INTELLIGENT SYSTEM FOR PERSIAN CALLIGRAPHY LEARNING

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF APPLIED SCIENCES
OF
NEAR EAST UNIVERSITY

By
BEHZAD SOROUSH BARHAGHI

In Partial Fulfillment of the Requirements for the Degree of Master of Science
in
Information Systems Engineering

NICOSIA, 2016

Behzad Soroush Barhaghi: Intelligent System for Persian Calligraphy Learning

Approval of the Graduate School of Applied Sciences

Prof. Dr. İlkay Salihoğlu
Director

We certify this thesis is satisfactory for the award of the Degree of Master of Science in Information Systems Engineering

Examining Committee in charge:

Prof. Dr. Fahreddin Sadikoğlu, Supervisor, NEU

Prof. Dr. Rahib Abiyev, Computer Eng. Dept., NEU

Assist. Prof. Dr. Kamil Dimililer, Automotive Eng. Dept., NEU

Assist. Prof. Dr. Boran Şekeroğlu, Department of Inf. Systems Eng., NEU

Assist. Prof. Dr. Yönay Kırsal Ever, Software Eng. Dept., NEU

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Behzad Soroush Barhaghi

Signature:

Date:
18.7.16

ACKNOWLEDGEMENTS

It is not possible to thank everybody who has had an involvement with me during the course of my studies for the degree of MSc. However, there are some people who must be thanked.

Firstly, I would like to thank my supervisor Prof. Dr. Fahreddin Sadikoglu for his guidance and encouragement throughout the thesis. He has walked me through all the stages of the writing of my thesis. Without his consistent and illuminating instruction, this thesis could not have reached its present form.

I would like to thank especially to my lovely wife whose encouragement and support has helped me achieve beyond my greatest expectations. Without her help and support throughout the years it could not have been possible for me to come this far.

I would like to thank Ayub Kouhpayeh Hagi as a one of the leaders of Iranian Calligraphy Association who has been very helpful through the duration of my thesis.

Furthermore, I would like to thank Dr Şahin Özgur Çeri for his patience and support that helped me overcome many crisis situations and finish this dissertation. Also, I would like to thank my parents for their encouragement and prays.

To my devoted wife SAHAR …

ABSTRACT

Persian Nastaliq calligraphy is one of the most famous oriental arts. Last few years the number of the Nastaliq learners increase dramatically. Too many learners in various ages and level (30 students to a teacher), lack of specialist in this filed, too less time (2 h per week) and modern life troubles (e.g. Traffic, Working hours, etc.) lead to the ineffective learning. Learners can't exercise their writing practices sufficiently.

In this thesis, an interactive intelligent computer-based grading and guiding system is main proposed.

This study proposes for the first time an intelligent tutor system for automation of the Persian calligraphy learning process. The system incorporates image processing and machine learning technologies. Digitization, filtering, segmentation and feature extraction are image-processing techniques that prepare the appropriate input for the training phase of the multi multiclass SVMs machine learning phase. Displaying suitable feedback on screen for learners is another aim of the study. In this regard, the system provides facilities for Persian Nastaliq calligraphy learners to reduce errors that are inherent in traditional education methods, makes the process more efficient and allows people to take advantage of learning possibilities whenever and wherever they choose.

The developed system has been tested for more than 192 character and the experimental results revealed that approximately over than 90.05 percentages is working satisfactorily.

Keywords: Image Processing; Intelligent Tutor; Machine Learning; Persian Nastaliq Calligrapy; SVM

ÖZET

Farsça Nastaliq kaligrafisi en ünlü doğu kökenli sanat dallarından biridir. Son birkaç yılda Nastaliq eğitimi almak isteyen öğrencilerinin sayısı önemli ölçüde artmıştır. Çeşitli yaş ve seviyelerdeki orangi sayısı, bu alanda uzman eksikliği, eğitim için ayrılan zaman kısıtı ve modern yaşam sorunları bu çalışmanın yapılmasının esas nedenleridir. Bu tezin temel amacı, Nastaliq eğitim için bilgisayar tabanlı değerlendirme ve yönlendirme sistemi tasarlanmasıdır.

Bu çalışmada ilk kez Pers kaligrafi öğrenme otomasyonu için akıllı bir öğretmen sistemi önerilmektedir. Bu sistem, görüntü işleme ve bilgisayarla öğrenme teknolojilerini içermektedir. Tezdakı kaligrafi imajların temel görüntü işleme teknikleri olan; sayısallaştırma, filtreleme, segmentasyon ve özellik çıkarma gibi sinyal ve görüntü işleme yöntemleri incelenmiştir.

Tasarlanan bu sistem ile, geleneksel eğitim yöntemlerinde karşılaşılan hataların azaltılanması, eğitim sürecinin daha etkin ve adaylarda istedikleri zaman ve mekanda öğrenme imkanı sağlanmıştır.

Geliştirilen sistem 192' den fazla karakter için test edilmiş ve sonuçlarda yaklaşık olarak yüzde 90,50 güvenilirlik elde edilmiştir.

Anahtar Kelimeler: Akıllı eğitmen; Bilgisayar Öğrenmesi; Destekçi Vektör Makinesi (DVM); Farsça Nastaliq kaligrafi; Görüntü İşleme

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURE

LIST OF ABBREVIATIONS

PeNCIL:	Persian Nastaliq Calligraphy Intelligent Learning
ARG:	Attributed Relation Graph
DTW:	Dynamic Time Warping
SVM:	Support Vector Machine
SVMs:	Multiclass Support Vector Machine
Min:	Minimum
Max:	Maximum
CCV:	Color Coherence Vector
CM:	Color Moments
AI:	Artificial Intelligence
ANN:	Artificial Neural Networks
2D:	2 Dimension
ICC:	Isolated Calligraphy Character
RGB:	Red-Green-Blue
ROI:	Region of Interest
RBF:	Radial Basis Kernel Functions

CAHPTER 1

INTRODUCTION

1.1 Introduction

Looking at our daily lives, we can find the footprints of improving computer science and engineering in every aspect of life such as education, tourism, health, and transportation. One of the most rapidly developing areas in the field of engineering is that of intelligent art. Nastaliq calligraphy is the art of writing Persian characters (Hamidi & Rashvand, 2015). It was developed in Iran in the It originated in 14th and 15th centuries in Iran. Now, this kind of calligraphy as an art lesson sill teach interactively in traditional form of class by specific calligraphy art teachers. Students learn how to hold and wield the "Galam" (specific Nastaliq pen) to write graceful characters. In order to write a beautiful calligraphic character, learners need learn isolated characters in advance which called "Mofradat" (Rouhfar, 2014). Mofradat is the elementary skills in learning Nastaliq calligraphy. Moreover, in the writing skills, there are four material components which called 'four treasures of study' namely Galam, ink, paper and inkwell used in writing calligraphic characters. In addition, the standard samples are very essential importance for novices. Novices Learners imitate the characters from samples to improve their hand handwriting. Vast amounts of this effort called "Mashq". Teachers which called "Ostad" check all written characters and grade them then give the improving guidelines to learners interactively in traditional class. Last few years the number of the Nastaliq learners increase dramatically. Too many learners in various ages and level (30 students to a teacher), lack of specialist in this filed, too less time (2 h per week) and modern life troubles (e.g. Traffic, Working hours, etc.) lead to the ineffective learning. Learners can't exercise their Mashq sufficiently. In this thesis, an interactive intelligent computer-based grading and guiding system is main proposed.

According to the experts' experience, the assessment for Persian calligraphic characters includes the structure of characters (global features) and the smoothness of strokes (local features) (Mohsen, 2009).

The Persian Nastaliq Calligraphy Intelligent Learning (PeNCIL) is the topic of this thesis. The PeNCIL tool is presented as an intelligent standard tutorial tool to overcome the traditional method's insufficiencies. The developed system provides facilities for learners to practice individually and, in addition, the PeNCIL is capable of evaluating and providing appropriate feedback automatically on the learners' work. This system was developed based on using image processing and machine learning with character assessment abilities.

1.2 The Aim of the Thesis

The development of an intelligent platform for Persian calligraphy learners to practice individually and receive suitable feedback is the main aim of the current study. In this regard, the PeNCIL tool was developed to overcome the shortages in the classical learning method.

1.3 The Importance of the Thesis

This study proposes for the first time an intelligent tutor system for automation of the Persian calligraphy learning process. The system incorporates image processing and machine learning technologies. In the assessment phase by comparing the filtered and normalized inserted character by its sample character the developed system. The PeNCIL system provides facilities for Persian Nastaliq calligraphy learners to reduce errors that are inherent in traditional education methods, makes the process more efficient and allows people to take advantage of learning possibilities whenever and wherever they choose.

1.4 Limitation of the Study

1. This study is limited by the date that starts from February 2015 until June 2016 depending on the application design, development and testing in this study. This thesis was based on the Persian isolated characters Nastaliq calligraphy.
2. This study explains the features of intelligent tutors and its advantages.
3. This study discusses image processing and machine learning techniques
4. This application is limited to Persian Isolated Characters calligraphy.

1.5 Overview of the Thesis

This section expresses the thesis sections briefly:

Chapter One: It explains the main aim and proposes of the thesis and the importance of the current study.

Chapter Two: This chapter demonstrates previous studies which are conclude in the intelligent calligraphy tutors, image processing and machine learning.

Chapter Three: This chapter discusses Persian Calligraphy, Image Processing, Machine Learning and Image Feature extraction.

Chapter Four: This chapter provides current system design and implementation

Chapter Five: This chapter provides current study experimental results.

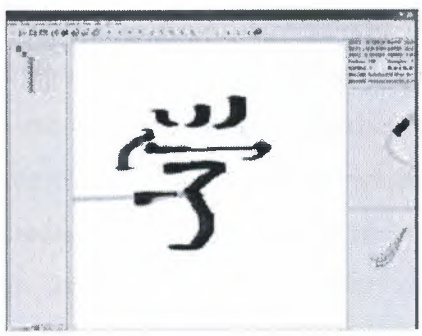
Chapter Six: This chapter displays how the current system works.

Chapter Seven: The last chapter of this thesis includes conclusion and suggestions for future study.

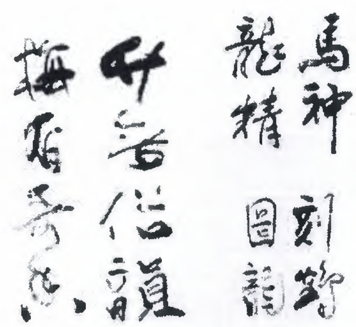
CHAPTER 2
RELATED RESEARCH

2.1 Intelligent Calligraphy Tutors

Xu, Jiang, Lau and Pan (2007) designed and implemented the Chinese calligraphy system by collaboration of Artificial Intelligence techniques and machine learning algorithms as seen in Figure 2.1. Existent calligraphy strokes extrication was carried out by two-phase semi-automatic mechanism. Firstly, algorithmic techniques combination, conducted to do extraction. Secondly, the developed system provide intelligent interface in order to input Chinese characters. Supervised learning technique was used in this study. They claimed that a parametric representation of Chinese calligraphy is the main reason behind this chosen of this technique. They claimed that the developed system would be useful to "aesthetic calligraphy decomposition, visual quality evaluation, automatic calligraphy generation, and calligraphy writing tutoring and correction".



(a) A screenshot of intelligent calligraphy writing tutoring system



(b) Some calligraphy created interactively using the system

Figure 2.1: Intelligent calligraphy writing tutoring system

Bezine and Alimi (2013) developed the Arabic handwriting automated system. The developed educational system presented facilities to learning and detect the errors automatically. They described that how student' illegible handwriting is caused teachers' wasting times. In the traditional methods students need to spend more time to improve their hand writing which is not advantage of this method. The developed system is capable to check mistakes of the inserted handwriting and provide a feedback immediately to the students in order to correct mistakes. This system was developed by applying the Attributed Relation Graph (ARG) method by using graph matching algorithm to detect difference between inserted pattern and template in order to find out optimal graph matching.

Taele and Hommond (2015) presented intelligent sketch education BopoNoto application based on The Zhuyin phonetic in order to provide platform for Zhuyin language students form. The developed system included two main parts: Firstly, the suitable interface which provides facilities to student for practicing the symbols, secondly, the recognition and assessing the inserted sketch system facilities in both technical and visual correctness as seen in Figure 2.2. Recognition system which includes four main techniques: Symbol segmentation, tone symbol recognition, symbol classification by using two different metrics (Hausdorff distance-base and stroke points coverage)assessment classification by using stroke count, stroke direction and stroke order correctness tests were performed to design the developed system intelligent sketching interface. They carried out symbol segmentation and They claimed that BopoNoto application successfully improved students zhuyin phonetic symbols understanding and writing.

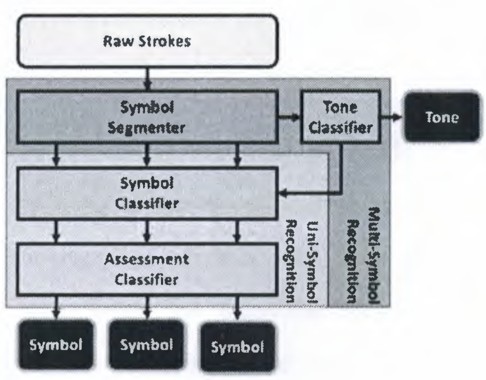


Figure 2.2: BopoNoto's recognition system for classifying both individual and multiple zhuyin phonetic symbols

Albu, Hagiescu, Puica and Vladutu (2015) developed English language intelligent tutoring system for pupils. The developed system included two main section. First section the developed system evaluated students' handwritten symbols quality automatically which acquired by using a digital-pen. A binary image representation technique which called the Dynamic Time Warping (DTW) was carried out to compare the template with inserted one. They claimed that the teaching handwriting improves pupils' personality, communication skills and abilities of basic coordination. Thus, emotion recognition which consist of speech evaluation and face emotion evaluation were conducted in the study. Positive, negative and neutral are three basic emotional expressions was recognized by the developed system which mentioned pupils' interest level for the lesson.

Han, Chou and Wu (2008) study targets were primary schools students. Fuzzy inference and image processing techniques were performed in the study to evaluate students' Chinese calligraphy and automatically scoring. Accurate location and appropriate widths and size were used as assessing statistical features. In addition the developed system presented some instructions to users in order to improve their handwriting.

2.2 Image Processing

Sano, Ukida and Yamamoto (2010) proposed a visualization and measurements methods to guide and recognize the pupils Japanese handwriting skills by using image processing techniques. The learners used the brush pen and paper then used a camera to take the photo of the strokes written and inputted in to the developed system. The developed system measured and visualized the speed of brush pen, line width and path of the pen. They found that the Japanese hand writing learners by using characters feature parameters and profiles which displayed on screen helped them to recognize the right way of Japanese handwriting.

Liang, Bao and Liu (2010) presented the new methods in order to classifier Chinese handwriting and Chinese Painting. They mentioned that in the previous study there were various methods to do this classification; however, they presented a novel method in the paper. In the study, after completed the classical Morphological method they used the specific novel threshold formula to obtain the best threshold values by using below formula.

$$threshold1 = 0.3^{threshold2} = 0.4^{threshold3} = 0.9^{0.75} <=^{comlexity} <= 1.55, 0.35 <=^{occupyratio} <= 0.85 \quad (2.1)$$

The study experimental results showed that however the novel algorithm is simple but it has satisfactory efficiency results in the study.

Tsai, Lin and Chang (2003) applied a novel authentication methods for Chinese calligraphy by using image processing techniques. The aim of the study is provided the digitalized Chinese calligraphy in order to preserves the culture property from malicious manipulations. Their novel method detected location of the malicious manipulations and determined accurately. The authentication message which provided by results of image processing techniques such as compressed feature images was used as a verification way by users. In addition, they claimed that the proposed scheme was robust and efficient method for digital calligraphic images authentication.

2.3 Machine Learning

Qi (2008) presented an unlimited size assessment method for Chinese character writing. Author used the SVM machine learning method to identify the stroke and stroke order of written character. Structure normativeness assessment method was described in the paper which exposed that the handwriting is good or bad by identifying the measure of calligraphy. The nature Chinese characters' assessment features were used as structural assessment features in order to assess characters such as: symmetry, uniformity and compaction. Support Vector Machine method was adapted to identify the written character and determine the difference between the original and inserted one and at last grade the level of the writing. The experiment showed that the current method well reflect the writing quality.

Azmi, Omer, Nasrudin, Muda and Abdullah (2011) applied a novel technique to digital Jawi Paleography. They said that classification of ancient manuscripts revealed paleographers useful information. Features from characters, tangent value and features known as Grey-Level Co-occurrence Matrix which are basic structural of the Jawi (Arabic writing in Malay language) characters were three main parts of the novel techniques which called triangle method. Seven Unsupervised Machine Learning algorithms and one (Support Vector Machine (SVM)) Supervised Machine Learning conducted in the study to compare are Diwani, Farisi, Nasakh, riq 'ah and Thuluth calligraphy. 12 features extracted from the formed triangles. 711 numbers of models used for the Unsupervised Machine and 1019 for the Supervised Machine. The Supervised Machine Learning testing and training results which shown in Table 2.1 revealed that the Diwanigive class results is very significant however other classes results need more improvements.

Table 2.1: Classification result using SVM

Image	Number of Image	Result
Diwani	29	69%
Farsi	25	12.9032%
Riqah	27	14.2857%
Thuluth	27	25.9259%

CHAPTER 3
THEORETICAL FRAMEWORK

3.1 Persian Calligraphy

Persian alphabet contains 32 characters which is written from right to left on horizontal line. Persian Nastaliq is one of the best (Mohsen, 2009) and beautiful (Sattar et al., 2008) Persian calligraphies version.

Such as other types of languages, learning isolated characters is the first step of the Persian language witting skill. In the traditional method of the Nastaliq calligraphy teaching, learners need to learn how to use ink and specific calligraphy pen which called "Galam" on coating paper in advance then they must practice the sample isolated character which called "Mashq" several times which were assigned by teachers before. Learners' Mashqs were evaluated by teacher one by one carefully with specific roles and feature structures which related to each isolated character stroke proportion specifically that need more time and efforts to asses and give feedbacks on learners' Mashqs (Hamidi, & Rashvand, 2015; Rouhfar, 2014). The lack of specialists in this area as a teacher is another insufficiencies in the traditional Nastaliq teaching methods which leads learners to discourage and leave courses (Rouhfar, 2014). In this thesis the Persian Nastaliq Calligraphy Intelligent Learning (PeNCIL) tool was presented as an intelligent standard tutorial tool to overcome the traditional methods insufficiencies. The developed system provides facilities for learners to practice individually as well as the PenCIL is capable to evaluate and provide suitable feedbacks automatically on learners' works.

3.2 Image Enhancement

Enhancement of the image sharpens the image features which can be used further for specific application includes image analysis, feature detection like edges, boundaries and so on. (Patel and Goswami, 2014) explained that due to enhancement of the features, the dynamic range of the features increases without changing the inherent content of the image. The main aim of the image enhancement is increase the quality of the original image and convert to the high quality image (Gonzalez and Woods, 2002).

The frequency and the spatial domain methods are two main category of image enhancement. Image processing based on the frequency approach is carried out the transformation of the image based on the Fourier transformation.

3.3 Image Processing

In our discussion of Image segmentation science, we should make it clear that it is an area that covers different fields such as character, face, ear, finger recognition etc. This rapidly-developing technology needs to utilize image processing techniques or diverse kinds of solutions based on computer picture analysis (Hu & Ji, 2009). In order to enhance raw images which were received by cameras/sensors, satellites, or normal pictures in normal daily life, the Image Processing techniques were applied to use these images for various applications. Over the last several decades, various Image Processing techniques have been developed. Various applications which use the image processing techniques are listed in below:

Medical Imaging (e.g. MRI), Remote Sensing, Forensic Studies, Non-destructive Evaluation, Textiles, Military, Film industry (e.g. Special effects), Arts, Document processing, Printing Industry. Image Processing steps are categorized in image scanning, storing, enhancing and understanding.

3.3.1 Image processing methods

Image processing consists of two main methods based on type of the processing: Analog Image Processing and digital Image Processing.

- ***Analog Image Processing***: Television image is common example of analog image processing techniques which refers to electrical means of the image.
- ***Digital Image Processing***: in order to process these types of images digital computers are used. In this case, after converting image to digital format via scanner or digitizer, then image is ready to process. In order to obtain a desired processing result, the objects are defined as numerical representations to a series of operations. Generally, The digital image processing term refers to a two-dimensional picture processing by a digital computer (Gonzalez and Woods, 1992; Aussem, Murtagh and Sarazin, 1995). It means that, the digital image processing simply implies any two-dimensional data by digital processing. A digital image contains of a real numbers array which represented by a finite bits number.

3.3.2 Digital image processing algorithms

The fundamental operations of the digital image processing will described in this section. Generally can be divided these operations into four main categories namely: operations based on the image histogram, operations based on simple mathematics, operations based on convolution and operations based on mathematical morphology.

3.3.2.1 Image processing operations based on the image histogram

According to Maini and Aggarwal (2010) study, the manipulation of an image or region histogram is an important point of these operations. Contrast stretching and Equalizations are two common operations based on the image histogram (Figure 3.1). The image histogram represents the relative frequency of various gray levels rate in the images.

- Contrast stretching: Normally, the brightness value of any scanned image don not use the available dynamic range totally. Thus in order to overcome this insufficiencies the histogram stretching is conducted over the available dynamic range.

If *brightness 0 to brightness 2^B–1* ***then***
brightness = 0% Min value
Else
brightness = 100% Max value
End If

Figure 3.1: Image processing operations based on the image histogram

The appropriate transformation is given by:

$$b[m,n] = (2^B - 1) \cdot \frac{a[m,n] - \text{minimum}}{\text{maximum} - \text{minimum}} \tag{3.1}$$

Generally, Contrast stretching increases the difference between an image maximum intensity value and the minimum intensity value. As a result all the intensity values are extend in two range (Min and Max). A one-to-one relationship are exists in intensity values between the source image and the target image. It means that, the original image is restorable from the contrast-stretched image.

- ***Equalizations:*** In order to compare two images two or more images on specific basis (e.g. texture) standard histogram normalization is applied in advance. This operation is very useful for the images which have taken under different circumstances. The histogram equalization is the common histogram technique of normalization. As seen in Figure 3.2 the histogram equalization corresponds of a distribution of brightness where probably all values are equal. According to Senthilkumaran and Thimmiaraja (2014) study this technique focus on the contrast of the original image and it may be enhanced in the image contrast. In order to reduce or redistribute the intensity of distributions, histogram equalization can be applied to the original image or on part of the image which selected before. The histogram equalization is based on principle in being a point process, the image does not receive new intensities. Therefore, new values of the equalization will map to the existing values. The value of the pixels count in advance initially start with a zero valued array. Afterwards, the histogram values were stored in a separate array. Element 1, this array holds the values of the 1 and 0, the histogram elements. (Luong, 2006)

  Unfortunately, for an arbitrary image, one can only approximate this result. For a "suitable" function $f(\bullet)$ the relation between the input probability density function, the output probability density function, and the function $f(\bullet)$ is given by:

$$p_b(b)db = p_a(a)da \Rightarrow df = \frac{p_a(a)da}{p_b\,(b)} \qquad (3.2)$$

It should be mentioned that once histogram equalization is performed, there is no way of getting back the original image.

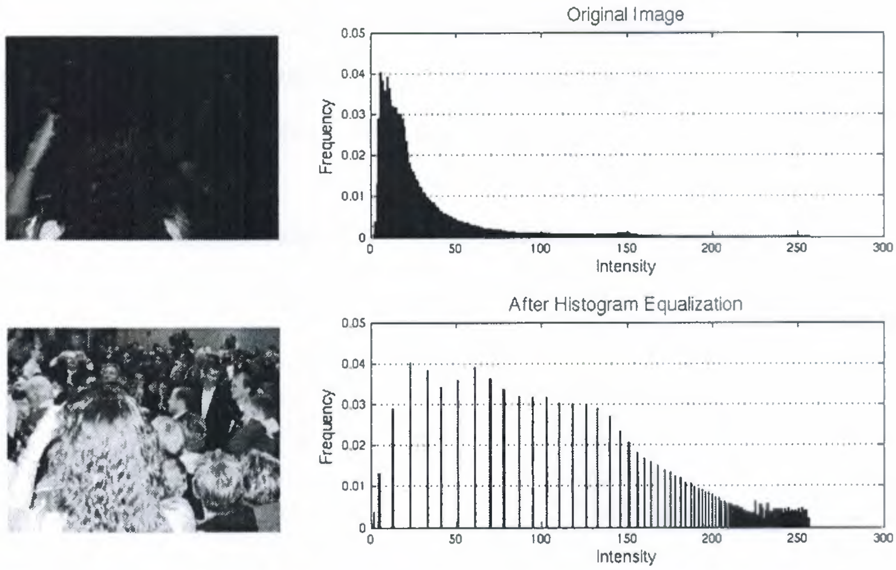


Figure 3.2: Histogram of equalized image

3.3.2.2 Image processing based on mathematics-based operations

The mathematics-based operations were distinguished difference between binary and ordinary arithmetic. There are two "0" and "1" brightness values in the binary case, but 2^B brightness values or levels are exists in the ordinary case. In this regards, the image processing can generate easily in many more levels. Thus, in order to avoid arithmetic overflow problems, most of the software systems provide representations bits (16 or 32 bit) for pixel brightnesses. Binary and Arithmetic-based are two common operations.

- ***Binary operations:*** Binary (Boolean) based operations arithmetic describes point operations and a variety of efficient implementations. These operations include simple look-up tables. Operations are applied on a pixel-by-pixel basis. The binary operations and standard notation are listed in below:

$$
\begin{array}{ll}
NOT & c = \bar{a} \\
OR & c = a + b \\
AND & c = a \odot b \\
XOR & c = a \oplus b \\
SUB & c = a \setminus b = a - b = a \odot \bar{b}
\end{array}
\tag{3.3}
$$

- ***Arithmetic-based operations:*** The image processing basis is based on ordinary mathematics with the gray-value point operations and ordinary mathematics. Table 3.1 explains the Arithmetic-based operations.

Table 3.1: Arithmetic-based operations

Operation	Definition	Preferred data type
ADD	$c = a + b$	integer
SUB	$c = a - b$	integer
MUL	$c = a \odot b$	Integer or floating point
DIV	$c = a / b$	floating point
LOG	$c = log(a)$	floating point
EXP	$c = exp(a)$	floating point
SQRT	$c = sqrt(a)$	floating point
TRIQ	$c = sin/cos/tan(a)$	floating point
INVERT	$c = (2^B - 1) - a$	integer

3.3.2.3 Image processing based on Convolution-based operations

According to Costa (2011) study, digital images are widely used a convolution based mathematical operation with filters. This operation is applied when a window of finite size and shape is scanned across the image. The output pixel value is the weighted sum of the input pixels within the window where the weights are the values of the filter assigned to every pixel of the window itself. The window with its weights is called the convolution kernel.

3.3.2.4 Image processing based on Smoothing operations:

In order to prepare inserted image for the next steps (e.g. segmentation), the smoothing algorithms are applied. These algorithms reduce noise of inserted images. In this regards two main algorithms namely linear and non-linear are conducted. Linear algorithms are applicable to analysis in the Fourier domain and non-linear algorithms are not. Furthermore, rectangular support and circular support for there are two implementation methods (Costa, 2004).

3.3.2.5 Image processing based on Morphological-based operations:

Mendozaa, Melinb and Liceaa (2009) explained that a wide set of image processing operations which can process the images based on their shapes called morphology. Theses operation sets create an output image through applying a structuring element to an input image. As a result input and output image have same size. Each pixels' value in the output image is outcome of the corresponding pixel in the input image with its neighbors comparing results. Modifying the neighborhood pixel size and shape create a morphological operations which are conductible to specific shapes. Image enhancement, shape analysis, segmentation, image analysis, computer vision problems, texture analysis, etc. are just some of the major applications of this technique (Chen and Haralick, 1995; Demin et al., 1995). Morphological method consists of various operations such as: Dilation, Erosion, Opening, Closing, Hit and Miss Transform, Thinning, Thickening and Skeletonization/Medial Axis Transform. Dilation and Erosion are two main basic morphological operations. Dilation fills the pixels to the boundaries of objects in an input image, while erosion operation removes the pixels on input image object boundaries. The shape and size of the structuring element have a direct effect on the number of added or removed pixels during the image process.

- ***Dilation and Erosion***

As we mentioned before the dilation and erosion are two main techniques of morphological which applied on binary image. The values of pixel are adjusted by the black equal 0 and white equal 1. The simple counting and storing these pixels value work as per predetermined rules depending on the characteristics of the neighboring pixels. The original image is used to analyze the value the each pixel in the image and in practice, the original image is replaced after every few lines of binary codes are assigned back to the original image. The pixel values are used to evaluate the rest of the image and not the entire image is used up for the process (Russ, 2011).

Haralick et al. (1987) explained that erosion use the subtraction vector of the set elements to firm the dual to dilation). Erosion by switching OFF pixels in a feature that were originally ON, this simplest approach is to select a threshold value and switching off pixels which are different from this threshold value and thereby isolating the values or pixels which conform to the object being studied. The advantage is that when there is bright point in a dark region, simple enumeration can isolate this pixel, while this technique uses the average brightness of the region to determine whether a pixel is to remain ON or is to be switched OFF.

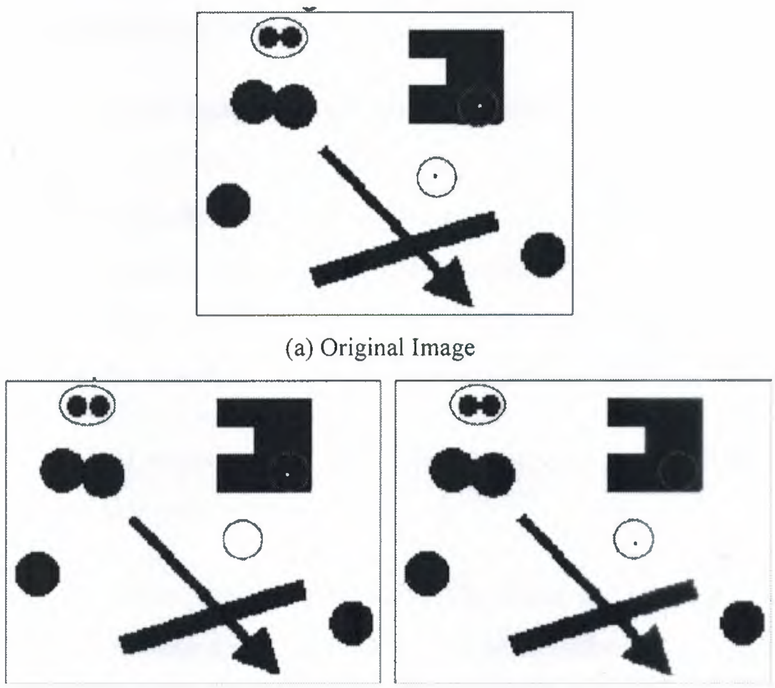
The results of this technique is removing any pixel that is in contrast and is touching the background. Furthermore, applying Erosion can remove a layer of pixels around the desired object and it can have side effect too.

According to Haralick et al. (1987) explanation the operate to add more pixel instead of removing pixels, is called dilation in the morphology techniques. The dilation technique add background pixels which are on the study object or beside another pixel in the foreground. Opposite of the Erosion this operation adds a layer of pixels around the object of the study. These added layer cause appear clearer object. Plating and etching or growing and shrinking are other names of the Erosion and Dilation process based on their function behavioral which are add or subtract the foreground.

- ***Opening and Closing***

When an erosion process is followed by a dilation, it is called an opening since this combination of actions has the potential to create gaps between features which are touching (just touching), as indicated in the Figure 3.3 isolating pixel noise and removing fine lines in binary images is usually performed through the process of opening.

When the same operations are performed in the opposite order it can produce a different result. This is called closing, since it can close openings in the features of the images. Adjustment of the erosion and dilation can be performed through a number of parameters, especially pertaining to the adding or removing neighboring pixels and how many times the operations have to be performed. Mostly, the erosion and dilation sequences are kept in equal numbers, but this again is dependent on the nature and composition of the image. Opening of separate touching features can sometimes be done using multiple erosion operations till the point where the features have been separated and yet stopping short of completely erasing the pixels. When the separation is completed, the dilation process can be used to grow back the object to the original size and this has to be stopped before they eventually merge to negate the impact of the erosion process and this is accomplished by the rule which stops the closing process from merging a pixel that has been separated so that the separation is maintained. Additional logic is needed at each stage of the operation, since feature identification of the pixels must be performed at each stage of dilation. An additional rule also has to prevent the increase in the size of the object beyond the original size. Russ (2011) explained that if there are multiple objects in the image and with different sizes, this technique can make some features disappear for the sake of identifying some other features.



(a) Original Image

(b) Opening (c) Closing

Figure 3.3: Closing and opening operations.

Generally, Dilation expands objects while erosion shrinks them the other hand opening operations can delete islands and cut narrow isthmuses while the closing operations fills the thin gulf and the small holes (Demin et al., 1995).

3.4. Image Feature Extraction

Semantic understanding is the first step for a computer program. In this regards, extracting low-level visual features of images and consider the kind of extracted features which will be play main roles in image processing (Li and Shawe-Taylor, 2005; Leung and Chen, 2002). An attribute of the image or a primitive characteristic of the image is called an image feature. Literature review revealed that retrieval systems and image annotation have been used visual features such as color, texture and shape, etc. (Shih et al., 2001; Stanchev, Green and Dimitrov, 2003; Yang et al., 2008; Islam, Zhang and Lu, 2008; Arivazhagan and Ganesan, 2003). Generally, global, region-based and block-based are three methods which represent the feature.

The images unique features such as location, brightness, shape and size can provide the unique attribute for each images. Within each of these features, different measurements can be made of a pinpointed nature through a number of different means (Russ, 2011).

A set of feature that is good contains distinguishable elements can help in identifying and detecting features clearly. Objects in the same class categorized in the same feature group to do robust understanding. A small set which they values represent the different features of the image must be extract new features (Mesanovic et al., 2012). Reza et al. (2011) demonstrated that feature extraction is crucial in any kind of image classification performing and affect the results of the classification significantly.

Mesanovic et al. (2012) explains that the most crucial step is the identification of the meaningful features in the image, by below reasons:

1. From the initial set, all the possible subsets must be found and this is a laborious task;
2. Some of the discriminations apply to at least some of the subsets;
3. Variations between intra and inter class features within the image is narrow;
4. The inclusion of more and more features can reduce the utility of the model.

3.4.1 Color Feature

One of the important images' features which are defined subject to a particular model or color is named color feature. Specifying the color spaces such as RGB, LUV, HSV and HMMD which were mentioned by previous studies Stanchev, Green and Dimitrov (2003) is the first step for color feature extracting. Color histogram(Jain and Vailaya, 1996) color coherence vector (CCV) (Pass and Zabith, 1996), color correlogram (Huang et al. 1997), color moments (CM) (Flickner, 1995), and etc. are important color features which have been proposed in previous studies(Liu et al., 2007). CM is the simplest and very effective features. A summary of different color methods with their advantages and weakness are defined in Table 3.2 (Zhang, Islam and Lu, 2012).

Table 3.2: Color methods contrast (Zhang, Islam and Lu, 2012)

Color Method	Pros.	Cons.
Histogram	Simple to compute, intuitive	High dimension, no spatial info, sensitive to noise
CM	Compact, robust	Not enough to describe all colors, no spatial info
CCV	Spatial info	High dimension, high computational cost
Correlogram	Spatial info	Very high computational cost, sensitive to noise, rotation and scale
DCD	Compact, robust, perceptual meaning	Need post-processing for spatial info
CSD	Spatial info	Sensitive to noise, rotation and scale
SCD	Compact on need, scalability	No spatial info, less accurate if compact

3.4.2 Texture Feature

Texture is useful characterization for images in wide range (Tsai, 2007). Literature review showed that a texture for recognition and interpretation has been used by human visual systems too. It should be mentioned that texture feature measured a group of pixel while usually color feature measure a pixel property. A large number of techniques based on the texture feature domain have been proposed to extract texture features. In this regards, firstly in original image domain the pixel statistics computing or the local pixel structures finding, then in the next step image transforms into frequency domain afterwards calculates feature from the transformed image (Tian, 2013). Table 3.3 summarizes the Spatial and Spectral texture features pros. and cons.

Table 3.3: Texture feature contrast (Tian, 2013)

Texture method	Pros.	Cons.
Spatial texture	Meaningful, easy to understand, can be extracted from any shape without losing info.	Sensitive to noise and distortions
Spectral texture	Robust, need less computation	No semantic meaning, need square image regions with sufficient size

3.4.2 Shape Feature

Shape is an important cue for human to recognize and identify the objects through simple geometrical forms (e.g. different directions straight lines). Shape feature extraction techniques classified into two main methods (Zhang and Lu, 2004): Region based and contour based. Firstly, the boundaries of shape are calculated then features extract from the entire region (Mezaris, Kompatsiaris and Strintzis, 2003). Shape feature includes two main cases: absolute spatial location of regions (Yang, Dong and Fotouhi, 2005) and relative locations of regions (Mezaris, Kompatsiaris and Strintzis, 2003). Figure 2 explains an example of a 2D string representation.

d		
	b	c
a	a	

a

Object Symbols	a,b,c,d

b

Relationship Symbols	**Meaning**
$<$	Left-right or below-above
$=$	At the same spatial location as

c

$(a = d < a = b < c, a = a < b = c < d)$ d

Figure 3.4: Shape feature extraction

Yang, Kpalma and Ronsin (2008) presented a diagram of the shape-based feature extraction existing approaches. Figure 3.5 explains the hierarchy of the shape feature extraction classification approaches.

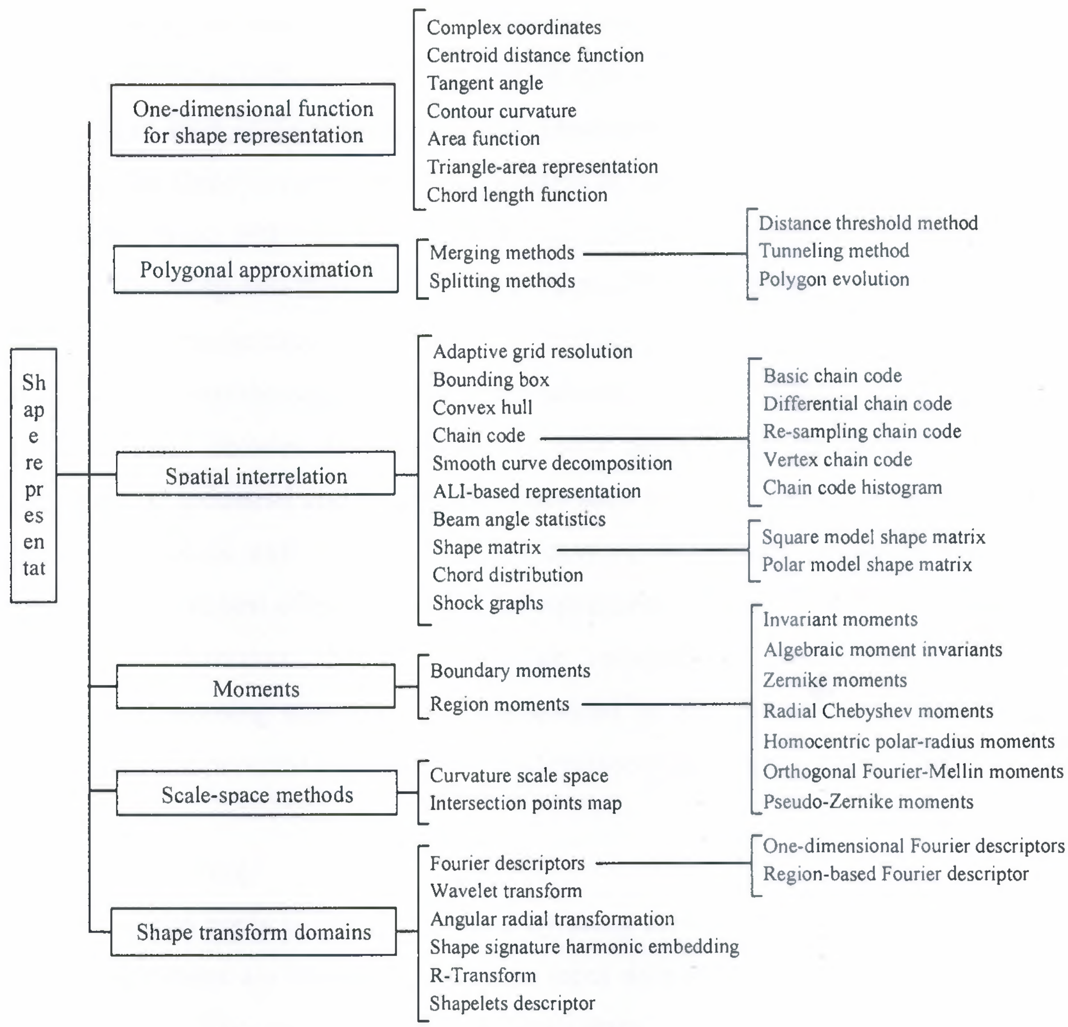


Figure 3.5: Hierarchy of the shape feature extraction classification approaches (Yang, Kpalma and Ronsin, 2008)

3.5. Classification and Machine learning

Create various class by identifying and collecting similar set of the data in one class is called classification (Thomas and Kumar, 2014). Classification is the significant component of image analysis. These classification techniques are usable in various areas which machine learnings techniques are applied. A type of the artificial intelligence (AI) which is capable to learn without being explicitly programmed is called Machine learning. In this case the developed system provides facilities to applications to teach themselves and respond when faced with new data. Data mining and machine learning process are so similar such as searching data through look for patterns. The data mining case extracts the data for human comprehension while machine learning uses human-free techniques to search and detect patterns through adjust program actions (Schapire, 2003).

Over the last several decades machine learning base applications are used in various aspects of daily life, industrial and theoretical issues such as: Fraud detection; Web search results; Real-time ads on web pages and mobile devices; Text-based sentiment analysis; Credit scoring and next-best offers; Prediction of equipment failures; New pricing models; Network intrusion detection. Pattern and image recognition; Email spam filtering. Generally machine learning algorithms are categorized in two main group: Supervised, Unsupervized, Semi-supervised learning and Reinforcement learning (Schapire, 2003).

3.5.1 Supervised learning

Supervised algorithms apply to search new data by using the data which has been learned before. These algorithms are trained the previous input data by using labeling in order to obtain desired output. The supervised learning algorithms' inputs are a set of inputs data with the correct corresponding outputs data. These algorithms learn by comparing its real output with correct outputs to find out errors. Classification, prediction and gradient boosting, regression are some methods of supervised learning. These methods are capable to learn patterns and predict the additional unlabeled data values. Generally, supervised learning algorithms are conducted in the applications which predict feature events based on historical data (Rasmussen and Williams, 2005).

3.5.1.1 Artificial Neural Networks (ANN)

Artificial neural networks (ANN) is a learning systems that inspired the biological in nature in the form of complex interconnected neurons. According to Mitchell (1997) study

ANN is compound of simple units that are densely interconnected in sets and producing a single real-valued output while taking many real-valued inputs.

Jensen, Qiu, and Patterson (2001) claimed that ANN no need for normal distribution and ability of simulation of non-linear and complex patterns in an adaptive manner are two main advantages of this supervised learning machine.

- ***Architecture of Neural Networks***

Hidden, input and output are the names of the layers which are the main structure of the ANN (Jain, Mao, and Mohiuddin, 1996).

In networks that are feed-forward, flow of the signal is in the feed-forward direction or from the input to the output units. While the processing of the data can move over many layers of units, there are no feedback connections available. Networks that are incorporating feedback connection are recurrent networks. The properties of the network that are dynamic in nature are very important, which is in contrast to the feed-forward networks. In some instances there is a relaxation process for the activation values of the units so that the evolution of the network to the stable states wherein there are no changes in the activations. In some other uses, the activate values change in the neurons of the output values are high so that the output of the network is constituted by the dynamical behavior. The multispectral reflectance values of the individual pixels combined with their surface roughness, texture, slope, terrain elevation, aspect, etc. The application of the hidden layer neurons helps the non-linear simulation of the neurons in the data that is input (Abraham, 2005).

- ***Artificial Neural Network learning***

In the ANN context, the learning process can be stated to involve the intricacies of studying the updates of the architecture of the network and the weights of the connection so that tasks can be performed efficiently by the network. The connection weights must be learnt by the network from the patterns that are available (Jain, Mao, and Mohiuddin, 1996). When there is iterative updating of the connection weights, the network must automatically learn to improve performance. The ability of ANNs to learn intuitively through experience enables them to be considered exciting. ANNs learn through their own experience by learning the underlying rules, like for example, the input-output

relationships, instead of following a fixed set of rules that are coded by the humans. This is the major advantage of ANNs over traditional systems (Jain, Mao, and Mohiuddin, 1996).
There are three different learning situations in ANNs, namely, reinforcement, supervised and unsupervised learning. In supervised learning, the inputs are provided in the form of an input vector along with the responses that are desired at the output layer, one for each node. The actual and the desired responses are compared and the discrepancies are found through a forward pass. This learning is used to then determine changes in the weights as per the subscribed rule. Since the desired signals in the output is provided by an external source (teacher) and therefore, this is called supervised learning (Abraham, 2005).
ANNs need testing and training like supervised classification to ensure that the information extracted from the ancillary and remotely sensed data to be useful information (Jensen, 2005).
Various classification problems are widely solved using a supervised ANN using the back-propagation technique. Figure 3.6 depicts the structure of the topology of the back-propagation ANN.

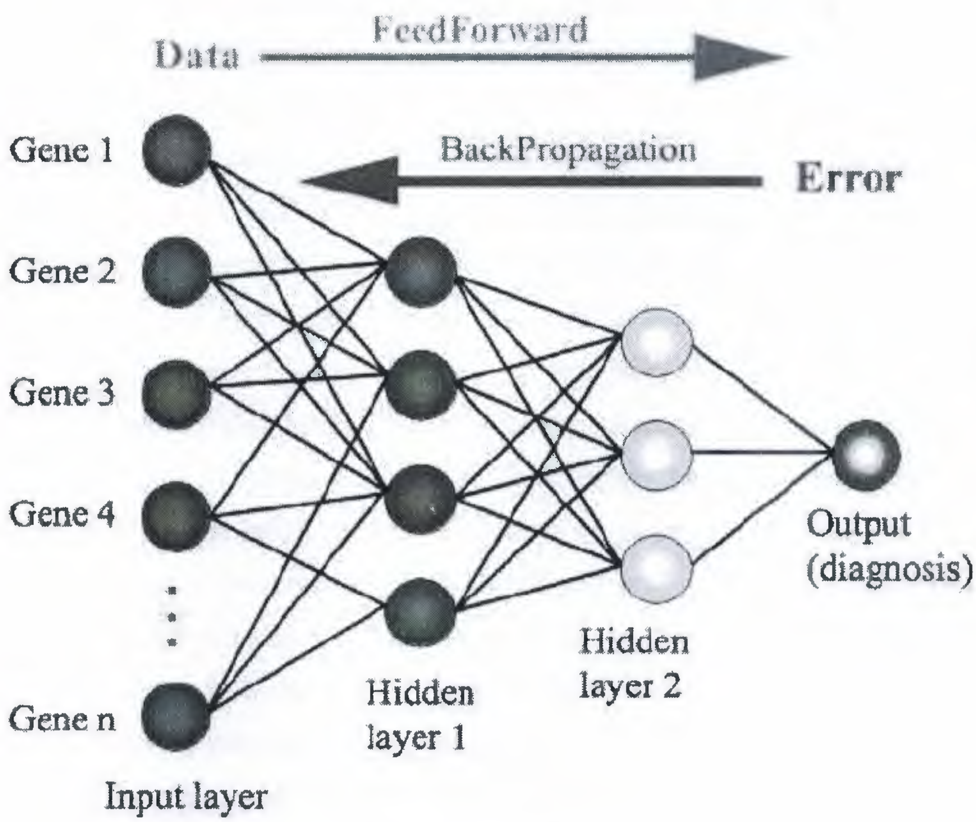


Figure 3.6: Typical ANN main components

Back-propagation ANN comprises of the typical components of hidden, input and output layers. Individual training pixels comprising spectral reflection throughout the spectrum and other descriptors like slope, elevation, etc. are present in the input layer. Each layer comprises of interconnected nodes. This state of connectedness provides for the flow of information in different directions at the same time or in other words, back propagation is allowed. The weight of the connections is stored and then learnt by the ANN. During the classification procedure these weights are used. As the representativeness of the training data increases, the probability that the ANN will better mirror the reality and produce classification that is more accurate. The individual thematic map classes such as forest or water may be represented by the output layer (Jensen, Qiu, and Ji, 2001).

- ***Training***

During training, x, y locations are specified by the analyst in the input image with attributes that are known used as the training sites (Jensen, Qiu, and Ji, 2001). The spectral information per pixel and the surrounding contextual information for the sites of training are accumulated and conveyed to the ANN input layer. The class value or the true target value is relayed to the output layer by assigning the value of 1 to this class at the same time and the rest of the neurons are denoted by the value 0.

Training of the neural network from an image as an example at a certain point of time and place may represent the state of things at the vicinity of the image and may be so for a particular reason and therefore, this cannot be extended through time and space.

When the weights are adjusted through the algorithm of back propagation, the learning is accomplished. Each time training happens, the true target value and the output of the network are compared. The difference between these values is considered as an error value and the feedback of the same is passed to the previous layers for the updating of the connection weights. The adjustment that is made to the network is proportional to the level of the error. Further improvements in the network will not be possible when the root mean square (RMS) error diminishes after a number of iterations of such feedback happen (Jensen, Qiu, and Ji, 1999) and at this stage it can be considered that the training process is accomplished and the network has achieved convergence. The inherent rules are stored in the network as example weights and they are used in the testing phase.

- *Testing (Classification)*

During this stage, the textual and/or the spectral characteristics of the scene in the form of individual pixels are passed on as input neurons irrespective of whether they originate from the rural or urban geography. The weights stored in the network are compared with the input neurons to produce an output value for the output layers. The fuzzy membership grade between 0 and 1 is assigned to every pixel in the output neuron representing the class of the neuron. The fuzzy classification map of the entire study is obtained through the value of every output neuron of every class. When these maps are defuzzified, a hard classification map is obtained through using a local maximum function by unique classification of each pixel through the fuzzy membership highness (Jensen, Qiu, and Ji, 2001).

3.5.1.2 Support Vector Machine (SVM)

SVMs are one of the powerful machine learning techniques for classification (Ganesan, Subashini, and Jayalakshmi, 2014). More than estimation and linear operation inversion, the SVMs are capable to prove a novel approach to pattern recognition problems and can establish connections with learning theories from statistics very clearly (Burges, 1998). SVMs have provided a number of successful applications in a various areas such as: pattern recognition, supervised classification techniques, biometrics, image analysis and bioinformatics (Ma and Guo, 2014).

Analysis, statistics, machine learning and optimization are some disciplines which are carried out the SVM theoretical and algorithmic analysis (Cristianini and Shawe-Taylor, 2005). Binary classification is base approach of SVMs also it can be extended to include scenarios which are multi class. According to Mathur and Foody (2008) study SVM achieved through the dividing the problem (multiclass) into sequential analyses which are binary and the same can be solved using SVM (binary) using the one-against-all or the one-against-one approach.

Figure 3.7 shows the Hyperplanes separation. The objective of SVM is to find an approach to find the most optimal plane or the plane which provisions the largest possible distance margin of the two classes' nearest points. This is called the functional margin. In general, this approach ensures that when the margin is larger, the classifier returns a lower generalization error (Thome, 2012).

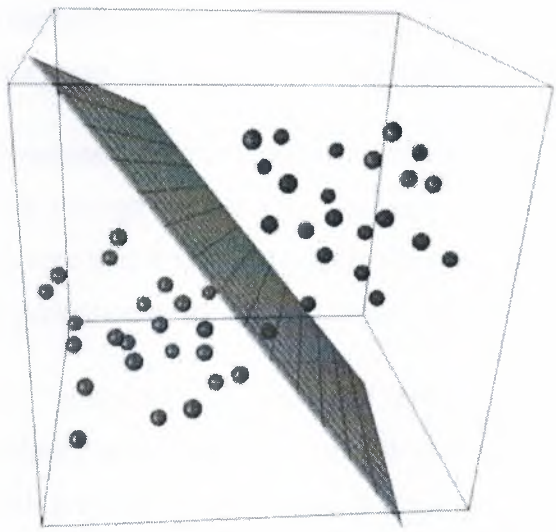


Figure 3.7: 3D view of Hyperplanes separation

The maximum-margin hyperplane is found through the algorithm of this approach into the feature space that is transformed. There may be a non-linear transformation and/or high dimension may characterize the transformed space. The hyperplane is drawn by the classifier in the feature space which is a curve that is non-linear separation in the space that was original. Gausian radial basis function is the ideal kernel in this SVM. Well regularized maximum margin classifiers can be infinite dimension and therefore, the results are intact and effective. Thome (2012) demonstrated some of the common kernels are as below:

$$Polynomial\ (homogeneous)\colon K(x_i, x) = (x_i * x)^d \tag{3.4}$$

$$Radial\ Basis\ Function\colon K(x_i, x) = \exp(-\lambda \|x_i, x\|^2)\ ; \lambda > 0 \tag{3.5}$$

$$Gaussian\ Radial\ basis\ function\colon K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\alpha^2}\right) \tag{3.6}$$

The RBF kernel can be a reasonable choice by default, often as it can map non-linearly the samples into a space of higher dimension and hence is different from the linear kernel in being able to handle relations between attributes and class labels that are non-linear. (Hsu, Chang, and Lin, 2010) state that the linear kernel is to be taken as a variant of the RBF kernel since penalty parameter C in the RBF and linear models have similar performance with the added parameters (C, γ). Additionally, the RBF and the sigmoid kernels behave similarly under certain conditions (Hsu, Chang, and Lin, 2010). The polynomial is another Kernel functions includes greater number of hyperparameters than the RBF model.

C and γ are the basic parameters for the RBF and should be initial at the beginning. Good C and γ were identified through testing data and can be predicted accurately by the classifier. It needs to be note that it may not always ideal to try to make the training very accurate or the accurate prediction of the training data by known classifiers.

It is recommended that a "grid search" is performed through cross validation on C and γ. Different (C, γ) pair values can be experimented with and the specific set with the cross validation accuracy which is highest can be found. It was observed that the combination of C and γ in exponentially growing sequences produced the best results for good parameter identification.

3.5.2 Unsupervised learning

Drawing inferences from datasets is unsupervised algorithms strategies. Unsupervised algorithms are applied when there are no historical labels data. The algorithm must discover what is being shown. Exploring the data and find suitable relative structures are the goal of the algorithms. Unsupervised learning algorithms are work well on transactional data. Self-organizing maps, k-means clustering, nearest-neighbor mapping and singular value decomposition are popular unsupervised learning techniques. Moreover, these algorithms are used for text segmentation topics, items recommendation and data identifications (Hofmann, 2001).

3.5.3 Semi-supervised learning

According to Zhu and Goldberg (2009) study Semi-supervised learning algorithms are conducted to applications which are used for supervised learning too. In this case, semi-supervised learning algorithms use both type of data for training: labeled (small amount) and unlabeled (large amount) data for training. The advantages of this combination are the less expensive unlabeled data and take less effort to obtain results (Chapelle, Scholkopf and Zien, 2009). Prediction, classification and regression are some methods which can be used semi-supervised learning. When the cost of labeling association is too high to complete labeled training process, Semi-supervised learning is useful to apply.

3.5.4 Reinforcement learning

Robotics, navigation and gaming are used the reinforcement learning. In this case, the algorithms find out via trial and error that actions yield the greatest rewards. Reinforcement learning includes three primary components: Agent, environment and actions. The agent means learners or decision maker the environment refers to everything that the agent interacts with, and actions mean what the agent can do. Generally the main goal of the reinforcement learning is learning the best policy (Stone and Veloso, 2000).

CHAPTER 4
STUDY SIMULATION

4.1 The Thesis Framework

Persian calligraphic characters are written on glossy coated paper (Rouhfar, 2014) with cream or tan color. In order to identify and evaluate the students' handwriting, image processing and machine learning methods are conducted in the current thesis. The current study framework consists of six main sections as shown in Figure 4.1 Input data, image processing, feature extraction, character recognition, assessment and feedback.

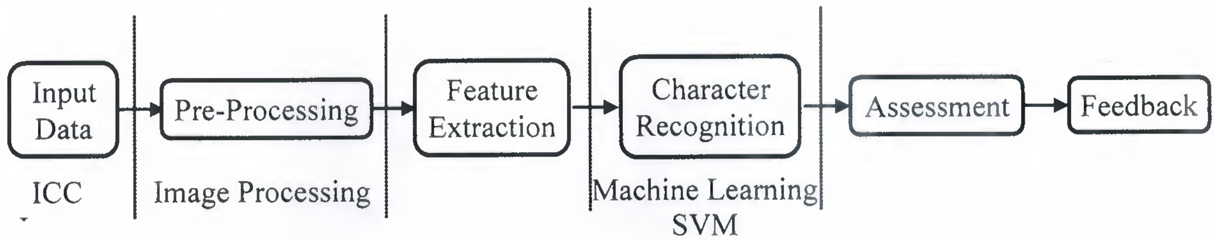


Figure 4.1: Framework of the study

4.2 Input Data

Written Isolated Calligraphy Character (ICC) images were applied as input data in the current study. This recognition system applies the photos of the ICC one at a time. The Offline method (Patel, 2013) was carried out to complete the input process by converting the written paper-based ICCs into images (JPG and BMP formats), either by scanning or using a camera. Figure 4.2 displays some sample of the Persian Isolated Calligraphy Characters.

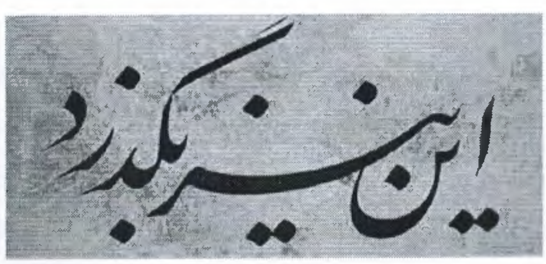


Figure 4.2: Sample Persian Nastaliq calligraphies

4.3 Pre-Processing

In order to delete unnecessary artifacts in the inserted image that may be available due to scanning or any unpredicted problems while writing the characters by ink the pre-processing operation are applied in the current study. The main result of the pre-processing step was the preparation of a suitable image for the feature extraction phase, the steps of which are shown in Figure 4.3. Pre-processing consists of the Grayscale conversion, Binarization, Noise Removal, Cropping and Normalization, Thresholding and Morphology operation (Dilation) steps.

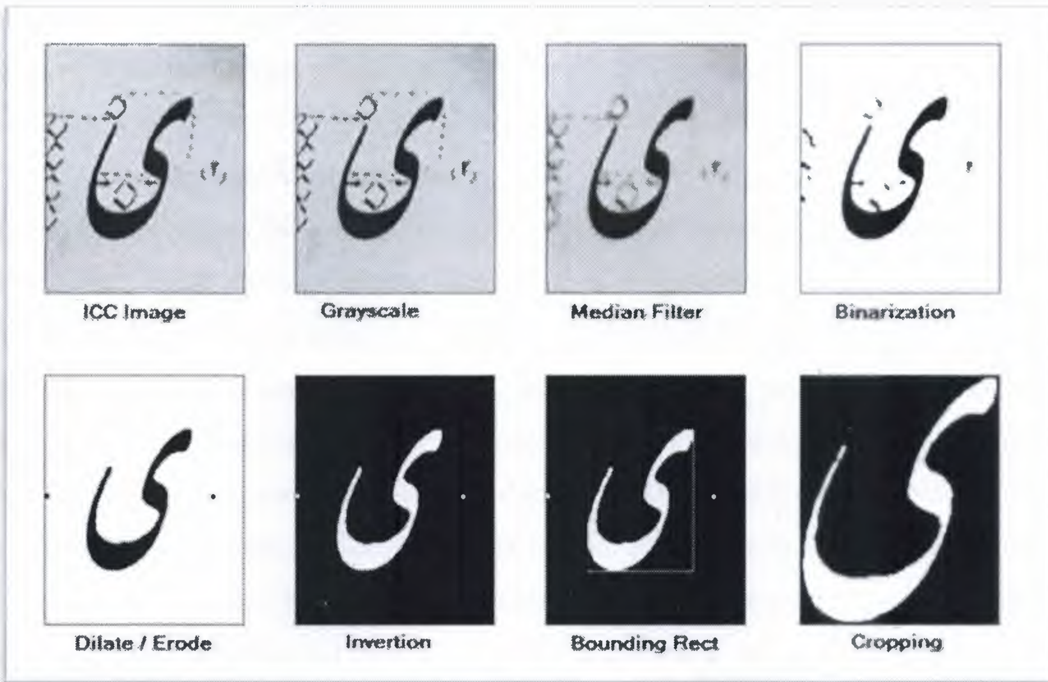


Figure 4.3: Image processing steps

- ***Binarization***

The Grayscale conversion function converts the inserted image from Red-Green-Blue (RGB) format to Grayscale. In order to keep edges due to removing noise of inserted character image Bilateral Filtering is highly effective. The OTSU method as a global Binarizition (Otsu, 1979; Gupta, Jacobson and Garcia, 2007; Chen, Li, Hu and Liang, 2014) operation converts the Grayscale image to a binary format of pixels $x_i \in \{0,1\}$, where zero refers to a black pixel and 1 refers to a white one. The Grayscale filter and

Binarizition operations were used to decrease the size of the input image (Gupta, Jacobson and Garcia, 2007).

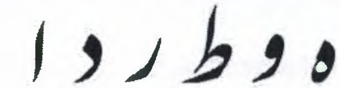
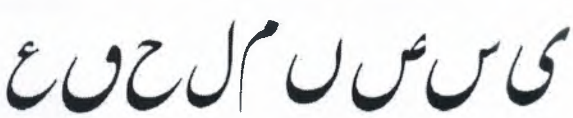
- ***Noise-Removal***

In order to acquire the clear image without unnecessary artifacts Noise-Removal step conducted in the study. In this step, many pixels are lost and image clarity decreases. As we mentioned in chapter 3 Morphology offers a set of image processing functions. Dilation and Erosion are two common operation of Morphology but because of the inserted character images background color (white) Dilation is a preferred choice and conducted in this study. Afterwards, Noise-Removal by using a Bilateral filter (Srinivasan and Ebenezer, 2007) (to smoothing image) removed speckle, salt and paper noise and provided a clear image for the next phase.

- ***Cropping and Normalization***

In order to eliminate the extra portion of the free noise inserted image, cropping process is inevitable. Firstly the top-leftmost black pixel of the inserted character image is identified and stored in a variable then top-rightmost, bottom-leftmost and bottom-rightmost black pixels are identified and stored. The cropping function is used these data to crop and extract only the inserted character. According to Han, Chou and Wu (2008) study these recognized cropped area called Region of Interest (ROI). The last phase before final noise removing of pre-processing called Normalization. According to Persian Isolated character shape and ROI value, Nataliq Calligraphic isolated characters are categorized in three main group as shown in Table 4.1. Normalization was applied to resize the craped image (size of ROI) depends on these three group of characters. Figure 4.4 explains the Pseudocode of Normalization.

Table 4.1: Calligraphic character normalization

Group	Calligraphic character sample	Normalization
A		192×128
B		160×128
C		128×160

```
α = ROI_Height / ROI_Width
Begin
        IF α > 1.40 && α < 1.60 Then
            Image resize to 192 × 128 pixels
              ELSEIF α > 1.15 && α < 1.35
                  Image resize to 160 × 128 pixels
                    ELSEIF α > 0.70 && α < 0.90
                        Image resize to 128 × 160 pixels
                   ENDIF
              ENDIF
        ENDIF
STOP
```

Figure 4.4: Normalization pseudocode

- ***Thresholding***

The matrix of Normalized Gray-Scale character image is subjected to Thresholding. In this regard, after extracting and storing the maximum values of column the maximum value of raw will be extracted and stored too. The gray levels of the pixels of the background and the gray levels of the pixels of the object are totally different in many applications of pattern recognition and image processing. According to Hossain et al. (2011) study to separate the object from the background, Thresholding is the easiest and simplest approach The value of 1 is assigned to the object and 0 assigned to the background pixels. The assignment is based on T a threshold value, which may be the value of a color or an intensity that is desired to be separated (Varshney et al., 2009). Global Thresholding is an intuitive approach, where one threshold value is selected and applied to the entire image being processed and therefore, the Thresholding is stated to be global in nature.

- ***Final Noise-Removal***

Dilation operation is performed in this study as a last phase of image pre-processing. Dilation includes structuring elements namely flat and non-flat. In the current study non-flat structing function outputs are fed to the Dilation function.

4.4 Feature Extraction

As seen in Figure 4.1 the results of previous step (pre-processing) feed to the feature extraction phase. Feature extraction was used to convert the pre-processing phase results' segments to a relevant feature vector by using structural, statistical and Global transformations features (Han, Chou and Wu, 2008). It means that the shape of the ICCs has a direct effect on feature vector numbers. The other word, all ICCs which categorized into 3 groups (Table 4.1) have different feature vector numbers. Figure 4.5 explains the Pseudocode of Feature Extraction and Figure 4.6 shows the framework of each group feature vector creation process.

```
Begin
        IF α > 1.40 && α < 1.60 Then
                Feature vector (Group A)
                ELSEIF α > 1.15 && α < 1.35 || α > 0.70 && α < 0.90
                        Feature vector (Group BC)
                ENDIF
        ENDIF
STOP
Arr[1311] Feature vector (Group A) {

                int Statistical (Zoning, Profile, Crossing)
                int Structural(Histograms)
                int Moment (Hu)

                }
Arr[1179] Feature vector (Group BC) {
                int Statistical (Zoning, Profile, Crossing)
                int Structural(Histograms)
                int Moment (Hu)

                }
```

Figure 4.5: Feature extraction pseudocode

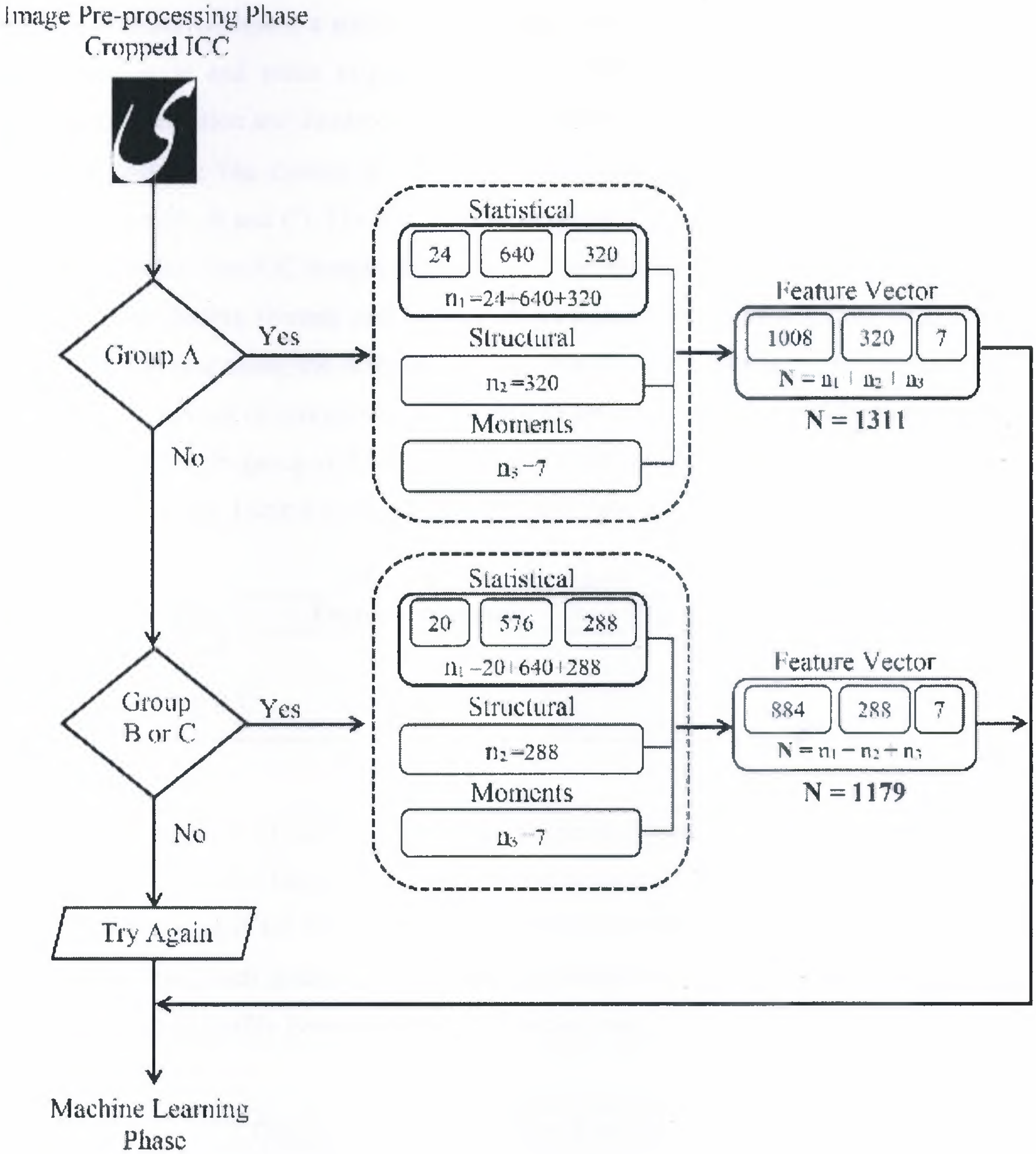


Figure 4.6: Feature extraction frame work

4.4.1 Statistical features

Statistical features represent a statistical distribution of character image based on images variations of style and some extent. The statistical features which normally used for character representation are: Zoning, Profiles and Crossings.

- ***Zoning feature:*** The Zoning feature extraction method is applied to all characters in each groups (A, B and C). The ICC images in group A are divided into 24 equal zones (32x32 pixels). The ICC images in group B, C are divided into 20 equal zones (32x32 pixels). The Zoning (Kumar and Bhatia, 2014) features are extracted from each zone pixels by moving along the diagonals of its respective 32 x 32 pixels. In this case the normalized number of foreground pixels in each cell of the ICC is considered a Zoning feature. In total, in group A 24 and groups B, C 29 features are extracted through the Zoning technique. Table 4.2 explains all the details about both feature vectors.

Table 4.2: Zoning feature

Group	Features Number	Pixels Number
Group A	24	32*32
Group B,C	20	

- ***Profile Feature:*** This feature counts the pixel (Upper, Lower, Left and Right) numbers between The ICC image edge and bounding box. Furthermore the Profile feature is useful for the presentation of the ICC external. According to the ICC categorizing each group has its specific numbers of the Profile features. Table 4.3 explains the Profile Feature numbers for each group.

Table 4.3: Profile feature

Group	Features Number
Group A	640
Group B,C	576

- ***Crossings Feature:*** Counting the ICC images transitions pixels from background to foreground onwards horizontal and vertical is results of the Crossing feature. Each group of the ICC has specific numbers in this feature as seen in Table 4.4.

Table 4.4: Crossings feature

Group	Features Number
Group A	640
Group B,C	576

4.4.2 Structural feature

ICC can be represented by the structural features. This features calculate high tolerance to style variations and distortions. In this case may also encode some information about the objects' structure or may deliver some information as to what kind of components make up that object.

- ***Histograms of Horizontal and Vertical Projection:*** Counting the number of column and row pixels of an ICC image is this feature process. These features work independently of deformation and noise. Each group of the ICC has specific numbers in this feature as seen in Table 4.5.

Table 4.5: Histograms of H & V features

Group	Features Number
Group A	320
Group B,C	288

4.4.3 Global transformations-Moment

- ***Hu Moment Invariants:*** Generally, in order to recognize shape the Hu Moment invariants have been used. Moments provide uniquely objects' characteristics to represent its shape. Seven of these unique shapes descriptor values computed independently of object scale, translation and orientation from central moments. Each group of the ICC has same numbers in this feature as seen in Table 4.6.

Table 4.6: Hu Moment Invariants

Group	Features Number
Group A,B and C	7

Totally, 1311 future vector elements consist of 984 Statistical, 320 Structural and 7 Moment in group A and 1179 future vector elements consist of 884 Statistical, 288 Structural and 7 Moment in group B and C. Table 4.7 briefly explains all feature details.

Table 4.7: The study feature extraction

Group	Features	Methods	Elements	Vector
A	Statistical	Zoning: Density Features	24	1311
		Profiles	640	
		Crossings	320	
	Structural	Histograms of Horizontal &Vertical	320	
	Moment	Hu Moment Invariants	7	
B, C	Statistical	Zoning: Density Features	20	1179
		Profiles	576	
		Crossings	288	
	Structural	Histograms of Horizontal &Vertical	288	
	Moment	Hu Moment Invariants	7	

4.5 Recognition Phase (Classification)

The Multiclass Support Vector Machine (SVMs) method as a one of the supervised learning machine, which includes the training and testing sections, was carried out in the recognition phase. SVMs applying a maximum voting on all possible and available binary classes' combination which called "one versus all". The main reason that the SVM method was chosen is because of its high recognition rate in comparison with other classification methods[17]. Furthermore, it's easy to use, good performance in generalization and capable to adapt various type of the problem with little tuning. SVM present various types of kernel function such as: linear, polynomial, quadratic, and radial basis kernel functions (RBF). The experimental study results revealed that RBF as a function to kernelized learning function and in this regard, Euclidean chosen as a distance function for this study. Moreover C and ξ parameters are determined as Table 4.8. The configured SVMs machine now are ready to complete the training and in continue testing phase in this study. According to feature extraction phase there are three main group of the characters, so in order to obtain the best results in classification it should be use three multiple SVMs based on these ICC three groups. Figure 4.7 explains the block diagram of this phase of the study and Figure 4.8 shows the Pseudocode. By using this method multi multiclass SVMs voting rate of the machine are decreasing significantly and accuracy recognition rate are increase dramatically.

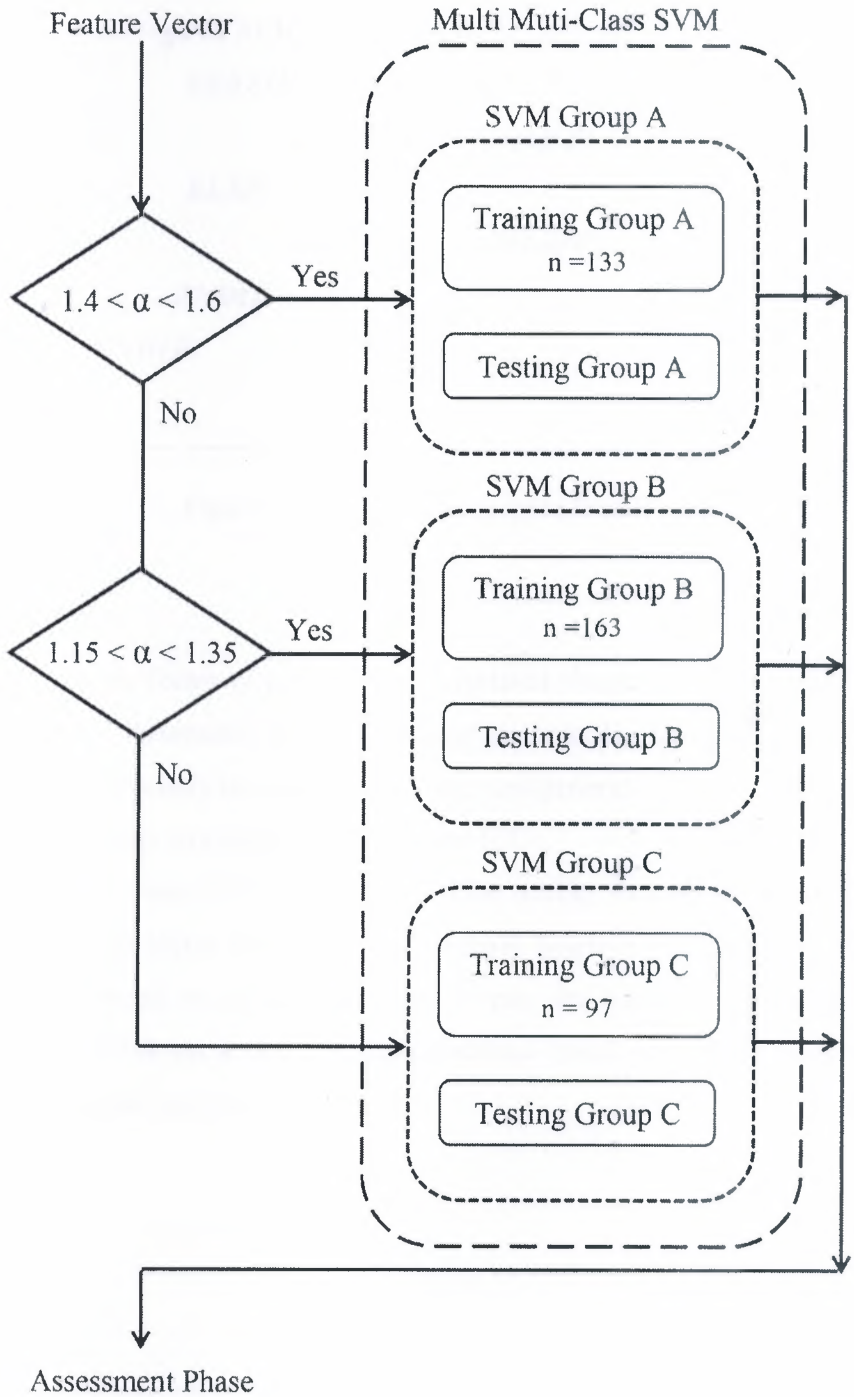


Figure 4.7: Classification chosen frame work

```
Begin
        IF α > 1.40 && α < 1.60 Then
                goTo SVM_ A (Group A)
                    ELSEIF α > 1.15 && α < 1.35
                                goTo SVM_B (Group B)
                    ELSE
                                goTo SVM_ C (Group C)
                ENDIF
        ENDIF
STOP
```

Figure 4.8: Classification pseudocode

4.5.1 Training phase

All 275 valid possible forms of each of the 17 isolated characters' image were created in the training database. Afterward, the classification method (Zhang, Berg, Maire and Malik, 2006) was applied to classify the samples database and generate a final model to be used in the testing phase in order to recognize the inserted ICCs.

The data was divided into 50% training and 50% testing data by choosing the images randomly. The Radial Basis Function (RBF) kernel function was used to implement a estimate technique based on simple Heureistic to gain the optimum C and ξ parameters. Implementing the SVM on a larger sample database resulted in more accurate results (Kazemi, Yousefnezhad and Nourian, 2015).

Table 4.8: Multiclass SVMs general properties

Parameters	Properties	Formula
Kernel Function	RBF	$K(\mathbf{x},\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right)$ (4.1)
Distance Function	Euclidean	$\|\mathbf{x}-\mathbf{x}'\|^2$ (4.2)
Sigma	6.22	ξ
Complexity	1.5	C
Tolerance	0.2	

4.5.2 Testing Phase

The main role of the study's testing phase was identifying the inserted isolated character and provide suitable feed to the assessment phase. SVM (Zhang, Berg, Maire and Malik, 2006; Sani, Ishak and Samad, 2010) was applied to identify the character by using the inserted character feature vector. The current study training and testing phase block diagram is shown in Figure 4.9.

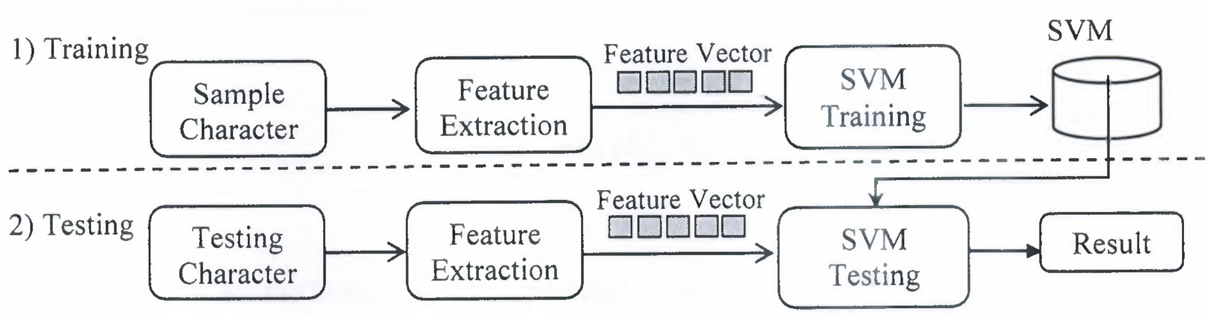


Figure 4.9: SVM training and testing block diagram

4.6 Assessment and Feedback phase

Providing a suitable practice platform for Persian calligraphy learners is the main aim of the PeNCIL. The last phases of the current study were assessing and displaying appropriate feedback on screen. In this study, the histogram of character contours distributions in polar coordinates was used as a feature descriptor. The contours are well-known shape representations that embody visual information by using a limited set of object points. In addition, the contours are widely used to reduce the number of points dramatically; moreover, the contours preserve the visual information shape (Belongie, Malik and Puzicha, 2002). Features were extracted by calculating the distances and angles of pixels inside a circular layout located at the shape centroid Figure 4.10.

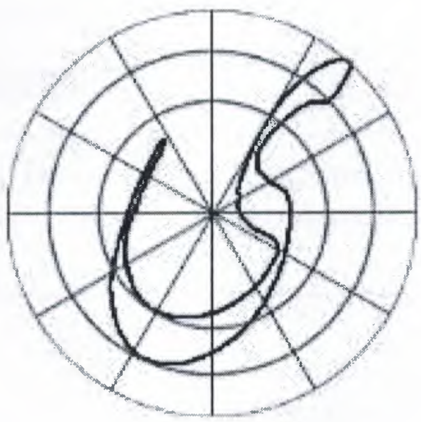


Figure 4.10: Feature extraction layout

The similarity between the identified template (I_1) and the ICC (I_2) is expressed by the Histogram Intersection measure:

$$S = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} min\,(H_{ij}^1, H_{ij}^2) \tag{4.3}$$

Where H^1 and H^2 are the histograms corresponding to template (I_1) and the ICC (I_2), and M and N are the histogram dimensions. The corresponding histogram of identified template (H^1) is fetched from the template dataset, where all templates' features had previously been extracted and saved.

This descriptor is efficient and global, which makes it conform more to human perception. This result creates feedback for the learners, which is displayed as a progress percentage on screen. Moreover, the correct identified template and inserted ICC image are shown immediately on screen in order to guide users to find out their mistakes. Figure 4.11 explains the assessment phase by using feature matching.

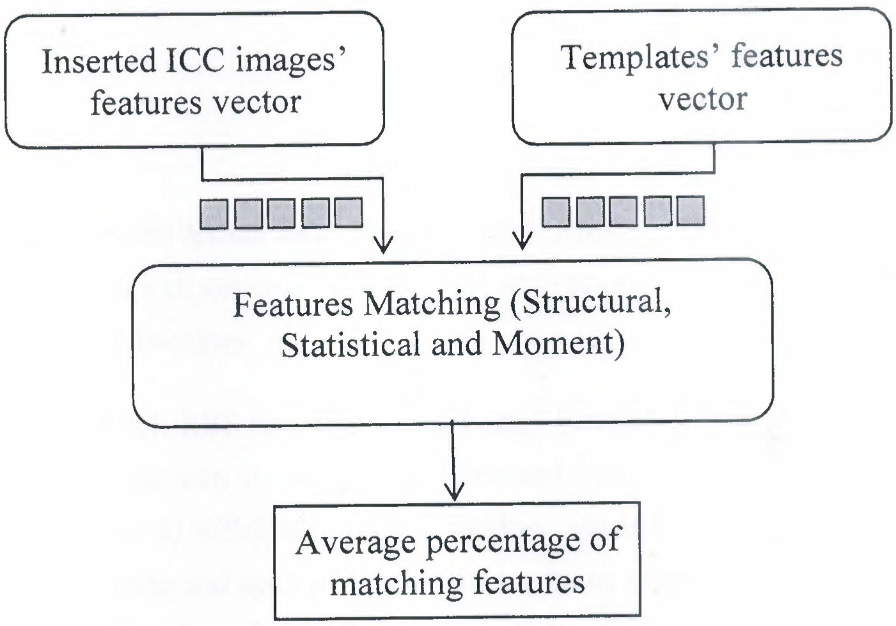


Figure 4.11: The assessment phase of developed system

CHAPTER 5
THE EXPERIMENTAL RESULTS AND ANALYSIS

5.1 Experimental Findings

In order to obtain results OpenCV 3.1 and Visual Studio 2013 based on C# coding are conducted in the study to implement the PeNCIL. The current thesis experimental results are out comes of training and testing each group of the characters individually:

- Group A
- Group B
- Group C

In order to training and testing data set, Multi (n=3) Multiple SVMs are applied in this study. Table 5.1 shows all variables which should set in advance. All these details are same in these three SVMs machines.

According to Table 5.1, RBF was chosen as a Kernel function of this machine. The main reason behind this selection is Radial Base Functions because of feature vector size. It means that RBF has hi efficiency in the machines which big size of feature vector. Furthermore, the sigma and optimum value are results of estimate techniques based on simple Heurestic. Also 1-vs.-1 general strategy was conducted in this study in order to reduce the problem of multiclass classification to multiple binary classification problems.

Table 5.1: The developed system SVMs properties

Parameters	Properties	Formula
Kernel Function	RBF	$$K(\mathbf{x},\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right) \quad (5.1)$$
Distance Function	Euclidean	$$\|\mathbf{x}-\mathbf{x}'\|^2 \quad (5.2)$$
Sigma	6.22	ξ
Complexity	1.5	C
Tolerance	0.2	
General Strategy	1-vs.-1	k*(k-1)/2 (5.3)

In rest of this chapter all results of experimental study for each groups are shown in Table 5.2 (Group A), Table 5.3 (Group B) and Table 5.4 (Group C) in order.

5.2 The experimental results of the study for group A

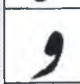
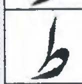
According to Table 5.2 (Group A), the amount of the Ratio (H*W) of this group members are equal 192* 128 and α is equal 1.5. Group A consists of 6 elements wich ctageorazed before based on the size of ratio. These data led the recognition phase to choice correct SVM machine and Correct Feature vectors in advance. The feature vectors which related to the Group A is consists of three main features Statistical, Structural and Moment. Statistical feature consists of three main feature vector namely Zoning, Profile and Crossing. The result of the Zoning which calculate the density of the ICC is 24 vector elements. Profile create the 640 vector elements and Crossing results is 320 vector elements. The second feature is Structural which used Histograms of Horizontal and Vertical. These feature product the 230 feature elements and at last Moment or Topological feature vector which used Hu Moment Invariants function provide 7 feature vector elements. All in all the Group A feature vector is results of these feature combination, it means that 1311 the relative feature vector is consists of 1311 number.

After feature extraction phase these vectors are ready to feed the Machine learning phase in both attempts: Training and Testing. 133 ICC samples of the Group A were used in the training phase. As shown in Table 5.2 each classified ICC indexed by number. $Class_0$ contains 17 training data, $Class_1$ contains 17 training data, Class2 contains 23, $Class_3$ contains 20, $Class_4$ contains 20, $Class_5$ contains 13. After complete the training phase, Testing phase was conducted. Testing phase which consists of 95 ICC were tested with these statistical details: $Class_0$ contains 14 testing data, $Class_1$ contains 12 testing data, $Class_2$ contains 19 testing data, $Class_3$ contains 14 testing data, $Class_4$ contains 14 testing data, $Class_5$ contains 10 testing data. Each ICC has own recognition as fallow: $Class_0$ Recognition rate is 94%, $Class_1$ Recognition rate is 92%, $Class_2$ Recognition rate is 95%, $Class_3$ Recognition rate is 86%, $Class_4$ Recognition rate is 89%, $Class_5$ Recognition rate is 93%. Due to these revealed results the recognition rate of Group A is averagely 91.50%.

Totally, the SVMs which was performed for Group A consists of 15 machine. In order to solve sub-problems of each machine $Class_1$ vs. $Class_0$, $Class_2$ vs. $Class_{0\&1}$, $Class_3$ vs. $Class_{0-2}$, $Class_4$ vs. $Class_{0-3}$, $Class_5$ vs. $Class_{0-4}$ and $Class_6$ vs. $Class_{0-5}$. Furthermore, the Group A

SVMs consist of 374 support vector points as below details: $Class_0$ contains 0 point, $Class_1$ contains 23 points, $Class_2$ contains 48 points, $Class_3$ contains 75 points, $Class_4$ contains 113 points and $Class_5$ contains 115 points.

Table 5.2: The experimental results of the study for group A

Group A								
Ratio (H×W)		192×128			α (ROI_H / ROI_W)		1.5	
Feature Vector	Statistical	Zoning: Density Features			24		1311	
		Profiles			640			
		Crossings			320			
	Structural	Histograms of Horizontal &Vertical			320			
	Moments	Hu Moment Invariants			7			
Machine Learning SVM (K=15)	ICC	Class Index	Training Dataset	Testing Dataset	Recognition Rate	Machine	Sub Problem C.Index vs C.Index	Support Vectors
		0	17	14	94%	-	-	-
		1	17	12	92%	1	7-vs-6	23
		2	23	19	95%	2	8-vs-(6,7)	48
		3	20	14	86%	3	9-vs-(6,7,8)	75
		4	20	14	89%	4	10-vs-(6,7,8,9)	113
		5	13	10	93%	5	11-vs-(6,7,8,9,10)	115
Total	6	6	133	95	-	15	15	374
Average Recognition Rate (%) For Group A						91.50		

5.3 The experimental results of the study for group B

According to Table 5.3 (Group B), the amount of the Ratio (H*W) of this group members are equal 160* 128 and α is equal 1.25. Group B consists of 9 elements which categorized before based on the size of ratio. These data led the recognition phase to choice correct SVM machine and Correct Feature vectors in advance. The feature vectors which related to the Group B is consists of three main features Statistical, Structural and Moment. Statistical feature consists of three main feature vector namely Zoning, Profile and Crossing. The result of the Zoning which calculate the density of the ICC is 20 vector elements. Profile create the 576 vector elements and Crossing results is 288 vector elements. The second feature is Structural which used Histograms of Horizontal and Vertical. These feature product the 288 feature elements and at last Moment or Topological feature vector which used Hu Moment Invariants function provide 7 feature vector elements. All in all the Group B feature vector is results of these feature combination, it means that 1179 the relative feature vector is consists of 1179 number.

After feature extraction phase these vectors are ready to feed the Machine learning phase in both attempts: Training and Testing. 163 ICC samples of the Group B were used in the training phase. As shown in Table 5.3 each classified ICC indexed by number. $Class_6$ contains 17 training data, $Class_7$ contains 21 training data, $Class_8$ contains 23, $Class_9$ contains 20, $Class_{10}$ contains 16, $Class_{11}$ contains 13, $Class_{12}$ contains 17, $Class_{13}$ contains 24, and $Class_{14}$ contains 12 data training. After complete the training phase, Testing phase was conducted. Testing phase which consists of 114 ICC were tested with these statistical details: $Class_6$ contains 12 testing data, $Class_7$ contains 12 testing data, $Class_8$ contains 18 testing data, $Class_9$ contains 13 testing data, $Class_{10}$ contains 13 testing data, $Class_{11}$ contains 13 testing data, $Class_{12}$ contains 14 testing data, $Class_{13}$ contains 10 testing data and $Class_{14}$ contains 7 testing data. Each ICC has own recognition as fallow: $Class_6$ Recognition rate is 93%, $Class_7$ Recognition rate is 82%, $Class_8$ Recognition rate is 84%, $Class_9$ Recognition rate is 86%, $Class_{10}$ Recognition rate is 94%, $Class_{11}$ Recognition rate is 94%, $Class_{12}$ Recognition rate is 92%, $Class_{13}$ Recognition rate is 81% and $Class_{14}$ Recognition rate is 93%. Due to these revealed results the recognition rate of Group B is averagely 89%.

Totally, the SVMs which was performed for Group B consists of 36 machine. In order to solve sub-problems of each machine $Class_7$ vs. $Class_6$, $Class_8$ vs. $Class_{6\&7}$, $Class_9$ vs. $Class_{6\text{-}8}$, $Class_{10}$ vs. $Class_{6\text{-}9}$, $Class_{11}$ vs. $Class_{6\text{-}10}$, $Class_{12}$ vs. $Class_{6\text{-}11}$, $Class_{13}$ vs. $Class_{6\text{-}12}$ and $Class_{14}$ vs. $Class_{6\text{-}13}$. Furthermore, the Group B SVMs consist of 1040 support vector points as below details: $Class_6$ contains 0 point, $Class_7$ contains 27 points, $Class_8$ contains 60 points, $Class_9$ contains 87 points, $Class_{10}$ contains 132 points, $Class_{11}$ contains 138 points, $Class_{12}$ contains 162 points, $Class_{13}$ contains 196 points and $Class_{14}$ contains 238 points

5.4 The experimental results of the study for group C

According to Table 5.4 (Group C), the amount of the Ratio (H*W) of this group members are equal 128*160 and α is equal 0.8. Group C consists of 2 elements which categorized before based on the size of ratio. These data led the recognition phase to choice correct SVM machine and Correct Feature vectors in advance. The feature vectors which related to the Group C is consists of three main features Statistical, Structural and Moment. Statistical feature consists of three main feature vector namely Zoning, Profile and Crossing. The result of the Zoning which calculate the density of the ICC is 20 vector elements. Profile create the 576 vector elements and Crossing results is 288 vector elements. The second feature is Structural which used Histograms of Horizontal and Vertical. These feature product the 288 feature elements and at last Moment or Topological feature vector which used Hu Moment Invariants function provide 7 feature vector elements. All in all the Group B feature vector is results of these feature combination, it means that 1179 the relative feature vector is consists of 1179 number.

After feature extraction phase these vectors are ready to feed the Machine learning phase in both attempts: Training and Testing. 97 ICC samples of the Group C were used in the training phase. As shown in Table 5.4 each classified ICC indexed by number. $Class_{15}$ contains 52 training data and $Class_{16}$ contains 45 training data. After complete the training phase, Testing phase was conducted. Testing phase which consists of 69 ICC were tested with these statistical details: $Class_{15}$ contains 31 testing data and $Class_{16}$ contains 38 testing data. Each ICC has own recognition as fallow: $Class_{15}$ Recognition rate is 93% and $Class_{16}$ Recognition rate is 89%. Due to these revealed results the recognition rate of Group C is averagely 91%.

Totally, the SVMs which was performed for Group C consists of 2 machine. In order to solve sub-problems of each machine $Class_{16}$ vs. Class15. Furthermore, the Group C SVMs consist of 18 support vector points as below details: $Class_{16}$ contains 18 point.

Table 5.3: The experimental results of the study for group B

Group B								
Ratio H×W		160×128			α (ROI_H / ROI_W)		1.25	
Feature Vector	Statistical		Zoning: Density Features		20			
			Profiles		576			
			Crossings		288		1179	
	Structural		Histograms of Horizontal &Vertical		288			
	Moments		Hu Moment Invariants		7			
Machine Learning SVM(k=36)	**ICC**	**Class Index**	**Training Dataset**	**Testing Dataset**	**Recognition Rate**	**Machine**	**Sub Problem C.Index vs C.Index**	**Support Vectors**
	ی	6	17	12	93%	-	-	-
	س	7	21	12	82%	1	7-vs-6	27
	ص	8	23	18	84%	2	8-vs-(6,7)	60
	ں	9	20	13	86%	3	9-vs-(6,7,8)	87
	م	10	16	13	94%	4	10-vs-(6,7,8,9)	132
	ل	11	13	13	94%	5	11-vs-(6,7,8,9,10)	138
	ح	12	17	14	92%	6	12-vs-(6,7, … ,11)	162
	ن	13	24	10	81%	7	12-vs-(6,7, … ,12)	196
	ع	14	12	7	93%	8	13-vs-(6,7, … ,13)	238
Total	9	9	163	114	-	36	36	1040
Average Recognition Rate (%) For Group B					**89**			

Table 5.4: The experimental results of the study for group C

Group C								
Ratio H×W	128×160			α (ROI_H / ROI_W)		0.8		
Feature Vector	Statistical	Zoning: Density Features			20		1179	
		Profiles			576			
		Crossings			288			
	Structural	Histograms of Horizontal &Vertical			288			
	Moments	Hu Moment Invariants			7			
Machine Learning SVM	ICC	Class Index	Training Dataset	Testing Dataset	Recognition Rate	Machine	Sub Problem C.Index vs C.Index	Support Vectors
		15	52	31	93%	-	-	-
		16	45	38	89%	1	15-vs-16	18
Total	2	2	97	69	-	1	1	18
Average Recognition Rate (%) For Group C			91					

5.5 Results Comparison

According to Table 5.5 the average recognition rate the developed system (PeNCIL) is equal 90.50. In addition, each group average recognition rate showed that Group B has less accuracy rate than the others. In excess of common failures in working with ink and image capturing insufficiencies, the Group B ICCs' shapes similarity are another problems which can cause this result. Considering to the shape of the Group B characters and experimental results in the Group B, we can claim that this developed system confuse when inserted ICCs are one of these characters as listed in below:

- Similar Group 1:  , and .
- Similar Group 2: , .

Similar Group 1 consists of three ICCs. These three ICCs are in the same group (Group B) with same ratio size 1.25 and same feature vector size. The other hand group 1 ICCs are so similar in shape and common beginner learners failures in working with ink and image capturing insufficiencies sometimes PeNCIL confused in recognition phase.

Similar Group 2 which consists of two ICCs are in the group B with same ratio size 1.25 and same feature vector size. Similar with the Group 1 these two ICCS are similar in shape and common beginner learners failures in working with ink and image capturing insufficiencies sometimes PeNCIL confused in recognition phase.

Table 5.5: The experimental results of the study

Average Recognition Rate	Training Data	Testing Data	Recognition Percentage
Group A	133	95	91.50
Group B	163	114	89.00
Group C	97	69	91.00
Average Recognition Rate of PeNCIL			**90.50**

CHAPTER 6
SYSTEM IMPLEMENTATION

6.1 Data Acquisition

In this chapter different Panels and their properties of the designed software for identifying ICC and assessing them by using template are demonstrated. OpenCV 3.1 and Visual Studio 2013 based on C# coding are conducted in the study to implement the PeNCIL. The PeNCIL is consists of a main page which provide facilities to insert the scanned or captured image. Figure 6.1 shows the main page of this system.

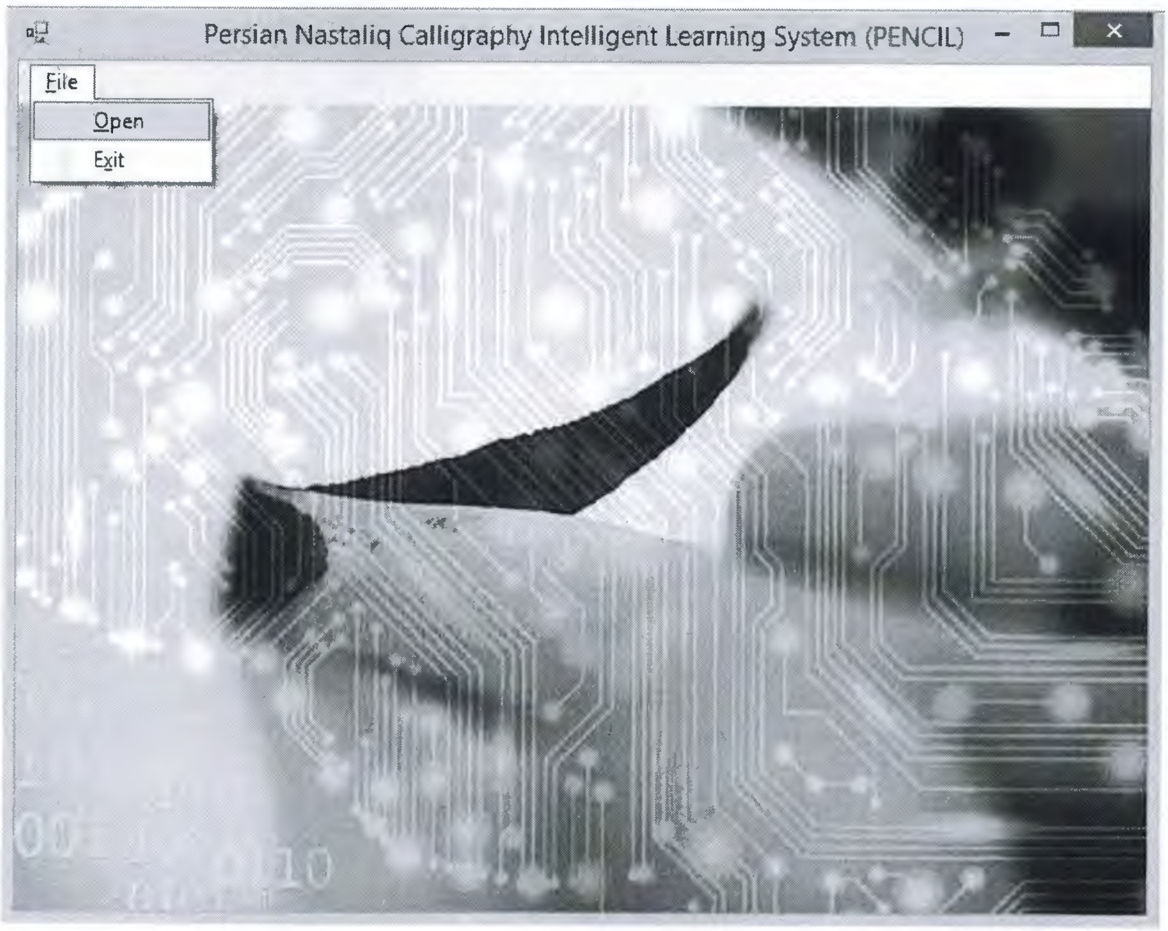


Figure 6.1: Snapshot of the PeNCIL main page

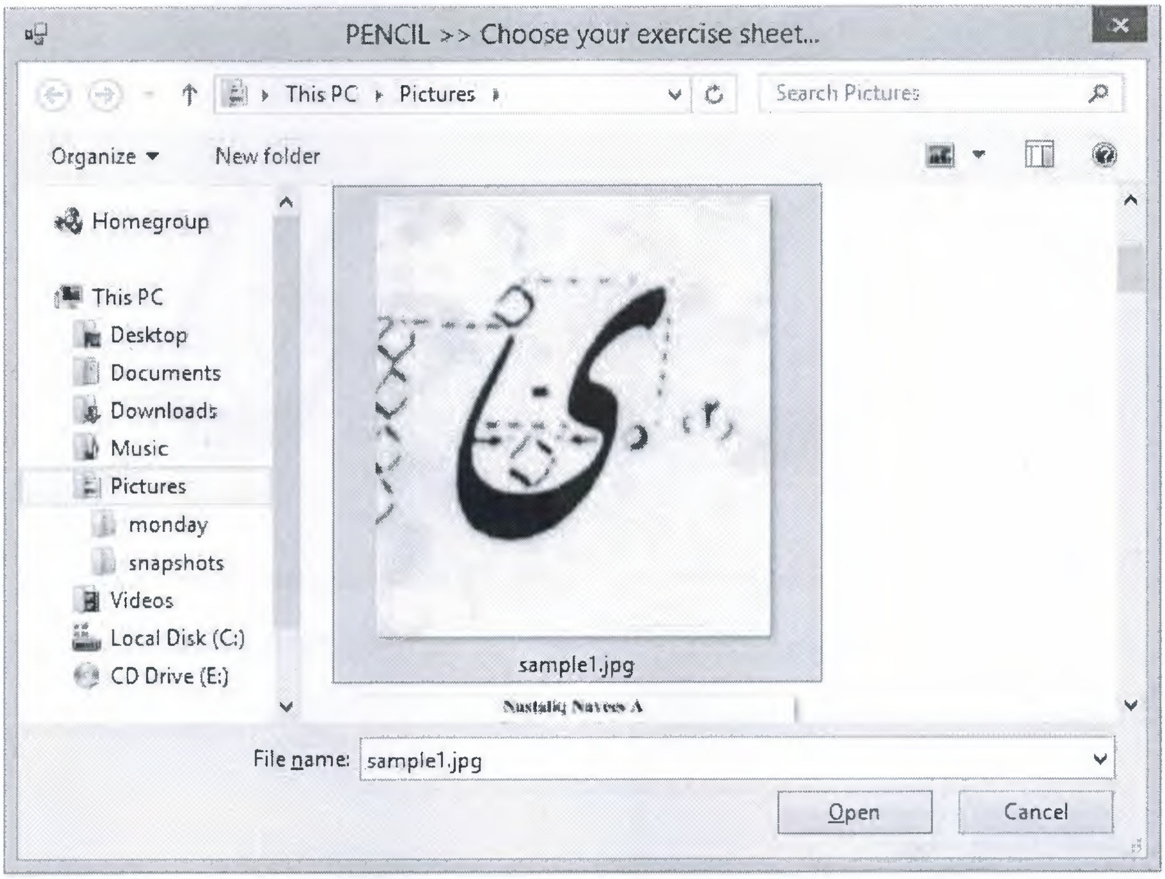


Figure 6.2: Snapshot of image browsing

6.2 Image Processing Phase

As we mentioned before in the chapter 4 image processing phase of this thesis consists of 5 pre-processing steps in order to prepare inserted image to send recognition phase. Figure 6.3 shows all steps snapshots.

a) Snapshot of raw image

b) Convert to GrayScale

c) Snapshot of Denoised

d) Snapshot of Binarization

Figure 6.3: Continued on the next page….

Figure 6.3: Continued…

e) Snapshot of Dilate

f) Snapshot of Erode

g) Snapshot of Invert Image

h) Snapshot of ROI Detection

Figure 6.3: Continued on the next page….

Figure 6.3: Continued...

i) Snapshot of Crop

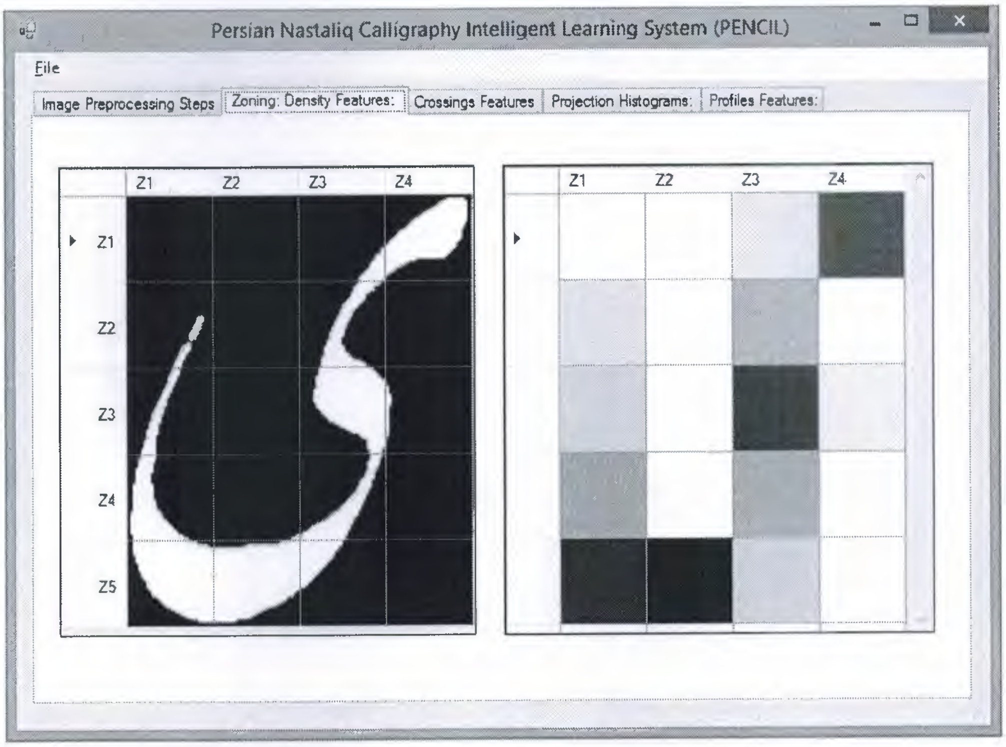
j) Snapshot of Final Noise Removal

k) Snapshot of Normalization

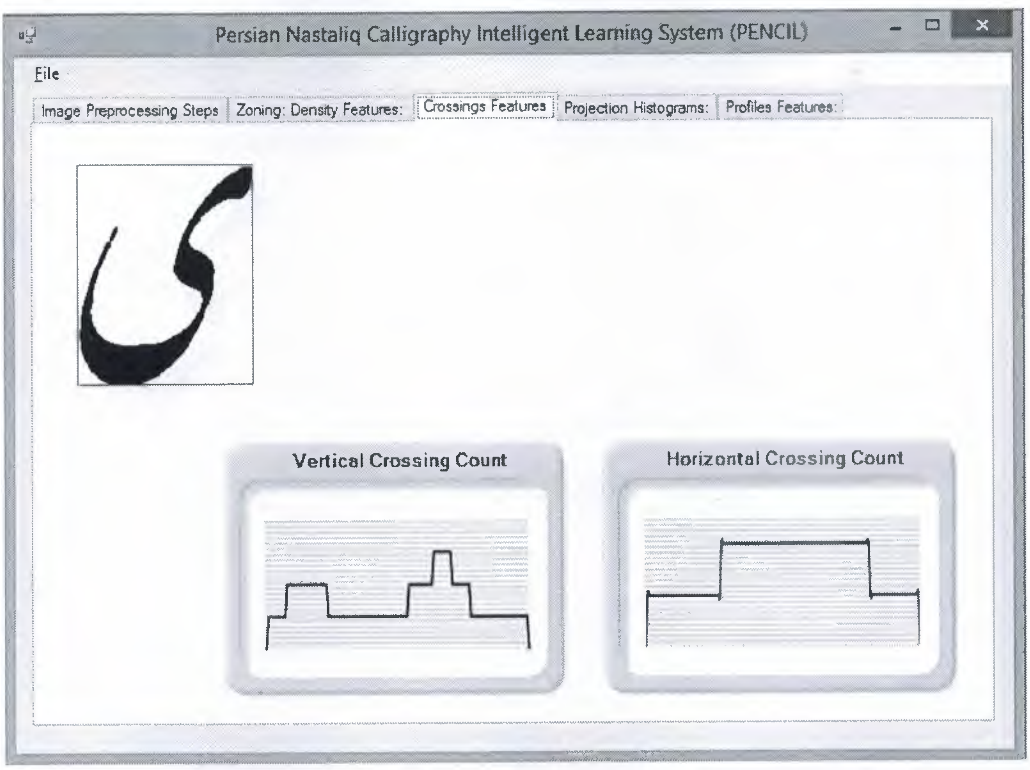
Figure 6.3: Snapshots of image pre-processing of the study

6.3. Feature Extracting

In order to provide the inserted images to understandable format for machine learning phase, Feature Extraction convert the images to relative feature vector is applied. Figure 6.4 in all parts (a), (b), (c) and (d) show the results of each feature steps.



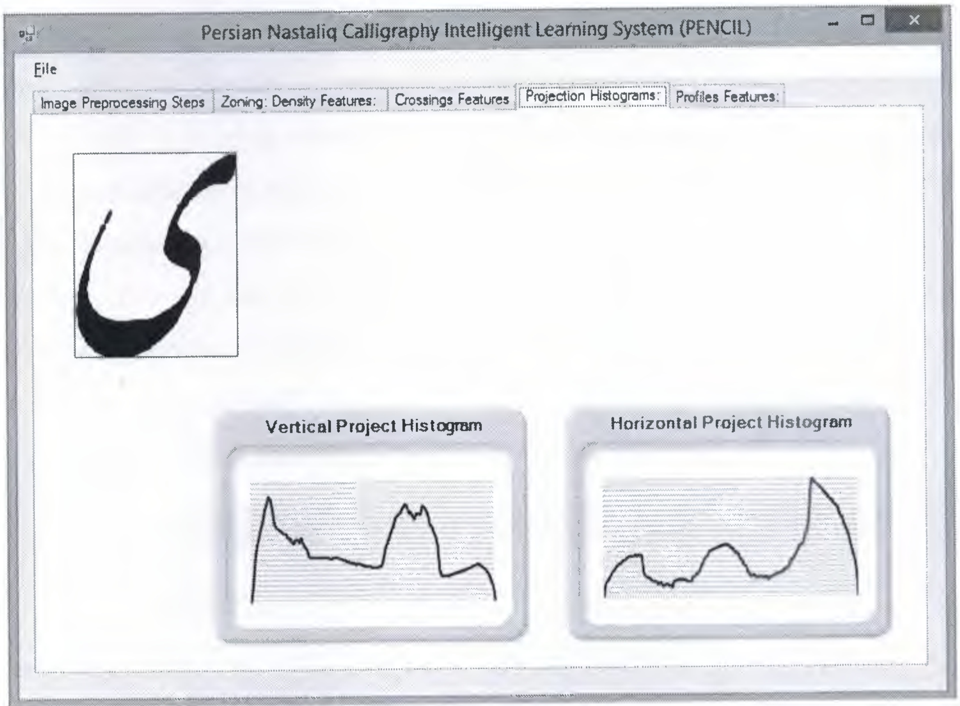
(a) Snapshot of Zoning: Density Feature



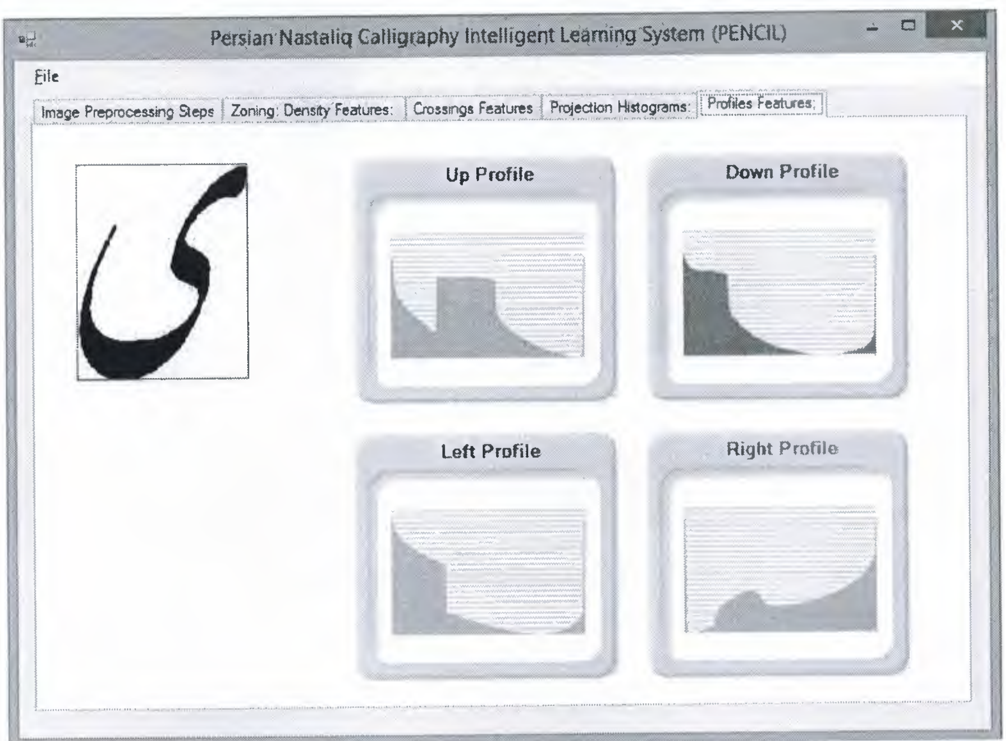
(b) Snapshot of Crossing Feature

Figure 6.4: Continued on the next page….

Figure 6.4: Continued…



(c) Snapshot of Projection Histogram



(d) Snapshot of Profile Feature

Figure 6.4: Feature extraction snapshots

6.4. Machine Learning

After passing the image pre-processing phases and convert to the feature vector now ICCs are ready to feed the machine learning. The Multi Multiclass SVM are conducted in this study which need a training phase in advance and testing phase. Figure 6.5 shows the training phase snapshot of sample character of Group B. As seen in Figure 6.5 the right side of menu is consists of SVMs Machine initialization data such as type of Kernel, Sigma and Tolerance. The left side of window displays the Group B sample data set to training. After set all variables and insert the sample data, "Start training" button applied to train the machine. "Classify by Voting" button provide facilities to get information about the rate of the voting results of this training.

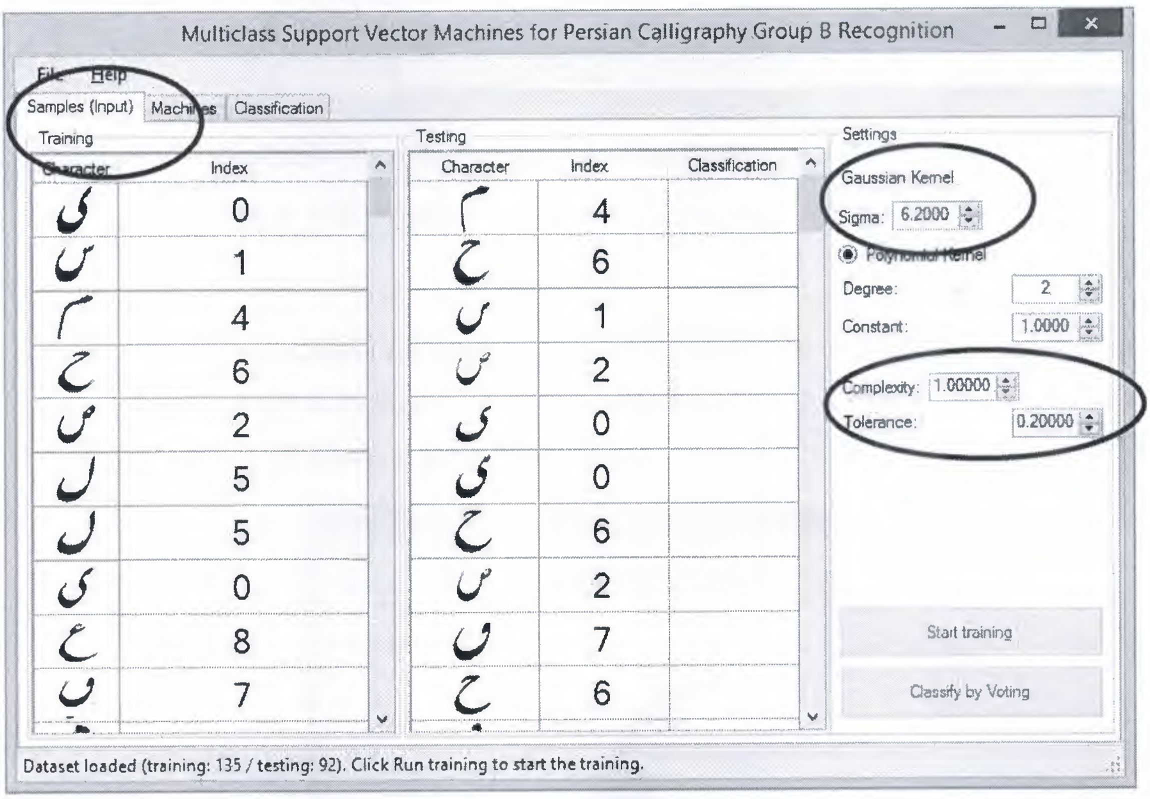


Figure 6.5: The snapshot of machine training phase of the PeNCIL

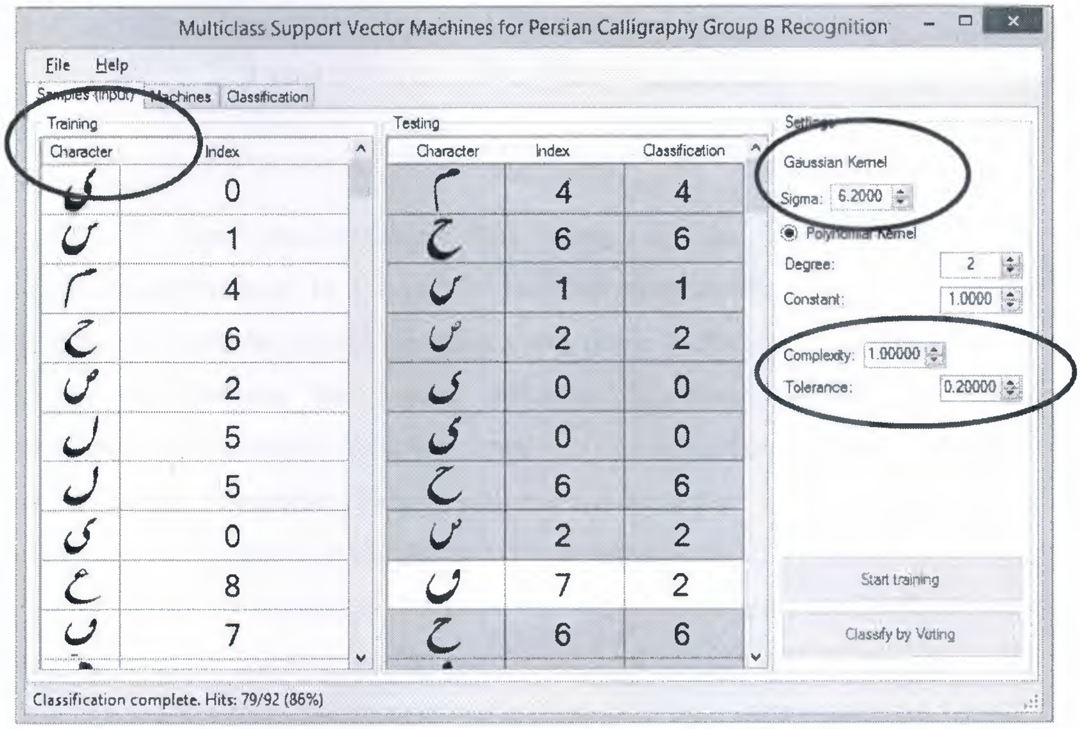


Figure 6.6: The snapshot of machine training phase of the PeNCIL

6.5. Assessment Phase

The next step after successful recognition of the inserted ICC in testing phase, assessment process will be applied to evaluate the inserted ICC by its relative sample and display the accuracy percentage on screen as seen in Figure 6.7.

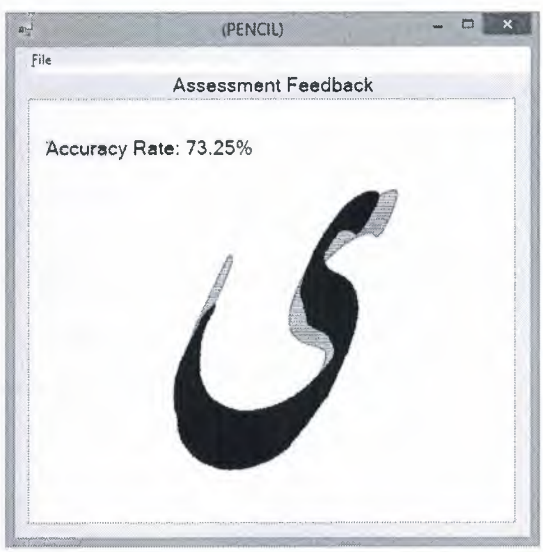


Figure 6.7: The Snapshot of assessment process output

CHAPTER 7
CONCLUSION AND FEATURE STUDIES

7.1 Conclusion

The PeNCIL was developed in this study as the first intelligent Persian Nastaliq calligraphy tool. In this study, the two main methods applied to create the intelligent tool were the image processing and multi multiclass SVM machine learning techniques. In addition, the feature matching technique was used to assess the compatibility of the learners' work with standard Persian Nastaliq calligraphy features. The developed system is capable of identifying and assessing the learners work automatically, which leads to a reduction in human errors that occur in the traditional method and also saves time. By using Structural, Statistical and Moment features, the PeNCIL provides a teacherless standard platform for Persian Nastaliq calligraphy enthusiasts and learners to use and practice wherever and whenever they like.

In order to obtain the desired results the ICC are categorized in three main groups (A, B and C). These categories are based on each ICC ratio size. The experimental result of this study revealed that averagely the testing and recognizing characters phase have 90.50 percentage accuracy.

7.2 Future Studies

The current system was developed based on isolated character of Nastaliq. So, it seems that, in order to cover the complete words one has to add more traing data in the recognition phase in the system.

On the other hand, there are various types of platform in intelligent tutors systems. The developed system can be implement in other platforms such as mobile platform where learners don't need to use ink in the practicing Mashq.

REFERENCES

Abraham, A. (2005). Artificial Neural Networks. In P.H. Sydenham and R. Thorn (Eds.), *Measuring System Design*. New York, NY, United State: John Wiley & Sons, Ltd.

Albu, F., Hagiescu, D., Puica, M. & Vladutu, L. (2015). Intelligent tutor for first grade children's handwriting application. *In proceeding of the 9th International Technology, Education and Development Conference* (pp. 3708-3717). Madrid, Spain: IATED.

Arivazhagan, S. & Ganesan, L. (2003). Texture classification using wavelet transform. *Pattern Recognition Letters, 24,* 1513-1521.

Aussem, A., Murtagh F. & Sarazin M. (1995). Dynamical recurrent neural networks towards environmental time series prediction. *International Journal of Neural Systems, 6*, 145–170.

Belongie, S., Malik, J. & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(4), 509-522.

Bezine, H. & Alimi, A. M. (2013). Development of an Arabic handwriting learning educational system. *International Journal of Software Engineering & Applications, 4*(2), 33-49.

Biemond, J. & Mersereau, R. M. (1997). Image and video processing. In V. Madisetti, and D.B. Williams (Eds.). *The Digital Signal Processing*, Florida, United State: CRC Press.

Burges, J. C. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery, 2*, 121–167.

Chapelle, O., Scholkopf , B. & Zien Eds, A. (2009). Semi-Supervised Learning. *IEEE Transactions on Neural Networks, 20*(3), 542.

Chen, S. & Haralick, R. M. (1995). Recursive Erosion, Dilation, Opening, and Closing Transforms. *IEEE Transactions, 4*(3), 335-345.

Chen, X., Li, S., Hu, J. & Liang, Y. (2014). A Survey on Otsu image segmentation methods. *Journal of Computational Information Systems, 10*, 4287-4298.

Costa, F. M. (2011). Image processing. Application to the characterization of thin films Manuel. *Journal of Physics: Conference Series, 274*, 1-9.

Costa, M. F. M. (2004). Application of image processing to the characterization of nanostructures. *Journal Mater Science, 6*(1), 12–20.

Demin, W., Haese-Coat, V., Bruno, A. & Ronsin, J. (1995). Some statistical properties of mathematical morphology. *Signal Processing, IEEE Transactions, 43*(8), 1955- 1965.

Patterson, K. (2001). A Neural Network image interpretation system to extract rural and urban land use and land cover information from remote sensor data, Geocarto International. Retrieved 22 June 2016 from http://www.utdallas.edu/~ffqiu/

Flickner et al. (1995). Query by image and video content: the QBIC system. *IEEE Computer*, *28*(9), 23-32.

Ganesan, S., Subashini, T. S. & Jayalakshmi, K. (2014). Classification of X-rays using statistical moments and SVM. *In Proceeding of International Conference on Communications and Signal Processing* (pp. 1109- 1112). Melmaruvathur: IEEE.

Gonzalez, R. C. & Woods, R. E. (1992). Digital Image Processing. Reading, Massachusetts, United State: Addison-Wesley.

Gonzalez, R. C. & Woods, R. E. (2002). Digital Image Processing (2ndEdition). Upper Saddle River, New Jersey, United State: Prentice Hall, Inc.

Gupta, M. R., Jacobson, N. P. & Garcia, E. K. (2007). OCR binarization and image pre-processing for searching historical documents. *Pattern Recognition, 40*, 389-397.

Hamidi, B. & Rashvand, Z. (2015). Visual capacity of Siah-MashqNastaliq calligraphy, in contemporary typography posters of Iran. *International Journal of Arts and Commerce, 4*(9), 124-131.

Han, C., Chou, C. & Wu, C. (2008). An interactive grading and learning system for Chinese calligraphy. *Machine Vision and Applications, 19,* 43–55.

Haralick, R. M., Sternberg, S. R. & Zhuang, X. (1987). Image analysis using mathematical Morphology. *IEEE Transactions Pattern Analysis and Machine Intelligence, 9*(4), 532-550.

Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning, 42*(1), 177–196.

Hossain, S. S., Maiti, A. & Chaki, N. (2011). Image binarization using iterative partitioning: A global Thresholding approach. *In Proceeding of International Conference of Recent Trends in Information Systems, doi: 10.1109/ReTIS.2011.6146882*

Hsu, C. W., Chang, C. C. & Lin, C. J. (2010). A practical guide to support vector classication. Department of Computer Science, National Taiwan University, Taipei 106 Taiwan. Retrieved June 22, 2016, from http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Hu, Y. & Ji, H. (2009). Research on image median filtering algorithm and its FPGA implementation. *In Proceeding of the Global Congress on Intelligent Systems* (pp. 226-230). Xiamen: IEEE.

Islam, M. M., Zhang, D. & Lu, G. (2008). A geometric method to compute directionality features for texture images. *In Proceeding of International Conference on Multimedia and Expo* (pp. 1521-1524). Hannover: IEEE.

Jain, A. K. & Vailaya, A. (1996). Image retrieval using colour and shape. *Pattern Recognition, 29*(8), 1233-1244.

Jain, A. K., Mao, J. & Mohiuddin, K. M. (1996). Artificial Neural Networks: A tutorial. IEEE. Retrieved June 22, 2016, from http://www.cogsci.ucsd.edu/~ajyu/Teaching/Cogs202_sp12/Readings/jain_ann96.pdf

Jensen, J. R. (2005). Introductory digital image processing: A remote sensing perspective. New Jersey, USA: Prentice Hall Series in Geographic Information Science.

Jensen, J. R., Qiu, F. & Ji, M. (1999). Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *International Journal of Remote Sensing, 20*(14), 2805-2822.

Kazemi, M. Yousefnezhad, M. & Nourian, S. (2015). A new approach in Persian handwritten letters recognition using error correcting output coding. *Journal of Advances in Computer Research, 6*(4), 107-124.

Kumar, G. & Bhatia, P. K. (2014). A detailed review of feature extraction in image processing systems. *In Proceeding of the Fourth International Conference on Advanced Computing & Communication Technologies* (pp. 5-12). IEEE.

Leung, W. H. & Chen, T. (2002). Trademark retrieval using contour-skeleton stroke classification. *In Proceeding of International Conference on Multimedia and Expo* (pp. 517-520). IEEE.

Li, S. & Shawe-Taylor, S. (2005). Comparison and fusion of multi-resolution features for texture classification. *Pattern Recognition Letters*, *26*(5), 633-638.

Liu et al. (2007). A shape ontology framework for bird classification. *In Proceeding of 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications* (pp. 478-484). Glenelg, Australia: IEEE.

Ma, Y. & Guo, G. (2014). Support Vector Machines applications. Springer Science & Business Media.

Maini, R. & Aggarwal, H. (2010). A comprehensive review of image enhancement techniques. *Journal of Computing, 2*(3), 8-13.

Mathur, A. & Foody, G. M. (2008). Multiclass and binary SVM classification: Implications for training and classification users. *IEEE Geoscience and Remote Sensing Letters,* doi: 10.1109/LGRS.2008.915597, 241- 245.

Mendozaa, O., Melinb, P. & Liceaa, G. (2009). A hybrid approach for image recognition combining type-2 fuzzy logic, modular neural networks and the Sugeno integral. *Information Sciences, 179*(13), 2078–2101.

Mesanovic, N., Mujagic, S., Huseinagic, H. & Kamenjakovic, S. (2012). Application of lung segmentation algorithm to disease quantification from CT images. *In Proceeding of International Conference of System Engineering and Technology* (pp.1-7). Bandung: IEEE.

Mezaris, V., Kompatsiaris, I. & Strintzis, M. G. (2003). An ontology approach to object-based image retrieval. *In Proceeding of International Conference on Image Processing,* (pp. 511-514). IEEE.

Mitchell, M. T. (1997). Machine learning. McGraw-Hill International Editions

Mohsen, S. (2009). The problem of stretching in Persian calligraphy and a new type postscript Nastaliq font, Unpublished Master Thesis, University of Waterloo, Canada.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *System Man Cybern, 9*(1), 62-66.

Pass, G. & Zabith, R. (1996). Histogram refinement for content-based image retrieval. *In Proceeding of Workshop on Applications of Computer Vision* (pp. 96-102). Sarasota: IEEE.

Patel, S., & Goswami, M. (2014). Comparative analysis of histogram equalization techniques. *In Proceeding of International Conference Computing and Informatics* (pp. 167-168). Mysore: IEEE.

Patel, U. (2013). An introduction to the process of optical character recognition. *International Journal of Science and Research, 2*(5), 155-158.

Qui, F. & Jensen, J.R. (2004). Opening the black box of neural networks for remote sensing image Classification. *International Journal of Remote Sensing, 5,* 1366-5901.

Rasmussen, C. E. & Williams. C. K. I. (2005). Gaussian processes for machine learning. Cambridge, Massachusetts, United States: MIT Press.

Reza, Z. M., Mueen, A., Seng, W. C. & Awedh, M. H. (2011). Combined feature extraction on medical X-ray images. *In Proceeding of 3th International Conference on Computational Intelligence, Communication Systems and Networks* (pp.264-268). Bali: IEEE.

Rouhfar, A. (2014). A new, easy, and effective method to learn Persian Calligraphy (Nastaliq). Amazon: Create Space Independent Publishing Platform.

Russ, J. C. (2011). The image processing Handbook. 6th Edition, Taylor and Francis Group, LLC.

Sani, M. M., Ishak, K. A. & Samad, S. A. (2010). Classification using adaptive Multiscale Retinex and support vector machine for face recognition system. *Journal of Applied Sciences, 10*, 506-511.

Sattar, S. A., Haque, S., Pathan, M. K. & Gee, Q. (2008). Implementation challenges for nastaliq character recognition. *Communications in Computer and Information Science, 20*, 279-288.

Schapire, R. E. (2003). The boosting approach to machine learning: An overview. *Nonlinear Estimation and Classification* (pp 149-171). Lecture Notes in Statistics

Senthilkuamaran, N. & Thimmiaraja, J. (2014). Histogram equalization for image enhancement using MRI brain images. *In Proceeding of World Congress of Computing and Communication Technologies* (pp. 80-83). Trichirappalli: IEEE

Shih, T. K. et al. (2001). An intelligent content-based image retrieval system based on color, shape and spatial relations. *National Science Council Physical Science and Engineering, 25*(4), 232-243.

Srinivasan, K. S. & Ebenezer, D. (2007). A new fast and efficient decision-based algorithm for removal of high-density impulse noises. *IEEE Signal Processing Letters, 14*(3), 189-192.

Stanchev, P. L., Green, Jr. D. & Dimitrov, B. (2003). High level colour similarity retrieval. *International Journal of Information Theories and Applications*, *10*(3), 363-369.

Stone, P. & Veloso, M. (2000). Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots*, *8*(3), 345-383.

Taele, P. & Hammond, T. (2015). BopoNoto: An intelligent sketch education application for learning Zhuyin phonetic script. *In Proceedings of the 21st International Conference on Distributed Multimedia Systems* (pp. 101-107), Vancouver, Canada, DOI: 10.18293/DMS2015-031.

Thomas, R. A. & Kumar, S. S. (2014). Automatic detection of lung nodules using classifiers. control, instrumentation, communication and computational technologies, doi: 10.1109/ICCICCT.2014.6993051

Thome, A. C. G. (2012). SVM classifiers concepts and applications to character recognition. Retrieved June 22, 2016, from http://dx.doi.org/10.5772/52009

Tian, D. P. (2013). A review on image feature extraction and representation techniques. *International Journal of Multimedia and Ubiquitous Engineering, 8*(4), 385.

Tsai, C. F. (2007). Image mining by spectral features: A case study of scenery image classification. *Expert Systems with Applications, 32*(1), 135-142.

Varshney, S. S., Rajpal, N. & Purwar, R. (2009). Comparative study of image segmentation techniques and object matching using segmentation. *Methods and Models in Computer Science,* doi: 10.1109/ICM2CS.2009.5397985

Xu, S., Jiang, H., Lau, F. C. M. & Pan, Y. (2007). An intelligent system for Chinese calligraphy. *In proceeding of the 22nd national conference on Artificial Intelligence* (pp.1578-1583). Vancouver, Canada: IEEE.

Yang, C., Dong, M. & Fotouhi, F. (2005). Image content annotation using Bayesian framework and complement components analysis. *In Proceeding of International Conference on Image Processing* (pp. 1193-1196). IEEE

Yang et al. (2008). A fast MPEG-7 dominant colour extraction with new similarity measure for image retrieval. *Journal of Visual Communication and Image Retrieval, 19,* 92-105.

Yang, M., Kpalma, K. & Ronsin, J. (2008). A survey of shape feature extraction techniques. *Pattern Recognition, 5*, 43-90.

Zhang, D. & Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition, 37*(1), 1-19.

Zhang, D. S., Islam, Md. M. & Lu, G. J. (2012). A review on automatic image annotation techniques. *Pattern Recognition, 45*(1), 346-362.

Zhang, H., Berg, A. C., Maire, M. & Malik, J. (2006). SVM-KNN: Discriminative nearest neighbour classification for visual category recognition. *Computer Vision and Pattern Recognition, 2*, 2126-2136.

Zhu, X. & Goldberg, A. B. (2009). Introduction to semi-supervised learning synthesis lectures on artificial intelligence and machine learning. Florida, United State: Morgan & Claypool Publishers.