International Journal of Future Generation Communication and Networking
Vol.14, No. 1, (2021), pp. 4189 – 4198

Devanagari Character Recognition Using Convolutional Neural Network

Priyanka Deshpande[1] , Ankita Bondre[2], Vaishnavi Bhise[3], H.B.Ghorpade[4] ,S.S.Bhosale[5], C.G.Patil[6]

Department of E&TC Engineering, SKNCOE, SPPU, Pune, India

[1]*priya.desh1304@gmail.com*

[2]*bondreankita22@gmail.com*

[3]*vaishnavibhise04@gmail.com*

Abstract

In the last few years, we can see many articles have published on optical character recognition (OCR). Nowadays there are many commercial OCR systems present in the market, most of these systems work for Roman, Chinese, Japanese, and Arabic characters. In India, we have 12 different scripts but there is no sufficient work is done on Indian language character recognition. In this paper, we are focusing on the Devanagari script which is used in different languages like Hindi, Sanskrit, Nepali, Marathi, Sindhi, and Konkani. Recently deep learning is replacing other pattern recognition techniques. Deep learning is the right choice to handle the challenges are involved in applications like character recognition, which involve a large amount of the database and variability in the data, we are using the Convolutional Neural Network for the feature extraction and the classification followed by pre-processing. This approach has been proven effective and accurate in identifying characters in a digital format.

Keywords*--Optical Character Recognition, Image pre-processing, MATLAB, Classification, Convolutional Neural Networks.*

I. INTRODUCTION

Due to so many reasons like, writing style variations, a lack of ordering information of strokes, overlapping of wide strokes, pen-type and also the large similarities between the characters handwritten character recognition is a challenging task. Deep neural network requires less engineering by hand. By using different algorithms Neural network learns the features from the applied data on its own. As the amount of data and variations increases neural network's learning gets better and better. If we provide the samples with variations in data, it will learn those variations, otherwise it will misclassify the new samples. There are many types of neural networks present in the market like ANN,KNN,RNN but according to project requirements Convolutional neural network gives best results in image classification with good accuracy.A convolutional neural network is also used for document analysis purpose. For printed as well as handwritten characters and word segmentation recognition many researchers have used convolutional neural network. A Devanagari character recognition system consists of following stages, namely, pre-processing, feature extraction and classification. In this paper ResNet50 which is the pre-trained model of CNN for the classification of the characters is used.

II LITERATURE SURVEY

In[2] this work, they propose a technique to recognize handwritten Devanagari characters using deep convolutional neural networks (DCNN) which are one of the recent

ISSN: 2233-7853 IJFGCN

techniques adopted from the deep learning community. They experimented the ISIDCHAR database provided by (Information Sharing Index) ISI, Kolkata and V2DMDCHAR database with six different architectures of DCNN to evaluate the performance and also investigate the use of six recently developed adaptive gradient methods. They experimented with a handwritten Devanagari character database with six different DCNN network architectures as well as six different optimizers. The highest recognition accuracy 96.02% was obtained using NA-6 network architecture and RMSProp—an adaptive gradient method (optimizer). Further, we again trained DCNN layer-wise, which is also adopted by many researchers to enhance the recognition accuracy, using NA-6 network architecture and the RMSProp adaptive gradient method. Using DCNN layer-wise training model, our database obtained 98% recognition accuracy, which is the highest recognition accuracy of the database. In [3] they discussed extraction of features from handwritten compound characters using Zernike moment feature descriptor and proposes SVM and k-NN based classification system. The proposed classification system preprocess and normalize the 27000 handwritten character images into 30x30 pixels images and divides them into zones. The pre-classification produces three classes depending on presence or absence of vertical bar. Further Zernike moment feature extraction is performed on each zone. The overall recognition rate of proposed system using SVM and k-NN classifier is upto 98.37%, and 95.82% respectively. they presented a system for offline handwritten simple and compound character recognition for Marathi derived Devanagari script. Devnagari has given better result for compound character. The proposed system gives improved recognition rate of 0.37% than other handwritten character recognition system. The system has been evaluated on a huge amount of Handwritten Character Database i.e. 12000 basic and 15000 compound character dataset created in our laboratory. In [4] they propose an OCR for printed Hindi text in Devanagari script, using Artificial Neural Network (ANN), which improves its efficiency. They found that the input matrix of size 48x57 gives better results than other choices. The test set used in this experiment is of 77 characters of five different types of fonts. This can be tested for a greater number of fonts. The toughest phase in the experiment is getting a good set of characters for classification. This highlights the need for generation of a large ground-truthed set of characters of various resolutions so that more research can be performed for recognition of languages from Indian subcontinent. The other future enhancement that can be incorporated in the work is to use a dictionary of words to correct the output. Certainly this will improve the performance. Further speech synthesizer can be integrated with the OCR with the aim of making a system for reading aids to the blind. We can also implement the neural network method for classifying hand-written texts. In hand-written documents, the fragmentation of characters and the variation in shape of characters are considerably greater compared to printed documents. The higher levels can be used to provide clues for a hypothesization system, which learns from the text it recognizes. The implemented the work for the scripts with only Hindi characters. However, we can extend it to classify the document with characters of more than one script.

The English character recognition is easier than the Devanagari characters, that huge data set is collected which consists more than 45000 samples and efficiently classified with the CNN .superior results and robust data set handling is done with 95.7% accuracy in [5]. For evaluation MNIST,CIFAR-10 and CIFAR-100 data sets are used. The NIN and SPP-net are incorporated in a single model. Different architecture model are introduced which influence the performance of new model formed with 59.85% accuracy by [6] .By using more techniques the final model of [6]should be checked for real time applications. Spatial Pyramid Pooling in Deep Convolutional Networks (SPP-net) is a flexible solution to handle different sizes, aspect ratios. It is faster than R-CNN method. They suggested the solution to train the deep network

ISSN: 2233-7853 IJFGCN

[7]. In [8] they proposed Deeply Supervised nets (DSN) method that method minimizes the error in classification and it improves the directness and transparency of hidden layer learning process. By using techniques extended from stochastic gradient method they analyzed their algorithm in [8]. The results of [8] showing state-of-the-art performance on MNIST, CIFAR-10, CIFAR-100 and SVHN. In [9] they proposed a model which can adapt itself to detect the rigid objects with long or small local deformations. On the different datasets like PASCAL VOC 2007, VOC 2010, ImageNet it has achieved 41.7% ,39.7%, and 14.7% precision respectively .The way of proposing bounding boxes can be improved in terms of recall and speed in [9]. By using a combination of classical tools from computer vision and deep learning (bottom up region proposals and convolutional neural networks) results are obtained in [10]. They have given the simple and scalable object detection algorithm which gives 30% improvement over previous results of PASCAL VOC 2012 in [10]. Problems faced during feature extraction and the recognition of handwritten digits are focused by using the Black box approach and Support Vector Machine is used to enhance the ability with LeNet5 architecture [12]. They performed the experiment on MNIST database. In [13] they presented a review of the OCR work done on Indian language scripts. There, at first, they briefly discussed different methodologies applied in OCR development in international scenario and then different work done for Indian language scripts recognition. Finally, They discussed steps needed for better Indian script OCR development. They believe that their survey will strongly encourage activities of automatic document processing and OCR of Indian language scripts. For faster training of dataset non saturating neurons and very efficient GPU implementation of convolution operation is done in [14]. For reduction of the overfitting in fully connected layers dropout method is efficient as proved in [14]. Different types of neural networks are compared in [15].

III. DEVANAGARI CHARACTER DATASET

Devanagari Script - Devanagari is part of the Brahmic family of scripts of Nepal, India, Tibet, and South-East Asia .The script is used to write Hindi, Sanskrit, Nepali, Marathi, Sindhi, and Konkani. The Nepalese writing system adopted from Devanagari script consists of 12 vowels, 36 base forms of consonant, 10 numeral characters and some special characters. Referred dataset is an image database of Handwritten Devanagari characters. There are 36 classes of characters with 2000 examples each. It is split into training set(85%) and testing set(15%).

ISSN: 2233-7853 IJFGCN

Fig.1. Sample Dataset images

Dataset specifications are as follows :-

All data set images are of Gray Scale data type present in the .png format. Images in dataset are having 32 by 32 resolution. Characters are centered within 28 by 28 pixel, padding of 2 pixel is added on all four sides of actual character.

IV. PRAPOSED METHODOLOGY

The basic OCR system consists of 4 stages as shown in figure

A. Devanagari Characters

B. Pre-processing

C. CNN

D. Output Prediction

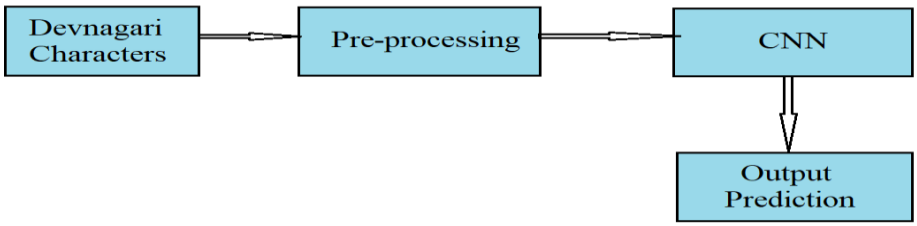

Figure :2.Block Diagram

A) INPUT

Database as shown below The input can be given directly using a digital pad or by using scanned copy of the text. In our case we have used standard Devanagari numeral

Figure : 3. Devanagari Data base

B) PRE PROCESSING

Image pre-processing is the term for operations on images at the lowest level of abstraction. These operations do not increase image information content but they decrease it if entropy is an information measure. The aim of pre-processing is an improvement of the image data that suppresses undesired distortions or enhances some

ISSN: 2233-7853 IJFGCN

image features relevant for further processing and analysis task. Here character image is resized to 224x224. As the input layer size of the ResNet-50 model is 224 x224x3 so we have to do the resizing operation on the scanned image.

C) CNN (CONVOLUTIONAL NEURAL NETWORK)

One of the classes of deep learning neural networks is CNN or the convolutional neural network (CNN). CNN is a machine learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and can differentiate one from the other. CNN works by extracting features from the images. Any CNN consists of the following:

1.The input layer is a grayscale image.
2.The Output layer is a binary or multi-class labels.
3. Hidden layers consisting of CNN are such as convolution layers, ReLU (rectified linear unit) layers, the pooling layers, and a fully connected Neural Network

ConvNets consist of multiple layers of overlapped tiling collections of small neurons to achieve a better representation of the original image. For image and video recognition ConvNets are widely used.
CNN Layers:-
1.Convolutional Layer
2. MaxPooling Layer
3.ReLU Layer
4. Fully Connected Layer
5. Softmax Layer

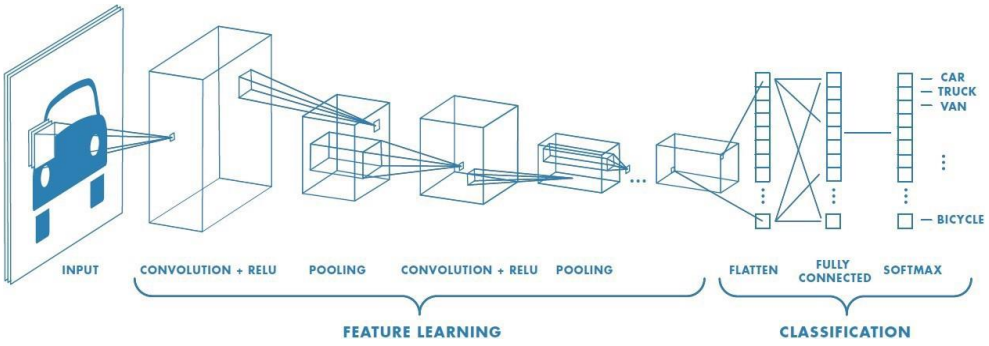

Figure: 4.CNN Architecture

1) Convolutional Layer- To extract high-level features from the image the convolution operation is important. while building the neural network, We can always add more than one convolution layer, where the first Convolution Layer is responsible for capturing gradients whereas the second layer captures the edges. According to the complexity of the image more. of layers are added.

2) MaxPooling Layer:- Max Pooling, which returns the maximum value from the portion of the image covered by the Pooling Kernel. This is mainly to reduce the computational complexity required to process the huge volume of data linked to an image.

ISSN: 2233-7853 IJFGCN

3) ReLU Layer:- A ReLU implements the function y = max(x,0), so the input and output sizes of this layer are the same. The network trains many times faster because of ReLU Layer.

4) Fully Connected Layer:- In the case of a fully connected layer, all the elements of all the features of the previous layer get used in the calculation of each element of each output feature. Figure 13 explains the fully connected layer L. Layer L-1 has two features, each of which is 2x2, i.e., has four elements. Layer L has two features, each having a single element.

5) Softmax Layer:- It is a function that turns a vector of K real values into a vector that sum to 1. For this reason, usually, a softmax function as the final layer of the neural network.Here CNN architecture is used for transfer learning. Resized image dataset has been applied to this CNN architecture. It consists of 8 layers. Basic features from input images are extracted by convolutional layer followed by RELU and Max-pooling layer. All neurons at the end are connected in the fully connected layer. Multiclass support vector machine is fitted using extracted features and predicted labels. Image is read and their features are extracted and 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.

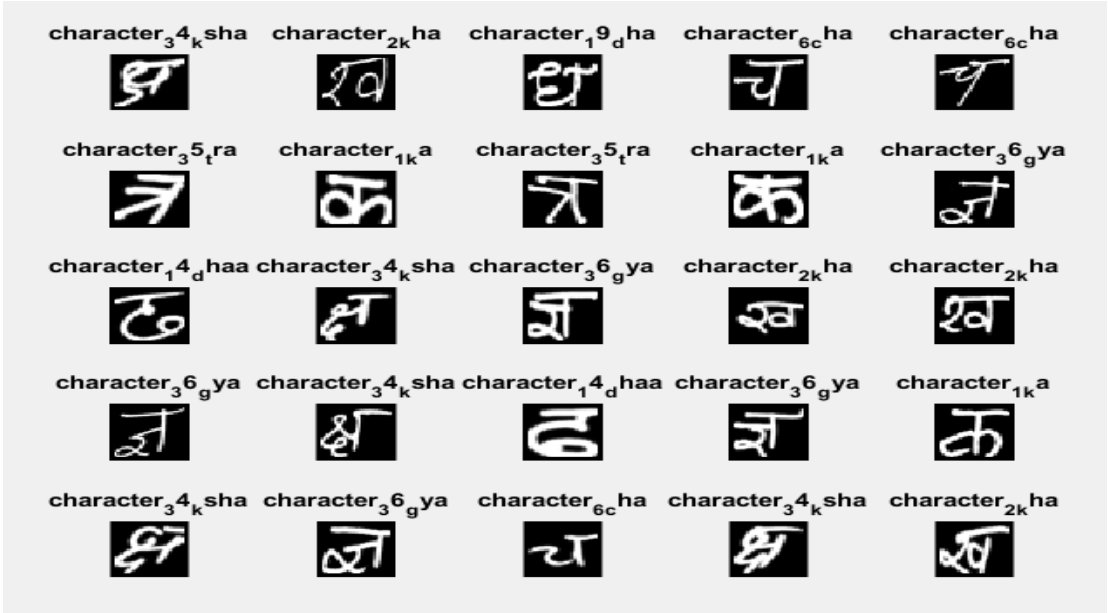

Figure:5.Predicted labels of characters

V. EXPERIMENTAL SETUP

Character Recognition is an image classification problem, and this paper proposes the transfer learning and ResNet-50 pre-trained model of convolution neural network for this task. In the paper, instead of designing a new deep convolutional neural network with the random initialization, we exploited existing trained CNN on images as a starting point to identify Devanagari characters. In CNN architecture, C is a convolution layer, R is a Rectified linear unit layer, N is the Normalization layer implementing local response normalization, P is pooling layer implementing max pooling, FC is the Fully Connected Layer.

ISSN: 2233-7853 IJFGCN

A) Data Preparation:-

The dataset of Handwritten Devanagari characters has 36 classes with 2000 images of each class. Dataset of 92,000 images has partitioned into a training set of 78,200 images (0.85) and a testing set of 13,800 (0.15). The images have a resolution of 32 by 32 in the grayscale png format with the original character in the centre. Along the sides, 2 pixels padding is used.

The dataset generated by UCI Machine Learning Repository had the handwriting of different people in wide variation that can be seen in Figure. The authors scanned handwritten documents and cropped each character manually to create this dataset of handwritten characters.`

Figure:6 Sample images to demonstrate variation of Handwriting

VI. RESULT AND DISCUSSION

For this experiment, MATLAB 2019b and Neural Network toolbox is used to implement CNN. To train the network, the dataset is divided into 3 parts (Training 56%, Validation 24%, testing 20%)

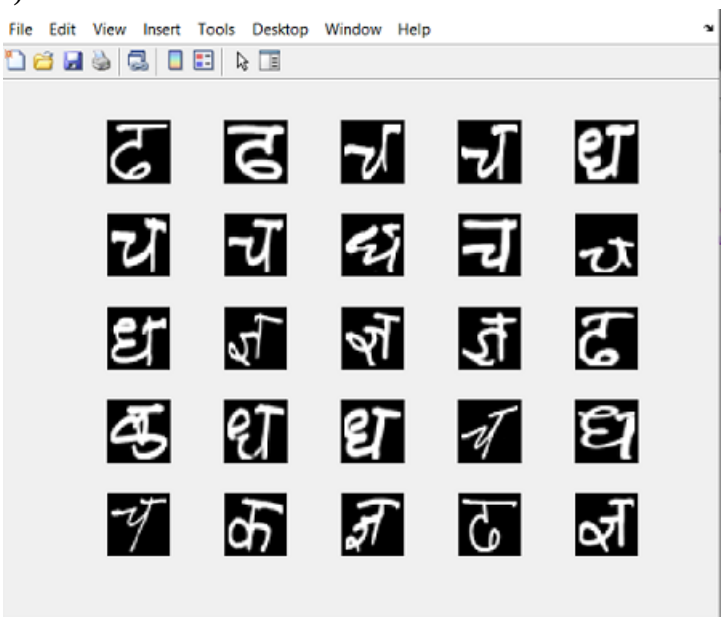

Figure:7. Random selected images from the dataset

ISSN: 2233-7853 IJFGCN

In figure 7 some random images are selected and displayed from the given dataset

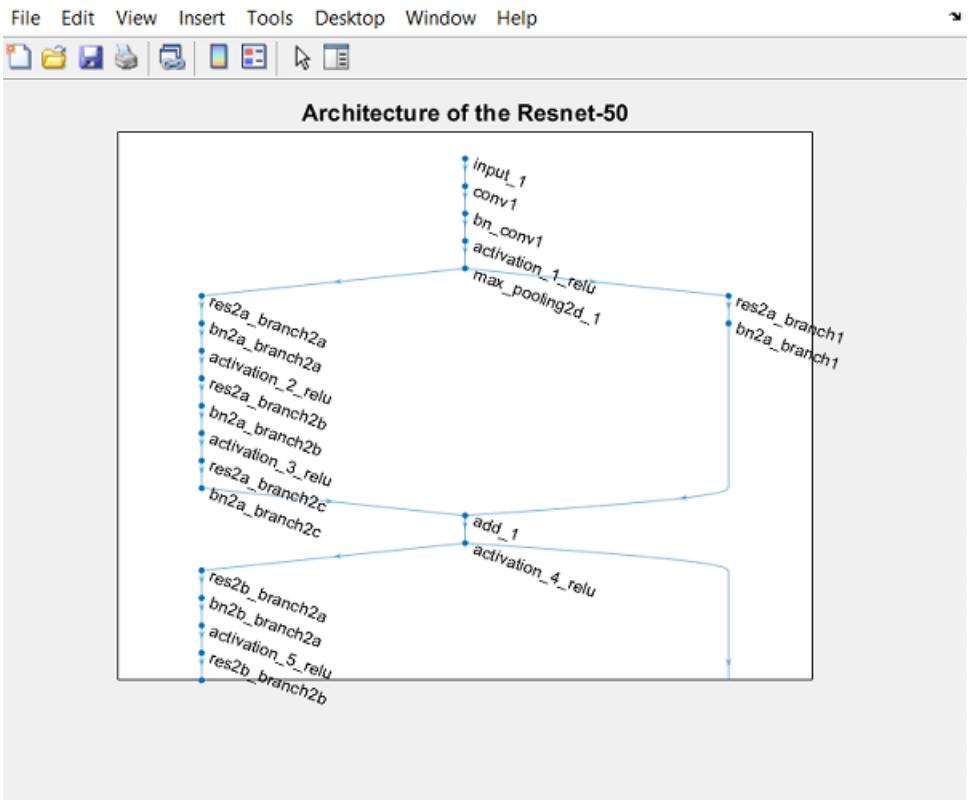

Figure 8.ResNet-50 Architecture

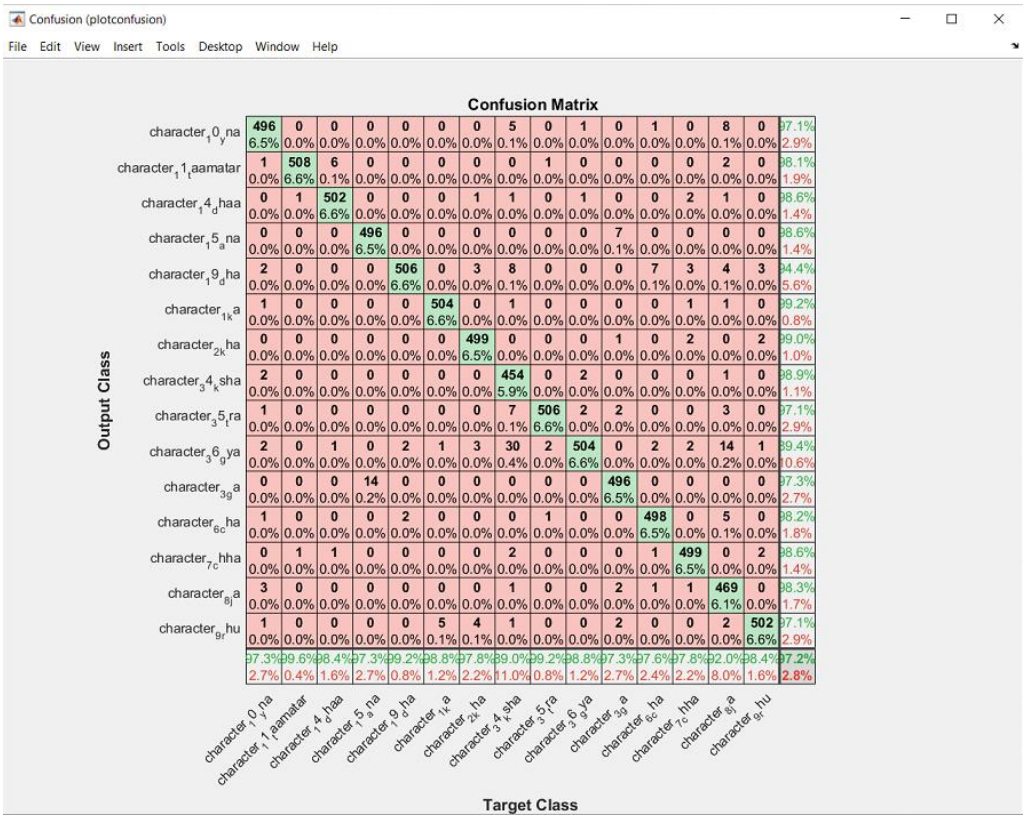

Figure9 Confusion Matrix

ISSN: 2233-7853 IJFGCN

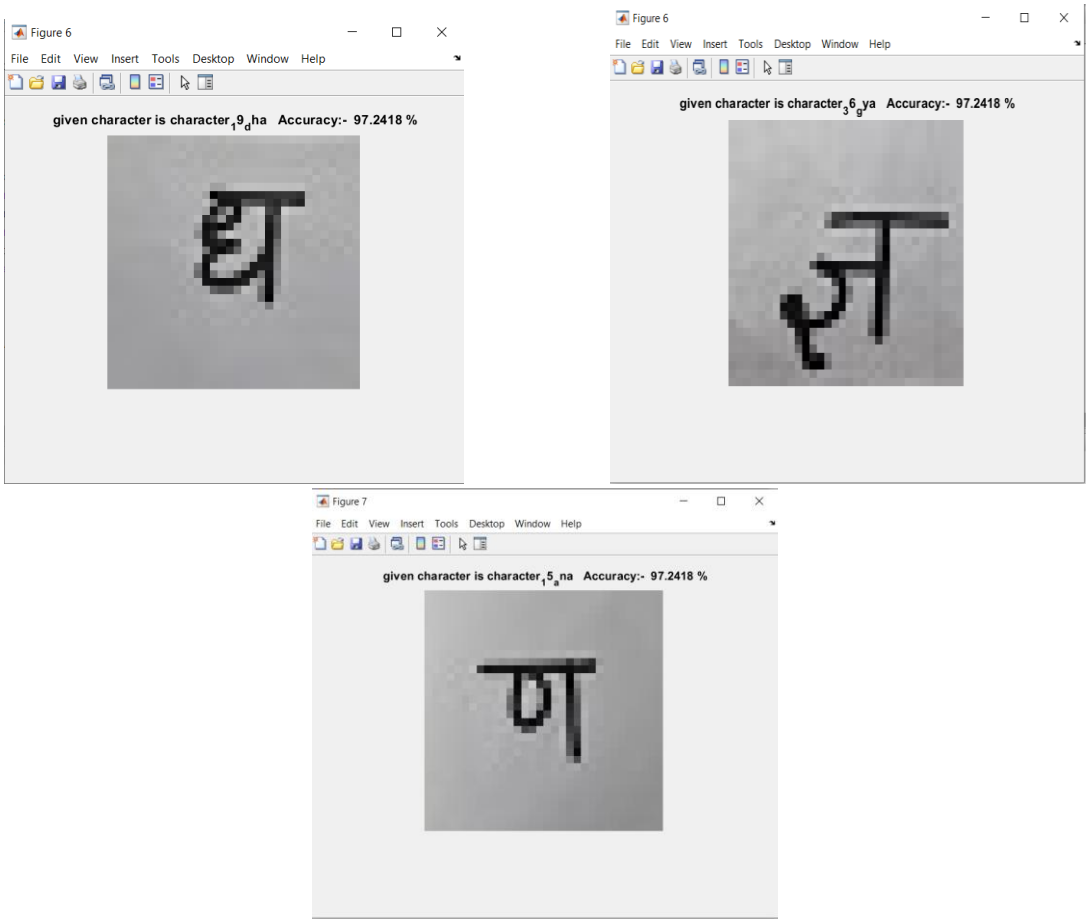

Figure 10.Classified images with accuracy

VII. CONCLUSION

In this paper a Devanagari character recognition system using Convolutional Neural Networks. By using ResNet-50, pre-trained network model the classification was done with an accuracy of 95-99% . CNN is hence a reliable technique for the classification of handwritten characters and can be implemented for applications including home schooling and automated answer checking. Devanagari character recognition is not only a major application in itself, but is also a significant step in the implementation of other applications.

REFERENCES

[1] Aarati Mohite ,Sushama Shelke “Handwritten Devanagari Character Recognition using Convolutional Neural Network”. 2018 4th International Conference for Convergence in Technology (I2CT) SDMIT
Ujire, Mangalore, India. Oct 27-28, 2018.

[2] Mahesh Jangid ,and Sumit Srivastava, “ Handwritten Devanagari Character Recognition Using Layer-Wise Training of Deep Convolutional Neural Networks and Adaptive Gradient Methods”. 13 February 2018

[3] Karbhari V. Kale, Prapti D. Deshmukh, Shriniwas V. Chavan, Majharoddin M. Kazi, Yogesh S. Rode. “ Zernike

ISSN: 2233-7853 IJFGCN

Moment Feature Extraction for Handwritten Devanagari (Marathi) Compound Character Recognition". (IJARAI)
International Journal of Advanced Research in Artificial Intelligence, Vol. 3, No.1, 2014

[4] Divakar Yadav, Sonia Sánchez-Cuadrado, and Jorge Morato, " Optical Character Recognition for Hindi Language
Using a Neural-network Approach" J Inf Process Syst, Vol.9, No.1, March 2013.

[5] Haider A. Alwzwazy1 , Hayder M. Albehadili , Younes S. Alwan , Naz E. Islam"Handwritten Digit Recognition UsingConvolutional Neural Networks" in International Journal of Innovative Research in Computer and Communication Engineering · February 2016

[6] Hayder M. Albeahdili, Haider A. Alwzwazy, Naz E. Islam." Robust Convolutional Neural Networks for Image
Recognition". (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 11, 2015

[7] Kaiming, He and Xiangyu, Zhang and Shaoqing, Ren and Jian Sun " Spatial pyramid pooling in deep convolutional
networks for visual" arXiv:1406.4729v4 [cs.CV] 23 Apr 2015.

[8] Chen-Yu Lee, Saining Xie, Patrick Gallagher, Zhengyou Zhang, Zhuowen Tu, " Deeply-Supervised Nets " NIPS 2014.

[9] Xiaoyu Wang Ming Yang Shenghuo Zhu Yuanqing Lin. " Regionlets for Generic Object Detection".ICCV2013.

[10] R. Girshick, J. Donahue, T. Darrell, and J. Malik. "Rich feature hierarchies for accurate object detection and
semantic segmentation". CoRR abs/1311.2524, 2013.

[11] Qasim Mahdi Haref a, Mohsin N. Srayyih AL-Maliki b, Maytham Mohammed Tuaama c ,Hayder M. Albehadili d.
"Arabic digit recognition using robust deep convolution neural network".Journal of Physics Conference Series ·May 2020.

[12] Fabien Lauer, Ching Y. Suen, and G´erard Bloch "A trainable feature extractor for handwritten digit recognition‖", Journal Pattern Recognition,Elsevier, 40 (6), pp.1816-1824, 2007.

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

[14] Krizhevsky, Alex, Sutskever, Ilya, and Hinton, Geoffrey. "ImageNet classification with deep convolutional neural
networks".

Website:

[15] https://www.analyticsvidhya.com/blog/2020/02/cnn-vs-rnn-vs-mlp-analyzing-3-types-of-neural-networks-
in-deep-learning/

ISSN: 2233-7853 IJFGCN