JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A NOVEL DEEP LEARNING ARCHITECTURE FOR DEVANAGARI CHARACTER RECOGNITION

[1]Vadapalli Kameswara Kiran, [2] Durga Ganga Rao Kola

[1]Student, M.Tech(CE&SP), Department of Electronics and Communication Engineering, UCEK(A), JNTU Kakinada, Andhra Pradesh, India, 533003.

[2]Assistant Professor, Department of Electronics and Communication Engineering, UCEK(A), JNTUKakinada, Andhra Pradesh, India, 533003.

Abstract : Devanagari is one of the most widely spoken languages in India, an ancient nation with many varied languages and civilizations. It is used in several other Asian nations in addition to India.Devanagari manuscripts provide a wealth of knowledge that can be used to advance understanding in a variety of subjects. Due to their age, writing style, ink stains, and damaged pages, these manuscripts are decaying. These manuscripts can provide information when they are digitalized. In order to recognize the characters in Devanagari manuscripts with better optimization, optical character recognition (OCR) is vital. Convolutional neural network, a potent technique in the field of pattern recognition, is utilised to classify these identified characters. Convolutional neural networks are used in this paper as a feature extractor and classifier for identifying these documents. 46 classes of Devanagari characters are taken into account, testing and validation data are separated.With the suggested CNN architecture, recognition accuracy of 98.26% is achieved using 75% training data and 25% testing data.

***Index Terms*- Devanagari manuscripts, Optical Character Recognition, Convolutional Neural Networks, Feature extraction, Classification.**

1. INTRODUCTION

Along with India, many other Asian Countries also use Devanagari. Ancient Devanagari manuscripts provide a wealth of knowledge. But because they are old, these texts are deteriorating day by day. So, one of the greatest approaches will be to digitize these documents in order to obtain information from the manuscripts. These manuscripts are scanned, binarized, and then segmented during the OpticalCharacter Recognition (OCR) process of digitization. For feature extraction and classification, a set of 46 classes of Devanagari characters is taken into consideration. The most popular deep learning technique, convolutional neural network, is utilized to extract features from the Devanagari characters. CNN is chosen because it works with raw pixel data and does not require any additional features for Image classification. Data validation and training are done using a novel CNN architecture. In order to classify data as well as extract features, CNN is used. Classification methods include Support Vector Machines, which divide the data into hyperplanes.

The Brahmi script gave rise to Devanagari. Scholars have puzzled about the meaning of the word Devanagari; one theory is that it may be a mixture of the Sanskrit words "Deva" (God, ruler, or Brahmans) and "Nagari" (city). Literally, they become "City of Gods" and "Script of Gods."[1]. Handwriting recognition is crucial to information processing in the contemporary era of digitization [2]. Applications of optical character recognition include reading license plates, sorting postal mail, scanning ancient documents for preservation in archaeological departments, automating the processing of old documents in libraries and banks, etc., [2]. There aren't many hierarchical learning-based OCR methods available, especially when there is noise. In addition, there are not many relevant databases that could be used to create and test these algorithms [3]. Numerous studies demonstrate that utilizing CNN as a feature extractor improves accuracy over other cutting-edge methods [4]. SVMs are a feasible solution for both classification and regression issues. It is an algorithm for supervised learning. It operates according to the idea of margin calculation. Eachpiece of data is shown as a point in n-dimensional space in this procedure (where n is the number of features). Each feature's value is the matching coordinate's value. It divides the training datasets into classes by locating a line (hyper plane) that divides the data into those classes. It operates by increasing the distances between the closest data point (in both classes) and the so- called margin hyperplane [5].

The rest of the paper is organized as follows:

In the Section 2, Attributes of Devanagari Script are discussed.
Data Collection and Pre-processing are discussed in Section 3.
In the Section 4, Proposed Methodology is discussed.
In the Section 5, Experimental Setup and Results are discussed.
Finally, Conclusion and Future Scope is given in Section 6.

2. ATTRIBUTES OF DEVANAGARI MANUSCRIPT

Devanagari Script is part of the Brahmic family of scripts of countries like India, Nepal, Tibet, and Southeast Asia, etc., Devanagari has been widely adopted across India and Nepal and used to write Sanskrit, Marathi, Hindi, Central Indo-Aryan languages, Konkani, Boro, and various Nepalese languages. The 36 consonants, 10 numerical characters, 12 vowels, and a few unique letters make up the Devanagari script. Shirorekha, or the headline, is the horizontal line that runs over the top of each character. Two or more consonants can be combined to create compound letters. Conjucts are characters created through the use of modifiers. Tables-1 and 2 show vowels and consonants, respectively.

Table-1: Vowels

Vowels	अ	आ	इ	ई	उ	ऊ	ऋ	ए	ऐ	ओ	औ

Table-2: Consonants

Consonants				
क	ख	ग	घ	ङ
च	छ	ज	झ	ञ
ट	ठ	ड	ढ	ण
त	थ	द	ध	न
प	फ	ब	भ	म
य	र	ल	व	स
ष	श	ह	क्ष	त्र
ज्ञ				

3. DATA COLLECTION AND PRE-PROCESSING

Preparing the dataset is the most important stage in using CNN to recognize Devanagari characters.The dataset must be created by binarizing and segmenting Devanagari manuscripts before creating the CNN network architecture. The firststep is to take the degraded Devanagari manuscripts, scan them, and then transform them into binary images. Binary pictures have pixel values of 0 and 1, respectively, and are composed of only black and white pixels. To create the dataset, these binary images are segmented and then labelled. The procedure of histogram projection technique is used to segment the image into lines, words, and characters during the segmentation phase. Finally, a dataset with 46 classes and 92,000 samples of handwritten Devanagari characters is obtained [7]. The flow of Data collection and Preprocessing is shown in Fig-1.

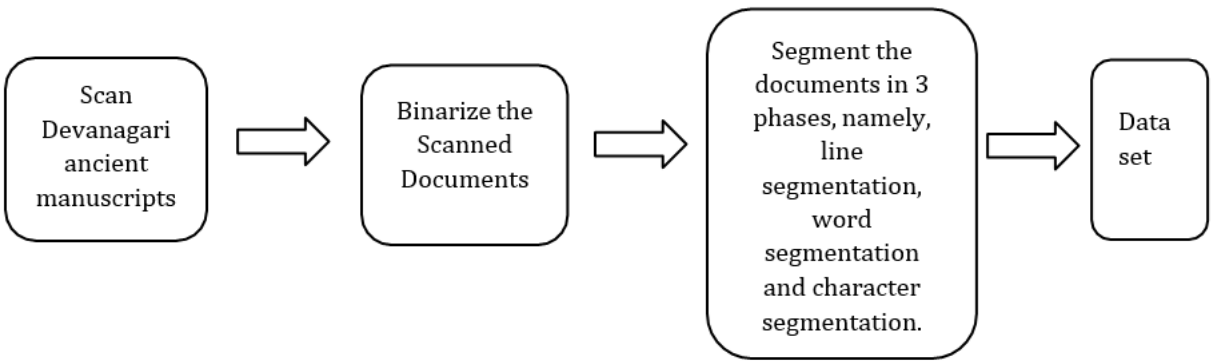

Fig-1: Flow diagram for dataset preparation

4. PROPOSED APPROACH

Convolutional neural networks are a certain kind of neural network that are specifically created to operate on image pixel data without the need of any particular feature for image classification of neural network that are specifically created to operate on image pixel data without the need of any particular feature for image classification. It takes Images as input of the network and considers various aspects/objects so that it differentiates one image with other. CNN consists of different layers in its architecture namely Image Input layer, Convolutional layer, Max Pooling layer, Batch Normalization layer, ReLU layer, Dropout layer, Fully Connected layer.Devanagari Dataset is spitted into training and validating data and training data is considered during the training of the network and it is been validated with considered testing data. The ratio of testing and Validating data is considered in two ratios of 75/25 and 80/20. They are been trained with 20,25 and 30 epochs each ratio and respective accuracies are obtained. The flow of the Propose approach is shown in Fig -2.

The considered CNN Network architecture is been designed using Convolutional layer, ReLU layer, Batch Normalization layer, Max Pooling layer, DropoutLayer and Fully connected layer.

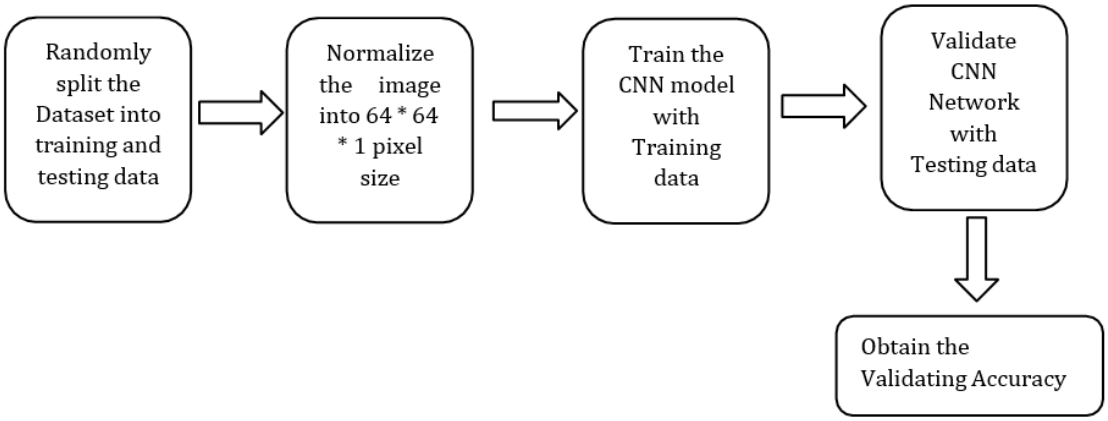

Fig -2: Flow of the proposed architecture.

4.1 Convolutional Layer

The images in Convolutional layer are reduced in a form which are useful to process without losingits features for getting good Prediction. A Kernel is covered over entire image by taking stride value as 1. The kernel value is been multiplied bythe value in the respective pixel and finally, the convoluted value is replaced at the original valueof the pixel. 5 * 5 Kernel is considered for convoluting entire image. The Image after passing through Convolution filters of the convolutional layer can be shown as Fig-3.

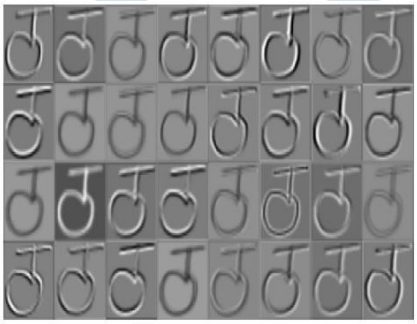

Fig-3: Character ठ after passing throughConvolutional Layer

4.2 Rectified Linear Activation Layer (ReLU)

It is one of the activation layers which works based on the function defined and activates the respective values if the function is satisfied. With ReLU layer they are identity, binary step, arctan are also there which are also activations layers. The function considered is given as:

$$f(x) = 0, x<0$$

$$f(x) =1, x\geq 0$$

4.3 Batch Normalization Layer

The minibatch data resulting from all the observations of each channel is normalized independently in this Layer. In order to increase the training of the CNN network and reduce the sensitivity that initialize network usually Batch Normalization Layer is been used between Convolutional Layer and nonlinearities.

4.4 Max Pooling Layer

To reduce the Spatial size of the feature map and remove spatial information, Convolutional layer sometimes followed by down sampling operations. Max Pooling Layer is used mainly for two reasons. To minimize the number of pixels by downsampling and transfer the important information to the next layer.

4.5 Dropout Layer

For a considered probability, the input elements are set to Zero randomly in the Dropout Layer. Usually, its probability value considered as 0.5. A dropout mask is initiated to set the input elements to zero randomly. In order to prevent the network from overfitting, this operation effectively modifies the network's basic design between iterations. Higher number elements are dropped during training. At the time of prediction, the layer's output and input are equal. No learning takes place in this layer which is similar to max or average pooling layer.

4.6 Fully Connected Layer

The proposed method includes a convolutional layer, a downsampling layer, and one or more fully connected layers, where each neuron is connected to every neuron in the layer before it. The number of classes for the considered dataset should be defined in this layer and probability of classification score is defined in this layer.

5. EXPERIMENTAL SETUP AND RESULTS

The experimental setup of the proposed work can be explained in the steps.

A Devanagari Dataset of 46 classes with 92000 characters are considered.
The images are normalized to a pixel sizeof 64 *64*1. Here 1 represents the channel of the image while RGB images are considered with 3 as the value. Binary images are those which consists of only black and white as their pixels.
The Considered Dataset is spitted into Training and Validating data set with ratios of 75/25% and 80/20%.
The Convolutional filters considered in Convolutional layer in CNN architecture are 32. The kernel size considered as 5 *5 with stride of 1.
Max Pooling layer is considered with kernel size of 2 * 2 with stride of 2.
The Network architecture is trained with Stochastic gradient descent (SGD) with momentum of 0.1.
The min batch size considered for training is 200, Validating frequency is 50 and initial learning rate 0.005.
The Proposed Network architecture is shown in Fig -4.

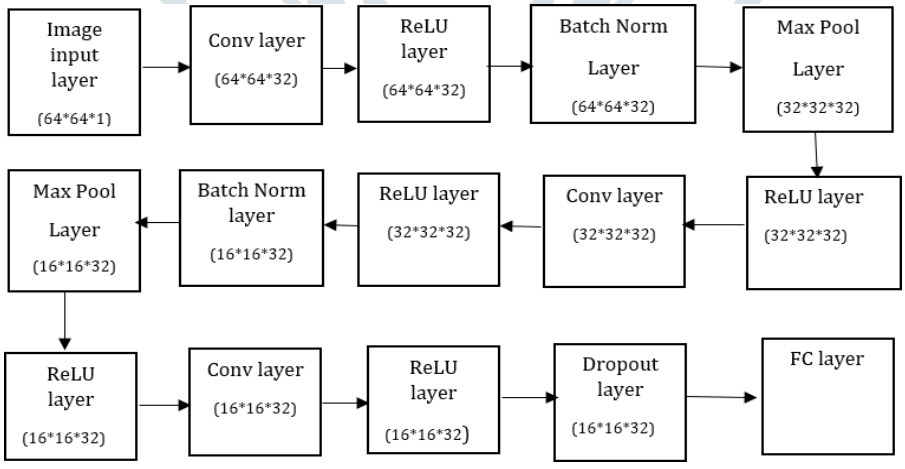

Fig-4: Architecture of CNN Network

The dataset is trained and validated by considering the Fig -4 CNN architecture. Firstly, the dataset is spitted into 75% of training data, 25% of validating data and they are trained by considering number of epochs as 20, 25, 30 at respective times. The accuracy obtained for 75- 25% partition on 30 epochs is 98.26% and for 80-20 % partition on 30 epochs is 98.05%. The experimental results are shown in table 3. For Comparison purpose Support Vector Machine (SVM) is also considered as a Classifier and respective accuracy are observed for 75-25% and 80-20% partitions.

Table-3: Experimental results on respective partitions with different number of epochs

Test/Train	Epochs	Proposed approach with SVMas Classifier	Proposed approach with CNN as Classifier
80/20	20	98.26%	98.19%
80/20	25	98.26%	98.22%
80/20	30	98.26%	98.05%
75/25	20	98.16%	98.16%
75/25	25	98.16%	98.23%
75/25	30	98.16%	98.26%

The training plots for 80-20 % and 75-25% partitions with 25,30 epochs each are shown in Fig-5, 6, 7 and 8. It is noticed that proposed methodology obtains good accuracy.

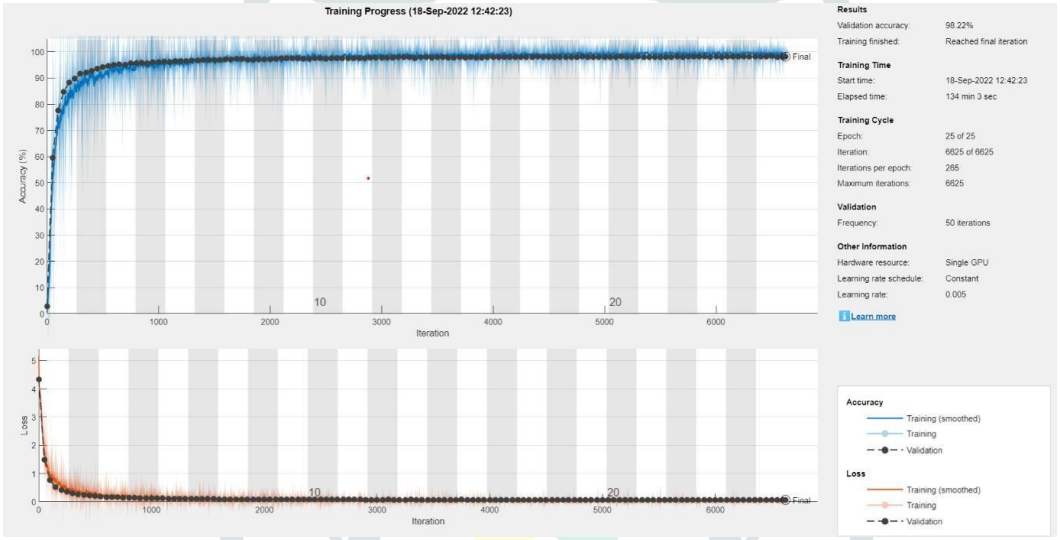

Fig -5: Training plot with 80-20% Train/Test partition with 25 epochs

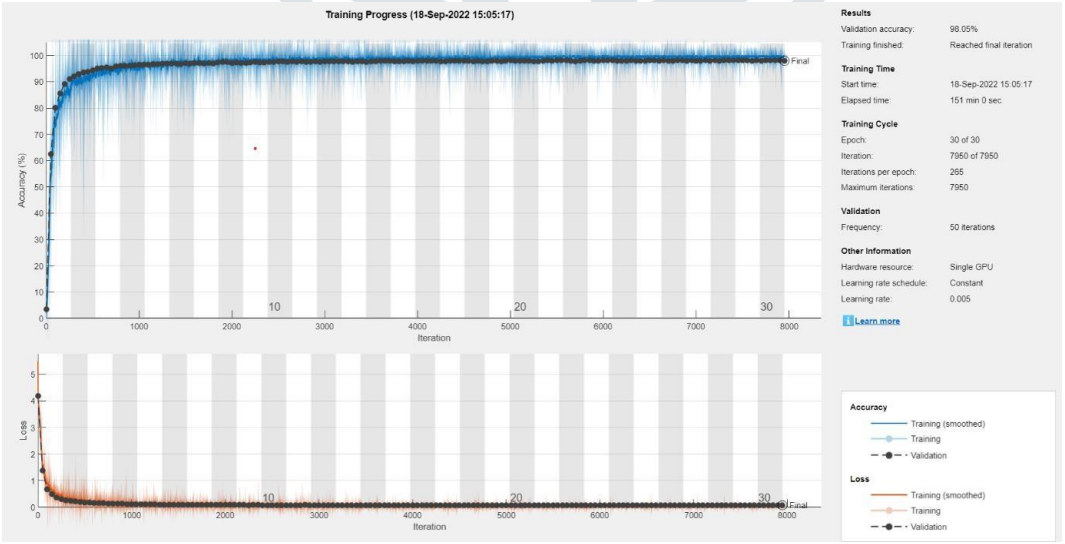

Fig-6: Training plot with 80-20% Train/Test partition with 30 epochs

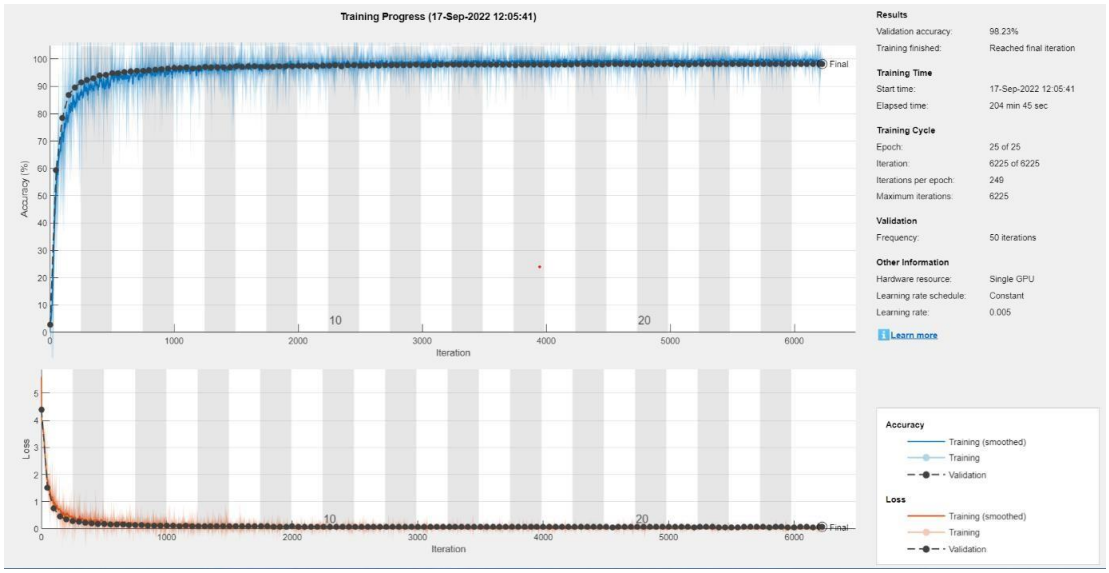

Fig-7: Training plot with 75-25% Train/Test partition with 25 epochs

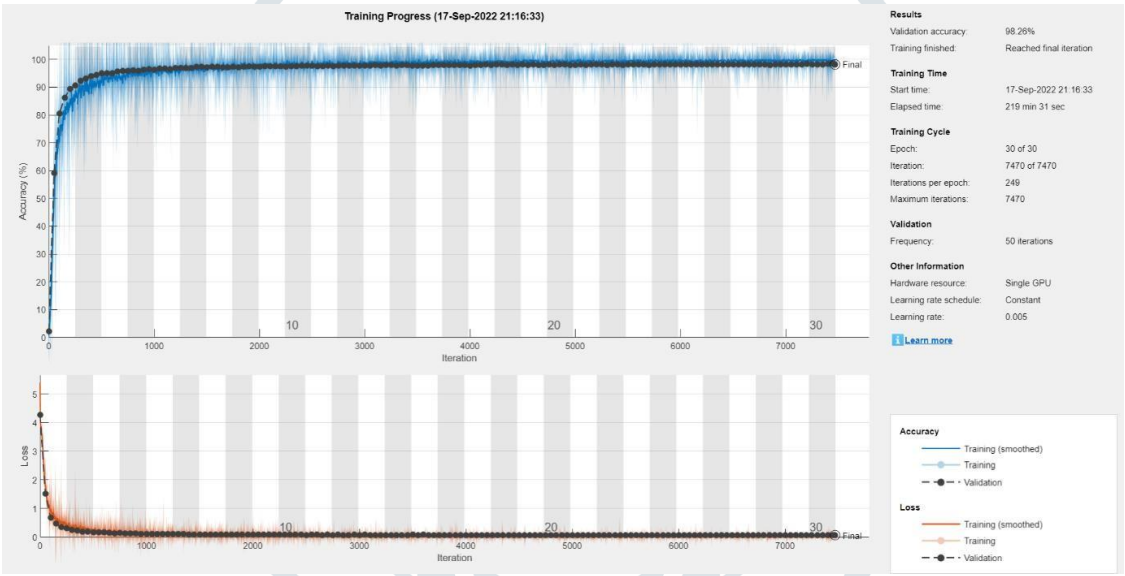

Fig -8: Training plot with 75-25% Train/Test partition with 30 epochs

6. CONCLUSION AND FUTURE SCOPE

This paper presents a novel deep learning architecture for Devanagari Character Recognition. A dataset of 92000 characters with 46 classes are considered for this work. Those handwritten characters are segmented from Devanagari Manuscripts which are about to degrade. Convolutional Neural Network (CNN) isused for feature extracting and classification of the Devanagari ancient characters since Deep Learning is observed as new fashion in the field of pattern recognition and character recognition.An accuracy of 98.05% is obtained with 80%training and 20% testing data and 98.26% of accuracy is obtained with 75% training and 25% testing data for 30 epochs each. In the present work, a dataset of 46 classes only is considered. In the future, some more characters in the data which can be defined from individual characters when wrapped up with the vowels are introduced to obtain the better accuracy.

REFERENCES:

[1] Avadesh M and Goyal N (2018). Optical character recognition for Sanskrit using convolution neural networks.Proceedings of the 13th IAPR International Workshop on Document Analysis Systems (DAS), Vienna, 447- 452.

[2] Ahlawat S, Choudhary A, Nayyar A, Singh S, Yoon B (2020) Improved handwrittendigit recognition using convolutional neural networks (CNN). Sensors20(12):3344.

[3] Alizadehashraf B and Roohi S (2017).Persian handwritten characterrecognition using convolutional neural network Proceedings of the 10th Iranian Conference on Machine Visionand Image Processing, 247-251.

[4] Sonika Rani Narang, Munish Kumar, M.K. Jindal, "DeepNetDevanagari: a deeplearning model for Devanagari ancient character recognition", Multimedia Tools and Applications, 2021.

[5] Alzubi J, Nayyar A, Kumar A (2018) Machine learning from theory to algorithms: an overview. J Phys Conf Ser 1142(1):012012.

[6] S. Acharya, A.K. Pant and P.K. Gyawali â€œDeep Learning Based Large Scale Handwritten Devanagari Character Recognitionâ€•,In Proceedings of the 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA),pp. 121-126, 2015.

[7] A. K. Pant, S. P. Pandey and S. P. Joshi "Off-line Nepali Handwritten Character Recognition using Multilayer Perceptron and Radial Basis Function neural networks," Third Asian Himalayas International Conference on Internet(AH-ICI), pp 1-5, 2012.

[8] Create a simple Deep Learning Network for image Classification MathWorks. https://in.mathworks.com/help/deeplearning/ug/create-simple-deeplearning network-for- classification.html.

[9] V. J. Dongre, V. H. Mankar "A Review of Research on Devnagari Character Recognition," International Journal of Computer Applications Vol. 12, No.2, pp.0975 – 8887, November 2010.

[10] M. G. Quiles, R. Romero, "A ComputerVision System based on MultiLayer Perceptrons for Controlling Mobile Robots," ABCM Symposium series in Mechatronics- Vol 2 pp. 661-668.

[11] D. W. Ruck, S. K. Rogers, M. Kabrisky, "Feature Selection Using a Multilayer Perceptron," Journal of Neural Network Computing, Volume 2, pp 40-48, November 1990.

[12] J. Yang, K. Shen, C. Ong, X. Li, "Feature Selection for MLP Neural Network: The use of Random Permutation of Probabilistic Outputs," IEEE Transactionon Neural Networks, vol. 20, issue 12, pp. 1911-1922, October 2009.

[13] Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Adv Neural Inf Proces Syst 25:1–9 21. Kumar S (2016). A study for handwritten Devanagari wordrecognition. Proceedings of the international conference on communication and signal processing (ICCSP), 1009–1014.

[14] Kumar M, Sharma RK, Jindal MK, Jindal SR (2018) Character and numeral recognition for non-Indic and Indic scripts: a survey. Artif Intell Rev52:2235–2261

[15] Lee H, Grosse R, Ranganath R and Ng AY(2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. Proceedings of the 26th Annual International Conference on Machine Learning, 609-616.

[16] Narang SR, Jindal MK, Kumar M (2019) Devanagari ancient documents recognition using statistical featureextraction techniques. Sādhanā 44(6):0141.

[17] Narang SR, Jindal MK, Ahuja S, Kumar M(2020a) On the recognition of Devanagari ancient handwritten characters using SIFT and Gabor features. Soft Comput 24:17279–17289.

[18] Narang SR, Jindal MK, Kumar M (2020b)Ancient text recognition: a review. Artif Intell Rev 53:5517– 5558.

[19] Narang SR, Jindal MK and Sharma P (2018) Devanagari ancient character recognition using HOG and DCT features. Proceedings of the 5th IEEEInternational Conference on Parallel,Distributed and Grid Computing (PDGC-2018), Solan, India.

[20] Phan TV, Nguyen KC, Nakagawa M (2016) A nom historical document recognition system for digital archiving. Int J Document Anal Recognit 19(1):49–64.