International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

CNN is Preferred over Other Methods for Handwritten Digit Recognition System

Akriti Singh

Galgotias University, Greater Noida, Uttar Pradesh, India

Abstract: ***Handwritten digit recognition is an important problem in today's world scenario because there are millions and millions of people across the globe with millions of different handwriting styles that could be a trouble for recognition by a human being because there are digits that are quite confusing. This paper presents an approach to offline handwritten digit recognition based on CNN and why It is preferred over other methods. For this we are using the mnist data sets for training our data and testing the outcomes to obtain the desired result to settle to the final decision that CNN is a better method than others. The only problem lies in developing an efficient algorithm for recognizing handwritten digits mainly the ones that are a little bit confusing. This method can be very useful for the future like the monitoring of the vehicles that break the traffic rules, and in banks for detecting fraud cheques.***

Keywords: Pattern recognition, handwritten digit recognition, CNN

1. Introduction

What is a Handwritten Digit Recognition System?

In this the machine trains itself to recognize human written digits from various sources like emails, bank cheques, etc. in real world scenarios [1]. The inputs are then taken and fed in a model where they are processed and converted in an array. The main problem lies in developing an efficient algorithm for recognizing handwritten digits which are submitted by the users. There are more than millions of people in the world and each individual has their own way of writing digits which can either be understood or can even make you confuse, in order to reduce the complexity of understanding digits this system can be of help to people and this system can be implemented in many different ways like KNN, CNN, Machine learning, OCR and the svm. [2] It will not only save time but also increase the efficiency of the work that is to be performed. Here we will be focusing on the Convolutional neural network approach and why is it better. The very first reason that comes to my mind is variety of image feature that we will get through the convolution layer cannot be obtained from anywhere else which can help in the best way to train the data. Using CNN for handwritten digit recognition has been so advanced that it has nearly reached a human perfection level. And with the passing time CNN is being used for many other things like video analyzing, robotics and pattern recognition [3, 4]. This pattern recognition model that is used in this system was introduced by Fukushima in 1980 [5,6]. And the CNN model framework that we will be working on was designed in 1998 by Lecun et a l [7]. in the model that we are using for the handwritten digit recognition the digits can be classified directly from its image pixels [8].

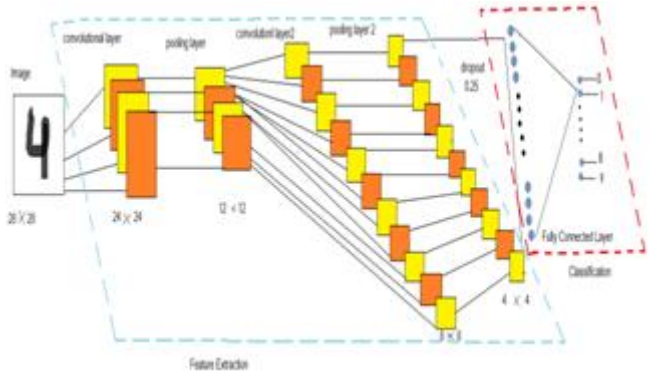
The algorithms that will be used for the processing in this system are back propagation and gradient descent [9]. The dataset that we are using for training and testing our data is the mnist data datasets which consist of a total of 70,000 datat sets [10]. And that dataset comprises of approx 10,000 testing dataset and 60,000 training dataset [11]. the digits that are used in this dataset are3 already normalized and are represented by an array size of 28*28. And the total pixels of the array is 784 whose values range from 0 to 255, where if the pixel value is 1 it means the background colour of the image is white and if it is 0 then the background colour is black [12]. Here the data that I am using has pixel 0 that is why the background colour is black of the images.

Talking about the other methods that has been earlier used for this recognition were SVM that was used by parveen kumar, Arun rana and Nitin Sharma where they used different kinds of svm classifiers like polynomial kernel, linear kernel etc. While using these 25 features were extracted from every character which were then used for training the svm which resulted in 94. 8 percent accuracy [13]. Later in 2002 Lipika Dey along with Reena Bajaj and Santanu Chaudhary decide to present a systematic built method for the recognition of handwritten numbers for which they used the connections network [14]. Zen and Ming in 2010 did a comparison to find out which is the optimum way to recognize handwritten numbers where after the comparison they concluded that KNN has the least error rate [15]. so that is the optimum way to perform this recognition. Now the method that I have used is CNN whose architecture is almost same to ANN but inn CNN the neurons are connected to the local field rather than being fully connected [9]. CNN used here even shows very less error rate even to that what the knn showed in that comparison done by Zen and Ming [15]. As the error rate that I achieved was approx 0. 02 percent which is nearly to perfect hence I think CNN is the best way to perform the Handwritten digit Recognition and should be preferred over other methods. [16] Even in this research the accuracy that they achieved was 98. 4 percent using CNN which was best among the other two methods KNN and SVM.

2. CNN-Convolutional Neural Network

CNN is a neural network that consists of connected layers of neuron used for classification and fault detection [20]. A neuron consist of a number known as activation and the connections are known as weights which help the network to determine which neuron is to be activated in which layer and what should be the activation popularly known as the feed forward process. [18] This process continues until the output

neurons get activated. The whole CNN process can be divided into two parts the first one is the feature extraction part and the second one is the classification part. The feature extraction is done by combining the two layers; the Convolutional layer and the pooling layer whereas the classification part is performed by the dense layers the image prediction is handled by the input layer and the output layer focuses on the different classes that we are trying to predict.



1) **Convolution layer**- this layer basically combines a filter with the prior layer, and the number of filter combined can range [17] from one to many because when the number of the filter increases in the layer the features of the image become more distinctive and you know what the best part is we don't even have to select which filter is to be used as that work is done by the neural network itself as they even learn the features themselves while training the data. And we select the kernel size which helps in determining the locality of the filter and choose the stride value which only determines how many pixels we need to advance while combining our filter to the layer.
2) **Pooling layer**-this layer is applied to down-sample the inputs for reducing the computational complexity of the model and for avoiding the over fitting problem, the most common technique used is max pooling.
3) **Dense layer**-this layer randomly exclude the number of neurons in the present layer in order to reduce the over fitting problem and then later converts the 2d matrix data to a vector called flatten.

Handwritten Digit Recognition Data

The dataset used are the MNIST datasets where each image is a 28×28 pixel grey scale square image. [1,7] There are 60,000 training datasets and 10,000 testing data sets. For example-

Calculating the accuracy of the trained model using CNN

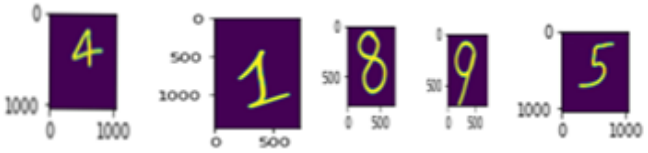
```
Train on 60000 samples, validate on 10000 samples
Epoch 1/10
 - 118s - loss: 0.1695 - accuracy: 0.9464 - val_loss: 0.0381 - val_accuracy: 0.9879
Epoch 2/10
 - 137s - loss: 0.0435 - accuracy: 0.9865 - val_loss: 0.0275 - val_accuracy: 0.9914
Epoch 3/10
 - 117s - loss: 0.0301 - accuracy: 0.9903 - val_loss: 0.0244 - val_accuracy: 0.9921
Epoch 4/10
 - 119s - loss: 0.0232 - accuracy: 0.9929 - val_loss: 0.0271 - val_accuracy: 0.9911
Epoch 5/10
 - 125s - loss: 0.0177 - accuracy: 0.9943 - val_loss: 0.0268 - val_accuracy: 0.9923
Epoch 6/10
 - 134s - loss: 0.0161 - accuracy: 0.9948 - val_loss: 0.0233 - val_accuracy: 0.9932
Epoch 7/10
 - 145s - loss: 0.0128 - accuracy: 0.9958 - val_loss: 0.0262 - val_accuracy: 0.9929
Epoch 8/10
 - 130s - loss: 0.0113 - accuracy: 0.9965 - val_loss: 0.0260 - val_accuracy: 0.9923
Epoch 9/10
 - 126s - loss: 0.0095 - accuracy: 0.9970 - val_loss: 0.0257 - val_accuracy: 0.9932
Epoch 10/10
 - 129s - loss: 0.0095 - accuracy: 0.9970 - val_loss: 0.0253 - val_accuracy: 0.9928

 712/10000 [=>.....] - ETA: 1s
 808/10000 [=>.....] - ETA: 1s
 904/10000 [=>.....] - ETA: 1s
 000/10000 [=>.....] - ETA: 1s
 096/10000 [=>.....] - ETA: 1s
 192/10000 [=>.....] - ETA: 1s
 288/10000 [=>.....] - ETA: 1s
 384/10000 [=>.....] - ETA: 1s
 480/10000 [=>.....] - ETA: 0s
 544/10000 [=>.....] - ETA: 0s
 640/10000 [=>.....] - ETA: 0s
 736/10000 [=>.....] - ETA: 0s
 832/10000 [=>.....] - ETA: 0s
 928/10000 [=>.....] - ETA: 0s
 624/10000 [=>.....] - ETA: 0s
 120/10000 [=>.....] - ETA: 0s
 248/10000 [=>.....] - ETA: 0s
 344/10000 [=>.....] - ETA: 0s
 440/10000 [=>.....] - ETA: 0s
 536/10000 [=>.....] - ETA: 0s
 632/10000 [=>.....] - ETA: 0s
 728/10000 [=>.....] - ETA: 0s
 824/10000 [=>.....] - ETA: 0s
 920/10000 [=>.....] - ETA: 0s
 000/10000 [==============================] - 6s 626us/step
st loss: 0.025287489897097476
st accuracy: 0.9927999973297119
```

As seen the accuracy and the loss calculated of the handwritten digit recognition model using CNN is approx 99. 28 and 0. 02 simultaneously.

Creating training data

First we download the datasets from Keras by using from keras. datasets import mnist, and then we use the sequential model for the model building which is prebuilt in Keras and then the dense layers are imported which will be later used for predicting the labels and then dropout layer is added that will reduce the over fitting problem. After that the flatten layer is added that converts the 3d array to 1d. After doing all this we have finally imported all the layers the Convolutional, the pooling layer and the numpy. Some class variables are assigned as we have 10 classes and also because of which the batch size becomes 128. After that we load the data and reshape it using the. Reshape (60000,28,28,1) function where 60000 is the number of images data is being trained on and all of this is done because CNN accepts only 4-d vector. After all of this is done we prepare our test data to test upon the trained data.

After training the data the result is-



On testing the test data we get the correct output-

```
(10, 28, 28, 1)
4
1
8
9
5
```


3. Conclusion

Using CNN we not only get the correct output but also the accuracy is more as compared to others and even the loss is approximately 0. 02 %. If the same procedure is done without using CNN then the error rate obtained is 2. 68% which is more hence CNN helps in giving the optimum result then other methods with less error rate. And this is why because while using CNN we get more distinct image features for training our datasets using the Convolutional layer than we can obtain through any other method. That is why CNN should be preferred over other methods for the digit recognition.

References

[1] Wan Zhu, Classification of MNIST Handwritten Digit Database using Neural Network.
[2] Norhidayu binti Abdul Hamid, Nilam Nur Binti Amir SjarifHandwritten Recognition Using SVM, KNN and Neural Network, Advance Informatics School Universiti Teknologi Malaysia.
[3] Y. LeCun et al. , "Backpropagation applied to handwritten zip code recognition," Neural computation, vol. 1, no. 4, pp. 541- 551, 1989
[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.
[5] D. Cireşan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," arXiv preprint arXiv:1202. 2745, 2012.
[6] K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," in Competition and cooperation in neural nets: Springer, 1982, pp. 267-285.
[7] Y. LeCun et al. , "Handwritten digit recognition with a backpropagation network," in Advances in neural information processing systems, 1990, pp. 396-404.
[8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradientbased learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998
[9] Fathma Siddique1 , Shadman Sakib2 , Md. Abu Bakr Siddique2, Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for Various Hidden Layers
[10] Li Deng, "The MNIST Database of Handwritten Digits images for Machine Learning Research", MIT Press, November 2012.
[11] Y. LeCun, “The MNIST database of handwritten digits,”
[12] AnujDutt, AashiDutt, “Handwritten Digit Recognition Using Deep Learning”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 6, Issue 7, July 2017, ISSN: 2278 – 1323.
[13] Parveen Kumar, Nitin Sharma, and Arun Rana. Article: Handwritten Character Recognition using Different Kernel-based SVM Classifier and MLP Neural Network (A COMPARISON). International Journal of Computer Applications 53(11):25-31, September 2012.
[14] Reena Bajaj, LipikaDey, and S. Chaudhury, “Devnagari numerals recognition by combining decision of multiple connectionist classifiers”, Sadhana, col. 27, part I pp-59-72, 2002.
[15] Wu, Ming & Zhang, Zhen. (2019). Handwritten Digit Classification using the MNIST Data Set.
[16] Meer Zohra, D. Rajeswara Rao, A Comprehensive Data Analysis on Handwritten Digit Recognition using Machine Learning Approach
[17] Hinton, Geoffrey E. , Zoubin Ghahramani, and Yee Whye Teh. ”Learning to Parse Images. ” neural information processing systems (2000): 463-469.
[18] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. , Imagenet large scale visual recognition challenge, International Journal of Conflict and Violence (IJCV) 115 (3) (2015) 211–252.
[19] Saad Albawi ; Tareq Abed Mohammed ; Saad Al-Zawi, Understanding of a convolutional neural network
[20] K. B. Lee, S. Cheon, and C. O. Kim, "A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes," IEEE Transactions on Semiconductor Manufacturing, vol. 30, no. 2, pp. 135-142, 2017