1. **Proposed Recognition System**
   1. **Problem definition**

The purpose of this project is to take handwritten digits as input, process the digit, train the neural network algorithm to recognise the digit and compare the merits and demerits of optical character recognition and neural network-based recognition.

One of the primary means by which computers are endowed with humanlike abilities is through the use of a neural network. Neural networks are particularly useful for problems that cannot be expressed as a series of steps, such as recognizing patterns, classifying them into groups, series prediction and data mining.

Pattern recognition is perhaps the most common use of a neural networks. The neural network is presented with a target vector and also a vector which contains the pattern information, this could be an image and handwritten data. The neural network then attempts to determine if the input data matches a pattern that the neural network has memorized.

A neural network for classification is designed to take input samples and classify them into groups. These groups may be fuzzy, without clearly defined boundaries. This project concerns detecting free handwritten digits.

* 1. **Existing system**

In the existing systems, Support Vector Machine algorithm is used for classification which is slower. So, in the proposed system, neural networks are used for classification and identification as this system is faster and it is a more efficient than Support Vector Machine. In the existing system using neural networks, the accuracy is limited to 90%, but in the proposed system the accuracy is increased to 97%.

* 1. **Proposed system**

In this system, the proposed recognition system is described. It consists of data set extraction, pre-processing, classification, recognition and post-processing stages. The block diagram of the proposed system is shown. Initially an image of the format jpg or gif is provided as input to the system. The image must be noise free and in black and white. Prior to segmentation we can select a part or whole of the image. This process is called cropping.

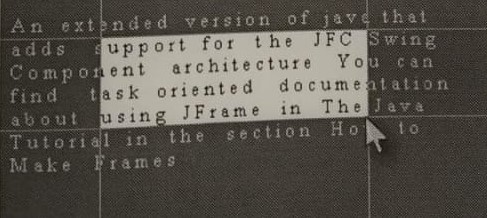


Figure 2.1 Cropping of an image

This project mainly consists of two parts:

1. Scanned document recognition
2. Handwritten character recognition

The various stages involved in scanned document recognition are segmentation is used to find the bounds of the text in the image, recognition involves down sampling and pattern matching, training involves down sampling and pattern storing and editing is the final step which involves cut, copy and paste and search and replace operations.

* **Segmentation –** Initially the boundaries of the cropped image is to be found. By staring from left most part of the image we continue to scan through a vertical line until we encounter a black pixel on the line. The x-coordinate specify the left boundary of the image. By similar means the four boundaries of the cropped image are obtained. After finding the four boundaries the lines present in the paragraph are found followed by the individual characters in that line are found.

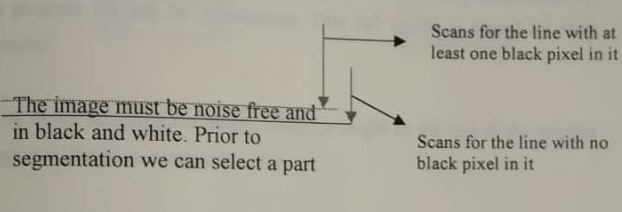


Figure 2.2 Line Segmentation

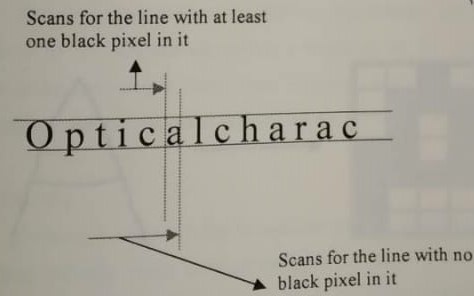


Figure 2.3 Character segmentation

* **Recognition –** This is the key phase wherein the segmented image is recognized character by character. This phase consists of two steps:

1. **Down sampling** – It is a process of converting an image into low resolution image. In this program we use 5\*7 resolutions. This 5\*7 matrix is stored for each character. By down sampling the image, whatever might be the size of the original image it is stored in a consistent size. The entire image is divided into 35 blocks. If any pixel in a block of the original image is black, then the corresponding value in the 5\*7 matrix is true.

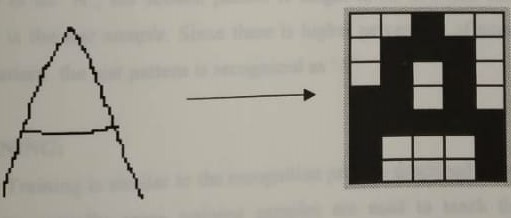


Figure 2.4 Down sampling

1. **Pattern matching** – It is a process wherein the already trained patterns of characters are compared with the test characters and the best match is returned. The accuracy of pattern matching depends on how well the neural network is trained.

The training samples are shown below. The first pattern is taught to be ‘A’ and the second pattern is taught to be ‘E’. The one present below is the test sample. Since there is higher percentage of match to the first pattern the test pattern is recognized as ‘A’.

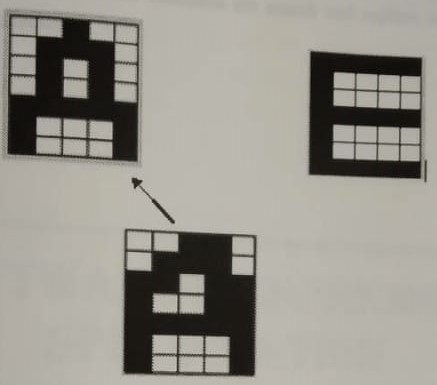


Figure 2.5 Pattern matching

* **Training –** Training is similar to the recognition process described above except that here initially some training samples are used to teach the neural network. Training process is implemented by evaluating the error of the weight matrix (an array comprising of the weights so that the error is output neurons) and then adjusting the weights so that the error is minimized as much as possible. At the end of the training process the final weight matrix obtained is used to classify the patterns.
* **Editing –** This is the phase where the recognized text can be manipulated so as to correct the errors present in the original scanned document. The functions provided in the editor are cut, copy, paste which are used for addition and deletion of text. Some other functions are search and replace which are used as spell checkers.

The various steps involved in Neural network-based digit recognition are:

* + 1. **Data extraction**

The input to the system is a scanned image of the digit with a suitable format like JPEG. In this proposed system, MNIST dataset is used which was constructed from a number of scanned images from the National Institute of Standards and Technology. Handwriting samples are taken from 250 people for training and 250 other people for testing. The data set consists of 70,000 samples of scanned handwritten images. Out of this, 60,000 are used to train the network and the remaining 10,000 samples are used for testing. The set of images given for training are completely different from the images given for testing so as to know how better the system recognizes the digits.

* + 1. **Pre-processing**

In this stage, the images are centralised and normalised in position and a sequence of operations are performed on it to make it suitable for classification like dilation of edges, noise removal, etc. It involves data cleaning, data integration, transformation and reduction.

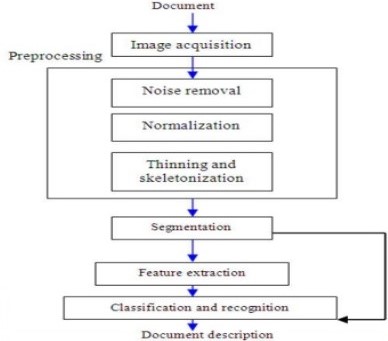


Figure 2.6 Block diagram of proposed recognition system

* + 1. **Classification**

A three-layered neural network is used for the purpose of classification and recognition as shown in the figure. It consists of an input layer, a hidden layer and an output layer. The first layer is input layer. The scanned data images contain 24\*24 pixels. So, the input layer consists of 784 neurons with each one representing a pixel value in the input image. The scanned images are grey scale images. So, an input 0 indicates white and an input 1 indicates black and any in between values represent the grey shade intensity which increases from 0 to 1. The middle layer is known as hidden layer. There can be any number of neurons in this layer and accuracy increases by increasing the number of hidden neurons because the system can learn and recognize the digits in a better way as the hidden neurons are increased. The final layer is output layer and in consists of 10 neurons each indicating the digits from 0 to 9 respectively. Similarly, for character recognition the output layer consists of 26 neurons each indicating the alphabets from a to z respectively.

* + 1. **Identification**

This stage is responsible for recognizing the digits based on the hidden neurons and output neurons. For this, to minimize the error in detection, a cost function is introduced. This cost function is minimized using stochastic gradient descent in which the data is divided into mini batches and then the gradient of each batch is evaluated. The main algorithm used in classification and recognition is feed forward back propagation algorithm. The output layer contains 10 neurons indicating from 0 to 9. Based on the outputs from the neurons in the hidden layer, if the first neuron is fired, it recognises the digit as zero, if the second one is fired, it indicates one and similarly if the last neuron is fired, it indicates 9.

* + 1. **Post-processing**

The last stage after identification is post-processing. In this stage the recognized digit is printed in a structured format.