Machine Learning for Handwriting Recognition

Preetha S[a]*, Afrid I M[b], Karthik Hebbar P[c], Nishchay S K[d]

[a,b,c,d] *Department of ISE,B.M.S. College of Engineering/ VTU, India, {preetha.ise,1bm16is006, 1bm16is044, 1bm16is060}*

[a]*Email: preetha.ise@bmsce.ac.in*

Abstract

With the knowledge of current data about particular subject, machine learning tries to extract hidden information that lies in the data. By applying some mathematical functions and concepts to extract hidden information, machine learning can be achieved and we can predict output for unknown data. Pattern recognition is one of the main application of ML. Patterns are usually recognized with the help of large image data-set. Handwriting recognition is an application of pattern recognition through image. By using such concepts, we can train computers to read letters and numbers belonging to any language present in an image. There exists several methods by which we can recognize hand-written characters. We will be discussing some of the methods in this paper.

Keywords: CNN Zoning; Incremental; Handwriting recognition.

1. Introduction

In the field of Machine Learning, recognition of objects has become most sought one. Some of the examples of object recognition are Face recognition, Hand write recognition, Disease detection etc. All these things can happen through large set of image data set. These image data set will contain both positive and negative data regarding that domain. This helps the algorithm to classify the unknown data in better ways. Hand write recognition is a new technology that will be useful in this 21st century. It can act as base functionality for the birth of new requirements. For example, a blind man cannot read news paper unless braille format exists. In this case we can train the algorithm to recognize characters in the news paper, store them as text and convert the text to speech. This can help lot of blind people to ease their daily work. The second application of hand write recognition could be language translation. In this case when a person is dealing with non-native language, he can just take a image of a document and send it to the hand write recognition algorithm. This algorithm can recognize the characters in image and convert them to text. Then the text can be converted to desired language of choice.

----

* Corresponding author.

One more application of hand write recognition would be, processing of large set of paper document like answer scripts. With the help of hand-write recognition and AI, the answer scripts can be evaluated without human involvement. For all above mentioned scenarios, hand write recognition acts as base case to be resolved. Hand write recognition is one of the type of Optical Character Recognition(OCR). OCR is identification of text, which may be printed or hand-written. In OCR, the document is captured via camera as image and can be converted to desired formats like PDFs. Then the file is fed to the algorithm for character recognition. This can drastically reduce human involvement in certain scenarios.
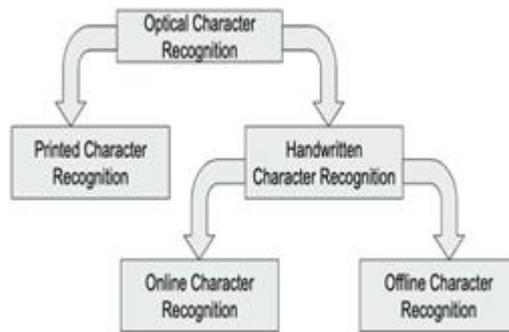

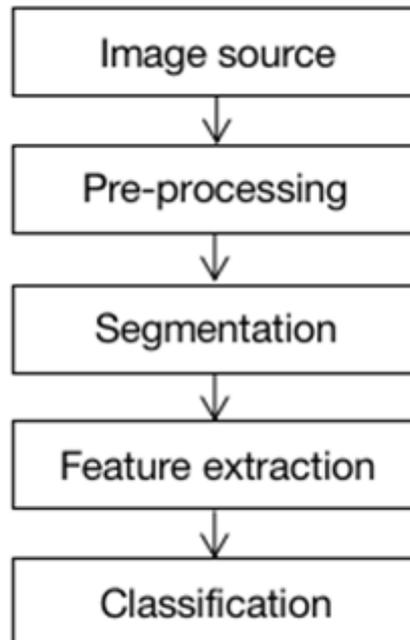**Figure 1:** classification of Optical Character Recognition

Figure 2: Phases of hand-write recognition

The two derivatives of OCR are Printed character recognition and Hand-written character recognition. Printed character recognition is as the name suggests, recognition of characters in the image of news paper or any other form of printed document. Hand-written character recognition involves recognition of characters written by human or has human involvement in it. It is divided into two types namely Online character recognition and Offline character recognition. Offline character recognition involves parsing image of document to series of texts and words. Online character recognition is a bit complicated process. It is a dynamic process. It involves of recognizing character data at the time of writing itself. It needs specialized writing pad and an electric pen. On the basis of movement of pen, the written character is recognized. The visual classification of Optical Character Recognition is shown in figure 1 and Phases of hand-write recognition in figure 2.

1. Image source

This phase comes in offline hand-written character recognition. Image source can be from any digitized tool. A scanner or a camera captures the image and is sent to next phase.

2. Pre-processing

Pre-processing is sequence of operation, that improves the quality of image and hence increases the accuracy of image. For the hand-written character recognition process, the following pre-processing techniques are followed.

   a) Noise-removal:- It is the process of removing noise from image. This also refers to smoothening the image by reducing the unwanted signals in the image. There exists many algorithms to remove noises in the image. Some of them are Gaussian filtering method, Min-max filtering method, Median filter etc.

   b) Binarization:- It is a mechanism of converting gray-scale or colored image to binary image. Binary images contain only 0's and 1's. The pixels in images are partitioned to 0's or 1's based on some constant value. If the pixel value is less than the constant, it is replaced with 0 or else with 1.

   c) Morphological operation: This is the process of increasing or decreasing the size of an image. This is done mainly because the algorithm would expect the constant image size. To increase the size of an image, we can add pixels to the boundary of image. To decrease the size of an image, we can remove pixels from the boundary of image.

3. Segmentation

Segmentation is a mechanism that extracts individual characters in the image. There exists two types of segmentation. They are Implicit segmentation and Explicit segmentation. In implicit segmentation the words are recognized without segmentation process. But in explicit segmentation, words are predicted by extracting individual character.

4. Feature-extraction

This is the important phase in the recognition process, and the algorithm of recognition starts from here. Each character contains its own features. It contains group of rules where each rule explains feature of a character. Extraction of such features is done in this phase.

5. Classification

By this time, the training would have completed and the testing of input data starts. The testing data would pass all the above process and the varying probabilities are assigned to the matching rules. The rule with highest probability is selected and the corresponding class-label is made recognized character.

2. Literature Survey

Using CNN

CNN is abbreviated form of Convolution Neural Network. Convolution refers to twisted or coiled. Any neural network is similar to human brain. Neural networks are designed by taking inspiration from brain. CNN is mainly used for Image classification. CNN consists of many layers depending on the requirements. Ahmed Mahdi Obaid and his colleagues [9] proposed an effective handwritten text recognition system using two different learning algorithms. Considering similar configuration, Scaled Conjugate Gradient algorithm proved to perform well in terms of accuracy and training time when compared Resilient Back-propagation algorithm. Salma Shofia Rosyda and his colleagues [2] discussed three main layers of CNN are

1. Convolution layer
2. Pooling layer
3. Fully connected layer

Convolution layer

The idea of image classification is because of capability of pattern detection in convolution layer. Jagan Mohan Reddy D., and his colleagues [4] proposed a method to improve recognition rate of individual characters for Telugu language using Deep Learning (DL). The input to this layer is the image matrix of form width*height*depth. Depth in the matrix refers to the number of channels in the image. For gray-scale image, number of channels would be one, where as for RGB image number of channels would be three. We can convert the RGB image to gray-scale image and perform CNN operation over it.

Table 1: Example for a filter

-1	-1	-1
1	1	1
0	0	0

The convolution layer uses matrix called kernel or filter. This filter tells the pattern that we want to recognize. For recognizing the top horizontal edge, following filter is used as represented in below table.The matrix representation of character image is extracted. The pre-processing stage is performed on the image to remove noise and other unwanted data if present. We can also perform dimensionality reduction on the image as part of pre-processing. The filter is multiplied with a sub-matrix of same size in image. The multiplication starts from top left corner. The multiplication is dot product, not the usual matrix multiplication. The resultant of the dot product is stored in the top left corner and the filter is moved to next sub-matrix. The filter is multiplied with every possible sub-matrix in image.

Figure 3: Images After processing and Convolution

Figure 5 shows two images. Left image represents image after pre-processing. Right image shows image after convolution. The bright edges in right image is obtained by applying the filter mentioned above.

Pooling layer

The proposed algorithm tries to address both the factors and well in terms of accuracy and time complexity. The proposed algorithm tries to address both the factors and well in terms of accuracy and time complexity. S. M. Shamim and his colleagues [10] discussed an off-line handwritten digit recognition method based on several machine learning techniques. Using WEKA; Bayes Net, Support Vector Machine, Multilayer perception, Naïve Bayes, Random Forest, Random tree and J48 are used to recognize digits. Layers are mainly used for dimensionality reduction. The output of Convolution layer can be of large size (200*200). This cannot be taken as input to next layers, because the next layer is fully connected layer. So the matrix size needs to be reduced. The pooling layer can extract dominant or submissive feature in the image. These are called max-pooling, min-pooling and average-pooling. In hand-write recognition Max pooling is used. In pooling a fixed m*n size of empty batch move along the image. When a batch overlaps on particular portion of convoluted matrix, the maximum/minimum/average value of all the pixel is taken. And the new matrix is constructed by filling the pooling values.

Fully connected layer

The main aim of fully connected layer is to take the result of pooling layer and use them to classify the image into a label. The output of pooling is usually matrix. This matrix is converted to a single dimensional array. This process is called flattening. The values in vector represents probability of some features of the object. Polaiah Bojja., et.al. [8] proposed a model to convert handwritten text to Voice and Computer Text. The model

analyzed various applications of Consumer and Health Sectors. The flattened value is passed through multiple layers of perceptrons. The inputs are multiplied by weights and passed to activation function. This activation function is mostly ReLU. This function would typically remove negative values present in the input. The definition of ReLU is given as $f(x)=max(0,x)$. The equation tells that if the input to ReLU is negative, it returns 0 otherwise it returns the input. This is the process in single iteration of CNN. Usually CNN goes through many such iteration. These iteration are called Epoch. As the number of epochs increases to certain extent, the accuracy rate also increases. When the threshold is crossed, the accuracy rate would decrease.

Handwriting recognition using ML.

There are many algorithms to recognize the handwriting. There is a technique called OCR (Optical Character Recognition) which is used to recognize the handwritten and paper documents. Yuval Netzer and his colleagues [7] discussed problems of digit recognizing systems using unsupervised feature learning methods. Paper documents refers to documents that should be scanned and some of those are typed digitally. Handwritten documents refer to those which are written by hand.

Surya Nath RS and his colleagues [1] presented two types of recognition methods of different natural languages.

1) Offline Handwriting recognition
2) Online Handwriting recognition

Offline Handwriting recognition: Here the documents which are already written and stored is used to recognize the characters present in the documents. The characters can be alphabets or numbers or any other symbols. This method can be used for recognizing mathematical expressions. This is used in many mobile apps and using this students can scan the document and Mathematical expression recognizer will recognize the equations and provide solutions. *Online Handwriting recognition:* Here the data is not scanned from document instead the characters are written using some electronic pen and character is recognized in real time. Here strokes are considered for recognition of characters.

In [3,5] many methods used for recognition of Handwriting are listed.

- Convolutional Neural network
- Semi Incremental Recognition
- Incremental Recognition
- Line and word segmentation
- Part based Method
- Slope and Slant correction method
- Ensemble method

1)   Incremental Recognition method:

Meenu Mohan et.al [6] proposed a method used to recognize the characters incrementally. Here

- First the stroke is received from user
- Geometric feature is updated
- Symbol is recognized
- Updating CYK table

2)   Line and Word segmentation

Salma Shofia Rosyda and his colleagues [2] discussed methods of CNN having highest accuracy along with Slope and Slant Correction Method with lowest accuracy. One of the major step in recognition of characters is segmentation. Without segmentation it is difficult to recognize the characters or symbols accurately. So the steps in segmentation process are:

- Finding area of interest by viewing scanned images and crop imagery
- Noise should be removes from image by method called subtraction. Pre-processing is done to remove the noise on the image
- Image is then converted to binary by mining text and by removing background.
- Hough transform is used for correction and skew detection
- Line Segmentation is done using Horizontal Projection.

3)   Part Based Method

This method is used to recognize the object. Properties of this method is given below:

- Single image should be represented by multiple keypoints
- Similarity of image will depend on image has key point or not.
- Sometimes Each class is represented by key points collection.

The benefits of Part-Based Method are:

- Difficult to normalize with pre processing but characters are recognized well
- It is independent of global structure
- It can be applied to cursive script

4)   Slope and Slant correction method

Meenu Mohan and his colleagues [6] used a method to reduce the style variation. Here by using this method slope of text can be estimated based on the slope of baseline. There is no contribution of ascenders and descenders in the initial formation, they are removed as possible.

5)   Ensemble Method

S. M. Shamim and his colleagues [10] presented new classifier, which is introduced in the Machine learning.

From this method multiple classifiers are generated from one base class automatically. The explanation is given below:

- Feature Search Ensmble method
- o Feature selection algorithm is used to identify features.
- Feature Selection Ensmble method

6)   Zoning method

P. Shankar Rao and his colleagues [11] implemented zoning method. The processed image taken is divided into zones and then the feature extraction of the symbol or characters is done

It has two methods:

a. Static Zone
b. Dynamic Zone

a)   Static Zone

Here the image is divided into uniform zones. During complete recognition the zones is fixed. This is done without having any prior information on feature extraction or feature distribution. Static zoning is done by having information on experimental evidences or by the years of experience of developer.

b)   Dynamic Zone:

Processed image is divided into zones which are non uniform. Here the images is divided into many zones which can be resized dynamically. And it is not static. Based on the adjacent zones the zone size is resized. And using this method whenever some zones are recognized and some other zones need extra space they can be extended by resizing the zone. This is the very important advantage. Having the results of optimization procedures, this dynamic zones are designed. In order to extract the features from the processed image, based on local information pattern we have to adjust the positions of every zones. This adjustment is done by moving dynamic zones near pattern body and the offset for adjustment of dynamic zone is calculated by maximizing pixel density.

3. Conclusion

There are many approaches for hand writing recognition. Some of them are incremental, Zoning, Convolutional Neural Network (CNN), semi-incremental segmentation, slope and slant correction. Among these methods, highest accuracy is achieved from Convolutional Neural Network (CNN) and the least accuracy is achieved from Slope and Slant Correction method. When the images are trained with CNN, we will achieve good accuracy and this is one of the successful method for hand writing recognition and only disadvantage with this method is that training time of the model is too high because lot of image samples are included. In Zoning

method, if zones which are achieved after dividing input image and if the count of these zones are lesser then accuracy will decrease. Main disadvantage of this method is that developers will face lot of problems while segmentation process but this method is too simple for hand writing recognition. This method only sees the Lat and which makes it simple. Hand writing recognition is very challenging because all the individuals have different hand writing and it becomes more complex to detect when these are compared to that of computer.

References

[1]. Surya Nath RS, and S. Afseena. "Handwritten Character Recognition–A Review." International Journal of Scientific and Research Publications (2015).

[2]. Salma Shofia Rosyda, and Tito Waluyo Purboyo. "A Review of Various Handwriting Recognition Methods." International Journal of Applied Engineering Research 13.2 (2018): 1155-1164.

[3]. Younus, S. B. S., S. Shajun Nisha, and M. Mohamed Sathik. "Comparative Analysis of Activation Functions in Neural Network for Handwritten Digits." Studies in Indian Place Names 40.71 (2020): 793-799.

[4]. Jagan Mohan Reddy D, A Vishnuvardhan Reddy "Recognition of Handwritten Characters using Deep Convolutional Neural Network".

[5]. Hruday M. "Implementation of Handwritten Character Recognition using Neural Network".

[6]. Meenu Mohan, and R. L. Jyothi. "Handwritten Character Recognition: A Comprehensive Review on Geometrical Analysis."

[7]. Yuval Netzer, Tao Wang Adam Coates Alessandro Bissacco Bo Wu Andrew Y. Ng. "Reading digits in natural images with unsupervised feature learning." (2011).

[8]. Polaiah Bojja., Naga Sai Satya Teja Velpuri, Gautham Kumar Pandala,S D Lalitha Rao Sharma Polavarapu "Handwritten Text Recognition using Machine Learning Techniques in Application of NLP".

[9]. Ahmed Mahdi Obaid, IIHazem M. El Bakry, IIIM.A. Eldosuky, IVA.I. Shehab "Handwritten text recognition system based on neural network." Int. J. Adv. Res. Comput. Sci. Technol.(IJARCST) 4.1 (2016): 72-77.

[10]. S. M. Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana, Abdullah Al Jobair "Handwritten digit recognition using machine learning algorithms." Global Journal Of Computer Science And Technology (2018).

[11]. P. Shankar Rao, J. Aditya. "Handwriting Recognition – "Offline" Approach".