A NEW CLASSIFICATION APPROACH BASED ON GEOMETRICAL MODEL FOR HUMAN DETECTION IN IMAGES

By

Malek Al-Nawashi

Supervised By

Dr Mohammad Saraee

School of Computing, Science & Engineering

University of Salford

Submitted in Partial Fulfilment of the Requirements of the Degree of Doctor of Philosophy

2019

A NEW CLASSIFICATION APPROACH BASED ON GEOMETRICAL MODEL FOR HUMAN DETECTION IN IMAGES

ABSTRACT

In recent years, object detection and classification has gained more attention, thus, there are several human object detection algorithms being used to locate and recognize human objects in images. The research of image processing and analysing based on human shape is a hot topic due to its wide applicability in real applications. In this research, we present a new shape-based classification approach to categorise the detected object as human or non-human in images. The classification in this approach is based on applying a geometrical model which contains a set of parameters related to the object's upper portion. Based on the result of these geometric parameters, our approach can simply classify the detected object as human or non-human. In general, the classification process of this new approach is based on generating a geometrical model by observing unique geometrical relations between the upper portion shape points (neck, head, shoulders) of humans, this observation is based on analysis of the change in the histogram of the x values coordinates for human upper portion shape. To present the changing of X coordinate values we have used histograms with mathematical smoothing functions to avoid small angles, as the result we observed four parameters for human objects to be used in building the classifier, by applying the four parameters of the geometrical model and based on the four parameters results, our classification approach can classify the human object from another object.

The proposed approach has been tested and compared with some of the machine learning approaches such as Artificial Neural Networks (ANN), Support Vector Machine (SVM) Model, and a famous type of decision tree called Random Forest, by using 358 different images for several objects obtained from INRIA dataset (set of human and non-human as an object in digital images). From the comparison and testing result between the proposed

approach and the machine learning approaches in term of accuracy performance, we indicate that the proposed approach achieved the highest accuracy rate (93.85%), with the lowest miss detection rate (11.245%) and false discovery rate (9.34%). The result achieved from the testing and comparison shows the efficiency of this presented approach.

Publications

2

- Obaida M. Al-Hazaimeh1&Malek Al-Nawashi &Mohamad Saraee **(Geometrical-based approach for robust human image detection)** Multimedia Tools Application, received: 22 September 2017 / Revised: 27 June 2018 / Accepted: 11 July 2018#Springer Science +Business Media, LLC, part of Springer Nature 2018

- Malek Al-Nawashi1, • Obaida M. Al-Hazaimeh1,• Mohamad Saraee1(**A novel framework for intelligent surveillance system based on abnormal human activity detection in academic environments)** Neural Computer & Application (2017) 28 (Suppl 1):S565–S572 , 2016. This article is published with open access at Springerlink.com

ACKNOWLEDGEMENT

I am grateful to the Almighty Allah for giving me the opportunity to complete my PhD thesis.

I would like to thank my supervisors Dr. Mohamad Saraee, and Dr. Obaida M. Al-hazaimeh who have encouraged and helped me very much in doing this research.

I owe special dedication to the memory of my mother, who always believed in my ability to be successful in the academic arena. You are gone, but your belief in me has made this journey possible thanks to the spirit of my mother.

I would like to thank my family for their constant support, and a special thanks to my wife for her support and encouragement.

I would like to thank the staff members at the University of Salford for their constant Support.

Contents

List of figures

List of tables

CHAPTER ONE

1 MOTIVATION AND INTRODUCTION

1.1 MOTIVATION AND INTRODUCTION

Nowadays, human object detection and recognition is a key ability which is required by most image processing and computer vision algorithms. The term human object detection can be introduced as a process of localizing all objects that are human in the images by detecting and identifying human features [7]. To detect the human object in images, we need the power of computer vision and image processing algorithms in terms of accuracy, efficiency, and flexibility to be able to identify and extract the human object features among different objects and then classify the detected features as human or non-human. In particular, human object detection and recognition has become a challenge for researchers nowadays in computer vision and image processing areas due to the fact that different objects tend to share many features and properties which are usually used for human object detection [120].

In computer vision, the human object detection algorithms are known as process-based multitasking, which means a set of processes can be executed simultaneously and concurrently, such as motion behaviour detection, human detection, object classification in images, face recognition, tracking and so many more [121]. Generally, human object detection algorithms can be used in several applications in different ways to control access to sensitive and public areas such as streets, train stations, parks, airports, malls and many other public areas [13]. Despite of the all advantages, the human object detection and classification algorithms are facing challenges and difficulties when detecting and then classifying objects as human or non-human from the images background changing and illumination, and object viewpoint. In order to ensure the quality service of object detection and classification, appropriate image processing algorithms in terms of accuracy, efficiency, and flexibility are needed[122].

The researchers in the area of computer vision and image processing have adopted object detection and classification in their research. Some researchers focus on faces (i.e. Direct face) and bodies (i.e. Visibility) at high image spatial resolution. Some of these proposed algorithms which are implemented in the large systems and real-time applications are costly and require significant development time and computer resources for matching based classification (i.e. High complexity). In computer vision, there are four major categories for both object detection and classification, the categories for object detection include: flow analysis, dynamic threshold, temporal differencing, and background subtraction and the categories for object classification include: colour, texture, motion, and shape [2].

One of the human object detections appreciates, which is called "State-of-the-Art detectors based on HOG features", but some of the HOG features cannot cover all the diversity and changing in the presented object model. In fact, the human detection methods based on the rich colour cues are not commonly used, due to the variety of clothing colours that the human can wear. The experimental result of the "State-of-the-Art detectors based on HOG features", shows an efficient way to detect the human specially in images, but it has a limitation to detect the human in videos because it used an off line fixed data set so that it is difficult to manipulate different factors in video such as inconstant background, changing of camera pose and the light and shadow challenging, [2]. Furthermore, it is built based on a generic object class using the [2] HOG features, so that it is cannot take the advantage of the special scene information provided in video at different frames, such as the static colour pattern of the foreground and background in case of implementing a special video. The detection accuracy of these kinds of methods seem to be inadequate and not efficient in cases where the human face is not visible or not clearly recognized, which may be caused if the person is too far away or out of view of the camera. Also, in somecases, the human body elements may be hidden or partially occluded, which affects the detection result and the detection accuracy. An example of false detection can be seen in July 2015, the dailymail.co.uk published By Richard Gray for mail online reported that Jacky Alcine claimed that Google Photos recognized him and his girlfriend as Gorillas, forcing Google to apologise after its image recognition software mislabelled photographs of some people as gorillas. [4]

Figure 1.1 Google recognizes some people as gorillas

From the observation of human detection vision approaches and the limitation of these approaches when detecting human objects in images and videos there are some issues, also by understanding the important role of human object detection in our day to day lives there are many applications therefore, the contribution of this research aims to present a new computer vision approach to classify the detected object in images as human or non-human, in high accuracy and sufficient ways. The classification process of this approach is based on a new geometrical model which can classify the detected object as a human by extracting some features of human shape.

1.2 RESEARCH QUESTIONS

During this study, the main question which has been identified is; how to develop a robust human object detection approach in images based on human shape. To answer the main question, we have to address the following questions:

1. How to detect the objects from the images.
2. How to extract a unique human feature from the human upper portion shape.
3. How to develop an effective approach to classify the detected object in images as human and to be insensitive to illumination conditions and occlusion.

1.3 RESEARCH AIMS AND OBJECTIVES

The main objective of this research is to propose a new shape-based approach to improve the performance of human object detection in terms of human or non-human in images. To achieve this goal, there are sub-objectives that have been set out as follows:

1. To identify and extract the unique features of human object shape from images.
2. To set up a geometrical model based on the extracted unique features for human object shape.
3. To design and build a new classification approach based on the mathematical model to detect the human object in images based on human upper portion shape.

1.4 CONTRIBUTION

The process of human detection is to extract and localise all human objects in images, this process of detection requires finding certain features of the human object, these must be special features that can specify human objects from other objects in the image.

Despite all the benefits of research, the performance of human object detection vision approaches is still far from what could be efficient and used reliably under the restrict environment. However, there are some factors can affect the performance of detection in term of the vision process, these factors refer to the unconstrained environments such as the pose of the camera, illumination in images, and fully and partially occluded objects.

Therefore, the contribution of this study is to propose a new approach that can enhance the performance of human object detection for automatically identifying and detecting human object in images. The new approach will enhance the performance of human object detection by classifying the detected object based on it is own geometrical model, this geometrical model consists of four parameters which are extracted from the human upper portion shape and to be used in classifying the detected object as human or non-human. Classifying the object based on the upper portion shape play a major part when addressing many challenging factors in detecting and classifying objects, such as the changes in the illumination, the camera pose and the object occluded, because the upper portion are usually clear and hard to cover.

1.5 ORGANIZATION OF THESIS

All in all, this thesis consists of six chapters. The thesis is organized as follows: **In Chapter 1**, an argument as to why the current study needs to be conducted has been discussed. In particular, this study seeks to propose a new algorithm that could improve the classification accuracy and performance. **Chapter 2** presents the background and related work which is a background for this study and other peoples works so far. The next chapter, i.e. **Chapter 3,** describes the design and implementation of the new approach. **Chapter 4** elaborates on the experiment results. **Chapter 5** discusses the performance evaluation and the comparison between the proposed approach and other approaches in terms of classification accuracy and the computation time. **Chapter 6** concludes the thesis. In this chapter, we present some of the contributions of this research. Future directions and future works are also demonstrated.

CHAPTER TWO

2 BACKGROUND AND RELATED WORK

2.1 INTRODUCTION

Computer vision has several active research subjects, but detecting human objects is considered one of the most active ones, due to the wide implementation in real applications. Human detecting could be interpreted as, the process of determining all the human objects in the sequences of videos and images, by ascertaining the human qualities present in videos and images. For a robust detection, we need the good capabilities of the approaches of computer vision, which must have the ability to bring out the mutual qualities between various humankinds from the videos and images, so they can be localised and separated from the background. This task has become quite challenging for researchers in computer vision areas due to the fact that different people tend to have different features which are usually caused by their variable appearances and postures. This task of human detection also corresponds to determining the region that contains human objects in the images or sequence videos, which leads the computer vision approach to start detecting objects in the images and video sequences and then to identify and classify the detected objects as humans or non-humans based on the human features depending on the system goals.

In the last decade, the task of human detection has risen to be an integral part of various real life applications especially in areas that required surveillance [7, 8], due to the large amount of visual data that the outcome of these applications need to be processed and managed. Similarly, the computational ability of tagging or labelling the images or the sequence videos based on multimedia visual content dataset will provide an efficient way for enabling subsequence search or retrieval, therefore the manual approach to do the same task is labour intensive and may be prone to ambiguity and errors. The main functions in the human detection sequence process are used to detect the human, while there are many post-processing functions used to achieve additional goals according to real applications.

These goals may vary depending on the advance task of the process such as counting the number of humans in images [9], recognize the person based on his face [10] and motion and behaviour detection, for example abnormal behaviour detection [11]. In general, these types of functions may use different types of real applications and address various aspects of purpose, such as a security in high sensitive areas such as airports, train stations, supermarkets, and many surveillance application systems, the cost of these systems tends to be high due to the cost of setting up the surveillance system equipment and infrastructure aspects of communication and computer processing [12]. There are different types of methods being used for human detection based on the material of the multimedia, such as detecting objects in images or in video because the appropriate method for detecting objects in an image can be different from the appropriate method used to detect the object in videos, this difference is due to changes in the background [13] [14]. Both of these methods can extract the features from images however, because the video is a sequence of images, it can utilise the motion features using extracting methods, for example, optical flow method [15], and foreground extraction method, [16] but these methods are not corresponding to extract objects from a static image base.

The task of recognizing and detecting objects such as humans in images and video sequences point research attention to the fields of computer vision and machine learning, around the world due to its wide applicability, scope and for the large potential applications that can be acquired, such as assistance system for auto- drive, monitoring systems, efficient graphic user interface, motion personification, and so on. The term for detecting humans in images and video sequences are defined as the process of identifying and localizing human existence by extracting its features and then distinguishing them from other inhumane objects. However, the task of detecting humans in images and video sequences, is a very difficult and complex task compared to other functions of detecting objects, because the human object may affect or contain some factors which make it different than other objects, such as: varying human pose, appearance and clothing, changing camera positions, changing the background dynamically, and sudden or gradual changes in lighting.

Nowadays there are numerous human detection methods, but most of them require a special condition to achieve intensive accuracy such as direct high-precision face, or for the whole body to be visible. Also, for classification and matching some other methods require an extensively huge vision content database. Many specific human detection approaches classify the human based on his shape by searching for circle shape to detect the head of the human from images or video sequences. Most of these human detection methods seem to be insufficient due to viewpoint appearance, for example, if the human object location is not close or not clearly visible to the camera which tends to make the face not visible. Also, in case of many objects, the component-based detection rises to be insufficient because of partial object occlusion. In many cases the identification or detection of humans in images or video is based on human head shape matching with circular or oval patterns that may lead to inaccurate detection in cases where the picture frame consists of other objects rather than humans that have the same circular shapes. Similarly the shape of a human head may not be round or oval all the time [17].

For human object class, there are specific additional difficulties or challenges in human detection. Firstly, the appearance of humans in images or video sequences can be varied rather than changing the viewpoint, but also change of pose and juxtaposition of body parts that are due to the ability of the human body to articulate at many joints. Secondly, human beings, are unlike others object because they can wear a variety of clothes and accessories which makes the detection task more complex in most human detection methods because clothing and wearing accessories, which have different colours and mixed colours and in some cases the clothing may contain different textures. Thirdly, human beings are difficult to detect because they can handle or carry other objects, or they can ride on another object like a bicycle or they can be obstructed in different ways. For these reasons it is a complicated process to recognize and detect humans in images and videos or to find a unique pattern to represent the humans. [18]. Figure 2.1 shows some examples of pedestrians in different poses taken from surveillance videos.

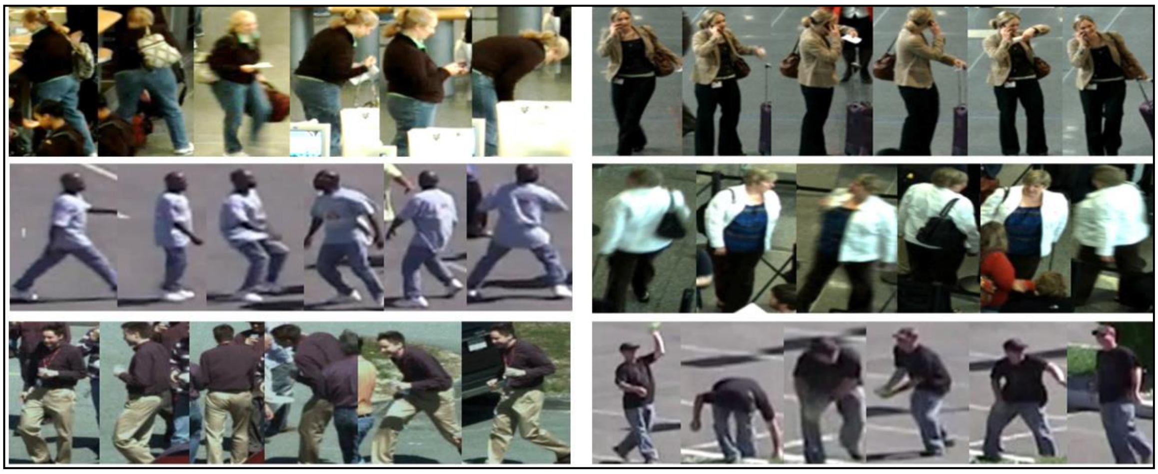


Figure 2.1 Some samples of pedestrians in different poses taken from surveillance video

The occlusion or partial occlusion by other objects is another problem for detecting interest and it is a big challenge for detecting a single human object or separate human object when they very close to other objects or are partially occluded. Figure 2.2 shows some examples of partial occlusion in surveillance videos [18].

Figure 2.2 Examples of partial occlusion in surveillance video

Generally, there are two main sequence steps in the human object detection process: the object detection step, and the object classification step, are as shown in the following Figure 2.3.

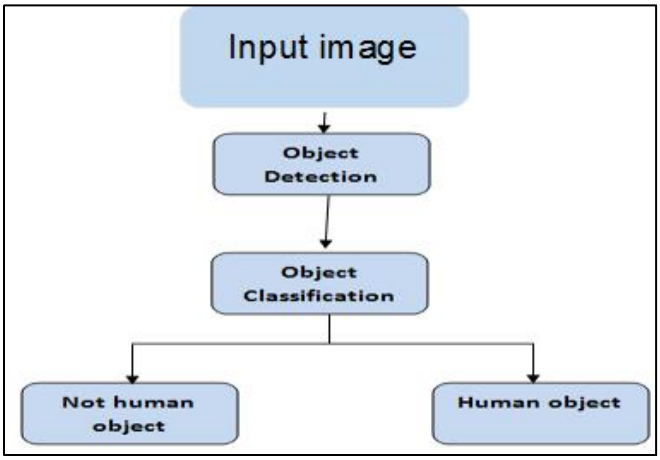


Figure 2.3 The two main sequence steps for human object detection process

In this chapter, a full description of the various phases of object detection and object classification has been presented, and the explanation of the most commonly available methods that apply these phases have been presented in detail. The explanation of these phases show and mention the advantages and the limitations of these methods, furthermore highlighting the applicability of these methods in real life applications and systems.

2.2 OBJECT DETECTION

The object detection task is considered one of the main and important steps within the process of extracting information from videos and images, because of the great importance in obtaining information (useful elements) from video and images and elimination of unwanted and non-important elements. The object detection task aims to find the location (region) of the interested object in images in order to localise it post process. The localisation process of the interested object can be done by grouping the pixels of the interested object (wanted object) and then clustering the object pixels together. The object detection process is defined as a sequence of processes to determine the objects of interest in the video sequence or images by grouping the pixels of the elements together as a block. This process can be executed through several methods, such as frame differencing, optical flow and background subtraction [19].

.

2.3 OBJECT DETECTION METHODS

There are many methods being used in object detection, each method has a different technique to extract the useful information from the object of interest as shown in Figure 2.4.

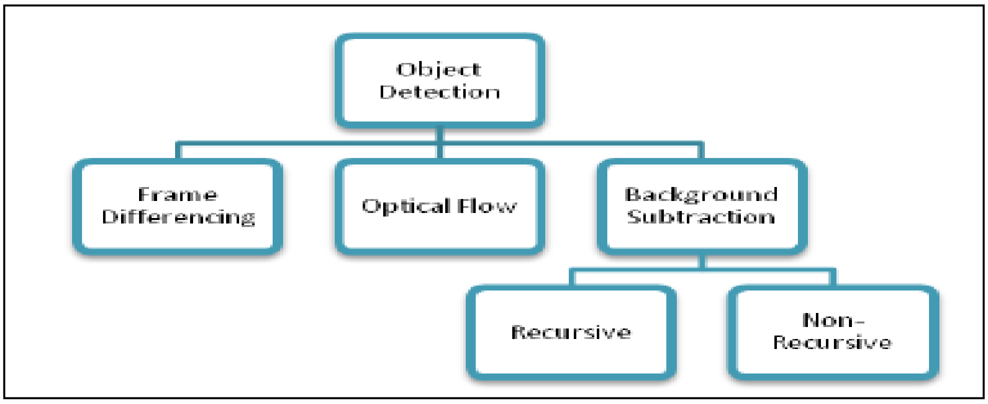


Figure 2.4 The several methods for object detection

The aim of these different object detection methods is to a locate the region of the interested object, therefore each one of these methods can achieve a high performance of object detection, compared with others, based on how to deal with different characteristics of the interested object, and the characteristics of the environment such as the background features (static background or not static background) and other visually noticeable parameters including illumination, cluttered background, and object occlusion. An explanation of the common methods being used in object detection is given below. [19]

2.3.1 Frame differencing

The frame differencing method is one of the useful object detection methods that can be used to detect the object of interest from the video stream. The process of this method is based on frame subtraction, because the position of moving objects within the video will change with the sequential time, which means that, the location of the moving object in one frame is different from the object location in the next consecutive frames. Detection of the moving object by this method is very simple and it is accrued by calculating the difference between two consecutive frames. By this calculation extracting and localising the moving object in video is very easy to implement spatially in case there is a static background. This method is efficient and adaptable to detect moving objects in a variety of dynamic environments, with very low computational complexity. In general, this method is difficult and low in accuracy to obtain and detect a complete outline of moving objects, this difficulty refers to the empty phenomenon which may appear and leads to miss detection of such objects. [20].

An examples of efficient object detection approaches that use this way of detection are presented below [92]. This approach is a robust and novel approach for detecting an object, in this approach the researcher presents a new prototype based on distortion template models. The deformable template model generated transforms the deformation parameters to its prototype, thus to derive an interpretation the presented prototype-based template combines both the local image cues and global structure information. Similarly, another approach presented by P. Felzenszwalb, et al [93] presented a new approach to detect objects, the process of this detection approach aims to use, the deformable part-based model. This model is based on two root filters and several part filters, as well as on using the deformable model which is used to weigh the configurations of the parts. Al-Najdawi et al [94] present a new approach to detect human objects and then adapt a tracking process to track humans in real time application. The process of this approach is based on extracting the continuous and non-continuous features of humans, using the Kanade-Lucas-Tomasi (KLT) technique, the use of this feature in the classification process.

[95] presently a novel approach for object detection in single videos is based on semiautomatic segmentation. In this method the classification problem is addressed by extracting the objects from the background of the video by dividing the single frame into a small uniform size of blocks and then based on manual segmentation for the first frame to use it as a training sample for classifying the object from the background. The implementation of this approach shows a high performance in detection of the interested objects however, the limitation of this approach is the inability to recognize and detect the boundary of the objects.

2.3.2 Optical Flow

The optical flow method is one of the object detection methods, and it is very efficient for detecting moving objects in videos. The process of this method is typically based on the motion of the object. This method detects the moving object by extracting useful information from the motion pattern, such as the surfaces and edges of object motion that are apparent. This motion pattern information can be used in this method to generate an image optical flow field, by calculating the point velocity of objects within one frame, which can be used to estimate the location of this point (object point) in consecutive frames. This estimation can be done by implementing such clustering process based on the distribution characteristics of the frame optical flow. This method has a wide use in several areas of computer vision, and it's efficient in extracting and detecting moving objects at high accuracy, however, calculating the point velocity of objects requires a high computational cost which and the performance of this method is very sensitive to noise, these reasons make this method unsuitable for the applications that can be affected by noise or are not required in a computational complexity [21].

One of the common approaches that are used for object detection, called (CAPOA) is presented in [62]. In this approach the objects can be detected even in cases where the object is partially occluded by adapting object occlusion analysis processes, this process can analyse the occlusion situation in the region of interest and then present a template that can be used as a mask to find the similarity matching regions corresponding to mask templates however, in another way this approach has a difficulty in the detection of a reappearing target.

Mirmehdi. et al [88] presented an efficient approach to detecting and recognizing an object. The main idea of this approach is based on the feedback control strategies used to improve the established single pass hypothesis generation and verification approaches. The improvement of the single pass hypothesis generation and verification approaches were acquired by extracting object generic class for the interested object instances, the extracting process of object instances had been done to recognise the object and to reduce the number of hypotheses by adapting a low-level optimal set of features. Similarly, in this approach to avoid any missing recognition, the feedback control had been adapted for a top-down recognition search.

Zhang et al. [89] present a new object presentation approach, the main idea of this object presentation approach is to locate the interested object by using a multi-block that can specify the binary pattern of the object in the image. The results of this approach can encode rectangular region intensities provided by LBP. Gupta et al. [90] present a robust approach to detect objects, the detection process in this approach had been completed, by employing a change detection method for analysing temporal information in the successive frames. The performance evaluation of this method shows a low complexity and computational loads with high accuracy.

Bar-Hillel et al [91] present a suggestion for a new approach to detect objects and classify the detected object as a human using a learning technique. In this approach the detection process is conducted by using feature synthesis. The suggested method is useful to improve pedestrian recognition in automotive applications and real time monitoring systems.

2.3.3 Background subtraction

The background subtraction method is one of the most common object detection methods, because it is very useful to detect the object of interest in both static images and videos. The main idea of this method is based on the possibility of separating the foreground from the background of the image or the frame. Foreground typically contains the useful information such as the objects of interest (wanted objects), to extract and localise these objects in an image or video farm, a background modelling process has to be done to subtract the foreground from the background. Through this process the unwanted data, such as the background data can be easily removed from the image or video frame, to leave just the necessary and useful data which is located in the foreground. The extraction of the object can then be done by finding the difference between the current image and the subtracted background. Thus, background modelling has to be performed to generate a reference model before background subtraction. Modelling must be sensitive to yield a reference model in an efficient way to help for object recognising. In this method the image has to be compared with the reference model, and in case of videos, each frame of the video sequence has to be compared with the reference model in order to determine the possible difference between the current video frame and the next consecutive frames. A suitable filter such as mean and median filters can be used to understand the background modelling and localise and detect the objects in image or videos. [22]

The process of this method is very simple, and it is at the same time very efficient in detecting objects of interest in a static image or video, especially in cases where the background is easy to extract and recognise, which leads to higher accuracy, the disadvantages of this method are the sensitivity to the dynamic changes in the external environment and the ability for anti- interference. There are mainly two approaches that can be used to perform the process of detecting objects using the background subtraction method, these two approaches aim to separate the foreground objects from the background of the image, and then localise the regions of the interested objects. These two approaches are the recursive approach and non-recursive approach, a full description of these two approaches and the entire process of each approach can be presented as follows:

2.3.3.1 Recursive approach:

Recursive approach is one of the background subtraction approaches; the core idea of this approach is based on updating the background model for each frame recursively, instead of maintaining a buffer for background estimation. In this approach the past input frames may cause effects on the current background model, therefore in case there are any errors accrued in the background model they can linger for a much longer period of time. The main advantage of this approach is that it requires less storage compared with others, and it includes different filters such as Gaussian of mixture, approximate median, and adaptive background. [23] [24].

2.3.3.2 Non-recursive approach

A non-recursive approach is one of the background subtraction approaches, the process of this approach is based on sliding-window for background estimation, which requires a large size of storage to store a buffer of window for the previous video frames, this buffer used is used for finding and storing the temporal variation of each pixel within the window-slide, in order to estimate and extract the background and then subtract it from the image or video frame to obtain the foreground which contains the objects. Non-recursive approach is a very efficient approach with strong ability to detect objects however, it's efficient if there is a need for a large buffer to dealing with slow-moving traffic [23] [24].

Many researchers were interested in this task in order to improve the process of access to information and useful elements of video and images and most of their interest was in trying to find a way to develop the use of one of these methods in real time applications, because the detection of such objects using this technique requires a less computational time compared with other techniques [28]. There are many researchers who use this approach to detect and classify objects. [62] A new approach called Content-Adaptive Progressive Occlusion Analysis (CAPOA)has been proposed, in this approach the occlusion situation in certain regions has been analysed and then a corresponding template mask has been created. However, the detection was difficult in some cases using this approach. Kirt Lillywhite et al [63] proposed a new approach to detect and recognise objects. The core idea of this approach is based on Evolution-Constructed (ECO) features. In this method no need for an expert human being to set up the features or tune the feature parameters because it's based on a basic image. Similarly, this method provides the ability to generate the features set for different types of objects under no limitation to image source types. These factors give this method advantages over others. G.L. Foresti et al.[64] present a new method, for detecting objects, the process of this method is based on the background subtraction technique, furthermore the shadow has been removed by this approach in subsequence phases.

C. Wohler et al [65] present a robust method for real time object classification. The process of this method aims to detect and classify the object of interest by completing the image sequences by time delay, instead of single image, in this method the neural network – time delay adapted to achieve high accuracy result. Papageorgiou and Poggio [66] propose a new approach for human object detection, the main idea of this approach is to represent the human object in the regions using the Haar wavelets, by choosing the 16- and 32-pixel scale with the 75% overlap. The performance of this method was efficient compared with others, even in case of low frequency changes contrast. [67] A new approach to detect objects aims to enhance the performance of using some of the previous approaches separately, such as background subtraction and temporal differencing by combining these two approaches together. The implementation result of this method shows an improvement in the object detection performance, compared with the implementation result and the detection performance for each one separately. Celik et.al [68] provide a new method for automatic dominant object detection in real time video sequences.

In [69] a new method proposed for moving object detection in video sequences based on frame differences which use the pixel wise differences between two sequential frames in videos. The advantage of this method is the ability to detect the moving object in dynamic environments however, it cannot extract all the relevant pixels, which may lead to some holes being left behind , to overcome this poor result the use of three-frame differencing is more suitable in many cases. Similarly, [70] a new method presented to detect objects aims to detect objects by using temporal difference methods in a low resolution of the video. Likewise, [11] the authors of this work present a new approach to detect a moving object in video sequences based on the frame difference method.

A [72] new approach proposed for detecting a salient object can be used, the process of this approach aims to extract object attentions to generate object template prototypes in order to classify and track the salient during the assessing object saliency in a video stream, using [73] a background subtraction method to detect and track vehicles in the traffic surveillance system. In this approach the Histogram-based filtering procedure had been used to present the reliable instances for the actual background at pixel level, by scattering background information and then carrying in the next sequential frames, this presented approach can deal with background instance under any traffic conditions. Levin et.al [74] developed a new supervised learning system for object detection based on using co-training, in this system two different classifiers have been used for the purpose of training each other, which leads to enhancements in the detection performance. Dirk Walther et al [75], proposed a multiple object recognition method for cluttered scene recognition. The process of this method aims to recognise the clutter using the bottom-up visual attention to technique. The evaluation of the experiments results compared with David Lowe's approach in recognising objects, shows that it has greater efficiency and a higher accuracy than David Lowe's method. Manuele Bicego et al [76], presented a new method for detecting and recognising 3D objects by using a Hidden Markov Model approach and the raster scan fashion being used to obtain overlapped sub-image objects. Furthermore, the Wavelet coefficients have been applied in HMM's to present the sequence vector model. The classification phase has been completed based on using the nearest neighbour rule. The implementation result of this method achieves a high accuracy in object recognition compared with others, even in the cases where the object is fully in occlusion or the appearance of objects is not clear. Bijan Shoushtarian et al, [77], present a new approach to detect objects such as humans in images. The process of this approach is based on using an efficient background subtraction method, in this approach three different methods for a dynamic background subtraction are used to achieve the efficiency of detection without missing object detection. The implementation result shows high accuracy in the performance of this method for detecting humans in images. Hui Chen et al, [78] proposed a new method for 3D object detection in Images. The process of this method combines two different approaches, which include the feature embedding approach and SVM to detect the objects. The implementation of this method shows a highly efficient performance in the detection of 3D objects, compared with the GH method. Carlos Cuevas et al [79], present a new method for moving object detection based on background modelling. The proposed work tends to detect the moving objects from a complex image by making a combination between the background model and foreground model, the implementation result shows high quality compared with others in detecting moving objects from images taken by non-static video cameras.

Ling Cai et al [80], proposed a new method to detect many objects. The main idea of this method is based on presenting a stereo vision model, this model overcomes many object detection issues such as illumination, shadows and the occlusion of objects.

Bangjun Lei et al [81] present a new method for detecting and tracking outdoor humans in real time application. The process of this method aims to achieve high accuracy detection based on presenting an efficient model to divide and separate the video frame to foreground and background, and then extract and detect the human object from the foreground, to be used further in the tracking process. Moctezuma et al. [82] presented a robust approach for human object identification and counting based on using the HOG and Gabor filter. As a result of evaluation of this method, it demonstrates a high-performance of accuracy in human identifying and counting compared with other methods. Anuj Mohan, et al [83] presented a new framework, to detect objects from a static image. The presented framework can be used to develop a different system with the purpose of identifying and counting human beings in cluttered scenes. Fortenberry Lei et al [84], proposed a new learning model for detecting objects against backgrounds based upon using image generation, derived optimal inference approaches are also boosting methods being used for learning. This model was very efficient when used to detect objects.

Christian Goerick et al [86], present a new approach for real time car detection and tracking. The detection and tracking process in this approach is based on using the power of artificial neural network, the advantage of this method is reducing the complexity and the design cost. Lee et al. [87] developed a new effective method to extract objects in images based on using a background subtraction method, the average implementation of this method shows a highly effective way to extract objects of interest in images with sensitivity to illumination changes.

A brief comparison between the most common methods being used for object detection is presented in Table 2.1.

Table 2-1 Comparing between the most common methods been used for object detection

Methods		Accuracy	Computational Time	Comments	
Background Subtraction	Gaussian Of Mixture	Moderate	Moderate	+	Low memory requirement
				-	It does not cope with multimodal background
	Approximate Median	Low to Moderate	Moderate	+	It does not require sub sampling of frames for creating an adequate background model.
				-	It computation requires a buffer with the recent pixel values
Optical Flow		Moderate	High	+	It can produce the complete movement information
				-	Require Large amount of calculation
Frame Differencing		High	Low to Moderate	+	Easiest Method. Perform well for static background.
				-	It requires a background without moving objects

2.4 OBJECT CLASSIFICATION

There are two main phases for multimedia extraction in computer vision applications, the first phase aims to localise and extract the object from an image or video which is called the detection phase, detection of the object singularly, does not provide useful information because the image and videos may contain several types of objects, and most of the computer vision applications implemented deal with a specific type of object based on its main purpose. The second phase is to identify the types of detected objects, and this is called the classification phase. The process of classification aims to identify the type of detected objects obtained from the detection phase, this identifying process can be done by extracting some features that correspond to one type of object only, which means finding unique and exclusive features in such type of objects, and modelling these exclusive futures to classify the interested object from other detected objects, since there are many different types of objects that most recent computer vision applications deal with such as human, ships, vehicles, dogs, and so on.

Object classification is a sequence of steps to identify the detected object in images or video as objects of interest, the classification process can be completed based on different parameters such as motion, colour, shape, and texture. Therefore, we can define and perform the classification method as motion-based classification, colour-based classification, shape-based classification, and texture-based classification. As shown in Figure 2.5.

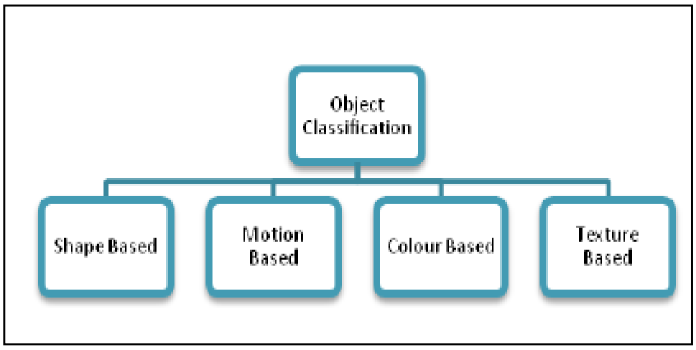


Figure 2.5 The various classification methods

2.5 OBJECT CLASSIFICATION METHODS

There is a verity of methods used to classify the detected objects. The process of these methods aims to identify the types of detected objects such as trees, human, vehicles, and varying other types. Typically, each classification method can identify the type of object by extracting some unique characteristics obtained from the detected objects called features. The kinds of features can be related to certain specifications of the wanted object, such as shape features, colour features, texture features, and motion features, which have all been used within the classification methods. [19]

2.5.1 Shape-based classification

The shape-based object detection approach is one of the detection approaches that aims to segment the object of interest in the image, and one of its complex problems is being able to detect and characterize the object, due to problems of shading and occlusion which may appear with many objects [29] [30]. There are many researchers who use this approach to detect and classify objects, for example, in [31] present a new method for tracking and detecting hockey players based on their location and size, by using a coordinate system to estimate and identify the hockey players in video sequence synchronization, in this approach the shape-based features and texture features have been used for classifying the hockey players.

Many researchers also have turned their attention to proposing new approaches for object detection based on the object's shape for example, Chih-Hsien Hsia et al [32] present a new object detection method for detecting the directional lifting based on discrete wavelet. The new methods (MDLDWT), detect the moving object using a shape-based approach which is an efficient approach when being used detect the moving objects in videos, and the implementation result of this method shows a high accuracy in multiple moving object detection and addressing the low-resolution configuration and object shape issues. S. Belongie et al [33], presented a new human object detection method, the process of this method is based on the shape of the interested object. The detection of this method aims to extract the object shape contexts, and then adapt the shape contexts in the classification process by matching the similarity between the object shape contexts in the next image. The similarity matching is acquired by localising the region points that contained the object and estimating the corresponding points in the next frame. Wang et al. [34] present a new approach for characterizing and recognizing the motions for the purpose of human object detection, this work aims to investigate the effect of motion in creating distortions in human shape, and then present these shape distortions as discriminate features to find the matching regions that have the same shape distortion in order to detect the motion of objects in the videos. Similarly, Liang Wang et.al [100] presented an efficient method for human object detection. The process of this method is based on a comparison between three categories; motion based, shape based, and component based. The performance of this work shows a periodic property based on a non-rigid articulated human body, thus it gives a very good cue for motion-based classification. Longbin Chen et.al [36] provide a new method, for human classification, the process of this method is based on a shape-based classifier, it works by extracting the skin segments and finding the computational parameters for some parts of the human body then presenting the shape using blob or silhouettes to classify the object as a human.

Zhe Lin et.al [37] present a robust approach for human object detection and pose estimation. The aim of this approach is based on using the shape-based approach to detect the human shape, by extracting the human shape feature, and then running the similarity matching process for human classification purpose. Similarly, this approach has a part-template tree being used to find the similarity between two matching images within hierarchy. Jorge Garcıa et.al [38] present a new technique for human object detection in images and videos, this technique represents a shape model for human heads, this model is used as a template by comparing the human head shape with any circular shaped pattern in images, using a robust classifier. Rusi Antonov Filipov et.al [39] present an efficient approach to detecting the human object in images. This approach is based on the detection of a human’s head in a variety of images obtained by a vertically oriented camera, by using the power of the shape-based approach to analysing 3D range data.

Li M., Zhang Z. et.al [40] present a new approach for fast human detecting and tracking. The main idea of this approach is based on using omega-shape features, to present the human shape by utilising, the omega-shape features of humans being, the performance evaluations of this method show high accuracy compared with other classifier methods such as HOG feature based. Huazhong Xu et.al [41] provide a new method for human detection, the process of this method is based on the shape classification, in this approach a bank of annular patterns have been used for representing a 2D correlation and then the SVM classifier is used to detect human objects by finding matches between people's head and the circular patterns. Mun Wai Lee et.al [42] present a new approach for estimating the human body pose. The estimation process in this approach, is based on a hierarchical technique, in this approach the detection process is based on, searching for components of interested objects. Furthermore, the proposed method can detect a different shape model, the implementation of this method shows an efficient result to detect humans based on their shape.

KON G Xiao-fang1. et.al [29] present a robust approach, for human detection and classification, the core idea of the process for this approach, is based on the head and shoulder contour edge, in this approach the updated background subtraction has been used to detect the foreground objects, also the shift- mean and the edge detection approach has been combined with this approach to detect the head and shoulder edges. The experimental result of this approach shows its efficiency for human recognition compared with other approaches. Xiaobai Liu et al. [43] present a new template to detect objects, by creating a hybrid template. This hybrid template extracts the necessary object features of different types of detection approaches and for each detection approach the template extracts the applicable features, such as texture features for face recognition, edge contour for human edge detection and the flatness regions for detection of objects based on part shapes or mask scaling. This method is efficient in detecting different types of objects in various detection approaches however, the experimental result shows that there is a limitation in detecting objects based on the changing acquired when the object moves. Therefore, it needs to be modified to search for the similarity between the old detection features and the new features after the object moved. In [44] a robust approach for video monitoring systems for detecting a moving object based on the gradient direction masking and the enhancement of edge localization mechanism, the implementation of this method achieved a high accuracy in detecting moving objects in video sequences compared with other methods. Similarly, another approach presented by Yuhua et al, in [45] for detecting objects in the video and then classifying the detected object as human. In this method a set of parameters has been extracted from the human object to build a classifier that can classify the detected object as human against other objects, the classifier parameters are generated by extracting human features from a set of sample data, and then the learning techniques are used to understand the classifier for true detection by feeding the classifier with both negative and positive samples. This method achieves a high accuracy for object classification. In [2] developed a new method for human body detection in images and videos, the process of this method is based on the texture of the objects by extracting object edge features, and then for the classification phase to classify objects as human from others, the SVM is applied to locate the human regions. This approach is highly a efficient approach for gradient orientation in localizing portions in images. William Robson Schwartz et.al [46] propose a new method for human object detection, the detection process of this method runs in an efficient way, based on combining both the edge-based features and the colour and texture information for the purpose of detecting human objects in images and videos. Similarly, Lowe et al [47], present a new approach for object recognition, the process of this approach is based on using the local scale-invariant features. This local scale feature is presented by using an efficient filtering approach to extract the object features and create index keys for the image. The index keys will then be used for matching the similarity in the next image to detect the wanted object. The implementation of this approach demonstrates good performance when detecting objects, as this approach adapts any changes in an image such as image rotation and scaling. Similarly, this approach can deal with the illumination that may be a cured in images.

Mohan et al. [49] presented a robust framework for human detection, for example-based detectors to recognize and localize some parts of humans, such as head, legs, and arms, in term of detecting human objects from images with high accuracy. Andriluka et al. [50] present a new single framework for human object detection, they also present a robust pictorial structure model. Krystian Mikolajczyk et.al [51] proposed a new approach, the process of this approach is based on, modelling human object parts as flexible assemblies for the detection of human objects in images and videos, the model representation presented by the advantage of using a co-occurrence of local features. Yi Yang et.al [52] present a simple yet efficient model for human object detection, and other types of objects, based on part model and using local mixtures of parts. In this model the object can be partially divided to extract the intensive part for the detection process, the articulation used in this model is to find the accuracy of changing appearances, furthermore the evaluation of the presented model shows a pose estimation in the criteria of human detection.

2.5.2 Motion-based classification

Motion based classification is one of the classification methods that can classify the object based on its motion. This method is very efficient in detecting moving objects rather than static objects. The core idea of this method is to identify the moving objects by finding periodicity of the motion. To address both rigid and non-rigid objects, the periodic property for non-rigid, articulated object motion, such as humans, show a higher average residual flow compared with the other rigid objects expected to present little residual flow. Thus analysing the rigidity and periodicity of the interested object to obtain the residual flow for that object is very useful for classifying an object based on the periodicity in motion. [15] [25] [117].

Detecting objects based on the motion of these objects is one of the approaches used to classify the object as a moving object or a static object. L. Han, M. Haleem, M. Taylor [85] propose an automatic detection and diagnosis and severity assessment of crop diseases using image pattern recognition. By developing a two-stage crop disease pattern recognition system which can automatically identify crop diseases and assess severity based on a combination of marker-controlled watershed segmentation, super pixel-based feature analysis and classification. This approach can accurately detect crop diseases and assess the disease severity with efficient processing speed.

Viola, Jones and Snow proposed new human detectors based on an Ad boost approach, this approach dealt with a large set of possible weak classifiers by selecting a small number of the large classifier sets and combining the selected weak classifiers to present an effective classifier that can detect objects in an efficient way. [48].

N. Murray, et.al [118] presents the results of the initial work that tested if focus of gaze could be more accurately gauged if eye movement was tracked, adding to that the head of an avatar observed in an immersive VE. The results of the experiment show that eye gaze is of vital importance to the subjects correctly identifying what a person is looking at in an immersive virtual environment. In [119] introduce EyeCVE, the world's first tele-presence system that allows people in different physical locations to not only see what each other are doing but follow each other's eyes, even when walking about. Projected into each space is avatar representations of remote participants, that reproduce not only the body, head and hand movements, but also those of the eyes. Spatial and temporal alignment of remote spaces allows the focus of gaze as well as activity and gesture to be used as a resource for non-verbal communication.

The changing in the sequence frame acquired by moving objects can be removed by extracting the changes between the video sequence frames. The efficient approach for finding the difference sequence frames is the temporal differences approach which has the ability to detect objects efficiency in a video stream environment, which is a very fast and dynamic environment, the limitation of this approach is the inability of a full detection and trace of the moving objects due to the empty phenomenon [92].

There are a variety of common techniques used in the motion detection approach, some of these techniques are described as follows:

2.5.2.1 Thresholding technique over interframe difference

This technique is one of the motion detection techniques, that can detect moving objects in videos. This technique is very efficient in detecting objects, based on finding the temporal changes that can be acquired in block or even in single pixel of the interested objects. The process of this technique, can be done by referencing the first detected frames, and then subtracting the next frame based on its reference, and applying a threshold value for more detection accuracy. In this type of approach, the objects cannot be detected in case of any changes in the sequential frames, so that it assumes the object must be continually moving.

Dhar et al**.** Propose a new method to utilize the detection of objects, the main idea of this method is to detect the object in images using a manual threshold selection. Foresti et al. [96] present a robust approach, with high accuracy for detecting human objects in monitoring scenes. This approach is based on the theory of the segmentation process. Elarbi-Boudihir et al[98] propose a new approach for object detection, this approach implemented a new monitoring system, to detect interested objects in low power by removing the unwanted video recording. Johnsen et al. [99] present a new method for background modelling based on approximated median filter by scaling the absolute differences between the current pixel and the median-modelled background pixel which is higher than a threshold. This method shows a better result when implemented.

2.5.2.2 Optical Flow technique

Optical flow is one of the motion detection techniques, that can detect moving objects in videos. This technique is a very common technique, that can detect objects in an efficient way. The main idea of this technique is based upon analysis and defines the interested region pixels, and then computes the direction and velocity for the region of interest. In order to detect objects by similarity matching the direction and velocity between the region of interest and the next video sequence frames, this technique is widely used in many tracking and surveillance systems, due to its high detection accuracy and ability to detect the moving objects in cases where the camera is not constant.

An example of using this type of technique is presented by Dalal et al. [2]. They developed a new technique for detecting if the human is moving., as much as the background and camera are moving in the scene. This technique is very useful for detecting pedestrians from moving cars as well as analysing the TV or film contents, the new technique is based on optical flow and background differencing combined with (HOG) which will calculate the gradient vectors in order to convert them to angles . The Optical flow technique is based on the process of clustering the image features, the experiment tests of this method shows the efficiency of detecting moving objects however, it has a limitation for real applications due to the complexity of its process and the huge calculation which is required for the detection process [28].

Augustin et al. [101] present a simple approach for moving object detection, and sequence tracking for the detected object, the performance result of this approach demonstrates its ability to detect the moving object with little delay, by extracting the changes in video frames. K. Hati et al. [102] provide a new approach for moving object detection, the core idea of this approach is based on using a temporal differencing approach, the experimented result of this approach shows a better performance in detecting moving objects as well as considering a fast detection method compared with other interim of low false detection. Antonakaki et al. [103] present an efficient approach to detect the objects in images, based on the use of the temporal differencing approach, in this approach modelling activities are based on using statistical activity recognition.

[104] The authors proposed a new technique incorporating the background model for increasing the accuracy of shadow detection in grey scale video sequences. Liu in [105] presents a new approach to detecting a moving object, the approach is based on using background subtraction methods, by comparing the differences pixel-by-pixel between the reference background image and the current frame. As a result, this approach is very sensitive to the changes in dynamic scenes as lighting and extraneous events etc. Collins et al., [106] in their project VSAM (Video Surveillance and Monitoring) provide a new hybrid technique for moving object detection by detecting the moving region based on the combination of the three-frame differencing and the adaptive background subtraction model. This method acquired a high success in the segment regions of moving objects in video sequences without the impurity of using a temporal differencing method or the background subtraction method separately.

Furthermore, in [107] the authors proposed a new technique based on background elimination technique and background registration technique. For moving object identification in a video clip. The implementation result of this technique shows that, the result can be affected in cases where the image contains a lot of noise. Bobick et al. [71] present a new method with the purpose of human recognition and detection. The process of this method is based on constructing vector image templates, by extracting the binary motion energy in the image, and then compressing it with the image motion history, as two temporal projection operators. TCutler et al. [108] present a new approach for human classification based on the main idea of this approach, which is to adapt a self-similarity-based time frequency technology. The implementation result of this approach shows that there is a restriction to periodic motions. Asif Ansari et al.[6] present a new robust approach for detecting a moving object. The main idea of this approach is based on motion detection and providing the monitor system with an audio alarm signal. This approach achieved high accuracy in detecting a moving object and it can be used in several monitoring systems for security purposes.

Liang Wang et.al [35] presented an efficient method for human object detection. The process of this method is based on a comparison between three categories motion based, shape based, and component based. The performance of this work shows a periodic property based on non-rigid articulated human body, thus it gives a very good cue for motion-based classification. Ko et al. [109] present a new approach for human detection, the process of this approach is based on background modelling. In this approach a new model is presented to subtract the foreground of the image, this model is unlike other models for the detection of human objects in images, the idea of this model works to find the differences between the object intensity variability of pixels in image location's background and the motion in the background. The experimental result of this approach exceeds the true detection rate.

Sebastien et al. [110] proposed a new approach which dealt with object learning by using colour information. In the new approach he develops the GHOSP (Genetic Hybrid Optimization & Search of Parameters) approach which contains objects to be learnt by using multidimensional observations taken from RGB colour images.

2.5.2.3 Gaussian mixture technique

Gaussian Mixture is another type of motion detection technique that can detect moving objects in videos, based on using Gaussian Mixture function. An example of real applications that use this type of technique is done by Bodor et al. As presented in [111]. They proposed an initial monitoring system to detect human objects in video sequences based on a mixture of Gaussians background segmentation in high performances in cases where the brightness of the light is constant. This monitoring system has a limitation when detecting and tracking the moving object in an environment with fast light changes. Another new method proposed to address the adaptive background penalty and the occlusion reasoning is separating the foreground and background regions, in order to detect objects in videos using frame differencing methods [112].

In [113] a new background model proposed for detecting the interested object, this model is based on using the Gaussian mixture model to enhance obtaining objects from images. At [114] present a new technique for pedestrian detection as a moving object in a scene, the process of this technique is based on estimating the background using the median function. And a combination of another two-pass approaches being used in this technique for classifying pedestrians and noise removal.

Tao Zhao et al [115] proposed a new approach for segmenting and tracking a human object, this model approach is based on presenting a new model, that can deal with segmenting the human object if partially occluded. Nicoletta Noceti et al [116], present a new on-line 3D approach for human object detection in a video sequence, this approach is based on Spatio-temporal constraints. In this method the local scale-invariant features being used for object modelling and for the recognition phase, the Ad hoc matching procedure is used. The performance of this method was very good for 3D object recognition in video sequences.

2.5.3 Colour-based classification

Colour based classification is a method that can classify the object based on the object colour, the process of this method aims to extract unique colour features for the detected object in images or video frames, these colour features are represented by using RGB colour space to present a colour histogram that can well describe the colour distribution in the image in order to segment the image into the background and foreground. Although the colour is not an appropriate classification technique for classifying such objects, it has found a usable way to classify detected objects because the colour has constant viewpoint changes and the classification process using this method is very simple and easy to acquire with low computational cost. There are two main ways to represent the colour features for classification purposes, the illumination spectral power distribution and the object's surface reflectance property, these two ways of representing the colour features can be used to classify the object by presenting a colour histogram to describe the colour distribution in the image, and then extract the foreground which contains objects by segmenting the image into the background and foreground. Gaussian Mixture Model can be used for presenting a colour histogram and the occlusion buffer for object occlusion, because the RGB is not a uniform colour space it is possible to use the "HSV (Hue, Saturation, and Value)" as a uniform colour space. [26]

There are many researchers who use this approach to detect and classify objects, for example, Sébastien et al. [53] proposed a detection approach called GHOSP, this approach is based on image colours. The process of this approach used object learning techniques based on the colour information to observe a multidimensional colour histogram for the image RGB colours, by this observation the learning techniques can learn from the extracted object colour features for a post detection process with similarities matching. Walk et al. [54] present a new method for human object detection, the efficiency of detection in this method is achieved by using a colour self-similarity feature for human object pattern, the colour feature captures pair wise information about spatial colour distributions, and it is considered a useful complement to the HOG feature. Yanjiang Wang et.al [55] presented a innovative method to detect human objects in coloured images based on making a detection for human skin like pixels and allocating the human face as a region, this method shows an advantage in human detection even under complex backgrounds.

Another example in, [56] proposed a new approach to detect objects, the process of this approach is based on generating object colour histograms, and then presenting a confident distance map, to find the similarity matches between the current frame and the next sequence frame in video. The mean shift technique is adaptive to segmenting the current objects position and localising the peak points of a confidence map for the object position, then localize and smooth the map boundary that contains the object using manual refinement in order to detect the interested objects. Saravana Kumarin [57] propose a method for representing the objects by using the HSV colour space. The main idea of this method is to cluster the object based on centroid colour values, this clustering process had been done using k-means clustering technique. The result of this clustering process aims to obtain the co-ordinate values that present the clustered object in order to represent the interested objects. In [58] Present a new approach called JSEG for images and video object detection, by using unsupervised segmentation based on colour-texture regions to extract the wanted elements from images and sequence videos.

Christian Wohler [59] proposed a new approach for detecting and recognizing the pedestrian, this approach aims to detect and recognize the pedestrians in real-time monitoring applications. The process of this approach to classify and recognize the pedestrian is based on generating a new model for the classification purpose. This model is generated by applying supervised training approaches for both dark and bright pedestrians wearing clothes. The test shows a better recognition capability. Shao-Yi Chien et al. [60] present a new approach to detect humans, this approach used the threshold decision approach with multi background modelling to segment the human object in a video sequence, he then used a composition between the diffusion distance that measure the similarity of object colour histogram and the segmentation of object motion clue, to present a particle filter with a likelihood function which was used to generate such a model for a human object tracking framework. Rana Farah et al. [61]. Proposes a new efficient approach for tracking an object and extracting the rodent from video frames under uncontrolled normal laboratory conditions, in this method a combination of three weak features for roughly tracking the target and then, adjusting the boundaries to extract the rodent.

2.5.4 Texture-based classification

Texture based classification is one of the classification methods that can classify the object based on the surface structures and their relationship to the surrounding environment. The process of this method aims to compute the dense grid of regularly spaced cells by counting the occurrences of gradient orientation in localized portions of an image, which lead to useful information that may be used for classifying an object [27] [1].

To make a clear comparison between the four of them, they are represented in Table 2.2.

Table 2-2 Comparing between the most common methods been used for object classification

Methods	Accuracy	Computational Time	Comments
Shape-Based	Moderate	Low	Simple pattern-matching approach can be applied with appropriate templates. It does not work well in dynamic situations and is unable to determine internal movements well.
Motion-based	Moderate	High	Does not require predefined pattern templates but struggles to identify a non-moving human.
Texture-based	High	High	Provides improved quality with the expense of additional computation time.
Color-based	High	High	It creates a Gaussian Mixture Model to describe the color distribution within the sequence of images and to segment the image into background and objects

In Table 2.3 is an analysis of the most common approaches used for object classification.

Table 2-3 Analysis of object classification approaches

Approach	Methods/Approach Used	Authors	Single(S)/ Multiple(M)
Shape based (Feature based)	1. PCA-HOG descriptor, hybrid Hidden Markov Model 2.Gaussian filters, Geodesic Active Contour model 3. Bootstrapping approach, shape-based active contour	1.Wei-Lwun Lu et al. [28] 2.Huabo Sun et al. [110] 3. John et al. [111]	S S S
Color based (Feature based)	1.HMM model, GHOSP (Genetic Hybrid Optimization & Search of Parameters) approach 2. color histogram (HC) bins and gradient orientation histogram (HOG) bins, kalman filter 3.Scale Invariant Feature Transform (SIFT), Kalman filter 4. Particle filters and multi-mode anisotropic mean shift. 5. K-Means clustering,HSV color space, Connected component analysis	1.Sebastien et al. [29] 2. Zhenjun et al. [113] 3. Junda Zhu et al. [45] 4. Zulfiqar et al. [114] 5.Saravanakumar et al. [46]	S S S M S
Fixed Template based	1. Content-adaptive progressive occlusion analysis (CAPOA) approach, Local best match authentication (LBMA) approach, Kalman filter	1. Jiyan et al. [30]	S

2.6 OBJECT REPRESENTATION

There are various representations of object shapes and appearances which are commonly used to represent the interested objects such as" (a) Centroid, (b) multiple points, (c) rectangular patch, (d) elliptical patch, (e) part-based multiple patches, (f) object skeleton, (g) control points on the object contour (h) complete object contour, (i) object silhouette" [40]. As shown in Figure 2.6.

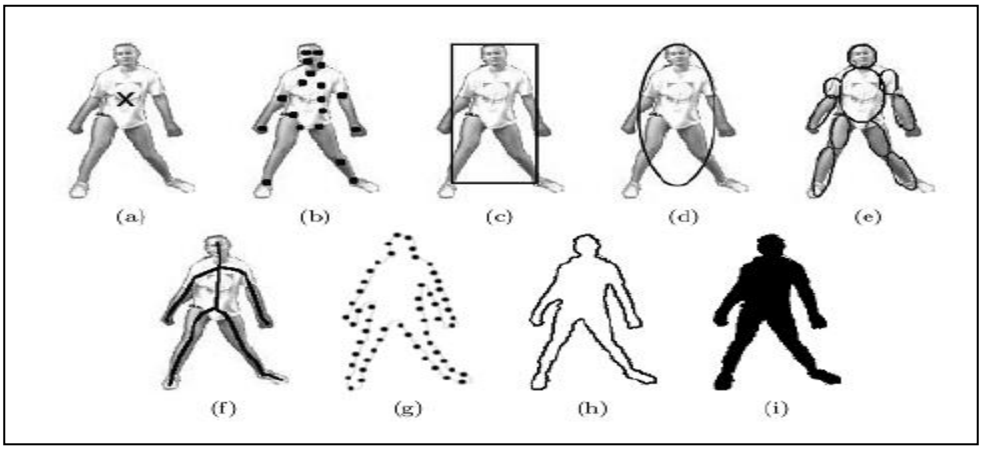


Figure 2.6The various representations of objects shapes

CHAPTER THREE

3 DESIGN AND IMPLEMENTATION

3.1 INTRODUCTION

Human object detection plays an important role due to its reliability detecting human objects in intensive applications and it's increase as a complicated task as humans can have different appearances and they can adopt a variety of poses. This opens new avenues for computer vision researchers facing these challenges by presenting and generating new algorithms for efficient and fast human detection systems, that can easily detect human objects in images and videos.

The target of detecting and recognizing humans within images is very useful due to the variety of applications and systems that require this process of detection. The process of human detection is to extract and localise all human objects in images, this process of detection requires a presence of certain features of the human object, these features must be special features that can specify a human object from another within an image. Generally, in order to detect humans robustly, we need to run computer vision approaches. To make it clear to understand, when trying to locate and count the number of humans in an image based on a computer vision approach, the initial task of the vision approach is to solve this task and start to identify and detect the human objects in the scene, then the counting process will become straightforward. This is one simple example for the needs of human detection, however, generally, the task of human detection in images or video sequences play an important role in security, law enforcement and military applications and therefore, it is an important computer vision problem and a big challenge for researchers due to more and more surveillance cameras being deployed in facilities or areas. The huge amount of multimedia content and the demand for automatic methods for multimedia management processing is increasing [5]. Recently there is a fast development in images and videos capturing device technology and it has become more easily available and cheaper, this development in video and image capturing devices have made the integrity of multimedia content in different application systems very easy.

3.2 THE NEW APPROACH

Human object detection approaches are a very challenging and complicated task as compared to other objects due to certain factors (i.e. Varying camera positions, dynamically changing background, and sudden). The major criterion for human detection processes is the knowledge of peoples features among different shapes. It is therefore, important to extract the correct feature from the given image or video sequence. Several approaches can be employed to detect human objects, based on experiment tests, some of these approaches processes take a long and significant time, which is not suitable for real time applications in visual surveillance systems.

The recent revelation that computer vision and image processing are using machine learning to detect the object in the images and then classify the objects into two groups: human and non-human [18]. Artificial intelligence, support vector machines, random forest, and artificial neural network are all approaches used for machine learning, these approaches allow the image processing and the computer vision to learn from the dataset to find a statistical relationship (i.e. Statistical regularities). The statistical relationship is used to detect the objects and then classify the detected object as human or non-human by means of previous experience. In other words, they are capable of learning from experience to reduce the human efforts for detection and classification processes [93].

As mentioned before, the human object detection and classification algorithms are facing some challenges and difficulties in detecting and classifying the detected objects. In certain situations, a geometrical approach is proposed that can imitate intelligent human behaviours and machine learning algorithms for object detection and classification. The main idea of the proposed approach is using the objects upper portion (shape features) to extract some geometrical parameters to classify the detected object as human or non-human. The upper portion is very important in the proposed approach because it is always visible and not so easy to disappear. Similarly, the proposed approach can reduce the challenges and difficulties in detecting and classifying the detected objects.

In this new approach, a new classifier has been designed based on a mathematical model which has been generated and built on extracting a unique and discriminating feature from the human upper portion shape. The classifier of this approach is used as a filter to classify the detected object as human or not human based on its' constrictors. There are two main steps in this new approach used to detect human objects in images and video sequences, these two steps are:

1. Detecting the objects from the images;
2. Classifying the objects as human or non-human based on the new approach.

The flow diagram of our proposed shape-based human detection approach can be seen in Figure 3.1.

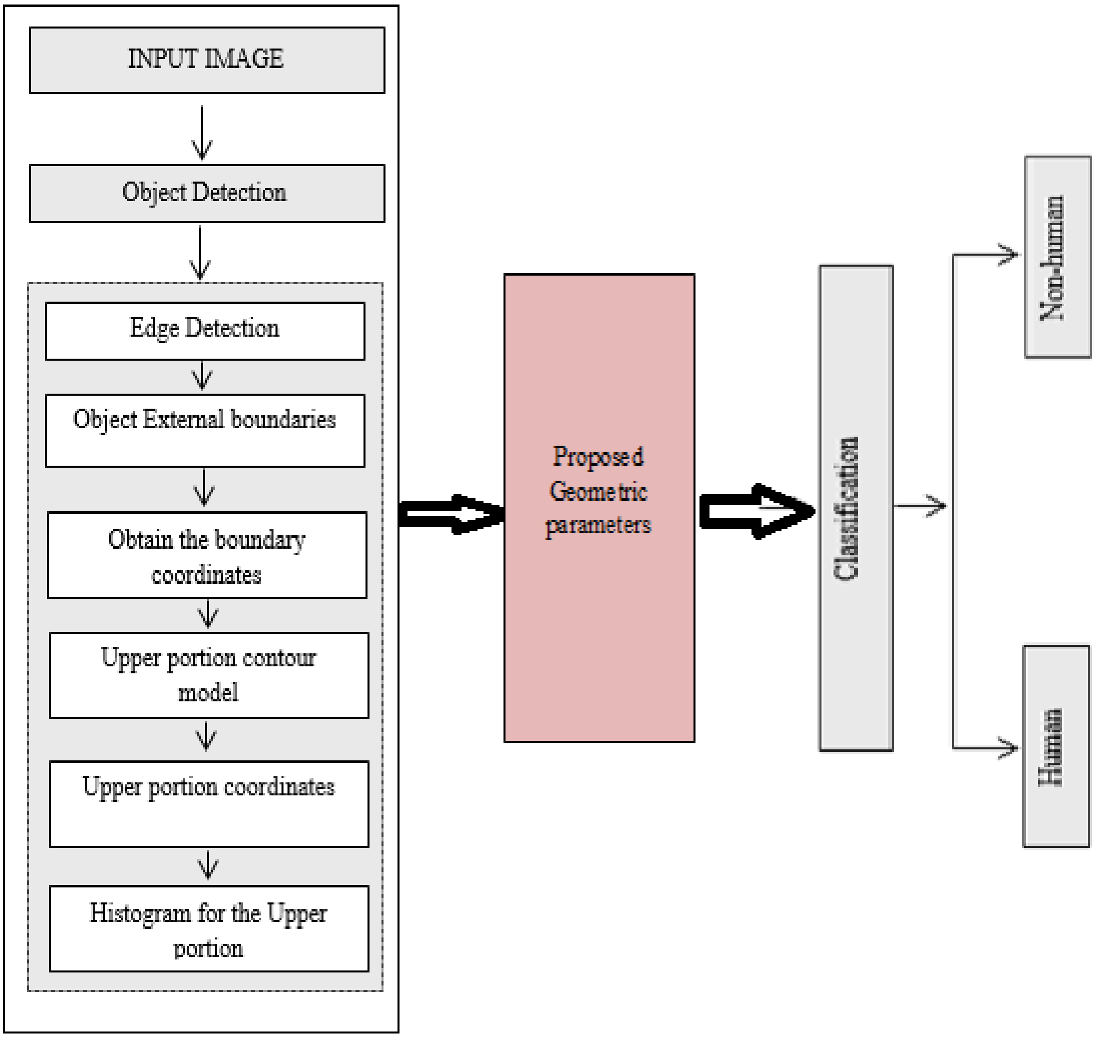


Figure 3.1 The Flow diagram of the proposed approach architecture

To make it clear, the proposed shape-based human detection approach is based on a set of parallel and sequential steps as shown in Table 3.1.

Table 3-1 Set of parallel and sequential steps for the proposed shape-based human detection approach.

Steps of the proposed approach:	
Step 1:	Background subtraction using histogram-based techniques with global threshold).
Step 2:	Object edge detection using Canny edge detection approach.
Step 3:	Extract the boundary edge using the boundary function.
Step 4:	Obtain [X, Y] coordinates for the boundary contour.
Step 5:	Extract the upper portion of the contour (i.e. Selected object).
Step 6:	Obtain [X, Y] coordinates for the upper portion of the contour.
Step 7:	Represent the obtained X values coordinate in a histogram.
Step 8:	Smooth the histogram using mathematical smooth functions.
Step 9:	Find the upper and lower peaks.
Step 10:	Perform the geometric model parameters.
Step 11:	Classify the object (i.e. Human or non-human).

In the following sub-section, discussed are the phases of the new approach in detail.

3.2.1 Object detection

As mentioned earlier, object detection task is one of the main and important steps in the process of extracting information from videos and images and it is considered the first step in this process because of the great importance in obtaining information and useful elements of video and images and the elimination of unwanted and non-important elements. This process is defined as a sequence to determine objects of interest in videos or images by grouping the pixels of the elements and interested groups of that block element together.

The first step of the object detection process is to determine the object from the images which is known as a region of interest (ROI) [123]. The object edges and boundaries refer to the region of interest [98], [123]. Images are different from video sequences because in the image is just a single image, but in the case of videos we have a sequence of frames, which contained a lot of single images, therefore, extracting objects from video requires a special approach such as temporal differencing, which is different from extracting objects from a single image. In general, there are a lot of techniques for object detection such as:

edge, random MARKOV, histogram, hybrid, and region and many other techniques [43]. For object detection and extraction from the images in this piece if work, histogram techniques were performed with a global threshold.

3.2.1.1 Histogram-based techniques

A digital image is a collection of small elements called pixels. Each of these elements has a value or set of value coding for the density level in each position. A digital image can be obtained using a large number of different devices such as a digital-camera, MRI machines or any type of device that can capture light intensity.

In the context of image processing, the image histogram usually refers to the pixel density graph. Which shows the number of pixels in an image at different intensities within that image. For example, in the 8-bit grayscale image there are 256 different possible densities therefore, the histogram will display 256 numbers showing the pixel distribution between those grey values. This process requires a conversion of the coloured image to greyscale firstly, and then an appropriate threshold must be determined before being used when converting a grey image to a binary image. If the image is suitable for the threshold, the repeating graphic consists of two colours. This means that the pixel density is two separate values. The histogram process scans the image in one path and the number of pixels that are found at each intensity value is retained. This to be used to create an appropriate graph.

There are two types of histograms, Intensity histogram and Colour Channel Histogram. In intensity histograms, the image needs to be converted to the greyscale firstly, and then presented to the histogram for the grayscale level of that image. In the colour channel histogram, there is no need to convert the image to greyscale level, it displays the histogram based on the different RGB colours as shown in Figure 3.2. Figure 3.3 shows an example of an image Intensity histogram.

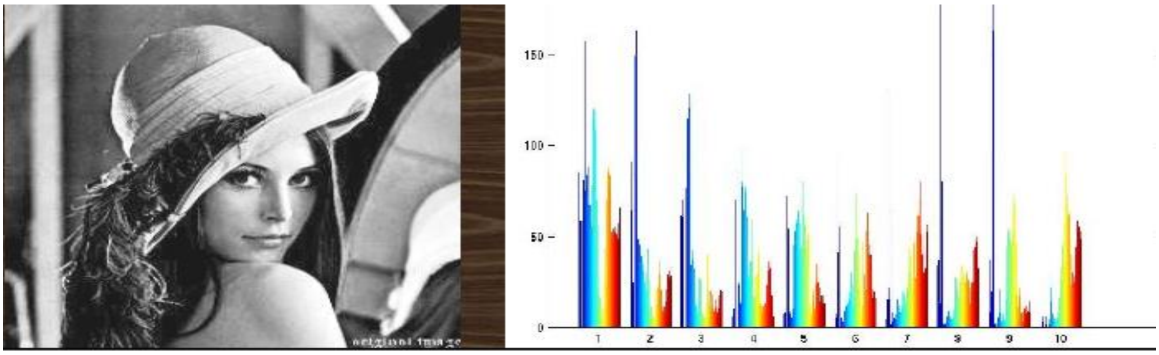


Figure 3.2 An example of image Colour Channel Histogram

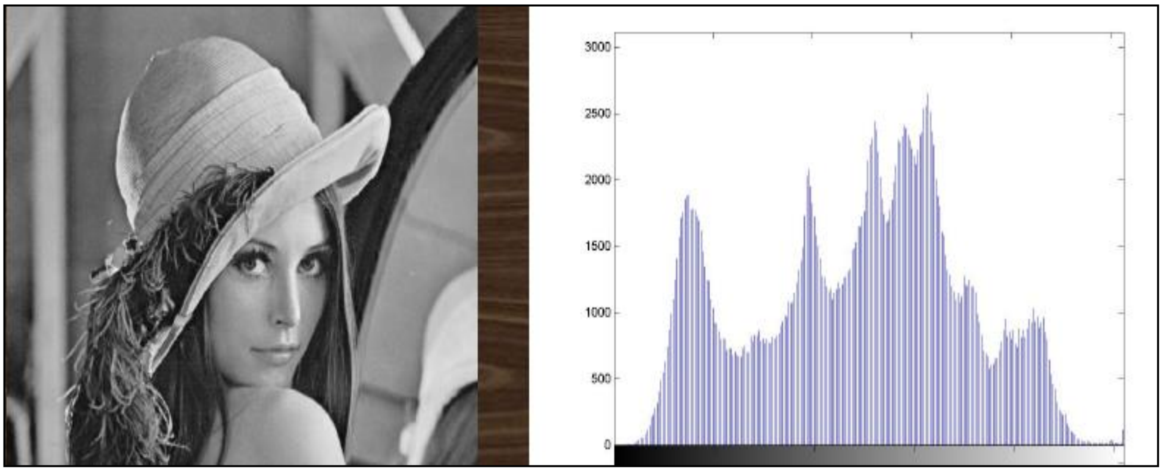


Figure 3.3 An example of image Intensity histogram.

Distinguished objects from images depend on the distinct intensity values in the grayscale of that image, this distinct intensity values cause the contrast of images. However, the image has a high contrast, making the object easier to distinguish than the images which have low contrast. The contrast can be calculated using the below Equation:

$$C_M(I) = \frac{\max(I) - \min(I)}{\max(I) + \min(I)} \tag{3.1}$$

Where: ***I*** is the luminance. Image histogram are very important in image processing due to the wide applications that it can be used for, such as image analysis, brightness and contrast image classification, image equalization, and histogram thresholding. The greyscale image is a set of small elements called pixels, each one of these pixels stores one value which is the value of its intensity.

The number of potential levels (intensity values) depends on the digital type that symbolizes the image. For example, the possible intensity for an image encoded with 8 bits equal 256 (28), representing a range from 0 -255. The histogram can be presented by plotting pr(rk) on the greyscale level by Equation 3.2 [124]:

$$p_r(r_k) = \frac{n_k}{N} \quad 0 \leq r_k \leq 1,\ k = 0,1,2 \ldots L-1 \tag{3.2}$$

Where: r_k = the intensity value, L= Number of grey levels, n_k = Number of pixels with grey level r_k and ***N*** = Total number of pixels. In this approach an image process has been conducted to obtain the greyscale image level, binary images and the histogram of greyscale images in order to extract the object and detect the contour edge of the detected objects, as shown in Figures 3.4 and 3.5.

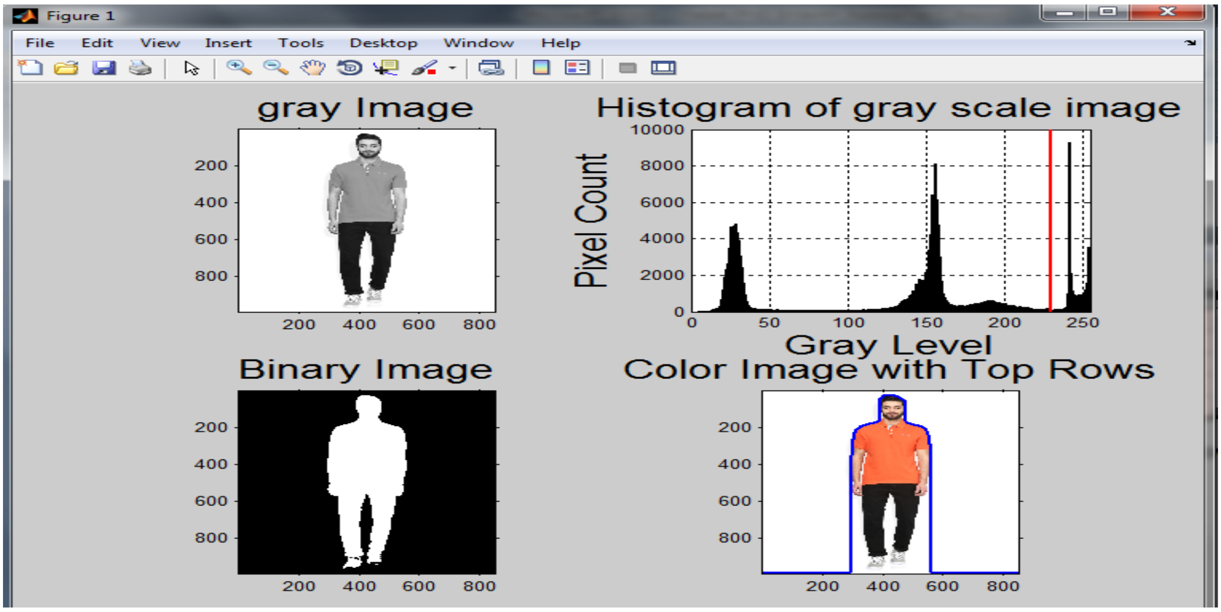


Figure 3.4 Obtaining the grey level, binary image and the histogram of grey scale image in order to detect the contour edge of the human detected objects

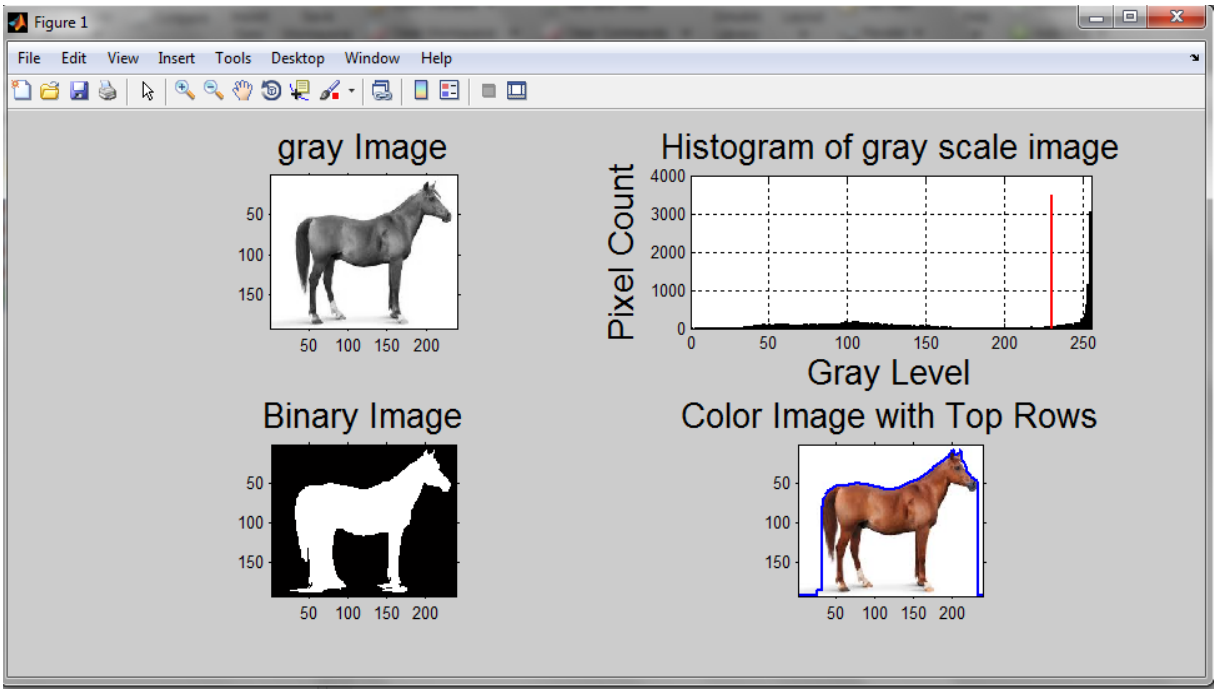


Figure 3.5 Obtaining the grey level, binary image and the histogram of grey scale image in order to detect the contour edge of the non-human detected objects.

3.2.2 Object extraction

Object extraction is the process of detecting and determining the wanted and interested elements in images by grouping the pixels of the interested elements. After detecting the object in the previous image processing steps, a thresholding function performed with histogram-based techniques is used to subtract images and determine the region of interest (Object) from that image as shown in Figure 3.6.

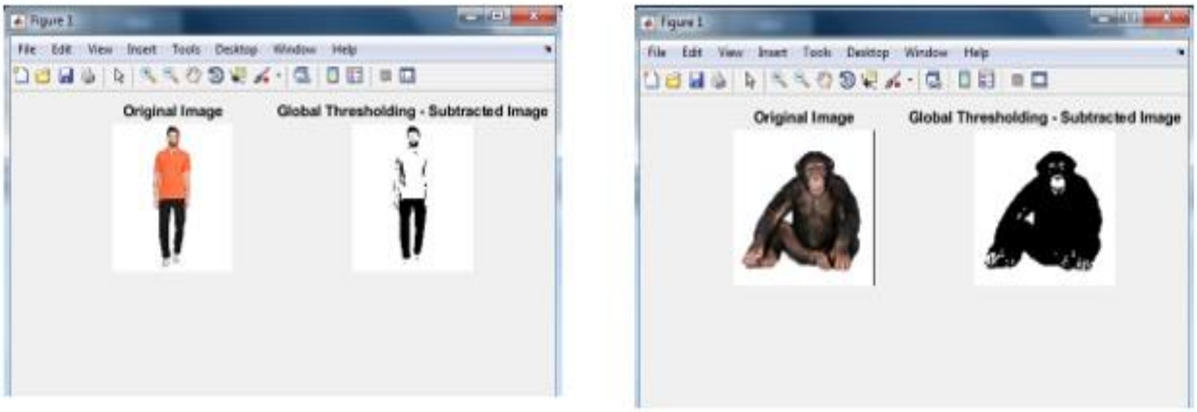


Figure 3.6 Object detection based on histogram techniques with global threshold

3.2.3 Edge detection

Edge detection is a computer vision and image processing technique for locating object edges or boundaries within images. The boundaries/edges can be found in the image by sharp changes in intensity (i.e. brightness) [95]. There are several approaches for edge detection. [45]. These approaches are shown in Figure 3.7.

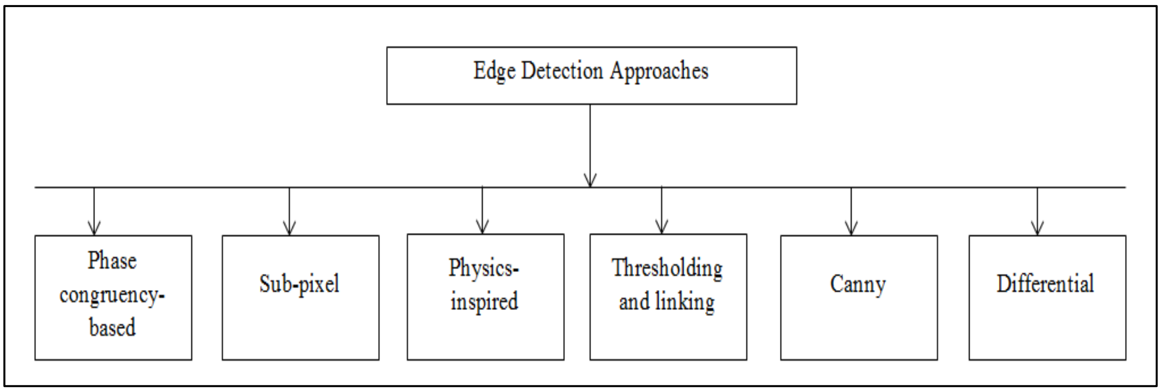


Figure 3.7 Edge detection approaches

In the proposed approach, to satisfy the object edge detection requirements in terms of good localization, good detection, and minimal response one of the edge detection approaches such as the canny edge detection was performed. Canny edge detection is one of the edge detections approaches and it is defined as a set of mathematical operations used to identify points in images by determining the change of brightness and sharpness in an image, the obtained points segmented as a set of curved lines called edges.

Canny edge detection function aims to detect the edges at a low error rate, which means that the discovery should retain as much resolution as possible for the edges of the image. The location of the obtained edge point must be located on the centre of the edge, and where possible avoid the image noise to create false edges. There are five steps in Canny edge detection function process: [125]

1. Smoothing the image to remove noise by applying a Gaussian filter.

Most edge detection results are easily affected by noise and can create a false edge detection. Gaussian filter is one of the most efficient methods that can be used to filter out noise, and it can be performed by the following formulas:

$$g(m,n) = G_\sigma(m,n) * f(m,n) \tag{3.3}$$

Where,

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{m^2+n^2}{2\sigma^2}\right) \tag{3.4}$$

2. Compute image intensity gradients:

   The edge gradient can be performed by Equation 3.5:

$$M(n,n) = \sqrt{g_m^2(m,n) + g_n^2(m,n)} \tag{3.5}$$

And,

$$\theta(m,n) = tan^{-1}[g_n(m,n)/g_m(m,n)] \tag{3.6}$$

Where, $\mathbf{g_{(m)}}$ the horizontal direction, $\mathbf{g_{(n)}}$ the vertical direction. Applying a thresholding to the edge gradient(**M**), to suppress the noise and keep the element of the detected edge (**T**) by Equation 3.7:

$$M_T(m,n) = \begin{cases} M(m,n) & \text{if } M(m,n) > T \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

3. Non-maximum pixels in the edge's suppression

   This step aims to thin the edge ridges in $\mathbf{M_T}$, by comparing the non-zero M_T (m, n) value with its two neighbours' values along the gradient direction. In cases where the M_T (m, n) is not the greatest value, then the M_T (m, n) is set to zero, or it keeps the value of M_T(m,n) without any change.

4. Result threshold:

   In this step a threshold is applied to further suppress and filter the noise and gap in the result, by keeping the edges that have a high gradient and filter out the weak edges that have low gradient values.

5. Segment and linking the edges:

   This step aims to bridge the gaps between the edge points to get a continuous edge by applying blob analysis.

Figure 3.8 shows some experimental result after performing the Canny edge detection approach, it is clear from Figure 3.8, that the object boundaries/edges have been detected successfully from the images.

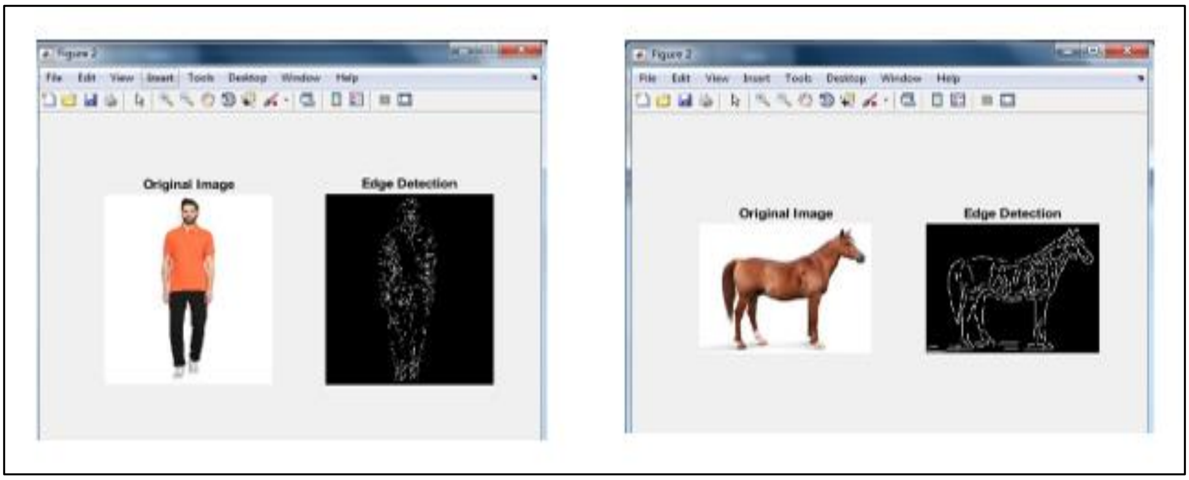
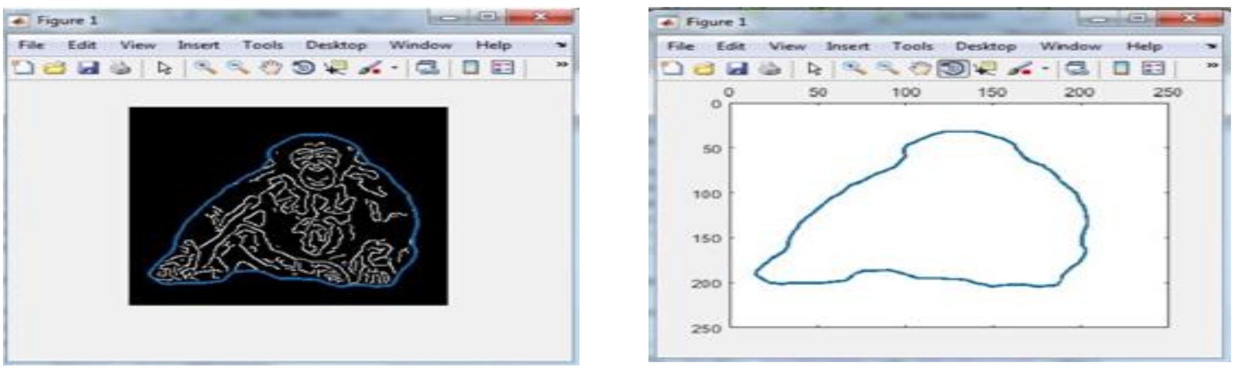


Figure 3.8 Some experiments result after performing canny edge detection approach

3.2.3.1 Boundary edge extraction

After object boundary edges have been detected and obtained from the images (external and internal boundaries), the external boundaries of the detected objects are extracted. In the proposed approach, a column vector of point was used (boundary functions) which returns a vector of point (x, y) as a 2-D boundary around the point, then the exterior boundaries are obtained (x (k), y (k)) as shown in Figure 3.9.



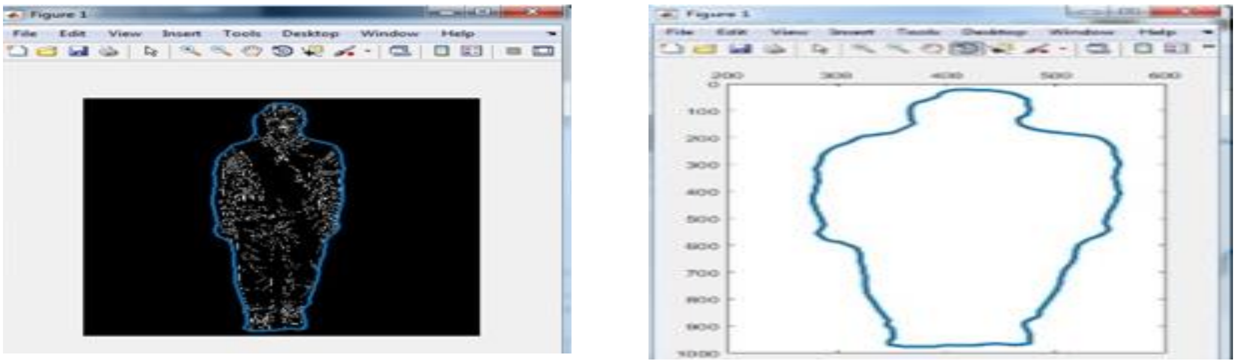


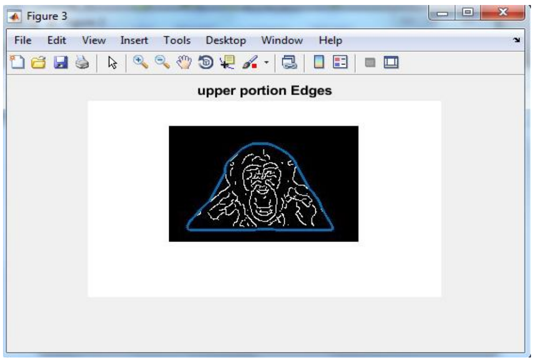
Figure 3.9 Extract boundaries of the detected object

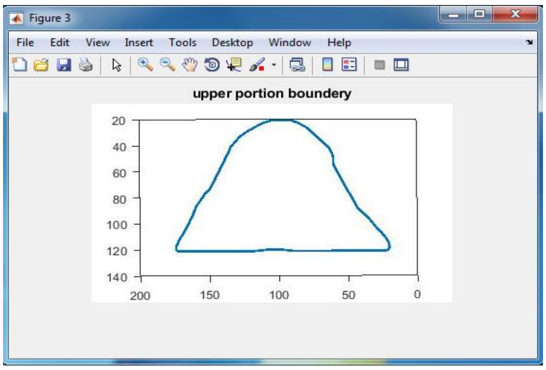
3.2.3.2 Extract the object upper portion

After performing the boundary function to extract the object boundary, the processes of extracting the upper portion of the contour takes place. In this approach, the focus of the study is on the upper portion of the object. The upper portion is very important for human object classification processes, as it contains some human features, i.e. head, neck, shoulders, etc. [28],[48] [113]. Furthermore, this upper portion part, in most cases, clearly appears because it is located on the upper portion of the object and it is not easy for it to be completely occluded from other objects.

To establish the upper portion of the contour, the first (position) pixel value is the top of the head, by scanning the array of row projection from the top to the bottom. The width of the neck is the first minimum value scanning from the top of the head, the width of the head is the maximum value scanning back from the minimum value. In general, the width of the shoulder equals the width of the human body, according to the knowledge of the human body [16]. Lastly, the upper portion contour model will be established. Figure 3.10 demonstrates the object upper portion which is extracted based on a combination of partial steps, these steps are summarized as the following.

1. Obtain the row and column projections from the binary image of the detected contour.
2. Smooth the projection curves.
3. Scanning the smoothed Row projection to perform the following
    A. Find the first non-zero pixel which specify the top of the head.
    B. Find the minimum value after the top of the head to specify the neck width.
4. Scanning the smoothed column projection to perform the following
    A. Find the height of the neck which correspond to the first minimum from the top of the head.
    B. Find the head width which is the maximum value in scanning back from the minimum value and the corresponding height from the head top.
5. Determining the shoulder width as 2.5 – 3 times of the head width.





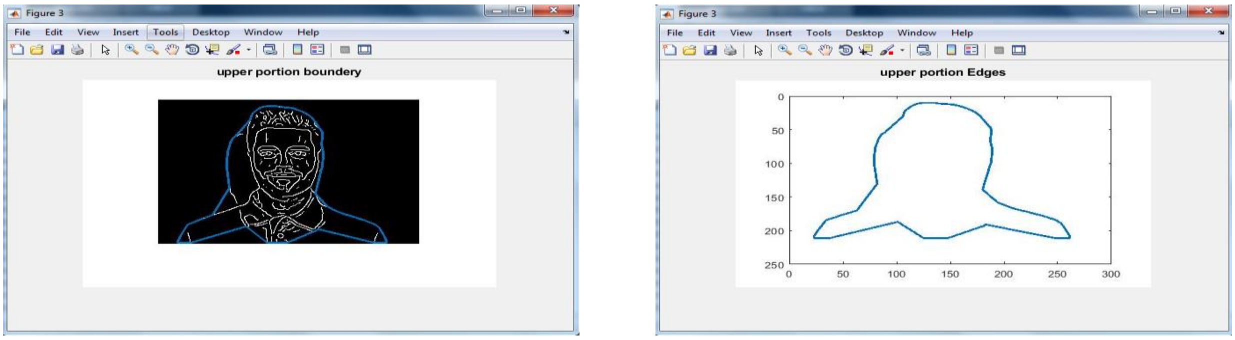


Figure 3.10 Object upper portion extraction

3.2.4 Shape base approach

This is the core step within the research, as the aim is to provide a shape based approach to classify the detected object as human or non-human, which means observing unique features in the human shape which are not applied to other objects. In the approach the focus was on the upper portion of the human shape in order to observe the unique features, these features are geometrical features in the human upper shape. The compilation of these feature can then be used to build up the classifier model of this approach.

To start observing the geometrical features from the human upper shape, the result of obtaining and extracting the upper portion shape from the whole human shape was used by applying the previous steps in sections 3.2.1 to 3.2.3, all these steps are used as a pre phase to establish building the classifier model.

After extracting the upper portion shape of humans from the previous steps, the next step is to obtain the X, Y coordinates for the upper portion shape, and this can be acquired by executing the coordinate function which is built in MATLAB software. Figure 3.11 shows an example of human upper portion shape presented in the cartesian coordinate system.

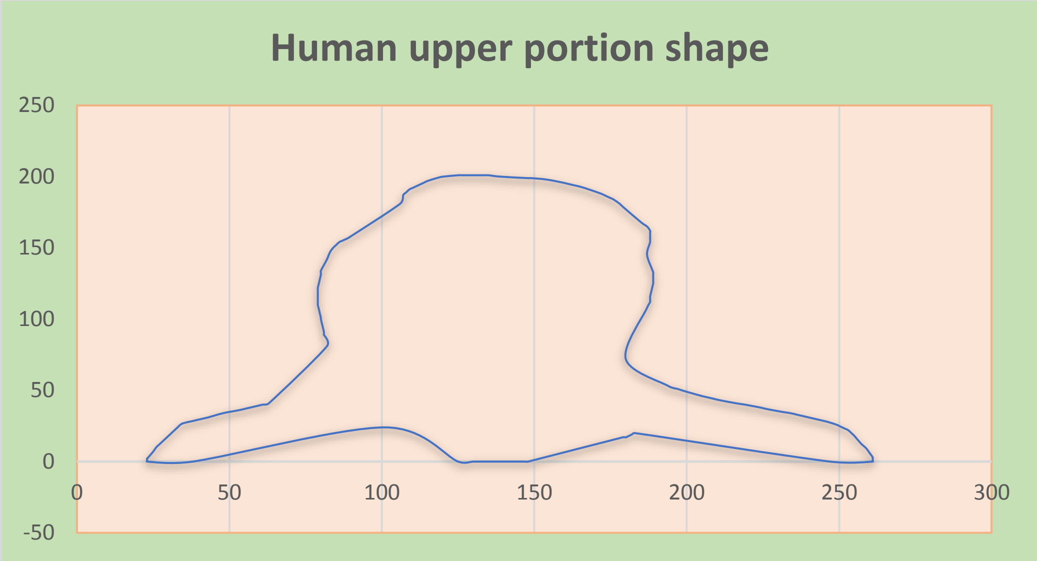


Figure 3.11 an example of human upper portion shape presented in the cartesian coordinate system

The result of executing the coordinate function in MATLAB software provides a matrix of X and Y values for each point in the human upper portion shape, as shown in Table 3.2.

Table 3-2 example of X,Y values for human upper shape coordinates

X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
23	2	79	111	111	193	175	185	187	108	259	9
25	7	79	112	112	194	176	184	180	72	260	6
26	10	79	113	113	195	178	181	194	53	261	3
27	12	79	114	114	196	179	179	197	51	261	2
28	14	79	115	115	197	185	168	200	49	261	1
29	16	79	116	118	199	187	165	205	46	261	0
30	18	79	117	120	200	188	162	207	45	247	0
31	20	79	118	125	201	188	161	209	44	183	20
32	22	79	119	126	201	188	160	211	43	182	19
33	24	79	120	127	201	188	159	216	41	181	18
34	26	79	121	128	201	188	158	219	40	180	17
35	27	79	122	129	201	188	157	224	38	179	17
37	28	80	131	130	201	188	156	226	37	148	0
39	29	80	132	131	201	188	155	231	35	147	0
41	30	80	133	132	201	188	154	234	34	142	0
43	31	80	134	133	201	187	145	236	33	140	0
48	34	81	138	134	201	189	133	238	32	139	0
53	36	82	142	135	201	189	132	240	31	138	0
55	37	83	147	139	200	189	131	242	30	130	0
57	38	84	150	148	199	189	130	244	29	125	0
59	39	85	152	149	199	189	129	246	28	101	24
61	40	86	154	154	198	189	128	249	26	38	0
63	41	87	155	159	196	189	127	250	25	23	0
82	81	88	156	161	195	189	126	251	24	23	1

81	89	89	157	163	194	189	125	252	23	23	2
81	90	106	181	165	193	188	116	253	22		
81	91	107	187	168	191	188	115	254	20		
80	100	108	189	172	188	188	114	255	18		
80	101	109	191	173	187	188	113	257	13		
79	110	110	192	174	186	188	112	258	11		

In this step, the coordinates of x and y were obtained for each pixel in the objects upper portion and presented in Table 3.2. In this study, an intensive investigation and analysis have been conducted by the coordinate values for x and y for large numbers of objects upper shape until the outcome of interest is observed (Proposed geometrical model) to classify the detected object as human or non-human. Figure 3.12 shows the presentation of the X and Y coordinates for the human upper shape which is shown in Table 3.2.

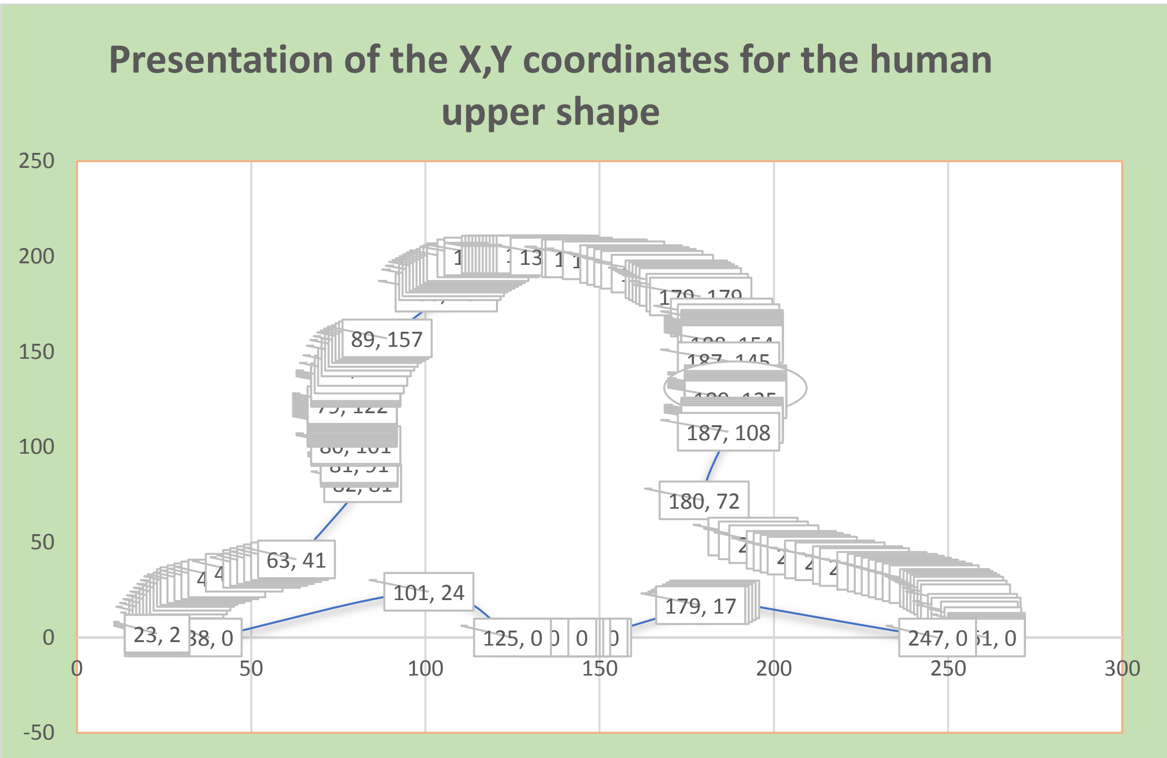


Figure 3.12 presentation of the X,Y coordinates for the human upper shape which shown in Table 3.2

The main observation based on the analysis and intensive study, is that the coordinate values of X for human objects appears not like non-human objects because some of the human shape features include neck, head and shoulders. To present the changes of the X values for humans, Figure 3.13 shows the increases in the X values coloured with green lines and the decreasing of X values which is coloured with red lines.

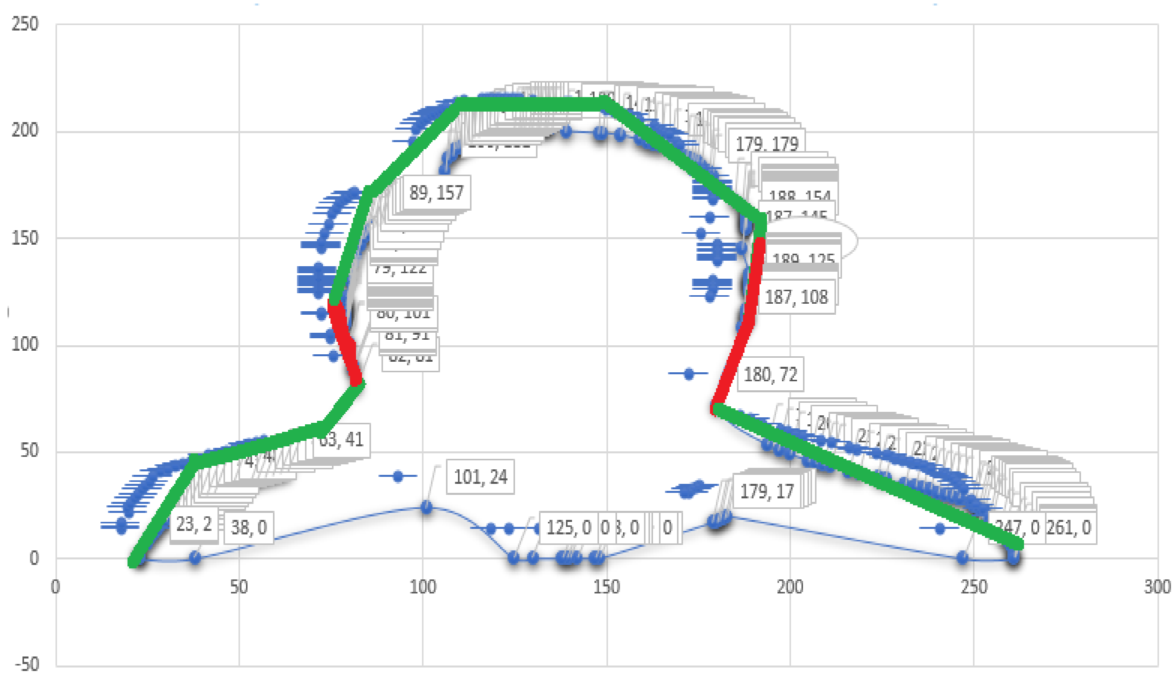


Figure 3.13 Present the changes acquired from the X values

From the above figure (Figure 3.13 and Table 3.2), it can be seen that the X value of the first point is 23, then the X value increases in the next sequence points coloured with green lines until it reaches the point that has the coordinate (82,81) where 82 is the X value. After this point the value of X starts to decrease, in the next sequence points are coloured with red lines until they reach the point that hold the coordinate (79,121), then the behaviour of the X value starts to increases in the next sequence points (presented by the green line) until reaching the point that has the coordinate (188,125).After, the X value start to decrease again (presented by the red line) until reaching the point that holds the coordinate (180,72), in the remaining points the value of X starts to increase until the maximum value of X, presented in the shape is at the point that has the coordinate is (261,0). These changes of the X values allows the start point to be observed and analyses the behaviour of X values and the changes in the X values indicate there are unique features that we can observe it, spatially from the points that the X value starts to increase or decrease.

To analyse the behaviour of X values and the changes acquired in the X values, the sequence X values where extracted from the X, Y coordinates of the upper shape points. Table 3.3 shows the sequence X values obtained from the previous Table 3.2.

Table 3-3 Extract the sequence X values only from the X,Y coordinates of the human upper shape points.

Point Number	X value	Point Number	X value	Point Number	X value	Point Number	X value	Point Number	X value	Point Number	X value
1	23	31	79	61	111	91	175	121	187	151	259
2	25	32	79	62	112	92	176	122	180	152	260
3	26	33	79	63	113	93	178	123	194	153	261
4	27	34	79	64	114	94	179	124	197	154	261
5	28	35	79	65	115	95	185	125	200	155	261
6	29	36	79	66	118	96	187	126	205	156	261
7	30	37	79	67	120	97	188	127	207	157	247
8	31	38	79	68	125	98	188	128	209	158	183
9	32	39	79	69	126	99	188	129	211	159	182
10	33	40	79	70	127	100	188	130	216	160	181
11	34	41	79	71	128	101	188	131	219	161	180
12	35	42	79	72	129	102	188	132	224	162	179
13	37	43	80	73	130	103	188	133	226	163	148
14	39	44	80	74	131	104	188	134	231	164	147
15	41	45	80	75	132	105	188	135	234	165	142
16	43	46	80	76	133	106	187	136	236	166	140
17	48	47	81	77	134	107	189	137	238	167	139
18	53	48	82	78	135	108	189	138	240	168	138
19	55	49	83	79	139	109	189	139	242	169	130
20	57	50	84	80	148	110	189	140	244	170	125
21	59	51	85	81	149	111	189	141	246	171	101
22	61	52	86	82	154	112	189	142	249	172	38
23	63	53	87	83	159	113	189	143	250	173	23
24	82	54	88	84	161	114	189	144	251	174	23
25	81	55	89	85	163	115	189	145	252	175	23
26	81	56	106	86	165	116	188	146	253		
27	81	57	107	87	168	117	188	147	254		
28	80	58	108	88	172	118	188	148	255		

29	80	59	109	89	173	119	188	149	257		
30	79	60	110	90	174	120	188	150	258		

After extracting the X values, the aim was to plot each one of these values with its sequence point number, in order to present a histogram of the X values, which means the X value of the first point starts from the left side of the human shape corresponding with number one, and the second X value will correspond with number two, and so on until the end point, as shown in table 3.3. The grey column is the sequence order and the green column is the corresponding X value,as each pixel in the shape will be presented as a point, and any small change in the pixels will be reflected in the histogram, by applying a mathematical smoothing function to avoid small angles. Figure 3.14 shows the histogram of the X value with it is corresponding sequence order.

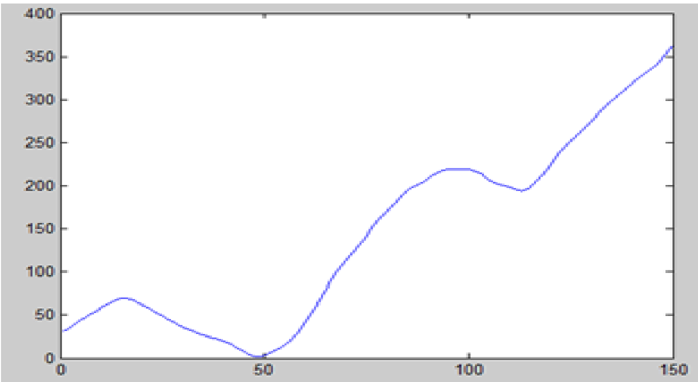


Figure 3.14 the histogram of the X value with it is corresponding sequence order

3.2.5 Geometrical model

This research aims to present a shape-based approach to classifying the objects as human or non-human, to achieve this aim the focus was to observe geometrical features of humans from their upper shape, so as to build up the classifier model. In order to extract geometrical features for humans, a random dataset was obtained from INRIA dataset, the selected dataset contains two labelled groups for both human and non-human, and then all the previous steps are applied until the histogram are received for each X value with its corresponding sequence order, for all objects in the selected dataset. Figure 3.15 shows examples of the X values

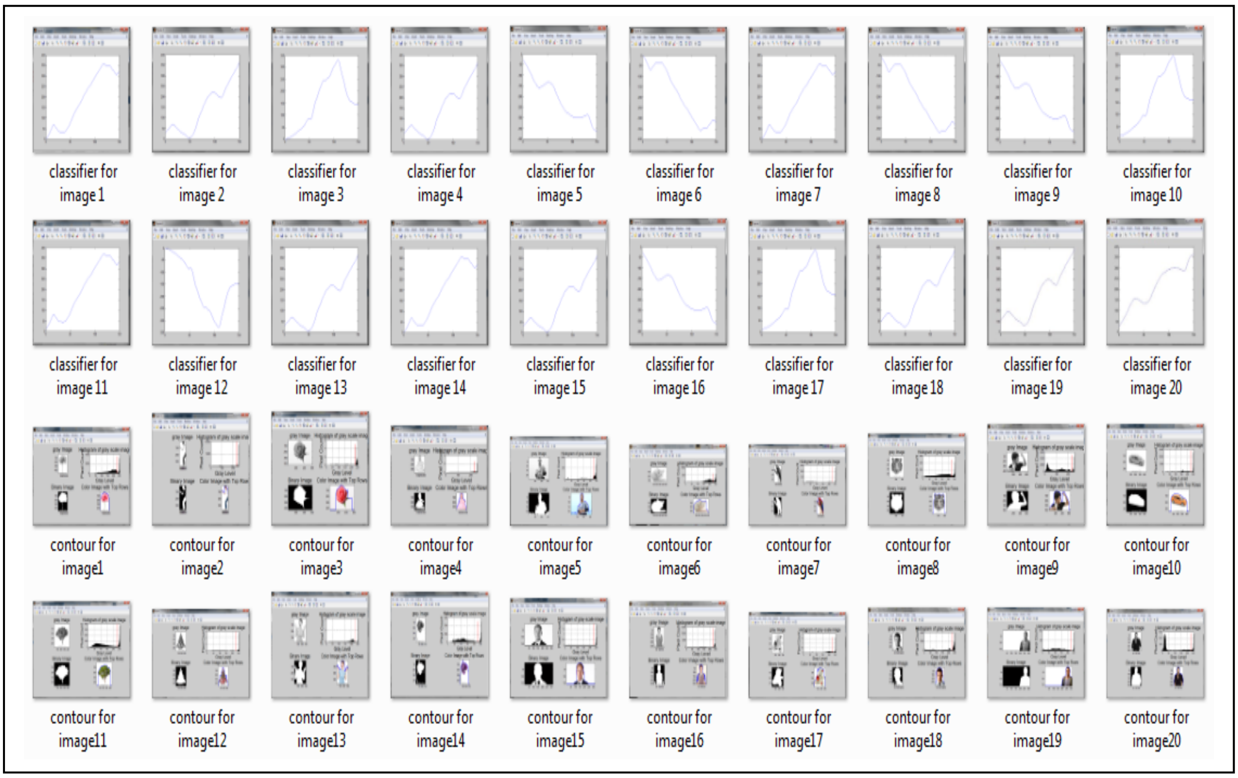


Figure 3.15 examples of the X values histogram for the selected dataset.

histogram for the selected dataset.

After obtaining the X values histogram for the selected dataset, the analysis of these histograms can begin using histogram analyses tools such as curvature function and find peek function. These functions are very useful when used to analyse the plot or histogram in terms of finding the number of peeks and the curve similarity [21], [111],[67], both of these function are integrated with MATLAB software. Figures 3.16 shows examples of found peek points in the plot histogram and the position of these peek points, Figure 3.17 shows examples of smoothing histograms and finding the number of peeks.

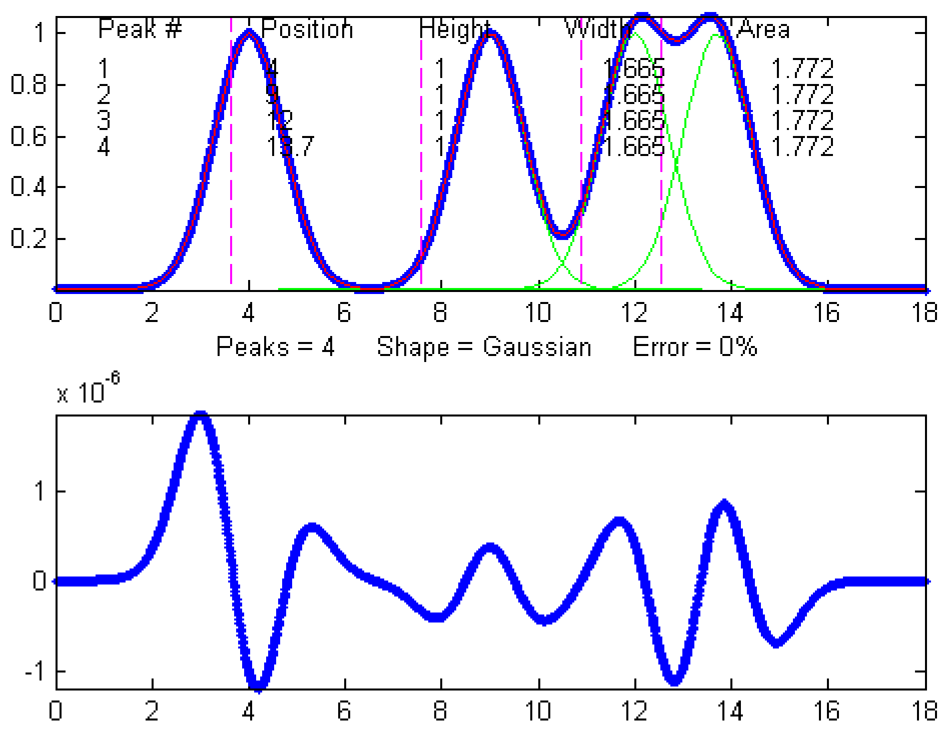


Figure 3.16 Example of find peeks points in plot histogram.

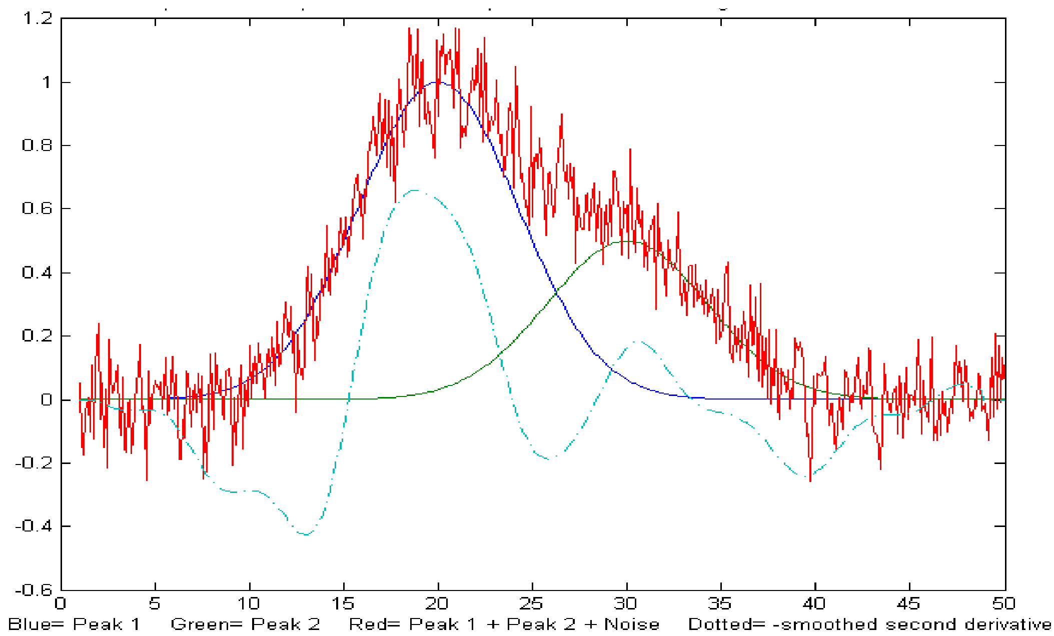


Figure 3.17 Example of smoothing histogram and finding number of peeks

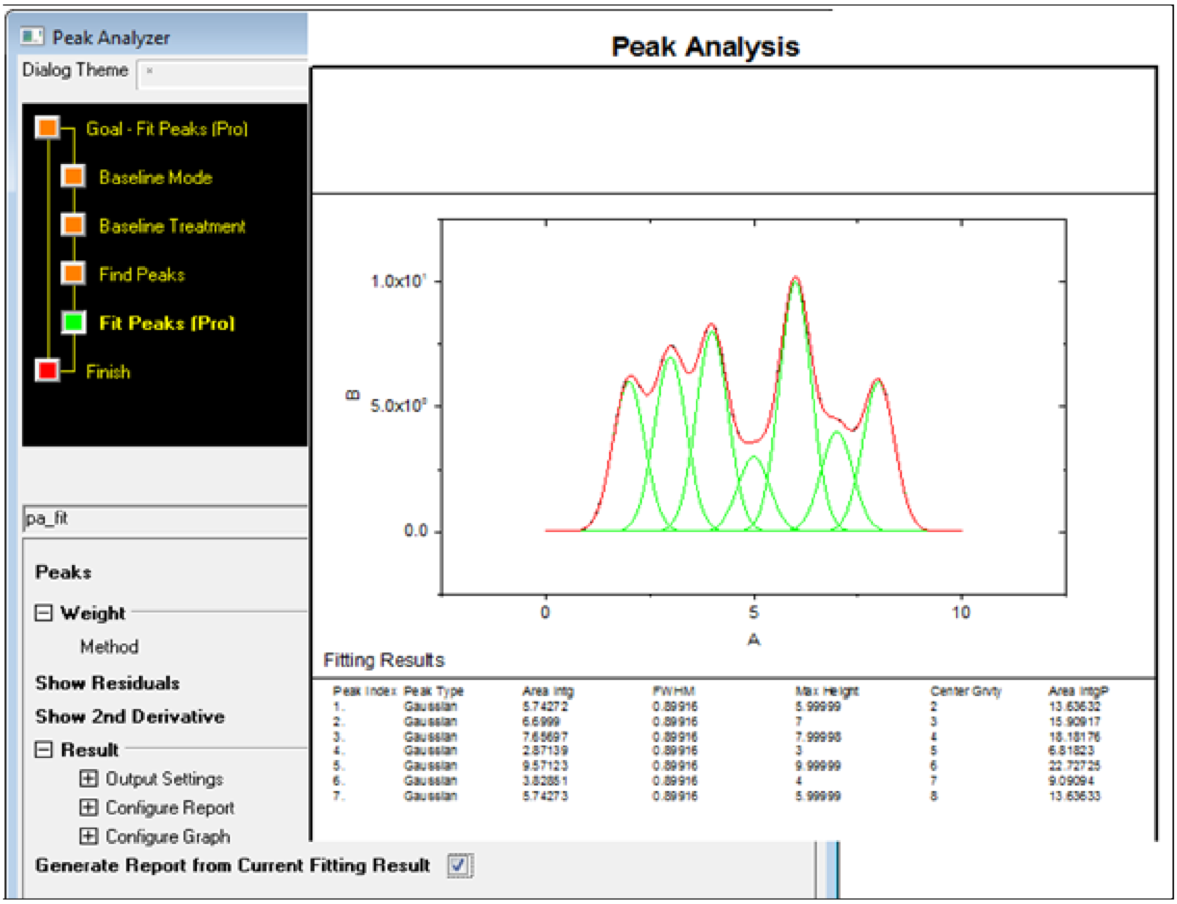


Figure 3.18 Example of peeks analysing

After analysing the plots of X value histograms using peek analysis software and finding the number and location of peek points for each object. Figure 3.19 shows the report of peek analysis for the selected samples of dataset.

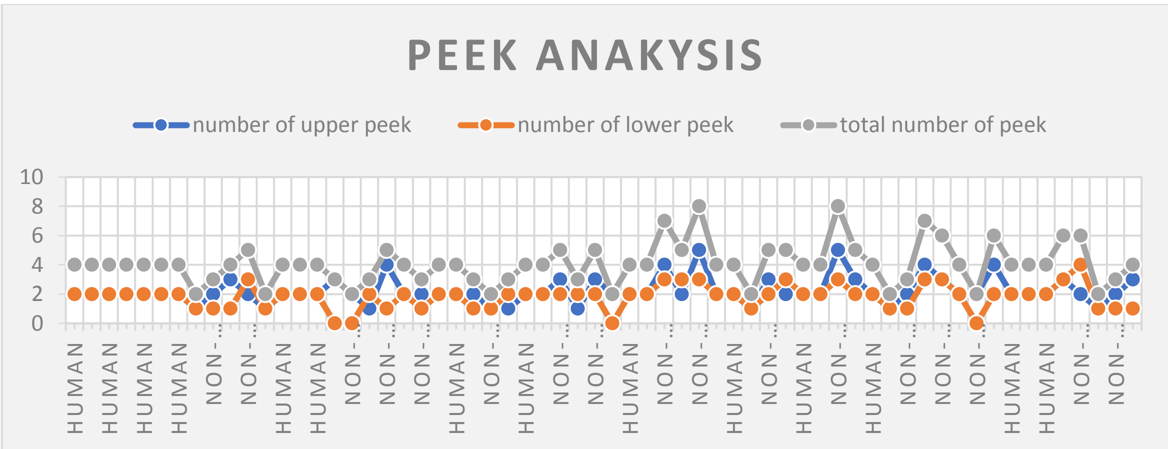


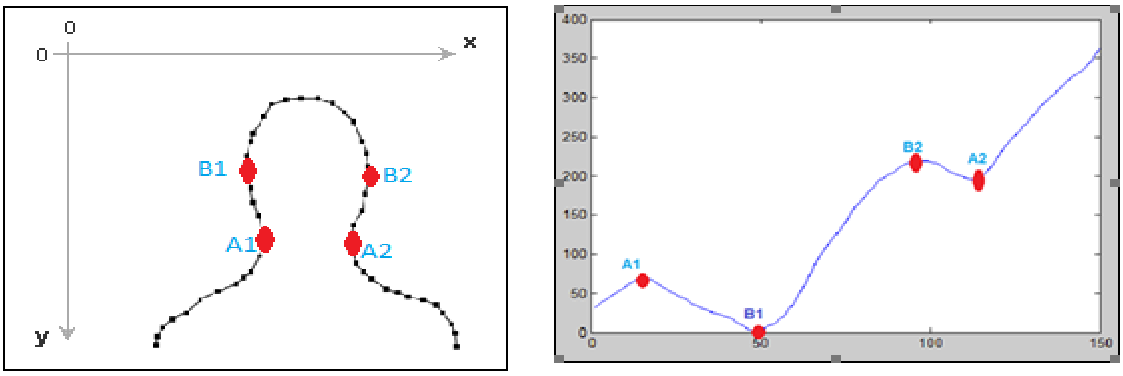
Figure 3.19 Report of peek analysis for the selected samples of dataset

Based on the results of the peek analysis, a newly developed object classifier based on object upper portion shapes by generating a geometrical mode has been created. This geometrical model has four parameters:

Parameter 1: Histogram observation, the histogram of the coordinate values of X for the human object which detracted from the upper portion has two upper peaks denoted by A1 and B2 and two lower peaks denoted by B1 and A2, which is not found on the histogram of the coordinate values of X for the non-human object as shown in Figure 3.20.

The number of the lower peak points for human object = 2

The number of the upper peak points for human object = 2



(a) Human object *(b) Upper and lower peaks for human object*

Figure 3.20 Describe the location of the Upper and lower peak points

After observing the number and location of upper peeks and lower peeks for human objects, and finding the location of each peek point, the rest of the geometrical parameters were observed by finding geometrical relations between these peek points.

These geometrical relations are based on the distances between the peek points, however, the location of each peek point is obtained with its corresponding X, Y coordinates, by applying the Euclidean distance formula, to calculate the distance between two peek points.

Based on the Euclidean distance formula, the distance between two points in the plane with coordinates (X_2, Y_2) and (X_1, Y_1) is given by the following formula

$$\text{Distance } ((X_2, Y_2), (X_1, Y_1)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Parameter 2: D1 and D2 observation, for human object we have observed that, the distance of D1 between A1 and A2 is less than the distance of D2 between B1 and B2 as shown in Figure 3.21.

The distance D1 <The distance D2

Parameter 3: D3 and D4 observation, for human objects it was observed that the distance of D3 between A1 and B2 is equal to the distance of D4 between B1 and A2 as shown in Figure 3.21, the distance threshold ts1 should take a small value, which leads to more accuracy.

The distance D3 = the distance D4 ± (ts1), | D3 – D4| = ts1

Parameter 4: After an anatomical science observation, the fourth parameter is based on the calculations and the measurements between the coordinates of the lower and upper peek points and between the minimum and maximum [x] values as shown in Figure 3.21. It was found that for human objects the distance of **D2** between the first lower peak point (*denoted by B1*) and the second upper peak point (*denoted by B2*) is more than one third of the distance of **D5** between the start point (*the minimum value of [X] denoted by C1*) and the end point of the shoulder (*the maximum value of [X] denoted by C2*), and less than two thirds of the distance of **D5,** this is often true in anatomical science [24] [28].

$$\frac{\mathbf{1}}{\mathbf{3}}(\boldsymbol{D5}) < \boldsymbol{D2} < \frac{\mathbf{2}}{\mathbf{3}}(\boldsymbol{D5})$$

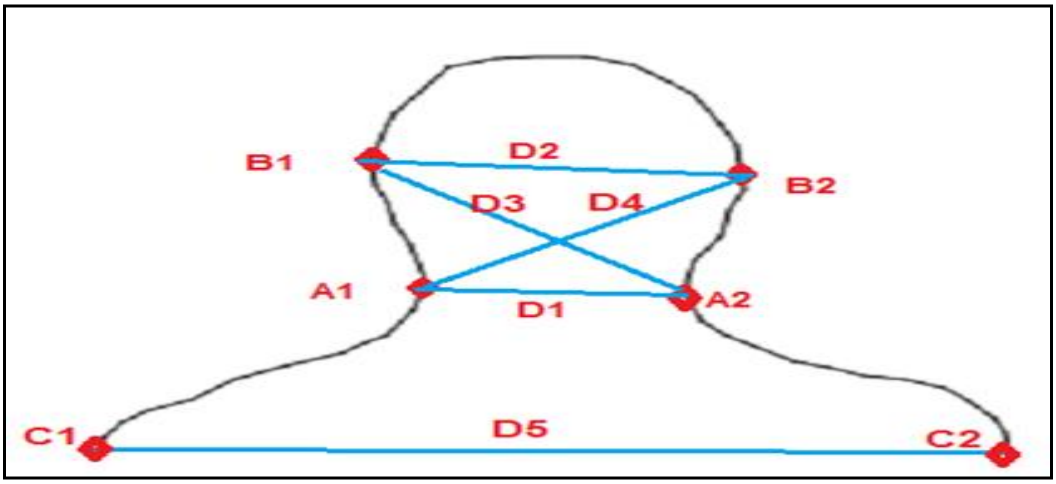


Figure 3.21 Description of classifier parameters

In general, the parameters of the proposed geometrical model for object classification can be summarized as seen in Table 3-4.

Table 3-4 Parameters of the proposed geometrical model for object classification

Parameter 1:	The number of upper peak point = 2, The number of lower peak point = 2
Parameter 2:	The distance D1 < the distance D2
Parameter 3:	The distance D3 = the distance D4 ± (ts1), | D3 – D4| = ts1 where the ts1 is a threshold.
Parameter 4:	The distance D2 > 1/3 of distance D5 And The distance D2 < 2/3 of distance D5, $\frac{1}{3}\mathrm{D5} < D2 < \frac{2}{3}\mathrm{D5}$

The parameters of the proposed geometrical model are implemented to detect human objects in images. Therefore, the proposed approach can classify the detected object as human if the parameter values are true**,** otherwise the detected object will be classified as non-human.

3.2.6 Proposed approach steps

The proposed approach is based on a set of parallel and sequential processes, which are summarized as the following:

1. Background subtraction using histogram-based techniques with global threshold).
2. Object edge detection using CANNY edge detection approach.
3. Extract the boundary edge using the boundary function.

Extract the upper portion of the contour.

4. Extract the boundary *{xi, yi}* coordinate matrix points for the contour which obtained from background subtraction/boundary edge detection.
5. Obtain *Y min***,** *Y max***,** X min**,** X max values from the boundary's points obtained in step4
6. Obtain the row and column projections from the binary image of the detected contour.
7. Smooth the projection curves using smooth function *S*.
8. Scan the smoothed Row projection to perform the following:
   8.1. Find the first non-zero pixel to specify the top of the head (*th*).
   8.2. Find the minimum value after the top of the head to specify the neck width (*nw*).
9. Scan the smoothed column projection to perform the following:
   9.1. Find the height of the neck which corresponds to the first minimum from the top of the head (*hn*).
   9.2. Find the head width which is the maximum value in scanning back from the minimum value and the corresponding height from the head top (*hw*).
10. Determine the shoulder width as 2.5 – 3 times of the head width (*hw*).
11. Extract the upper portion of the contour (i.e. selected object).

Establishing the geometric model

12. Obtain *{xi, yi}* coordinates for the upper portion of the contour.
13. Represent the obtained X values coordinate in a histogram.
14. Smooth the histogram using mathematical smooth functions *S*.
15. Find the upper peak points (*up*) and lower peaks points (*lp***)** in the histogram obtained from step 14.
    Parameter #1(*P1*): See Figure 3.13 &3.14.
16. Find the number of upper peak point and the number of lower peak points.
    16.1. Take a decision
        IF (The number of the upper peak point (*up*) = = 2)
        AND

(The number of the lower peak point (*lp*) = = 2) Then, P1= =1

Else

P1= = 0.

Parameter #2(*P2*): See Figure 3.14.

17. Find the distance (*D1*) between the first upper peak point (*up1*) and the second lower peak point (*lp2*).
18. Find the distance (*D2*) between the second upper peak point (*up2*) and the first lower peak point (*lp1*)
19. Take a decision

IF (*D1* < *D2***)** Then, P2= =1

Else

P2= = 0.

Parameter #3(*P3*): See Figure 3.14.

20. Find the distance (*D4*) between the two upper peak points (*up1, up2*).
21. Find the distance (*D3*) between the two lower peak points (*lp1, lp2*).
22. Take a decision

IF (| *D3 – D4***| =** *ts1* where the ts1 is a threshold. Then, P3 = = 1

Else

P3 = = 0.

Parameter #4(*P4*): See Figure 3.14.

23. Find the distance (*D5*) between the start point and the end point of the shoulder (*C1, C2*).
24. Get the distance (*D2*) which obtained in step 18.
25. Take a decision

IF ($\frac{1}{3}D5 < D2 < \frac{2}{3}$D5) Then, P4= =1

Else

P4 = = 0.

Classification decision

26. **IF (P1=1 And P2=1 And P3=1**

AND

P4 = =1) Then,

The detected object is Human

Else

The detected object is Non-human

END

CHAPTER FOUR

4 WEIGHT BASED DESIGN FUNCTION

4.1 Introduction

In this chapter, a modification of the proposed shape-based approach is carried out to classify the detected objects in different positions as human or non-human. The core idea of the modification rising is based on the discussion of some experimental results, by providing the proposed approach with a weight-based design function.

The structure of this chapter is organized as the following, dataset and justification, pre-experiments, shape-based approach modification, pro-experiments and conclusion of this chapter.

4.2 Datasets and justification

There are many datasets which are available nowadays, but the task of choosing the sufficient dataset is a very important task, spatially when dealing with object detection and recognition algorithms. This is because, for a robust detection we need the good capabilities of computer vision approaches , which must have the ability to bring out the mutual qualities of the interested object in different conditions, in order to detect and recognise the object with a high level of accuracy. The researchers in the area of computer vision rely heavily on evaluating and testing the performance of their new algorithms, in order to compare the new algorithm with other related algorithms, to achieve this goal, the use of benchmark data sets becomes necessary.

Dollar et al. [126] provided an efficient summary of most datasets which are freely available, and can be used in evaluating the object detection approaches. A comprehensive dataset was published under the name of Caltech Pedestrian dataset. [127] Presented a summary for various public datasets in term of object action recognition in video sequences. Similarly, making a categorisation of these public datasets based on the type of detected object, such as single human object dataset, movement objects datasets and social objects interaction.

Recently, there have been many benchmark datasets for human detection, which are public and available to evaluate the performance of new algorithms. These benchmark datasets are collected from several scenarios, and under different conditions such as viewpoint appearance, partial occlusion, and posture.

The different variety of these benchmark datasets are related to the wide range of real applications that it can be used in. For example, some of these benchmark datasets contain only images however, others may contain images and videos. Furthermore, these datasets are classified in different categories based on different purposes of use. For example, some of these datasets are used to detect humans in general purposes such as ("INRIA, MIT, USC-A, USC-C, and Penn-Fudan") datasets. Other datasets can be used for monitoring purposes such as ("USC-B, and CAVIAR") datasets. For pedestrian detecting there are ("Caltech, TUD, Daimler Chrysler, the ETH, and CVC") datasets.

Most of these benchmark datasets contain two folders, one is called train folder which contains some images or videos to be used for machine learning algorithms, and the second is called test folder, which contains images or videos for testing the performance of algorithms, each one of these folders are divided into two categories; positive samples and negative samples.

These datasets also have different specifications for images and videos, such as the resolution of the images (pixel format), the length of videos, and so on.

In this study, images were collected from **INRIA** datasets and from the **Caltech 101** dataset, these were chosen for testing and evaluating the proposed approach, because these two datasets have some characteristics and properties that make them more sufficient and applicable for testing and evaluating this approach. This approach aims to classify the human object based on it is upper shape, and **INRIA** dataset for example was collected as part of the research work on the detection of upright people in images, also both of these datasets contained labelled images which make the use of this kind of dataset very useful in terms of classification purposes Furthermore, these two datasets are widely used in testing and evaluating human detection algorithms, because these two datasets contain images from several different sources, all the images have very good resolution, and highlight the people, also many humans in these datasets are bystanders, so ideally there is no particular bias in their pose. Moreover, these datasets contains 101 categories of different objects and each category contains 40 to 800 images in the size 200 *300-pixels, which makes the selected images from both datasets very useful in terms of testing and evaluating the proposed approach. Both of these datasets contain two group formats, the original images and positive images, which are in a normalized 64x128 pixel format, these two group formats provide power in evaluating any new algorithm with machine learning approaches because the first step in the machine learning approach is to learn, and this can be easily based on these two group formats.

The original folder has two sub folders; the train folder and test folder and each one of these two folders are divided into two categories which incude positive images and negative images. A comparison between the public available human object datasets are shown in Table 4.1 [104].

Table 4-1 publicly available human object dataset

	Properties				Training		Testing			Current Performance	
Dataset	static image	video	stereo	occlusion	#humans	#neg images	#humans	#neg images	#full images	miss rate at 10^{-4} FPPW	miss rate at 1 FPPI
MIT [61]	√						924			≈ 0.0 [22]	
INRIA [66]	√				2416	1218	1126	453	228	0.02 [165, 77]	0.15 [26]
CAVIAR [1] [62] *corridor view*		√		√			5614		1590		
USC-A [63]	√						313		205		0.03 [170]
USC-B [63]	√			√			271		54		0.03 [92]
USC-C [63]	√						232		100		0.05 [171]
Penn-Fudan [68]	√						345		170		0.25 [113]
DC [73]	√				14400	1200 + 15000 (samples)	9600	10000 (samples)		≈ 0.6 [137]	
TUD [2] [70]		√			3552 × 2	192 × 2 + 26 × 2	1326		508 × 2		0.35 [27]
Caltech [75]		√		√	192000	61000	155000	56000	65000		0.8 [26]
CVC [3] [69]	√						2000	6175 (samples)			
ETH [4] [76]			√				12000		1804		0.3 [27]

4.3 Experiments before modifying the proposed approach

In this section, the presented approach for human detection in images is evaluated based on a mathematical model, the presented approach is performed in order to test the validation of the presented approach. The performance of this phase runs before making the modification of the proposed approach. The performance of the presented approach can be summarised in two experimental results. In these two experiments different images were used for several objects obtained from INRIA datasets (set of human and non-human as an object in digital images) and Caltech 101dataset in order to distinguish the accuracy level. The selected images were labelled and contained a single object in different camera poses and viewpoints.

The Caltech 101 dataset contains 101 categories of different objects and each category contains 40 to 800 images in sizes of 200 *300 pixels [128], whereas the INRIA contains two group formats, the original images and positive images, cropped in different sizes such as 64 × 128 pixels, and 214 × 320 – 648 × 486 pixels [41]. For a homogeneous dataset (same size of pixels) the selected images were cropped for the experiment into 64 x 128 pixels using an image cropper approach [104]. The proposed approach was implemented using MATLAB R2017b and tested on 1.8 GHz core i7 (IV), 16 GB memory and 512 GB hard drive. The performance analysis of these experiments where performed in two tests: accuracy matrix, and confusion matrix.

4.3.1 Experimental number one

This experiment is the first to evaluate performance of the proposed approach in order to classify the detected object in different positions as human or non-human based on object shape. Figure 4.1 shows samples of objects in the dataset.

Figure 4.1 Some sample of objects in the dataset

This experiment is based on 450 images in total divided into two classes, the human class which contains 150 images and the non-human object's class which contains 300 images. The non-human object's class contains images for varied types of objects such as monkeys, horses', dogs, cars, and other types of non-human objects.

As mentioned before in chapter three, the proposed approach is a shape-based object detection approach which can classify the human object based on it is mathematical model, there are a sequence of steps needed in order to apply the mathematical model, this sequence of steps starts from the object extraction step, edge detection, contour detection, upper portion extraction. The projection of X coordinates for the upper portion contour are then presented in order to perform the proposed approach classifier based on the mathematical model in order to classify the object as human or non-human. Figure 4.2 and Figure 4.3 show some samples of the experimental result of these steps.

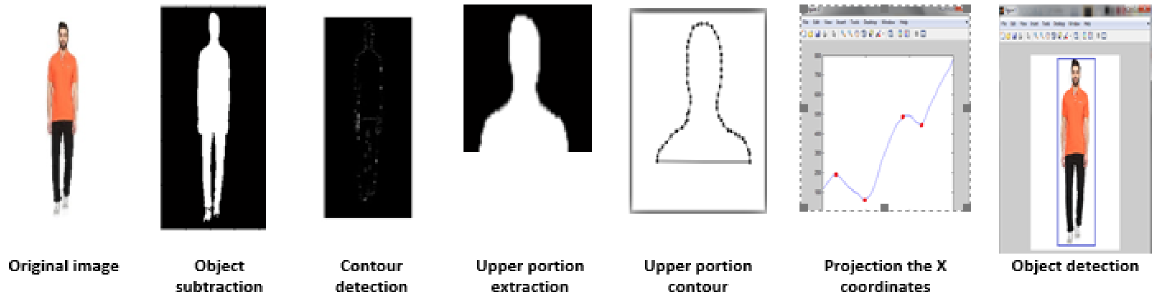


Figure 4.2 The experimental result of sequence steps for human object

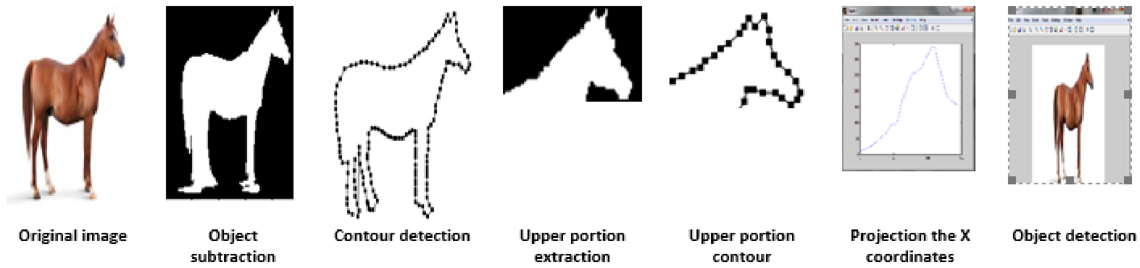


Figure 4.3 The experimental result of sequence steps for Non-human object

As shown in figure 4.2 and figure 4.3, the classifier of the proposed approach classifies the detected object as human by surrounding the human object with a blue rectangle.

After implementing the experiment based on the selected sample of images which contained 450 images in total divided into two classes, the human class which contains 150 images and the non-human object's class which contains 300 images . The experimental results of the detection performance of the proposed approach can be presented in a Confusion matrix, this Confusion matrix is a specific table that is used to visualise the approach performance, this table includes two rows which represents the instances in a predicted class, and two columns which represents the instances in an actual class. From this matrix, the performance analyses of the approach can be reported by presenting the number of *false positives*, *false negatives*, *true positives*, and *true negatives*. The confusion matrix test of the proposed approach for this experiment can be seen in Table 4.2.

Table 4-1 The confusion matrix test for the proposed approach

a	b	Classified as		
259	41	a	=	NH
21	129	b	=	Hu

From the confusion matrix table, the performance accuracy of the proposed approach can be obtained from this experiment. The confusion matrix accuracy is given by Equation 4.1:

$$Accuracy = \frac{\sum \text{Diagonal Sample of confusion matrix}}{\text{Total Sample}} \tag{4.1}$$

By performing the above equation to calculate the performance accuracy of the proposed approach in this experiment, it was found that the proposed approach detects 259 objects as non-human from 300, and detects 129 objects as human from 150. The overall performance detection accuracy in this experiment is 388 of 450, where 388 is $\sum$ Diagonal Sample of the confusion matrix, and 450 is the total number of the dataset samples, this means the accuracy of this approach in this experiment is equal to 86.2%.

From the confusion matrix report, the performance analysis of the proposed approach, indicates that for the human class 129 images from the actual number of human (150)) images are detected truly as human (True positive), and 21 images from the actual number of human images are detected as non- human (False negative), while 259 images from the actual number of non-human (300)) images are detected truly as non-human (True Negative**)**, and 41 images from the actual number of non-human images are detected as human (False Positive).

For the non-human class of 259 images the actual number of non-human images are detected as truly non-human (True positive), and 41 images from the actual number of non-human images are detected as human (False negative), while 129 images from the actual number of human images are detected truly as human (True Negative), and 21 images from the actual number of human images are detected as non- human (False Positive). Table 4.3 shows a summary for the performance analysis of the proposed approach in this experiment for each class.

Table 4-2 A summary for the performance analysis of the proposed approach in experiment number one.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	129	41	259	21
Non-Human	259	21	129	41

4.3.1.1 Statistical analysis of the Performance

For a more statistical measurement of the analysis of the performance of the proposed approach in this experiment, some of the most common statistical measures function have been calculated for deep analysis of the proposed approach performance, such as the Sensitivity, Precision, Negative predictive value, Specificity, Miss rate, Fall-out, False discovery rate, False omission rate, and the Accuracy. Table 4-3 shows the corresponding formula and description for these statistical functions which is taken from Wikipedia.

Table 4-3 The corresponding formula and description for these statistical functions.

Function name	**The Formula of the function**	**Description**
Sensitivity (True positive rate)	Sensitivity (TPR) $=\frac{TP}{TP+FN}$	The Sensitivity also called as Recall or True positive rate (TPR): this is statistical measures that present a measurement of the proportion of actual positives detection that are correctly identified.
Precision (positive predictive value)	Precision (PPV) = $\frac{TP}{TP+FP}$	The Precision also called as positive predictive value (PPV)
Negative predictive value	Negative predictive value (NPV) = $\frac{TN}{TN+FN}$	Negative predictive value is the proportion of individuals with a **negative** test result who are free of the target condition
Specificity (True Negative rate)	Specificity (TNR) = $\frac{TN}{TN+FP}$	The Specificity, also called as selectivity or True Negative rate (TNR), measures the proportion of actual positives that are correctly identified
Miss rate also called False Negative Rate	Miss rate (FNR) = $\frac{FN}{FN+TP}$	The Miss rate also called False Negative Rate (FNR), The fraction or **percentage** of accesses that result in a **miss**
The Fall-out also called False Positive Rate	Fall-out (FPR) = $\frac{FP}{FP+TN}$	The Fall-out also called False Positive Rate (FPR), is the probability of falsely rejecting the null hypothesis for a particular test.
False discovery rate	False discovery rate (FDR) = $\frac{FP}{FP+TP}$	is a method of conceptualizing the **rate** of type I errors in null hypothesis testing when conducting multiple comparisons.
False omission rate	False omission rate (FOR) = $\frac{FP}{FP+TP}$	False omission rate is measures the proportion of **false** negatives which are incorrectly rejected.
The Accuracy	The Accuracy (ACC) = $\frac{TP+TN}{TP+TN+FP+FN}$	The **accuracy** of a measurement is how close a result comes to the true value.

The statistical analysis for the performance of the proposed approach in this experiment provide a full description for the validity of the proposed approach in classifying the detected object as human or non-human. From this description it can be noted that the influence of this approach in terms of positively classifying the detected objects and the weaknesses of this approach in terms of false classification of detected objects. Each one of these statistical analyses describes the performance of the approach in different aspects in order to evaluate and validate the approach. The result of the statistical analysis of the performance of the proposed approach is summarised in table 4-4

Table 4-4 The summarise result of the statistical analysis for the performance of the proposed approach

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	86%	75.88%	92.5%	86.33%	14%	13.66%	24.11%	7.5%	86.22%
Non-Human	86.33%	92.5%	75.88%	86%	13.66%	14%	7.5%	24.11%	86.22%

Table 4-4 presents the distribution result of the statistical analysis of the performance of the proposed approach for this experiment as shown in Figure 4.5.

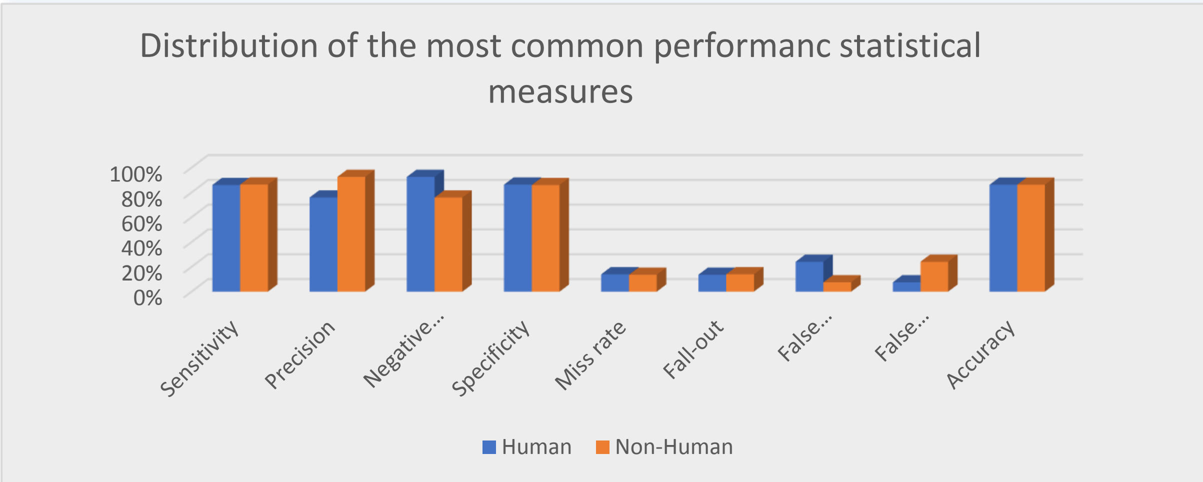


Figure 4.4 The distribution result of the statistical analysis for the performance of the proposed approach for this experiment

The experimental result for the performance of the proposed approach based on the performance statistical analysis as shown in Table 4-4 and figure 4.5 indicating that the accuracy of the proposed approach in classifying the detected objects as human is 86.22% with a miss-detection rate of about 14%, and the false discovery rate for human class is 24.11% and 7.5% for the non-human class. This result shows the success of the proposed approach in terms of classifying the detected object as human or non-human however, it does not achieve as high efficiency compared with the stat of the art approaches such as machine learning approaches (Artificial Neural Networks (ANN), Support Vector Machine (SVM) Model, and a famous type of decision tree called Random Forest). For this reason, a deep analysis of the performance of the proposed approach was performed spatially for false detected objects in order to improve the performance of the proposed approach to acquire state-of-the-art efficiency.

The results of the analysis found falsely detected objects, for example, humans classified as non-human (False Negative) or non-human classified as human (False Positive), it was found that some non-human objects were classified as human in the proposed approach, these objects were falsely classified because they have a similar shape to a human such as monkeys. The classifier of the proposed approach performs based on the shape of an object. This leads therefore, to another experiment in order to discover the weaknesses in the proposed approach classifier.

4.3.2 Experiment number two

After performing experiment number one and analysing the performance of the proposed approach, it was indicated that the performance result shows the success of the proposed approach in terms of classifying the detected object as human or non-human. However, it does not achieve the highest efficiency compared with others, and after performing a deep analysis of the proposed approach spatially for the false detection of objects in order to improve the performance of the proposed approach to achieve higher efficiency. It was found that some non-human objects were classified as human in the proposed approach, these objects were falsely classified because they have a similar shape to humans, the classifier of the proposed approach performance is based on the shape of the object therefore, in this experiment the aim is to evaluate the performance of the proposed approach to reduce challenges in data set samples which include images of humans and monkeys only.

In this experiment we different images for humans and different kinds of monkeys were used, these images were obtained from INRIA dataset and Caltech 101dataset in order to distinguish the accuracy level. Figure 4.6 shows a sample of objects within the dataset

Figure 4.5 Some samples of objects in the dataset

The performance analysis of this experiment was performed in two tests: accuracy matrix, and confusion matrix. The experiment was based on 240 images in total divided into two classes, the human class which contains 160 images and the non-human object's class which contains 80 images. The non-human object's class contains images of a variety of different kinds of monkey. Figure 4.7 and Figure 4.8 show some samples of the experimental result of these steps, where the proposed approach classifies the human object by surrounding the human object detected with a blue rectangle as shown in figure 4.8.

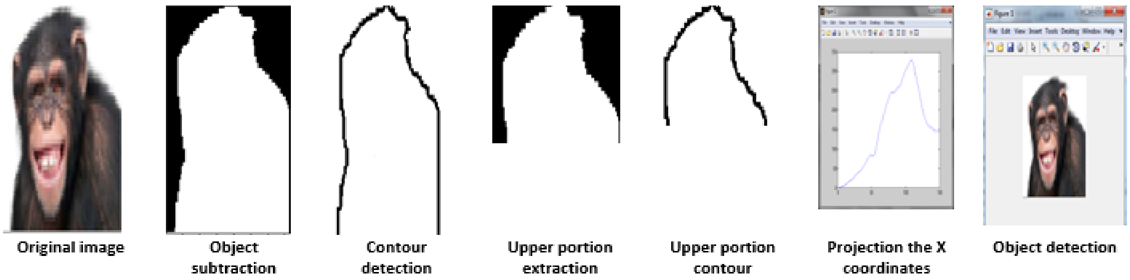


Figure 4.6 The experimental result of sequence steps for Non-human object

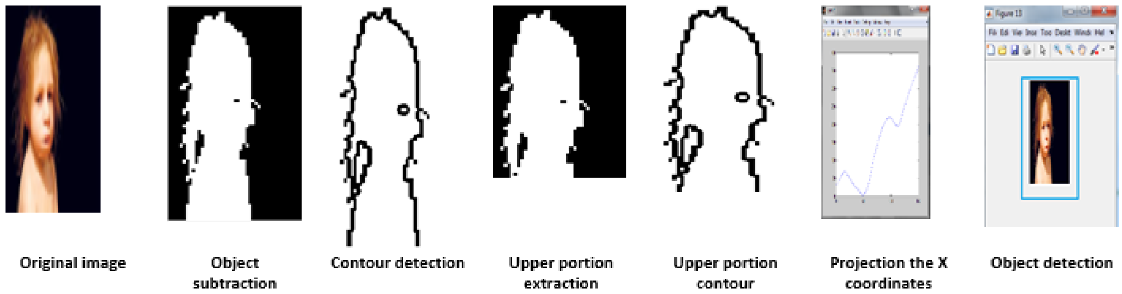


Figure 4.7 The experimental result of sequence steps for human object

The confusion matrix test of the proposed approach for this experiment is reported in Table 4.2.

Table 4-2 The confusion matrix test for the proposed approach in experiment number two

a	b	← Classified as		
51	29	a	=	Monkeys
56	104	b	=	Human

From the confusion matrix table, the performance accuracy of the proposed approach in this experiment can be obtained. The confusion matrix accuracy is given with Equation 4.1:

$$Accuracy = \frac{\sum \text{Diagonal Sample of confusion matrix}}{\text{Total Sample}} \tag{4.1}$$

By performing Equation 4.1 to calculate the performance accuracy of the proposed approach in this experiment, it was found that the proposed approach detects 51 objects as non-human (monkeys) of 80, and detects 104 objects as human of 160. The overall performance detection accuracy in this experiment is 155 of 240, where 155 is ∑ Diagonal Sample of the confusion matrix, and 240 is the total number of dataset samples, meaning the accuracy of this approach within this experiment is equal to 64.58%.

From the confusion matrix report, the performance analysis of the proposed approach, indicates that for the human class of 104 images from the actual number of human (160) images are detected truly as human (True positive), and 56 images from the actual number of human images are detected as non- human False negative), while 51 images from the actual number of non-human (80) images are detected truly as non-human (True Negative**)**, and 29 images from the actual number of non-human images are detected as human (False Positive).

For the non-human (Monkeys) class 51 images of the actual number of non-human (80) images are detected truly as non-human (True positive), and 29 images from the actual number of non-human images are detected as human (False negative), while a 104 images from the actual number of human (160) images are detected truly as human (True Negative), and 56 images from the actual number of human images are detected as non- human (False Positive). Table 4.3 shows a summary for the performance analysis of the proposed approach within this experiment for each class.

Table 4-5 A summary for the performance analysis of the proposed approach in experiment number two.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**104**	**29**	**51**	**56**
Monkeys	**51**	**56**	**104**	**29**

4.3.2.1 Statistical analysis of the Performance

For a more statistical measurement, an analysis of the performance of the proposed approach, some of the most common statistical measure functions have been calculated for deep analysis of the proposed approach performance, such as the Sensitivity, Precision, Negative predictive value, Specification, Miss rate, Fall-out, False discovery rate, False omission rate, and Accuracy, a full descriptions of these functions are shown in table 4-3.

The statistical analysis for the performance of the proposed approach in this experiment provide a full description for the validity of the proposed approach in classifying the detected object as human or non-human. From this description we can note the power of this approach in terms of positive classification of the detected object and the weakness of this approach in terms of falsely classifying the detected object. Each one of these statistical analysis defines the performance of the approach in different aspect in order to evaluate and validate the approach. The result of the statistical analysis of the performance of the proposed approach is summarised in table 4-6

Table 4-6 The summarise result of the statistical analysis for the performance of the proposed approach

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	**65%**	**78.19%**	**47.66%**	**63.75%**	**35%**	**36.25%**	**21.8%**	**52.33%**	**64.58%**
Monkeys	**63.75%**	**47.66%**	**78.19%**	**65%**	**36.25%**	**35%**	**52.33%**	**21.8%**	**64.58%**

Table 4-6 presents the distribution results of the statistical analysis for the performance of the proposed approach for this experiment is shown in Figure 4.5

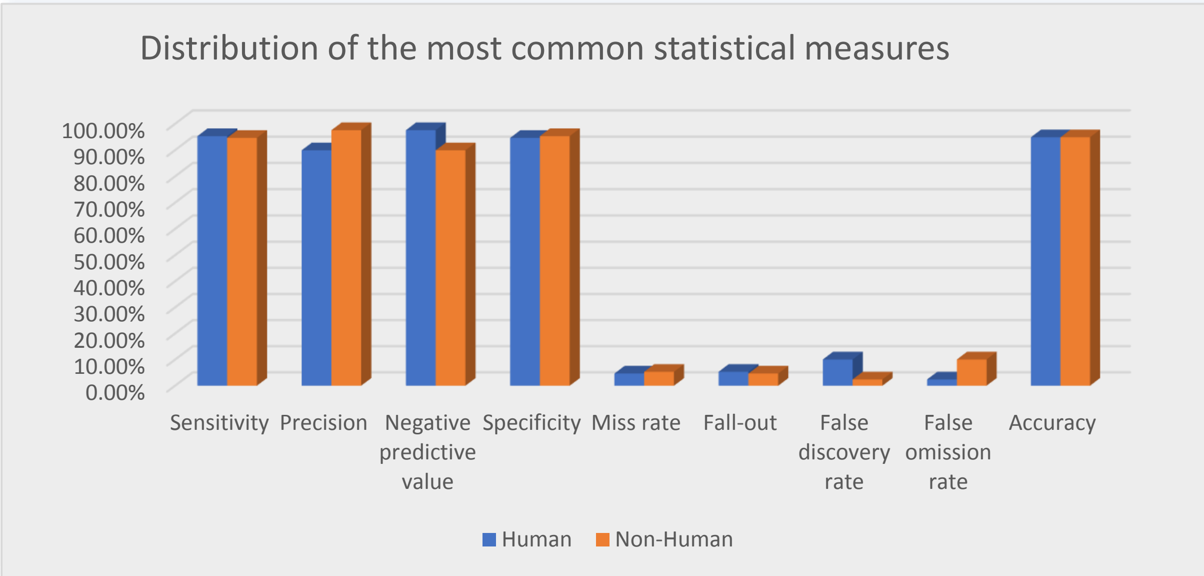


Figure 4.8 The distribution result of the statistical analysis for the performance of the proposed approach for experiment number two.

Experiment two's result for the performance of the proposed approach based on the performance statistical analysis as shown in Table 4-6, and figure 4.10 indicates that the proposed approach's accuracy to classifying the detected object as human or monkeys is 64.58% with a miss -detection rate of about 35%, and the false discovery rate for human class at 21.8% and 52.33% for the monkeys class. This result shows the decrease of the accuracy of the proposed approach compared with the accuracy result in experiment number one, this change of accuracy was acquired when the proposed approach was performed using the most challenging data set samples which are monkeys objects, as a monkeys shape is very similar to the human shape and the classifier of the proposed approach is a shape based classifier.

The low accuracy rate of the proposed approach in terms of classifying the detected object as a human or monkey, does not achieve the highest efficiency compared with other approaches. For this reason, we performed a deep analysis for the performance of the proposed approach spatially for the mathematical model classifier in order to improve the performance of the proposed approach classifier and to acquire state of the art efficiency.

4.4 Modify the proposed approach

After performing experiment number one for the proposed approach, the results of the experiment demonstrates the success of the proposed approach in terms of classifying the detected object as human or non-human. However, it does not achieve as high efficiency compared with state of the art approaches. For this reason, a deep analysis for the performance of the proposed approach spatially was performed for the falsely detected objects in order to improve the performance of the proposed approach to acquire state-of-the-art efficiency.

As a result of the analysis the false detected objects, for example, humans classified as non-human (False Negative) or non-human classified as human (False Positive), it was found that there are some non-human objects classified as human in the proposed approach. These objects are falsely classified because they have a similar shape and the classifier of the proposed approach performs based on the shape of the object. This therefore, requires another experiment in order to discover the weakness in the proposed approach classifier.

In experiment number two, a dataset sample for human and monkeys was selected, to evaluate the performance of the proposed approaches' challenges in cases that have dataset samples very similar in shape, in order to obtain the weakness of the proposed approach classifier. After performing experiment number two which have human and monkey objects as shown in section 4.3.2, the result of experiment number two shows the proposed approach accuracy to classify the detected objects as human or monkey is 64.58% which is ultimately very low and the miss-detection rate is very high about at about 35%, and the false discovery rate for human class is 21.8% and 52.33% for the monkeys class. This result demonstrates the decrease of accuracy in the proposed approach compared with the accuracy result in experiment number one, this change of accuracy proves the weaknesses of the proposed approach classifier spatially when the proposed approach preforms using the most challenging data set samples which are monkeys, because the monkey shape is very similar to the human shape and the classifier of the proposed approach is a shape based classifier. Figure 4.11shows an example of a false detection acquired by the proposed approach.

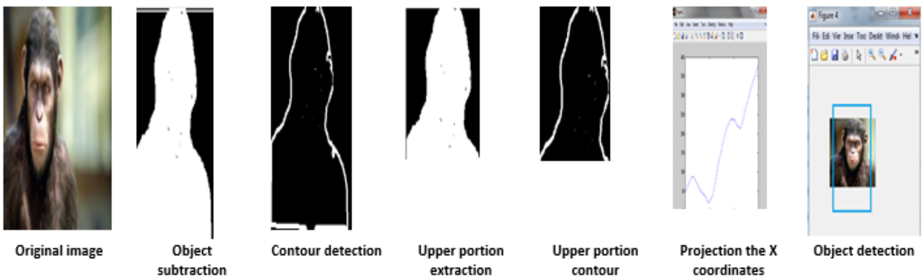


Figure 4.9 An example of false positive detection acquired by the proposed approach

As we can see from Figure 4.11 the classifier of the proposed approach acquired a false positive detection by classifying the monkey object as a human object by surrounding it with a blue rectangle. Another example of the false detection of the proposed approach can be seen In Figure 4.12, where the classifier of the proposed approach acquired a false negative detection by classifying the human object as a non-human object by leaving it without a blue rectangle around it.

Figure 4.10 An example of false negative detection acquired by the proposed approach

For the above reason, the aim is to make a thorough analysis for the proposed approach classifier which is based on its own mathematical model, in order to extract and obtain the weakness of the classifier which causes false detection, in order to make the suitable modifications for the classifier-mathematical model to increase detection performance of the proposed approach.

As mentioned in chapter three, the classifier of this proposed approach is based on its own mathematical model, this mathematical model has four parameters, each one of these parameters conduct its value based on some geometrical calculation for the upper portion of the object shape. In terms of obtaining the limitation and weakness of the proposed approach and to increase the performance accuracy, an analysis of the inner process of the classifier was conducted to obtain the value of each mathematical model parameters for each false detection case. As a result of this thorough analysis, we found that the classifier of this approach based on its own mathematical model provides a false detection result in cases were the objects upper portion shape has the similarity of human upper portion shapes such as. To solve this issue a sensitive analysis for human and monkey's upper portion shapes was conducted to extract the differences between these objects shapes based on the result of each mathematical model parameters. This analysis leads us to find some features that can correspond to human upper portion shapes and not correspond to other object upper portion shapes such as monkeys, this can enhance the performance of the proposed approach classifier by making some changes in the mathematical model parameters.

4.4.1.1 Threshold modifies

After analysing the results of experiment number two, some features were found to be able to correspond to human upper portion shape and not correspond to other object upper portion shapes such as monkeys, and in term of enhancing the performance of the proposed approach classifier a suitable and simple change needs to be modified on the mathematical model parameters by adding a specific value (threshold) for some of the mathematical model parameters, figure 4.13 shows the four parameters of the mathematical model

Parameter 1:	The number of upper peak point = 2, The number of lower peak point = 2
Parameter 2:	The distance D1 < the distance D2
Parameter 3:	The distance D3 = the distance D4 ± (ts1), | D3 – D4| = ts1 where the ts1 is a threshold.
Parameter 4:	The distance D2 > 1/3 of distance D5 And The distance D2 < 2/3 of distance D5, $\frac{1}{3}D5 < D2 < \frac{2}{3}D5$

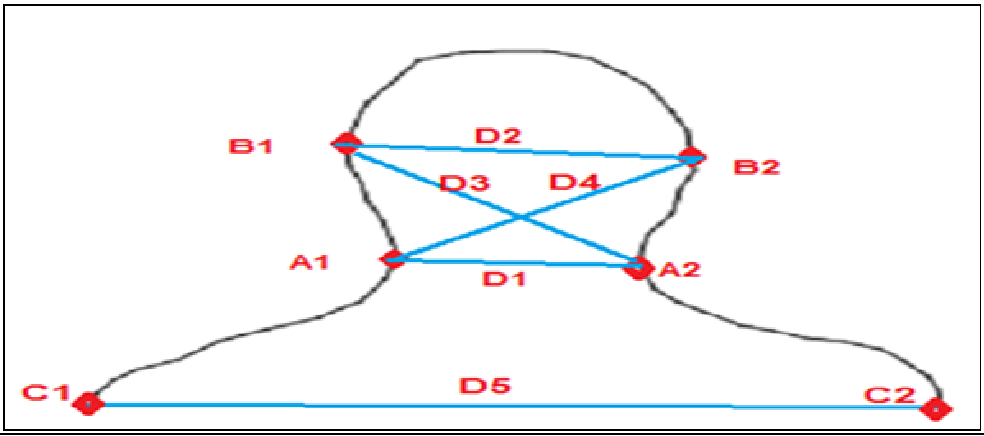


Figure 4.11 The four parameters of the mathematical model

4.4.1.1.1 Thresholding Parameter number two

As we can see from Figure 4.13 parameter number two indicates that the distance D1 is less than the distance D2, and this is true for humans, but after analysing the inner result of experiment number two it was found in the monkey case this is true as well, but because there is no specific values of difference between distance D1 and distance D2. After a sensitivity analysis based on several tests, It was found that parameter number two can be modified by adding a specific value threshold which corresponds to human shape and does not correspond to any other object shape even the monkey's shape. The threshold can be calculated by the 4.10 formula

$$TH_1 = (\sum(D1 + D2)/2) - D1 \tag{4.10}$$

The parameter number two is modified o that

the distance D1 is less than the distance D2 for at least TH_1

The formula for the parameter number two can be presented as

$$D2 - D1 >= TH_1 \tag{4.11}$$

4.4.1.1.2 Thresholding parameter number three

As we can see from Figure 4.13 parameter number three indicates that the distance D3 is equal to the distance D4 with a small difference mentioned by the threshold ts1 (where ts1 calculated by 10% of the distance D3), this is true for humans however, after analysing the inner result of experiment number three it was found in the case of the monkey that this is true. After a sensitivity analysis based on several tests, it was found that the threshold of parameter number three can be modified by changing the specific value of the threshold ts1, this change corresponds to human shape and does not correspond to another objects shape even the monkey's shape, the threshold can be changed as presented in Formula 4.12.

The threshold $(TH_2) = |D3-(\sum(D3 + D4)/2) + D3 * 2\%|$ (4.12)

Parameter number three is modified to be the same distance that D3 is equal to the distance D4 ± TH_2. The formula for the parameter number three is presented as follows

$$|D3\text{-}D4| = 0 \pm TH_2 \qquad (4.13)$$

These changes in the mathematical model parameters can enhance the performance of the proposed approach classifier which leads to increased accuracy.

4.4.1.2 Weighting modifies

The result of experiment number two shows that there is a high percentage of false detection in this proposed approach, by analysing the inner process of the mathematical model for the proposed approach it was found that there are some parameters classifying the detected object as a human (have true value), but in fact this object is not human. For this reason and based on several tests and analysis, it was observed that the proposed approach can enhance the classifier rather than the thresholding modifiers by providing the mathematical model parameters with different weight values.

As mentioned in chapter three the proposed approach classifier is based on its own mathematical model and can classify the detected object as a human, if the results of all the parameter values for the mathematical model are all true otherwise, the detected object will be classified as non-human. This means that all the parameters have the same weight, but after analysing the false detection cases, it was observed that the classifier can be modified by providing different weight values for the mathematical model parameters, these weight values can be varied for the parameters based on analysing the true detection performance and the false detection performance. The result of analysing the performance of detection it was observed that the four parameters must have the true value to classify the human from other objects however, for the objects that have a very similar shape such as monkeys, the classifier algorithm needs to be modified by adding a weight scale condition rather than the first condition, the value of these parameters must be true. This weigh condition will be a constant value for parameters one and four and are rare for parameters two and three.

4.4.1.2.1 Weighting parameter number two

As mentioned above, to improve the detection accuracy of the proposed approach the proposed approach classifier must be modified based on providing some of its parameters with a thresholding value, this threshold value can specify the human object from other objects.

For effective accuracy, we aimed to modify the classifier algorithm by adding a weight-based decision function for the classifier model parameters.

As described in chapter three and as seen in Figure 4.13, parameter number two indicates that the distance D1 is less than the distance D2, and because there is no specified value of the difference between the distance D1 the distance D2, the parameter was modified by adding a specific value threshold which is corresponding to human shape and does not correspond to another objects shape even the monkey's shape, this threshold can then be calculated by the Formula 4.10.

$$TH_1 = (\sum(D1 + D2)/2) - D1 \quad (4.10)$$

Which means that, the distance D1 is less than the distance D2 for at least TH_1

The formula for the parameter number two can be presented as

$$D2 - D1 >= TH_1 \quad (4.11)$$

By modifying parameter number two with a specific threshold the classifier performance is enhanced. oHwever, as we can see in formula 4.10, the threshold TH_1 will have a range of values based on the subtracted average value of the summation for (D1, D2) the distance D1. For development this parameter is given a weight value, this weight is based on the value of the threshold TH_1.However, from the extensive testing, it was observed that the accuracy of detecting human increases when the value of TH_1 is higher therefore, the parameter is provided with a weight based on the value of its threshold TH_1. The weight value can be given by using the following formula.

$$W_2 = S + (S * TH_1) / S \quad (4.14)$$

Where

S is a static value for all parameters; TH_1 is the parameter number two threshold.

From the formula 4.14, it can be seen that the weight value for parameter number two will be high if the value of the TH_1 is high, likewise the weight value for parameter number two will be high if the difference between distance D1 and the distance D2 is high, and the weight value will be small if the difference between distance D1 and distance D2 is small.

4.4.1.2.2 Weighting parameter number three

After parameter number two is given a specific weigh value based on its threshold value, in order to improve the detection accuracy of the proposed approach the aim is to provide parameter number three with a specific weight value as well.

As described in chapter there and as can be seen from figure 4.13, parameter number three indicates that distance D3 is equal to distance D4 with a small difference mentioned by the threshold ts1 (where ts1 calculated by 10% of the distance D3). After sensitive analysis based on several tests, it was found that the threshold of parameter number three can be modified by changing the specific value of the threshold ts1, this change corresponds to human shape and does not correspond to any another object shape even the shape of the monkey's, the threshold can be changed as presented in Formula 4.12.

The threshold $(TH_2) = |D3-(\sum(D3+D4)/2)+D3*2\%|$ (4.12)

Parameter number three modified to be as

The distance D3 is equal to the distance D4 ± TH_2

The formula for parameter number three can be presented as follows.

$|D3-D4| = 0 \pm TH_2$ *(4.13)*

By modifying parameter number three with a specific threshold the classifier performance can be enhanced, but as seen in Formula 4.12, the threshold TH_2 will have a range of values based on the subtraction between the distance D3 and the average value of the distances (D3, D4) ± 2% of the distance D3. For more enhancement this parameter is given a weight value, this weigh is based on the value of the threshold TH_2, however, from the extensive testing it was observed that the accuracy of detecting humans increases when the value of TH_1 is smaller therefore, parameter is provided with weight based on the value of its threshold TH_2. The weight value can be given by the following formula.

$$W_3 = S + (S/TH_2) \quad (4.15)$$

Where

S is a static value for all parameters; TH_2 is parameter number three threshold.

From formula 4.15, the weight value for parameter number three will be indicated as high, if the value of the TH_2 is small the weight will be lower if the value of TH_2 is high. in Similarly, the weight value for parameter number two will be high if the difference between distance D1 and distance D2 is small, and the weight value will be small if the difference between distance D1 and distance D2 is high.

4.4.1.3 Weight-based decision

Improving the detection accuracy of the proposed approach, the classifier is modified based on providing some of its parameters with a thresholding value, these threshold values can specify the human object from other objects in terms of increasing the classifier accuracy. Similarly, for a more efficient accuracy, the aim to modify the classifier algorithm is done by adding a weight-based decision function for the classifier model parameters, the weight-based decision function formula can be given as follows.

$$\mathbf{F_{(w)} = P_1W_1 + P_2W_2 + P_3W_3 + P_4W_4} \quad \mathbf{(4.16)}$$

Where

P_1, P_2, P_3, P_4 are the classifier parameters

W_1, W_2, W_3, W_4 are the respective weight for the classifier parameters

The classifier of the proposed approach can classify the detected object as human if

$\mathbf{F_{(w)} >= F_{(th)}}$ where $\mathbf{F_{(th)}}$ is the critical weight threshold.

The improvement of the proposed approach by indicating the corresponding modifies for the classifier model parameters is shown in table 4-8.

Table 4-7 The improvement of the proposed approach by indicate the corresponding modifies for the classifier parameters

Constrictors of the proposed object classification:	
Parameter 1:	The number of upper peak point = 2, The number of lower peak points = 2 The weight value W_1= S where is S is a critical weight value
Parameter 2:	The distance D1 < the distance D2 for at least TH_1 D2-D1 >= TH_1 The weight value W_2 =S+(S*TH_1) / S where is S is a critical weight value, TH_1 is a respective threshold.
Parameter 3:	The distance D3 = the distance D4 ± TH_2 $\lvert D3-D4 \rvert = 0 \pm TH_2$ where the TH_2 is a threshold. The weight value W_3 = S+(S/TH_2) where is S is a critical weight value, TH_2 is a respective threshold
Parameter 4:	The distance D2 > 1/3 of distance D5 And The distance D2 < 2/3 of distance D5, $\frac{1}{3}\mathrm{D5} < D2 < \frac{2}{3}\mathrm{D5}$ The weight value W_4= S where is S is a critical weight value

After modifying the classifier parameters of the proposed approach, by specifying threshold values for some of the classifier parameters and by providing a weight-based decision function for the classifier parameters, the classifier of the proposed approach based on its mathematical model parameters can classify the human object by acquired two conditions. If the result of all parameters are true and if the value of the weight based-decision function is more or equal to the $\mathbf{F_{(th)}}$ the critical weight threshold.

To evaluate the performance of the proposed approach in classifying the human object from other objects after executing the suitable modifies, the aim is to reperform the previous two experiments (experiment number one and experiment number two) in order to indicate the changes of the proposed approach accuracy results.

4.5 Experiments after modifying the proposed approach

After modifying the classifier parameters of the proposed approach by specifying threshold values for some of the classifier parameters and by providing a weight-based decision function for the classifier parameters. The aim to evaluate the performance of the proposed approach in classifying the human object from other objects, after executing the suitable modifiers by reperforming the previous experiments number one and two in order to indicate the changes on the proposed approach's accuracy results.

4.5.1 Experimental number three

In this experiment the same dataset used in experiment number two was used, in order to reperform the experiment number two to evaluate it after the modifications. Figure 4.14 shows some samples of objects in the dataset.

Figure 4.12 Some samples of objects in the dataset

The performance analysis of this experiment was conducted in two tests: accuracy matrix, and confusion matrix. The experiments based on 240 images in total where divided into two classes, the human class which contains 160 images and the non-human object's class which contains 80 images, the non-human object's class contains images for a variety of different kinds of monkeys. Figure 4.15 and Figure 4.16 show samples of the experimental result of these steps, where the proposed approach classifies the human object by surrounding the detected human object with a blue rectangle as shown in Figure 4.16.

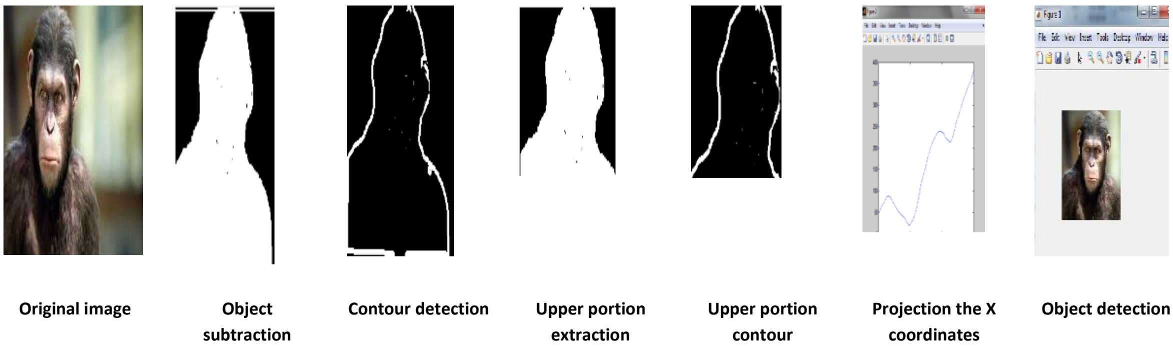


Figure 4.13 The experimental result of the sequence steps for Non-human object

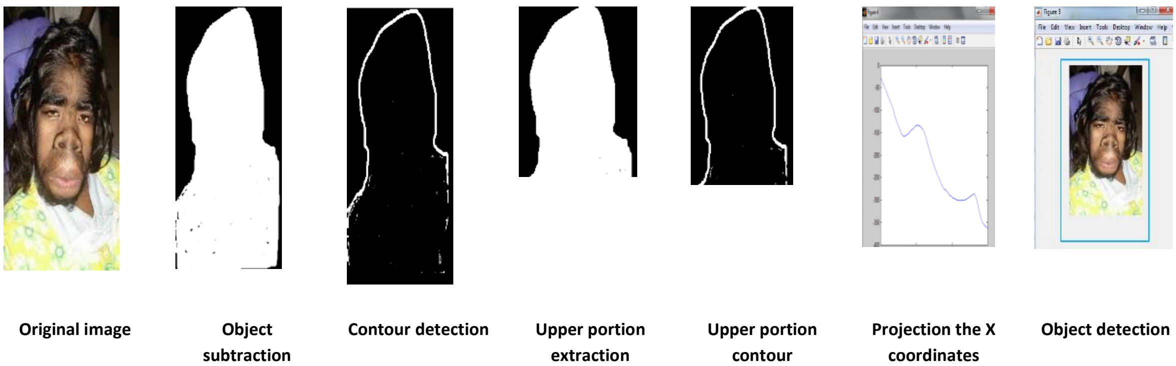


Figure 4.14 The experimental result of the sequence steps for human object

In this experiment the previous experiment number two was reperformed to evaluate the proposed approach after modification, this experiment is based on the selected samples of images which contains 240 images in total divided into two classes, the human class which contains 160 images and the non-human objects (monkeys only) class which contains 80 images. After implementing this experiment, the experimental result of the detection performance can be presented in a confusion matrix. The confusion matrix test of the proposed approach for this experiment is reported in Table 4.8.

Table 4-8 The confusion matrix test for the proposed approach in experiment number three

a	b	← Classified as		
67	13	a	=	Monkeys
8	152	b	=	Human

From the confusion matrix table, we can obtain the performance accuracy of the proposed approach in this experiment. The confusion matrix accuracy is given by Equation 4.1:

$$Accuracy = \frac{\sum \text{Diagonal Sample of confusion matrix}}{\text{Total Sample}} \tag{4.1}$$

By performing Equation 4.1 to calculate the performance accuracy of the proposed approach in this experiment, it was found that the proposed approach detected 67 objects as non-human (monkeys) of 80, and detects 152 objects as human of 160. The overall performance detection accuracy in this experiment is 219 of 240, where 219 is $\sum$ Diagonal Sample of the confusion matrix, and 240 is the total number of dataset samples, meaning the accuracy of this approach in this experiment is equal to 91.25%.

From the confusion matrix report, the performance analysis of the proposed approach indicates that for the human class 152 images from 160 actual images of humans where detected as truly human (True positive), and 8 images from the actual number of human images) are detected as non- human (False negative), while 67 images from the actual number of 80 images of non-humans where detected as non-human (True Negative), and 13 images from the actual number of non-human images where detected as human (False Positive).

For the non-human (monkey) class 67 images from 80 images of actual non-humans images where detected as truly non-human (True positive), and 13 images from the actual number of non-humans images where detected as human (False negative), while 152 images from 160 images of the actual number of humans where detected as truly human (True Negative), and 8 images from the actual number of humans where detected as non- human (False Positive). Table 4.10 shows a summary of the performance analysis for the proposed approach in this experiment for each class.

Table 4-9 A summary for the performance analysis of the proposed approach in experiment number three.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**152**	**13**	**67**	**8**
Monkeys	**67**	**8**	**152**	**13**

4.5.1.1 Statistical analysis of the performance

The statistical analysis for the performance of the proposed approach in this experiment provides a full description for the validity of the proposed approach in classifying the detected object as human or non-human. From this description the influence of this approach can be seen from the positive classification of the detected objects and the weakness of this approach in terms of the false classification of the detected objects. The results of the statistical analysis of the performance of the proposed approach is summarised in table 4-10.

Table 4-10 The summarise result of the statistical analysis for the performance of the proposed approach

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	95%	92.12%	89.33%	83.75%	5%	16.25%	7.87%	10.66%	91.25%
Monkeys	83.75%	89.33%	92.12%	95%	16.25%	5%	10.66%	7.87%	91.25%

From Table 4-10, the distribution results of the statistical analysis is presented, while the performance of the proposed approach for this experiment is shown in Figure 4.18.

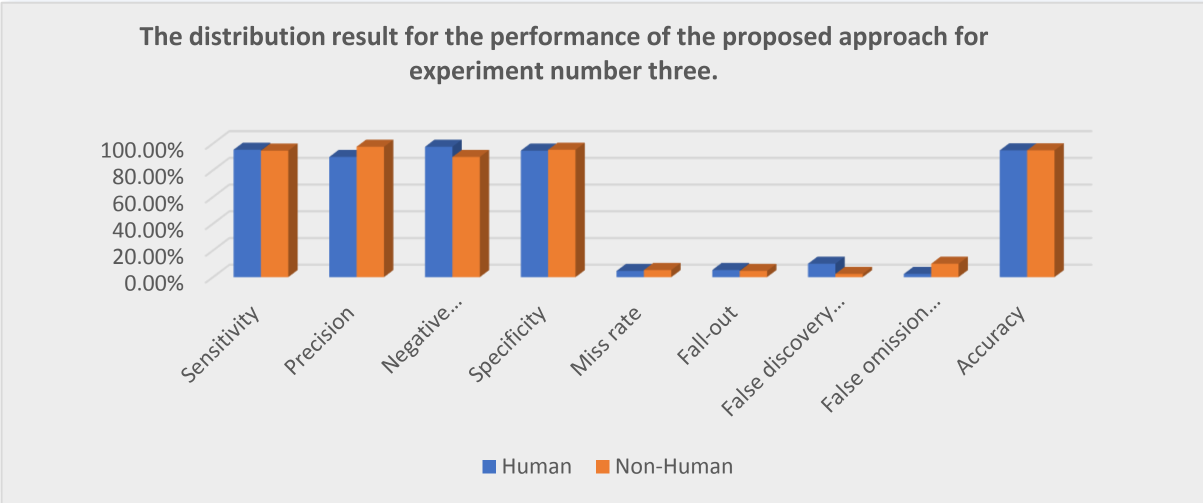


Figure 4.15 The distribution result of the statistical analysis for the performance of the proposed approach for experiment number three.

The results of experiment number three of the proposed approach based on the performance statistical analysis is shown in Table 4-11 and figure 4.18. It indicates that the proposed approach's accuracy to classifying the detected object as human or monkey is 91.25% with a miss-detection rate of about 10.66%, and the false discovery rate for the human class is 7.87% and 10.66% for the monkey class, while the accuracy in experiment number two is 64.58% with a miss-detection rate of about 35%, and the false discovery rate for human class at 21.8% and 52.33% for the monkey class.

This result shows a 26.67% increase in the accuracy of the proposed approach after implementing the classifier modification, compared to the accuracy result in experiment number two. Both experiments were performed under the same dataset and processor characteristics. This change of accuracy was acquired when preforming the proposed approach after modifying the classifier by making the appropriate changes to the threshold values of the classifier parameters and by applying the weight-based decision function for the classifier parameters.

This accuracy rate of the proposed approach in terms of classifying the detected object as human or non-human indicate the efficiency improvement of the classifier performance compared with the accuracy rate for experiment number two.

4.5.2 Experimental number four

In this experiment, the aim was to re-evaluate the performance of the proposed approach in classifying human objects from other objects after executing the suitable modifiers by reperforming previous experiment number one, in order to see the changes on the accuracy results of the proposed approach. Figure 4.19 shows samples of objects taken from the dataset

.

Figure 4.16 Some sample of objects in the dataset

The performance analysis of this experiment was conducted based on 450 images in total divided into two classes, the human class which contained 150 images and the non-human object's class which contained 300 images. The non-human object's class contains images for varied types of objects such as monkeys, horses, dogs, cars, and other types of non-human objects. Figure 4.20 and Figure 4.21 show some samples of the experimental results of these steps.

Figure 4.17 The experimental result of sequence steps for human object

Figure 4.18 The experimental result of sequence steps for Non-human object

As shown in Figure 4.20, the classifier of the proposed approach classifies the detected object as human by surrounding the human object with a blue rectangle. After implementing the experiment based on the selected samples of images, the confusion matrix test of the proposed approach for this experiment is reported in Table 4.11.

Table 4-11 The confusion matrix test for the proposed approach in the experiment number four

a	b	← Classified as

284	**16**	**a**	**=**	**Non-Human**
7	**143**	**b**	**=**	**Human**

From the confusion matrix table, it was found that the proposed approach detects 284 objects as non-human from 300, and detects 143 objects as human from 150, the overall performance detection accuracy in this experiment is 427 of 450, where 427 is ∑ Diagonal Sample of the confusion matrix, and 450 is the total number of dataset samples, meaning the accuracy of the approach within this experiment is equal to 94.88%.

From the confusion matrix report, the performance analysis of the proposed approach, indicates that for the human class 143 images from 150 of the actual number of humans where detected as truly human (True positive), and 7 images from the actual number of human images where detected as non- human (False negative), while 284 images from 300 images of of non-humans where detected as truly non-human (True Negative**)**, and 16 images from the actual number of non-human images where detected as human (False Positive).

For the non-human class 284 images from 300 images of actual non-humans where detected as truly non-human (True positive), and 16 images from the actual number of non-human images where detected as human (False negative), while a 143 images from 150 images of the actual number of humans where detected as truly human (True Negative), and 7 images from the actual number of human images are detected as non- human (False Positive). Table 4.12 shows a summary for the performance analysis of the proposed approach in this experiment for each class.

Table 4-12 A summary for the performance analysis of the proposed approach in experiment number four.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	143	16	284	7
Non-Human	284	7	143	16

4.5.2.1 Statistical analysis of the Performance

In this section the same functions used in the previous experiments where used for a deep analysis of the proposed approach's performance, more details of these functions are shown in table 4-3. The result of the statistical analysis of the performance of the proposed approach is summarised in table 4-13

Table 4-13 The summarise result of the statistical analysis for the performance of the proposed approach

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	95.33%	89.93%	97.59%	94.66%	4.66%	5.33%	10.06%	2.4%	94.88%
Non-Human	94.66%	97.59%	89.93%	95.33%	5.33%	4.66%	2.4%	10.06%	94.88%

Table 4-13 presents the distribution results of the statistical analysis, and the performance of the proposed approach for this experiment is shown in Figure 4.19.

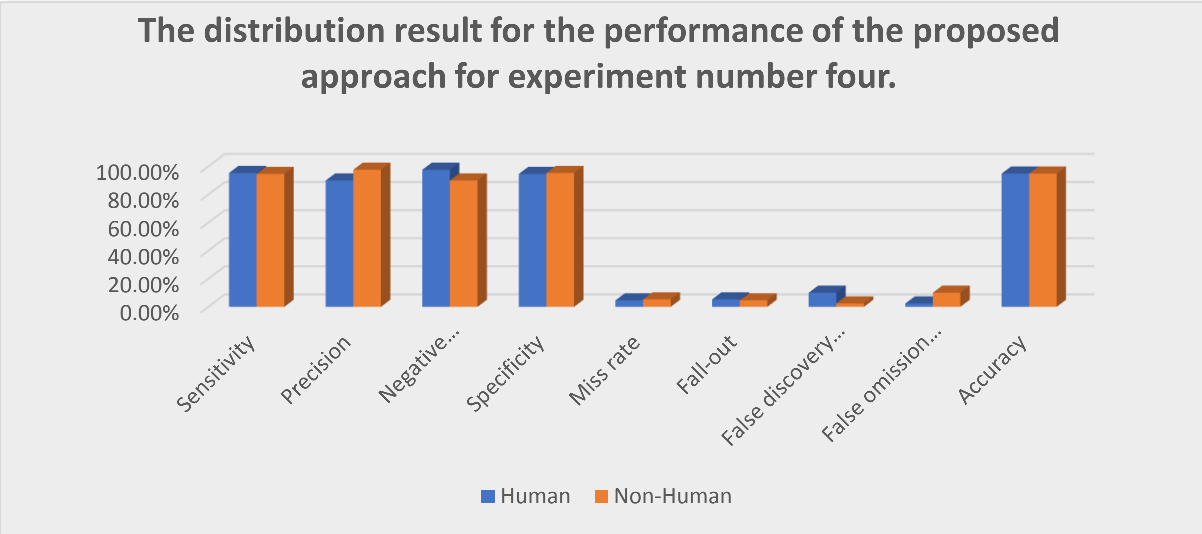


Figure 4.19 The distribution result of the statistical analysis for the performance of the proposed approach for experiment number four.

The results of experiment number four are based on the performance statistical analysis as shown in Table 4-13, Figure 4.23 shows that the proposed approach's accuracy in classifying the detected objects as human or non-human is at 94.88% with a miss detection rate of about 5%, and false discovery rate for the human the class is 10.06% and 2.4% for the non-human class. While the accuracy in experiment number one was 86.22% with a miss- detection rate of about 14%, and the false discovery rate for the human class being 24.11% and 7.5% for the non-human class.

This result shows 8.66% increase in the accuracy of the proposed approach after implementing the classifier modification compared with the accuracy result in experiment number one, both experiments where performed under the same dataset and processor characteristics. This change in the accuracy was acquired when the proposed approach was performed after modifying the classifier by making the appropriate changes to the threshold values of the classifier parameters and by apply the weight-based decision function for the classifier parameters.

This accuracy rate of the proposed approach in terms of classifying the detected object as human or non-human indicates the improvement in efficiency of the classifier performance compared with the accuracy rate for experiment number one.

4.6 Conclusion

In this chapter, a description of the main experiments carried out with the proposed approach to classify the detected objects in different positions as human or non-human on INRIA and Caltech 101 datasets (set of human and non-human as an object in digital images) in order to know the accuracy level. The proposed approach was implemented in these experiments using MATLAB R2017b and tested on 1.8 GHz core i7 (IV), 16 GB memory and 512 GB hard drive.

Experiment number one was performed based on 450 images in total collected from INRIA and Caltech 101 datasets, these 450 images where divided into two classes, the human class which contains 150 images and the non-human object's class which contained 300 images. The non-human object's class contained images for various types of objects such as monkeys, horses, dogs, cars, and other types of non-human objects. The results of experiment number one shows that, the proposed approaches accuracy in classifying the detected object as human is 86.22% with a miss- detection rate of about 14%, and a false discovery rate for the human class at 24.11% and 7.5% for the non-human class. This result shows the success of the proposed approach in terms of classifying the detected objects as human or non-human however, it did not achieve the highest level of efficiency compared with other approaches. For this reason, a deep analysis for the performance of the proposed approach spatially was performed for the falsely detected objects in order to improve the performance of the proposed approach to acquire state-of-the-art efficiency.

As the result of the analysis for the falsely detected objects, it was found that there were some non-human objects classified as human in the proposed approach. These objects where falsely classified because they have a similar shape to humans, such as monkeys, and the classifier of the proposed approach performs based on the shape of the object. This lead to another experiment being performed in order to discover the weaknesses in the proposed approach classifier.

Experiment number two was performed on 240 images in total collected from INRIA and Caltech 101 datasets, these 240 images where divided into two classes, the human class which contained 160 images and the non-human objects class which contained 80 images. The non-human objects class contained images for a variety of different kinds of monkeys. The result of experiment number two shows that, the proposed approaches accuracy to classifying the detected object as human is 64.58% with a miss- detection rate of 35%, and a false discovery rate for the human class at 21.8% and 52.33% for the Monkey class. This result shows the decrease in the accuracy of the proposed approach compared with the accuracy result of experiment number one. This change in the accuracy acquired when performing the proposed approach using the most challenging data set samples which are monkey objects, as the monkey shape is very similar to the human shape and the classifier of the proposed approach is a shape-based classifier.

Based on the results of experiment number one and two, the aim was to perform a deep analysis for the proposed approach's classifier which is based on its own mathematical model, in order to extract and obtain the weaknesses of the classifier which results in a false detection, in order to make the suitable modifications to the classifier-mathematical model to increase the approach's detection performance. As a result of this deep analysis, it was found that the classifier for this approach is based on its own mathematical model providing a false detection results in cases where the object's upper portion shape has the similarity of human upper portion shapes. To solve this issue an analysis for human and monkey upper portion shapes was conducted in order to extract the differences between these objects shapes based on each result of the mathematical model parameters. This analysis lead to modifications being made to the classifier of the proposed approach by finding some features that are able to correspond to human upper portion shapes and not correspond to other objects upper portion shapes such as monkeys, this can enhance the performance of the proposed approach's classifier by making some changes to the mathematical model parameters.

After modifying the classifier parameters of the proposed approach, by specifying threshold values for some of the classifier parameters and by providing a weight-based decision function for the classifier parameters, the classifier of the proposed approach based on its mathematical model parameters can classify the human object by acquiring two conditions, if the result of all parameters are true and if the value of the weight based-decision function is more or equal to $\mathbf{F}_{\mathbf{(the)}}$ the critical weigh threshold.

To evaluate the performance of the proposed approach in classifying the human object from other objects after executing the suitable modifications, the aim was to reperform the previous two experiments (experiment number one and experiment number two) in order to show the changes to the proposed approach's accuracy results.

In experiment number three experiment number two was reperformed using the same dataset and processor characteristics, the results of experiment number three for the performance of the proposed approach shows that, the proposed approach's accuracy when classifying the detected objects as human or monkeys at 91.25% with a miss- detection rate of about 10.625%, and false discovery rate for the human class at 7.87% and 10.66% for the monkey class. Meanwhile, the accuracy in experiment number two was 64.58% with a miss-detection rate of about 35%, and a false discovery rate for the human class at 21.8% and 52.33% for the monkey class.

This result shows a 26.67% increase in the accuracy of the proposed approach after implementing the classifier modifications compared with the accuracy results in experiment number two. This change of accuracy was acquired when performing the proposed approach after modifying the classifier by making the appropriate changes to the threshold values of the classifier parameters and by applying the weight-based decision function to the classifier parameters. This accuracy rate of the proposed approach in terms of classifying the detected object as human or non-human indicates the efficiency of the improvements to the classifier's performance compared with the accuracy rate for experiment number two.

In experiment number four, experiment number one was reperformed using the same dataset and processor characteristics, the result of experiment number four for the performance of the proposed approach shows that, the proposed approach's accuracy when classifying the detected objects as human or non-human is 94.88% with a miss- detection rate of about 5%, and the false discovery rate for the human class at 10.06% and 2.4% for the non-human class. While the accuracy in experiment number one was 86.22% with a miss- detection rate of about 14%, and the false discovery rate for the human class being 24.11% and 7.5% for the non-human class.

This result shows an 8.66% increase in the accuracy of the proposed approach after implementing the classifier modification compared with the accuracy result in experiment number one. This change in the accuracy was acquired when performing the proposed approach after modifying the classifier by making the appropriate changes to the threshold values of the classifier parameters and by applying the weight-based decision function for the classifier parameters.

The accuracy rate of the proposed approach in terms of classifying the detected objects as human or non-human indicate an improvement in the efficiency of the classifier's performance compared with the accuracy rate for experiment number one. Table 4-15 shows a comparison between these four experiments results.

Table 4-14 A comparison between the four experiments results.

	Total number of Instances	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy	Average of Miss rate	Average of False discovery rate
Experiment #1	450	388	62	86.22%	13.83%	15.8%
Experiment #2	240	155	85	64.58%	35.62%	37.06%
Experiment #3	240	219	21	91.25%	10.625%	9.265%
Experiment #4	450	427	23	94.88%	4.99%	6.23%

From Table 4-15 the improvement of the proposed approach classifier can be seen after the modifications where applied. Experiment number one and two where performed before the modification where made while experiments number three and four where reperformed from experiment number one and two after the modifications. From the comparison between experiment number one and four (same dataset) it can be seen that the increasing accuracy rate is 8.66%, while the decreasing average of miss- detection rates are 8.84%, and the decreasing average of the false discovery rate is 9.57%. Similarly, from the comparison between experiment number two and three (same dataset) it can be seen that the increasing accuracy rate is 26.67%, while the decreasing average rate of miss- detection is 24.995%, and the decreasing average of the false discovery rate is 27.795%.

The experimental results show the improvements of the classification accuracy, and indicates that the proposed approach is efficient in classifying humans from other objects, even with objects that have a similar shape such as monkeys. For the global evaluation and validation in terms of performance, complexity, and accuracy, this new approach will be compared in the next chapter with other global and state of the art approaches using the same environment and datasets.

CHAPTER FIVE

5 EVALUATION

5.1 INTRODUCTION

In this chapter, the aim is to evaluate the performance of the proposed approach in classifying the detected objects as human or non-human by comparing it with some of the common machine learning approaches such as support vector machine, random forest, and artificial neural network. The machine learning approaches are widely used in areas of image processing and computer vision for evaluating researcher approaches. The evaluation is carried out based on performing the proposed approach and the common machine learning approaches using the same dataset and under the same processor conditions.

5.2 DATA ACQUISITION

For evaluating the proposed approach it was performed as well as some of the common machine learning approaches, such as support vector machine, random forest, and artificial neural network using different images for several objects obtained from INRIA dataset[70] (set of human and non-human as an object in digital images) in order to distinguish the accuracy level. The selected images were labelled and contained a single object in different camera poses and viewpoints.

The INRIA contained two group formats, the original images and the positive images, cropped in different sizes such as 64×128 pixels, and $214 \times 320 - 648 \times 486$ pixels [41]. For a homogeneous dataset (same size of pixels) the selected images where cropped into 64 x 128 pixels using an image cropper approach [104].

5.3 MACHINE LEARNING APPROACH

The machine learning approaches are widely used in areas of image processing and computer vision and play a major rule in evaluating and measuring the accuracy of researcher approaches. The process for the machine learning approach have two stages, the first stage being a learning stage (training) and the second stage being for testing and evaluating.

For the training stage the machine needs to learn from the dataset by splitting it into two classes, (human and non-human) and then extracting features for each class. Based on these features the machine will find a statistical relationship (i.e. Statistical regularities) for each class, this statistical relationship will learn from the machine in the future to detect and classify the objects based on the different dataset classes. After educating the machine to classify the objects based on extracting spatial features for each dataset class, the machine is then be able to classify the objects in terms of dataset classes, and it is ready to be used in testing and evaluating the classification using any new samples within the same dataset classes. Figure 5.1 shows the general architecture of the machine learning approaches.

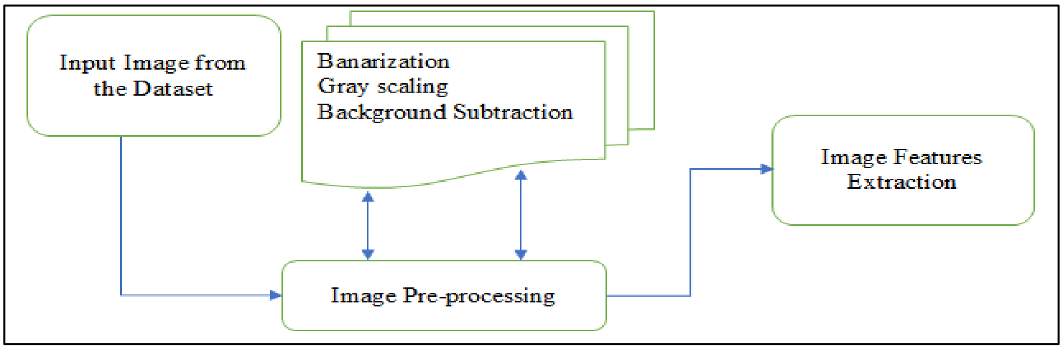


Figure 5.1 The general architecture of the machine learning approaches

In evaluating the proposed approach and comparing it with machine learning approaches, 358 images where obtained from the INRIA dataset, and then 150 images where selected randomly from the obtained 358 images. These 150 selected images are divided into two classes; human class and non-human class, they are then used to learn the machine approaches in terms of extracting spatial features for each class in order to let the machine classify the object as human or non-human, then they perform the experiment on the 358 images obtained. For learning machine approaches,11 optimal features where selected for each class as shown in Table 5.1. In this work, the machine learning toolkit (WEKA) is used for testing and training to increase the classification accuracy level [69] [105] [106].

Table 5-1 The optimal 11 features for learning the machine approaches

Feature	Formula	Description
Circularity	$S_{CR} = 4\pi \frac{Area}{perimeter^2}$	The shape of the detected object.
Mean	$S_M = \bar{b} = \sum_{b=0}^{L-1} bp(b)$	The average intensity values of the pixels.
Standard Deviation	$S_D = \sigma_b = \left[\sum_{b=0}^{L-1} ((b - \bar{b}))^2 p(b)\right]^2$	The deviation or the variance between the pixels in the input image.
Contrast	$S_C = \sum_i \sum_j (i - j)^2\, p(i,j)$	The local contrast of the image in terms of gray levels.
Energy	$S_N = \sum_{b=0}^{L-1} [p(b)]^2$	The uniformity of the texture.
SKEWNESS	$S_S = \frac{1}{\sigma_b^3} \sum_{b=0}^{L-1} ((b - \bar{b}))^3 p(b)$	The asymmetry of the probability distribution of a real-valued random variable (i.e. Positive or negative).
Kurtosis	$S_K = \frac{1}{\sigma_b^4} \sum_{b=0}^{L-1} ((b - \bar{b}))^4 p(b) - 3$	The peakedness of the probability distribution of a real-valued random variable.
Entropy	$S_E = -\sum_{b=0}^{L-1} p(b) \log_2\{p(b)\}$	The randomness of a gray level distribution.
Homogeneity	$S_H = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(\frac{P_{ij}}{(1 + \lvert i - j \rvert)}\right)$	The closeness element distribution in GLCM to the diagonal GLCM.
Correlation	$S_o = \frac{\sum_i \sum_j (ij)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	The gray level linear dependence between the relative pixels.
Diameter	$D = 2r$	Diameter of the edge pixels.

As mentioned previously the INRIA dataset was used in this study, and a MATLAB tool was utilised to extract the human (HU) and non-human (NH) feature values as shown in Table 5.2. The experiment was based on 358 images in total as seen in the excel sheet, they were then converted to a CSV file which is identified by WEKA toolkit.

Table 5-2 Human and non-human features values

Class	Contrast	SKWENESS	Kurtosis	Entropy	Mean	STD	Circulation	Energy	Correlation	Homogeneity	Diameter
NH	0.4096313	1.4647412	3.68919358	2.4586934	0.713589	0.10497	0.170500657	0.45427	0.934805284	0.942542972	77.25822421
NH	1.543092	0.6327101	1.73922325	3.2139312	0.601251	0.26009	0.519347116	0.212383	0.915807358	0.88422887	26.96555684
NH	0.3634939	1.52722	3.88071689	2.3389425	0.72396	0.099062	0.268591532	0.634812	0.890879927	0.940385611	25.34866285
NH	0.8382716	0.255393	1.70289185	4.7602913	0.522904	0.261673	0.974510188	0.185554	0.515996847	0.770823045	7.664083887
NH	0.0697674	0.1292439	1.70976018	4.5640954	0.506209	0.280004	0.815340875	0.501085	0.839330855	0.965116279	8.475885149
NH	1.0309966	0.0943804	1.72407625	5.6132643	0.507041	0.282991	0.764440758	0.085073	0.906028085	0.823847803	58.34654096
NH	0.1013036	0.2145565	1.68338518	4.7371289	0.517438	0.268583	0.763436059	0.451547	0.921573758	0.956244784	94.06135716
NH	0.1221504	0.1518394	1.69952556	5.0986174	0.509636	0.278057	0.759209005	0.424858	0.811680164	0.948126907	79.15356731
HU	0.0759527	0.0015947	1.79877976	5.0321695	0.50063	0.293954	0.788735455	0.373156	0.96935851	0.976240812	85.86579201
HU	0.1463341	0.0060509	1.79229953	5.8292385	0.500485	0.293107	0.687587736	0.236915	0.94951322	0.948412698	75.88524314
HU	0.1518828	0.0338739	1.77452882	5.5845508	0.502124	0.290003	0.60910958	0.27102	0.957502599	0.94182095	83.77675236
HU	0.1796053	0.0135222	1.77979291	5.6617798	0.500324	0.292691	0.614101161	0.246377	0.953511134	0.93569349	81.31811464
HU	0.1754386	-0.001136	1.79884816	5.6168371	0.499668	0.293077	0.711405994	0.23134	0.910003404	0.929006544	35.64520799
HU	0.1124818	0.1390353	1.70625791	5.3618748	0.511466	0.27722	0.44672996	0.256375	0.933455605	0.95308311	94.91928442
HU	0.1018853	0.1390269	1.70471295	5.3245961	0.511393	0.276944	0.390734792	0.276487	0.937241685	0.959387494	106.4762046
HU	0.0463755	0.1656096	1.7094672	4.8644615	0.514252	0.274189	0.682144486	0.475984	0.920752753	0.978265257	43.26714196
HU	0.1343467	0.0092822	1.79246547	5.8303768	0.500543	0.292455	0.69110382	0.264841	0.95075864	0.946510865	55.02971645

To describe the distribution of the optimal 11 features for the example images where divided into two classes; human class and non-human class. Figure 5.2 shows this distribution of the selected features.

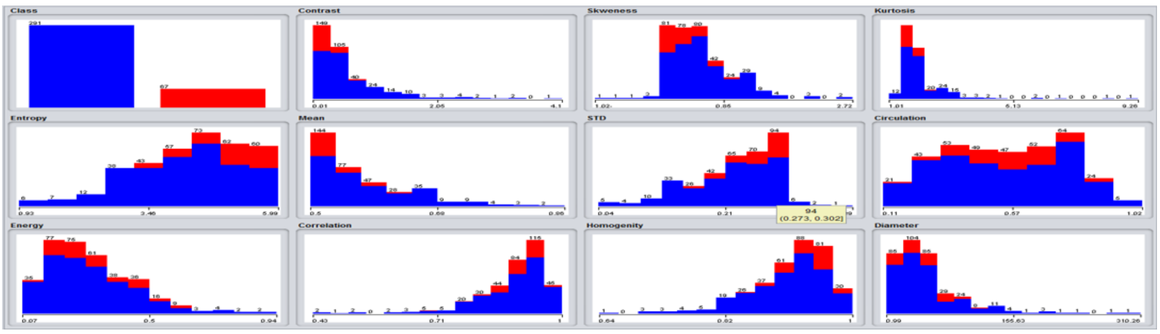


Figure 5.2 The instances distributions of the selected features for human and non-human classes

5.4 PERFORMANCE ANALYSIS

This section provides a performance analysis for this experiment of the proposed approach and some of the traditional machine learning methods such as support vector machine, random forest, and artificial neural network in order to evaluate the classification accuracy for the detected object as human or non-human. The performance analysis of this experiment was performed in two tests: accuracy matrix, and confusion matrix using different images for several objects obtained from INRIA dataset (set of human and non-human as an object in digital images) in order to recognise the accuracy level. The selected images were labelled and contained a single object in different camera poses and viewpoints. The number of selected samples are 358 images in total divided into two classes, the human class which contained 67 images and the non-human object's class which contained 291 images. The non-human object's class contains images for varied types of objects such as monkeys, horses, dogs, cars, and other types of non-human objects, and it was implemented using MATLAB R2017b and tested on a 1.8 GHz core i7 (IV), with a 16 GB memory and 512 GB hard drive.

After implementing the experiment based on the selected samples of images, the experimental result of the detection performance of our proposed approach and the machine learning approach is presented in a confusion matrix. The confusion matrix test for the proposed approach for this experiment is reported in Table 5-3.

Table 5-3 The confusion matrix test for the proposed approach

a	b	← Classified as		
282	9	a	=	NH
13	54	b	=	Hu

The confusion matrix test of the support vector machine approach for this experiment is reported in Table 5-4.

Table 5-4 Describe the Confusion matrix test for SVM based on human and non-human classes

a	b	← Classified as		
287	**4**	**a**	**=**	**NH**
66	**1**	**b**	**=**	**Hu**

The confusion matrix test of the artificial neural network approach for this experiment is reported in Table 5-5.

Table 5-5 Describe the Confusion matrix test for artificial neural network based on human and non-human classes

a	b	← Classified as		
275	**16**	**a**	**=**	**NH**
21	**46**	**b**	**=**	**Hu**

The confusion matrix test of the random forest approach for this experiment is reported in Table 5-6.

Table 5-6 Describe the Confusion matrix test for a random forest approach based on human and non-human classes

a	b	← Classified as		
282	**9**	**a**	**=**	**NH**
27	**40**	**b**	**=**	**Hu**

From the confusion matrix table, the performance accuracy in this experiment can be obtained for the proposed approach and the machine learning approaches. The confusion matrix accuracy is given by Equation 4.1:

$$Accuracy = \frac{\sum \text{Diagonal Sample of confusion matrix}}{\text{Total Sample}} \tag{4.1}$$

By performing the Equation 4.1 to calculate the performance accuracy, it was found that the proposed approach detects 282 objects as non-human of 291, and detects 54 objects as human of 67, the overall performance detection accuracy in this experiment is 336 of 358, where 336 is ∑ Diagonal Sample of the confusion matrix, and 358 is the total number of the dataset samples, meaning the accuracy of the approach in this experiment is equal to 93.85%.

in order to calculate the performance accuracy of the support vector machine approach in this experiment, it was found that the support vector machine approach detects 287 objects as non-human of 291, and detects 1 object as human of 67, the overall performance detection accuracy in this experiment is 288 of 358, where 288 is ∑ Diagonal Sample of the confusion matrix, and 358 is the total number of the dataset samples, meaning the accuracy of this approach in this experiment is equal to 80.446.

The performance accuracy of the artificial neural network approach in this experimentfound that the artificial neural network approach detects 275 objects as non-human of 291, and detects 46 objects as human of 67, the overall performance detection accuracy in this experiment is 321 of 358, where 321 is ∑ Diagonal Sample of the confusion matrix, and 358 is the total number of the dataset samples, meaning the accuracy of the approach in this experiment is equal to 89.664%.

Meanwhile, the random forest approach in this experiment detected 282 objects as non-human of 291, and detects 40 objects as human of 67, the overall performance detection accuracy in this experiment is 322 of 358, where 322 is ∑ Diagonal Sample of the confusion matrix, and 358 is the total number of the dataset samples, meaning the accuracy of the approach in this experiment is equal to 89.944%.

From the confusion matrix report, the performance analysis of the proposed approach, indicates that for the human class 54 images from the actual number of human (67) images were detected as truly human (True positive), and 13 images from the actual number of human (67) images were detected as non- human (False negative), while 282 images from the actual number of non-human (291) images were detected as truly non-human (True Negative), and 9 images from the actual number of non-human (291) images were detected as human (False Positive).

For the non-human class 282 images of the actual number of non-human (291) images were detected as truly non-human (True positive), and 13 images from the actual number of non-human (291) images were detected as human (False negative), while 54 images from the actual number of human (67) images were detected as truly human (True Negative), and 13 images from the actual number of human (67) images were detected as non- human (False Positive). Table 5-7 shows a summary for the performance analysis of the proposed approach in this experiment for each class.

Table 5-7 A summary for the performance analysis of the proposed approach.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**54**	**9**	**282**	**13**
Non-Human	**282**	**13**	**54**	**9**

The performance analysis of the support vector machine approach is based on the confusion matrix report, it indicates that for the human class 1 image from the actual number of human (67) images were detected as truly human (True positive), and 66 images from the actual number of human images were detected as non- human (False negative), while 287 images from the actual number of non-human (291) images were detected as truly non-human (True Negative**)**, and 4 images from the actual number of non-human images were detected as human (False Positive).

For the non-human class 287 images of the actual number of non-human (291) images were detected as truly non-human (True positive), and 4 images from the actual number of non-human images were detected as human (False negative), while 1 image from the actual number of human (67) images were detected as truly human (True Negative), and 66 images from the actual number of human images were detected as non- human (False Positive). Table 5-8 shows a summary for the performance analysis of the support vector machine approach in this experiment for each class.

Table 5-8 A summary for the performance analysis of the support vector machine approach.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**1**	**4**	**287**	**66**
Non-Human	**287**	**66**	**1**	**4**

Based on the confusion matrix report, the performance analysis of the artificial neural network approach, indicates that for the human class 46 images from the actual number of human (67) images were detected as truly human (True positive), and 21 images from the actual number of human images were detected as non- human (False negative). While 275 images from the actual number of non-human (291) images were detected as truly non-human (True

Negative**)**, and 16 images from the actual number of non-human images were detected as human (False Positive).

For the non-human class 275 images from the actual number of non-human (291) images were detected as truly non-human (True positive), and 16 images from the actual number of non-human images were detected as human (False negative). While 46 images from the actual number of human (67) images were detected as truly human (True Negative), and 21 images from the actual number of human images are detected as non- human (False Positive). Table 5-9 shows a summary for the performance analysis of the artificial neural network approach in this experiment for each class.

Table 5-9 A summary for the performance analysis of the artificial neural network approach.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**46**	**16**	**275**	**21**
Non-Human	**275**	**21**	**46**	**16**

From the confusion matrix report, the performance analysis of the random forest approach, indicates that for the human class 40 images from the actual number of human (67) images were detected as truly human (True positive), and 27 images from the actual number of human images were detected as non- human (False negative). While 282 images from the actual number of non-human (291) images were detected as truly non-human (True Negative**)**, and 9 images from the actual number of non-human images were detected as human (False Positive).

For the non-human class 282 images of the actual number of non-human (291) images were detected as truly non-human (True positive), and 9 images from the actual number of non-human images were detected as human (False negative). While 40 images from the actual number of human (67) images were detected as truly human (True Negative), and 27 images from the actual number of human images were detected as non- human (False Positive). Table 5-10 shows a summary for the performance analysis of the random forest approach in this experiment for each class.

Table 5-10 A summary for the performance analysis of the random forest approach.

Class	True Positive TP	False Positive FP	True Negative TN	False Negative FN
Human	**40**	**9**	**282**	**27**
Non-Human	**282**	**27**	**40**	**9**

For a more statistical measure analysis for the performance of the proposed approach and the machine learning approaches, some of the most common statistical measure functions have been calculated for a deep analysis of the proposed approach's performance, such as the Sensitivity, Precision, Negative predictive value, Specificity, Miss rate, Fall-out, False discovery rate, False omission rate, and the Accuracy. These statistical analyses of the performance of the proposed approach provide a full description for the validity of the proposed approach in classifying the detected object as human or non-human. From this description the influence of this approach in terms of positive classification of the detected object and the weakness of this approach in terms of false classification of the detected objects. Each one of these statistical analysis describes the performance of the approach in different aspects in order to evaluate and validate the approach.

The corresponding formula for each one of most common statistical functions can be found in Table 5-11.

Table 5-11 The corresponding formula for the most common statistical functions

Function name	**The Formula of the function**
Sensitivity (True positive rate)	Sensitivity (TPR) =$\frac{TP}{TP+FN}$
Precision (positive predictive value)	Precision (PPV) = $\frac{TP}{TP+FP}$
Negative predictive value	Negative predictive value (NPV) = $\frac{TN}{TN+FN}$
Specificity (True Negative rate)	Specificity (TNR) = $\frac{TN}{TN+FP}$
Miss rate also called False Negative Rate	Miss rate (FNR) = $\frac{FN}{FN+TP}$
The Fall-out also called False Positive Rate	Fall-out (FPR) = $\frac{FP}{FP+TN}$
False discovery rate	False discovery rate (FDR) = $\frac{FP}{FP+TP}$
False omission rate	False omission rate (FOR) = $\frac{FP}{FP+TP}$
The Accuracy	The Accuracy (ACC) = $\frac{TP+TN}{TP+TN+FP+FN}$

After calculating the most common statistical functions for both classes (human and non-human) based on the respective formula for each function presented in Table 5-11, it was found that, for the proposed approach, the sensitivity for the human class is 80.59% and for the non-human class is 88.74%. The precision for the human class is 85.71% and for the non-human class is 95.59%, the Negative predictive value for human class is 95.59% and for the non-human class is 85.71%, the Specificity for human class is 96.90% and for non-human class is 80.59%, the Miss rate for human class is 19.40% and for the non-human class is 3.09%, the Fall-out for the human class is 3.09% and for the non-human class is 19.40%, the false discovery rate for the human class is 14.28% and for non-human class is 4.40%, the false omission rate for the human class is 4.40% and for the non-human class is 14.28% , the accuracy for the human class is the same as the non-human class and is equal to 93.85%. The results of the statistical analysis of the performance of the proposed approach can be summarised in Table 5-12.

Table 5-12 The result of the statistical function for human and non-human classes for the proposed approach.

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	**80.59%**	**85.71%**	**95.59%**	**96.90%**	**19.40%**	**3.09%**	**14.28%**	**4.40%**	**93.85%**
Non-Human	**88.74%**	**95.59%**	**85.71%**	**80.59%**	**3.09%**	**19.40%**	**4.40%**	**14.28%**	**93.85%**

Table 5-12 presents the distribution result of the statistical analysis of the performance of the proposed approach for this experiment as shown in Figure 5.3.

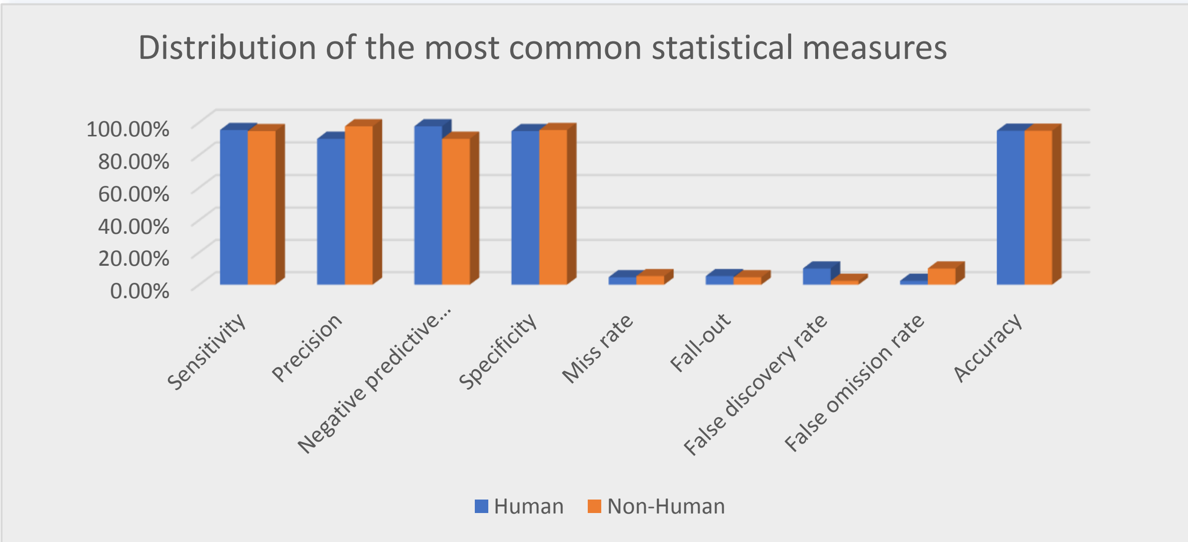


Figure 5.3 The distribution of the most common statistical measures for the proposed approach

Similarly, after calculating the most common statistical functions for both classes (human and non-human) based on the respective formula for each function for the support vector machine, it was found that, the sensitivity for the human class is 1.49% and for the non-human class is 98.62%, the precision for the human class is 20% and for the non-human class is 81.30%. The negative predictive value for human class is 81.30% and for the non-human class is 20%, the specificity for the human class is 98.62% and for non-human class is 1.49%, the miss rate for the human class is 98.5% and for the non-human class is 1.374%.The fall-out for the human class is 1.374% and for the non-human class is 98.5%, the false discovery rate for human class is 80% and for the non-human class is 18.696%, the false omission rate for the human class is 18.696% and for the non-human class is 80%, the accuracy for the human class is the same as the non-human class and is equal to 80.446%. We can summarise the results of the statistical analysis of the performance of the support vector machine approach in Table 5-13.

Table 5-13 The result of the statistical function for human and non-human classes for support vector machine approach.

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	**1.49%**	**20%**	**81.3%**	**98.62%**	**98.5%**	**1.374%**	**80%**	**18.696%**	**80.446%**
Non-Human	**98.62%**	**81.3%**	**20%**	**1.49%**	**1.374%**	**98.5%**	**18.696%**	**80%**	**80.446%**

Based on Table 5-13 the distribution result of the statistical analysis of the performance of the support vector machine approach for this experiment is shown in Figure 5.4.

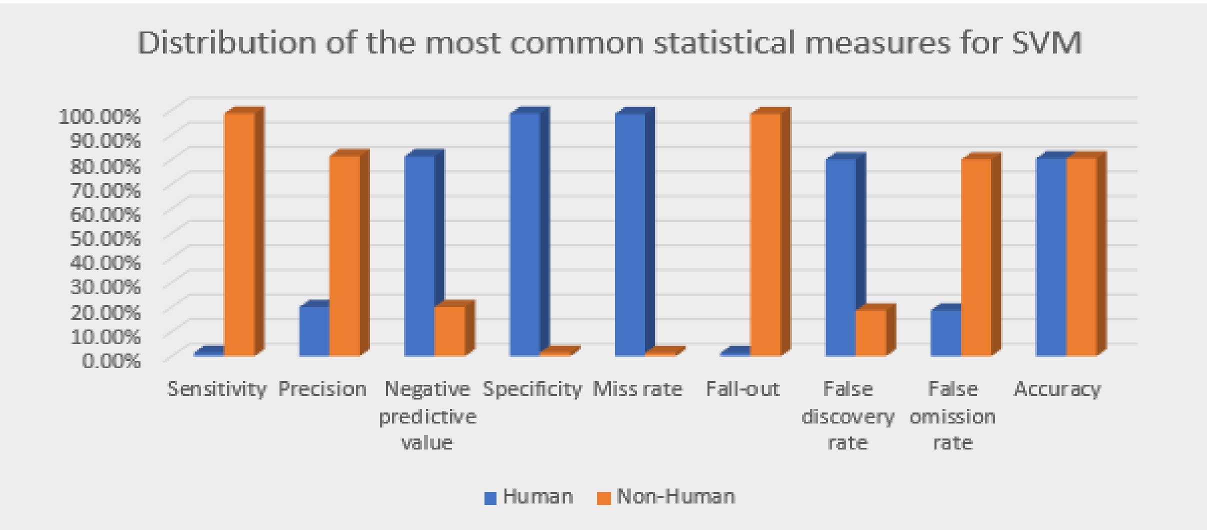


Figure 5.4 The distribution of the most common statistical measures for the support vector machine approach

For the artificial neural network approach, the statistical analysis for both classes (human and non-human) indicated that, the sensitivity for the human class is 68.656% and for the non-human class is 94.5%, the precision for the human class is 74.193% and for the non-human class is 92.9%. The Negative predictive value for the human class is 92.9% and for the non-human class is 74.193%, the specificity for the human class is 94.5% and for the non-human class is 68.656%, the miss rate for the human class is 31.34% and for the non-human class is 5.498%, the fall-out for human class is 5.498% and for the non-human class is 31.34%, the false discovery rate for the human class is 25.8% and for the non-human class is 7.09%, the false omission rate for the human class is 7.09% and for the non-human class is 25.8%, the accuracy for the human class is the same as the non-human class and is equal 89.664%. The results of the statistical analysis of the performance of the artificial neural network approach is summarised in Table 5-14

Table 5-14 The result of the statistical function for human and non-human classes for artificial neural network approach.

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	68.656%	74.193%	92.9%	94.5%	31.34%	5.498%	25.8%	7.09%	89.664%
Non-Human	94.5%	92.9%	74.193%	68.656%	5.498%	31.34%	7.09%	25.8%	89.664%

The distribution result of the statistical analysis of the performance of the artificial neural network approach for this experiment can be presented as shown in Figure 5.5.

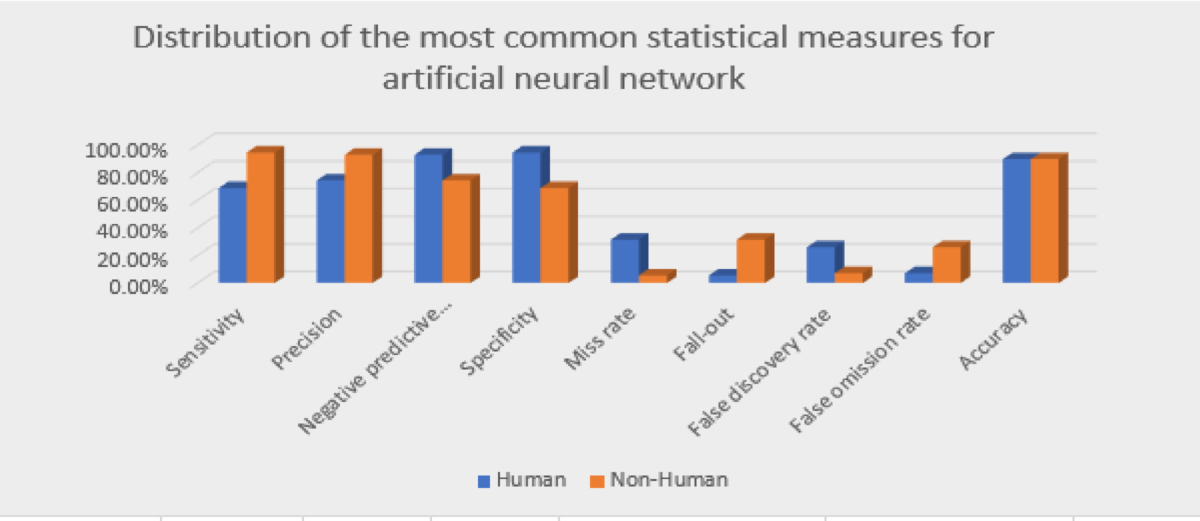


Figure 5.5 The distribution of the most common statistical measures for the artificial neural network approach

While the statistical analysis for both classes (human and non-human) for the random forest approach, shows the sensitivity for the human class is 59.7% and for the non-human class is 96.9%, the precision for human class is 81.632% and for the non-human class is 91.262%. The negative predictive value for the human class is 91.262% and for the non-human class is 81.632%, the specificity for the human class is 96.9% and for the non-human class is 59.7%, the miss rate for human class is 40.298% and for the non-human class is 3.092%, the fall-out for human class is 3.092% and for the non-human class is 40.298%, the false discovery rate for the human class is 18.367% and for the non-human class is 8.737%, the false omission rate for the human class is 8.737% and for the non-human class is 18.367%, the accuracy for the human class is the same as the non-human class and it equal 89.944%. We can summarise the result of the statistical analysis of the performance of the random forest approach in Table 5-15.

Table 5-15 The result of the statistical function for human and non-human classes for random forest approach.

Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Human	**59.7%**	**81.632%**	**91.262%**	**96.9%**	**40.298%**	**3.092%**	**18.367%**	**8.737%**	**89.944%**

Non-Human	**96.9%**	**91.262%**	**81.632%**	**59.7%**	**3.092%**	**40.298%**	**8.737%**	**18.367%**	**89.944%**

From Table 5-15 we can present the distribution results of the statistical analysis of the performance of the random forest approach for this experiment as shown in Figure 5.6.

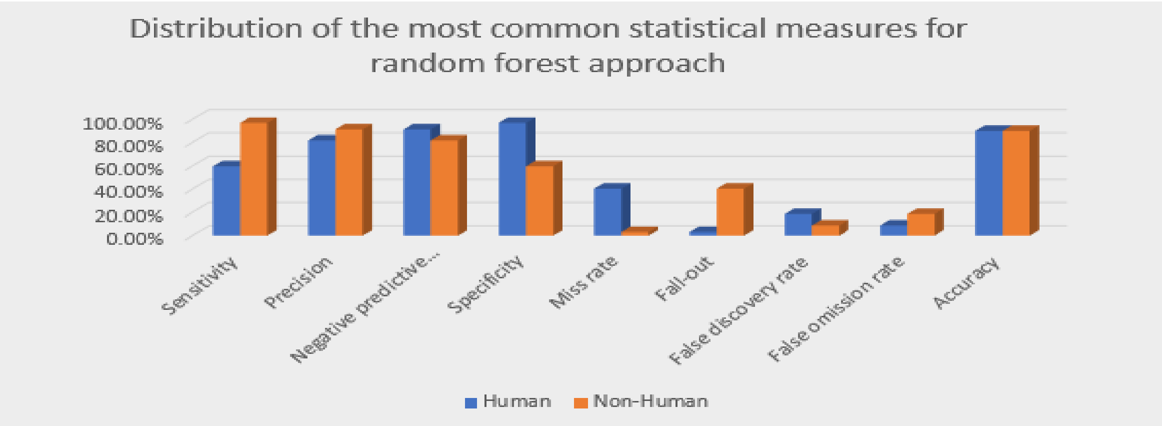


Figure 5.6 The distribution of the most common statistical measures for the Random forest approach

5.5 Evaluation

In this chapter, we aimed to evaluate the performance of the proposed approach in classifying the detected object as human or non-human by comparing it with some of the common machine learning approaches such as support vector machine, random forest, and artificial neural network. The performance analysis of these approaches was performed in two tests: accuracy matrix, and confusion matrix using the same dataset and they were implemented using MATLAB R2017b and tested on a 1.8 GHz core i7 (IV), 16 GB memory and 512 GB hard drive.

After carrying out the experiment based on the selected samples of images, the results of the detection performance for these approaches are presented in a confusion matrix, from the confusion matrix, the performance accuracy for these approaches were obtained, the accuracy of the proposed approach was 93.85%, while the accuracy of the support vector machine was 80.44%, the accuracy of artificial neural network was 89.664 and the accuracy for random forest was 89.944%. Figure 5.7 presents the comparing accuracy results of these approaches.

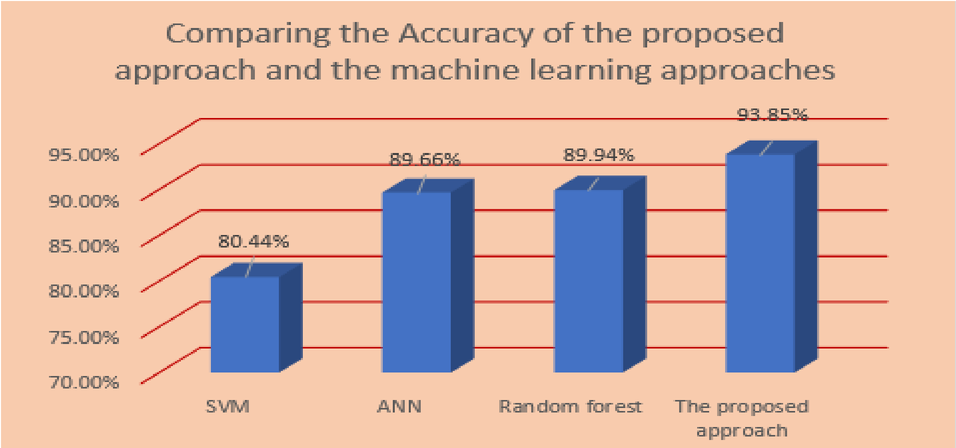


Figure 5.7 Comparing the Accuracy of the proposed approach and the machine learning approaches

As seen from Figure 5.7, the proposed approach achieved the highest accuracy rate (93.85%) compared with the other machine learning approach and the support vector machine achieved the lowest accuracy rate (80.44%). For more details in comparison between the proposed approach and the machine learning approach. Table 5-16 presents a comparation summary of the results of these approaches in terms of the number of correctly classified instances, the number of incorrect classified instances, the average of missed detection rates, and the average false discovery rate.

Table 5-16 A comparation summary for the results of the proposed approach and the machine learning approaches

	Total number of Instances	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy	Average of Miss rate	Average of False discovery rate	Computational Performance (Time-*ms*)
Proposed approach	358	336	22	93.85%	11.245%	9.34%	13475
Support vector machine	358	288	70	80.44%	49.937%	49.348%	13387
Artificial neural network	358	321	37	89.664%	18.419%	16.445%	13889
Random forest	358	322	36	89.944%	21.695%	13.552%	14240

Table 5-16 indicates that the proposed approach achieved the highest number of correctly classified instances (336 of 358), then the Random forest approach (322 of 358), then the Artificial neural network approach (321 of 358), and the Support vector machine approach achieved the lowest number of correctly classified instances (288 of 358). Figure 5.8 shows a Comparing for instance classification performance between the proposed approach and the machine learning approaches.

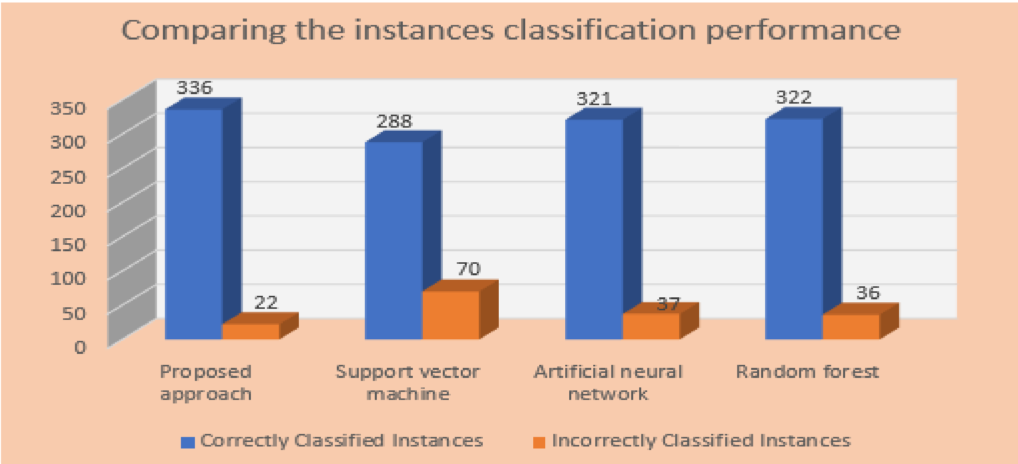


Figure 5.8 Comparing the instances classification performance between the proposed approach and the machine learning approaches

For the average of miss detection rate, the proposed approach achieved the lowest percentage rate of miss- detection (11.245%), the artificial neural network approach (18.419%), than the random forest approach (21.695%), while the support vector machine approach achieved the highest average miss -detection (49.937%).

For the average false discovery rate, the proposed approach achieved the lowest percentage rate of false discovery (9.34%), then the random forest approach came next (13.552%), then the rtificial neural network approach (16.445%), while the support vector machine approach achieved the highest average of false discovery rate (49.937%).

For the computational performance time (speed), the Support vector machine approach achieved the minimum computational performance time (the speediest) completing the process of classifying in13387 ms, the proposed approach comes next (13475 ms), then the artificial neural network approach (13889 ms), and the random forest approach achieved the maximum computational performance time by completing the process of classifying the instances in 14240 ms. Figure 5.9 shows a comparison between the proposed approach and the machine learning approach in the computational performance time.

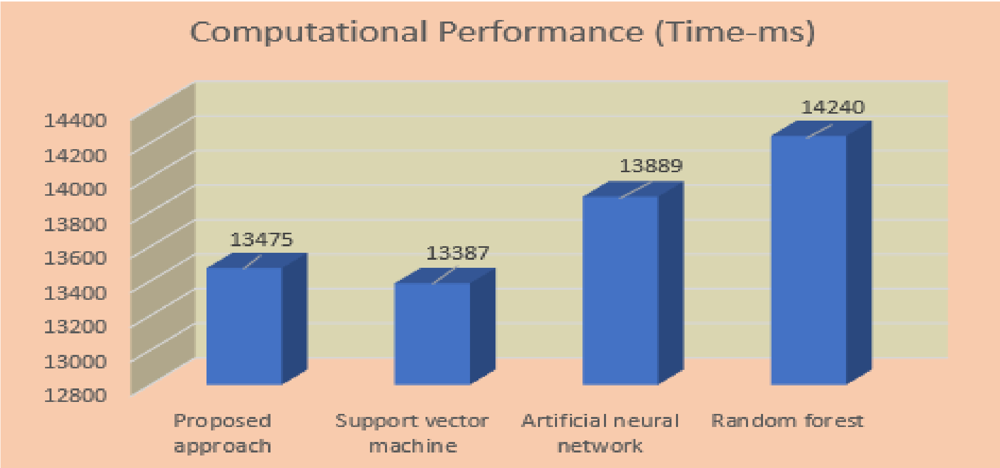


Figure 5.9 Comparing between the proposed approach and the machine learning approach in the computational performance time.

As we mentioned before the Machine learning approach such as the support vector machine, random forest, and artificial neural network have been introduced and performed in order to evaluate the classification accuracy of the proposed approach. For more of a comparison of the classification accuracy, Table 5-17 presents a full comparison between the

Table 5-17 A full comparing between the proposed approach and the machine learning approaches based on the result of the statistical functions for each approach.

proposed approach and the machine learning approaches based on the result of the statistical functions for each approach.

Approach	Class	Sensitivity	Precision	Negative predictive value	Specificity	Miss rate	Fall-out	False discovery rate	False omission rate	Accuracy
Random forest	Human	59.7%	81.632%	91.262%	96.9%	40.298%	3.092%	18.367%	8.737%	89.944%
	Non-Human	96.9%	91.262%	81.632%	59.7%	3.092%	40.298%	8.737%	18.367%	89.944%
	Average	78.30%	86.45%	86.45%	78.30%	21.70%	21.70%	13.55%	13.55%	89.94%
Artificial neural network	Human	68.656%	74.193%	92.9%	94.5%	31.34%	5.498%	25.8%	7.09%	89.664%
	Non-Human	94.5%	92.9%	74.193%	68.656%	5.498%	31.34%	7.09%	25.8%	89.664%
	Average	81.58%	83.55%	83.55%	81.58%	18.42%	18.42%	16.45%	16.45%	89.66%
Support vector machine	Human	1.49%	20%	81.3%	98.62%	98.5%	1.374%	80%	18.696%	80.446%
	Non-Human	98.62%	81.3%	20%	1.49%	1.374%	98.5%	18.696%	80%	80.446%
	Average	50.06%	50.65%	50.65%	50.06%	49.94%	49.94%	49.35%	49.35%	80.45%
The proposed approach	Human	80.59%	85.71%	95.59%	96.90%	19.40%	3.09%	14.28%	4.40%	93.85%
	Non-Human	88.74%	95.59%	85.71%	80.59%	3.09%	19.40%	4.40%	14.28%	93.85%
	Average	84.67%	90.65%	90.65%	88.75%	11.25%	11.25%	9.34%	9.34%	93.85%

Table 5-17 demonstrates the proposed approach accuracy to classify the detected object as human is 93.85% with the average miss- detection rate of 11.245%, and a false discovery rate for human class at 14.28% and 4.4% for the non-human class with an average of 9.34%. While the support vector machine approach accurately classified the detected objects as human at 80.446%, with the average miss -detection rate of 49.937%, the false discovery rate for the human class is 80% and 18.696% for the non-human class on average 49.35%. The artificial neural network approach accuracy of classifying the detected object as human is 89.664% with the average miss- detection rate of 18.419%, and a false discovery rate for the human class is 25.8% and 7.09% for the non-human class on average is 16.45%, and the Random forest approach accuracy of classifying the detected object as human is 89.944% with the average miss- detection rate 21.695%, and the false discovery rate for the human class is 18.367% and 8.737% for the non-human class on average is 13.55%.

From the above comparison between the proposed approach and the machine learning approaches in term of accuracy performance, it can be seen that the proposed approach achieved the highest accuracy rate (93.85%), with lowest miss- detection rate (11.245%) and false discovery rate (9.34%). The support vector machine approach achieved the minimum computational performance time (the speediest) by completing the process of classifying the instances (13387 ms) with the lowest accuracy rate (80.446%) and the highest average of miss-detection rate (49.937%). and the highest average false discovery rate (49.35%). While the proposed approach comes after the support vector machine by completing the process of classifying the instances (13475 ms) with the highest accuracy rate (93.85%) and the lowest average miss -detection rate (11.245%) and the lowest average false discovery rate (9.34%). Figure 5.10 shows the comparative results of the accuracy performance.

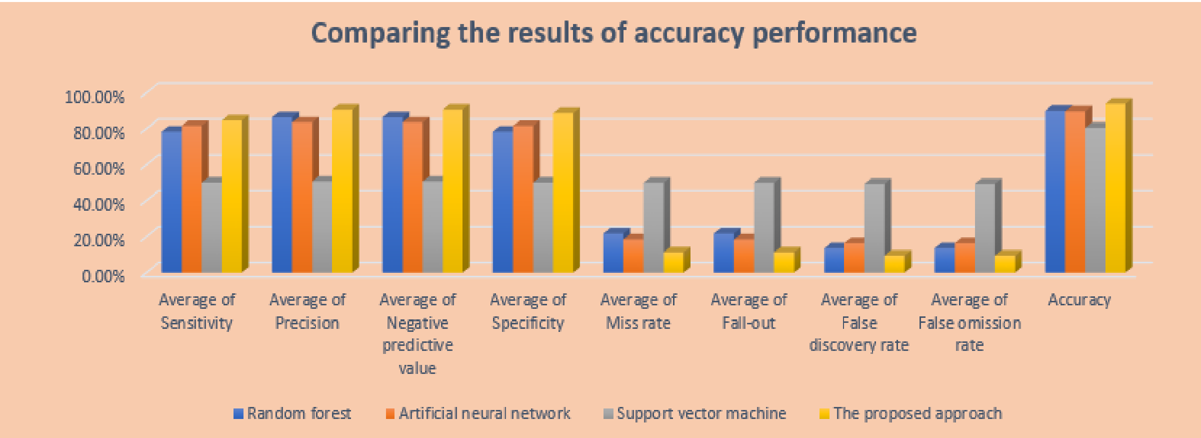


Figure 5.10 Comparing the results of the accuracy performance between the proposed approach and the machine learning approaches

Figure 5.10 shows the efficiency of the proposed approach's performance in classifying the detected object as human or non-human by comparing it with the machine learning approaches such as the support vector machine, random forest, and artificial neural network.

5.6 CONCLUSION

In this chapter the Machine learning approaches such as the support vector machine, random forest, and artificial neural network have been introduced and performed in order to evaluate the classification accuracy of the proposed approach. The evaluation carried out by performing the proposed approach and the machine learning approaches using the same dataset samples and under the same processor characteristics.

The experimental results indicate the proposed approach accuracy to classifying the detected object as human is 93.85% with an average miss -detection rate of 11.245%, the false discovery rate for the human class is 14.28% and 4.4% for the non-human class with an average of 9.34%. While the support vector machine approach's accuracy to classifying the detected object as human is 80.446% with the average miss- detection rate of 49.937%, and the false discovery rate for human the class is 80% and 18.696% for the non-human class on average is 49.35%, the artificial neural network approach accuracy to classifying the detected object as human is 89.664% with the average miss- detection rate at 18.419%, and the false discovery rate for the human class is 25.8% and 7.09% for the non-human class on average is 16.45%.The Random forest approach accuracy to classifying the detected object as human is 89.944% with the average a miss -detection rate at 21.695%, and the false discovery rate for the human class is 18.367% and 8.737% for the non-human class on average is 13.55%.

From the above comparison between the proposed approach and the machine learning approaches in term of accuracy performance, the proposed approach shows the highest accuracy rate achieved (93.85%), with lowest miss detection rate (11.245%) and false discovery rate (9.34%). The support vector machine approach achieved the minimum computational performance time (the speediest) by completing the process of classifying the instances (13387 ms) with the lowest accuracy rate (80.446%) and the highest average miss-detection rate (49.937%), also the highest average false discovery rate (49.35%). While the proposed approach comes after the support vector machine by completing the process of classifying the instances (13475 ms) with the highest accuracy rate (93.85%) and lowest average miss- detection rate (11.245%) and the lowest average false discovery rate (9.34%).

This therefore, indicates that the proposed approach is efficient with low calculating complexity and achieves a higher classification accuracy than machine learning approaches. Thus, the proposed approach provides a standard way or a good alternative for real-time applications.

CHAPTER SIX

6 APPLICATIONS CASE STUDY

The task of recognising and detecting objects such as humans in images and videos sequences turn research attention to the fields of computer vision and machine learning around the world, due to its wide applicability scope and for the large potential applications that can be acquired, such as assistance systems for auto- drive, monitoring systems, efficient graphic user interface, motion personification, and so on. In the following section a brief description of two examples for the applicability of human detection in real time systems is presented.

6.1 Human detection in surveillance system

In the last decade, the task of human detection rises to be an integral part in various real applications especially in areas that require surveillance [7, 8], due to the large amount of visual data that the outcome of these applications produce which need to be processed and managed.

In video surveillance systems used to identify and detect human objects there must be someone or something monitoring the video sequence. This is usually done by a human operator, this operator has to monitor the stream of records captured from the surveillance cameras and displayed on many screens, in order to detect the abnormal behaviour. Because human operators are very good and efficient at recognising positions, it will do so as long as the operators are able to focus and watch all the screens in a short time [6]. Clearly there is a limit to how much one person can effectively follow and watch all at the same time, and with the installation of more cameras, more human resources are needed. For example, human abnormal behaviour detection in surveillance systems is widely used in many real time applications, and it has become a crucial need for security purposes, because detecting abnormal actions robustly, increases the opportunity to avoid accidents and may be acquired by triggering such alarms or signals to the surveillance system operators.

Identifying abnormal behaviour can be different in many applications, that's because every application environment has its specifications of abnormal behaviour. These abnormal behaviours or actions can be such as people running in a specific place at the same time, someone holds illegal items in their hand, or someone jumping in a secure section, and so on. Figure 6.3 below show some example of abnormal human behaviours.

Figure 6.1 Some example of human abnormal behaviours.

The first step in detecting human abnormality behaviours using surveillance systems, is to detect the human object in an image or video frame, in order to classify the behaviour as normal or abnormal, so that the needs of such an approach with high accuracy for classifying the located object as a human is very important for further process abnormality detection or tracking.

An example of human abnormality detection is to detect the abnormal behaviours of humans (students) in academic scenarios, this case study can be implemented in real time applications by following a sequence of steps in order to detect the human and then classify the human behaviour as normal behaviour or abnormal.

The below figure shows the general follow diagram for abnormality detection.

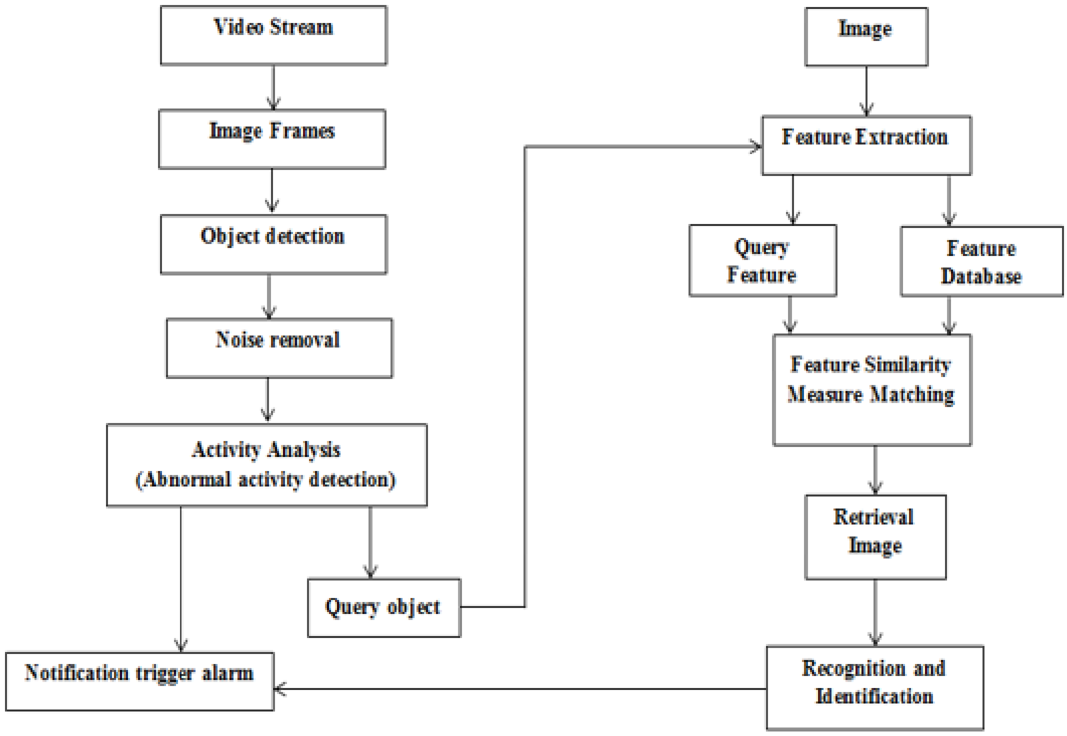


Figure 6.2 The general follow diagram for the abnormality detection system.

In this case study, the aim is to detect the abnormal behaviour of the students and to classify the identity of the student who did the abnormal behaviour, to do this, it requires a pre-phase to collect the student pictures and their details from the students records, and then from the students pictures, the unique features are extracted for each student and it is stored with the corresponding student details in a database to be used in identifying the student.

From the general flow diagram, the first step is to obtain the video stream, and this can be captured by several kinds of surveillance system cameras, then this video stream is divided into 32 frames as the normal number of frames.

The second step is to locate the region of the abnormality, and detect the objects that are causing this abnormal behaviour so that a temporal differencing approach can been used to detect the regions of the abnormality and detect the object who causes it.

After that a binary statistical erosion function was applied to remove any noise that can affect the detection.

By using this the objects can be detected however, to classify the detected object as human or non-human, the similarity pattern matching was used by applying the Omega equation which presents a spatial pattern called S pattern. Using this S pattern, the similarity pattern matching process runs in order to classify the human and ignore other objects. The following figure show the flow diagram used to localise and detect objects.

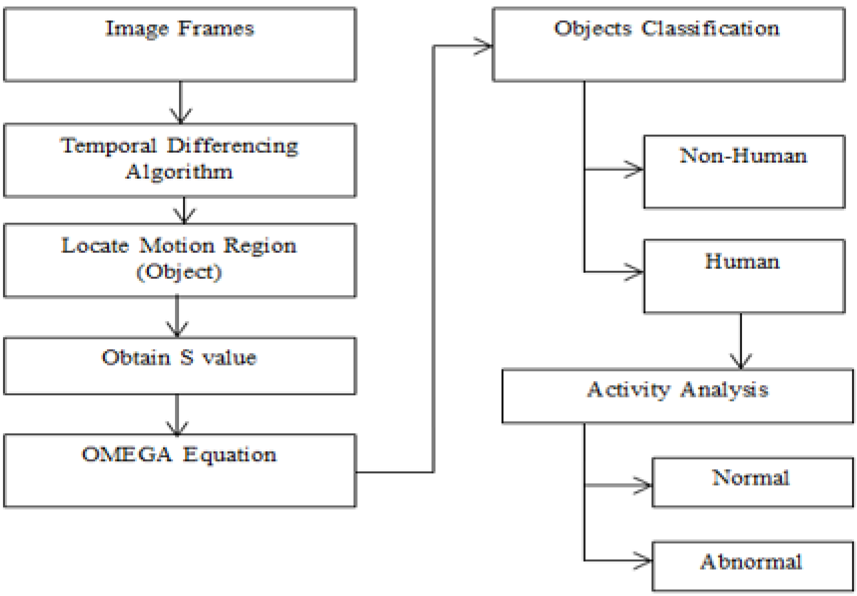


Figure 6.3 the follow diagram to localize and detect object.

The main steps of the shape model is summarised based on the OMEGA equation as the following:

Steps of the shape model based on OMEGA Equation:	
Step I:	The motion region box is designed to include the object of interest and whose axes are aligned with the image axes
Step II:	Based on the set of boundary points obtained (i.e. Motion region box), coordinates (C_x, C_y) are calculated.
Step IV:	Obtain the distance $d = (C_y - Y_{min})$
Step V:	Obtain H= half of distance, where H is the window height for extracting the head and shoulder portion of the human object.
Step VI:	Extract the set of co-ordinates from the boundary of the upper-segmented contour.
Step VII:	These values are then substituted in the corresponding Equation to obtain the pattern for *S*.

The results of the above steps is to present the S pattern which can be used to classify the detected object as human or non-human based on the similarity of shape matching. Figure 6.4 shows the presented S pattern.

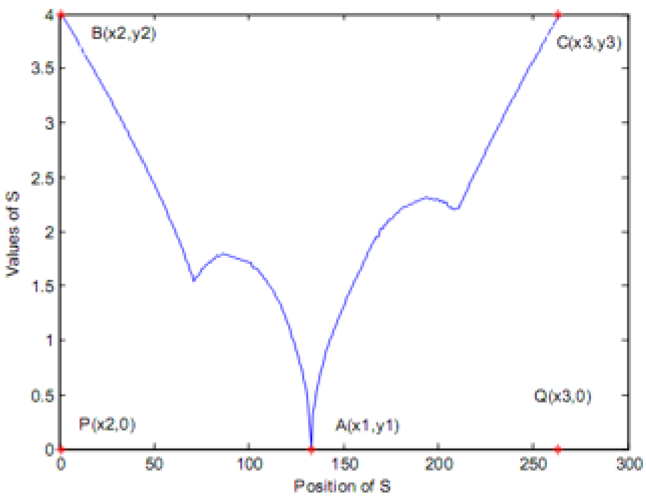


Figure 6.4 The presented S pattern.

After classifying the detected object as human, the next step is to analyse the activity of the human in terms of what's normal activity or abnormal activity, in order to this to do this the activity features are extracted by using the support vector machine approach to classify the abnormal activity of a human, the general steps for analysing the human activity shown in Figure 6.5.

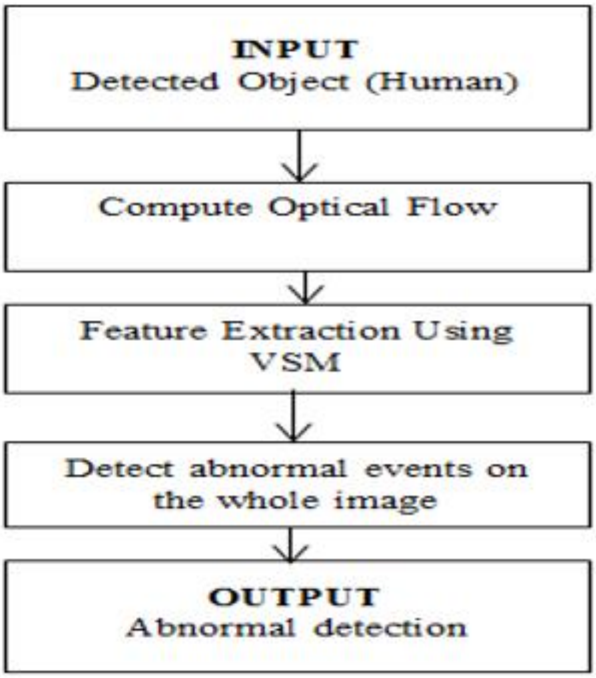


Figure 6.5 General steps for analysis the human activity

The main idea of the support vector machine is to divide the data set into different groups based on finding the HYPERPLANE and then the furthest group with the closest points to the class. Figure 6.6 shows thedistance of group using the HYPERPLANE based on support vector machine.

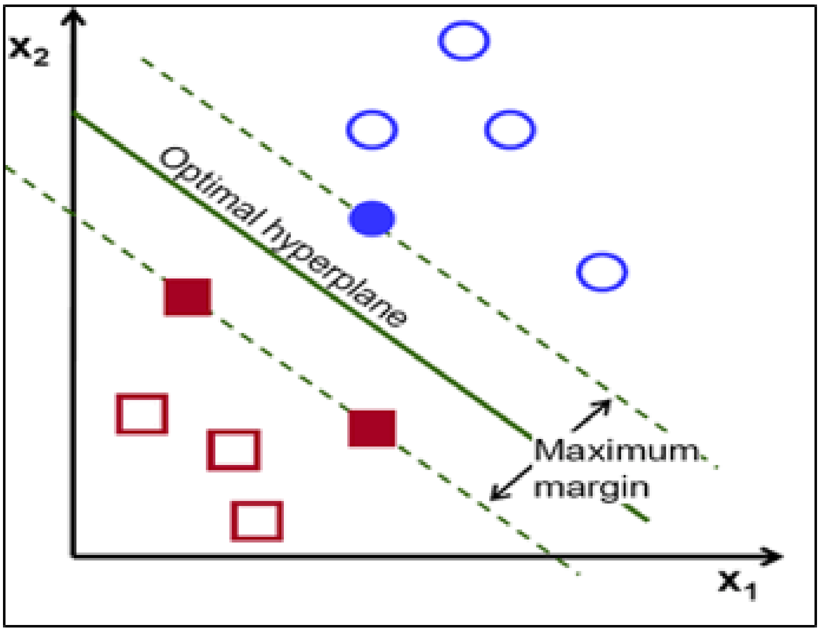


Figure 6.6 The choosing distance of group using the HYPERPLANE based on support vector machine.

The next step after classifying the human activity, in cases were the human activity is classified as abnormal, a trigger alarm will be sent directly to the security team for security purposes, and a picture of the person who is causing the abnormal activity will be obtained and send forward to the information retrieval process.

In the information retrieval process, the obtained image of the person's features will be extracted to find the matching features between this person features and the dataset of all the students features in order to identify the person.

Figure 6.7 shows the flow of steps for the retravel of information in order to extract the details and identity the person who is causing the abnormal activity.

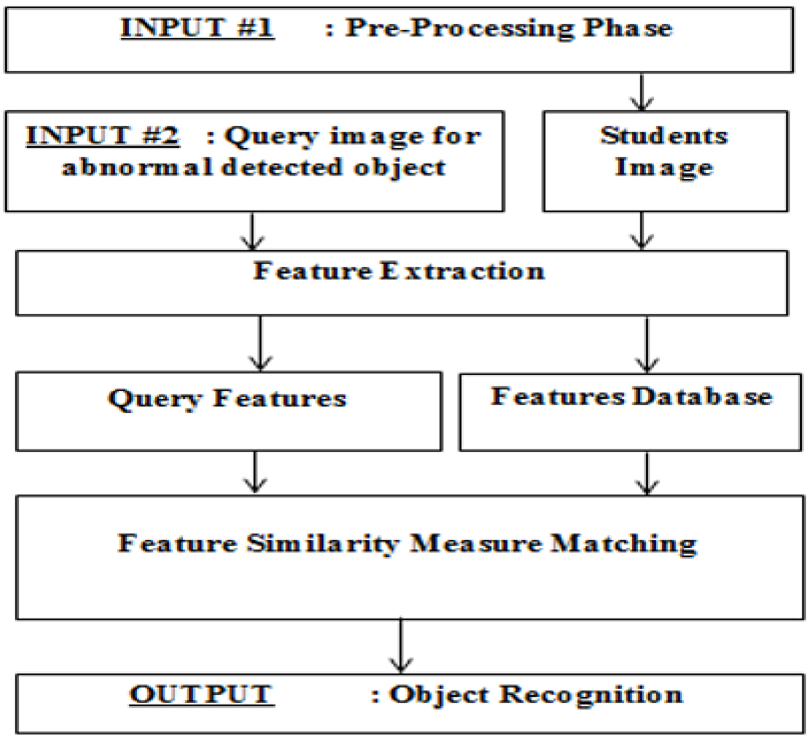


Figure 6.7 The flow steps for information retravel in order to extract the details and identity of the person who causes the abnormal activity.

6.2 Advance driver assistance system:

Driver assistance systems have risen to be one of the hottest topics in computer vision and machine learning areas, and it rivets the attention of both the computer science community and the automobile industry to develop a variety of efficient systems that can improve traffic safety, this due to the rising number of road accidents and the popularity of vehicles over the last century.

Driver assistance system aims to reduce the number of accidents for traffic safety, this can be achieved by developing many mechanisms and systems that can anticipate accidents and provide the driver with assist to avoid the anticipated accidents or to reduce the accident severity. Some of these mechanisms and systems aim to monitor the driver behaviour during the driving process and assist the driver with such alarms or signals in case of any fault acquired by the driver, for example driver head pose monitoring, driver eye gaze monitoring.

Another driver assistance mechanism and systems focus on monitoring and analysing the environmental infrastructure such as road and lane detection, analyse the traffic signs and more. Furthermore, some of these driver assistance mechanisms focused on monitoring the safety in vehicles, for example antilock braking systems, electronic stabilisation programs, and airbags. Human detection or pedestrian detection plays a major challenge in driver assistance systems, and it aims to detect the presence of a human in a specific area of interest, in terms of warning the driver to avoid accidents. This challenge is to rise up because the pedestrian not like other objects, the pedestrian can have varying appearances, such as different clothes, changing sizes, and they can be located in unstructured environments. The general flow diagram of a pedestrian detection system is shown in Figure 6.1.

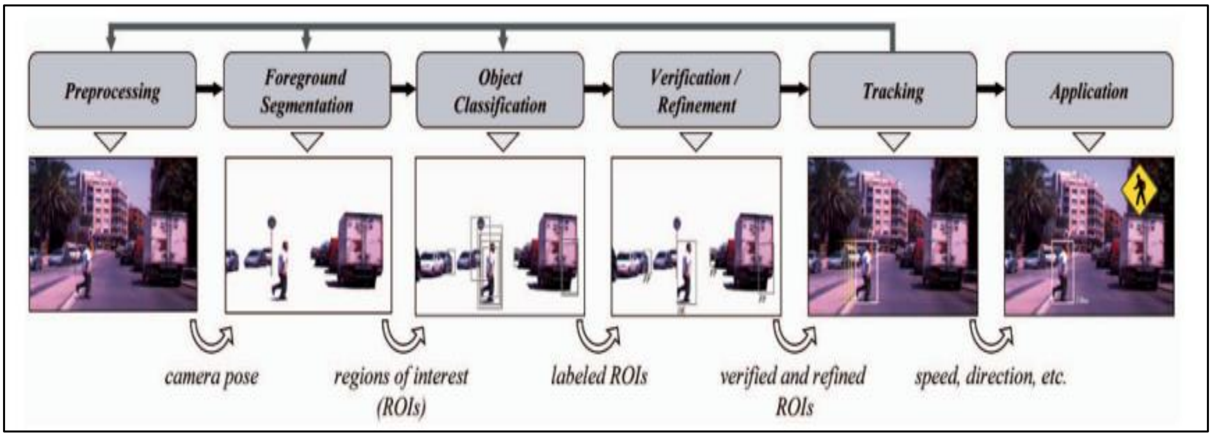


Figure 6.8 The general flow diagram for the process of pedestrian detection system

The statistical accidents for pedestrians indicates that 70% of pedestrian accidents acquired at the front of the vehicle, typically, this led to it being used in front sensors for pedestrian detection.

Similarly, human detection is very important in a self-driving car, which is a new machine learning industrial challenge, where there is no driver in the car, the car tries to sense the presence, and identify the objects on the road that are ahead of the car, to analyse the scenario and provide the machine with the optimal decision for traffic safety.

Figure 6.2 shows examples of self-driving car scenarios, and the optimal decision.

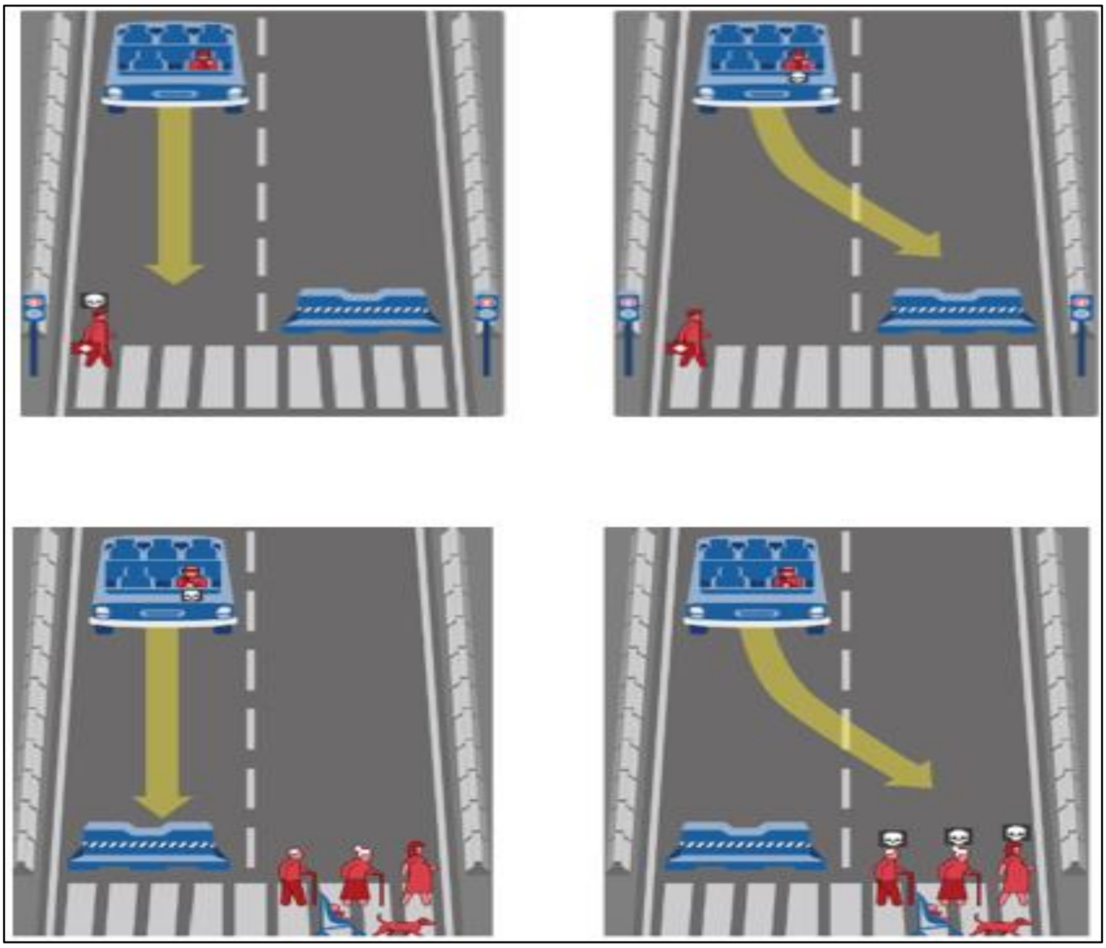


Figure 6.9 Examples of self-driving car scenario, and the optimal decision.

In fact, there are several approaches for object detection however ,for driver assistant and pedestrian protection systems, the need for efficient approaches is increasing, that's because some object detection approaches require a full-face detection and another is based on colour or texture detection, which shows limitations to use such of these approaches, due to challenges of pedestrian appearance, and the occluded problem. The proposed approach is efficient for being pedestrian protect systems, because it can detect and classify the detected object as human based on the upper portion part of the object, which is typically visible and not easy to occlude.

CHAPTER SEVEN

7 CONCLUSION AND FUTURE WORKS

7.1 CONCLUSION

In this thesis, we have achieved our objectives of presenting a new shape-based classification approach which concentrates on improving object classification accuracy. The proposed approach has made an important contribution by providing an increased accuracy in object classification of human or non-humans from images, which has received large attention within the literature. In the proposed approach, the objects under different conditions can be accurately detected and classified by combining the features that are extracted from the objects upper portion and the proposed geometrical model parameters.

The machine learning approaches, such as a random forest model, artificial neural network model, and support vector machine model was introduced and performed to test and evaluate the classification accuracy of the proposed approach, by making a comparison between the proposed approach and these machine learning approaches in order to test the efficiency. However, the analysis and conclusions would have been stronger and more generalised if the dataset were larger. Therefore, in the experiments public dataset was used, which contains an object such as a human or non-human in the images (358 images) known as INRIA dataset. INRIA dataset contains human and non-human images cropped in different sizes such as 64 × 128 pixels, and 214 × 320 – 648 × 486 pixels. The experimental results show the classification performance for the machine learning approaches and the proposed approach are as follows:

The proposed approach accuracy to classify the detected object as human is 93.85% with the average miss detection rate at 11.245%, the false discovery rate for human class is 14.28% and 4.4% for non-human class in average of 9.34%.While the support vector machine approach accuracy used to classify the detected object as human is 80.446%, with the average mis-detection rate at 49.937%, the false discovery rate for human class is 80% and 18.696% for non-human class on average is 49.35%, the Artificial neural network approach accuracy to classify the detected object as human is 89.664% with the average misdetection rate being 18.419%, and the false discovery rate for human class is 25.8% and 7.09% for non-human class on average is 16.45%, and the Random forest approach accuracy to classify the detected object as human is 89.944% with the average miss-detection rate at 21.695%, and the false discovery rate for human class at 18.367% and 8.737% for non-human class on an average of 13.55%.

From the above comparison between the proposed approach and the machine learning approaches in term of accuracy performance, we can indicate that the proposed approach achieved the highest accuracy rate (93.85%), with lowest miss-detection rate (11.245%) and false discovery rate (9.34%).

For the computational performance time, the support vector machine approach achieved the minimum computational performance time (the quickest) by completing the process of classifying the instances (13387 ms) with the lowest accuracy rate (80.446%) and the highest average miss-detection rate (49.937%), also the highest average false discovery rate (49.35%). While the proposed approach comes after the support vector machine by completing the process of classifying the instances (13475 ms) with the highest accuracy rate (93.85%) and lowest average miss-detection rate (11.245%), also with the lowest average false discovery rate (9.34%).

This indicates that the proposed approach is efficient with low calculating complexity and achieves a higher classification accuracy than machine learning approaches. Thus, the proposed approach provides a standard way, or a good alternative for real-time applications.

7.2 FUTURE WORKS

This study provides a new approach for object detection in images and to classifying the detected object as human or non-human. The implementation of the new approach shows the efficiency of this approach, when compared with other approaches in terms of complexity, efficiency and the overall performance.

There are some work that can be addressed in the future:

- Improving this approach to detect and classify an object in videos, since this approach can detect and classify objects in static images.
- Improving this approach to detect and classify many objects in an image, since this approach can detect and classify a single object in static images.

REFERENCES

[1] Duc Thanh Nguyen, Wanqing Li, Philip O.Ogunbona. Human Detection from Images and Videos. Article in Pattern Recognition · September 2015.

[2] Navneet Dalal, Bill Triggs and Cordelia Schmid," Human Detection using Oriented Histograms of Flow and Appearance "2011.

[3] Subra Mukherjee, Karen Das .A Novel Equation based Classifier for Detecting Human in Images International Journal of Computer Applications (0975 – 8887) Volume 72– No.6, May 2013

[4] https://www.dailymail.co.uk/sciencetech/article-3145887/Google-apologises-Photos-app-tags-black-people-gorillas-Fault-image-recognition-software-mislabelled-picture.html

[5] Duc Thanh Nguyen. Human detection from image and video.PhD Thesis ,University of Wollongon ,2012.

[6] A. Ansari, T. Manjunath, and C. Ardil, "Implementation of a Motion Detection System," *International Journal of Electrical and Computer Engineering*, vol. 3, 2008.

[7] Paul M., Haque S. M. and Chakraborty S. Human detection in surveillance videos and its applications – a review. Advances in Signal Processing, 2013(176):1–16, 2013.

[8] Su F., Fang G. and Kwok N. M. Adaptive colour feature identification in image for object tracking. Mathematical Problems in Engineering, 2012(509597):1–18, 2012.

[9] Hou Y. L. and Pang G. K. H. People counting and human detection in a challenging situation. Systems, Man and Cybernetics, 41(1):24–33, 2011.

[10] Jafri R. and Arabnia H. R. A survey of face recognition techniques. Journal of Information Processing Systems, 5(2):41–68, 2009.

[11] *Revathi A. R. and Kumar D. A review of human activity recongnition and behavior understanding in video surveillance. Advanced Computer Science and Information Technology, pages 375–384, Chennai, India, 2012.*

[12] *Chowdhury M. S., Kuang Y. C. and Ooi M. P. Fast and accurate human detection for video applications using edgelets. Computer Applications and Industrial Electronics, pages 74–79, Kuala Lumpur, Malaysia, 2010.*

[13] *Ogale N. A. A survey of techniques for human detection from video. Master Thesis, University of Maryland, 2006.*

[14] *Santhanam T., Sumathi C. P. and Gomathi S. A survey of techniques for human detection in static images. Computational Science, Engineering and Information Technology, pages 328–336, Coimbatore, India, 2012.*

[15] *Figueira D., Moerno P., Bernardino A., Gaspar J. and Victor J. S. Optical flow based detection in mixed human robot environments. Visual Computing, pages 223–232, Las Vegas, USA, 2009.*

[16] *Zhang L. J. and Liang Y. L. Motion human detection based on background subtraction. Education Technology and Computer Science, pages 284–287, Wuhan, China, 2010.*

[17] *Subra Mukherjee, Karen Das .A Novel Equation based Classifier for Detecting Human in Images International Journal of Computer Applications (0975 – 8887) Volume 72– No.6, May 2013*

[18] *Guang shu human detection, tracking and segmentation in surveillance video M.S. Shanghai Jiaotong University, 2009*

[19] Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking: A Survey on Object Detection and Tracking Methods. International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 2, February 2014

[20] Rupali S.Rakibe, Bharati D.Patil, "Background Subtraction Approach Based Human Motion Detection",International Journal of Scientific and Research Publications, May 2013

[21] Abhishek Kumar Chauhan, Prashant Krishan, "Moving Object Tracking Using Gaussian Mixture Model And Optical Flow", International Journal of Advanced Research in Computer Science and Software Engineering, April 2013.

[22] J.Joshan Athanesious, P.Suresh, "Systematic Survey on Object Tracking Methods in Video",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) October 2012, 242-247.

[23] Sen-Ching S. Cheung and Chandrika Kamath, "Robust techniques for background subtraction in urban traffic video".

[24] K.Srinivasan, K.Porkumaran, G.Sainarayanan, "Improved Background Subtraction Techniques For Security In Video Applications"

[25] Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking Using Kalman Filter", International Journal of Computer Science and Mobile Computing,April 2013, pg.326 – 332.

[26] M.Sankari, C. Meena, "Estimation of Dynamic Background and Object Detection in Noisy Visual Surveillance", International Journal of Advanced Computer Science and Applications, 2011, 77-83.

[27] *Jae-Yeong Lee; Wonpil Yu, "Visual tracking by partition-based histogram backprojection and maximum support criteria," Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on , vol., no., pp.2860,2865, 7-11 Dec. 2011*

[28] *Sanjivani Shantaiya, Keshri Verma, Kamal Mehta, A Survey on Approaches of Object Detection, International Journal of Computer Applications (0975 – 8887) Volume 65– No.18, March 2013.*

[29] *Huabo Sun, Lei Yan, Peter Mooney and Rui Liang," A new method for moving object detection using variable resolution bionic compound eyes " ,International Journal of the Physical Sciences Vol. 6(24), pp. 5618-5622, 16 October, 2011.*

[30] *Wei-Lwun Lu, James J. Little," Simultaneous Tracking and Action Recognition using the PCA-HOG Descriptor "Proceedings of the 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*

[31] *Lan Wu," Multiview Hockey Tracking with Trajectory Smoothing and Camera Selection"*

[32] *Chih-Hsien Hsia and Jing-MingGuo., "Efficient modified directional lifting-based discrete wavelet transform for moving object detection", Signal Processing, article in press.*

[33] *S. Belongieet al, "Shape Matching and Object Recognition Using Shape Contexts", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24 Issue 4, Pg. No. 509-522, April 2002*

[34] *L Wang, X Geng, C Leckie, R Kotagiri, "Moving shape dynamics: a signal processing perspective", in IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008) (IEEE, Piscataway, 2008), pp. 1–8*

[35] *Liang Wang, Weiming Hu, Tieniu Tan, "Recent developments in human motion analysis"-Journal of pattern recognition Society,vol.36, pp-585-601, 2003.*

[36] *Longbin Chen, Julian J. Mc Auley, Rogerio S. Feris, Tib´erio S. Caetano, Matthew Turk, "Shape Classification Through Structured Learning of Matching Measures", Proceedings of IEEE conference on Computer Vision and Pattern Recognition (CVPR), 2009.*

[37] *Zhe Lin and Larry S. Davis, "Shape-Based Human Detection and Segmentation via Hierarchical Part-Template Matching", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, No. 4,pp-604-618, 2010.*

[38] *Jorge Garcıa, Alfredo Gardel, Ignacio Bravo, Jos´e Luis L´azaro, Miguel Mart´ınez and David Rodr´ıguez, " Directional People Counter based on Heads Tracking", IEEE transaction on Industrial Electronics, Vol. PP, Issue:99,2012.*

[39] *Rusi Antonov Filipov, Flavio Luis Cardeal Padua, Marco Aurelio Buono carone, "Pylon grid: A fast method for human head detection in range images",Journal of Neurocomputing(Elsevier),Vol.100, pp-74-85, 2013.*

[40] *Li M., Zhang Z. X., Huang K. Q. and Tan T. N. Rapid and robust human detection and tracking based on omega-shape features. Image Processing, pages 2545–2548, Cairo, Egypt, 2009.*

[41] *Huazhong Xu, Pei Lv, Lei Meng, "A People Counting System based on Head-shoulder Detection and Tracking in Surveillance Video", Proceedings of International Conference On Computer Design And Appliations (ICCDA), 2010.*

[42] *Sebastien Lef evre and Emmanuel Bouton and Thierry Brouard and Nicole Vincent," A new way to use hidden markov models for object tracking in video sequences "2003.*

[43] Xiaobai Liu, Liang Lin, Shuicheng Yan, Hai Jin, and Wenbin Jiang," Adaptive Object Tracking by Learning Hybrid Template Online ",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 21, No. 11, November 2011.

[44] Pranab Kumar Dhar et al., "An Efficient Real Time Moving Object Detection Method for Video Surveillance System", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 5, No. 3, September, 2012.

[45] Yuhua Zheng1 and Yan Meng," Object Detection And Tracking Using Bayes-Constrained Particle Swarm Optimization ", ISBN 978-1-60021-992-4, 2007 Nova Science Publishers, Inc

[46] William Robson Schwartz, Aniruddha Kembhavi, David Harwood, Larry S. Davis, "Human Detection Using Partial Least Squares Analysis", proceedings of IEEE 12th International Conference on Computer Vision,pp-24-31, 2009.

[47] Lowe, D.G, "Object recognition from local scale-invariant features", Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, Pg. No. 1150 – 1157.

[48] Paul Viola, Michael Jones, Daniel Snow," Detecting Pedestrians Using Patterns of Motion and Appearance ", Proceedings of the International Conference on Computer Vision (ICCV), October 13.

[49] A.Mohan,C.Papageorgiou , and T.Poggio, Example-based object detection in images by components, IEEE Trans.Pattern Anal.Mach.Intell.,vol 23.no.4, pp.349-361,Apr.2001.

[50] M. Andriluka, S. Roth, and B. Schiele. People-tracking-by-detection and people-detection-bytracking. In CVPR, 2008.

[51] *Krystian Mikolajczyk, Cordelia Schmid, Andrew Zisserman, "Human Detection based on a Probabilistic Assembly of Robust Part Detector", Proceedings of 8th European Conference on Computer Vision (ECCV),pp-68-82, 2004.*

[52] *John Chiverton, Xianghua Xie and Majid Mirmehdi," Automatic Bootstrapping and Tracking of Object Contours ", IEEE Transactions on Image Processing, Vol. 21, No. 3, March 2012.*

[53] *Sebastien Lef evre and Emmanuel Bouton and Thierry Brouard and Nicole Vincent," A new way to use hidden markov models for object tracking in video sequences "*

[54] *S. Walk, N. Majer, K. Schindler, and B. Schiele. New features and insights for pedestrian detection. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2010.*

[55] *Yanjiang wang, Baozong Yuan, "A novel approach for Human Face detection from color images under Complex Background"- Pattern Recognition, Elsevier, Volume 34, issue 10, pp- 1983-1992, 2010.*

[56] *Junda Zhu, Yuanwei Lao, and Yuan F. Zheng, "Object Tracking in Structured Environments for Video Surveillance Applications ", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 20, No. 2, February 2010.*

[57] *S. Saravanakumar, A. Vadivel and C.G. Saneem Ahmed," Human object tracking in video sequences "ICTACT Journal on Image and Video Processing, August 2011, Volume: 02, Issue: 01.*

[58] *Yining Deng, Member, IEEE, and B.S. Manjunath, Member, IEEE, "Unsupervised Segmentation of Color-Texture Regions in Images and Video," IEEE Trans. Pattern Analysis And Machine Intelligence, vol.23, no. 8, pp. 800-810, Aug. 2001.*

[59] Christian Wohler, *"Autonomous in situ training of classification modules in real-time vision systems and its application to pedestrian recognition", Pattern Recognition Letters Vol.23, Pg. No. 1263–1270 .*

[60] *Shao-Yi Chien, Member, IEEE, Wei-Kai Chan, Yu-Hsiang Tseng, and Hong-Yuh Chen, "Video Object Segmentation and Tracking Framework With Improved Threshold Decision and Diffusion Distance," IEEE Trans. Circuits and systems for video technology, vol. 23, no. 6, pp. 921-934, Jun. 2013.*

[61] *Rana Farah, Student Member, IEEE, J. M. Pierre Langlois, Member, IEEE, and Guillaume-Alexandre Bilodeau, Member, IEEE, "Catching a Rat by Its Edglets," IEEE Trans. Image Process., vol. 22, no. 2, pp. 668–678, Feb. 2013.*

[62] *Jiyan Pan, BoHu, and Jian Qiu Zhang," Robust and Accurate Object Tracking Under Various Types of Occlusions ", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 2, February 2008*

[63] *Kirt Lillywhite et al, "A feature construction method for general object recognition", Pattern Recognition Vol.46, Pg. No. 3300–3314, 2013.*

[64] *G. L. Foresti, C. Micheloni, L. Snidaro, P. Remagnino, and T. Ellis, "Active video-based surveillance system: the low-level image and video processing techniques needed for implementation," Signal Processing Magazine, IEEE, vol. 22, pp. 25-37, 2005.*

[65] *C. Wohler and J.K. Anlauf , " real time object recognition on image sequences with adaptable time delay neural network approach- applications for autonomous vehicles", image and computer vision, Vol.19, Pg. No. 593-618.*

[66] *C. Papageorgiou and T. Poggio. (2000). A trainable system for object detection. IJCV, 38(1):15–33.*

[67] *Shahbe Mat Desa and Qussay A Salih. Image subtraction for real time moving object extraction. In Computer Graphics, Imaging and Visualization, 2004. CGIV 2004. Proceedings. International Conference on, pages 41–45. IEEE, 2004.*

[68] *H. Celik, A. Hanjalic, and E. A. Hendriks. Unsupervised and simultaneous training of multiple object detectors from unlabeled surveillance video. In Computer Vision and Image Understanding, 2009.*

[69] *Alan J Lipton, Hironobu Fujiyoshi, and Raju S Patil. Moving target classification and tracking from real-time video. In Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on, pages 8–14. IEEE.*

[70] *Budi Sugandi, Hyoungseop Kim, Joo Kooi Tan, and Seiji Ishikawa. Tracking of moving objects by using a low resolution image. In Innovative Computing, Information and Control, 2007. ICICIC'07. Second International Conference on, pages 408–408. IEEE, 2007.*

[71] *AF Bobick, JW Davis, "The recognition of human movement using temporal templates". IEEE Trans. Pattern Anal. Mach. Intell. 23(3), 257–267 (2001)*

[72] *Victoria Yanulevskaya, Jasper Uijlings, Jan-Mark Geusebroek, "Salient object detection: From pixels to segments", Image and Vision Computing Vol. 31, Pg. No. 31–42, 2013.*

[73] *Nicholas A. Mandellos et al, "A background subtraction approach for detecting and tracking vehicles", Expert Systems with Applications, Vol.38, Pg. No. 1619–1631, 2011 .*

[74] *A. Levin, P. Viola, and Y. Freund. Unsupervised improvement of visual detectors using co-training. In CVPR, 2003.*

[75] Dirk Walther et al, "Selective visual attention enables learning and recognition of multiple objects in cluttered scenes", Computer Vision and Image Understanding Vol. 100, Pg. No.41–63, 2005.

[76] Manuele Bicego et al, "A Hidden Markov Model approach for appearance-based 3D object recognition", Pattern Recognition Letters Vol.26, Pg. No. 2588–2599 .

[77] Bijan Shoushtarian and Helmut E. Bez, "A practical adaptive approach for dynamic background subtraction using an invariant colour model and object tracking", Pattern Recognition Letters Vol. 26, Pg. No. 5–26, 2005.

[78] Hui Chen and Bir Bhanu, "Efficient Recognition of Highly Similar 3D Objects in Range Images", IEEE Transactions on Pattern analysis and machine intelligence, vol. 31, no. 1, Pg. No. 172-179, January 2009.

[79] Carlos Cuevas and Narciso García , "Improved background modeling for real-time spatio-temporal non-parametric moving object detection strategies", Image and Vision Computing Vol.31, Pg. No. 616–630,2013.

[80] Ling Cai et. al., "Multi-object detection and tracking by stereo vision", Pattern Recognition Vol.43, Pg. No. 4028–4041, 2010.

[81] Bangjun Lei and Li-Qun Xu, "Real-time outdoor video surveillance with robust foreground extraction and object tracking via multi-state transition management", Pattern Recognition Letters Vol.27, Pg. No. 1816–1825, 2006 .

[82] D Moctezuma, C Conde, IM Diego, E Cabello, "Person detection in surveillance environment with HoGG: Gabor filters and histogram of oriented gradient", in ICCV Workshops (IEEE, Piscataway, 2011), pp. 1793–1800.

[83] A.Mohan,C.Papageorgiou , and T.Poggio, Example-based object detection in images by components, IEEE Trans.Pattern Anal.Mach.Intell.,vol 23.no.4, pp.349-361,Apr.2001.

[84] Ian Fasel, Bret Fortenberry and Javier Movellan, "A generative framework for real time object detection and classification'', Computer Vision and Image Understanding Vol. 98, Pg.No.182–210

[85] L. Han, H. Dong, J. McClelland, L. Han, D. Hawkes, et al. (2017). A Hybrid Patient-Specific Biomechanical Model Based Image Registration Method for the motion estimation of Lungs.Medical Image Analysis. 39, pp.87-100.

[86] Christian Goerick and Detlev Noll, "Artificial neural networks in real-time car detection and tracking applications", Pattern Recognition Letters Vol.17, Pg. No. 335-343.

[87] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax approach for mixed subgaussian and supergaussian sources," Neural computation, vol. 11, pp. 417-441, 1999.

[88] Mirmehdi, M. et al, "Feedback control strategies for object recognition", IEEE Transactions on Image Processing, Vol. 8, Issue: 8 , Pg. No. 1084 – 1101 .

[89] L Zhang, SZ Li, X Yuan, S Xiang, "Real-time object classification in video surveillance based on appearance learning", in IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2007 (CVPR 2007) (IEEE, Piscataway, 2007),pp. 1–8

[90] M. M. V. Gupta and S. Sawarkar, "Change Detection based Real Time Video Object Segmentation," in International Journal of Engineering Research and Technology, 2012.

[91] A. Bar-Hillel, D. Levi, E. Krupka, and C. Goldberg (2010). Part-based feature synthesis for human detection. In ECCV.

[92] Yu Zhong, Anil K. Jain, M.-P. Dubuisson-Jolly," Object Tracking Using Deformable Templates ",IEEE transactions on pattern analysis and machine intelligence, vol. 22.

[93] P. Felzenszwalb, R. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. In PAMI, 2010.

[94] Nijad Al-Najdawi et al., "An Automated Real-Time People Tracking System Based on KLT Features Detection",The International Arab Journal of Information Technology, Vol. 9, No. 1, January 2012.

[95] Yi Liu, Yuan F. Zheng, "Video Object Segmentation and Tracking Using ψ -Learning Classification ", IEEE transactions on circuits and systems for video technology, vol. 15, no. 7, July 2005.

[96] G. L. Foresti, C. Micheloni, and C. Piciarelli, "Detecting moving people in video streams," Pattern Recognition Letters, vol. 26, pp. 2232-2243, 2005.

[97] M. Elarbi-Boudihir and K. A. Al-Shalfan, "Intelligent Video Surveillance System Architecture for Abnormal Activity Detection," in The International Conference on Informatics and Applications (ICIA2012), 2012, pp. 102-111.

[98] S. Walk, N. Majer, K. Schindler, and B. Schiele. New features and insights for pedestrian detection. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2010.

[99] Swantje Johnsen and Ashley Tews," Real-Time Object Tracking and Classification Using a Static Camera "Proceedings of the IEEE ICRA 2009 Workshop on People Detection and Tracking Kobe, Japan, May 2009.

[100] Bahadir Karasulu and Serdar Korukoglu, "Moving object detection and tracking by using annealed background subtraction method in videos: Performance optimization", Expert Systems with Applications Vol. 39, Pg. No. 33–43, 2012 .

[101] M. B. Augustin, S. Juliet, and S. Palanikumar, "Motion and feature based person tracking in surveillance videos," in *Emerging Trends in Electrical and Computer Technology (ICETECT), 2011 International Conference on*, 2011, pp. 605-609.

[102] *K. K. Hati, P. K. Sa, and B. Majhi, "LOBS: Local background subtracter for video surveillance," in Microelectronics and Electronics (PrimeAsia), 2012 Asia Pacific Conference on Postgraduate Research in, 2012, pp. 29-34.*

[103] P. Antonakaki, D. Kosmopoulos, and S. J. Perantonis, "Detecting abnormal human behaviour using multiple cameras," Signal Processing, vol. 89, pp. 1723-1738, 2009.

[104] J Cezar Silveira Jacques, Claudio Rosito Jung, and Soraia Raupp Musse. Background subtraction and shadow detection in grayscale video sequences. In Computer Graphics and Image Processing, 2005. SIBGRAPI 2005. 18th Brazilian Symposium on, pages 189–196. IEEE, 2005.

[105] Ya Liu, Haizhou Ai, and Guang-you Xu. Moving object detection and tracking based on background subtraction. In Multispectral Image Processing and Pattern Recognition, pages 62–66. International Society for Optics and Photonics, 2001.

[106] Robert T Collins, Alan Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt, et al. A system for video surveillance and monitoring, volume 102. Carnegie Mellon University, the Robotics Institute Pittsburg, 2000.

[107] *Vibha L, Chetana Hegde, P Deepa Shenoy, Venugopal K R, L M Patnaik," Dynamic Object Detection, Tracking and Counting in Video Streams for Multimedia Mining ",IAENG International Journal of Computer Science,35:3,IJCS_35_3_16,21 .*

[108] *R Cutler, LS Davis, "Robust real-time periodic motion detection, analysis, and applications". IEEE Trans. Pattern Anal. Mach. Intell. 22(8), 781–796 (2000)*

[109] *T Ko, S Soatto, D Estrin, "Warping background subtraction", in 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010) (IEEE, Piscataway, 2010), pp. 1331–1338 .*

[110] *Sebastien Lef evre and Emmanuel Bouton and Thierry Brouard and Nicole Vincent," A new way to use hidden markov models for object tracking in video sequences "*

[111] *Robert Bodor, Bennett Jackson, Nikolaos Papanikolopoulos," Vision-Based Human Tracking and Activity Recognition",*

[112] *Qian Zhang and King Ngi Ngan," Segmentation and Tracking Multiple Objects under Occlusion from Multiview Video ", IEEE Transactions On Image Processing, Vol. 20, No. 11, November 2011.*

[113] *Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., volume 2. IEEE.*

[114] *Sourabh Khire & Jochen Teizer," Object Detection and Tracking ", 2008.*

[115] *T.Zhao,R.Nevatia,and B.Wu, Segmentation and tracking of multiple humans in crowded environments, IEEE Trans.Pattern Anal.Mach.Intell.,vol30,no.7,pp.1198-1211,Jul.2008 .*

[116] *Nicoletta Noceti et al, "Spatio-temporal constraints for on-line 3D object recognition in videos", Computer Vision and Image Understanding Vol.113 Pg. No. 1198–1209, 2009.*

[117] *L. Han, M. haleem, M. Taylor (2016). Automatic Detection and Severity Assessment of Crop Diseases using Image Pattern Recognition. Studies in Computational Intelligence. 647, pp.283-300.*

[118] *N. Murray, D. Roberts, A. Steed, P. Sharkey, P. Dickerson, J. Rae and R. Wolff, Eye Gaze in Virtual Environments: Evaluating the need and initial work on implementation, In Concurrency and Computation: Practice and Experience, Wiley Publishers, ISSN 1532-0626, 2009.*

[119] *R. Wolff, D. Roberts, Alessio Murgia, Norman Murray, John Rae, Will Steptoe, Anthony Steed, and Paul Communicating Eye Gaze across a Distance without Routing Participants to the Spot, Sharkey, Proc. 11th IEEE ACM Int. Symp. Distributed Simulation and Real Time Applications DSRT, 2008, 2nd in best paper award and shortlisted for journal.*

[120] *M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," IEEE transactions on pattern analysis and machine intelligence, vol. 28, pp. 657-662, 2006.*

[121] *Benayed Y, Fohr D, Haton JP, Chollet G (2003) Confidence measures for keyword spotting using support vector machines. In: Acoustics, Speech, and Signal Processing. Proceedings.(ICASSP'03). 2003 IEEE International Conference on, pp. I-I*

[122] *Broggi A, Bertozzi M, Fascioli A, Sechi M (2000) Shape-based pedestrian detection,^ in Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE, pp. 215–220*

[124] University of Utah, CS 4640: Image Processing Basics, Spring 2012

[125] Umbaugh, Scott E (2010). Digital image processing and analysis : human and computer vision applications with CVIPtools (2nd ed.). Boca Raton, FL: CRC Press. ISBN 978-1-4398-0205-2

[126] P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: an evaluation of the state of the art.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 4, pp. 743–61, Apr. 2012.

[127] M. A. R. Ahad, J. Tan, H. Kim, and S. Ishikawa, "Action Dataset – A Survey," in SICE Annual Conference, 2011, pp. 1650–1655.

[128] L. Fei-Fei, R. Fergus and P. Perona. Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. IEEE. CVPR 2004, Workshop on Generative-Model Based Vision. 2004