Artificial Neural Network Based On Optical Character Recognition

Sameeksha Barve
Computer Science Department
Jawaharlal Institute of Technology, Khargone (M.P)

Abstract

The recognition of optical characters is known to be one of the earliest applications of Artificial Neural Networks. In this paper, a simplified neural approach to recognition of optical or visual characters is portrayed and discussed. OCR (Optical Character Recognition) System or to improve the quality of an existing one. The use of artificial neural network simplifies development of an optical character recognition application, while achieving highest quality of recognition and good performance. 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.

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

1. Introduction

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic conversion of scanned images of handwritten, typewritten or printed text into machine-encoded text. It is widely used as a form of data entry from some sort of original paper data source, whether documents, sales receipts, mail, or any number of printed records. It is crucial to the computerization of printed texts so that they can be electronically searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision. Early versions needed to be programmed with images of each character, and worked on one font at a time. "Intelligent" systems with a high degree of recognition accuracy for most fonts are now common. Some systems are capable of reproducing formatted output that closely approximates the original scanned page including images, columns and other non-textual components. Optical Character Recognition, or OCR, is the process of translating images of handwritten, typewritten, or printed text into a format understood by machines for the purpose of editing, indexing/searching, and a reduction in storage size. Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic translation of images of handwritten, typewritten or printed text (usually captured by a scanner) into machine-editable text. OCR is a field of research in pattern recognition, artificial intelligence and machine vision. Though academic research in the field continues, the focus on OCR has shifted to implementation of proven techniques. Optical character recognition (using optical techniques such as mirrors and lenses) and digital character recognition (using scanners and computer algorithms) were originally considered separate fields. Because very few applications survive that use true optical techniques, the OCR term has now been broadened to include digital image processing as well.

One of the most classical applications of the Artificial Neural Network is the Character Recognition System. This system is the base for many different types of applications in various fields, many of which we use in our daily lives. Cost effective and less time consuming, businesses, post offices, banks, security systems, and even the field of robotics employ this system as the base of their operations. Wither you are processing a check, performing an eye/face scan at the airport entrance, or teaching a robot to pick up and object, you are employing the system of Character Recognition.

Most OCR systems decompose the process into several stages:

1. Format Analysis (Document Image => Character String Image)

2. Character Segmentation (Character String Image => Character Image)

3. Feature Extraction (Character Image => Character Properties)

4. Classification (Character Properties => Character ID)

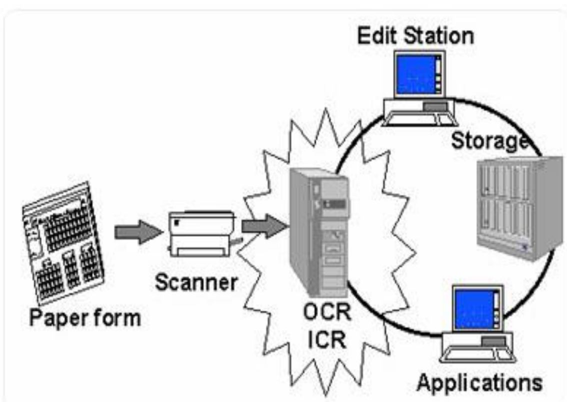

Fig.1 Working of OCR

2. Types of Recognition Engines

Recognition Engines are -

1. Optical Character Recognition (OCR)
2. Intelligent Character Recognition (ICR)
3. Optical Mark Recognition (OMR)
4. Magnetic Ink Character Recognition
5. Barcode Recognition

1. Optical Character Recognition (OCR):-

OCR engines turn images of machine-printed characters into machine-readable characters. Images of machine-printed characters are extracted from a bitmap. Forms can be scanned through an imaging scanner, faxed, or computer generated to produce the bitmap. OCR is less accurate than optical mark recognition but more accurate than intelligent character recognition.

2. Intelligent Character Recognition (ICR):-

ICR reads images of hand-printed characters (not cursive) and converts them into machine readable characters. Images of hand-printed characters are extracted from a bitmap of the scanned image. ICR recognition of numeric characters is much more accurate than the recognition of letters. ICR is less accurate than OMR and requires some editing and verification. However, proven form design methods outlined later in this paper can minimize ICR errors.

3. Optical Mark Recognition (OMR):-

OMR technology detects the existence of a mark, not its shape. OMR forms usually contain small ovals, referred to as 'bubbles,' or check boxes that the respondent fills in. OMR cannot recognize alphabetic or numeric characters. OMR is the fastest and most accurate of the data collection technologies. It is also relatively user-friendly. The accuracy of OMR is a result of precise measurement of the darkness of a mark, and the sophisticated mark discrimination algorithms for determining whether what is detected is an erasure or a mark.

4. Magnetic Ink Character Recognition:-

MICR is a specialized character recognition technology adopted by the U.S. banking industry to facilitate check processing. Almost all U.S. and U.K. checks include MICR characters at the bottom of the paper in a font known as E-13B. Many modern recognition engines can recognize E-13B fonts that are not printed with magnetic ink. However, since background designs can interfere with optical recognition, the banking industry uses magnetic ink on checks to ensure accuracy.

Fig.2 Magnetic Ink Character Recognition

5. Barcode Recognition:-

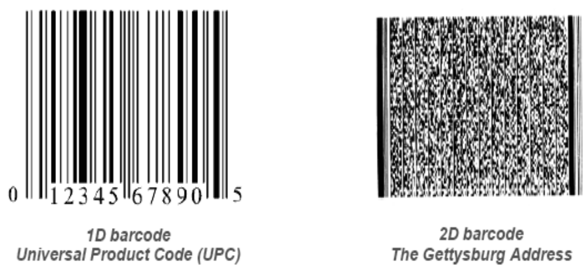

Fig.3 Barcode Recognition

A barcode is a machine-readable representation of information. Barcodes can be read by optical scanners called barcode readers or scanned from an image using software. A 2D barcode is similar to a linear, one-dimensional barcode, but has more data representation capability.

3. Artificial Neural Network

An **artificial neural network (ANN)**, usually called "neural network" (NN), is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. In more practical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Neural networks have seen an explosion of interest over the last few years, and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology and physics. Indeed, anywhere that there are problems of prediction, classification or control, neural networks are being introduced. This sweeping success can be attributed to a few key factors:

Power: Neural networks are very sophisticated modeling techniques capable of modeling extremely complex functions. In particular, neural networks are nonlinear. For many years linear modeling has been the commonly used technique in most modeling domains since linear models have well-known optimization strategies. Where the linear approximation was not valid (which was frequently the case) the models suffered accordingly. Neural networks also keep in check the curse of dimensionality problem that bedevils attempts to model nonlinear functions with large numbers of variables.

Ease of use: Neural networks learn by example. The neural network user gathers representative data, and then invokes training algorithms to automatically learn the structure of the data. Although the user does need to have some heuristic knowledge of how to select and prepare data, how to select an appropriate neural network, and how to interpret the results, the level of user knowledge needed to successfully apply neural networks is much lower than would be the case using (for example) some more traditional nonlinear statistical methods.

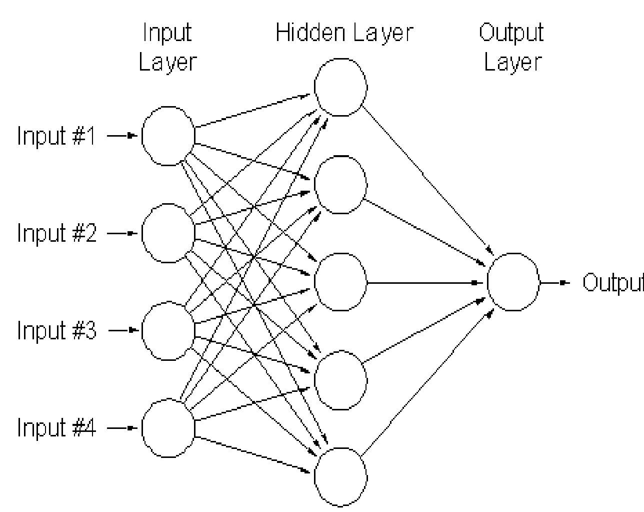

Fig.4. Structure of Artificial Neural Network

4. Algorithm Used

One of the most typical problems to which a neural network is applied is that of optical character recognition. Recognizing characters is a problem that at first seems extremely simple- but it's extremely difficult in practice to program a computer to do it. And yet, automated character recognition is of vital importance in many industries such as banking and shipping. The U.S. post office uses an automatic scanning system to recognize the digits in ZIP codes. We may have used scanning software that can take an image of a printed page and generate an ASCII document from it. These devices work by simulating a type of neural network known as a back propagation network.

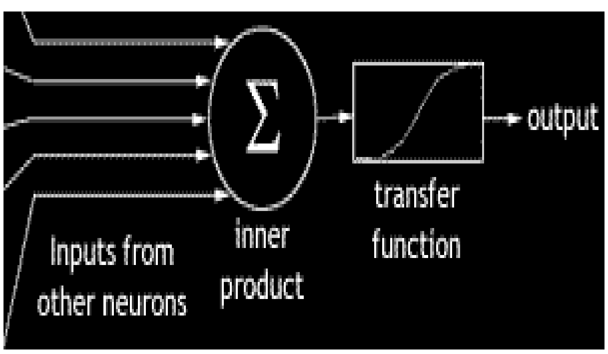

Fig.5 Back propagation

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.

$$\mathrm{EA}_j = \frac{\partial E}{\partial y_j} = y_j - d_j$$

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.

$$\mathrm{EI}_j = \frac{\partial E}{\partial x_j} = \frac{\partial E}{\partial y_j} \times \frac{dy_j}{dx_j} = \mathrm{EA}_j y_j \left(1 - y_j\right)$$

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.

$$\mathrm{E}W_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial W_{ij}} = \mathrm{EI}_j y_i$$

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.

$$\mathrm{EA}_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial y_i} = \sum_j \mathrm{EI}_j W_{ij}$$

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.

5. 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.

6. References

[1] Introduction to Artificial Intelligence OCR using Artificial Neural Networks by Kurt Alfred Kluever (kurt@kloover.com) Department of Computer Science Golisano College of Computing and Information Sciences Rochester Institute of Technology February 18, 2008

[2] 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

[3] Predicting Geotechnical Parameters of Sands from CPT Measurements Using Neural Networks by C. Hsein Juang,* Ping C. Lu & Caroline J. Chen

[4] Optical character recognition by The Association for Automatic Identification and Data Capture Technologies

[5] Character Recognition Using Neural Networks by Near East University, North Cyprus, Turkey via Mersin-10, KKTC