International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-1, November 2019

Recognition of Handwritten Digits using Convolutional Neural Network and Linear Binary Pattern

Prashanth Kambli, Amruthalakshmi, Naresh E

Abstract: ***Over the past few years there has been a tremendous developments observed in the field of computer technology and artificial intelligence, especially the use of machine learning concepts in Research and Industries. The human effort can be further more reduced in recognition, learning, predicting and many other areas using machine learning and deep learning. Any information which has been handwritten documents consisting of digits in digital form like images, recognizing such digits is a challenging task.***

The proposed system can recognize any handwritten digits in the document which has been converted into digital format. The proposed model includes Convolutional Neural Network (CNN), a deep learning approach with Linear Binary Pattern (LBP) used for feature extraction. In order to classify more effectively we also have used Support Vector Machine to recognize mere similar digits like 1 and 7, 5 and 6 and many others. The proposed system CNN and LBP is implemented on python language; also the system is tested with different images of handwritten digits taken from MNIST dataset. By using proposed model we could able to achieve 98.74% accuracy in predicting the digits in image format.

Keywords **: *CNN, LBP, RBF, SVM***

I. INTRODUCTION

Artificial Intelligence (AI) plays an indispensable role in current technologies. The godfather of AI is John McCarthy, defines AI as "the science and engineering of making machines" [1]. Machine Learning (ML) is the subset of Artificial Intelligence. All AI are does not counts as ML, but all ML is considered as AI. Machine Learning is dynamic and without human intervention it has ability to make some changes when exposed to more data [2]. Deep Learning (DL) is a subset of ML and they are referred to deep artificial neural network, they have set of new records in accuracy for many major problems like image recognition, sound recognition and text recognition [3].

In the past, before the existence of computers the information was stored in the paper was not much efficient, as it may have lost or destroyed after some time. But in modern technology we can store the information safely in the computer for long time as well as we can have multiple copies of it in the cloud storage [4]. Anytime anywhere we can retrieve the information. Instead of hiring the labor for converting this information in the paper into digitalized information, machine learning technology can be used for conversion without human intervention [5].

Handwritten digit dataset are not much clear in nature because the digits are not always perfectly straight lines and sharp [6]. Feature extraction process will help to remove this redundancy from data. Also digits may be written in different sizes and orientation so that the proposed system can overcome this limitation [7].

Handwritten Digit Recognition is a challenging task because of the difference in writing style from person to person. There are two important steps to develop a reliable handwritten digit recognition system [8]. Step one helps to extract the discriminative feature from the handwritten image [9]. Step two helps to classification, when a new digit image arrives for testing. The classification should take less training time and must be able to classify digit with high accuracy [10].

II. LITERATURE SURVEY

Leo Pauly et al. [11] proposed a system which could able to recognize digits of South Indian languages. They were used histogram oriented gradient (HOG) and artificial neural network (ANN). Initially they gathered image and then removed noise in the image, segmented those images and extracted the discriminative features from the image, got an accuracy up to 83.4%.

R Kurniawan Nur et al.[12] used the handwritten digit recognition system in the general election commission of Indonesia for c1 form to their election purpose which face issues in recognizing hand written digits. For extracting the feature they have used Local Binary Pattern and for classification K Nearest Neighbor (KNN) such that the experiment have achieved the accuracy of 89.65% in recognizing the images.

Michio Yasuda el at.[13] proposed a system in which they used pre-trained Convolutional Neural Network (CNN) and support vector machine(SVM), to recognize the handwritten digits.

Revised Manuscript Received on November 30, 2019.
***** Correspondence Author
Prashanth Kambli*, Department of Information Science and Engineering, Ramaiah Institute of Technology, Bengaluru, India. Email: prash.kambli@msrit.edu
Amruthalakshmi, Department of Information Science and Engineering, Ramaiah Institute of Technology, Bengaluru, India. Email: amrithalakshmibs@gmail.com
Naresh E, Department of information Science and Engineering, Ramaiah Institute of Technology, Bengaluru, India. Email: nareshkumar.e@gmail.com

The result of this work measured with distortion and without distortion and they got error rate as 0.93% and 1.03% respectively.

Dr. Hak KeungLam et al.[14] presented a work with Ensemble Technique. So Ensemble Neural Network and Ensemble Decision Tree combination is used to recognize the handwritten digit and it yields 84% accuracy. Here the goal of ensemble learning is not achieved.

One of the applications of handwritten digit recognition is in processing the Arabic cheque. Challenges faced by the Arabic handwritten digits is to identify the stoke direction, which is done using Gabor filter and also Extreme Learning Machine (ELM) and Sequential minimal optimization (SOM) helps in classification [15].

Muhammad Usman Ghani et al.[16] worked on the handwritten digit recognition concept by using Hidden Markov Model(HMM) an statistical markov model to classify the digits based on their extracted features of the handwritten digits images. Baum welch algorithm is used as trainer for the HMM parameter.

Prashanth Kambli et al. [17] survey work on related topic concludes that it is better to use Convolutional Neural Network, ecause by comparison with different methods as in the Figure 1in that CNN able to achieve highest accuracy.

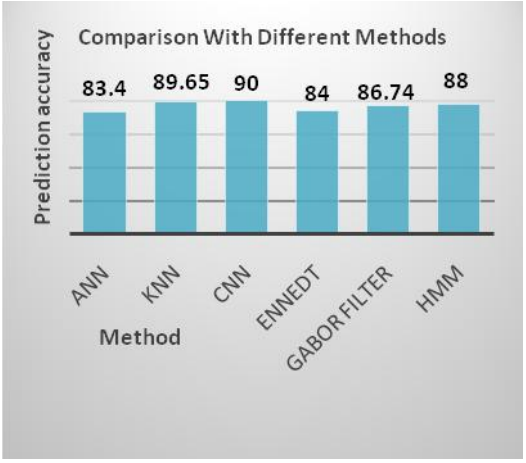

Fig 1: Comparison with different methods

III. METHADOLOGY

Basically it includes two models one for extracting the features and another for classification. For feature extraction we are using Convolutional Neural Network and Linear Binary Pattern. And SVM using RBF kernel is used as classifier. In this session we discuss more about these methods.

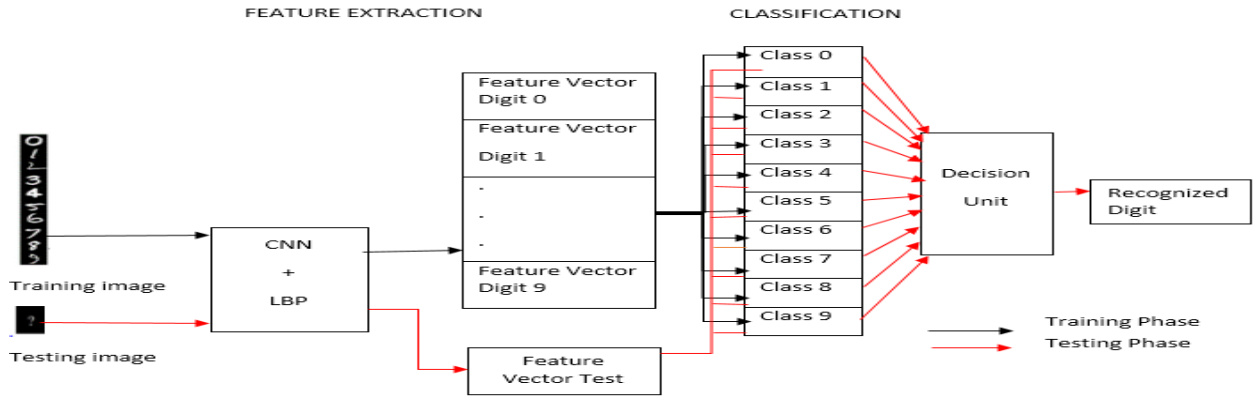

Fig 2: Proposed model

Feature Extracter

For feature extraction we use the combination of both CNN and LBP to get higher accuracy. With reference to our survey paper "A Survey: Handwritten Digit Recognition Approaches" we concluded that CNN is the best method for feature extraction. Convolutional Neural Network which is type of Deep Neural network, it works similar to barin of human. into the brain in different layers.

Take an example of handwritten digit image of number 6 and it is considered as input image and it goes through several layers. It takes a patch of from the input, further it will send the patch to the convolution layer and applies set of filters. Apply group of filters to those images. That activated data goes into further layer named pooling layer, it is a nonlinear down sampling layer where the image is reduce to half the size. The same patch will go into the pooling layer. This process keeps on with Convolutional and pooling layers [18] where it undergoes with bunch of activation filters in convolution layer. The information from the image is extracted and further down sampled it. Finally at the end of the network the processed data is connected altogether in the layer called Fully Connected Layer.

Different layers of CNN are Convolutional Layer, Pooling Layer, Rectified Linear Unit (ReLU) Layer and Fully Connected Layer [19]. The detected features from the filters are then convolved and dot product is calculated in activation layer. Next comes is pooling layer, which is basically used for down-sampling which chooses max pooling [20]. For dimensionality reductions use the max pooling. It will help to overcome the over-fitting problems by providing an abstract form of the representation.

The nest layer ReLU refers to Rectified Linear Unit layer. ReLU layer has ability to train a huge data is faster by not making significant difference to accuracy. The two keywords in this layer are stride and padding. Around the input volume the stride will control the filter convolution process by shifting one unit at a time. So stride refers to number of filter shifted. Padding could take the value like valid (no padding), casual (dilated convolution) and same (output has same length as original input).The last layer of CNN is fully connected layer, where layers are inter connected to its previous layer activation function.

Retrieval Number: A5045119119/2019©BEIESP
DOI: 10.35940/ijitee.A5045.119119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

After applying CNN, for betterment in predicting digits the proposed method suggests to apply the Linear Binary Pattern (LBP). By binary derivation, Linear Binary Pattern will describe the neighborhood pixel. Later, recognizing the pattern the LBP is used. To form a short code to describe the pixel neighborhood these binary derivatives are used. We can use it in many research areas like pattern recognition, text analysis, etc. Linear Binary Pattern works very simple, by taking a pixel surrounded by 3 by 3 neighboring pixel. Then the center pixel value is compared with neighboring pixel. If the neighbor of the center pixel has larger value than the center pixel will generate a binary as 1. Otherwise it will generate binary as 0. Later the eight neighbor of that center can be represented with an 8-bit number.

Feature Classifier

For feature classification SVM using RBF kernel is used. If you search for classification algorithm, you might get lots of methods like Naïve Bays, K means, SVM, Apriori, ANN and many more. But our challenge is to choose the suitable classifier and it should be more efficient. Based on my survey, support vector machine suits well for my work. Since it gives high accuracy, well theoretical guarantee for over-fitting and if data are not linearly separable then it can be solved using RBF kernel concept. In machine learning RBF refers to Radial Basis Function, it is one of the well-known kernel function used in many kernalized learning algorithm. Usually we use linear SVM, where the data are linear the hyper-plane can easily separate the classes. Where as in case of nonlinear data RBF kernel can be used to separate the classes. The advantages of using RBF kernel are

i. Comparatively polynomial kernel RBF kernel has less hyper parameters which will reduce the complexity of model selection.
ii. Also it has fewer nonlinear difficulties

Because of these reasons we prefer SVM using RBF kernel method for classification of features.

MNIST Dataset

The MNIST database (Modified National Institute of Standard and Technology database) is actually a huge dataset of images and these images represents digits that might be drawn or handwritten by other people. The size of each image is 28 by 28 pixel which includes six thousand training dataset, and one thousand testing dataset. The dataset includes ten different classes for each digits from 0 to 9. Figure 3 and Figure 4 are the samples of MNIST dataset.

Fig 3: Sample1

Fig 4: Sample2

IV. IMPLEMENTATION

A. CNN Algorithm

CNN is constructed with combination of 4 layers namely

- Convolution
- ReLU (Activation layer)
- Pooling
- Fully connected layer

Convolution has four steps:-

Step 1:- With every image pixel corresponding feature pixel is multiplied.

Step 2:- By sum up all the values find the summation.

Step 3:- The sum is then divided by number of pixel in the feature.

Step 4:- The final value obtained is placed at the center of the filtered image.

$$Convolution = \frac{\sum_{p=0}^{n} I_p \times F_p}{n} \qquad equation\ (1)$$

Where,

I_p is Image pixel

F_p is Feature pixel

n is total number of pixel

Rectified Linear Unit (ReLU) has two steps:-

This activation function is used to remove all the negative values from the convolution layer. Linear relationship between dependent and independent variable is as shown in equation (4).

Step 1:- Take the input and check is it above zero

Step 2:- if no output is zero else output is same as in put

$$ReLU = f(y) = \begin{cases} 0 \ if \ y < 0 \\ y \ if \ y \geq 0 \end{cases} \qquad equation\ (2)$$

Pooling layer has four steps:-

Step 1:- choosing window size

Step 2:- choosing a stride

Step 3:- walk the window across filtered images

Step 4:- By every window, choose the maximum value

$$Pooling = max\_pool(V, W, S) \qquad equation\ (3)$$

Where,

V is value

W is window

S is stride

Fully connected layer has two step:-

Step 1:- flattening, which converts pooled feature matrix into linear vector

Step 2:- Apply softmax() to normalize the output

$$Flatten = con(2D, linear\ vector) \qquad equation\ (4)$$

$$FC = softmax(flatten) \qquad equation\ (5)$$

B. Working of CNN

The data from the MNIST dataset is taken as the input for the CNN feature extraction process. Set of filters are applied in convolution layer, then max pooling is done in pooling layer, after that activation function named ReLU is applied. Later we flatten the data means converting 2D vector to 1D array. Then we apply Fully-connected which will have 1000 output neurons. One more time we apply ReLU andfully connected layers. After that we will get CNN result.

C. Working of LBP

From input database MNIST a pixel point (C),it is compared with its 3 cross 3 neighborhood (N). Such that if C>=N then it will give binary 1 otherwise binary 0. The resulted matrix is then multiplied with mask matrix and gives out LBP result.

D. LBP Algorithm

Step 1:- In the image I, for every pixel (m, n) choose N adjacent pixels at the radius of R.
Step 2:- The intensity difference between the present pixels C with N adjacent pixels is calculated.
Step 3:- The threshold is calculated such that, if current pixel C greater than or equal to neighboring pixel N then 0 is assigned to all the negative differences otherwise all assigned with 1.
Step 4:- Convert the N-bit vector to its respective decimal value and replace the intensity value at (m, n) with this decimal value.
The equation (1) gives the LBP descriptor for every pixel, C and N represents current/center and neighboring pixel respectively. First we need to find the threshold T as in the equation (2), which is then multiplied with 2^n where n is 0, 1, 2…

$$LBP(M,N) = \sum T(C,N)\, X\ 2^n \qquad equation\ (6)$$

Where, n = 0, 1, 2….

$$T[C,N] = \begin{cases} N = 0 & if\ C \geq N \\ N = 1 & Otherwise \end{cases} \qquad equation\ (7)$$

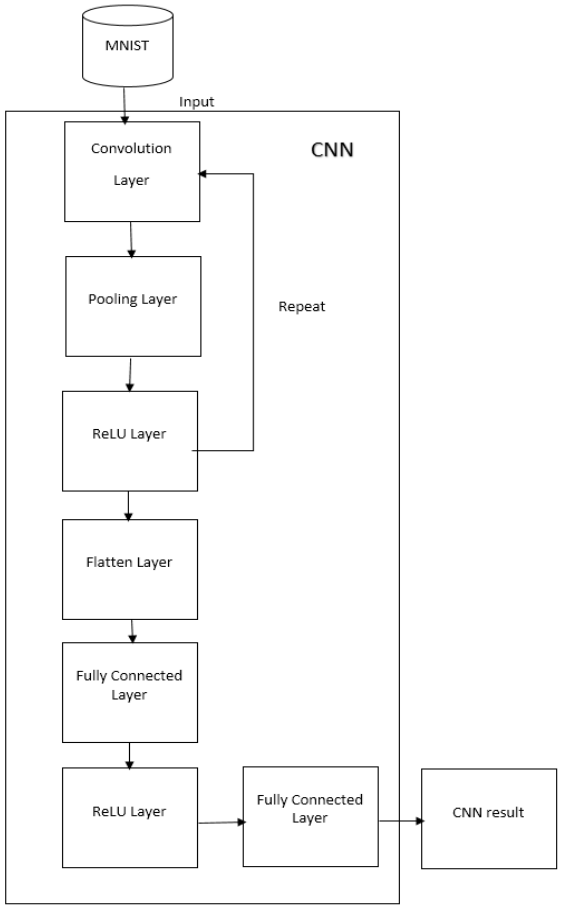

Fig 5: CNN working model

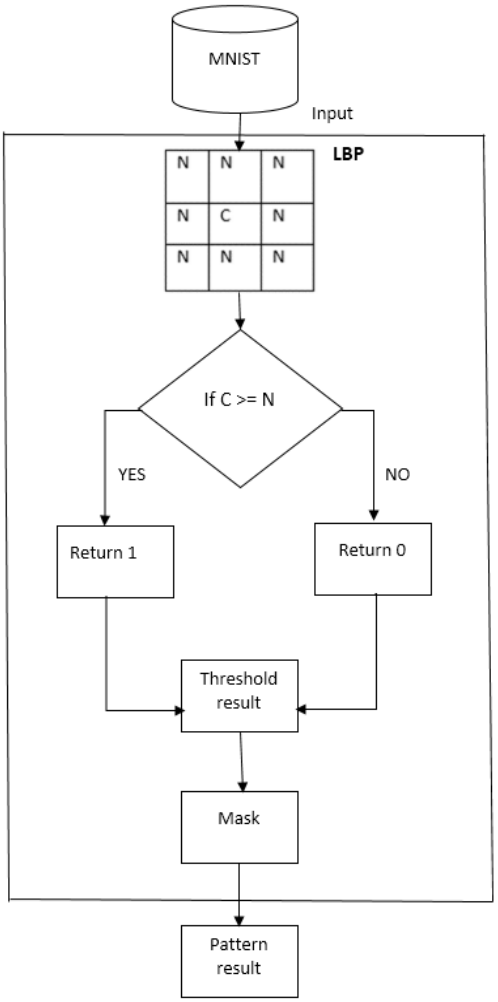

Fig 6: LBP working model

V. RESULTS

The Table 1 shows the result of training accuracy of 55000 training data from MNIST dataset which has been trained in ten iteration.

TABLE 1: Training result

Iteration	Training accuracy (in %)	Time taken (in sec)
1	60.12	80
2	77.18	70
3	78.66	72
4	87.77	70
5	89.09	61
6	95.74	62
7	96.34	62
8	96.75	62
9	97.08	62
10	97.35	63

For testing the 10000 dataset from same MNIST database is taken which gave results of accuracy 98.74% by SVM using RBF kernel classifier. The experimental goal has been successfully achieved improving the prediction and classification of handwritten digits.

IV. CONCLUSION AND FUTURE IMPROVEMENT

The paper Recognition of Handwritten Digits using CNN and LBP is presented. According to the proposed model Convolutional Neural Network and Linear Binary Pattern helps in extracting the discriminative features from the input image of MNIST dataset. To separate the classes based on feature of nonlinear data Radial Bias Function of Support Vector Machine is used. And this system achieved objective of the project in prediction and classification of the handwritten digit in image format with an accuracy of 98.74%. So combination of CNN with LBP helps in improvising the learning capability of the machines.

The proposed system could able to recognize the handwritten digit only, but in the future it can be extended to recognize the handwritten character as well as symbols.

REFERENCES

1. Hemant Kumar, Rishabh, "Comprehensive Review on Supervised Machine Learning Algorithms", 2017 International Conference on Machine Learning and Data Science (MLDS), 2017.
2. Vanita Jain, Arun Dubey ,Amit Gupta, " Comparative analysis of machine learning algorithms in OCR", 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016.
3. R.R. Malyan , S. Sunthankar, "Machine learning for hand printed character perception", IEE Colloquium on Machine Learning, 2014.
4. Sonia Kuwelkar, Mithili N Mayekar, "Implementation of machine learning algorithm for character recognition on GPU", International Conference on Computing Methodologies and Communication (ICCMC), 2017.
5. Priya,Rajendra Singh,Soni Cha Priya,Rajendra Singh,Soni Changlani, "Recognition of Handwritten Digits using Proximal Support Vector Machine", Journal of Emerging Technologies and Innovative Research (JETIR), Volume 4, Issue 04, 2017.
6. Anish Kumar Jha1 , Anupam Ghosh1, "Hand Written Digit Recognition using Convolutional Neural Network (CNN)", International Journal of Innovations & Advancement in Computer Science, Volume 6, Issue 5, 2018.
7. "Handwritten Bangla Digit Recognition Using Chemical Reaction Optimization", 9th ICCCNT ,2018.
8. Bappy Shahriar Abir, H.M. Mehedi Hasan, "Handwritten Bangla Digit Recognition Using Chemical Reaction Optimization", 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018.
9. Shruti R. Kulkarni, Bipin Rajendran,", 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2017.
10. Amitava Choudhury, Tanmay Bhowmik, "Handwritten Bengali Numeral Recognition using HOG Based Feature Extraction Algorithm", 5th International Conference on Signal Processing and Integrated Networks (SPIN), 2018.
11. Leo Pauly, Dr.Binu paul, R,"Hand written Digit Recognition System for South Indian Languages using Artificial Neural Network", *th International Conference on Contemporary Computing, 2015.
12. R Kurniawan Nur, Nurul ,"Handwriting Digit Recognition using Local Binary Pattern Variance and K-Nearest Neighbor Classification", 4th International Conference on Information and Communication Technology , 2016.
13. Michio Yasuda, Yumi Nakashima, "Pattern Augmentation for Handwritten Digit Classification based on Combination of Pre-trained CNN and SVM", 2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT), 2017.
14. Dr Hak KeungLam, Retno Larasati, "Handwritten Digit Recognition using Ensemble Neural Network and Ensemble Decision Tree", International Conference on Smart Cities, Automation & Intelligent Computing System Yogyakarta, Indonesia, 2017.
15. Hanadi Hassen ; Ali Hamdi,"Arabic Bank Cheque Words Recognition Using Gabor Features", 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR), 2018.
16. Muhammad Usman Ghani, Syed, "Handwritten Digit Recognition using DCT and HMMs", 12th International Conference on Frontiers of Information Technology,2014.
17. Prashanth Kambli, Amruthalakshmi "Investigating the Various Approaches towards Handwritten Digit Recognition ", 5th National Conference on Advancement In Information Technology, 2019, Journal of Network Security Computer Networks.
18. Ratan Kumar Basak, Shalini Guha, Saptarshi De, Sucheta Nag, "Hindu Arabic character recognition using mathematical morphology", IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2016.
19. Nikolaos Toulgaridis, Theodore Antonakopoulos, "RBM-based hardware accelerator for handwritten digits recognition", Panhellenic Conference on Electronics and Telecommunications (PACET), 2017.
20. Po-Yuan Shih,Wei-Bin Liang," Integration of Orthogonal Feature Detectors in Prameter Learningof Artificial Neural Networks to Improve robustness and the Evaluation on Hand-Written Digit Recognition Tasks", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016.

AUTHORS PROFILE

Prashanth Kambli is a M.Sc. (Engineering) by research degree holder and is serving as an assistant professor in IS&E department of MSRIT. Membership of professional societies: ISTE, IEEE, ACM.

Amruthalakshmi is completed her bachelor degree in the stream of ISE and now pursuing her Master degree in Software Engineering.

Naresh E is a M.Tech degree holder and is serving as an assistant professor in IS&E department of MSRIT. Membership of professional societies: ACM, ISTE, IAE.