International Journal of Engineering &Technology, 7 (3.12) (2018) 341-343

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Neural Network with Regression Algorithms for Optical Character Recognition

J.V.Vidhya[1*], Archit Kumar[2], Yashwardhan Parakh[3]

[1]Assistant Professor, [1, 2, 3]Dept. of Computer Science and Engineering
SRM University, Chennai, Tamil Nadu, India
Corresponding Author E-mail: [1]vidhya.j@ktr.srmuniv.ac.in, [2]archit4k@gmail.com
[3]yashwardhan1995@gmail.com

Abstract

In today's automatic and robust modern world, possibilities of optical character recognition is endless. Previously OCR was used in postal service to read address from mail, car number plate tracking, automation of bank check transfer but today it has taken document management system to whole new level. Using OCR we can convert normal hardcopy document into Searchable text. We will use deep Neural network to train systems to recognise characters in a precise manner, basically we have proposed neural network model combined with machine learning technique like gradientDescent, regression, softmax normalization which will help to increase the efficiency of the OCR. Computer will able to recognise hand written digit. We will be using Google's advanced TensorFlow to create an OCR system which will be efficient and robust in action.

Keywords *- Optical Character Recognition, deep Neural network, TensorFlow, patten matching, machine learning, regression, gradient descent, softmax regression*

1. Introduction

In todays world there are so many writing tools available like touchpad, touchpen etc still many people still follow the old method of taking notes with there bare hands. Taking notes with old methods is very difficult process. Firstly you have store it. Accessing it also is a huge task. Sharing it efficently is tedious. Many times important documents are lost in this process.
Now in our research paper we have taken up this problem and we believe that accessing, storing, sharing of digital format documents is much easier compared to hand written notes. The aim of the project is to classify different hand written notes and convert them into digital form.
In handwritten text there are many fields where we can research so we are narrowing our research scope by specifiying the meaning of handwritten text. Our challenge is to classify images which are in cursive or block writing. Now in our research we will combine many algorithms that can extract words from a given line or that can extract lines from a given document. So by this combination our research can take a deliverable form and can be used by end user.But the most challenging part is classification part which we will tackle by breaking lines into words and documents into lines. We are also using CNNs but we know that it works well on raw input pixels. We are aiming at extracting characters from each word image and then classifing each character to form a whole document.
According to A. Rajavelu, M. T. Musavi and M. V. Shirvaikar, feature extraction is the most important step in an optical character recognition system because the features such as loops, holes, strokes, lines, etc. are important to identify a text by humans, but computers need a better technique.
They wanted to reduce time taken in image processing while maintaining efficiency and versatility.
Thus, they adopted Walsh functions for feature extraction and neural network with 40 hidden nodes.
But Convolutional Neural Network (CNN) and spatial sparsity are better for feature extraction when used with slower max-pooling to retain more spatial information (Benjamin Graham, 2014).

2. Traditional Vs. Our Proposed Technique

1. Our techniques combines neural network with machine learning techniques efficiently.
2. For example, just for changing the bias of neural model we are using gradient descent exponentially normalized with softmax regression machine learning technique.
3. This greatly should improve the accuracy.
4. This technique uses more computation power but todays computer with high number of teraflops graphics compute engine can easily do the job.
5. We will use GPU (not CPU) for parallel computation.

3. Challenges Involved

1. Selecting the deep learning model for OCR and compromising between speed and accuracy.
2. For speed: Softmax Regression (or Decision Tree, etc.).
3. For accuracy: Convolutional Neural Network (or Random forest).
4. Deciding the number of outputs at each level of neural network (again a compromise between speed and accuracy).
5. Using the concept of slow pooling.
6. Improving the performance of the system.

4. Softmax Regression

Softmax regression is also called as multinomial logistic regression. It involves to process. The first process is all the evideces of our input is summed up. The second process is all the evidences which we have collected from input are then converted into probabilities.
To match up that all evidences are of same class, what we do is we perform weighted addition of pixel intensities. If pixel has a high intensity against an image, weight will be negative and vice-versa.

5. Convolutional Neural Network

In In machine learning,CNN is a neural network which is forward feeded.In CNN there are multilayer perceptrons which makes minimal preprocessings.CNNs other names are shift invariant or space invariant NNs.The development of CNNs was based on nuerons.CNNs use less preprocessing when it is compared with other classification algorithms.CNNs has input layers and output layers.Convolutional layers are the name given to hidden, pooling, fullyconnected and normalization layers.

6. Pooling

Pooling is known as non-linear down-sampling.Now in pooling also max pooling is the most commom.What it does is it breaks images into many non-overlapping rectangles and for every regions finds the maximum output.It helps to reduce the spatial size of representation and it reduces both parameters and computation in network.

7. Cross Entropy

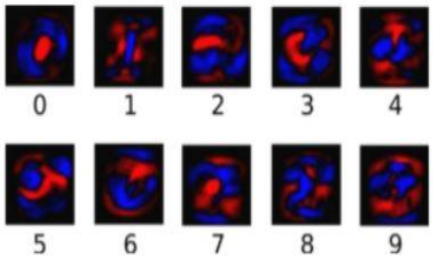
The main work of cross entropy is it measures the average number of bits needed to check an event from a set, that is lying between two probabilities distributions.
It makes the system better in generalizing training data set.Also it helps to choose weights in the network and set heuristics to choose hyper-parameters that a good for the network.

8. Issues in Using the Existing System

According to A. Rajavelu, M. T. Musavi and M. V. Shirvaikar, feature extraction is the most important step in anoptical character recognition system because the features such as loops, holes, strokes, lines, etc. are important to identify a text by humans, but computers need a better technique. They wanted to reduce time taken in image processing while maintaining efficiency and versatility. Thus, they adopted Walsh functions for feature extraction and neural network with 40 hidden nodes. But Convolutional Neural Network (CNN) and spatial sparsity are better for feature extraction when used with slower max-pooling to retain more spatial information (Benjamin Graham, 2014).

9. Proposed Technique

1. Convert images into grayscale for faster processing.
2. In order to recognise patterns we need data set, we have choosen MNIST data set which is alarge collection of 28x28 images of hand written digit. Images are greyed out because colour values are not needed for computation. Dataset will be divided into 2 parts, training data and test data. Each image has its label of it's digital value. Training dataset has 60,000 images where as testing data has 10,000 images. We will feed the training data into deep Neural network to change the weights of the edges. After each iteration of training we need to calculate the correctness of recognised value so that we can do feedback and change the bias for more better results.
3. The first process is all the evideces of our input is summed up. The second process is all the evidences which we have collected from input are then converted into probabilities.
4. Example to show the weights one model learned. Red color denotes negative weights and blue color denotes positive weights.



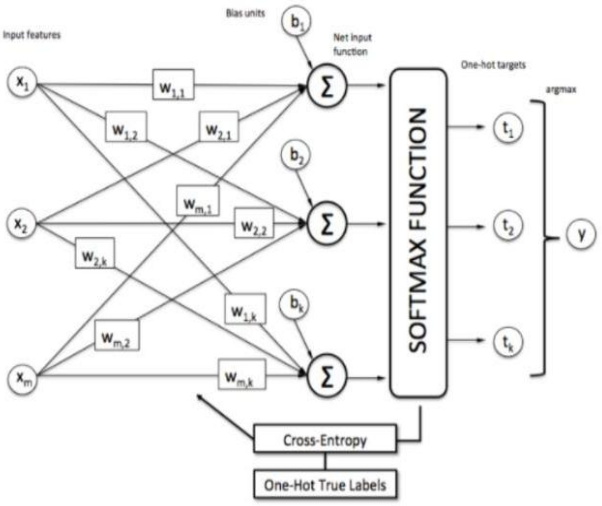
5. Bias are the extra evidences which we will add.By this we can show that there are things that are independent of the system.
6.

$$\text{Evidence(i)} = \sum_j^n W(i,j)x(j) + bias(i)$$

7. We will use deep neural network with 4 internal layers to increase accuracy of adjusted weights.
8. Passing of evidences to softmax function.
9. Y = Softmax(evidence)
10. In order to get the output we want we are using softmax which is acting as link.
11. We can normalize by taking the exponential value, this will improve accuracy.

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$$

12. Then implement the regression so that it adapt with the new unknown images.
13. If we use 0.5 as a learning rate to reduce the cross entropy ,using GDA that will give us the best possible result.



10. Evaluation

1. Each image in mnist data has it's desired label.
2. tf.argmax function gives you the index of the highest entry ina tensor on an axis. For example, tf.argmax(x,2) is the label our model.Our modal thinks tf.argmax(x,2)is mostlikely for each input, while tf.argmax(x_,2) is the correct label.Tf.equal fuction can be used to match the correctness of prediction.

3. Basically, match with correct answer with generated answer and put as flag 1 in the array otherwise 0.
4. Example: [1011] has 75% accuracy.

11. Conclusion and Future Work

A CNN is efficient for recognising the features of the text with slow max-pooling. More the nodes in neural network, more is accuracy and processing time.
A system needs to be built to set the weights and biases according to the training data. Also, the input images should be converted to greyscale images as a pre-processing step. Along with them, a regression model should be used to categorize the images as one of the numbers between 0-9.

1. Character Recognition for various orientations.
2. Character Recognition for multiple font styles.
3. To convert handwritten and printed text to digital format to save trees.
4. Converting printed documents to searchable text.
5. Increasing the efficiency of character recognition.
6. During parking, usually paper is used to calculate the bill, automating the process by
7. automatically reading the number plate to save paper.

References

[1] A Neural Network Approach to Character Recognition (Rajavelu, Musavi and Shirvaikar).
[2] Spatially sparse Convolutional neural network (Benjamin Graham).
[3] www.medium.com
[4] www.wikipedia.com
[5] www.tensorflow.com
[6] https://www.tensorflow.org/
[7] https://pythonprogramming.net/
[8] https://github.com/