

**THE REPUBLIC OF TURKEY
BAHCESEHIR UNIVERSITY**

**IR IMAGE EDGE DETECTION USING NEURAL
NETWORK AND CLUSTERING**

Master's Thesis

TALA MOHAMMADZADEH MEYMANDI

ISTANBUL, 2018

**THE REPUBLIC OF TURKEY
BAHCESEHIR UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED
SCIENCES
COMPUTER ENGINEERING**

**IR IMAGE EDGE DETECTION USING NEURAL
NETWORK AND CLUSTERING**

Master's Thesis

TALA MOHAMMADZADEH MEYMANDI

Supervisor: ASSIST. PROF. TARKAN AYDIN

İSTANBUL, 2018

**THE REPUBLIC OF TURKEY
BAHCESEHIR UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
COMPUTER ENGINEERING**

Name of the thesis: IR image edge detection using Neural Network and clustering
Name/Last Name of the Student: Tala Mohammadzadeh Meymandi
Date of the Defense of Thesis: May.13.2018

The thesis has been approved by the Graduate School of Natural and Applied Sciences.

Assist. Prof. Yücel Batu SALMAN
Graduate School Director
Signature

I certify that this thesis meets all the requirements as a thesis for the degree of Master of Science.

Assist. Prof. Tarkan AYDIN
Program Coordinator
Signature

This is to certify that we have read this thesis and we find it fully adequate in scope, quality and content, as a thesis for the degree of Master of Science.

Examining Committee Members

Signature

Thesis Supervisor
Assist. Prof. Tarkan AYDIN

Thesis Co-supervisor
Assist. Prof. Pınar SARISARAY BÖLÜK

Member
Assist. Prof. Mürüvvet Aslı AYDIN

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my Supervisor Asst. Prof. Dr. Tarkan Aydin for his advice, encouragement, guidance and continuous feedback throughout this thesis.

I would like to thank my family for their constant support throughout my life.

May, 2018

Tala Mohammadzadeh MEYMANDI

ABSTRACT

IR IMAGE EDGE DETECTION USING NEURAL NETWORK AND CLUSTERING

Tala Mohammadzadeh Meymandi

Computer Engineering

Thesis Supervisor: Assist. Prof. Tarkan Aydin

May 2018, 87 pages

Nowadays image processing and feature extraction methods provide significantly important knowledge about images. The first step for identifying objects in an image is extracting the image properties. Edge detection is one the common features of image processing, because edges include useful information about an image. Although general public may not deal with Infrared images directly, this field is widely benefited in many sciences. Therefore, a proper infrared image edge detection method could result in thorough comprehension. In this study, infrared images are selected for edge detection due to their application in various technologies such as medical, military fields and surveillance purposes. According to the structure of these images, it is not possible to extract their edges using common methods. Therefore, a new method is proposed for edge detection of infrared images. In the proposed method, first the image is segmented by a clustering algorithm. Then, Neural Network algorithm is selected to extract the region of interest among the segmented clusters. In the last step, morphological operators are used to extract the edges from the Region of Interest. For segmentation, two K-means and Mean Shift clustering methods are applied separately, and their cluster features are used as the Neural Network inputs. Pursuant to the advantage of Mean Shift clustering algorithm in cluster number determination this method may be favorable in many cases. The evaluation results of the proposed method and comparison with other available methods indicate the method's good performance for infrared image edge detection.

Keywords: Infrared images, K-means clustering, Mean Shift clustering, Neural Network, Region of Interest

ÖZET

KIZILÖTESİ GÖRÜNTÜDE SİNİR AĞI VE KÜMELEME İLE KENAR BELİRLEME

Tala Mohammadzadeh Meymandi

Bilgisayar Mühendisliği

Tez Danışmanı: Dr. Öğr. Üyesi Tarkan AYDIN

Mayıs 2018, 87 sayfa

Günümüzde görüntü işleme ve özellik çıkarma yöntemleri, görüntüler hakkında önemli bilgiler sağlamaktadır. Bir görüntüdeki nesnelere tanımlamak için ilk adım görüntü özelliklerini ayıklamaktır. Kenar belirleme, görüntü işlemenin ortak özelliklerinden biridir, çünkü kenarlar bir görüntü hakkında faydalı bilgiler içerir. Genel halk doğrudan Kızılötesi görüntülerle ilgilenmese de bu alan birçok bilimlerinde yaygın olarak kullanılmaktadır. Bu nedenle, uygun bir kenar belirleme metodu ile kızılötesi görüntüde, kapsamlı anlayış sonuçlanabilir. Bu çalışmada, kenar belirleme metodu kızılötesi görüntüler için seçilmiştir; çünkü bu görüntüler çeşitli teknolojilerde mesela medikal, askeri alanlar ve gözetim amaçları için uygulanmaktadır. Bu görüntülerin yapısına göre, kenarlarını ortak belirleme metodlarla mümkün değildir. Bu nedenle, kızılötesi görüntülerin kenar bulması için yeni bir yöntem önerilmiştir. Önerilen yöntemde, önce görüntü bir Kümeleme algoritması ile bölümlere ayrılır. Ardından, ayrılmış bölümler arasında ilgi bölgesi çıkarmak için Sinir Ağı algoritması seçilir. Son adımda, ilgi bölgesinde Morfolojik işletmeciler kenarları çıkarmak için kullanılır. K-Ortalamalar Kümeleme ve Ortalama Kaydırma yöntemleri ile görüntü bölünür ve kümelerin özellikleri Sinir Ağının girişleri olarak kullanılır. Küme sayısının belirlenmesi için Ortalama Kaydırma ile Kümeleme algoritmasının avantajına göre, bu yöntem birçok durumda uygun olabilir. Önerilen yöntemin değerlendirme sonuçlarına göre ve diğer mevcut yöntemlerin karşılaştırmanın neticelerine göre, yöntemin kızılötesi görüntü kenarı belirleme için iyi performansını göstermektedir.

Anahtar Kelimeler: Kızılötesi görüntüler, K-Ortalamalar Kümeleme, Ortalama Kaydırma ile Kümeleme, Sinir Ağı, İlgi Bölgesi

CONTENTS

| | |
|--|-------------|
| TABLES..... | ix |
| FIGURES..... | x |
| ABBREVIATIONS..... | xii |
| SYMBOLS..... | xiii |
| 1. INTRODUCTION..... | 1 |
| 1.1 INTRODUCTION..... | 1 |
| 1.2 PROBLEM STATEMENT..... | 2 |
| 1.3 THESIS OBJECTIVES..... | 5 |
| 1.4 NEW ASPECTS AND RESEARCH INNOVATION..... | 5 |
| 1.5 THESIS STRUCTURE..... | 5 |
| 2. LITERATURE REVIEW..... | 7 |
| 2.1 INTRODUCTION..... | 7 |
| 2.2 IMAGE PROCESSING..... | 7 |
| 2.2.1 Image Processing Applications..... | 8 |
| 2.2.2 Imaging..... | 14 |
| 2.2.3 Pre-processing..... | 14 |
| 2.2.4 Feature Extraction from Images..... | 16 |
| 2.3 THERMAL IMAGING..... | 19 |
| 2.3.1 Thermal Imaging Components..... | 20 |
| 2.3.2 Different Generations of Thermal Cameras..... | 21 |
| 2.3.3 Effective Factors on Image Quality..... | 22 |
| 2.3.4 Accuracy and Recognition Factors of Images..... | 22 |
| 2.3.5 Main Effective Factors in Thermal Imaging..... | 23 |
| 2.3.6 Selection of Wavelength Region for Thermal Cameras..... | 24 |
| 2.3.7 Atmospheric Effects on Thermal Cameras' Performance.... | 24 |
| 2.4 IMAGE SEGMENTATION..... | 26 |
| 2.4.1 Clustering..... | 26 |
| 2.4.1.1 K-means clustering..... | 27 |
| 2.4.1.2 Mean shift clustering..... | 27 |
| 2.4.1.2.1 Mean shift applications..... | 27 |

| | |
|---|-----------|
| 2.4.2 Classification..... | 29 |
| 2.4.2.1 Artificial neural network | 29 |
| 2.5 RESEARCH BACKGROUND..... | 33 |
| 2.6 CONCLUSION..... | 34 |
| 3. DATA AND METHOD..... | 35 |
| 3.1 INTRODUCTION..... | 35 |
| 3.2 APPLIED TOOLS..... | 35 |
| 3.3 OUTLINE OF THE THESIS..... | 36 |
| 3.4 PRE-PROCESSING..... | 37 |
| 3.5 IMAGE SEGMENTATIONS..... | 39 |
| 3.5.1 K-means Clustering..... | 39 |
| 3.5.1.2 Optimal number of clusters..... | 42 |
| 3.5.2 Mean Shift Clustering..... | 43 |
| 3.6 ROI EXTRACTION..... | 47 |
| 3.6.1 Feature Extraction from The Image Regions..... | 47 |
| 3.6.2 Artificial Neural Network..... | 49 |
| 3.7 EDGE DETECTION..... | 51 |
| 3.7.1 Post-Processing..... | 53 |
| 3.7.2 Edge Detection Using Morphological Operators..... | 55 |
| 3.8 CHAPTER SUMMARY | 56 |
| 4. FINDINGS..... | 57 |
| 4.1 INTRODUCTION..... | 57 |
| 4.2 EVALUATION..... | 57 |
| 4.2.1 Dataset..... | 57 |
| 4.3 EVALUATION RESULT..... | 57 |
| 4.3.1 Evaluation Method..... | 58 |
| 4.3.1.2 Confusion matrix..... | 69 |
| 4.3.1.3 Running time..... | 72 |
| 4.4 CONCLUSION..... | 73 |
| 5. DISCUSSION..... | 74 |
| 5.1 COMPARING THE RESULTS WITH OTHER METHODS..... | 74 |
| 5.2 COMPARING THE RESULTS WITH OTHER WORKS..... | 76 |

| | |
|--|----|
| 5.3 COMPARING BOTH CLUSTERING METHODS..... | 79 |
| 6. CONCLUSION..... | 86 |
| 6.1 INTRODUCTION..... | 86 |
| 6.2 THESIS RESULTS..... | 86 |
| 6.3 PROPOSAL FOR FUTURE RESEARCH..... | 87 |
| REFERENCES..... | 88 |
| APPENDICES | |
| Appendix A.1 Running time..... | 94 |
| Appendix A.2 Confusion matrix..... | 95 |

TABLES

| | |
|--|----|
| Table 2.1: Pre-processing categories..... | 15 |
| Table 2.2: Correspondence between ANN and BNN..... | 33 |
| Table 3.1: Extracted Quantile | 45 |
| Table 3.2: Extracted features for each region using K-means clustering..... | 48 |
| Table 3.3: Extracted features for each region using Mean Shift clustering..... | 49 |
| Table 3.4: Region labeling for training the NN (K-means clustering)..... | 51 |
| Table 3.5: Region labeling for training the NN (Mean Shift clustering)..... | 51 |
| Table 4.1: Confusion matrix of Figure 4.1 | 70 |
| Table 4.2: Confusion matrix of Figure 4.3..... | 70 |
| Table 4.3: Total results of all 16 images..... | 70 |
| Table 4.4: Performance comparison of both methods..... | 72 |

FIGURES

| | |
|---|----|
| Figure 1.1: Prewitt edge detection method..... | 3 |
| Figure 2.1: Motion tracking..... | 9 |
| Figure 2.2: Medical image Processing applications..... | 9 |
| Figure 2.3: Distinction of various tissues from each other..... | 10 |
| Figure 2.4: Image processing application for ultrasound measurement..... | 10 |
| Figure 2.5: Computer-assisted surgeries using image processing technique..... | 11 |
| Figure 2.6: Intercept targets in military applications using image processing..... | 12 |
| Figure 2.7: Image processing in security systems..... | 12 |
| Figure 2.8: Image processing in geographic systems for identifying cover crops..... | 13 |
| Figure 2.9: Separation of the defective fruit surface by image processing..... | 14 |
| Figure 2.10: Feature extraction steps..... | 16 |
| Figure 2.11: Electromagnetic waves..... | 19 |
| Figure 2.12: Object tracking for surveillance..... | 28 |
| Figure 2.13: Object tracking in a soccer game..... | 28 |
| Figure 2.14: Structure of the humans' brain..... | 30 |
| Figure 2.15: Simple structure of the Neural Network..... | 32 |
| Figure 3.1: General steps of the proposed method..... | 37 |
| Figure 3.2: Median Filter..... | 38 |
| Figure 3.3: Applying Median filter..... | 39 |
| Figure 3.4: Block diagram explaining the k-means clustering steps..... | 40 |
| Figure 3.5: Clusters extracted using k-means clustering | 41 |
| Figure 3.6: Elbow criterion..... | 43 |
| Figure 3.7: Flat kernel..... | 44 |
| Figure 3.8: Image density structure..... | 45 |
| Figure 3.9: The resulting clusters by Mean Shift clustering..... | 46 |
| Figure 3.10: Clusters extracted using Mean Shift Clustering..... | 47 |
| Figure 3.11: Expected object for edge detection in the IR image..... | 48 |
| Figure 3.12: Structure of Neural Network in this study | 50 |
| Figure 3.13: Pre-processing..... | 52 |
| Figure 3.14: Clustering using both methods..... | 52 |
| Figure 3.15: Binary image rebuilt from the extracted ROI | 53 |

| | |
|--|----|
| Figure 3.16: Binary image after Erosion..... | 54 |
| Figure 3.17: Binary image after Dilation..... | 55 |
| Figure 3.18: Detected edges..... | 55 |
| Figure 4.1: Detected edges (b,c) of image img_00003 (a) in folder 0002..... | 58 |
| Figure 4.2: Detected edges (b,c) of image img_00014 (a) in the folder 0002..... | 59 |
| Figure 4.3: Detected edges (b,c) of image img_00028 (a) in the folder 0002..... | 60 |
| Figure 4.4: Detected edges (b,c) of image img_00001 (a) in the folder 0004..... | 61 |
| Figure 4.5: Detected edges (b,c) of image img_00009 (a) in the folder 0004..... | 62 |
| Figure 4.6: Detected edges (b,c) of image img_00018 (a) in the folder 0004..... | 63 |
| Figure 4.7: Detected edges (b,c) of image img_00001 (a) in the folder 0006..... | 63 |
| Figure 4.8: Detected edges (b,c) of image img_00009 (a) in the folder 0006..... | 64 |
| Figure 4.9: Detected edges (b,c) of image img_00018 (a) in the folder 0006..... | 65 |
| Figure 4.10: Detected edges (b,c) of image img_00001 (a) in the folder 0007..... | 65 |
| Figure 4.11: Detected edges (b,c) of image img_00011 (a) in the folder 0007..... | 66 |
| Figure 4.12: Detected edges (b,c) of image img_00022 (a) in the folder 0007..... | 66 |
| Figure 4.13: Detected edges (b,c) of image img_00001 (a) in the folder 0008..... | 67 |
| Figure 4.14: Detected edges (b,c) of image img_00012 (a) in the folder 0008..... | 67 |
| Figure 4.15: Detected edges (b,c) of image img_00024 (a) in the folder 0008..... | 68 |
| Figure 4.16: Detected edges (b,c) of image img_00032(a) in the folder 0009..... | 68 |
| Figure 5.1: Comparison of proposed method with the common algorithms..... | 75 |
| Figure 5.2: Comparison of proposed method with the common algorithms..... | 76 |
| Figure 5.3: Comparison of proposed methods with CNN_DGA method..... | 77 |
| Figure 5.4: Comparison of proposed method with CNN_DGA method..... | 77 |
| Figure 5.5: Comparison of the proposed method with Mix_model..... | 78 |
| Figure 5.6: Comparison of the proposed method with Mix_model..... | 79 |
| Figure 5.7: Comparison between k-means and Mean Shift clustering..... | 80 |
| Figure 5.8: Comparison between k-means and Mean Shift clustering..... | 81 |
| Figure 5.9: Comparison between k-means and Mean Shift clustering..... | 82 |
| Figure 5.10: Comparison between k-means and Mean Shift clustering..... | 83 |
| Figure 5.11: Comparison between k-means and Mean Shift clustering..... | 84 |
| Figure 5.12: Comparison between k-means and Mean Shift clustering..... | 85 |
| Figure 5.13: Comparison between k-means and Mean Shift clustering..... | 85 |

ABBREVIATIONS

| | | |
|--------|---|---|
| ACO | : | Ant Colony Optimization |
| ANN | : | Artificial Neural Network |
| BNN | : | Biological Neural Network |
| CCD | : | Charge-Coupled Device array |
| CNN | : | Cellular Neural Networks |
| DGA | : | Distributed Genetic Algorithms |
| FOV | : | Field of View |
| GPS | : | Global Positioning System |
| HgCdTe | : | Mercury Cadmium Telluride Detectors |
| HSL | : | Hue Saturation Intensity |
| InSb | : | Indium Antimonite Semiconductors |
| IR | : | Infrared Radiation |
| KDE | : | Kernel Density Estimation |
| MLP | : | Multilayer Perceptron |
| MRTD | : | Minimum Resolvable Temperature Difference |
| NETD | : | Noise Equivalent Temperature Difference |
| NN | : | Neural Network |
| RGB | : | Red Green Blue |
| ROI | : | Region of Interest |
| SSE | : | Standard Square Error |
| USA | : | United States of America |

SYMBOLS

| | | |
|---------------------------------|---|---------------|
| Centigrade | : | C |
| Degree | : | ° |
| Extracted Region from the image | : | I_n |
| micrometer | : | μm |
| millimeter | : | mm |
| nanometer | : | nm |

1. INTRODUCTION

1.1 INTRODUCTION

The knowledge of image and photo is a great science. Nowadays, acquiring and analyzing images have a great impact on many different sciences. Because of this great influence various methods for image acquisition and processing methods are being developed. Each method has its own advantages and disadvantages and they're being used in their own fields. It is important to consider that due to the variety of techniques available for image processing, in many cases applying a particular method is not enough. One of the basic utilization of image processing is edge detection. All image edges include useful information that could be very helpful for object detection. Infrared image is one of the most popular and vital image types in today's life. Infrared Radiation (IR) images are widely used in medical and military industries and also in surveillance applications. The temperature of human body is a great sign of humans' well-being. Therefore, IR images could provide considerable information about health and therefore, Infrared imaging has a significant role in diagnosing many illnesses and disorders. Many warfare weapons and smart devices use this technology for capturing and identifying objects. Breast cancer, diabetes, kidney transplantation, dermatology, heart diseases, fever screening and brain imaging are some of the examples that indicate the success of using IR imaging. Accordingly, IR images are applied in many different technologies due to their characteristics and a lot of researchers are focused on this subject. In this study, a new method is proposed to study edge detection of infrared images which is one of the most important tasks about infrared images. For this purpose, image segmentation methods with machine learning tools are applied.

In the beginning of this chapter, problem statement and the initial definitions of the research tools are presented, then the research objectives are expressed. Finally, new aspects and research innovation are expressed.

1.2 PROBLEM STATEMENT

In recent years, (Wang et al. 2014; Lahiri et al. 2012) various methods have been proposed for image processing such as Cellular Neural Networks (CNN), genetic algorithms and wavelet transforms. One of the most important aspects of image processing is edge detection. Edge detection technology is used to extract edge features. Edge feature is one of the most basic features of the image and it could be used to display the image. Edge detection is a sensitive task for target tracking. Therefore, tracking objects in images and movies is one of the most significant tasks. There are various image types for processing and the type of image determines the operation that must be performed on them. In the next section, a brief explanation of the infrared images is given.

Infrared energy from all objects with a temperature above zero ° Kelvin (the absolute temperature or -273 ° Centigrade) is emitted. IR is a part of the electromagnetic spectrum with the frequency between the color spectrum and the radio waves. The IR wavelength in the electromagnetic spectrum is between 0.7 μm and 1000 μm (one millimeter). In this band the waves with the wavelengths between 0.7 micrometers and 20 μm are used for measuring the temperature. Cameras' imaging sensors convert this energy to electrical signals which are displayed on monitor as a thermal monochromatic image. These images also show different values of heat (Duarte et al. 2014). Sir William Herschel (1800) discovered Infrared thermography. Sir But (1940s) invented the first infrared imaging system. Since 1960s Infrared imaging has been used in medical science. Over the past 20 years, there have been significant improvements in the quality of imaging equipment, the standardization of techniques and clinical imaging protocols (Duarte et al. 2014).

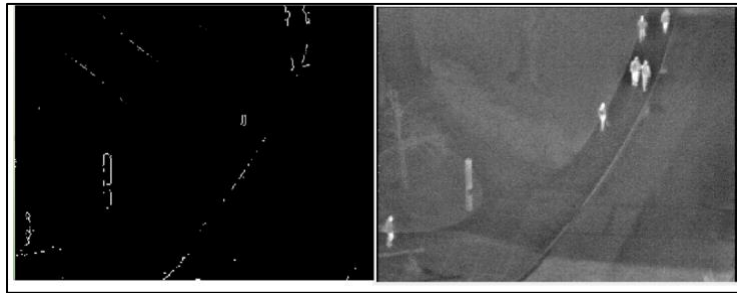
The following features could be noted as some of the IR features (Zhou et al. 2011) :

- a. Random changes of the external environment and thermal imaging systems' flaws could cause different infrared image noises such as thermal noise. The existence of these noises causes reduction of the signal quality.
- b. The infrared image determines the temperature distribution in the image; this image is black and white (gray); it is not a color or three-dimensional image, so it has low resolution for the human eye.

- c. Due to the structure of thermal images their imaging systems also have low recognizing ability of objects, hence its spatial precision is lower than the visible light in the Charge-Coupled Device array (CCD), which makes the infrared image resolution lower comparing with other image types.

Infrared image has many uses in medical sciences (Duarte et al. 2014) and military technologies (Abdulmunim et al. 2012, Sun 2003). Due to lower contrast and resolution of IR images comparing with color images, common image processing algorithms are not suitable. The most common edge detection algorithms are Prewitt, Canny, Sobel and Roberts edge detection algorithms which could not be applied (Wang et al. 2014). Because most of the common image edge detection methods work by extracting high frequency signals. As common operators are sensitive to noise it is difficult to distinguish between image noise and edges. Figure 1.1 shows the result of IR image edge detection using Prewitt edge detection method as a first order derivative filter.

Figure 1.1: Prewitt edge detection method



Therefore, other edge detection methods are required. One of these methods are based on machine learning algorithms and clustering methods. Clustering algorithms segments the images according to the images' features. K-means and Mean Shift clustering are two popular clustering algorithms that are used in this study. In this approach each clustering method is applied separately with the machine learning algorithm due to their characteristics and their results are compared.

K-means algorithm is an unsupervised algorithm which clusters the pixels based on some similar features such as gray levels. In this method the number of clusters (k) is constant which should be defined initially by the user. Some random points are selected as centers of the clusters (centroids). This algorithm has a loop with two parts:

- i. Assignment: Each point is assigned to the cluster with the closest centroid.
- ii. Computation: computing the centroids of all clusters in each loop.

The computation continues until convergence happens (Dhanachandra et al. 2015). For assigning and computing the clusters and their centroids many different distance measures are used. Mean Shift is a hierarchical non-parametric clustering algorithm, unlike k-means clustering method the algorithm itself figures out the number and the location of the clusters. The main concept of this algorithm is to find the densest region by computing the mean within a chosen bandwidth. In each iteration first, the points within the radius of the mean is calculated then the new mean is computed. These two steps are computed in a loop until convergence happens. In many cases this algorithm is preferable as it figures out the optimized number of clusters. Unlike k-means it doesn't require to define the number of clusters. That is a great advantage, as in k-means with incorrect cluster number the algorithm doesn't work properly (Cheng 1995).

After clustering the image, the Region of Interest (ROI) should be found and extracted using a proper method. for this purpose, proper machine learning algorithm could be implemented. In this study, ROI is extracted using Neural Network (NN) algorithm which is one of the common machine learning algorithms. The NN is inspired by the biological behavior of neural systems of the human's brain. This network is a combination of a large number of connected processing elements (nerves). The NN is a combination of three layers (Gonzalez et al. 2008):

- Input layers
- Hidden layers
- Output layers

Some of the Neural Network's advantages and disadvantages are listed as following (Gonzalez et al. 2008):

Advantages:

- a. Widespread application in many different fields.
- b. Very flexible because the user decides about its structure.
- c. Finds the complex relationship between inputs and outputs.

Disadvantages:

- a. It's hard to interpret, so it's difficult to explain it.
- b. Awareness is limited about the fundamental connections.

- c. Need to be designed and preprocessed accurately with predictive variables.

In this study, in order to detect the edges both k-means and Mean Shift clustering algorithms are used for IR segmentation and Neural Network algorithm is used for ROI extraction.

1.3 THESIS OBJECTIVES

The main aim of this thesis is to provide a new method for the edge detection of IR images. In this thesis, the combinations of NN with two different clustering methods (k-means and Mean Shift clustering) are studied and a new method is provided that properly detects the edges of infrared images. The results of this research could be used in the medical industry, military, and in general wherever the infrared image is used.

1.4 NEW ASPECTS AND RESEARCH INNOVATION

Considering the earlier work, although extensive work was done on infrared image edge detection, there was no such proposed method for infrared image edge detection. On the other hand, given that both K-means and Mean Shift clustering methods have high results in image segmentation and Artificial Neural Networks (ANN) also functions as well, so the proposed method is expected to perform well.

1.5 THESIS STRUCTURE

The structure of this thesis is defined as following:

- a. In the first chapter, the generalities of the study, the expected objectives of the research and the aspects of innovation are discussed.
- b. The second chapter explains a general overview of available methods in this field. It also explains the algorithms used in this thesis.
- c. In the third chapter, the proposed method is described, and the tools used in this study are described in detail.
- d. In the fourth chapter, the proposed method is evaluated.
- e. In the fifth chapter, the proposed methods are compared with other methods, works and also with each other.

f. In the sixth chapter, the conclusion and future work proposals are illustrated.

2. LITERATURE REVIEW

2.1 INTRODUCTION

Digital images have a great role in today's life. Over recent decades, many industries and applications require special imaging techniques. Each of these special imaging techniques needs its own tools for processing while working with the tools requires understanding the concepts of image analysis. Infrared imaging is one of the imaging techniques that is not generally applicable, but its value is due to its application in many sensitive technologies. Unfortunately using the common visible and color imaging tools could not extract useful information by processing the infrared images. There are available methods for processing these images. One useful method is based on machine learning algorithm. Therefore, this study proposes a method to process infrared images which is based on machine learning algorithms.

In this chapter, first, concepts of infrared images with definitions of image processing techniques are presented. In the second part of this chapter, the image segmentation methods used for clustering part of this study are explained, then the classification algorithms necessary for extracting the ROI are described. Finally, part of the works done in this area is also introduced as the literature of this study.

2.2 IMAGE PROCESSING

There are two types of image processing: analog and digital image processing. Nowadays image processing is rather referred to digital image processing than analog part. In general, there are two types of images: analog images and digital images.

The digital image processing is a field of computer science working on digital images acquired by digital cameras or scanners. This field includes two branches: Image Enhancement and Computer Vision. A digital image is an input of the functions performed in digital image processing image and it's usually a two-dimensional image. The output of the function depending on the purpose could be either an image or the extracted feature of the image. Image enhancement advantages the acquisition tools such as filters for eliminating the noises, better visualization and adjusted contrast within an image. On the other hand, computer vision includes the techniques for analyzing and manipulating an image for better perception of the structure and the content of an image.

The extracted characteristics could be benefited in different technologies such as Robotics (Gilbert et al. 2005; Gonzalez 2002).

There are three main tasks in the image processing: preprocessing, enhancement and displaying the image or the features of an image.

The main operations in digital image processing (Marques, 2011) :

- a. Geometric Transformations: such as resizing, rotation .
- b. Arithmetic and Logic Operations: The arithmetic operations are used for different purposes such as extracting the differences between images or finding the mean of two images.
- c. Color Enhancement: Brightness and contrast enhancement and adjustment of the color space.
- d. Aliasing and Image Enhancement: The aim is to filter the signals with the frequencies above the sampling rate.
- e. Compression: Compression techniques are used to decrease the size of an image.
- f. Image Segmentation: Segmenting the image into meaningful parts.

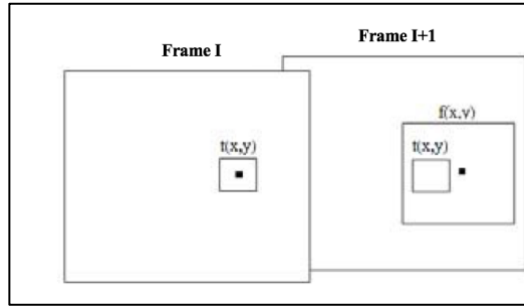
2.2.1 Image Processing Applications

Image processing methods have been applied in many different sciences such as industry, medical fields, security and surveillance monitoring. Some of the applications are mentioned in this section briefly (Iskan et al. 2009):

Pattern Recognition: The goal is to identify and extract a pattern with specified features and categorize the data. Identification of letters or numbers of a text or a license plate are some of the common examples about utilization of pattern recognition (Bhanu, 2005).

Motion Tracking: There are various ways to track a moving object in a video sequence. One of the common methods is the correlation function in two consecutive frames. In the first frame one or more points with a window around them are selected. While a search window in the second frame is determined. By selecting the windows' correlation around each point and the correct determination of the search window in the next frame, the window's correlation around each point in the next frame could be calculated and the location of the maximum correlation could be defined as the new pixel location.

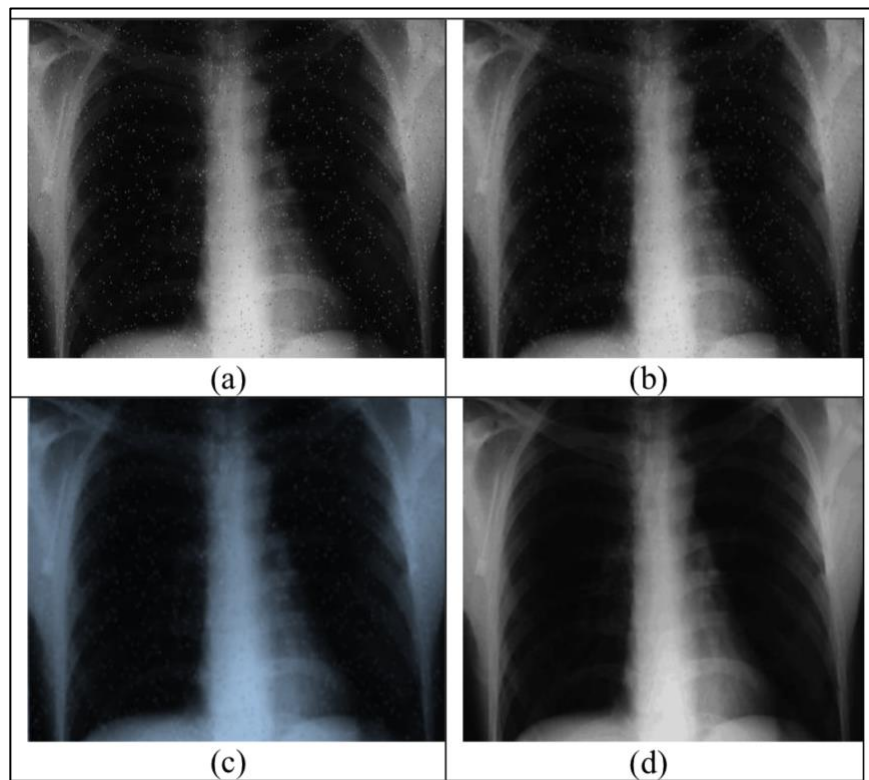
Figure 2.1: Motion tracking



Medical Applications: Image processing knowledge is applied in many different medical fields (Garge et al. 2009; Iscan et al. 2009). Some of the common medical applications are listed below:

- a. Quality Enhancement of Thermal Images: Figure 2.2 indicates this process.

Figure 2.2: Medical image Processing applications

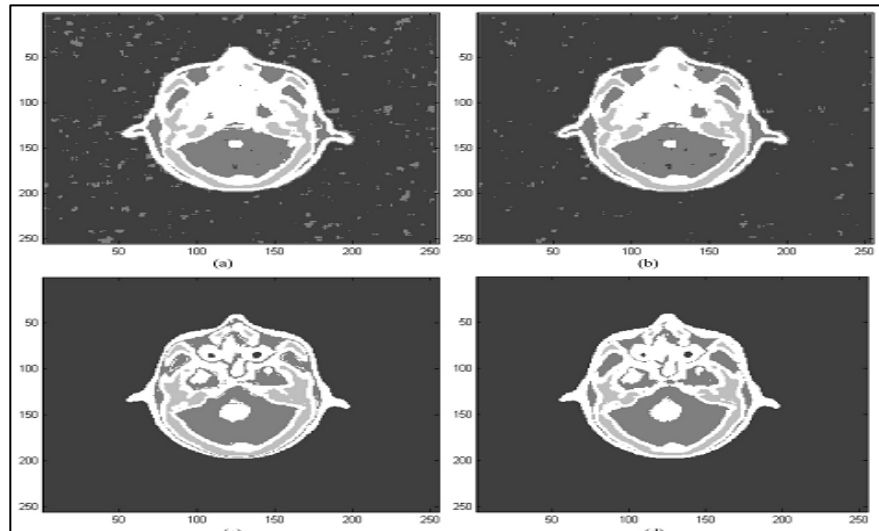


Source: Jambhorka, Sagar et al., 2012 p.310

- b. Separating the distinctive tissues from each other: Due to distinct characteristics of different tissues such as permeability the distinction of different tissues is possible with image segmentation techniques. By the use of image processing technique, the

identification of cancerous tissues and locating the exact place of brain tumors are practical (Iscan et al. 2009).

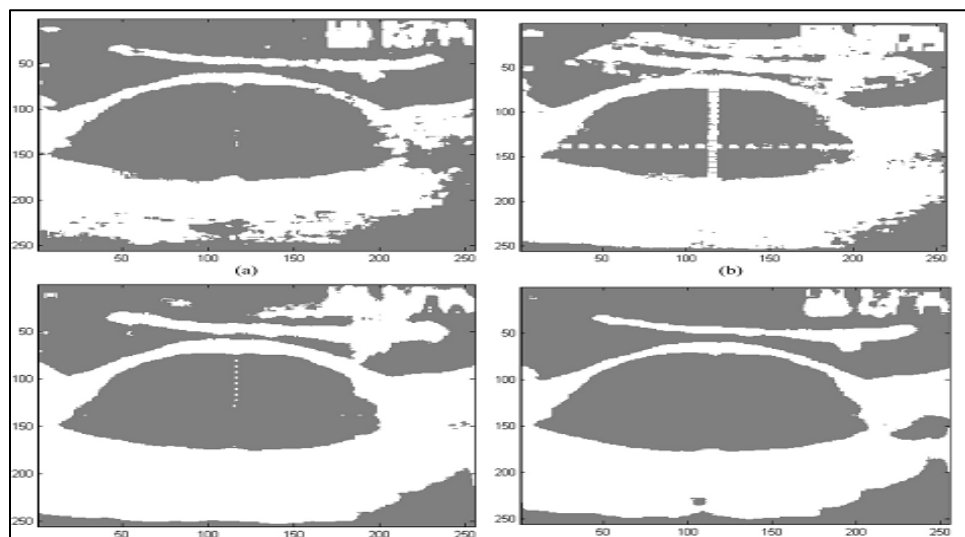
Figure 2.3: Distinction of various tissues from each other



Source: Iscan Zafer et al 2009 p. 897

- c. Measurements of Sonographic Images: Image Processing is used to calculate the distance and surface values of ultrasound Images.

Figure 2.4: Image processing application for ultrasound measurement



Source: Iscan Zafer et al 2009 p. 895

- d. Computer-assisted Surgeries: By using computers as surgeons' assistants two/three dimensional models of tissues or organs are obtained and surgeons could be guided throughout operations.

Figure 2.5: Computer-assisted surgeries using image processing technique

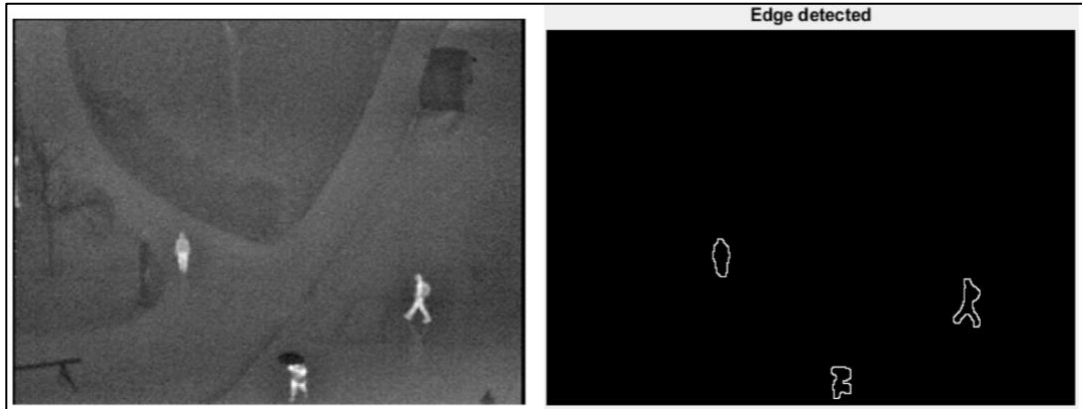


Source: <http://www.futuretechnology500.com/index.php/future-medical-technology/robotic-surgery-advantages-and-disadvantages/>

Military Applications: Currently, many military systems are equipped with cameras and image processing techniques. Some common utilizations of this knowledge are explained below (Wang et al. 2014; Dimitris et al. 2003):

Long range precision missiles apply image processing techniques with the use of GPS (Global Positioning System) data. Systems that lock on the target with predetermined specifications (aircraft, tanks, ...). Unmanned aerial vehicles driven by image processing techniques are used for missile shooting and launching purposes.

Figure 2.6: Intercept targets in military applications using image processing

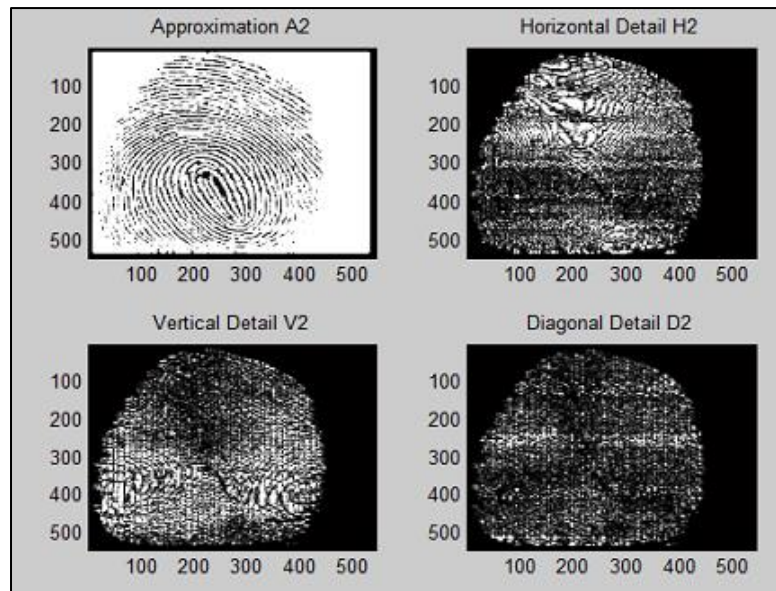


Industrial Applications: Image processing knowledge is applied rapidly over the past few years in this field:

- a. Control and guidance of the manipulators
- b. Separation of chemicals with different colors
- c. Measurement of leather surfaces
- d. Quality control of the factory products

Identification and security systems: Fingerprint recognition, face recognition and iris recognition are some of the common image processing applications (Kambli et al.2010).

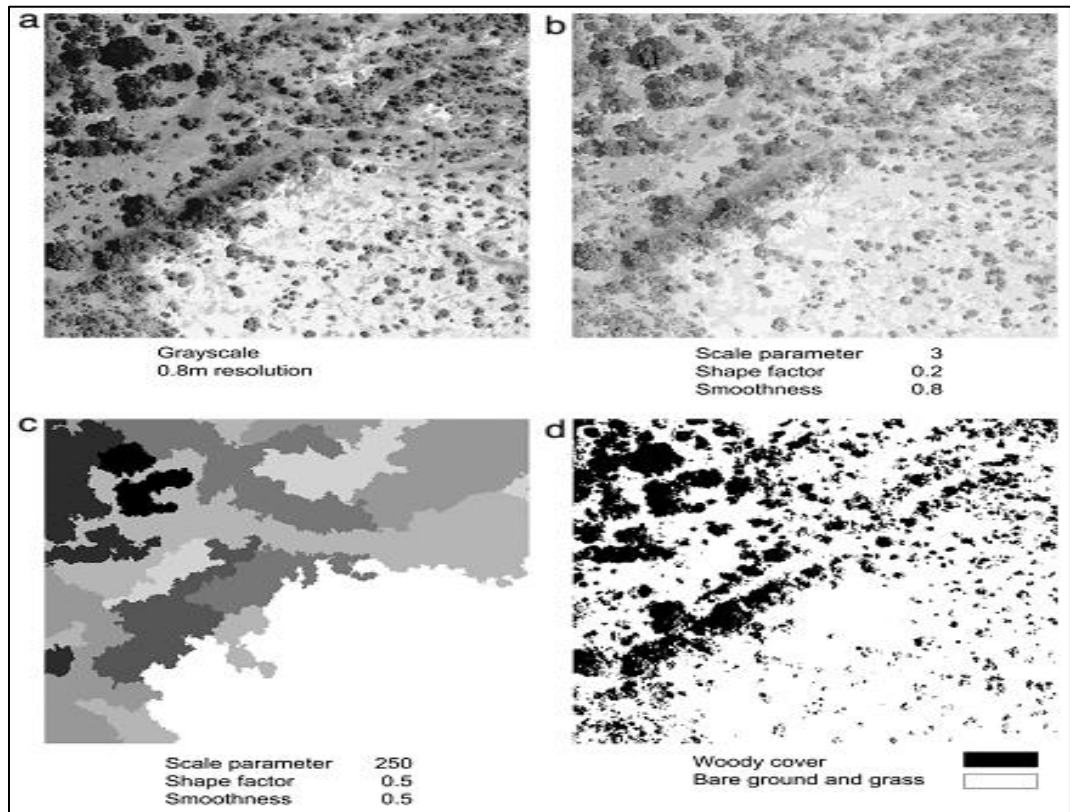
Figure 2.7: Image processing in security systems



Source: Kambli Mansi et al. , p.920

Remote Sensing Systems: Image Processing methods are benefited to extract meaningful information from satellite images. The separation of different graphical zones (sea, land, farms, mountains) are some of the related instances (Blaschke 2010).

Figure 2.8: Image processing in geographic systems for identifying cover crops

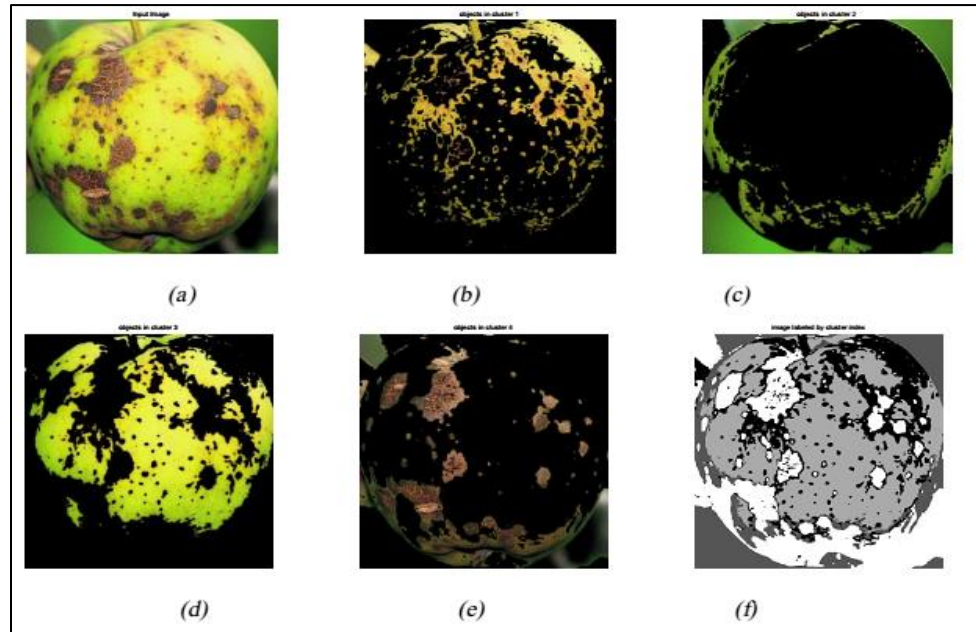


Source: Blaschke, T. p.8

Agricultural Image Processing Applications: The food industry is one of the important industries that mainly use machine learning algorithms:

- Classification of the agricultural products
- Separation of the defective agricultural products
- Packaging the agricultural products
- Identification of plant pests
- Calculating the agricultural crops

Figure 2.9: Separation of the defective fruit surface by image processing



Source: Dubey, et al. p.8

The main steps for defective surface detection of a fruit are:

- a) Imaging
- b) Feature extraction (Image Processing)
- c) The extracted features are applied in proper algorithms

2.2.2 Imaging

According to the variety of image processing applications, the proper imaging technique relevant to the utilized application is required:

Imaging Methods:

- a) Imaging with conventional cameras (Morcol et al.2010)
- b) Satellite imaging (Blaschke 2010)
- c) Imaging using sound waves (Iscan et al. 2009)
- d) Imaging using X-rays (Garge et al. 2009)
- e) Imaging with infrared cameras (Wang et al. 2011)

2.2.3 Pre-processing

The raw images obtained from the imaging device have many problems. The imperfection of the imaging devices is the reason for the errors and low quality of the images. Pre-processing consists of four general methods. By the use of Pre-processing methods, the

visibility of the image is enhanced. Any pre-processing operation requires information about the image, camera and also the surrounding environment (Gonzalez et al. 2002). The Image preprocessing methods classified in different categories are described in the Table 2.1:

Table 2.1: Pre-processing categories

| Type | Description |
|---------------|--|
| Brightness | The enhancement of the pixel brightness |
| Local Binary | Considering small neighborhood around an image |
| Geometric | The aim is to correct the geometric falsification due to various coordinates |
| Global Binary | Using overall information of the whole image |

- a. Brightness: This amendment includes the grayscale and the pixel intensity.
- b. Geometric Transformation: In general, there are two types of Geometric transformation: One is relevant to the system such as camera angle, while the others are related to random noises and could be related to the sensors.
- c. Local/Global pre-processing: Local pre-processing is due to the processes done on each pixel with its neighborhood pixels. But Global pre-processing requires data of the entire image.

There are other pre-processing categories for image pre-processing. One category includes two types of Geometric and Radiometric transformations. The radiometric transformation refers to the type of transformation in which the main target is the elimination of atmospheric and sensor noises. Image Enhancement is also considered as one type of the image pre-processing methods. In general, there are two types of Image Enhancement: Spatial and Spectral. Spatial Enhancement is Filtering, and Spectral Enhancement is Stretching. Noise removal itself is also a separate category.

Preprocessing procedures:

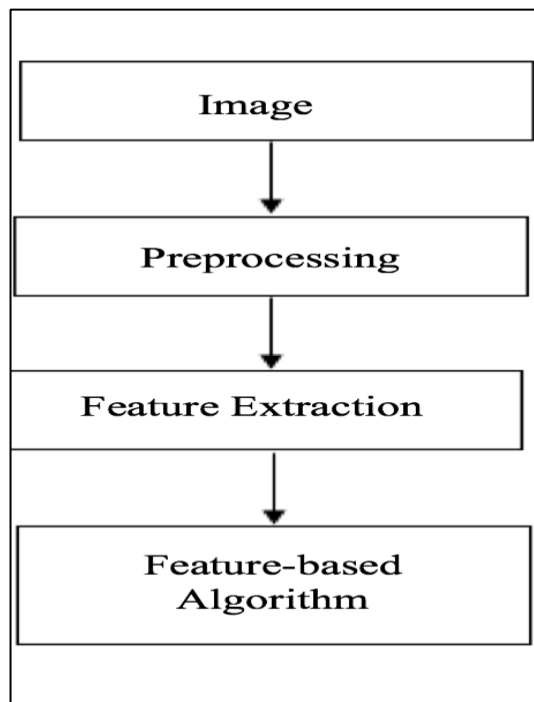
- a. Data Cleaning: The purpose of this section is to eliminate the existing noise and removing the existing conflicts between the data.
- b. Data Reduction: Because of the use of different databases, additional and sometimes duplicate information may be created among the data. By using correlation and clustering algorithms this redundancy could be removed.

- c. **Data Transformation:** The range of the attributes may not be the same. For example, the value of one attribute may be between one to ten while the other one could be between one to one thousand. Therefore, the normalization is required.

2.2.4 Feature Extraction from Images

The most important image processing task is the extraction of the proper features for each application. The steps of working on the image are as follows:

Figure 2.10: Feature extraction steps



Source: Seema, et al, p.50

After imaging step, the image is sent to the preprocessor to remove unwanted data and noise from the images. The purpose of extracting attributes is to reduce image data by using certain properties such as color, texture, or shape. Some of the shape properties include coexistence matrix, fast Fourier transform, and rapid wavelet transformation for fruit recognition. Some of the color properties includes mean, variance, skewness and elongation. Textural features are entropy, energy, contrast, and correlation (Mohammad et al. 2016). These Features are introduced briefly:

- a. **Edge-based features:** In this method, the map and plot of the image edges determine the objects' features. The advantage of using edges is the steadiness. Using the edges as features has advantages over other features. The edges have steady features

and they are resistant to light conditions, color changes of objects and outer texture of objects. Their changes do not affect the edges as well. Edges also define the boundaries well. Therefore, feature extraction is done with a great precision especially in crowded backgrounds with many objects. Among numerous edge detection algorithms Canny, Sobel, Roberts are the popular algorithms (Seena et al. 2015).

b. Morphological features: Morphological features play a great role in classification purposes. The analysis of the morphological features starts with fruit range detection. there are a lot of morphological features for extraction. Fruit border as a morphological feature is connected to the fruit dimensions (Seema et. al 2015). The features are categorized into two main groups. The first groups' components are round fruits such as orange and apples, while the other one consists of banana and carrot. Area is one of the morphological features. the area is calculated by the following formula:

$$Area = \sum_{i=1}^m \sum_{j=1}^k f(i, j) \quad (2.1)$$

k is the number of columns, m is the number of rows and the f function is calculated by this formula (Mercol et al. 2008):

$$f(i, j) = \begin{cases} 1 & \text{if } (i, j) \in I_n \\ 0 & \text{o.w} \end{cases} \quad (2.2)$$

c. Morphological Image Processing (Morphology): The aim of applying morphology is to reduce the flaws of an image by the use of shape features (Seema et. al 2015). These algorithms process the binary images. There is a structural element moved across all the image pixels. Generally, there are two types of operations that will affect the resulting image:

i. Erosion: The erosion function is described by simple formula if the structural element could fit the image or pixels, the output value is 1 otherwise it's zero.

$$g(x) = \begin{cases} 1 & \text{if structural element fits the image} \\ 0 & \text{else} \end{cases} \quad (2.3)$$

ii. Dilation: The dilation function is described by simple formula if the structural element could hit the image or pixels, the output value is 1 otherwise it's zero.

$$g(x) = \begin{cases} 1 & \text{if structural element hits the image} \\ 0 & \text{else} \end{cases} \quad (2.4)$$

d. Color-based features: color is one of the basic features that human eye uses to distinguish the objects from each other. Morphological features may cause misinterpretation according to the similarity between the fruits in the same group. The similarity between banana and carrot is an instance. Therefore, color models such as Hue Saturation Intensity (HSL), Red Green Blue (RGB) could be used to separate these objects (Seema et. al 2015).

e. Textural features: These features are extracted according to statistical concepts. The main applied matrices are gray level and co-occurrence matrix. In this method the neighboring points with the equal gray level are compared throughout the image. (Mercol et al. 2008)

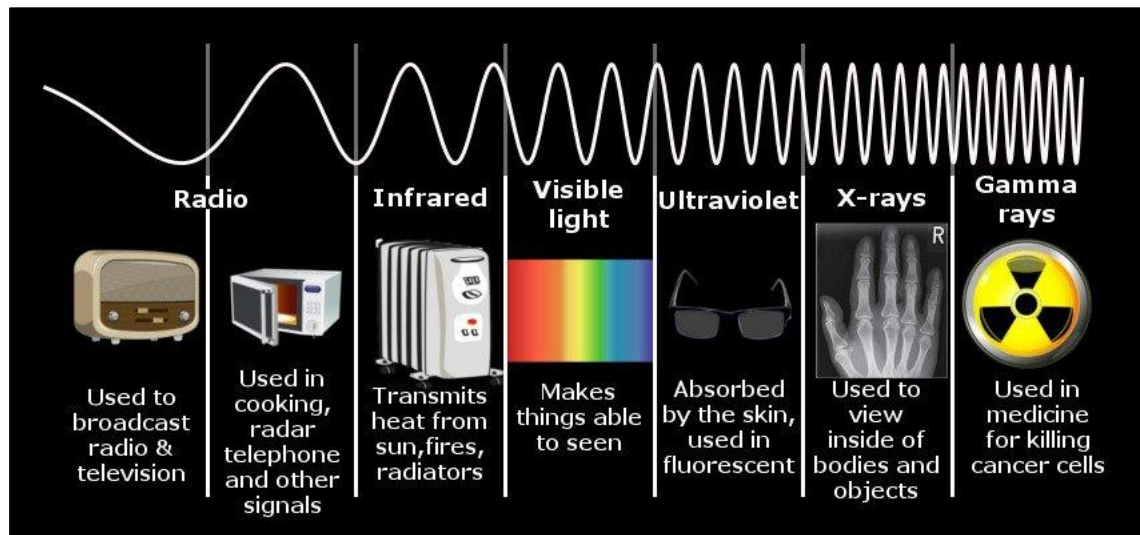
Statistical concepts:

- i. **Contrast:** The contrast of an image (often known as variance) calculates the contrast level between any point (pixel) and its neighbors. The occurrence matrix is used for calculation.
- ii. **Correlation:** It calculates the relation between a pixel and its neighbors.
- iii. **Energy:** often known as "uniformity", "energy uniformity" or "second order torque", which is the sum of the squared related components in co-occurrence matrix.
- iv. **Homogeneity:** This value indicates the closeness level of matrix components to the original diameter.
- v. **Skewness:** It measures the asymmetry level around the mean and it's standardized according to third torque. The zero value expresses symmetry (non-skidding), the positive value shows that the skewness is to the right and if the value is negative, the skewness is to the left.
- vi. **Kurtosis:** It measures the distance of the data from the normal distribution, and its value equals to the fourth-order central torque of a distribution. The value of normal distribution is equal to three. If the Kurtosis is greater than three, it means that the distribution is smooth and if it's less than three, the distribution is inverse.

2.3 THERMAL IMAGING

The temperature of the human body is a great source of a persons' health. Because first of all the human body has a specific temperature different than its surrounding environment. There are two temperatures related to the body: the inner body (core) temperature and the outer body temperature. The alteration of the body temperature (33-42 °C) is an obvious sign of an abnormality. Thermal systems are applied for the purpose of creating and improving the operational capabilities of military forces in night combat. It is also applied for detecting and tracking goals that are visually hidden and camouflage. Thermal imaging systems are part of the passive systems that operate in the mid-infrared region of the electromagnetic spectrum. All objects emit electromagnetic wave which is directly related to their temperature (Pasagic et al. 2008). Infrared wave is an electromagnetic wave with the wavelength between 700 nm and 1mm. This radiation is between microwave and visible light. According to the Plank law any object with the temperature above absolute zero (-273 °C) emits energy that could be recorded by thermal cameras as a black and white image (Pasagic et al. 2008).

Figure 2.11: Electromagnetic waves



Source: <https://www.scienceabc.com/pure-sciences/why-are-infrared-waves-associated-with-heat.html>

Thermal imaging systems are divided into two kinds of cooled and uncooled cameras. Cooled thermal cameras have higher temperature resolution and higher temperature sensitivity so the images have better quality than non-cooler cameras and they have higher prices. The uncooled camera sensors are working at the room temperature while the

temperature of the cooled thermal camera sensors is decreased to cryogenic temperature (-32 degrees F) (Pasagic et al. 2008). The reason for high resolution is because of this low temperature of the working unit. Working at low degree in small dimensions need high quality. In general, thermal cameras require some time to adjust the sensors' temperature. The amount of required time depends on the environment temperature. The thermal images could be obtained during the day and night. As the earth's temperature is stabilized throughout a night, the distinction between the object are better at earliest dawn. Spain and United States of America (USA) are the first countries that used these systems at World War two.

2.3.1 Thermal Imaging Components

- a. Composite thermal system: This unit is responsible for collecting thermal radiation of the object, focusing it at a point, and creating a thermal image. Thermal cameras, like the night vision cameras, consist of several lenses and mirrors, but their structures are different. In these cameras, materials transparent to infrared radiation (such as germanium and silicon) are used (Pasagic et al. 2008).
- b. Detectors: Detector is an element that absorbs infrared radiation collected by a set of objects that changes one of its electrical properties (electrical conductivity, resistivity or volte formation) by absorbing this radiation, and this alteration causes the creation of an electrical signal. After transforming the infrared photon into electrical signals, these signals are amplified and processed by the camera's electronic component, then by the means of devices such as light emitting diodes, liquid crystal diodes or micro-monitors the signals are converted into photons with the visible light wavelength. Each detector element can only transform one point of the object into a visible image. Therefore, in order to obtain a two-dimensional high-quality image, the dimensions of these numerous elements and the distances between them must be very small. Because of these tiny element structures, it is very difficult to construct detectors and they're often produced as linear arrays instead of two-dimensional arrays. A linear array can only represent a line of the target thus a scanner is used to have a two-dimensional image. Detectors are divided into two groups of thermal and photonic or quantic detectors according to the electrical signal production method. So far, most military thermal imaging systems have used Mercury Cadmium Telluride Detectors (HgCdTe) (8-12 mm) for hot and dry area detection and Indium Antimonite

Semiconductors (InSb) (3.0-3.5 mm) for the detection of moist regions such as beaches and seas. Although cooled thermal cameras have great efficiency, due to their high price, size and weight these technologies are preferably replaced by uncooled cameras with the same function (Pasagic et al. 2008).

- c. Scanners: In some thermal imaging systems, there is a scanner whose task is to transfer the target plate information to the detector. In fact, the scanner transfers different points of data in time and line-by-line to the detector.
- d. Electrical circuits: The circuits include power supplies, biases, amplifiers, processors and displays.
- e. Opto-mechanical (eyepiece) system: The eyepiece system enables the observer to see the image.

2.3.2 Different Generations of Thermal Cameras

Generation zero (Woolfson, 2012): A thermal camera built with a single-element detector or a linear array with a small number of elements, is called the zero generation. In this system, two horizontal and one vertical scanners are required.

First generation: If a thermoset camera is built with a very long linear array, it's called the first generation (Woolfson, 2012): In this system, only a horizontal scanner is needed, that is the reason for manufacturing most of the black and white cameras using this technology.

Second generation (Woolfson, 2012): It includes cameras with a long-term, multi-linear array. In this system, only a horizontal scanner is required. The image of these cameras is not significantly different from generation one.

Third generation (Woolfson, 2012): It refers to a camera with a two-dimensional array of detectors with a high number of elements. This system no longer needs a scanner. This generation is the latest generation of thermal cameras that are fully developed in military systems and used in large industrial countries. The distinctive image of this generation of cameras and the ability of separating the element colors from each other have distinguished them from their previous generations.

2.3.3 Effective Factors on Image Quality

Factors such as noise (system noise, background, etc.), atmospheric environments, system technical specifications, distance, dimensions, etc., cause restrictions on the operation of the camera and therefore selection and design are more complicated (Woolfson, 2012).

In general, some of the main factors that affect the quality of the image can be summarized as follows (Woolfson, 2012):

- a. Monitor: The factors that affect the monitor are related to the radiation, contrast and distance from observer.
- b. Page Elements: factors such as target specifications, background specifications, movements and reflections.
- c. Specifications of the thermal image system: factors such as the resolution, sensitivity, noise and output of the camera.
- d. Atmosphere Transmission Factor: factors such as haze, rain and dust.
- e. It should be noted that in the discussion of picture quality, much of the research is done on two issues of spatial resolution and temperature sensitivity.

2.3.4 Accuracy and Recognition Factors of Images

The level of utilization of images depends on the image quality and the ability to obtain information from it. For the image and its information, four precision or diagnostic steps are defined, which are (Pasagic et al. 2008):

- a. Detection: Sensation or detection of an object that may be a target. (Usually observing the object as a spot)
- b. Orientation: Detects the overall dimensions of the system. (Latitude and Width Detection)
- c. Recognition: The diagnosis of the target category and the ability to determine the group of objects. For example, showing that the object is an airplane or a helicopter.
- d. Identification: Identification and distinguishing the target among the objects belonging to its group. (the type of the object's aircraft).

In fact, from the first to the fourth step, the number of pixels increases, and the image quality will improve.

2.3.5 Main Effective Factors in Thermal Imaging

As mentioned, thermal imaging systems have different generations, and each generation has its own features that are used in the design of cameras. Some characteristics are also common in system performance and the result is important for the user.

The most important features of thermal cameras are (Woolfson, 2012):

- a. **MRTD (Minimum Resolvable Temperature Difference):** The thermal imaging system response depends on the sensitivity and spatial resolution. In order to assess the image quality in terms of sensitivity, resolution, dependency and interaction between them, a feature called MRTD is defined. That is the lowest target black body temperature difference from the background which could be measured by the system. The MRTD is limited to the sensitivity of the system, which means when the temperature difference is less than a minimum value, the object cannot be detected. In plain language, the MRTD is a camera feature that determines at least which sensitivity (or temperature difference) is required at any frequency (Rayleigh Criterion).
- b. **Resolution:** In many cases, spatial resolution is considered to be the only determining factor of the image quality. In fact, spatial resolution is the system's smallest receivable part. The resolution is sometimes expressed by the instantaneous eyesight. The interpretation of resolution power depends on the type of application. On the other hand, spatial resolution includes the effects of the system's target and noise contrast. By the way, there is a clear difference between spatial resolution (ability to see detail) and the ability to see anything (detect).
- c. **Sensitivity:** Sensitivity is the smallest signal that can be detected by the system, or in other words, a signal that produces a signal-to-noise ratio at the output of the system. The sensitivity depends on the ability to capture the optics and detectors' ability of noise detection, and it's independent of resolution.
- d. **Field of View (FOV):** The maximum angular field (horizontally and vertically) in any functional position visible on the display is called the field of view. The choice of field of view usually depends on the type of application, technology, detector and scanner properties.
- e. **Instantaneous FOV:** This property is the angular component that let the systems receive information and it determines the system's resolution. The smaller feature

results in better receiving, as long as it could provide enough energy for detection. The larger field of view causes the larger Instantaneous field of view, and this decreases the resolution of the system.

- f. Noise Equivalent Temperature Difference(NETD):** This feature indicates the temperature sensitivity of the system and it's the least variation between body temperature and background, which produces noise in the output signal. This attribute depends on the detector's specification, the optical-atmospheric transmittance, and the noise of the system.
- g. F-Number:** The focal number is the ratio of the focal length to the diameter of the lens in an image forming system. In fact, the focal number expresses how much light is collected by the speed of a lens.
- h. T-number:** A T-number expresses the speed of a lens, assuming that the lens transmits all the light emitted from the subject. In fact, various lenses have different T-numbers. Lenses with the same focal numbers may actually have different speeds. For two lenses with the same T-number, the resulting images are the same with equal brightness.

2.3.6 Selection of Wavelength Region for Thermal Cameras

According to the level of object's emitted radiation at normal temperatures and atmospheric constraints, only two regions (3-5 μm) and (8-12 μm) can be used for passive photography. If targets are hotter than the intended environment (such as the exhaust of the missile), since the maximum wavelengths of these targets are shorter and the radiation power in that area is high, it is better to use a camera with a wavelength of three to five Micrometer. In general, without taking into account the specific application, one cannot make a region superior to another. Of course, the wavelength range of eight to twelve microns is a fine range (Woolfson, 2012).

2.3.7 Atmospheric Effects on Thermal Cameras' Performance

To see objects on the surface of the earth, the electromagnetic waves emitted from the surface of the object pass through the air and reach the camera. Since air is a mixture of different gases, water vapor and particulates, it absorbs and also spreads some of these waves depending on the wavelength. Considering different conditions, such as the amount of gas, wind, temperature and other atmospheric conditions during the emission

pathway, the assessment of these effects is complicated. The amount of contrast (sharpness) of the object or background is the factor that affects the visibility of objects. The atmospheric conditions reduce the contrast according to the distance from the viewing point. The atmosphere imposes important constraints on the performance of electro-optical systems. In fact, the environment could be considered as one of the most important components of an optical system. Today, due to the fact that the quality and capabilities of detection systems and radiation sources (such as lasers) have enhanced significantly, the most important limitation on system performance is usually the atmospheric environment (Woolfson, 2012).

Most of the atmospheric disturbances that affect the radiation transmission and performance of the thermal imaging system are:

- a. Slippage of radiation (absorption, dispersion) that has the greatest effect and limits the range of systems.
- b. Radiation of the environment and the infrared region
- c. The deviation of the actual target location
- d. Radiating Modulation

The atmospheric passage rate is not the same for all wavelengths, and the objects at ambient temperature with only a perceptible spontaneous radiation in two regions (3-5 and 8-12 microns), are not absorbed by atmospheric influences and are suitable for thermal imaging. As infrared radiation is less absorbed than visible light in the presence of mist and smoke of the Earth's atmosphere, therefore these cameras can be used in adverse weather conditions.

Thermal cameras have different applications in industrial, military and nonmilitary cases, the most important applications of which are:

- a. Observations and operations at night
- b. Guided missiles
- c. Intelligent and identification operations
- d. Helping the planes for landing and taking off
- e. Photography during Nighttime and adverse weather conditions
- f. Photographing of camouflage and hidden objects
- g. Fire control system usages

It should be noted that all unmanned aerial reconnaissance aircraft flyable at night-time and any type of weather conditions are equipped with thermal cameras for identification purposes. This system is critical for identifying enemy forces in the battlefield during operations such as displacement, expansion, division, camouflage, hiding and so on.

Advantages of thermal cameras instead of night vision cameras

- a. Ability to create a picture at night and day
- b. Online image transformation to the receiving data stations. (The image is visible both by eye and the image simultaneously is sent to an external monitor).
- c. Failure to reveal by night vision systems: Some night vision cameras require a source of help for the purpose of seeing the target (Active System), which can be seen by night vision systems. But the thermal camera does not need an external source and it observes objects' own radiation.
- d. Thermal cameras do not require light for their functions, therefore even placement of a 18,000-watt projector in front of the desirable object doesn't alter the output image(Pasagic et al. 2008).

2.4 IMAGE SEGMENTATION

Image segmentation has a great value in image processing field. Because Segmenting an image into some meaningful components help interpretation. The segmentation methods that are applied in this study are as follows:

- a. Clustering
- b. Classification

2.4.1 Clustering

Clustering algorithms are unsupervised learning system in which data is not labeled. In this type of clustering algorithms data is clustered in a way that objects in each group are similar together while they are different from the objects of other groups. The similarity between the clusters' objects is obtained according to the extracted features. There are many different clustering algorithms but according to this study two applied clustering methods are briefly explained here:

2.4.1.1 K-means clustering

In this clustering method data objects are gathered according to k which is the predefined number of clusters and also the selected initial centroids. The algorithm consists of two iterations and it continues until convergence happens. Data objects are clustered according to the initial centroids. The new centroid is then calculated for each cluster this loop continues until convergence happens which means the centroid doesn't move.

The most important problem of this algorithm is the number of clusters which should be defined in advance that has a great impact on the performance of the algorithm and it could highly affect the results. The other disadvantage of this algorithm is due to the impact of centroid selection which also could affect the algorithm because sparse centroids could cause unwanted results (Mohammad et al 2016).

2.4.1.2 Mean shift clustering

Mean Shift is a hierarchical non-parametric clustering algorithm, unlike k-means clustering method the algorithm itself figures out the number and the location of the clusters. The important factor of algorithm that should be determined is the radius around the data point and sometimes also referred as bandwidth. Radius is defined around each data point determining the bandwidth in which the points of that cluster is located.

In the second step the mean of these data points is calculated as new cluster center and with this new point a new cluster is considered. The process continues until the mean doesn't change, which means the mean is optimized and converged. In Mean Shift clustering although all points were started as one clusters by continuing the algorithm the mean of the cluster shifts that means the cluster center moves to the point of convergence (Comaniciu et al. 2002).

2.4.1.2.1 Mean shift applications

Mean Shift algorithm is highly applied in discontinuity preserving smoothing segmentation and object tracking, this is very useful in different fields such as military, industry, sports and surveillance cameras. For example, the object could be tracked as moving in by the missiles.

Figure 2.12: Object tracking for surveillance



Source: Comaniciu, D., et. al. 2003. p. 578

The other application is tracking the players in the sport field. In the Figure 2.13 the green and blue rectangle indicate the movement of the player in different directions. The Mean Shift is done for all the frames and the centroid is computed (Comaniciu et al. 2003).

Figure 2.13: Object tracking in a soccer game



Source: Comaniciu, D., et. al. 2003. p. 576

2.4.2 Classification

Classifications are supervised learning algorithms in which data is labeled. Once the proper image features are extracted, the gathered features should be classified in a proper order, so that any instance of the problem space could be placed in the correct group. This step involves methods for matching each of the patterns derived from the feature extraction stage with one of the problem space classes. Number of attributes must be minimized according to relativeness between each input feature vector to one of the reference vectors. The reference vectors are the training dataset variables which are extracted previously from some training samples. Neural Network is one of the most common classification algorithms. The point is that these structures are not necessarily separable and sometimes are used as a combined concept (Singh 2011). ANN algorithm is explained below:

2.4.2.1 Artificial neural network

In general, ANN solve complex problems by using the human brain function method. Common computational methods use the same algorithm. They follow a set of preset commands to solve problems. The processing ability of conventional computers are restricted to define and solve problems, but neural networks are able to find patterns in information that no one ever knew about their existence (Singh 2011).

Neural networks have opened up a new and distinct answer to the problems instead of common methods. Common computational methods follow an algorithm. The algorithms are a set of preset commands used to solve problems except in special cases where the computer needs a series of information, and this limits the processing ability of ordinary computers to the solved problems. These algorithms are beneficial for processing the large number of instances and they increase the analyzing rate. The ANN is used to analyze the problem and find the best possible solution for any unsolved situation.

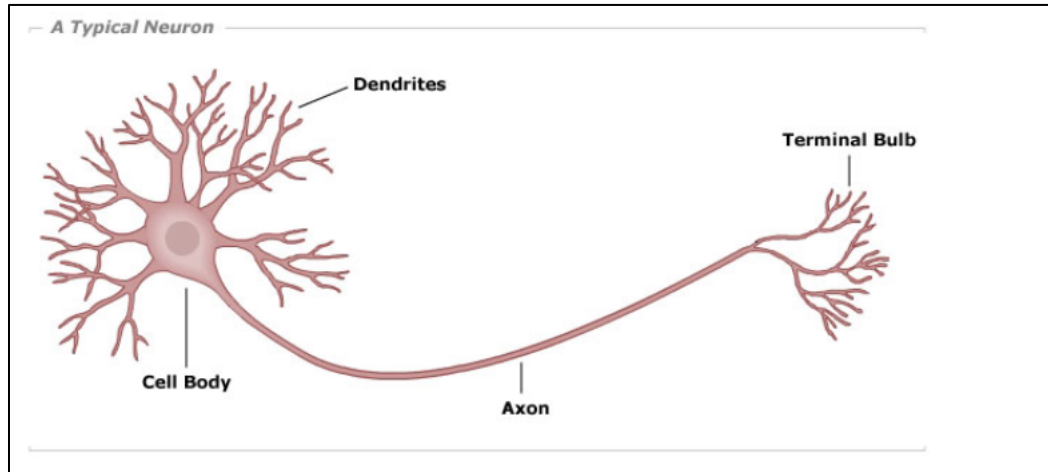
Neural networks and common computational methods are not in competition, but they complete each other. There are tasks that are more suitable for algorithmic methods and there are tasks that are more suitable for neural networks. Furthermore, there are issues requiring a system that is obtained by combining both methods with high precision (Singh 2011).

Artificial Neural Networks provide a different method for processing and analyzing information. But it should not be inferred that neural networks can be used to solve all computational problems. Common computational methods continue to be the best option for solving specific groups of issues such as accounting, warehousing, and so on. Neural Networks are moving in the direction that the tools have the ability to learn and plan. Neural Network structures are capable of solving problems without the help of an expert and external planning. In fact, the Neural Network is able to find patterns in the information that no one ever knew about it.

Since ANNs are developed based on the human brain, their structure which is a bio-network is explained:

Neuron: Neuron is the humans' brain cell which is defined as the main structural element of the brain. In Figure 2.14, a simplified representation of the structure of a neuron is shown. A bio-neuron, after receiving the input signals (in the form of an electrical pulse) from other cells, combines these signals together and, after performing another operation on a hybrid signal, the output appears (Singh 2011).

Figure 2.14: Structure of the humans' brain



Source: https://online.science.psu.edu/bisc004_activewd001/node/1907

As shown in Figure 2.14, the neurons are made up of four main parts: Dendrites, Soma (Cell Body), Axon and Synapse (Nerve Ending). Dendrites are the same components that are scattered around longitudinal fibers from the center of the cell. Dendrites play the role of communication channels for transmitting electrical signals to the cell center. At the end of the dendrites, there is a special biological structure called synapse, which plays the role of connecting gateways for communication channels. In fact, various signals are

transmitted through the synapses and dendrites to the cell center, where they are combined. The mentioned combining operation could be obtained by a simple algebraic action. In principle, even if this is not the case, by using mathematical modeling it can be considered as an ordinary summation action. For this purpose, a special function is applied to the signal, and the output as a different form of electrical signal is transmitted from an axon (and its synapse) through other cells.

The human brain has the most evolved structure among living organisms, the reasons are as follows:

- a. A human brain consists of at least ten to the power of eleven neural cells or neurons.
- b. It is one of the biggest brains among all living creatures. The intelligence factor is not just due to the size of the brain, otherwise the elephant or the whale should be smarter than humans. Human brain intelligence is due to the number of connections between brain neurons. Neuron is the smallest unit of a Neural Network that forms the function of neural networks, and each neuron has several parts:
 - i. Soma or Body: It is modeled as a mathematical function.
 - ii. Dendrites: Function inputs
 - iii. Axon: Function output

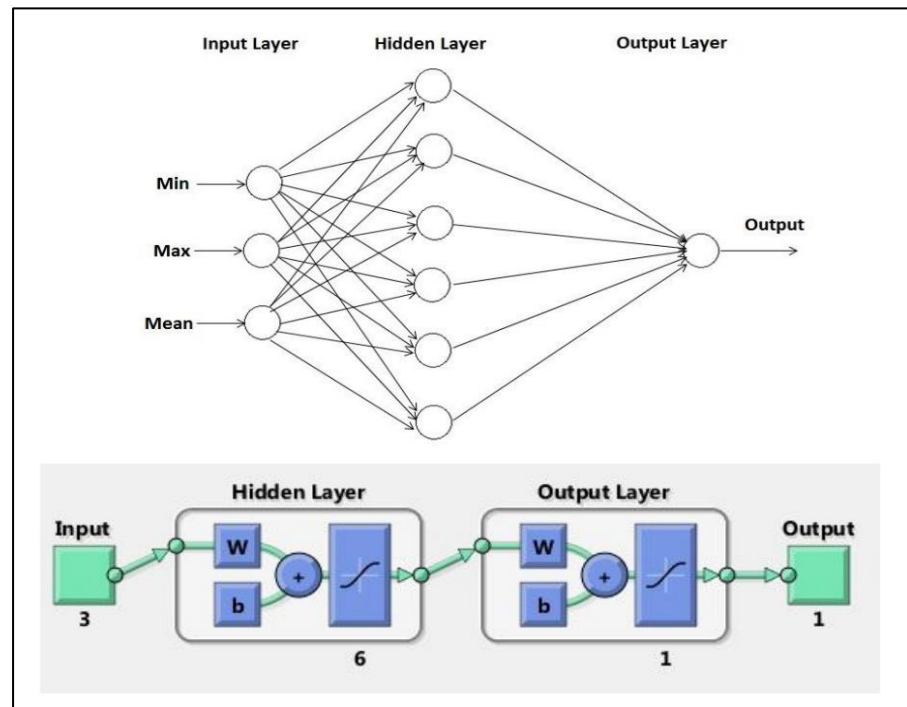
Artificial neural networks process information in a manner similar to the human brain. They consist of a number of superficially interconnected processing elements (the neural cell) that work together in parallel to solve a particular problem and cannot be programmed to perform a specific task. Examples should be carefully selected; otherwise, the useful time is lost, or even the network may work incorrectly. The Neural Network score is according to its ability of solving unknown problems and its performance is unpredictable.

An ANN forms a collection of neurons. The most important factors that differentiate types and applications of the Neural Network include the applied type of neurons, the layout or structure and the input/output intervals. Artificial Neural Networks are a combination of neuronal complexes which are very similar to biological neurons. Therefore, it takes a lot of inputs with different weights and produces an input-dependent output. Biomedical neurons can be either causing it or not.

The structure of the cells in the network is called network architecture. In the architecture of a network, the number of layers and connections between them are important. Network

inputs called "input layer" and network outputs called "output layer" and, if necessary, layers between these two layers are called hidden layers (Gonzalez et al. 2008). Figure 2.15 represents a simple structure of the Neural Network.

Figure 2.15: Simple structure of the Neural Network



- a. Input layer: This layer receives inputs and sends the input signal to the next layer based on its power connection with the next layer. The relationship power of each neuron with another neuron is called the weight of that neuron.
- b. Middle (hidden) layer: The number of interlayers and the number of their neurons is arbitrary. The middle layers must be carefully selected to produce the proper output.
- c. Output layer: Another group of neurons also forms the outside world through its outputs. The Neural Network acts like a function. This function accepts outputs and inputs which are exactly same as the number of input and output neurons accordingly. Among different types of neural networks, some of the common ones are (Sinngh 2011):
 - i. Multi-Layer Perceptron
 - ii. Hopfield Network proposed by Hopfield (1982)
 - iii. Kohonen Feature Map (Kohonen 1997)
 - iv. Adaptive Resonance Theory

The following table shows the correspondence between Artificial Neural Network and Biological Neural Network (BNN):

Table 2.2: Correspondence between ANN and BNN

| Artificial Neural Network | Biological Neural Network |
|---------------------------|---------------------------|
| Neuron | Soma |
| Input | Dendrite |
| Output | Axon |
| Weight | Synapse |

2.5 RESEARCH BACKGROUND

Wei Wang et al. (2011) presented a method based on Cellular Neural Networks and Distributed Genetic Algorithms (DGA) for infrared edge detection. They trained the network using CNN format and distributed genetic algorithm. CNN can be used to process infrared images with special modifications. The results of their experiments showed that the edges detected by CNN-DGA were highly accurate. Similarly, compared to the way CNN has been trained by the Particle Swarm Optimization algorithm, the speed of the proposed method has been significantly improved.

Qingju et al. (2016) Provided an edge detection method in infrared images based on an Ant Colony Optimization algorithm. Edge extraction is one of the most important tasks in detecting infrared images. The Ant Colony Optimization algorithms have properties that can enhance the efficiency of the edge detection system, control the noise with high precision and can extract the right information from the edge. Along with these points, he compared the Ant Colony Optimization (ACO) with the classical Canny Edge detection algorithm. The results of the experiments showed that ACO had high efficiency for the edge detection of infrared images .

Wang et al. (2011) Provided an ACO-based method and a Sobel operator to identify the edge in infrared images. His method used a Sobel operator to control the primary position of the ants in the ACA. This method has been able to detect thin edges well and improve the overall performance of the algorithm. According to the report, the results of the tests indicated a good performance of the proposed method.

Qingju et al. (2016) Using a morphological-Canny compound, provided an algorithm for edge detection in infrared imagery. Effective extraction of the edge curves of infrared images helped detecting the geometric properties of defects. The results of the experiments show that the algorithm has a high anti-noise effect and it has recognized better geometric properties of the edge.

2.6 CONCLUSION

Working with infrared images is very important. One of the important actions in the processing of infrared images is edge detection. Because of providing useful data edge detection has a great value.

Accordingly, in this chapter, first explanations about infrared images were provided. Next, the concepts and definitions of image processing are described along with the image segmentation. Then, the concepts and stratified algorithms were expressed and last, some part of the work done in this field were expressed.

Next in this thesis:

In the third chapter, the proposed method will be explained, and the tools used in this study will be described in detail. In the fourth chapter, the proposed methods will be evaluated according to other cases and will be compared with each other.

In the fifth chapter, the results will be discussed and proposals for future works will be presented.

3. DATA AND METHOD

3.1 INTRODUCTION

In the previous chapters, the importance of image processing and infrared images were explained, and some of the major methods used in image processing were considered. After reviewing relevant methods and identifying existing research deficiencies a new approach is proposed to detect infrared image edges. Therefore, in this chapter, a general research plan is presented. Then, tools used in this research, image segmentation, extraction of expected area using Neural Network and edge detection are discussed. Finally, the method of evaluating the proposed method is expressed.

3.2 APPLIED TOOLS

To implement the proposed method, MATLAB software is used (MathWorks 2015). In this software, based on the demands of the research, Neural Network, clustering and image processing toolboxes were used. The name of this software is derived from the English MATrix LABoratory Label (MATLAB). MATLAB was first designed for better accessibility of matrix software designed by LINPACK¹ and EICPACK² projects. MATLAB is a high-level language initially developed based on C language. It has a high technical capability for computing, and visualization. So, MATLAB is a modern programming environment that includes high level data structures, error-detection tools. It also supports object-oriented programming, and so on. These factors are a great tool for learning and research purposes. This programming language has many advantages over common programming languages such as C for solving technical problems. MATLAB is an interactive language in which the initial data is an array that does not require a dimension and its software package has been commercially available since 1984 and now serves as a standard tool in many universities and industries around the world. It also provides easy-to-use matrix, computational or functional operations, various algorithms and easy communications with other programming languages. MATLAB software has a wide range of applications, including image processing, Neural Networks, control design, Artificial Intelligence, and so on. As stated, MATLAB is written in C for

¹ MATrix LABoratory

² LINear system PACKage

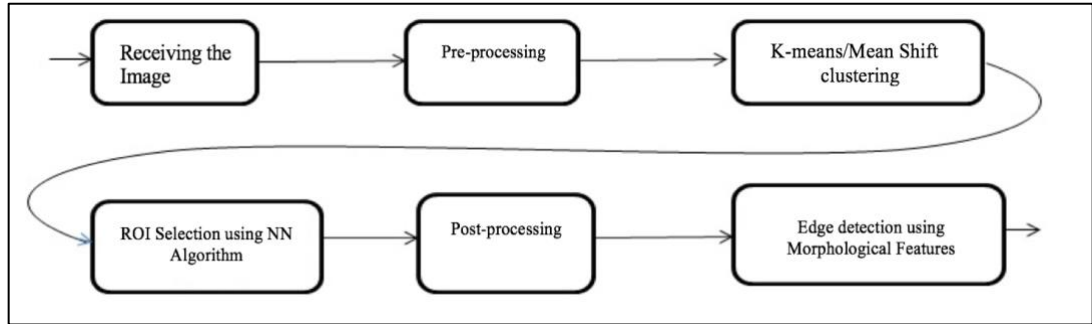
speed and high performance, but its graphical interface is implemented with Java. One of the main advantages of MATLAB is the easy learning ability and the various accessible documents for learning and use. The other programming language applied in this study is R language (R core Team 2017). R is a programming language for statistical computing and graphics supported by the R Foundation for Statistical Computing.

3.3 OUTLINE OF THE THESIS

In general, working with raw images, identifying and extracting desirable properties of any image require fundamental step-by-step processes. This study also focuses on image preparation and edge extraction. In this proposed method the first step includes image segmentation which is applied by two clustering algorithms (K-means and Mean Shift clustering). First, by using K-means clustering, the image is taken from the input unit and segmented into K pieces. Then, using the Multilayer Perceptron (MLP) neural network, which was previously trained with the training dataset, the ROIs are extracted among all available k pieces. ROI is the cluster which includes edges. Then the extracted region, which is converted to the binary image is sent to the edge detection unit. Edge detection unit extracts the edges using the morphological operators. The same procedure is done by using Mean Shift Clustering instead of K-means clustering algorithm. The reason of this choice is due to the shortcoming of k-means in which the number of clusters should be defined in advance and it could cause time-wasting specially in test part in which the test image is sent to the Neural Network algorithm. Mean Shift Clustering has this advantage as it could figure out the number of clusters automatically. This time, after receiving the image and pre-processing step in which possible noises are eliminated image is segmented by Mean Shift clustering algorithm. Then, using the MLP neural network, which was previously trained with the training dataset, the ROI is extracted among all available pieces. Then the extracted region, which is converted to the binary image is sent to the edge detection unit. Edge detection unit extracts the edges using the morphological operators.

The diagram of the proposed method is drawn in Figure 3.1:

Figure 3.1: General steps of the proposed method



In this chapter the Figure 3.1 is explained by selecting two images from the dataset. One sample image is selected to explain training part and one sample image is selected to explain the ROI extraction. All above steps and processing methods one by one are expressed with resulting images and tables. Figure 3.3.a is the example image selected to illustrate the training part.

3.4 PRE-PROCESSING

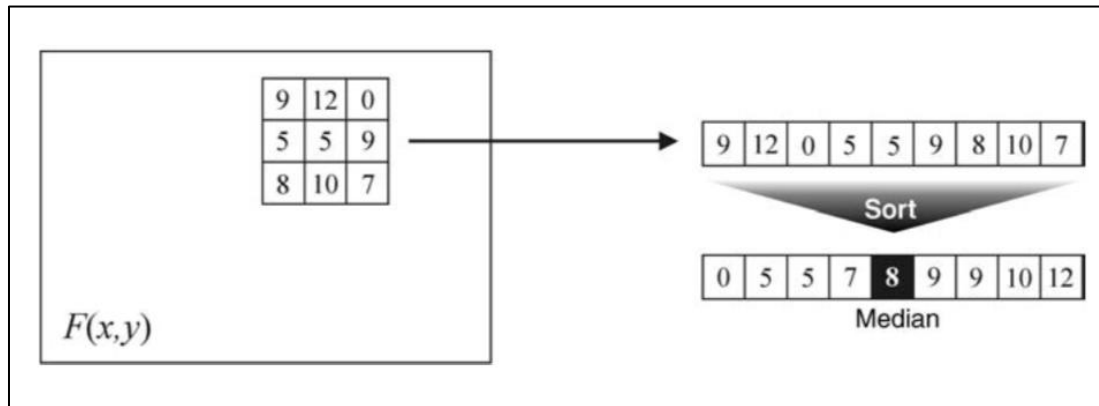
As stated in the previous chapters, infrared (thermal) images have low precision and are disposed to occurrence of various noises, especially to pepper and salt noise. Noises are removed to improve image quality for better image visualization. Removing the noises are also necessary because this noise is a problem for one of the most basic stages of image processing, which is segmentation. In fact, if the image has a noise, the segmentation step is applied with much lower success rate. So, it is best to reduce noise as much as possible before the segmentation step. In this section, salt and pepper noise is introduced, then the removal or reducing method is explained.

If the image contains salt and pepper noise, then black and white spots appear on most parts of the image. These black dots fall on the pixels of the original image and lower the quality of the original image. Based on the percentage of noise on the image, the dispersion of these black and white points is low or high.

One of the ways to remove noise from a digital image is to filter a noisy image. Therefore, if a suitable filter is used to remove noise, then by using this filter, a processed image with less noise or in the best case without any noise will be obtained. The desirable filter is a small image. Usually this image is a square image whose number of rows and columns is an odd number. For example, 3 rows and 3 columns, or 5 rows and 5 columns, and so

on. To remove salt and pepper noise from a digital image, a filter called the median filter is used. This filter has the ability to remove salt and pepper noises. The median filter's job is to arrange all the neighbors of a central pixel ascending and select the middle element of the ordered numbers and replace the central pixel. The process of implementing the median filter is shown in the Figure 3.2:

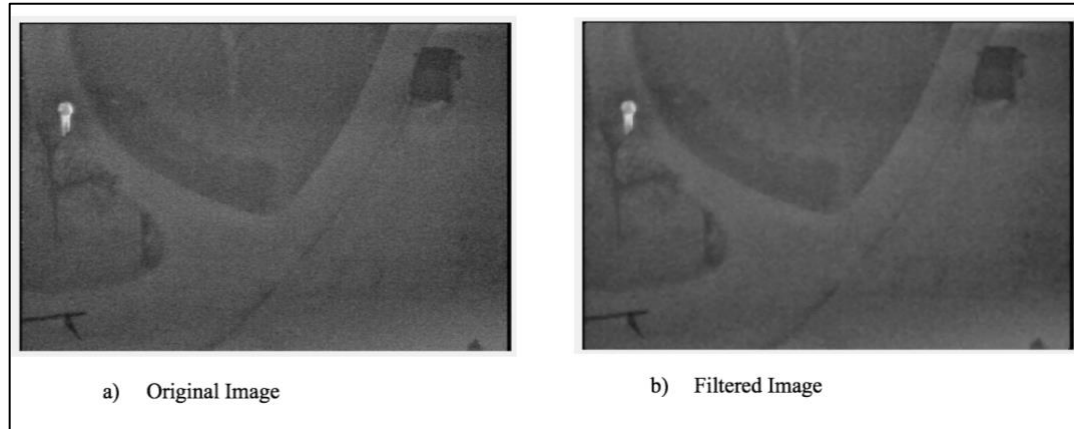
Figure 3.2: Median Filter



Source: Marques, O. 2011, p.217

Based on what is shown in Figure 3.2 the filtering operation is performed. The structure of this filter is based on sorting. Figure 3.3 shows applying the median filter on the infrared image. The image on the left side is the original image and the image on the right side is the filtered one. In this example 3 by 3 window is selected then all the pixel values are sorted then the middle one is selected and the value 5 which is located at the center of the window is replaced by number 8. By this simple idea a lot of noises such as salt and pepper are removed easily. the reason is due to the structure of salt and pepper noises which are unwanted irrelevant noises so smoothing the image could eliminate their effects (Marques et al. 2011).

Figure 3.3: Applying median filter



The applied filter is a three by three matrix.

3.5 IMAGE SEGMENTATIONS

Due to the significance of the segmentation step of this research, the segmentation process of the image in this research is described. After performing the filtering operation (noise removal), the image is prepared for segmentation and divided into several regions. The segmentation in this study is used to separate the different clusters of the image. K-means and Mean Shift clustering methods are separately used for segmentation in this research. For this purpose, two separate programs are developed one using k-means while the other one benefited Mean Shift clustering algorithm. The important point is that for both Neural Network training algorithm and also for testing the dataset the same clustering method is applied. The basic structure of this program includes two part which are training the neural network, testing and evaluation of the method. In this chapter the aim is to explain the method for training and testing. The results are explained in the next chapter. For this purpose, both algorithms will be expressed using a sample image. The NN is trained by 29 images which is selected from the relevant dataset¹.

3.5.1 K-means Clustering

The general k-means clustering steps in this study is presented in Figure 3.4. In this research, k-means clustering algorithm is applied using Euclidean distance.

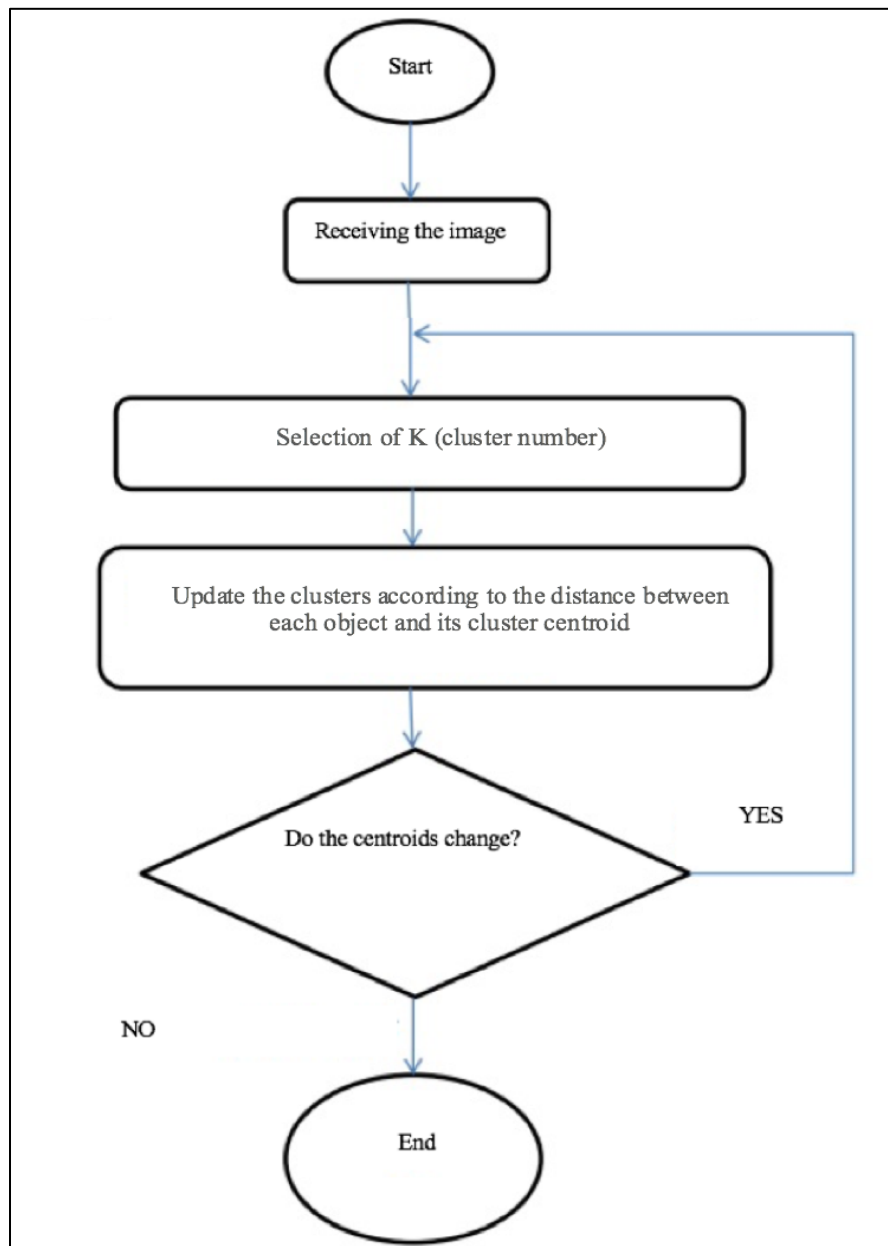
The Euclidean distance criterion is calculated by Formula 3.1:

¹ J.-Davis and M.-Keck, *A two stage approach to person detection in thermal imagery*

$$\text{Distance}(X_j, C_K) = \sqrt{\sum_{i=1}^d (X_{j,i} - C_{k,i})^2} \quad (3.1)$$

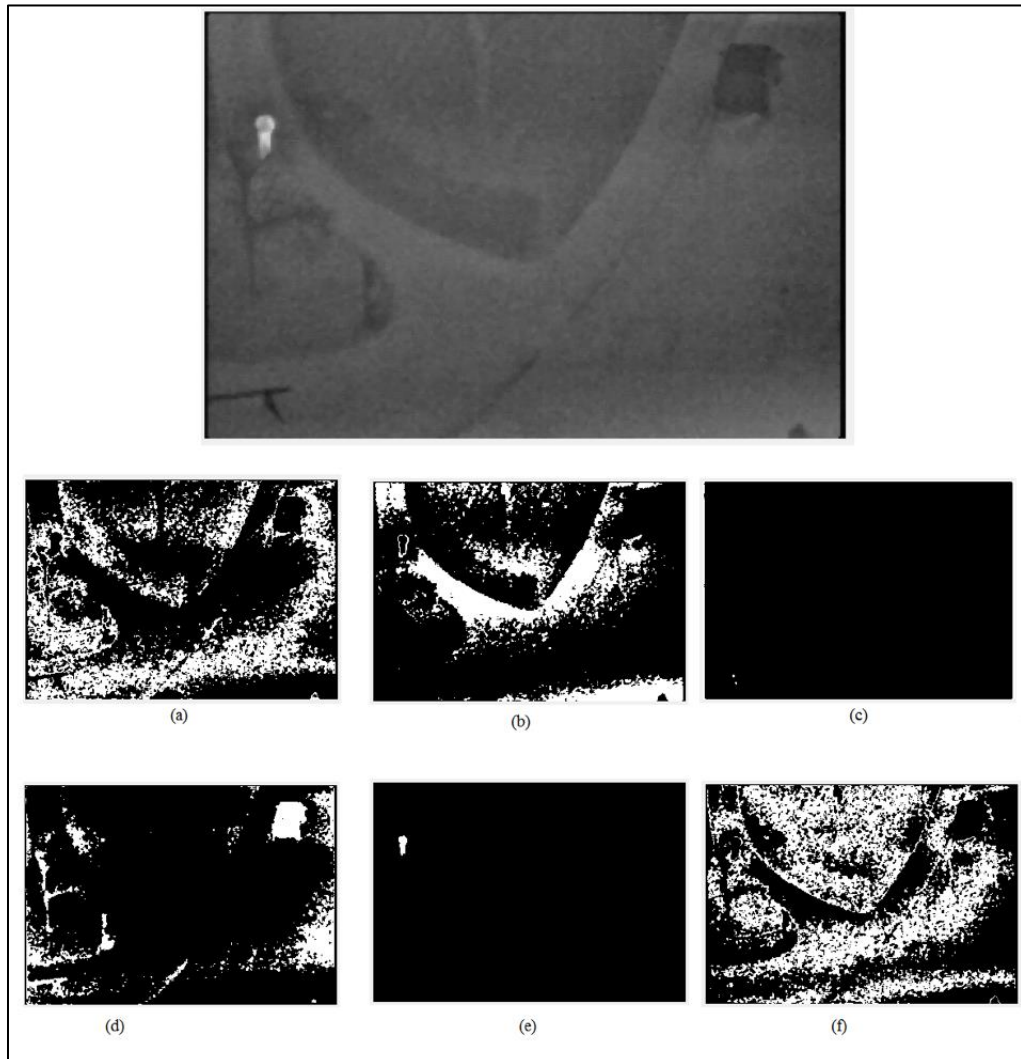
In Formula 3.1, j represents the number, k in this study is the image pixels. C represents the center of the clusters and the X is the desired data. D representing the dimensions of the data is equal to one in this study. Because the image used in this research is gray and it only has a value between zero and 255. It should be noted that the number of clusters in this method should be determined by the user. According to the elbow criterion optimal number of clusters could be calculated.

Figure 3.4: Block diagram explaining the k-means clustering steps



In this part, the results of the image clustering using K-means clustering and segmentation of the image into several regions are presented. Figure 3.5 shows the result of clustering on the sample infrared image. In this Figure, the upper image (gray image) is the pre-processed image and the rest of the images are resulting clusters.

Figure 3.5: Clusters extracted using k-means clustering



Based on Figure 3.5, the original image is segmented into six clusters. One point to note is that, among these six resulting images, only one image should be used for edge detection so that the person on the corner side of the original image could be detected. The object to be found in this image is shown as a red circle in Figure 3.11. Now, by looking at the images (a to f) in Figure 3.5, the cluster e has the desired object. Based on

the Figures 3.5-e the white areas indicate the pixels of the original image. Accordingly, the identifiable object is located in the e-cluster. Unlike humans, computers could not figure out the right cluster automatically. Therefore, Neural Network algorithms need to be applied to enable the computers to extract the correct regions.

3.5.1.2 Optimal number of clusters

There are many different machinery algorithms that tries to find the right number of algorithm. Although this is still an ongoing research but in this study elbow criterion is applied which fortunately could estimate most of the clusters correctly.

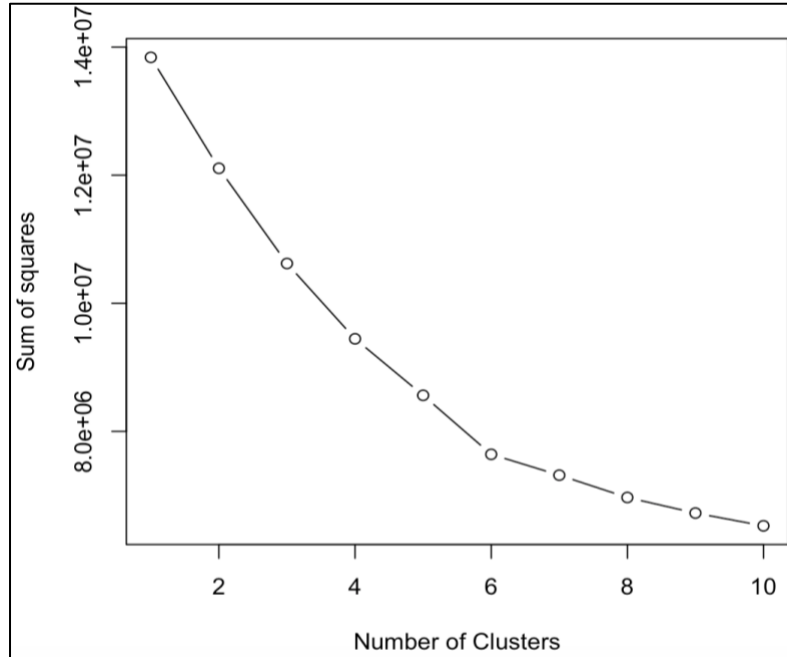
The main structure of elbow criterion is to consider a number as maximum number for clusters. Then for each of this numbers the (Standard Square Error) SSE is computed. it is obvious by considering each point as a cluster, of course the number of clusters are optimized because the SSE is zero as each point which is a centroid emerge to its own cluster. Therefore, the goal is to increase the number of clusters while the SSE decreases the elbow point is the location where k is optimized number with lower value of SSE and highest number of clusters. The slope line before elbow point is steeper than the slope line after this point.

This idea may not be useful in all cases specially in datasets which could not be clustered well.

$$SSE = \sum_i (Y - \bar{Y}) \quad (3.2)$$

The number of clusters for the following example is selected as the result of elbow criterion. The elbow point is 6.

Figure 3.6: Elbow criterion



As stated at the beginning of this chapter, each step is simulated by all the training images as inputs. Therefore, the optimized cluster number of each image is also estimated using elbow criterion.

Although elbow criterion is applied to find the optimized number of clusters, but this function may not always be able to find the correct number of clusters. Therefore, in this study Mean Shift Clustering is also applied.

3.5.2 Mean Shift Clustering

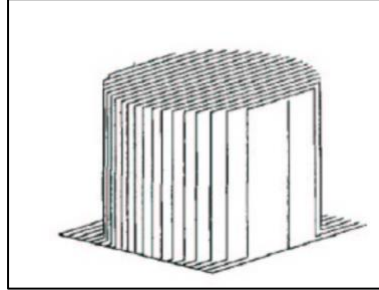
Mean Shift is a hierarchical non-parametric clustering algorithm, unlike k-means clustering method the algorithm itself figures out the number and the location of the clusters. The main idea of Mean Shift is based on the Kernel Density Estimation (KDE). KDE is benefited to find the distribution associated with the dataset by using bandwidth kernel. The kernel is actually the weighting function and various kernels result in different clusters. One of the popular kernel is Gaussian kernel.

$$K(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{\|x\|^2}{\sigma^2}} \quad (3.3)$$

Another popular kernel is Flat kernel. This the kernel used in this thesis.

$$k(x) = \begin{cases} 1 & ||x|| \\ 0 & otherwise \end{cases} \quad (3.4)$$

Figure 3.7: Flat kernel



Source: Yizong 1995. p. 792

Mean Shift considers the feature space as a sample of underlying probability density function. Mean Shift finds the densest region of the feature space. The densest region is the mode which determines the clusters. The aim is to find the densest region. Mean Shift computes the mean of the window considered for each data point and the center is shifted to the computed mean and this repeat until optimization which means the mean doesn't change and converges. and the centroid moves to the denser region in each step.

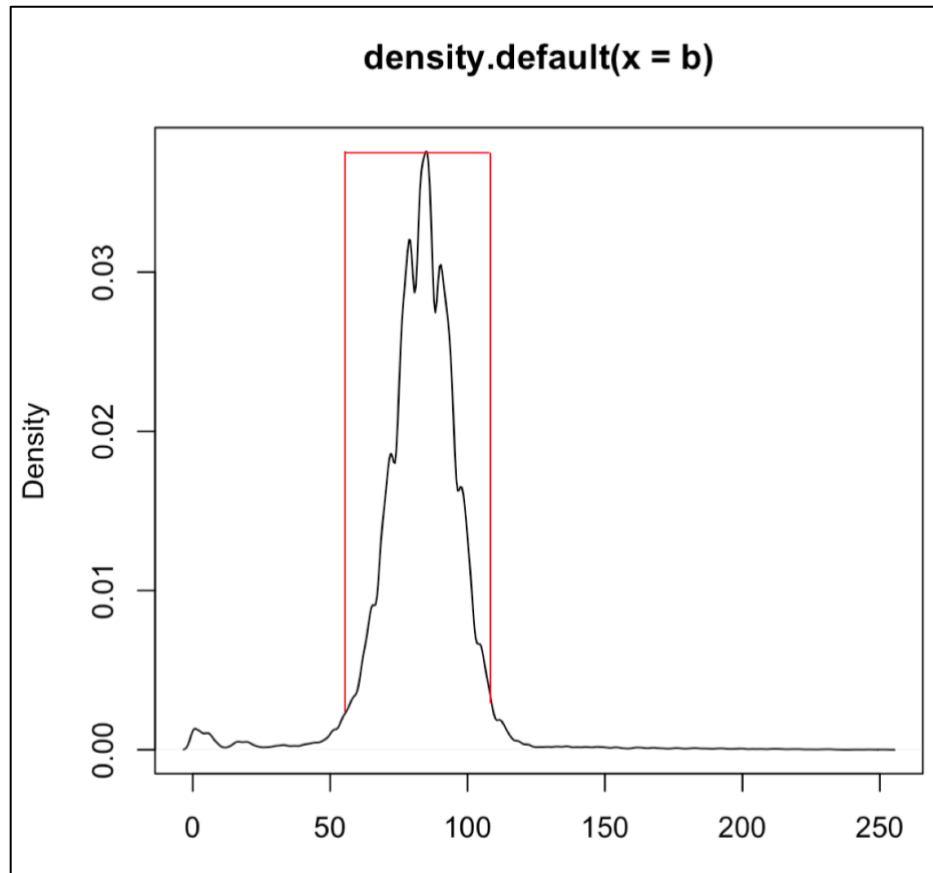
The Mean Shift function is as follows:

$$m(x) = \frac{\sum_{i=1}^n g\left(\frac{x-x_i}{h}\right)x_i}{\sum_{i=1}^n g\left(\frac{x-x_i}{h}\right)} - x \quad (3.5)$$

the $g(x)$ is the gradient of kernel ($k'(x)$) and h is the bandwidth. Kernel density function estimates the density.

After noise reduction the image should be segmented into optimized number of clusters for this purpose the image is sent to the clustering unit. As mentioned in the last part bandwidth is the initial value that must be initialized, the flat kernel is used for computing the bandwidth. Therefore, the bandwidth is estimated from the Gaussian estimation. The Figure 3.8 is the density structure of the image.

Figure 3.8: Image density structure



The quantile structure of the image is:

Table 3.1: Extracted Quantile

| 0% | 25% | 50% | 75% | 100% |
|----|-----|-----|-----|------|
| 0 | 77 | 84 | 92 | 255 |

This value for bandwidth is used for clustering. The resulting clusters are shown in Figure 3.9 and Figure 3.10. Figure 3.9 indicates the clusters and the values of the members which is between 0 and 255.

Figure 3.9: The resulting clusters by Mean Shift clustering

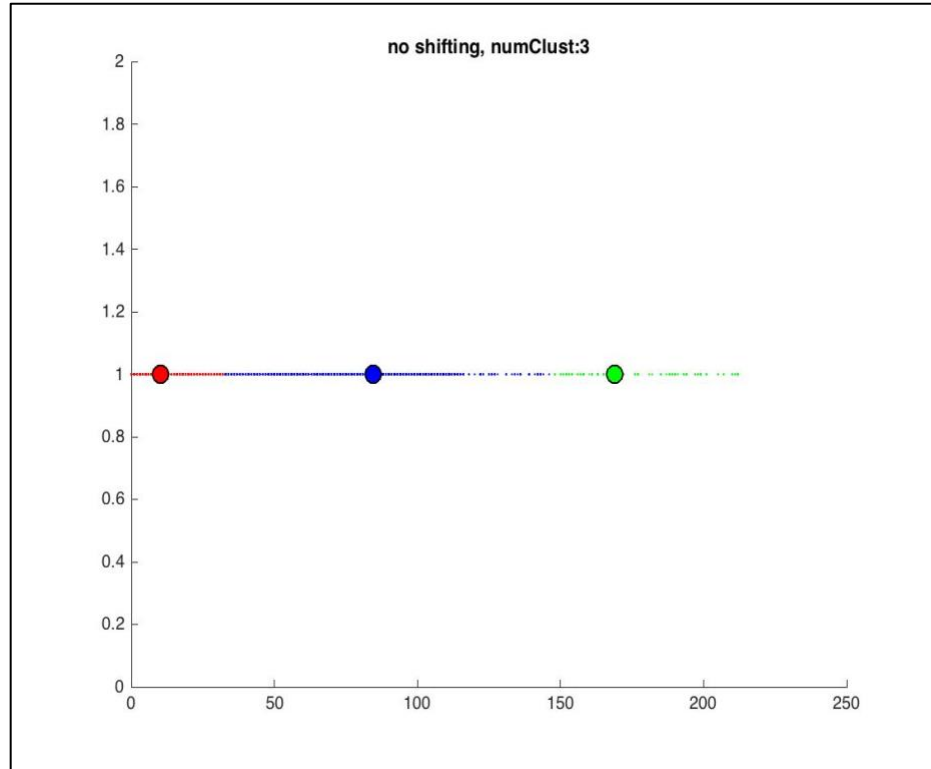
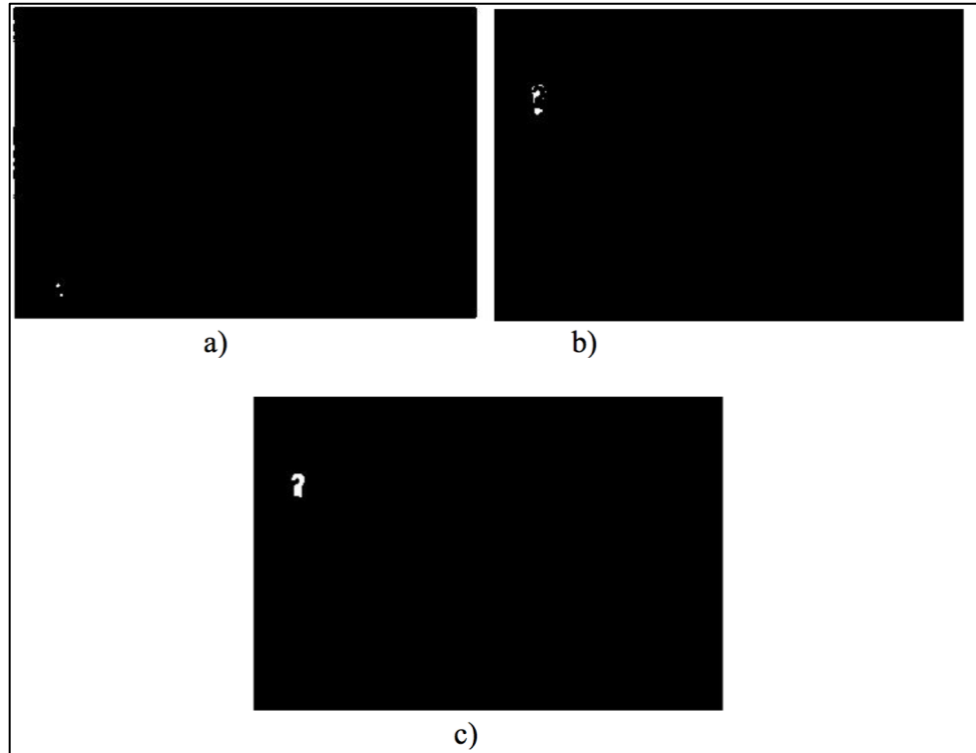


Figure 3.10.a, Figure 3.10.b and Figure 3.10.c represent the extracted clusters. Between these 3 clusters only one cluster includes the ROI (c).

Figure 3.10: Clusters extracted using Mean Shift Clustering



3.6 ROI EXTRACTION

The previous section described the image segmentation. Also, the target region is visually identified. The identifying method for the program is described in this section.

The Neural Network is used to identify the region of interest. The Neural Network is used to classify the areas and the ROI is Figure 3.10.c . For applying the neural network, features must be extracted from the regions of the image.

3.6.1 Feature Extraction from The Image Regions

In this study, simple statistical properties are used as features of regions. These features include:

- i. The minimum pixel value of each cluster
- ii. The maximum pixel value of each cluster
- iii. The average pixel value of each cluster

These properties are used according to the simplicity of acquirement and their high separating properties as they could separate and distinct the clusters easily. Obtaining these features is easy and reduces the implementation complexity of the method.

Figure 3.11: Expected object for edge detection in the IR image



K-means clustering: Considering the regions shown in Figure 3.5, Table 3.2 shows the characteristics of these areas. As stated above all the tables and images are representing the results of the figure 3.3.a. According to these defined features, each region has three characteristics.

Table 3.2: Extracted features for each region using K-means clustering

| Cluster number | Min | mean | Max |
|----------------|-----|-------|-----|
| a | 83 | 86.97 | 92 |
| b | 40 | 67.92 | 73 |
| c | 0 | 11.25 | 39 |
| d | 74 | 87.67 | 89 |
| e | 135 | 170.9 | 212 |
| f | 93 | 97.5 | 134 |

As the table indicates that the region e has the largest values.

Mean Shift clustering: Considering the region shown in Figure 3.10, Table 3.3 shows the characteristics of these areas clustered by Mean Shift clustering. According to the features defined above, each region has three characteristics. Figure 3.10.c is the desired cluster, unlike k-means this time the desired region doesn't have the highest values because means shift clustering is based on density.

Table 3.3: Extracted features for each region using Mean Shift clustering

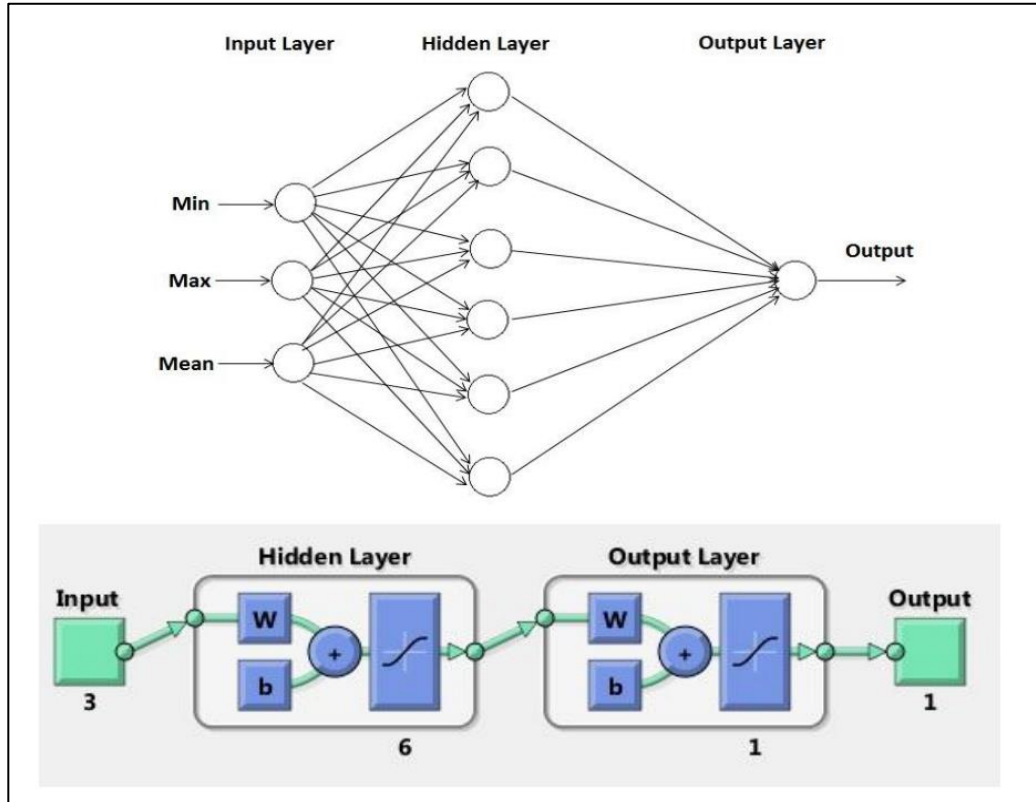
| Cluster | a | b | c |
|---------|------|-------|-------|
| min | 35 | 36 | 36 |
| max | 133 | 231 | 215 |
| mean | 84.2 | 83.99 | 84.47 |

3.6.2 Artificial Neural Network

After extracting the features, the next step is training the ANN. Unlike humans, computers could not directly find the ROI among all regions and they need to be trained for identifying the ROI. The structure of ANNs are based on simulating the human brain, thus the ANN after training are able to automatically find the ROI. The Neural Network used in this study is the multilayer perceptron and trained by Error Back Propagation method. Its structure is shown in Figure 3.12.

As shown in Figure 3.12, the Neural Network has three layers. it means that it has a hidden layer. In order to build a NN algorithm input data is needed. The inputs are the features extracted from all clusters of each image. The number of inputs of the Neural Network is three which is equal to the number of properties extracted (The minimum pixel value of each cluster, The maximum pixel value of each cluster, The average pixel value of each cluster). The number of hidden neurons is six. And the network has one output. If this output is greater than one, this means that the area under consideration is the ROI, otherwise the area is not the one with edges.

Figure 3.12: Structure of Neural Network in this study



By selecting the correct data for Neural Network training and the right training method each image will be extracted correctly. With proper training of the neural network, the ROI of each image could be extracted. For this purpose, after extracting three mentioned features for each image from all clusters among training dataset the features are gathered as inputs for the neural network and the output clarifies whether the region is the intended one or not, which is done by labelling.

K-means clustering: The result of the K-means clustering method is written as Table 3.4, where the last column represents the label of the area. The labels are used for training the neural network. With proper training the ROI of each image is extracted well. Label with value of 1 represents the ROI while labels with -1 value represents the unintended region. These labels are determined by the users so with proper training the ROI of each image is extracted well. An important point about the labels is the property values of the intended region. As the table indicates all three features of the intended region is higher comparing to the other regions. That is expected because ROI is the brightest region.

Table 3.4: Region labeling for training the NN (K-means clustering)

| Cluster Number | min | mean | Max | Label |
|----------------|-----|-------|-----|-------|
| a | 83 | 86.97 | 92 | -1 |
| b | 40 | 67.92 | 73 | -1 |
| c | 0 | 11.25 | 39 | -1 |
| d | 74 | 87.67 | 89 | -1 |
| e | 135 | 170.9 | 212 | 1 |
| f | 93 | 97.5 | 134 | -1 |

Mean Shift Clustering: In the Mean Shift algorithm, the same step is also implemented. The result is written as Table 3.5, where the last column represents the label of the area. Label with value of 1 represents the ROI while labels with -1 value represents the unintended region. These labels are determined by the users so with proper training the ROI of each image is extracted well. An important point about the labels is the property values of the intended region. In this algorithm unlike k-means the intended cluster doesn't have the highest values.

Table 3.5: Region labeling for training the NN (Mean Shift clustering)

| Cluster | a | b | c |
|---------|------|-------|-------|
| min | 35 | 36 | 36 |
| max | 133 | 231 | 215 |
| mean | 84.2 | 83.99 | 84.47 |
| Label | -1 | -1 | 1 |

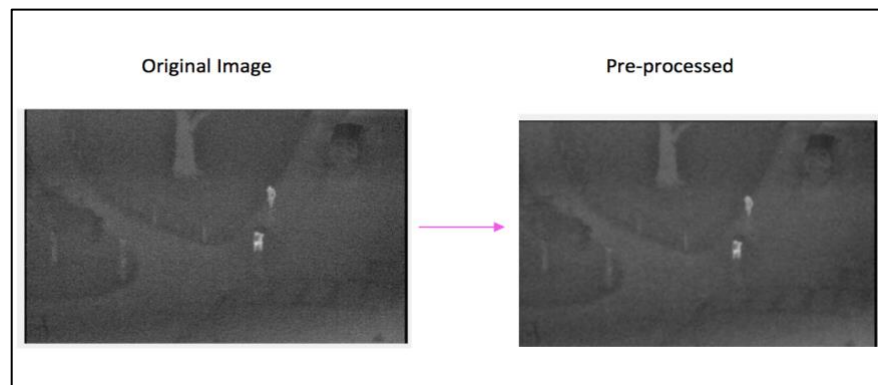
3.7 EDGE DETECTION

After training the ANN the next step is testing the algorithm and finally the edges are detected in this part. The first two steps are similar as the previous part, but instead of labelling this time the trained ANN is used for ROI extraction. The selected image first should be segmented using a proper clustering method. Then by the use of ANN the ROI

is extracted. The NN is tested by various examples from different datasets. The results are stated in the next chapter. In this part the structure of the test part is defined. The basic structure of edge detection includes these steps:

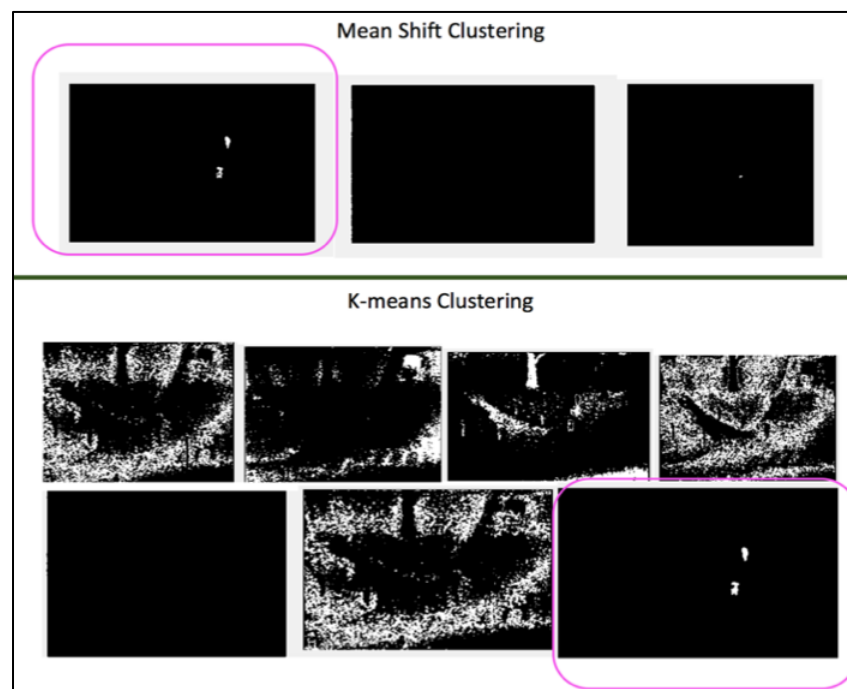
- i. First each image is preprocessed using median filter.

Figure 3.13: Pre-processing



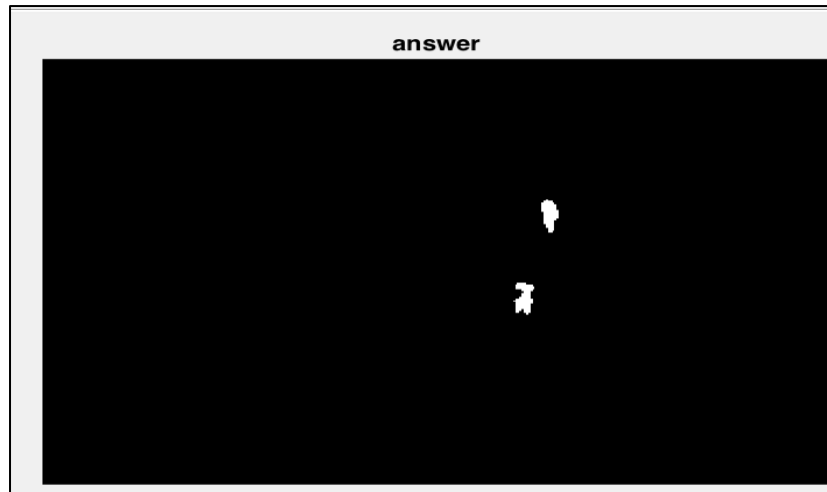
- ii. The filtered image is clustered using one of the clustering algorithms. It means that if K-means clustering is selected as the segmentation method the clusters are segmented by this method. Again, optimal number of clusters for this method is estimated by elbow criterion. For Mean Shift first, the bandwidth is calculated then the image is segmented by Mean Shift clustering method. Figure 3.14 indicates clustering results for both methods.

Figure 3.14: Clustering using both methods



- iii. Now the clusters are sent to ANN unit in which the ROI is extracted accordingly. This time by simulation using the trained ANN, the region of interest is extracted. After ROI extraction the binary image (answer) is built. Figure 3.13 shows the built image:

Figure 3.15: Binary image rebuilt from the extracted ROI



- iv. In last step the extracted ROI was sent to edge detection unit. At this stage, it is time to determine the edges of the infrared image from the ROI that was extracted previously. This section includes two main parts:

- a. Post-processing step
- b. Edge detection using morphological features

3.7.1 Post-Processing

Once the Region of Interest has been selected among all regions, it should be sent to the edge detection unit. In this part the image is post processed to improve the images' quality for better visualization and better edge extraction. In this step, the aim is to delete unwanted points and white noise that is missed in the pre-processing step. Post-processing unit includes two parts:

- i. **Erosion:** After ROI extraction of all images the ROI must be sent to the edge detection unit. In ROI presence of tiny unwanted points is probable so post-processing operation is required. This step is done using morphological erosion operator. It is used

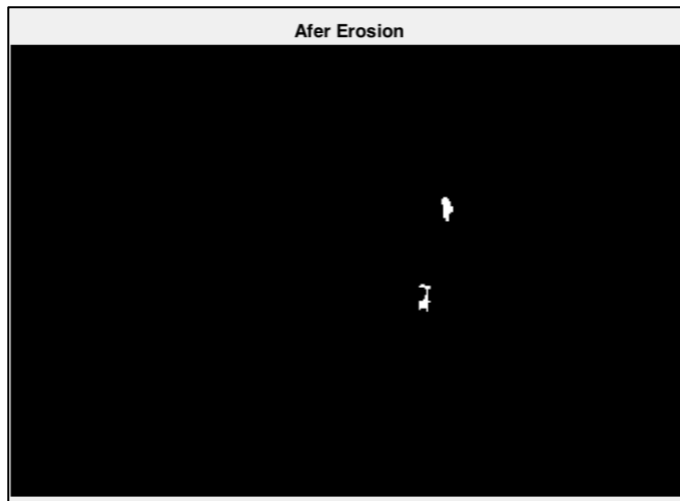
to eliminate points with the value less than 3x3. The SEE matrix is used as a structural element for this operation. This equation is used to eliminate the unwanted tiny points:

$$\text{Image, SEE: ErodeImage} = \text{Image} \ominus \text{SEE} = \{z | (\text{SE})_z \cap (\text{Image}^c) \neq \emptyset\} \quad (3.6a)$$

$$\text{SE} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.6b)$$

The resulting image is eroded. With this step the white noise is eliminated.

Figure 3.16: Binary image after Erosion



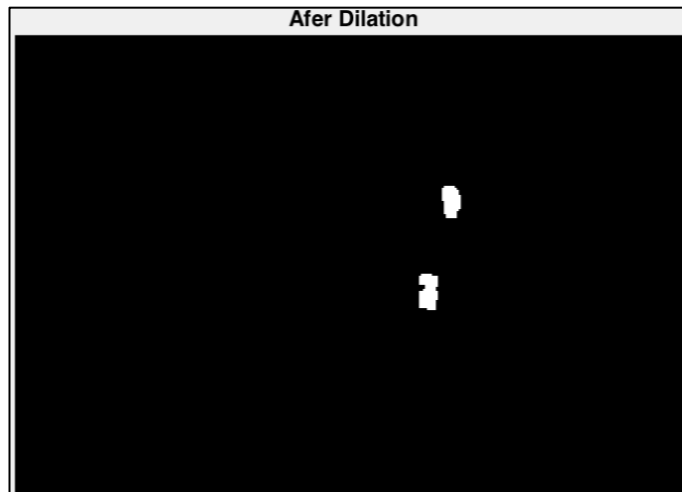
- ii. Dilation:** In this part, the resulting image should be dilated for better visualization. The remaining image is returned to its original state. The dilation equation:

$$\text{Image} = \text{ErodeImage} \oplus \text{SED} = \{z | (\widehat{\text{SE}})_z \cap \text{Image} \neq \emptyset\} \quad (3.7a)$$

$$\text{SED} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (3.7b)$$

In the last step, the image is narrower so in this step the inverse operation is done in order to dilate the image. This image is ready for edge detection.

Figure 3.17: Binary image after Dilation

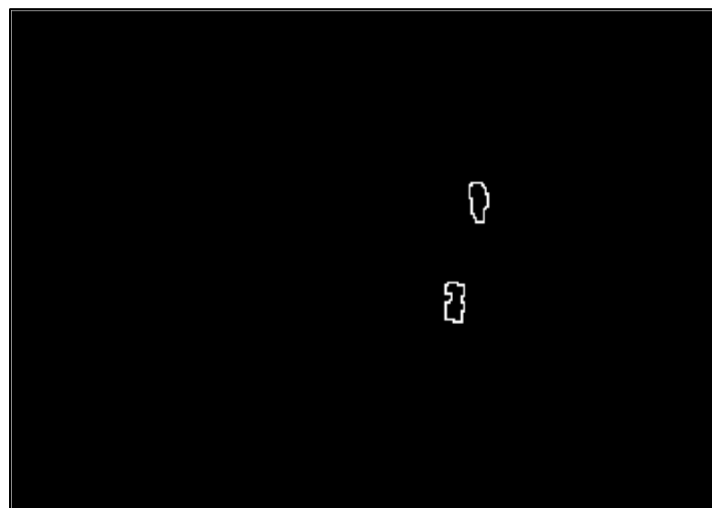


3.7.2 Edge Detection Using Morphological Operators

After receiving the image, the edges are detected. The obtained image is binary. In this step, instead of using common edge operators edge detection is performed by morphological operators. The morphological equation is written in the Equation 3.7:

$$ImageEdge = Image - (Image \ominus SEE) \quad (3.8)$$

Figure 3.18: Detected edges



3.8 CHAPTER SUMMARY

The edge detection of infrared images has a great importance. In this chapter, based on the research that was done from previous works first we presented a new way for edge detection of the infrared images. So, at the beginning of this chapter, the overall research process has been explained. Next, the pre-processing method of infrared images was expressed. Then, the image segmentation and extraction of the region of interest were described. Last, the infrared image edge detection based on morphological operators were defined.

In the fourth chapter, the proposed method will be evaluated.

4. FINDINGS

4.1 INTRODUCTION

In the previous chapters the proposed method of research was clearly described. In this chapter, the methodology for evaluation and the dataset selected in the research will be described first. Then, the results of the proposed method are presented.

4.2 EVALUATION

To validate and evaluate the proposed method, a proper IR image dataset is selected according to the proposed research method. Then, based on the available information of the dataset, the efficiency of the proposed method is evaluated.

4.2.1 Dataset

The dataset used in this study is the OSU Thermal Pedestrian Database. This dataset has 10 different files including images. The general info about the dataset are provided by J Davis and Keck (2005)¹:

Data Details:

Pedestrian intersection on the Ohio State University campus

Number of sequences = 10

Total number of images = 284

Format of images = 8-bit grayscale bitmap

Image size = 360 x 240 pixels

Sampling rate = non-uniform, less than 30Hz

Environmental information for each sequence provided in subdirectories

Ground truth provided in subdirectories as list of bounding boxes (with approximately same aspect ratio) around people.

For the ground truth data, we selected only those people that were at least 50% visible in the image (i.e., highly occluded people were not selected).

4.3 EVALUATION RESULT

In this study, a new method was proposed for edge detection of infrared images. Next in this chapter, the evaluation and results are described. Due to the explanations given in the dataset section and according to the use of dataset and expectations of edge detection methods it is expected that the proposed method is capable of identifying the edges properly.

¹ J.-Davis and M.-Keck, *A two stage approach to person detection in thermal imagery*

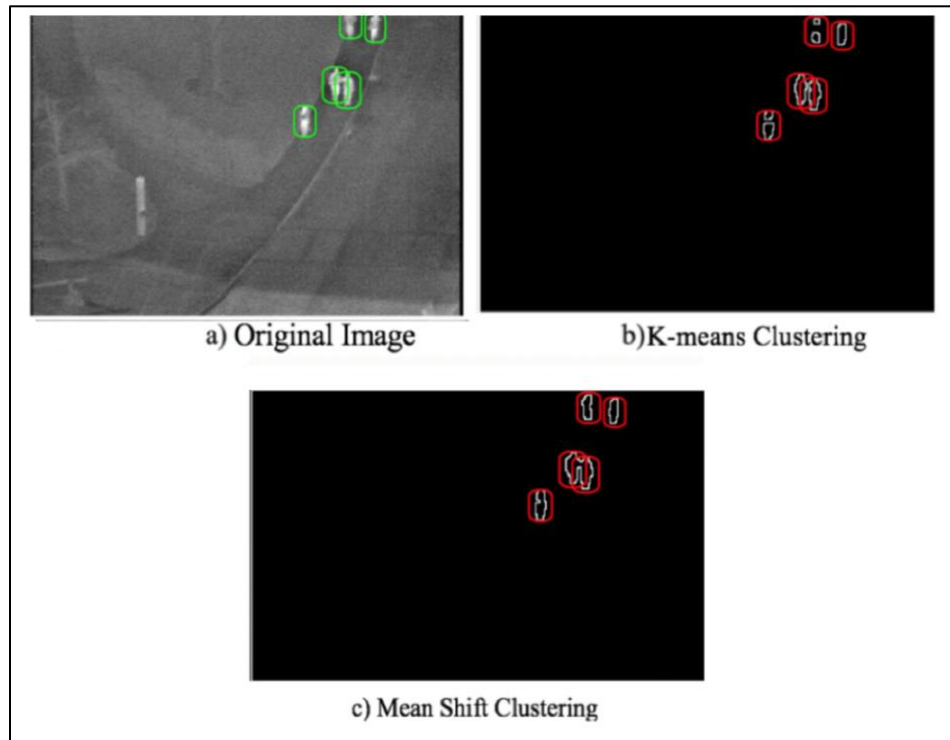
4.3.1 Evaluation Method

In this part clustering and ANN are also used. K-means clustering depends on cluster number determination and therefore, elbow criterion is benefited for this purpose. For Mean Shift clustering proper bandwidth should be defined. The results are expressed for both clustering methods bellow. In this part, 16 different images are selected randomly from the different files of the dataset and the results of the proposed method are given. Each image is evaluated by both k-means and Mean Shift clustering methods and the result is expressed one by one. The green and red ellipses are drawn according to the Ground Truth text file.

Images are selected randomly from all folders and the edges are extracted by two different clustering methods. Among all figures, a is the original image b is the edges extracted by K-means clustering while c is the edges extracted by Mean Shift clustering method.

Image 1:

Figure 4.1: Detected edges (b,c) of image img_00003 (a) in folder 0002



As the results indicate, both methods detected the edges perfectly. But in the Figure 4.1.c the edges of the person at the top right-corner is detected with better precision . The reason is due to the cluster numbers that are determined by the program itself and it is the advantage of this program.

Image 2:

Figure 4.2: Detected edges (b,c) of image img_00014 (a) in the folder 0002

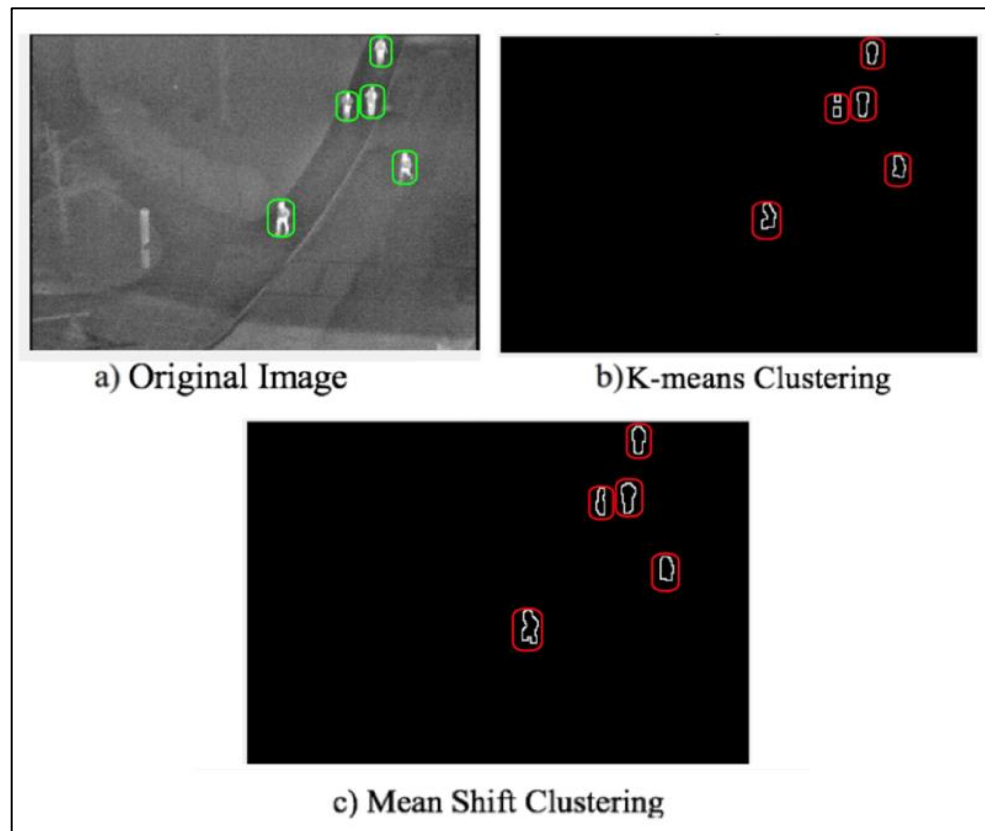
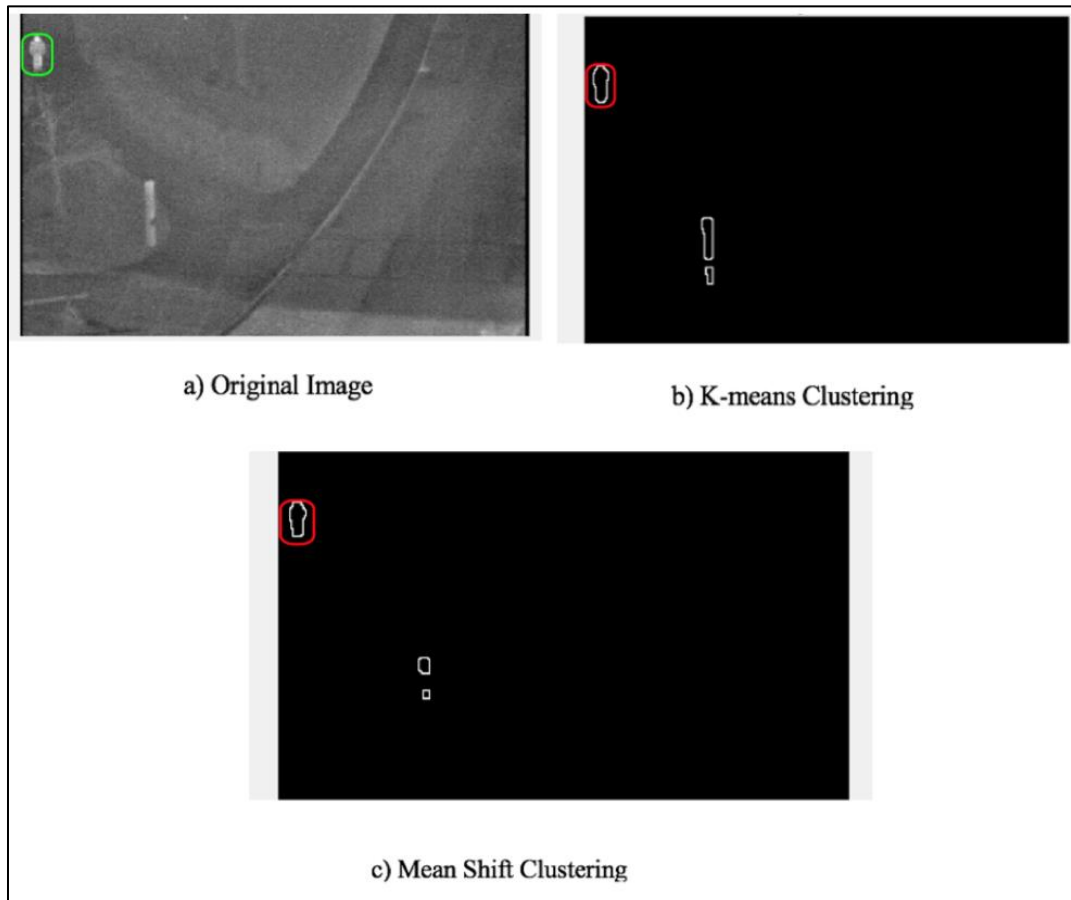


Figure 4.2.a is also detected correctly by both methods. But again, like the previous example the edges of two objects in the Figure 4.2.c are detected better by Mean Shift method.

Image 3:

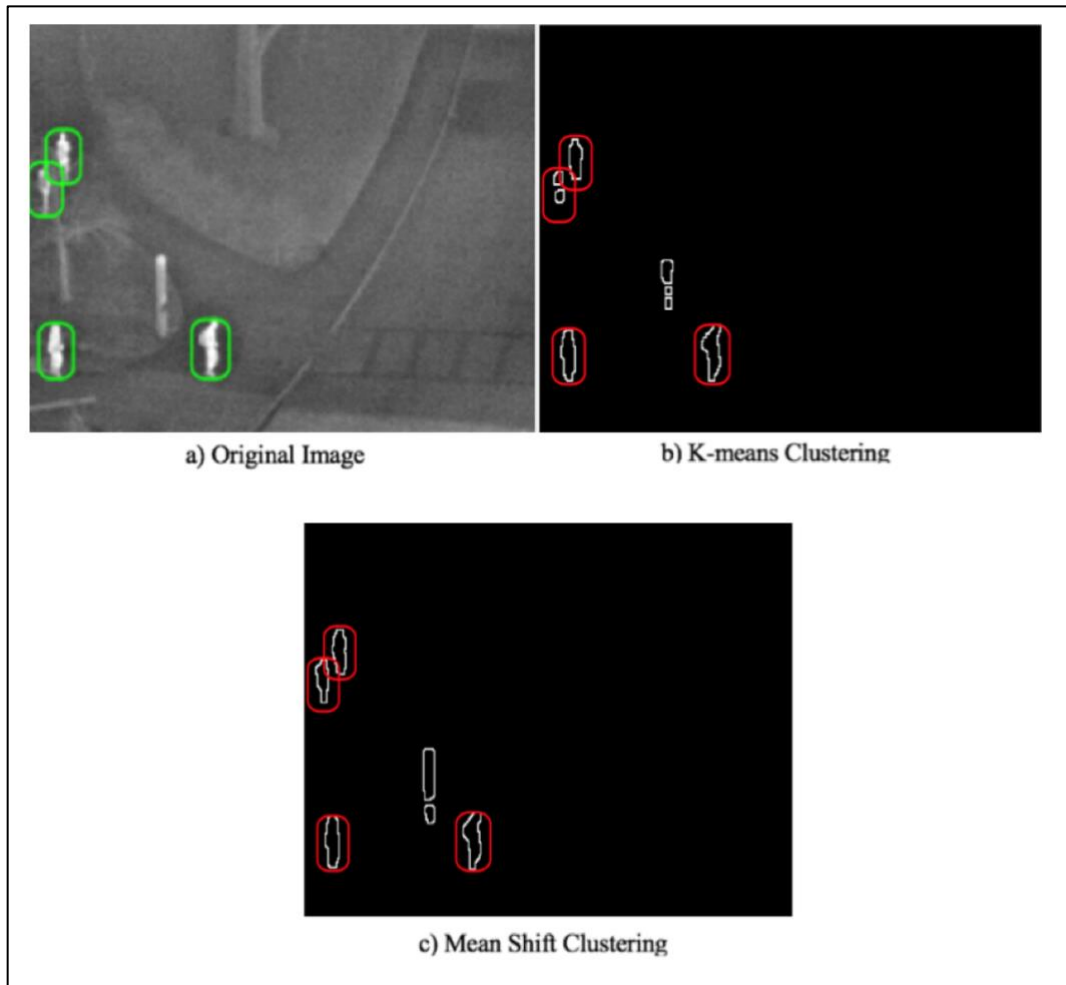
Figure 4.3: Detected edges (b,c) of image img_00028 (a) in the folder 0002



As the results of Figure 4.3 indicate both methods have detected the edges correctly. Both methods have detected an object (lamp post light) which wasn't defined as an identifiable object in the ground truth file. The reason for this matter is according to the structure of the program. Because the program is implemented in a way that it could recognize the objects based on their infrared radiation. So, it's not the programs' error but as the Ground Truth file clarified that the recognizable objects are humans. This matter is recognizable in next examples too.

Image 4:

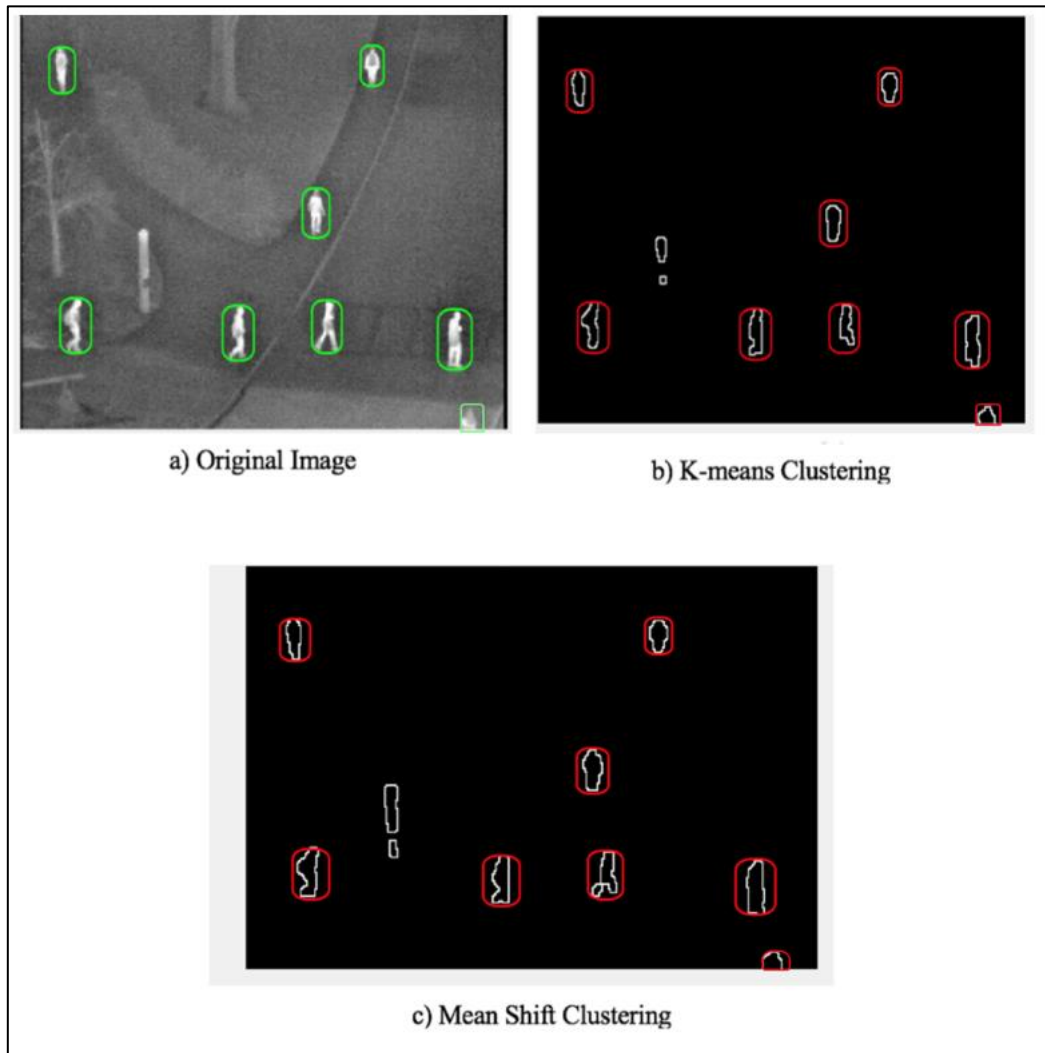
Figure 4.4: Detected edges (b,c) of image img_00001 (a) in the folder 0004



The edges of Figure 4.4.a are detected better by Mean Shift clustering and the lamp post light is detected by both methods.

Image 5:

Figure 4.5: Detected edges (b,c) of image img_00009 (a) in the folder 0004



The edges of Figure 4.5.a are detected correctly by both methods. The lamp post light is detected by both methods.

Image 6:

Figure 4.6: Detected edges (b,c) of image img_00018 (a) in the folder 0004

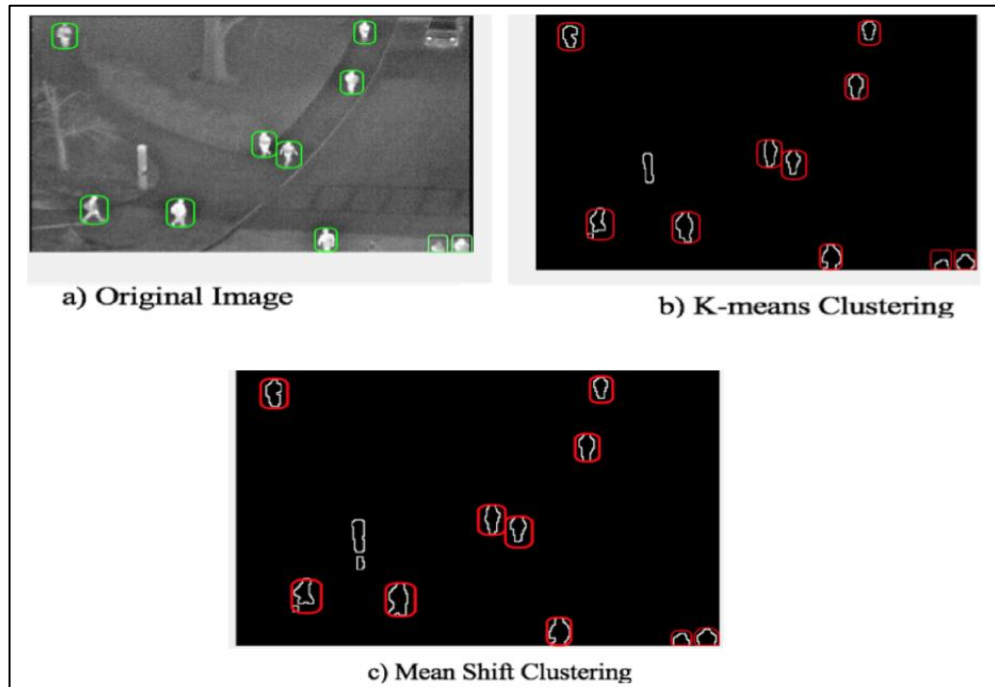
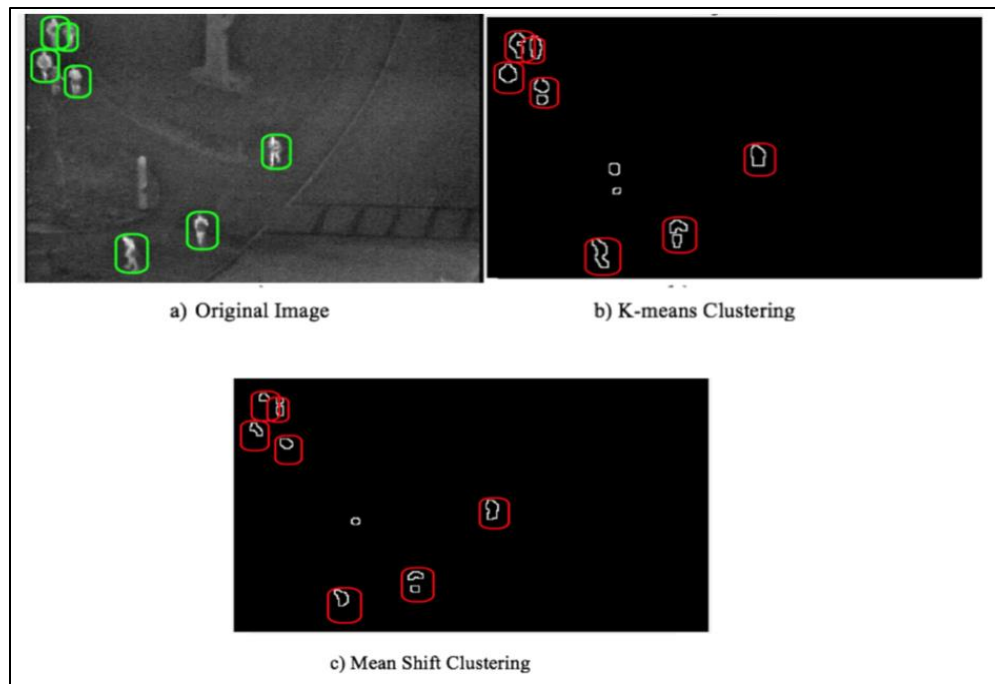


Image 7:

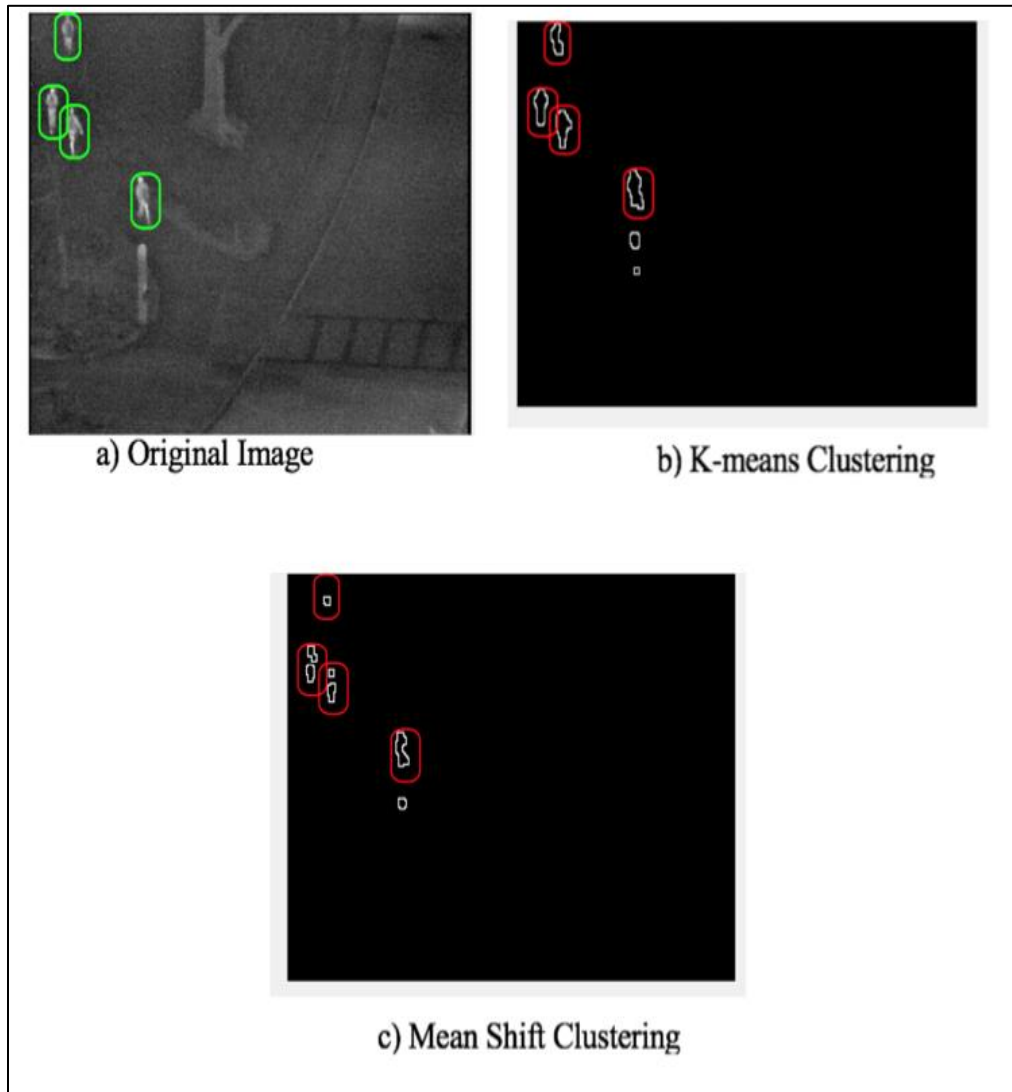
Figure 4.7: Detected edges (b,c) of image img_00001 (a) in the folder 0006



For Figure 4.6 edges are detected correctly by both methods. Figure 4.7.b has better results comparing with Figure 4.7.c. Figure 4.7.b and 4.7.c include the lamp post light.

Image 8:

Figure 4.8: Detected edges (b,c) of image img_00009 (a) in the folder 0006



For this example, k-means detected the edges better.

Image 9:

Figure 4.9: Detected edges (b,c) of image img_00018 (a) in the folder 0006

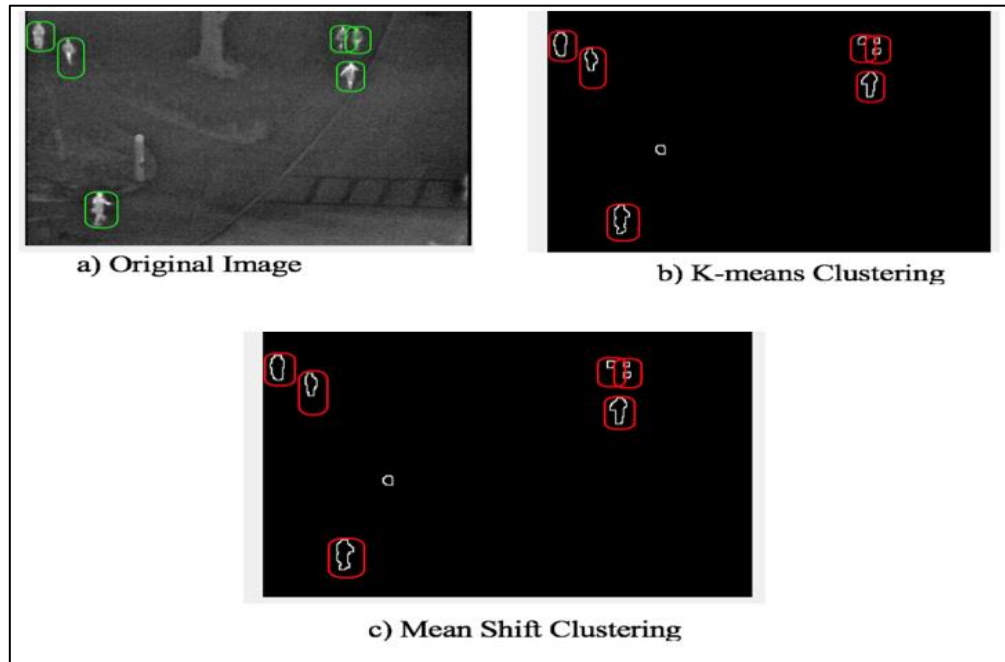
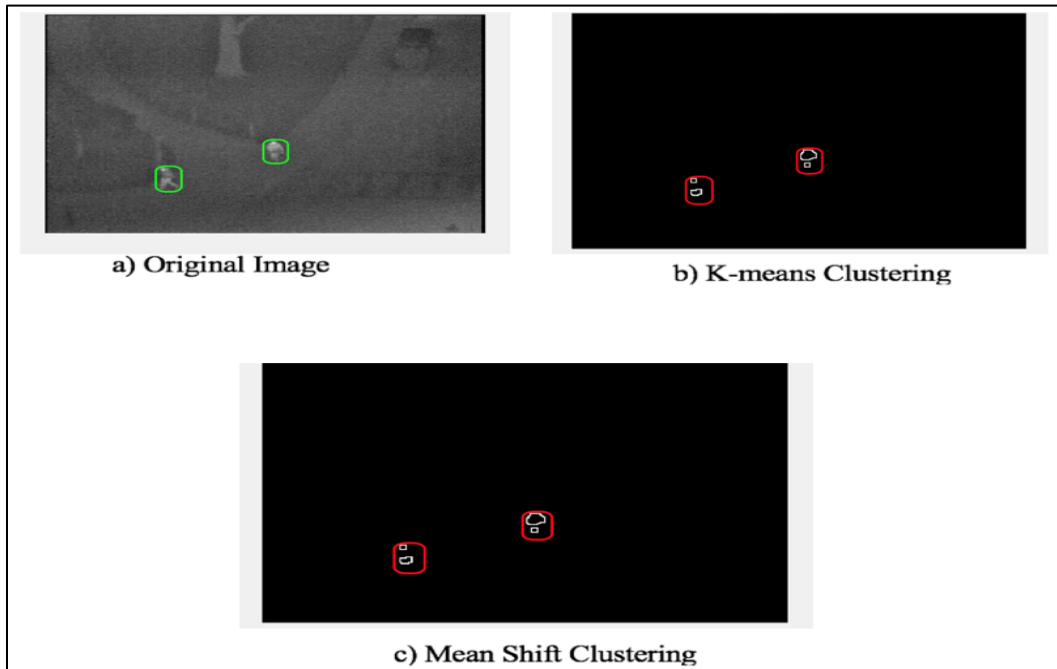


Image 10:

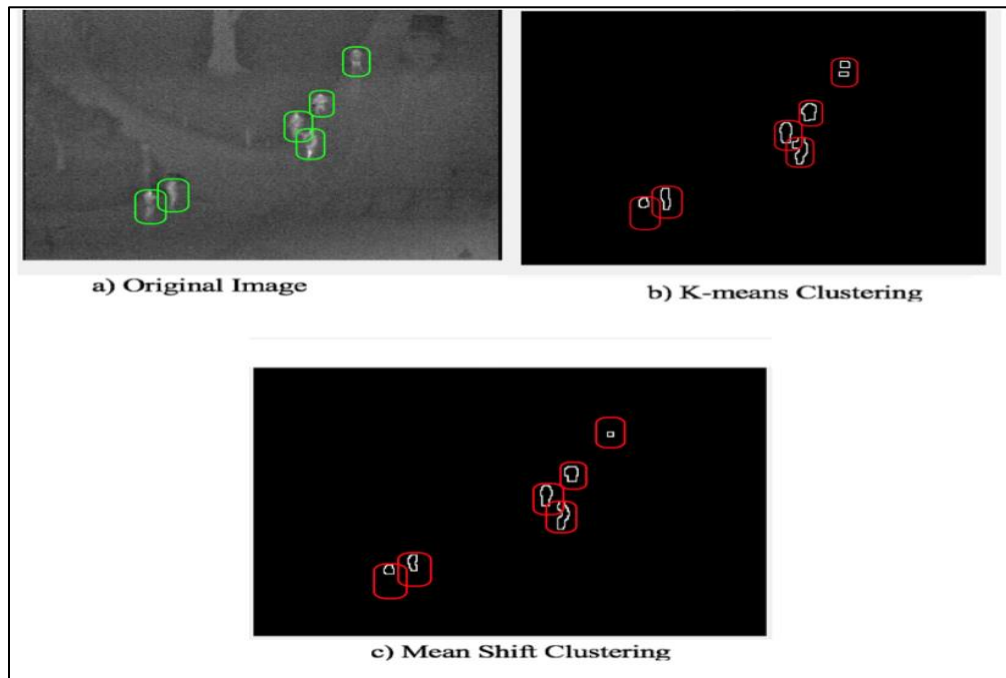
Figure 4.10: Detected edges (b,c) of image img_00001 (a) in the folder 0007



both examples (Figure 4.9, Figure 4.10) have the same results for both clustering methods.

Image 11:

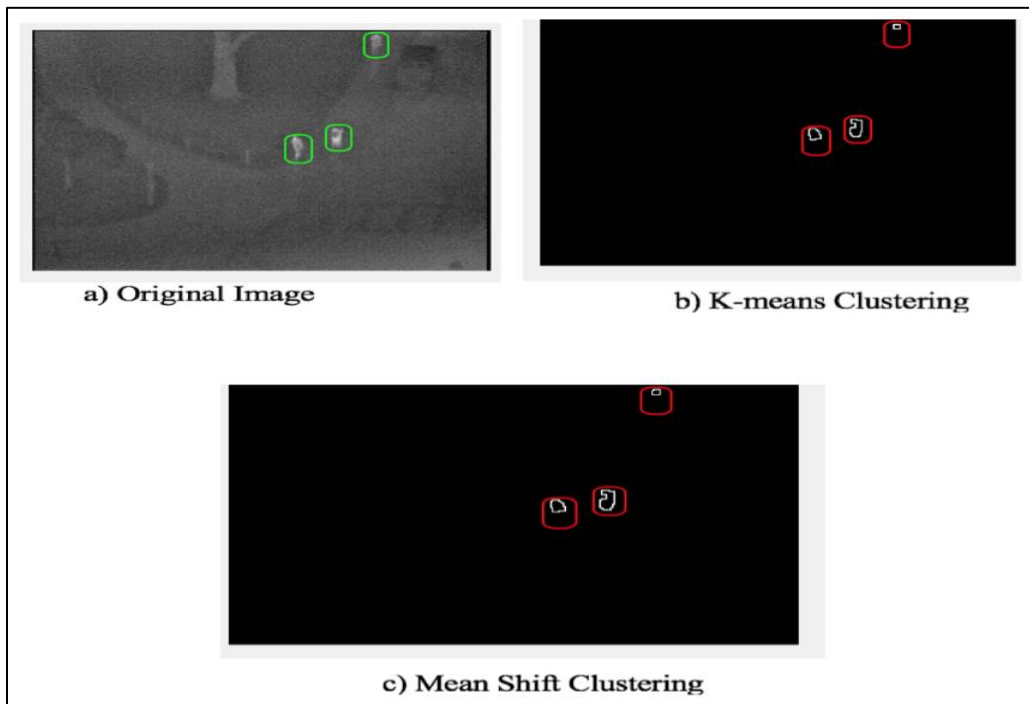
Figure 4.11: Detected edges (b,c) of image img_00011 (a) in the folder 0007



Results of k-means method especially for the top right-corner pedestrian is better.

Image 12:

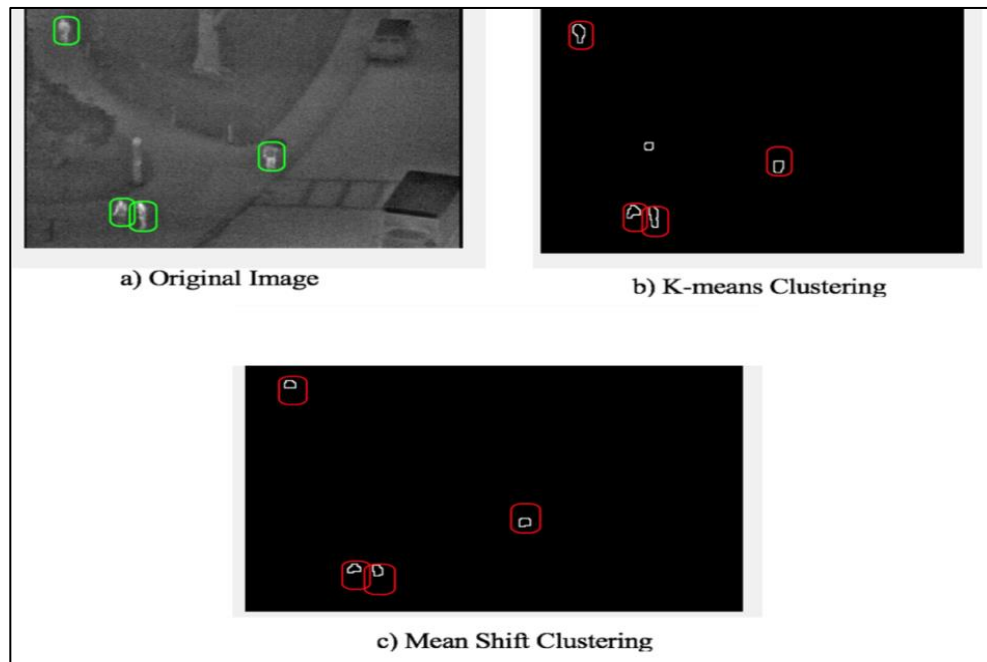
Figure 4.12: Detected edges (b,c) of image img_00022 (a) in the folder 0007



The results of both methods are close.

Image 13:

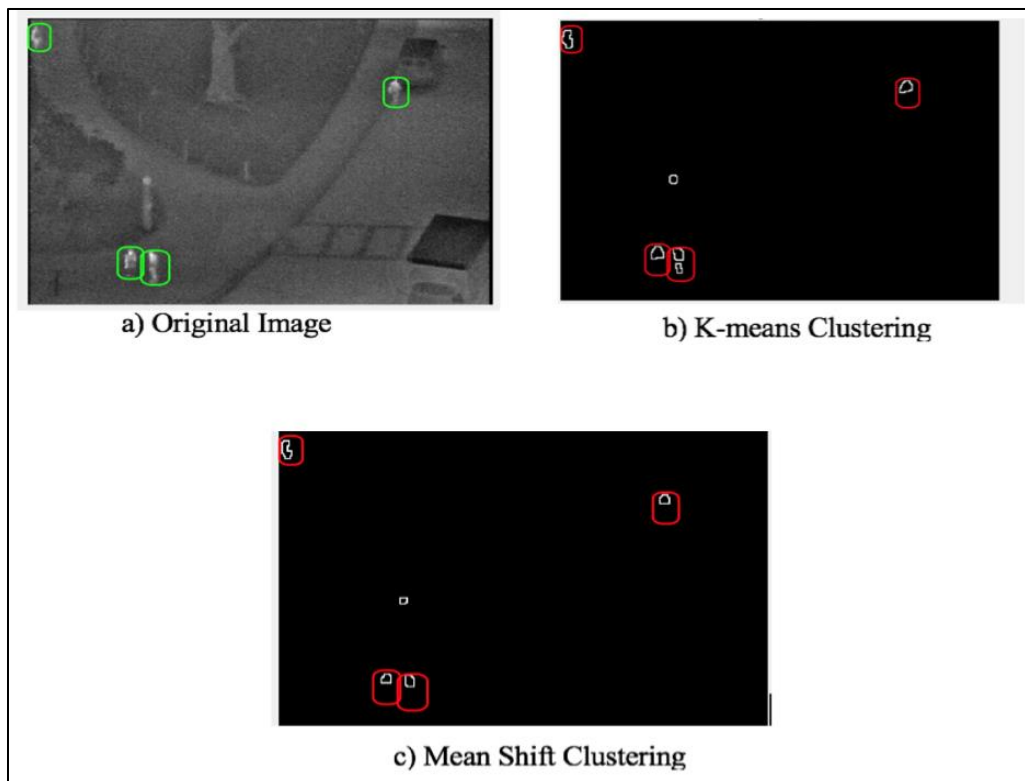
Figure 4.13: Detected edges (b,c) of image img_00001 (a) in the folder 08



The results of k-means method are better, but the lamp post light is not detected by Mean Shift.

Image 14:

Figure 4.14: Detected edges (b,c) of image img_00012 (a) in the folder 0008



The results of both methods of Figure 4.14 are close.

Image 15:

Figure 4.15: Detected edges (b,c) of image img_00024 (a) in the folder 0008

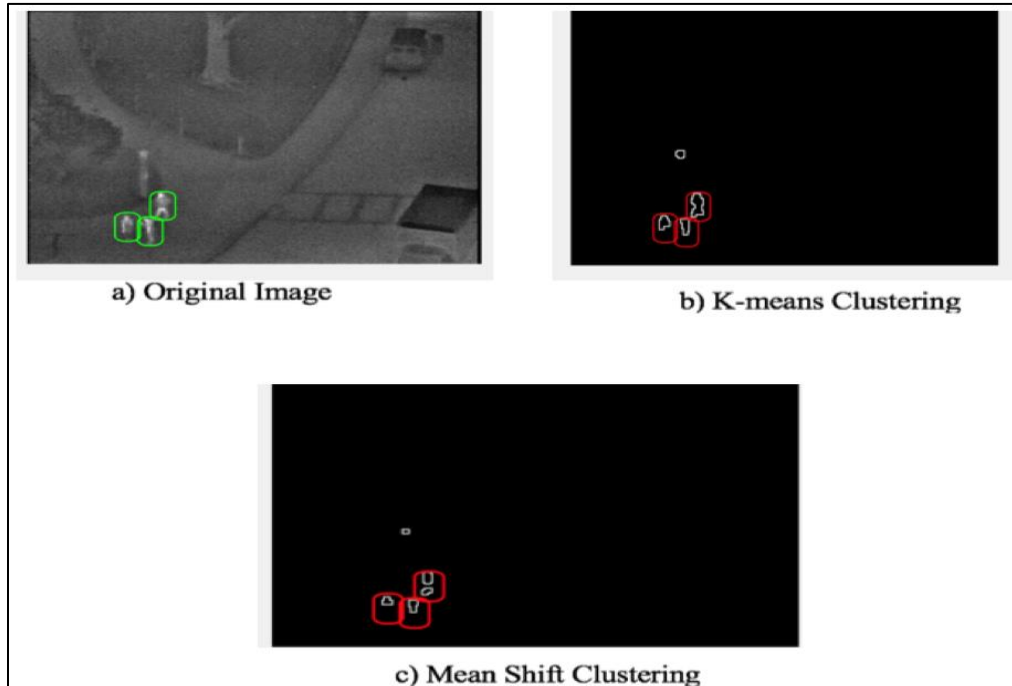
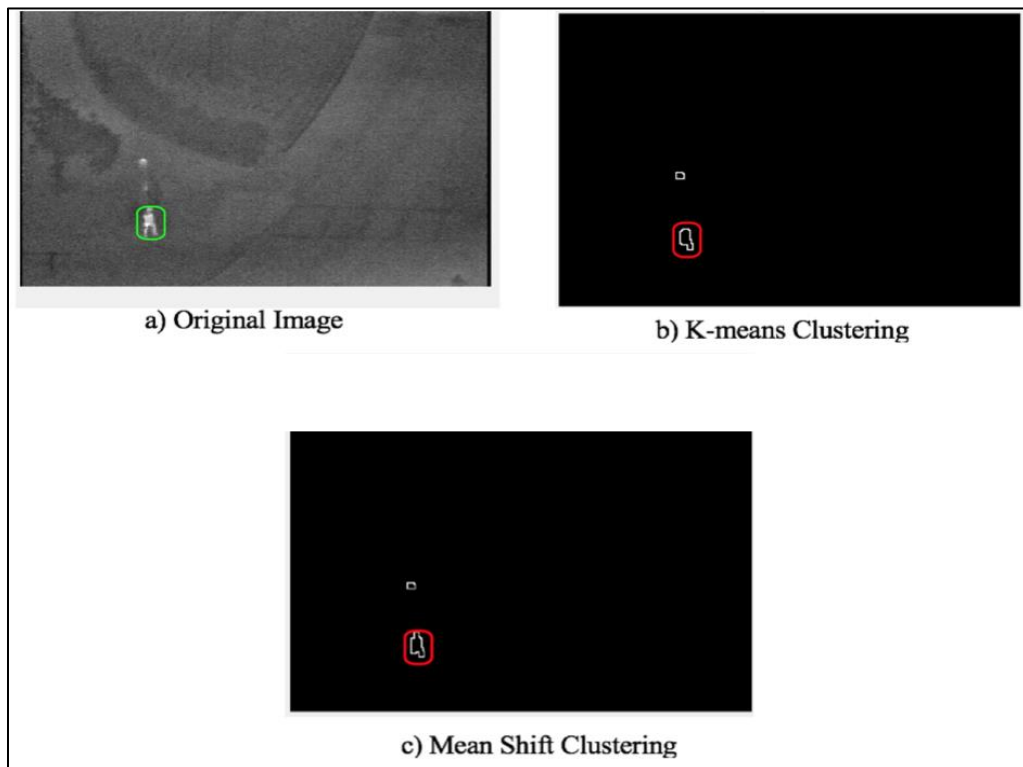


Image 16:

Figure 4.16: Detected edges (b,c) of image img_000032 (a) in the folder 0009



The results of both methods for Figure 4.15 and Figure 4.16 are close.

The above images are initialized using these values, the values are obtained by elbow criterion for K-means, and using quantile of the density for Mean Shift clustering:

i. K-means Clustering Initialization:

The number of clusters for the first folder is eight. The number of clusters for the second folder is six. The number of clusters for the fourth folder is five. The number of clusters for the sixth folder is six. The number of clusters of the seventh folder is seven. The number of clusters of the eighth folder is seven. The number of clusters for the K-means algorithm in ninth folder is ten.

ii. Mean Shift Clustering Initialization:

The selected bandwidth for the images of the second folder is nineteen. The bandwidth for the images of the fourth folder is twenty. The selected bandwidth for the images of the sixth folder is sixteen and for seventh folder is fifteen. The selected bandwidth for the images of the eighth and ninth folder is seventeen and sixteen.

4.3.1.2 Confusion matrix

In this section the performances of both clustering methods are computed using confusion matrix. For this purpose, the ground truth of the dataset is benefited to calculate this assessment. In the confusion matrix, pixels represent total number of pixels and TP is true positive values which are the total number of edge pixels that are correctly detected. TN is true negative values which are the total number of non-edge pixels that are correctly detected. FP is false positive values which are the total number of non-edge pixels that are incorrectly detected as edges. FN is false negative values which are the total number of edge pixels that are incorrectly detected as non-edge values.

In this part the resulting confusion matrices of all 16 images (Figure 4.1 to 4.16) are calculated: Table 4.1 and Table 4.2 illustrate the confusion matrices of two images (Figure 4.1 and Figure 4.3). The results of all the other confusion matrices (Table-2 to Table-15) are available in Appendix-2 section.

Table 4.1: Confusion matrix of Figure 4.1

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=3080 | FN=0 | K-means Clustering |
| Actual=No | FP=0 | TN=83320 | |
| Actual=yes | TP=3014 | FN=66 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=83320 | |

Table 4.2: Confusion matrix of Figure 4.3

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=638 | FN=0 | K-means Clustering |
| Actual=No | FP=485 | TN=85277 | |
| Actual=yes | TP=638 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=312 | TN= 85450 | |

Table 4.3 is the total results that are calculated for all 16 images (Figure 4.1-4.16). For this purpose, all of the corresponding tables are acquired. Tables of other fourteen images are presented in Apendix-2 (Table-2 to Table-15).

Table 4.3: Total results of all 16 images

| Pixels= 1382400 | Detected =Yes | Detected =No | |
|------------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 41915 | FN=138 | K-means Clustering |
| Actual=No | FP=2962 | TN=1337385 | |
| Actual=yes | TP= 40623 | FN=414 | Mean Shift Clustering |
| Actual=No | FP=2883 | TN=1338480 | |

Following rates are calculated for both clustering methods. The inputs are obtained according to Table 4.3:

k-means Clustering

$$\text{Accuracy: } \frac{(TP+TN)}{Total} = \frac{(41915+1337385)}{1382400} = 0.997757523 \quad (4.1)$$

$$\text{Misclassification Rate: } \frac{(FP+FN)}{Total} = \frac{(2962+138)}{1382400} = 0.002242477 \quad (4.2)$$

$$\text{Sensitivity/Recall: } \frac{(TP)}{Actual\ yes} = \frac{(41915)}{(41915+138)} = 0.99671843 \quad (4.3)$$

$$\text{False Positive Rate: } \frac{(FP)}{Actual\ yes} = \frac{(2962)}{(2962+1337385)} = 0.00220988 \quad (4.4)$$

$$\text{Specificity: } \frac{(TN)}{Actual\ yes} = \frac{(1337385)}{(2962+1337385)} = 0.99779012 \quad (4.5)$$

$$\text{Precision: } \frac{(TP)}{Detected\ yes} = \frac{(1337385)}{(2962+1337385)} = 0.93399737 \quad (4.6)$$

Mean Shift Clustering:

$$\text{Accuracy: } \frac{(TP+TN)}{Total} = \frac{(40623+1338480)}{1382400} = 0.99761502 \quad (4.7)$$

$$\text{Misclassification Rate: } \frac{(FP+FN)}{Total} = \frac{(2883+414)}{1382400} = 0.00238498 \quad (4.8)$$

$$\text{Sensitivity/Recall: } \frac{(TP)}{Actual\ yes} = \frac{40623}{(40623+414)} = 0.98991154 \quad (4.9)$$

$$\text{False Positive Rate: } \frac{(FP)}{Actual\ yes} = \frac{(2883)}{(2883+1338480)} = 0.002149306 \quad (4.10)$$

$$\text{Specificity: } \frac{(TN)}{Actual\ yes} = \frac{(1338480)}{(2883+1338480)} = 0.997850694 \quad (4.11)$$

$$\text{Precision: } \frac{(TP)}{Detected\ yes} = \frac{(41915)}{(41915+2962)} = 0.933733278 \quad (4.12)$$

Table 4.4 shows the result of comparison between both methods:

Table 4.4: Performance comparison of both methods

| | K-means clustering | Mean Shift clustering | Comparison |
|------------------------|--------------------|-----------------------|--|
| Accuracy | 0.997757523 | 0.99761502 | K-means has higher accuracy rate |
| Misclassification Rate | 0.002242477 | 0.00238498 | K-means has lower misclassification rate |
| Sensitivity | 0.99671843 | 0.98991154 | K-means has higher sensitivity |
| False Positive Rate | 0.00220988 | 0.002149306 | Mean shift has lower False Positive Rate |
| Specificity | 0.99779012 | 0.997850694 | K-means has higher specificity |
| Precision | 0.93399737 | 0.933733278 | K-means has higher precision |

As Table 4.4 indicates rates are very good for both methods. Comparison of both methods illustrates that k-means clustering method has better accuracy, sensitivity, misclassification rate and precision, but the results of specificity rate is better for Mean Shift clustering method.

This section is intended to assess the effectiveness of the proposed method itself and the method has not been compared with other methods. The simulation results indicate methods' good performance and it has been able to recognize the edges in most images. The important point about k-means clustering method is that the method depends on the number of clusters. By selecting the correct number of clusters in the k-means algorithm, the proposed method will perform well.

In the next chapter, the results of the proposed method are compared with other edge detection methods.

4.3.1.3 Running time

In this part the running part of both methods for some random images from different folders are calculated:

The running time of the final program for the image `img_00002.bmp` from the second folder using the trained Neural Network algorithm and k-means clustering is 3.600170

seconds. The program's running time using the trained Neural Network algorithm and mean shift clustering method is 2.397038 seconds.

The running time of the final program for the image `img_00014.bmp` from the sixth folder using the trained Neural Network algorithm and k-means clustering is 3.200960 seconds. The program's running time using the trained Neural Network algorithm and mean shift clustering method is 4.003568 seconds.

The running time of the final program for the image `img_00001.bmp` from the fourth folder using the trained Neural Network algorithm and k-means clustering is 3.545127 seconds. The program's running time using the trained Neural Network algorithm and mean shift clustering method is 3.186451 seconds.

The running time of the final program for the image `img_00011.bmp` from the seventh folder using the trained Neural Network algorithm and k-means clustering is 3.545127 seconds. The program's running time using the trained Neural Network algorithm and mean shift clustering method is 4.328679 seconds.

The running time of the final program for the image `img_00022.bmp` from the ninth folder using the trained Neural Network algorithm and k-means clustering is 3.919369 seconds. The program's running time using the trained Neural Network algorithm and mean shift clustering method is 2.476575 seconds.

As results indicate the Mean shift clustering program's running time for all sample images except the second one is less than k-means clustering method. It's due to the clustering part. Because the number of clusters in k-means method in most examples is more than mean shift cluster numbers, therefore it takes more time for the program to extract the edges. The running time of all images of chapter four and five are calculated and written in Appendix-1.

4.4 CONCLUSION

The IR image edge detection is highly practical in image processing. The IR images have many uses in medical and military fields. According to their special structure, traditional methods of processing are not useful. Therefore, there are other ways to process these images. In this study, a new method for processing these images was presented for processing. In the previous chapters, the proposed method for implementation is discussed. In this chapter results of two clustering methods are evaluated.

5. DISCUSSION

As mentioned in last chapters IR images are significantly valuable for various applications and because of the obtained information from edge detection as an image processing knowledge. In this study a new method is proposed to detect and extract the edges. In the previous sections the findings and outputs of this method were discussed. In order to evaluate the application of this study the efficiency of the proposed method is compared with commonly used edge detection operators and it is also compared with other works done so far which are briefly introduced in the Literature Review chapter. The results indicated a good performance of the proposed method.

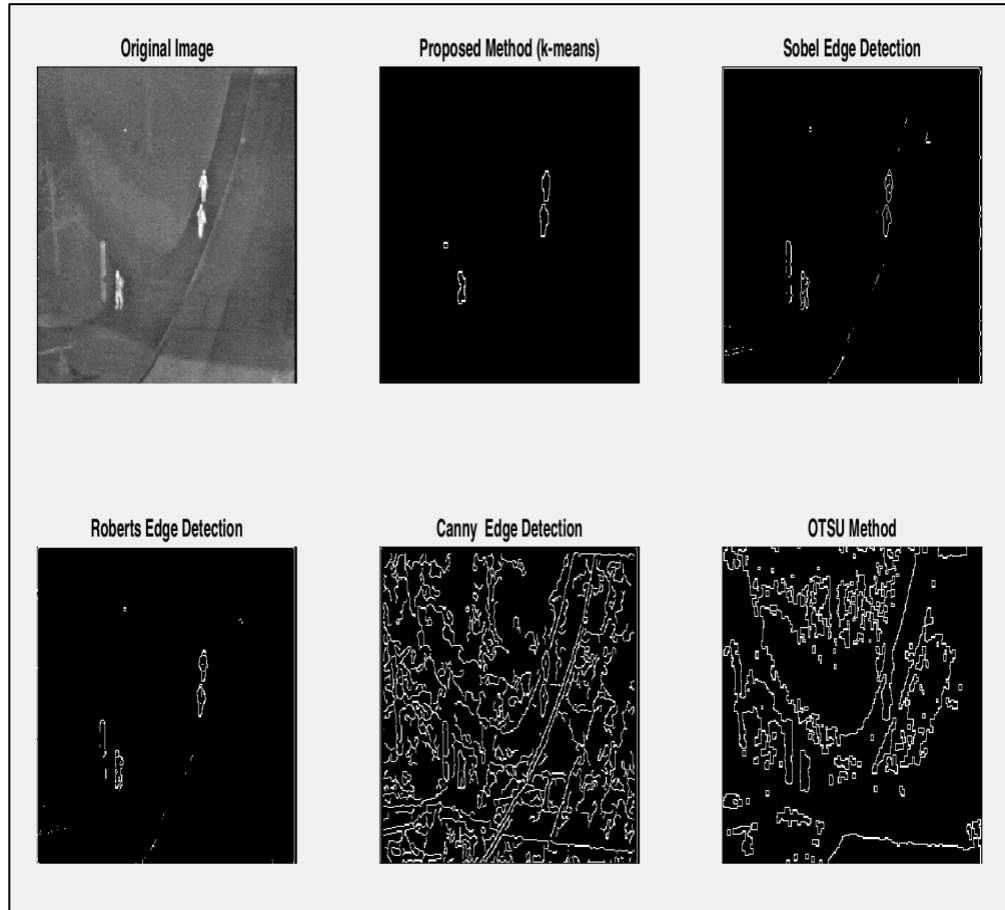
In this section the performances of both clustering algorithms are also evaluated and compared with each other by some random IR images. The images are selected from different IR datasets.

5.1 COMPARING THE RESULTS WITH OTHER METHODS

In this section the performance of the proposed method is compared with common edge detection algorithms that are Prewitt, Canny, Roberts and Sobels algorithms. The reason for this selection is because of the popularity and high performance of these algorithms for edge detection of visible light images. The proposed method is also compared with edge detection using OTSU segmentation algorithm. For this purpose, the same input image is used as input for both k-means and also Mean Shift clustering method. As mentioned before for each clustering algorithm ANN is trained using the same clustering method. It means that the same clustering is applied in both training and also test parts. Both algorithms show significantly better result comparing with the common methods.

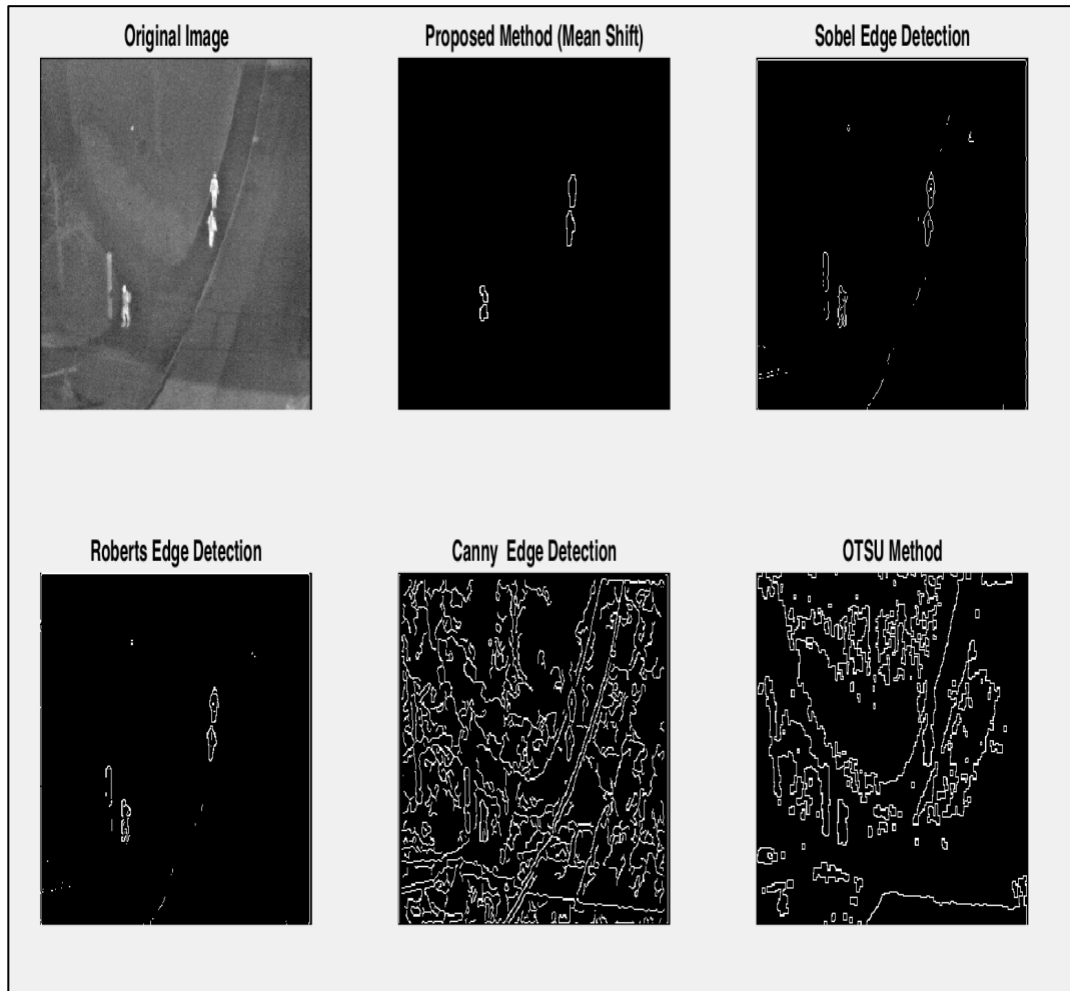
- i. K-means clustering:

Figure 5.1: Comparison of proposed method with the common algorithms



ii. Mean Shift clustering:

Figure 5.2: Comparison of proposed method with the common algorithms



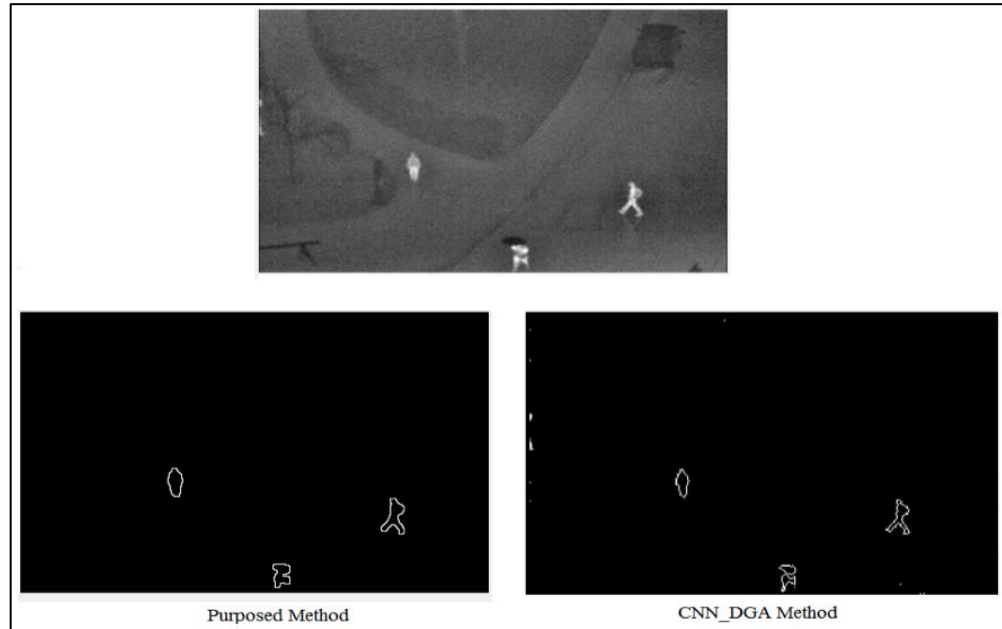
As the images indicate the outputs of both proposed methods have better results comparing with all the other algorithms.

5.2 COMPARING THE RESULTS WITH OTHER WORKS

In this section, the results from the proposed method are compared with the results of CNN-DGA method which is proposed by Wang et al. (2013). The algorithm is also compared with the result of Mixture Edge Detection Method proposed by Abdulmunim et al. (2012). For this purpose, the images and the results used in these studies reported in the papers are used for the comparison.

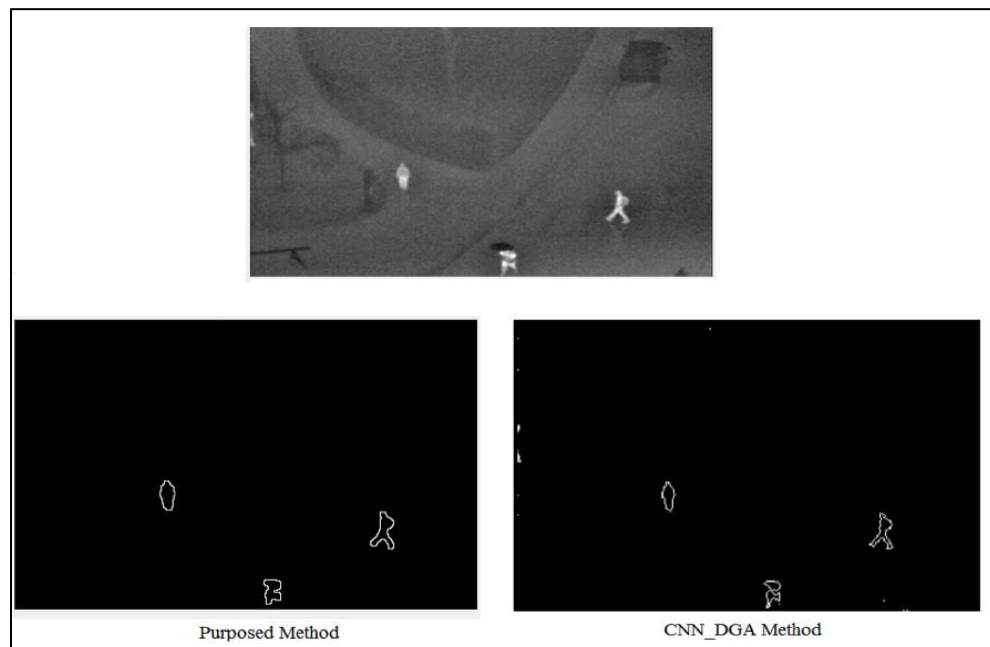
i. K-means Clustering:

Figure 5.3: Comparison of proposed method with CNN_DGA method



ii. Mean Shift Clustering

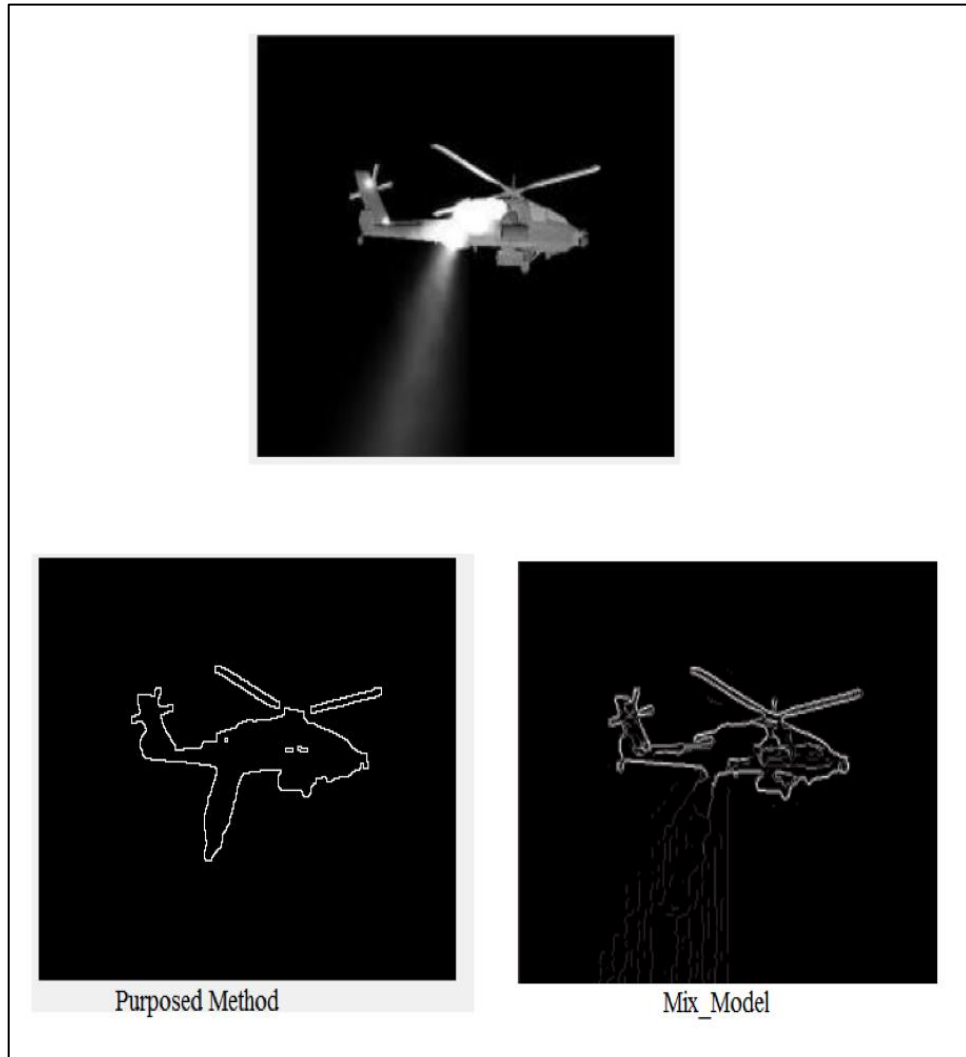
Figure 5.4: Comparison of proposed method with CNN_DGA method



The CNN-DGA result has some white noises in the output which don't exist in the output of both proposed methods. The selected clusters for K-means clustering method is three.

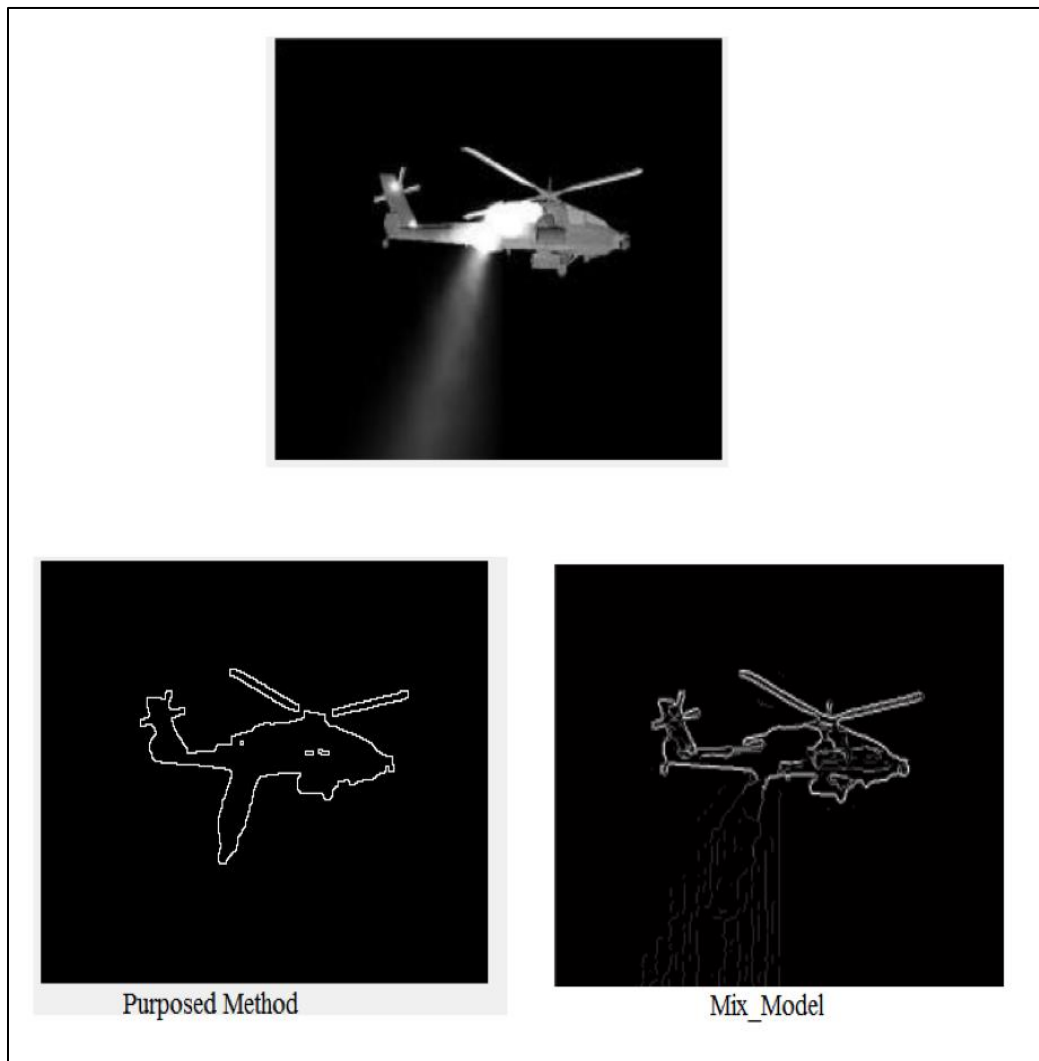
i. K-means Clustering

Figure 5.5: Comparison of the proposed methods with Mix_model



ii. Mean Shift Clustering:

Figure 5.6: Comparison of the proposed methods with Mix_model



The selected cluster number for K-means clustering method for this comparison is two. As indicated above, the proposed method has been able to perform well in detecting edges in infrared images.

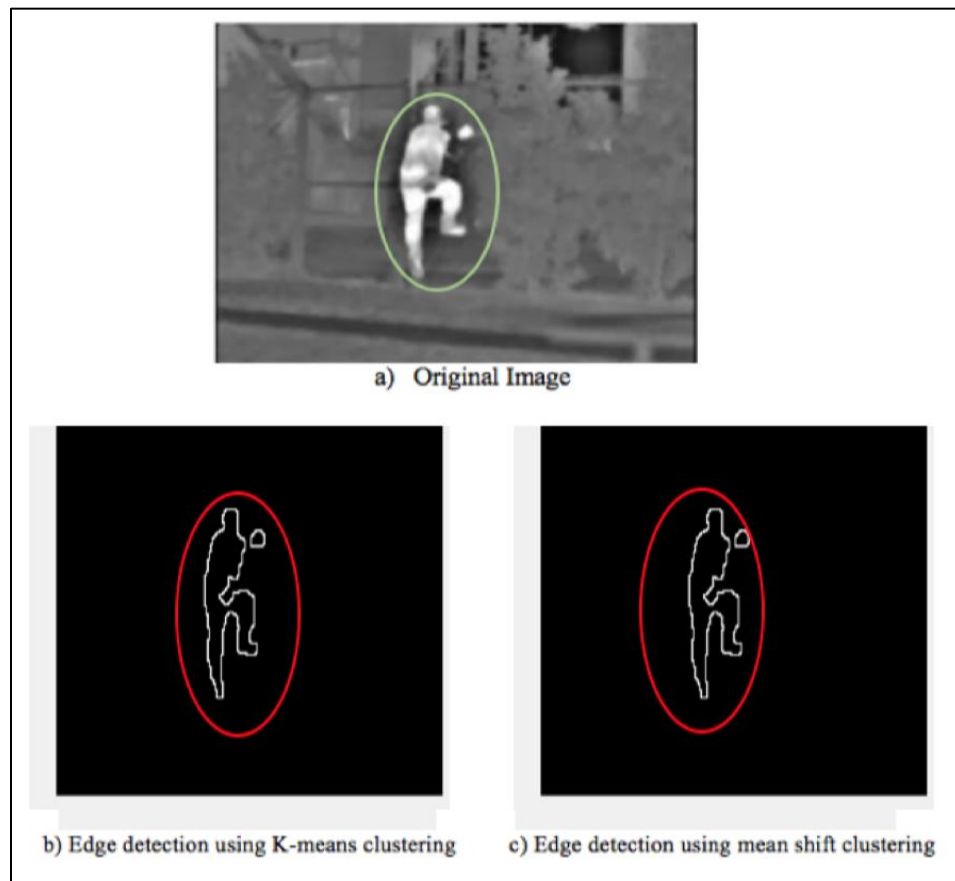
5.3 COMPARING BOTH CLUSTERING METHODS

In the last chapter the edge detection results using both methods for some random images from different datasets are selected and the edges are extracted using both methods. As mentioned in previous part the problem of k-means clustering method is finding the correct number of clusters and Mean shift is proposed to solve this problem. It has a great advantage of cluster number auto determination. For the following examples the correct

number of clusters for K-means could be determined. In this section some examples from different datasets are selected and the edges are detected using both clustering methods. For each example the Neural Network algorithm is trained using the same clustering method.

Figure 5.7.a shows the original image the green and red circles define the detectable edges. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 20 and the selected number of clusters for K-means clustering is 4.

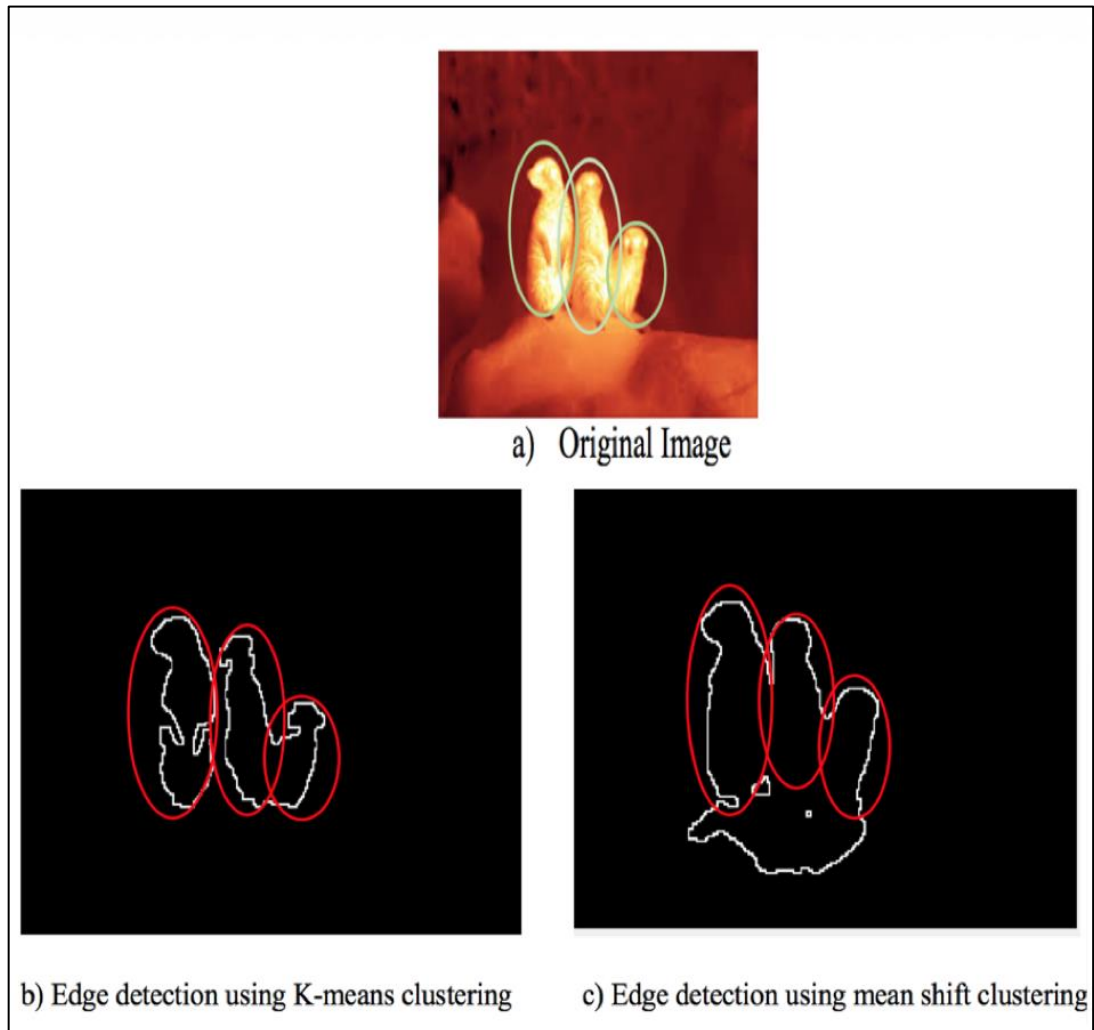
Figure 5.7: Comparison between k-means and Mean Shift clustering



As the results indicates both methods detect the edges correctly.

Figure 5.8.a shows the original image; the green and red circles define the detectable edges. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 10 and the selected number of clusters for K-means clustering is 4.

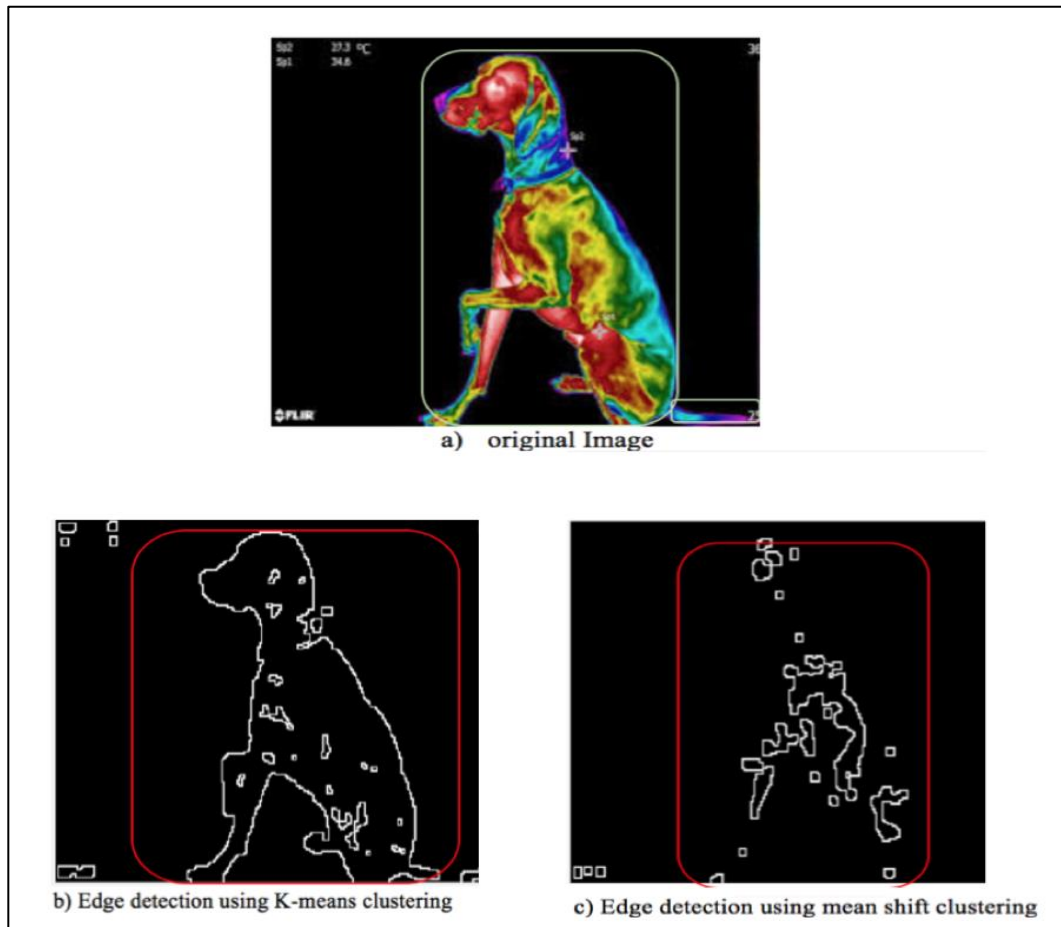
Figure 5.8: Comparison between k-means and Mean Shift clustering



As the results indicates in this example k-means has a better result rather than Mean Shift clustering method.

Figure 5.9.a shows the original image; the green and red circles define the detectable edges. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 10 and the selected number of clusters for K-means clustering algorithm is 2.

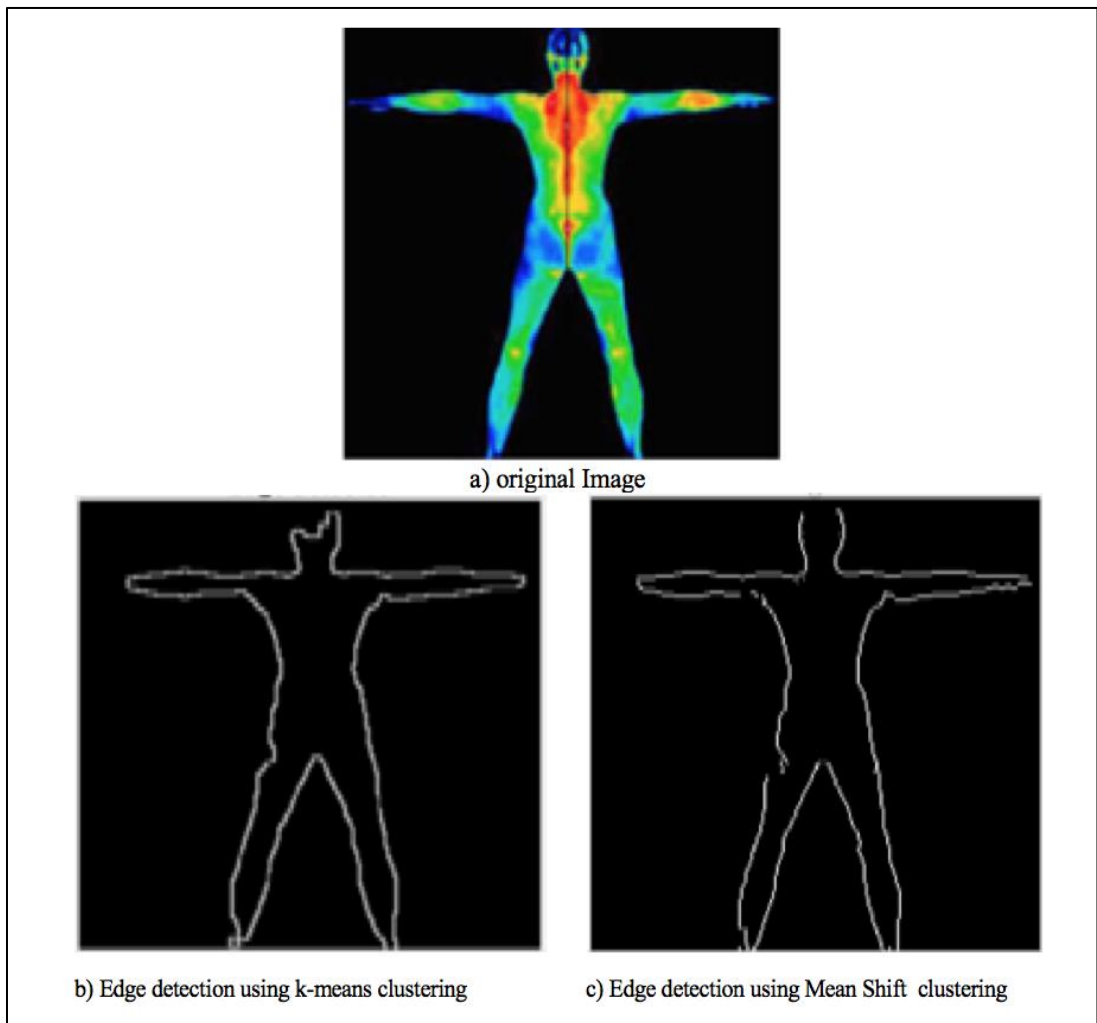
Figure 5.9: Comparison between k-means and Mean Shift clustering



As the results indicates in this example k-means clustering has a better result rather than Mean Shift method.

Figure 5.10.a shows the original image. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 20 and the selected number of clusters for K-means clustering algorithm is 2.

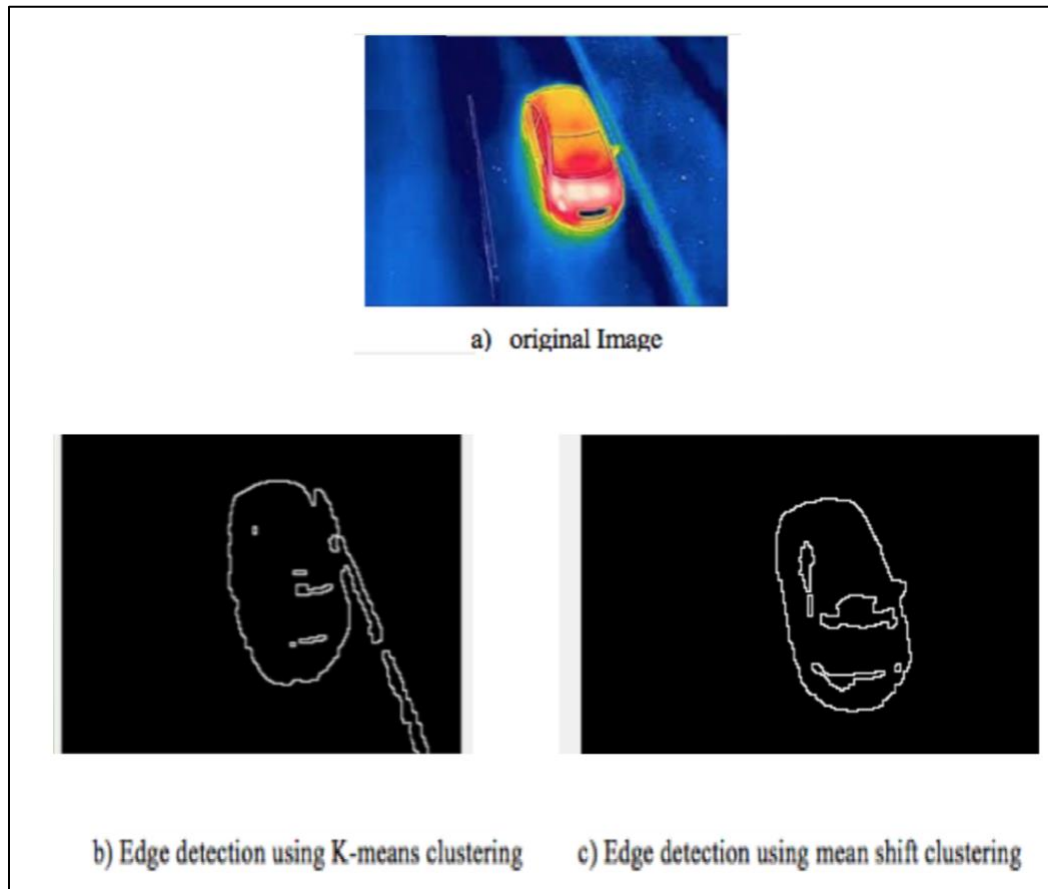
Figure 5.10: Comparison between k-means and Mean Shift clustering



In this example both methods perform well.

Figure 5.11.a shows the original image. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 8 and the selected number of clusters for K-means clustering algorithm is 4.

Figure 5.11: Comparison between k-means and Mean Shift clustering

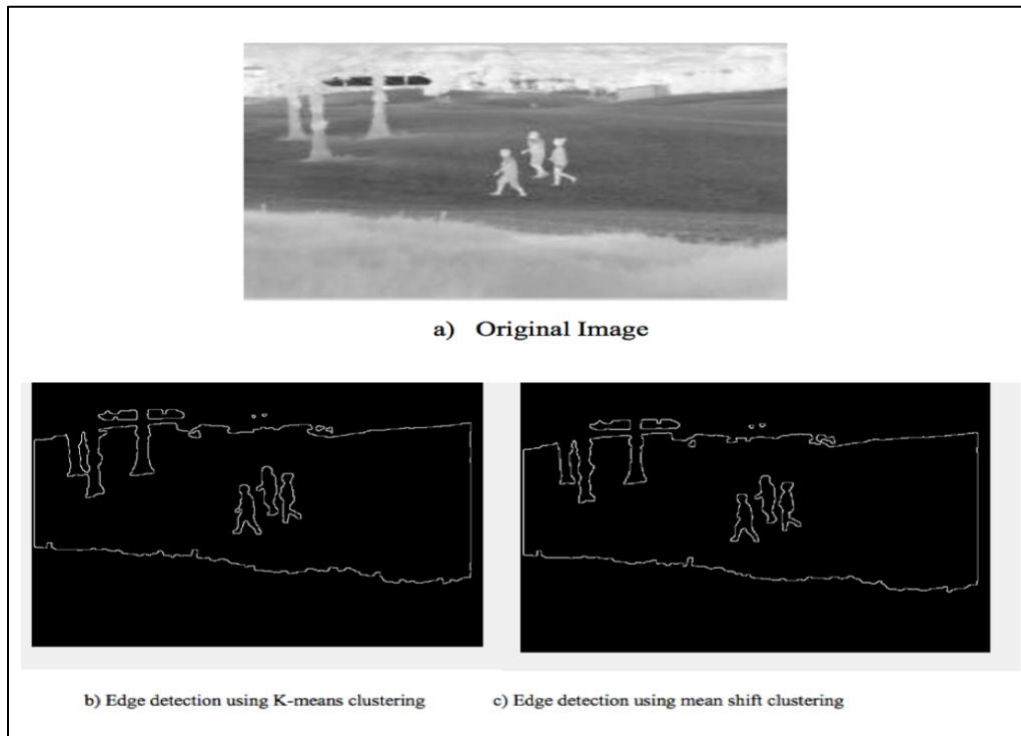


In this example if the aim is only detection of the car's edges, Mean Shift clustering has better results.

In general, both algorithms perform well. The important point is initialization step which is number of clusters for k-means clustering method and bandwidth selection for Mean Shift clustering algorithm. Variation of these values cause the different results.

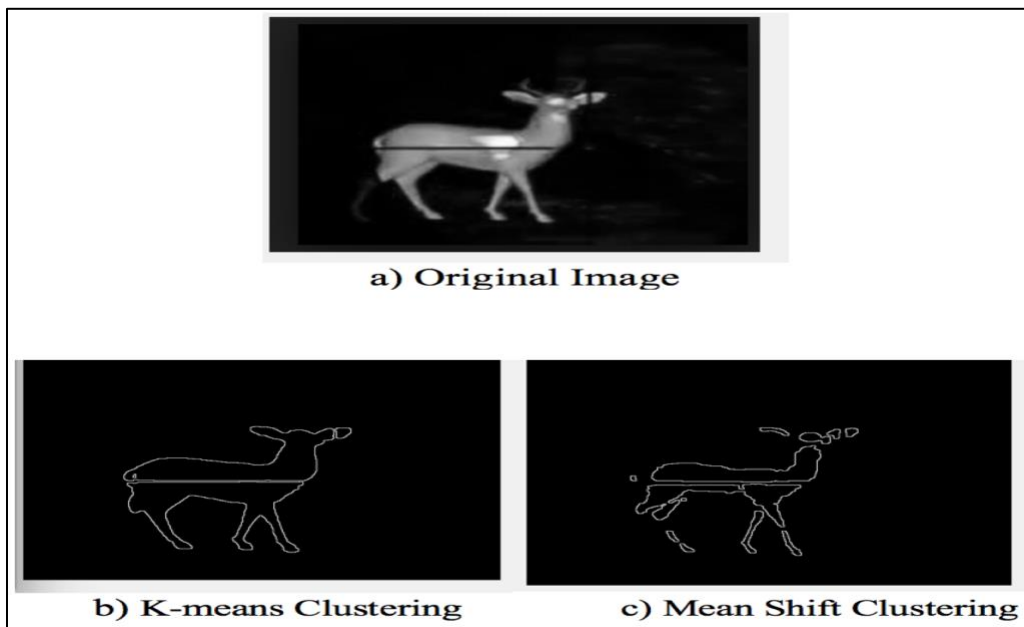
Figure 5.12.a shows the original image. Two different clustering methods are implemented. The bandwidth for Mean Shift clustering is 8 and the selected number of clusters for K-means clustering algorithm is 2.

Figure 5.12: Comparison between k-means and Mean Shift clustering



Both methods detected the edges well.

Figure 5.13: Comparison between k-means and Mean Shift clustering



In this example K-means clustering has a better result

6. CONCLUSION

6.1 INTRODUCTION

Infrared images are special images. These images have their own applications. Regarding the structure of these images, common methods are not suitable for processing and editing these images. It should be considered that image edge detection is one of the most important actions in image processing. The image edges provide important information for identifying objects in the image. In this study, due to the importance of infrared imaging, a new method is proposed for image processing. Therefore, in the first chapter of this research, the goals and generalities of the research were expressed. In the second chapter, literature and research background were studied. In the third chapter, the proposed method of research was described in detail and the implementation method was also explained thoroughly. In the fourth chapter, the results of the method were expressed, and the performance of this method was analyzed. In fifth chapter, the results of the proposed method were compared with other common edge detection algorithms, other studies and also themselves. In this chapter, the results of the research will be expressed and, finally, suggestions for future work will be presented.

6.2 THESIS RESULTS

In this thesis, the research is done to represent a proper method of IR image edge detection based on machine learning algorithms and image processing methods. The edges are detected by applying image segmentation concepts and extracting the Regions of Interest using the MLP, as well as the morphological operators of the infrared image edges. The general procedure in this study is as follows:

First, preprocessing was done on images to clear the pepper and salt noises on the image, using the median filter. Next, using k-means algorithm or Mean Shift clustering, the image was segmented into several clusters. Then, the region of interest was extracted using the Neural Network algorithm. After extracting the region of interest, using the morphological operators, the image holes were filled and the small unwanted points in the image were also deleted. Finally, the edges were extracted by using morphological operators.

In this study, for evaluation of the proposed method a standard dataset was used to survey the efficiency of method for the infrared image edge detection. The results of the proposed method were compared with common edge detection methods. Comparing the results with Prewitt, Canny, Roberts, and Sobel operators showed that the proposed method has been able to extract the edges of the image well and have better performance than these operators. In order for further evaluation of the effectiveness of the proposed method, the results of the proposed method were examined with other works, in each case the results also showed that the proposed method was able to show good performance. Since two different clustering methods were applied these two methods were compared by confusion matrices, running time and different random IR images. In general, proposed method based on clustering and extraction of the region interest, has shown good performance and could be used in different relevant applications.

6.3 PROPOSAL FOR FUTURE RESEARCH

Certainly, any research work done has some advantages and disadvantages. This research is not an exception. So, here are some suggestions for future work:

- a. Applying other techniques and image segmentation methods.
- b. Applying other classification methods (simpler neural network methods) to extract the expected area.
- c. As stated in the proposed method, k-means clustering is one of the applied segmentation algorithms. Although elbow criterion and Mean Shift clustering method are applied to solve the problem of cluster number manual determination, but other methods could be applied to extract the edges.
- d. This method could be extended in video sequences especially for surveillance purposes.

REFERENCES

Books

- Bhanu, B., & Pavlidis, Ioannis. 2005. *Computer Vision Beyond the Visible Spectrum* Advances in Pattern Recognition. London: Springer London.
- Gilbert Strang, 2005, *Linear Algebra and Its Applications*, Brooks Cole,
- Gonzalez, R., & Woods, Richard E. 2008. Chapter 2: Moving Object Detection & Tracking in Videos. *Digital image processing*. 3rd edn. Upper Saddle River, N.J.: Prentice Hall.pp.15-39
- Gonzalez, R., Woods, & Richard, E, 2002. *Digital image processing* Issue no:2. Upper Saddle River, N.J.: Prentice Hall.
- Iftexharuddin, Khan, & Awwal, Abdul. 2012. *Field Guide to Image Processing*. SPIE Press.
- Kohonen, T. 1997. *Self-organizing maps*, Springer series in information sciences; 30). Berlin; New York: Springer.
- Marques, O. 2011. *Practical Image and Video Processing Using MATLAB*. Hoboken, NJ, USA: John Wiley & Sons.
- Qiu, Peihua. 2005. *Wiley Series in Probability and Statistics*. Hoboken, NJ, USA: John Wiley & Sons.
- Szeliski, R. 2011. *Computer Vision : Algorithms and Applications* (Texts in computer science). London: Springer-Verlag London Limited.
- Woolfson, M. 2012. *The fundamentals of imaging : From particles to galaxies*. London : Hackensack, NJ: Imperial College Press ; Distributed by World Scientific Publishing.
- Zhang, W. 2005. Computational ecology: *Artificial neural networks and their applications*. Singapore; Hackensack, NJ; London: World Scientific.

Periodicals

- Abdulmunim, Matheel, E. M, Suhad, 2012. Propose a Mixture Edge Detection Method for Infrared Image Segmentation. *British Journal of Science*. **6** (2)
- Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, **65**(1), pp. 2-16.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*. **65**(1), pp. 2-16.
- Comaniciu, D., Ramesh, V., & Meer, P. 2003. Kernel-based object tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **25**(5), pp. 564-577.
- Comaniciu, Meer, 2002. Mean Shift a robust approach toward feature space analysis. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE PAMI*, **24** (5), pp. 603-619.
- Davis, J., Keck, M., 2005. A two-stage approach to person detection in thermal imagery *IEEE OTCBVS WS Series Bench*. In Proc. Workshop on Applications of Computer Vision pp. 364-369
- Dhanachandra, Manglem & Chanum, 2015. Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm. *Procedia Computer Science*. **54**, pp. 764-771.
- Dimitris, Manolakis, Marden, David & Shaw, Gary A, 2003. Hyperspectral image processing for automatic target detection applications. *Lincoln laboratory journal* **14** (1), pp. 79-116.
- Duarte, Carrão, Espanha, Viana, Freitas, Bártolo, . . . Almeida, 2014. Segmentation Algorithms for Thermal Images. *Procedia Technology*. **16** (C). pp, 1560-1569.
- Dubey, Shiv Ram, Anand, Singh Jalal, 2012. Adapted approach for fruit disease identification using images. *International Journal of Computer Vision and Image Processing*, **2**(3), 44-58.
- Garge, D.M. Bapat, V.N. 2009. A low cost wavelet based mammogram image processing for early detection of breast cancer. *Indian Journal of Science and Technology*. **2** (9)
- Iscan, Yüksel, Dokur, Korürek, & Ölmez. 2009. Medical image segmentation with transform and moment based features and incremental supervised neural network. *Digital Signal Processing*. **19** (5), pp. 890-901.

- Jambhorka, Sagar, G.N.Sarage, 2012. Enhancement of Chest X-Ray images Using Filtering Techniques. *International Journal of Advanced Research in Computer Science and Software Engineering* **2** (5), PP.308-312
- Kannan, Ramathilagam, Devi, & Sathya, 2011. Robust kernel FCM in segmentation of breast medical images. *Expert Systems With Applications*, **38** (4), pp. 4382-4389.
- Lahiri, Bagavathiappan, Jayakumar, & Philip. 2012. Medical applications of infrared thermography: A review. *Infrared Physics and Technology*, *Infrared Physics and Technology*. **55** (4)
- Mohammad, M. B, R. N, Srujana, A. J. N., Jyothi & P. B. T, Sundari. 2016. Disease Identification in Plants Using K-means Clustering and Gray Scale Matrices with SVM Classifier. *International Journal of Applied Sciences, Engineering and Management*.**5** (2), pp. 84 – 88.
- Pasagic, V., Muzevic, M., & Kelenc, D. 2008. Infrared Thermography In Marine Applications. *Brodogradnja*, *59*(2), 123-130.
- Qingju ,Tang, Chiwu, Bu, Yuanlin, Jiansuo, Liu Zang, & Li Dayong, 2016. Infrared Image Edge Detection Based on Morphology-Canny Fusion Algorithm. *International Journal of U- and E- Service, Science and Technology*, **9**(3), pp. 259-268.
- Seema. Kumar, A. & Gill, G. S, 2015. Computer Vision based Model for Fruit Sorting using K-Nearest Neighbor classifier. *Int, Journal of electrical &Electronics Engg.* **2** (1), pp. 49-52
- Singh, S. 2011. Artificial Neural Network. *Nature Precedings*, Nature Precedings.
- Sun, S, 2003. Automatic target recognition using boundary partitioning and invariant features in forward-looking infrared images. *Optical Engineering*. **42** (2), pp. 524-533.
- Wang, Yang, Xie, & An, 2014. Edge detection of infrared image with CNN_DGA algorithm. *Optik - International Journal for Light and Electron Optics*. **125** (1), pp. 464-467.
- Yizong Cheng. 1995. Mean shift, mode seeking, and clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **17**(8), 790-799.

Other Publications

- Dang Hongshe, Song, Jinguo & Guo, Qin, 2010. A Fruit Size Detecting and Grading System Based on Image Processing. *Second international conference on intelligent human-machine and cybernetics*. 2010 School of Electric and Information Engineering, Shaanxi University of Science and Technology China.
- Khuwaja, G., & Tolba, A. 2000. Fingerprint image compression. *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop*. **2**, pp. 517-526.
- MATLAB and Statistics Toolbox Release 2015b, The MathWorks, Inc., Natick, Massachusetts, United States
- Mercol. Juan Pablo, Gambini María Juliana, & Juan Miguel Santos, 2008. Automatic classification of oranges using image processing and data mining techniques. *XIV Congreso Argentino de Ciencias de la Computación*.
- Pennestate Eberly College of Science Human Body, Form & Function 2018, https://online.science.psu.edu/bisc004_activewd001/node/1907 [retrieval date 11 Feb 2018].
- Peshin, Akash., 2017, Why Are Infrared Waves Associated With Heat? [online]. Science ABC. www.scienceabc.com/pure-sciences/why-are-infrared-waves-associated-with-heat.html [accessed 4 January 2018].
- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Robotic surgery-advantages and disadvantages, Future medical technology 2011 <http://www.futuretechnology500.com/index.php/future-medical-technology/robotic-surgery-advantages-and-disadvantages> [accessed 5 February 2018].
- Seyedarabi, 2010, Image processing, *Image processing*, Image processing applications. University of Tabriz. Nasri.
- Shirazi, M., & Morris, B, 2015. Vision-based vehicle queue analysis at junctions. *Advanced Video and Signal Based Surveillance (AVSS), 2015 12th IEEE International Conference on*, pp. 1-6.

- Shunyong, Zhou., Yang, Pingxian, 2011. Infrared image segmentation based on Otsu and genetic algorithm. *Multimedia Technology (ICMT), 2011 International Conference on*. IEEE. 30 August 2011. Hangzhou, China, pp. 5421-5424.
- Sivaramakrishnan, R., Antani, Sameer. Candemir, Sema. & Xue, Zhiyun, Abuya, Joseph, et al. , 2018. Comparing deep learning models for population screening using chest radiography. *Medical Imaging: Computer-Aided Diagnosis*. **10575**. International Society for Optics and Photonics. February 2018, Houston, Texas, United States, pp. 2-12
- Song, Yuheng, Hao, Yan, 2017. Image Segmentation Algorithms Overview. *eprint arXiv preprint*.
- Wang, Dong, & Zhang, Jingzhou. 2011. Infrared image edge detection algorithm based on sobel and ant colony algorithm. *Multimedia Technology (ICMT), 2011 International Conference on*, 4944-4947.

APPENDICES

Appendix A.1 Running time

In this section the running time of all images whose edges are extracted in fourth and fifth chapter are calculated:

Table-1: Calculated Running Time for Images of fourth and fifth chapter

| Figure Number | Running time of K-means Clustering | Running time of Mean Shift Clustering |
|----------------------|---|--|
| Figure 4.1 | 3.609888 sec | 3.072522 sec |
| Figure 4.2 | 3.767448 sec | 3.644015 sec |
| Figure 4.3 | 3.502037 sec | 3.501053 sec |
| Figure 4.4 | 3.620084 sec | 3.280337 sec |
| Figure 5.5 | 3.664335 sec | 3.243459 sec |
| Figure 4.6 | 3.817510 sec | 3.162980 sec |
| Figure 4.7 | 3.157077 sec | 3.728421 sec |
| Figure 4.8 | 3.507092 sec | 3.221922 sec |
| Figure 4.9 | 3.180126 sec | 3.796464 sec |
| Figure 4.10 | 3.157077 sec | 2.607492 sec |
| Figure 4.11 | 4.541129 sec | 2.475525 sec |
| Figure 4.12 | 3.936479 sec | 2.600794 sec |
| Figure 4.13 | 3.869416 sec | 3.234204 sec |
| Figure 4.14 | 3.790868 sec | 3.138017 sec |
| Figure 4.15 | 4.501476 sec | 3.518253 sec |
| Figure 4.16 | 4.181246 sec | 2.417631 sec |
| Figure 5.1-5.2 | 3.723121 sec | 3.803506 sec |
| Figure 5.3-5.4 | 3.101607 sec | 2.009334 sec |
| Figure 5.5-5.6 | 2.995475 sec | 2.263715 sec |
| Figure 5.7 | 2.995456 sec | 2.293981 sec |
| Figure 5.8 | 2.953164 sec | 2.151352 sec |
| Figure 5.9 | 2.935648 sec | 2.128833 sec |
| Figure 5.10 | 2.696574 sec | 2.008255 sec |
| Figure 5.11 | 2.991042 sec | 2.080714 sec |
| Figure 5.12 | 3.250820 sec | 2.671819 sec |
| Figure 5.13 | 4.850709 sec | 5.616439 sec |

Appendix A.2 Confusion matrix

Following Tables (Table-2 to Table-15) are confusion matrices of chapter fourth images. For all fourteen images confusion matrices are calculated for both methods. Total result of all confusion matrices including Table 4.1 and Table 4.2 are calculated as Table 4.3 .

Table-2: Confusion matrix of Figure 4.2

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=2515 | FN=9 | K-means Clustering |
| Actual=No | FP=0 | TN=83876 | |
| Actual=yes | TP=2524 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=83876 | |

Table-3: Confusion matrix of Figure 4.4

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=2970 | FN=14 | K-means Clustering |
| Actual=No | FP=405 | TN=83011 | |
| Actual=yes | TP=2984 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=521 | TN= 82895 | |

Table-4: Confusion matrix of Figure 4.5

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 3996 | FN=0 | K-means Clustering |
| Actual=No | FP= 308 | TN= 82096 | |
| Actual=yes | TP=2984 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=521 | TN= 82895 | |

Table-5: Confusion matrix of Figure 4.6

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=4467 | FN=4 | K-means Clustering |
| Actual=No | FP=308 | TN=81621 | |
| Actual=yes | TP=4467 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=521 | TN= 81412 | |

Table-6: Confusion matrix of Figure 4.7

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=4440 | FN=99 | K-means Clustering |
| Actual=No | FP=308 | TN=81553 | |
| Actual=yes | TP=4364 | FN=175 | Mean Shift Clustering |
| Actual=No | FP=168 | TN=81693 | |

Table-7: Confusion matrix of Figure 4.8

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 2664 | FN=0 | K-means Clustering |
| Actual=No | FP= 308 | TN=83428 | |
| Actual=yes | TP=2631 | FN=33 | Mean Shift Clustering |
| Actual=No | FP=168 | TN=83568 | |

Table-8: Confusion matrix of Figure 4.9

| Pixels= 86400 | Detected =Yes | Detected =No | K-means Clustering |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 3878 | FN=12 | |
| Actual=No | FP=168 | TN=82342 | |
| Actual=yes | TP= 3878 | FN=12 | Mean Shift Clustering |
| Actual=No | FP=168 | TN=82342 | |

Table-9: Confusion matrix of Figure 4.10

| Pixels= 86400 | Detected =Yes | Detected =No | K-means Clustering |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 1205 | FN=0 | |
| Actual=No | FP=0 | TN=85195 | |
| Actual=yes | TP= 1205 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=85195 | |

Table-10: Confusion matrix of Figure 4.11

| Pixels= 86400 | Detected =Yes | Detected =No | K-means Clustering |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 3517 | FN=0 | |
| Actual=No | FP=0 | TN=82883 | |
| Actual=yes | TP=3477 | FN=40 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=82883 | |

Table-11: Confusion matrix of Figure 4.12

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 1725 | FN=0 | K-means Clustering |
| Actual=No | FP=0 | TN=84675 | |
| Actual=yes | TP=1685 | FN=40 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=84675 | |

Table-12: Confusion matrix of Figure 4.13

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 2379 | FN=0 | K-means Clustering |
| Actual=No | FP=168 | TN=83853 | |
| Actual=yes | TP=2359 | FN=20 | Mean Shift Clustering |
| Actual=No | FP=0 | TN=84021 | |

Table-13: Confusion matrix of Figure 4.14

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP=2155 | FN=0 | K-means Clustering |
| Actual=No | FP=168 | TN=84077 | |
| Actual=yes | TP=2127 | FN=28 | Mean Shift Clustering |
| Actual=No | FP=168 | TN=8077 | |

Table-14: Confusion matrix of Figure 4.15

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 1648 | FN=0 | K-means Clustering |
| Actual=No | FP=168 | TN=84584 | |
| Actual=yes | TP= 1648 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=168 | TN=84584 | |

Table-15: Confusion matrix of Figure 4.16

| Pixels= 86400 | Detected =Yes | Detected =No | |
|----------------------|----------------------|---------------------|------------------------------|
| Actual=yes | TP= 638 | FN=0 | K-means Clustering |
| Actual=No | FP=168 | TN= 85594 | |
| Actual=yes | TP= 638 | FN=0 | Mean Shift Clustering |
| Actual=No | FP=168 | TN= 85594 | |