International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-6, April 2019

# An Efficient Offline Hand Written Character Recognition using CNN and Xgboost

**Joseph JamesS , C.Lakshmi, UdayKiran P, Parthiban**

***Abstract:*** ***The purpose of this paper is to legitimize and implement the usage of Convolutional neural networks (CNN) in parallel with XGBoost model to improve handwriting Recognition systems. The usage of CNNs in recognizing handwritten characters is a broadly researched project yet the inclusion of different types of classification models along with CNN is sparse. The learning model proposed in this paper is based on (CNN) as a feature extraction tool and XGBoost as an accurate prediction model. The XGBoost gradient boosting model is evaluated for loss function and regularization and an appropriate objective function is decided. With the proposed method in which CNN and XGBoost are used together there is an expected increase in accuracy rate and total computation time. The model is trained and evaluated using the NIST special database 19 dataset which consists of 810,000 isolated character images including lower case, upper case and digits in the english language. The improvement in accuracy is in comparison with the handwriting recognition model which uses CNN alone and is augmented with the use of tree ensembles model which is XGBoost. The improved accuracy percentages are specified separately for lowercase letters, uppercase letters and numeral characters.***

***Index Terms:- XGboost-Extreme Gradient Boosting, CNN-Convolutional Neural Networks, NIST- National Institution of Standards and Technology.***

## I. INTRODUCTION

Hand writing recognition is a prominent field under the image recognition application's umbrella. With many existing applications already improving, handwriting recognition hasn't been research to its depths. There is a even further scope for improvement in the field of handwriting recognition methods. The demanding sub-field of this topic is offline handwriting recognition where already handwritten scripts and documents are scanned and perceived by the system. Unlike the online handwriting recognition where additional inputs such as stroke speed, stylus on touch time and tip pressure analysis augment the recognition, offline handwriting recognition systems rely purely on the captured image and its pixels. The input format also hugely impacts the system's efficiency. Some of the common problems faced by the system are complexity of the handwritten characters, distorted input images, highly disconnected characters, prominent but highly ambiguous characters and language specific training of the system.

**Revised Manuscript Received on April 07, 2019**.

**Joseph James S,** Software Engineering, SRMIST, Kancheepuram, India.
**Dr. C. Lakshmi,** Software Engineering, SRMIST, Kancheepuram, India.
**Uday Kiran P**, Software Engineering, SRMIST, Kancheepuram, India.
**Parthiban**, Software Engineering, SRMIST, Kancheepuram, India.

As the system involves object recognition, the use of best feature extraction tools should be the subject of study. Along with a feature extraction tool a more efficient classification tool has to be used so as to improve the overall efficiency of the system. There are two schools of thoughts where on one side the integrated usage of classifier system and feature extraction system is carried out and on the other side these two systems are used separately. Typically feature extraction tools

use variety of parameters to properly capture image information in the appropriate format and at the same time apply constraints to expedite the situation. After this step the classifier is manually trained using the previously extracted data in a particular format. However, the use of integrated handwriting systems promise a betterment in efficiency and ease of training.

CNN has been modified in many ways for handwriting recognition systems. The major changes occur in the architecture of CNN where the number of neurons in the hidden layer are increased or decreased to match the target efficiency. Also the number of hidden layers that take in input from the max-pooling and the convolution layers are modified.

Some research increase the complexity of the system in order to improve the efficiency.

In [1] propose a sparse CNNs or called DeepCNet for online handwriting character recognition. The pictures are enhanced with signature information, and thesparsity is used to increase the depth of network. It makes aslow speed that allows for the retention of more spatialinformation. Similarly other research have base on fewer phases for better results. The most popular classifiers that are usually used with CNN in place of the conventional CNN classification layer are SVM, Gradient Boosting and logistic Regression. In [2] the system Involves four major steps Data acquisition, preprocess (input image to binary image), feature extraction,classification (SVM classifier is used). Hence to broaden the spectrum of the applicability of the system better classifiers can be used.

Elleuch [3] has conducted research which slightly different from the approaches previously mentioned. In [3] combines a CNN with another end classifier. The kernel SVM approaches using as an end classifier for handwritten Arabic character recognition.

This research deals with the usage of Extreme gradient Boosting, which uses weak tree ensemble model for gradient boosting, along with CNN as a strong feature extraction tool to classify and recognize the 26 distinct English letters both uppercase and lowercase. The model is first tested for base validation with the MNIST digits dataset for handwriting recognition.

## II. LITERATURE SURVEY

There are many works related to handwriting recognition that follow different methodology other than CNN. The majormethod progress and advantages have been discussed below.Anuja et al [7] developed an offline English handwritten word recognizer. The work aimed at extracting features from handwritten words and comparing it with database. It eliminates theadditional splitting of word into characters or sub words. This method is faster than older methods.Yang et al [8] developed an improved DCNN for recognition of Chinese handwriting. Their work focuses on non Linear Normalization and Domain Specific knowledge contribution to DCNN. They found that HSP-DCNN significantly outperforms than other widely used ensemble strategies, .indicating a relative error rate reduction of 42%.The work of Gupta et al [9] shows that their proposed CR system could recognize 21 word images successfully out of the 26 word images that were used.Fuzzy theory has been used for recognition of Hand-Printed English Character in [10]. Fuzzy theory is simpler than neural network and soft computingtechniques, but performs high prediction rates. Results claimed that this approach predicted more than 90% of a set of a hundred test characters.

In [11] the method uses document capturing, preprocessing, segmentation of zones containing ICR cells to individual components, the features extraction form the components.The classification system takes less computation timethan others methods.

## III. DATASET AND METHODOLOGY

*A. Data*

In this section the dataset used and the proposed method are mentioned.

The dataset being using is the EMNIST dataset which is the extension to the original MNIST dataset. The EMNIST dataset consists of handwritten letters and digits extracted from the NIST database 19. The dataset images are converted to a 28x28 pixel image format and the structure of the dataset that matcheswith the MNIST dataset. The full complement of the NIST Special Database 19 (EMNIST) is available in the ByClass and ByMerge splits. These two datasets have the same image information but differ in the number of images in each class. Both datasets have an uneven number of images per class and there are more digits than letters. The number of letters roughly equate to the frequency of use in the English language.

Fig 1: Data Set example

***B.*** *Dataset Preprocessing*

Before training the model with the dataset, we have applied few image preprocessing methods like contrast enhancement, image sharpening and scaling on our data to make the data set more real and applicable to our proposed algorithm.

***C.*** *Segmenting and Padding images*

Dataset consists of images of words, so individual character must be extracted from the words using segmentation Connected Component Labeling (CCL) method. The result of segmentation gives images of different size and shape due the variable height and width of English letters. Carping images will not be suitable because we may lose some portion of images. Also resizing may cause poor image quality due to different aspect ratios of different images. So padding the original images in the white space of size which maximum size in our data set.

***D.*** *Rotating Images*

Though our dataset contains images of various letters and digits, few images were rotated with some degree to serve two purpose, first one is by tilting images we got an another images which adds number of training samples. Second one is by tilting the images we will get image data of real time handwriting, so that our model will more robust to any test data.

***E***. *Proposed Method*

The basic generic methodology has been shown in the following diagram:-

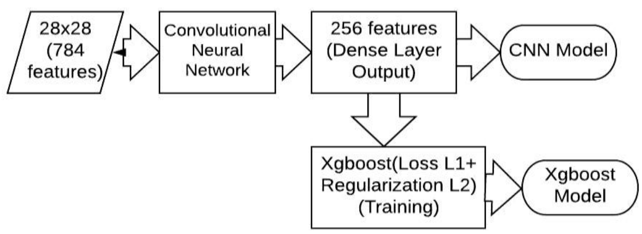


Fig 2: Method Flow of Handwriting recognition

*F. XGBoost*

Extreme Gradient Boosting (Xgboost) where the term Gradient Boosting is used for improving the prediction accuracy. The Xgboost is a set of classification and regression trees (CART).The mathematical representation of the tree model which has loss function and the regularization function is:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F}$$

where KK is the number of classifiers, ff is a cost function in functional space, and FF is the set of all possible CARTs. The objective function is represented as:

$$\text{obj}(\theta) = \sum_{i}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Based on the objective function training for the model can be carried out. It is difficult task to train all the classifiers at once. So we used an additive technique,find what has been learned, and add one more new classifier. The value of prediction at step t is $\hat{y}_i^{(0)} = 0$ , so we have finally:

$$\text{obj}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t} \Omega(f_i)$$

In XGBoost the regularization complexity can be defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

The outputs received from CNN's extraction layer are fed to the XGboost model for training and testing. During training after each batch of images accuracy is calculated. Finally after completion of all the epochs the final accuracy is the mean value of individual accuracies.

***G**. Experiment*

In the proposed method the CNN is initially used to extract features of the dataset images. CNN acts as a powerful feature extraction tool. The training data is divided into 4 parts- (train_images, train_labels) and (test_images, test_labels). Then each set is passed into the designed CNN.The first layer in a CNN is always a convolutional layer. It takes input images of format (28, 28, 1) where image is 28 x28 pixels wide and monochrome. The filter dimension for input layer is set to 784(28 x28).
Convolutional Neural Network contains maxpooling orsub-sampling layer. The pooling layer is placed followed by theconvolutional layer. Sub-sampling layersimplifies the convolutional layer output information. The max pooling layer simplifies the image data by 4 folds.The output of the first convolutional layer has six featuremaps of size 24x24 pixels and the second convolutional layercontains 12 feature maps of size 8x8 pixels. The kernel size whichwas used for convolution is 5x5. We used ReLu activation function in our hidden layers to provide convergence of hidden layer output which is input to the next layer. We used flattening and one hot encoding methods to make classification easier and efficient. Softmax layer is used as the fully connected layer to classify the characters.

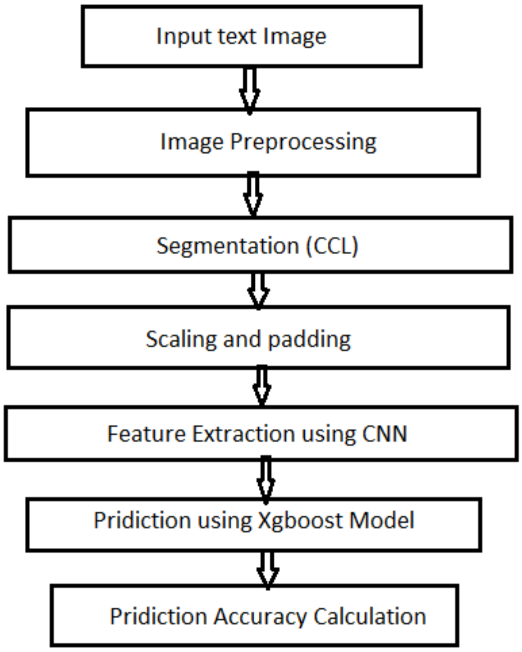

Fig.3.Flow Diagram of Proposed method

Though the softmax layer classifies the characters, the transfer learning technique with Convolutional neural network can provide better classification accuracy. Hence we replaced softmax layer with XGboost classifier technique.The output of the firstsub-sampling layer is 12x12 with six feature maps and thesecond is 4x4 with 12 feature maps. The output of the secondsub-sampling layer is transposed into vector feature. The extracted feature maps from the hidden layers of Convolutional neural network are stored as separate data and they used as training data for XGboost classification technique.

## IV. RESULT AND DISCUSSION

In this research work we made two experiments, one is Convolutional neural network with our architecture for the NIST dataset combined with our prepared dataset and evaluated the accuracy of prediction. Secondly, the extracted features from CNN hidden layers outputs are given to Xgboost classification method and evaluated the classification accuracy. We divided the dataset into 75% for training and 25% for testing in our experiment.The CNN model without Xgboost with softmax layer shows an accuracy of 93.42% and CNN with Xgboost gives an accuracy of 97.18 for letter and digit recognition. Whereas English letter recognition systems that are not based on CNN have an even lesser accuracy rate. The accuracy results are shown in below table.

Table.1. Accuracy Comparison

| Method | Accuracy |
|---|---|
| CNN | 93.42% |
| CNN with Xgboost | 97.18% |

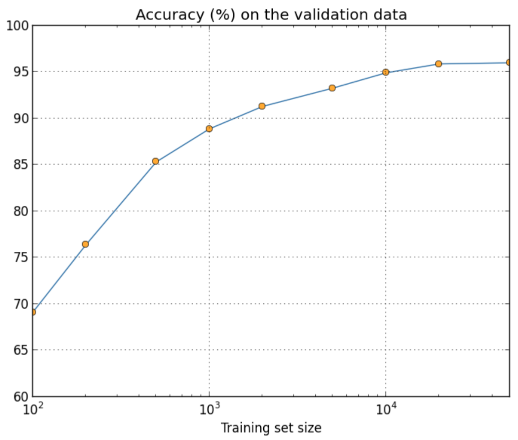


Fig.4. CNN + XGBoost Accuracy

## V. CONCLUSION AND FUTURE WORK

In this proposed work, we used CNN with Xgboost classification technique where CNN for feature extraction and Xgboost for prediction. From our experiment results on NIST data for training and our data for testing, our proposed method achieves an accuracy rate comparatively better than CNN with softmax layer as classifier. In future research, few more CNN architecture can be tried to extract features. Better preprocessing technique like other segmentation methods and noise removal methods can be applied on raw text images for better training and prediction on joining letters. Further the research work can be done on reducing computation time with better GPU systems or parallel processingsystem and CNN with LSTM.

## REFERENCES

1. Benjamin Graham, "Sparse arrays of signatures for online character recognition," pp. 1–10,2013. White paper.
2. Shruthi A., Patel M S., "Offline Handwritten Word Recognition using Multiple Features with SVM Classifier for Holistic Approach", International Journal of Innovative Research in Computer and Communication Engineering, vol.3, 2015
3. Elleuch M.,Maalej R and Kherallah M.,"A New Design Based - SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition", Elsevier, Vol.80, pp-1712-1723, 2016.
4. Lecun Yann and Bengio Y,"Convolutional Networks for Images, Speech, and Time-Series," Handbook of Brain Theory and Neural Network, 1998.
5. Jerome H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine." IMS 1999 Reitz Lecture. White paper, February 24, 1999.
6. Darmatasia and Mohamad Ivan Fanany, " Handwriting Recognition on Form Document UsingConvolutional Neural Network and Support VectorMachines (CNN-SVM)",International Conference on Information and Communication Technology (ICoICT),2017.
7. Anuja Naik, Patel M S.,"Offline English Handwritten Word Recognizer Using Best Feature Extraction", irdIndia, Vol.3, 2014.
8. Weixin Yang, Lianwen Jin+, Zecheng Xie, Ziyong Feng, "Improved Deep Convolutional Neural Network for Online Handwritten Chinese Character Recognition using Domain-Specific Knowledge", ICDAR, 2015.
9. Anshul Gupta, Manisha Srivastava, Chitralekha Mahanta, "Offline Handwritten Character Recognition Using Neural Network", IEEE Conference on Computer Applications and Industrial Electronics, 2011.
10. Puttipong Mahasukhon, Hossein Mousavinezhad,Jeong-Young Song, "Hand-Printed English Character Recognition based on Fuzzy Theory",IEEE Explore, 2012.
11. F. Farooq, K. Sridharan, and V. Govindaraju, Identifying Handwritten Text in Mixed Documents, ICPR 2006, 18th International Conference on Pattern Recognition, v. 2, pp. 1142 - 1145, 2006
12. Abhishek Jindal,Mohd Amir, "Automatic Classification of Handwritten and Printed Text in ICR Boxes", IEEE Explore, 2014.
13. S. Violante, R. Smith, and M. Reiss, A Computationally Efficient Technique for Discriminating Between Hand-Written and Printed Text, IEEE Colloquium on Document Image Processing and Multimedia Environments, 2 Nov.,pp. 17/1 - 17/7, 1995.
14. URL; http://ufldl.stanford.edu/tutorial/supervised/ConvolutionalNeuralNetwork/
15. URL:https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/
16. Rafael C.Gonzalez, "Digital Image Processing" 3rd edition.
17. URL: https://cs231n.github.io/convolutional-networks/
18. Harpreet Kaur and Simple Rani,"Handwritten Gurumukhi Character Recognition Using Convolution Neural Network", IJCIR, vol.13, pp.933-943, 2017.
19. Weixin Yang and Lianwen Jin," Improved Deep Convolutional Neural Network For Online Handwritten Chinese Character Recognition using Domain-Specific Knowledge", white paper.

## AUTHORS PROFILE

Dr.C.Lakshmi holds Ph.D in Computer Science and Engineering from SRM University and is Professor in the department of Software Engineering SRM Institute of Science and Technology. Her main area of research interest includes pattern Recognition, image processing, machine learning and web services. She is a Life time member of the Indian Society for Technical Education (ISTE), International Association of Computer Science and Information Technology, Singapore and International Association of Engineers-IAENG, Hong Kong. She has published several papers in well-known peer-reviewed journals.

Joseph James S holds a M.E degree in Software Engineering from Anna University and is a research scholar in the department of Computer Science and Engineering, SRM Institute of Science and Technology. His main area of research interest includes pattern Recognition, image processing and machine learning. He is member of IEEE and Indian Science Congress. He has published several papers in well-known peer-reviewed journals.