

# ARTIFICIAL NEURAL NETWORK BASED CHARACTER RECOGNITION USING BACKPROPAGAT

Madhup Shrivastava  
NITTTTR Shamla Hills Bhopal  
(M.P.) India  
[08madhup@gmail.com](mailto:08madhup@gmail.com)

Monika Sahu  
NITTTTR Shamla Hills Bhopal  
(M.P.) India  
[monikasahu7@gmail.com](mailto:monikasahu7@gmail.com)

Dr. M.A.Rizvi  
NITTTTR Shamla Hills Bhopal  
(M.P.) India  
[marizvi@nittrbpl.ac.in](mailto:marizvi@nittrbpl.ac.in)

## ABSTRACT

Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data format. OCR is the translation of optically scanned bitmap of printed or written text character into the character codes, such as ASCII. This is an efficient way to turn hard copy material into digital data files that can be edited or manipulated. The optical character recognition refers to the branch of computer science that involves reading text from paper and translating the images into a form that the computer can manipulate. The potential of this technology is typically used for general character recognition which includes the transformation of anything humanly readable to machine manipulatable representation. OCR systems are enormous because they enable users to harness the power of computers to access printed documents.

The aim of this paper is to find a means by which the database entry from handwritten forms can be automated. Firstly the paper deals with the technology scanning hard copy data. Secondly describes machine learning process for training the system for converting hard copy into soft copy.

**Keywords**— OCR (Optical Character Recognizer), ANN (Artificial Neural Network), MLP (Multi Layer Perceptron), Optical Language Symbols, Unicode.

## 1. INTRODUCTION

OCR is the acronym for Optical Character Recognition. This technology allows a machine to automatically recognize characters through an optical mechanism. Based on many factors human beings recognize many objects in the manner our eyes and brain perceives. The ability to comprehend these signals varies in each person. By considering these variables, we can understand the challenges faced by the technologist developing an OCR system.

Optical character recognition, OCR is the mechanical or electronic conversion of scanned images of handwritten, typewritten or printed text into machine-readable text. OCR is the process of translating scanned images of typewritten text into machine editable or manipulatable information [1]. 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. Sometime it is crucial to computerized printed texts so that they can be electronically stored, searched 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 [2].

All OCR systems include an optical scanner for reading text, and sophisticated software for analyzing images. Most OCR systems use a combination of hardware device called scanner and

supporting software to recognize characters, although some inexpensive systems do it entirely through software. Advanced OCR systems can read text in large variety of fonts, but they still have difficulty with handwritten text. OCR is widely used in legal profession, where searches that once required hours or days together can now be searched accomplished in a few seconds [3].

In OCR processing, the scanned-in image or bitmap is analyzed for light and dark areas in order to identify each alphabetic letter or numeric digit. When a character is recognised, it is converted into an ASCII code. Hardware designed expressly for OCR are used to speed up the recognition process. OCR is being used by libraries to digitize and preserve their holdings. OCR is also used to process checks and credit card slips and sort the mails. Billions of magazines and letters are sorted every day by OCR machines, considerably speeding up mail delivery.

In this paper a new concept and implementation of the machine learning is introduced in OCR which will open the field for researcher. So Machine learns whenever it changes its structure, program or data in such a manner that it's expected future performance improves. Machine learning usually refers to the changes in systems that perform tasks associated with Artificial Intelligence.[13]

## 2. METHODOLOGY

In this work the approach used in the OCR based on the Artificial Neural Network (ANN) Back propagation Algorithm and Multilayer Perceptron Neural Network Model i.e.

Artificial Neural Network: An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well [4].

Modelling system and functions using neural mechanism is a relatively new and developing science in computer technologies. The particular area derives its basis from the way neurons interact and function in the neural animal brain, specially humans. Neurons in the brain communicate with one another across special electro chemical links known as synapses. At a time one neuron can be linked to as many as 10,000 other although links as high as hundred thousands are observed to exist [5].

An artificial neural network is a system based on the operation of biological neural networks, in other words, is an emulation of biological neural system.

Unlike an animal brain, the traditional computer works in serial mode, which is to mean instructions are executed only one at a

time, assuming a uni-processor machine. A modern processor such as the Intel Pentium-4 or AMD Opteron making use of multiple pipes and hyper-threading technologies can perform up to 20 MFlops (Millions Floating Point executions) in a single second.[6]

An ANN is typically defined by three types of parameters:

1. Interconnection pattern between different layers of neurons
2. The learning process for updating the weights of the interconnections
3. The activation function that converts a neuron's weighted input to its output activation

The training process can be seen as an optimization problem, where we wish to minimize the mean square error of the entire set of training data. This problem can be solved in many different ways, ranging from standard optimization heuristics like simulated annealing, through more special optimization techniques like genetic algorithms to specialized gradient descent algorithms like backpropagation. In computational networks, the activation function of a node defines the output of that node given an input or set of inputs. A standard computer chip circuit can be seen as a digital network of activation functions that can be "ON" (1) or "OFF" (0), depending on input. This is similar to the behaviour of the linear perceptron in neural networks. However, it is the *nonlinear* activation function that allows such networks to compute nontrivial problems using only a small number of nodes.

*Characteristics of Artificial Neural Network*

- A neural network can perform tasks which a linear program will not be able.
- When an element of the neural network fails, it can continue without any problem by their parallel nature.
- Whatever a neural network learns earlier will not be required to be reprogrammed.
- It can be implemented in any application.
- It can be implemented without any problem.
- The neural network needs training to operate.
- The architecture of a neural network is different from the architecture of microprocessors therefore needs to be emulated.
- Requires high processing time for large neural networks.[7]

**Strengths of ANN:** Neural networks are very sophisticated modeling techniques capable of modeling extremely complex function. 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 also keep in check the curse of dimensionality problem that bedevils attempts to model nonlinear functions with large numbers of variables.

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

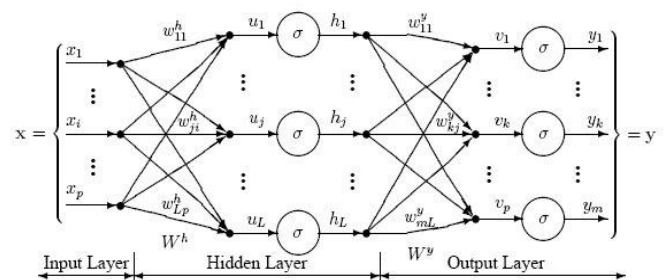
the case using some more traditional nonlinear statistical methods.

**Backpropagation:** It is a common method for training the ANN. There are 2 phases in the training of ANN by this method.

1. *Propogation:* Each propagation involves the following steps:
  - a. Forward propagation of a training pattern's input through the network in order to generate the output
  - b. Backward propagation of the propagation's output activations through the network using the training pattern's target in order to generate the deltas of all output and hidden neurons
2. *Weight updates:* According to desired output we want the weight will be updated using some functions that can be differ from each other in some sense.

Multilayer perceptron neural network also works on this theory to achieve their goal and output.

**Multilayer Perceptron Neural Network Model:** The following diagram illustrates a perceptron network with three layers:



**Fig1: Multilayer Perceptron Model**

This network has an input layer (on the left) with three neurons, one hidden layer (in the middle) with three neurons and an output layer (on the right) with three neurons.

There is one neuron in the input layer for each predictor variable. In the case of categorical variables, *N-1* neurons are used to represent the *N* categories of the variable.

**Input Layer** — A vector of predictor variable values ( $x_1...x_p$ ) is presented to the input layer. The input layer (or processing before the input layer) standardizes these values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the *bias* that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

**Hidden Layer** — Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight ( $w_{ji}$ ), and the resulting weighted values are added together producing a combined value  $u_j$ . The weighted sum ( $u_j$ ) is fed into a transfer function,  $\sigma$ , which outputs a value  $h_j$ . The outputs from the hidden layer are distributed to the output layer.

**Output Layer** — Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight ( $w_{kj}$ ), and the resulting weighted values are added together producing a combined value  $v_j$ . The weighted sum ( $v_j$ ) is fed into a transfer function,  $\sigma$ , which outputs a value  $y_k$ . The  $y$  values are the outputs of the network.

If a regression analysis is being performed with a continuous target variable, then there is a single neuron in the output layer, and it generates a single y value. For classification problems with categorical target variables, there are N neurons in the output layer producing N values, one for each of the N categories of the target variable. Most common activation functions are the logistic and hyperbolic tangent sigmoid function. Following hyperbolic tangent were used:

$$f(x) = (2 / (1 + e^{-2x})) - 1$$

and derivative:

$$f'(x) = f(x)(1 - f(x))$$

**Optical Language Symbols:** All languages are categorized by their own characters. These characters are either a delegate of a specific audio glyph, accent or whole words in some cases. In structure world language characters manifest various levels of organization. With respect to this structure there always is an issue of compromise between ease of construction and space conservation. Highly structured alphabets like the Latin set enable easy construction of language elements while forcing the use of additional space. Medium structure alphabets like the Ethiopic consume space due to representation of whole audio glyphs and tones in one symbol, but dictate the necessity of having extended sets of symbols and thus a difficult level of use and learning. Some alphabets, namely the oriental alphabets, exhibit a very low amount of structuring that whole words are delegated by single symbols. Such languages are composed of several thousand symbols and are known to need a learning cycle spanning whole lifetimes.

Representing alphabetic symbols in the digital computer has been an issue from the beginning of the computer era. The initial efforts of this representation (encoding) was for the alphanumeric set of the Latin alphabet and some common mathematical and formatting symbols. It was not until the 1960's that a formal encoding standard was prepared and issued by the American computer standards bureau ANSI and named the ASCII Character set. It is composed of and 8-bit encoded computer symbols with a total of 256 possible unique symbols. In some cases certain combination of keys were allowed to form 16-bit words to represent extended symbols. The final rendering of the characters on the user display was left for the application program in order to allow for various fonts and styles to be implemented.

At the time, the 256+ encoded characters were thought of suffice for all the needs of computer usage. But with the emergence of computer markets in the non-western societies and the internet era, representation of a further set of alphabets in the computer was necessitated. Initial attempts to meet this requirement were based on further combination of ASCII encoded characters to represent the new symbols. This however led to a deep chaos in rendering characters especially in web pages since the user had to choose the correct encoding on the browser. Further difficulty was in coordinating the usage of key combinations between different implementers to ensure uniqueness.

It was in the 1990s that a final solution was proposed by an independent consortium to extend the basic encoding width to 16-bit and accommodate up to 65,536 unique symbols. The new encoding was named Unicode due to its ability to represent all the known symbols in a single encoding. The first 256 codes of this new set were reserved for the ASCII set in order to maintain compatibility with existing systems. ASCII characters can be extracted from a Unicode word by reading the lower 8 bits and ignoring the rest or vice versa, depending on the type of ending (big or small) used.

The Unicode set is managed by the Unicode consortium which examines encoding requests, validate symbols and approve the final encoding with a set of unique 16-bit codes. The set still has a huge portion of it non-occupied waiting to accommodate any upcoming requests. Ever since its founding, popular computer hardware and software manufacturers like Microsoft have accepted and supported the Unicode effort.

### 3. DATA FLOW DIAGRAM FOR ANN BASED ON OCR

Step 1: Input the document into Optical Character Recognizer.

In this step input the hard document to convert into the soft.

Step 2: Scan the image to convert.

Step 3: Train the network for the appropriate working.

In this we train our network using the concept of Artificial Neural Network using backpropagation algorithm of training.

Step 4: Load the image into the network.

In this we load the scanned image into the tool which we developed to process it.

Step 5: Identify the lines present in the image.

Step 6: Detection of characters presented in the image.

In this we detect the characters present in the document and match it with our database i.e. Unicode for which we trained our Artificial Neural Network.

Step 7: Map the character image into pixel matrix.

If the match is found we map those characters into pixel matrix to show their existence.

Step 8: Get the output document which is in soft and editable format.

Finally we get the soft copy of our hard document in the editable and manipulatable format.

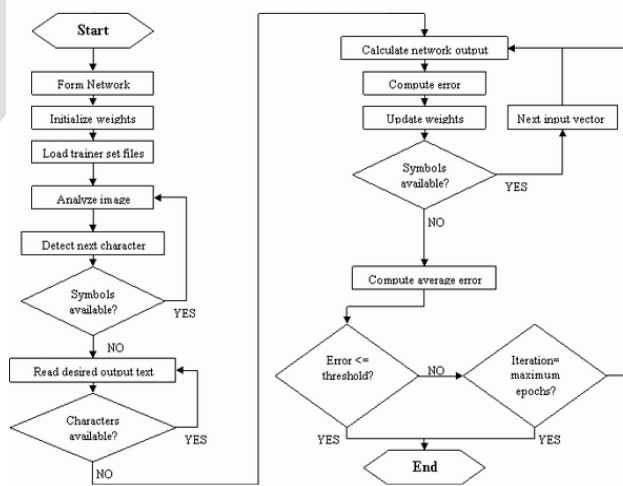


Fig 2: Data flow in OCR

## I. ALGORITHM FORMULATED FOR ANN BASED OCR

OCR technology requires both hardware and software. In addition, sophisticated OCR systems require an additional hardware i.e. scanner which can scan various printed data. An optical scanner scans the text on a page, then breaks the fonts down into a series of dots called a bitmap. The software can read most common fonts and distinguish where lines start and stop. This bitmap is then translated into computer text [11].

Newly developed algorithm on which OCR works after completion of document scanning is illustrated below:

1. Start with the input of a document.
2. Formation of the network.
3. Initialize the weights.
4. Load the trainer set files into machine.
5. Start analysis of image
6. Detect the characters
  - If(symbol available)
  - Goto step 5
  - Else
  - Read desired output text
  - If (character available)
  - Read the output text
  - Else
  - Goto step 7
7. Calculate the network output
8. Compute the errors and update the weights
  - If(errors <= threshold)
  - End of process
  - Else if(Iteration= maximum epochs)
  - Goto step7
  - Else
  - End of process

The necessary steps are:

1. Preparing the sequence of input symbol images in a single image file
2. Typing the corresponding characters in a text file (\*.cts [character trainer set] extension)
3. Saving the two in the same folder

The application will provide a file opener dialog for the user to locate the \*.cts text file and will load the corresponding image file by itself.

The tool which is developed as part of this research is based on Machine learning concept of “Training of ANN” i.e. OCR for conversion of hard copy to soft copy. In the training of ANN with a set of input and output data, we wish to adjust the weights in the ANN, to make the ANN give the same outputs as seen in the training data. On the other hand, we do not want to make the ANN too specific, making it give precise results for the training data, but incorrect results for all other data. When this happens, we say that the ANN has been over-fitted.

## II. RESULT AND DISCUSSION

The tool has been developed using .NET language, the developed tool can be used to convert hard copy document into digital format which is editable and manipulatable in the computer readable format. The newly developed tool is modular and simple to test and implement. Programme is developed in such a way that same modules are used in training, working and testing phase.

OCR is mainly used in libraries to digitize and preserve precious documents. OCR is also used to process checks and credit card slips and sort the mail. Billions of magazines and letters are sorted every day by OCR machines, considerably speeding up mail delivery.

The basic steps in testing input images for characters can be summarized as follows:

### Algorithm:

Step 1: loading of image file

Step 2: analyse image lines for character recognition and map symbol vectors

Step 3: for each character line detect consecutive character symbols

3.1: analyse and process symbol image to map into an input vector

3.2: feed input vector to network and compute output

3.3: convert the Unicode binary output to the corresponding character and render to a text box

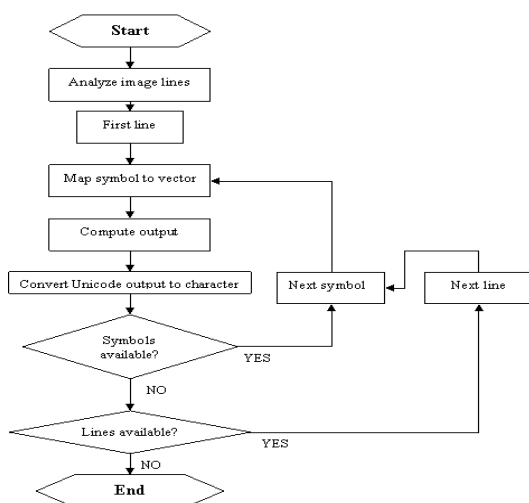


Fig 3: Flowchart for testing in OCR mechanism

## III. INTERFACE SCREENSHOTS

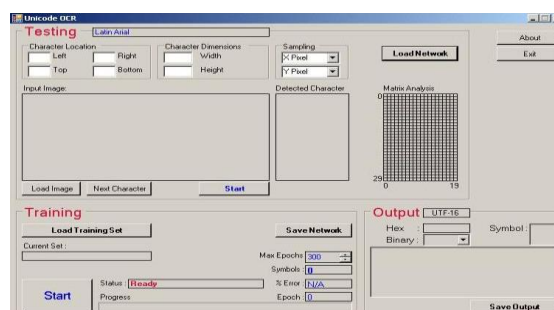


Fig 4: First input screen of system



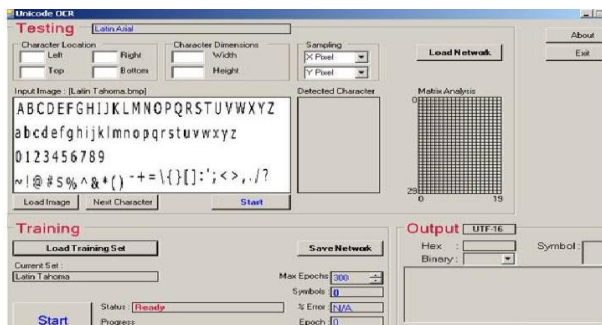


Fig 5: Screen after first input in OCR

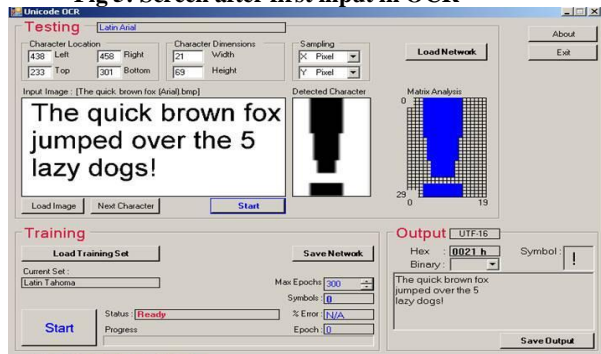


Fig 6: Output Screen of OCR

#### IV. CONCLUSION AND FUTURE WORK

OCR technology provides fast, automated data capture. Proper scanning, good form design, sufficient data validation, and targeted manual review deliver accurate results with huge savings over manual processes. Errors can be avoided through solid planning and good follow-through. In the present work the artificial neural network has been trained and tested for a number of widely used font type in the Latin alphabet. In the current work a software tool is developed in .NET which is scalable, i.e. more number of features can be added into the existing research like more number of fonts from any typed language alphabet.

The tool which is developed gives about 85% correct result in handwritten documents, which is very good in this area. The effort are continued to increase the effectiveness of results. If the

developed tool is not able recognizing letters it means that quality of written documents is poor. .

The results achieved in this work are good but every work has certain limitations that can be improve in the future. In this work can extend this work like non bold character can also be recognized. Non proper spacing sentences can also be accepted in future. Confusion between 0 and D, 1 and I can be removed.

Now we are using scanner to scan the images but in future the existing toll will modified is such a way that it can use laser to scan the images. Now tool is trained to recognizing images of characters but in future it will be trained in such a way that it will able understand the string of characters also.

#### REFERENCES

- [1] Optical Character Recognition (OCR) *What You Need to Know* By Phoenix Software International
- [2] [http://en.wikipedia.org/wiki/Optical\\_character\\_recognition](http://en.wikipedia.org/wiki/Optical_character_recognition)
- [3] [http://www.webopedia.com/TERM/O/optical\\_character\\_recognition.html](http://www.webopedia.com/TERM/O/optical_character_recognition.html)
- [4] [http://www.doc.ic.ac.uk/~nd/surprise\\_96/journal/vol4/cs11/report.html#Introduction%20to%20neural%20networks](http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Introduction%20to%20neural%20networks)
- [5] Artificial Intelligence and cognitive science  
<http://ai.stanford.edu/~nilsson> [6] Offline Handwritten Recognition Using Artificial Neural Network  
<http://wilsonat@mrs.umn.edu>
- [7] <http://www.learnartificialneuralnetworks.com/>
- [8] Using Neural Networks to Create an Adaptive Character Recognition System By Alexander J. Faabor, Cornell University, Ithaca NY (May 14, 2002)
- [9] Hand Printed Character Recogniser using Neural Network  
Shahzad Malik April 2000 [10] Neural Networks and Fuzzy Logic  
By V. Rao, H.Rao MIS Press, New York
- [11] Neural Networks  
<http://www.statsoft.com/textbook/stneunet.html>
- [12] <http://www.wisegeek.com>
- [13] <http://robotics.stanford.edu>