NEURAL NETWORK (NN) BASED METHOD FOR OPTICAL CHARACTER RECOGNITION (OCR)
4.1 Introduction

Feature extraction is an important stage in OCR and plays a significant by considering the character representation as rotation, translation and scale change. Invariant and Zernike moments are used for justify the above said representation. Complex polynomials set are used to make Zernike moments targeting the formation of orthogonal set.

Discrimination of document and offline-handwritten characters features is generally maintained by ‘feature extraction’. Effectiveness of extraction process and suitable classifier choice decides the OCR performance. Some of the features that are unique to specific signature will be collected in feature extraction process. Then these features will be used as script classifier input. Such features will fulfill the characteristic preservation of characters by focusing on relevant information. Feature extraction processes are found to vary with respect to robustness, properties of invariance and distortions/variability of characters sets.

In character recognition system, feature extraction will be applied immediate to the pre-processing stage. Pattern recognition aim is to consider the given pattern and assigning it as possible output. This will be done in two phases: feature selection and classification. Feature selection has vital role in the whole process because the classifier cannot to recognizable from bad/poor feature selection. Criteria to choose features given by Lippman are: “Features should contain information required to distinguish between classes, be insensitive to irrelevant variability in the input, and also be limited in number, to permit, efficient computation of discriminant functions and to limit the amount of training data required”.

During the pattern classification construction, feature extraction play significant role aiming at relevant information extraction by characterizing individual class. Within the process, feature vectors will be formed by extracting relevant features from alphabets/objects. To recognize the input unit with target output unit, classifiers will use these vectors. As these features allow fairly easy to distinguish, classification between different classes will be easier for classifier by looking at them. Retrieving most important data from raw data is the primary objective of feature extraction. It finds the set of parameters which will precisely and uniquely define the character shape. Feature vector becomes the identity of each character in feature extraction phase. It extract a set of features thus maximizing recognition rate through generating set of similar features for same symbol with variety of instances. Most popular feature extraction methods are Deformable templates, Zernike Moments, Template matching, Zoning, Gradient feature, Unitary Image transforms, Projection Histograms, Graph description, Contour profiles, Gabor features, Fourier descriptors, Spline curve approximation and Geometric moment invariants.

4.2 Importance of Feature Extraction

Features will be obtained by applying feature extraction technique to segments and then followed by applying classification/post-processing methods after pre-processing and effective segmentation is. Recognition system efficiency is impacted by feature extraction phase and hence it is essential to focus on it. High recognition performance can be achieved by considering feature selection. It is given as “extracting from the raw data information that is most suitable for classification purposes, while minimizing the within class pattern variability and enhancing the
between class pattern variability”. Utmost care should be taken while considering suitable technique for feature extraction and should be based on the input to be applied. Looking into various available techniques is essential covering several possible cases.

### 4.3 A Detailed Review of Feature Extraction in Image Processing Systems

An Artificial Neural Network is inspired by brain which is a biological nervous system. It comes under the information processing domain to solve specific problems and comprises large number of heavily interconnected processing elements such as neurons working in unison. It is configured for data classification and pattern recognition via process of learning. It processes the information as similar as brain do. It learns through examples. Synaptic connections existed between the neurons will be adjusted for learning in biological systems. Basic neuron has multiple inputs and single output. It contains two modes of operation: training and using mode. In former case, neuron will be trained based on specific patterns of input. In later case current output is the output associated with the input detected from taught patterns of input. The training rule is utilized when the given input pattern is not in the list of taught patterns of input. Neural network has many applications out of which: reasoning, association and classification are widely used. A feed-forward neural network is used for implementing pattern recognition. It is trained for association between outputs and patterns of input. It is then identifies the input pattern to deliver the output with respect to the output pattern.
Effectively, four different approaches to pattern recognition can be highlighted significantly.

*Statistical pattern recognition:*

In statistical pattern recognition, the given exercise is depicted as a part of composite hypothesis testing of the datum which is generated from a particular class or delivered from the space of measurements to the space of classes’ regression. Here each hypothesis is pertained to a particular premise. The main challenge in this method is the way of solving which involves statistical computation of various class conditional probabilities although such method is one of the oldest and widely used.

*Syntactic pattern recognition:*

Here each pattern is depicted as a composition of primitives or sub-patterns tied together with respect to the rules of a grammar string generated from class association. Automata with respect various grammars will be used for parsing the operations in class identifications. Main hurdles in this method which are responsible for limited applicability are grammatical inference and parser design.

*Knowledge-based pattern recognition:*

It is derived from the advances evolved in artificial intelligence (AI) with rule-based system. Every rule in AI is in the form of a clause reflecting the proof of specific class presence. This method will generate sub-problems as follows: way of construction of rule-based method and integration mechanism of evidence outputted by the used rules.
Neural Pattern Recognition:

Artificial Neural Network (ANN) is an advanced method of working for pattern recognition. It encompasses various models and contains two important points to be noted: (i) these models are composition of neurons which are large in number and similar units structurally and functionally. These neurons will connect various configurations by weighted links. (ii) Parameters of Ann’s model will be generated from I/O paired data sets supplied by training which is an estimation process.

Efficient characters recognition method based on a neural network is proposed. The proposed method learns from the training data set leading to cent percent accuracy in recognition of character. OCR technology enables a machine to automatically recognize characters verified by optical system. Humans identify several things through a mechanism. By reviewing these variables, the challenges faced in technologies of OCR system can be verified. The end aim is to simulate the human-perceptive abilities leading to a computer trained to perform fundamental document objectives for a given input document. A representative image diagram of the
typical OCR system is illustrated in Figure 4.1. Each stage of its processing is known to be inferred for numerous problems, while it can effects the overall system’s efficiency. An OCR system when integrated into one phase would be advantage to improve the efficiency and this is what present research proposes.

4.4 Feature Extraction Model (Training Phase)

The profile information is used for the preliminary classification of a text document image into constituent lines using Horizontal Profile. These lines are further segmented into words using Vertical Profile and the words. Around 100 document samples with different font size are considered for experimentation. In the classification process, all the characters are extracted and a database is created with these characters. Features are extracted for these components in training phase.

The present work is experimentally analyzed in two stages, training phase and testing phase.

Figure 4-2 Procedure for Extracting Feature from the Characters [152]

In a diagonal feature extraction method [152], every character image having 90x 60 pixels size will be separated in to
fifty four equal zones having 10x10 pixels each. From each zone pixels, features will be extracted by moving across diagonals of respective sets. Single sub-feature will be identified by summation of 19 diagonal lines in each zone and foreground pixels across each diagonal line. Then such values will be averaged to get a single feature value. This will be repeated for all zones.

Few zones will be exists with empty diagonals with respect to foreground pixels and are having zero feature values. In the end, each character is represented with 54 extracted features. 9 and 6 features are obtained additionally by averaging respective row and column-wise zones placed values. Finally, each character is represented with a 69-length feature vector that is nothing but 54 + 15 features.

Once trained, the neural network should be able to recognize not only to obtain the perfect pattern, but also the corrupted or noisy versions. In fact to improve and validate the recognition ability, deliberately added some noisy versions of the patterns into the training dataset for training the network (say one in five). There could be other patterns of noise addition to generate the training dataset so as to introduce the randomness in training.

![Network error diagram](image)

**Figure 4-3 Use of Validation Sets [153]**
Local Minima problem may occur in Back-propagation algorithm. Weights will be changed so that the error will be minimum and hence the problem occurs. As shown in Figure 4.3, error may have to shoot-up with in a fall. The procedure may be gets stuck and the corresponding error may not will not reduce. One simple solution to this problem is to weights resetting to random numbers and then again training.

For better performance of NN, the hidden layer plays crucial role depending upon number of neurons. They will not interact with outside environment directly and they show influence on overall result. Trail & error method and Rule of thumb method are used to get number of hidden nodes in each layer. *Trial & Error Method* is characterized by varied and repeated attempts till the success are achieved or agent ends try. Forward approach starts with two hidden neurons and training/testing of NN is done. No. of hidden neurons is shoot-up and it will be repeated till training/testing is improved. Backward approach is opposite to that of the forward approach where an initial guess is assumed of no. of hidden neurons and again trained/tested NN. No. of hidden neurons will be decreased and trained/tested NN again. This will be repeated till the training/testing is improved. *Rule of thumb method* is identifying neurons number for using in the hidden layers. No. of hidden neurons will be between the input/output layer size. This number should be two-thirds of size of the input + output layer. Also, hidden neurons numbers should be less than input layer size times twice.
4.5 Proposed Algorithm for Character Recognition

Figure 4-4 Block Diagram of Proposed Algorithm

Step 1: Given pattern is applied to network input layer and is processed via network before reaching output units.
Step 2: These inputs are propagated through the network until they reach the output units.
Step 3: Forward propagation produces predicted output due to the fact that it is a controlled learning method.
Step 4: Predicted outputs are subtracted from expected outcome and an error will be generated correspondingly.
Step 5: Errors are feedback through network using contribution computation of every processing unit (hidden) and the respective adjustments necessary to generate expected output are found.

Step 6: Based on the learning experience of NN, weights of connection will be updated. Once the NN is experienced then the expected outcome for any of the input patterns can be generated effortlessly.

Step 7: Finite number of patterns containing an input and targeted output are used for supervised training of NN.

Step 8: Activations of patterns will be passed by neurons to the later layer (inside hidden layer).

Step 9: Hidden layer neurons outputs are generated using the weights that will process patterns with a threshold function and an optional bias.

Step 10: Activations from the output layer will determine the final output of NN.

Step 11: Above steps will be repeated with every pattern used for NN training and each iteration is known as a cycle/epoch. These iterations will be repeated till the error is within an acceptable tolerance range.

Step 12: Threshold value in the output layer adjustment is done by multiplication of output calculated error and the learning rate.

\[
\phi_k = \alpha e \tag{4.2}
\]

Where,

\( \alpha \) is the learning rate

\( e \) is the error function

\( \phi_k \) is the threshold function

\[
Y = f(i) = f\left\{\sum_{i=1}^{n} X_i \ W_i - \phi_k \right\} \tag{4.3}
\]
Where,

\( X_i \) is the input

\( W_i \) is the weight

\( F(I) \) is the activation function

**Step 13:** To analyze the performance of classification of untrained patterns by second set of inputs, they will be tested only after NN has learned the right classification based on the training set.

### 4.6 Related work

Printed character recognition techniques are used in different areas by using Artificial Neural Networks. With the back propagation technique used in the present study, the efficiency was 98\% for white background and 96.87\% for black background. A discussion on contemporary researchers’ work is given in Table 4.1.

### 4.7 Conclusion

In this work, back propagation algorithm for character recognition is adopted. The presented approach computes the rate of error efficiently and increases the character recognition accuracy. Offline character recognition is a difficult problem, not only because of the great quantity of variations in font type and size, but also, because of the overlapping of the neighbor characters.

Recognition strategies strongly based on the topology of the data to be recognized. Using BP-ANN in OCR gives good results. However, the OCR affects by many factors like noise, brightness, coloured background. So pre-processing is necessary to be used in documented images as an initial step for character recognition systems to remove the effects of these factors. Each image requires different pre-processing techniques depending on the effect of the factor that may affect the quality of it.
Table 4.1 Discussion on contemporary researchers’ work

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Proposed by</th>
<th>Method based on</th>
<th>Recognition Efficiency (%)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rao &amp; Ajitha [1995]</td>
<td>Feature Based Approach</td>
<td>78 to 90, 91 to 95</td>
<td>Across different subjects. Reference and test sets from the same subject.</td>
</tr>
<tr>
<td>3</td>
<td>Negi, et al. [2001]</td>
<td>Template matching</td>
<td>92. 97.3 70.1</td>
<td>With no post processing for Hemalatha font for the chosen newspaper font.</td>
</tr>
<tr>
<td>4</td>
<td>Lakshmi &amp; Patvardhan [2002]</td>
<td>Using the k-nearest neighbor algorithm</td>
<td>92</td>
<td>For different fonts with font sizes 20, 30 and 35.</td>
</tr>
<tr>
<td>5</td>
<td>Jayaram [2015]</td>
<td>Artificial Neural Networks</td>
<td>93.87 97.9</td>
<td>For Vemana font. For Pothana font.</td>
</tr>
<tr>
<td>7</td>
<td>Srinivas, et al. [2007]</td>
<td>Computation using Zero-crossing features and classification using a 2-stage classifier</td>
<td>93.2</td>
<td>for a single font</td>
</tr>
</tbody>
</table>