iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VIII **Month of publication:** August 2020

DOI: https://doi.org/10.22214/ijraset.2020.30815

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VIII Aug 2020- Available at www.ijraset.com

Recognition of Handwritten Characters using Deep Convolutional Neural Network

S. G. K. Abhishek[1], Sangeeta Adike[2], Varun Totakura[3], Padma Rajani[4]
[1, 2, 3]*UG Scholar,* [4]*Assistant Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, R.R. District, Hyderabad, India*

Abstract: ***One of the most researched topics with wide implementations in different fields is Handwritten Character Recognition (HCR). To utilize this technology in automated data-entry applications, this is one of the fields on which past and recent works have been done to focus on a variety of languages. Individual characters in a word are recognized through a set of images, which are then studied by the Deep Neural Network. It is with the help of this recognition that results are ranked. This ranking is performed based on the client's request. A Convolutional Deep Neural Network Model is considered to acknowledge the handwritten characters in this paper. It learns, using local receptive areas and heavily connected neural network layers, a useful set of ordination which are required to generate an optimum result.***

Keywords: Handwritten Character Recognition, Deep Neural Network (DNN), Deep Convolutional Neural Network (DCNN).

I. INTRODUCTION

Handwritten Character Recognition is a field of research in image contrast, computer vision and style recognition. An average computer will gain the ability to make a distinction between the characters on photographs, paper records, touch screen gadgets from different resources and then convert into machine-readable notations. This helps in a wide variety of computer applications dealing with optical character reception and framework development for character handling [1]. One of the notable issues with HCR is Picture Rating. It is the process of sending input pictures to analyse from an invariable arrangement of surrounding reliant pictures. In Optical Character Recognition (OCR), calculations are conducted on a data collection consisting of recognized characters and interprets how to group these characters that are found from the test set [2].

Formerly, a set of calculations were devised for categorizing letters or digits. Digit Recognition was the task that rounded up many of the complex computations that were done in this field [3]. Notably, format coordination and fundamental calculations were highly used during the onset of OCR. In these calculations, some tests of letters and numbers are averaged to create models for acknowledgement. In a significant measure of trials, the results generated were sparse as these computations for OCR were easy to digest such that distinctive types would generate similar results.

Since the late 1980's, Neural Networks have been used because larger datasets and arrangement strategies were being deployed [4]. A large fraction of these frameworks nowadays utilize neural network systems for manually handwritten character recognition (HCR). Neural Systems are learning techniques linked to character recognition in Machine Learning. The motivation behind neural networks is to imitate the learning functions that occur in living beings like humans' neural network. Being a standout amongst the most path breaking learning models, neural networks are mostly used where there is a need for mechanization of activities where the goals cannot be achieved by an individual alone or are not precise in nature. In such scenarios, neural systems can be swift at finding solutions and may also uncover connections between noted instances of information that us humans cannot view [5].

A Neural Network can be deployed such as a Deep Neural Network (DNN), which makes use of in surplus of a hidden layer. The distinction between neural and deep neural networks is the number or depth of the hidden layers used in the network. DNN can be a feed forward neural network which consists of more than one hidden layers as depicted in Fig. 1 [6].

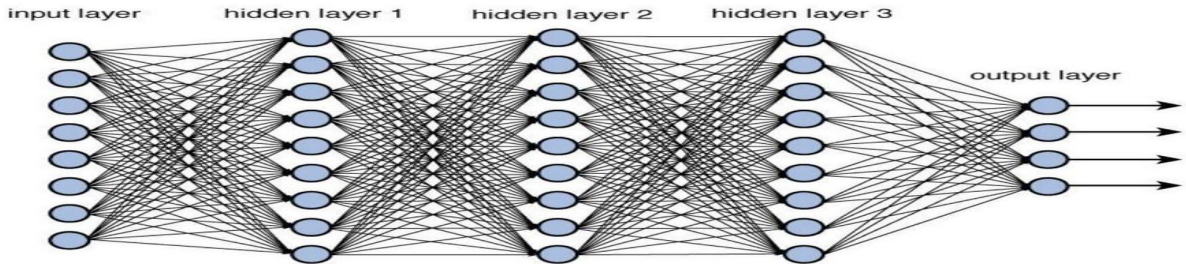


Fig. 1: DNN Architecture [6]

A DNN comprises an input layer, output layer and many intermediary hidden layers. Therefore, the amount of connections and adaptable components are very huge. The deep neural networks require significant collection of examples to impede over fitting and several regular signs have type-setting arrangements [7]. In images, closeby arrangements of edges make up the frame themes, themes collect into components and components mould subjects. Corresponding improvements exist from audio in addition to phonemes, syllables, words and sentences. The collection enables a synopsis of the data to transfer to next layers when data in the prior layers change in position and form. One section of the Deep Neural System which is overall simple to prepare and consists of smaller arrangements is the Convolutional Neural Network (CNN) [8][9]. It is a natural divulgence of multilayer perceptron (MLP). Fukushima proposed for a multilayer neural network system [10], and has also taken handwritten character recognition and other computer vision issues into consideration. LeCun [11], has applied the Convolutional Neural Network system to classify the ImageNet dataset. Present progress on CNN has been concentrated on Computer Vision issues such as picture division [12], picture inscribing [13], and picture grouping [14]. There has also been a keen interest about manually written character digits [15] and identification of characters in different languages and dialects. This concern has grown due to the reasonable dilemma and likeness of written characters by hand and liberality in classes. In this review, hand-written characters will be examined and character recognition is checked with the help of DNN at first, and then with CNN techniques.

II. ALTERNATIVE TECHNIQUES

Researchers have developed many approaches for handwritten character recognition systems. Significant and essential systems are indicated in this work. Character recognition frameworks have been engineered utilizing different rationale [16]. The frameworks that are created by some research can be built by using hardware with VLSI (Very Large Scale Integration) circuitry. The input character recognition of this framework is resilient to dynamic motion. Other studies and investigations made use of hamming error correcting codes from communication theory with neural network systems in their framework. Another technique was developed to acknowledge the written hand characters in different dialects in its' Neural network System [17]. Even though the framework has generated efficient and accurate results, some variances were also observed if the handwritten characters are in intense format. One of the researchers has even offered a strategy to relate the dependence between hand writers and their penmanship [18]. In their methods the researchers have mostly used the Multi-layer feed forward neural network system.

III.REVIEWED TECHNIQUES

Convolutional Neural Network approach is fundamentally a neural network structure that makes use of convolution in place of basic network systems with an identical number of layers. It has comprehensive applications in image and video acknowledgement, handling of characteristic dialect and suggested frameworks. From a design perspective, CNN is a neural network that employs a minimum of one convolution operation in at least one of its layers. This convolution process is performed in the convolution layer. The convolution layer consists of three primary procedures. These are - Convolution, Sub-sampling or pooling, and actuation. A CNN embodies a load of convolution layers and a peak pooling layer followed by a whole actuation layer. The convolution task, which is the most essential layer of the system, is performed by the convolution layer. The sampling layer is next. This layer is paramount in cases where the neural network has to deal with larger images, the size of trainable components can be increasing in an exponential manner. This, consequently increases the time for neural network system preparation and this is not suitable for practical use. The scale of the image is reduced with the help of the pooling layer. In this study, the Modified National Institute of Standards and Technology (MNIST) database released by the United States Department of Commerce is used. This database consists of an enormous number of images of handwritten characters. Reduction in the size of the images can thus reduce the overall time taken to ready the neural network system to function. A broad view of layers in a CNN is shown in Fig. 2.

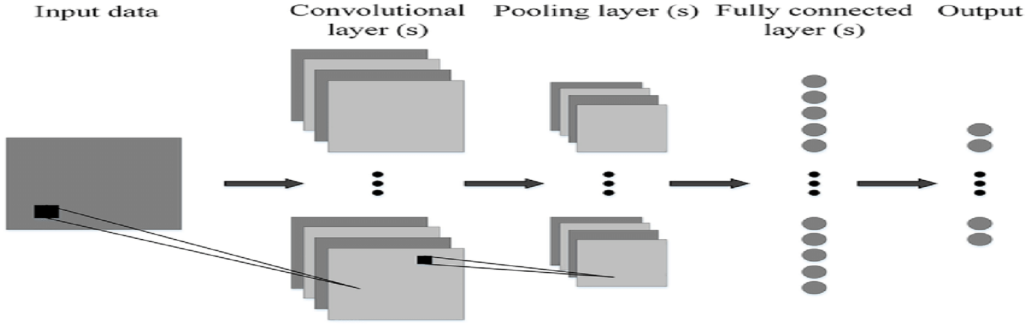


Fig. 2: CNN Architecture [19]

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VIII Aug 2020- Available at www.ijraset.com

In this study, the Modified National Institute of Standards and Technology (MNIST) database released by the United States Department of Commerce is used. This database consists of an enormous number of images of handwritten characters. Reduction in the size of the images can thus reduce the overall time taken to ready the neural network system to function. A broad view of layers in a CNN is shown in Fig. 2.

1) *Input Layer:* The input layer is the layer for the primary image which adds up to the operation at the start of the system architecture. This input is the character image as shown in Fig 3. This input image can be a Grayscale or Color (RGB) image. Dimensions WxHxD can be present for the input layer where WxH is the Width and Height of the image and depth of the image is represented by D. Depth for grayscale and RGB images are 1 and 3 respectively. Therefore, the input layer as demonstrated in Fig. 3a has dimensions of 32x32x3 for an RGB image, while 32x32x1 is the dimensions for a grayscale image as illustrated in Fig. 3b

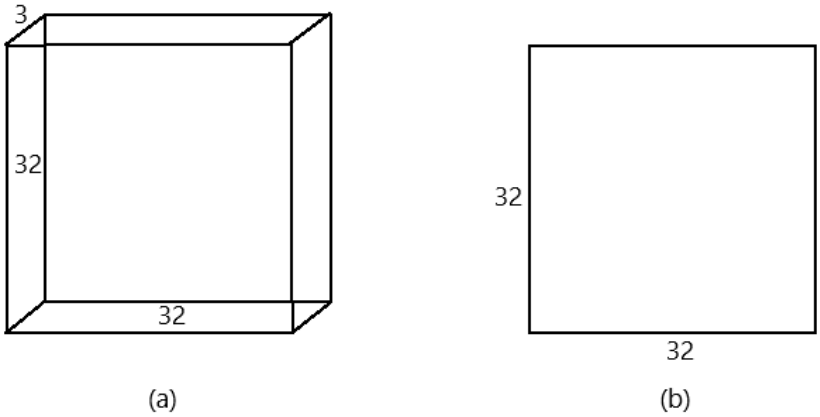


Fig. 3: Input Images

2) *Convolution Layer:* This is the layer where most of the computation is done. The Convolutional Layer consists of channels with learning abilities and called elements of this layer. Each layer is detected by a square matrix of dimensional width, length and depth in pixels. This square matrix is called a filter or a channel. These filters cover the full data volume. A typical channel can be 5 pixels wide and long, and 3 pixels deep (i.e. 5x5x3 pixels). These filters are designated as shaded channels and use RGB images. Filters will have a depth of 1 pixel unit and a size of 5x5x1 in this study. The character images used are grayscale images (non-color). Throughout the forward propagation of the neural network, each channel is checked for its width and length and compared with other channels in the layer to create a 2-D information measure. The intensity of pixels' information of the shapes of characters are taken into consideration while the other areas are taken as 0 pixels in each channel of the Convolution Layer. As each filter is filtered through a channel sample which is of a specific width and length of the information measure, a 2-dimensional fractional character online is supplied from each channel. This can give the image's local position as the response of that channel.

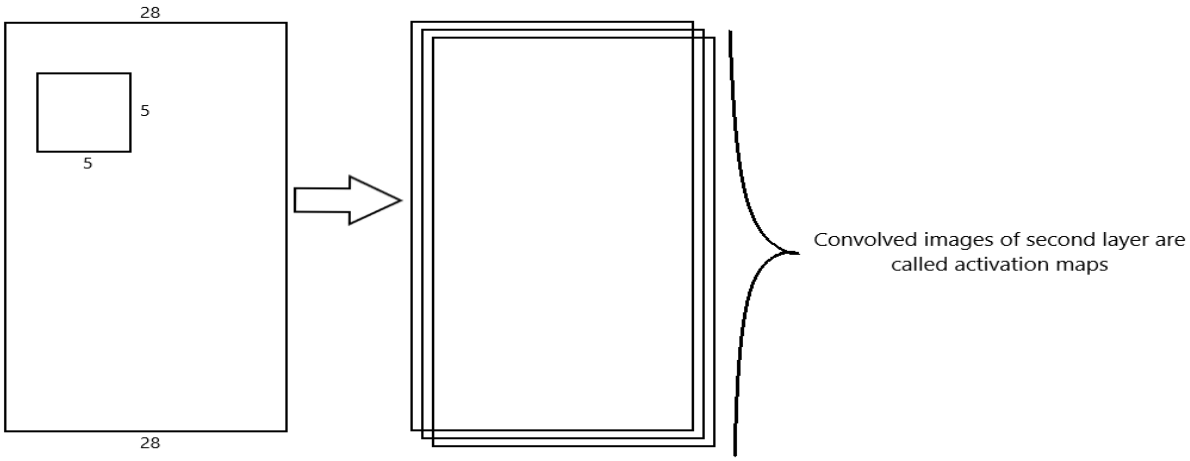


Fig. 4: Initial Maps from convolution layer

Inherently, the filter section is wrapped over the entire image and the produced output is called initial map or activation maps as shown in Fig. 4. These are generated from the 2nd Convolutional Layer. The testing conditions decide the size and amount of filters to be used, as there is no clearly defined approach to determine them. Firstly, filters may contain minor amounts of noisy values. But these values can be updated with each learning phase of the network. The input layer accepts an image of a certain dimension, i.e. Width(W) x Height(H) x Depth(D). During the Convolution operation two more hyper parameters namely, Filter Size (F) and Stride (S) are issued for the purpose of generating an input for the upcoming layers with W1xH1xD1 as dimensions. The formulas for the calculation W1 and H1 are expressed in Equations (i) and (ii). The depth remains identical as D. P is called the Padding which introduces rows and columns of zeros for each aspect of the image.

$$W\_1 = (F-W+P)/(S+1) \quad (1)$$

$$H\_1 = (F-H+P)/(S+1) \quad (2)$$

32 filters were used in this study, each 5x5x1 in size, and values of P and S were taken as 1 and 0 respectively. Therefore, the measurements of the resultant second layer image were 32(28x28).

3) Pooling Layer: The Pooling layers are situated in between the Convolutional Layers in the CNN Architecture. The purpose of the pooling layers is to compress the components quantity when the images are very large. Moreover, Elimination of unused elements in each image is introduced through down inspecting or sub testing, in order to conserve critical data. Pooling Layer works openly on the information section at every depth and changes this information at every spatial dimension. Maximum Pooling (MAX), Normal Pooling and Whole Pooling are some of the various types of spatial pooling, with MAX pooling being the most commonly used. MAX pooling as a pooling layer with each filter of 2x2 dimension with a phase of 2 sub poolings. Each of these Maximum pooling tasks has to select a maximum value out of 4 numbers. Depth-dimension here stays constant. In this study, a maximum size of 3x3 P = 1 and S = 2 are sent to the first pooling layer. 32 (14x14) was the output dimension for this layer. In this layer, several other filters as previously mentioned can also be used.

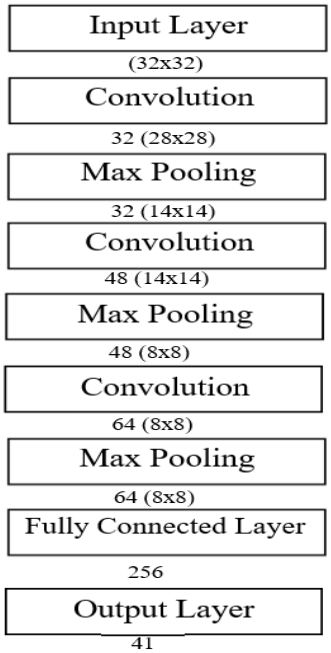


Fig. 5: CNN Architecture used for the Experiment

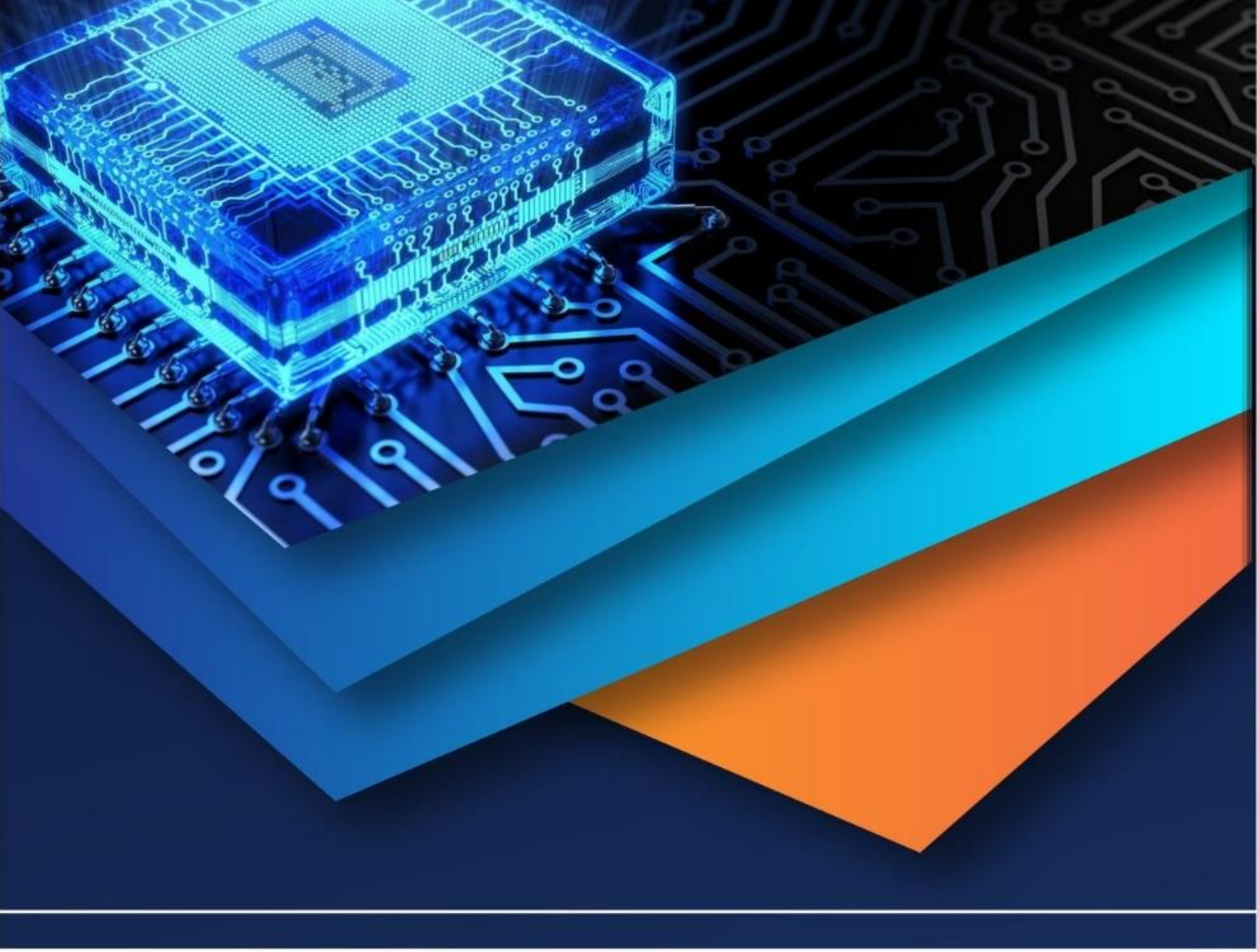
4) Fully Connected Layer: This is the last layer of the neural network called as the FC Layer. The final image matrix output from the pooling layer 2 with width and height W and H respectively are converted into a 1-dimensional vector form and employed into the system. The implementation of this layer can hence improve the expansion of the framework followed by a capability to balance. A huge number of fully connected layers can be present based upon the application architecture. An output layer with 41 neurons was generated from this study as it was expected that the total number of character classes were 41. The complete linked layer has 256 neurons. The number assigned for these neurons is experimentally determined. The Matlab Neural Network Toolbox is used in the experiments. The whole architecture used in this study is displayed in Fig. 5

IV. CONCLUSIONS

A convolutional neural system is assessed for readers' attention. This system is exceedingly dynamic for discerning handwritten characters. The work depends on the acceptance of the characters as CNN inputs. Compared to other deep learning models, CNN has desirable execution in both big data and images. The main intention to utilize deep learning was to edge the strength of CNN that are capable of managing large dimensions of input and sharing their weights. There are thousands of elements and Hyper elements to tune in CNN architecture. It is still unclear why convolutional networks are flourishing while general back-propagation algorithms fail. This might be due to the fact that convolutional networks utilize a hierarchy and provide a solution to a complex framework as compared to simple ones.

REFERENCES

[1] R. Vaidya, D. Trivedi, S. Satra, M. Pimpale, "Handwritten Character Recognition Using Deep-Learning". Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 772-775, 2018
[2] G. S. Budhi and R. Adipranata, "Handwritten Javanese Character Recognition Using Several Artificial Neural Network Methods", J.ICT Res. Appl., vol. 8, no. 3, pp. 195–212, 2015
[3] A. Rajavelu, M.T. Musavi, and M.V. Shirvaikar, "A neural network approach to character recognition", Neural Netw., vol. 2, no. 5, pp. 387– 393, 1989.
[4] S. Mori, C. Y. Suen, and K. Yamamoto, "Historical review of OCR research and development," Proc. IEEE, vol. 80, no. 7, pp. 1029 –1058, 1992
[5] J. Pradeep, E. Srinivasan and S. Himavathi. "Neural Network based Handwritten Character Recognition system without feature extraction", International Conference on Computer, Communication and Electrical Technology ICCCET 2011
[6] Ravindra Parmar, "Training Deep Neural Networks", Towards Data Science, Sept. 2018 https://towardsdatascience.com/training-deep-neural-networks-9fdb1964b964
[7] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning", Nature, Vol. 521, pp. 436-444, 2015
[8] Y. Liang, J. Wang, S. Zhou, Y. Gong, and N. Zheng, "Incorporating image priors with deep convolutional neural networks for image super resolution", Neurocomputing, Vol. 194, pp. 340- 347, 2016
[9] R. Nijhawan, H. Sharma, H. Sahni, and A. Batra, "A deep learning hybrid CNN framework approach for vegetation cover mapping using deep features", 13th International Conference on Signal Image Technology & Internet-Based Systems (SITIS), pp. 192-196, 2017
[10] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biological Cybernetics., vol. 36, no. 4, pp. 193–202, 1980
[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in neural information processing systems, pp. 1097–1105, 2012
[12] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling", IEEE Trans. Pattern Anal. Mach. Intel., Vol. 35, no. 8, pp. 1915–1929, 2013
[13] O. Vinyals, A. Toshev, S. Bengio, and D. Ethan, "Show and tell: A neural image caption generator", Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, pp. 3156–3164, 2015
[14] D. C. Ciresan, U. Meier, J. Masci, L. Maria Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification", Proceedings in 22nd International JointConference on Artificial Intelligence, Vol. 22, pp. 1237- 1242, 2011
[15] E. Kussul and T. Baidyk, "Improved method of handwritten digit recognition tested on MNIST database", Image Vis. Compute., vol. 22, no. 12, pp. 971–981, 2004
[16] W. Lu, Z. Li, B. Shi." Handwritten Digits Recognition with Neural Networks and Fuzzy Logic", IEEE International Conference on Neural Networks, Vol. 3, pp.1389-1392, 1995
[17] P. Banumathi, G. M. Nasira, "Handwritten Tamil Character Recognition using Artificial Neural Networks", International Conference on Process Automation, Control and Computing, 2011
[18] B. V. S. Murthy, "Handwriting Recognition Using Supervised Neural Networks", International Joint Conference on Neural Networks, 1999
[19] Yi Wang, Zhice Fang, Haoyuan Hong, "Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China", Science of the Total Environment, 2019
[20] Varun Totakura, Mohana Krishna Janmanchi, Rajesh Durganath, M. I. Thariq Hussan (2020), "Prediction of Animal Vocal Emotions Using Convolutional Neural Network", International Journal of Scientific & Technology Research, VOLUME 9, pp. 6007-6011.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY