Bengali Handwritten Compound Characters Recognition by Utilizing Transfer Learning

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

***Abstract*—Being the fifth most spoken language of the world, Bengali is spoken by 265 million people globally. Therefore, the recognition of handwritten Bengali characters has been an area of interest for a decade now. Previously, researchers have contributed to the recognition of Bengali numerals and basic characters e.g. vowels and consonants. However, apart from basic Bengali characters, Bengali has a wide range of compound characters as well. Not many remarkable contributions have been yet contributed to this domain. With a focus on contributing to this field, in this research, we started with 171 Bengali compound characters. We applied a modified DenseNet-201 architecture for the recognition of the compound characters. After result analysis, we achieved an overall accuracy of 96.89% which outperformed all previous studies by a remarkable margin. Hence, we concluded that our design is more capable of accurate recognition of the Bengali compound characters.**

***Index Terms*—Bengali Handwritten Characters Recognition, Compound Characters, Transfer Learning, Modified DenseNet-201**

I. Introduction

Handwritten character identification, an essential element of pattern identification and artificial intelligence, has been crucial in numerous divisions like commercial OCRs, understanding of bank checks, postal cards, etc. [1], [2]. In practically all these applications, offline characters are identified which is typically acknowledged as optical character recognition (OCR). Beforehand, investigations have been conveyed on the handwritten characters but due to a shortage of good quality cameras for obtaining photographs, the results were not competent. Lack of benchmark databases was also a dilemma for the identification of the handwritten characters. The improvement in cameras and benchmark datasets has expanded the expanse of investigation in the domain of handwritten character identification in various applications like recognition of medical prescripts and business cards [3].

The area of the handwritten character identification dilemma is extremely reliant on the script. For example, different language has different alphabets or characters, hence, the script-specific characteristics contribute to the improvement of the identification. Being one of the most expressed languages worldwide, Bengali handwritten character identification has been a region of concern for the researchers. But like other popular languages, Bengali not only has primary vowel and consonant characters, it also has a huge number of compound characters. The Bengali language has approximately 400 characters in total which makes the Bengali handwritten character identification more complex. Very few researches have been directed in the field of Bengali handwritten character identification compared to other popular languages [4], [5], [6], [7]. Furthermore, most of the researches were centered on the identification of the primary Bengali handwritten characters. A very shorter quantity of researches has been conveyed on the identification of Bengali compound handwritten characters which incorporate approximately 85% of the total characters of the Bengali language [1], [8]. More about earlier investigations have been discussed in the literature review segment.

In this study, we began with the CMATERdb Bengali compound isolated characters dataset. Firstly, a few preprocessing rounds were directed on the dataset. Next, we implemented proposed modified DenseNet-201 paradigm for training. After that, the performance was measured in terms of class-specific and overall accuracy. Our strategy produced an overall accuracy of 96.89% which outperformed all previous studies by a noteworthy margin.

II. Literature Review

Although contributions in recognition of Bengali compound characters have not received remarkable contributions, still many investigations and studies have been conducted on the CMATERdb dataset to recognize the compound characters. In 2014, a convex hull and quadtree-based feature set were used to achieve an overall accuracy of 79.35% [9]. In 2016, an overall accuracy of 86.65% was achieved with the assistance of the Support Vector Machine classifier [10]. A research conducted in 2017 produced an accuracy of 90.33% by utilizing the deep convolutional neural network [11]. Meanwhile, another research reported 93.68% accuracy with the assistance of a deep convolution network also [12]. In 2018, a study of hybrid-hog and convolutional neural networks achieved an overall accuracy of 92.57% and 92.05% respectively [13]. DenseNet was utilized to achieve 81.83% accuracy in 2019 [14]. A5 Model was applied to achieve 93.90% accuracy in one research of 2019 also [15]. Another research suggested an overall accuracy of 95.5% where deep convolution neural network was utilized [16]. The highest achieved accuracy for the dataset was observed in the most recent work of 2019 which achieved 95.70% accuracy [17]. However, we achieved an overall accuracy of 96.89% which outperformed all the previous studies.

978-0-7381-1102-5/20/$31.00

TABLE I: Class-wise Accuracy (Acc.), Precision (Pre.) and Recall (Rec.) illustration

Class	**Acc.**	**Pre.**	**Rec.**	**Class**	**Acc.**	**Pre.**	**Rec.**	**Class**	**Acc.**	**Pre.**	**Rec.**
1	1.00	1.00	1.00	58	0.96	1.00	0.96	115	1.00	0.98	1.00
2	0.96	0.96	0.96	59	0.70	0.79	0.70	116	0.96	1.00	0.96
3	1.00	1.00	1.00	60	1.00	1.00	1.00	117	1.00	0.98	1.00
4	1.00	1.00	1.00	61	1.00	0.95	1.00	118	0.97	1.00	0.98
5	0.97	0.98	0.98	62	0.97	1.00	0.98	119	1.00	1.00	1.00
6	1.00	1.00	1.00	63	0.97	0.98	0.98	120	1.00	1.00	1.00
7	1.00	1.00	1.00	64	0.91	1.00	0.92	121	1.00	1.00	1.00
8	1.00	1.00	1.00	65	1.00	1.00	1.00	122	0.90	1.00	0.91
9	1.00	1.00	1.00	66	0.98	1.00	0.98	123	1.00	0.95	1.00
10	1.00	1.00	1.00	67	0.97	0.96	0.98	124	0.89	0.95	0.89
11	1.00	1.00	1.00	68	1.00	0.98	1.00	125	0.94	1.00	0.94
12	0.98	0.91	0.98	69	0.98	0.94	0.98	126	1.00	0.98	1.00
13	1.00	0.96	1.00	70	0.98	0.98	0.98	127	0.92	1.00	0.93
14	1.00	1.00	1.00	71	1.00	1.00	1.00	128	1.00	1.00	1.00
15	1.00	1.00	1.00	72	1.00	1.00	1.00	129	0.92	0.96	0.93
16	0.90	0.98	0.90	73	1.00	1.00	1.00	130	1.00	0.93	1.00
17	1.00	1.00	1.00	74	1.00	1.00	1.00	131	1.00	1.00	1.00
18	0.90	0.95	0.91	75	1.00	0.98	1.00	132	1.00	0.98	1.00
19	0.93	0.86	0.94	76	1.00	1.00	1.00	133	1.00	1.00	1.00
20	0.78	1.00	0.78	77	1.00	0.96	1.00	134	1.00	0.98	1.00
21	0.95	1.00	0.96	78	1.00	0.98	1.00	135	0.96	1.00	0.97
22	0.98	1.00	0.98	79	1.00	1.00	1.00	136	1.00	0.94	1.00
23	0.91	0.96	0.92	80	1.00	1.00	1.00	137	0.98	1.00	0.98
24	0.98	1.00	0.98	81	0.97	1.00	0.98	138	1.00	1.00	1.00
25	0.95	0.90	0.96	82	0.75	0.51	0.75	139	0.94	0.91	0.94
26	1.00	0.98	1.00	83	0.98	0.98	0.98	140	0.98	0.93	0.98
27	0.97	0.98	0.98	84	0.97	1.00	0.98	141	0.39	0.62	0.39
28	1.00	0.98	1.00	85	1.00	0.98	1.00	142	0.97	0.98	0.98
29	0.97	1.00	0.98	86	0.88	0.94	0.88	143	1.00	0.96	1.00
30	0.78	0.89	0.78	87	1.00	0.93	1.00	144	0.82	0.82	0.83
31	1.00	0.98	1.00	88	0.90	0.79	0.90	145	0.94	0.98	0.95
32	0.96	0.89	0.96	89	0.98	1.00	0.98	146	0.97	1.00	0.98
33	1.00	0.96	1.00	90	0.97	1.00	0.98	147	0.98	1.00	0.98
34	0.98	0.97	0.98	91	0.95	0.95	0.95	148	1.00	0.98	1.00
35	0.96	0.98	0.97	92	0.98	0.94	0.98	149	0.90	0.97	0.90
36	1.00	1.00	1.00	93	1.00	1.00	1.00	150	0.93	0.92	0.94
37	0.98	0.96	0.99	94	0.96	1.00	0.96	151	0.97	0.98	0.98
38	0.97	0.95	0.97	95	0.97	0.91	0.98	152	0.84	0.88	0.85
39	0.98	0.98	0.98	96	0.98	0.98	0.98	153	1.00	1.00	1.00
40	0.86	0.93	0.87	97	0.97	1.00	0.98	154	1.00	1.00	1.00
41	1.00	0.98	1.00	98	0.97	1.00	0.98	155	0.98	0.98	0.98
42	1.00	1.00	1.00	99	1.00	1.00	1.00	156	1.00	1.00	1.00
43	0.98	0.96	0.98	100	0.93	0.89	0.93	157	0.98	0.98	0.98
44	0.88	0.86	0.88	101	1.00	0.98	1.00	158	0.97	0.98	0.98
45	0.93	1.00	0.94	102	0.98	1.00	0.98	159	0.95	0.96	0.96
46	1.00	1.00	1.00	103	1.00	1.00	1.00	160	0.98	0.98	0.98
47	0.95	1.00	0.95	104	1.00	0.99	1.00	161	0.96	0.97	0.97
48	1.00	0.95	1.00	105	0.87	1.00	0.88	162	1.00	0.98	1.00
49	0.97	0.96	0.98	106	1.00	0.96	1.00	163	1.00	0.98	1.00
50	0.98	1.00	0.98	107	0.98	0.98	0.98	164	0.95	1.00	0.96
51	0.95	1.00	0.96	108	0.98	1.00	0.98	165	1.00	1.00	1.00
52	0.96	0.98	0.96	109	1.00	0.96	1.00	166	1.00	0.98	1.00
53	1.00	1.00	1.00	110	0.96	0.98	0.96	167	0.98	1.00	0.98
54	0.97	0.98	0.98	111	1.00	1.00	1.00	168	0.98	1.00	0.98
55	1.00	0.96	1.00	112	0.98	0.94	0.98	169	1.00	1.00	1.00
56	1.00	1.00	1.00	113	0.94	0.98	0.94	170	0.93	0.82	0.94
57	1.00	0.98	1.00	114	1.00	1.00	1.00	171	1.00	1.00	1.00

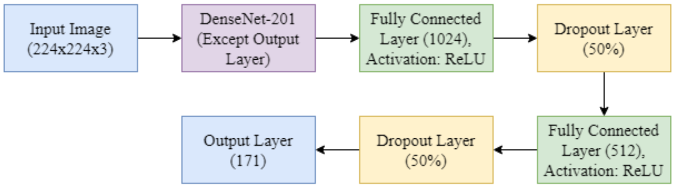

Fig. 1: Proposed modified DenseNet-201 design

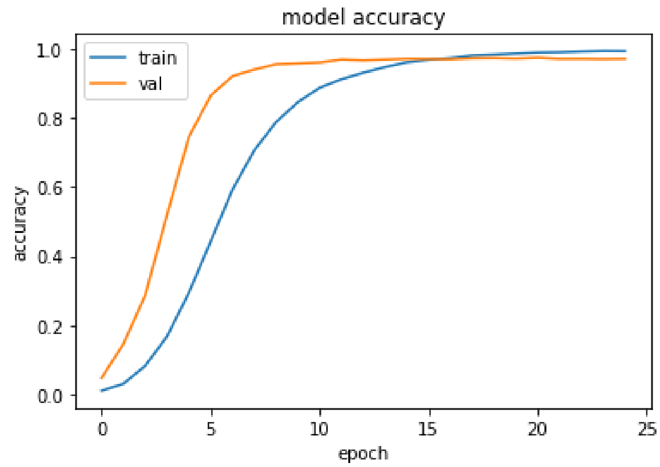

Fig. 2: Training accuracy of our proposed model

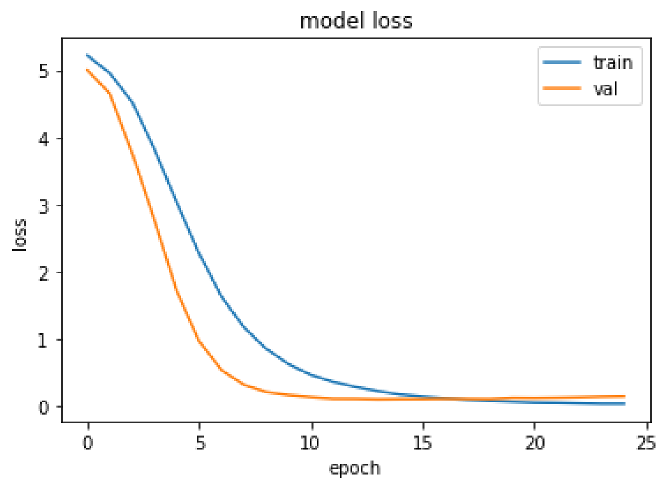

Fig. 3: Training loss of our proposed model

III. Materials and Methods

In this segment, first, the dataset information has been presented. After that, a discussion on transfer learning and modified DenseNet-201 has been reported.

A. Dataset Description

In this investigation, we examined the openly available CMATERdb Bengali compound isolated dataset holding a total of 171 classes [1]. There are a total of 34,439 photographs in the train set and 8,520 photographs in the test set. The dataset is imbalanced and the pictures are not fixed-sized which makes the identification more complex. Furthermore, for some of the Bengali compound characters, there are several writing fashions. Placing the same class label for two separate writing fashions turns the identification into a more complex one.

B. Transfer Learning

Transfer learning focuses on gathering knowledge acquired while determining one barrier and performing it to another but similar perplexity [18]. For example, the knowledge gained while learning to distinguish cars could be employed for the identification of trucks. This area of research bestows an exceptional connection to the permanent archives of cerebral investigation on the transfer of learning, though certain relations among the two fields are incompetent.

C. Modified DensNet-201 Architecture

The widespread foundation of Convolutional Neural Network (CNN) can be located in [19]. Nevertheless, in conventional CNN, all layers are constantly correlated which makes the network difficult to stretch wider and deeper, as it may evolve beyond difficulties of either collapsing or gradient missing conditions. After that, ResNet proposed an approach to engaging the shortcut attachment by jumping at least two layers. Then, DenseNet additionally developed the model by concatenating all the characteristic graphs sequentially instead of summation of the output characteristic charts from all former layers. In this study, we've introduced a modified DenseNet-201 architecture which is represented in Figure-1. After the DenseNet-201 basic layers, we implemented a fully connected layer of size 1024 supported by a dropout layer of 50%. Then, another fully connected layer of size 512 was joined supported by another dropout layer of 50%. Finally, an output layer is attached having a size of 171. While training by the design, no layer was held frozen.

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

Classification Methodology	Overall Accuracy
CH+QTLR [9]	79.35%
SVM [10]	86.65%
Deep CNN [11]	90.33%
Deep CNN [12]	93.68%
Hybrid-Hog [13]	92.57%
DenseNet [14]	81.83%
A5 Model [15]	93.90%
Deep CNN [16]	95.50%
CNN [17]	95.70%
Proposed Modified DenseNet-201	**96.89%**

IV. Experimental Analysis

In this segment, firstly, the preprocessing steps have been discussed. Next, the design of operations has been presented. Finally, result analysis is performed with proper evidence.

A. Preprocessing

Because of providing images to a convolutional neural network, a massive preprocessing of the images was jumped as CNN is a compelling network that can recognize relevant characteristics from raw photos. Nevertheless, some preprocessing levels were expected. The input pictures were in various shapes; therefore, they were reshaped to 224x224x3.

B. Experimental Settings

The model was trained for 25 epochs with a batch size of 24 as after that the validation loss grew almost constant for the rest of the epochs. 'Adam' optimizer [20] with the learning rate of 0.000008 was employed to maximize the error function. Categorical cross-entropy function was employed for the loss or error function. For avoiding overfitting, the dropout technique was functioned.

C. Result Analysis

The train data and test data were provided separately in the original dataset. At first, the train data was divided into the train set, validation set. 88% of the train data was stored in the train set and 12% of the data were stored in the validation set. Then, the proposed modified DenseNet-201 structure was employed to the train set. Figure-2 represents the training accuracy and validation accuracy of our proposed architecture. On the other hand, Figure-3 represents the training loss and validation loss of our proposed architecture. TABLE I represents the class-wise accuracy, precision and recall of individual classes of this study. Our proposed design produced an overall accuracy of 96.89% for the Bengali Compound Characters dataset. TABLE II illustrates the comparison between our proposed work and distinguished former investigations. From Table-2 it can be noticed that our proposed design outperformed all the previous methods by a distinguished boundary, therefore, our model is proficient in identifying the recognized characters more precisely.

V. Conclusion

Despite being a popular language Bengali character recognition has not gained many contributions so far. Previously, studies have been conducted for the recognition of Bengali numerals and basic characters e.g. consonants and vowels. But there are a massive number of compound characters in Bengali literature that covers almost 85% of all the characters. To recognize these characters we started with a dataset of 171 distinct compound characters of Bengali literature. We applied a modified DenseNet-201 architecture and achieved an overall accuracy of 96.89% which outperformed all the previous studies by a noteworthy boundary.

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, K. Acharya, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "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.

[10] 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.

[11] 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.

[12] A. Ashiquzzaman, A. K. Tushar, S. Dutta, and F. Mohsin, "An efficient method for improving classification accuracy of handwritten bangla compound characters using dcnn with dropout and elu," in *2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 2017, pp. 147–152.

[13] S. Sharif, N. Mohammed, S. Momen, and N. Mansoor, "Classification of bangla compound characters using a hog-cnn hybrid model," in *Proceedings of the International Conference on Computing and Communication Systems*. Springer, 2018, pp. 403–411.

[14] M. M. Hasan, M. M. Abir, M. Ibrahim, M. Sayem, and S. Abdullah, "Aibangla: A benchmark dataset for isolated bangla handwritten basic and compound character recognition," in *2019 International Conference on Bangla Speech and Language Processing (ICBSLP)*. IEEE, 2019, pp. 1–6.

[15] P. Keserwani, T. Ali, and P. P. Roy, "A two phase trained convolutional neural network for handwritten bangla compound character recognition," in *2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR)*. IEEE, 2017, pp. 1–6.

[16] A. Fardous and S. Afroge, "Handwritten isolated bangla compound character recognition," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2019, pp. 1–5.

[17] P. Keserwani, T. Ali, and P. P. Roy, "Handwritten bangla character and numeral recognition using convolutional neural network for low-memory gpu," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 12, pp. 3485–3497, 2019.

[18] 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.

[19] Y.-D. Zhang, C. Pan, X. Chen, and F. Wang, "Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling," *Journal of computational science*, vol. 27, pp. 57–68, 2018.

[20] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. IEEE, 2018, pp. 1–2.