

RECOGNITION NORMAL AND ABNORMAL HUMAN ACTIVITIES BY IMPLEMENTATION K-NEAREST NEIGHBOR AND DECISION TREE MODELS

¹WALA'A N. JASIM, ²ESRA J. HARFASH

¹Department of Computer Science, College of Science, Basra University, Iraq

²Department of Computer Science, College of Information Technology, Basra University, Iraq

E-mail: ¹walaa1_1979@yahoo.com, ²esra_jasem_2011@yahoo.com

ABSTRACT

Recently the advancement of technology has led to develop human activities recognition (HAR) system in video. HAR system is one of the important fields in computer vision (CV). This paper presents a system for helping people who suffer from health problems and are stayed alone for long times especially the elderly, by recognizing three normal activities: (walking, drinking and eating) and six abnormal activities: (headache, vomiting, fainting, renal colic, intestinal colic, angina) from videos monitoring that are chosen from the daily life activities of elderly people. In this paper, we proposed iterative thresholding (global thresholding) to subtract background from foreground, and utilized four different features of extraction techniques: (Radon Transform, DWT, SURF and STIP). Finally, K-Nearest Neighbor (KNN) and Decision Tree (DT) are used to classify normal and abnormal activities in video. When the system recognizes any abnormal activities, the alarm system is activated by sending SMS email to the person who is concerned with the status of the patient. The system evaluates (Radon Transform, DWT, SURF and STIP) with K-NN and DT. LDA is used to select the best features from the four feature extraction techniques and PCA is used with DWT and SURF to apply dimension reduction to obtain the best results. In addition, this paper presents the combination between (STIP & Radon Transform) and (STIP & SURF) in order to rise the evaluation of the STIP technique and others features extraction techniques with DT. After combination, the performance of combined techniques was improved.

Keywords: *Human Activity Recognition(HAR), Normal and Abnormal Activity, Feature Extraction, Classification, K-NN and DT.*

1. INTRODUCTION

Human activities recognition is a challenging and an important area in computer vision. Among them, abnormal activity recognition plays a main role in an intelligent monitoring [1], especially monitoring the people who suffered from a health problem and that stayed alone for long times such as the elderly.

The aim of activities recognition is to observe the actions of one or many agents from sequences of observation on the actions for agents and the environment conditions [2]. Recognition of activity and recognition of action are the most widespread utilized expressions. The expression action is predominately confused with the expression activity. Action, commonly, points to a series of primary movements done by a single object, that is represented in an atomic movement that has

ability for describing by the limbs, for example a walking step. As for, activity includes a number of successive actions. For example, dancing activity includes sequential repetitions of many actions, such as (waving hand, walking and jumping,) etc. Activities can be placed on a higher level than actions [3].

The term "human activity" means "something that people do or cause to happen" or "activities are considered more complex and include coordinated actions among number of persons" [4]. Recently, human activity recognition (HAR) has attracted a lot of interests in the video analysis technology area because of the increasing requests in a lot of applications. The rising demand for security and safety has resulted in many researches of monitoring intelligent. It has a vast range of applications, for example: entertainment environments, surveillance environments, healthcare systems (hospitals,

eldercare, home-nursing) , monitoring people in big waiting rooms, centers of shopping, observing vehicles inside/outside cities or in campus , on highways, bridges and in tunnels etc. [3][5] . Majority of these applications request an automated recognition of high-level activities, that include multi simple (atomic) actions of human [6].

The goal of this paper is to design HAR system has ability recognition normal activities: (walking , drinking and eating) and abnormal activities : (headache , vomiting , fainting , renal colic , intestinal colic and angina) from the patient's monitoring video especially which it is need emergency helping in order to serves the healthcare area. All activities in our dataset are taken from the daily life activities for people especially of elderly people living lonely . HAR systems are needed to include the safety of elderly persons and let them live an life independently in their place residence.

This paper proposes approaches to complete the recognition of human activities recognition preprocessing step by using iterative thresholding method for subtracting background . In this step, we can solve the problem of similarities between different activities that seem similar , (Radon Transform , STIP , SURF and DWT) for extracting features and combination(STIP & Radon Transform) and (STIP & SURF) which are experimented only in this paper, LDA and PCA are used for dimension reduction . Finally, K-NN and DT are used for classification of the types of activities which are found in our dataset.

The alarm system is activated when the system recognizes one of the abnormal activities by sending SMS email to the person who is concerned with the status of the patient. Whilst when the system recognizes one of the normal activities , the alarm step doesn't need to activate .

The rest of the paper is organized as follows; after showing the introduction in section 1, the related work is given in section 2. Section 3 shows how the challenges for HAR system. Section 4 explains the database that used for monitoring human activities. Section 5 shows the main steps and all techniques used to complete the proposed system whilst Section 6 provides the results and performance evolution of the proposed system. Section 7 gives suggestion for future works . Finally, the conclusion is given in the last section.

2. RELATED WORKS

In this section, the methods which are applied in this research paper were state of the art in background subtraction , feature extraction and classification in order to get the best results for recognizing normal and abnormal human activities. In the past, most of abnormal HAR systems presented only fainting or falling activities recognition systems [8,9] but recently, focuses became more on other abnormal activities such as headache , vomiting , fainting , renal colic , intestinal colic and angina which highlighted this paper. In [8] Derek Anderson et al. In 2009 proposed a technique in order to recognize human activity from linguistic summarizations of temporal fuzzy inference curves representing the states of a 3-D object called voxel person. The fuzzy logic hierarchy is utilized where the producing from every level is summarized then fed into the next level. They presented a two level model for falling detection. First level inferred the states of the human(person) at every image. Second level is operated on linguistic summarizations of voxel person's states and inference regarding activities performed. The rules utilized for detecting falling were designed under the supervision of healthcare committee to ensure that they reflect the way in that elderly perform these activities. In [9] Caroline et al. in 2011 presented a new health care system in order to service the elderly person at home . This is done by propos a new health care system in order to serve the elderly person at home by detecting the falling status via analyzing shape of individual deformation during sequence of videos. Matching of shape method is utilized for tracking human's silhouette along the sequence of videos .Then, the deformation of shape is quantified from these silhouettes based on shape analysis techniques. Finally, the fallings are detected from normal activities by utilizing a Gaussian mixture model. Zafar et al. 2013 in [10] presented a hierarchical HAR system for recognizing abnormal activities from the daily activities of elderly person who are living lonely. This system contained two levels of features extraction and activities recognition. The first level consisted of KDA ,R-transform, K-means and HMM for recognizing the activity from video. As for the second level consisted of KDA, k-means and HMM, that is applied selectively to recognize activity from the first level when it belongs to the selected group. The validation of the system via six abnormal activities; backward falling, forward falling, fainting, paining chest, headache, and

vomiting and the normal activity is walking. In [11] J.L.Mazher Iqbal et al. 2015 proposed system for recognizing the abnormality activities and presented effective solution for detecting these abnormalities. The aim of this system is to serve the elderly people . This system first captures video utilizing webcam to detect the abnormality activities . The captured video splitted into frames and extracted the features for each frame like boundaries and edges by utilizing scale invariant feature transform (SIFT). These extracted feature vectors are used with HMM to recognize the abnormal activities. The proposed algorithm is tested on six different activities (forward fall, regressive fall, midsection agony, swoon, regurgitation, and migraine).

3. CHALLENGES FOR HUMAN ACTIVITIES RECOGNITION SYSTEMS

Regarding human activity recognition, there are still several challenges and issues that spur the evolution of activity recognition methods to get better the accuracy under more realistic status. We will shed some light on the most important challenges related to this topic in the following :

- ✓ Collection of data: collection the data of training under real conditions[12]
- ✓ The behavior of human : performing many tasks and at the same time making the recognition of activity process more complicated [13] .
- ✓ Multiple People : More than one person can exist in the same environment(place) [12] .
- ✓ Intraclass variability: the same activity is done by different persons differently [12].
- ✓ Interclass similarity: statuses that are fundamentally various, but that is seemed very similar [12].
- ✓ Data loss that is resulted from projection of a three dimensional scenes to a two dimensional image noise (e.g.: salt and pepper noise, fixed pattern noise, banding noise) .
- ✓ The cluttered and dynamic background .
- ✓ Conditions of lighting (like: outdoor, indoor , night ,morning, etc) .
- ✓ Conditions of weather (like: foggy, rainy, windy, sunny, etc) .
- ✓ Change of illumination, waving clothes in the air, light reflections, shadows flying birds, etc), change of object appearance, change of object shape
- ✓ Humans number in the scene .

4. DATABASE

The video monitoring dataset contains nine classes of human activities that are recorded with seven persons and are divided into two types : (normal and abnormal) activities. The normal activities contains three classes (walking , drinking and eating) and the abnormal activities contains six classes (headache , vomiting , fainting , renal colic , intestinal colic and angina),that are performed in different scenarios , several times (morning , afternoon and evening) , sunny and humid weather , and different distance between camera and person.

The activities are performed in room with good lighting condition , homogeneous and limited variation background indoor with a static camera (canon 700D) with 25 fps frame rate . The resolution of each video 640 × 480 pixels and has a length of three seconds for each video . All videos are stored using AVI file format . Table (1) below summarizes the specifications of dataset ,and in figure (1) is shown some examples of activities that are depended in this paper are shown.

TABLE 1: Described the database

	Normal Activities	Abnormal Activities
Classes	- Number of activities = 3 - Verbs: walking , drinking ,eating	- Number of activities = 6 - Verbs: headache , vomiting , fainting , renal colic , intestinal colic , and angina
The total of classes = 9		
Video	- Number of videos = 216	- Number of videos = 432
	The total of videos = 648	
	- Resolution = 640 × 480 pixels - Length of each video = 3 seconds - Frame rate = 25 frames / seconds - Static camera (canon 700D)	
Properties	- Homogeneous and limited variation background indoor - Performed by seven persons with various clothes, appearance, body and skin color .	



Fainting

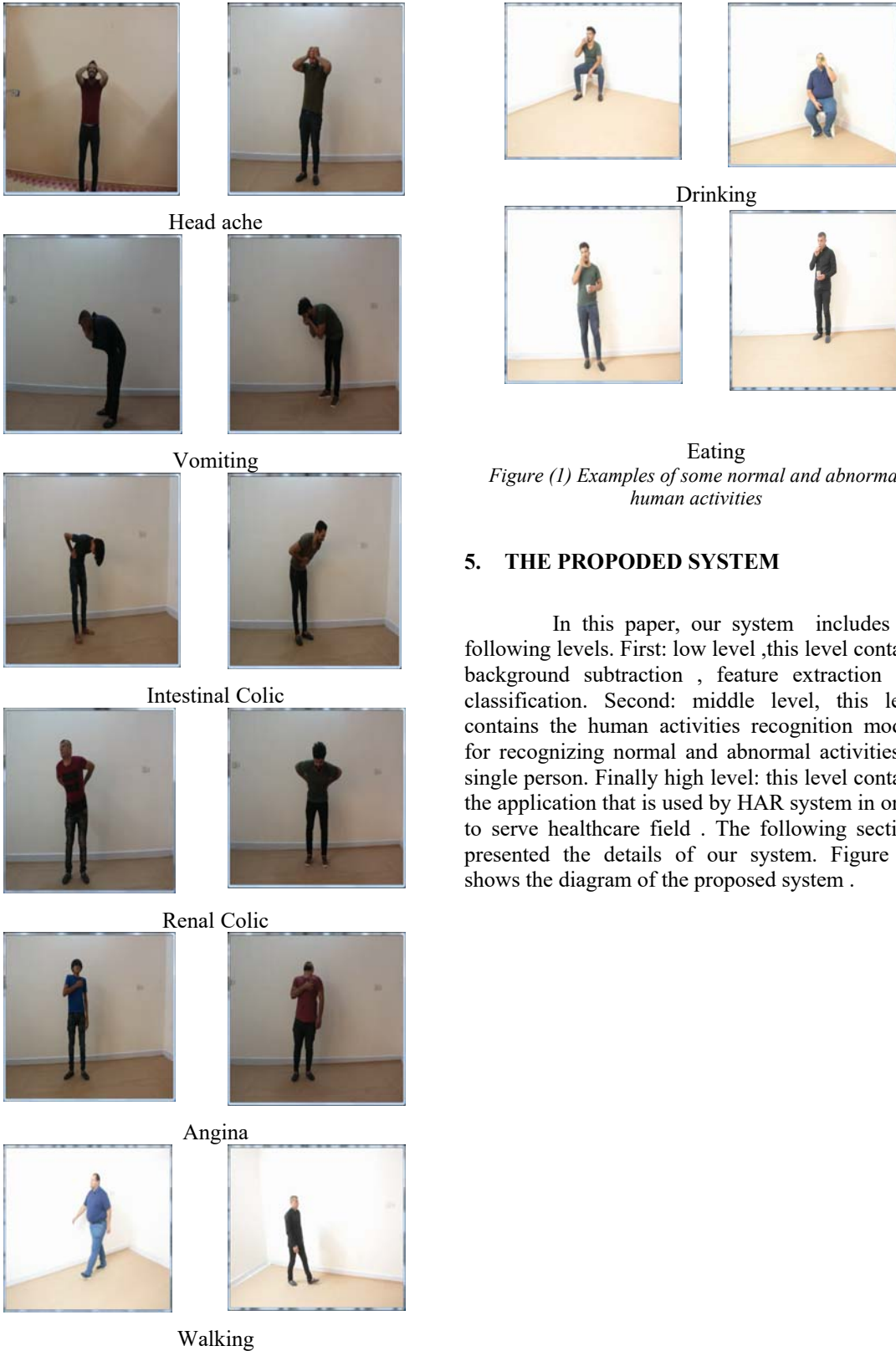


Figure (1) Examples of some normal and abnormal human activities

5. THE PROPOSED SYSTEM

In this paper, our system includes the following levels. First: low level, this level contains background subtraction, feature extraction and classification. Second: middle level, this level contains the human activities recognition models for recognizing normal and abnormal activities of single person. Finally high level: this level contains the application that is used by HAR system in order to serve healthcare field. The following sections presented the details of our system. Figure (2) shows the diagram of the proposed system.

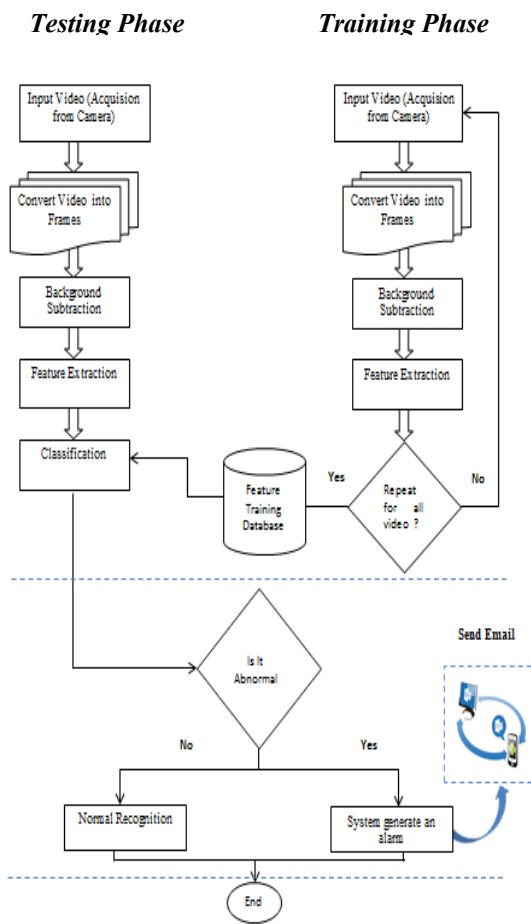


Figure (2) The block diagram of the proposed system

5.1 Pre-processing

This step represents a critical step because of the success of the next phases of the system depends on it as a strong base. Therefore, this step is very significant and it is fundamental. The bellow paragraphs show the details of this stage in the presented system.

5.1.1 Read Video

Each read video consists of sequence of frames (which are number 90 frames). Every frame of the video sequences is considered as a color image. Figure(3) and figure(4) show four frames of different positions of inputted video.



Figure (3) Frames sequence of video for two abnormal activity ,(a) Angina , (b) Fainting





Figure (4) Frames sequence of video for two normal activity , (a) Walking , (b) Drinking

5.1.2 Background Subtraction

In order to detect the foreground object efficiently in video, you must search for robust techniques for background modeling. The diverse background modeling techniques for background subtraction are more commonly in usage . The background subtraction is a process of dissociating the foreground (object) from background of the image. The Iterative thresholding method is one kind of global thresholdings which is proposed in this system to subtract background from foreground easily and accurately.

The Global Thresholding :

When T only depends on gray-level values , this thresholding technique is called “global thresholding” technique[14] . When selecting the suitable single threshold value T to input image ,the image 's pixels are divided into two classes background and foreground, that is the global thresholding explained in the following equation

$$g(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq T_{global} \\ 0, & \text{if } f(x,y) < T_{global} \end{cases} \quad (1)$$

Where:

$f(x,y)$ is an original image , $g(x,y)$ is a binary image, T is a threshold value[15].

There are some of global thresholding techniques like : "minimum thresholding, histogram concave analysis, Otsu, optimal thresholding, entropy-based thresholding, iterative thresholding, MoM-keeping Thresholding" [14]. This paper is used “iterative thresholding” that is one the most famous global thresholding techniques .

Iterative Thresholding Method:

Iterative thresholding is a global thresholding method. It is comparatively easy

technique that does not require a lot of specific knowledge of the image and is strong against noise of image [16]. In [17] procedure to compute the intensity threshold T automatic which is described by Gonzalez and Woods in (2002). In the following, we summarized the iterative thresholding method that is applied in our system to separate background from foreground.

Input : For all RGB images (frames) of inputted video

Output : Cell array contain binary level images.

Step1 : Convert images to gray level images : $F = \text{grayscale}(\text{image})$.

Step2 : Estimate the threshold T of image F .

Step3: Segment the F to two group $G1$ according to the conditions $F < T$ and $G2$ according to the conditions $F \geq T$.

Step4: Calculate the mean of $G1$ and $G2$, $MG1 = \text{mean}(G1)$, $MG2 = \text{mean}(G2)$.

Step5 : Calculate the new threshold T_{NEW} by the following equation (2):

$$T_{NEW} = 0.4 * (MG1 + MG2) \quad (2)$$

Step6 : Find $D = |T - T_{NEW}| < 0.1$.

Step7 : Repeat steps from 3 until 6 if the difference D is less than 0.1 .

Step8: Converts the grayscale image F to a binary image.

$$R = \begin{cases} 1, & \text{if } F \geq T \\ 0, & \text{if } F < T \end{cases} \quad (3)$$

Figure (5) show the result of background removed according to the above steps.

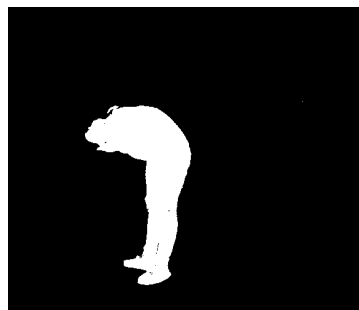


Figure (5) An example of background removed subtraction.

After the background subtraction step , the index of the binary images are calculated which has three colors and then take the coordinate of pixels that equals one. This step solved the problem of high similarities in postures of various activities because some various activities seemed very similar ,as in figure (6) .



Figure (6) The colored image without background subtraction.

5.2 Background of Feature Extraction Techniques

Extraction of feature played an important role in image processing[18]. It is the main step in HAR system; features that are extracted useful for pattern recognition and/or classification [19]. It is desirable to extract features which focus on discriminating between classes. In this stage, the characteristic of segmented object is extracted such as colors, motions, silhouette and shape. In general, features can be classified into four classes: space-time information, local descriptors, frequency transform, body modeling [5]. This paper introduces four various types of feature extraction techniques which are used in image processing, in addition to the combination of some of these techniques, where reviews the definitions and notations and using of these techniques.

5.2.1 Radon transform

Radon transform is famous for its vast domain of imaging application. It has been defined in numerous various of notations and forms. In mathematics, 2D radon transform is transform consists of the integral of a function over a set of lines in each directions, that is approximately equal for finding the projection of a shape on any given line. The following equation is the most used and widespread one[20] [21].

$$TRf(p, \Theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \Theta - y \sin \Theta - p) dx dy = \text{Radon}\{f(x, y)\} \quad (4)$$

Where :

$\Theta \in [0, \pi]$, $p \in [-\infty, \infty]$ and $\delta(\cdot)$ is the Dirac delta function,

$$f(\omega) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{Otherwise} \end{cases} \quad (5)$$

Where :

D is the binary silhouette, p and Θ are the position, $R_f(p, \Theta)$ is integral of f over the line

$$p = x \cos \Theta + y \sin \Theta.$$

On along lines various directions, the Radon transform consists the summation of intensities of pixel. The center point of the silhouette is selected as referencing point. Via utilizing the radon transform of silhouettes, a map between coordinate system(x,y) and the system of the radon transform (p, Θ). Figure (7) shows the radon transform definition. Giving a defined direction, the silhouette is projected to ρ axis. In other words, pixels along a group of parallel of lines to the line that is perpendicular to ρ axis are summed together [20].

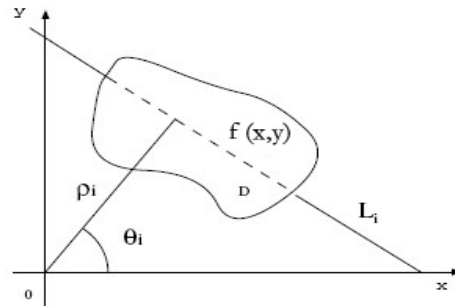


Figure (7) The Radon transform definition

Figure (8) shows some silhouettes and their corresponding radon transform. The original silhouettes are little various because of shadow and noise. Nevertheless, their radon transforms are not much various.

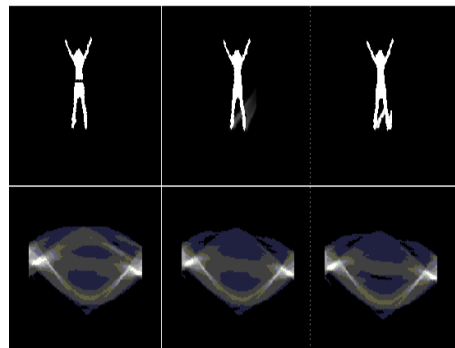


Figure (8) Silhouettes and their the Radon transform

Radon transform taken with regard to center of the silhouette enjoys translation invariance and it is also not difficult to scale. As for the rotation

invariant, Radon transform doesn't guarantee it. But in recognition of action, the rotation invariant isn't required and occasionally the rotation invariant is not good for recognizing the action. As well, Radon transform itself is not speedy sufficient for achieving recognition in real time [20]. On the other hand, for geometry transformation like: rotation, scaling, and translation, the radon transform has the above characteristics for a scaling factor α , and it is sensitive to them, as in the (6), (7) and (8) equations [21].

✓ Scaling [21]

For a scaling factor α ,

$$\text{Radon}\{f(\frac{x}{\alpha}, \frac{y}{\alpha})\} = \frac{1}{\alpha} T_{\text{Rf}}(\alpha p, \Theta) \quad (6)$$

✓ Translation [21]

For translation of vector $\vec{\mu} = (x_0, y_0)$,

$$\text{Radon}\{f(x - x_0, y - y_0)\} = T_{\text{Rf}}(p - x_0 \cos \Theta - y_0 \sin \Theta, \Theta) \quad (7)$$

✓ Rotation [21]

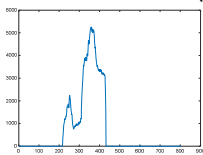
For rotation of Θ_0

$$\text{Radon}\{f \Theta_0(x,y)\} = T_{\text{Rf}}(p, \Theta + \Theta_0) \quad (8)$$

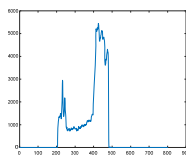
This paper used this transform to extract the features of input image by calculating the projection of the image intensity along the radial line directed at the determined angle(Θ). In this work, the theta ($45^\circ, 90^\circ, 135^\circ$ and 180°) are selected. When theta equals 90° , the result is the best from other angles in both classifier models (K-NN and DT), and the less result when theta equals 180° . In figure (9), example of the features of radon transformation for one frame according to each selected angle is shown.



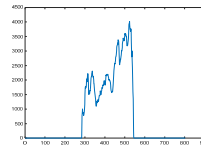
(a)



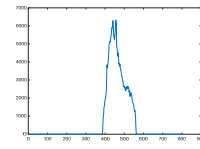
(b)



(c)



(d)



(e)

Figure (8) Figure (9) The extract feature using Radon transform (a) Image after background subtraction, (b) $\Theta=45^\circ$, (c) $\Theta=90^\circ$, (d) $\Theta=135^\circ$, (e) $\Theta=180^\circ$

5.2.2 Discrete Wavelet Transform (DWT)

Wavelet is considered more common method in image processing and computer vision. Several applications, like recognition, detection, image retrieval and compression have been implemented by it. The major causes that made Wavelet transforms more popular are its complete theoretical, the large pliability for selecting bases and the low computational intricacy[22]. In general, the Wavelet transform can be illustrated in the following equation [23]:

$$F(a,b) = \int_{-\infty}^{\infty} f(x) \Psi_{(ab)}^*(x) dx \quad (9)$$

Whereas: "*" is complex conjugate symbol and " Ψ " is some function [23].

There are two kinds of Wavelet transform: first, Continuous Wavelet Transform (CWT) and second, Discrete Wavelet Transform (DWT) [22]. This paper will use the second kind of Wavelet transform in extracting feature stage.

Discrete Wavelet Transform (DWT) is common kinds of Wavelet Transform that can be used for feature extraction such as face recognition because its insensitivity to great variation in light direction and face pose [22]. The Discrete Wavelet Transform (DWT) is an executing of the Wavelet transform utilizing a discrete group of the wavelet translations and scales obeying known rules. This type of transform decomposes the signal into interchangeably orthogonal set of wavelets, that is the major variance from the Continuous Wavelet Transform (CWTs) [23].

The basis of DWT get back to "1976" when the techniques for decomposing the discrete time signals were designed. The DWT is calculated via sequentially low-pass and high-pass filtering of the discrete time-domain signal that is shown in figure (10). This algorithm is called Mallat-tree decomposition or Mallat algorithm, that represents the famous and oldest algorithm [24] [25]. In figure (10), the signal is indicated via sequence $x[n]$, where n is an integer. The low-pass filter is referred to via G_0 whilst the high-pass filter is referred to

via H_0 . In each level, the high-pass filter produces detail information, $d[n]$, whilst the low-pass filter is associated with scaling function produce coarse approximation, $a[n]$ [24].

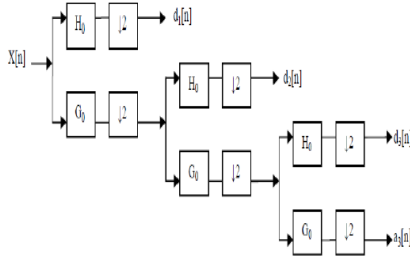


Figure (10) Three-level wavelet decomposition tree

On the other hand, the major advantages of the DWT the space frequency decomposition of image is provided by it, Unlike, DCT Fourier transform and discrete cosine transform that merely equips frequency decomposition. Via providing space frequency decomposition, the DWT lets energy compaction at the low frequency sub bands and the space localization of edges at the high frequency sub bands. Moreover, the DWT doesn't present a blocking effect at the low data rates[27]. DWT is based on the sub-band coding that is found to produce a fast calculation of wavelet transform. It is easy for implementing and reducing the calculation time and required of resources [26]. Finally, DWT can be used for easy and fast denoising of noisy signals [23].

In our system, the PCA is implemented on this type of features(DWT) to reduce the high dimensions (242×322) of DWT features to (1×242) vector feature for each image. Figure (11) shows the coefficient approximate example of the DWT.



Figure (11) The DWT feature ,(a) Image after background subtraction, (b) CA coefficient

5.2.3 Speeded Up Robust Features (SURF)

The SURF is a robust and fast method for local that is proposed by Herbert Bay[28]. The SURF is based on the same stages and principles of the SIFT, but it uses a various scheme and it must provide better results and faster[29]. The

authenticity of the SURF method is that these operations are speeded up via using the box filters techniques[30]. Box filters techniques which are included exchange each pixel of an image with the average in the box. When it is expanded in some simple ways, it becomes an effective popular purpose tool for the image processing [31].

The SURF algorithm is based on two successive steps (the feature detection and the feature description). The figure (12) is shown the entire SURF structure, where it is splitted between interest point detection and interest point description [28].



Figure (12) Steps of SURF algorithm

5.2.3.1 Interest Point Detection [28] [32]

The first step for interest point detection generates the integral image from the original image as in equation (12). The previous equation calculates the summation of all pixel values within a region of rectangular from beginning point to a location of pixel.

$$I_{\Sigma}(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(x, y) \quad (10)$$

Where the point of x, y refers to a location of pixel. The typical technique is computed from many pixels in original image for the sum of region of the rectangular but this method uses integral image only needing 4 points, Figure (13) indicates this method.

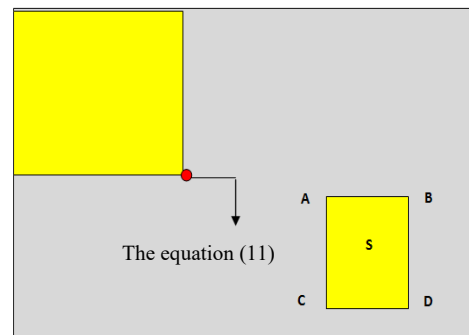


Figure (13) The way calculated area in the integral image

$$S = \Sigma(D) - \Sigma(B) - \Sigma(C) + \Sigma(A) \quad (11)$$

The second step extracts interest point using equation (2.23) of Hessian matrix. The approximate second order Gaussian simplified Gaussian filter is utilized to facilitate

calculations . Figure (14) shows the filters of approximated Hessian matrix. These point to D in equation (13).

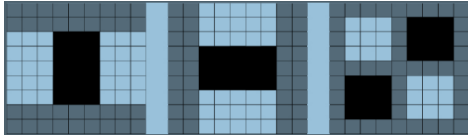


Figure (14) Filters of approximated Hessian matrix(D_{xx} , D_{yy} , D_{xy})

5.2.3.2 Interest Point Description [29] [32]

The aim of descriptor is to equip a unique and strong description of feature, that means each interest point must have a unique description that doesn't rely on the features rotation and scale.

A descriptor can be generated based on the surrounding of area an interest point. The descriptor of SURF is based on Haar Wavelet responses and can be computed efficient with integral images. The scheme of "Haar Wavelet " is needed to choose the orientation. Via determining a unique orientation for an interest point, it is possible for achieving rotation invariance. Before the descriptor is computed the interest area surrounds the interest point is rotated to its direction. The SURF descriptors are strong to rotations and an upright version, U-SURF, must be strong for rotations $\pm 15^\circ$ without performing an orientation assignment .

The SURF descriptor describes an interest area with size $20s$. The interest area is splitted into " 4×4 " sub-areas which are described via the values of a wavelet response in the x and y directions . The wavelet response in the x and y directions is indicated to as dx and dy consecutively. The wavelets utilized to computed the response is shown in Figure (15).

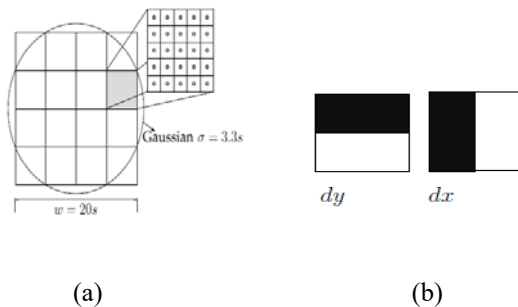


Figure (15) (a) $20s$ areas is divided into 4×4 subareas that are sampled 5×5 times to get the wavelet response in (b)

The white and black areas correspond to a weight 1 and -1 for the Haar kernels. They are utilized with a filter size of $2s$.

$$v = \{ \sum dx, \sum |dx|, \sum dy, \sum |dy| \} \quad (12)$$

In Figure (15) , for each subarea, a vector v is computed via the (13) equation relied on " 5×5 " sample. The descriptor for an interest point is the sixteen vectors for the subareas concatenated . Finally the descriptor is normalized, in order for achieving invariance to contrast difference which will represent themselves as a linear scaling of the descriptor.

The SURF size in this work is 128 while the default is 60 .The 128 size is selected because it gives more accuracy than if the value is less. As for the Metric Threshold (strongest feature threshold) is 600 while the default is 1000 . We reduce the value default of this threshold to obtain more blob . Figure (16) shows an example of SURF feature .

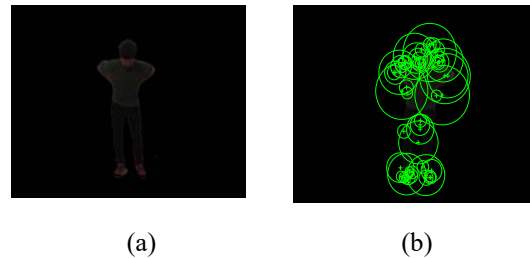


Figure (16) SURF feature , (a) Image after background subtraction (b) The strong points

5.2.4 Space-Time Interest points (STIP)

Laptev and Lindeberg proposed the 3D space-time interest points (STIPs) via extending Harris detectors. The first using of the STIP was recognition of action [33] . Spatial interest points in images are expanded to spatio-temporal local structure in video where the values of image have big local variations in both time and space [34], i.e. The features of local image are providing compacted and abstracted representations of images, for example : corners , expand the concept a spatial corner detector to a spatio-temporal corner detector [33] .

STIP has special interest because of it is a simple and strong low-level feature providing an effective description of moving object within videos. Generally, STIPs are based on luminance, and color has been largely ignored [35].

5.2.4.1 Interest Points in the Spatial Domain

An Image $f^{sp} : R^2 \rightarrow R$ can be modeled by its linear scale representations $L^{sp} : R^2 \times R_+ \rightarrow R$ define via convolution of f^{sp} with "Gaussian kernels" of variance σ_f^2 .

$$L^{SP}(x,y;\sigma_f^2) = g^{SP}(x,y;\sigma_f^2) * f^{SP}(x,y) \tag{13}$$

$$g^{SP}(x,y;\sigma_f^2) = \frac{1}{2\pi\sigma_f^2} e^{-\frac{(x^2+y^2)}{2\sigma_f^2}} \tag{14}$$

The concept of the Harris interest point detector is for finding spatial locations where f^{sp} has significantly changes in both directions. To search for interest points, one analyzes the matrix of second moments matrix integrated over the Gaussian window with variance σ_f^2 .

$$\mu^{SP}(\cdot; \sigma_f^2; \sigma_f^2) = g^{SP}(\cdot; \sigma_f^2) * \begin{pmatrix} (L_x^{SP})^2 & L_x^{SP} L_y^{SP} \\ L_x^{SP} L_y^{SP} & (L_y^{SP})^2 \end{pmatrix} \tag{15}$$

L_x^{SP} and L_y^{SP} are Gaussian derivatives computed at local scale σ_f^2 . Two significant large values of λ_1, λ_2 refer to the presence of an interest points. For detecting such points, the corners are chosen in the image. Therefore detect positive maxima of the following function (17):

$$H^{SP} = \det(\mu^{SP}) - k[\text{tr}^2(\mu^{SP})] = \lambda_1\lambda_2 - k(\lambda_1 + \lambda_2)^2 \tag{16}$$

The result of detecting Harris interest points in an outdoor image sequence of a walking person is shown in figure (17) [33].

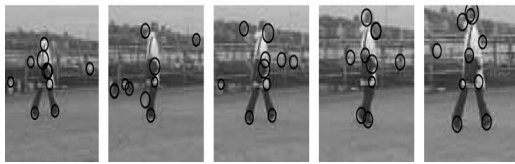


Figure (17) Pure spatial interest point detector selects both moving and stationary points in the image sequence

5.2.4.2 Interest Points in the Spatio-Temporal Domain [33]

Here has developed an operator which responded to events in temporal image series at specified locations and with specified extents in space time. The concept is to expand the interest points notion in the spatio-domain via demanding the values of image in local spatial-temporal volumes to have large divergences along both the spatial & temporal directions.

To model a spatio-temporal image sequence, the function $f : R^2 \times R \rightarrow R$ is utilized

and builds its linear scale-space representation $L : R^2 \times R \times R_+^2 \rightarrow R$ via convolution of f with an anisotropic Gaussian kernel with independent spatial variance and temporal variance. Consecutively σ_f^2, τ_f^2 .

$$L(\cdot; \sigma_f^2, \tau_f^2) = g(\cdot; \sigma_f^2, \tau_f^2) * f(\cdot) \tag{17}$$

$$g(\cdot; \sigma_f^2, \tau_f^2) = \frac{1}{\sqrt{(2\pi)^3 \sigma_f^2 \tau_f^2}} e^{-\frac{(x^2+y^2)}{2\sigma_f^2} - \frac{z^2}{2\tau_f^2}} \tag{18}$$

To search for interest points, one analyzes the matrix of second moments in the following equation (19), and the maxima of the following spatio-temporal corner function in equation (20).

$$\mu = g(\cdot; \sigma_f^2, \tau_f^2) * \begin{pmatrix} (L_x)^2 & L_x L_y & L_x L_t \\ L_x L_y & (L_y)^2 & L_y L_t \\ L_x L_y & L_y L_t & (L_t)^2 \end{pmatrix} \tag{19}$$

$$H = \det(\mu) - k \text{tr}^3(\mu) = \lambda_1 \lambda_2 \lambda_3 - k(\lambda_1 + \lambda_2 + \lambda_3)^3 \tag{20}$$

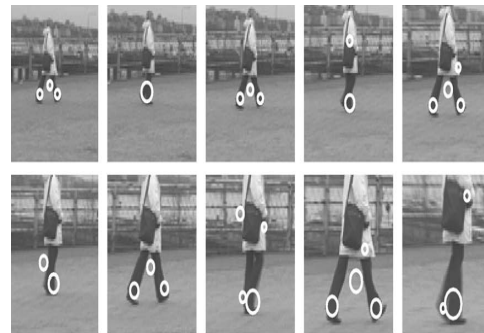


Figure (18) Spatio-Temporal interest point detector for a zoom-in sequences of a walking person

5.3 Dimension Reduction (DR) And Feature Selection

DR is an important task which played a major role in statistics, machine learning and others. DR can be done via either reducing the number of features; a task called feature selection, or via reduce the number of patterns; called data reduction [36]. Dimensionality reduction of data produces a compacting low-dimensional encoding for a given high-dimensional dataset [37]. Besides DR can minimize the amount of storage needed via reducing the size of the data sets. DR can help to understand the data sets via discarding any irrelevant features, and to focus on the major

important features, and DR can enable the detecting of rich information[38].

Feature extraction and dimension reduction can be combined in one step. There are different techniques used to accomplish this purpose such as principal component analysis (PCA), linear discriminant analysis (LDA), Singular Value Decomposition etc. [39]. This paper is utilized LDA and PCA methods to minimize the dimensionality of the feature space .

5.3.1 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is one of the most important supervised dimension reduction techniques [7] , which represents a very major tool in a wide difference of problems. It is commonly utilized in machine learning problems such as : face and gesture recognition ,pattern recognition, , feature extraction , data dimensionality reduction and data classification [40][41].The LDA is proposed by R. Fisher in 1936, so it is also known as Fisher discriminant analysis (FDA) [42] .

The LDA is a “classical” method in the patterns recognition where it is utilized for finding linear features combination that is characterized or split up two or more classes of objects or events. The result of combination may be utilized as a linear classification or, more common, for dimensions reduction before it can be classified . The LDA is worked on only top Eigen values of small datasets. It doesn't properly work on large dataset because of its performing scalability reduces on large datasets [43]. The idea of LDA can be summarized in the following steps [44] :

- ✓ Training set $X = \{ x^{(n)} \in R^d \}_{n=1}^N$, $Y = \{ y^{(n)} \in L \}_{n=1}^N$, $L = \{ l_1, l_2, \dots, l_c \}$.
- ✓ Define X_i as the feature set contain all features samples with label l_i , and the number of sample of label l_i in X_i is denote as N_i .
- ✓ Define the class mean vector as , the following equation (21)

$$\mu_i = \frac{1}{N_i} \sum_{x^{(n)} \in X_i} x^{(n)} \tag{21}$$

- Define the total vector mean as , the following equation (22)

$$\mu = \frac{1}{N} \sum_{n=1}^N x^{(n)} \tag{22}$$

- ✓ Define the between-class scatter matrix as the following equation (23)

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) (\mu_i - \mu)^T \tag{23}$$

- ✓ Define within-class scatter matrix as the following equation(24)

$$S_w = \sum_{i=1}^c \sum_{x^{(n)} \in X_i} (x^{(n)} - \mu_i) (x^{(n)} - \mu_i)^T \tag{24}$$

Transform model of LDA [44]

- ✓ Desire to transform the original d-dimensional vector into p-dimensional ($p \leq d$).
- ✓ Find $d \times p$ transformation matrix W, then $z = W^T x$, where $z \in R_p$. Each column of W is a projection dimension .
- ✓ Get a transform set : $Z = \{ z^{(n)} \in R^p \}_{n=1}^N$, denote Z as a matrix. $p \times N$ matrix.
- ✓ LDA doesn't search to reconstruct the d-dimensional vectors.

5.3.2 Principle Component Analysis (PCA)

PCA is one of the most common technique for dimensionality reduction and feature extraction because of its simplicity is used in various applications. It uses orthogonal linear combination to obtain linear transformation space of the data sets but it doesn't work good with nonlinear data sets [37][38]. Sometimes ,the analyzing of patterns in a high dimensional data is complicated , so the PCA represents a strong tool to deal with such data[45].

For a set of data on n dimensional, PCA is aimed for finding a linear subspace of dimension d minimum than n such that the data points lie mainly on this linear subspace . The linear subspace can be specified via d orthogonal vectors which form a new coordinate system, called the principal components . The principal components are orthogonal, linear transformations of the original data points , therefore ,there can be no more than n of them [37]. The following is the PCA algorithm for given training set Y_1, Y_2, \dots, Y_N : [45] :

- ✓ Calculate the mean vector $Me = \frac{1}{N} \sum_{n=1}^N Y_n$ (25)

- ✓ Calculate the covariance matrix $C = \frac{1}{N} \sum_{n=1}^N (Y_n - Me) (Y_n - Me)^T$ (26)

- ✓ Calculate the eigenvalue/eigenvector pairs (λ_i, u_i) of C, $1 \leq i \leq N$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$.

- ✓ Calculate the first k principal components $z_j^{(i)} = Y_j^T u_i$, for each observation Y_j , 1

$\sum_{j=1}^n$, along the direction u_i , $i= 1, 2, \dots, k$.

5.4 The Classification

After choosing suitable features from image or video, classification algorithms are the next step under consideration for recognizing activities of human for achieving good recognition execution. It is major to select an appropriate classification algorithm utilizing the selected features representation[5].

The general concept of classification is one of data analyses forms which extracts models characterizing significant data classes. This model, called classifier, predicts categorical (discrete, unordered) class labels. The classification process of data consists of two-steps : "The learning step" (where is built a classification model) and "the classification step" (where the model is utilized for predicting labels of class for data)[46]. Next are two approaches for classification used in this work to find the decision boundaries between the classes which are K-NN and Decision tree models.

5.4.1 K-Nearest Neighbor (KNN)

K-Nearest Neighbors (K-NN) method is one of the most significant machine learning algorithms that is often utilized to classify objects based on the most similarity samples of training in the feature space, i.e. K-NN is based on closest training examples in the feature space. The classification is based on the distance between a set of inputted data points and training points. Different metrics can be utilized for determining the distance (Euclidean distance , Spearman distance Mahalanobis distance and etc.) [47]. The Euclidian distance is the most common metric which utilized to measure distance , which is computed as the (28) equation. If we have two points x and y where each point is an n -dimensions vector, i.e. $x = \{x_1, x_2, \dots, x_n\}$, $y = \{y_1, y_2, \dots, y_n\}$. According to Euclidean measuring , the distance function $d_E(x, y)$ between two points is defined [48].

$$d_E(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{27}$$

The K-NN is a supervised learning algorithm . If no model is learning from the data of training that is called " lazy learning " ,K-NN from this type . In contrast , before testing, the eager learning method learn data models [49].

The idea of K-NN is very easy and quite efficient in a lot of applications. It works as the following : when D is the training dataset ,

nothing on the examples of training is done. When the example of test is presented, the K-NN is comparing d with each example of training in D to calculate the distance or similarity between them . The k most (similar) example in D are then chosen. This set of examples is named the "K-nearest neighbors" of d . Then d takes the most recurrent class among the k closest neighbors .

For example figure (19) shows two classes of data ,the filled squares is " positive" and the empty circle is " negative" . If $k=1$, the test data point will classify as negative , and if $k=2$, the model cannot be decided ,if $k=3$ the class will be positive .

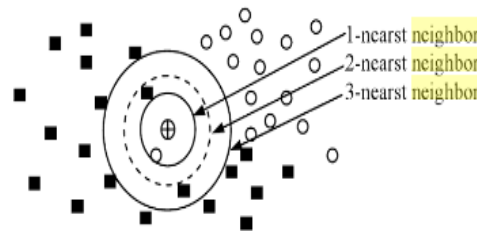


Figure (19) Example of K-NN classification

5.4.2 Decision Tree (DT)

Decision Tree (DT) is one of the most widespread classification algorithms presently that is used in machine learning and data mining . It is considered to be one of the powerful common supervised classification techniques which is constructed from the dataset of training [50]. DT is a hierarchical data structure consists of node and directed edges[51].

Generally, in the classification, the aim is to learn the decision tree that represents data of training such that labels for new examples that can be specified. Decision tree is a classifier for instances represented as vectors of feature (e.g.: shape=?; color=?; label=?;). Nodes are tests for values of feature , leaves assign the labels, and at each node there should be one branch for each feature value. Figures (20) and (21) show simple examples of DT . The structure of DT has three type of nodes [51] :

- ✓ The node is called "root node " when it has no incoming edges and zero or more outgoing edges .
- ✓ The node is called "Internal node" when it has one incoming edge and two or more outgoing edges .
- ✓ The node is called "leaf "or "terminal" nodes when it has one incoming edge and no outgoing edges .

Each leaf node specified a label of class. The non-terminal node that contains the root and others internal nodes includes the conditions of attribute test to split records which have various characteristics.

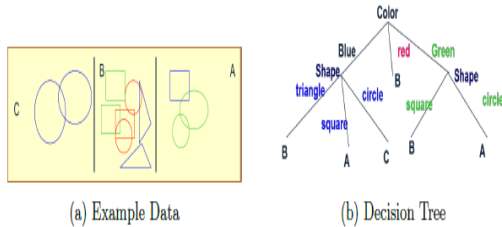


Figure (20) Example1 Decision Tree

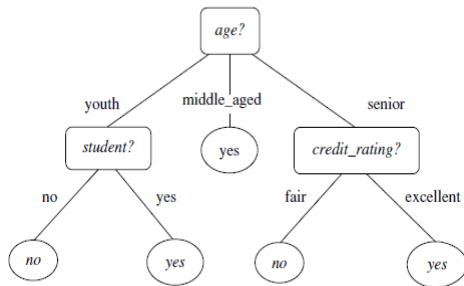


Figure (21) Example2 Decision Tree

Some decision tree algorithms are producing only a binary trees (where every internal node branches to punctually two other nodes), whereas others can produce non binary trees. In this paper the DT is Binary Decision Tree(BDT). The binary tree gives an interesting and often illuminating way of looking at data in regression or classification problems. It adds a flexible nonparametric tool to the data analyst’s arsenal[45] [52].

5.5 Alarm System

Alarm is an important step that is active only when the abnormality human activity happened, i.e. when an abnormal action is recognized, alert system is activated in the form of alarm, referring to need of helping [22]. It plays an important role in securing and safety, increasing awareness, and giving an alert of a potentially dangerous situations. Most of people interacted with alarms on a daily basis and it is important to depend on systems that are both reliable and well suited to specific needs. We know that the alarm systems are many and varied types such as : burglar alarm, fire alarm, car alarm and emergency alert

system (ESA)... etc. [53]. The alarm is usually used to alarm using a buzzer, lights, or recorded sounds. This paper utilized “Email” as type of alarm systems rather than hardwires, and that is simple and speedy because immediate destination person is alerted whether (doctor, nurse, or any person responsible for the patient). When any abnormal human activity is detected, an email will send right away.

Generally, the most medical alarm systems use body sensors in the event of any abnormal activity which are put on patients’ bodies and in their environment. Once aid is needed, the system notifies the medical personnel about the situation to help the patient quickly. For example, some HAR systems that are used wearable sensor called 'zigbee' is automatically detect an activity such as falling status, In the case of an emergency, the concerned party is notified by short message service (SMS) and email [54].

The e-mail is the public internet application, fundamental tool for business and communication and the most widespread way of communication today that is utilized more day after day. The Email is used for sending documents, text, video, audio and data of tables at home, work and several areas [55]. In relation to our work e-mail is used as an alert tool when abnormal activity occurred to the person by sending short message to the destination person.

6. PERFORMANCE EVOLUTION OF THE PROPOSED SYSTEM

The proposed methods have been run on the data of the experiment which it discussed in section 4. This section illustrates the results of the proposed system implementation which used the video in the indoor to recognize the health type of human activities from these videos. Different techniques are applied in order to obtain the best results. The success rate is calculated by using the following equation (28):

$$Success\ Rate = \frac{Number\ of\ success\ states}{Total\ number\ of\ states} \times 100\% \tag{28}$$

The proposed system tests the inputted video to recognize the class of health human activities by implementing different features with K-NN and DT models.

6.1 The results of K-NN model

The results of the accuracy recognition that are obtained of implement K-NN model on the current healthcare human dataset is discussed. The

results are shown for different features that are utilized in this work. Table (2) shows the recognition rate that is obtained before applying the method LDA technique with each type of feature .

Table (2) : The results accuracy of K-NN for normal and abnormal activities for each class and the total classes before using LDA

Class	DWT	SURF	Radon $\Theta=45^\circ$	Radon $\Theta=90^\circ$	Radon $\Theta=135^\circ$	Radon $\Theta=180^\circ$
Faint	100	36.11	86.11	94.44	77.77	86.11
Vomit	91.66	47.22	88.88	88.88	97.22	97.22
Headache	86.11	75	69.44	88.88	80.55	47.22
Angina	94.44	86.11	86.11	91.66	69.44	52.77
Intestinal Colic	94.44	77.77	94.44	94.44	100	91.66
Renal Colic	88.88	22.22	83.33	88.88	83.33	75
Walk	94.44	100	100	100	100	100
Drink	69.44	100	100	100	100	100
Eat	50	100	100	100	100	100
Total	85.49	71.60	89.81	94.13	89.81	83.33

In figures (22) is the charts that presented the measure of the differences among features total classes.

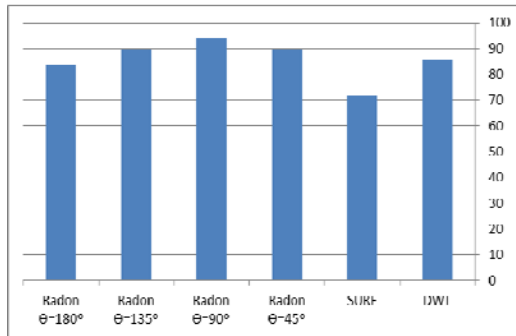


Figure (22) The results accuracy of K-NN for normal and abnormal activities for the total classes before using LDA.

From the above results, many of the types of features before implement the LDA give good results. STIP's results are ignored with K-NN .

Table (3) shows the recognition rate that is obtained after applying the LDA method . Figures (23) presents the chart of the difference in the results accuracy of classes with K-NN model and LDA .

Table (3): The results accuracy of K-NN for normal and abnormal activities for each class and the total classes after using LDA

Class	DWT	SURF	Radon $\Theta=45^\circ$	Radon $\Theta=90^\circ$	Radon $\Theta=135^\circ$	Radon $\Theta=180^\circ$
Faint	86.11	58.33	63.88	88.88	75	72.22
Vomit	61.11	58.33	61.11	36.11	75	72.22
Headache	69.44	72.22	44.44	69.44	50	38.88
Angina	66.66	75	47.22	75	38.88	13.88
Intestinal Colic	72.22	72.22	77.77	75	69.44	80.55
Renal Colic	63.88	58.33	38.88	86.11	55.55	25
Walk	94.44	100	100	100	100	100
Drink	69.44	100	100	100	100	100
Eat	50	100	100	100	100	100
Total	70.36	77.15	70.36	81.17	73.76	67

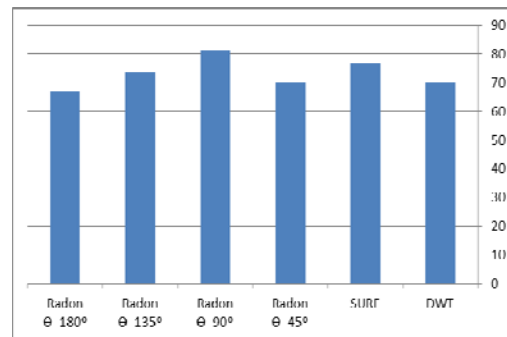


Figure (23) The results accuracy of K-NN for normal and abnormal activities for the total classes after using LDA

The results of accuracy after applying LDA show that the success rates of SURF is increased but of DWT and Radon are decreased. Figure (24) clearly shows the difference in accuracy of recognition before and after the implementation of LDA with K-NN models . Here too the STIP's results are ignored because they are not encouraging.

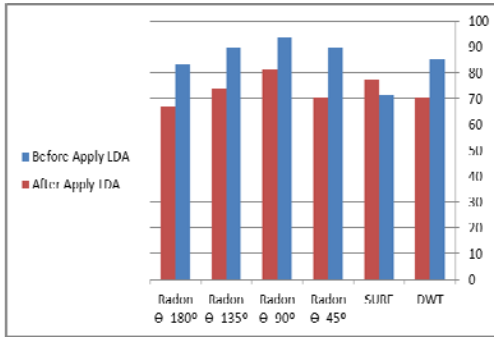


Figure (24) The comparative for accuracy of K-NN for normal and abnormal activities for the total classes before and after using LDA

6.2 The results of decision tree (DT) model

DT is a well-known method that is used in various kinds of domains, such as machine learning, pattern recognition, artificial intelligence, and data mining. Here we will analyze the most important results obtained with this model and with all types of features extracted. Table (4) show the overall accuracy results of DT and without implementing the LDA.

Table (4): The results accuracy Of Decision Tree (DT) for normal and abnormal activities for each class and the total classes.

Class	DWT	STIP m=20	STIP m=40	STIP m=60	SURF	Radon $\theta=45^\circ$	Radon $\theta=90^\circ$	Radon $\theta=135^\circ$	Radon $\theta=180^\circ$
Angina	55.55	55.55	58.33	52.77	38.88	58.33	66.66	38.88	22.22
Headache	77.77	44.44	58.33	58.33	55.55	47.22	55.55	61.11	38.88
Vomiting	77.77	52.77	36.11	38.88	52.77	69.44	88.88	41.66	83.33
Fainting	100	44.44	42.22	50	33.33	75	94.44	61.11	69.44

Intestinal Colic	58.33	58.33	38.88	38.88	38.88	33.33	66.66	58.33	69.44	77.77
Renal Colic	58.33	5.55	41.66	33.33	27.77	44.44	58.33	44.44	33.33	33.33
Walking	88.88	94.44	94.44	94.44	88.88	100	97.22	91.66	100	100
Drinking	86.11	55.55	75	88.88	97.22	100	100	97.22	88.88	88.88
Eating	75	94.44	69.44	66.66	91.66	91.66	100	100	97.22	97.22
Total	75.30	54.00	57.71	58.01	57.71	72.52	80	67.28	68	68

Where m in the above tables represents the maximum number points in the STIP method. Figures (25) is presented the chart of the results of accuracy of DT model for normal and abnormal activities for total classes.

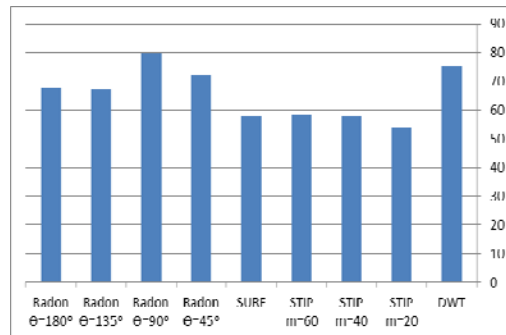


Figure (25) The results accuracy of Decision Tree (DT) for normal and abnormal activities for the total classes

From the above results are noticed. The best results are with DWT and Radon(theta=90). The results of STIP have also improved slightly compared to that with KNN but the general accuracy with DT model is less than K-NN. We

observe that results for recognition base on (radon transform + K-NN) and radon transform + DT) when angle equal 90° are the best from the rest (45 °, 135 ° and 180 °) .

6.3 The Results Of Combination Features

In this research, the combination of certain features is experimented in order to obtain better results of recognition rate especially with features that gave low accuracy results. So, these features have been identified and combined together (mean the features of the accuracy that is not encouraging) as: STIP , Radon and SURF features. The STIP is the worst, so it is combined with Radon and SURF..The combination of features was as follows :

- 1- (STIP + Radon when $\Theta = 135^\circ$ and 180°) .
- 2- (STIP + SURF).

It is worth mentioning here that the results of recognition accuracy (after the merger) with K-NN model are not mentioned here because they gave very bad results so they are not displayed. Table (5) and (6) are shown the results of combining STIP with radon features when $\Theta=135^\circ$ and $\Theta=180^\circ$ consecutively.

Table (5): The results accuracy of Decision Tree (DT) when combine STIP and Radon ($\Theta=135^\circ$)

Class	STIP m=20 + Radon $\Theta=135^\circ$	STIP m=40 + Radon $\Theta=135^\circ$	STIP m=60 + Radon $\Theta=135^\circ$
Fainting	55.55	63.88	55.55
Vomiting	52.77	72.22	61.11
Headache	55.55	55.55	61.11
Angina	38.88	50	55.55
Intestinal Colic	75	69.44	83.33
Renal Colic	41.66	38.88	38.88
Walking	100	100	100
Drinking	97.22	88.88	88.88
Eating	100	88.88	91.66
Total	68.51	70	70.67

Table (6): The results accuracy of Decision Tree (DT) when combine STIP and Radon ($\Theta=180^\circ$)

Class	STIP m=20 + Radon $\Theta=180^\circ$	STIP m=40 + Radon $\Theta=180^\circ$	STIP m=60 + Radon $\Theta=180^\circ$
Fainting	58.33	58.33	47.22
Vomiting	80.55	80.55	36.11
Headache	69.44	63.88	52.77
Angina	33.33	50	69.44

Intestinal Colic	77.77	77.77	69.44
Renal Colic	25	27.77	27.77
Walking	100	100	100
Drinking	88.88	100	100
Eating	97.22	100	100
Total	70.05	73.14	67

The figures (26) , (27) ,(28) and (29) are presented charts for the results of accuracy of decision tree (DT) of combination STIP and Radon ($\Theta=135^\circ$) and STIP + Radon ($\Theta=180^\circ$) for each class and total classes .

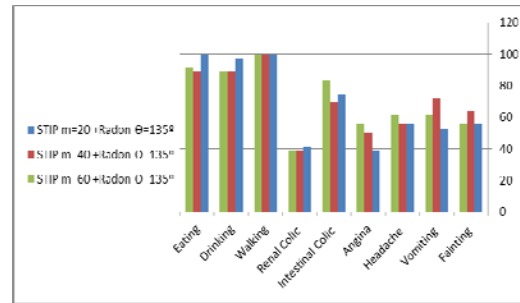


Figure (26) The results accuracy of decision tree (DT) when combine STIP and Radon ($\Theta=135^\circ$) for each class .



Figure (27) The results accuracy of Decision Tree (DT) when combine STIP and Radon ($\Theta=135^\circ$) for total classes

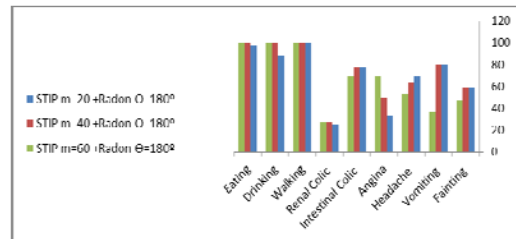


Figure (29) The results accuracy of Decision Tree (DT) when combine STIP and Radon ($\Theta=180^\circ$) for each class

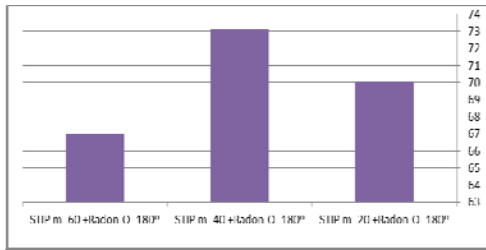


Figure (29) The results accuracy of Decision Tree (DT) when combine STIP and Radon ($\Theta=180^\circ$) for the total Classes

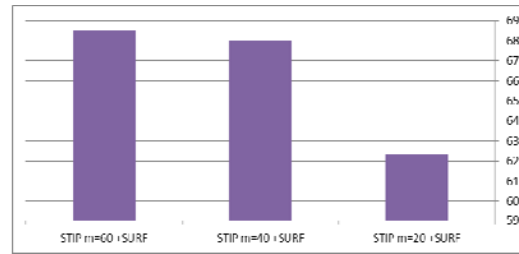


Figure (31) The results accuracy of Decision Tree (DT) When combine STIP and SURF for the total classes

Through the above results it is noticed that the rate of recognition has been increased compared to the previous of features, also it is noticed that all the results from combining STIP with Radon (180 or 135) are closed.

The following merged features are between STIP and SURF. Table (7) shows the results and Figures (30) and (31) present the results of accuracy of decision tree (DT) when STIP and SURF are combined for each class and the total classes, consecutively .

Table (7) The results accuracy of Decision Tree (DT) when combine STIP and SURF

Class	STIP m=20 + SURF	STIP m=40 + SURF	STIP m=60 + SURF
Fainting	22.22	58.33	55.55
Vomiting	61.11	69.44	69.44
Headache	55.55	58.33	61.11
Angina	36.11	47.22	55.55
Intestinal Colic	44.44	41.66	58.33
Renal Colic	41.66	36.11	16.66
Walking	100	100	100
Drinking	100	100	100
Eating	100	100	100
Total	62.34	68.00	68.51

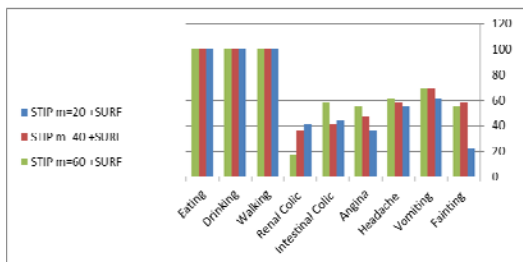


Figure (30) The results accuracy of Decision Tree (DT) when combine STIP and SURF for each class

Figure (32) presented chart for the comparative among combination features with DT model. presented chart for the comparative among

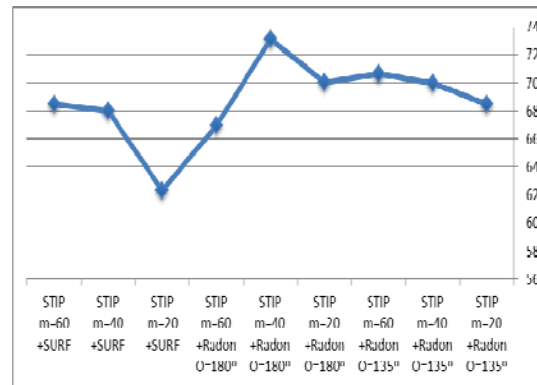


Figure (32) Presented the comparative for the combination features

6. CONTRIBUTIONS

This research provides a set of contributions, namely, first the use of Iterative Thresholding method for subtraction background from foreground . The second is the use of Radon transform , discrete wavelet transform , Speeded Up Robust Features and Space-Time Interest points techniques for extracting features . The third is to implement dimension reduction on extracted features using Linear Discriminant Analysis and Principle Component Analysis . The fourth is the use K-Nearest Neighbor and Decision Tree as classifiers with the features extraction techniques in order to classify and recognize normal and abnormal activities in our dataset which will mention in section 4 .Finally is the use SMS email as alarm step when the system recognizes any abnormal activities.

7. SUGGESTIONS FOR FUTURE WORKS

There are three suggestions for future works. First: construct large database that has many

activities in our daily life in several places with different backgrounds (indoor and outdoor). Second: construct HAR system for monitoring suspicious cases in order to enhance security and safety, for example, Stealing, Some crimes, and cases of loitering in the roads. Finally, recognize epilepsy statuses in real time with all kinds such as petit mal and grand mal ... etc.

8. CONCLUSION

This research work is proposed in order to recognize the normal and abnormal human activities for serving the healthcare fields. Our system is implemented by registered database that has three classes of normal activities (and six classes of abnormal activities. A new technology has been presented by combining STIP features extraction technique, once with Radon transform and other with SURF. This combination is used only with DT model. In addition, the Radon transform, SURF and STIP were used separately with K-NN and DT. The summary of our HAR system is just as follows. First: implementing background subtraction from each frame in order to obtain images without background and is almost limited to the person's activity. This step is important to increase accuracy and speed processing. Second: applying the feature extraction by using (Radon transform, DWT, SURF, STIP and combination techniques). Third: implementing K-NN and DT models to classify the classes and recognize the activities of abnormal then active the alarm system. Generally, the implementation of this system has given very satisfactory experimental results in terms of correct recognition of human activities rate whether with K-NN and DT models. Nevertheless as comparative between two models K-NN achieved better results than DT.

REFERENCES:

- [1] K. G. M. Chaturamali, S. Ramasinghe, and R. Rodrigo, 2014, "Abnormal Activity Recognition Using Spatio-Temporal Features," in Proceedings of 7th International Conference in Information and Automation for Sustainability (ICIAfS), IEEE, pp. 1-5.
- [2] B. Gottfried and H. K. Aghajan, 2009, "Behaviour Monitoring and Interpretation-BMI: Smart Environments," Vol. 3, Ios Press.
- [3] A. Taha, H. H. Zayed, M. Khalifa, and E.-S. M. El-Horbaty, 2015, "Human Activity Recognition for Surveillance Applications," in Proceedings of the 7th International Conference on Information Technology, pp. 577-586.
- [4] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, 2008, "Machine Recognition of Human Activities: A Survey," IEEE Transactions on Circuits and Systems for Video technology, Vol. 18, issue 11, pp. 1473-1488.
- [5] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, 2013, "A Review on Video-based Human Activity Recognition," Computers, Multidisciplinary Digital Publishing Institute, Vol. 2, issue 2, pp. 88-131.
- [6] J. K. Aggarwal and M. S. Ryoo, 2011, "Human Activity Analysis: A Review," ACM Computing Surveys (CSUR), ACM, Vol. 43, issue 3, pp. 1-47.
- [7] Z. A. Khan and W. Sohn, 2011, "Abnormal Human Activity Recognition System Based on R-Transform and Kernel Discriminant Technique for Elderly Home Care," IEEE Transactions on Consumer Electronics, Vol. 57, issue 4, pp. 1843-1850.
- [8] D. Anderson, R. H. Luke, J. M. Keller, M. Skubic, M. J. Rantz, and M. A. Aud, 2009, "Modeling Human Activity from Voxel Person Using Fuzzy Logic," IEEE Transactions on Fuzzy Systems, Vol. 17, issue 1, pp. 39-49.
- [9] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, 2011, "Robust Video Surveillance for Fall Detection Based on Human Shape Deformation," IEEE Transactions on Circuits and Systems for Video technology, Vol. 21, issue 5, pp. 611-622.
- [10] Z. A. Khan and W. Sohn, 2013, "A hierarchical Abnormal Human Activity Recognition System Based on R-Transform and Kernel Discriminant Analysis for Elderly Health Care," Computing, Springer, Vol. 95, issue 2, pp. 109-127.
- [11] J. L. M. Iqbal, J. Lavanya, and S. Arun, 2015, "Abnormal Human Activity Recognition Using Scale Invariant Feature Transform," International Journal of Current Engineering and Technology, Vol. 5, issue 6, pp. 3748-3751.

- [12] A. Bulling, U. Blanke, and B. Schiele, 2014, "A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors," *ACM Computing Surveys (CSUR)*, ACM, Vol. 46, issue 3, pp. 0-33.
- [13] M. Alzahrani, and S. Kammoun, 2016, "Human Activity Recognition: Challenges and Process Stages," *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 4, issue 5, pp. 1111-1118.
- [14] W.-X. Kang, Q.-Q. Yang, and R.-P. Liang., 2009, "The Comparative Research on Image Segmentation Algorithms," in *Proceedings of the First International Workshop in Education Technology and Computer Science*, IEEE, , pp. 703-707.
- [15] R. F. Yaser, 2016, "Multiobject Extraction from Color Image," M.Sc. Thesis, Dept. of Computer Science, University of Al-Mustansiriya.
- [16] D. M., 2012, "Wavelet and Curvelet Based Thresholding Techniques for Image Denoising," *International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE)*, Vol. 1, issue 10, pp: 77-81.
- [17] G. , R.C. and R.E. Woods, 2002, "Thresholding. Digital Image Processing," Pearson Prentice Hall.
- [18] A. B. ,Verina, and Y. B. N., 2016, "Feature Extraction Techniques for Video Processing in MATLAB," *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 4, issue 4, pp. 5292-5296.
- [19] E. B. T. Al-Abadi, 2014, "Gait Recognition System Using Two techniques: Support Vector Machine and Neural Network," M.Sc. Thesis, Department of Computer Science, University of Basra.
- [20] S.R. Deans, 2007, "The Radon Transform and Some of its Applications," Courier Corporation.
- [21] Y. Wang, K. Huang, and T. Tan, 2007, "Human Activity Recognition Based on R Transform," in *Proceedings of the IEEE Conference in the Computer Vision and Pattern Recognition, (CVPR'07)*, IEEE, pp. 1-8.
- [22] S. Kakarwal, and R. Deshmukh, 2010, "Wavelet Transform Based Feature Extraction for Face Recognition," *International Journal of Computer Science and Application Issue*, pp. 100-104.
- [23] P. Klapetek, D. Necas, and C. Anderson, 2004, "Gwyddion User Guide," Czech Metrology Institute.
- [24] Z.G.Sh. and V.M. Thakare, 2016, "Wavelet Based Feature Extraction Technique for Face Recognition and Retrieval: A Review," in *Proceedings of the National Conference on Recent Trends in Computer Science and Information Technology*, pp. 49-54.
- [25] R., Polikar, 1996, "Fundamental Concepts & an Overview of the Wavelet Theory," *The Wavelet Tutorial Part I*, Rowan University, College of Engineering Web Servers.
- [26] Ch.-L., L., 2010, "A Tutorial of the Wavelet Transform," Department of Electrical Engineering, National Taiwan University (NTUEE).
- [27] Sh. , F.Y., 2010, "Image Processing and Pattern Recognition: Fundamentals and Techniques," John Wiley & Sons.
- [28] J.-S. Sh. , J.-H. Yoo , and D.-S. Kang, 2012, "A Study on Moving Object Tracking Algorithm Using SURF Algorithm," and Depth Information. *Advances in Systems Theory, Signal Processing and Computational Science*, pp. 90-93.
- [29] P, J.T., 2011, "Study Group SURF: Feature Detection & Description" Department of Computer Science, Aarhus University.
- [30] E. Oyallon, and J. Rabin, 2015, "An Analysis of the SURF Method," *Image Processing On Line*, Vol. 5, pp. 176-218.
- [31] M. J. McDonnell, 1981, "Box-Filtering Techniques," *Computer Graphics and Image Processing*, Elsevier, Vol. 17, issue 1, pp. 65-70.
- [32] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, 2008, "Speeded-up Robust Features (SURF)," *Computer vision and image understanding*, Elsevier, Vol. 110, issue 3, pp. 346-359.
- [33] I. Laptev, 2005, "On Space-Time Interest Points," *International journal of computer vision*, Springer, Vol. 64, issue 2-3, pp. 107-123.
- [34] S. Zhang, Z. Wei, J. Nie, L. Huang, S. Wang, and Z. Li, 2017, "A Review on Human Activity Recognition Using Vision-Based Method," *Journal of Healthcare Engineering*, Hindawi, Vol. 2017, pp: 1-31.
- [35] A. El Oualkadi, F. Choubani, and A. El Moussati, 2016, "Proceedings of the Mediterranean Conference on Information &

- Communication Technologies 2015", MedCT 2015. Vol. 2, Springer.
- [36] Z. Voulgaris and G.D. Magoulas. 2008, "Dimensionality Reduction for Feature and Pattern Selection in Classification Problems," in Proceedings of the Third International Multi-Conference in Computing in the Global Information Technology, IEEE, pp. 160-165.
- [37] A. Ghodsi, 2006, "Dimensionality Reduction a Short Tutorial," Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada, Vol. 37, pp. 1-21.
- [38] S. A. Najim, 2014, "Information Visualization by Dimensionality Reduction: A Review," Journal of Advanced Computer Science & Technology, Vol. 3, issue 2, pp. 101-112.
- [39] B. Shaw, and T. Jebara., 2009, "Structure Preserving Embedding," in Proceedings of the 26th Annual International Conference on Machine Learning, ACM, pp. 937-944.
- [40] P.M.C. Guerreiro, 2008, "Linear Discriminant Analysis Algorithms," M.Sc. Thesis, Technical University of Lisbon, Portugal.
- [41] S. Balakrishnama, and A. Ganapathiraju, 1998, "Linear Discriminant Analysis - a Brief Tutorial," Institute for Signal and information Processing, Vol. 18, pp. 1-8.
- [42] G. Zhong, L.-N. Wang, X. Ling, and J. Dong, 2016, "An overview on Data Representation Learning: from Traditional Feature Learning to Recent Deep Learning," The Journal of Finance and Data Science, Elsevier, Vol. 2, issue 4, pp. 265-278.
- [43] N.K. Kamila, 2015, "Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing," IGI Global.
- [44] R. O. Duda, P. E. Hart, D. G. Stoke, 2012, "Pattern Classification", 2nd ed., John Wiley & Sons.
- [45] E. J. HARFASH, 2016, "Face Recognition System Using PCA, LDA, Kernel PCA and Kernel LDA," International Journal of Computer Science Engineering and Information Technology Research (IJCEITR), Vol. 6, issue 5, pp. 9-20.
- [46] J. Han, M. Kamber and J. Pei, 2011, "Data Mining: Concepts and Techniques," Elsevier.
- [47] X. Yang and Y. Tian, 2017, "Super Normal Vector for Human Activity Recognition with Depth Cameras," IEEE Transactions on pattern analysis and machine intelligence, Vol. 39, issue. 5, pp. 1028-1039.
- [48] S. Kaghyan , and H. Sarukhanyan, 2012, "Activity Recognition Using K-nearest Neighbor Algorithm on Smartphone with TRI-AXIAL Accelerometer," International Journal of Informatics Models and Analysis (IJIMA), Vol. 1, issue 2, pp. 146-156.
- [49] B. Liu, 2007, "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data," Springer Science & Business Media.
- [50] L. Hawarah, A. Simonet, M. Simoneth, 2006, "Dealing with Missing Values in a Probabilistic Decision Tree during Classification," in proceedings 6th IEEE International Conference on Data Mining.
- [51] P.-N. Tan, M. Steinbach, and V. Kumar, 2006, "Classification: Basic Concepts, Decision Trees, and Model Evaluation," Introduction to Data Mining, Addison-Wesley Boston, Mass, Vol. 1, pp. 145-205.
- [52] L. Breiman, 2017, "Classification and Regression Trees," CRC press.
- [53] H.W. Trimmer, 1999, "Understanding and Servicing Alarm Systems," Butterworth-Heinemann.
- [54] S. Ranasinghe, F. Al Machot, and H.C. Mayr, 2016, "A Review on Applications of Activity Recognition Systems with Regard to Performance and Evaluation," International Journal of Distributed Sensor Networks, SAGE Publications Sage UK, Vol. 12, issue 8, pp. 1-22.
- [55] G.S. Chhabra, and D.S. Bajwa, 2015, "Review of E-mail System, Security Protocols and Email Forensics," International Journal of Computer Science & Communication Networks, Vol. 5, issue 3, pp. 201-211.