Optical Character Recognition Using Artificial Neural Networks Approach

Siddhi Sharma¹, Neetu Singh²

¹M.Tech 2nd Year Student, ²Assistant Professor, Dept. (C.S.E), MUST – FET, Lakshmangarh, India

Abstract—The recent advances in computer technology many recognition task have been automated. OCR, Optical Character Recognition is a scheme of converting the images of typewritten or printed text into a format that is understood by machine. The goal of OCR is to classify the given character data represented by some characteristics, into a predefined finite number of character classes. For the recognition to be precise certain topological and geometrical properties are calculated, on the basis of which characters are classified and recognized. These properties are called features and, the collection of such features is called vectors which help in defining the character uniquely. The main aim of this attempt is to explore the utility of Artificial Neural Networks based approach to the recognition of characters. A unique multilayer perception of neural network is built for classification using BackPropagation learning algorithm.

Keywords—Artificial Neural Network, BackPropagation Algorithm, Multilayer Feed Forward Architecture, Optical Character Recognition, Pattern Recognition.

I. INTRODUCTION

A Pattern is a set of objects or phenomenon or concepts where the elements of sets are similar to one another in certain ways or aspects [1]. Pattern Recognition is one of important and vast fields of computer intelligence, which focuses on the recognition of patterns and regularities of data. It is an act of taking input of raw data and extract specific or unique features to recognize that data. 1985(Watanabe) said that pattern recognition can be looked as categorization problem, as inductive process, as structure analysis, as discrimination method and so on [2]. Optical Character Recognition is an application of pattern recognition. OCR is a method of converting a scanned image of typewritten, handwritten or printed text into computer readable format [3]-[5] i.e. a format that is understood by machine. Now when a page is scanned, it is stored in TIF format. When an image is displayed on the monitor, humans can read it, but to a computer, it’s just a sequence of black and white pixels. Computer does not recognize any word in image. The main aim of OCR is to look at the image and try to decide if these pixels are representing a particular letter or not.

II. ARTIFICIAL NEURAL NETWORK

An Artificial Neural is basically an engineering approach of biological neuron [6]-[7] i.e. Neural Networks basically aim at mimicking the structure and functioning of the human brain, to create intelligent behavior.

A Neural Network is a massively parallel distributed processor made up of simple processing units which have natural propensity for storing experiential knowledge and making it available for use. It resembles to brain in two aspects. First, Knowledge is acquired by the Network from its environment through a learning process. Second, Interneuron connection strength is used to store acquired knowledge.

In Neural Network, each node perform some simple computation and each connection conveys a signal from one node to another labeled by a number called “connection strength” [8].

![Simple Neuron Diagram](attachment:neuron_diagram.png)

1. A Simple Neuron [8]

Linear Combination \( U_k \), \( U_k = \sum w_{kj} * x_j \)
Induced Local Field \( V_k \), \( V_k = U_k + b_k \)
Activation function defines the value of output \( Y_k \), \( Y_k = \phi (V_k) \)

The Activation function used here are of different type, Threshold Activation Function, Piecewise Linear Activation Function, Sigmoid Activation Function, Signum Activation Function etc.

Learning is formally defined as a process by which free parameters of a Neural Networks are adapted through a process of simulation by the environment in which the network is embedded.
Once the system begins to learn containing some initial weight values, as the learning process increase weight values keeps on changing and provide the final output at end. Learning can be classified as: First, Supervised Learning i.e. learning with Teacher, Second, Unsupervised Learning i.e. learning without Teacher.

Typical pattern recognition systems are designed using two pass. The first pass is a feature extractor that finds features within the data which are specific to the task being solved. The second pass is the classifier, which is more general purpose and can be trained using a neural network and sample data sets.

III. BACKPROPAGATION ALGORITHM

As Optical Character Recognition is defined as a Multiclass Problem, amongst various classification methods [9]-[10], we have used, Multilayer Feed Forward Architecture, which contains an Input Layer, an Output Layer and one or more Hidden Layer [11]. As the number of Hidden Layer increases the complexity of network also increases.

Back-Propagation Neural Network (BPNN) algorithm is the most popular and the oldest supervised learning multilayer feed-forward neural network algorithm proposed by Rumelhart, Hinton and Williams [12]-[14].

Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by you. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Steps of Back Propagation Algorithm are as follows [18]-[20].

A. Weight Initialization

Set all weights and Node threshold to some small random values.

B. Calculation of Activation

1) Input Unit: The Activation Level of the input unit is determined by the instances presented to the Network.

2) Hidden unit and Output unit: The Activation Level $O_j$ of Hidden unit and Output Unit are determined by:

\[ O_j = F \left( \sum w_{ji} * O_i - \theta_j \right) \]

Where $w_{ji}$ = weight from input $O_i$ to unit $j$

$\theta_j$ = Node threshold at unit $j$

F = Activation Function

C. Weight Training

1) Weight Change: Start at output unit and work backward to hidden layer, recursively adjust the weight by-

\[ w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} \]

2) Weight Change Computation: The weight change is computed by-

\[ \Delta w_{ji} = \eta \delta_j O_i \]

Where $\eta$ = learning rate, $\delta_j$ = error gradient

The error gradient is given as follows at Output Unit

\[ \delta_j = O_j (1 - O_j) (T_j - O_j) \]

And for Hidden Unit

\[ \delta_j = O_j (1 - O_j) \sum_k \delta_k w_{kj} \]

Where $T_j$ = Target Value, $O_j$ = Actual Output Value, $\delta_k$ = Error Gradient at unit k to which a connection point at unit j.

D. Repeat Iterations until convergence.

IV. METHODOLOGY

The entire Process is broken-down into Pre-processing, Feature Extraction and then it is passed through the Artificial Neural Network for training and recognition. The steps are shown in figure 3.
A. Pre-processing

In this stage the acquired image is passed through various phases, because the image cannot be directly passed to the recognition system [15].

1) Conversion to Gray Image: Converts the truecolor image RGB to the Grayscale intensity image by using a function called rgb2gray ()

2) Segmentation: Converts any image into a series of black text written on a white background. Thus, it induces uniformity to all the input images. This also reduces computational power as it to deal with two colours i.e. Black and White.

3) Normalization: The process of normalization is used to resize the acquired images into same size so that further processing is applied.

4) Noise Reduction: The process is used to reduce the noise if present before the image is subjected to ANN.

5) Complement Image: The process of converting the white pixels of an image to black pixels and vice versa.

The result for pre-processing is shown in figure 5.

B. Feature Extraction

It is the process which is used to serve different ideas [16], the process is used to extract properties that are used to recognize the character uniquely, and on the same time it is used to extract properties that are required to differentiate between similar characters.
6. Feature Extraction Process

3) Vertical Detail Image: - For an Input Image the Vertical Points of image are displayed in this portion.

4) Diagonal Detail Image: - For an Input Image the Diagonal Points of image are displayed in this portion.

7. The output of stage 2 Feature Extraction showing different plots of Low-pass Approximation, Horizontal Detail Image, Vertical Detail Image, and Diagonal Detail Image at Level 1 and at level 2.

V. TRAINING AND CLASSIFICATION

Before the recognition to be done, the Artificial Neural Network must be ‘trained’ so that the network gets a potential of mapping various inputs to their corresponding output [17], so that the system classify various characters. For training the Neural Networks, we have different feature vectors obtained from the database using the above feature extraction technique. The technique uses 6 different geometrical features to extract 48 parameters are fed into ANN. There are 25 samples taken for each character to train the Network. Thus a Matrix of 48X250 values is fed to the network to receive 10 different values at output one for each character in database. The experimental results are shown for 10 characters.

8. Input and Output Parameters of ANN during the phase of Training and Classification.

From a set of 250 samples 175 samples of different characters are chosen to train the Neural Networks and the rest of the untrained samples are taken to test the Network. It shows a recognition rate of 84.8% for 10 class problem in which out of 75, 65 samples are correctly recognized. The figure below shows the performance of the system in Training, Testing, and Validation phases.


This surge in performance results from false recognitions due to similarities between ambiguous characters such as E & B, I & J etc.

The figure 10 below shows the plots the training state from a training record (tr) returned by train.
The below figure shows the error plot for the samples provided for testing, training and validation.

Recognition Rate is defined as a ratio of Number of character (Patterns) found correctly to the total Number of Patterns.

\[
RR = \frac{\text{No: of characters found correct}}{\text{Total Number of Patterns}} \times 100
\]

\[
RR = \frac{212}{250} \times 100 = 84.8\%
\]

12. Confusion Matrix obtained for 10 Class Problem, recognition rate obtained 84.8%.

VII. CONCLUSION AND FUTURE WORK

The experimental results illustrates that the Artificial Neural Network can be applied successfully to solve the Optical Character Recognition problem with English characters. The recognition rate OCR with the characters is quite high as shown in the above section.

The system can be enhanced for the recognition of all 26 characters. However, other kinds of pre-processing and neural network models may be tested for a better recognition rate in the future. The test set used in this experiment is of at least five different types of fonts. This can be tested for a greater number of fonts. Better feature extraction techniques must be used for better performance.

Similarly a system for Hand–written characters can be established. In hand-written character recognition system, the fragmentation of characters and the variation in shape of characters are considerably greater compared to printed documents, accordingly the system built must be good enough to perform better and provide better precision of results.
REFERENCES


