2020 23rd International Conference on Computer and Information Technology (ICCIT), 19-21 December, 2020

Bengali Handwritten Isolated Compound Characters Recognition by Applying Transfer Learned Deep Convolutional Neural Network

Md. Mehedi Hasan*, Azmain Yakin Srizon†, Abu Sayeed‡ and Md. Al Mehedi Hasan§
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh
Email: *mmehedihasann@gmail.com, †azmainsrizon@gmail.com, ‡abusayeed.cse@gmail.com, §mehedi_ru@yahoo.com

2020 23rd International Conference on Computer and Information Technology (ICCIT) | 978-1-6654-2244-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICCIT51783.2020.9392699

Abstract—**Optical character recognition (OCR) has been an area of interest for researchers for decades. Many researchers have contributed largely to the development of OCR for script-specific recognition. Handwritten characters recognition has been a big part of this field. Although some literatures have received well-established outcomes, others still haven't experienced remarkable outcomes yet. Despite being the fifth most spoken language of the world by 228 million people, Bengali has not yet received a remarkable contribution to handwritten characters recognition. Some researchers have offered some promising results for basic Bengali handwritten characters recognition but very few researches offered the recognition of Bengali compound handwritten characters. These few researches have applied support vector machine classifier and some deep neural network classifiers for classification but the outcomes were not much satisfactory. In this research, we considered 171 Bengali compound handwritten characters and apply a modified ResNet-50 model for recognition. We achieved an overall accuracy of 96.15% which outperformed the previous best result of 90.33% by a remarkable margin.**

Index Terms—**Bengali Handwritten Compound Isolated Characters, Transfer Learning, Modified ResNet-50**

I. Introduction

Handwritten characters recognition, an important part of pattern recognition and artificial intelligence, has been essential in various sectors like commercial OCRs, reading of bank checks, postal addresses, etc. [1], [2]. In almost all these applications, offline characters are recognized which is typically known as optical character recognition (OCR). Previously, researches have been conducted on the handwritten characters but due to a lack of good quality cameras for capturing images, the results were not satisfactory. Lack of benchmark databases was also a problem for the recognition of the handwritten characters. The development in cameras and benchmark datasets has broadened the scope of research in the field of handwritten character recognition in different applications like reading of medical prescriptions and business cards [3].

The domain of the handwritten character recognition problem is highly dependent on the script. For example, different language has different alphabets or characters, hence, the script-specific features contribute to the enhancement of the classification. Being one of the most spoken languages of the world, Bengali handwritten character recognition has been an area of interest for the researchers. But like other common languages, Bengali not only has basic vowel and consonant characters, it also has a large number of compound characters. The Bengali language has almost 400 characters in total which makes the Bengali handwritten characters recognition more complicated. Very few researches have been conducted in the area of Bengali handwritten characters recognition compared to other languages [4], [5], [6], [7]. Moreover, most of the researches focused on the recognition of the basic Bengali handwritten characters. A very smaller amount of researches have been conducted on the recognition of Bengali compound handwritten characters which covers almost 85% of the total characters of the Bengali literature [1], [8], [9], [10], [11]. More about previous researches has been addressed in the literature review section.

In this research, we started with the CMATERdb Bengali compound isolated dataset. At first, some preprocessing steps were conducted on the dataset. After that, we applied the ResNet-50 model for training. Finally, the performance was measured in terms of class-specific and overall precision, recall, f-score, and support. Our approach achieved an overall accuracy of 96.15% which outperformed all previous researches by a noticeable margin.

II. Literature Review

Bengali handwritten character recognition has been a popular area of research for a decade now [1], [8], [9], [10], [11]. Many efforts have been provided to recognize the core Bengali handwritten letters [12], [13]. One of the best examples is the MNIST dataset having 10 classes with 60,000 training images [14]. Having a large training set with fewer classes made the process easier to recognize the Bengali characters in various researches [4], [5], [6], [7], [12], [13]. However, the complexity arises when compound Bengali handwritten characters were taken under consideration because of having a large number of classes with complicated and various writing styles. In 2007, a research introduced the recognition of 138 Bengali compound characters by applying a modified quadratic discriminant function classifier [8]. A multilayer perceptron classifier was employed to recognize the quad tree-based features in 55 frequently used Bengali compound

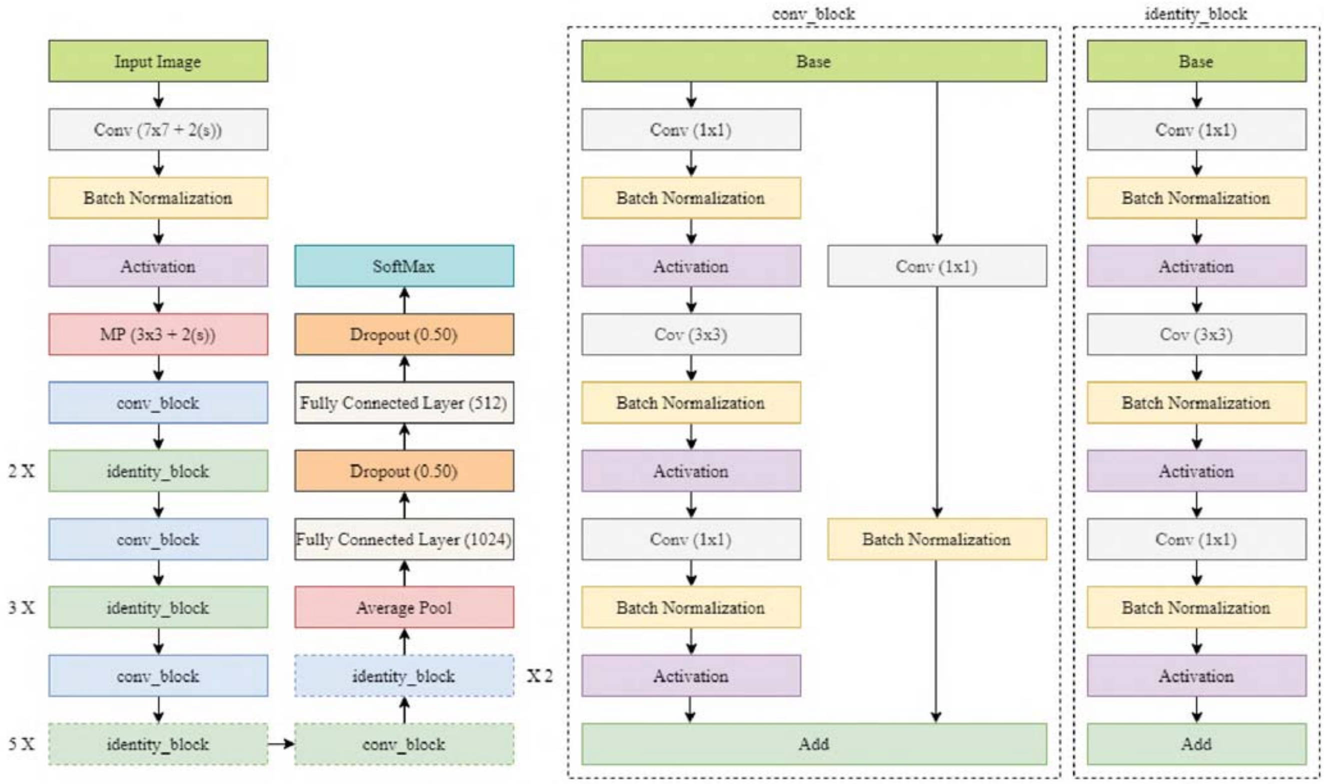

Fig. 1: Proposed modified ResNet-50 architecture along with the elaboration of the convolution block and identity block

characters in 2009 [1]. In 2010, the research was upgraded by considering 93 Bengali characters (50 basic and 43 frequently occurred Bengali characters) [9]. This time, both quad tree-based and shadow features were considered along with support vector machine and multilayer perceptron classifiers. Meanwhile, a technique was proposed to enhance the recognition of by considering topological features calculated from convex shapes of different strokes [15].

In 2015, the two-pass technique was introduced to classify the Bengali handwritten characters which dramatically improved the recognition process [16]. A total of 171 Bengali compound characters were considered in this research along with other basic Bengali characters. Longest run-based and convex-hull based features were considered and a support vector machine classifier was employed to achieve an overall accuracy of 87.26% in this research. In another investigation, 86.65% classification accuracy was achieved by following two different sorting optimization techniques to recognize the most valuable features of the images [17]. Most recently another research suggested a deep convolutional neural network-based approach which achieved an overall accuracy of 90.33% where they considered all 171 Bengali compound classes [18]. In this research, we considered the same dataset used in [1], [8], [9], [16], and [18] to recognize the compound Bengali characters.

III. Materials and Methods

In this section, first, the dataset description has been provided. After that, a description of transfer learning and the residual neural network has been introduced. Finally, our proposed architecture has been presented.

A. Dataset Description

In this research, we considered the publicly accessible CMATERdb Bengali compound isolated dataset having a total of 171 classes [1]. There are a total of 34,439 images in the train set and 8,520 images in the test set. The dataset is imbalanced and the images are not fixed-sized which makes the recognition more challenging. Moreover, for some of the Bengali compound characters, there are different writing styles. Setting the same class label for two different writing styles makes the recognition more complicated.

B. Transfer Learning

Transfer learning concentrates on collecting information obtained while resolving one obstacle and implementing it to another but similar dilemma [19]. For instance, the information obtained while discovering to identify cars could utilize during the recognition of trucks. This field of investigation shows a remarkable relationship to the enduring history of cerebral research on the transfer of learning, though confirmed relations among the two areas are inadequate. From a pragmatic viewpoint, transferring or conveying knowledge from earlier accomplished assignments for the training of new jobs has the potentiality to dramatically enhance the individual performance of an agent [20].

TABLE I: Class-wise precision, recall, f1-score and support illustration

Class	Precision	Recall	F1-Score	Support	Class	Precision	Recall	F1-Score	Support
1	0.87	1.00	0.93	52	87	0.95	1.00	0.97	54
2	0.98	1.00	0.99	55	88	0.82	0.78	0.80	51
3	1.00	1.00	1.00	48	89	1.00	1.00	1.00	51
4	0.98	0.98	0.98	48	90	0.98	0.96	0.97	49
5	0.91	0.98	0.94	41	91	0.93	0.98	0.95	41
6	0.98	1.00	0.99	53	92	0.93	1.00	0.96	51
7	1.00	0.98	0.99	44	93	0.98	1.00	0.99	51
8	1.00	0.98	0.99	43	94	0.96	0.91	0.93	54
9	0.93	1.00	0.97	42	95	0.92	0.90	0.91	49
10	1.00	0.98	0.99	43	96	0.95	0.97	0.96	59
11	1.00	1.00	1.00	48	97	1.00	1.00	1.00	41
12	0.89	0.94	0.91	50	98	1.00	1.00	1.00	42
13	1.00	1.00	1.00	51	99	1.00	1.00	1.00	53
14	0.98	1.00	0.99	45	100	0.84	0.95	0.89	44
15	0.98	0.95	0.97	44	101	0.98	0.93	0.95	55
16	0.96	0.94	0.95	51	102	0.98	0.98	0.98	52
17	0.98	1.00	0.99	48	103	0.96	0.98	0.97	52
18	0.86	0.98	0.91	43	104	0.97	0.99	0.98	72
19	0.88	0.91	0.90	47	105	1.00	0.97	0.99	40
20	0.96	0.92	0.94	51	106	1.00	0.98	0.99	53
21	0.96	1.00	0.98	49	107	0.98	0.98	0.98	54
22	1.00	0.98	0.99	51	108	0.98	1.00	0.99	50
23	0.92	0.92	0.92	48	109	0.98	1.00	0.99	51
24	1.00	0.97	0.98	59	110	0.93	0.98	0.96	56
25	1.00	0.94	0.97	48	111	1.00	1.00	1.00	40
26	1.00	0.98	0.99	47	112	0.92	0.96	0.94	50
27	1.00	0.98	0.99	42	113	1.00	1.00	1.00	50
28	1.00	1.00	1.00	50	114	1.00	1.00	1.00	51
29	1.00	0.98	0.99	42	115	0.96	0.98	0.97	50
30	0.80	0.80	0.80	50	116	1.00	0.94	0.97	50
31	1.00	0.96	0.98	50	117	0.91	1.00	0.95	51
32	0.94	0.94	0.94	53	118	0.85	1.00	0.92	45
33	0.93	0.93	0.93	44	119	1.00	1.00	1.00	52
34	1.00	0.95	0.97	58	120	0.98	1.00	0.99	51
35	1.00	0.98	0.99	61	121	0.98	1.00	0.99	51
36	1.00	1.00	1.00	49	122	0.97	0.91	0.94	43
37	0.94	0.97	0.96	78	123	1.00	0.94	0.97	52
38	0.95	0.93	0.94	40	124	0.97	0.83	0.90	47
39	0.97	0.95	0.96	59	125	1.00	0.98	0.99	50
40	0.88	0.84	0.86	45	126	1.00	1.00	1.00	41
41	0.94	1.00	0.97	46	127	0.97	0.95	0.96	40
42	1.00	1.00	1.00	45	128	0.98	0.96	0.97	49
43	0.98	0.98	0.98	50	129	0.89	0.93	0.91	54
44	0.73	0.94	0.82	50	130	0.98	1.00	0.99	51
45	1.00	1.00	1.00	49	131	1.00	1.00	1.00	42
46	1.00	1.00	1.00	48	132	0.98	0.98	0.98	50
47	0.98	0.98	0.98	44	133	1.00	1.00	1.00	50
48	1.00	1.00	1.00	40	134	0.98	0.96	0.97	50
49	1.00	0.92	0.96	49	135	1.00	0.90	0.95	58
50	1.00	0.98	0.99	51	136	1.00	0.97	0.98	66
51	1.00	1.00	1.00	48	137	1.00	0.98	0.99	52
52	0.98	0.96	0.97	50	138	0.98	0.96	0.97	51
53	0.98	0.94	0.96	67	139	1.00	0.90	0.95	51
54	0.98	0.96	0.97	49	140	1.00	0.95	0.97	56
55	0.98	0.98	0.98	45	141	0.62	0.93	0.75	46
56	1.00	0.95	0.98	42	142	0.95	0.93	0.94	45
57	0.98	0.93	0.96	58	143	0.92	0.89	0.91	54
58	0.96	1.00	0.98	51	144	0.87	0.70	0.78	64

Class	Precision	Recall	F1-Score	Support	Class	Precision	Recall	F1-Score	Support
59	0.65	0.75	0.69	44	145	0.89	0.91	0.90	55
60	1.00	0.96	0.98	45	146	1.00	0.92	0.96	48
61	0.93	0.98	0.95	41	147	0.94	1.00	0.97	51
62	1.00	1.00	1.00	47	148	0.96	1.00	0.98	52
63	0.94	1.00	0.97	49	149	1.00	0.97	0.99	40
64	0.98	0.90	0.93	48	150	0.98	0.90	0.94	49
65	1.00	1.00	1.00	49	151	0.98	1.00	0.99	45
66	0.98	1.00	0.99	50	152	0.92	0.67	0.78	52
67	0.98	0.98	0.98	48	153	1.00	0.98	0.99	51
68	1.00	1.00	1.00	52	154	1.00	1.00	1.00	50
69	0.98	0.98	0.98	50	155	0.98	1.00	0.99	53
70	1.00	0.98	0.99	51	156	1.00	0.98	0.99	51
71	0.98	0.98	0.98	52	157	0.96	0.96	0.96	50
72	0.98	1.00	0.99	56	158	1.00	0.98	0.99	41
73	0.97	0.98	0.97	58	159	0.94	0.98	0.96	48
74	0.98	1.00	0.99	48	160	1.00	1.00	1.00	51
75	0.89	0.98	0.93	51	161	0.97	1.00	0.98	58
76	0.98	0.98	0.98	54	162	1.00	0.98	0.99	52
77	0.96	0.96	0.96	49	163	0.98	0.98	0.98	50
78	0.98	1.00	0.99	54	164	1.00	0.94	0.97	49
79	1.00	1.00	1.00	51	165	1.00	0.98	0.99	47
80	0.96	0.96	0.96	54	166	0.95	1.00	0.97	52
81	1.00	1.00	1.00	48	167	0.98	1.00	0.99	54
82	0.82	0.35	0.49	40	168	0.98	0.98	0.98	52
83	0.89	0.98	0.94	52	169	1.00	0.98	0.99	61
84	1.00	0.98	0.99	43	170	0.96	0.98	0.97	49
85	0.95	0.95	0.95	44	171	1.00	1.00	1.00	51
86	0.89	0.96	0.92	50	Avg.	0.96	0.96	0.96	49.82

C. Residual neural network

The residual neural network, popularly known as ResNet was developed from the phenomenon of the pyramidal cells of the cerebral cortex. ResNets utilize the method of jumping or skipping several layers. ResNets may skip two or three layers which most of the time hold ReLU along with batch normalization [21]. Two types of special networks are introduced while applying ResNets. HighwayNets use an additional weight matrix to learn the jump weights [22] and DenseNets are used for parallel jumps [23].

One urge for jumping the layers is to bypass the problem of disappearing gradients, by reusing ReLU and additional activations from a former layer till the neighboring layer absorbs its weights. While training, the weights adjust to silence the upstream layer and magnify the earlier jumped layer. Jumping efficiently clarifies the arrangement, employing several layers in the primary training steps. This advances learning by lessening the influence of disappearing gradients because of fewer layers to scatter through. The system then progressively reconstructs the jumped layers as it determines the feature expanse. At the edge of the training, while each layer is extended, it stays more alike and therefore learns quicker.

D. Our Proposed Architecture

In our study, we applied ResNet-50, a convolutional neural network which is 50 layers deep. First, the input image is passed into a convolutional layer. Then, after employing the batch normalization layer, the activation layer is applied. After that, there is a max-pooling layer. Then, a convolution block is added followed by two identity blocks. Again, one more convolution block is added followed by three identity blocks. After that, another convolution block is added followed by five identity blocks. Finally, the last convolution block is added followed by two identity blocks. Then, the average pool was applied followed by a fully connected layer of size 1024, a dropout layer of 50%, another fully connected layer of size 512, another dropout layer of 50%, and a SoftMax layer. Convolution block and identity block are made of some convolution layers, batch normalization layers, and activation layers. The architecture and expansion of the convolution block and identity block can be discovered in Figure-1.

IV. Experimental Analysis

In this section, first, preprocessing steps have been introduced. After that, the design of experiments has been discussed. Finally, result analysis is presented with proper evidence.

A. Preprocessing

Because of providing images to a convolutional neural network, a heavy preprocessing of the images was skipped as CNN is a powerful network that can detect useful features from raw images. However, two stages of preprocessing were performed on the Bengali handwritten compound characters

TABLE II: Comparison of performances among our proposed approach and previous approaches

Classifiers	No. of Classes	Overall Accuracy
SVM [9]	160	80.51%
SVM [11]	171	79.35%
SVM [16]	171	87.26%
DCNN [18]	171	90.33%
Proposed	171	96.15%

dataset. As the raw images of the dataset were different in resolution, we had to resize the image into a fixed size. We resized the images into 224x224x3 as most of the images were in this range. Moreover, we encoded the class levels in order to align with the nature of the output of our proposed network. Also, we considered different writing styles of a class as a single class for this research.

B. Design of Experiment

The model was trained for 25 epochs with a batch size of 24 as after that the validation loss became nearly constant for the rest of epochs. 'Adam' optimizer with the learning rate of 0.0001 was used to maximize the error function. Categorical cross-entropy function was employed for the loss or error function. For bypassing the overfitting, the dropout technique was practiced.

C. Result Analysis

The dataset did not require train-test sets splitting as the train and test set was provided separately from the data source. The test set had no influence over the training. After applying the preprocessing steps described earlier, our proposed architecture was applied to the processed dataset. Figure-2 illustrates the training accuracy of our proposed model. On the other hand, Figure-4 illustrates the training loss for our proposed model. Finally, the trained model was utilized to predict the overall accuracy of the test set provided separately with the CMATERdb Bengali compound isolated dataset and our model obtained an overall accuracy of 96.15% which outperformed all other studies. Table-1 illustrates class-wise precision, recall, f1-score and support. Table-2 showcases the difference in performance among our proposed model and performances achieved in previous researches. From Table-2, we concluded that our proposed architecture outperformed all previous researches by a significant margin in terms of CMATERdb Bengali compound isolated characters dataset.

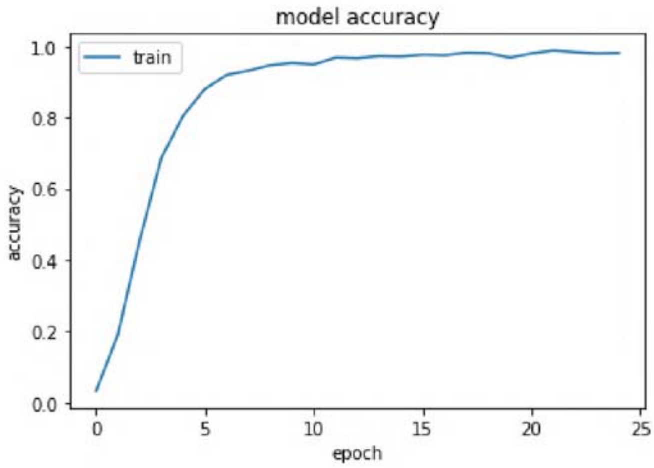

Fig. 2: Training accuracy of our proposed model

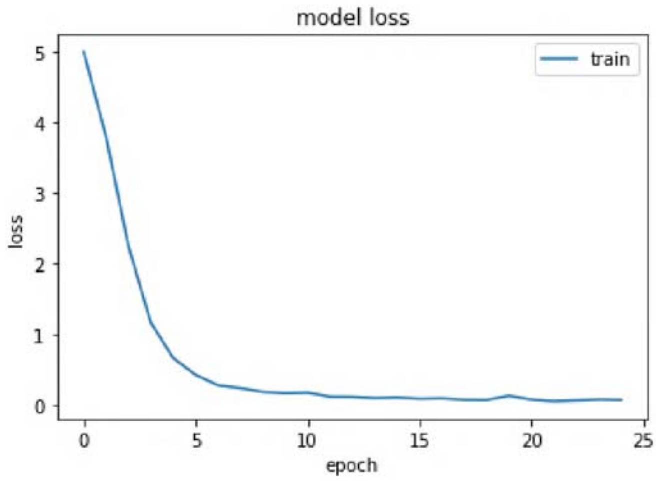

Fig. 3: Training loss of our proposed model

V. Conclusion

Previously, many researchers have contributed to the recognition of handwritten characters but a few pieces of research have been conducted on the Bengali handwritten dataset. Moreover, the recognition of the Bengali compound handwritten dataset is even rarer. Hence, in this research, we started with the CMATERdb Bengali compound isolated characters dataset. After applying some preprocessing, we applied a modified ResNet-50 architecture to train our model. After that, the trained model was utilized to predict the classes of the target test set which was exclusively provided by the data source. Finally, we measured the performances and showed that our result outperformed all previous results by a remarkable margin.

References

[1] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, and D. K. Basu, "Recognition of numeric postal codes from multi-script postal address blocks," in *International Conference on Pattern Recognition and Machine Intelligence*. Springer, 2009, pp. 381–386.

[2] S. N. Srihari, "Recognition of handwritten and machine-printed text for postal address interpretation," *Pattern recognition letters*, vol. 14, no. 4, pp. 291–302, 1993.

[3] A. F. Mollah, S. Basu, and M. Nasipuri, "Handheld device-based character recognition system for camera captured images," *International Journal of Image and Graphics*, vol. 13, no. 04, p. 1350016, 2013.

[4] A. Sufian, A. Ghosh, A. Naskar, F. Sultana, J. Sil, and M. H. Rahman, "Bdnet: bengali handwritten numeral digit recognition based on densely connected convolutional neural networks," *Journal of King Saud University-Computer and Information Sciences*, 2020.

[5] S. Ahmed, F. Tabsun, A. S. Reyadh, A. I. Shaafi, and F. M. Shah, "Bengali handwritten alphabet recognition using deep convolutional

neural network," in *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*. IEEE, 2019, pp. 1–4.

[6] S. Chatterjee, R. K. Dutta, D. Ganguly, K. Chatterjee, and S. Roy, "Bengali handwritten character classification using transfer learning on deep convolutional network," in *International Conference on Intelligent Human Computer Interaction*. Springer, 2019, pp. 138–148.

[7] S. Mondal and N. Mahfuz, "Convolutional neural networks based bengali handwritten character recognition," in *International Conference on Cyber Security and Computer Science*. Springer, 2020, pp. 718–729.

[8] U. Pal, T. Wakabayashi, and F. Kimura, "Handwritten bangla compound character recognition using gradient feature," in *10th international conference on information technology (ICIT 2007)*. IEEE, 2007, pp. 208–213.

[9] N. Das, B. Das, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "Handwritten bangla basic and compound character recognition using mlp and svm classifier," *arXiv preprint arXiv:1002.4040*, 2010.

[10] N. Das, K. Acharya, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "A novel ga-svm based multistage approach for recognition of handwritten bangla compound characters," in *Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012*. Springer, 2012, pp. 145–152.

[11] ——, "A benchmark image database of isolated bangla handwritten compound characters," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 17, no. 4, pp. 413–431, 2014.

[12] R. Ghosh, C. Vamshi, and P. Kumar, "Rnn based online handwritten word recognition in devanagari and bengali scripts using horizontal zoning," *Pattern Recognition*, vol. 92, pp. 203–218, 2019.

[13] M. M. Hoque *et al.*, "Developing a fuzzy feature-based online bengali handwritten word recognition system," in *International Joint Conference on Computational Intelligence*. Springer, 2019, pp. 577–590.

[14] M. S. Rahman *et al.*, "Towards optimal convolutional neural network parameters for bengali handwritten numerals recognition," in *2016 19th International Conference on Computer and Information Technology (ICCIT)*. IEEE, 2016, pp. 431–436.

[15] S. Bag, P. Bhowmick, and G. Harit, "Recognition of bengali handwritten characters using skeletal convexity and dynamic programming," in *2011 Second International Conference on Emerging Applications of Information Technology*. IEEE, 2011, pp. 265–268.

[16] N. Das, R. Sarkar, S. Basu, P. K. Saha, M. Kundu, and M. Nasipuri, "Handwritten bangla character recognition using a soft computing paradigm embedded in two pass approach," *Pattern Recognition*, vol. 48, no. 6, pp. 2054–2071, 2015.

[17] R. Sarkhel, N. Das, A. K. Saha, and M. Nasipuri, "A multi-objective approach towards cost effective isolated handwritten bangla character and digit recognition," *Pattern Recognition*, vol. 58, pp. 172–189, 2016.

[18] S. Roy, N. Das, M. Kundu, and M. Nasipuri, "Handwritten isolated bangla compound character recognition: A new benchmark using a novel deep learning approach," *Pattern Recognition Letters*, vol. 90, pp. 15–21, 2017.

[19] J. West, D. Ventura, and S. Warnick, "Spring research presentation: A theoretical foundation for inductive transfer," *Brigham Young University, College of Physical and Mathematical Sciences*, vol. 1, no. 08, 2007.

[20] T. George Karimpanal and R. Bouffanais, "Self-organizing maps for storage and transfer of knowledge in reinforcement learning," *Adaptive Behavior*, vol. 27, no. 2, pp. 111–126, 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[22] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," *arXiv preprint arXiv:1505.00387*, 2015.

[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.