Fundamental Recognition of ADL Assessments Using Machine Learning Engineering

2 authors:

Mouazma Batool
Air University of Islamabad

Madiha Javeed
Air University of Islamabad

Fundamental Recognition of ADL Assessments Using Machine Learning Engineering

Mouazma Batool
Dept. of Computer Science
Air University
Islamabad, Pakistan
181788@students.au.edu.pk

Madiha Javeed
Dept. of Computer Science
Air University
Islamabad, Pakistan
191880@students.au.edu.pk

***Abstract*—This paper describes an RGBD (Red Green Blue Depth) image daily life activity identification system that can monitor and detect human activities without the need of optical markers or motion sensors. In this paper, we have developed a practical methodology for detecting human activity in the daily environment. For feature extraction technique, we suggested a noble method for recognition of ADL activities based on symmetry principle. By extracting silhouette from depth images and performing mapping operations over RGB images, we have been able to extract the skeleton information of the human body in RGBD images and identify ADL activities using four critical parameters: angle formation between hand and upper half of the body, angle between center body point and hand, angle formation between hand and lower half of the body, and angle between two hands of the single silhouette. We employed the linearly dependent concept (LDC) and long short-term memory-recurrent neural networks (LSTM-RNN) for feature selection and classification, compared the results to existing approaches. The objective of our research is to not only find an effective and useful collection of features from the silhouette, but also to outperform current approaches. Finally, the suggested method's testing results showed a 2.5 percent increase in accuracy with a 92.83 percent success rate, as well as a reduction in relative error to 2.47 percent of the original dataset.**

Keywords— activities of daily living, linearly dependent concept, recurrent neural network, symmetry principle.

I. Introduction

Activities of daily living (ADL) identification has gotten importance in a wide range of applications including surveillance, human-computer interaction, pedestrian detection systems, healthcare systems, homo-robotics, and smart homes [1-7]. Most ADL recognition systems utilize grey or RGB video sequences [8-11]. Here, the extraction of elements of human actions in clutter background, light variations, and camera motion is a big difficulty here [12]. Moreover, the appearance information given by RGB images is not resilient against typical fluctuations such as illuminance shift, making deployment of RGB-based vision systems in real-world settings difficult [13]. Depth sensors have become widely employed in recent years in a variety of applications, including object recognition, object detection, scene labelling, and action recognition [14-24]. Because depth information provides extra information about the object's spatial dimension, which is invariant to color fluctuation and light, it can considerably increase object identification ability [25]. Meanwhile, it has been demonstrated that utilizing RGB-D information for ADL improves classification accuracy significantly [26-27].

Most academics have focused their efforts on developing increasingly complex algorithms, while some have taken a different approach in quest of a new sort of representation that can better interpret the world. Cartas et al. [28] The ADLEgoDataset of Daily Living, which contains 105,529 annotated photos, was launched. The dataset was examined using the CNN+LSTM algorithm and attained an accuracy of 80.12%. Yu et al. [29] have proposed a knowledge-driven multisource fusion architecture for recognizing egocentric behaviors in everyday life (ADL). The Dezert–Smarandache theory and a convolutional neural network has trained to provide a collection of textual tags to distinguish ADL activities. Experiments indicate that the proposed technique correctly classified 15 preset ADL classes with an average accuracy of 85.4 percent. Heidarivincheh et al. [30] has combined classification and regression recurrent model. The recurrent voting node has been then forecasted the frame's relative location to display the completion moment. The system has been tested on RGBD-AC dataset and has achieved an accuracy of 89% for both complete and incomplete sequences.

Miron et al. [31] presented temporal convolutional neural network (Res-TCN) in assessing the correctness of a human motion or action. For every movement type, a different model has been trained with the explicit purpose of differentiating between optimal and non-optimal movements. This paper produced 71.2% results for certain actions but can also fall into the trap of classifying an incorrect execution of an action as a correct execution of another action used. Wang et al. [32] presented stacked auto encoder (SAE) deep learning model as a fundamental building blocks for ADL activities identification. The SAE has been then trained using greedy layer-wise. In order to fine-tune the network settings, a fully connected layer and a classification layer have been connected on top of the SAE. The model had achieved an accuracy of 90.43% on the publicly accessible ADL-based dataset. Xiao et al. [33] presented a deep convolutional model and a self-attention model-based end-to-end framework for recognizing human action from skeletal sequences. To extract discriminative deep features, DCM has used CNN. The SAM algorithm has been made to learn attention weights from variations in motion. Additionally, the attention weight for skeleton-based action categorization learned attentive deep characteristics. Over the SYSU-3D dataset, the efficacy of the suggested model has attained an accuracy of 80.36 %.

In this study, we have presented a novel type of symmetry principle based on the silhouette of a human. We have showed that the camera's location and the distance between the camera and the target can change. Each kind of skeleton will be processed once a skeleton has been retrieved from each RGBD (Red Green Blue Depth) frame. The upper half of the body, lower half of the body, center of the body, the position of the hands, and the variation of the hands are the four parameters discussed in this work.

The following is how the rest of the paper is organized: The solution framework, which includes silhouette detection, feature extraction, optimization, and classification, is described in Section II. The experimental results are presented in Section III, along with a comparison to other state-of-the-art systems. Finally, part IV brings the paper to a close.

II. SYSTEM DESIGN

The RGB and depth sequences from the RGBD-AC (RGBD-Action-Completion-2016) [34] dataset has been used in the proposed ADL system. The replacement operation, scaling operation, median filter, canny edge filter, and mapping depth image across RGB images have been employed in the preprocessing step to extract silhouette from RGBD images. These RGBD silhouettes have been then used to extract symmetry principal [35-42] features. With linearly dependent concept (LDC) [43] and long short-term memory-recurrent neural networks (LSTM-RNN) [44], these properties have been further improved and categorized. Finally, the leave one subject out (LOSO) method has been utilized to test and train the sequence of all ADL tasks. The basic diagram of our suggested ADL framework is shown in Fig. 1.

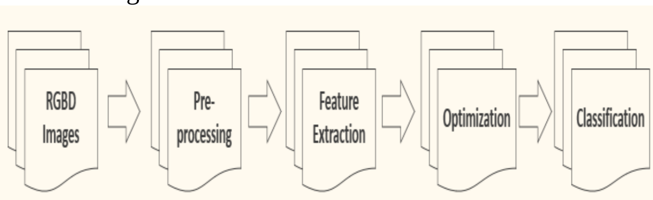

Fig. 1. Architecture diagram of the presented methodology for activities of daily living (ADL) system.

A. *Silhouette extraction of RGB-D images using depth silhouettes*

The pre-processing step accomplished through five phases.

1) Substitution operation: In the first step, missing depth values are interpolated by substitution operation [45]. The substitution operation for missing values of the depth images search through neighboring data on both left and right pixels [46]. Later, missing value is replaced with the larger of the two.

2) Scaling operation: In the second step, depth image has been linearly scaled [47] within a range of 0 to 255 as:

$$d' = \frac{d - d_{min}}{dmin_{max}\omega_{t\,arg\,et} + d'_{min}} \qquad (1)$$

Where, d depicts depth values of images before scaling. While d' is the depth values of images after scaling. Moreover, d_{min} and d_{max} represents minimum and maximum values in depth images before scaling. The d'_{min} depicts the lower bound of image, which is 0, and $\omega_{t\,arg\,et}$ depicts the upper bound of depth image which is 255.

3) Median Filter: In the third step, noise of the rescaled depth image has been removed using median filter [48].

4) Canny Edge Filter

In the fourth step, canny edge filter has been applied that use multi-stage algorithm to detect and correct depth silhouette edges [49-50].

5) Mapping over RGB image

In the last step, resulting depth silhouette is mapped to RGB images using affine transformation to extract silhouette from the RGB frames [51-57].

The advantage of extracting silhouette from depth images is that border information of silhouette can be detected even under poor light conditions [58-60]. The results of RGB and depth silhouette is illustrated in Fig. 2.

Fig. 2. Examples of silhouette extraction and background removal in RGBD images through five phases

B. *Feature Extraction*

The technique of extracting important information from raw data is known as feature extraction [61]. Detecting human operations including drink, open, pick, plug, switch, and pull requires several essential places in the human body [63]. Angles between two hands, angle between hand and upper half of body, angle between hand and lower half of body, angle between hand and center point of the body, and other aspects can all be used as motion indicators [64-68]. Some of the indicators are more stable than others. By stability, we mean that some points in the human body do not change because of movement [69-70]. For example, we found that the position of the head fluctuates relatively little during cyclic motions in all our samples [71-72]. In this paper, five features of the human body have been extracted: (1) The thirteen key points of the body has been detected; (2) The angle of hand and upper half of the body; (3) The angle between centers line of the body and hand; (4) The angle between lower half of the body and hand; (5) The angle formation between two hands. The implementation technique for our feature extraction approach has been described below.

1) Thirteen key points detection: Fig. 3 depicts a comprehensive overview of the human body key point detection model, which includes thirteen human body points that are classified as three key skeleton fragments: lower body, mid-point of the body, and upper body, and has been based on link of neck, head, shoulders, wrist, hand points, and elbow with knees, hips, foot, and ankle. Each part of the body plays a role in the completion of a specific task.

The central torso point has been calculated using the human silhouette's outer shape for the identification of human body points. The point 1/4 between the foot and the knee points has been used to determine the human ankle posture [73]. To estimate the wrist point, we divided the distance between the hand and elbow points by ¼ . The comprehensive explanation of human body part detection has been seen in Algorithm 1.

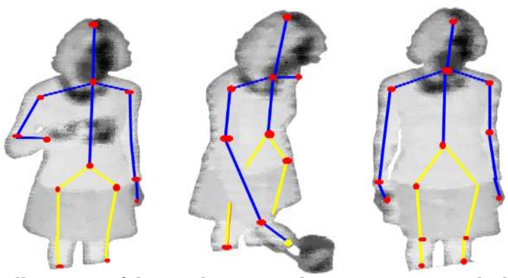

Fig. 3. Illustration of thirteen key points detection on extracted silhouette.

Algorithm1: Human silhouette key point detection

Input: Human silhouette
Output: Nineteen key points detection as neck, head, shoulders, wrists, hands, and elbows, knees, hips, foot, ankle, and torso.
H_{BS} = Human Body Silhouette
H = Height
W = width
L = Left
R = Right
J_h = Head
J_n = Neck

do
    For i = 1 to N
        I_h = *Get_Head_Point*(J_h)
        I_n = *Get_Upper_Point*(J_n)
        I_m = *Get_Mid_Body_Point*(H, W)
        I_f = *Get_Bottom_Body_Point*(H_{BS})
        I_{knp} = *Get_Midpoint*(I_m, I_f)
        I_{hnp} = *Get_Midpoint*(I_h, I_n)
        I_{elp} = *Get_Midpoint*(I_{hnp}, I_h)
        I_{wrp} = *Get_Midpoint*(I_{hnp}, I_{elp})/2
        I_{hip} = I_m
        I_{anp} = *Get_Midpoint*(I_{knp}, I_f)/4

While ((search (H_{BS}) & search (L, R))!=NULL)

return 13_Body_Points (head, neck, shoulders, elbows, wrists, hands, hips, torso, feet, knees, and ankles).

2) Angle between upper half of the body and hand: In RGBD-AC dataset, the angle formation in some activities have been formed between hand and upper half of the body such as switch, open and drink. Therefore, angle forms between upper half of the body and hand, is the most important feature of measuring activities performed by silhouette. It can be seen as meeting the decision requirement for the occurrence of the event at the given time. Because ADL activities require a quick transition from start to finish in RGBD-AC dataset, the time required for event detection has been set at once per fifteen adjacent frames, with a 0.5s time interval. The angle detection process [74] has been depicted in Fig. 4. The coordinates of the upper body point and hand is $I_m(t) = (x_{t1}, y_{t1})$ and $I_{hnd}(t) = (x_{t2}, y_{t2})$. At time t, the angle is expressed as;

$$\theta_t = arctan\left|\frac{y_{t1}-y_{t2}}{x_{t1}-x_{t2}}\right| \qquad (2)$$

where θ_t represents the satisfying the decision condition of the occurrence of the event at time *t*.

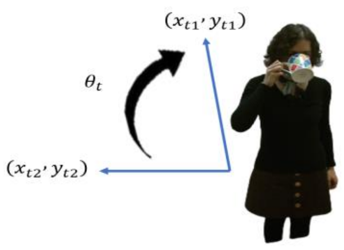

Fig. 4. Angle formation between upper body part and hand illustration on the silhouette.

3) Angle between center body point and hand: The angle creation between the center body point and the hand is the most visible aspect of silhouette in the process of ADL activities [75] in the RGBD-AC dataset as shown in Fig. 5. The characteristics of the degree of angle creation between the central body point and the hand while performing activities such as open and pull have been set at once every twenty-five adjacent frames. The coordinates of the center body point and hand is $I_m(t) = (x_{t3}, y_{t3})$ and $I_{hnd}(t) = (x_{t2}, y_{t2})$. At time t, the angle between center body point and the hand is expressed as:

$$\theta_t = arctan\left|\frac{y_{t3}-y_{t2}}{x_{t3}-x_{t2}}\right| \qquad (3)$$

where θ_t represents the satisfying the decision condition of the occurrence of the event at time *t*.

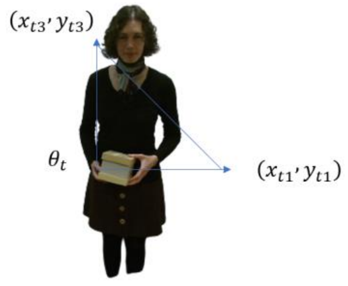

Fig. 5. Angle formation between center body point and hand illustration on the silhouette.

4) Angle between hand and lower half of the body: Angle formation has been generated between hand and lower part of the body in several activities in the RGBD-AC dataset, such as plug, switch, and pick. As a result, the most important element of assessing actions conducted by silhouette is the angle established between the lower half of the body and the hand as shown in Fig. 6. In the RGBD-AC dataset, the time necessary for event detection has been adjusted to once every fifteen adjacent frames, with a 0.5s time gap, because ADL activities require a rapid transition from start to conclusion. The lower body point and hand coordinates are expressed in (4). The coordinates of the lower body point and hand is $I_m(t) = (x_{t4}, y_{t4})$ and $I_{hnd}(t) = (x_{t2}, y_{t2})$. At time t, the angle is expressed as;

$$\theta_t = arctan\left|\frac{y_{t4}-y_{t2}}{x_{t4}-x_{t2}}\right| \qquad (4)$$

where θ_t represents the satisfying the decision condition of the occurrence of the event at time *t*.

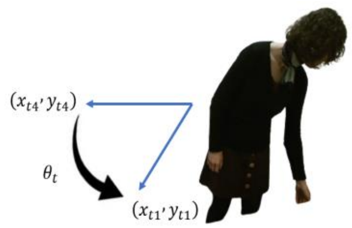

Fig. 6. Angle formation between hand and lower half of the body illustration on the silhouette.

5) Formation of angle between hands: The position of hands in each class of activities change systematically [76]. Therefore, formation of angle between hands postures can be used as a motion indicator. The coordinates of the right and left hands is $I_m(t) = (x_{t0}, y_{t0})$ and $I_{hnd}(t) = (x_{t2}, y_{t2})$. At time t, the angle is expressed as;

$$\theta_t = arctan\left|\frac{y_{t0}-y_{t2}}{x_{t0}-x_{t2}}\right| \qquad (5)$$

where θ_t represents the satisfying the decision condition of the occurrence of the event at time *t*. The angle formation between hands has been shown in Fig. 7.

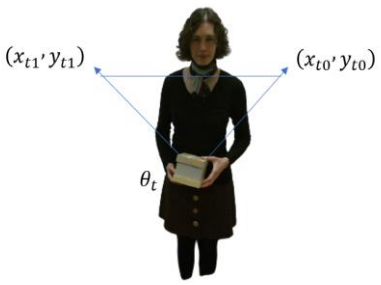

Fig. 7. Angle formation between hands illustration on the silhouette

C. *Linearly dependent concept (LDC) optimization*

In this paper, linearly dependent concept (LDC) has been used to reduce redundant features and to improve the accuracy of classifier [77]. There have been 793 features that have been considered as vectors. First, the vector set of features has been set as $x_1, x_2, \ldots, x_k \in V$ where V is a vector set of features. The V has been considered as a homogeneous linear combination system for all vectors and has been expressed as;

$$S = \sum_{i=1}^{k} c_i x_i \tag{6}$$

where, c_i is the element of real set and x_i is an element of V. The Gaussian elimination approach has been used to obtain the constant values. After implementing the suggested technique, 365 features have been chosen from the features vector, while 428 have been eliminated.

D. *Long short-term memory-recurrent neural networks (LSTM-RNN)*

Due to the lack of memory components in traditional neural networks, they may be unable to relate earlier knowledge to the current task to conduct reasoning about previous occurrences [78]. Recurrent neural networks (RNNs) are designed to processing sequential incoming data by allowing information to survive across loops in the network topology. Most of these achievements may be ascribed to the usage of LSTMs, a type of RNN capable of learning long-term dependencies [79]. As a result, we employed LSTM-RNNs to predict probable behaviors based on sequential sensor data observation.

We employed a simple LSTM model with one input layer, one hidden layer, and one output layer in our research. For the input, hidden, and output layers, the number of neurons was set to 10, 42, and 9, respectively. During training, the learning rate and batch size were set at 0.005, 1600, respectively. The results of LSTM-RNN classifiers have been depicted in Fig. 8.

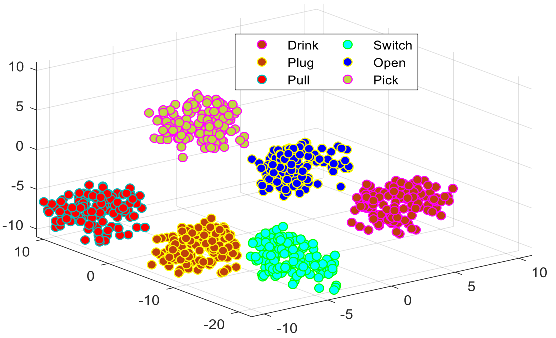

Fig. 8. Long short-term memory-recurrent neural networks (LSTM-RNN) visual results over RGBD-AC dataset.

III. Experimental Results

The dataset description, experimental findings, recognition accuracy, and a comparison of our technique to existing state-of-the-art ADL recognition systems are all presented in this part.

A. *RGBD-AC Dataset*

The RGBD-AC [34] dataset contains both complete and incomplete examples of various activities. A Microsoft Kinect v2 was used to create 414 sequences in the collection. Plug: plug socket, switch: turn off the light, open: open the jar, drink: drink from a cup, pull: pull the drawer, and pick: choosing item from the desk are the six actions captured in the sequences. The RGBD-AC dataset requirements for each action such that it could not be performed as; switch: subjects were asked to act as if they had forgotten to turn off the light, plug: subjects were given a plug that did not fit the socket, open: a lid was glued to the jar to prevent it from being opened, yank: a drawer that was locked could not be yanked. A total of eight individuals (3 females and 5 men) conducted at least four full and four partial sequences for each activity.

B. *Experimental Results*

With training and testing data, the suggested system was tested using the leave one subject out (LOSO) cross validation approach. To distinguish distinct postures and movements, the human activity classification system was tested using recall, precision, and F-measure. Table I depicts the confusion matrix of the RGBD-AC dataset for six different activities. The F-measure combines recall and precision respectively. The classification result of 92.83% has been achieved on F-measure over RGBD-AC (RGBD-Action-Completion-2016), reported in Table II. Finally, Table III shows a comparison of the proposed ADL approach with existing framework techniques on the RGBD-AC dataset. Overall, the findings revealed that our proposed strategy outperformed existing state-of-the-art methods.

TABLE I. CONFUSION MATRIX RESULTS OF SIX DIFFERENT ADL ACTIVITIES OBTAINED OVER RGBD-AC DATASET

RGBD-AC Dataset	***Drink***	***Open***	***Pick***	***Plug***	***Switch***	***Pull***
Drink	95.50	0	4.5	0	0	0
Open	0	92.00	3.5	2.5	1.5	0.5
Pick	2.5	0	94.50	0.5	0.5	2.0
Plug	0.5	1.5	0	93.50	4.5	0
Switch	0	0.5	2.5	5.5	90.00	1.5
Pull	0.5	0.5	3.5	3.5	0.5	91.50
Mean Accyuracy = 92.83%						

TABLE II. PRECISION, RECALL, AND F-MEASURE RESULTS OBTAINED BY PROPOSED METHOD USING LOSO VALIDATION SCHEME

RGBD-Action Completion-2016 dataset	***Precision***	***Recall***	***F-measure***
Drink	0.964	0.947	0.955
Open	0.931	0.910	0.920
Pick	0.951	0.940	0.945
Plug	0.950	0.922	0.935
Switch	0.900	0.902	0.900
Pull	0.915	0.917	0.915

TABLE III. COMPARISON OF RECOGNITION ACCURACY ON RGBD-AC DATASET WITH OTHER EXISTING FRAMEWORKS AND METHODS

Methods	Recognition Accuracy (%)
Res-TCN [31]	71.20
CNN+LSTM [28]	80.12
Deep Convolutional Model and a Self-Attention Model [33]	80.36
Dezert–Smarandache theory [29]	85.40
Classification and regression recurrent model [30]	89.00
Stacked auto encoder deep learning model [32]	90.43
Proposed approach	**92.83**

IV. CONCLUSION

Using RGBD silhouettes, we suggested an effective symmetry principle-based feature extraction methodology for ADL. The method we have suggested for extracting features from RGBD silhouettes is a combination of points and body form. On the publicly available RGBD-AC dataset, we trained and tested an ADL system uses LOSO and have achieved a mean recognition rate of 92.83%. Many applications can benefit from the suggested feature extraction method, including smart homes, autonomous video monitoring, and health care. We intend to construct our own ADL dataset in future for an online continuous environment that will make it more useful for ADL applications. We will also employ these characteristics for exceedingly compound situations.

REFERENCES

[1] A. Jalal, S. Lee, J. Kim, and T. Kim, "Human activity recognition via the features of labeled depth body parts," in Proceedings SHHT, pp. 246-249, 2012.

[2] A. Jalal, J. T. Kim, and T.-S Kim, "Development of a life logging system via depth imaging-based human activity recognition for smart homes," in Proc. of the ISSHB, pp. 91-95, 2012.

[3] A. Jalal, J. T. Kim, and T.-S. Kim, "Human activity recognition using the labeled depth body parts information of depth silhouettes," in Proc. of the ISSHB, pp. 1-8, 2012.

[4] A. Jalal, N. Sharif, J. T. Kim, and T.-S. Kim, "Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart homes," Indoor and Built Environment, vol. 22, pp. 271-279, 2013.

[5] A. Jalal, Y. Kim, and D. Kim, "Ridge body parts features for human pose estimation and recognition from RGB-D video data," in Proc. of the ICCCNT, pp. 1-6, 2014.

[6] A. Jalal and Y. Kim, "Dense depth maps-based human pose tracking and recognition in dynamic scenes using ridge data," in Proc. of the AVSS, pp. 119-124, 2014.

[7] F. Farooq, A. Jalal, and L. Zheng, "Facial expression recognition using hybrid features and self-organizing maps," in Proc. of the ICME, July 2017.

[8] M. Mahmood, A. Jalal, and H. A. Evans, "Facial expression recognition in image sequences using 1D transform and Gabor wavelet transform," in Proc. of the ICAEM, 2018.

[9] A. Jalal, Majid A. K. Quaid, and A. S. Hasan, "Wearable sensor-based human behavior understanding and recognition in daily life for smart environments," in Proc. of the FIT, 2018.

[10] M. Gochoo, I. Akhter, A. Jalal, and K. Kim, "Stochastic remote sensing event classification over adaptive posture estimation via multifused data and deep belief network," Remote Sensing, 2021.

[11] A. Jalal, M. A. K. Quaid, and M. A. Sidduqi, "A triaxial acceleration-based human motion detection for ambient smart home system," in Proc. of the ICAST, 2019.

[12] A. Jalal, M. Mahmood, and A. S. Hasan, "Multi-features descriptors for human activity tracking and recognition in Indoor-outdoor environments," in Proc. of the ICAST, 2019.

[13] A. Jalal and M. Mahmood, "Students behavior mining in E-learning environment using cognitive processes with information technologies," Education and Information Technologies, 2019.

[14] A. Jalal, A. Nadeem, and S. Bobasu, "Human body parts estimation and detection for physical sports movements," in proc. of C-Code, 2019.

[15] A. A. Rafique, A. Jalal, and A. Ahmed, "Scene understanding and recognition: statistical segmented model using geometrical features and Gaussian naïve bayes," in proc. of ICAEM, 2019.

[16] M. Batool, A. Jalal, and K. Kim, "Sensors technologies for human activity analysis based on SVM optimized by PSO algorithm," IEEE ICAEM conference, 2019.

[17] A. Ahmed, A. Jalal, and A. A. Rafique, "Salient segmentation based object detection and recognition using hybrid genetic transform," IEEE ICAEM conference, 2019.

[18] A. Jalal, M. A. K. Quaid, and K. Kim, "A wrist worn acceleration based human motion analysis and classification for ambient smart home system," JEET, 2019.

[19] K. Kim, A. Jalal and M. Mahmood, "Vision-based human activity recognition system using depth silhouettes: A Smart home system for monitoring the residents," JEET, 2019.

[20] A. Ahmed, A. Jalal, and K. Kim, "Region and decision tree-based segmentations for multi-objects detection and classification in outdoor scenes," in proc. of FIT, 2019.

[21] A. A. Rafique, A. Jalal, and K. Kim, "Statistical multi-objects segmentation for indoor/outdoor scene detection and classification via depth images," in proc. of IAST, 2020.

[22] A. Ahmed, A. Jalal, and K. Kim, "RGB-D images for object segmentation, localization and recognition in indoor scenes using feature descriptor and Hough voting," in proc. of IAST, 2020.

[23] M. Mahmood, Ahmad Jalal, and K. Kim, "WHITE STAG model: Wise human interaction tracking and estimation (WHITE) using spatio-temporal and angular-geometric (STAG) descriptors," Multimedia Tools and Applications, 2020.

[24] M. Quaid and Ahmad Jalal, "Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm," Multimedia Tools and Applications, 2019.

[25] A. Nadeem, A. Jalal, and K. Kim, "Human actions tracking and recognition based on body parts detection via Artificial neural network," in proc. of ICACS, 2020.

[26] S. Badar, A. Jalal, and M. Batool, "Wearable sensors for activity analysis using SMO-based random forest over smart home and sports datasets," IEEE ICACS, 2020.

[27] S. Amna, A. Jalal, and K. Kim, "An accurate facial expression detector using multi-landmarks selection and local transform features," IEEE ICACS, 2020.

[28] A. Cartas, P. Radeva and M. Dimiccoli, "Activities of daily living monitoring via a wearable camera: Toward real-world applications," in IEEE Access, vol. 8, pp. 77344-77363, 2020.

[29] H. Yu, W. Jia, Z. Li, et al., "A multisource fusion framework driven by user-defined knowledge for egocentric activity recognition," EURASIP Journal on Advances in Signal Processing, vol. 14, 2019.

[30] F. Heidarivincheh, M. Mirmehdi, and D. Damen,"Action completion: A temporal model for moment detection," arXiv, 2018.

[31] A. Miron, and C. Grosan, "Classifying action correctness in physical rehabilitation exercises," arXiv, 2021.

[32] A. Wang , S. Zhao, C. Zheng , J. Yang , G. Chen, C. Y. Chang, et al., "Activities of daily living recognition with binary environment sensors using deep learning: A comparative study," IEEE Sensors Journal, vol. 21, no. 4, 2021.

[33] R. Xiao, Y. Hou, Z. Guo, C. Li, P. Wang, W. Li, et al., "Self-attention guided deep features for action recognition," in Proceedings of the 2019 IEEE International Conference on Multimedia and Expo (ICME), pp. 1060–1065, 2019.

[34] F. Heidarivincheh, M. Mirmehdi, and D. Damen, "Beyond action recognition: Action completion in RGB-D Data," British Machine Vision Conference (BMVC), 2016.

[35] S. Badar, A. Jalal, and K. Kim, "Wearable inertial sensors for daily activity analysis based on Adam optimization and the maximum entropy Markov model," Entropy, vol. 22, no. 5, pp. 1-19, 2020.

[36] A. Jalal, N. Khalid, and K. Kim, "Automatic recognition of human interaction via hybrid descriptors and maximum entropy markov model using depth sensors," Entropy, 2020.

[37] A. Ahmed, A. Jalal, and K. Kim, "A novel statistical method for scene classification based on multi-object categorization and logistic regression," Sensors, 2020.

[38] M. Batool, A. Jalal, and K. Kim, "Telemonitoring of daily activity using accelerometer and gyroscope in smart home environments," JEET, 2020.

[39] A. Jalal, M. A. Quaid, S. B. Tahir, and K. Kim, "A study of accelerometer and gyroscope measurements in physical life-log activities detection systems," Sensors, 2020.

[40] A. Rafique, A. Jalal, and K. Kim, "Automated sustainable multi-object segmentation and recognition via modified sampling consensus and Kernel sliding perceptron," Symmetry, 2020.

[41] A. Jalal, M. Batool, and K. Kim, "Stochastic recognition of physical activity and healthcare using tri-axial inertial wearable sensors," Applied Sciences, 2020.

[42] A. Jalal, I. Akhtar, and K. Kim, "Human posture estimation and sustainable events classification via Pseudo-2D stick model and K-ary tree hashing," Sustainability, 2020.

[43] A. Jalal, M. Batool, and K. Kim, "Sustainable wearable system: Human behavior modeling for life-logging activities using K-Ary tree hashing classifier," Sustainability, 2020.

[44] S. B. Tahir, A. Jalal, and K. Kim, "IMU sensor based automatic-features descriptor for healthcare patient's daily life-log recognition," in proc. of IAST, 2021.

[45] I. Akhter, A. Jalal, and K. Kim, "Pose estimation and detection for event recognition using sense-aware features and Adaboost classifier," IEEE IBCAST, 2021.

[46] A. Jalal, M. Batool and B. Tahir, "Markerless sensors for physical health monitoring system using ECG and GMM feature extraction," IEEE IBCAST, 2021.

[47] J. Madiha, A. Jalal, and K. Kim, "Wearable sensors based exertion recognition using statistical features and random forest for physical healthcare monitoring," IEEE ICAST, 2021.

[48] K. Nida, M. Gochoo, A. Jalal, and K. Kim, "Modeling two-person segmentation and locomotion for stereoscopic action identification: A sustainable video surveillance system," Sustainability, 2021.

[49] P. Mahwish, A. Jalal, and K. Kim, "Hybrid algorithm for multi people counting and tracking for smart surveillance," IEEE IBCAST, 2021.

[50] A. Ahmed, A. Jalal, and K. Kim, "Multi-objects detection and segmentation for scene understanding based on texton forest and kernel sliding perceptron," JEET, 2020.

[51] J. Madiha, M. Gochoo, A. Jalal, and K. Kim, "HF-SPHR: Hybrid features for sustainable physical healthcare pattern recognition using deep belief networks," Sustainability, 2021.

[52] S. Amna, A. Jalal, M. Gochoo and K. Kim, "Robust active shape model via hierarchical feature extraction with SFS-optimized convolution neural network for invariant human age classification," Electronics, vol. 10, no. 4, 2021.

[53] A. Jalal, A. Ahmed, A. Rafique and K. Kim "Scene semantic recognition based on modified fuzzy c-mean and maximum entropy using object-to-object relations," IEEE Access, 2021.

[54] A. Jalal, M. Mahmood, and M. A. Sidduqi, "Robust spatio-temporal features for human interaction recognition via artificial neural network," IEEE Conference on FIT, 2018.

[55] A. Hira, A. Jalal, M. Gochoo, and K. Kim, "Hand gesture recognition based on auto - landmark localization and reweighted genetic algorithm for healthcare muscle activities," Sustainability, 2021.

[56] N. Amir, A. Jalal, and K. Kim, "Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy markov model," Multimedia Tools and Applications, 2021.

[57] I. Akhter, A. Jalal, and K. Kim, "Adaptive pose estimation for gait event detection using context-aware model and hierarchical optimization," JEET, 2021.

[58] M. Gochoo, S. Badar, A. Jalal, and K. Kim, "Monitoring real-time personal locomotion behaviors over smart indoor-outdoor environments via body-worn sensors," IEEE Access, 2021.

[59] P. Mahwish, G. Yazeed, M. Gochoo, A. Jalal, S. Kamal et al., "A smart surveillance system for people counting and tracking using particle flow and modified SOM," Sustainability, 2021.

[60] M. Gochoo, S. R. Amna, G. Yazeed, A. Jalal, S. Kamal et al., "A systematic deep learning based overhead tracking and counting system using RGB-D remote cameras," Applied Sciences, 2021.

[61] Y. Ghadi, I. Akhter, M. Alarfaj, A. Jalal, and K. Kim, "Syntactic model-based human body 3D reconstruction and event classification via association based features mining and deep learning," PeerJ Computer Science, 2021.

[62] W. Manahil, A. Jalal, M. Alarfaj, Y. Ghadi, S. Tamara, S. Kamal, and D. Kim, "An LSTM-based approach for understanding human interactions using hybrid feature descriptors over depth sensors," IEEE Access, 2022.

[63] Y. Ghadi, J. Madiha, M. Alarfaj, S. Tamara, A. Suliman, A. Jalal, S. Kamal, and D. Kim, "MS-DLD: multi-sensors based daily locomotion detection via kinematic-static energy and body-specific HMMs," IEEE Access, 2022.

[64] A. Usman, Y. Ghadi, S. Tamara, A. Suliman, A. Jalal, and J. Park, "Smartphone sensor-based human locomotion surveillance system using multilayer perceptron," Applied Sciences, 2022.

[65] A. Ayesha, Y. Ghadi, M. Alarfaj, A. Jalal, S. Kamal, and D. Kim, "Human pose estimation and object interaction for sports behaviour," CMC, 2022.

[66] B. Mouazma, Y. Ghadi, A. Suliman, S. Tamara, A. Jalal, and J. Park, "Self-care assessment for daily living using machine learning mechanism," CMC, 2022.

[67] Y. Ghadi, W. Manahil, S. Tamara, A. Suliman, A. Jalal, and J. Park, "Automated parts-based model for recognizing human-object interactions from aerial imagery with fully convolutional network," Remote Sensing, 2022.

[68] Y. Ghadi, R. Adnan, S. Tamara, A. Suliman, A. Jalal, and J. Park, "Robust object categorization and Scene classification over remote sensing images via features fusion and fully convolutional network," Remote Sensing, 2022.

[69] S. Tamara, J. Madiha, M. Gochoo, A. Suliman, Y. Ghadi, A. Jalal, and J. Park, "Student's health exercise recognition tool for E-learning education," IASC, 2022.

[70] Y. Ghadi, W. Manahil, M. Gochoo, A. Suliman, S. Chelloug, A. Jalal, and J. Park, "A graph-based approach to recognizing complex human object interactions in sequential data," Applied Sciences, 2022.

[71] A. Alam, S. Abduallah, A. Israr, A. Suliman, Y. Ghadi, S. Tamara, and A. Jalal, "Object detection learning for intelligent self automated vehicles," IASC, 2022.

[72] Y. Ghadi, A. Israr, A. Suliman, S. Tamara, A. Jalal, and J. Park, "Multiple events detection using context-intelligence features," IASC, 2022.

[73] H. Sadaf, Y. Ghadi, M. Alarfaj, S. Tamara, A. Jalal, S. Kamal, and D. Kim, "Sensors-Based Ambient Assistant Living via E-Monitoring Technology," CMC, 2022.

[74] M. Jamil, A. Shaoor, A. Suliman, Y. Ghadi, S. Tamara, A. Jalal, and J. Park, "Intelligent Sign Language Recognition System for E-learning Context," CMC, 2022.

[75] S. Tamara, A. Israr, A. Suliman, Y. Ghadi, A. Jalal, and J. Park, "Pedestrian Physical Education Training over Visualization Tool," CMC, 2022.

[76] M. Alarfaj, W. Manahil, Y. Ghadi, S. Tamara, A. Suliman, A. Jalal, and J. Park, "An Intelligent Framework for Recognizing Social Human-Object Interactions," CMC, 2022.

[77] Y. Ghadi, B. Mouazma, M. Gochoo, A. Suliman, S. Tamara, A. Jalal, and J. Park, "Improving the Ambient Intelligence Living using Deep Learning Classifier," CMC, 2022.

[78] R. Hammad, M. Muneed, A. Suliman, Y. Ghadi, A. Jalal, and J. Park, "Home Automation-based Health Assessment along Gesture Recognition via Inertial Sensors," CMC, 2022.

[79] M. Mushhood, S. Shizza, Y. Ghadi, A. Suliman, A. Jalal, and J. Park, "Body Worn Sensors for Health Gaming and e-learning in Virtual Reality," CMC, 2022.