Bangla Handwritten Digit Recognition Using Deep Convolutional Neural Network

Conference Paper · January 2020

DOI: 10.1145/3377049.3377077

Rabeya Basri
Bangladesh University of Business and Technology

Md. Reduanul Haque
Daffodil International University

Mohammad Shorif Uddin
Jahangirnagar University

Bangla Handwritten Digit Recognition Using Deep Convolutional Neural Network

Rabeya Basri
Department of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
rabeyabasri3@gmail.com

Mohammad Reduanul Haque
Department of Computer Science and Engineering
Daffodil International University
Dhaka, Bangladesh
reduan.cse@diu.edu.bd

Morium Akter
Department of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
moriumcse@gmail.com

Mohammad Shorif Uddin
Department of Computer Science and Engineering
Jahangirnagar University
Savar, Dhaka, Bangladesh
shorifuddin@gmail.com

ABSTRACT

Handwritten Bangla digit recognition is one of the most challenging computer vision problems due to its diverse shapes and writing style. Recently deep learning based convolutional neural network known as deep CNN finds wide-spread applications in recognizing different objects due to its high accuracy. This paper investigates the performance of some state-of-the-art deep CNN techniques for the recognition of handwritten digits. It considers four deep CNN architectures, such as AlexNet, MobileNet, GoogLeNet (Inception V3), and CapsuleNet models. These four deep CNNs have been experimented on a large, unbiased and highly augmented standard dataset, NumtaDB and confirmed that the AlexNet showed the best performance on the basis of accuracy and computation time.

CCS CONCEPTS

Computing methodologies → Artificial intelligence → Computer vision → Computer vision problems → Object recognition

KEYWORDS

Handwritten Bangla digit recognition, deep convolutional neural network, AlexNet, MobileNet, GoogLeNet, CapsuleNet, NumtaDB.

ICCA 2020, January 10–12, 2020, Dhaka, Bangladesh

ACM ISBN 978-1-4503-7778-2/20/01…$15.00
https://doi.org/10.1145/3377049.3377077

ACM Reference Format:

Rabeya Basri, Mohammad Reduanul Haque, Morium Akter and Mohammad Shorif Uddin**.** 2019. Bangla Handwritten Digit Recognition Using Deep Convolutional Neural Network. In ICCA 2020: International Conference on Computing Advancements, January 10–12, 2020, Dhaka, Bangladesh. ACM, Dhaka, Bangladesh, 7 pages. https://doi.org/10.1145/3377049.3377077

1 Introduction

In recent years, handwritten digit recognition receives great attention to the researchers due to its numerous application potentials such as national ID number recognition, automatic license plate recognition for vehicles, postal code recognition on envelopes, amount recognition on bank cheques and many more. It enables a machine to detect and classify the sample images of digits into 10 classes from 0 to 9. However, Bangla numerals recognition research lacks concentration although it is one of the most popular languages of the world [1]. Moreover, in a consequence of putting emphasis on digitalization by the Governments, importance of Bangla document analysis has increased more than ever before.

Meanwhile, several factors, e.g. variations in writing styles of individual digits, similarities of different character in shapes, overlaps, and the interconnections of the neighboring characters etc. make the recognition process more complex and challenging.

To overcome these difficulties researchers [1-23] attempted different methods for better recognition of handwritten Bangla digits. Recently, deep neural network, especially convolutional neural network (CNN) based solutions shows notable accuracy in image recognition, classification, annotation and various other fields that learns, extracts and classifies features automatically [1,10,13,14,16,17,19,20,22,23].

Pal et al. [6–9] worked for Indian postal automation using water reservoir scheme and achieved the accuracy of 94.13% and 93%, for the handwritten Bangla and English numerals, respectively, without specifying the response time and the recognition reliability. Sharif et al. [10] proposed a hybrid model combining deep CNN with HOG features to classify the handwritten Bangla numerals with lesser epochs. They obtained 99.02% accuracy for ISI dataset [11] and 99.17% accuracy for CAMTERDB dataset which consists of 6000 images (600 images per digit) of unconstrained handwritten Bangla digits, among them for training and testing 4000 and 2000 images are used respectively [12]. But integrating this hybrid model with multiple handcrafted features need further research.

Shopon et al. [13] used deep CNN with unsupervised pre-trained auto encoder for handwritten Bangla digit recognition. They gained an accuracy of 99.50% for CMATERDB but unable to pre-train larger datasets for better outcome.

Rabby et al. [14] presented lightweight CNN model that has applied on CAMTERDB BanglaLekha-Isolated [15] dataset provided validation accuracy of 99.74% and 98.93% respectively. But it was unable to eliminate the confusions for overwriting and erroneous labeling of images.

Alom et al. [16] achieved an accuracy of 98.78% applying CNN with Gabor feature using CMATERDB dataset. They also worked on handwritten Bangla character recognition using different deep learning models and showed that DenseNet provided the highest recognition accuracy of 98.31% for alphabet, 99.13% for digits, and 98.18% for special characters [17]. They also provided a very detailed study on different state-of-the-art deep models with different learning techniques. They used the AlexNet and GoogLeNet models to classify ImageNet dataset [18].

Akhand et al. [19] introduced an artificial pattern-based CNN method for better supervised learning and accuracy, which provided an accuracy of 98.98% using ISI dataset.

Shaha and Shaha [20] obtained an accuracy of 97.21% after applying a new deep CNN model over Bangla handwritten isolated character dataset. They put special attention on classifying basic and compound Bangla handwritten characters and numerals. Recently, NumtaDB dataset [21] has been widely used as a standard dataset, containing more than 85000 sample images. It is a challenging dataset as it contains no preprocessed data like MNIST and CMATERDB dataset.

Shawon et al. [22] proposed deep CNN model obtained testing accuracy of 92.72% using NumtaDB dataset. But their system perform poor for translation and rotational scenarios.

Zunair et al. [23] presented an unconventional transfer learning approach using VGG-16 deep CNN model to classify images of diverse Bangla digits and achieved an accuracy of 97.09% on NumtaDB dataset by freezing intermediate layers with less epochs and parameters.

There is a confusion in using a particular CNN model, as extensive comparative works on the effectiveness of these models on Bangla digit recognition is not yet performed. For these reasons, who are new in the field seeking a good comparative understanding of the available techniques surrounding each conclusion are in a dilemma. To cope with this issue, the current paper tries to perform an extensive experimentation on four widely used state-of-the-art deep CNN architectures, like AlexNet, MobileNet, Inception V3 and CapsuleNet. This leads us to the following contribution:

It investigates the performance of some state-of-the-art deep CNN techniques to find the best one for the recognition of handwritten Bangla digits.

The rest of the paper is organized as follows: the general architectures of four deep CNN models are described in Section 2. Section 3 presents the experimental results and discussions. Finally, Section 4 presents the conclusion along with future research scopes.

2 Digit Recognition Method

Several methods exist to solve image recognition tasks including SVM (support vector machine), KNN (K-nearest neighbor), and CNN (convolutional neural network). Among them CNN especially deep CNN has exhibited an outstanding performance on the largest image classification dataset, ImageNet [24] and has enormous applications for natural language processing (NLP) [25]. This research focuses on the evaluation of the individual performance of four widely used deep CNN models including AlexNet, MobileNet, Inception V3 and CapsuleNet and provides a comparative evaluation among these four models to find the one that gives the highest recognition accuracy.

2.1 Convolutional Neural Network

Convolutional Neural Network (CNN) is a deep appearance of traditional Artificial Neural Network (ANN) architecture. In 1980, CNN was first proposed by Fukushima [26]. A basic CNN architecture is composed of a huge number of self-optimistic neurons with learnable weights and biases. CNN is trainable with the gradient-based learning algorithm for minimizing error. LeCun et al. [27] proposed a gradient-based learning algorithm to train CNN (LeNet-5) for digit recognition and achieved an excellent result. Being a supervised method, CNN is able to reduce the number of parameters in order to solve complex image recognition tasks.

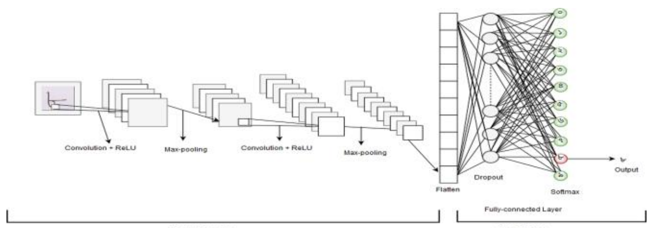

Figure 1: Basic Schematic Diagram of CNN-based Digit Recognition System.

A basic CNN architecture consisting of two main parts such as feature extractor and classifier and is shown in Figure 1. The feature extraction unit consists of convolution layer, an activation function and pooling layer, where the output from the previous layer is served as the input to the next layer. The classification unit consists of a fully-connected layer.

Convolution is the basic building block of CNN which is a mathematical combination of two functions that merges two sets of information to produce a third function. The convolution is performed on the input data using a filter (kernel) to produce a feature map. A 2D activation map is generated via convolving this filter over the input image. It reduces the complexity of the model by optimizing the output through three hyper parameters including depth, stride and zero-padding [28]. The next layer is the activation layer known as Rectified Linear Unit (ReLU) [29].

The pooling layer mainly causes down-sampling over the input dimension for reducing the complexity for further layers, which is referred to as the resolution reduction in image processing. It shortens the training time and controls the over-fitting. The most popular pooling method known as Max-pooling divides the image into sub-region rectangles and picks the maximum value among other values inside of that sub-region [29].

After the completion of convolution and pooling layers features are extracted and the number of parameters compared to the original images are reduced. A fully-connected layer is required to produce the final output containing classes equal to the number of input classes to determine whether an image belongs to a particular class or not.

A normalized exponential function, Softmax is added to the end of the architecture, used for multi-class classification. It generates the output sum between 0 and 1, so the output can be interpreted as probability [30]. Finally, an optimization technique is used to calculate an accurate and optimum result by training the models in a more effective and efficient manner. Here**,** we use the Adam optimizer to minimize the loss function as it converges very fast and increases the learning speed of the models [31].

2.2 Deep Convolutional Neural Network

Deep Convolutional Neural Network (deep CNN) is an extension of CNN, require more convolutional and pooling layers in the architecture (usually the number of layers must be more than three). In the last few years, deep CNN has become very popular for Bangla handwritten digit recognition as being an efficient learning technique. Remarkable examples of such networks include LeNet, AlexNet, VGGNet, ResNet, MobileNet, GoogleNet, CapsuleNet etc. Among them, MobileNet and GoogleNet are particularly designed for large scale implementation, but AlexNet and CapsuleNet are for general purpose.

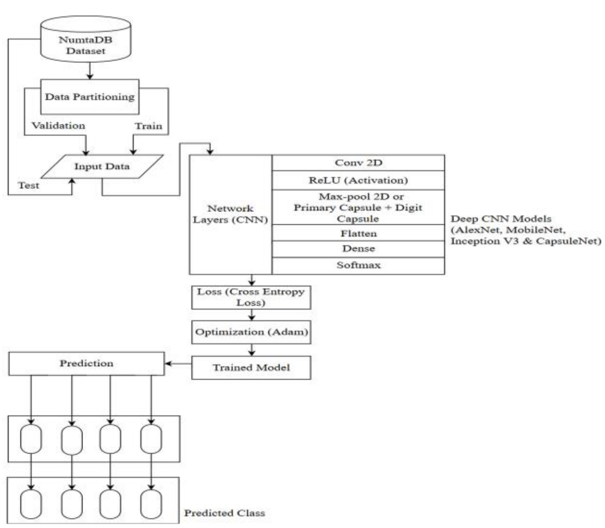

Figure 2: Flow Diagram for Implementing Handwritten Digit Recognition.

In this research, we have implemented AlexNet, MobileNet, GoogleNet (Inception V3) and CapsuleNet models. After that we have used NumtaDB dataset to evaluate the accuracy for Bangla handwritten digit recognition. For the convenience of understanding, an overview of these four deep CNN models is described as follows:

2.2.1 AlexNet

AlexNet proposed by Alex Krizhevesky [32] was the winner of ILSRVC 2012. AlexNet consists of 5 convolutional layers and 2-fully-connected (FC) Layers with ReLU activation and dropout**.** A single convolutional layer of this architecture contains multiple kernels of same size to extract interesting features from the input.

2.2.2 MobileNet

MobileNet is an efficient CNN architecture, based on a factorized form of convolutions with good accuracy for classifying images on mobile devices for embedded vision applications [33]. MobileNet generally uses 3×3 depthwise separable convolutions.

2.2.3 GoogLeNet (Inception V3)

GoogLeNet proposed by Christian Szegedy, was the winner of ILSVRC 2014 and the 1st runner up of ILSVRC 2015 for image classification and detection. This model used average pooling instead of fully connected layers for reducing computation complexity by removing a large amount of parameters compared to the traditional CNN [34]. GoogLeNet is a 22 layers deep CNN architecture containing a basic Inception module, which operates some convolution operations and later sums up the results to improve the state-of-the-art recognition accuracy. This Inception architecture was refined in various ways by performing batch normalization, additional factorization using activation.

2.2.4 CapsuleNet

A structured model called CapsuleNet consisting of two different capsule layers. A primary capsule layer that represents a group of neurons, working together as a capsule unit and a digit capsule layer that is obtained through a mutual agreement among different primary capsules by dynamic routing. CapsuleNet uses squashing nonlinearity instead of using softmax or sigmoid to generate the output as a vector [35].

A brief architectural overview of these four deep CNN models is summarized in Table 1.

Table 1: Brief Architecture of Deep CNN Models.

	Model Name			
Properties	**AlexNet**	**MobileNet**	**Inception V3**	**CapsuleNet**
Input Size	224×224	224×224	227×227	28×28
Conv. Layers	5	28	21	1
Filter Size	3, 5, 11	1, 3	1, 3, 5	9
Stride	2	2	1, 2	1, 2
Parameter	60M	4.2M	23M	8.2M
FC Layer	2	1	1	3

The overall implementation for recognizing handwritten Bangla digits is illustrated by a flow diagram in Figure 2.

3 Experimental Results and Discussions

For experimentation, we have used NumtaDB dataset, consisting of six different datasets that were collected from different sources and at different times. This dataset contains approximately 89,671 images of Bangla handwritten digits. All of these six datasets have been separated into training and testing sets. Both training and test sets contain individual subsets depending on the source data such as training-a, testing-a etc. to avoid the presence of duplicate images in both. The sources are labeled from 'a' to 'f'. The digits in dataset ‘a’ are collected from Bengali Handwritten Digits Database (BHDDB), similarly digits in dataset ‘b’ are collected from BUET101 Database (B101DB), digits in ‘c’ are collected from OngkoDB, digits in ‘d’ are collected from ISRTHDB, digits in dataset ‘e’ are collected from curated version of BanglaLekha-Isolated Numerals and digits in ‘f’ are collected from UIUDB. Dataset-f is added only to the testing set, labeled as testing-f, as it contains no corresponding metadata for contributors. Two augmented datasets connected to the testing set are augmented from test images of dataset 'a' and 'c' including the changes in spatial transformations i.e. rotation, translation, shear, shift and zoom along with the changes in brightness, contrast, saturation, noise associated with occlusions and superimposition [21]. Datasets statistics are shown in Table 2.

Some sample images from both training and test sets are shown in Figure 3 to Figure 5. Figure 6 shows the experimental results along with test images.

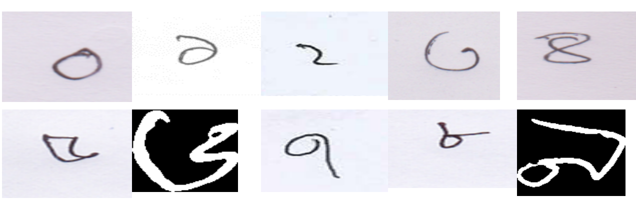

Figure 3: Sample images from different training sets (training a to e).

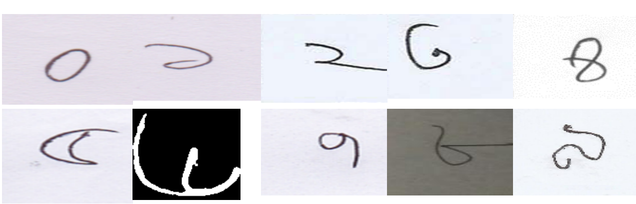

Figure 4: Sample images from different testing sets (testing a to f).

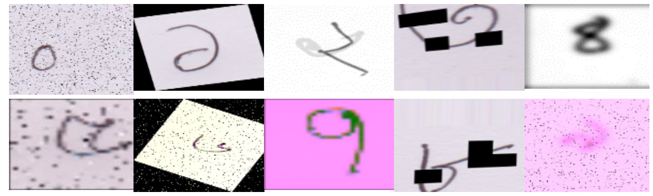

Figure 5: Sample images from augmented sets (aug-a and aug-c).

Table 2: Dataset Statistics (for NumtaDB).

Types of Dataset	Dataset	Quantity	Total
Training	a	19,702	72,045
	b	359	
	c	24,298	
	d	10,908	
	e	16,778	
Testing	testing-a	3,489	17,626
	testing-b	69	
	testing-c	4,381	
	testing-d	1,948	
	testing-e	2,970	
	testing-f	495	
	Aug-a	2,168	
	Aug-c	2,106	

Table 3: Recognition Accuracy under Different Deep CNN Models using NumtaDB Dataset. *[N: Normal; A: Augmentation].

Method	Accuracy	
	N	N+A
AlexNet	92%	99.01%
MobileNet	79.7%	83.58%
GoogLeNet (Inception V3)	93%	98.51%
CapsuleNet	91.3%	92.24%

Table 4: Confusion Matrix for AlexNet.

	Predicted Class									
Actual Class	0	1	2	3	4	5	6	7	8	9
0	100	0	0	0	0	0	0	0	0	0
1	5	95	0	0	0	0	0	0	0	0
2	0	0	100	0	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	0	0	100	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0
7	0	0	0	0	0	0	0	100	0	0
8	0	0	0	0	0	0	0	0	100	0
9	0	28.6	0	0	0	0	0	0	0	71.4

Accuracy = 99.01%

Table 5: Confusion Matrix for MobileNet.

	Predicted Class									
Actual Class	0	1	2	3	4	5	6	7	8	9
0	95	0	5	0	0	0	0	0	0	0
1	0	95	0	5	0	0	0	0	0	0
2	0	0	100	0	0	0	0	0	0	0
3	5	0	0	95	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	30	60	10	0	0	0	0
6	0	0	0	5	30	35	30	0	0	0
7	5	0	0	0	25	5	0	65	0	0
8	10	0	0	0	20	0	0	5	65	0
9	0	14.3	4.8	0	0	0	0	0	0	80.9

Accuracy = 83.58%

Table 6: Confusion Matrix for Inception V3.

	Predicted Class									
Actual Class	0	1	2	3	4	5	6	7	8	9
0	100	0	0	0	0	0	0	0	0	0
1	15	85	0	0	0	0	0	0	0	0
2	0	5	95	0	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	0	0	100	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0
7	0	0	0	0	0	0	5	95	0	0
8	10	5	5	0	0	0	5	0	75	0
9	0	4.8	0	0	0	0	0	4.8	0	90.4

Accuracy = 98.51%

Table 7: Confusion Matrix for CapsuleNet.

	Predicted Class									
Actual Class	0	1	2	3	4	5	6	7	8	9
0	100	0	0	0	0	0	0	0	0	0
1	15	83	0	0	0	0	0	2	0	0
2	0	5	90	5	0	0	0	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	0	0	100	0	0	0	0
6	0	0	0	0	0	0	100	0	0	0
7	7	0	0	0	0	0	5	88	0	0
8	7.8	5	5	0	0	0	5	0	77.2	0
9	0	4.8	6.2	0	0	0	0	4.8	0	84.2

Accuracy = 92.24%

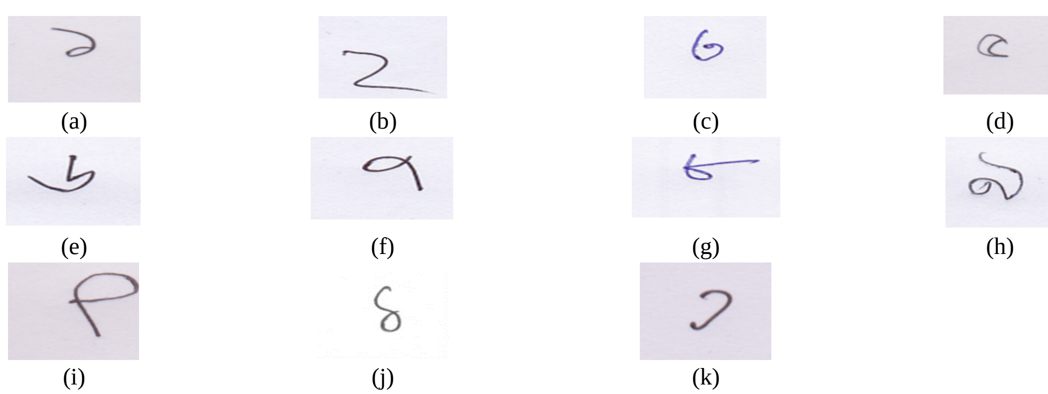

Figure 6: Examples of images of 11 digits (from the NumtaDB Dataset. Digit images in (a), (c)-(g) are recognized by all CNNs correctly; (b) is recognized by InceptionV3 and AlexNet but MobileNet, CapsuleNet failed to recognize; (h) is recognized by only AlexNet. However, all the techniques failed to recognize (i), (j) and (k) due to their inappropriate size and shapes.

Table 8: Computation Time for Four Deep CNN Models.

Model Name	Computation Time (second)
AlexNet	1.14
MobileNet	12.52
GoogLeNet (Inception V3)	22.53
CapsuleNet	3.86

4 Conclusion and Future Scope

As stated above, recognizing handwritten digits is one of the most challenging tasks that has already occupied a great importance due to its various practical applications in daily life.

This research is reported on the implementation of four state-of-the-art deep CNN architectures on NumtaDB dataset and reports the evaluation of the performance of each model. After investigating the individual performance, it can be observed that among the four models, AlexNet provides the highest recognition accuracy of 99.01% over the normal and augmented data, which can be said the best recognition accuracy compared to previous related works. The implementation is much-time consuming process, requiring high configurable machine with GPU. The future endeavor for this research is to implement these deep CNN models through a transfer learning approach using an inference rule in order to improve accuracy.

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] Y. Qiao. The Mnist Database of handwritten digits. Available online: http://www.gavo.t.u-tokyo.ac.jp/~qiao/database.html

[3] U. Bhattacharya, M. Shridhar, S. K. Parui, P. K. Sen and B. B. Chaudhuri (2012). Offline recognition of handwritten Bangla characters: an efficient two-stage approach. Pattern Analysis and Applications, 15(4), 445-458.

[4] B. B. Chaudhuri and U. Pal (1998). A complete printed Bangla OCR system. Pattern Recognition, 31(5), 531–549.

[5] U. Pal (1997). On the development of an optical character recognition (OCR) system for printed Bangla script. Ph.D. Thesis. Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India.

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

[7] 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.

[8] 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.

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

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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

[17] 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.

[18] 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

[19] M. A. H. Akhand, M. Ahmed and M. M. H. Rahman (2016). Convolutional neural network training with artificial pattern for Bangla handwritten numeral recognition. 5th International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, Bangladesh, 625-630.

[20] 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.

[21] S. Alam, T. Reasat, R. M. Doha and A. I. Humayun (2018). NumtaDB - Assembled Bengali Handwritten Digits. Available online: http://arxiv.org/abs/1806.02452

[22] 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.

[23] 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.

[24] J. Deng, W. Dong, R. Socher, L. Li, Kai. Li and Li. Fei-Fei (2009). ImageNet: A large-scale hierarchical image database. IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 248-255.

[25] A. Gelbukh (2005). Natural language processing. 5th International Conference on Hybrid Intelligent Systems (HIS'05), Rio de Janeiro, Brazil, 1.

[26] K. Fukushima (1980). A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36(4), 193–202.

[27] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner (1998). Gradient-based learning applied to document recognition. In Proceedings of the IEEE, 86(11), 2278–2324.

[28] K. O'Shea and R. Nash (2015). An Introduction to Convolutional Neural Networks. Available online: http://arxiv.org/abs/1511.08458

[29] S. Albawi, T. A. Mohammed and S. Al-Zawi (2017). Understanding of a convolutional neural network. International Conference on Engineering and Technology (ICET), Antalya, Turkey, 1-6.

[30] R. B. Arif, M. A. B. Siddique, M. M. R. Khan and M. R. Oishe (2018). Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Convolutional Neural Network. 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), Dhaka, Bangladesh, 112-117.

[31] V. Ayumi, L. M. R. Rere, M. I. Fanany and A. M. Arymurthy (2016). Optimization of Convolutional Neural network using Microcanonical Annealing Algorithm. International Conference on Advanced Computer Science and Information Systems (ICACSIS), Malang, Indonesia, 506-511.

[32] A. Krizhevsky, I. Sutskever and G. E. Hinton (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

[33] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Available online: http://arxiv.org/abs/1704.04861

[34] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich (2015). Going Deeper With Convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 1-9.

[35] B. Mandal, S. Dubey, S. Ghosh, R. Sarkhel and N. Das (2018). Handwritten Indic Character Recognition using Capsule Networks. IEEE Applied Signal Processing Conference (ASPCON), Kolkata, India, 304-308.