Review on Deep Learning Handwritten Digit Recognition using Convolutional Neural Network

Akanksha Gupta, Ravindra Pratap Narwaria, Madhav Singh

Abstract: ***In this digital world, everything including documents, notes is kept in digital form. The requirement of converting these digital documents into processed information is in demand. This process is called as Handwritten digit recognition (HDR). The digital scan document is processed and classified to identify the hand written words into digital text so that it can be used to keep it in the documents format means in computerized font so that everybody can read it properly. In this paper, it is discussed that classifiers like KNN, SVM, CNN are used for HDR. These classifiers are trained with some predefined dataset and then used to process any digital scan document into computer document format. The scanned document is passed through four different stages for recognition where image is pre-processed, segmented and then recognized by classifier. MNIST dataset is used for training purpose. Complete CNN classifier is discussed in this paper. It is found that CNN is very accurate for HDR but still there is a scope to improve the performance in terms of accuracy, complexity and timing.***

Keywords: Handwritten digit recognition (HDR),Deep learning, Convolutional Neural Networking (CNN),Artificial Neural Network (ANN), Object Character Recognition (OCR), Modified National Institute of Standards and Technology (MNIST), SVM (Support Vector Machine),KNN (K nearest Neighbors), Rectified Linear Unit(ReLU), NN (Neural Networks).

I. INTRODUCTION

Convolutional Neural Networks (CNNs) is very well known deep learning algorithm which can be used to process image. It assigns weights and biases to various parts of the image and very capable of differentiating one image from another same kind of image. Good accuracy has been achieved for handwritten digits recognition by using Convolutional Neural Networks. Mammalian system of visualization is taken into consideration to create CNN architecture. CNN is created by D. H. Hubel in 1962 [1]. Two algorithms with name gradient descent & back propagation are utilized to train the model. Character images of handwritten digits are used as an input. Artificial neural network (ANN) consists of one input layer, one output layer and some layers which exist in between input layer and output layer, these middle layers are hidden layers. CNN and ANN are very similar to each other.CNN deep learning algorithm worked on analysis of visual images. CNN can be used in applications like detection of object, identification of face, in the field of robotics, video processing, segmentation, in the field of pattern recognition, processing of natural language, detection of spam, categorization, speech identification, classification of digital image etc[4].

Manuscript received on January 19, 2021.
Revised Manuscript received on January 23, 2021.
Manuscript published on January 30, 2021.

Akanksha Gupta, M.E student from Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India
Dr. Ravindra Pratap Narwaria, Assistant Professor in MITS, Gwalior, Madhya Pradesh, India
Prof. Madhav Singh, Assistant Professor MITS, Gwalior, Madhya Pradesh, India

Object Character Recognition (OCR) is used on printed or documented letters to convert them into text. Digital documents can be created by extracting and storing information using OCR scanning tools.OCR can be implemented by pattern recognition and through segmentation method. Handwritten digit recognition (HDR) is similar to OCR, just in place of complete object image HDR recognize digits. HDR is light and faster than OCR. HDR plays an important role in medical documentation, banking documentation, student record keeping, taxation[2]. Many methods has been used for HDR like Neuro-Fuzzy Systems (NFS), Artificial Neural Network (ANN), Support Vector Machine (SVM) and deep learning-based classifiers [3]. All the classifiers are providing good accuracy but still there is a lot to explore in the field of HDR to further improve the performance. The performance parameters used to find the performance of classifiers are accuracy, running time and computational complexity. The important factor in CNN model is that it fully use the topological information as well as it is invariant to basic transformations like rotation, translation etc. English Handwriting datasets like Modified National Institute of Standards and Technology (MNIST) are used to calculate the performance of HDR technique [8].

II. HDR MODELS

The popular models which are used for HDR are KNN (K nearest Neighbors), SVM (Support Vector Machine), NN (Neural Networks).

A. KNN (K nearest Neighbors)

KNN is used to solve regression problems and also used as a classifier. In KNN classifier, since computations are calculates up to the end stage that's why it is also called as late learning classification algorithm. And since all the computation occur locally; it is also called as instance-based classification algorithms. There is no training required earlier in KNN classifier, as well as there is no generalization is performed on training data. KNN algorithm describes categorical value by making use of majority of votes of K - nearest neighbors, the K value used to differ here. It is found that votes value changes with change in K value as shown in Fig. 1.

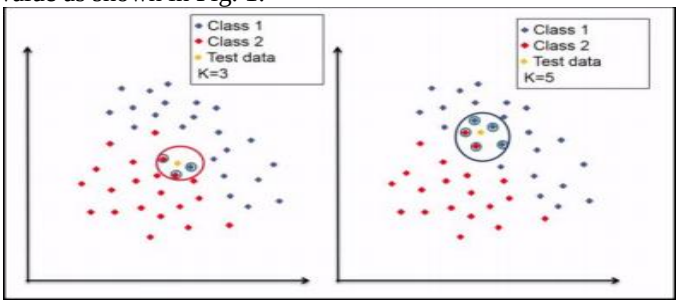

Fig.1. KNN [1]

B. SVM (Support Vector Machine)

Support Vector Machine is a kind of supervised learning. It can be used for both regression problems and for classification purpose. SVM make use of optimal hyper plane that can be utilized to divide it into multiple categories. For 2D spaces, independent variable data points are plotted which are corresponds to dependent variables [3]. After it, classification is started to find hyper plane or linear/nonlinear plane which is used to categorize class as shown in Fig. 2.

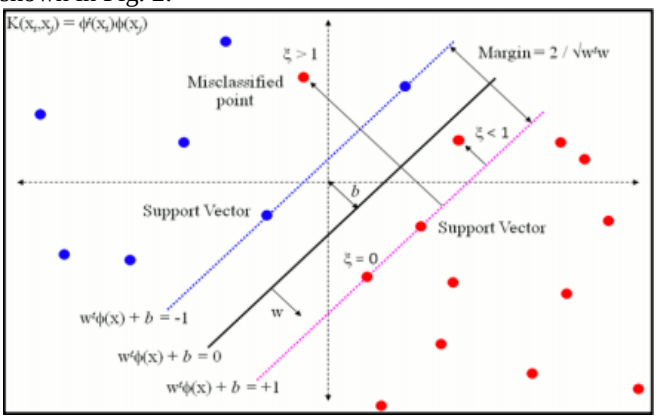

Fig.2. SVM [1]

C. NN (Neural Networks)

The Neural Networks methodology is inspired from working of brain. It became popular in the field of computational power. NN is termed as Deep learning where multilayers are connected together to form a network. Nodes are formed by using these layers. Each node is subjected to execute some computation. This then input into node's activation function, in order to show context signal progress into the network for classification purpose [7].

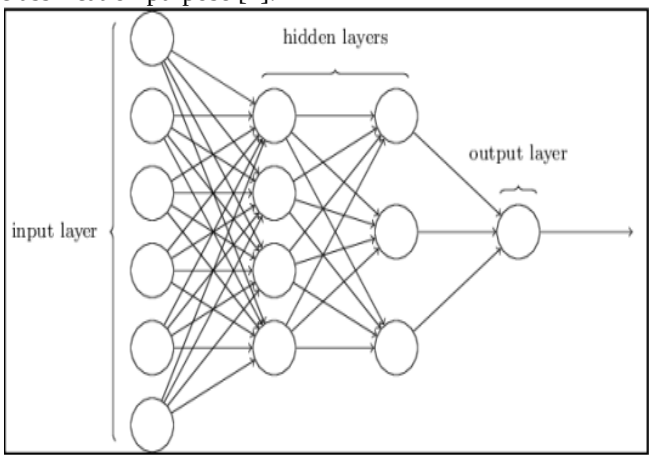

Fig.3. NN [1]

III. MNIST DATASET FOR HDR

The standard database used for HDR is MNIST database (Modified National Institute of Standards and Technology database). It contains images of handwritten digits which are very frequently used to train classifiers for image processing applications. Machine learning algorithms also use this data base very frequently. This database has training image database of 60,000 images and testing image database of 10,000 images. Some of them are shown in Fig. 4 below. Where 50% of the images are taken from NIST's database.Many researches have been done on it to achieve good accuracy [9].

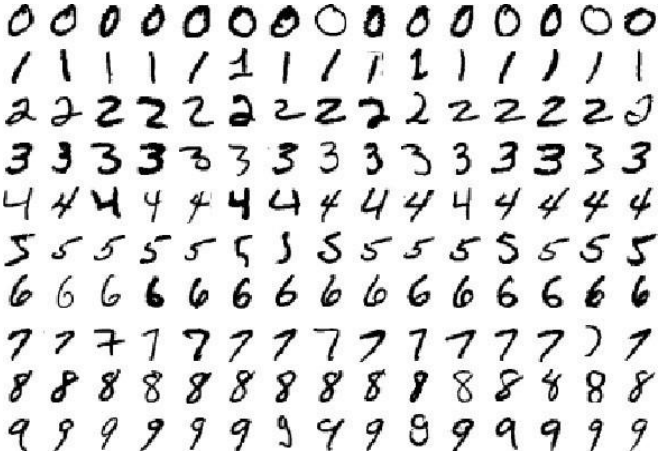

Fig.4. some images of MNIST

IV. CONVOLUTIONAL NEURAL NETWORKING (CNN)

The CNN classifier is very popular to perform HDR. CNN is a 7 layer convolutional network where there is one input layer then five layers which are hidden and at last one output layer.The size of input layer is 28 by 28 pixel that is 784 neurons can be sent to its input. All the input images are greyscale in nature where intensity value vary from 0 to 255, here 0 represents black and 1 represents white.

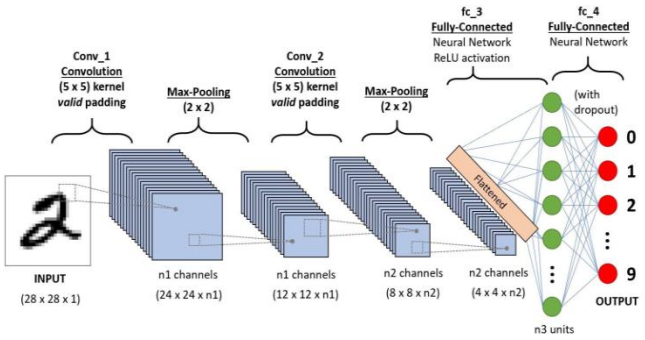

Fig.5. CNN steps

CNN is kind of ANN with feed forward method. Connectivity in CNN is inspired by the organization of the animal visual cortex. CNN has many neurons that carry two parameters which are learnable weights & biases. Some input is provided to each neuron and it performs operations like dot product and performs it with non-linearity.

A. Layers of CNN

There are many layers in CNN. Deep leaning is a use of all these layers in iteration. The 3 category of CNN layers can be explained as:

1) Input Layer: the raw pixel values of digital image are carried by input layer.
2) Convolutional Layer: A result of input block neuron layer is supplied into convolution layer. There are many filters defines for this layer by user. The filter with window size of 5x5 is used for input pixels and it gives out the highest intensity pixels in output.

3) Rectified Linear Unit Layer: the function of this layer to pass image pixel into activation function element wise. Back propagation is applied in CNN, this leads to change in values of image pixel, in order to rectify it ReLU function is applied.
4) Pooling Layer: down sampling is performed on spatial dimensions like width and height by this layer. The output comes in the form of volume.
5) Fully Connected Layer: computation of score of classes is performed at this layer. It computes maximum of score secured by input digits[6].

V. STAGES FOR HDR

There are following stages to perform HDR using CNN. It has 4 stages as given below.

1) Pre-Processing: in this stage, multiple operations are performed on input image. This stage is to make image ready for segmentation. The operations like noise filtering, standardization and image smoothening is performed on the input image. All the images like gray scale or colored image is first converted into binary image. This operation of binarization is performed to limit the quantity of data by performing threshold operation on it. The images will look like as shown in Fig. 4 MNIST dataset.
2) Segmentation: after the completion of pre-processing process, some sub-images are created from the original image data. Now this pre-processed image is divided into small sub images where each sub-image represents a single digit. All the singular digit images are resized into pixels. Edge detection method is applied to perform segmentation process.
3) Feature Extraction: at the end of segmentation process, each output image is shown as a separate matrix where each value represents a single pixel. These matrix representing individual single digit is an easiest way of representation for further stages. This operation of representation each sub-image into the form of matrix is termed as feature extraction stage. This stage is used to remove the data redundancy from the image data.
4) Classification & Recognition: in this stage, the formed matrix is taken as input to classifiers to recognize the digit present in the image. The output of extraction is input into classifiers like KNN, CNN, SVM. These classifiers used trained data to find the possible digit present in the image [5].

VI. CONCLUSION

Handwritten digit recognition has immense applications in the field of medical, banking, student management, and taxation process etc. Many classifiers like KNN, SVM, CNN are used to identify the digit from the handwritten image. as per the review, CNN is providing better performance than others. Stages of HDR using CNN classifier is discussed in this paper. MNIST dataset consist of handwritten numbers from 0-9 and it is a standard dataset used to find performance of classifiers. HDR consists of three different stages. First is preprocessing where dataset is converted into binary form and image processing has been applied on it. Second stage is segmentation where the image is converted into multiple segments. Third stage is feature extraction where features of image are identified. Last stage is classification where classifiers like KNN, SVM, CNN are used. Results of HDR is improved a lot by using CNN classifier but it can be improved further in terms of complexity, duration of execution and accuracy of results by making combination of classifiers or using some additional algorithm with it.

REFERENCES

1. Mayank Singh, Rahul (2020). Handwritten Digit Recognition using Machine Learning. International Research Journal of Engineering and Technology, 7(7), 921-925.
2. Yellapragada SS Bharadwaj, Rajaram P, Sriram V.P, Sudhakar S, KollaBhanuPrakash (2020, April). Effective Handwritten Digit Recognition using Deep Convolution Neural Network. International Journal of Advanced Trends in Computer Science and Engineering, 9(2), 1335-1339.
3. SavitaAhlawata, AmitChoudhary (2019). Hybrid CNN-SVM Classifier for Handwritten Digit Recognition. Elsevier B.V, International Conference on Computational Intelligence and Data Science, Procedia Computer Science 167, 2554-2560.
4. FathmaSiddique, ShadmanSakib, Md. Abu BakrSiddique (2019). Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for Various Hidden Layers. 5th International Conference on Advances in Electrical Engineering (ICAEE).
5. Vijayalaxmi R Rudraswamimath, Bhavanishankar K (2019, June). Handwritten Digit Recognition using CNN. International Journal of Innovative Science and Research Technology, 4(6), 182-187.
6. AnujDutt, AashiDutt (2017, July). Handwritten Digit Recognition Using Deep Learning", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 6(7), 990-997.
7. M. M. A. Ghosh and A. Y. Maghari (2017). A Comparative Study on Handwriting Digit Recognition Using Neural Networks. International Conference on Promising Electronic Technologies.
8. Haider A. Alwzwazy, Hayder M. Albehadili, Younes S. Alwan, Naz E. Islam (2016, February). Handwritten Digit Recognition Using Convolutional Neural Networks. International Journal of Innovative Research in Computer and Communication Engineering, 4(2), 1101-1106.
9. Fabien Lauer, Ching Y. Suen, G´erard Bloch (2007). A trainable feature extractor for handwritten digit recognition. Journal Pattern Recognition, Elsevier, 40(6), 1816-1824.

AUTHORS PROFILE

Akanksha Gupta currently a M.E student from Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India. She has obtained B.E degree from ITM Group of Institute, Gwalior, Madhya Pradesh, India. Her current research interest includes Image Processing.

Dr.Ravindra Pratap Narwaria currently working as an Assistant Professor in MITS, Gwalior, Madhya Pradesh, India. He has received his PHD degree from Banasthali Vidhyapeeth, Rajasthan, India and M.E degree from RGPV, Bhopal, India. He has published 11 Conference papers and 57 Journal papers. His research interest includes Instrumentation Control and Bio-signal Processing.

Prof.Madhav Singh currently working as an Assistant Professor in MITS, Gwalior, Madhya Pradesh, India. He has currently pursuing his PHD degree from RGPV, Bhopal, Madhya Pradesh, India and also done his M.tech and B.E degree from RGPV, Bhopal, Madhya Pradesh, India. He has published 15 Journal papers. His research interest includes VLSI Design, Adhoc Network and Wireless Communication.