IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

A Robust Hybrid Human Activity Recognition System using Lbptop and Body Joint Features with Majority Voting for High Accuracy

Halakundi Chidananda[1]*, T Hanumantha Reddy[2]

[1]Computer Science & Engineering Department, R Y M Engineering College, Ballari, India
[2] Computer Science & Engineering Department, R Y M Engineering College, Ballari, India

*Corresponding Author: chidably999@gmail.com

Available online at: www.ijcseonline.org

Accepted: 23/Jul/2018, Published: 31/July/2018

Abstract**—** Human activity recognition has seen an enormous success in the last decade performing a dominant part in the field of ubiquitous computing. This rising demand can be attributed to several real life applications fundamentally dealing with human-centric applications like healthcare and eldercare systems. Many research experiments with data mining and machine learning procedures have been experiencing precisely to recognize human activities for healthcare systems. This work proposes a hybrid method to recognize the patient actions under care using a simple camera instead of multiple expensive sensors using machine learning with LBPTOP algorithm and body joint features with majority voting framework for real time monitoring applications with greater efficiency of recognition. This work uses different classifiers to achieve the experimental results approximately above 90% which clearly shows a remarkable recognition achievement compared to the other activity recognition techniques.

Keywords**—** Machine Learning, Human activity, Body joint features, Real time, LBP-TOP, Classifiers

1. Introduction

Monitoring human activities of daily living is an Essential way of describing the functional and health status of a human [1]. Hence, one of essential component is human activity recognition(HAR) in healthcare systems and personal life care; especially for the disabled and aged people [2] .In smart environments, video cameras can be deployed, smart environments like smart hospitals or smart homes to monitor daily activities of elderly people to acquire video clips of time-series activity; many studies support that healthcare services and personalized life care can reduce the rate of mortality [3] particularly for the disabled and elderly people in smart environments [4,5]. The kinematics study of human motion tells that a human action can be divided into a posture sequences. Repetition of the sequence by the similar subject at various subjects or various times with some changes. Methods introduced so far for human action recognition are based on the various silhouette positions and the model of sequence dynamics. In common, they are of two types based on the modelling of action dynamics: *explicit* and *implicit* models. In case of an implicit model, the silhouettes action sequences extracts the *action descriptors* such that the recognition action is changed into a static classification from a temporal classification problems.

Both temporal and spatial characteristics of the actions are supposed to be captured by the action descriptors. For instance, Davis and Bobick [9] introduced a strategy of stacking the postures into a moment-history images (MHI) and movement-energy images (MEI). Seven Hu movements [10] are excerpted by using MHI and MEI to perform as action descriptors. Action recognition is on the basis of the distance of Mahalanobis between every known actions moment descriptor and the input. Meng [11] expanded the MHI and MEI to a hierarchical type and used a support vector machine (SVM) to recognize the performance. A new strategy introduced from Chen *et al.* [8], to acquire the five extremities of the structure that related to the legs, head, and arms star figure units [12] are fitted to silhouettes. To acquire the spatial spreading of the five extremities over the time of an activities GMMs are used by ignoring the silhouettes temporal order in the performance sequence.
Recently, Shah and Yilmaz [6] treated a silhouettes sequence as a spatio–temporal volume and introduced to generate the properties of differential geometric surface, i.e., Mean curvature and curvature of Gaussian, to create a descriptor for every performance, called as performance sketch. Gorelick *et al.* [7], [13] generated features of space-time

including shape structure, saliency in space–time , orientation and dynamics of performance by utilizing the solution properties to the equation of Poisson and employed nearest neighbourhood (KNN) to divide the performances. The implicit modelling method has the uses that can manage few number of samples in training and the recognition is relatively simple. However, before the performance can be recognized it needs good temporary segmentation and it usually provides less strength encoding of the dynamics in performance.

On the other side, explicit model follows the method that the activity consists of sequence of postures and usually consists of two components: postures description and modelling of the postures dynamics. Several characters have been implemented to explain the postures. This work identifies a human actions on the basis of serious of silhouette postures. Especially, the work concentrates on the Identification of human actions, the least understandable semantically significant movement units, for example, walk, jump, and run. A performance identification system is desired to be subject independent who play out the activities, free from the speed at which the activities are executed, powerful against silhouettes generation of noise, measurable to maximum number of performances, and expandable with new performances. In spite of the impressive research in the previous couple of years, such a framework is yet to be established. Because of their practical applications several developers have been examining video-dependent HAR frameworks, and also because of the correct human activities identification is still considered as a challenging task where the correctness deficiency usually appears because of several problems such as the low variance among the characteristics of the action classes of different type, and high difference among the characteristics of the activity class of same type, extracting efficient features failure. Identifying activities of human from sequences of video or still images is a challenging task because of some problems, such as partial occlusion, background clutter, view point, scale changes, appearance, and lighting. This work introduces a strategy to analyze and understand the activity of patient under care using a cost effective camera rather than many costly sensors using body joint features with majority voting system and machine learning with LBPTOP algorithm for real time monitoring applications with maximum accuracy of recognition.

This work is organised as follows: Section I contains the introduction of human activity recognition in healthcare systems and personal life care, section II discusses about Local Binary Pattern from Three Orthogonal Planes (LBP-TOP), Section III explains about the extraction of spatial and temporal features, section IV contains the proposed work which includes two stages. In the first stage, Methodology-I uses body joint features, HMM and GDA for recognition of human activity and the methodology-II uses features of LBPTOP with different classifiers. The second stage is the maximum voting system, which recognises the class of the activity based on the maximum votes of the classes obtained at the first stage, section V discusses about the experimental results and section VI concludes the research work with future updates.

II. LOCAL BINARY PATTERN – TOP

Local binary pattern is an operator and is very simple efficient texture used for labelling of individual pixel of an image and by thresholding the neighbourhood of individual pixel and yields the response which are taken in the form of binary numbers. LBP texture operator has been found in several applications and it is considered to be an efficient approach because of discriminative strength and computational simplicity. In applications of real world, robustness is the important property of local binary pattern to changes in gray-scale occurred by lighting changes. Images are analysed in real time because of its computational effortlessness is shown in equation (1).

$$LBP_{P,R} = \sum_{p=0}^{p-1} S(g_p - g_c)\,2^n, \overset{1}{S}(X) = \begin{cases} 1, X \geq 0 \\ 0, X \leq 0 \end{cases} \qquad (1)$$

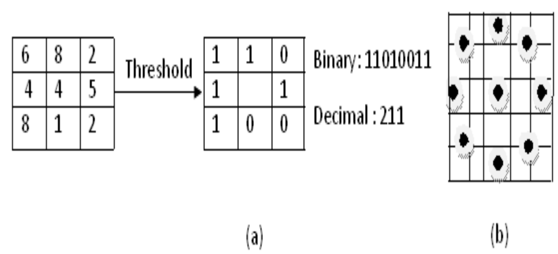

Figure 2.1. (a) Operator of basic LBP (b) Circular (8, 2) surrounding

Where, gp denotes gray values of P equally dispersed pixels on a circle of radius R, g_c denotes centre pixel grey value of local neighbourhood. By considering the variations between the centre pixel values and neighbourhood values, with respect to grey scale local binary pattern achieves invariance. The operator of basic LBP surroundings is shown in figure 2.1(a).

Occurrence of different binary pattern is stored by using a histogram. Definition of neighbor can be enhanced to include circular neighbourhoods with any number of pixels, as shown in Figure 2.1(b). So maximum scale primitives of texture or micro patterns like corners, lines and dots can be stored. Uniform patterns are used to decrease the feature vector length of the local binary pattern. Pattern are considered to be uniform if it has at least two bitwise transitions from 1 to 0 or vice-versa. For example 10101000 is non-uniform pattern because it has 4-bit transition.

Patterns in non-uniform are made to store in a single bin during histogram computation, while considering only uniform local binary pattern. In the analysis of facial image local texture descriptors have obtained tremendous attention due to their robustness to challenge such as changes in illumination and pose. In our approach local binary patterns used by a temporal texture recognition was generated from the Three Orthogonal Planes (LBP-TOP). Compared to ordinary LBP, LBP-TOP method is more efficient.

We are extracting features or information in two dimensions in ordinary LBP, and three dimensions in LBP-TOP i.e., T, X and Y. For LBP-TOP, the radii in temporal and spatial axes T, X, and Y, and the number of neighboring points in the TX, XY, and YT planes can also be distinct and can be marked as R_T, R_X and R_Y, P_{TX}, P_{XY}, P_{YT}. The LBP-TOP data is then served as LBP-TOP PTX, PXY, PYT, RT, RX and RY. In the event that the middle pixel's coordinates gtc, e are (tc, xc, yc), then the local neighbourhood coordinates in TX plane $g_{TX,p}$ are given by (t_c-$R_T$sin(2πp/P_{TX}), x_c+R_X cos(2πp/P_{XY}), y_c), the local neighbourhood coordinates in TY plane $g_{TY,p}$ are given by (t_c-$R_t$sin(2πp/P_{TY}),x_c, y_c –R_T cos(2πp/P_{TY})) and the local neighbourhood coordinates in XY plane g_{XYp} (t_c, x_c-$R_X$cos(2πp/P_{XY}),, y_c –R_Y sin(2πp/P_{XY})). Sometimes, the 3 axes radii are same thus do the number of neighboring focuses in TX, TY, and XY planes. In that case, we use LBP-$TOP_{P,R}$ for abbreviation where $R=R_T=R_X=R_Y$ and P=PTX=PTY=PXY. The movements of body is represented by using LBP-TOP, due to its capacity to explain the signals of spatiotemporal, robustness to monotonic gray-scale variations caused by illumination change. Taking in account the movement of body, the descriptors are gathered by concatenating patterns in local binary on 3 planes in orthogonal from the activity sequence: TY, TY, and XY, dealing only the co-occurrence statistics in these 3 directions. The histograms of LBP-TOP in every block volume are concatenated and computed into a one histogram. All characteristics from each block volume are linked to denote the motion and appearance of the body sequence. In this type, effective description of the body movement can be obtained on three separate levels of locality. Information contained in the histogram labels (bins) are from three planes in orthogonal, explaining temporal information and appearance at the pixel level.

To give information on a regional level the labels are added over a least block shows the features of the motion and appearance in time segments and specific locations, and global description of the body motion is built by concatenating all the information from the regional levels. Moreover, even though different activities are segmented into the equal number of block volumes and they have different length, so their characteristics vectors lengths are the equal to compare.

The human movements histogram can be defined as follows equation 2:

$$H_{r,c,d,j,i} = \sum_{x,y,t} I\{f_j(x,y,t) = i\} \quad i = 0,\dots,n_j - 1; j = 0,1,2. \tag{2}$$

where, n_j denotes different names number produced by the LBP operator in the j^{th} plane (j=0: TX, 1:TY, and 2:XY), f_j (t,x,y) represents the central pixel(t,x,y) LBP code in the j^{th} plane, t Є {RT…,T-1-RT}, x Є {RX…,X-1-RX}, y Є {RY…,Y-1-RY}. The figure 2.2 denotes the block volumes concatenated histograms.

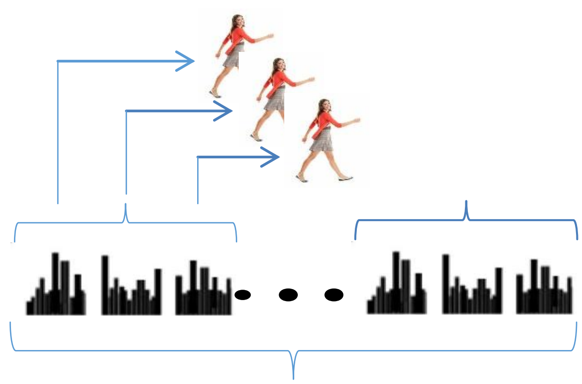

Figure 2.2.Concatenated histograms of body actions of each block

III. FEATURES EXTRACTION OF BODY JOINT

A. Generation of features

Using a trained random forest several body segments of human actions are first segmented. Spatial characteristics consisting of the average of the depth values, the 3-D pair angles in body joint, the depth values variance, and the every segmented body part area are mixed with the characteristics of motion denoting the direction and magnitude of every joint in the next frame to create the spatiotemporal characteristics in a frame.

Once the segmented depth body shape is available, the 3-D body skeleton unit having 16 joints which are represented by the 3-D centroids of every specific body parts. The joints of the body are represented and considered for the right foot, left foot, right knee, left knee, left palm, left elbow, right palm, right elbow, right hip, left hip, central hip, chest, right shoulder, left shoulder, neck and head. Corresponding divided body parts with joints and different colours, from both right hand waving, hand waving, right leg moving performance and sitting.

B. Extraction of spatial characteristics

The first characteristics of spatial are denoted by the 15 joint pairs. The joint pairs are denoted as neck-right shoulder, head-neck, left shoulder left elbow, right shoulder-right elbow, neck-left shoulder, left elbow-left palm, right elbow-right palm, chest-hip, , neck-chest, hip-left thigh, hip-right thigh, right thigh-right knee, left thigh-left knee, left knee-left foot, and right knee-right foot. Each body limb is signified in the round arranged framework that is having azimuthal angle and polar angle of each pair and make the skeleton invariant to length of the limb. Consequently, the azimuthal and polar point of each pair having QK and QJ joints can be taken as follows in equation 3 and 4:

$$\delta = \arctan\left(\frac{Q_{K_Y} - Q_{J_y}}{Q_{K_X} - Q_{J_X}}\right) \quad (3)$$

$$\emptyset = \arcsin\left(\frac{Q_{K_Z} - Q_{J_Z}}{p}\right) \quad (4)$$

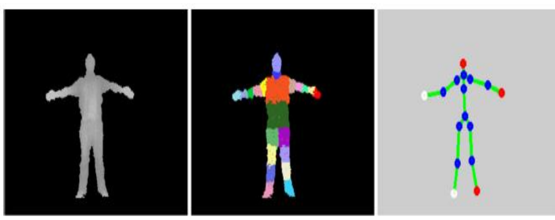

Figure 3.1 A sample action body structure, labelled body segments , and corresponding joints

Where ρ can be denoted as

$$\rho = \sqrt{(Q_{Kx} - Q_{Jx})^2 + (Q_{Ky} - Q_{Jy})^2 + (Q_{Kz} - Q_{Jz})^2}$$

In this regard, to produce the characteristics of invariant limb size, only azimuthal and polar angles are considered. The next characteristic is the mean depth parameter of each body segment as denoted by equation (5).

$$\overline{X_S} = \frac{1}{T_S} \sum_{i=1}^{T_S} X_{Si} \quad (5)$$

Where denotes the pixels number in the division, for the segment depth values. The equation 6 gives depth parameters variance of each body segment so that

$$C_S = \frac{1}{T_S} \sum_{i=1}^{T_S} \left(X_{Si} - \overline{X_S}\right)^2 \quad (6)$$

The equation 7 gives final feature of spatial includes the depth values area of every divided body part and can be denoted as

$$R_s = \sum_{i=1}^{T_S} (X_{Si}) \quad (7)$$

Hence, the each image size of the spatial feature becomes 1 × 96, considering the changes of the 22 body segments depth parameters (i.e., 1 × 22), the means for the 22 body segments profundity esteems (i.e., 1 × 22), the areas of the 22 body segments depth parameters (i.e., 1 × 22), and 15 three dimensional joint-pair angles (i.e., 1 × 30).

C. Extraction of temporal characteristics

Temporal characteristics denoting the parameters of motion (i.e., direction and the magnitude of the joints) in the following image frame are evaluated for 16 joints. A joint weight by two successive depth images is shown in equations 8, 9, 10, 11.

$$D = \sqrt{\left(Q_{Jx(i-1)} - Q_{Jx(i)}\right)^2 + \left(Q_{Jy(i-1)} - Q_{Jy(i)}\right)^2 + \left(Q_{Jz(i-1)} - Q_{Jz(i)}\right)^2} \quad (8)$$

$$G_{QJ(X,Y)} = arctan\left(\frac{Q_{Jy(i-1)} - Q_{Jy(i)}}{Q_{Jx(i-1)} - Q_{Jx(i)}}\right) \quad (9)$$

$$G_{QJ(Y,Z)} = arctan\left(\frac{Q_{Jz(i-1)} - Q_{Jz(i)}}{Q_{Jy(i-1)} - Q_{Jy(i)}}\right) \quad (10)$$

$$G_{QJ(X,Z)} = arctan\left(\frac{Q_{Jz(i-1)} - Q_{Jz(i)}}{Q_{Jx(i-1)} - Q_{Jx(i)}}\right) \quad (11)$$

The directional angles joints size becomes a dimension vector 148. Therefore, the size of the temporal characteristics for every image sequence becomes 1 × 64, which produces the characteristics of spatiotemporal for each depth shape 1 × 160 altogether. The Ii denotes the characteristics vector for the video frame. The characters of magnitude from the both hand waving activity and right hand performance, respectively, the significant values of the right hand joints magnitudes, while the important values of the two hand joints magnitudes, representing the significance of the characters of motion in these performances discrimination. It is also apparent that the parameters of spatiotemporal are invariant in position, the body shape added in the feature contains no positional information; hence, the position of the activity inside the frame does not matters, because the proposed characters can be successfully taken.

IV. PROPOSED SYSTEM

Below figure 4.1 shows the proposed system, which includes two stages. In the first stage, Methodology-I uses body joint features, HMM and GDA for recognition of human activities and the methodology-II uses features of LBPTOP with different classifiers. The second stage is the maximum voting system, which recognises the class of the activity based on the maximum votes of the classes obtained at the first stage.

A. Methodology I:

Below figure 4.2 shows the flow chart which explains the overall operation for detection of human activity for methodology- I.

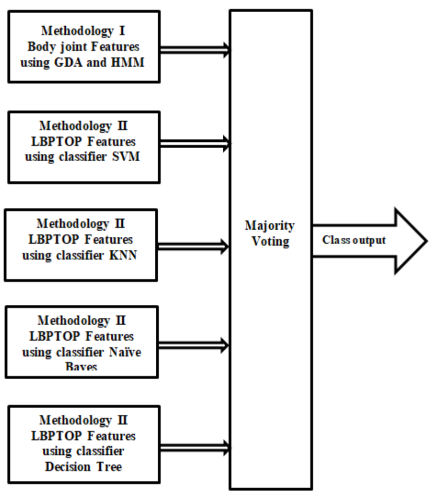

Figure 4.1: Diagram of Proposed Hybrid Block

B. Generalized discriminant analysis (GDA) on the characteristics of spatiotemporal

Generalized discriminant Generalized (GDA) is connected to isolate the features further and make them more vigorous and it is the last stage of the feature extraction procedure [23]. GDA is a generalized method of linear discriminant analysis (LDA), a conventional statistical method that has been successful on classification problems [19]. The strategy is essentially in light of the determination of eigenvalue that gives the greatest of the inertia result. The significant impediment of LDA is the disappointment for nonlinear issues. Input distance is mapped into a feature space of high dimensional with properties in linear by that GDA generalizes LDA and in a local way such as the LDA type it becomes possible to solve the problem in that space.

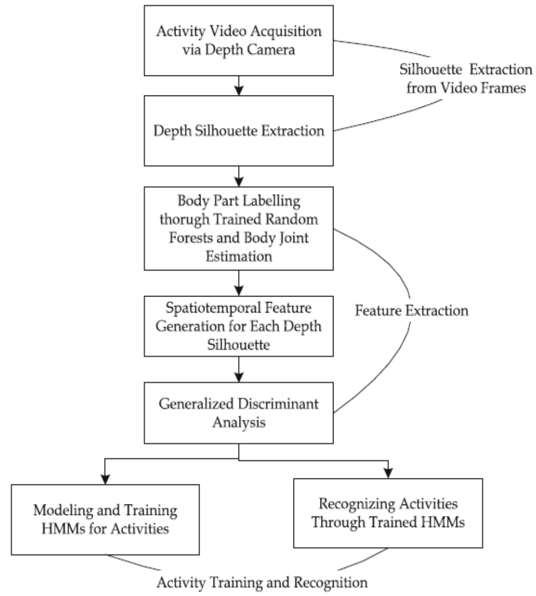

Figure 4.2: Flow Chart of Methodology-I

C. Activity modelling via HMM

There is a substantial volume of works on the Hidden Markov Models (HMMs) applications to a wide scope of pattern recognition tasks [20, 21 and 22]. An HMM is a Time successive data can be modelled by utilizing stochastic finite state machine. Basically, a typical HMM consists of four major parts: namely state transition matrix, state observation matrix, states, and initial state distribution. The probabilities among the states are characterized by the transition in state matrix and the observation probabilities by every state are available in the state observation matrix. HMM contains some condition or property that may have during a particular time are represented by states. Firstly the probability in state of an HMM is indicated by initial state distribution. When HMM design is characterized with the four segments, the following stage is training to the HMM. To train, initially features are classified into specific number of clusters, preparing a codebook which consists of the cluster centroids. At that point through vector quantization, pattern groupings are produced from the codebook. Later these symbol sequences are utilised to model sequential patterns in an HMM.

D. Methodology II:

Below figure 4.3 shows the flow chart which explains the overall operation for the detection of the human activity for methodology- II.

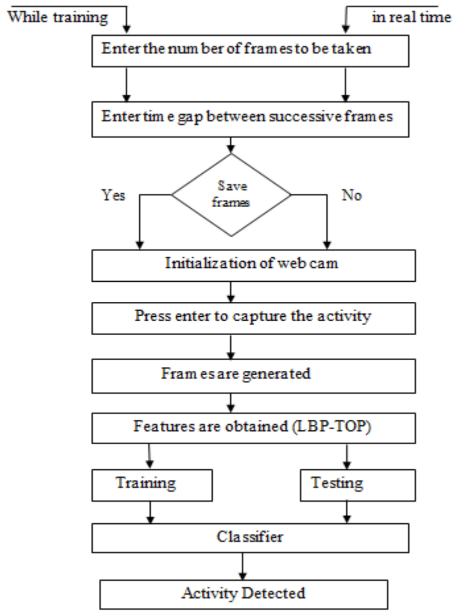

Figure 4.3: Flow Chart

E. Classifiers

Support Vector Machine (SVM): Hyper plane or a large spatial which exists in SVM, where hyper planes sets. For classification, hyper planes are used. Hyper planes which forms great separation among the two classes. In particular the separation to the convenient sample or each side from the hyper plane is increased. Different samples are divided by very large and clear gap. By using SVM mapping samples are converted into points in space [15].

K-Nearest Neighbor (KNN): The unknown samples classification is done by associating the sample to known from unknown on the basis of some criterion or some distance called classifier of KNN, which is a classifier based on an instance [16]. For k set of neighbors a common class of k-nearest neighbors is assigned to a sample having a small positive number are considered as the classifier training samples to know the correct classification.

Navies Bayes Classifier (NBC): NBC classifier is designed which assumes that existence of specific features of a class is not dependent or not related on the existence of any other features. Navies Bayed model which uses the high likelihood technique by that parameter estimation is done [17].

Decision Tree (DT): Training set separation is done by using DT classifier in a way in recursive, with every division containing superior samples by one class. Using DT classifier, a complicated process of action taken is split into measurable actions taken and a tree like structure output is generated in binary. The unknown classes are classified by using some rules which are evaluated from tree which are derived from the various classes [18].

V. RESULTS OF THE EXPERIMENT

This work focuses on the recognition of human activities like walking, Hand Waving, Idle, Sudden Fall. The results have been analysed thoroughly using the trained networks on several independent test patterns. Table 1 compares the accuracy obtained for recognition of human activities using the proposed method with various classifiers and the same is shown in fig 5.2.

The figure 5.3 shows the accuracy performance of the classifiers based on the number of training sets in the increasing order from train set1 to train set4 for the different human activities as discussed above.

VI. CONCLUSION

Recognition of human activities by using a simple inexpensive webcam instead of using multiple expensive sensors is the main objective of this work. Image acquisition initiates the overall operation by capturing video frames of the person body movements. The results obtained in the Methodology-I for human activity recognition are further enhanced by our proposed hybrid system with majority voting to conform the classification of the activity recognised. The overall results obtained shows an improvement of 10% compared to the existing methods up to date. Hence for patients and elders care takers this system shall be an important tool.

Upgrades of the future

Although the current results are prominent, this work still has room for enhancement. Future plan is to have research to improve the accuracy of recognition of human activity. First, automated techniques are essential to determine optimal features during training. Second, recognition rate could be enhanced using more sophisticated criteria. Third, faster algorithms and other classifiers would also be analysed and experimented.

Table 1. Accuracy of human activity recognition

Sl. No.	Classifier	Activity	Accuracy Recognition
1.	SVM	Walking	85%
		Hand waving	86%
		Idle	85%
		Sudden Fall	80%
2.	KNN	Walking	86%
		Hand waving	82%
		Idle	83%
		Sudden Fall	82%
3.	Naïve Bayes	Walking	82%
		Hand waving	84%
		Idle	80%
		Sudden Fall	86%
4.	Decision Tree	Walking	84%
		Hand waving	88%
		Idle	79%
		Sudden Fall	81%
5.	GDA,HMM	Walking	86%
		Hand waving	85%
		Idle	84%
		Sudden Fall	86%
6.	Hybrid (Proposed)	Walking	96%
		Hand waving	94%
		Idle	90%
		Sudden Fall	92%

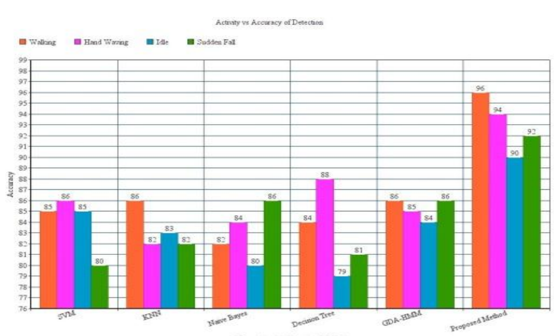

Figure 5.3 Accuracy of human activity recognition

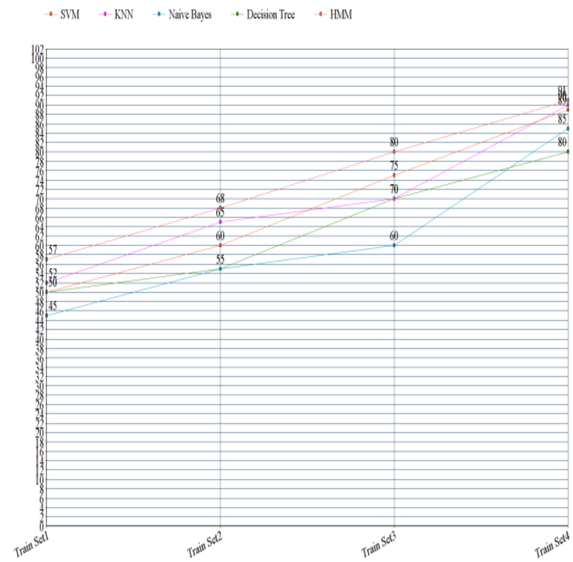

Figure 5.3 Accuracy vs. No. of Trainings

REFERENCES

[1]. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW: Studies of illness in the aged. The index of ADL: A Standardized Measure of Biological and Psychological Function: JAMA 1963; 185:914–919.

[2]. Chan M, Esteve D, Escriba C, Campo E: A review of smart homes-Present state and future challenges: Comput Methods Programs Biomed 2008;91:51–81.

[3]. Tarricone R, Tsouros AD: The solid facts: home care in Europe: World Health Organization (WHO), Geneva, 2008.

[4]. Lai JHK, Yik FWH: Perceived importance of the quality of the indoor environment in commercial buildings: Indoor Built Environ 2007;16(4):311–321.

[5]. Aylward C, Harris RS, Harrison PTC: Analysis of topics and trends in indoor environment research in Europe: Studies in Indoor Monitoring: Indoor Built Environ 2005;14(3-4):197–200.

[6]. A. Yilmaz and M. Shah, "Actions sketch: a novel action representation," in Proc. Comput. Vis. Pattern Recognit., 2005, vol. 1, pp. 984–989.

[7]. L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 12, pp. 2247–2253, Dec. 2007.

[8]. D.-Y. Chen, H.-Y. M. Liao, and S.-W. Shih, "Humnn action recognition using 2-D spatio-temporal templates," in Proc. Int. Conf. MultimediaExpo, 2007, pp. 667–670.

[9]. A. Bobick and J. Davis, "The recognition of human movement using temporal templates," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 3, pp. 257–267, Mar. 2001.

[10]. M. Hu, "Visual pattern recognition by moment invariants," IRE Trans. Inf. Theory, vol. 8, no. 2, pp. 179–187, Feb. 1962.

[11]. H. Meng, N. Pears, and C. Bailey, "A human action recognition system for embedded computer vision application," in Proc. Comput. Vis. Pattern Recognit., 2007, pp. 1–8.

[12]. H. Fujiyoshi and A. Lipton, "Real-time human motion analysis by image skeletonization," in Proc. 4th IEEE Workshop Appl. Comput. Vis., 1998, pp. 15–21.
[13]. M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," in Proc. Int. Conf. Comput. Vis., 2005, pp. 1395–1402.
[14]. L.Wang and D. Suter, "Learning and matching of dynamic shape manifolds for human action recognition," IEEE Trans. Image Process., vol. 16, no. 6, pp. 1646–1661, Jun. 2007.
[15]. Vapnik, V. Statistical learning theory (Adaptive and Learning Systems for Signal Processing, Communications and Control Series). John Wiley & Sons, New York (1998).
[16]. Larose, D. T. KNN. In: Discovering Knowledge in Data: An Introduction to Data Mining, 1st Ed., pp. 90-106. New Jersey; Wiley Interscience (2004).
[17]. Han, J., Kamber, M. Data Mining: Concepts and Techniques. Morgan Kaufmann (2006).
[18]. Larose, D. T. Decision Trees. In: Discovering Knowledge in Data: An Introduction to Data Mining, 1st Ed., pp. 108-126. New Jersey; Wiley Interscience (2004).
[19]. Lepetit V, Fua P (2006) Keypoint recognition using randomized trees. IEEE Trans Pattern Anal Mach Intell28:1465–1479
[20]. Baum E, Eagon J (1967) An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bull Am Math Soc 73:360–363
[21]. Lawrence R, Rabiner A (1989) Tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286
[22]. Nishimura Hand TsutsumiM (2001) Off-line hand-written character recognition using integrated1DHMMsbased on feature extraction filters. Sixth International Conference on Document Analysis and Recognition, p 417–421
[23]. Uddin, M.Z. (2017) Human activity recognition using segmented body part and body joint features with hidden Markov models. Multimed Tools Appl 76:13585–13614 DOI 10.1007/s11042-016-3742-2.

Authors Profile

Halakundi Chidananda, perceived B. E. (Electronics and Communication Engineering) from Siddaganga Institute of Technolgy,Tumakuru under Bengaluru University and Master of Technology (Computer Networking and Engineering) from Ballari Institute of . Technology and Management, Ballari under Visvesvaraya Technological University, Belagavi. He is currently a research scholar registered to V.T.U., Belagavi in the area of Image and Video processing working on the issues of Human Activity Recognition from Video Sequences. Currently, he is working as Assistant Professor in Department of Computer Science and Engineering, Rao Bahadur Y Mahabaleswarappa Engineering College, Ballari. His subjects of interests are Image and Video Processing, Machine Learning, Artificial Intelligence, Software Engineering, Programming Languages, Data Structures and Object Oriented Programming. He has published total of 09 publications in the area of Image and Video Processing.

Dr. T. Hanumantha Reddy, perceived B.E. (Electronics and Communication) from Gulbarga University, and Master of Engineering in Computer Engineering from Mysore University, Mysore. He has completed his Ph.D. from J. N. T. U., Hyderabad in the area of Image Processing. Currently he is working as Vice Principal and Professor and Head in the Department of Computer science and Engineering, R. Y. M. E. C., Bellary. His subjects of interest are Image and Video Processing, Machine Learning, Artificial Intelligence, Data Communications and Computer Networks, Computer Architectures and Multimedia computation and Communication. He has published total more than 40 publications in the area of Image Processing.

,