Abstract

The problem of handwritten recognition considered to be very important problem because of its numerous applications and theoretical values in the domain of pattern recognition.

In this research, models of Neural Networks are used to recognize written characters, applying Artificial Neural Network (ANN) of three types, which are: - Kohenen All Classes in One Network (ACON), Kohenen One Class in One Network(OCON), and Learning Vector Quantization (LVQ).

The feature extraction process made use of Haar Wavelet Transformation to extract the parametric features of the handwritten characters.

Also Geometrical features were also used to extract features (Moment and Complex Moment).

The system was implemented using Visual Basic Language, database of 130 persons was established, 70 samples from the database were used for training, and the all 130 samples were used for testing the system. The efficiency of the system was tested using the Recognition Rate.

The results show that the wavelet transformation with both Kohenen Learning Vector Quantization and Kohenen One Class One Network (OCON) achieves the highest recognition rate in which it scores 94%.

Acknowledgment

I would like to express my sincere appreciation to my research supervisor, Dr. Sattar B. Sadkan, for giving me the major steps to go on to explore the subject, shearing with me the ideas in my research "Handwritten Recognition Using Neural Network" And perform the points that I felt were important.

Also I wish to thank Dr. Venus W. Samawi my supervisor for their available advice and encouragement. Grateful thanks for the Head of Department of Computer Science Dr. Taha S. Bashaga.

I wish to thank the staff of Computer Science Department at the AL-Nahrain University for their help.

I would like to say "thank you" to my faithful friend for supporting and giving me advises.

You

Appendices

Appendix A

BMP File Format

The BMP file format divides the a graphical file into four major parts, these are:

1. *Bitmap File Header:* the bitmap file header is 14-bytes long and is formatted as follow:

UNIT	bfType	(holds the signature value0xd42,which Identify the file as BMP)
DWORD	bfSize	(holds the file size)
UNIT	bfReserved	(Not used)
UNIT	bfReserved	(Not used)
DWORD	bfOffBits	(specify the offset, relative to the beginning of the file, where the data representing the bitmap itself begins)

2. ***Bitmap Information Header***: this part contains the bitmap important information about the image ,the widows format for this header is :

DWORD	biSize	(holds the header length in bytes)
LONG	biWidth	(identify the image width)
LONG	biheight	(identify the image height)

WORD	biBitCount	(identify the number of bits/pixel in the image and thus the maximum number of colors that the bitmap can contain)
DWORD	biCompression	(identify the compression scheme that the bitmap employs)
DWORD	biSizeImage	(set to zero for uncompressed image,else it holds the size (in bytes) of the bit representing the bitmap image for the compressed image)
LONG	biXPelsPerMeter	
LONG	biYPelsPerMeter	
DWORD	biCIRImportant	

3. ***Palette (Color table containing RGB triple structure):*** the colors tables specify the colors used in the bitmap .the bitmap. The bitmap files comes in four color formats:

1- 2-color	one-bit per pixel
2- 16 - color	four -bit per pixel
3- 256 - color	eight-bit per pixel
4- 16.7million-color	24-bits per pixel

4. bitmap bits: the bitmap bits is the set of bits defines the image – the bitmap itself. In the 2-color,16-color,and 256-color bitmap format, each entry in the bitmap is an index to the color table. in the 16.7 million-color bitmap ,where there is no color table, each bitmap entry directly specify a color ,the first 3-bytes in each 24 bit entry specify the pixel colors red component ,the second specify the green component, the third specify the blue component.

Chapter 4

Results and Discussion

This chapter is concerned with discussion of the results produced by the designed system to perform the operation of handwritten recognition system for English characters. Calculations of the results mentioned in this chapter were based on samples taken from a group of persons who has been chosen to write some patterns of English characters.
The language used to program the training and recognizing software was Visual Basic 6.0.

4.1 Database Description

The total samples used in this thesis gathered from 130 different persons, each person where asked to write the 26 character in their capital and small cases on a white paper, after that, papers were scanned and saved as bitmap 24 colored images.

These samples were used for both training and testing of the system, 70 samples where used from the 130 sample in the training process of the constructed neural network of this system, and all 130 samples were used in the testing of this system.

Two databases were created from these images, one for capital case character and one for small case, each database have table of records which contains character instances from letter a to letter z in the small case database, and from letter A to letter Z in the capital case database each records contains the following fields :-

1- Character name :- What is this character(a, b, c,…).
2- Image number :- The number of the stored image sample.
3- Character image:-Character image of size 64*64.

Data entered to the database by using software designed for this purpose. This software reads the images that contains the character samples, perform the required preprocessing steps and segment the character sequence in each image into individual characters sub-image and store the sub- image in the character image field , and then stores the image file number in the image number field and what character is the one stored in the character image field (i.e. name of character) in the character name filed.

After constructing the database that contains the character samples, the system will be prepared to train the neural network on the this database.

4.2 System Training Process

After finishing the database construction, Neural Network training process will begin by using part of the samples in the constructed database to produce neural network weights that will be used by the system for the recognition process of the input handwritten characters in the test phase, 70 samples were used in the testing process. Fig 4.1 illustrates simple view of the training process.

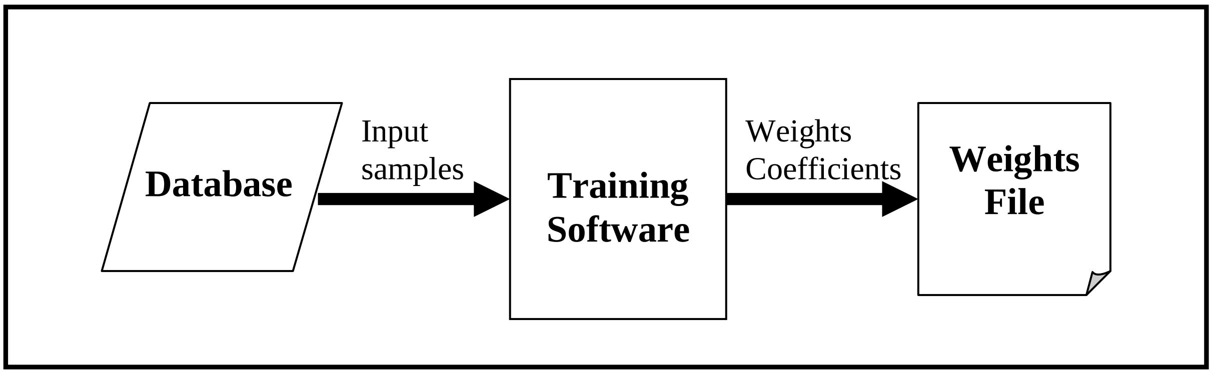


Fig (4.1) Training Process

A set of samples stored in the database delivered to the neural network for training. the weights of the neural net is first initialized and then will be updated in the training process by using the training algorithms mentioned in the previous chapter, at the end , the training process will stopped when there is no large changes in the values of the weights.

Weights saved into a file to use in the testing procedure.

The training software interface shown bellow:-

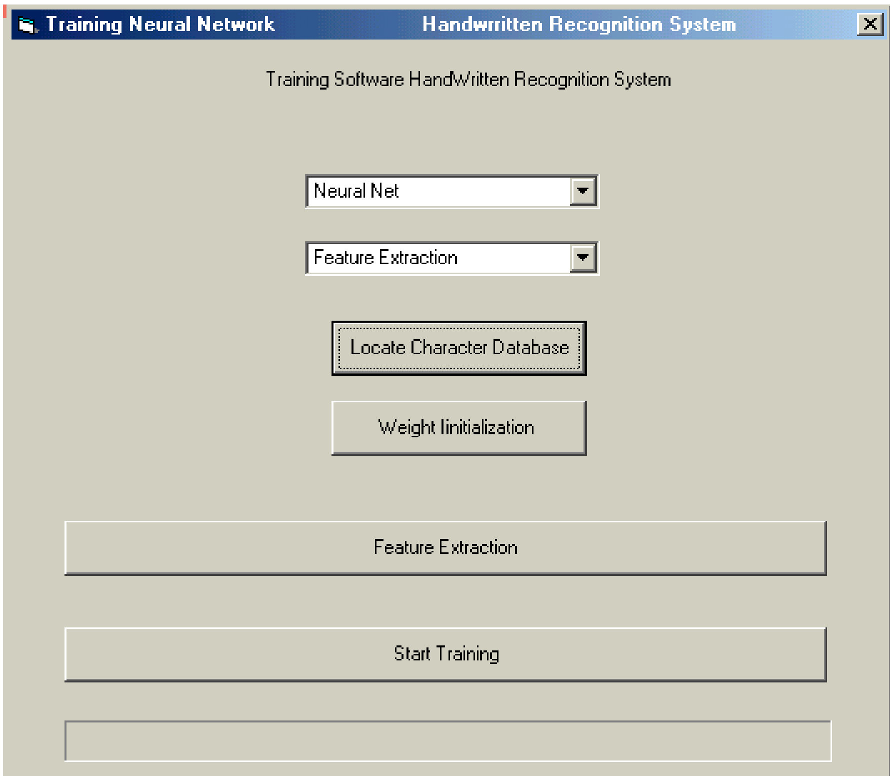


Fig 4.2 Training Software Interface

First, we specify the Neural Network type to train as shown (Fig 4.3):-

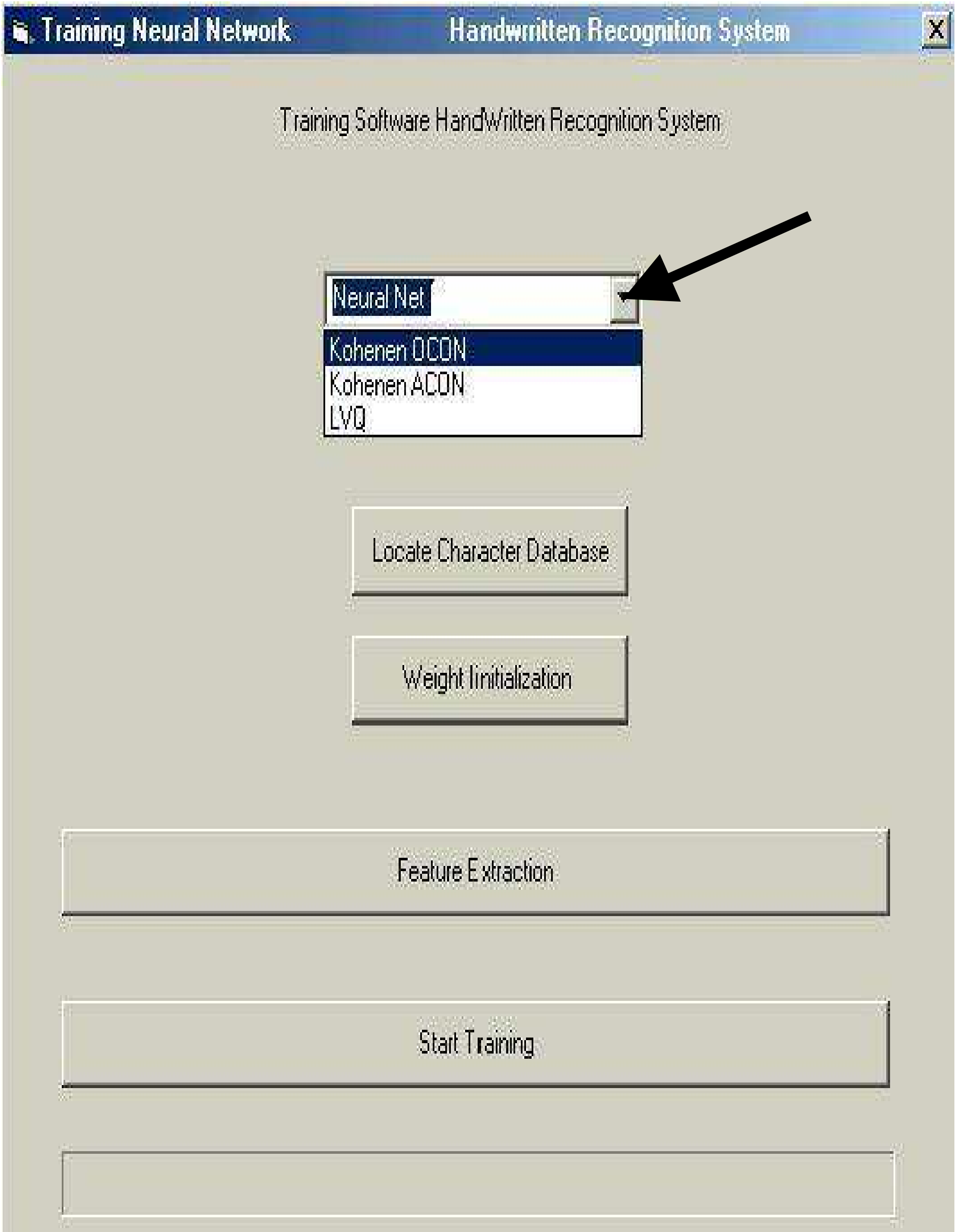


Fig 4.3 Specify the Neural Network Type

Then we specify the Feature Extraction method to use in the training as shown (Fig 4.4):-

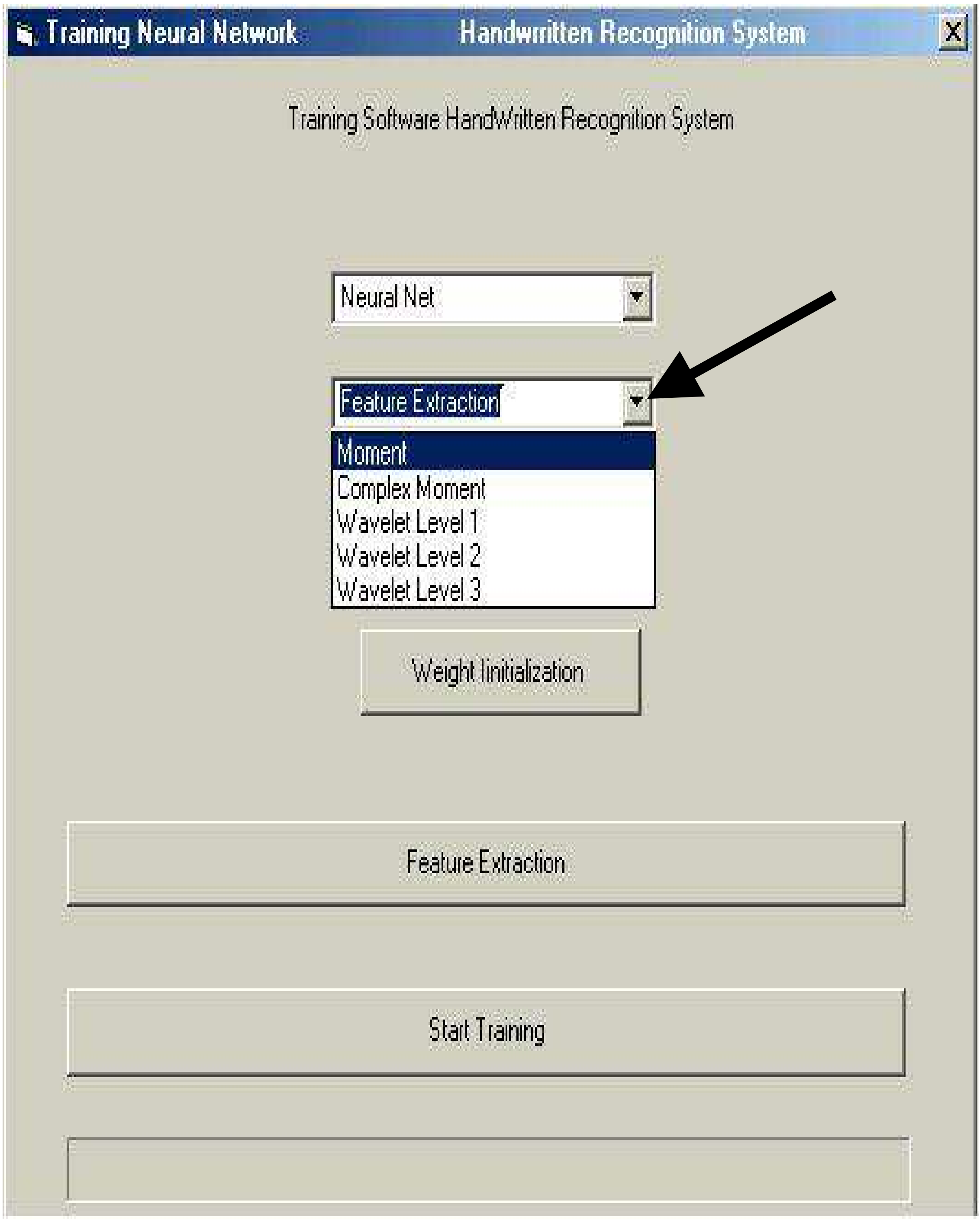


Fig 4.4 Specify Feature Extraction method Type

Then we locate the database through the browsing window as shown(Fig 4.5):-

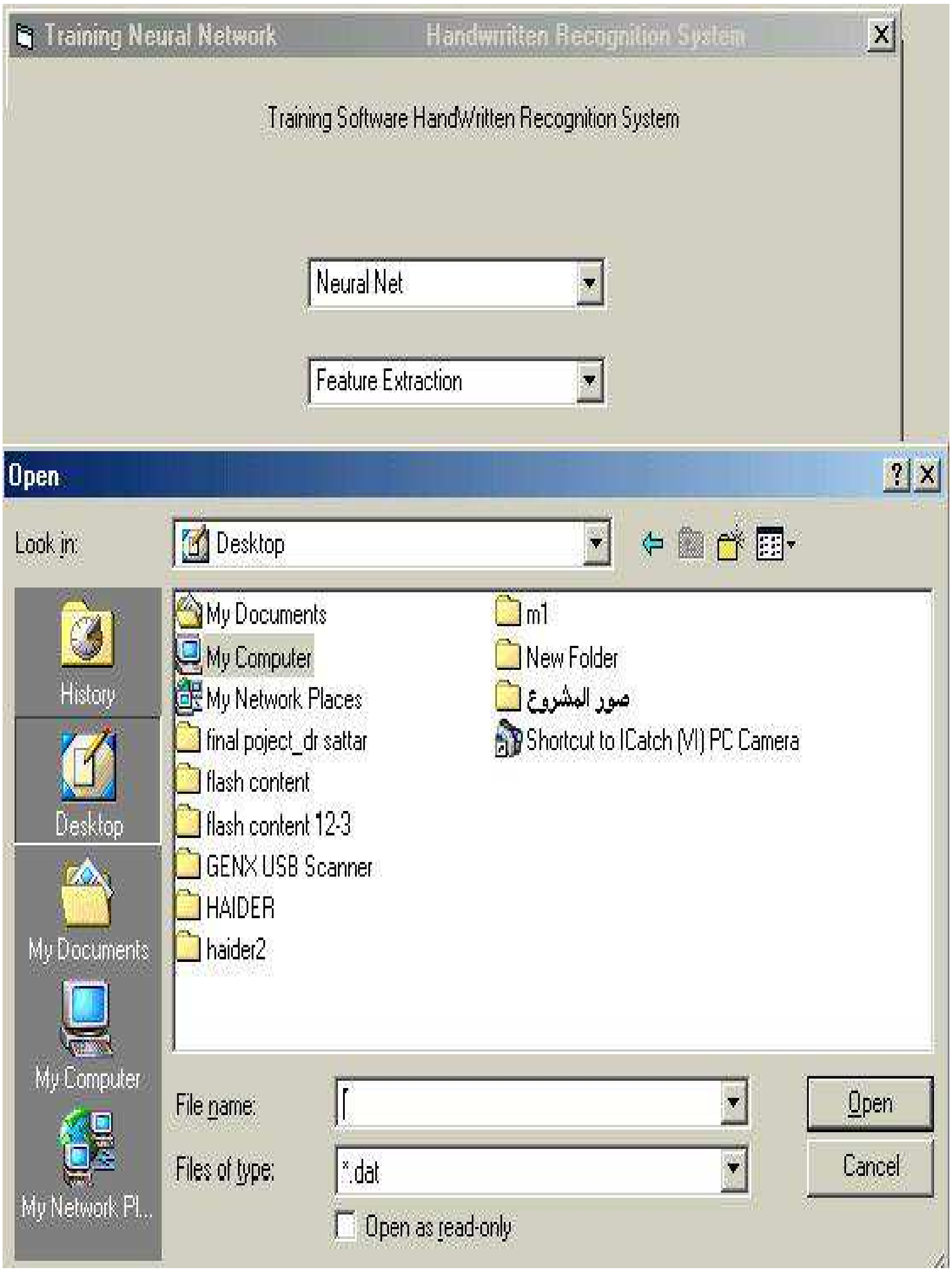


Fig 4.5 Locate the Database

Then we initialize the weights to begin the training as shown in (Fig 4.6):-

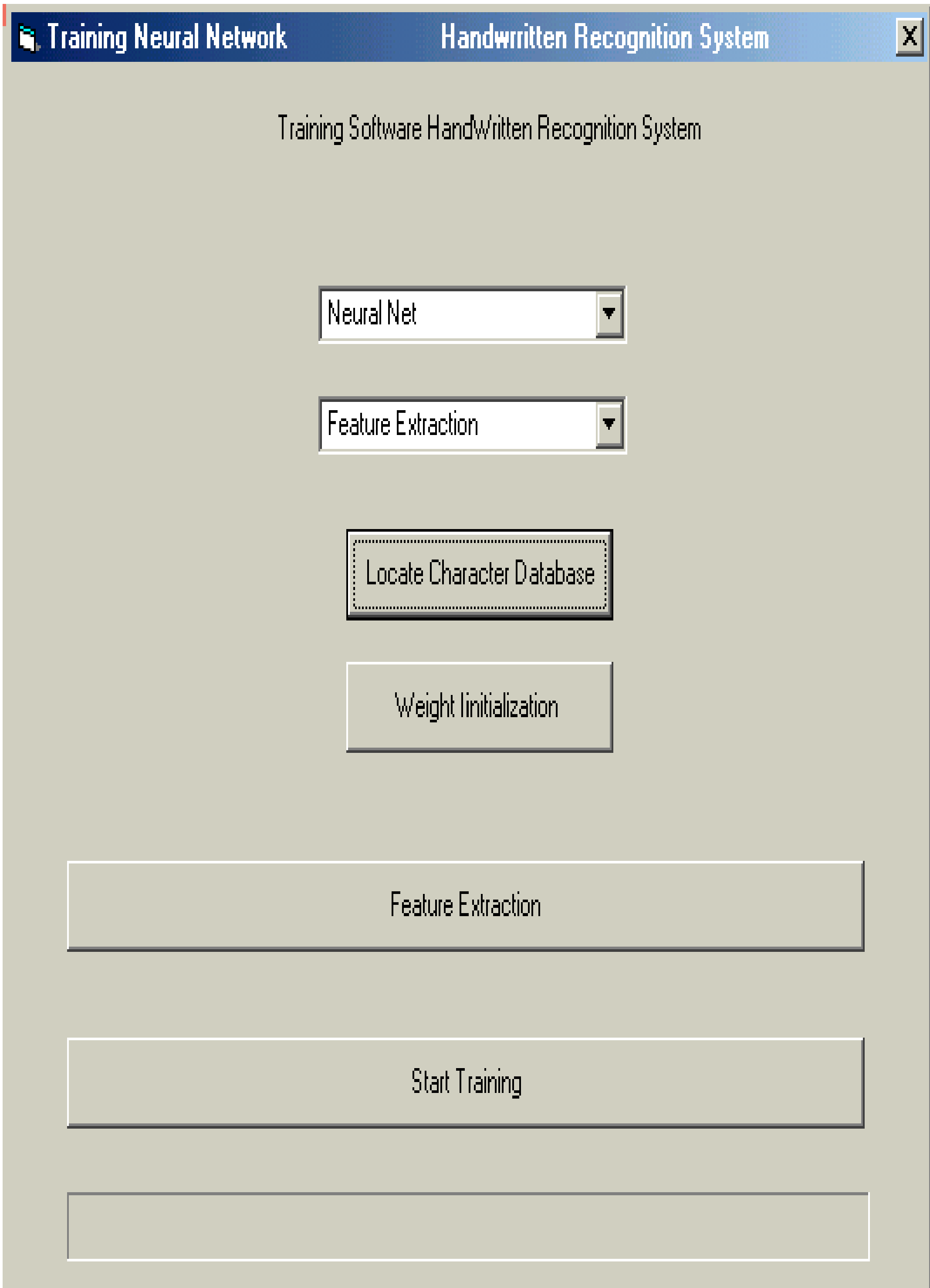


Fig 4.6 Initialize the Weights

Next, we start the feature extraction process as shown in (Fig4.7):-

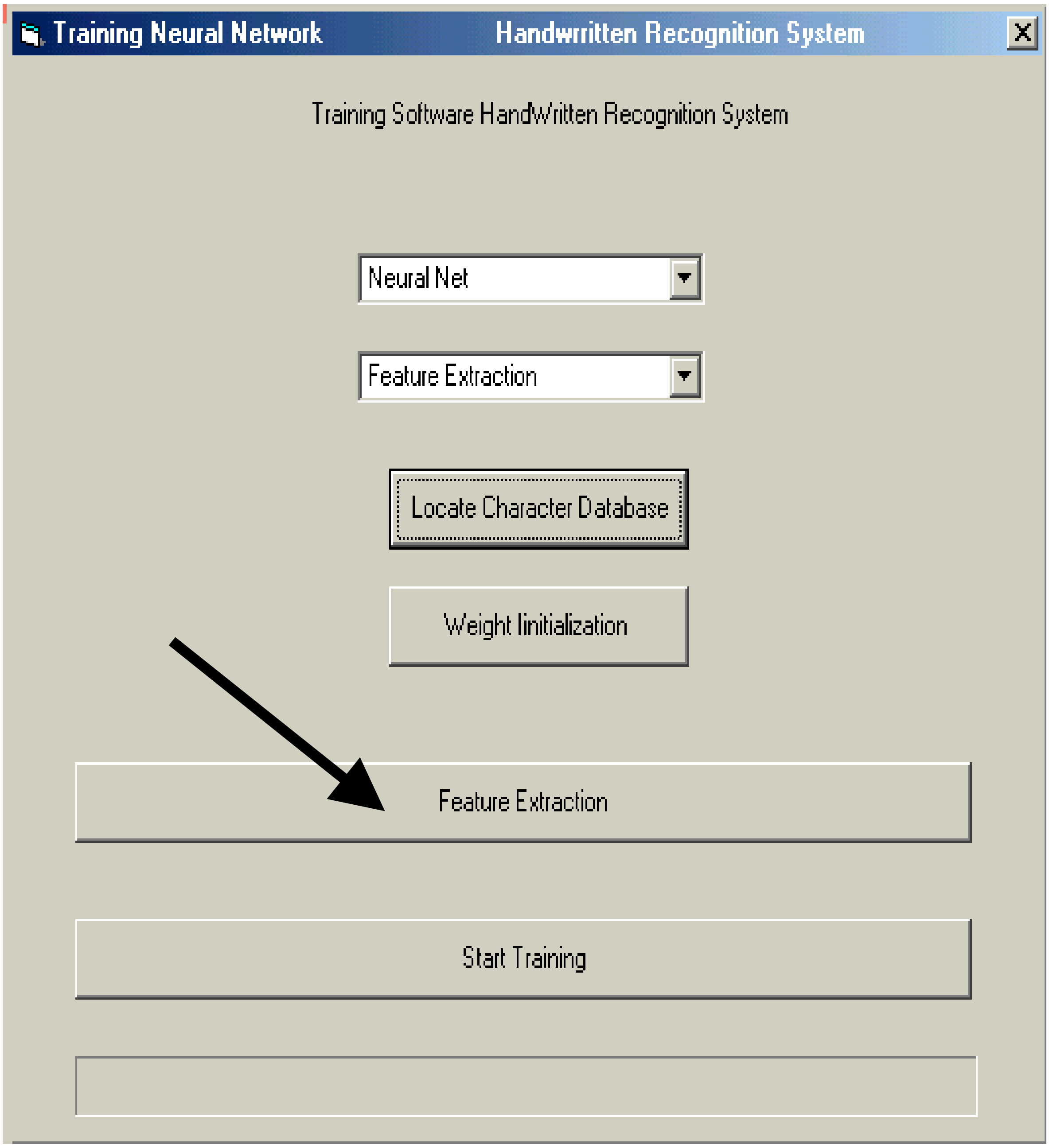


Fig 4.7 Initialize the Weights

Then we start the training process as shown in (Fig 4.8):-

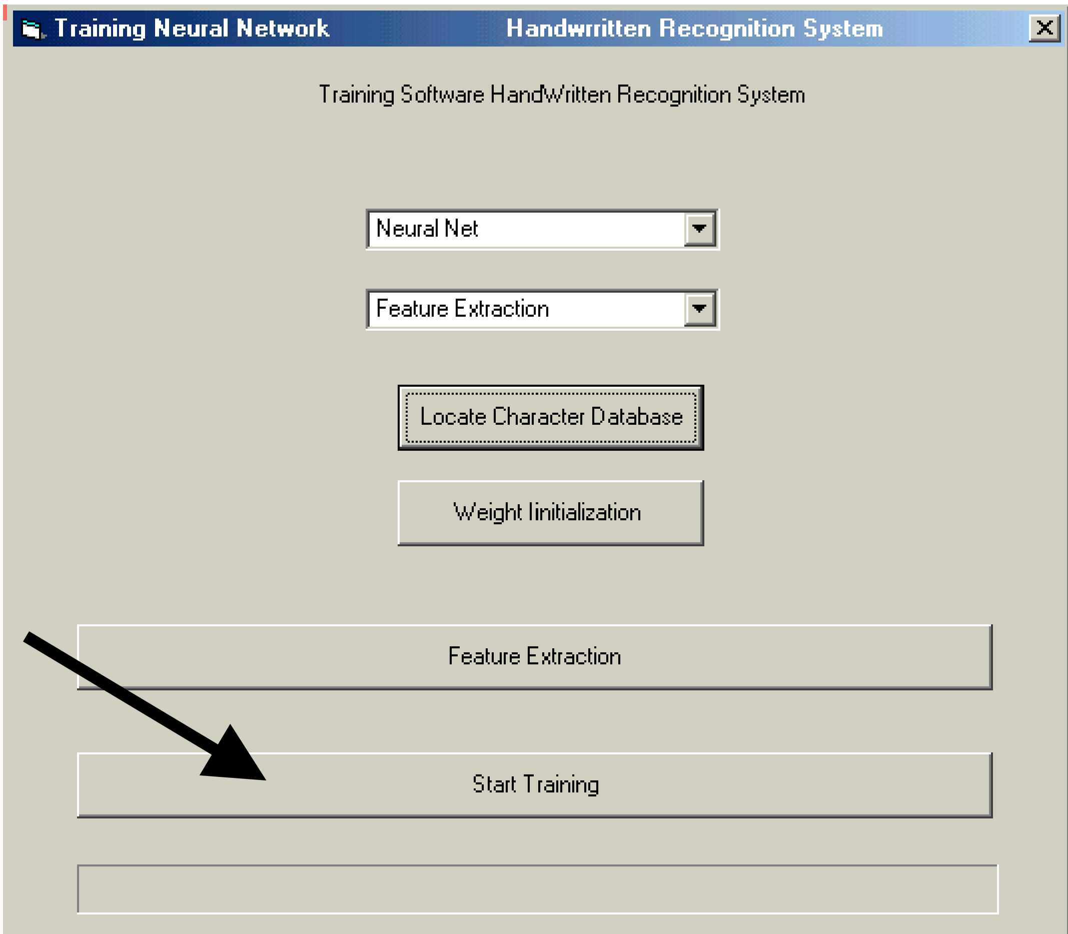


Fig 4.7 Start Training

4.3 System Testing Process

in the testing process, database that contains the character samples where all used to test the performance of the system against two main points :-

- Feature extraction techniques.
- Neural network types and architectures.

In addition, several study cases where taken to measure the performance of this system against different input styles.

Fig 4.9 shows the Handwritten Recognition system interface:-

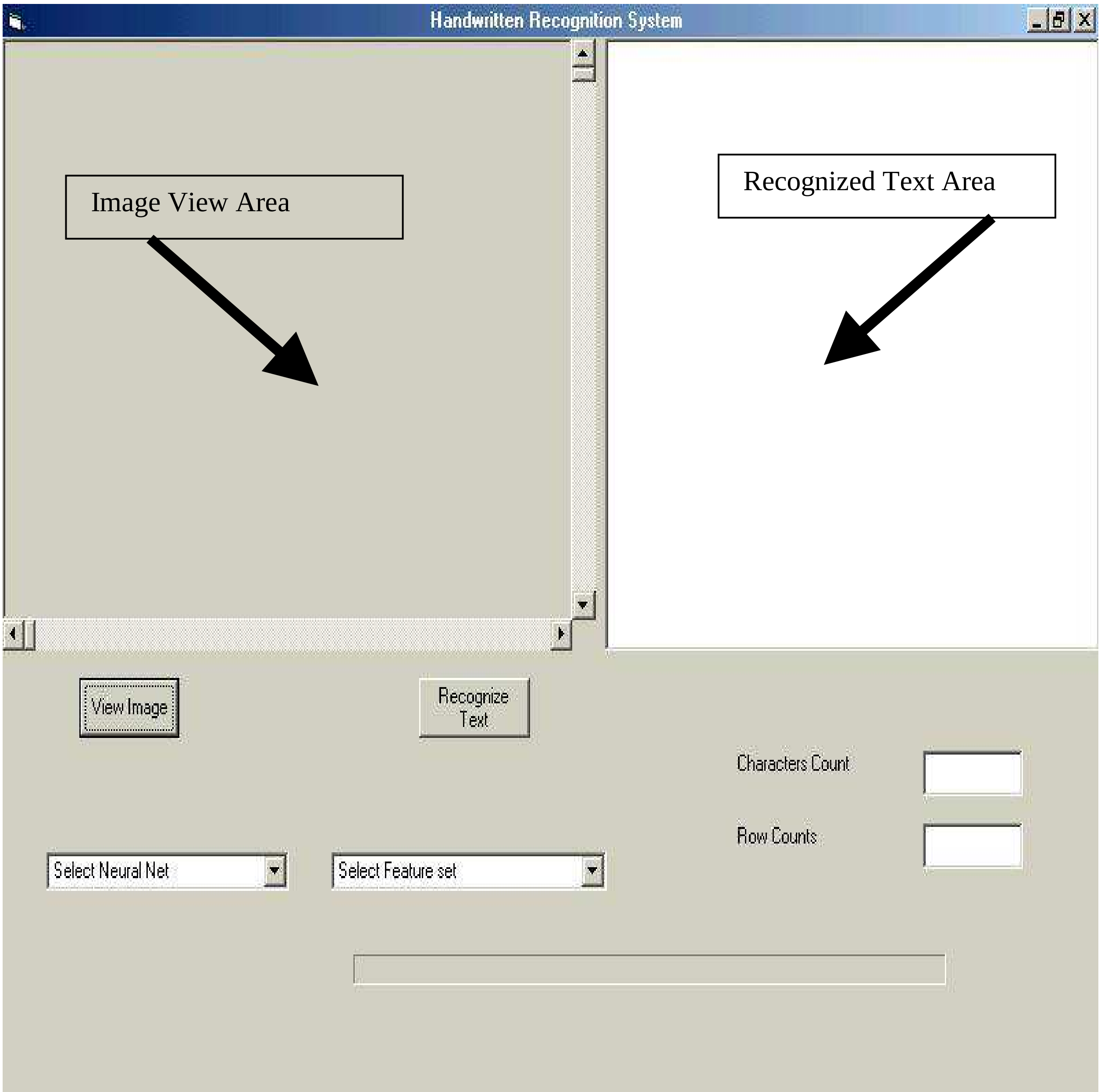


Fig 4.9 Handwritten Recognition Systems

First, we specify the Neural Network Type from the list of Neural nets supported by the system as shown in fig(4.9):-

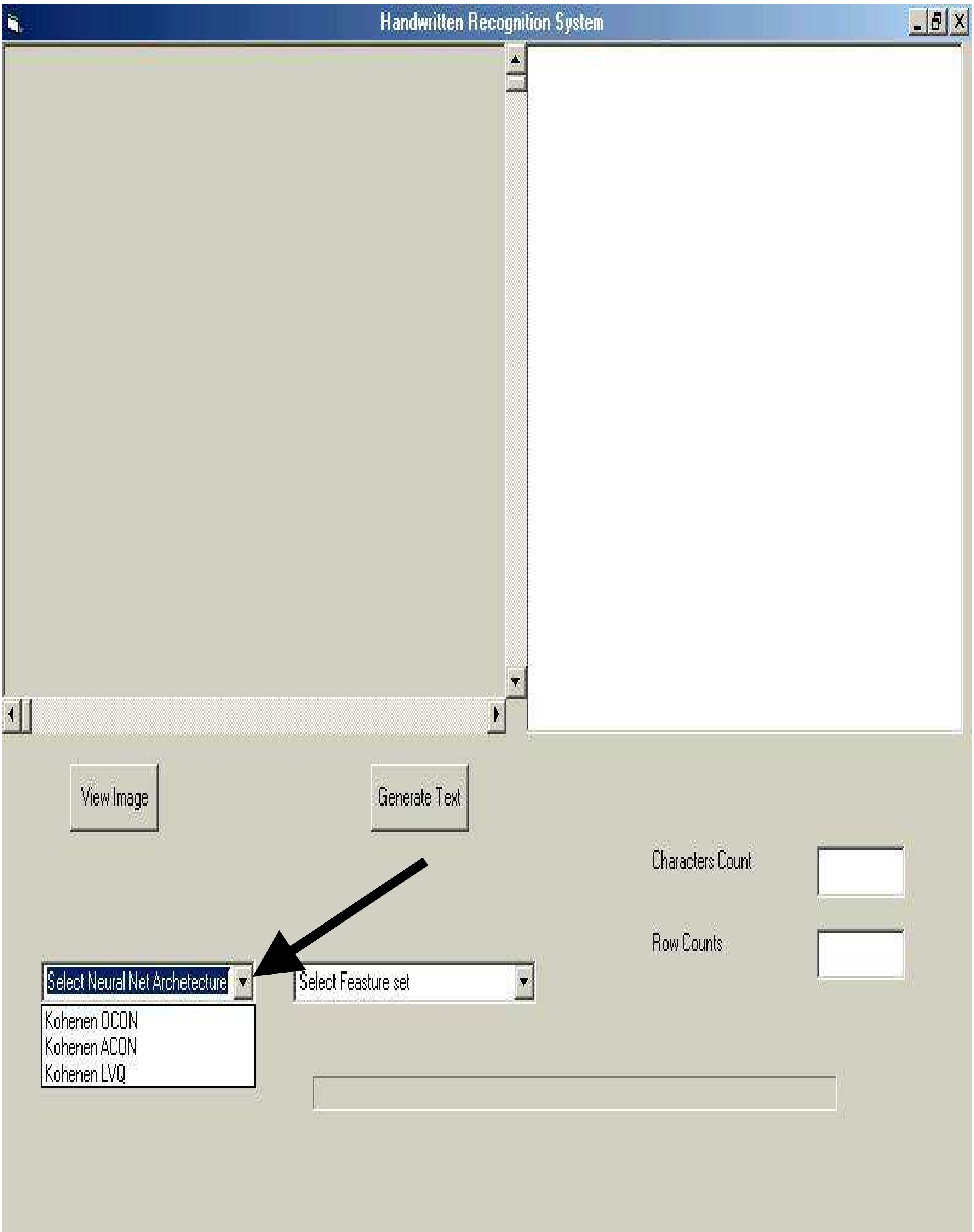


Fig 4.10 Specify the Neural Network Type

Next, as shown in fig (4.10) the Feature extraction method supported by the system:-

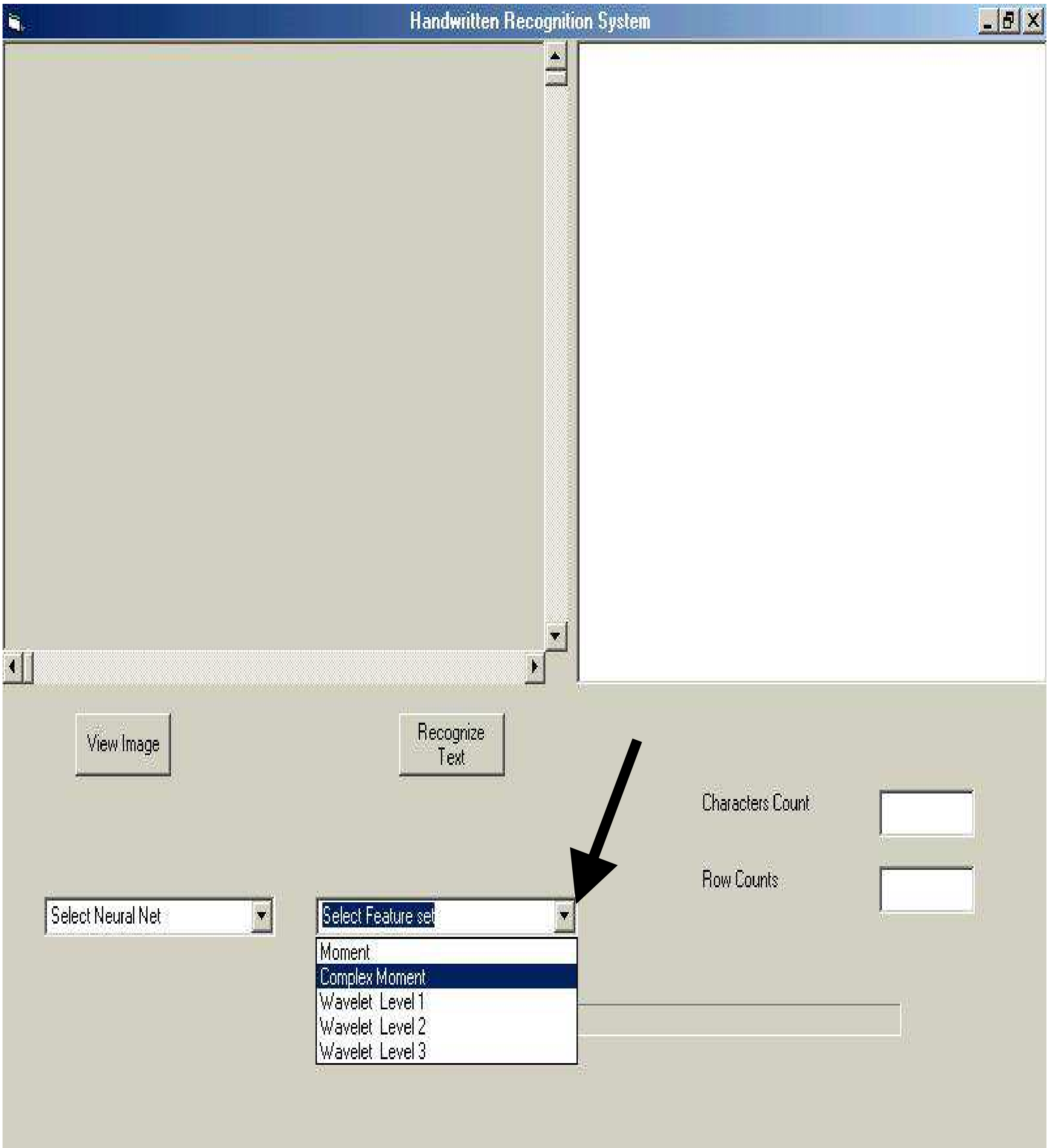


Fig 4.11 specify the Feature Extraction method

Then the image that contains the text to be recognized viewed by locating the image through the browsing window:-

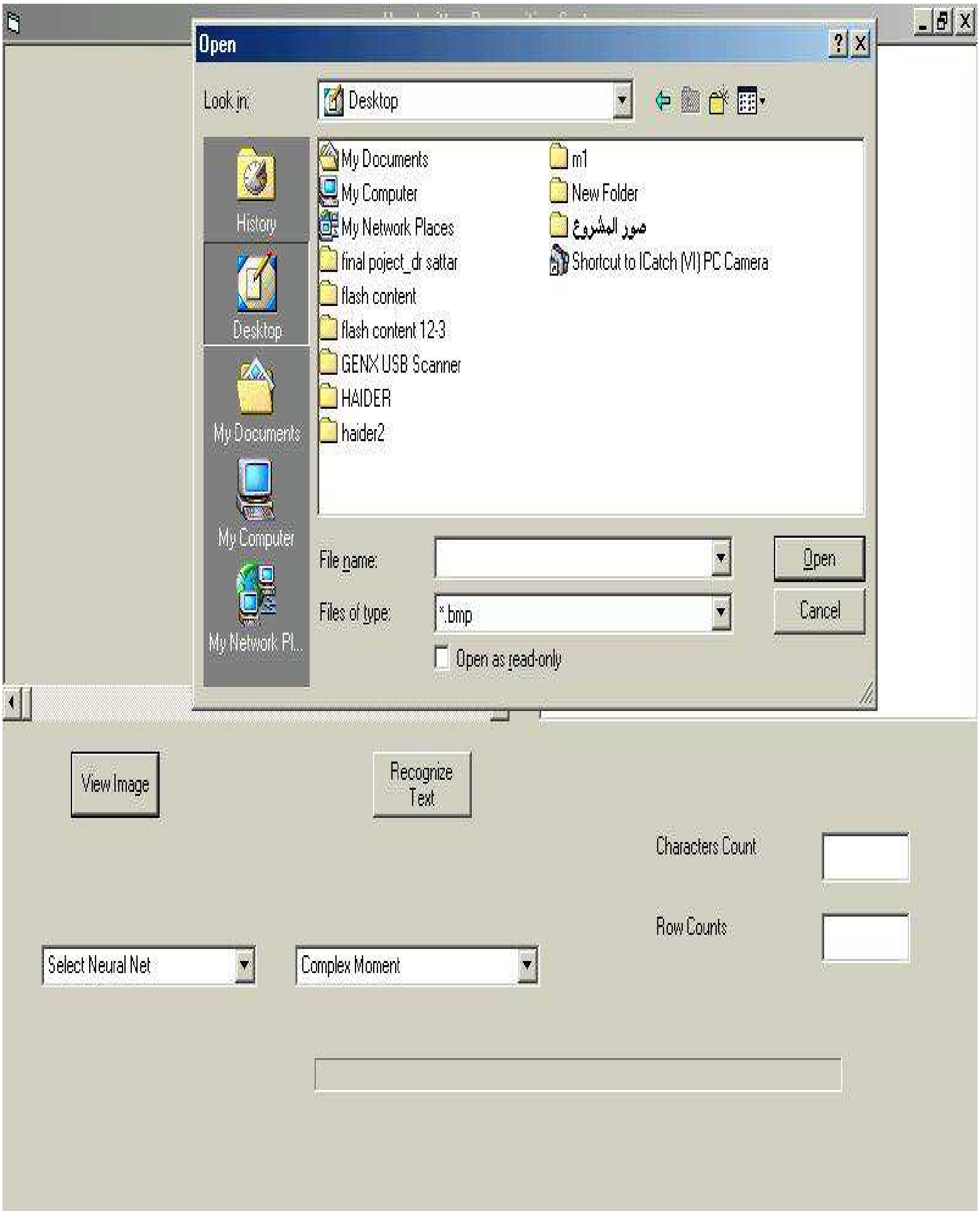


Fig 4.12 Image View

Finally, the text extracted from the image as shown in fig (4.9):-

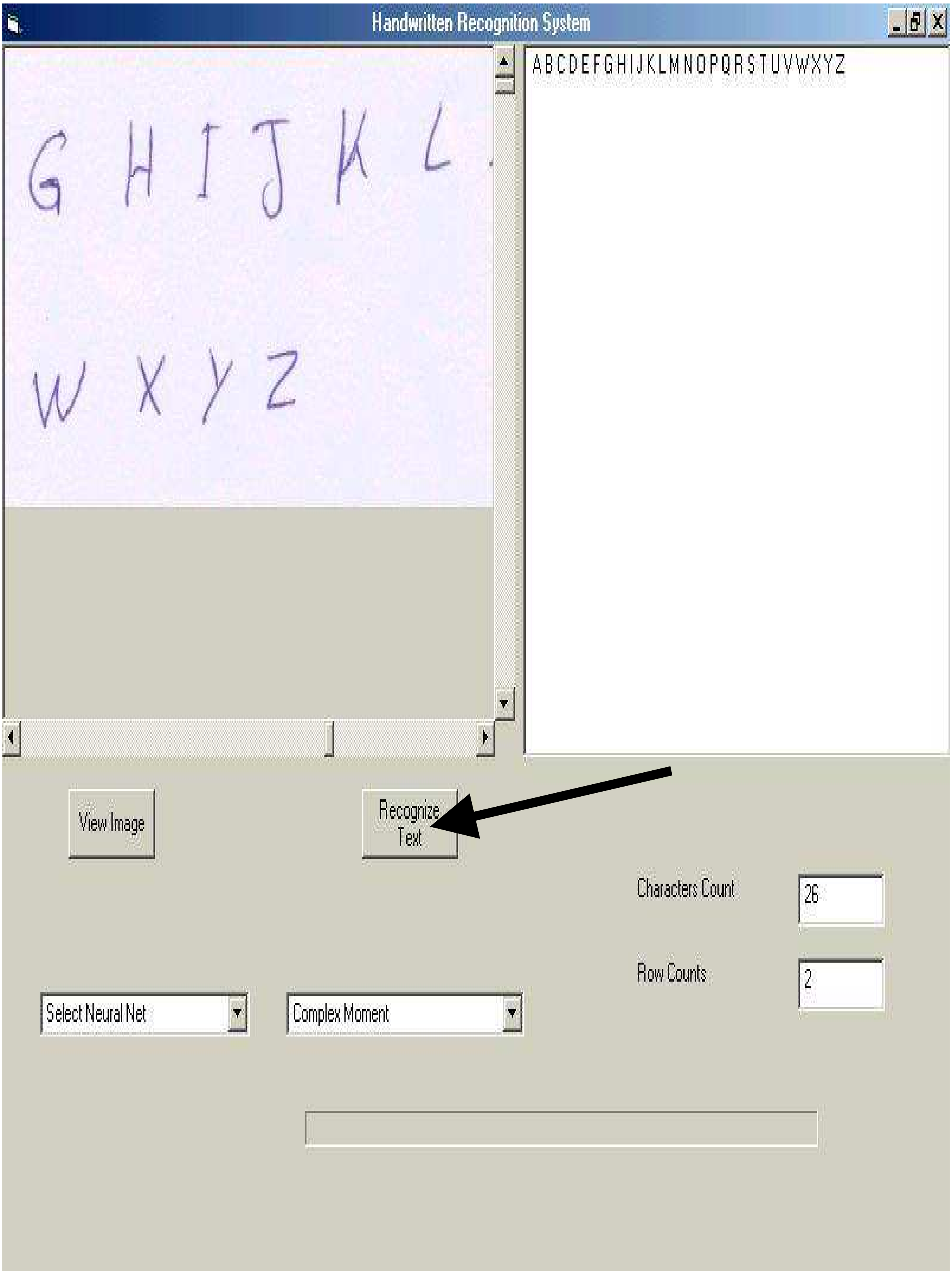


Fig 4.13 Text Generation

In addition, as shown in tables (4.1), (4.2), (4.3), the recognition rate are calculated for the developed methods of feature extraction techniques with neural network types used in this system.

The recognition rate calculated by the equation:-

Recognition Rate= $\frac{N}{M} * 100\%$ where N= number of correctly recognized characters, M= total number of characters.

Table 4.1 Recognition Rate against Moment Technique and Neural Network Architectures

Type of NN	Moment	Wavelet With Moment
Kohenen OCON	71.34%	71.98%
Kohenen ACON	33.14%	32.53%
LVQ	70.02%	71.34%

The results in table (4.1) shows the comparison of the calculated recognition rate of moments and moments with wavelet transformation against neural network types designed in this thesis, in which Kohenen with OCON architecture scores high recognition rates with moments and its performance increases when we combine Wavelet transformation with Moment, this combination also increases the recognition rate of LVQ neural net. also we notice that Kohenen with ACON architecture generally failed in recognition process with both Moment and combined Moment and Wavelet Approach.

Table 4.2 Recognition Rate against Complex Moment Technique and Neural Network Architectures

	Complex Moment	Wavelet with Complex Moment
Kohenen OCON	73.87%	74.66%
Kohenen ACON	23.05%	23.33%
LVQ	73.16%	74.25%

Table 4.2 shows the comparison of the calculated recognition rate of complex moments and complex moment with wavelet transformation against neural network types designed in this thesis, Recognition rates are at highest levels with Kohenen OCON and LVQ Neural Net, and the combined complex moment and wavelet transformation approach increases the recognition rates in these two architectures. Kohenen ACON also shows very low recognition rates with the two feature extraction techniques.

Table 4.3 Recognition Rate against Wavelet Transformation Techniques and Neural Network Architectures

	Wavelet 1	**Wavelet 2**	**Wavelet 3**
Kohenen OCON	84.37%	94.07%	89.36%
Kohenen ACON	34.43%	33.89%	31.77%
LVQ	85.55%	93.52%	88.69%

Table (4.3) shows the comparison of the calculated recognition rate of wavelet transformation feature sets (LL subband for each of the three levels of resolution of wavelet transformation) against neural network types designed in this thesis, Kohenen OCON and LVQ Neural Net have the best performance with wavelet transformation as feature extraction technique, we also notice that coefficient of wavelet transformation in the second level of resolution scores the highest recognition rates among other levels of resolution.

Study cases is also taken in our test process to measure the performance of the constructed system against different factors, the factors that is studied in this thesis is ***age*** and ***sex***, where 4 samples where taken from four different persons, the first two samples were taken from two women, One is 24 years old; the other is 42 years old. The other ample taken From two men, one is 22 years old, the other sample were 45 years old.

In the sex factor, table (4.4) and (4.5) shows the recognition rate for males and females taken as study cases:-

Table 4.4 Recognition rate for Females

	Moment	Moment with wavelet	Complex moment	Complex moment with wavelet	Wavelet level 1	Wavelet level2	Wavelet level3
Kohenen ACON	**84.61**	**88.46**	**92.30**	**94.23**	**96.15**	**100**	**96.15**
LVQ	**84.61**	**86.53**	**94.30**	**96.15**	**100**	**100**	**92.30**

Table 4.5 Recognition rate for Males

	Moment	Moment with wavelet	Complex moment	Complex moment with wavelet	Wavelet level 1	Wavelet level2	Wavelet level3
Kohenen ACON	**84.61**	**84.61**	**94.23**	**96.15**	**100**	**100**	**96.15**
LVQ	**84.61**	**88.46**	**92.30**	**92.30**	**98.05**	**98.05**	**96.15**

The recognition rates are similar between males and females in all feature extraction techniques and neural network architectures.

Table (4.6) and (4.7) shows the recognition rate for the two case studies of persons at different ages.

Table 4.6Recognition rate for 22 years old persons

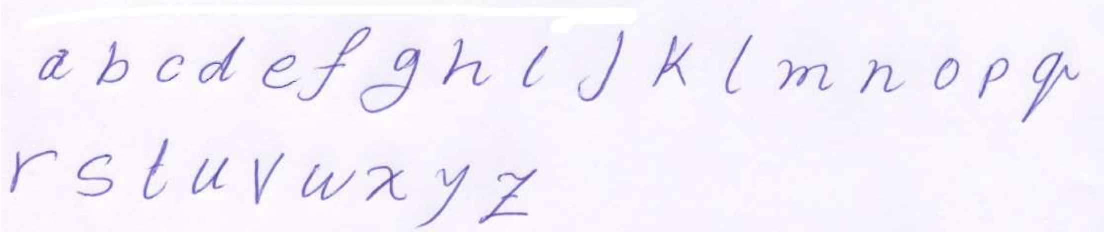
	Moment	Moment with wavelet	Complex moment	Complex moment with wavelet	Wavelet level 1	Wavelet level2	Wavelet level3
Kohenen OCON	85.57	86.53	87.50	88.46	96.15	99.93	94.23
LVQ	84.61	85.57	88.46	90.38	98.07	100	96.15

Table 4.7 Recognition rate for 45 years old persons

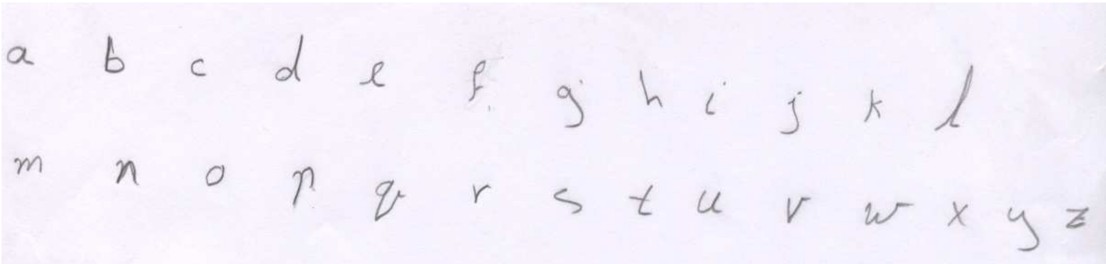
	Moment	Moment with wavelet	Complex moment	Complex moment with wavelet	Wavelet level 1	Wavelet level2	Wavelet level3
Kohenen OCON	86.53	87.50	88.46	88.46	98.07	100	96.15
LVQ	87.50	88.46	92.30	94.23	100	100	96.15

The recognition rates are similar in all feature extraction techniques and neural network architectures.

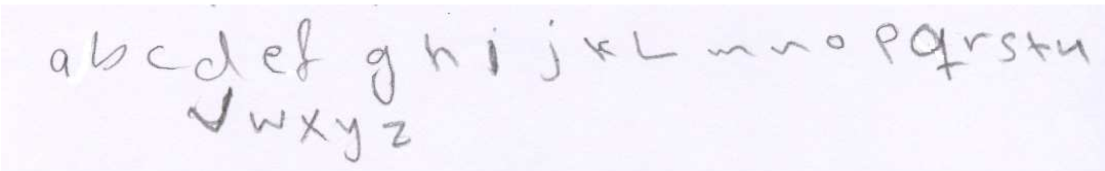
In fig (4.14), samples taken from the database, contains the small letters written by a person.



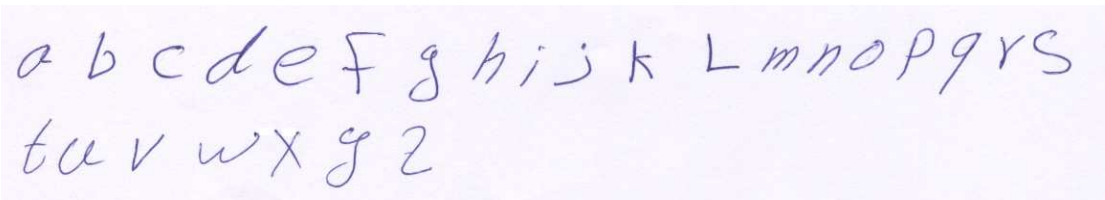
(A)



(B)



(C)



(D)

Fig 4.14 Database Samples

This sample contains an obvious variation in the writing of characters. The general shape of character itself changes (as in letters f, l , q ,p),also the bending degree of the writing style in each sample and for each letter.

These samples are some of the samples that miss-recognized some of there letters by the system.

See in Fig 4.15, some English characters are rotated and scaled versions of other characters.

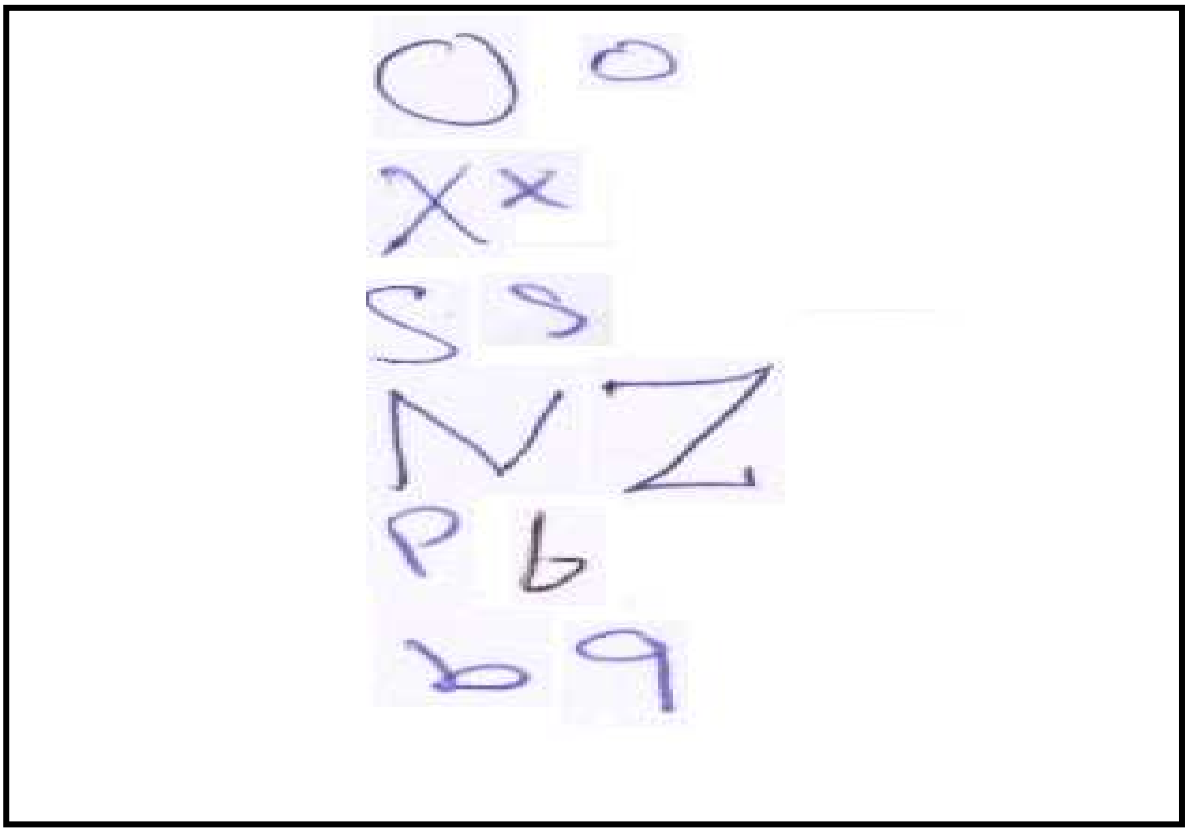


Fig 4.15

The Characters" "O", "x" "X", "s" "S", we notice that the capital cases of these characters are scaled versions of the small cases, although the system succeeds in recognition of those characters without missing the case of these characters.

Characters like “N” “Z”, we notice “Z” is a rotated version of “N”, also characters like “p” “b”, and “b” “d” have such a behavior, and the system succeeded in recognition of these characters.

Some problems raises in the early steps of the system and causes mistakes in recognition, Fig (4.16) shows some characters have a problem of disconnected edges; it may caused by the writer, or during the thresholding process.

Character “B” for example, will be presented to the feature process as two separated characters, and the same thing will happen to the characters “D”, ”E”, and “K”, so characters with such cases will not be recognized correctly.

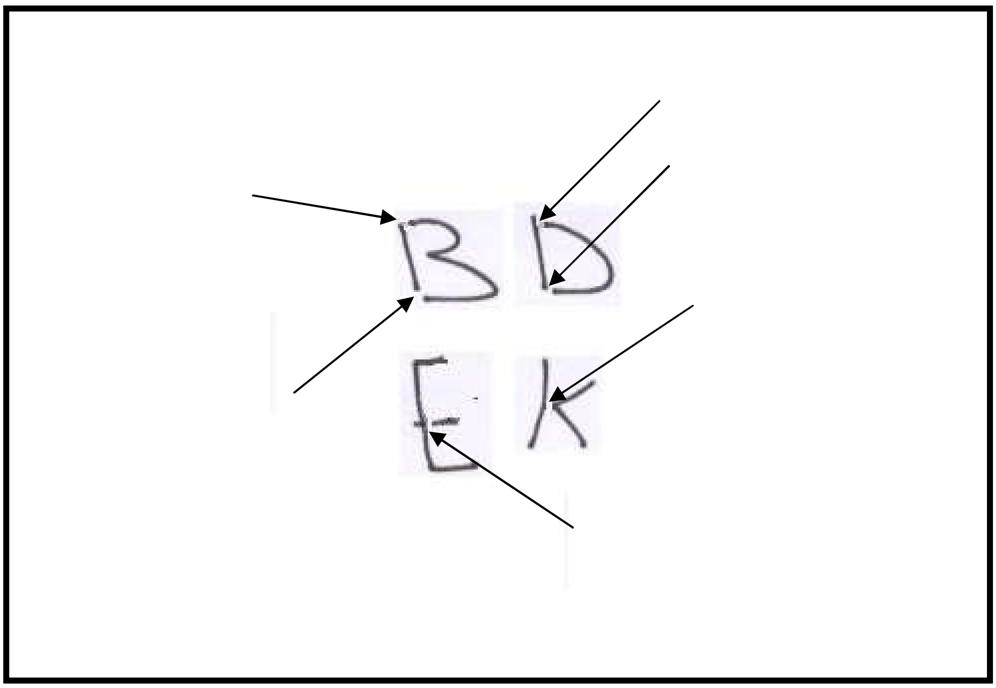


Fig 4.16

Chapter Five

Conclusions and Future Work

5.1 Conclusions

1- The recognition rate outlined in the previous chapter indicates that Kohenen Neural Network OCON architecture and LVQ scores the highest recognition rate with all feature sets presented to this model.

2- The weakness of the ACON neural network in term of recognition rate, is due to the overlapping of the clusters centers between the characters in the feature space, this makes the neural net miss recognize most of the characters, but in LVQ an OCON nets, the required separation is provided between the clusters, so the recognition rate rise up into the excepted level.

3- Results shows that wavelet transformation coefficients gives better recognition rates in comparison with other features extraction techniques, and the second level of resolution in wavelet transformation gives the best description to characters image and thus gives the highest recognition rates among all feature extraction techniques used in this thesis.

4- Wavelet transformation enhance the recognition rates with small amount when combined with geometrical features like moments and complex moment, this because the description level of moments and complex moments overcome the description provided by wavelet transformation.

5- Expansion of the database will increase the recognition rate with kohenen OCON Neural Network, as we can see in fig (5.1); the recognition rate is increasing as the number of samples increases.

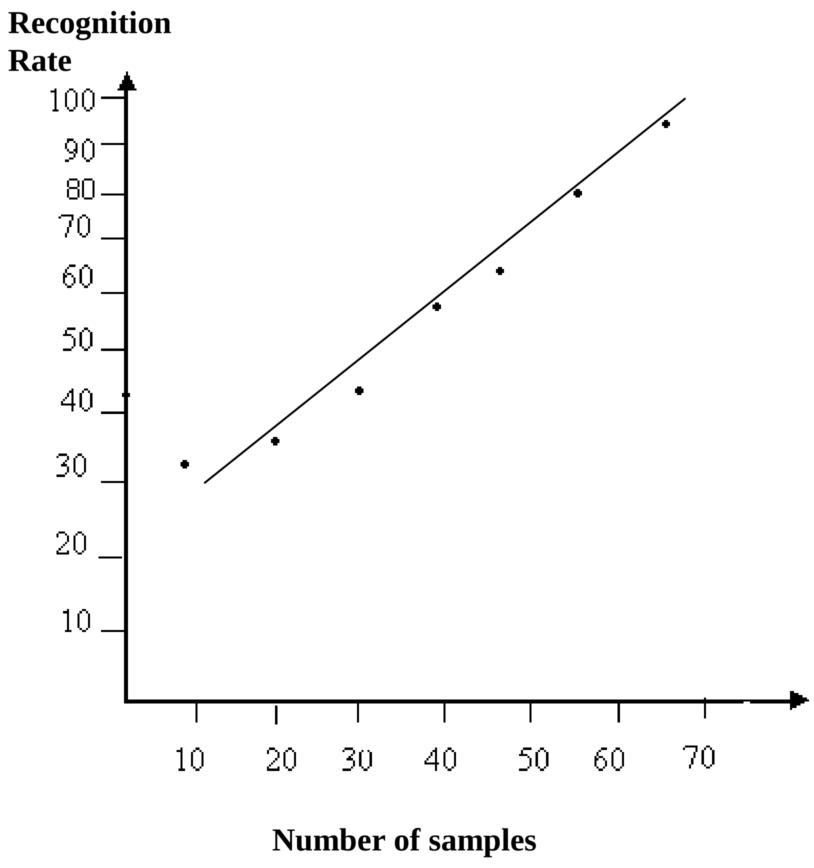
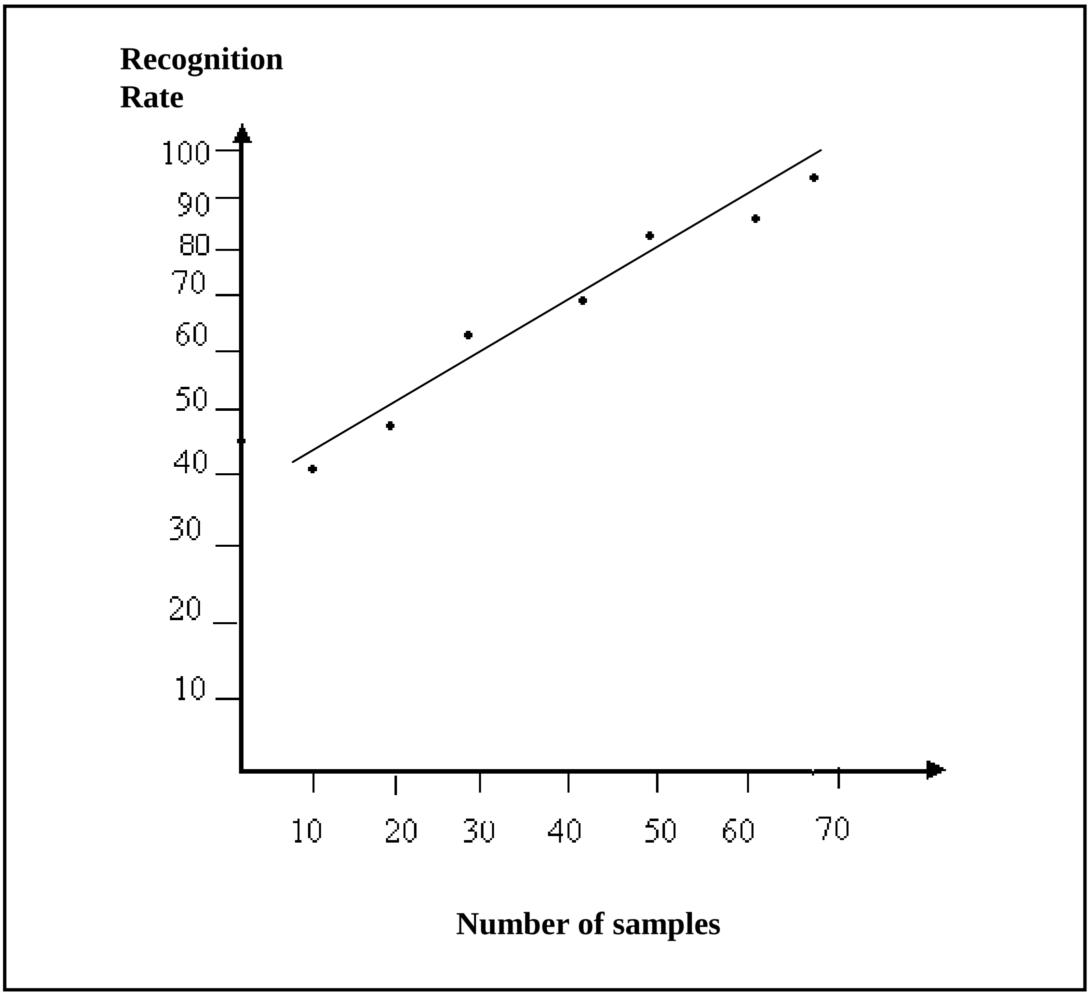


Fig 5.1 Number of training samples against recognition rate for OCON kohenen neural network

Some of the characters were written by different persons contains high difference in their general shape. This considered a drawback for the recognition rate, and it will be minimized if the number of training sample increases.

Fig (5.2) shows the increase of recognition rate as the number of samples increases for LVQ Neural Network.



Nu

Fig 5.2 Number of training samples against recognition rate LVQ neural network

6- The recognition rate is not affected by the age or gender of the person.

5.2 Future Work

1. Using Wavelet transformation as feature extraction techniques with other geometrical features may increases the recognition rate to expectable level.
2. Investigate other Neural Network architectures in the system to observe the recognition rate it may presented with these architecture.
3. Using other types of wavelet transformation like Symlet, Coiflet, and Biorthogonal.
4. Using Genetic algorithms to select the best feature set that describes the inputs to the system.

Chapter Four
Implementation of Web Site Search Engine

Chapter One

Introduction to Character Recognition

1.1 Overview

Character recognition problem is widely studied in pattern recognition field for the last decades because of both it theoretical value in pattern recognition and its numerous application like:- automatic processing of post addresses in mail letters, automatic money amounts determination in bank checks, processing and analyzing of handwritten documents[Mac96].

One of the most common divisions between character recognition systems lies whether the recognizer focused on ***handwritten text*** or ***machine printed text***. Recognition system of handwritten text can be further classified into two main classes depending on the way recognition system acquire data that are:-***on line recognition system*** in which data is presented to the system as a sequence of coordinates (x, y, t), where t is time. ***Off line recognition system*** where data is presented to the system as a bitmap image typically from a scanner [Mac96].

On the other hand, character recognition system can be classified according to the recognition level, in which there are two general approaches, the first approach involves segmenting the text into individual characters and recognizing each character separately. The second approach involves segmenting the text into words and recognizing each word as a complete entity using some global features of the word and using the semantic roles of the language [Gad00].

In general, pattern recognition systems have general style of work and have common phases in order to perform its required job. In this thesis, we will discuss handwritten character recognition systems in specific way; although we will mentions some of its features as pattern recognition system features in general since handwritten character recognition systems is pattern recognition system in the origin.

1.2 Problem Definition

The advances in the field of handwriting identification and recognition result from a better understanding of the fundamental biophysical and psychophysical processes involved in handwriting generation, and the application of this knowledge to various types of specific systems. From the view point of biophysics and psychophysics, handwriting can be represented as the output of a space-time variant system equivalent to the writer, where the input is a learned motor program, and described by the curve-linear displacement, the angular displacement, and the torsion of the trajectory according to the intrinsic properties of curves, certain neurons within the brain fire with a predetermined intensity and duration, and this nerve network activate the proper muscles in a predetermined order. The motion of pen on the paper resulting from muscle contraction and relaxation leaves a partial trace of the trajectory of the pen tip. How the motor program is constructed, used

and influenced by other biophysical and psychophysical mechanisms is still an open question, but the design of a handwriting identification or recognition system is based on the fact that people do not write according to a standard penmanship, and the deviation from the norm is individual dependent[Bed98].
Handwritten character recognition proves to be challenging problem due to the large variance the data may exhibit, not only there are changes and distortions from one writer to another, but even for samples produced by the same writer [Bed03].

Researches on recognition of handwritten character recognition have made impressive progress and many systems have been developed, however there is still a significant performance gap between handwritten and machine printed character recognition systems [Red03]. Machine printed characters are uniform in size, position and pitch for any given font. In contrast, handwritten characters are non-uniform; they can be written in many different styles and sizes by different writers and by the same writer. Therefore, the reading of machine printed writing is a much simpler task than reading hand-printed writing and has been accomplished and marketed with considerable success [Sin99]. In fact, several companies have produced low-cost software that can, when used with scanners, read in documents of several fonts. However, in spite of many years of research, optical readers that can recognize handwritten materials at a satisfactory rate are rare. Even if they exist, they typically can read only hand-printed digits that must be legible and consistently placed [Sin99].

Another issue arises in the context of handwritten text is that of both word and character segmentation, determination of the word boundaries, and which strokes are grouped together to form character. therefore researchers gives three major categories into which non cursive text may be grouped from a segmentation point of view [Mac96]:-

1. Box mode: - characters are written in predefined boxes.
2. Ruled mode: - characters and words are written in a predefined lines.
3. Unruled mode: - characters and words may be written anywhere on the input surface and may also slope arbitrary.

In box mode, segmentation is trivial. In ruled and unruled mode, segmentation problem could turn out to be very difficult. Therefore segmentation should not be under estimated. Incorrect segmentation will lead to poor recognition and affects the performance of the overall system [Mac96].

1.3 Literature Survey

Many techniques have been presented in the recent decades in the domain of handwritten recognition in the way of reaching best performance for a problem that been proved to be a challenging one ,and of course it's impossible to explore all the approaches to this problem ,but there is some interesting ones that score high recognition rates and adapted new techniques that proved it's abilities in other fields and other problems of the same field of pattern recognition.

The approaches presented to character recognition process focuses on two phases or sub-processes within character recognition system, that is:-

1. Feature extraction.
2. Character recognizer.

For each of the phases above, many techniques have been presented to improve the overall performance of the system and score high recognition rates, and because of that, each phase has been considered as research field of its own.

In the next sections of this chapter we will examine each of these phases and some abstract overview for some of the techniques presented for each phase.

1.3.1 Feature Extraction

Object recognition is generally performed on either the raw image in the image plane or on the feature representation in the feature space. In the earlier case, known the low level image recognition, the system learns and recognizes an object according to the information given by all the pixels in the image plane. In an N×N image plane, the object is described by an image vector which consists of N^2 pixel values. The size of the image vector increases as the resolution of the object image increases [Cha95].

One of the drawbacks of this approach is a huge dimensionality which depends the computational burden of the system. Moreover, the image vector of the shifted object image may be quite different from the original one. On the other hand, not all the pixels of the object image reveal crucial information of the object characteristics, and there is a large redundancy in the image vector [Cha95].

The feature-based recognition uses only the information that best characterizes the object. It extracts the important information conveyed by some pixels and processes it to obtain the feature representation. The object in the image plane is then represented by its feature vector in the feature space. In this case, the learning and the recognition is done in the feature space. Dimensionality of the input vector is greatly reduced, and the recognition can be invariant to some image transformations, such as image ***translation, rotation,*** and ***scaling***, if the object features are properly selected [Cha95].

Many methods have been introduced to extract features; generally it could be classified into two main categories:

1. Geometrical features: - those features that are directly extracted from the image (spatial domain). Many techniques have been introduced in this category that describes the shape of the pattern.

2. Transformation features: - those features that are extracted after a transformation is made to map the image from the spatial domain into another domain.

Most Researchers presentations in this field fall into those two groups.
In the next sections we will explore some of these presented approaches.

1.3.2 Geometrical Features

- [Mac95] proposed the use of ***endpoints, intersection points*** and ***loops*** found in characters as features, this fairly small collection gives a good recognition results, although all these features assumed to be extracted from a skeletonized bitmap, the main drawback of this approach is that it highly depending on the spatial resolution of the image, in case of a disconnection is found in character edge wrong recognition will happen for this character.

- [Cha95] used the ***angle sequence*** and ***vector contour*** representation as feature extraction technique. Angle sequence contains detailed local information of the object shape, since it records the orientation relationships between adjacent boundary pixel pairs. On the other hand, the vector contour representation developed provides global information of the object shape. Local change of orientation affects slightly the vector contour representation of an object. Thus, the angle sequence is sensitive to local changes, and that will make the vector counter also inadequate since it is a consequence computations made on angle sequence.

- [Sin99] presented two types of feature extraction techniques and compare them with raw date presented to the classifier; these are mesh-***based feature*** and ***directional distance distribution feature***. Strengths of mesh features include its simplicity and its generalization power. But the overall feature did not improve recognition rates beyond that of the non-feature based classifier. In the other hand directional distance distribution feature has been chosen for its generalization power and also for its simplicity. The computation time of the extraction algorithm is relatively short since its dominant operations are integer comparison and addition; it also has, in fact, proved to increase recognition results in system especially when compared with non feature classifiers.

- [Cov00] performed on-line recognition of Thai alphabetical characters by using geometrical feature; information extracted for each written character is encoded and gathered into several groups such as character's width/height ratio and directional code.

- [Bel02] made an approach for recognition of the Indian Bangla handwritten numerals in presenting a new technique called ***water overflow from the reservoir***, Reservoir concept was used both for touching numeral segmentation and isolated numeral recognition resulting a good recognition rate for this language characters.

- [Red03] proposed an approach for feature extraction named ***bar Mask encoding method*** bar mask used in the experiment is similar to the seven segment alpha numeric display used in the familiar to digital or electronic watches, although this approach suffer from the problem of rotation and bending of characters.

- [Cha03] used ***seven moment invariance*** as geometrical features in developing an OCR machine, the characteristics of moment's invariant and the results of this research seems to be promising to be used for handwritten characters recognition, noticing that

[Cha03] uses also ***Fourier transformation*** as feature extraction method for achieving of such promising results.

1.3.3 Transformation Features

- [Cha95] used of ***Fourier transformation*** to generate feature set from Fourier coefficients and points to the powerful characteristics of Fourier transformation of being invariant to image rotation and translation, it's main disadvantage is the neglecting of all spatial features and depends highly on frequency features.

- [Mac95] Notice the absence of spatial characteristics in Fourier descriptor, and presents ***Gabor transformation*** as feature extraction method that overcome the problem of Fourier transform mentioned above and pointing to the wavelet transformation as a promising technique for combining spatial and frequency details to produce an efficient feature set.

- [Bed98] use wavelet packet transformation approach and refer to the time-frequency localization and compression capability of wavelet packet transform using best-basis algorithm is used for feature extraction, enhancing the accuracy of recognition at pixel level.

1.3.4 Character Classification

In general term, classification is the process of assigning each object, from set of objects, to one of a set of classes, in pattern recognition, the object is a pattern extracted from the image, and the classes are various categories occurring in the image. the pattern in this step of processing not usually a set of points, but, it is a set of numerical features formed by feature extraction process[Nib86].

There are different models of classification system, the main two models are ***statistical***, and ***artificial neural network*** [Ven93]. Since most of neural networks approaches are based on statistical method, the concentration will be on statistical and neural network classifiers.

There are different types of neural network, some of these types have been applied to the problem of handwritten character recognition and some of these researches score high recognition rate [Sab01], and the tests shows that the recognition rate has been increased and the overall performance has been enhanced in comparison with other recognition techniques.

1.3.5 Statistical Classification

- [Day95] constructed a mixture of ***locally linear generative models*** of a collection of pixel-based images of digits, and used them for recognition. Different models of a given digit are used to capture different styles of writing, and new images are classified by evaluating their likelihoods under each model. [Day95] used an

EM-based algorithm in which the M-step is computationally straightforward ***principal components analysis (PCA)***.

- [Bel02] used **a *binary tree classifier*** for the numeral recognition**,** building the tree was made in the training phase that will generate a binary tree where a leaf node of the tree may contain up to 3 numerals. Next, he counts on more specific feature to identify numerals of different leaf nodes.

- [Hsi03] made a comparison between techniques of statistical classification and Neural Network classification, his results shows that the overall performance of the character recognition system increased in the domain of recognition rate and speed of recognition.

1.3.6 Neural Network Classification

- [Ver00] presented a description of an implemented system for the recognition of printed and handwritten postal addresses, based on Artificial Neural Networks (ANNs). Two classification methods were compared for the task of character and address recognition. [Ver97] compared two neural network techniques, measuring recognition rate and accuracy. feed forward, multi-layered neural network with ***Error Back propagation (EBP)*** algorithm were used for training the neural network, The presented results, clearly indicated that the back propagation ANN was the best classification method for the task.

- [Cov00] used the classification method of ***back-propagation neural network*** for classifying on-line Tahi characters, and shows that recognition rate is satisfactory in comparing with other systems working with the same language, reliability rate indicates that the overall accuracy of recognition tends to optimistically increased in this technique.

- [Red03] proposed an approach that combines the unsupervised and supervised learning techniques for unconstrained handwritten numeral recognition. This approach uses the ***Kohonen self-organizing neural network*** for data classification in the first stage and the ***learning vector quantization (LVQ) model*** in the second stage to improve classification accuracy, showing that supervised Neural Network perform better than unsupervised techniques.
- [Sab01] applied Neural Network approach as a classifier for character recognition problem, and uses back propagation Neural network model in different Network architectures, these are All Classes One Network (***ACON)*** and One Class One Network (***OCON)***, and compared the results between these architectures and show that using one network for each class has better performance and scores high recognition rate.

- [Cha03] proposed a simple ***voting scheme*** for off-line recognition of hand printed numerals. One of the main features of the proposed scheme is that this is ***not script dependent***. Another interesting feature is that it is sufficiently fast for real-life applications. In contrast to the usual practices, presented and studied the efficiency of a majority voting approach when all the classifiers involved are

multilayer perceptrons MLP of different sizes, on wavelet transforms at different resolution levels.

1.3.7 Combined Approaches

- [Par97] explored and tested handwritten digit recognition with the ***multilayer Perceptron*** and the ***Fuzzy ARTMAP,*** and compared the performance of these two neural networks. [Par97] saw that they yield similar recognition rates, but Fuzzy ARTMAP learning process is a lot faster. Moreover, Fuzzy ARTMAP is capable of incrementally stable learning.

- [Bed98] This paper presents a novel method for automatic handwritten character recognition by combining ***wavelet packet transform*** with ***neuro-fuzzy approach***. The result shows that WPT and neuro-fuzzy approach has more accurate recognition at pixel level than the one using only fuzzy logic.

- [Sue02] compared between ***genetic algorithms for feature subset selection***. The first approach considers a ***simple genetic algorithm (SGA)*** while the second one takes into account an ***iterative genetic algorithm (IGA),*** which is claimed to converge faster than SGA. The feature vector is based on a mixture of geometrical features while the classifier is a neural network trained with the ***back propagation algorithm***.

1.4 Aim Of Thesis

This research establishes a handwritten recognition system that considers three main feature extraction techniques which is ***Wavelet Transformation***, ***Moments***, and ***complex moment*** And uses ***Neural Network*** to recognize characters which are ***Kohenen Model*** in, both ***ACON*** and ***OCON*** architecture, and ***Learning Vector Quantization LVQ***.

1.5 Thesis Layout

Chapter two provides an overview for the theoretical background for handwritten character recognition system.

Chapter Three explain the practical part of the thesis and states the algorithms used in constructing the proposed system.

Chapter four presents the results and the user interface of the system and discusses certain study cases used to test the system.

Chapter five provides the conclusion and future work.

Chapter One Introduction to Character Recognition

Chapter Three
Proposed Handwritten Character Recognition System

Chapter Three

Proposed Handwritten Character Recognition System

3.1 Introduction

This thesis presents the application of artificial neural network and wavelet transformation into the problem domain of handwritten character recognition. The use of wavelet transformation coefficients is to produce features sets at different levels of resolution.

Basically, kohenen neural network model have been used in this thesis to compare between different feature sets, and two types of kohenen net are used, they are the ordinary kohenen network and learning vector quantization network (LVQ).

3.2 Proposed Character Recognition Systems

The designed character recognition system based on the pattern recognition system architecture described in chapter two with some variations in the order of preprocessing steps, some of them where shifted forward after the segmentation process because of its nature of being at character level processing.

3.2.1 Image Acquisition

Image that contains the text to be recognized is acquired by using scanner and a bitmap image with true colors will be saved for further processing as shown in fig (3.1).

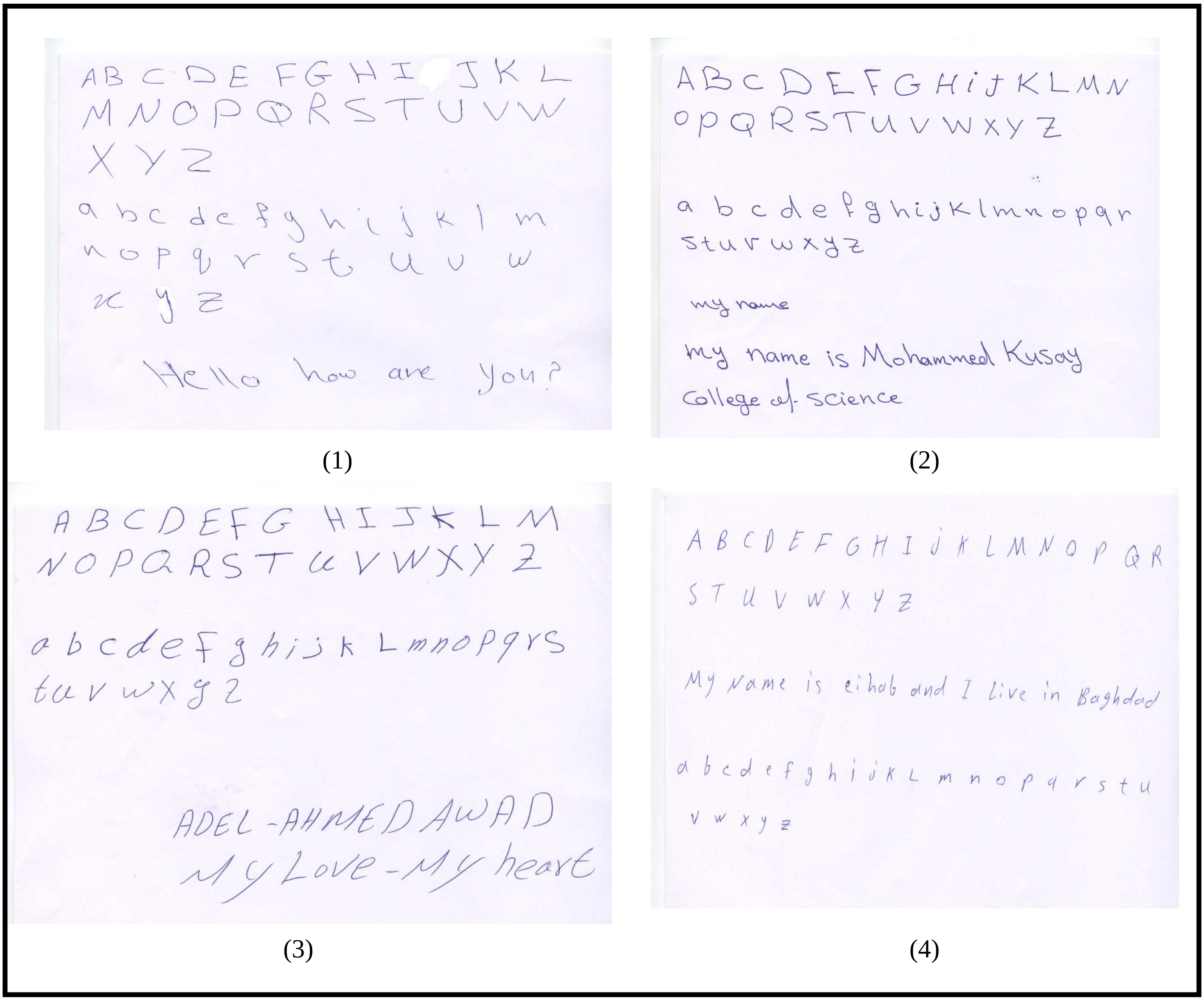


Fig (3.1) Four Samples taken from four different persons

Database is made to support the training and testing processes of the system, Samples are taken from 130 different persons, each person is asked to write all 26 English characters on a white paper in its capital and small cases as shown in Fig (3.1). Images are scanned with 150 dpi resolution, and saved as 24 bit per pixel bitmap.

3.2.2 Preprocessing

key function of preprocessing is to improve the input image in way that increase the chances for success of the following process, and using different image processing techniques for normalizing input patterns.

For character recognition, the preprocessing here will include **edge thinning** or **skeletonizing** and **noise removing**.

3.2.3 Noise Removing

Noise removing is one of the image enhancement applications in which the principal objective is to process an image so that the result is more suitable than the original for specific application. In this thesis we use smoothing filters which are used for blurring and noise removing; this filtering is useful for removal of the small details from an image prior to large object extraction and bridging of small gaps in lines and edges [Gon00].

Spatial lowpass filter is used, a 3*3 mask as shown in Fig(3-3) is to be moved over the image with coefficients have the value of 1, the response will be then the sum of the gray levels of the nine pixels divided be 9.

$$1/9 \times \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

3.2.4 Edge Thinning

Edge thinning is an attempt to represent the structural shape of a plane region and reduce it to a graph; this reduction may be accomplished by obtaining the skeleton of the region via a thinning algorithm [Gon00]. This step will reduce some of the variance that is usually found in characters like the width of their edges, therefore, it considered as a uniforming process to make all the characters with 1 pixel width [Gon00].

The algorithm will assume that edge points have 1 value and the background points have 0 value, the process consist successive passes of two basic steps applied to the contour points of the given region ,where a contour point is any pixel with value 1 and have at least one 8-neighbor valued 0, here is the algorithm [Gon00]:-

Algorithm 3.1

Let p_1 be a point in the image have value 1 and have 8-neighbor that is p_2 ,p_3 ,p_4 ,p_5,p_6 ,p_7 ,p_8 and p_9 ,Fig(3.2) shows the order of these terms.

Step (A) For each point p_1 in the image do

1) Find $N(p_1)$ which is the number of nonzero neighbor points that is: $N(p_1)= p_2+p_3+p_4+p_5+p_6+p_7+p_8+p_9$.
2) Find $S(p_1)$ which is the number of 0-1 transitions in the ordered sequence of $p_2,p_3,p_4,p_5,p_6,p_7,p_8,p_9$.
3) Test for the following conditions to be satisfied all :

   $2 \le N(p_1) \le 6$.

   a. $S(p_1)=1$.

   b. $p_2 . p_4 . p_6 = 0$.

   c. $p_4 . p_6 . p_8 = 0$.

   if all conditions satisfied then ,the point p1 is flagged .

Step (B) Delete the flagged points .

Step (C) For each point p_1 in the image do

1. Test for the following conditions to be satisfied all :

   a. $N(p_1) <= 6$.

   b. $S(p_1)=1$.

   c. $p_2 . p_4 . p_8 = 0$.

   d. $p_2 . p_6 . p_8 = 0$.

   If all conditions satisfied then ,the point p1 is flagged .

Step (D) Delete the flagged points .

p_9	p_2	p_3
p_8	p_1	p_4
p_7	p_6	p_5

Fig 3.2 Eight Neighbor Terms

Steps A and B is applied to every boarder pixel in the character region, if one or more of conditions (a) to (d) are violated ,the value of the point in question is not changed. If all conditions are satisfied the point is flagged for deletion. However, the flagged points are not deleted until all boarder points have been processed. This delay prevents changing the structure of the data during the execution of the algorithm.

In condition (a), p1 is deleted only if it has more than one pixel or less than seven pixel with value 1, having 1 pixels valued 1 implies that p1 is an endpoint of the skeleton and should not be deleted, also deleting p1 that have 7 seven pixels valued 1 will cause erosion into the region.

Condition (b) violated when it's applied to points on a stroke 1 pixel thick, deleting such point will make a disconnection in the character region.

Condition (c) and (d) are satisfied simultaneously by the minimum set of values (p_4 =0 or p_6=0) or (p_2 =0 and p_8 =0), thus, with the reference to neighborhood arrangement, the point that satisfies these condition, is an east or south boundary point or a northwest corner point in the boundary.

In either case, p1 is not part of the skeleton and should be removed. Similarly, conditions (c) and (d) given in step C satisfy simultaneously by the following minimum set of values (p_2 =0 or p_8 =0) or (p_4 =0 and p_6 =0). These points correspond to north or west boundary points, or a southeast corner point.

This algorithm will produce the same image but with all edges been thinned to 1 pixel width.

3.2.5 Segmentation

It is generally the first step in any attempt to analyze or interpret an image automatically. It is defined to be the process of partitioning the image into distinct regions that are meant to correlate strongly with objects or features of interests in the image. It is regarded as a process of grouping together pixels that have similar attributes [Gon00].

Image thresholding is a segmentation technique [Gon00], and since it classifies pixels into two categories, and there are two possible output values, therefore, threshold segmentation creates binary image, and the nature of the image depends on the property being thresholded.

Because the thresholding produces image with only two colored value, its seems better to make the segmentation process earlier than the preprocessing step because of data reduction that is made by segmentation process since it will produce an image with two possible values and that will make the preprocessing goes faster.

The algorithm used for Thresholding the image into two values 0 and 1 are stated bellow:-

Algorithm 3.2

let Image1 be the input image in which each pixel Image1 (x, y) has 3 colored values ImgGreen, ImgRed, and ImgBlue.
Let T be the threshold that will split the image pixels into two categories of pixels 0 and 1.

For each pixel in the image Img(x, y), do

1. Find gray scale for Img(x, y) :-
   Grayscale =(ImgGreen(x, y)+ImgRed(x, y)+ImgBlue(x, y))/3

2. IF Grayscale >=T then
   Set Img(x ,y)= 1
   Else
   Set Img(x, y)=0

3.2.6 Character Segmentation

In this process, the image is separated into number of sub-images, each of these sub-images contain a part of the image that may represent an English character instance, the separation or segmentation process based on ***each number of edges been related together and separated from adjacent edges is considered here to be a possible character instance***.

This is the algorithm for character segmentation:-

Algorithm 3.3

Let Img be the input image, and char() is array of records, each char(i) will store a sub image which represent a character, start point and end point of the character image .

Let Lns(i,j) be a two dimension array, where i is the number of line, and j is the number of characters found in line i.

Step (A) For each pixel (x, y) in the image Img do

1. If Img(x, y)=1 then
   a. Add this point to Char_Points list
   b. For each point p1 in the Char_Points list do
   c. Locate the 8-neighbors of p1 ($p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9$) .
   d. If p_i value =1 and p_i is not in Chat_Points list
   Then add p_i to Char_Points list .
   (i=2,3..........9)
   e. If no of points in Char_Points < 3 then
   Set found character flag to true .

2. If found character flag is true then save the character image And start and end position of the character.

Step (B) for each character char(i) in the character list do

1. if y value of start point char(i) > y of the end point of char(i-1) then

   char(i) is a start character in a new line.

2. After all character been processed ,we will have Lns(k, c).

Step (C) for each line, run a sort algorithm to put the character in it's correct occurrence in the image, sorting will depends on the x value of the start point of the characters.

In step A, image is scanned for finding characters, when a pixel is with value 1, it considered to be a possible edge point, so edge track procedure is executed, edge track procedure will start with this point an checks the neighbors of this point, neighbor points with 1 value will considered a edge points and added to the list of character points and the process will proceed through all points in the list until no more points is left.

The shape that have been extracted will be saved in the character list and will be deleted from the original image for not to be extracted twice, and the starting and ending points of the shape is also saved because of its necessity in other steps in the algorithm.

The scanning will continue for all image points. Step A will produce a list for all shapes detected in the image and considered to be possible characters.

In step B, extracted characters will be arranged into lines, the basic idea is that characters in the same line lies within specific horizontal range, so, each character will be checked if its horizontal value of its start points is larger than the end point of the previous character in the list. If it was larger then this character is belong to the next line, if its not then this character belong to the same line of the previous one. In this way the characters will be arranged into lines.

In step C the lines produced from the previous step will be processed each one by sorting its characters according to their position, in case that some characters where extracted not in the same order that they appears in the image, this case may occur in some writings when the line have a slope towards north, or when certain characters is higher than the others.
The sorting of each line is made by using a kind of sorting algorithm.
After these steps are finished, we will have a number of shapes that might represents characters, extracted from the image, arranged into lines, each of these shapes will be preprocessed and then features will be extracted for finding out if they are characters and classified into 26 character classes.

3.2.7 Further Preprocessing

After the characters been separated, each character need more processing before being ready for the feature extraction step. In this step the character will be ***scaled*** to standard size, and ***shifted*** into a standard position. This step is important because it will reduce the variance caused by shifting and scaling of the character. Here are the algorithms used for this step.

a - Scaling

Each character sub-image will be scaled to 64*64 size so to put in all character in a uniform size by using the ***Nearest Neighborhood Interpolation***, and here is the algorithm bellow:-

Algorithm 3.4

Let character image be Img1 with size h1 *w1 and the scaled image will be Img2 with size h2*w2

For each pixel (x2,y2)in Img2 do

1. Find the corresponding pixel (x1,y1) in Img1 to have the value of Img2(x1,y1) pixel

   X1 = Round (x2 * ((w1 - 1) / (w2 - 1)))

   Y1 = Round (y2 * ((h1 - 1) / (h2 - 1)))
2. Img2(x2,y2) =Img1(x1,y1)

In this thesis, the size of the scaled image h2 and w2 will be equal to 64.

b - Shifting

In this process, the character will simply be shifted toward the upper left corner of the image, so that all the characters will be almost at the same position of the image.

Here is the shifting algorithm:

Algorithm 3.5

Let the input image is Img1 which contains the characters and the shifted character image will be Img2:-

Find the row r at which the character starts.

1. Find the column c at which the characters start.
2. For each pixel (x2,y2) in Img2 assign the value taken from the corresponding pixel(x1,y1) in Img1 with row shift r and column shift c .

3.2.8 Feature Extraction

Feature extraction process will generate feature sets that represent characters occurrences in the image, in our work, features are extracted by using two different techniques, ***wavelet transformation*** and ***character moments***.

a - Wavelet Transformation

Each character image will be transformed by using wavelet transformation; coefficients produced in this transformation will act as features sets.

There are different types of wavelet transformations depending on the bases functions used in the transformation, in our work; Haar transformation is used to implement wavelet transformation. The Haar bases vectors are:-

Lowpass: $\frac{1}{\sqrt{2}}[1,1]$

Highpass: $\frac{1}{\sqrt{2}}[1,-1]$

The transformation will be implemented by convolving these vectors with rows and then with columns.

Wavelet transformation will be applied to all characters been extracted in the previous step, here is the algorithm of Haar transformation used in this work [Umb98]:-

Algorithm 3.6

For each character image in the character list do:

1. Convolve the low-pass filter with rows and save the results.
2. Convolve the lowpass filter with the columns (of the result from step 1) and subsample this result by taking every other value, this will give us lowpass/lowpass version of the image.

Convolve the result form step 1, the lowpassed filtered rows, with the highpass filter on the columns. Subsample by taking every other value to produce the lowpass /highpass image.

3. Convolve the original image with the highpass filter on the rows and save the results.
4. Convolve the result form step 4 with the lowpass filter on the columns, subsample to yield the highpass lowpass version of the image.
5. Convolve the columns from step 4 with highpass filter to obtain highpass version of the image.

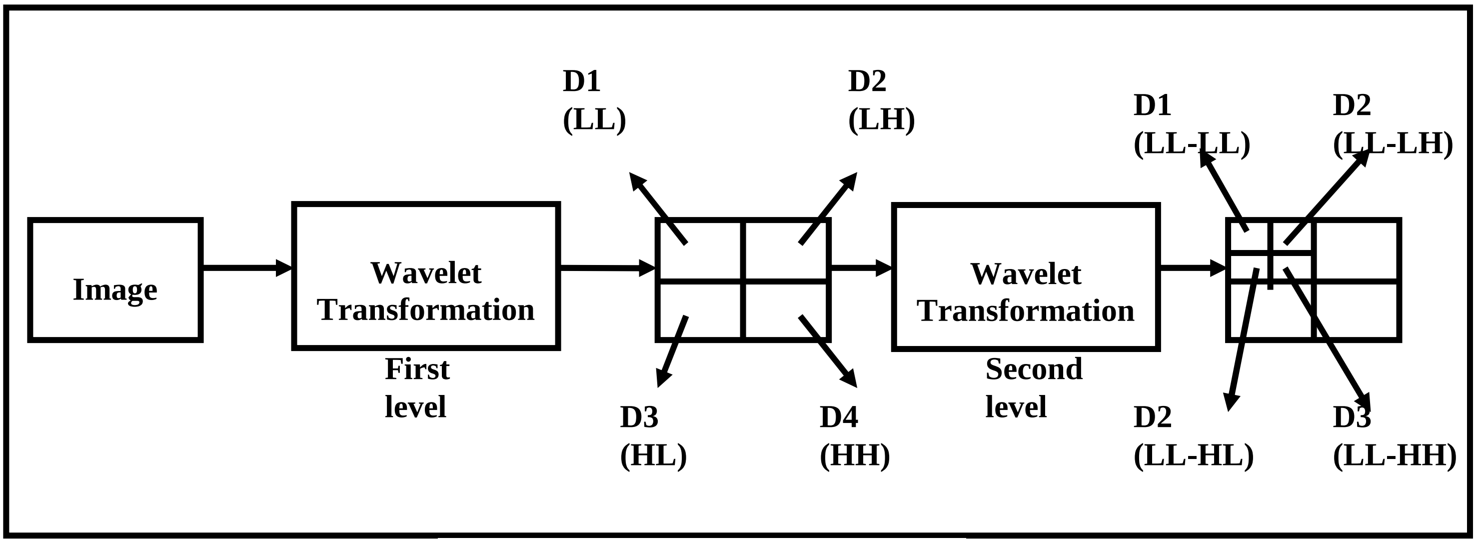


Fig 3.3 Wavelet transformation

In this work, wavelet transformation is applied character image with size 64*64 and with three level of resolution, in the first level 4 sub-bands are resulted , as shown in Fig(3.3) ,each of these subbands(D_1 , D_2 , D_3 and D_4) is 32*32 size ,each of these subbands can act as Feature set, because of the unique distribution and reduction of characteristics that describe the image. For the first band D_1, it corresponds to the lowest frequencies and contains the global characteristics of the image, D_2 will give the vertical high frequencies and contains the horizontal details, D_3 will give the horizontal high frequencies and contains the vertical details, and D_4 will give the high frequencies in both diagonal directions and contains the diagonal details. From that we see that each subband has a certain way of describing the character image, so, it's promising to use these bands as feature sets.

For the next level of resolution, the algorithm will be applied only on D_1 Since it contains the global characteristics and describes the general shape of the character, this will yields also four sub bands, with the same manner, with only one difference that the size of each band is reduced to 16*16.

For three level of resolution, this method will produced 3 feature sets that describe the character.

b - Moments

The shape of boundary segments can be described quantitatively by using moments. This method cannot deal with disjoint shapes where single closed boundary may not be available. For region-based invariants, all of the pixels of the image are taken into account to represent the shape. Because region-based invariants combine information of an entire image region rather than exploiting information just at the boundary points, they can capture more information regarding the image.

In this work, each character image is described by using the seven moments. And the algorithm applied in this work is as follow [Gon00]:-

Algorithm 3.7

Let p and q equal to 5(p and q =1,2,3....) ,for each character image Img(x, y) do:-

1.find m_{pq} from the equation :-

$$m_{pq} = \sum_x \sum_y x^p y^q \, Img(x, y)$$

2.find $\bar{x}$ and $\bar{y}$ from the equations :-

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad , \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

3.find μ_{pq} from the equation :-

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q \, Img(x, y)$$

4.finally calculate 7 moments from the equations:-

$$\phi_1 = \mu_{02} + \mu_{20}$$

$$\phi_2 = (\mu_{02} - \mu_{20})^2 + 4\mu_{11}^2$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{11} - \mu_{03})^2$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$$

$$\phi_5 = (\mu_{30} + 3\mu_{12})(\mu_{30} + \mu_{12})\ [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

$$\phi_6 = (\mu_{02} + \mu_{20})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$$

$$\phi_7 = (3\mu_{21} + \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{12} - \mu_{30})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$$

Each character image will be described by these 7 moments values.
The character image is also transformed by using wavelet transformation for one level of resolution, and the algorithm above is applied to each of the four subbands to produce 28 moment value that will represent the feature set that describe the character.

c - Complex Moment

Complex moment $C_{pq}^{(f)}$ of order (p + q) of the image f(x, y) is defined by equation 2-3 where the real and imaginary parts of the complex moments of the local power spectrum are proposed as features.

In this work, the order of the moments used is 3, and the procedure of extracting complex moments from character image is as follow:

Algorithm 3.8

For each character image do:

1. `Find the real and imaginary parts of moment term for the first order by applying the equations:

$$M_r^1 = \sum_y \sum_x x \operatorname{Im} g(y,x) \qquad M_i^1 = \sum_y \sum_x y \operatorname{Im} g(y,x)$$

2. Find the real and imaginary parts of moment term for the second order by applying the equations:

$$M_r^2 = \sum_y \sum_x (x^2 - y^2) \operatorname{Im} g(y,x) \qquad M_i^2 = \sum_y \sum_x 2xy \operatorname{Im} g(y,x)$$

3. Find the real and imaginary parts of moment term for the third order by applying the equations:

$$M_r^3 = \sum_y \sum_x (x^3 - 3xy^2) \operatorname{Im} g(y,x) \qquad M_i^3 = \sum_y \sum_x (3x^2 y - y^3) \operatorname{Im} g(y,x)$$

4. Find the complex Moment terms for the first and second and third order from the equations bellow :

$$M^1 = \sqrt{M_r^{1^2} + M_i^{1^2}} \qquad M^2 = \sqrt{M_r^{2^2} + M_i^{2^2}}$$

$$M^3 = \sqrt{M_r^{3^2} + M_i^{3^2}}$$

By applying this procedure to each character image, we will have 3 features for each character. An extension is made to this procedure by applying these equations on each of the four bands resulted from wavelet transformation for the first resolution level, therefor we will have another 12 features for each character.

3.3 Classification

The process of classification will operate on the feature sets produced by feature extraction process; each character here is recognized and classified according to its features. Neural Network classifier is used to classify characters instances, and self organization Neural Network model was used in this work.

3.3.1 Kohenen Neural Network

Kohenen Neural Network is one of the unsupervised competitive learning neural networks. It is widely used as unsupervised by using data clustering process, and bellow is the description of the Kohenen network used in this work.

The network architecture consists of two layers of nodes first layer consist of N nodes, where N is number of the features presented to the network. The output layer consists of 26 nodes representing the possible classes the input patterns (or characters instances) may fall into.

a - Classifier Architecture

The inputs domain in this work considered being the English characters in its capital and small cases, so, the classifier will have total

52 outputs representing the character classes. For each character case (capital and small) a classifier is made to deal with each case.

The ***Minimum Designating*** layer used to make the final decision by selecting the minimum output value from the two classifiers (minimum distance is the most probable answer about the class of the input).

Each of these classifiers is build as ***unsupervised Kohenen Neural Network*** and ***Supervised LVQ Neural Network***, and in Kohenen Neural Network, two types of architectures were used in the designing of the neural net that is ***OCON*** and ***ACON***.

b - OCON and ACON

Small case character classifier and Capital case classifier are designed each in this work with two architectures, ACON and OCON. In ACON, all character classes are within one neural net.

In OCON, each character has its own neural net that is each net represents a character class, in which the training algorithm is applied to each net alone, while in ACON; the training will include all characters at once.

C - Learning Phase

The network will be trained by using 130 sample taken from 130 different persons, each person have been asked to write the 26 characters from (a) to (z) and from (A) to (Z) in their capital and small form without any writing constrains. 70 of these samples have been used for training, and all 130 samples are used for testing.

The algorithm used for training the network is as follow [Sam99]:

Algorithm 3.9

Step (1) Let $x = \{x_1, x_2, ..., x_n\}$ be the training pattern vector.

Let n be the input vector length (number of features).

Let m be the number of output nodes (number of clusters or classes).

Step (2) initialize the weight vectors w_j for j=1,...,m .

Initialize α learning rate (0<α<=1) .

Step(3) while stopping condition is false ,repeat steps 4-10.

Step(4) for each input vector x ,do steps 5-7

Step(5) for each j compute

$$D_j = \sum_i (w_{ij} - x_i)^2 \quad \text{for } j=1,....,m$$

Step(6) find index j such that $D_j = \min_{j=1,\ldots,m} (D_j)$

Step(7)Update the winner weight vector(cluster)

$$w_{i,j}^{t+1} = w_{i,j}^t + \alpha(x_i^t - w_{i,j}^t)$$

Step(8)Update the learning rate , α .

Step(9)Find $E^t = \| W^{t+1} - W^t \|$.

Step(10)If $E^t \le \varepsilon$ stop ,else go to Step(3) .

d- Weights and Learning Rate Initialization

Weights and learning rate are initialized by using different kinds of procedures presented by researchers, in this work, initialization is done as followed [Sam99]:-

Algorithm 3.10

- learning rate α is initialized to a small positive number $<\alpha'$,where $0<\alpha'<0.1$,and $0<\alpha<\alpha'$.

- As the learning process progresses , α is increased slowly toward α' $(\alpha=\alpha+0.2*\alpha)$.
  This slow increasing in α forces the weight vectors to be close to the input vectors through gradual separation of the weight vectors according to the input clusters used for training .

- When α become $>=\alpha'$ and the learning proceeds , α is decreased $(\alpha=\alpha*0.999)$slowly until the stopping condition is reached .

- Weights vectors are initialized by choosing randomly 26 input vectors to represents the 26 classes ,these input vectors will represents the weight vectors that learning process will starts from.

e - Stopping Condition

The network training is stopped when the weight vectors changes are very small for two successive iterations [Sam99].

3.3.2 LVQ Neural Network

Learning vector quantization is a supervised learning extension of the Kohenen network, in which designated classes of the training samples are known in advance.

The network consist of N input nodes, representing the number of features representing each pattern, and 26 output class, each class represented by M number of nodes.

a - Learning Phase

The main difference between the ordinary Kohenen network and the LVQ network is that the training pattern is classified in advanced, so the training is said to be supervised.

The training process for LVQ rewards a winning neuron if it belongs to the correct class that specified in advance by moving it toward the input vector, and punishes it if the winning neuron does not belong to the correct class by moving it away from the input vector.

The algorithm used in this work is as follow [Sam99]:

Algorithm 3.11

Step(1) Let $x = \{x_1, x_2, ..., x_n\}$ be the training pattern vector.

Let T be the correct class for the training vector.

Let C_j be class represented by j-th output unit.

Let n be the input vector length.

Let m be the number of output nodes(number of clusters).

Step(2) Initialize the weight vectors w_j for j=1,...,m.

Initialize α learning rate (0<α<=1).

Step(3)While the stopping condition is false, repeat steps 4-9.

Step(4)for each input vector x, do steps 5-7

Step(5)for each j, compute

$$D_j = \sum_i (w_{ij} - x_i)^2 \quad \text{for j=1,....,m}$$

Step(6) find index j such that $D_j = \min_{j=1,...,m} (D_j)$.

Step(7)Update weight vectors w_j as follow:

$$w_{i,j}^{t+1} = w_{i,j}^t + \alpha(x_i^t - w_{i,j}^t) \text{ if } T = C_j$$

$$w_{i,j}^{t+1} = w_{i,j}^t - \alpha(x_i^t - w_{i,j}^t) \text{ if } T \neq C_j$$

Step(8)Reduce the learning rate α

Step(9)Find $E^t = \|W^{t+1} - W^t\|$.

Step(10)If $E^t \le \varepsilon$ stop ,else go to Step(3) .

b - Weight Initialization

Initialization of the learning rate for LVQ network is the same as ordinary Kohenen network. Weight vectors initiated by the weight vectors resulted from Kohenen network. This weight initialization will improve the convergence speed rather that using other initialization techniques.

c - Stopping Condition

The learning phase for the LVQ network will stop by using the same condition used in Kohenen Network learning.

Chapter Two
Theoretical Background

Chapter Two

Theoretical Background

2.1 Introduction

Many attempts have been made to make pattern recognition systems reaches a level of performance that is close to emulating human capabilities in performing general image analysis functions. research in biological and computational systems continually is uncovering new and promising theories to explain the human visual cognition, however, the state of art in computerized image analysis for the most part is based on heuristic formulations tailored to solve specific problems, for example some machines are capable of reading printed, properly formatted documents at a speeds that are orders of magnitude faster than the speed that the most skilled human reader could achieve. However the systems of this type are highly specialized and thus have a little or on extendibility. That is, current theoretic and implementation limitations in the field of image analysis imply solutions that are highly problem dependent [Gon00].

This chapter concerned with the pattern recognition systems that will be applied to our character recognition system, describing its phases, operations required within each phase, some feature extraction techniques and types of classification methods used in such systems.

2.2 Elements of Handwritten Recognition Systems

Techniques used in pattern recognition systems could be divided into three basic phases: (1) low level processing, (2) intermediate processing, and (3) high level processing. Although these subdivisions have no definitive boundaries, they do provide a framework for categorizing various processing that are inherent components of an autonomous pattern recognition system. Fig (2.1) illustrates these phases [Gon00].

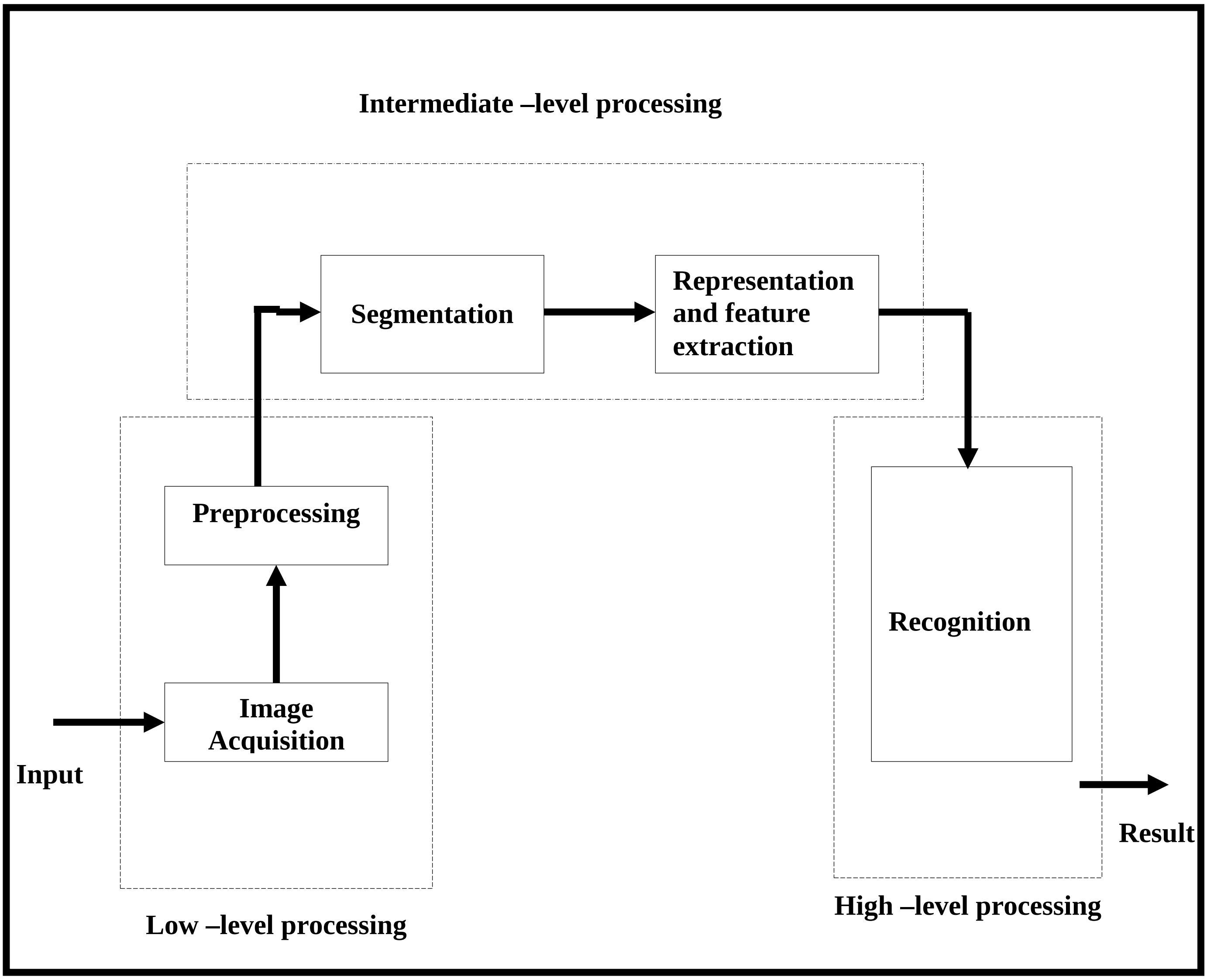


Fig (2.1) Areas of Pattern Recognition System

2.2.1 Low Level Processing Phase

This phase deals with functions that may be viewed as automatic reactions, requiring no intelligence on the part of the image recognition system. Image acquisition and preprocessing are considered here as low level functions, this area encompasses activities from the image formation process itself to compensations, such as noise reductions and image deblurring. Low level functions may be compared to the sensing and adapting processes that a person goes through when trying to find a seat immediately after entering a dark theater from bright sunlight. The intelligence process of finding an occupied seat cannot begin until a suitable image is available. The process followed by the brain in adapting the visual system to produce such an image is an automatic unconscious reaction [Gon00].

a. Image Acquisition

This process is to acquire digital image from physical world, by using scanner or digital camera. the image produced by the imaging device contain the data that is to be processed, and further information will be extracted from, therefore, image status, like resolution, brightness, and defects comes from low performance of the imaging device will highly affects the next processing steps and as consequence will define limits to the performance of the recognition system and the recognition rate.

b. Preprocessing

The key function of preprocessing is to improve the image in way that increase the chances for success of the other process, for character recognition, the preprocessing is usually deals with techniques of image enhancement, noise removing, and other techniques required.

2.2.2 Intermediate Level Processing

In this level of processing, techniques are made to characterize image components and extract information needed for recognition. Segmentation techniques is required here to segment the text into separated characters, and the feature extracting technique is required here in way that each character will be described by a set of features rather than it's raw representation, at which all characters will be recognized based on those features.

a – Character Segmentation

It is generally the first step in any attempt to analyze or interpret an image automatically. It is defined to be ***the process of partitioning the image into distinct region that are meant to correlate strongly with objects or features of interests in the image***. it also be regarded as a process of grouping together pixels that have similar attributes.

In general, autonomous segmentation is one of the most difficult tasks in image processing. This step in the process determines the eventual success or failure of the whole process of analysis. Therefore, considerable care should be taken in choosing and implementation of the segmentation techniques [Gon00].

In the context of handwritten character, segmentation has a crucial importance for its direct effect on the recognition rate and on the performance of the overall system. Therefore, from the segmentation point of view, there are three major categories into which noncursive text may be grouped:-

1. ***Box mode***: - characters are written in predefined boxes.
2. ***Ruled mode***: - characters and words are written in a predefined lines.
3. ***Unruled mode***: - characters and words may be written anywhere on the input surface and may also slope arbitrary.

In box mode, segmentation is trivial. In ruled and unruled mode, segmentation problem could turn out to be very difficult [Mac96].

b – Binarization

Segmentation algorithms for images are generally based on one of the two basic properties of gray level values: ***discontinuity*** and ***similarity***. In the first category, the approach is to partition an image based on abrupt changes in gray level. The principal area of this category ***is the detection of isolated points and detection of lines and edges in an image***. The principle approach in the second category is based on Thresholding, region growing, region splitting, and merging [Nib86].

The concentration here will be on the category of similarity and ***Thresholding segmentation*** techniques will be used in our system.

Thresholding is one of the most important approaches to image segmentation. Thresholding transforms a data set containing values that vary over some range into a new data set that containing just two values.

It does this by applying threshold to the input data, input values that fall below the threshold are replaced by one of the output values; input values at or above the threshold are replaced by the other output value[Gon00].

Image Thresholding is a segmentation technique because it classifies pixels into two categories, and because there are two possible output values, threshold segmentation creates binary image, and the nature of the image depends on the property being thresholded [Nib86].

Suppose that the gray level histogram shown in Fig(2.2) corresponds to an image, $f(x, y)$, composed on light object on dark background, in such a way that objects and background pixels have gray levels grouped into two dominant modes. One obvious way to extract the object from the background is to select a threshold T that separates these modes. then any point (x, y) for which $f(x, y) > T$ is called an object point; otherwise the point is called a background point [Gon00].

The success or fail of Thresholding techniques depends critically on the selection of an appropriate threshold, an obvious solution is to relay on intervention by a human operator, who can vary the threshold until acceptable results are achieved [Umb98]. However, this is not possible in cases where fully automatic segmentation is required. Alternatively, we might be able to determine in advance single, fixed threshold that will always give good results. this is feasible only in highly constrained imaging scenarios, when we have control over the lighting conditions and the degree of contrast that exist between different image features, Fig(2.2) shows the threshold T in an image histogram[Gon00].

Another approach is to make T equal to the mean gray level of the image. The idea here is that the mean lies between two extremes of gray level, one representing the features of interest and the other everything else. Clearly this works for images containing bright objects on a simple dark background, or vice versa [Gon00].

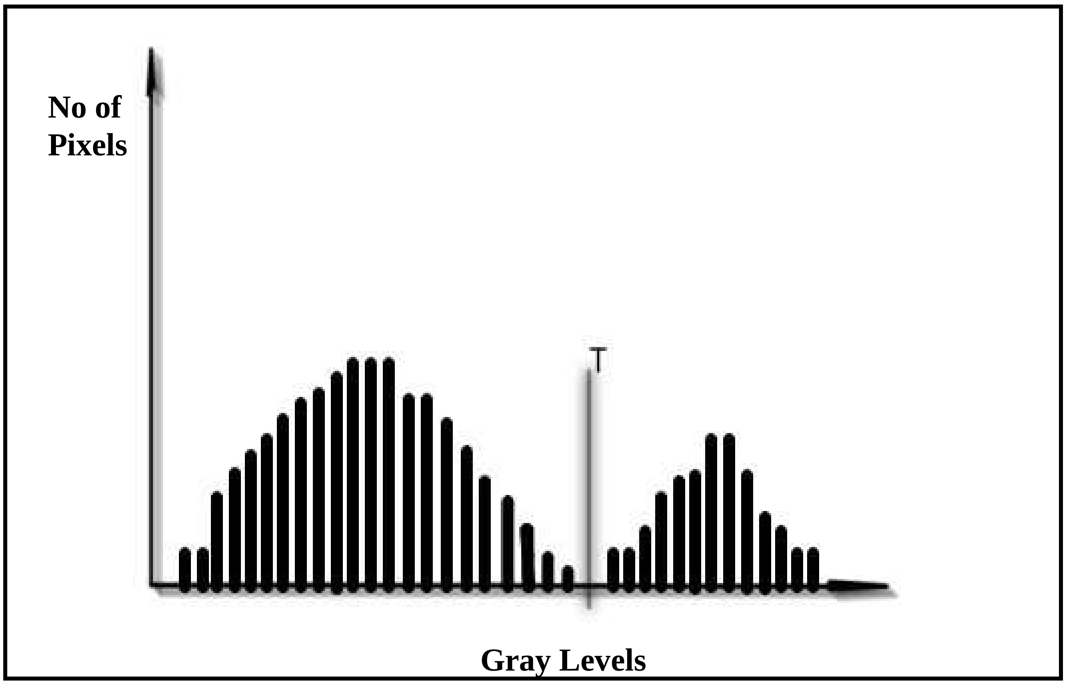


Fig (2.2) Gray level Histogram that can be partitioned by single Thresholding

A more general approach to the threshold selection involves analyzing the histogram of an image; this is based on an assumption, true only in certain situation, that different features in an image give rise to distinct peaks in its histogram. If the assumption hold true, then we may distinguish between two features of different gray level by Thresholding at a point between the histogram peaks corresponding to those two features.

c -Features Extraction

Object recognition is generally performed on either the raw image in the image plane or on the feature representation in the feature space. In the earlier case, known the low level image recognition, the system learns and recognizes an object according to the information given by all the pixels in the image plane. In an N×N image plane, the object is described by an image vector which consists of N2 pixel values. The size of the image vector increases as the resolution of the object image increases [Cha95].

One of the drawbacks of this approach is a huge dimensionality which deepens the computational burden of the system. Moreover, the image vector of the shifted object image may be quite different from the original one. On the other hand, not all the pixels of the object image reveal crucial information of the object characteristics, and there is a large redundancy in the image vector [Cha95].

The feature-based recognition uses only the information that best characterizes the object. It extracts the important information conveyed by some pixels and processes it to obtain the feature representation. The object in the image plane is then represented by its feature vector in the feature space. In this case, the learning and the recognition is done in the feature space. Dimensionality of the input vector is greatly reduced, and the recognition can be invariant to some image transformations, such as image ***translation, rotation,*** and ***scaling***, if the object features are properly selected [Mac96].

Many methods have been introduced to extract features; generally it could be classified into two main categories:

1. Geometrical features: - those features that are directly extracted from the image (spatial domain). Many techniques have been introduced in this category that describes the shape of the pattern; the most interesting technique is the Moment.

2. Transformation features: - those features that are extracted after a transformation is made to map the image from the spatial domain into another domain.

Now we will explain two important features, belong to the geometrical features, which are the ***moment*** and the ***complex moment.***

i. Moments

Moment-based invariant are the most common region-based image invariant which have been used as pattern features in many applications. Hu first introduced a set of moments-based invariant using nonlinear Combinations of regular moments. Hu's seven moment invariants have the desirable properties of being invariant under image translation, scaling, and rotation. Regular moments are defined as:

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy \quad \text{for p, q = 1, 2, 3.....} \qquad \textbf{\textit{(2.1)}}$$

A uniqueness theorem states that if f(x, y) is a piecewise continuos and has nonzero values only in a finite part of the xy plane, moments of all orders exist and the moment sequence (m_{pq}) is uniquely determined by f(x, y). Conversely, (m_{pq}) uniquely determines f(x, y). The central moments can be expressed as [Gon00]:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy \qquad \textbf{\textit{(2.2)}}$$

Where

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \text{and} \qquad \bar{y} = \frac{m_{01}}{m_{00}}$$

For digital image

$$\mu_{pq} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^p (y-\bar{y})^q f(x,y) \qquad \textbf{\textit{(2.3)}}$$

The central moment of order up to 3

$$\mu_{10} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^1 (y-\bar{y})^0 f(x,y)$$

$$= m_{10} - \frac{m_{10}}{m_{00}}(m_{10})$$

$$= 0$$

$$\mu_{11} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^1 (y-\bar{y})^1 f(x,y)$$

$$= m_{11} - \frac{m_{10}m_{01}}{m_{00}}$$

$$\mu_{20} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^2 (y-\bar{y})^0 f(x,y)$$

$$= m_{20} - \frac{2m^2{}_{10}}{m_{00}} + \frac{m^2{}_{10}}{m_{00}} = m_{20} - \frac{m^2{}_{10}}{m_{00}}$$

$$\mu_{02} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^0 (y-\bar{y})^2 f(x,y)$$

$$= m_{02} - \frac{m^2{}_{01}}{m_{00}}$$

$$\mu_{30} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^3 (y-\bar{y})^0 f(x,y)$$

$$= m_{03} - 3\bar{x}m_{20} + 2\bar{x}^2 m_{10}$$

$$\mu_{12} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^1 (y - \bar{y})^2 f(x,y)$$

$$= m_{12} - 2\bar{y} m_{11} - \bar{x}\ m_{02} + 2\bar{y}^2 m_{10}$$

$$\mu_{21} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^2 (y - \bar{y})^1 f(x,y)$$

$$\mu_{03} = \sum_x \sum_y (\mathrm{x} - \bar{\mathrm{x}})^0 (y - \bar{y})^3 f(x,y)$$

$$= m_{03} - 3\bar{y} m_{02} + 2\bar{y}^2 m_{01}$$

The normalized central moments, denoted η_{pq} are defined:-

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}{}^{\gamma}} \qquad \textbf{\textit{(2.3)}}$$

Where

$$\gamma = \frac{p+q}{2} + 1 \qquad \text{For p + q = 2, 3...}$$

A set of seven invariant moments can be derived from the second And third moments [Cha03]:

$$\phi_1 = \mu_{02} + \mu_{20} \qquad \textbf{\textit{(2.4)}}$$

$$\phi_2 = (\mu_{02} - \mu_{20})^2 + 4\mu_{11}^2 \qquad \textbf{\textit{(2.5)}}$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{11} - \mu_{03})^2 \qquad \textbf{\textit{(2.6)}}$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \qquad \textbf{\textit{(2.7)}}$$

$$\phi_5 = (\mu_{30} + 3\mu_{12})(\mu_{30} + \mu_{12})\ [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \qquad \textbf{\textit{(2.8)}}$$

$$\phi_6 = (\mu_{02} + \mu_{20})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \qquad \textbf{(2.9)}$$

$$\phi_7 = (3\mu_{21} + \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{12} - \mu_{30})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \qquad \textbf{(2.10)}$$

ii. Complex Moments

Complex moment $C_{pq}^{(f)}$ of order (p + q) of the image f(x, y) is defined by:

$$C_{pq}^{(i)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - iy)^p (x - iy)^q f(x, y)dxdy \qquad \textbf{(2.11)}$$

Where i represents the imaginary part of the equation. The real and imaginary parts of the complex moments of the local power spectrum are proposed as features. These features are translation invariant inside homogeneous texture regions and give information about the presence or absence of dominant orientations in the texture [Gri99].

While for the transformation features, we will use in our thesis the ***wavelet transformation***. Hence we will introduce some important (briefly) information about this transformation.

i. Wavelet Transformation

It is well known from Fourier theory that a signal can be expressed as the sum of a, possibly infinite, series of sines and cosines. This sum is also referred to as a Fourier expansion. The big disadvantage of Fourier expansion is that it has only frequency resolution and no time resolution. This means that although we might be able to determine all the frequencies present in a signal, we do not know when they are present. To overcome this problem in the past decades several solutions have been developed which are more or less able to represent a signal in the time and frequency domain at the same time [Umb98].

The *wavelet transform* or *wavelet analysis* is probably the most recent solution to overcome the shortcomings of the Fourier transform. In wavelet analysis the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations we can speak of a multiresolution analysis [Umb98].

A wavelet φ is a function of zero average:

$$\int_{-\infty}^{+\infty}\varphi(t)dt = 0 \qquad \textbf{\textit{(2.12)}}$$

This is dilated with a scale parameter s, and translated by u:

$$\varphi_{us}(t) = \frac{1}{\sqrt{s}}\varphi(\frac{t-u}{s}) \qquad \textbf{\textit{(2.13)}}$$

The wavelet transform of f at the scale s and position u. is computed by correlating f with a wavelet atom:

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\varphi^*(\frac{t-u}{s})dt \qquad \textbf{\textit{(2.14)}}$$

ii. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is a special case of the WT that provides a compact representation of a signal in time and frequency that can be computed efficiently. The DWT is defined by the following equation:

$$DWT(s,u) = \sum_{s}\sum_{u} f(u)2^{-s/u}\varphi(2^{-s}n-u) \qquad \textbf{\textit{(2.15)}}$$

Where φ(t) is a time function with finite energy and fast decay called mother wavelet.

Analysis can be performed using a fast, pyramidal algorithm related to multirate filterbanks. As a multirate filterbank the DWT can be viewed as a constant Q filterbank with octave spacing between the centers of the filters. Each subband contains half the samples of the neighboring higher Frequency sub bands. In the pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step. This is achieved by successive highpass and lowpass filtering of the time domain signal and is defined by the following equations:

$$Y_{high}[k] = \sum_{k} \chi[n]g[k-n] \qquad \textbf{\textit{(2.15)}}$$

$$Y_{low}[k] = \sum_{k} \chi[n]h[k-n] \qquad \textbf{\textit{(2.16)}}$$

Where $Y_{high}[k]$, $Y_{lo}[k]$ are the outputs of the highpass (g) and lowpass (h) filters, respectively after subsampling by 2. Because of the downsampling the number of resulting wavelet coefficients is exactly the same as the number of input points [Uns03].

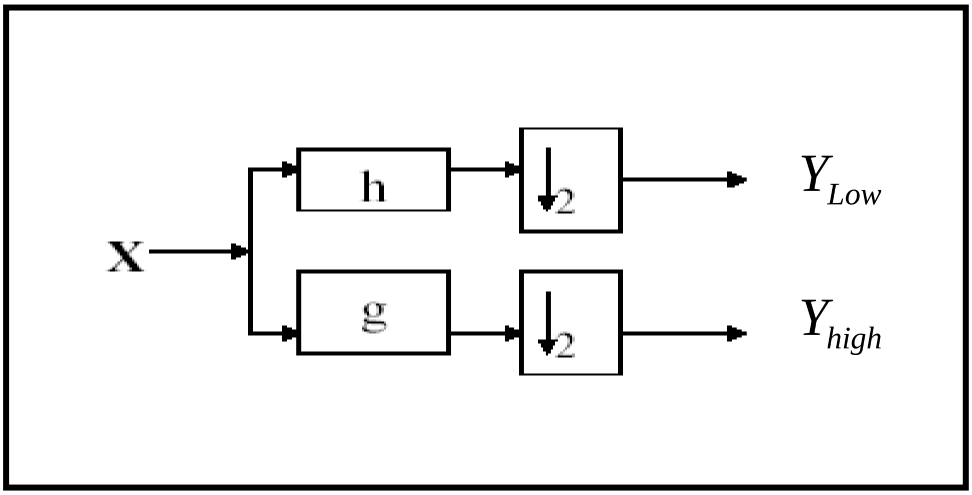


Fig (2-3) Highpass and lowpass Filters in Wavelet Transformation

iii. Wavelet Filters In Image Processing [Umb98]

Image transformations provide information regarding the spatial frequency content of the image. In general, a transformation maps image data into a different mathematical space via a transformation equation [Gon00].

The image in its native bit map representation, gives only features that describes the shape of the objects in that image, in other words, only spatial domain details. in some pattern recognition problems, frequency details of the image would serve as features on which recognition will be made on, by using spatial to frequency transformation like Fourier transformation, although, Fourier transformation expresses the image in term of sins and cosines, providing only the frequency features of the image and looses all spatial features, therefore, it seems intuitively obvious that features containing both spatial and frequency details would be superior to those limited to only one of them[Gon00].

Wavelet transformation can be described as a transform that has basis function that are shifted and expanded versions of them because of this, the wavelet transform contains not just frequency information but spatial information as well.

Wavelet transform function is basically two types of filters, high pass and low pass filters, these filters processes the image in both horizontal an vertical direction and break down it to four sub-bands or sub-sections.

One sub-band has been highpass filtered in both horizontal and vertical directions, one has been highpass filtered in vertical direction and lowpass filtered in the horizontal direction, one has been lowpass filtered in vertical direction and highpass filtered in horizontal direction, and one that has been lowpass filtered in both directions. D1 will corresponds to the lowest frequencies and contains the global characteristics of the image, D2 will gives the vertical high frequencies and contains the horizontal details, D3 will give the horizontal high frequencies and contains the vertical details, and D4 will gives the high frequencies in both diagonal directions and contains the diagonal details, As shown in the Fig (2.4) [Gon00].

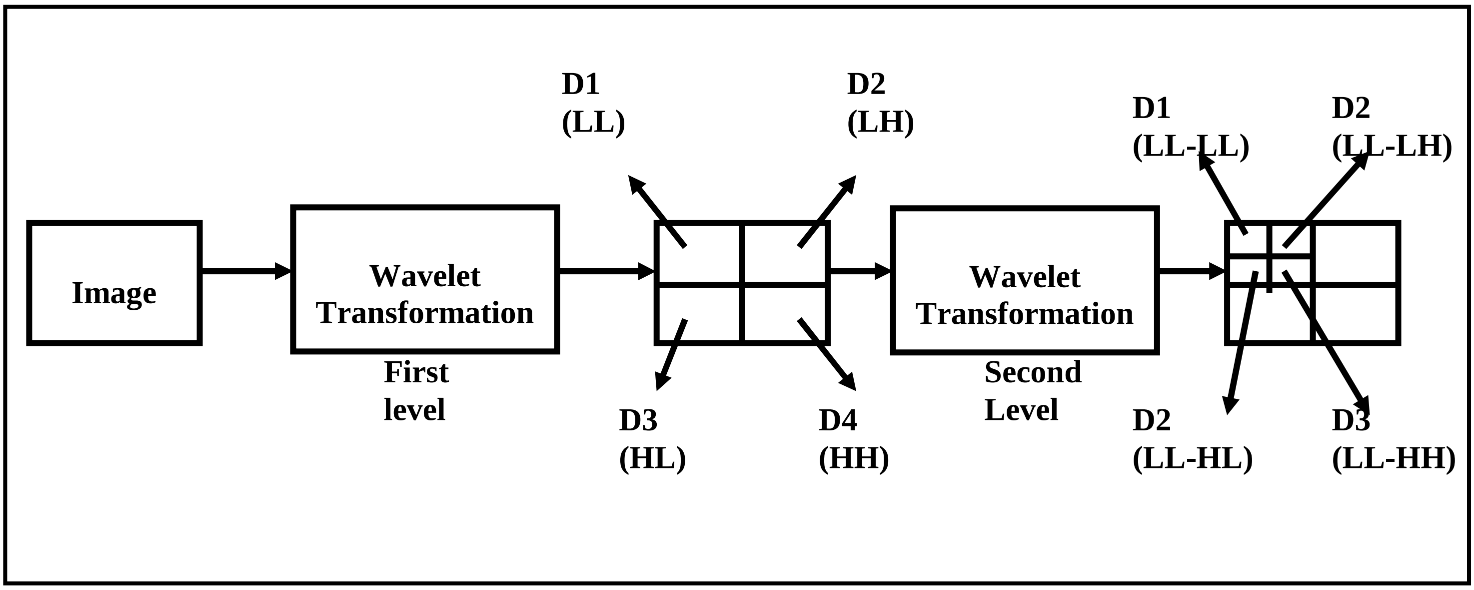


Fig (2.4) Wavelet Transformation

2.2.3 High Level Processing (Classification)

This level involves recognition and interpretation; these two processes have a strong resemblance to what generally meant by intelligence cognition. The majority of techniques used for low and intermediate processing encompasses a reasonably well defined set of theoretic formulations, however, as we venture into recognition, and especially into interpretation, our knowledge and understanding of fundamental principles becomes far less precise and much more speculative. This relative lack of understanding ultimately results in a formulation of constraints and idealization intended to reduce task complexity to a manageable level. The end product is a system with highly specialized operational capabilities [Gon00].

In general term, ***classification*** is the process of assigning each object, from set of objects, to one of a set of classes, in pattern recognition, the object is a pattern extracted from the image, and the classes are various categories occurring in the image. The pattern in this step of processing not usually a set of points, but, it is a set of numerical features formed by feature extraction process [Nibl86].

There are different models of classification system; the main three models are ***statistical***, ***syntactical***, and ***artificial neural network***. Since most of neural networks approaches are based on statistical method, the consternation will be on statistical and neural network classifiers [Ven99].

A. Statistical Parameter Based Classification

In this type of classifiers, the problem of pattern recognition is formulated as a statistical decision problem. Statistical pattern recognition is a relatively mature discipline and a number of commercial recognition systems have been designed based on this approach. In this approach, the problem is estimating density functions in a high-dimensional space and dividing this hyperspace into regions of categories or classes. Decision making in this case is preformed using appropriate discriminant function, thus, mathematical statistics forms the foundation of this approach [Mac95].

Statistical classification methods grouped into two categories, ***supervised classification*** and ***unsupervised classification***.

i. Supervised Classification

In supervised methods, the user supervises the process by initially selecting some pixels from each possible class, from these; the classification algorithm determines what each class looks like.

In these categories there are two popular supervised statistical classifications [Nib86]:-

- Bayesian Maximum Likelihood Classifier: - it's a well developed method from statistical decision theory that has been applied to problems of classifying image data.

- Minimum Distance Classification: - let M_p be the mean value of pixels in training class p. In an *n* features problem, M_p is a point (or vector) in *n* –dimensional space. The minimum distance classifier assigns a pixel to class p for which the distance from the pixel value v to M_p is minimum. Deferent distance measures may be used. the common one are :

Euclidean Distance: d (v, M_p) = $\sqrt{(v_1 - m_{i1})^2 + ... + (v_n + m_{in})^2}$ ***(2.16)***

This method is faster than the maximum Likelihood method but Less accurate.

ii. Unsupervised Classification

In this category, the classes are determined within the algorithm by locating clusters in pixel space (pixel space is the space in which a value v_1 in band 1, v_2 in band 2 ... is represented by the points, v_1, v_2 .....).

Although clusters are often fairly easy for human to identify, their centers and boundaries are difficult to be identified mathematically. One iterative clustering method is called K-Mean algorithm. K-Mean algorithm has the following steps [Nibl86]:-

1. Initially, the user should supply a set of means, or cluster centers, $M_1, M_2 \ldots M_U$ (for u clusters). Each M_p is a vector in n – dimensional space.

2. For each pixel in the image, assign the pixel to the class to whose mean is closest.

3. Recomputed the mean of each class as the average of the pixels assigned to it.

4. If any of the class means has changed significantly, go to step 2, else stop.

The advantage of the unsupervised classification is that it tends to identify clusters in the pixel space that are numerically separable; whereas the advantage of the supervised classification is that the classes that are used are meaningful to the user or the analyst.

B. Neural Network Based Classification

Several novel modes of computation have emerged that are collectively known as soft computing. They are used to exploit the tolerance for imprecision and uncertainty in real-world problems to achieve tractability, robustness, and low cost [Kun93].

Neural network is one of the major components of this approach. Neural Network can be defined as information processing systems that have certain performance characteristics in common with biological neural network that actually the basic principle of neural network working is abstract simulation of real neuron system [Fu94].

Classification systems are expected to automatically classify or cluster patterns based on there measured properties or features, in such systems, each input vector should be decided whether it belongs or not belong to a particular class or category. With this view point, a neural network (which can be defined as the system that recognizes patterns) can be used for designing a classification system [Ven93].

As with statistic classification methods, there are two basic categories of neural classifiers [Sam99]:-

i- Supervised Neural Network Classifiers: in these classifiers, learning process apply supervised learning algorithms that include a special case of reinforcement learning which deals with training instances (patterns) that already been classified into desired classes. The trained network often produces surprising results and generalizations in applications where explicit derivation of mappings and discovery of relationships is almost impossible.

ii- Unsupervised Neural Network Classifiers: in the unsupervised methods, learning algorithms deals with unlabeled training patterns, and the classes will be identified during the training by locating clusters in the pixels space, and assuming that each cluster corresponds to a class. The problem becomes a clustering identification.

This second type of classification will be discussed in the next section in more details since it will be the model that will be used in this thesis**.**

2.3 Artificial Neural Network

Neural networks are an abstract simulation of real neurons system that contains collection of neuron like processing elements communicating with each other via axon connections, these neuron units are adaptable nodes which, through the process of learning from task examples, store experimental knowledge and make it available for use, these knowledge are encoded in the neural network by the weights assigned for the neuron connections. Artificial neural network have been developed as a generalization of mathematical models of human cognition or neural biology based on the assumption that [Fu94]:

1. Information processing occurs at many simple elements called neurons.

2. Signals are passed between neuron over connections link (synaptic connections).

3. Each connection link has an associated weight which represents information being used by the net to solve a problem.

4. Each neuron applies an activation function (usually nonlinear) to its input (sum of weighted input signals) to determine its output signal.

Any neural network is characterized by:

1. **Architecture**: - which represent the connection pattern between neurons. The behavior of the network highly depends on how the neurons are arranged into layers (number of layers and number of neurons per layer), in addition to the type of connection (feed forward, feed backward).

2. **Learning or training algorithm**: - it's the procedure used for modifying synaptic weights. Learning algorithms divided into main groups ***supervised*** and ***unsupervised*** algorithms. Weight adjustments should follow certain learning rules such as **winner take all**, delta, and **perception**.

3. **Activation function:**-when the neural network is practically used to solve problem, the solution lies in the activation levels of the output units, these activation levels are calculated by the activation function used in the neural net.

There are different kinds of neural network models, and for a certain problem there exist a neural model that provides satisfactory results. Many resources have presented several taxonomy upon which neural network models can be categorize, mainly artificial neural network can by classified based on [Fu94][Kun93][Sam99][Kin95]:-

1. **Neural network architecture.**
2. **Application used in.**
3. **Learning rules**.

Fig (2.5) shows the basic classes of artificial neural network based on these three bases.

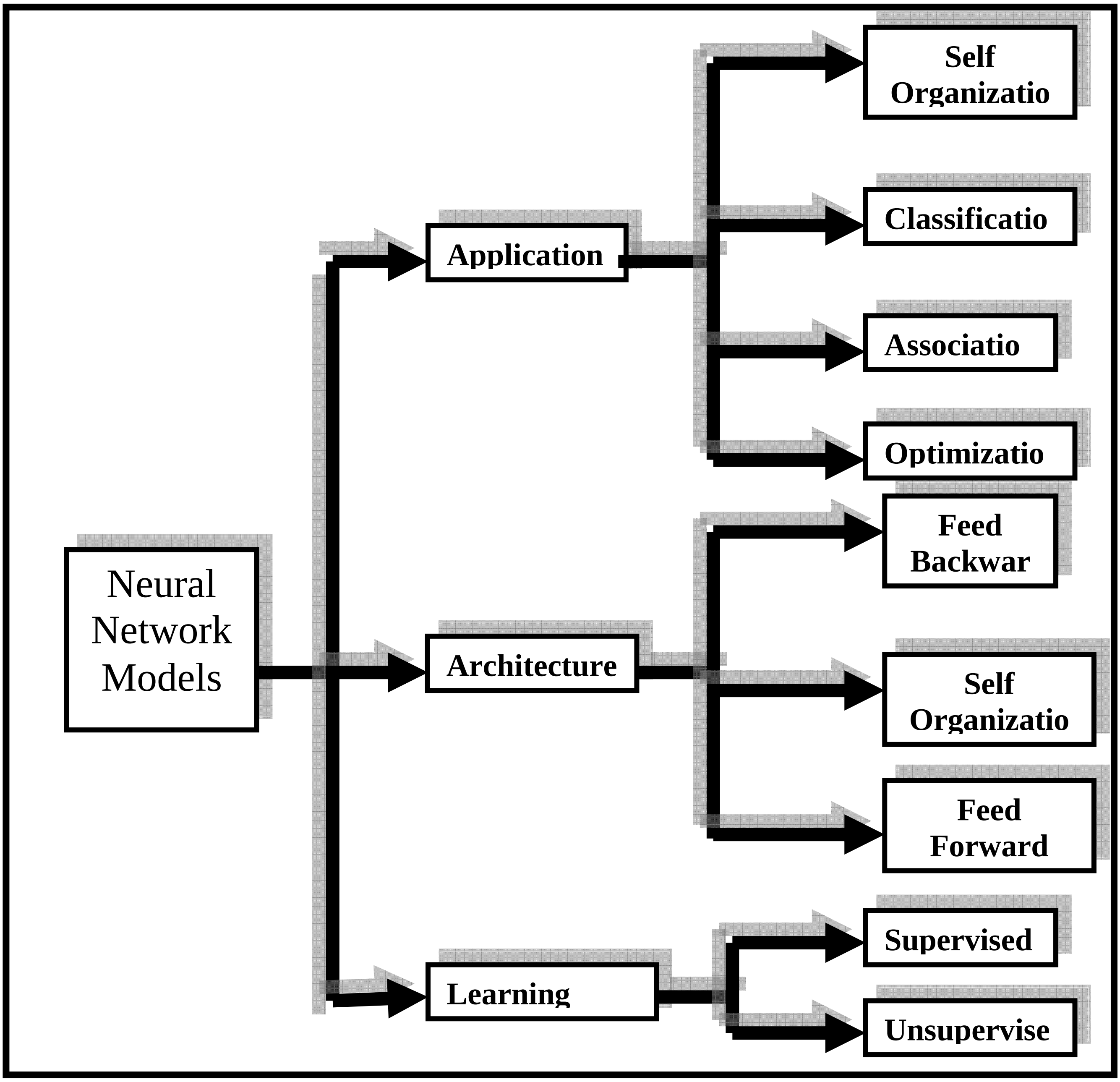


Fig (2.5) Neural Network Classification

2.3.1 Neural Network Structure

Two possible network structures were suggested to be used for handling multi-category classification problem, these are [Sam99]:-

- ***All-Classes-in-One-Network (ACON):*** the ACON structure is adapted by the conventional multilayer percepron (MLP), where all the classes are lumped into one super network. The super-net has the burden of having to satisfy all the teachers simultaneously, so the number of hidden units is expected to be very big. Empirical results confirm that the convergence rate of ACON degrades drastically with respect to the network size because the training is influenced by conflicting signals from different signal teachers. Therefore it's sometimes advantageous to decompose a huge network into many subnets. The most popular decomposition is the OCON structure.

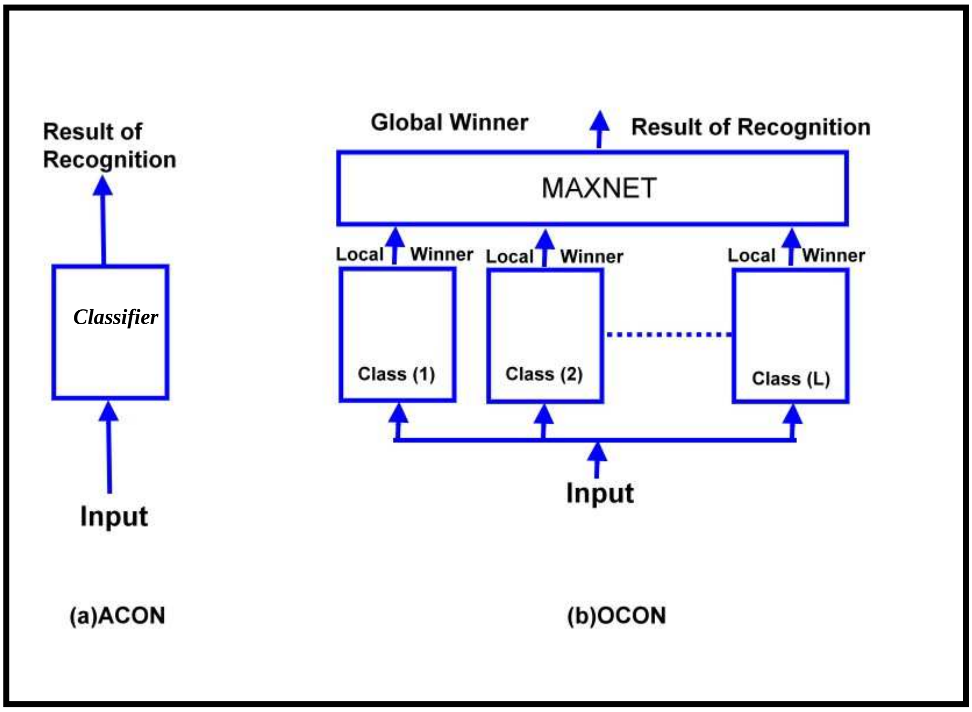


Fig (2.6) types of Network structures (a)ACON (b) OCON

- ***One-Class-in-One-Network (OCON):*** in the OCON structure, one subnet is developed to one class only.

Although the number of subnets in the OCON is relatively large, each individual subnet has considerably smaller size than the ACON super-net which may offer computational saving in the training phase and performance improvement in the retrieving phase. Fig (2.6) shows the structure difference between ACON and OCON.

2.3.2 Kohonen Neural Network Model

This type of neural network exhibit a significant behavior of the biological neural system that is the self organization, which defines the ability to discover the structure, patterns or features directly from their environment, by extracting and acting upon the regularity and similarity without benefit of a teacher (supervision)[Fu94].

Neural network apply this behavior on the basis of the clustering concept drown from the statistical pattern recognition and specifically the k-mean algorithm. the main idea of this algorithm is that the patterns of similar features clustered together, and the patterns similarity is measured by Euclidean distance between this patterns in the input pattern space, that is, patterns that have small Euclidean distance between each other is considered to be within the same cluster, clusters here represent classes of the patterns need to be classified [Mac95].

The Kohenen net architecture consists of two layers, an input layer and a kohenen layer. These two layers are fully connected. Each input layer neuron has a feed forward connection to each output layer neuron.

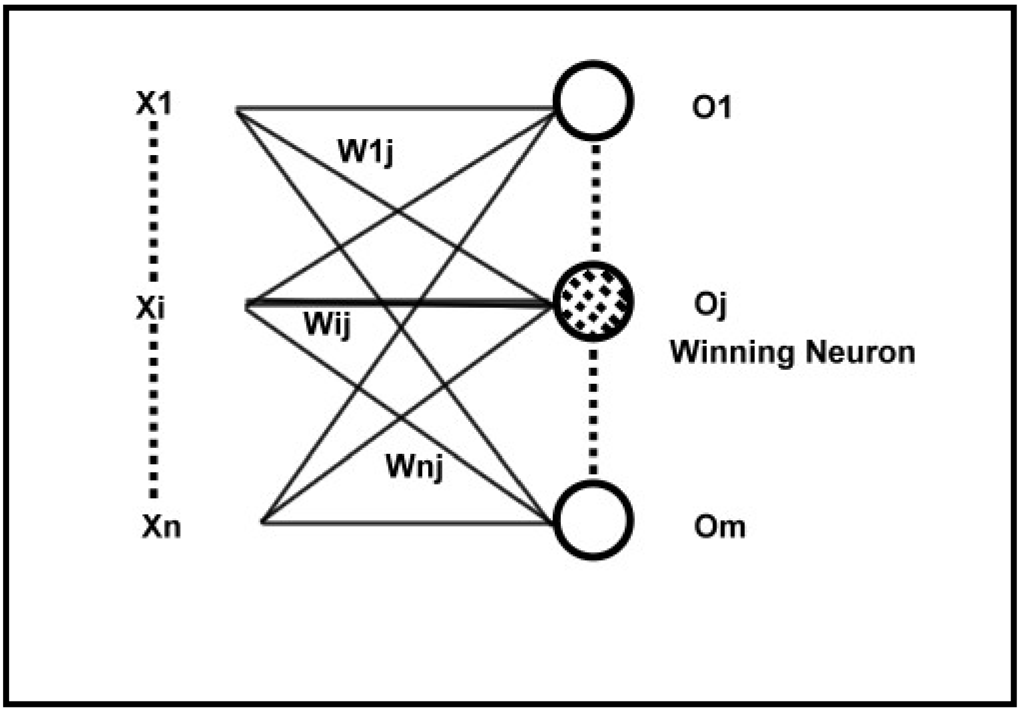


Fig (2.7) Kohenen Architecture

Kohenen network use competitive learning rule called ***winner takes all***, in essence, the node with the largest activation level is declared the winner in the competition, this nodes is the only node that will generate output signal, all other nodes are set to zero activation level, and this node is the only node that will learn from the current pattern, i.e. the weight connections of this nodes will be updated [Sam99].

The learning process here is about finding the best center point for the cluster that represent the class, this center point will be compared with each input pattern to find out how close this pattern form the cluster of this point.

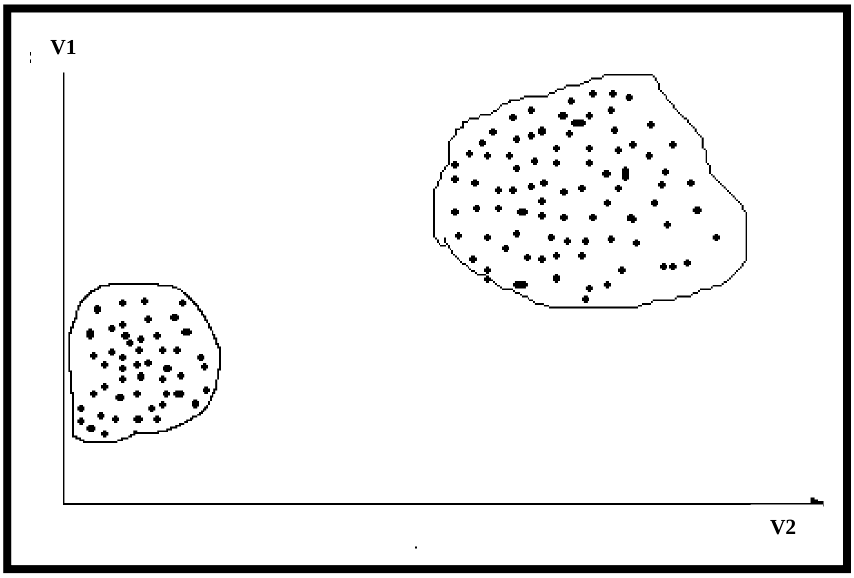


Fig (2.8) Clustering process in a two dimension feature space

2.3.3 Learning Vector Quantizer Model (LVQ)

This net is a supervised learning extension of the kohenen network method .it allow specification of the categories into which inputs will be classified .during the training phase, the output unit is positioned by adjusting their weights through supervised training, for that, it's required to know the classes (categories) of the training set in advanced and make them part of the training set [Sam99].

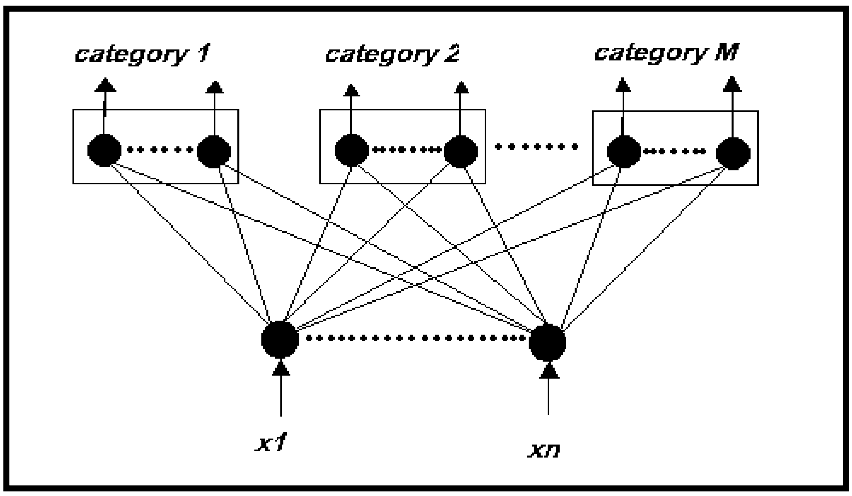


Fig (2.9) LVQ Architecture

The training procedure rewards the winning neuron if it belongs to the correct class by moving it toward the input vector, and punishes it if the winning neuron does not belong to the correct class. After the training, LVQ net classifies an input vector by assigning it to the class of the output unit that has its weight vector closest (minimum Euclidean distance) to the input vector [Fu94].
The LVQ architecture is the same with only one single exception that some neurons may be assigned to one class [Fu94].

The LVQ Neural Net like Kohenen net architecture consists of two layers, an input layer and a output layer. These two layers are fully connected. each input neuron has a feed forward connection to each output layer neuron, the only exception that in kohenen net, each output node represent a class into which patterns are assigned, while in LVQ net, each class represented by several output nodes, and that will increase the performance in case there is an overlapping between the features of different classes [Ven93].

LVQ network learning procedure has a difference from ordinary Kohenen learning procedure, in essence, the training patterns are known in advance which class it belongs to, so, when an output node is declared to be the winner, it will be checked to see if the current pattern is been assigned to the correct class, if it is, the weights will be updated to move the neuron closer to the cluster center ,if not, the neuron will be punished and move it away from the cluster center, this process will increase time of convergence to best solution[Sam99].

Chapter Five Conclusions and Future Work

Dedication

I dedicate my work to all the researchers and scientists who use the science to make the world a better place.

To all the people who sacrifice in their lives for a better future for their country and for their children.

To my country as a simple gift, to my Friends, to my Father, my Mother, and my Brothers.

Yousif

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

وَيَسْأَلُونَكَ عَنِ الرُّوحِ ۖ قُلِ الرُّوحُ مِنْ أَمْرِ رَبِّي

وَمَا أُوتِيتُم مِّنَ الْعِلْمِ إِلَّا قَلِيلًا ﴿٨٥﴾

صدق الله العلي العظيم

الأسراء-٨٥

LIST OF Abbreviation

ACON	All Classes One Network
DWT	Discrete Wavelet Transformation
LVQ	Learning Vector Quantization
OCON	One Class One Network
WT	Wavelet Transformation

List Of symbols

Symbol	Description
μ	Central Moment
C	Complex Moment
v	Feature Vector
g	High pass Filter
x	Horizontal Coordinates
f(x, y)	Image Function
h	Low pass Filter
M	Mean Value of Pixels
p	Moment Order
q	Moment Order
m	Moment Value
η	Normalized Central Moment
n	Number of Features
Y_{low}	Output of Low pass Filter
Y_{high}	Output of The High pass Filter
s	Scale Parameter
T	Threshold Value
t	Time
P	Training Class Index
u	Translation Parameter
y	Vertical Coordinates
ϕ	Wavelet Function
Wf	Wavelet Transformation
D	Wavelet transformation Sub-band

References

[Alk98] AL-Kattan A. R, Evaluation of Different Structures and Supervised Learning Parameters on the Performance of The Neocognitron ",, MSC Thesis College of Science of AL-Nahrain University, Baghdad, Iraq, 1998.

[Bed98] SCHWIEBERT S. S. & BEDI J. S," Wavelet Packet Transformation and Nero-Fuzzy Approach To Handwritten Character Recognition", www.cs.wayne.edu/~loren/papers_list.html, 1998.

[Bel02] Bellman1 C.J, "A MACHINE LEARNING APPROACH TO BUILDING RECOGNITION IN AERIAL PHOTOGRAPHS", Department of Geospatial Science, RMIT University, Melbourne 3000, Australia, 2002.

[Cha03] Chaudhuri B. B, "A Majority Voting Scheme for Multiresolution Recognition of Handprinted Numerals", www.csc.liv.ac.uk/ ~prima/ICDAR2003/Papers/0002_620_bhattacharya _u.pdf , 2003.

[Cha95] Chai S. K, "Multi Classifier Neural Network for Handwritten Character Recognition", PHD Thesis, 1995.

[Cov00] Apichart Sajjapong1, Wiwat Vatanawood and Nongluk Covavisaruch." On-line Handwritten Thai Characters Recognition", http://cgci.cp.eng.chula.ac.th/Publications.html,2000.

[Day95] Peter Dayan, "Recognizing Handwritten Digits Using Mixtures of Linear Models", Department of Computer

Science, University of Toronto, Ontario, Canada M5S 1A4, http://www.roles.nici.kun.nl/7th.iwfhr.2000/roceedings/pdf/sornal.pdf, 1995.

[Fu94] Fu L, "Neural Network in Computer Intelligence", University of Florida, McGraw Hill, Inc, 1994.

[Gad00] PAUL GADER, "Word Level Discriminative Training For Handwritten Word Recognition", University of Missouri-Columbia,http://unipen.nici.kun.nl/7th.iwfhr.2000/proceedings/pdf/paper-038-chen-gader.pdf

[Gon00] Gonzalez R. C, “Digital Image Processing ", Addison- Welsley publication Company, Inc, 2000.

[Hsi03] Hsien M. P, "Handwriiten Character Recognition", School of Electrical Engineering, University of Queens, 2003. www.wnd.wp.jp/html/hesin.browne/documents/ISRA03.pdf

[Kin95] Kinnebrock W, "Neural Network Fundamentals, Applications, Examples", Galgotia Publications, 1995.

[Kun93] Kung S. Y, "Digital Neural Network", Department of Electronic Engineering Princeton University, PTR Prentice Hall, Newjersey, 1993.

[Mac96] Macy R. B, "Pattern Recognition with Neural Network In C++ ", IEE Press, 1996.

[Nib86] Niblack W, "An introduction To Image Processing", Prentice Hall, UK, 1986.

[Red03] Reddy N. V, " Two-tier architecture for unconstrained handwritten character recognition", http://inspec.iee.org/inspec/currentawareness/pdfs/products/samples/neural_networks.pdf, 2003.

[Sab01] Suzete E. N. Correia, Jo˜ao M. de Carvalho, Robert Sabourin , "On the Performance of Wavelets for Handwritten Numerals Recognition" www.livia.etsmtl.ca/publications/2002/Suzete_icpr02.pdf, 2001.

[Sam99] Samawi V. W, "An Investigation Into the Use of Neural Networks In Texture Classification", Ph D Thesis, College Of Science AL-Nahrain University, 1999.

[Sin99] Sinha A, "Improved Recognition Module For The Identification of Handwritten Digits", Thesis, MSC, Massachusetts, Institute of Technology 1999.

[Sue02] Suen C. Y, Feature Subset Selection Using Genetic Algorithms for Handwritten Digit Recognition", www.chem.purdue.edu/.../Dongmao/List_of_reference/featureSelection/feature-subset-selection-using.pdf, 2002.

[Umb98] Umbaugh S. E, "Computer Vision And Image Processing", Ph. D, Prentice Hall, Inc, 1998.

[Uns03] Michael Unser, " Wavelet Theory Demystified", IEEE, http:// bigwww.epfl.ch/publications/unser0301.pdf

[Ven93] Venetsanopoulas A. N, "Artificial Neural Network Learning algorithms Performance, Evolution, and Application", Kluwer Academic Publishers, Massochusetts, USA, 1993.

[Ver00] Michael Blumenstein and Brijesh Verma." A Neural

Network for Real-World Postal Address Recognition", http://iteseer.csail.mit.edu/2922.html

Par[97] Martin Busque, Marc Parizeau," A Comparison of Fuzzy ARTMAP and Multilayer Perceptron for Handwritten Digit Recognition", www.gel.ulaval.ca/~mbusque/reports/artmap_digit.pdf, 1997.

Table of Contents

List Of Tables

List Of Figures

Republic of Iraq
Al-Nahrain University
College of Science

Handwritten Recognition Using Neural Network

A THESIS
SUBMITTED TO THE
COLLEGE OF SCIENCE, Al-NAHRAIN UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE IN
COMPUTER SCIENCE

By
Yousif Mohamed Basil
(B.Sc. 2002)

SUPERVISORS

Dr Eng. Sattar B. Sadkan **Dr. Venus W. Samawi**

March 2005 **Moharam 1425**

الخلاصة

تمييز خط كتابة اليد يعتبر من المواضيع المهمة و ذلك بسبب تطبيقاته المهمة إضافة إلى قيمته النظرية في مجال تمييز النماذج الصورية ،في هذا البحث استخدمت الشبكات العصبية لتمييز الأحرف المكتوبة ، مت م التطبيق لثلاثة أنواع:-

Kohenen All classes in one network, Kohenen one class in one network, and Learning vector quantization.

عملية استخراج الصفات استخدمت التحويل المويجي نوع Haar ، كما استخدمت الصفات الهندسية لاستخراج الصفات المميزة للحروف وهي العزم و العزم المعقدة ، تم بناء النظام باستخدام لغة فيجوال بيسك ٦ ، و تم بناء قاعدة بيانات مكونة من ١٣٠ نموذج اخذت من ١٣٠ شخص .

أظهرت النتائج إن التحويل المويجي مع OCON و LVQ قد حققت أعلى معدل تمييز و هو ٩٤% .

الجمهورية العراقية
وزارة التعليم العالي و البحث العلمي
جامعة النهرين

تميز كتابة اليد بأستخدام الشبكات العصبية

رسالة
مقدمة الى قسم علوم الحاسبات في جامعة النهرين
كجزء من متطلبات نيل درجة الماجستير في علوم الحاسبات

من قبل
يوسف محمد باسل
(بكالوريوس جامعة النهرين ٢٠٠٢)

أشراف

د. ستار بدر سدخان
د. فينوس وزير سماوي

آذار ٢٠٠٥
محرم ١٤٢٥