

Optical Character Recognition Using Artificial Neural Network

Sameeksha Barve

Abstract— In this paper, an Optical Character recognition system based on Artificial Neural Networks (ANNs). The ANN is trained using the Back Propagation algorithm. In the proposed system, each typed English letter is represented by binary numbers that are used as input to a simple feature extraction system whose output, in addition to the input, are fed to an ANN. Afterwards, the Feed Forward Algorithm gives insight into the enter workings of a neural network followed by the Back Propagation Algorithm which comprises Training, Calculating Error, and Modifying Weights.

Index Terms— Optical character recognition, Artificial Neural Network, supervised learning, the Multi-Layer Perception, the back propagation algorithm.

I. INTRODUCTION

Optical Character Recognition, usually referred to as OCR, is the process of converting the image obtained by scanning a text or a document into machine-editable format. Computer system equipped with such an OCR system can improve the speed of input operation and decrease some possible human errors. Recognition of printed characters is itself a challenging problem since there is a variation of the same character due to change of fonts or introduction of different types of noises. Difference in font and sizes makes recognition task difficult if pre-processing, feature extraction and recognition are not robust. There may be noise pixels that are introduced due to scanning of the image. Besides, same font and size may also have bold face character as well as normal one. Thus, width of the stroke is also a factor that affects recognition. Therefore, a good character recognition approach must eliminate the noise after reading binary image data, smooth the image for better recognition, extract features efficiently, train the system and classify patterns. A lot of people today are trying to write their own OCR (Optical Character Recognition) System or to improve the quality of an existing one. This paper shows how the use of artificial neural network simplifies development of an optical character recognition application, while achieving highest quality of recognition and good performance. OCR system is a complicated task and requires a lot of effort. Such systems usually are really complicated and can hide a lot of logic. The use of artificial neural network in OCR applications and improve quality of recognition while achieving good performance. There are two basic methods used for OCR: Matrix matching and feature extraction. Of the two ways to recognize characters, matrix matching is the simpler and

more common. Matrix Matching compares what the OCR scanner sees as a character with a library of character matrices or templates. When an image matches one of these prescribed matrices of dots within a given level of similarity, the computer labels that image as the corresponding ASCII character.

Feature Extraction is OCR without strict matching to prescribed templates. Also known as Intelligent Character Recognition (ICR), or Topological Feature Analysis, this method varies by how much "computer intelligence" is applied by the manufacturer. The computer looks for general features such as open areas, closed shapes, diagonal lines, line intersections, etc. This method is much more versatile than matrix matching. Matrix matching works best when the OCR encounters a limited repertoire of type styles, with little or no variation within each style. Where the characters are less predictable, feature, or topographical analysis is superior.

II. STRUCTURE OF OCR SYSTEM

OCR is the acronym for Optical Character Recognition. This technology allows a machine to automatically recognize characters through an optical mechanism. Human beings recognize many objects in this manner our eyes are the "optical mechanism." But while the brain "sees" the input, the ability to comprehend these signals varies in each person according to many factors. By reviewing these variables, we can understand the challenges faced by the technologist developing an OCR system. The ultimate objective of any OCR system is to simulate the human reading capabilities so the computer can read, understand, edit and do similar activities it does with the text.

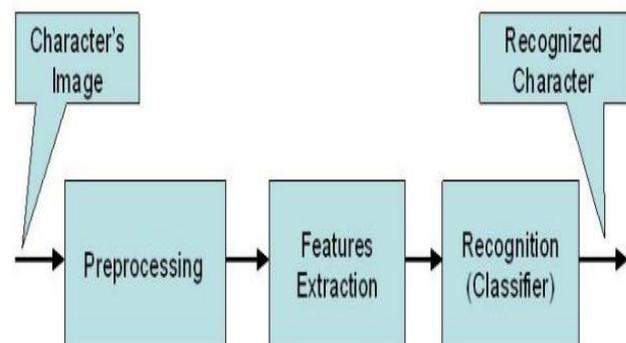


Fig1 Structure of OCR System

Block diagram of the typical OCR system. Each stage has its own problems and effects on the overall system's efficiency. Thus, to tackle the problems, either by solving each particular problem. OCR system by integrating all stages to one main stage, and this is what our research proposes. This paper presents new structure of OCR system which relies on the powerful proprieties. The algorithm is designed and tested in the related sections.

III. COMPONENTS OF OCR SYSTEM

Components of OCR system:-

Optical scanning:-

Through the scanning process a digital image of the original document is captured. In OCR optical scanners are used, which generally consist of a transport mechanism plus a sensing device that converts light intensity into gray-levels. Printed documents usually consist of black print on a white background. Hence, when performing OCR, it is common practice to convert the multilevel image into a bi-level image of black and white. Often this process, known as thresholding, is performed on the scanner to save memory space and computational effort.

Location and segmentation:-

Segmentation is a process that determines the constituents of an image. It is necessary to locate the regions of the document where data have been printed and distinguish them from figures and graphics. For instance, when performing automatic mail-sorting, the address must be located and separated from other print on the envelope like stamps and company logos, prior to recognition.

Pre-processing:-

The image resulting from the scanning process may contain a certain amount of noise. Depending on the resolution on the scanner and the success of the applied technique for thresholding, the characters may be smeared or broken. Some of these defects, which may later cause poor recognition rates, can be eliminated by using a pre-processor to smooth the digitized characters.

Feature extraction:-

The objective of feature extraction is to capture the essential characteristics of the symbols, and it is generally accepted that this is one of the most difficult problems of pattern recognition. The most straight forward way of describing a character is by the actual raster image. Another approach is to extract certain features that still characterize the symbols, but leaves out the unimportant attributes.

Post processing: -

Post processing are two types-

1. Grouping
2. Error-detection and correction

IV. ARTIFICIAL NEURAL NETWORK

A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform "intelligent" tasks similar to those performed by the human brain. Neural networks resemble the human brain in the following two ways:

1. A neural network acquires knowledge through learning.
2. A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights.

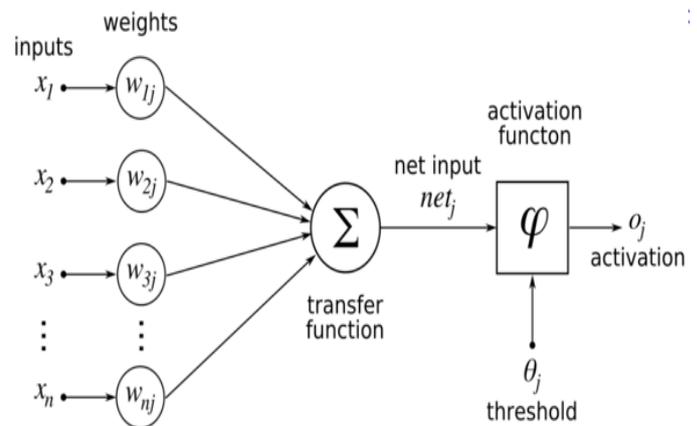


Fig 2 Structure of ANN

The most common neural network model is the multilayer perceptron (MLP). This type of neural network is known as a supervised network because it requires a desired output in order to learn. The goal of this type of network is to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown. A graphical representation of an MLP is shown below.

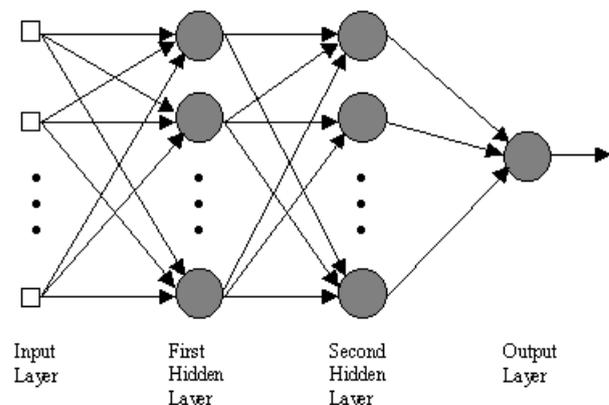


Fig 3 Block diagram of MLP

Block diagram of a two hidden layer multiplayer perceptron (MLP). The inputs are fed into the input layer and get multiplied by interconnection weights as they are passed from the input layer to the first hidden layer. Within the first hidden layer, they get summed then processed by a nonlinear function (usually the hyperbolic tangent). As the processed data leaves the first hidden layer, again it gets multiplied by interconnection weights, then summed and processed by the second hidden layer. Finally the data is multiplied by interconnection weights then processed one last time within the output layer to produce the neural network output.

V. BACKPROPAGATION ALGORITHM

The back-propagation algorithm consists of four steps:

1. Compute how fast the error changes as the activity of an output unit is changed. This error derivative (EA) is the difference between the actual and the desired activity.

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j$$

2. Compute how fast the error changes as the total input received by an output unit is changed. This quantity (EI) is the answer from step 1 multiplied by the rate at which the output of a unit changes as its total input is changed.

$$EI_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \times \frac{dy_j}{dx_j} = EA_j y_j (1 - y_j)$$

3. Compute how fast the error changes as a weight on the connection into an output unit is changed. This quantity (EW) is the answer from step 2 multiplied by the activity level of the unit from which the connection emanates.

$$EW_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial W_{ij}} = EI_j y_i$$

4. Compute how fast the error changes as the activity of a unit in the previous layer is changed. This crucial step allows back propagation to be applied to multilayer networks. When the activity of a unit in the previous layer changes, it affects the activities of all the output units to which it is connected. So to compute the overall effect on the error, we add together all these separate effects on output units. But each effect is simple to calculate. It is the answer in step 2 multiplied by the weight on the connection to that output unit.

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial y_i} = \sum_j EI_j W_{ij}$$

By using steps 2 and 4, we can convert the EAs of one layer of units into EAs for the previous layer. This procedure can be repeated to get the EAs for as many previous layers as desired. Once we know the EA of a unit, we can use steps 2 and 3 to compute the EWs on its incoming connections.

VI. CONCLUSION

Artificial neural networks are commonly used to perform character recognition due to their high noise tolerance. The systems have the ability to yield excellent results. The feature extraction step of optical character recognition is the most important. A poorly chosen set of features will yield poor classification rates by any neural network. At the current stage of development, the software does perform well either in terms of speed or accuracy but not better. It is unlikely to replace existing OCR methods, especially for English text. A simplistic approach for recognition of Optical characters using artificial neural networks has been described. Despite the computational complexity involved, artificial neural networks offer several advantages in back-propagation network and classification in the sense of emulating adaptive human intelligence to a small extent.

REFERENCES

- [1] Optical character recognition by The Association for Automatic Identification and Data Capture Technologies
- [2] Character Recognition Using Neural Networks by Near East University, North Cyprus, Turkey via Mersin-10, KKTC
- [3] Introduction to Artificial Intelligence OCR using Artificial Neural Networks by Kurt Alfred Kluever (kurt@klover.com) Department of Computer Science Golisano College of Computing and Information Sciences Rochester Institute of Technology February 18, 2008
- [4] Optical Character Recognition using Neural Networks (ECE 539 Project Report) Deepayan Sarkar Department of Statistics University of Wisconsin deepayan@stat.wisc.edu December 18, 2003
- [5] Predicting Geotechnical Parameters of Sands from CPT Measurements Using Neural Networks by C. Hsein Juang,* Ping C. Lu & Caroline J. Chen