 2021).
9. T. Ghosh, M. M. Abedin, S. Mahmud Chowdhury, and M. A. Yousuf, “A Comprehensive Review on Recognition Techniques for Bangla Handwritten Characters,” in 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), Sep. 2019, pp. 1–6, doi: 10.1109/ICBSLP47725.2019.202051.
10. C. Xu, D. Chai, J. He, X. Zhang, and S. Duan, “InnoHAR: A Deep Neural Network for Complex Human Activity Recognition,” IEEE Access, vol. 7, pp. 9893–9902, 2019, doi: 10.1109/ACCESS.2018.2890675.
11. R. R. Chowdhury, M. S. Hossain, R. ul Islam, K. Andersson, and S. Hos-sain, “Bangla Handwritten Character Recognition using Convolutional Neural Network with Data Augmentation,” in 2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR), May 2019, pp. 318–323, doi: 10.1109/ICIEV.2019.8858545.