# Handwritten ZIP code Classification Using

# Artificial Neural Networks

**K. Siva Kumar , D.Uma Devi**

## Abstract

A new mechanism proposes the use of Artificial Neural Network (ANN) models for classifying handwritten zip-codes and also presents new techniques for to extract digits from zip-code images. Remove the slope and slant from handwritten zip-codes and normalize the size of Digit images with our proposed methods. Experiments have been conducted on handwritten ZIP codes and the classifying rates up to 99% achieved.

**Keywords:** ZIP code, Handwritten ZIP code, Pattern Recognition, Neural Network, Artificial Neural Network, Classification, NPR Tool, Slope correction and Slant correction.

## 1. Introduction

ZIP codes are a system of postal codes used by the United States Postal Service (USPS) since 1963. The term *ZIP*, an acronym for Zone Improvement Plan, is properly written in capital letters and was chosen to suggest that the mail travels more efficiently, and therefore more quickly, when senders use the code in the postal address. The basic format consists of five decimal numerical digits. An extended ZIP+4 code, introduced in the 1980s, includes the five digits of the ZIP code, a hyphen, and four more digits that determine a more precise location than the ZIP code alone.

Recently, a lot of works was done by depending on the computer; In order to let the processing time to be reduced and to provide more results that are accurate, for example, depending on different types of data, such as characters and digits and the numbers are used frequently in normal life operation. In order to automate systems that deal with numbers such as postal code, Banking account numbers and numbers on car plates. And an automatic recognition number system is proposed in this study.

Digit recognition has been extremely found and studied. Various approaches in image processing and pattern recognition have been developed by scientists and engineers to solve the digit classification and recognition problems [1, 2]. That is because it has an importance in several fields and it may probably be used in checks in banks or for recognizing numbers in cars plates, or many other application. In this study, system for recognized of digits is built, which may benefit various fields, the system Concerning on isolated digits, the input is considered to be an image of specific size and format, the image is processed and then recognized to result of an edited digits.

## 2. 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. 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. The behaviour of an ANN (Artificial Neural Network) depends on both the weights and the input-output function (transfer function) that is specified for the units.

The Single Neuron is a model based on the biological neuron which is found in the

brain. It receives inputs and add's them up. The neuron shall then produce an output if the sum is greater than the "threshold value". The Neuron's output is either on or off. The output depends only on the inputs. A certain number must be on at any one time in order to make the neuron fire.
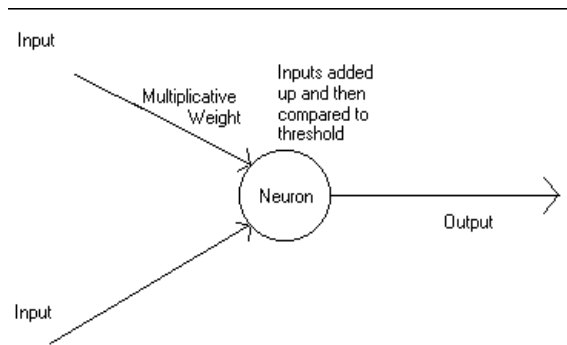


**Figure 2.1: A Simple Neuron**

## 3. Handwritten ZIP code Classification Using ANN
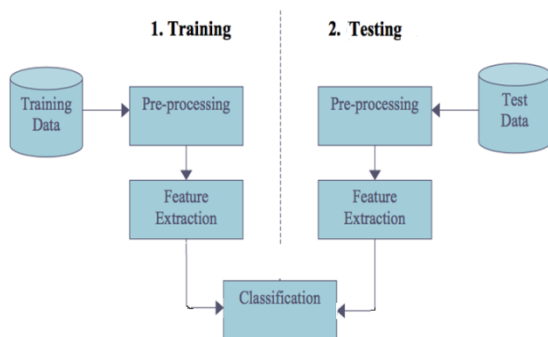
3.1 System Block Diagram



**Figure 3.1.1: Block diagram of ZIP code Classification system**

The System divided into three modules. These are Pre-processing, Feature Extraction and Classification.

3.2 Pre-processing:

The Pre-processing includes two steps. These are Skew (or) Slope correction and Slant correction.

3.2.1 Skew (or) Slope correction:

With the slope correction, the handwritten ZIP code is horizontally rotated such that the lower baseline is aligned to the horizontal axis of the image.
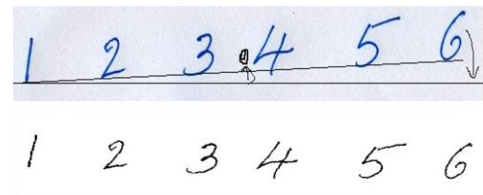


**Figure3.2.1.1: Slope correction**

3.2.2 Slant Correction:

**Slant correction** means that, Slant is the angle between the vertical direction and the direction of the vertical Digit strokes. Slant correction transforms the ZIP code into an upright position.
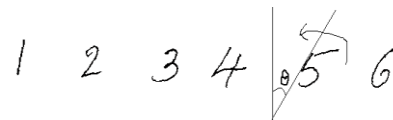


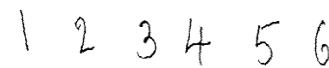**Figure 3.2.2.1: ZIP code contains slanted digits**



**Figure 3.2.2.2: Slant corrected image**

3.3 Feature Extraction:

Feature extraction includes two steps. These are *Digit extraction* and *Digit Normalization*

3.3.1 Digit Extraction:

Digit Extraction means that we can extract the digits from ZIP code image. For example (fig 3.3.1.1) we want to extracted the digits from below zip code image are:
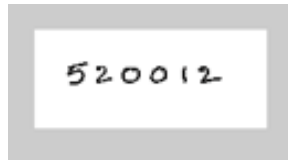
**Figure 3.3.1.1: Original Zip-code image**



**Figure 3.3.1.2: Extracted Digits from**

**Original Zip-code**

Procedure for Digit Extraction:

Input: Pre-processed ZIP code image

Output: Extracted digits

Method:

1. Scan each row from top to bottom until you get '0' value. Now consider the row in which we get 0 value as the 1st row.(fig:3.3.1.3)
2. Scan each column from left to right until you get '0' value. Now consider the column in which we get 0 value as the 1st column. (fig:3.3.1.3)
3. Scan each row from bottom to top until you get '0' value. Now consider the row in which we get 0 value as the 2nd row. (fig:3.3.1.3)
4. Scan each column from 1st column to right until you get all pixel values are '1's. Now consider the column in which we get all pixel values are 1's as the 2nd column. (fig:3.3.1.3)
5. Now we have the required four values 1st row,1st column,2nd row and 2nd column, using these values we can extract digits and will be placed as new separated image. (fig:3.3.1.3)
6. After creating the new separated image, The cropped area in the original image will be replaced with 1's.

7. Now, the above process (steps 1-6) is repeated until we should get all digits from original image.
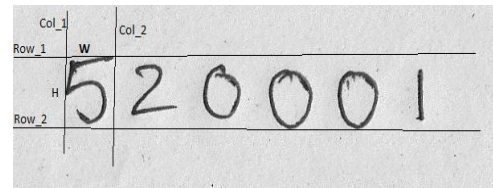


**Figure 3.3.1.3: Digits are extracted from ZIP code image.**

3.3.2 Digit Normalization

***Digit normalization*** tries to make the system invariant to the Digit size and to reduce the empty background areas. For example (fig 3.3.2.1) we want to normalize the following extracted digit is:
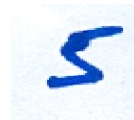


**Figure 3.3.2.1: Extracted Digit**



**Figure 3.3.2.2: Normalized Digit**

Procedure for Digit Normalization:

Input: Extracted image.

Output: Size normalized image.

Procedure:

1. Scan each row from top to bottom until you get '0' value. Now consider the row in which we get 0 value as the 1st row.
2. Scan each column from left to right until you get '0' value. Now consider the column in which we get 0 value as the 1st column.

3. Scan each row from bottom to top until you get '0' value. Now consider the row in which we get 0 value as the 2$^{nd}$ row.
4. Scan each column from 1$^{st}$ column to right until you get all pixel values are '1's. Now consider the column in which we get all pixel values are 1's as the 2$^{nd}$ column.
5. Now we have the required four values 1$^{st}$ row,1$^{st}$ column,2$^{nd}$ row and 2$^{nd}$ column, using these values we can extract digits and will be placed as new separated image. (fig:3.3.2.3)
6. Now the new separated called as normalized digit.
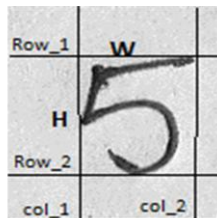


**Figure 3.3.2.3: Digit Normalization**

3.4 Classification

In the process of classification we utilize "**nprtool**" in order to train the data and then classify accordingly. In this process first we will digit normalization outputs as the input to the neural network which can have the hidden layers according to our interest. Once the data is given as input we will go for the training the network. If the error obtained is greater than what we expected then we can again retrain our data. Once the training is completed we can go for validation and testing.

In the process of classification we utilize *nprtool* in order to train the data and then classify accordingly. In this process first we will resize the digit normalization outputs as 15* 15 size. Now the resized normalized image having 225 pixels values. So we can take the each pixel value as one feature value for that normalized digit.

The input to the neural network which can have the hidden layers according to our interest. Using the Neural Network Pattern Recognition Tool we can done classification step.
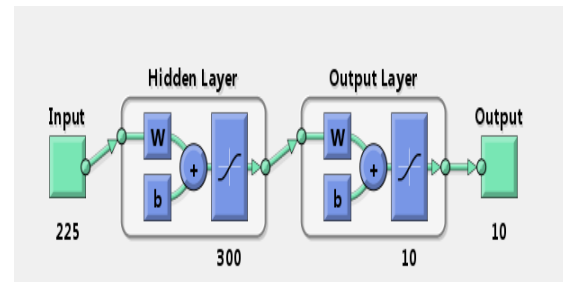


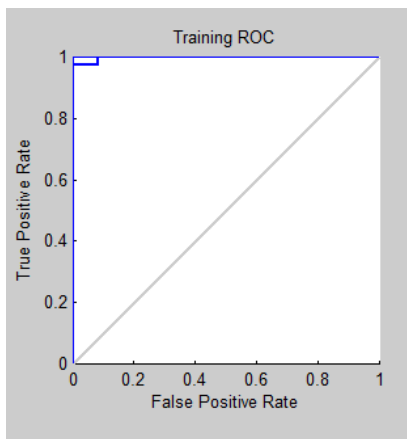**Figure 3.4.1: Neural Network Architecture**

## 4. Experimental Results:

After 2 training passes the error rate on the training set (102 handwritten zip codes having 619 handwritten digits) was 1.0% and the MSE was 2.3*e$^{-3}$



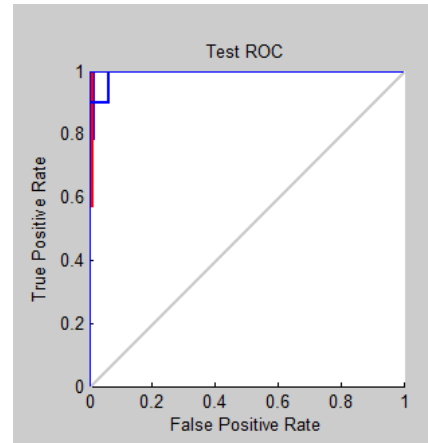| | Samples | MSE | %E |
|---|---|---|---|
| Training: | 433 | 3.94103e-4 | 2.29885e-1 |
| Validation: | 93 | 4.34679e-3 | 1.08695e-0 |
| Testing: | 93 | 7.35482e-3 | 4.34782e-0 |

**Table 4.1 Results after 2 training passes**

**Figure 4.1: Result Confusion Matrixes**
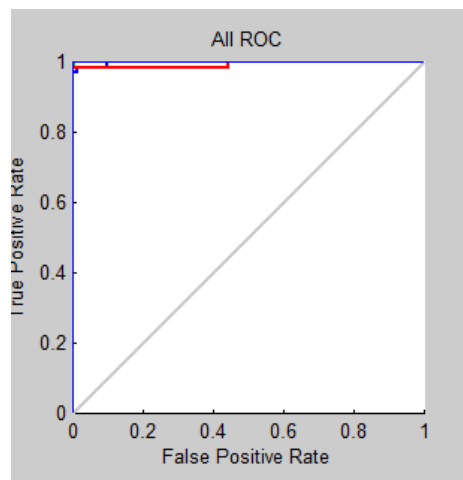


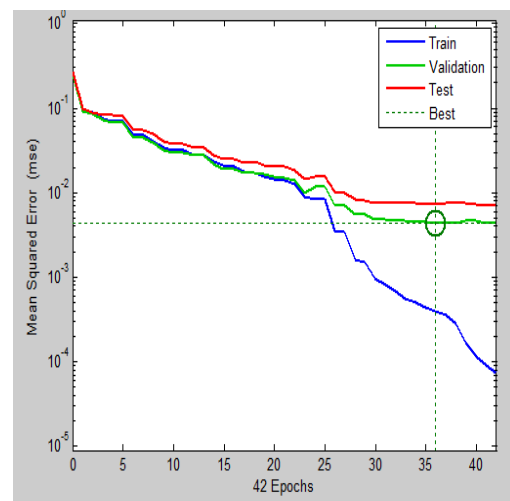**Graph 4.1: ROC for Training Data**
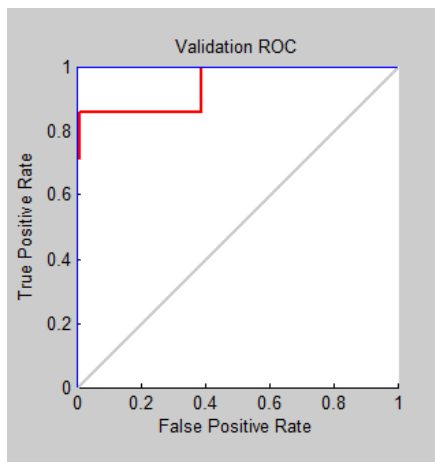


**Graph 4.2: ROC for Validation Data**



**Graph 4.3: ROC for Test Data**



**Graph 4.4: ROC for All Data**



**Graph 4.5: Performance curve**

## 5. Conclusion:

Neural Network system, back-propagation learning, is used to recognize handwritten

digits were very successful. An image, which contained 100 samples of each number, was trained and tested. The accuracy rate of classified digits was 99%. This accuracy rate is very high. From the training and testing results, it was concluded that the system had more trouble identifying numeral 5". This may be caused by the fact that the digit is running together or maybe it is not fully connected.

## References

[1]. Tzanakou, E.M., 2000. Classifiers: An Overview, from Supervised and Unsupervised Pattern Recognition: Feature Extraction and Computational Intelligence. CRC Press LLC., ISBN: 0-8493-2278-2, pp: 371.

[2]. Woogon Chung and Evagelia Micheli-Tzanakou, 2000. Classifiers: An overview, from Supervisedand Unsupervised Pattern Recognition: Feature Extraction and Computational Intelligence. Illustrated Edn., CRC Press, USA., ISBN: 0849322782, pp: 371.

Biography

K.**Siva Kuamr** Pursuing M.Tech (CSE) from Sri Mittapalli College of Engineering , Affiliated to JNTU, Kakinada. His Interested areas includes Artificial Neural Networks, Image processing.

**D.Uma Devi** received her B.Tech degree in Computer Science from GITAM,Andhra University,Vishakapatnam, M.Tech degree in Computer Science from Andhra University,Vishakapatnam. Currently, she is working as a Associate Professor CSE Department in Sri Mittapalli College of Engineering, Tummalapalem,Guntur. She has got 12 years of teaching experience.She has published 2 papers in international journals.