DETECTION OF MAN-MADE STRUCTURES IN AERIAL IMAGERY USING QUASI-SUPERVISED LEARNING AND TEXTURE FEATURES

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Electronics and Communication Engineering

by
Mesut GÜVEN

December 2010
İZMİR

We approve the thesis of **Mesut GÜVEN**

Assoc. Prof. Dr Bilge KARAÇALI
Supervisor

Assist. Prof. Dr. Şevket GÜMÜŞTEKİN
Committee Member

Assist. Prof. Dr. Türker İNCE
Committee Member

24 December 2010

Prof. Dr. Ferit Acar SAVACI
Head of the Department of Electrical
Electronics Engineering

Prof. Dr. Sedat AKKURT
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Assoc. Prof. Dr. Bilge KARAÇALI for his valuable guidance and support. I am very proud that I had the chance to work with him.

I would like to thank to the members of my Thesis Committee Asst. Prof. Dr. Şevket GÜMÜŞTEKİN and Asst. Prof. Dr. Türker İNCE for their useful comments. I also would like to thank to my friends Devrim ÖNDER, Başak Esin KÖKTÜRK and Tunca DOĞAN for their unfailing support.

I want to express my gratitude to my superior commanders for their support and permission in my education. Without their motivation, I will not be able to complete my education.

Finally, I am deeply thankful to my family and my lovely wife Tuba for their support and endless love through my life.

ABSTRACT

DETECTION OF MAN-MADE STRUCTURES IN AERIAL IMAGERY USING QUASI-SUPERVISED LEARNING AND TEXTURE FEATURES

In this thesis, the quasi-supervised statistical learning algorithm has been applied for texture recognitioning analysis. The main objective of the proposed method is to detect man-made objects or differences on the terrain as a result of habitating. From this point of view, gaining information about human presence in a region of interest using aerial imagery is of vital importance. This task is adressed using a machine learning paradigm in a quasi-supervised learning.

Eigthteen different sized aerial images were used in all computations and analysis. The available data was divided into a reference control set which consist of normalcy condition samples with no human presence, and a mixed testing data set which consisting images of habitate and cultivated terrain. Grey level co-occurrence matrices were then computed for each block and “Haralick Features” were extracted and organized into a texture vector. The quasi-supervised learning was then applied to the collection of texture vectors to identify those image blocks which show human presence in the test data set.

In the performance evaluatian part, detected abnormal areas were compared with manually labeled data to determine the corresponding reciever operating characteristic curve. The results showed that the quasi-supervised learning algorithm is able to identify the indicators of human presence in a region such as houses, roads and objects that are not likely to be observed in areas free from human habitation.

ÖZET

DOKU ÖZNİTELEKLERİ VE YARI GÜDÜMLÜ ÖĞRENME İLE HAVA GÖRÜNTÜLERİNDE İNSANA AİT YAPILARIN TESPİT EDİLMESİ

Bu tezde, doku özniteliklerinin tanınması ve analizi için istatistiksel "yarı güdümlü öğrenme" hava fotoğraflarına uygulanmıştır. Sözü edilen metodun ana hedefi; insana ait nesneleri ve arazideki değişimleri tespit etmektir. Bu bakış açısıyla; incelenenen bir arazi parçasında insan varlığına ilişkin bir bilgiye sahip olmak büyük önem arz etmektedir. Bu iş yarı güdümlü öğrenme yardımıyla yapılmaya çalışılacaktır.

Tüm hesaplamalarda ve analizlerde farklı boyuttaki 18 hava fotoğrafı kullanılmıştır. Mevcut resimler, insana ait izler bulunmayan referans kontrol grubuna ve insana ait izler içeren karışık test grubuna ayrılmıştır. Daha sonra gri seviyeli eş oluşum matrisleri hesaplanmış ve bu matrislerden "Haralick öznitelikleri" ile desen vektörleri elde edilmiştir. Sonraki adımda yarı güdümlü öğrenmenin, insana ait izler içeren blokları tespit edebilmesi için öğrenme algoritması desen vektörleri üzerinde koşturulmuştur.

Performans değerlendirme kısmında ise; yarı güdümlü öğrenmenin tespit ettiği anormal bölgeler, el ile etiketlenmiş bloklarla karşılaştırılarak sınıflandırma başarım eğrisi çıkartılmaktadır. Sonuçlar yarı güdümlü öğrenmenin evler, yollar gibi insana ait nesneler ile doğal yaşamda bulunması güç olan dokuları yüksek bir yüzdeyle otomatik olarak tespit edebildiğini göstermektedir.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

CHAPTER 1

INTRODUCTION

Today machine learning applications are being used in many fields. The aim of the machine learning is to automatically learn to recognize complex patterns and make intelligent decisions based on data. With the advent of air photography, unmanned air vehicle technologies and high-speed computers, it is becoming possible to perform learning algorithms on pictorial data. High resolution air photographs are used for reconnaissance efforts. Intelligence specialists try to gain information by visual observation or other detection methods, about the activities and resources of potential threats. They look for tangible structures, movements of opposing forces and any terrestrial abnormalities on a particular area.

In this thesis, an automated quasi-supervised learning algorithm is applied to air photographs in a reconnaissance scenario. The pictorial information was provided from a reconnaissance aircraft and all the images used in the experiments were captured using a high-resolution aerial camera. The resolution of the images was 0.43 meter per pixel. Aerial images were converted to grayscale image and computations were made on grayscale images. The study was carried-out on eighteen aerial images of different sizes extracted from a big aerial image, nine of these images have natural terrestrial conditions and there are not any human made structures or vehicles etc. The other nine images have terrestrial conditions such as cultivated lands, human made buildings and roads. Images were farther divided into 625 square-shaped grid blocks; each block representing the related regions on the images. Nine different block sizes were used in the experiments and sizes of these blocks are: 7, 10, 14, 20, 28, 40, 56, 63, and 80 pixels respectively. In other words, images were divided into blocks of $9m^2$, $18.5m^2$, $36.2m^2$, $74m^2$, $145m^2$, $296m^2$, $580m^2$, $734m^2$, and $1183m^2$ areas on land respectively. For every block size, co-occurrence matrices were computed from each block. There are two parameters used in computing co-occurrence matrices, first one is the distance between neighboring pixels and the second is the angular relationship of the neighboring pixels. Various alternatives were used to compute different co-occurrence matrices. From these

matrices, several Haralick features were calculated and organized into texture vectors to be used for recognition.

The strategy used for recognition is a quasi-supervised learning that requires the prior knowledge of only the presence or absence of abnormalities in the respective datasets and not the labeling of individual samples. Quasi-supervised statically learning depends on a reference control data set which consist of only normalcy condition with no human presence and a mixed testing data set which consist of human-made objects along with unhabitated land. The learning algorithm than detects the samples that are unique to the testing data set. By definition, those regions special to the testing data set are abnormal regions that we want to detect as regions of interest. For a reconnaissance scenario like this, the regions of interest on aerial images can be illustrated as human-made constructions, roads and cultivated terrains.

This thesis is organized as follows: in chapter 5, performance of the learning algorithm with different distance parameters and different block sizes are measured via using the receiver operating curve. In performance evaluation part, each block in the test image was needed to be labeled manually as normal or abnormal. The blocks which consist of completely or partially human-made objects were labeled as abnormal regions. These abnormal blocks have roads, cultivated soil or anything that shows human existence. After labeling testing data set, abnormally labeled regions were tried to be detected by using quasi-supervised learning. If those detected regions match with the abnormally labeled regions, we consider those regions as “true detection” areas and vice versa we consider as “false alarm” areas. The area under receiver operating curve gives the rate of true detections versus false alarms. The most successful texture profiles than determined via ROC curve. Experimental results showed that optimum parameters of the learning algorithm are 64 grey levels, 1 pixel distance neighborhood and $80^{x}80$ pixel block size.

CHAPTER 2

TEXTURE RECOGNITION

2.1. Image Texture

Texture is one of the important characteristics used in identifying objects or regions in images. There are many researchers in image processing and computer vision areas who have considered the concept of feature vectors to cope with texture classification. In texture segmentation, many algorithms partition the image into a set of regions which are visually distinct and uniform with respect to textural properties [9], [10], [11]. In remote sensing radar aplications, texture features have been used to identify forest regions and their boundaries and to identify and analyse variuos crops [12], [13]. In biomedical data analyse, texture features are used for identifying diseases [27], [28], [29]. In industrail vision inspection, texture features have been used to perform the classification of different surface materials [14]. Obviously, there are many other applications in which texture is used to carry-out a recognition or a classifcation task.

2.2. Texture Analysis

Texture analysis is important in many applications of computer image analysis for classification or segmentation of images based on local spatial variations of intensity or color. A successful classification or segmentation requires an efficient description of image texture. Important applications include industrial and biomedical surface inspection, for example for defects and disease, ground classification and segmentation of satellite or aerial imagery, segmentation of textured regions in document analysis, and content-based access to image databases. However, despite many potential areas of application for texture analysis, there are only a limited number of successful examples. A major problem is that textures in the real world are often not uniform, due to changes

in orientation, scale or other visual appearance. In addition, the degree of computational complexity of many of the proposed texture measures is very high.

2.2.1. Texture Segmentation

Texture segmentation is a difficult problem because one usually does not know a priori of what types of textures exist in an image, how many different textures there are, and what regions in the image have which textures. In fact, one does not need to know which specific textures exist in the image in order to do texture segmentation. All that is needed is a way to tell that two textures (usually in adjacent regions of the images) are different. The two general approaches to performing texture segmentation are analogous to methods for image segmentation: region-based approaches or boundary-based approaches. In a region-based approach, one tries to identify regions of the image which have a uniform texture. Pixels or small local regions are merged based on the similarity of some texture property. The regions having different textures are then considered to be segmented regions. This method has the advantage that the boundaries of regions are always closed and therefore, the regions with different textures are always well separated. It has the disadvantage, however, that in many region-based segmentation methods, one has to specify the number of distinct textures present in the image in advance. In addition, thresholds on similarity values are needed. The boundary-based approaches are based on the detection of differences in texture in adjacent regions. Thus boundaries are detected where there are differences in texture. In this method, one does not need to know the number of textured regions in the image in advance. However, the boundaries may have gaps and two regions with different textures are not identified as separate closed regions.

2.2.2. Texture Classification

Texture classification process involves two phases: the learning phase and the recognition phase. In the learning phase, the target is to build a model for the texture content of each texture class present in the training data, which generally comprises of images with known class labels. The texture content of the training images is captured with the chosen texture analysis method, which yields a set of textural features for each

image. These features, which can be scalar numbers or discrete histograms or empirical distributions, characterize given textural properties of the images, such as spatial structure, contrast, roughness, orientation, etc. In the recognition phase the texture content of the unknown sample is first described with the same texture analysis method. Then the textural features of the sample are compared to those of the training images with a classification algorithm, and the sample is assigned to the category with the best match. Optionally, if the best match is not sufficiently good according to some predefined criteria; the unknown sample can be rejected instead.

A wide variety of techniques for describing image texture have been proposed. Texture analysis methods were divided into four categories: statistical, geometrical, model-based and signal processing. In this part, a short introduction will be provided. For surveys on texture analysis methods, Haralick was proposed very usefull textural features [2].

Statistical methods analyze the spatial distribution of gray values, by computing local features at each point in the image, and deriving a set of statistics from the distributions of the local features. Depending on the number of pixels defining the local feature statistical methods can be further classified into first-order (one pixel), second-order (two pixels) and higher-order (three or more pixels) statistics. The basic difference is that first-order statistics estimate properties (e.g. average and variance) of individual pixel values, ignoring the spatial interaction between image pixels, whereas second- and higher-order statistics estimate properties of two or more pixel values occurring at specific locations relative to each other. The most widely used statistical methods are co-occurrence features [1] and gray level differences, which have inspired a variety of modifications later on. Other statistical approaches include autocorrelation function, which has been used for analyzing the regularity and coarseness of texture, and gray level run lengths, but their performance has been found to be relatively poor.

Geometrical methods consider texture to be composed of texture primitives, attempting to describe the primitives and the rules governing their spatial organization. The primitives may be extracted by edge detection with a Laplacian-of-Gaussian or difference-of-Gaussian filter, by adaptive region extraction [18], or by mathematical morphology. Once the primitives have been identified, the analysis is completed either by computing statistics of the primitives or by deciphering the placement rule of the elements [19].

Model-based methods hypothesize the underlying texture process, constructing a parametric generative model, which could have created the observed intensity distribution. Pixel-based models view an image as a collection of pixels, whereas region-based models regard an image as a set of subpatterns placed according to given rules. The observed intensity function is regarded as the output of a transfer function whose input is a sequence of independent random variables, i.e. the observed intensity is a linear combination of intensities in a specific neighborhood plus an additive noise term. Various types of models can be obtained with different neighborhood systems and noise sources. Random field models analyze spatial variations in two dimensions. Global random field models threat the entire image as a realization of a random field, whereas local random field models assume relationships of intensities in small neighborhoods. Widely used classes of local random field models type are Markov random field models, where the conditional probability of the intensity of a given pixel depends only on the intensities of the pixels in its neighborhood. In a Gaussian Markov random field model the intensity of a pixel is a linear combination of the values in its neighborhood plus a correlated noise term. Describing texture with the random field models is an optimization problem, the chosen model is fitted to the image, and an estimation algorithm is used to set the parameters of the model to yield the best fit. The obtained parameter values are then used in further processing, e.g. for segmenting the image. In contrast to autoregressive and Markov models fractals have high power in low frequencies, which enables them to model processes with long periodicities. An interesting property of this model is that fractal dimension is scale invariant. Several methods have been proposed for estimating the fractal dimension of an image.

There exist a number of classification algorithms. Among the most widely used are parametric statistical classifiers derived from the Bayesian decision theory, nonparametric k-nearest neighbor classifier, and various neural networks such as multilayer perceptrons. Given a texture description method, the performance of the method is often demonstrated using a texture classification experiment, which typically comprises of following steps;

- Selection of image data: the image data and textures may be artificial or natural, possibly obtained in a real world application. An important part of the selection of image data is the availability and quality of the ground truth associated with the images.

- Partitioning of the image data into sub images. Image data are often limited in terms of the number of original source images available, hence in order to increase the amount of data the images are divided into sub images, either overlapped or disjoint, of a particular window size.
- Preprocessing of the subimages and dividing available data into training and testing sets.
- Selection of the classification algorithm. In addition to classification algorithm this may involve other selections such as metrics or dissimilarity measures. Selection of classification algorithm can have great impact in the final performance of the texture classification procedure and no classifier can survive with poor features. Also good features can be wasted with poor classifier design.
- Definition of the performance criterion. Determining the proportion of true detections (classification accuracy) or false alarms (classification error) is used as performance criterion.

It is obvious that the final outcome of a texture classification experiment depends on numerous factors, both in terms of the possible built-in parameters in the texture description algorithm and the various choices in the experimental setup. Results of texture classification experiments have always been suspect to dependence on individual choices in image acquisition, preprocessing, sampling etc.

CHAPTER 3

PROBLEM DESCRIPTION AND PROPOSED METHOD

3.1. Problem Description

Today air reconnaissence efforts constitute the backbone of the military intelligence. Many countries use reconnaissance and surveillance aircrafts for military purposes. In addition, these aircrafts are used in many countries for civilian purposes too. But especially they are used for border surveillance (patrolling) or prevention of smuggling and illegal migrations. A photo reconnaissance aircraft has no armament and does not necessarily require high performance capacity. High resolution aerial images are available with the state-of-the-art aerial imaging technologies. Intelligence specialists try to find possible threats on these aerial images by visual observation or other detection methods. They search for the clues that prove the enemy activities or potential enemy.

Currently, unmanned reconnaissance aircrafts capture air images and transmit the aerial data to the hub. Experts in the hub scan the aerial data and search for anything unnatural. Without any machine learning application, this process is very exhaustive and it is a time consuming jop. The application of machine learning techniques to aerial data can be a usefull method in detecting human existence on the air photographs. Both supervised or unsupervised learning techniques can be used in solving human-existence detection problem on the air photographs. In aerial image reconnaissance tasks we search everything that is unnatural and proves human existence on images. From this point of view supervised learning algorithms are not suitable for this task. Because supervised learning needs pre-determined classes and the definition of a certain segment of data. However, in aerial images, we do not search certain shapes we look for anything that proves human existence. For unsupervised learning applications the target variable is unknown or has only been recorded for too small a number of cases. So, unsupervised learning or quasi-supervised learning is suitable for aerial reconnaissance task.

3.2. Problem Solving

In thesis, a quasi-supervised learning algorithm was used to recognize the abnormally defined regions that proves human existence. The specific implementation was constructed as follows: First, aerial image was divided into a reference control data set and a mixed testing data set. Eigthteen aerial images were extracted from two big aerial images. Nine of these images were belong to control data set and the other nine were belong to the test data set. In the second step, all images were then divided into 625 non-overlapping pixel blocks. In the third step, grey level co-occurrence matrices were computed from each block. From these matrices, several Haralick features were calculated and organized into texture vectors to be used for recognition. 42 texture profiles were generated by changing the block sizes and distace parameter. Finally, the quasi-supervised learning was applied to the collection of texture vectors to recognize the blocks which consist man-made structures and the most successful system parameters were determined by using ROC curve. The specific implementation will be explained in chapter 4.

3.3. Grey Level Co-occurrence Matrices (GLCM)

For extracting the textural information of a grey tone image the grey level co-occurrence matrices (also called the Grey Tone Spatial Dependency Matrix) are one of best known texture analyse methods in the literature. The studies showed that statistical computations on grey levels of images were able to give usefull descriptors of perceptual feeling of texture [1], [2]. Suppose that we have an n×m sized image to be analysed and gray tone appearing in each resolution cell is quantized to some levels. We make a gray tone comparison of each resolution cell to it is “d” distance pixel neigbours. There are 4 possible angular neigbourhoods.

Matematically a Co-occurrence matrice **C** is defined over an **n x m** image **I**, parameterized by an offset **(Δx, Δy)**, as:

$$\mathrm{C}\,\Delta x\,\mathrm{C}\,\Delta y\,(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, \text{ if } I(p,q) = \text{i and } I(p+\Delta d, q+\Delta d) = j \\ 0, \text{ otherwise} \end{cases} \qquad (3.1)$$

The grey tone of the resolution cell is compared with the "Δd" distance and "0º, 45º, 90º, 135º" degrees neighbours' grey tone. The above function takes the value of "1" if the argument is true and takes "0" otherwise. It is possible to generate a set of different co-occurrence matrices from the same image by changing distance parameter and angular neigborhood. The value of the image is a grayscale value quantized to some grey level. The GLCM is a tabulation of how often different combinations of pixel brightness values (grey levels) occur in an image. If the quantization level is N, than we will have a N×N dimensional co-occurrence matrix. Symmetrical property is an innate property of GLCM. Symetric matrix means that the same values occur in cells on opposite sides of the diagonal. This property and computation of a GLCM will be presented with an example below.

3.3.1 Grey Level Co-occurrence Matrice Example

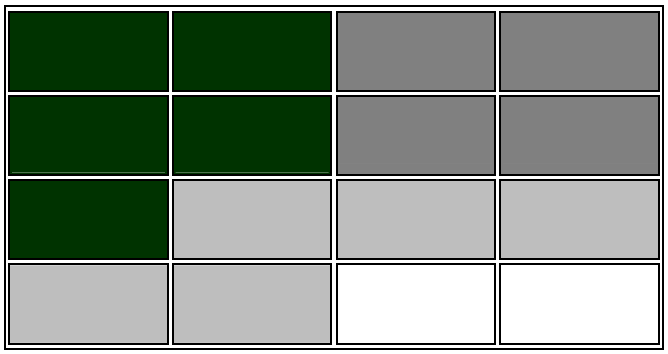


Figure 3.1. Sample Image, 4grey levels

Suppose that we have a sample image which was quantised to 4 grey levels and it's grey levels are:

Table 3.1. Grey Levels of Sample Image

0	**0**	**1**	**1**
0	**0**	**1**	**1**
0	**2**	**2**	**2**
2	**2**	**3**	**3**

Table 3.2. A Grey Level Co-occurrence Matrice Table

Neighbour pixel value → Reference pixel value	0	1	2	3
0	0,0	0,1	0,2	0,3
1	1,0	1,1	1,2	1,3
2	2,0	2,1	2,2	2,3
3	3,0	3,1	3,2	3,3

Table 3.3. 0º and 1 pixel distance Grey Level Co-occurrence Matrice Table

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

Table 3.4. 90º and 1 pixel distance Grey Level Co-occurrence Matrice Table

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

Table 3.5. 45º and 1 pixel distance Grey Level Co-occurrence Matrice Table

4	1	0	0
1	2	2	0
0	2	4	1
0	0	1	0

Table 3.6. 135º and 1 pixel distance Grey Level Co-occurrence Matrice Table

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

3.4. Quasi Supervised Statistical Learning Method

Supervised learning applications requires the definition of a certain segment of data. The ground truth data set are available in some cases but in aerial reconnaissance tasks target variables are not known. An alternative strategy can be used in aerail image reconnaissance tasks. In this thesis, the strategy used for recognitioning the man-made objects is a quasi-supervised stastical learning. The method used for quasi-supervised learning can be explained as follows: Available data is divided into two groups, one of which is known to be free of the objects of detection, and the other containing the objects of detection along with features of normalcy commonly shared with the first dataset. The objects of detection in the aerial image reconnaissance tasks are usually man-made structures or specific abnormalities on the ground. The first dataset can be referred to as the reference control dataset, while the second as the mixed testing dataset. Such a scenario describes a quasi-supervised learning setting that requires the prior knowledge of only the presence or absence of abnormalities in the respective datasets and not the labeling of individual samples. Since abnormal regions are unique to the testing data and do not exist in the reference control data, we expect the learning algorithm to detect the regions specific to testing data. The approach uses the ratio of the number of times a given block is assigned to the reference control and mixed testing datasets through the course of successive nearest neighbor classifications on it's tecture profile with randomly assembled reference sets as an estimate of the posterior probability of the respective classes for that block.

A reference set is generated with "2n" elements, n of them is taken from the control data set and n of them is taken from the testing data. The point "x" is assigned to the label of it's nearest neighbor. This classification is done repeatedly for N times.

After N times of nearest neighbor classification the posterior probabilities of the point is estimated. R = {xi, yi} "xi" represents the point, "yi" represents the class label (0 for control data and 1 for test data). Nearest neigbor classifier is defined by:

$$F_r(x) = y_i \text{ With } i = \arg\min i = 1,2,3,....l, \quad d(x, x_i) \tag{3.2}$$

Let the $\{R_x\}, x = 1,2,.....N(iid.)$ be a reference set consist of equal elements from each data set. Let control data set labeled with "0" (class 0) and test data set labeled with "1" (class 1). We have previosly mentioned the nearest neighbor classifier.

$$f0(x) = \frac{\sum_{j=1}^{N} 1(Fr(x)) = 0}{N} \tag{3.3}$$

$$f1(x) = \frac{\sum_{j=1}^{N} 1(Fr(x)) = 1}{N} \tag{3.4}$$

Equation (3.3) and Equation (3.4) takes the value of "1" when the inside argument is true and takes "0" otherwise. These two values estimates the class conditional probability densities for "class 0" and "class 1" respectively. The probability of assigning a point "x" to either of the two classes by a nearest neighbor classifier is directly proportional to the number of points of each class in a neighborhood of x. Supposing n points from each class are included in the reference set each time, the total number of distinct reference sets is the combination of all possible sets. Implementation of all possible sets is well beyond today's computation ability. But it is still possible to compute the average number times a given point would be assigned to either class. As a result quasi-supervised learning estimates the posterior probability of a given point "x" by the help of a reference set which consist of equally represented elements from each classes.

CHAPTER 4

IMPLEMENTATION PART

4.1. Introduction to Specific Implementation

First, aerial images were divided into two groups: one is the reference control data set and the other one is the mixed testing data set. All images were than divided into small non-overlapping blocks and texture features were computed from those blocks. In the classification part, habitated regions were recognized by using the quasi-supervised statistical learning algorithm. The most important advantage of this algorithm is manual segmentation of regions is not needed in learning phase. All the information required is the existence of normal and abnormal profiles in each image. Finally the most succesfull feature profile was determined with the performance evaluation metods. In this chapter the specific implementation of the quasi-supervised learning to the aerial images will be presented.

4.1.1. Materials

The images used in this thesis was provided from a reconnaissance aircraft belonging to the Turkish Armed Forces. All the images used in the experiments have the same resolution of 43 cm. per pixel. These images are usually used for mapping or geolocical tasks. In thesis, aerial data was divided into two classification group as mentioned beforehand. First group of images have natural characteristics and represents the normalcy conditions which is known to be free of the objects of detection, and the second group of images have residential areas, some roads, cultivated soil and man-made structures defining the habitation.

There are nine control images and nine test images of different sizes, ranging $175^{x}175$ pixels to $2000^{x}2000$ pixels. Also nine different size of blocks were used in the experiments. The image sizes and the block sizes are seen on the table below.

Table 4.1. Image and Block Sizes

Image size (pixel)	Block size (pixel)	Block Area (m²)
175ˣ175	7ˣ7	9m²
250ˣ250	10ˣ10	18.5m²
350ˣ350	14ˣ14	36.2m²
500ˣ500	20ˣ20	74m²
700ˣ700	28ˣ28	145m²
1000ˣ1000	40ˣ40	296m²
1400ˣ1400	56ˣ56	580m²
1575ˣ1575	63ˣ63	734m²
2000ˣ2000	80ˣ80	1183m²

4.1.1.1. Control Images

In experiments, totally 18 different images were used. First nine of the aerial images were belong to the **reference control data set,** representing the natural terrestrail conditions. Resolution of the control images are “43cm.” per pixel.

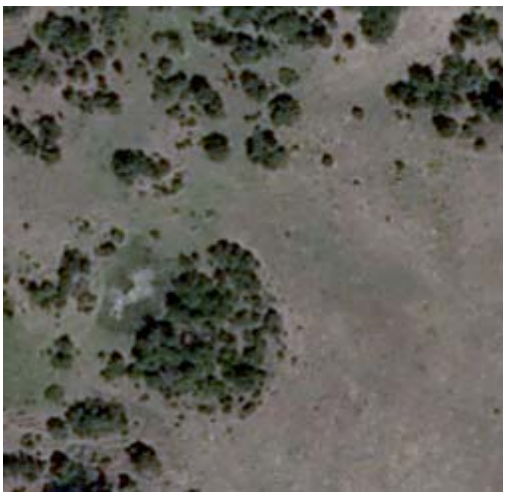


Figure 4.1. Control image, represents the normalcy terrestrial conditions. Size of 250ˣ250 pixels.

Figure 4.2. Control image, represents the normalcy terrestrial conditions. Size of 350^{x}350 pixels.

Control data set consist of nine aerial images. Each images were divided into 625 non-overlaping blocks. The block sizes used in experiments were 7^{x}7, 10^{x}10, 14^{x}14, 20^{x}20, 28^{x}28, 40^{x}40, 56^{x}56, 63^{x}63, 80^{x}80 pixels respectively. Proportional to block sizes, control image sizes chance in a range of 175^{x}175, 250^{x}250, 350^{x}350, 500^{x}500, 700^{x}700, 1000^{x}1000, 1400^{x}1400, 1575^{x}1575, 2000^{x}2000 pixels.

4.1.1.2. Test Images

Nine test images were used in the experiments. Resolution of the test images are “43cm.” per pixel. These images had the same sizes with control images. Test images consist of man-made structures along with unhabitated land. Man-made structures and the elements of habitated land were constituted the objects of detections in recognitioning human existence in aerial images.

Figure 4.3. Test image, consist of both natural areas and man-made structures, 350^{x}350 pixels

Figure 4.4. Test image, consist of both natural areas and man-made structures, 500^{x}500 pixels

4.1.2. Specific Implementation

In the first step, all images were quantised to 64 grey levels. Than images were divided into non-overlaping blocks as seen in Figure 4.5. In order to divide each images into 625 blocks nine different size of blocks were used. For example, 175x175 pixels sized image was divided into 7x7 pixels blocks. So 25 blocks were extracted in vertical axis, and 25 blocks were extracted in horizontal axis.

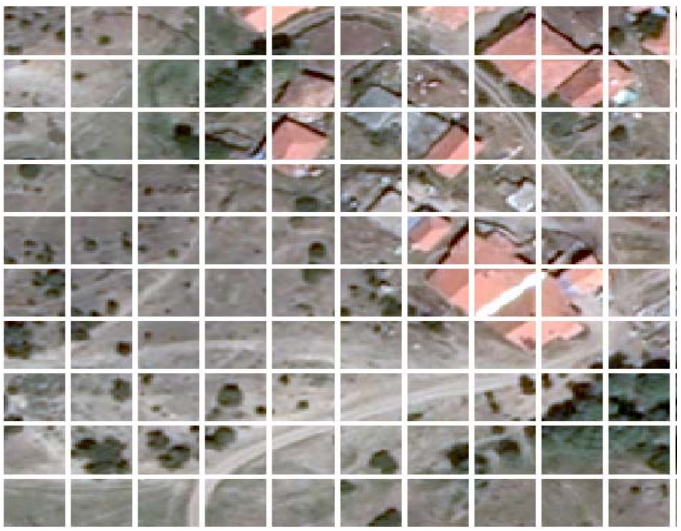


Figure 4.5. Test image, size 250x250 pixels. Control images and Test images divided into 100 non-overlapping blocks.

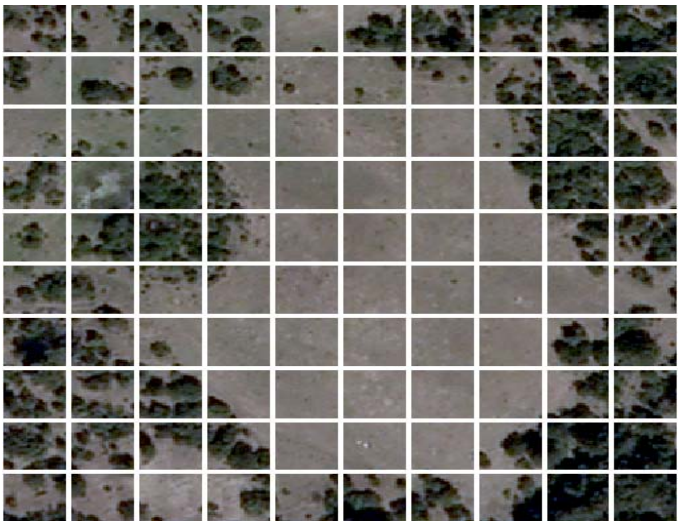


Figure 4.6. Control image, size 250x250 pixels. Control images and Test images divided into 100 non-overlapping blocks.

In the second step, the grey level co-occurrence matrices were computed and from these matrices four Haralick features were calculated and organized as a texture vector in order to recognize the man-made structures in the test images. Texture computation process will be defined with an example:

Figure 4.7 First block is represented with a blue colored grid.

Suppose that in our example we have the block size of 7^x7 pixel. First block, (first row first and first column element) represented with a **blue colored block** is the first element of the computation. Grey levels of that block is seen on matrix "B", size of 7^x7 pixels:

$$B = \begin{bmatrix} 119 & 116 & 86 & 50 & 75 & 119 & 146 \\ 105 & 111 & 82 & 62 & 72 & 107 & 131 \\ 103 & 103 & 90 & 91 & 99 & 108 & 119 \\ 102 & 111 & 119 & 122 & 127 & 136 & 137 \\ 113 & 125 & 116 & 121 & 134 & 139 & 140 \\ 127 & 106 & 63 & 47 & 77 & 108 & 128 \\ 135 & 90 & 37 & 25 & 56 & 95 & 113 \end{bmatrix}$$

$$G1 = \begin{bmatrix} 29 & 29 & 21 & 12 & 19 & 29 & 36 \\ 26 & 27 & 20 & 15 & 18 & 26 & 32 \\ 25 & 25 & 22 & 22 & 24 & 27 & 29 \\ 25 & 27 & 29 & 30 & 31 & 34 & 34 \\ 28 & 31 & 29 & 30 & 33 & 34 & 35 \\ 31 & 26 & 16 & 12 & 19 & 27 & 32 \\ 33 & 22 & 9 & 6 & 14 & 23 & 28 \end{bmatrix}$$

First block was then converted to 64 grey level image. And matrix G1 represents the 64 grey level image. Matrix G1 was generated after quantization of matrix B to 64 grey levels. We have previosly mentioned that co-occurrence matrices dimention is determined by the number of the grey levels of an image. So in this example, a 64^x64 dimensional grey level co-occurrence matrice was computed. In order to illustrate the grey level co-occurrence matrice, a smaller sized matrice was then generated by quantization of matrice B into 10 grey levels. Matrice G2 represents the 10 grey leveled image.

$$G2 = \begin{bmatrix} 4 & 4 & 3 & 2 & 3 & 4 & 5 \\ 4 & 4 & 3 & 2 & 3 & 4 & 5 \\ 4 & 4 & 3 & 3 & 3 & 4 & 4 \\ 4 & 4 & 4 & 4 & 4 & 5 & 5 \\ 4 & 4 & 4 & 4 & 5 & 5 & 5 \\ 4 & 4 & 2 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 1 & 2 & 3 & 4 \end{bmatrix}$$

Grey level co-occurrence matrices which were computed from the matrice G2 are seen below:

GLCM_0º =

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 6 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 6 & 4 & 8 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 8 & 24 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 & 6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

GLCM_45º =

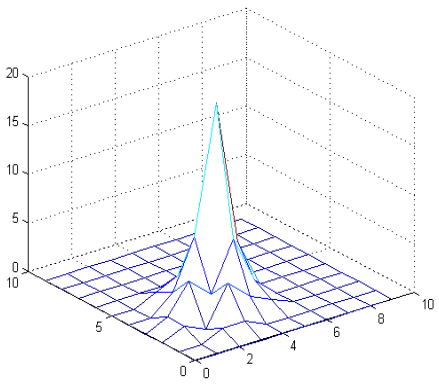
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 4 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 2 & 7 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 7 & 20 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 5 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

GLCM_90º =

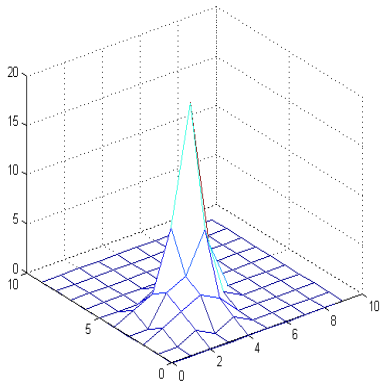
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 8 & 5 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 5 & 28 & 7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 7 & 8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

GLCM_135º=

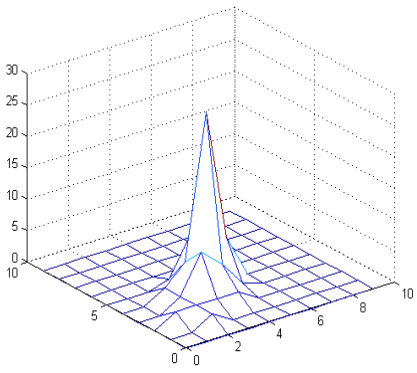
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 4 & 8 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 8 & 20 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



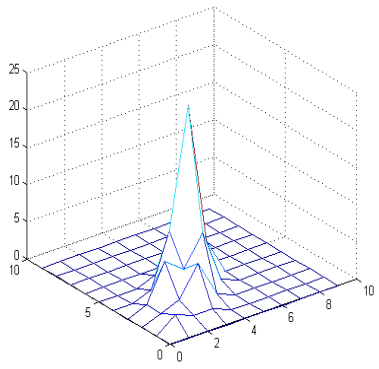
(1) 45º Degree



(2) 135º Degree



(3) 90º Degree



(4) 0º Degree

Figure 4.8. 0º, 45º, 90º, 135º degree and 1 pixel distance neighborhood GLCM of G2

Table 4.2. Four direction (0º- 45º- 90º- 135º) GLCM features.

Contrast	0.7619	Contrast	1.2778	Contrast	0.6667	Contrast	1.0556
Correlation	0.6403	Correlation	0.3296	Correlation	0.6327	Correlation	0.4164
Entropy	0.1267	Entropy	0.1200	Entropy	0.1545	Entropy	0.1246
Homogeneity	0.7143	Homogeneity	0.6505	Homogeneity	0.7619	Homogeneity	0.6875
0º degree features		45º degree features		90º degree features		135º degree features	

Table 4.3. Feature vector.

0.7619	Contrast	0º degree
0.6403	Correlation	
0.1267	Entropy	
0.7143	Homogeneity	
1.2778	Contrast	45º degree
0.3296	Correlation	
0.1200	Entropy	
0.6505	Homogeneity	
0.6667	Contrast	90º degree
0.6327	Correlation	
0.1545	Entropy	
0.7619	Homogeneity	
1.0556	Contrast	135º degree
0.4164	Correlation	
0.1246	Entropy	
0.6875	Homogeneity	

This vector represents the texture profile of the first block which was computed from four direction co-occurrence matrices and four of Haralick features. The test images and control images, firstly splited into 625 blocks. This texture vector was constructed from one pixel distance neigbourhood grey level co-occurrence matrice. And the same procedure was done with the distances of three, five, eight and ten pixel

neighbourhood grey level matrices. At the end of these exhaustive computation a set of texture profiles were collected. 42 different textural features were tested in the experiments for determining the optimal system parameters. These textural vectors were generating by using different distances of neighbourhoods and different sizes of blocks. All textural features are seen on the table below.

Table 4.4. Textural feature vectors used in experiments.

Distance / **Block size**	1pixel distance neighborhood GLCM	3 pixel distance neighborhood GLCM	5 pixel distance neighborhood GLCM	8 pixel distance neighborhood GLCM	10 pixel distance neigborhod GLCM
7x7	x	x	x	-	-
10x10	x	x	x	x	-
14x14	x	x	x	x	x
20x20	x	x	x	x	x
28x28	x	x	x	x	x
40x40	x	x	x	x	x
56x56	x	x	x	x	x
63x63	x	x	x	x	x
80x80	x	x	x	x	x

After generating the raw feature vectors, mean-variance normalization was then carried-out along the feature vectors. It is one of the most common approaches for feature normalization, especially when close to Gaussian distribution is assumed. It is subtraction of the population mean and scaling to achieve unit variance is seen on the equation 4.1. The $F(i)$ is the raw value of the i′th feature, $\mu_{F(i)}$ is the feature mean, $\sigma_{F(i)}$ is the standart deviation and $F'(i)$ is the normalized feature vector.

$$F'(i) = \frac{F(i) - \mu_{F(i)}}{\sigma_F(i)} \tag{4.1}$$

4.2. Labeling the Test Images

In order to evaluate the detection performance of the learning algorithm with different texture profiles, we need to determine the abnormal blocks manually beforehand. In our experiments there are nine different block size as mentioned before. These blocks and images are illustrated on figures below;

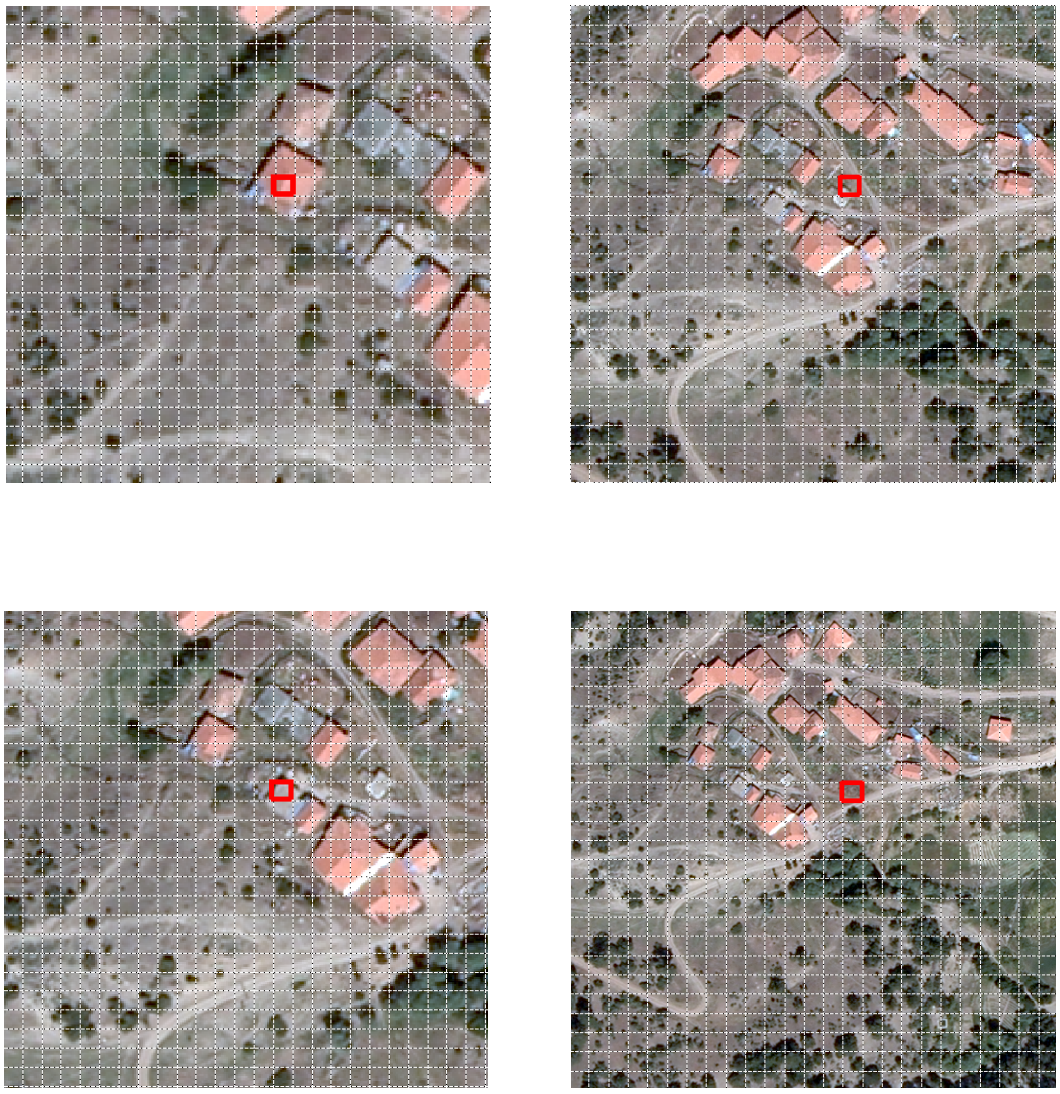


Figure 4.9. Test images and 7, 10, 14 and 20 pixel sized blocks.

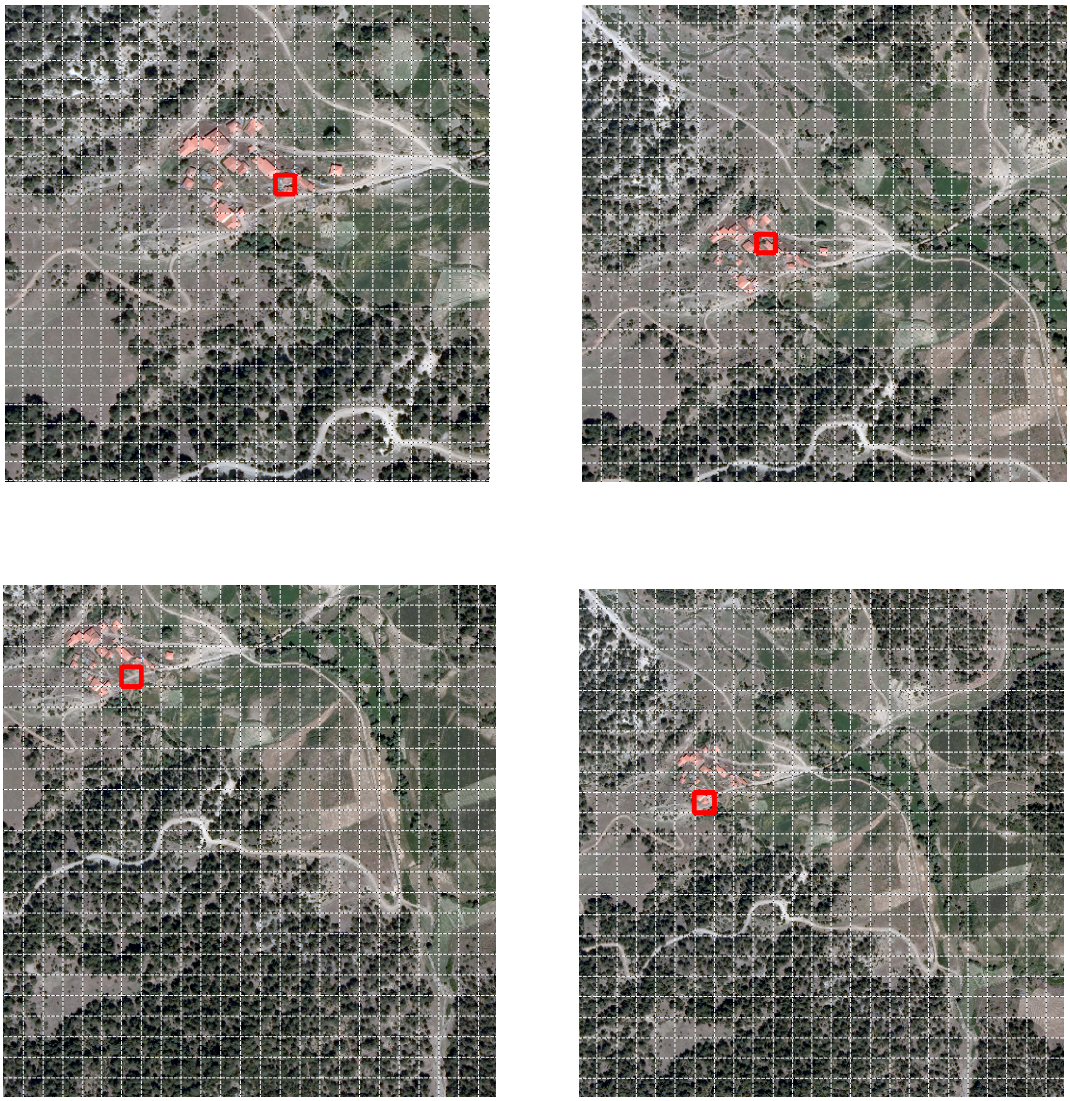


Figure 4.9. Test images and 28, 40, 63, 80 pixel sized blocks.

Every test image and each block was labeled as normal and abnormal with manually evaluation. The regions that we consider as abnormal are the blocks consist of man-made objects. There are two scenario of labeling a block as abnormal, in the first situation; man-made objects constitute the full area of the blocks or the majority of the block area, in the second situation; the object constitutes only a small part of the related block. It is a dilemma whether or not these small parts of structures is enough to determine a block as abnormal. If we consider these blocks as normal it wouldn't be a convinient decision, because these blocks have abnormal textural features too. On the other hand, we can not estimate the effect of these abnormalities to the texture, this would put the classifier under heavy constrain. As a result in the labeling strategy, the

blocks which have man-made structures was considered as abnormal regions. After classification, we expect the learning algorithm to recognize those areas. In the figure below the blue colored blocks represent the abnormally labeled regions.

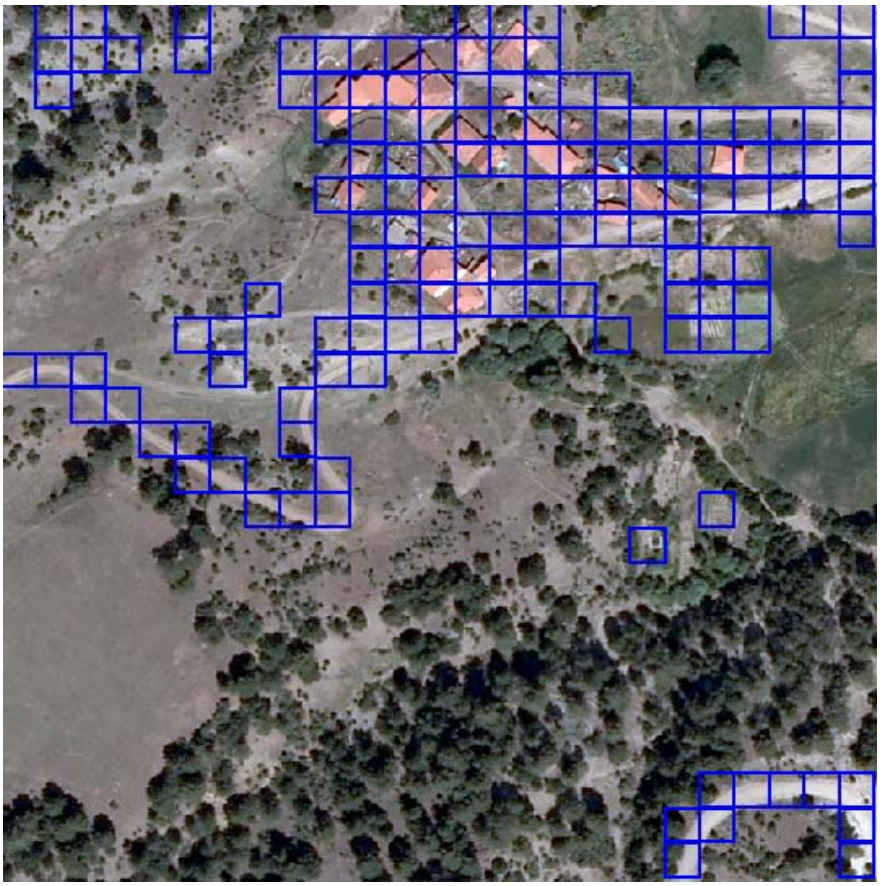


Figure 4.11. True detection ares, abnormally labeled grids.

Blue colored blocks represent the abnormally labeled regions. These regions have parts of foothpaths, roads, plowed land and buildings. The recognition performance of each texture profile was evaluated against manual labeling for determining the most successful texture vector and optimal system parameters.

4.2.1. True Detections and False Alarms

The aim of the quasi-supervised learning algorithm is to recognize the blocks considering man-made structures or objects which do not exist on unhabitated lands. The success of the learning algortihm was evaluated with the number of true detections

and false alarms. True detections are the regions where both learning algorithm and manually labeled data approves the abnormality. False alarms are the regions where learning algorithm finds a normalcy area as abnormal. The number of true detections and false alarms give usefull informations about the success of a specific texture profile. Reciever operating characteristics curve was generated by using the ratio of true detections versus false alarms. The area under reciever operating characteristics curve yield the performance evalutaion ratio of texture profiles.

.

CHAPTER 5

EXPERIMENTAL RESULTS

5.1. Detection Performance

The performance of the learning algorithm with different textural parameters was evaluated by the experimental results in this paper. Every man-made structures and objects of human existence on aerial images were expected to recognize with learning algorithm. In this chapter, the optimal system parameters are determined and the success of the learning algorithm under given textural properties will be defined.

5.1.1. Optimal Block Size

Experiments were carried-out with nine different block sizes. These blocks have the size of 7^x7, 10^x10, 14^x14, 20^x20 ,28^x28, 40^x40, 56^x56, 63^x63, 80^x80 pixels respectively. This also means that each block has an area of 9m², 18.5m², 36.2m², 74m², 145m², 296m², 580m², 734m², and 1183m² on land respectively. In the labeling part we have mentioned that the blocks which consider the objects of detections were labeled as abnormal regions, no matter how big the area of the object is.

5.1.1.1. Performance of 7^x7 Pixel (9 m²) Block Size

175^x175 pixel sized test image and control images were divided into 7^x7 pixel blocks. All the blocks in the test image were labeled as normal or abnormal beforehand and the labeled data is seen on the figure below as blue colored blocks. Texture vector was computed by one pixel distance grey level co-occurrence matrices. Than quasi-supervised learning algorithm was implemented on texture vectors and the regions of interest in aerial images were recognized. Detection performance was assessed with reciever operating characteristics curve.

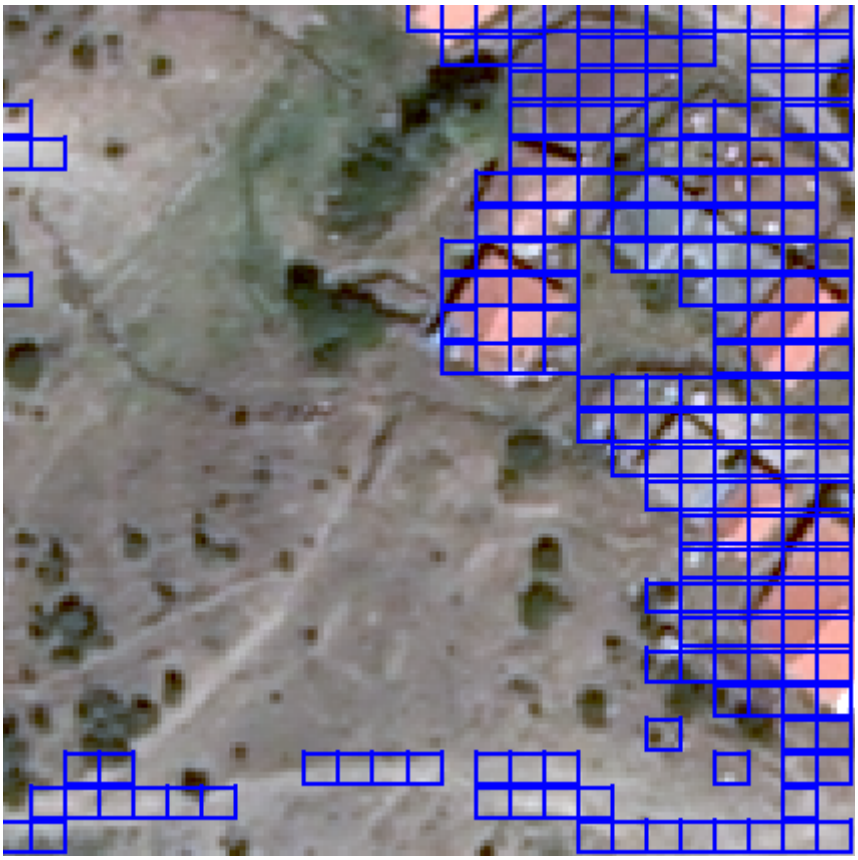


Figure 5.1. True detection ares, labeled grids for 175x175 pixel.

At the end of the recognition process, posterior probabilities of each block was computed in order to assign the related block to class 1 or class 0. $F0(i)$ represents the probability of assigning the i'th block to class 0, and $F1(i)$ represents the probabiltiy of assigning i'th block to class 1. We will decide labels of each block according to the comparison rule given below;

$$\text{If} \quad \begin{array}{ll} F1(i) \geq \text{Treshold Value}, & \text{assign class 1} \\ F1(i) \leq \text{Treshold Value}, & \text{assign class 0} \\ F0(i) \geq \text{Treshold Value}, & \text{assign class 0} \\ F0(i) \leq \text{Treshold Value}, & \text{assign class 1} \end{array} \tag{5.1}$$

This comparison rule simply defines that if the block's class probability is greater than the treshold value, related block will be assigned to the related class. Optimal treshold value is determined according to the number of true detections and the

number of false alarms. Optimal treshold value was selected as the sharpest point of reciever operating characteristics curve where true detection rate is optimal and false alarm rate is minimum.

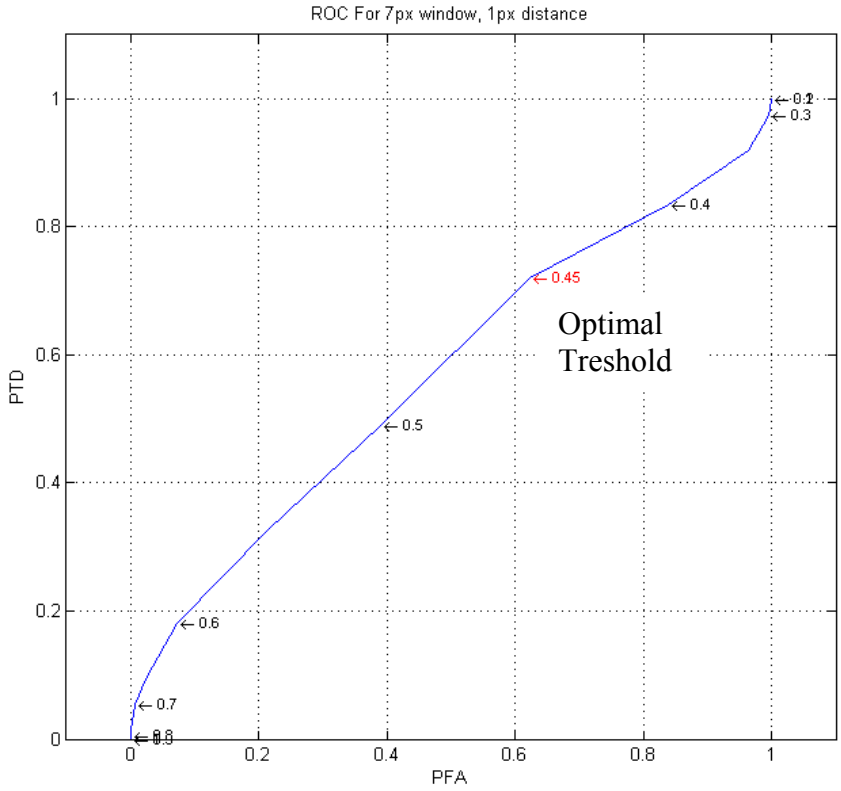


Figure 5.2. ROC curves for 7x7 pixel sized block and 1 pixel distances.

Reciever operating characteristics curve represents the detection performance of the learning algorithm with texture profile of 7x7 pixel block sizes and 1 pixel distance GLCM. The red colored treshold value (0,45) was considered as the optimal treshold, it is the sharpest point of reciever operating characteristics curve. Even the with the optimal treshold value, the learning algorithm could find only 72 percent of true detection areas. There are 200 blocks of true detection regions and 425 blocks of false error regions. With the optimal treshold value, learning algorithm was able to detect 144 ares of true detections.

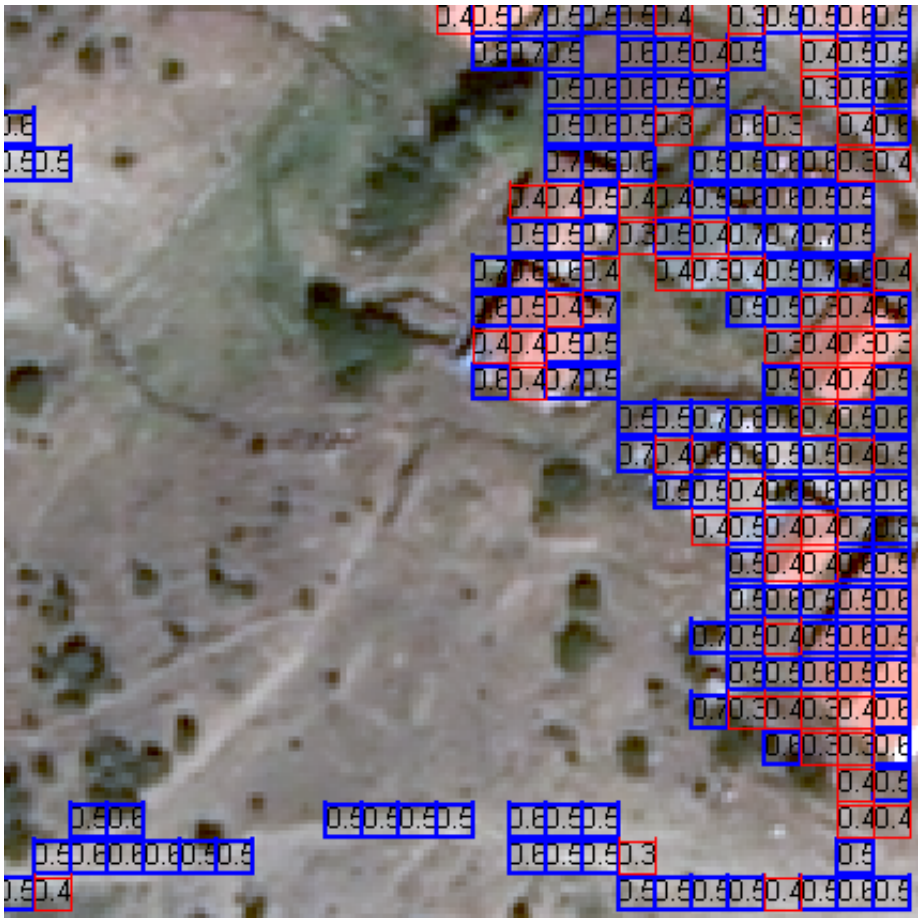


Figure 5.3. Blue grids and Red grids.

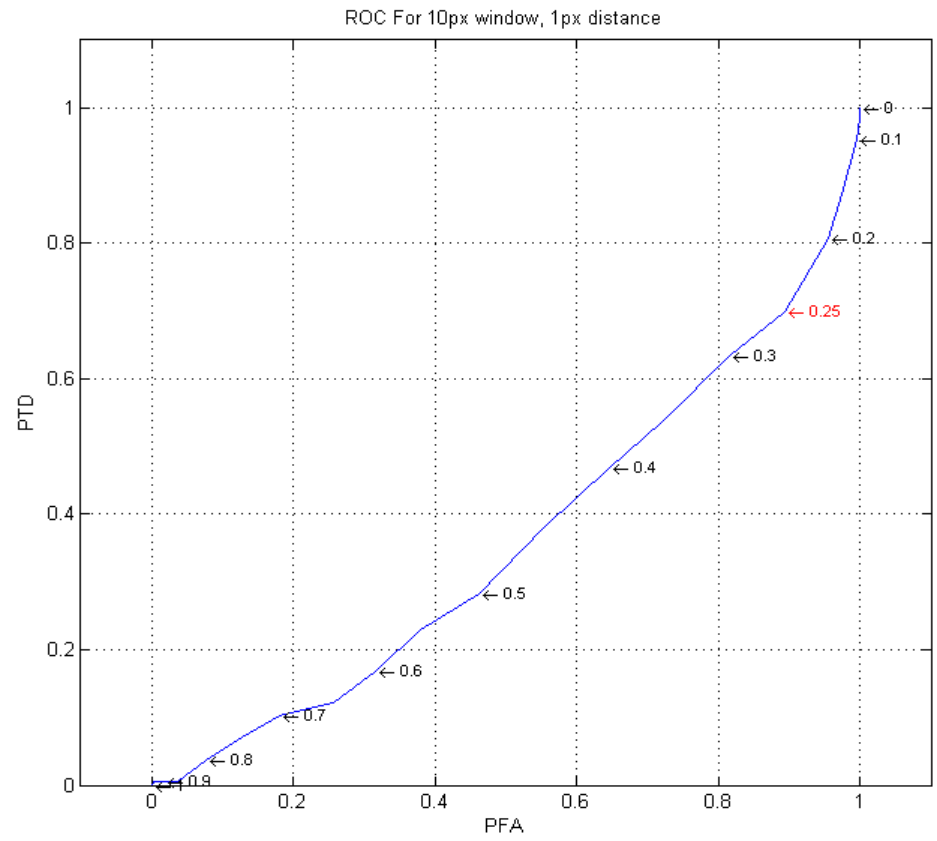
Figure 5.3. represents the detection performance with the texture profile of one pixel distance grey level co-occurrence matrice and 7x7 pixel blocks. Blue colored blocks are used to illustrate the regions where classifier had succeded and red colored blocks are used to illustrate the regions where classifier could not detected with the treshold value of 0,45. Important point in performance evaluation of the texture vector of 7x7 pixel blocks, learning algorithm was detected too many false regions.

5.1.1.2. Performance of 10x10 Pixel (18.5 m²) Block Size

250x250 pixel sized test image and control image was used in the experiment. Both images were divided into 10x10 pixel blocks. Texture vectors were computed from one pixel distance grey level co-occurrence matrice. The labeled data is seen as blue blocks on the figure below.

Figure 5.4. Blue colored blocks, the labeled data for 250x250 pixel sized test image.

The blue colored places are the regions of interest, we expect the classifier to find these blocks. There are 213 blocks marked as abnormal region. As seen on the figure 5.5. and figure 5.6. The textural feature of 10x10 pixel block is not succesfull and and even worse than 7x7 pixel sized block. Another textural feature which was used in this experiment was calculated from the 10x10 pixel blocks and three pixel distance neighbourhood grey level co-occurrence matrices. This feature profile had given a worse detection result than the feature of one pixel distance neighbourhood grey level co-occurrence matrice. Also another experiment was carried-out along the five pixel distance neighbourhood grey level co-occurrence matrices, this two textural profiles yielded that one pixel distance neighbourhood is the most informative textural property.



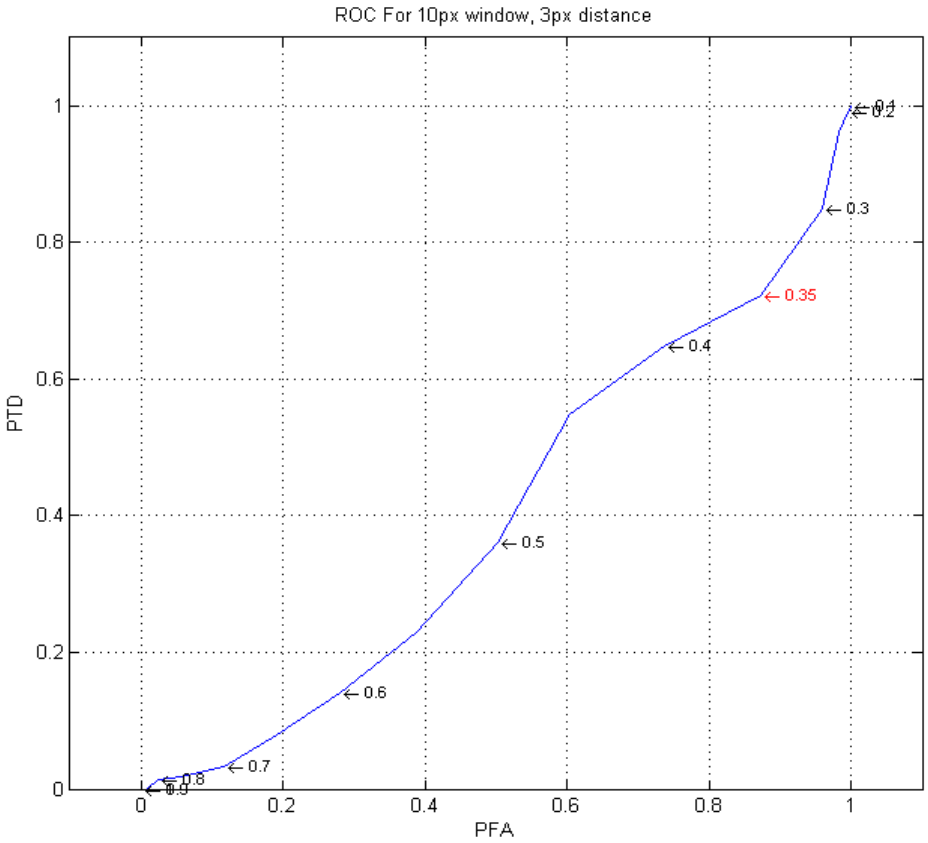


Figure 5.5. ROC curves for 10x10 pixel sized block and 1, 3 pixel distances

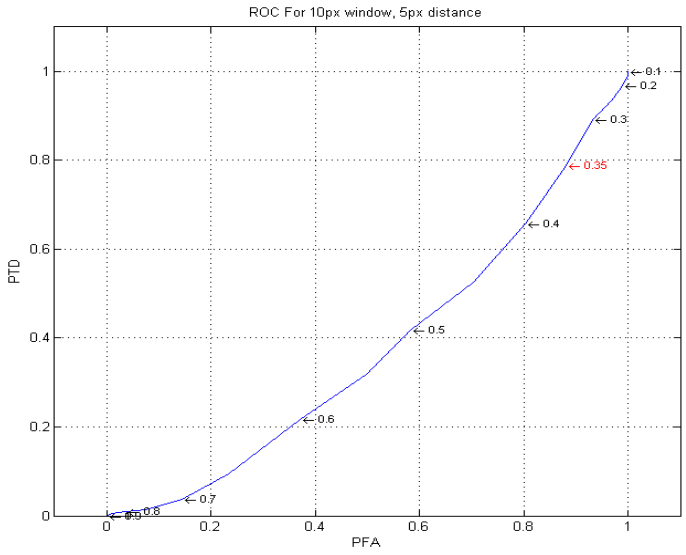


Figure 5.6. ROC curves for 10x10 pixel sized block and 5 pixel distances

5.1.1.3. Performance of 14x14 Pixel (36 m²) Block Size

350x350 pixel sized test image and control image was used in the experiment and both images were divided into 14x14 pixel blocks. The labeled data for 350x350 pixel sized test image is seen on the figure below. There are 246 abnormal places and 379 normal places.

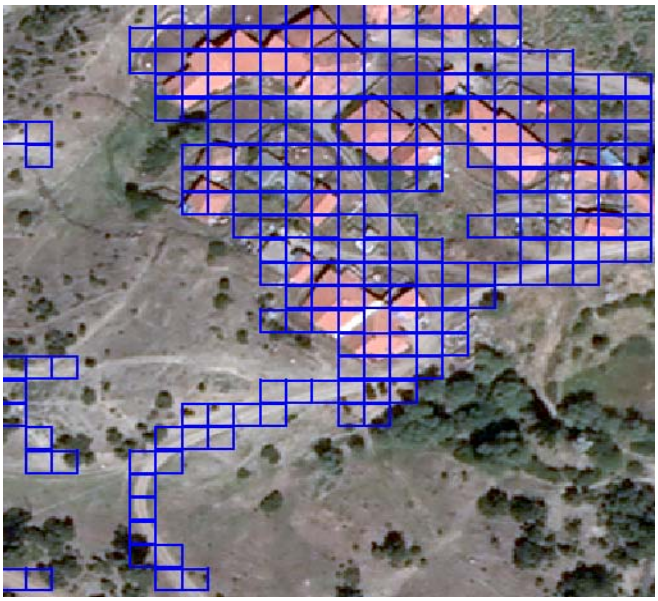


Figure 5.7. Blue colored blocks, the labeled data for 350x350 pixel sized test image

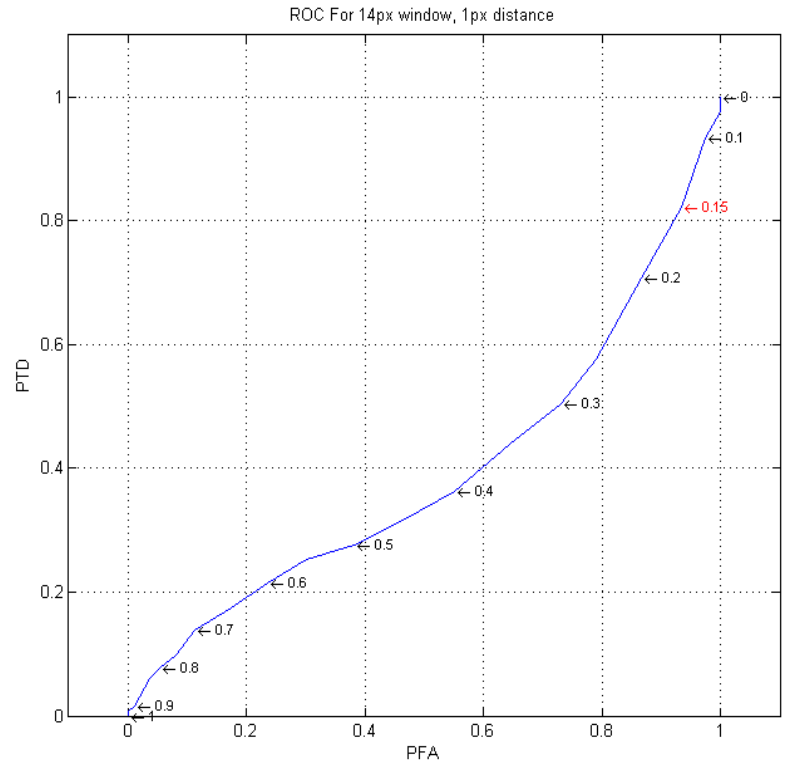


Figure 5.8. ROC curves for 14x14 pixel sized block and 1 pixel distances

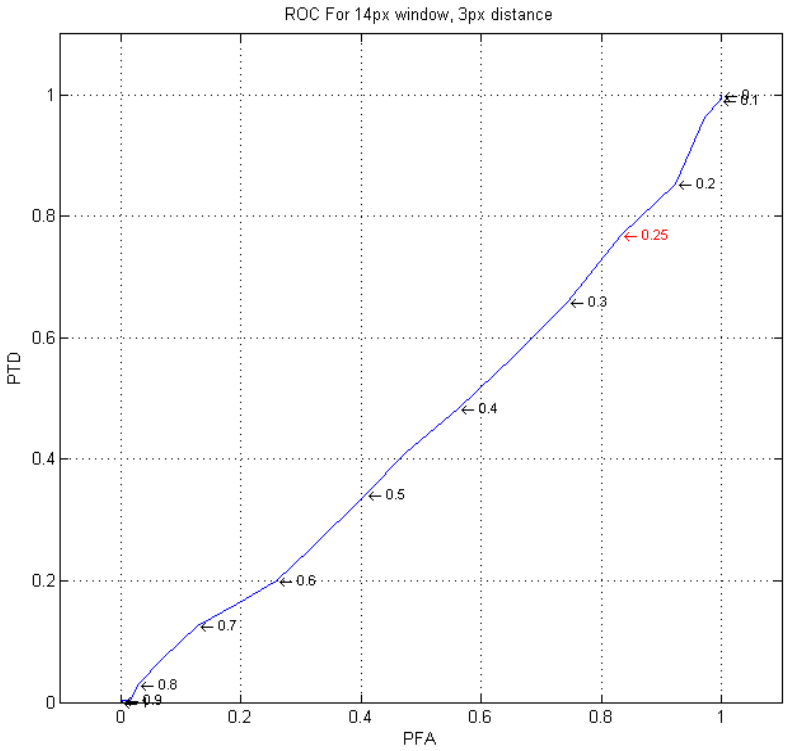


Figure 5.9. ROC curves for 14x14 pixel sized block and 3 pixel distances.

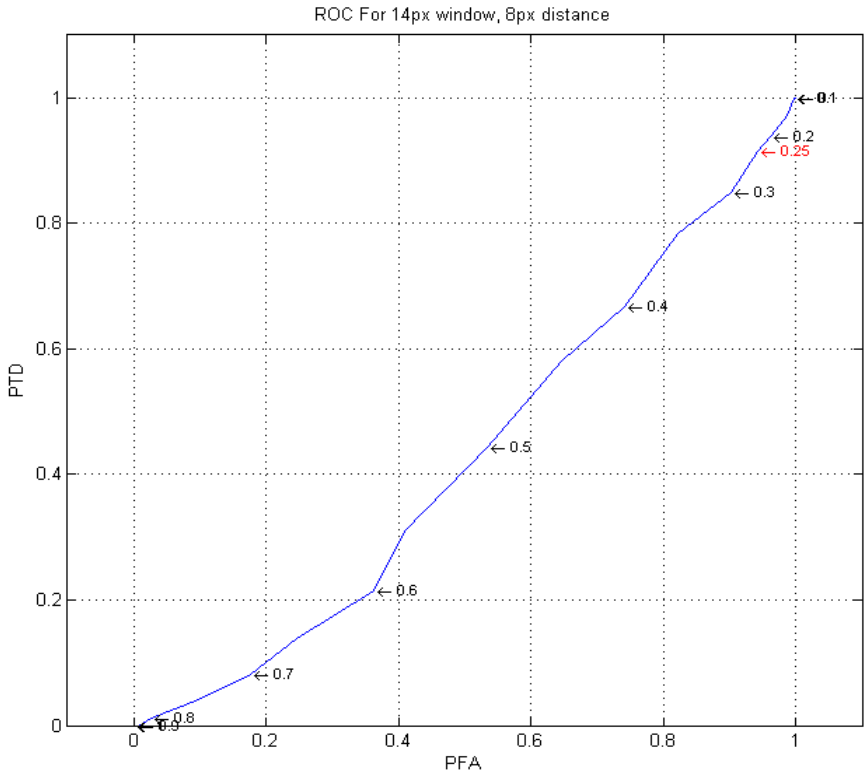


Figure 5.10. ROC curves for 14x14 pixel sized block and 8 pixel distances.

Even with a suitable treshold value, only 80 percent of abnormal regions were able to detected and learning algorithm had given too many false alarms with the 14x14 pixel blocks textural properties.

5.1.1.4. Performance of 20x20 Pixel (74 m²) Block Size

500x500 pixel sized test image and control image was used in the experiment and both images were divided into 20x20 pixel blocks. The test image was labeled manually. As a result, 203 regions were labeled as the regions of interest and the other 422 blocks were labeled as normal.

This block size has an area of 74 m² on land. The aim of using this block size was to detect the vehichles and some structures which have the size of five meters. But experimental results showed that the block size of 20x20 pixel is not suitable for recognitioning the materials like small vehicles and other structures. Detection performance of 20x20 pixel block was better than the smaller sized blocks. The

performance of 20x20 pixel block is seen with the reciever operating characteristics curve on the figures below.

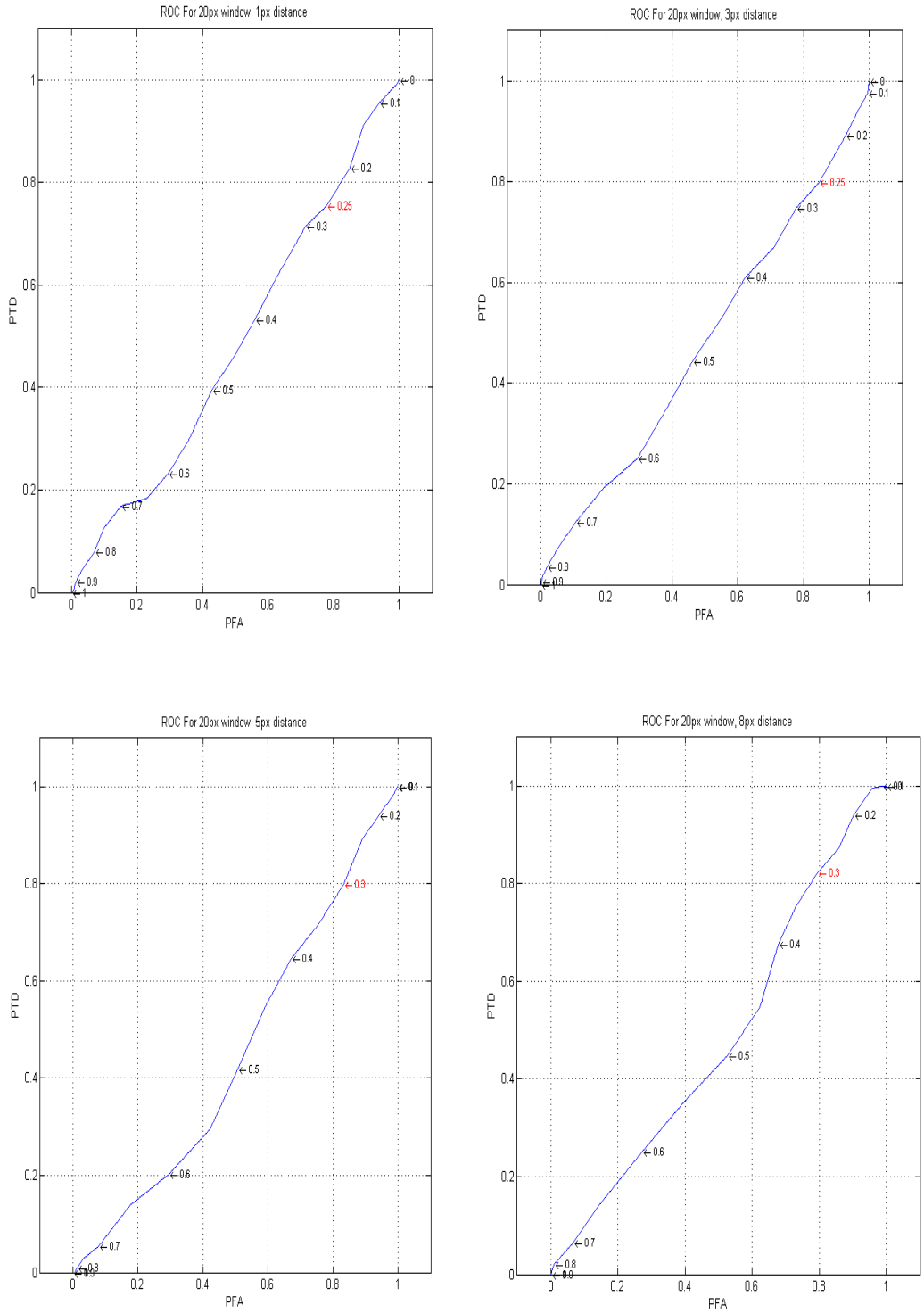


Figure 5.11. ROC curves for 20x20 pixel sized block and 1, 3, 5, 8 pixel distances

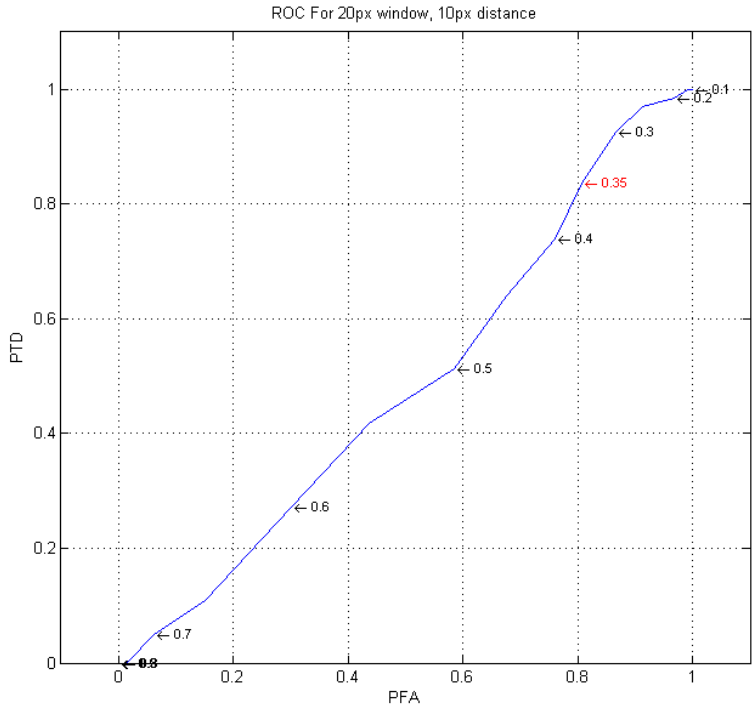


Figure 5.12. ROC curve for 20x20 pixel sized block and 10 pixel distances.

5.1.1.5. Performance of 28x28 Pixel (145m²) Block Size

700x700 pixel sized test image and control image was used in the experiment and both images were divided into 28x28 pixel blocks. There are 145 abnormal places and 480 normal places.

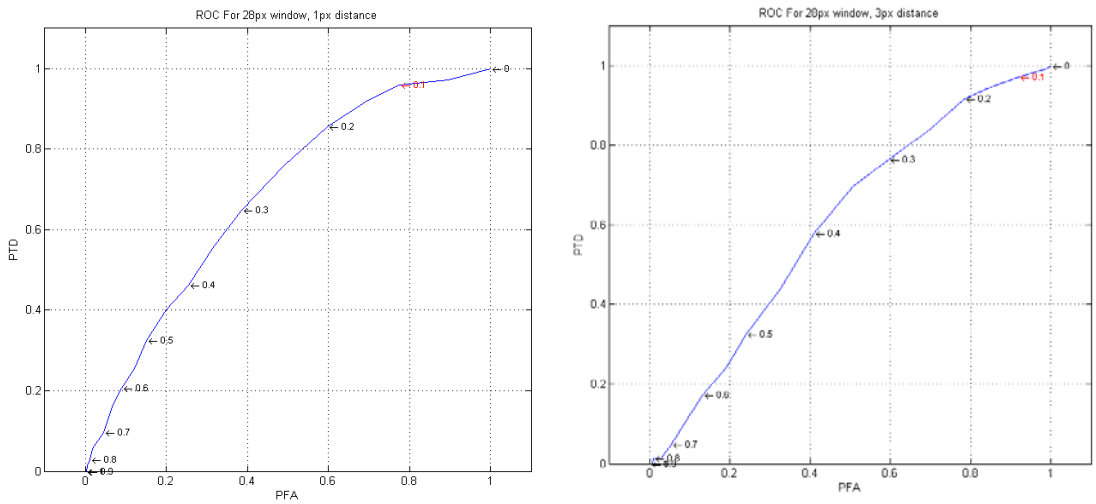


Figure 5.13. ROC curves for 28x28 pixel sized block and 1, 3 pixel distances.

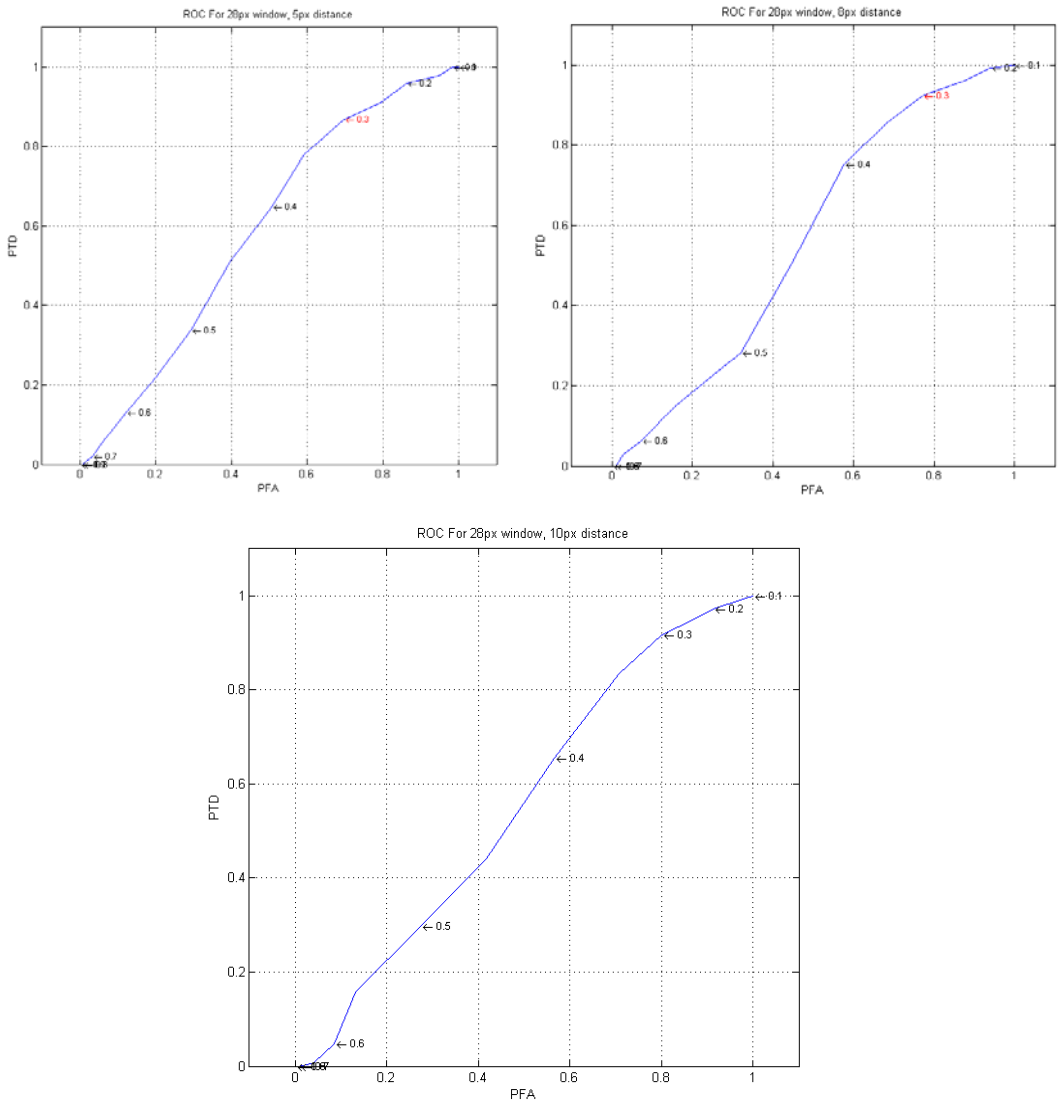


Figure 5.14. ROC curves for 28x28 pixel sized block and 5, 8, 10 pixel distances.

Detection performance was better than smaller sized blocks but the performance of the learning algorithm was not good enough. This experiment showed that the texture property of 28x28 pixel blocks is not suitable for recognitioning task.

5.1.1.6. Performance of 40x40 Pixel (296m²) Block Size

1000x1000 pixel sized test image and control image was used in the experiment and both image divided into 40x40 pixel blocks. After manually labeling 138 blocks were appeared as regions of interest and the other 487 blocks were appeare as normal. The labeled data is seen on the figures below.

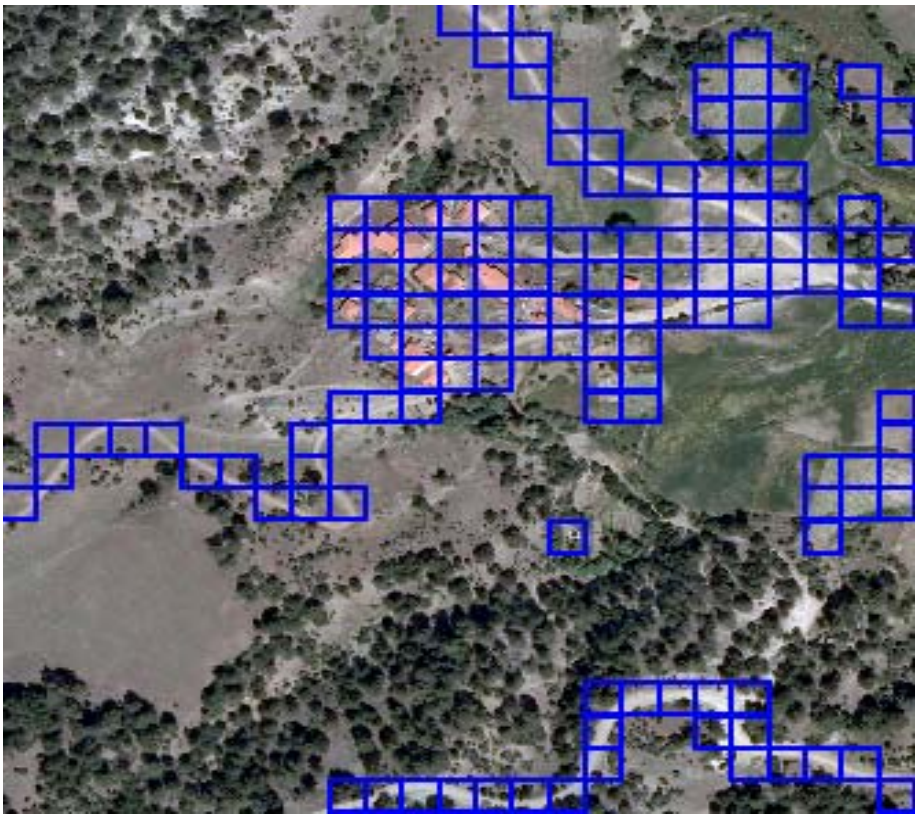


Figure 5.15. Blue colored blocks, the labeled data for 1000x1000 pixel sized test image

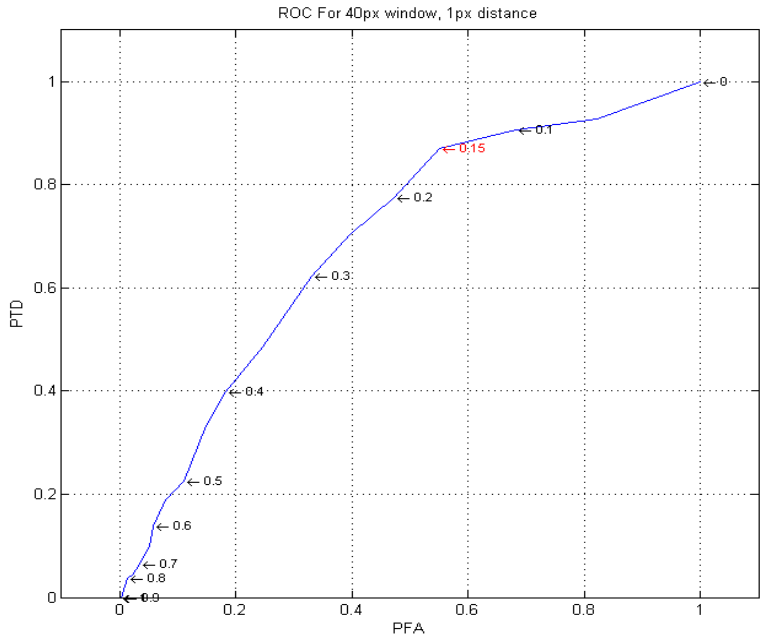


Figure 5.16. ROC curve for 40x40 pixel (296m²) Block Size and 1 pixel distance.

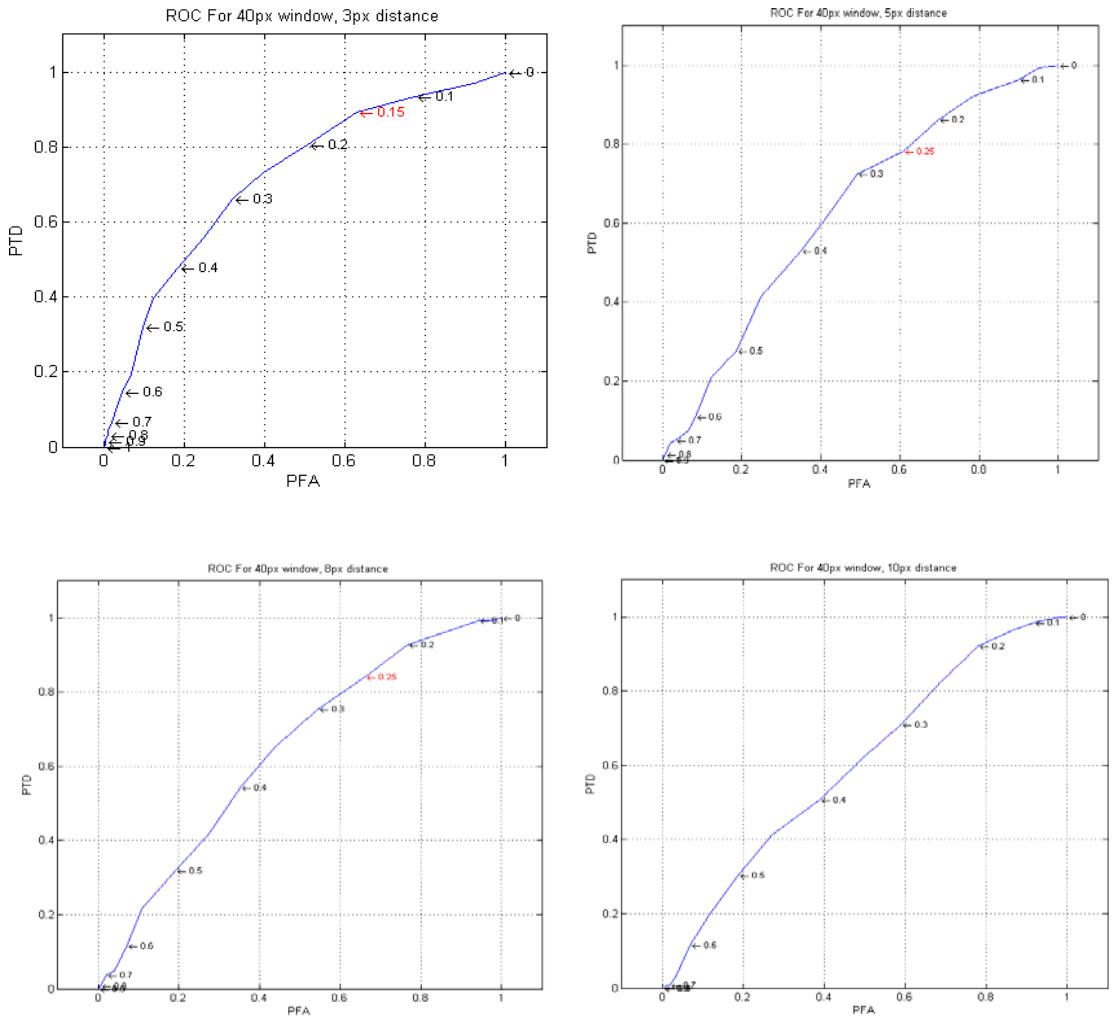


Figure 5.17. ROC curves for 40x40 pixel (296m²) Block Size and 3, 5, 8, 10 pixel distances.

This experiment showed that the results of the bigger sized blocks are better than small ones. Another result was noted that one pixel distance grey level co-occurrence matrices are the most informative textural property.

5.1.1.7. Performance of 56x56 Pixel (580m²) Block Size

1400x1400 pixel sized test image and control image was used in the experiment and both images were divided into 56x56 pixel blocks. 164 blocks were noted as abnormal blocks and 461 blocks were noted as normal. The reciever operating characteristics curves are seen on the figures below.

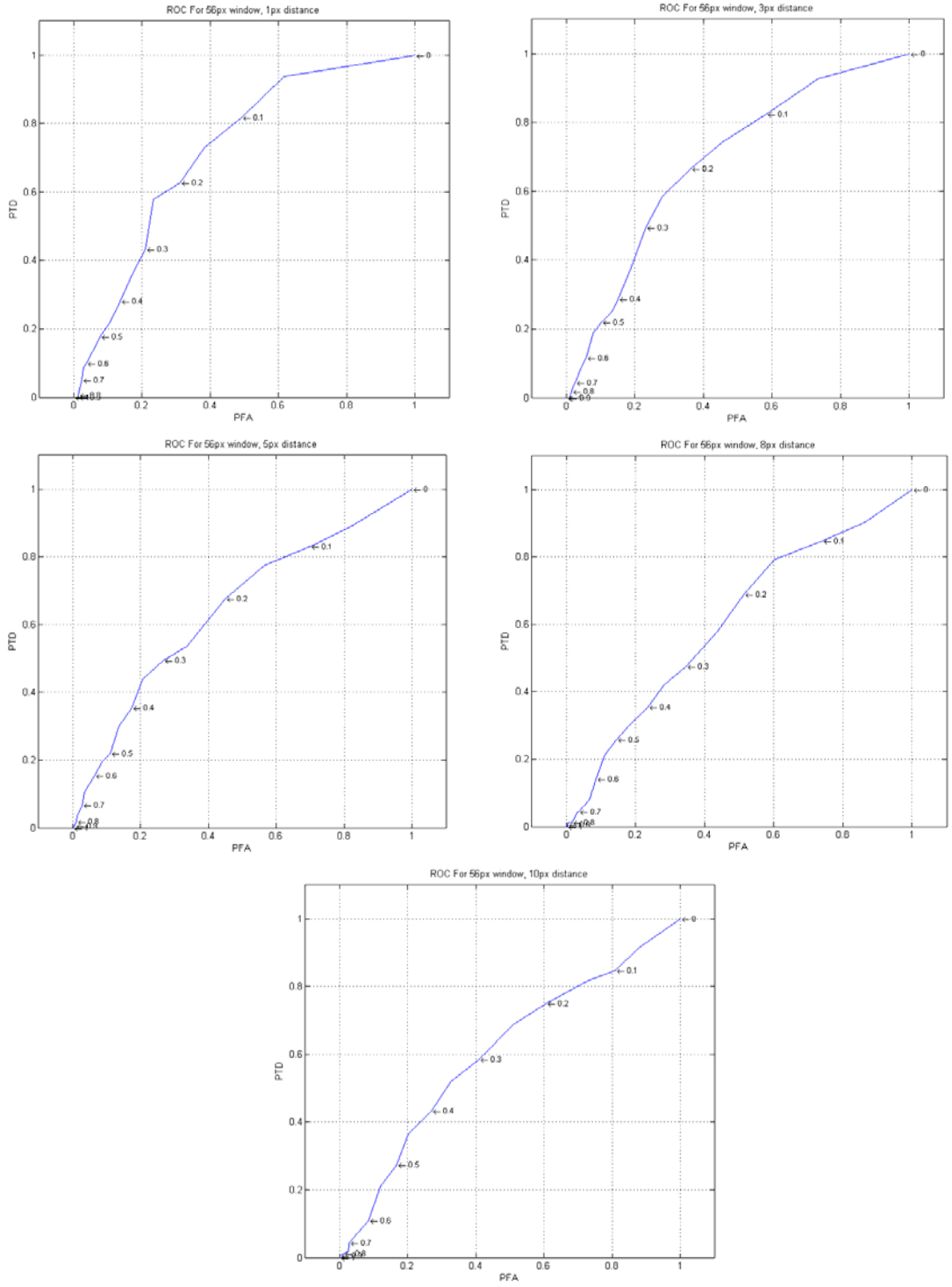


Figure 5.18. ROC curves for 56x56 pixel (580m²) Block Size and 1, 3, 5, 8, 10 pixel distances.

Detection performance of 56x56 pixel (580m²) block size textural features is better than 40x40 pixel sized blocks, and again the most informative feature is the one from 1 pixel distance neighbourhood.

5.1.1.8. Performance of 63x63 Pixel (734m²) Block Size

1575x1575 pixel sized test image and control image was used in the experiment and both images were divided into 63x63 pixel blocks. 175 blocks were noted as abnormal and the other 450 blocks were noted as normal. The reciever operating characteristics curves are seen on the figures below.

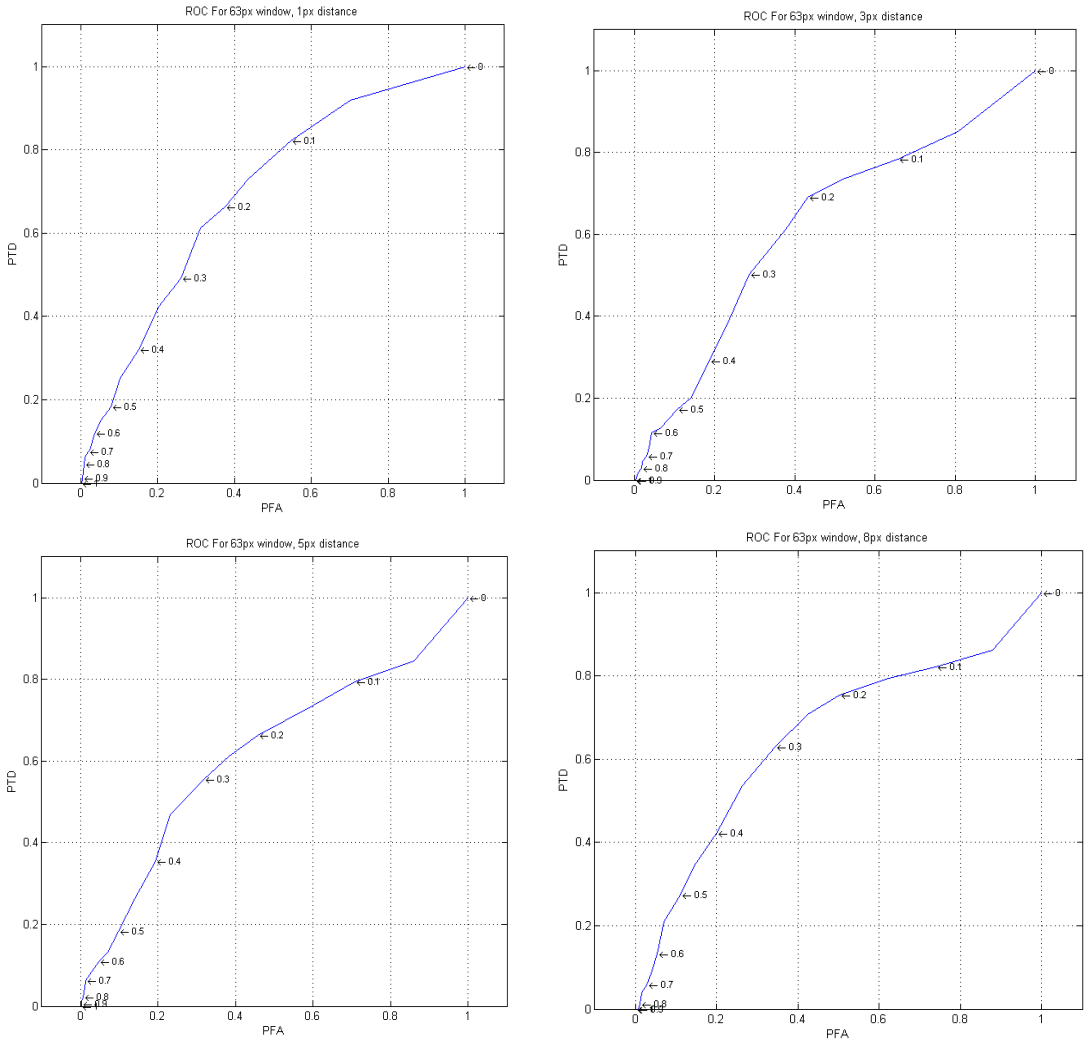


Figure 5.19. ROC curves for 63x63 pixel (734m²) Block Size and 1, 3, 5, 8 pixel distances.

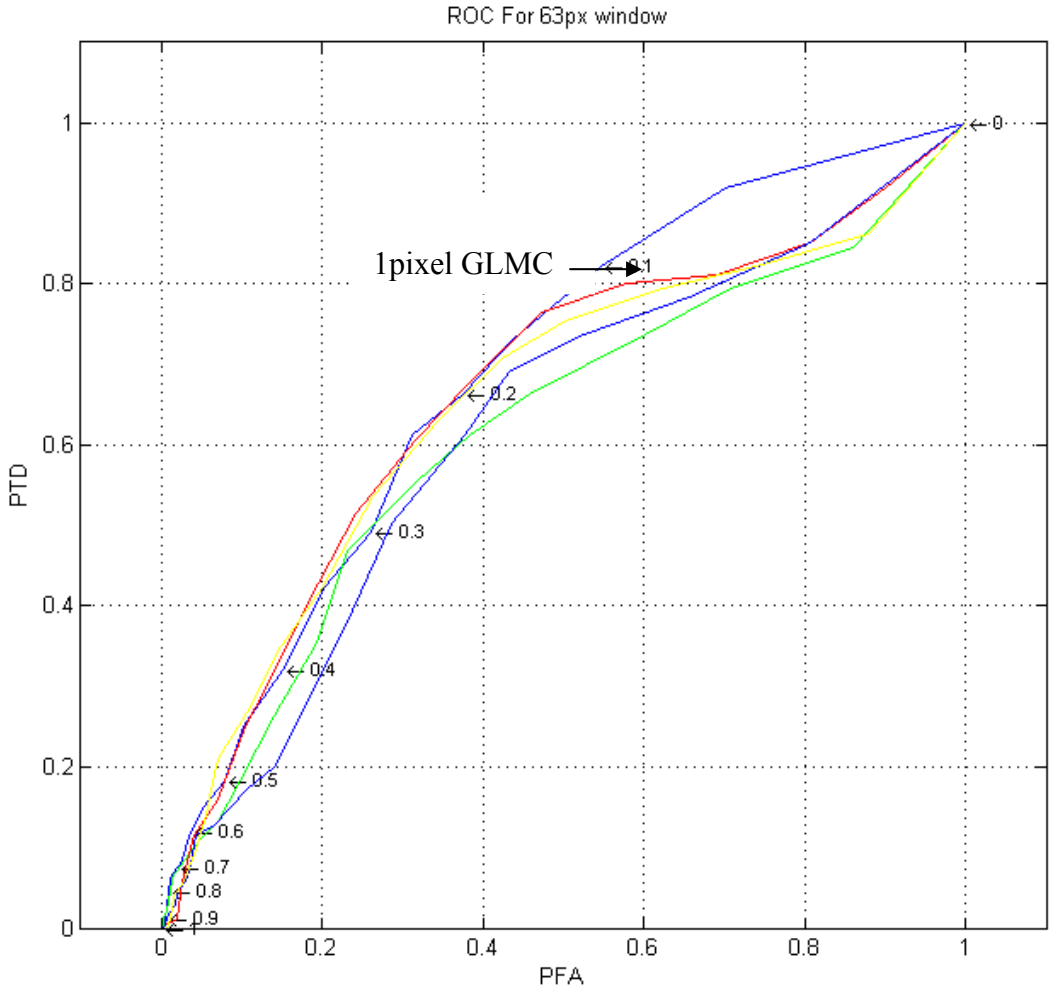


Figure 5.20. ROC curves for 63x63 pixel (734m²) Block Size and 1, 3, 5, 8, 10 pixel distances.

Figure 5.20. illustrates the performance of the textural features with different distances. The area under reciever operating characteristics curve was maximized with the one pixel distance GLCM feature, blue colored plot symbolizes the 1 pixel distance GLCM ROC curve. And black plot is for 3 pixel distance GLCM, green plot is for 5 distance pixel GLCM, yellow one is for 8 distance pixel GLCM, red plot is for 10 distance pixel GLCM.

These experiment showed that bigger block sizes give more accurate recognitions and generally one pixel distance grey level co-occurrence matrices yielded more textural information.

5.1.1.9. Performance of 80x80 Pixel (1183m²) Block Size

Along experiments for searching the optimal block size, 80x80 pixel (1183m²) sized block yielded the best detection performance of all. The area of the 80x80 pixel sized block is 1183m² and this is equal to a square having 34.4 meters side line. Along searching for optimal block size, some experiments carried- out for 100x100 pixel block and 120x120 pixel block. But after 80x80 pixel (1183m²) block size, detection performance was observed to decrease. Experiments showed clearly that block sizes of 100x100 pixel or more than 100x100 pixel are not convenient for identifying the objects of detections in the aerial images such as houses, roads and cultivated lands.

2000x2000 pixel sized test image and control image was divided into 625 non-overlapping grid blocks. Each block was marked as normal or abnormal with manually. The abnormally marked blocks are seen on the figure below. 174 blocks marked as abnormal and the rest 451 blocks marked as normal.

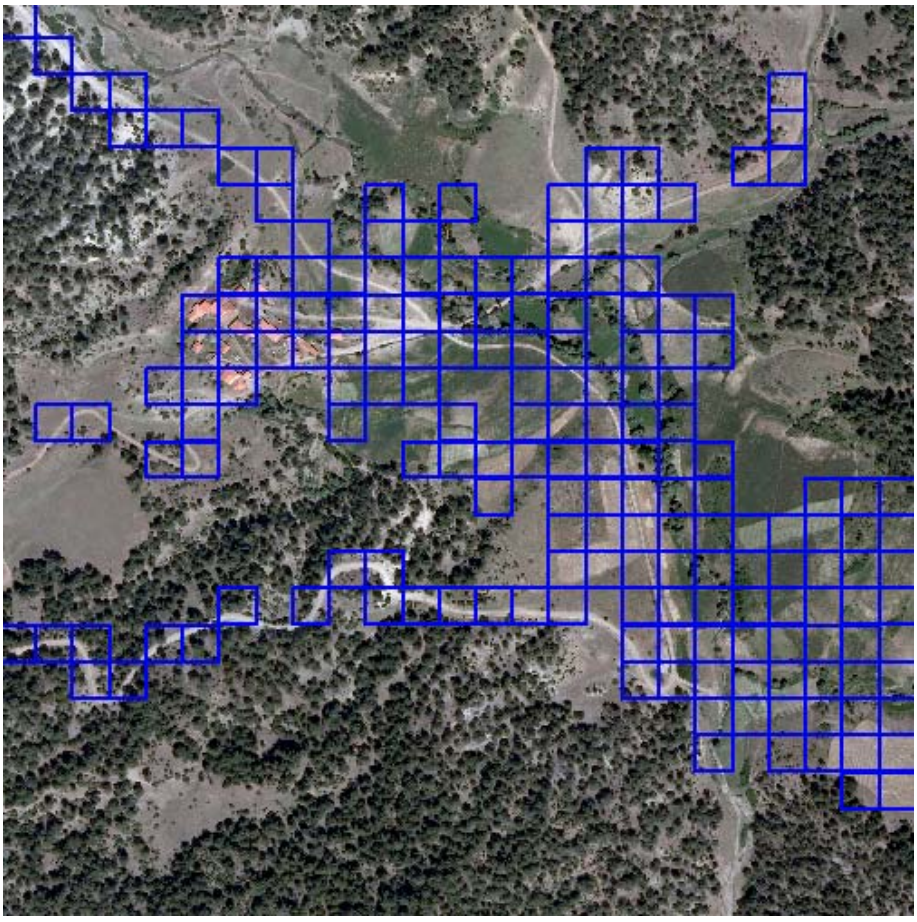


Figure 5.21. Blue colored blocks, the labeled data for 2000x2000 pixel sized test.

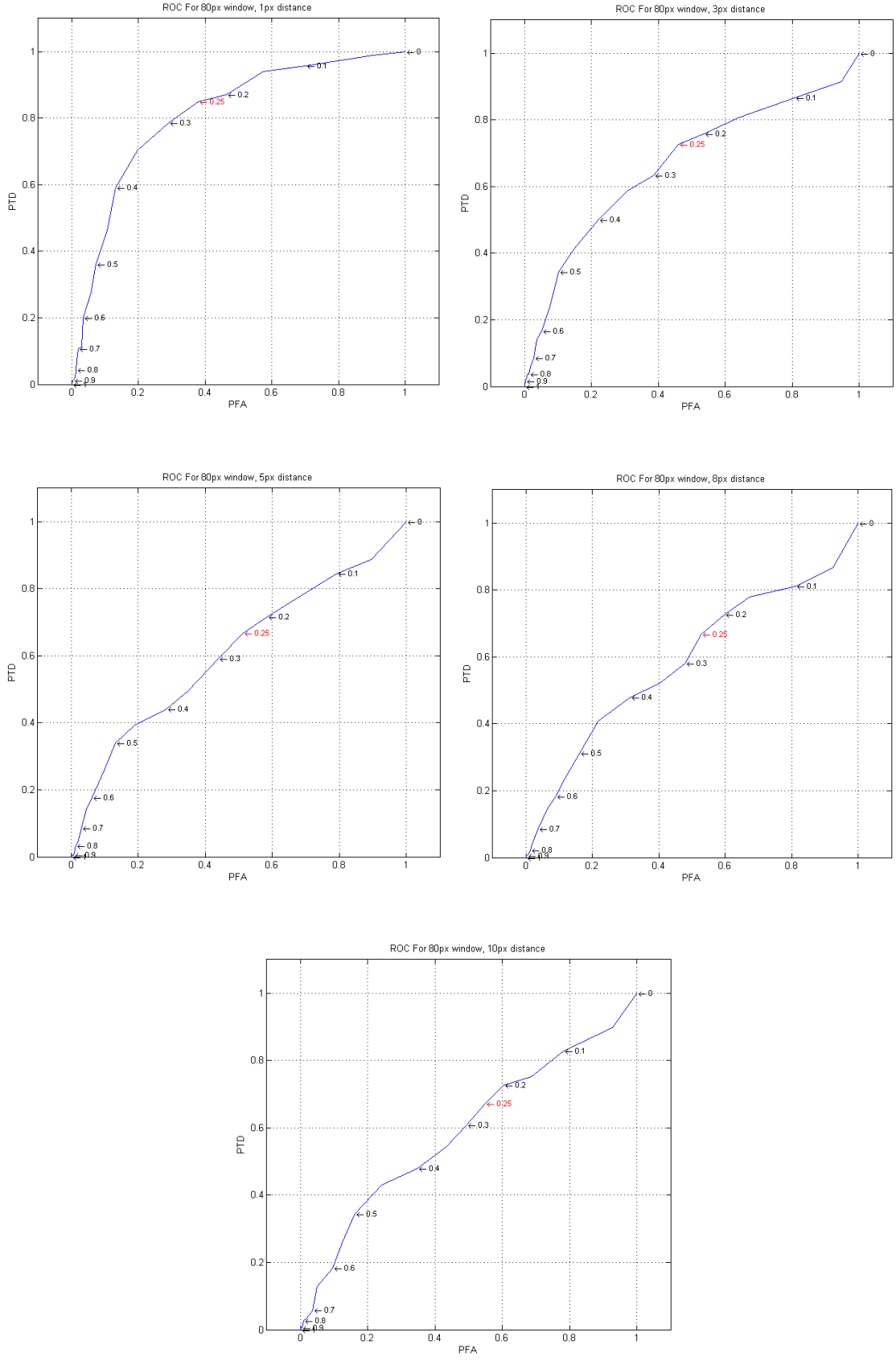


Figure 5.22. ROC curve for 80x80 pixel (1183m²) Block Size and 1, 3, 5, 8, 10 pixel distances.

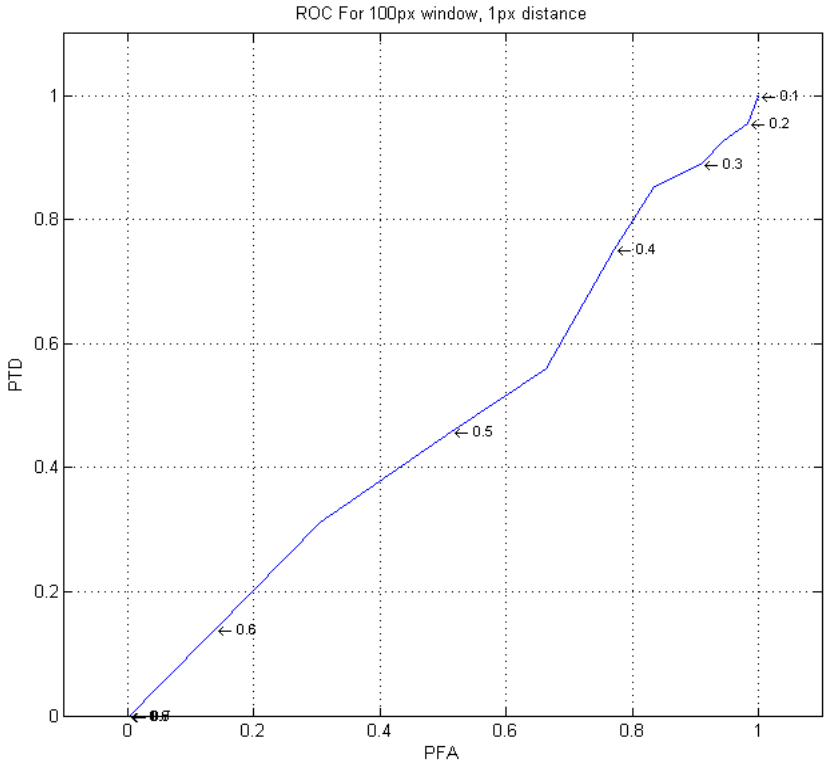


Figure 5.23. ROC curve for 100x100 pixel (1849m²) Block Size and 1, 3, 5, 8, 10 pixel distances.

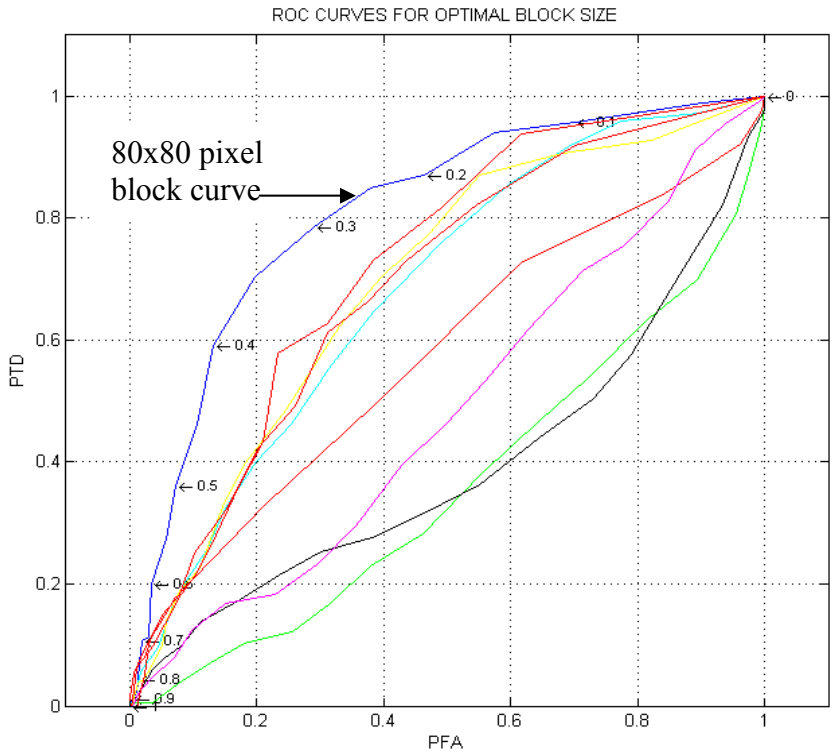


Figure 5.24. ROC curves for optimal block size.

On the Figure 5.24. The blue colored curve represents the 80x80 pixel blocks, red colored curves represents 63x63 pixel block, yellow colored curve represents 40x40 pixel block, cyan colored curve represents pixel 28x28 pixel block, magenta colored curve represents 20x20 pixel block, black colored curve represents 14x14 pixel block, green colored curve represents 10x10 pixel block and one pixel distance neighbourhood textural features.

From the reciever operating characteristics curve which is illustrated in figure 5.23, textures of the block sizes more than 80x80 pixels is not suitable for detecting the objects of interest in the aerial images. The blocks of 100x100 pixels and 120x120 pixels had given a bad detection performance. Among the blocks which were examined in the experiments, the 80x80 pixel sized block feature was the most effective of all. The area under reciever operating characteristics curve was maximum with the 80x80 pixel block and one pixel distance neighborhood.

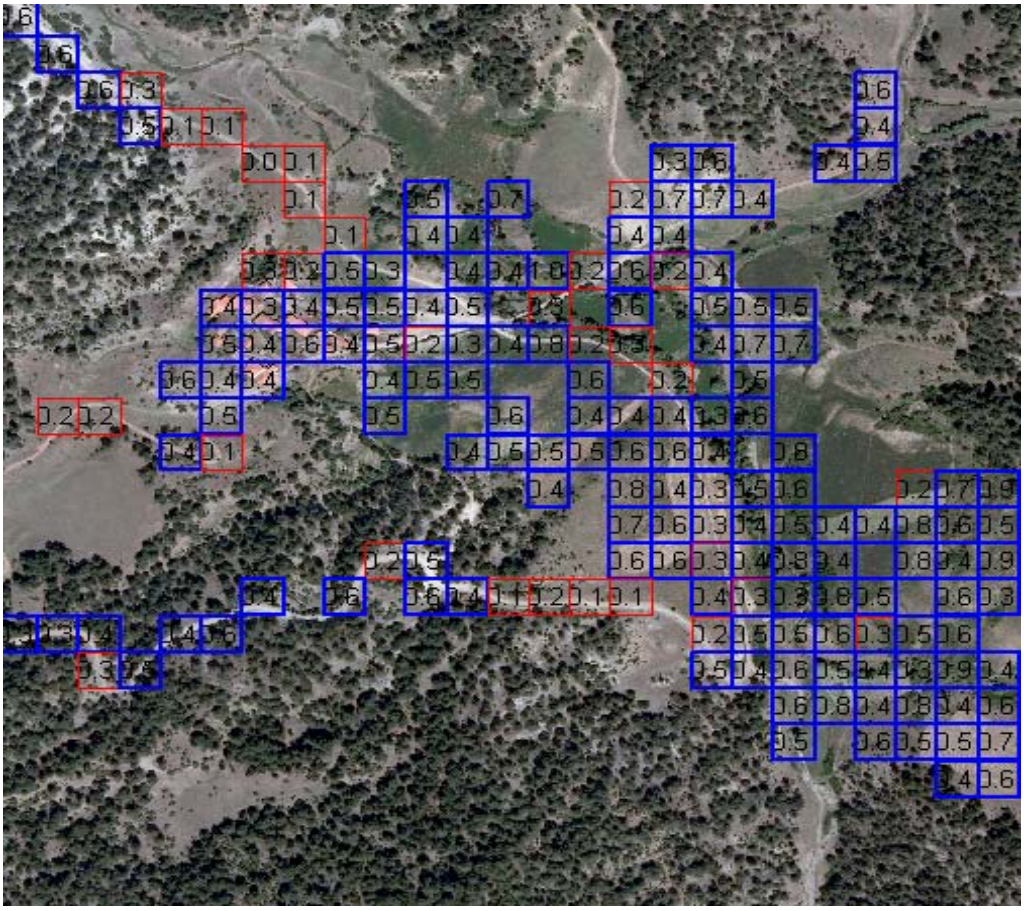


Figure 5.25. True detection regions and the regions where QSL failed with the textural feature of 80x80 pixel blocks and 1 pixel distance neighborhood GLCM.

A suitable treshold value of 0.32 the learning algorithm could detected over 80 percent of all objects of detections and it had given the minimum false alarm rate. Remember that, there were 174 blocks that we expected the learning algorithm to detect in the 2000x2000 pixel test image. As a result nearly 140 regions of interest were detected by the learning algorithm.

On the Figure 5.25. regions of true detections and the false alarms are illustrated. Blue colored blocks represent the true detection areas where learning algorithm was detected truly and red colored grids are the regions where learning algorithm could not detect (Figure 5.25). The common property of red blocks is there is a little abnormal structure along the whole area of the block. So texture feature vectors of those blocks are more similar to normal feature vectors. We have mentioned about classification algorithm, quasi-supervised learning algorithm tries to detect the abnormal ones via the distances between feature vectors in the feature space. For improving quasi-supervised learning algorithm's performance more samples should be used in the learning phase and more discriminative features can be used.

Consequently, 80x80 pixel blocks and 1 pixel distance neighborhood GLCM feature vector is the most succesfull block size in detecting the man-made structures on aerial images.

5.1.2. Optimal Distance for GLCM Feature Vectors

Each block was associated with a feature vector and distance measures that compute distances between these feature vectors were used to find similarities between blocks with the assumption that images that are close to each other in the feature space are also visually similar. Because of this assumption we should determine the most informative texture feature vector.

We have talked about computation of the grey level co-occurrence matrices and four Haralick features extracted from those matrices. In computing the grey level co-occurrence matrices, the distances of 1 pixel, 3 pixel, 5 pixel, 8 pixel and 10 pixel neighbourhood were used in the experiments. Distance between pixels is another important parameter in building the texture vectors. Experiments showed that the most informative texture feature vector is the one pixel distance neighborhood. Below, some reciever operating curves is given and it is clear that the area under reciever operating

curve is maximum with the one pixel distance neighborhood grey level co-occurrence matrice.

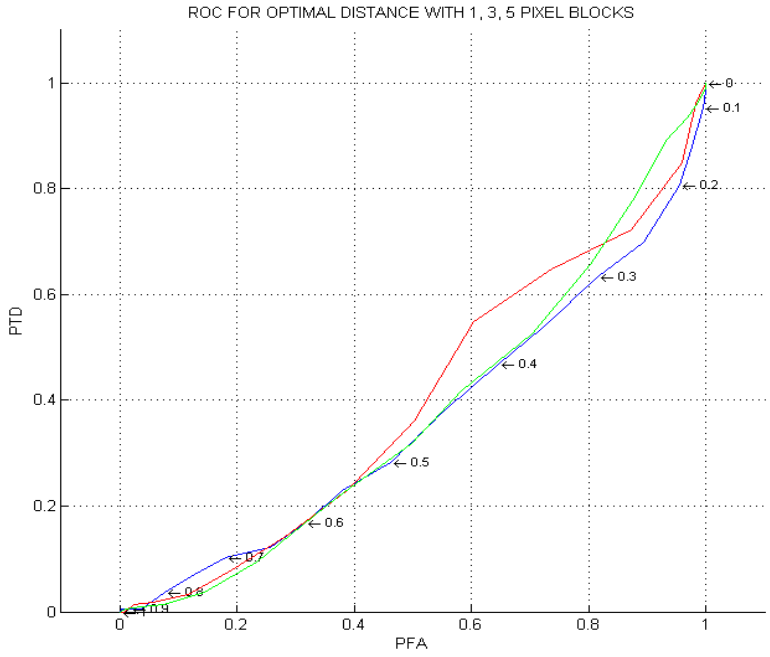


Figure 5.26. ROC curves for optimal distance neighborhood GLMC feature, 10x10 pixel blocks and 1, 3, 5 pixel neighborhoods respectively.

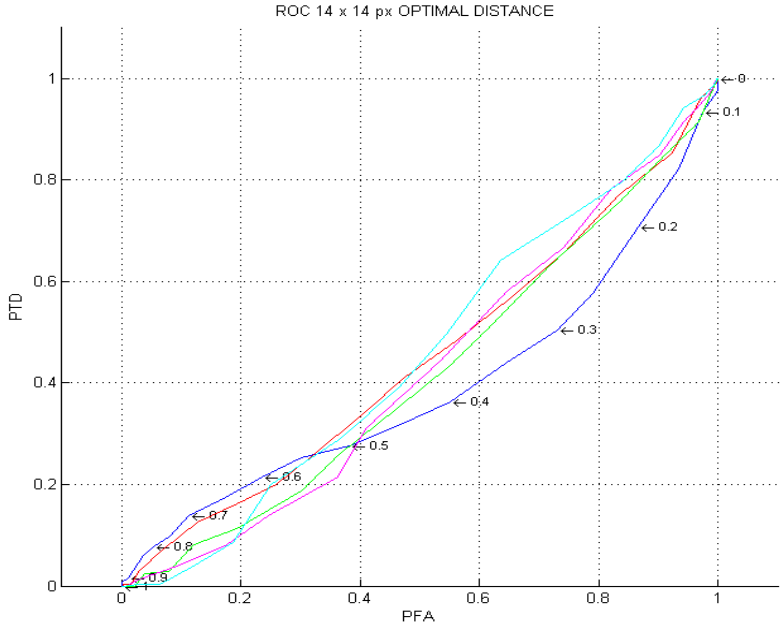


Figure 5.27. ROC curves for optimal distance neighborhood GLMC feature, 14x14 pixel blocks and 1, 3, 5, 8, 10 pixel distance neighborhoods respectively.

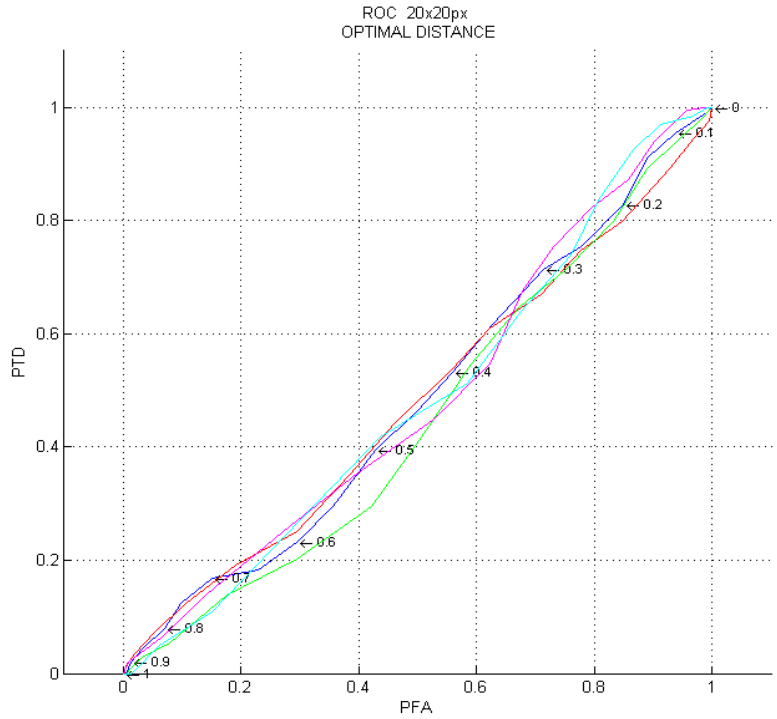


Figure 5.28. ROC curves for optimal distance neighborhood GLMC feature, 20x20 pixel blocks and 1, 3, 5 pixel neighborhoods respectively.

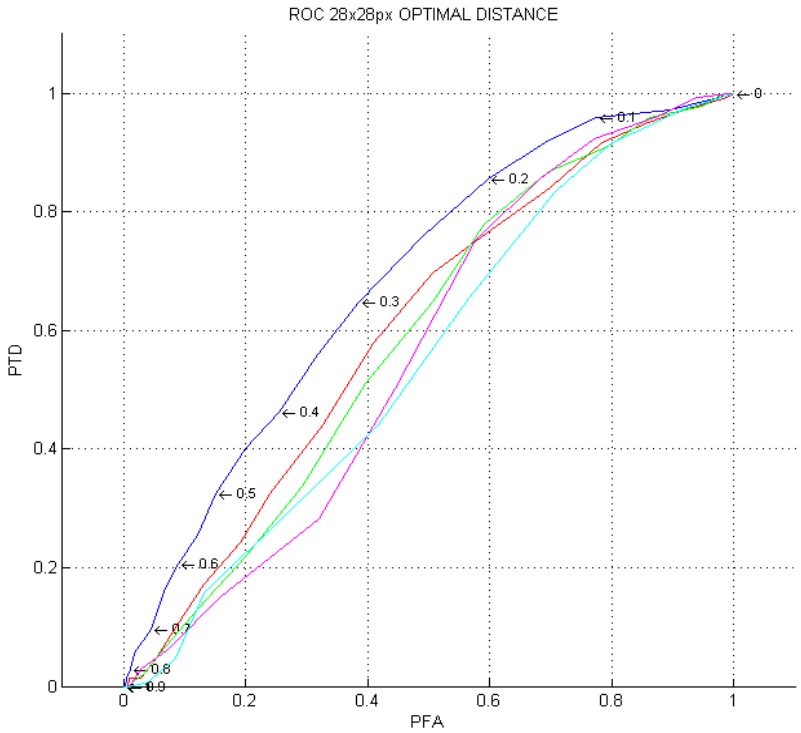


Figure 5.29. ROC curves for optimal distance neighborhood GLMC feature, 28x28 pixel blocks and 1, 3, 5 pixel distance neighbourhoods respectively.

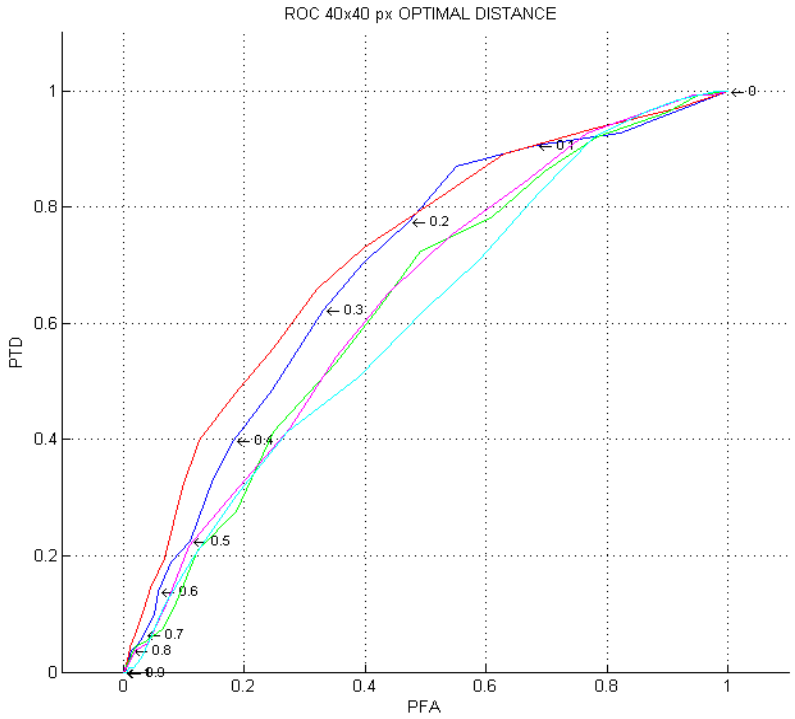


Figure 5.30. ROC curves for optimal distance neighborhood GLMC feature, 40x40 pixel blocks and 1, 3, 5 pixel distance neighbourhoods respectively.

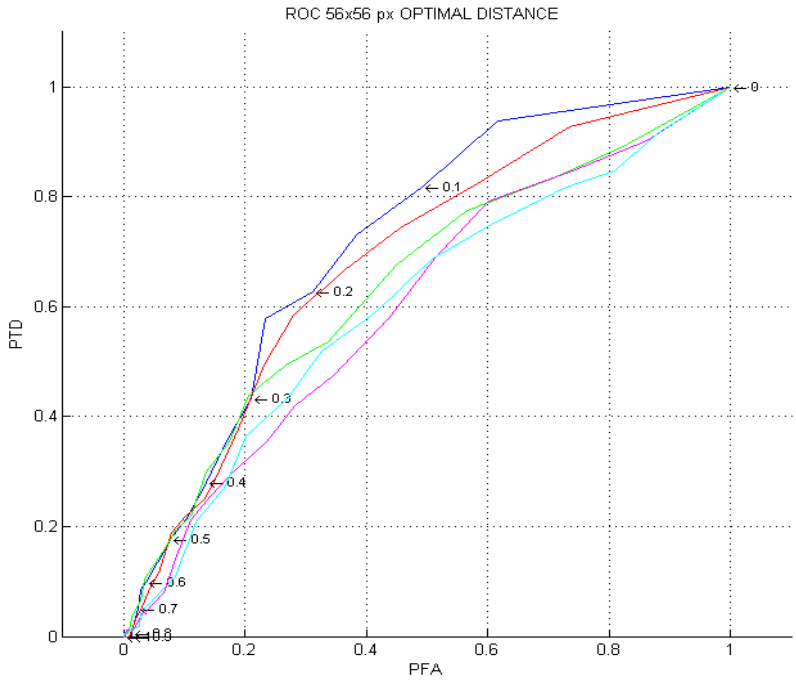


Figure 5.31. ROC curves for optimal distance neighborhood GLMC feature, 56x56 pixel blocks and 1, 3, 5 pixel distance neighbourhoods respectively.

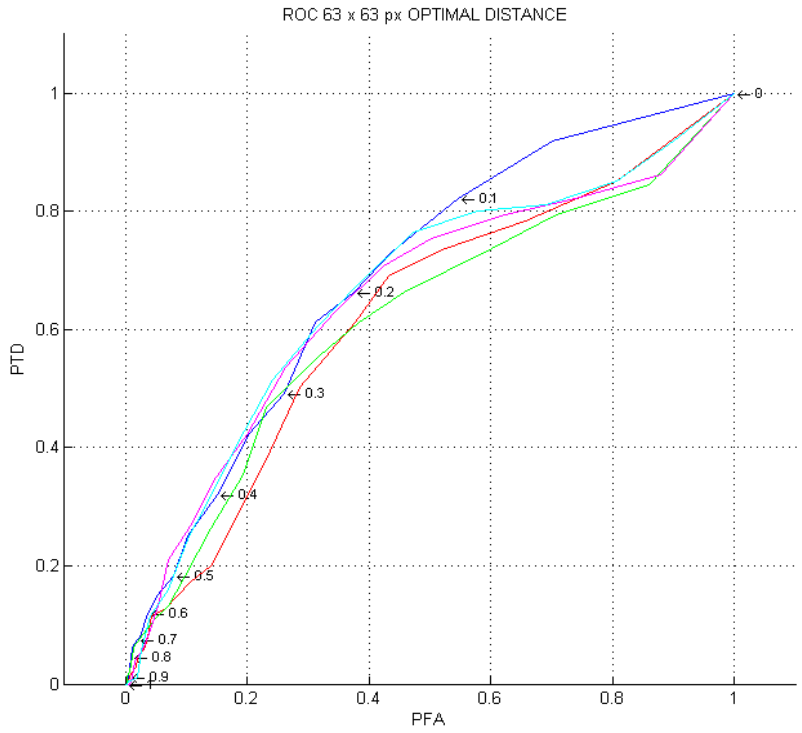


Figure 5.32. ROC curves for optimal distance neighborhood GLMC feature, 63x63 pixel blocks and 1, 3, 5 pixel distance neighbourhoods respectively.

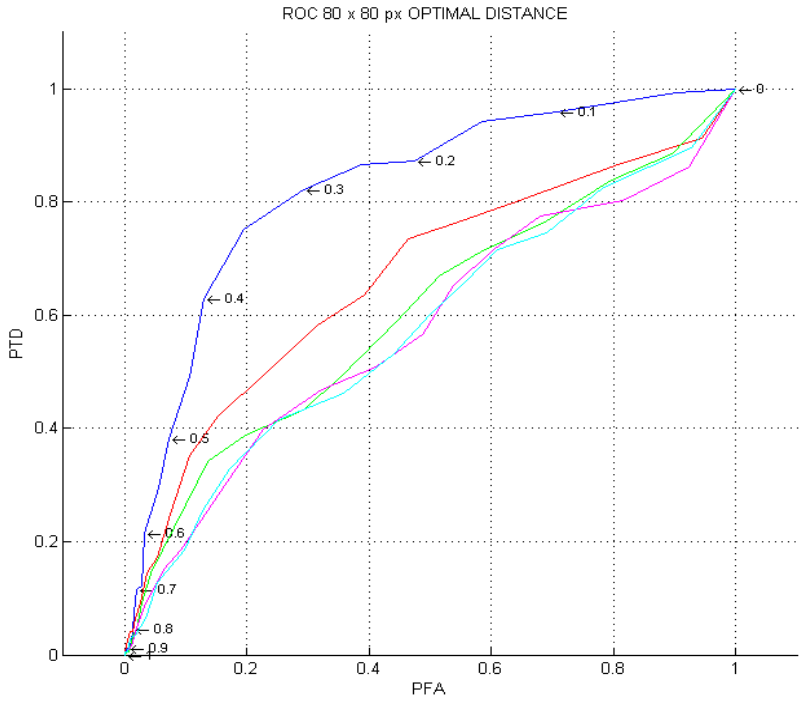


Figure 5.33. ROC curves for optimal distance neighborhood GLMC feature, 80x80 pixel blocks and 1, 3, 5 pixel distance neighbourhoods respectively.

In all figures for optimal distance above, the blue colored curves were used for representing the 1 pixel distance neighbourhood, red colored curves were used for representing 3 pixel distance neighbourhood, green colored curves were used for representing 5 pixel distance neighbourhood, magenta colored curves were used for representing 8 pixel distance neighbourhood and cyan colored curves were used for representing 10 pixel distance neighbourhood.

One pixel distance neighbourhood features had given bad results in small sized blocks like $10^{x}10$ pixel block and $14^{x}14$ pixel sized block. But the block sizes of $20^{x}20$ pixels or more, one pixel distance neighborhood became more succesfull in recognition. Remember that $80^{x}80$ pixel block size is the best one. As a result we can say that along all experiment, the $80^{x}80$ pixel sized blocks and one pixel distance grey level co-occurrence matrices were found as the optimal system parameters for quasi-supervised learning algorithm in detecting man-made structures on aerial images.

5.1.3. Quantization Level

The effect of quantization to detection performance was also tested. Experiments for quantization were carried-out along the $2000^{x}2000$ pixel sized test image and control image. Quantization is consist of seperation of RGB cube into equal sub cubes. The effect of grey levels was tested on four different cases: 32, 64, 128, 256 grey leveled images were used. In the feature extraction part, $2000^{x}2000$ pixels test image and control was used in for the experiments. Both images were divided into $80^{x}80$ pixel blocks. For each grey levels, one pixel distance neigbourhood co-occurrence matrices were then computed and organized into feature vectors.

Experiments on the effect of grey levels to the detection success resulted that 64 grey level is the most successful feature property of all. Actually 32 grey level textural features also had given a good detection performance and it was very close to the 64 level grey level. All four grey levels had different computation time. Reducing the grey levels has the same meaning to reduce the computation time. In aerial reconnaissance scenario the learning algorithm should respond in near real time. if it is necessary, the 32 grey levels can be used for extracting textural features. Because using 32 grey levels will reduce computation time and the performance of the learning will not be effected severely.

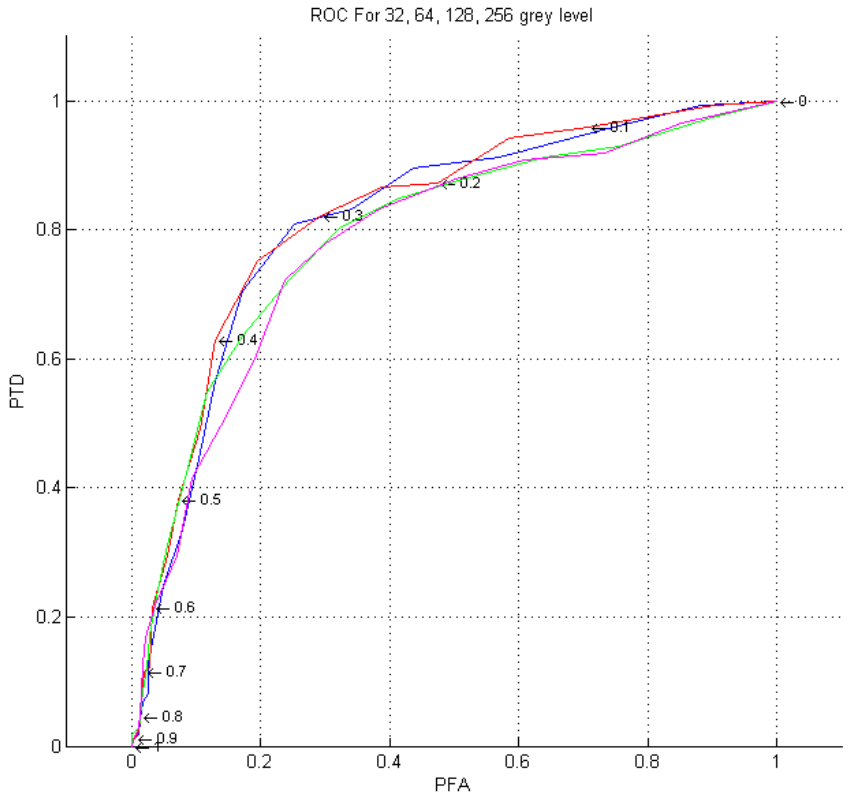


Figure 5.34. ROC curves for optimal quantization level(GLMC features are $80^{x}80$pixel blocks and 1 pixel distance neighborhoods and 32, 64, 128, 256 grey levels respectively).

Blue colored curve represents the 32 grey level feature, red colored curve represents the 64 grey level feature, green curve represents the 128 grey level feature and magenta colored curve represents the 256 grey levels. Detection performance of four different textural features are very close to each other as seen on the figure, but the most successful textural profile is 64 grey level feature vector.

CHAPTER 6

CONCLUSION

In this thesis, a quasi-supervised learning algorithm was implemented on eigthteen aerial images. Images were divided into reference control dataset and a testing data set, nine of the images were belong to the testing data set and the other nine images were belong to control data set. The elements of reference control data were normalcy terrestrial conditions and natural looking, but testing images consisted of both man-made structures like building, roads etc. and natural terrestrial land. Those images splited into small blocks and each block was associated with a textural feature vector. Totally 42 different texture profile were tested and selection of the most successful texture profile is presented.

Since image classification is based on textural features and texture is defined with feature vectors, we should determine the most informative textural properties. Learning algorithm needs a distance measure that computes the distances between the feature vectors. These distance measures are used to determine the similarities between images with the assumption of images close to each other in the feature space is also smilar. So in order to make a true recognition we should find the correct textural properties. Experiments were carried-out along 42 different textural features. These features were computed from the nine different pairs of test and control images, each image was quantized to grey levels and grey level information was used. The purpose of all the experiments are to detemine the optimal system parameters for learning algorithm. $80^{x}80$ pixel size block and one pixel distance neighbourhood grey level co-occurrence matrices were observed as the most efficent assets in detecting the man-made structures on the aerial images.

Supervised learning requires the definition of a certain segment of data. In some applications, the ground truth data are available and the target variables are well defined. But in our scenario of detecting man-made structures in aerial images, there is no pre-determined object and target variable is unknown. So we should use a quasi-supervised learning in aerial reconnaissance scenario. Quasi-supervised statistical learning algorithm is an appropriate tool for this task. Because it is based on a

classification method that divides available data into reference control data which has only normalcy conditions and a mixed testing data which has abnormal regions along with normals. Than identifiying the samples that are specific to testing data is enough for detecting the man-made structures in the aerial images.

The results of the experiments showed that abnormal regions can be identified accurately with the appropriate texture vectors. According to the experimental results and performance evaluation of those 42 texture profiles, one pixel distance neighbourhood grey level co-occurrence matrices and the block size of $80^{x}80$ pixels had given good detection results. Grey level information was used in all experiments and the most successfull textural profile in grey tones was 64 level quantization. It was noted that quantization level do not effect detection performance too much.

Consequently; quasi-supervised learning was observed as a successful technique for recognitioning man-made structures in aerial images. The $80^{x}80$ pixel block size, one pixel distance neighbourhood and 64 level quantization properties are the most succesfull system parameters for aerial images. In future works, number of the samples can be increased and color information can be used for improving the detection success.

REFERENCES

[1] Robert M. Haralick, K. Shanmugam, and Its'hak Dinstein, "Textural Features For Image Classification", IEEE Transactions On Systems, Man, And Cybernetics, vol. Smc-3, No. 6, November 1973.

[2] Haralick, R.M., "Statistical And Structural Approaches To Texture", Proceedings Of The IEEE, 67(5), pp. 786-804, 1979.

[3] Selim Aksoy and Robert M. Haralick, "Feature Normalization and Likelihood-based Similarity Measures for Image Retrieval" , 5 October 2000

[4] Karacali B., "Quasi-Supervised Learning for Biomedical Data Analysis", Pattern Recognition, 2009.

[5] Hui Zhang, Jason E. Fritts, Sally A. GoldmanHui Zhang, Jason E. Fritts, Sally A. Goldman, "A Fast Texture Feature Extraction Method for Region-based Image Segmentation"

[6] Andreas Stolcke, Sachin Kajarekar, Luciana Ferrer, "Feature Normalization and Likelihood-based Similarity Measures for Image Retrieval"

[7] A.P. Pentand, R.W. Picard, S Sclaroff. Photobook: Content-based manipulation of image databases. International journal of computer vision,18(3) 233-254,1996.

[8] C.Carson, S Belongie, H. Greenspan, J. Malik : Color and texture based image segmantation using em and it's application to content-based image retrieval. IEEE Transacntions on pattern analysis and machine intelligence, 24(8) :1026-1038, August 2002.

[9] A. K. Jain, F. Farrokhnia. Unsupervised texture segmentation using gabor filters. Patter Recognition, 24:1167-1186, December 1991

[10] M. Petrou. Classifying textures when ssen from different distances. In IAPR international conference on pattern recognition, pages 83-86 August 2002.

[11] Munoz. Inmage segmantation integrating color, texture and boundary information. PhD. Thesis university of Girona, February 2003

[12] S.Fukuda and H. Hirosawa. A wavelet-based texture feature set applied to classification of multifrequency polarimetric sar images. IEEE transactions on geosience and remote sensing, 37:2282-2286, May 1999

[13] D. We J. Linders A new texture approach to discrimination of forest cleancut, canopy and burned area using airborne c-band sar. IEEE transactions on geosience and remote sensing, 37(1), 555-563,1999

[14] F. Lumbreas, R. Baldrich. Multi resolution color texture representation for tile classification. Pages 145-153, Bilbao, Spain, 1999.

[15] Mihran Tuceryan, Anil K. Jain: Texture Segmentation Using Voronoi Polygons. IEEE Trans. Pattern Anal. Mach. Intell. 12(2): 211-216 (1990)

[16] Use of gray value distribution of run lengths for texture analysis Original Research Article Pattern Recognition Letters, Volume 11, Issue 6, June 1990, Pages 415-419 A. Chu, C.M. Sehgal and J.F. Greenleaf

[17] Tuceryan, M. and A. K. Jain, "Texture Segmentation Using Voronoi Polygons," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990

[18] Tomita, Fumiaki and S. Tsuji, *Computer Analysis of Visual Textures*, Kluwer Academic Publishers, Boston, 1990.

[19] Zucker, S. W., "Toward a model of Texture," Computer Graphics and Image Processing, 1976.

[20] Sklansky, J., "Image Segmentation and Feature Extraction," IEEE Transactions on Systems, Man, and Cybernetics, 1978.

[21] Hawkins, J. K., "Textural Properties for Pattern Recognition," Academic Press, New York, 1969.

[22] Jain, A. K., S. K. Bhattacharjee, and Y. Chen, "On Texture in Document Images," to appear in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1992.

[23] Khotanzad, A. and R. Kashyap, "Feature Selection for Texture Recognition Based on Image Synthesis," *IEEE Transactions on Systems, Man, and Cybernetics*,

[24] Eom, Kie-Bum and R. L. Kashyap, "Texture and Intensity Edge Detection with Random Field Models," In Proceedings of the Workshop on Computer Vision.

[25] Voorhees, H. and T. Poggio, "Detecting textons and texture boundaries in natural images," First International Conference on Computer Vision, London, 1987.

[26] Tuceryan, M. and A. K. Jain, "Texture Segmentation Using Voronoi Polygons," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990.

[27] Karacali, B. and A. Tozeren, "Automated detection of regions of interest for tissue microarray experiments: an image texture analysis", BMC Med Imaging, 7: pp.

[28] Esgiar, A.N., et al., "Microscopic Image Analysis for Quantitative Measurement and Feature Identification of Normal and Cancerous Colonic Mucosa", IEEE Transactions On Information Technology In Biomedicine, 2(3): pp. 197-203.

[29] A.P. Dhawan, Y. Chitre, C. Kaiser-Bonaso, "Analysis of mammographic microcalcifcations using gray-level image structure features", IEEE Trans. Med. Imag., 15 (3) pp. 246–259, 1996.