

Bangla Hand Written Characters and Digits Recognition Using Convolutional Neural Network

Military Institute of Science and Technology

Tahasin Mahmud : 201714061, Md Ishak : 201714031

Introduction

Convolutional Neural Network (CNN) is a deep learning application which is inspired by the structure of visual system. CNN is widely used in face recognition, image classification, action recognition, human pose estimation and document analysis. A simple, lightweight CNN model has been proposed in this project for classifying Bangla handwriting characters consisting of 11 vowels, 39 consonants and 10 digits. The objectives of this project is to build a neural network model that can classify new sample images correctly in real time after training on a moderately large volume of handwritten characters and numbers.

Data and Features

Dataset was collected from a online database 'Ekush' which contained 306,464 handwritten images of 120 Bangla simple and compound characters and digits.

15000 images of 50 Bangla alphabet and 10 digits were taken form this huge dataset for training and testing the CNN model. Fig.1 shows the sample images from the dataset. The dataset contains wide variation of distinct characters because of different people's writing styles. Some of these character images are very complex shaped and closely correlated with others.



Figure 1: Sample images from the dataset

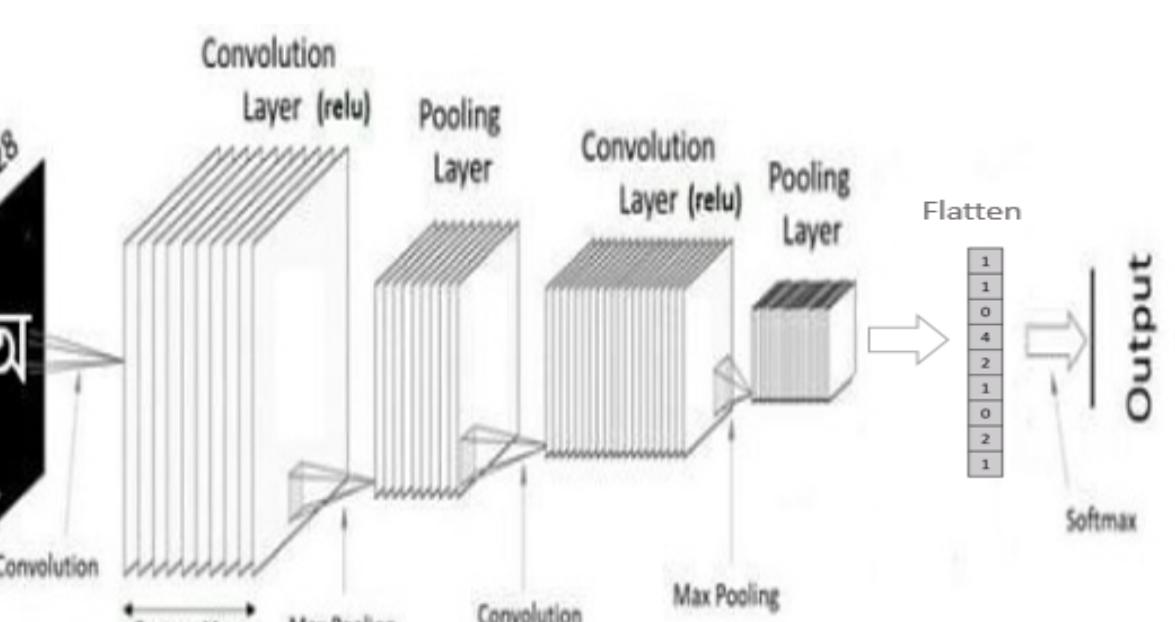


Figure 2: Steps involved in developing CNN model

These are the steps we followed for developing the CNN model:

- Step 1(a): Convolution Operation
- Step 1(b): ReLU Layer
- Step 2: Pooling
- Step 3: Flattening
- Step 4: Full Connection

Convolution Operation: In convolution operation we have an input image and feature detector. feature detector detects the feature of the images.

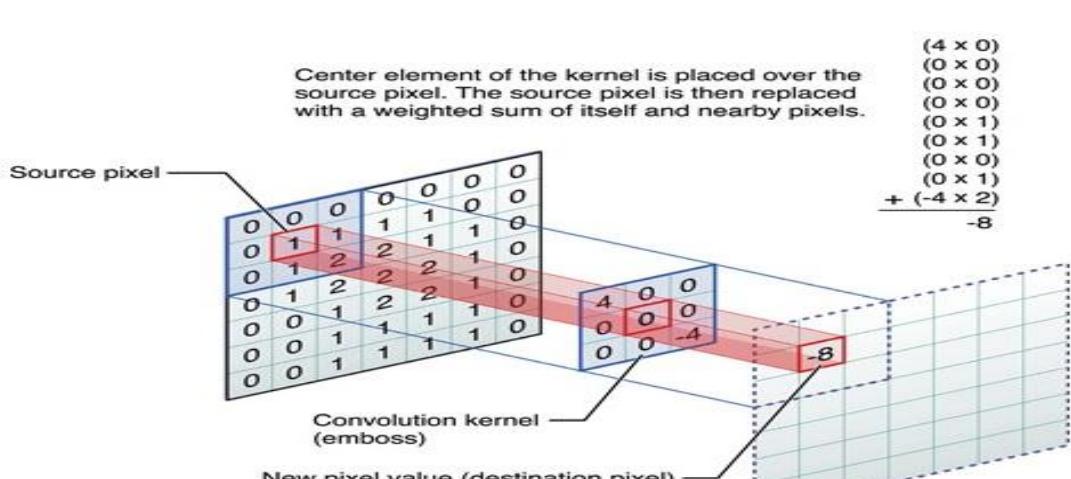


Figure 3: Convolution Operation using a 3x3 kernel

Models

Convolutional Neural Network are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other.

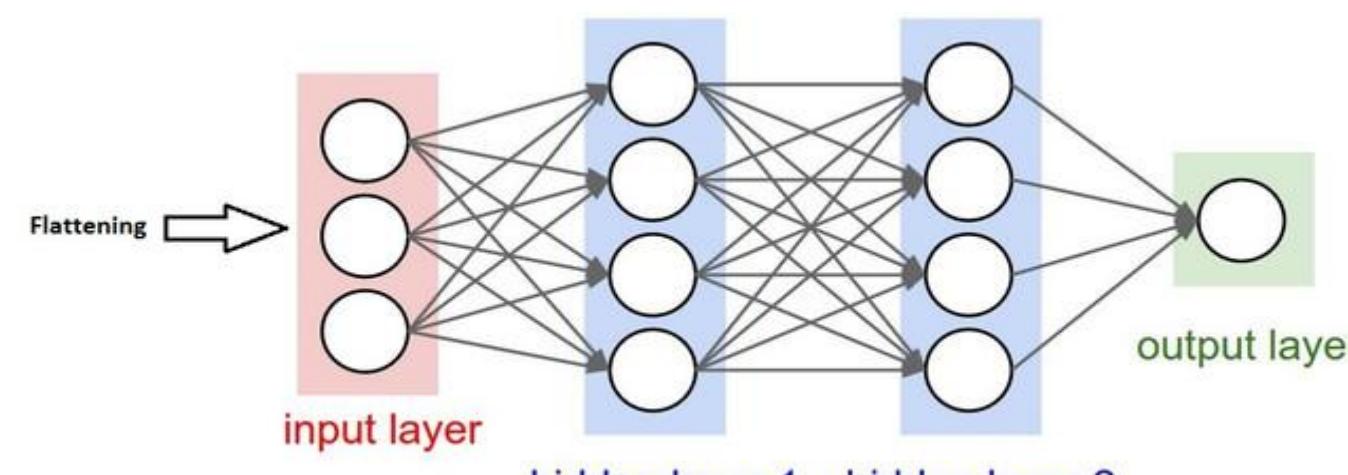


Figure 4: A fully connected CNN model

Conclusions

- CNN has ability to visual patterns directly from pixelimages with minimal pre-processing.
- The method is shown competitive performance with existing methods on the basis of the test set accuracy
- The presented result indicates that training of CNN might be improved hence get better performance

Future Work

- In the dataset of 'Ekush' database, the information of a person's age, gender, education level and area was attached with the image. Future work will include predicting person's age, gender, education level and location using CNN.
- Predict if the image is hand-written or typed image.

References

- [1] L. F. C. Pessoa and P. Maragos, Neural networks with hybrid morphological/rank/linear nodes: a unifying framework with applications to handwritten character recognition, Pattern Recognition, vol. 33, no. 6, pp. 945-960, June 2000.
- [2] Bellili, M. Gilloux and P. Gallinari, An MLP-SVM combination architecture for offline handwritten digit recognition, Document Analysis and Recognition, Springer-Verlag, vol. 5, no. 4, pp. 244-252, 2003.
- [3] G. Vamvakas, B. Gatos and S. J. Perantonis, Handwritten character recognition through two-stage foreground subsampling, Pattern Recognition, vol. 43, no. 8, pp. 2807-2816, August 2010.
- [4] U. Bhattacharya, M. Shridhar, S. K. Parui, P. K. Sen and B. B. Chaudhuri, Offline recognition of handwritten Bangla characters: an efficient two-stage approach, Pattern Analysis and Applications, vol. 15, no. 4 , pp. 445-458, 2012.
- [5] T. K. Bhowmik, P. Ghanty, A. Roy and S. K. Parui, SVMbased hierarchical architectures for handwritten Bangla character recognition, International Journal on Document Analysis and Recognition, vol. 12, no. 2, pp. 97-108,2009.

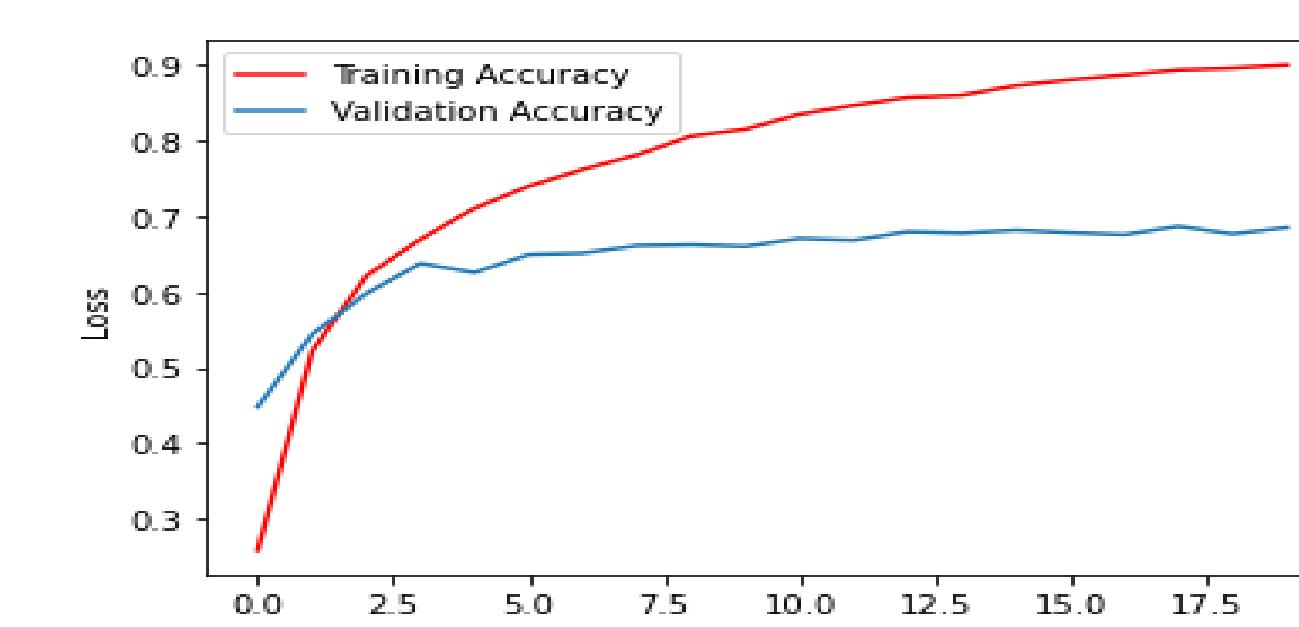


Figure 5: Train and Validation accuracy

It is evident from Fig.6 loss decreases with the increase of epoch which means model's performance improves gradually.

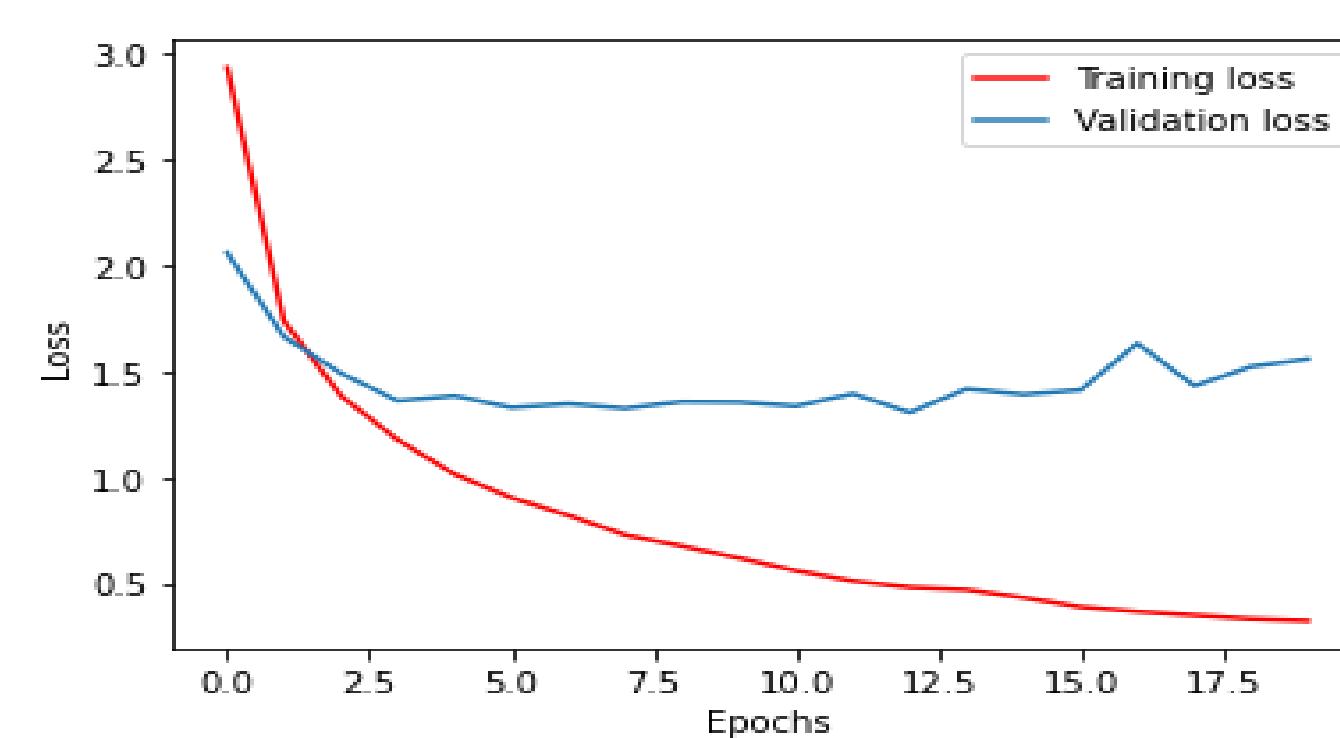


Figure 6: Train and Validation loss