Gujarati Handwritten Character Recognition Using Convolutional Neural Network

[1]Mikita R Gandhi, [2]Dr. Vishvjit Thakar , [3]Dr. Hetal Patel

[1]Assistanrt Professor, [2]Professor, [3]Professor
[1]Electronics and Communication Engineering Department,
[1]A.D.Patel Institute of Technology, New V.V.Nagar, Gujarat, India

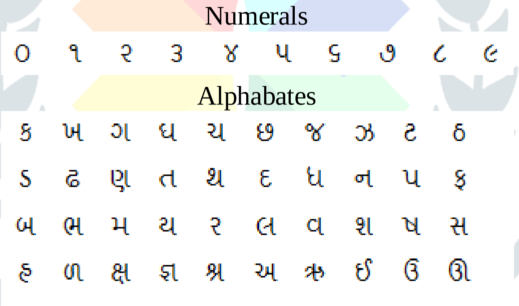
Abstract : **In the morden world, Handwritten character recognition is a method extensively used and still it is very essential challenge. Conventional machine-learning techniques need hand craft feature vector extraction from raw data and its required skill to convert input pattern into descriptor which could used in classifier to make a decision for a particular class where it belongs. To handle herewith dilemma, the Convolutional Neural Networks (CNN) was introduced recently, had not required traditional feature extraction method, the network learned itself and had given the better result than usual classification methods. In this paper, off- line Gujarati character was recognizing using Convolutional Neural Network. This method has been shown a new era for Gujarati handwritten OCR and gave satisfactory result in the field of Gujarati Handwritten Character Recognition. In this approach, Gujarati Handwritten database was created, 5000 samples for numerals and 10,000 samples for alphabets. Experiment result had shown 95.10% and76.25% accuracy for numerals and alphabets respectivly.**

IndexTerms - Handwritten Gujarati character Recognition, Convolutional Neural Network, Convolutional Layer, Pooling Layer, Fully connected Layer

I. Introduction

Nowadays, Recognition of handwritten characters in any language has been one of the most challenging and attractive research areas. To extract amount written on bank check, postal address and zip code recognitions are very important application for Off-line handwritten character recognition systems.

Gujarati is mother language of Gujarat state in India and most of the government and private agencies in Gujarat are used Gujarati language for official work. Most of research work down in this area was focused on Gujarati Numeral Recognition using traditional feature extraction method. In this paper, Convolutional Neural Network (CNN) was used for classification of handwritten Gujarati Numerals and alphabets. Figure 1 shows Guajarati numerals and alphabets used for handwritten classification work.



Numerals

૦	૧	૨	૩	૪	૫	૬	૭	૮	૯

Alphabates

ક	ખ	ગ	ઘ	ચ	છ	જ	ઝ	ટ	ઠ
ડ	ઢ	ણ	ત	થ	દ	ધ	ન	પ	ફ
બ	ભ	મ	ય	ર	લ	વ	શ	ષ	સ
હ	ળ	ક્ષ	જ્ઞ	શ્ર	અ	ઋ	ઈ	ઉ	ઊ

Figure 1 Gujarati Characters

A CNN is one of the methods of Deep Learning which acquired an input image, allot significance (learnable weights and biases) to different part/objects in the image and be intelligent to distinguish one from the other. The pre-processing task is very less necessary in a CNN as compared to other classification methods. While in primordial techniques filters are hand-engineered, with enough working out on feature extraction, CNN network has the skill to study these filters/characteristics itself [1].

CNN has unlimited accuracy as compare to traditional classifier like KNN, SVM etc, but it required very large database and high speed computational hardware.

The paper is organized as follows. Section 2 shows CNN used in different language by researcher next; Section 3 describes basic methodology and explains the CNN architectures in detail. Section 4 illustrates results and discussion. At last, section 5 provides a conclusion base on experiment result.

II. Related Work

Handwritten character recognition using Deep Neural Network is giving good accuracy compare to tradition classification techniques where feature extraction method totally depended upon algorithm designer. Handwritten Characters recognition of languages like English, Hindi, Devanagari, Arabic etc had given very stupendous accuracy using CNN.

Shibaprasad Sen et al [2] were worked on online Bengali handwritten character recognition where 1000 Bengali character samples were used for training and testing purpose. The algorithm was classified the handwritten samples in different 50 classes.Two stage CNN network was giving 99.40% accuracy of recognition. On the other handwritten [3] bengala digits standard dataset database had given 98.78% accuracy using CNN with Gabor filter Shrawan Ram et al [4] , recognize devanagari handwritten characters with 96.9% accuracy. This paper was described network architecture with detailed role of each layer. Akm Ashiquzzaman et al[5] used CMATERDB 3.3.1 Arabic handwritten digit dataset in CNN model and got 97.4 % accuracy with 10 different layer architecture. With help of the deep neural network and python programming language, Rohan Vaidya et al[6] had developed the android application which convert image of handwritten documents into text script. Karishma Verma[7] et al proposed CNN model for handwritten Hindi character recognition and achieved 99.9% accuracy.

III. METHODOLOGY

Gujarati character recognition was performed in three steps: database preparation, Formation of CNN Networks, classification of handwritten samples.

1. Database preparation

There is no standard database was available for hand written Gujarati characters. The database was created by5000 samples of handwritten Gujarati numerals where each 10 numerals consist 500 samples and 10,000 samples of alphabets where 250 samples for each 40 alphabets.

2. Formation Of CNN Network

The CNN architecture mainly consist convolutional layer, pooling layer, subsmpling layer, fully connected layer and classification layer. All these layers are connected to each other to framing the CNN architecture.

2.1 Convolutional Layer

Convolutional layer is the first layer of the CNN network. This layer is directly performed on the raw input data. Figure 2 was described the process of convolution layer. Consider 6 X 6 input image as shown in Figure 2(a) and on this image 3 X 3 filter mask was applied which was shown in Figure 2(b). The convolutional layer was operated on input image as follows:

$$y\left(a+\frac{m}{2}, b+\frac{n}{2}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} w(i,j) * x(a+i-1,\ b+j-1) \qquad (1)$$

Where x is input image of size m X n, w is the filter mask and y is the output image.

Figure 2(c) shown the resultant output of convolution layer using the stride=1 of filter mask. Figure 2 was shown only single filter output, but in CNN network multiple filters with different mask were applied on single input image and generates multiple output images

1	1	0	0	0	1
0	1	1	0	1	0
0	0	0	1	0	0
1	0	0	0	1	0
1	0	1	0	1	1
0	0	0	0	1	1

(a)

-1	-1	-1
1	1	1
-1	-1	-1

(b)

-1	0	1	-1
-3	-1	-2	-1
-2	-2	-1	-2
0	1	0	-1

(c)

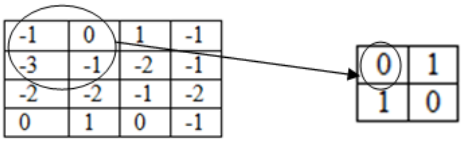
Figure 2 convolutional layer Process

2.2 Pooling Layer

The second layer which comes after the convolutional layer of the CNN network is pooling layer, also called sub sampling layer. The pooling operation can be done by two ways, maximum pooling or average pooling. The main purpose of pooling layer is reduction in feature map size. The maximum pooling function is performed on input y_i is follows:

$$M(y_i) = max\{y_{i+k,i+l}||k| \le \frac{m}{2}, |l| \frac{n}{2} \le k, l \in N\} \qquad (2)$$

Where k and l indicates numbers of rows and columns respectively. Consider value of k=l=2 and applied on Figure 2(c). Pooling Layer output was shown in Figure 3(b) where output image size was half of convolutional layer output image.



-1	0	1	-1
-3	-1	-2	-1
-2	-2	-1	-2
0	1	0	-1

0	1
1	0

(a) (c)

Figure 3 pooling layer output

+

2.3 Fully Connected Layer

Fully connected layer is the last layer of CNN network. This layer mapped output of pooling layers into number of classification layers.

2.4 ReLU Layer

ReLU called Rectified Linear Unit which perform non-linear operation where it execute element by element and substitute pixel which has negative values in the feature map by zero.

Output = Max(zero, Input) (3)

2.5 Softmax Layer

This layer emphasizes the biggest values of fully connected layers and minimizes values which are extensively below the maximum value and this way it transformed all values into rang of 0 to 1.

2.6 Classification Layer

Classification layer classify the input image in the particular class which obtain highest value in softmax layer.

3. Proposed CNN Architecture

In this proposed method, the CNN architecture contain total eight layers in which two convolutional layers, two maximum pooling layers, two ReLU layers, one fully connected layer and one Softmax layer.

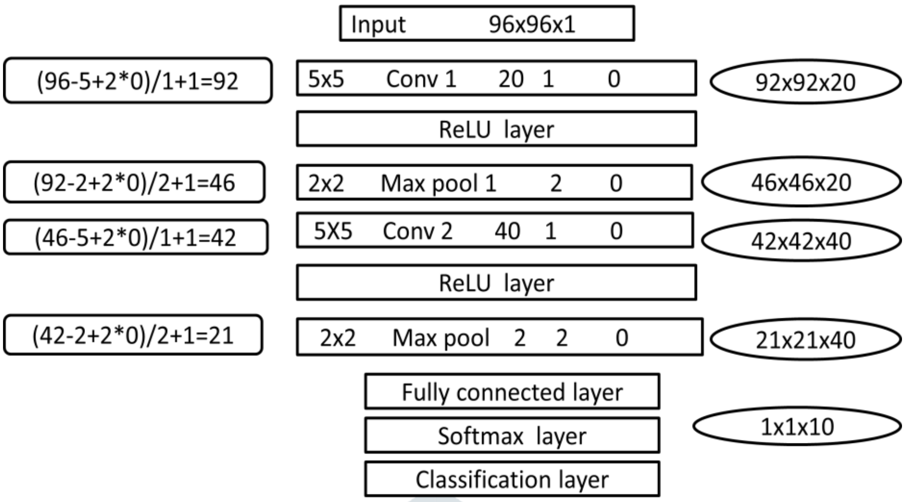


Figure 4 Proposed CNN Architecture

As shown in Figure 4, the proposed architecture was applied on two different input size data. The created Guajarati database has sample size 64 X 64. So, without resizing input image was applied to CNN network for training and testing. Then for better result sample images were resized into 96X96 in next time.

To understand CNN architecture considers the input image size was 96X96. This image was applied to convolutional layer 1. In conv1 layer, 20 indicate number of convolutional filter mask of size 5X5, 1 specifies number of stride and 0 shows no padding. Right side indicates output of 92X 92 size 20 images. These images given to ReLU layers where all negative values of pixels are replaced by zero. Then images are applied to maximum pooling layer which reduced image size into 46X46. Then repeat same procedure with 40 filters. For numerals fully connected layers used 10 class while for alphabets used 40 classes.

For implementation purpose this architecture was verify by two models : the fist model same as indicate in Figure 4 while in second model number of filters were changed in conv1 with 40 filters and conv2 with 80 filters.

IV. Result and Discussion

The experiment was conduct on Gujarati database of Numerals and alphabets using Matlab. The 80 % samples were given to train the CNN network and 20 % samples were used for testing purposed.

Figure 5 was showing each stage output of proposed method. Here Guajarati numeral '1' image was taken to demonstrate result of each layer. Convolutional layer 1 show 20 different images were given to ReLU and pooling layer. Pooling layers1 output show 20 images where size of images are half and became darken because of ReLU layer. Convolutional layer 2 and pooling layer 2 show result for 40 filters. Fully connected layer output shows how much digit '1' matched with other digits. Finally softmax layer indicate the input image was digit '1'.

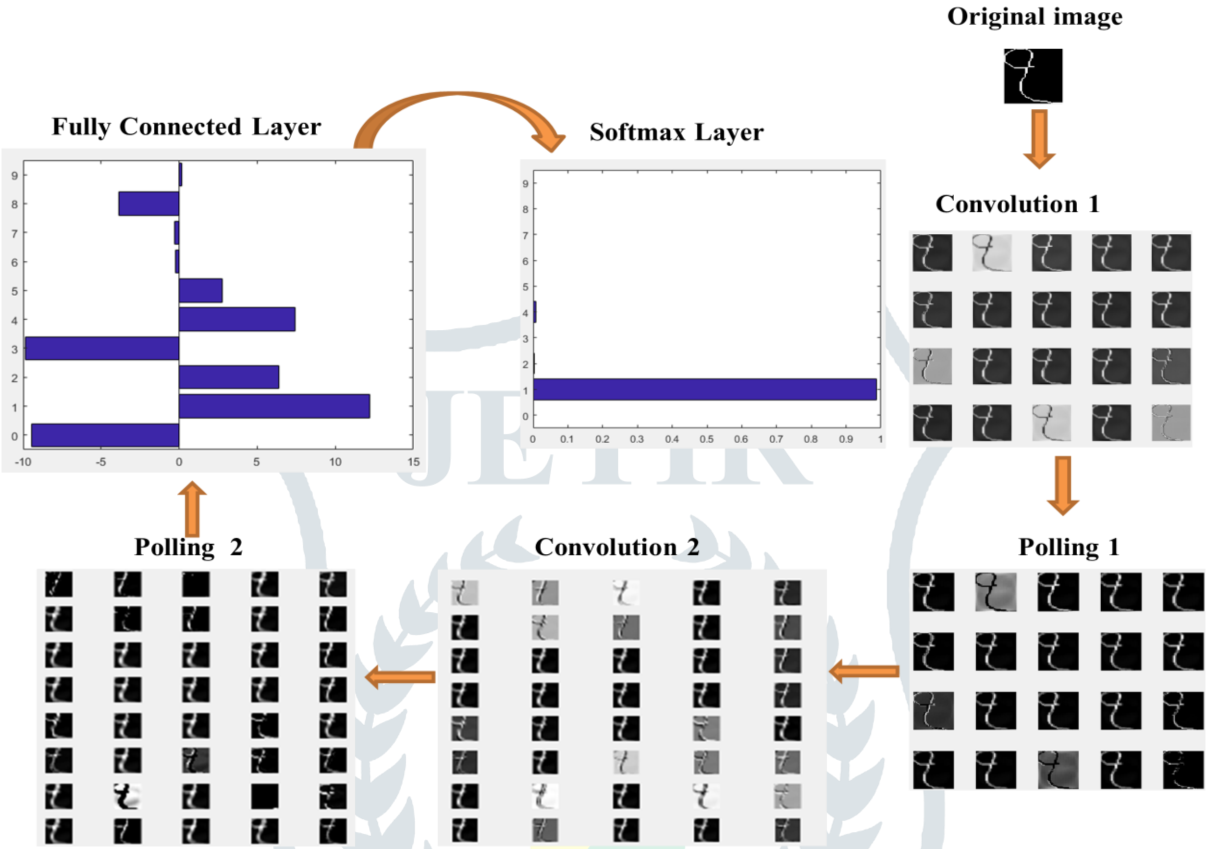


Figure 5 Each stage output images with conv1 with 20 filters and conv2 with 40 filters

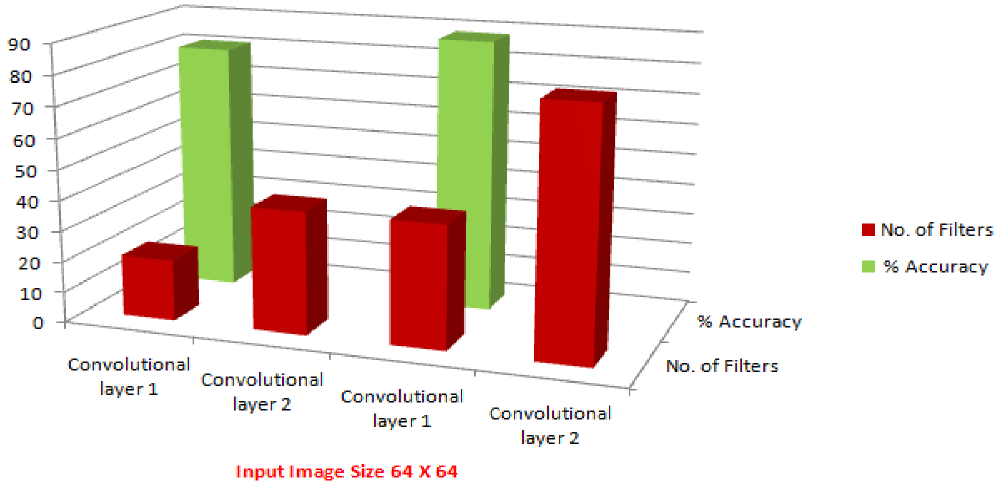


Figure 6(a) Gujarati Numerals accuracy with input size 64X64

Figure 6(a) and (b) shows accuracy for Gujarati Numerals with different input size of images. Using 40 and 80 filters in convolutional layers, Guajarati numerals were recognize with 95.10% maximum accuracy. Same way Figure 7(a) and (b) shows accuracy for Gujarati Alphabets and got 76.25% accuracy.

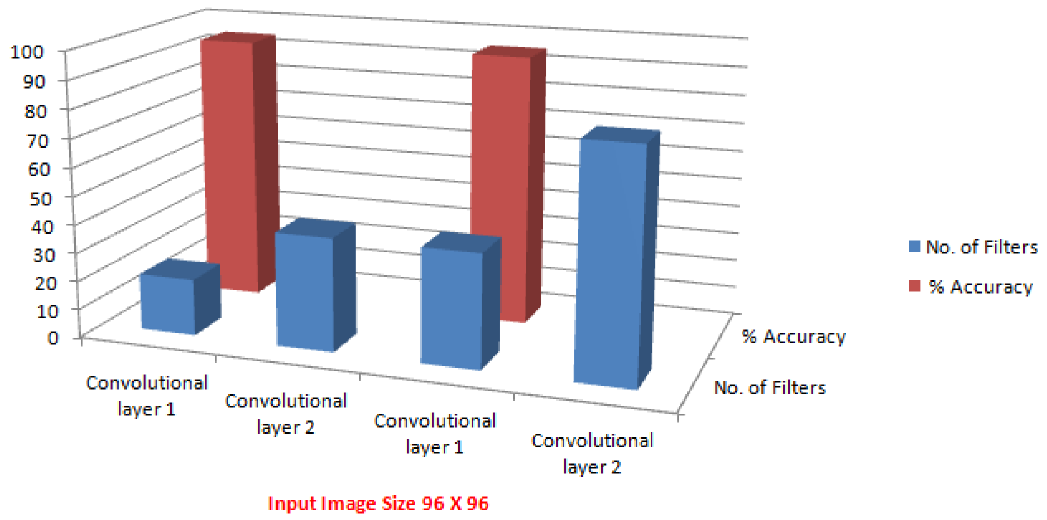


Figure 6(b) Gujarati Numerals accuracy with input size 96X96

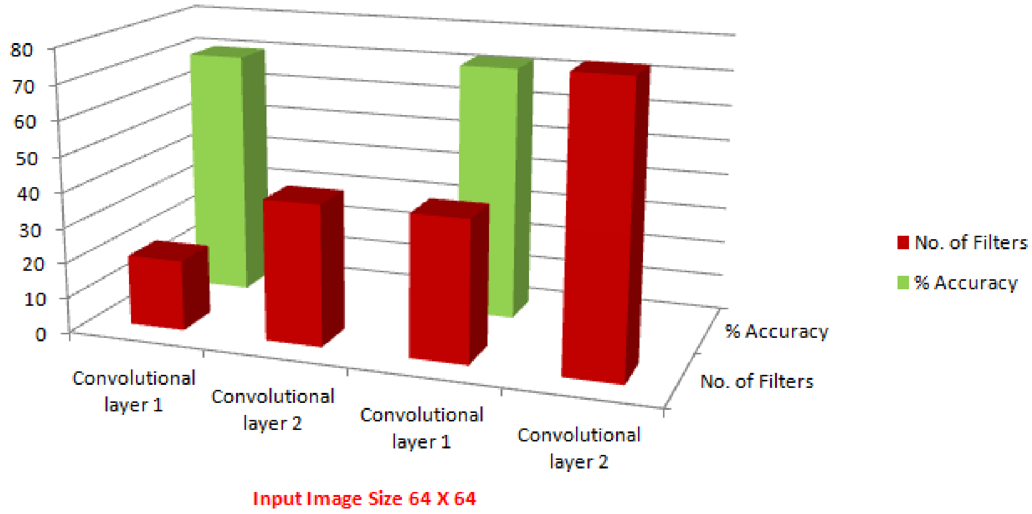


Figure 7(a) Gujarati Alphabets accuracy with input size 64X64

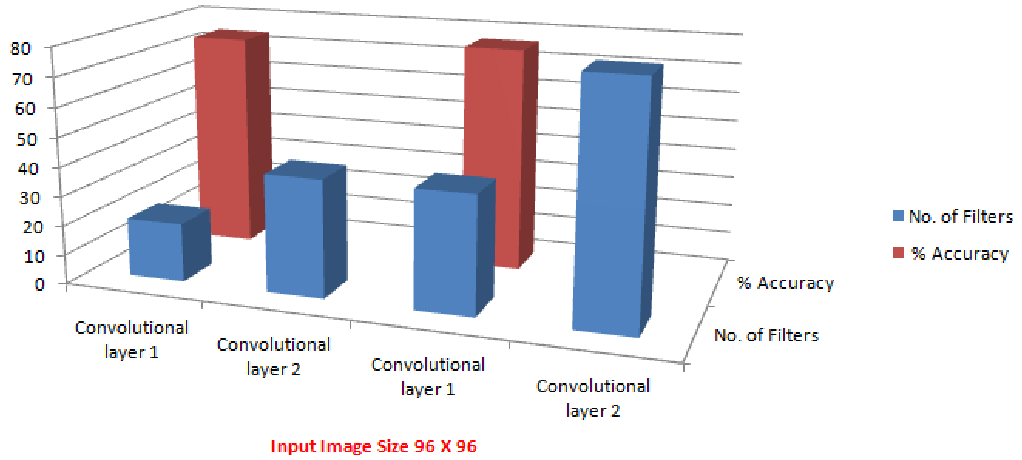


Figure 7(b) Gujarati Alphabets accuracy with input size

V. Conclusion and Future Work

In this paper, new approach for Handwritten Gujarati Character Recognition had introduced. CNN network had increased accuracy with increasing the no of filters and also with increasing input image size. The proposed model gives 95.10% accuracy for Gujarati numerals and 76% accuracy for Gujarati Alphabets.This accuracy can be increased by adding more stages of convolution layers and using transfer learning approach of pretrained network like Alexnet, Vgg16,GoogLeNet etc.

References

[1] Chaouki Boufenar, Adlen Kerboua, Mohamed Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition", Elsvevier Cognitive Systems Research**,**Volume 50, August 2018, Pages 180-195

[2] Sen S., Shaoo D., Paul S., Sarkar R., Roy K. ,"Online Handwritten Bangla Character Recognition Using CNN: A Deep Learning Approach", In: Bhateja V., Coello Coello C., Satapathy S., Pattnaik P. (eds) Intelligent Engineering Informatics. Advances in Intelligent Systems and Computing, vol 695. Springer, Singapore,2018.

[3] Alom, M.Z, Sidike, P., Taha, T.M., Asari, V.K.., "Handwritten Bangla Digit Recognition Using Deep Learning" Journal Neural Processing Letters; 45, pp: 703-725,2017

[4] Shrawan, R., Shloak, G., Basant, A., "Devanagri character recognition model using deep convolution neural network", Journal of Statistics and Management Systems Volume 21, Issue 4.,2018.

[5] Ashiquzzaman, A., Tushar, A. K.., "Handwritten arabic numeral recognition using deep learning neural networks.", IEEE International Conference on Imaging, Vision & Pattern Recognition (icIVPR), pp. 1–4. IEEE,2017

[6] R. Vaidya, D. Trivedi, S. Satra, M. Pimpale,"Handwritten Character Recognition Using DeepLearning". Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 772-775, 2018

[7] Karishma Verma, Manjeet Singh "Hindi Handwritten Character Recognition using Convolutional Neural Network", International Journal of Computer Sciences and Engineering, Vol.-6, Issue-6, pp.909-914,ZJune 2018

[8] Acharya, Shailesh Raj et al. "Deep learning based large scale handwritten Devanagari character recognition." 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA) (2015): 1-6.,2015

[9] H. Larochelle, Y. Bengio, J. Louradour and P. Lamblin "Exploring Strategies for Training Deep Neural Networks," The Journal of Machine Learning Research Vol. 10, pp. 1-40, 2009

[10] T.N. Sainath, A.–R. Mohamed, B. Kingsbury and B. Ramabhadran, "Deep convolutional Neural Networks for LVCSR," IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP), pp. 8614-8618, 2013

[11] Samad Roohi, Behnam Alizadehashrafi, "Persian Handwritten Character Recognition Using Convolutional Neural Network," 10th Iranian Conference on Machine Vision and Image Processing, Nov, 22-23, 2017.

[12] E.Kussul and T. Baidyk, "Improved method of handwritten digit recognition tested on MNIST database," Image Vis. Comput., vol. 22, pp. 971-981, 2004.

[13] Krizhevsky A, Sutskever I, Hinton G E. " Imagenet classification with deep convolutional neural networks". Advanees in neural information processing systems. Pp. 1097-1105,2012