2018 4th International Conference for Convergence in Technology (I2CT)
SDMIT Ujire, Mangalore, India. Oct 27-28, 2018

Handwritten Devanagari Character Recognition using Convolutional Neural Network

Aarati Mohite
Department of Electronics and Telecommunication
College of Engineering
Pune, India
mohitear17.extc@coep.ac.in

Sushama Shelke
Department of Electronics and Telecommunication
Dr. D. Y. Patil Institute of Engineering, Management & Research
Pune, India
shelke.sushama@gmail.com

***Abstract*—In any pattern recognition task, feature extraction and classification stages handle the responsibility of recognizing the patterns accurately. Deep learning relives the task of feature extraction and extracts them automatically, thus reducing the programmer's burden. Deep learning is replacing other pattern recognition techniques recently. In applications like character recognition, which involves large amount of database and variability in the data, deep learning is the right choice to handle the challenges involved. In this paper, a system for handwritten Devanagari character recognition using Convolutional Neural Network is discussed. The recognition accuracy obtained is 97.05% for Devanagari characters and 100% for Devanagari numerals.**

Keywords—handwritten Devanagari characters, pre-processing, classification, Convolutional neural network, Alexnet

I. INTRODUCTION

In the recent time, Deep learning field has undergone a radical change. Drastic changes in Artificial Neural Network, which is inspired from biological neuron network, has resulted in improved technology for image classification. Since image classification is the biggest challenge in pattern recognition applications, this has been resolved to a greater extent using Convolution neural network. Image classification in pattern recognition applications is difficult due to challenges such as viewpoint variation, occlusion, deformation and interclass variation of objects under classification.

In, Optical character recognition (OCR), where character images are converted into text, deep learning is becoming a suitable choice these days. The deep learning algorithms recognize the characters by designing its feature extraction technique. The text output converted by the OCR can be useful for storing and editing the script in applications like bank automation, postal automation, office automation etc.

Although, there are OCR systems available for English language, work for OCR systems in Indian languages is still in progress [1]. This is due to the complexity in these scripts. Devanagari character set has complex shape with various features like strokes, line segments, curves and holes.It has 36 consonants and 16 vowels. There are no upper-case and lower-case characterslike those found in English script. The script is written from left to right. Each character has a horizontal line at the top called as the header line. This line also joins the characters in a Devanagari word. The vowels are joined to consonants in various ways forming new shapes. Further, two or more consonants join together forming a complex structure called as the conjunct characters. A typical handwritten Devanagari character recognition system consists of following stages, namely, pre-processing, feature extraction and classification. Here we have assumed that the characters are already segmented. There are a large number of feature extraction and classification algorithms found during the literature survey.

Research for printed Devanagari character recognition began in 1970 [2]. Initially, the characters in Devanagari script were classified using the properties like the loops, lines, strokes, end points and their location in the characters [3, 4]. Then, when attempts were made for recognition of Devanagari handwritten numerals [5] and characters, these features failed to classify the characters accurately and there was a need of sophisticated feature extractor and classifier. As the work progressed to handwritten character recognition in Devanagari, attempts were made to find the features that could represent the characters efficiently [6-8]. Different researchers implemented different feature extraction and classification techniques for Devanagari character recognition. Researchers also tried a combination of several feature extraction and classification techniques in order to improve the performance [9]. Few researchers extracted multiple features and tested with one classifier and vice versa [10, 12]. The features extracted were regional features, boundary features and their combination, while classification techniques were simple as well as complex like ANN, SVM etc. Since there is no thumb rule for the choice of the feature extraction and classification, more advanced techniques like deep learning evolved recently; this eliminated the need for deciding the feature extraction techniques. Deep learning is a layered convolutional neural network (CNN) [13] that extracts the features automatically and classifies them.

Apart from character recognition application, CNNs are used in automatic facial expression recognition [14]. CNN is designed using two pooling layers, two fully connected layers and single Softmax regression layer. The feature extraction framework (Caffe) is used to extract features from faces. CNN is also used in prediction of temperature distribution in data centres. In this model CNN is used for learning local structure adaptively. This model can predict 20-minute future temperature distributions over 48 locations. Other application is classification of heart sound recordings [15], where CNN is used to predict whether heart sound recording sounds normal

978-1-5386-5232-9/18/$31.00 ©2018 IEEE

or not. Heart sound recordings are filtered using windowed-sinc hamming filter algorithm to remove noise. Filtered and segmented recordings are used to train CNN which further extracts features and gives classification function. Backpropagation with SGD is used to train CNN. Trained CNN assigns relative probability to each sound recording segment; further by combining these relative probabilities we can predict whether heart sound recording is normal or not. CNNs can also segment the tumor in a brain image. Multi-modality images are usedto train CNN network. This helps in early diagnosis of brain tumor. In this paper, CNN is used for recognition of unconstrained handwritten Devanagari characters.

The next part of the paper is organized as follows: Section II describes the deep learning. Section III presents the proposed environment. Section IV, presents the results obtained and finally, Section V discusses the conclusion.

II. Deep Learning

Deep learning is a part of machine learning algorithms which excels in the recognition of patterns. It is implemented using two or more layers of Artificial Neural Network where each layer extracts one or more features of the objects in images. These features may be edges, lines, blob, faces, and semantics. Each neural network consists of neurons, based on human brain analogy, where each of it takes input, performs operation and its result is followed to next neuron. There are four major architectures of deep network viz, Unsupervised Pretrained Networks (UPNs), Convolutional Neural Networks(CNNs), Recurrent Neural Networks, Recursive Neural Networks. We have used CNN in this paper.

Artificial Neural Network is computational system whose design is inspired from functioning of biological nervous system operation.

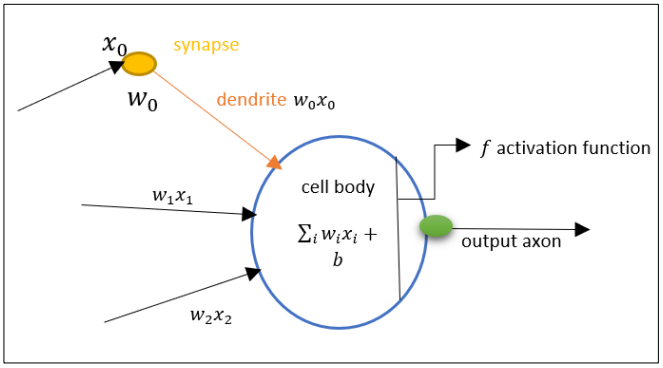

Fig. 1. Artificial neuron

In Fig. 1, neuron structure of ANN is shown. Function of this neuron is same like that of a biological neuron except here synapses are linear and have simple weights unlike ANN. Convolution neural network is very similar to Neural networks. It consists of learnable weights and biases. It takes raw image at input and converts it into class score function using CNN. CNN Consists of following layers:

A. *Convolutional Layer*

In Fig. 2, Convolutional layer of CNN is shown. It works like score function. Convolutional filter slides over special location of image and results in dot product between their weights and a small region. Summary of convolutional layer functioning can be described as below:

1. Accepts a volume of size W x H x D, where:
   - W width of image
   - H height of image
   - D depth of image

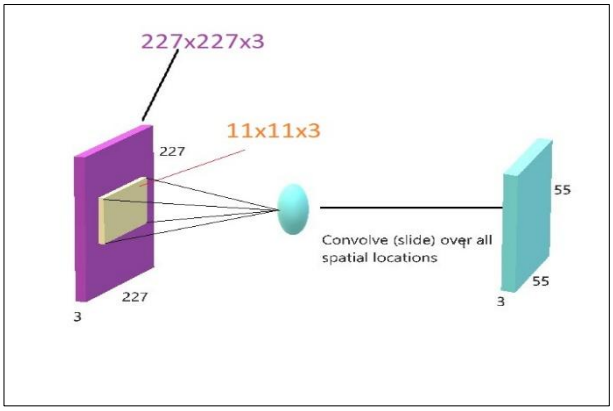

Fig. 2. Convolutional Layer

2. Requires four parameters:
   - Number of filters K
   - their spatial extent F
   - the stride S
   - the amount of zero padding P
3. Produces volume of size W_2x H_2x D_2, where:
   - W_2=(W_1-F+2P)/S + 1
   - H_2=(H_1-F+2P)/S + 1
   - D_2=K

B. *Rectification Linear Unit (RELU)*

It is an activation function. It applies element-wise and takes 0 at threshold value and hence its function is max (0, x).

C. *Pooling Layer*

Pooling layer is nothing but down sampler. It has been used in CNN to reduce image size and hence representation becomes smaller and more manageable. The aspect behind pooling layer is that it minimizes computation and prevents over-fitting. Summary of working of Pooling layer can be given as below:

Accepts a volume of size W_1X H_1XD_1

1. Requires three parameters:
   - their spatial extent F,
   - the stride S

2. Produces volume of size W_2X H_2XD_2 where:
   - $W_2 = (W_1 - F)$/S +1
   - $H_2 = (H_1 - F)$/S +1
   - $D_2 = D_1$

3. It introduces zero parameters since it computes fixed function of the input

D. *Fully Connected Layer:*

It consists of neurons which connects whole input volume as in ordinary neural network.

Let's consider some case studies of architecture of CNN. In this section we will consider some CNN architectures which has created drastic change in Deep learning [16].

TABLE I. CASE STUDIES OF CNN ARCHITECTURES

Archit-ecture	Comparison Factors					
	Year	Layer	Filter Size	Images used	Days and GPU Usage	Error Rate %
Alexnet	2012	8	11x11	15 million over 22000 categories	5-6 days /GTX 580 GPU	15.4
ZF Net	2013	8	7x7	1.3 million	12 days /GTX 580 GPU	11.2
VGG Net	2014	19/16	3x3	1.3 million	2 Week/ Nvidia Titan black	7.3
Google Net	2015	22	3x3/5 x5	1.2 million	Within week	6.7
Res Net	2015	152	3x3	1.28 million	8 GPU	3.6

III. PROPOSED ENVIRONMENT

The proposed environment for recognizing Devnagari characters is show in Fig. 3. In this system, the Devanagari character images undergo pre-processing at first. We have used transfer learning for Convolution neural network architecture, where existing architecture Alexnet(developed in 2012) [17] is used for our database.

A. *Devanagari Characters*

Handwritten Devanagari character and numeral database [18] developed by ISI Kolkata is used in this system. Few character samples are shown in Fig. 4.

B. *Pre-processing*

The characters are passed through various pre-processing steps before applying it to Deep learning (CNN). As we have used Alexnet it uses 227x227x3 image size. Hence, at first, the character image is resized to 227x227x3. The image dataset is divided into training and testing Datasets. The characters are now ready for feature extraction and for applying to CNN.

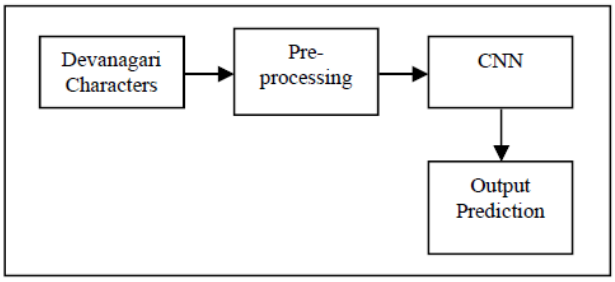

Fig. 3. System design.

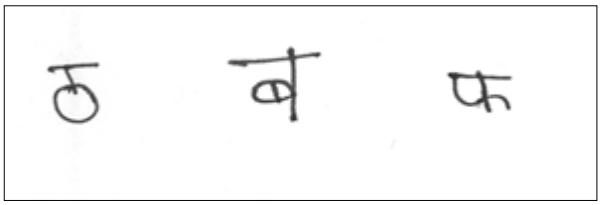

Fig. 4. Test characters of Devnagari script

C. *Feature Extraction using CNN*

Alexnet CNN architecture is used for transfer learning. Resized image dataset has been applied to Alexnet architecture. It consists of 8 layers. Convolutional layer extracts basic features from input images followed by RELU and Max-pooling layer. All neurons at the end are connected in fully connected layer. In this deeper layer, higher level features which are extracted using lower level features of earlier layers. Multiclass support vector machine is fitted using extracted features and predicted labels. Image is read and their features are extracted using Alexnet which are then given to classifier to predict the output class.

D. *Output Prediction*

Images from test dataset are used for prediction at output. Features are extracted using CNN from test images and are applied it to the classifier which predicts the label.

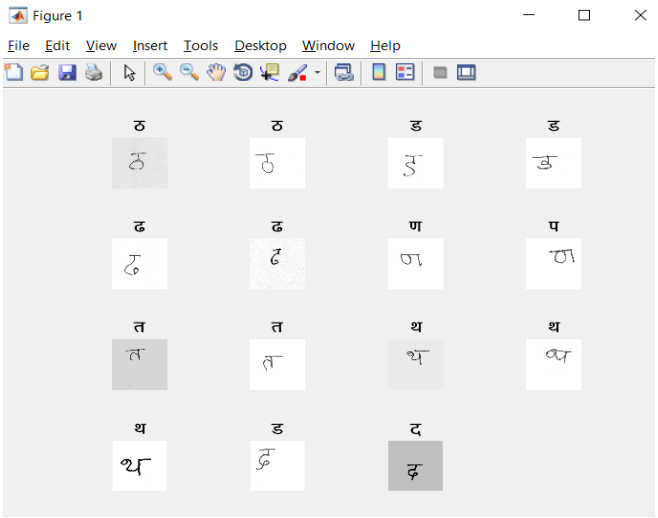

Fig. 5. Result of recognition in Matlab using Alexnet

IV. Results And Discussion

Around 40,000 Devanagari character images are used in database. The images are converted to 227x227x3 image size, 80% of which are used to train the Pre-trained architecture of Alexnet and 20% are used for testing. Fig.5 shows the result of sample test images. The images are at first pre-processed and the applied to CNN. This classifier eliminates the need for application of feature extraction technique. Accuracy of 91.23% is obtained in this transfer learning using Pre-trained architecture Alexnet, which is trained using NVIDIA GEFORCE GTX (RAM 4GB). After applying few more pre-processing techniques like image cropping the recognition accuracy improved to 97.05%. The image cropping algorithm eliminates the translation artifacts in the images. Further converting the images to 227 x 227, eliminates the effects of image scaling. Moreover, the rotation invariance is attained by the features extracted by the deep learning network. The numeral database of Devanagari was also tested on the proposed system. The recognition accuracy for this database was 100%.

V. Conclusion

This paper describes the use of transfer learning of pre-trained architecture like Alexnet, Googlenet, ResNet for the recognition of Handwritten Devnagari Characters script. Use of Pre-trained architecture saves time to develop CNN architecture from scratch. As the database increases, the architecture requires high performance GPU with high computational capability. Accuracy can be increased more samples are added in the database with high GPU computational capability for training images on architectures with more layers like Googlenet, ResNet.

References

[1] U. Pal and B. B. Chaudhuri, "Indian script character recognition: A survey," Pattern Recognit., vol. 37, pp. 1887–1899, 2004.

[2] R. M. K. Sinha and H. Mahabala, "Machine recognition of Devnagari script," IEEE Trans. Syst. Man Cybern., vol. 9, no. 8, pp. 435–441, Aug. 1979.

[3] V. Bansal and R. M. K. Sinha, "On How to Describe Shapes of Devanagari Characters and Use them for Recognition", Proceedings of 5th International Conference on Document Analysis and Recognition, 1999,pp. 410-413.

[4] V. Bansal, R.M.K. Sinha, "Partitioning and Searching Dictionary for Correction of Optically Read Devanagari Character Strings", Proceedings of 5th International conference on Document Analysis and Recognition,1999, pp. 653-656.

[5] P. M. Patil and T. R. Sontakke, "Rotation, Scale and Translation Invariant Handwritten Devanagari Numeral Character Recognition using General Fuzzy Neural Network", Pattern Recognition,vol. 40, no. 7, pp. 2110-2117, 2007.

[6] S. Kumar, "Performance comparison of features on Devanagari handprinted dataset," Int. J. Recent Trends in Engineering, vol. 1, no. 2, pp. 33–37, 2009.

[7] S. Arora, D. Bhattacharjee, M. Nasipuri, D. K. Basu, M. Kundu, and L. Malik, "Study of different features on handwritten Devanagari character," in Proc. 2nd Emerging Trends Eng. Technol., 2009, pp. 929–933.

[8] U. Pal, T. Wakabayashi, and F. Kimura, "Comparative study of Devanagari handwritten character recognition using different features and classifiers," in Proc. 10th Conf. Document Anal. Recognit., 2009, pp. 1111–1115.

[9] S. Arora, D. Bhattacharjee, M. Nasipuri, D. K. Basu, and M. Kundu, "Combining Multiple Feature ExtractionTechniques for Handwritten Devanagari Character Recognition",in Proc. IEEE Region Colloquium and the Third ICIIS, 2008, pp. 1-6.

[10] S. Shelke, and S.Apte, "A Novel Multi-feature Multi-classifier Scheme for Unconstrained Handwritten Devanagari Character Recognition", in Proc. 12th Int. Conf. Frontiers Handwrit. Recognit., 2010, pp. 215-219.

[11] S.Shelke, and S.Apte, "A Novel Multistage classification and Wavelet based Kernel Generation for Handwritten Marathi Compound Character Recognition", in Proc. Int. Conf. Communications and Signal Processing, 2011, pp. 193-197.

[12] Minhua Li, Chunheng Wang and Ruwei Dai, "Unconstrained Handwritten Character Recognition Based on WEDF and Multilayer Neural Network", in Proc. of the 7th World Congress on Intelligent Control and Automation, 2008, pp. 1143-1148.

[13] Andrej Karpathy and Justin Johnson, "CS231n: Convolutional Neural Networks for Visual Recognition", Stanford University online notes and lectures, Spring 2017 course.

[14] H. Mliki, N. Fourati, S. Smaoui and M. Hammami, "Automatic Facial Expression Recognition System",in Proc. International Conference on Computer Systems and Applications, 2013, pp. 1-4.

[15] G. D. Clifford C. Liu, B. Moody, D. Springer, I. Silva, Q. Li and R. G. Mark, "Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016", Proc. of Computing in Cardiology Conference, Vancouver, BC, 2016,pp.609-612.

[16] Adit Deshpande, "The 9 Deep Learning Papers You Need To Know about (Understanding CNNs Part 3)", https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html

[17] A Krizhevsky, I Sutskever and GE Hinton, "Imagenet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, pp. 1097-1105, 2012.

[18] U. Bhattacharya and B. B. Chaudhuri, "Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 3, pp. 444–457, Mar. 2009.