Intelligent Fine-Grained Daily Living Locomotion Prediction Based on Skeleton Modeling and CNN

Madiha Javeed[1*] , Naif Al Mudawi[2], Abdulwahab Alazeb[2], Hanan Aljuaid[3], Mohammed Hamad Alatiyyah[4], Khaled Alnowaiser[5], Asaad Algarni[6], Ahmad Jalal[1,7]

[1] Department of Computer Science, Air University, E-9, Islamabad 44000, Pakistan
[2] Department of Computer Science, College of Computer Science and Information System, Najran University, Najran 55461, Saudi Arabia
[3] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[4] Department of Computer Science, College of Sciences and Humanities in Aflaj, Prince Sattam bin Abdulaziz University, Al-Kharj 16278, Saudi Arabia
[5] Department of Computer Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[6] Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia
[7] Department of Computer Science and Engineering, College of Informatics, Korea University, Seoul 02841, South Korea

Corresponding Author Email: mjaved84@gmail.com

ABSTRACT

Activities of daily living are important for human locomotion prediction. Humans perform several activities while moving from one place to another. Some of these activities are coarse-grained while others are fine-grained. However, the challenges remain for efficient locomotion prediction for daily living activities. A couple of important challenges faced are modeling of the human skeleton along with extraction of the relevant feature. Therefore, this research focuses on locomotion prediction while performing daily tasks. For this, a benchmarked publicly available dataset known as Opportunity++ has been considered to conduct experiments. The paper proposed a system that converted the videos from dataset into sequences of frames and filtered them. Next, background of each frame is removed using background frame subtraction technique. Then, human is detected via the skeleton modeling method based on five skeleton points. Furthermore, stochastic features have been extracted from skeleton model. Using the quadratic discriminant analysis method, these extracted features are then optimized due to the large vector size. Lastly, locomotion detected has been validated via convolutional neural network. Moreover, the experimental results show that the proposed methodology has achieved improvement in terms of accuracy rate of 82.94%. Compared to other state-of-the-art systems in literature, the proposed system outperformed.

1. INTRODUCTION

Locomotion prediction is an important aspect of the computing world due to multiple locomotion modes in various conditions and terrains in humans' daily routine [1]. Accurate locomotion prediction is necessary to provide proper care and emergency services to elderly or patients. Though the activities of daily living (ADL) are of many types, thus multiple studies have focused on different aspects of daily locomotion prediction. Machine learning [2] and deep learning [3] have been utilized to predict daily locomotion. Likewise, using multiple types of sensors is also common for locomotion prediction. A few studies proposed the use of motion sensors like inertial sensors and electromyography (EMG) devices [4], while others preferred to utilize the vision-based data captured

from indoor-outdoor environments [5].

Current studies in the field of locomotion prediction [6-8] have focused over motion sensors or visual sensors-based ADL detection. Few of them has misclassifications while others did not treat complex and simple motion patterns separately. Vision-based researches focused on the accuracy while not focusing on extraction of human silhouette that limits the applications of proposed method. The proposed fine-grained daily locomotion prediction system has specific healthcare-related applications such as assisted living for elderly home care, patient's healthcare monitoring [9], smart home-based applications [10], and disability-risk recognition. These applications require finer-grained locomotion detection [11] to provide a critical platform for ADL-related safety and independence at home. Therefore, this research has focused on

fine-grained ADL for locomotion prediction. The major contributions of the proposed system are as follows.

1) An intelligent robust system has been proposed for locomotion prediction of ADL using state-of-the-art skeleton modeling and convolutional neural network (CNN). When compared to the existing systems, it is evident that the proposed locomotion prediction system has outperformed in terms of recognition results.

2) A benchmarked dataset called Opportunity++ [12] has been utilized for the experiments. There are a few researchers who have utilized this dataset, therefore the usage of such a novel dataset is important for the contribution to this research area.

3) A novel technique has been proposed to extract the human body through motion detection. This state-of-the-art technique has helped in extracting the human body in its excellent form shown through the skeleton points accuracies achieved.

4) Stochastic features like heat maps and saliency maps have been introduced into the system through the features extraction phase. These features have supported the system in terms of robustness towards the real-life application.

5) The system is more focused on finer-grained activities in the dataset, which will be helpful in recognizing the critical ADL detection in the system. Fine-grained actions are more significant to detect as these are used comprehensively to achieve the final ADL recognition.

The rest of the paper has been divided into multiple sections. Section II provides a literature review of the signal and video features-based locomotion prediction. Section III describes the proposed implemented system in detail. Section IV contains the experiments performed along with their results and a comparison with other state-of-the-art methodologies. Finally, section V elucidates the concluding remarks about the paper and some limitations, and future directions.

2. RELATED WORK

Automatic prediction of different gait events and ADL recognition for locomotion prediction is the primary purpose of this study. Multiple methodologies have been proposed in the past for this purpose. Some researchers focused on the features based on the video domain while others preferred to extract the features found on one-dimensional techniques. Following is a literature review of video features and signal features-based locomotion prediction models proposed in the past.

2.1 Signal features-based locomotion prediction

Different authors of research articles have proposed different works related to the signal features-based locomotion prediction of ADL. Table 1 gives a glimpse of such systems that are extracting signal-based features.

2.2 Video features-based locomotion prediction

Various researchers have looked into the methods used to predict locomotion based on video features. Table 2 summarizes such systems in detail and their limitations.

Table 1. Literature review for signal features-based locomotion prediction

Authors	Systems	Limitations
Jonic et al. [13]	Three machine learning techniques have been used to predict locomotion and ADLs, multilayer perceptron, an adaptive-network-based fuzzy inference system, and a combination of entropy minimization type of inductive learning technique along with the radial basis function type of artificial neural network. The authors have proposed this system to calculate the activity patterns and to determine a mapping mechanism between kinematics and patterns.	The proposed method has certain misclassifications because of the related randomization of multiple classes.
Gochoo et al. [14]	A system to predict locomotion through multiple steps. First, they denoised the sensory data followed by windowing of filtered data. Next, they utilized the time-domain, wavelet-domain, and time-frequency-domain features pool. Furthermore, they have selected the related features, and finally, a reweighted genetic algorithm has been applied to classify and learn the locomotion actions.	The system's performance is limited due to the same static and kinematic signals treatment.
Azmat and Jalal [15]	A smartphone-based locomotion prediction system consisted of pre-processing, windowing, segmentation, features extraction, and classification steps. Another step called template matching has been introduced in between the pre-processing and features extraction stages. It helped in determining the samples for complex and static locomotion behavior.	The system attained less accuracy due to misclassifications.
Zhou et al. [16]	The authors have proposed a system to detect gait intents. Time, frequency, and time-frequency domain-based features were used to extract features from raw data. Further, ensemble learning methods are introduced to classify the outputs.	The proposed methodology lacks the implementation of the initial filter.
Bejinariu et al. [17]	An evaluation method for rehabilitation centers has been proposed in this paper. Videos of patients performing certain actions have been used as input to the system. First, the walking patient's position was estimated via VGGNet CNN. Next, the joint angles were calculated followed by a study of variation in angles. Abnormal values of the angles told about the locomotion injuries and their recovery.	Only knee angles have served the purpose of statistical analysis resulting in low performance.
Pan et al. [18]	A more detailed action-level recognition is required for patients monitoring and elderly care centers. The heterogeneous multi-modal cyber-physical system had two phases. The first one contained the distributed vibration sensing-based event detection and activity recognition. The second one had the single-point electrical load sensing-based event detection and activity recognition. Based on both phases, the spatiotemporal aware event prediction ensemble for ADL and activity type has been recognized.	The detection accuracy of the system degrades in a complex situation, which limits the system's accuracy.

Table 2. Literature review for video features-based locomotion prediction

Authors	Systems	Limitations
Yan et al. [19]	A system to determine user movement intentions has been suggested by the authors. A system based on depth images has been proposed for locomotion recognition. Features extraction subsystem and finite-state machine subsystem have been used to make it possible to detect the locomotion mode. The depth features have been extracted using local average depth values and stair edge detection.	The system's accuracy has been decreased due to no filtration technique applied.
Moriya et al. [20]	An internet-of-things-based technology has been introduced to recognize ADL through machine learning. They have deployed a network in a home-based environment and used ECHONET Lite-ready appliances and motion sensors for features extraction. However, the system could achieve a classification accuracy of 68% only.	The model was not effective for real-time scenarios due to low performance.
Khalid et al. [21]	The RGB data was pre-processed through background subtraction and silhouette detection. The depth data was pre-processed using morphological closing, edge detection, area-based filter, and silhouette detection. Then, the features are extracted via geodesic distance, 3D Cartesian plane, joints MOCAP, and Way-point trajectory. Further, the features are combined and optimized using a particle swarm followed by the neuro-fuzzy classification technique.	The proposed method could not detect the overlapped silhouette, hence lower performance has been observed.
Zhang et al. [22]	The authors proposed an intelligent video surveillance system where the images are acquired followed by communication. Then, the moving targets are detected and binary images are post-processed. Further, the moving targets are tracked and sports behavior is analyzed. The feedback is given to the feedback control system.	Due to shape variability, environment changes, and real-time requirements, the moving target processes face many challenges and hence cause low accuracy rates.
Wang et al. [23]	A simple and efficient technique was proposed to address the gait recognition problem by identifying people by the way they walk. First, the background is subtracted and moving silhouettes are segmented and tracked. Then, principal component analysis based on Eigen space transformation was applied to reduce the dimensionality of the input features. Further, the patterns are classified using a supervised method.	The limitations of the proposed technique are the gait patterns need to be diversified and the lack of generality of viewing angles.
Guo et al. [24]	The paper introduced a two-step process to identify 3D pose and shape representations, where 3D pose sequences are processed and rendered by the shape representations. Lie algebraic theory was also used to represent human motion followed by extracting the 3D detailed shape from different views.	A small dataset was used that does not provide stable results from the proposed process.

3. PROPOSED IMPLEMENTED SYSTEM
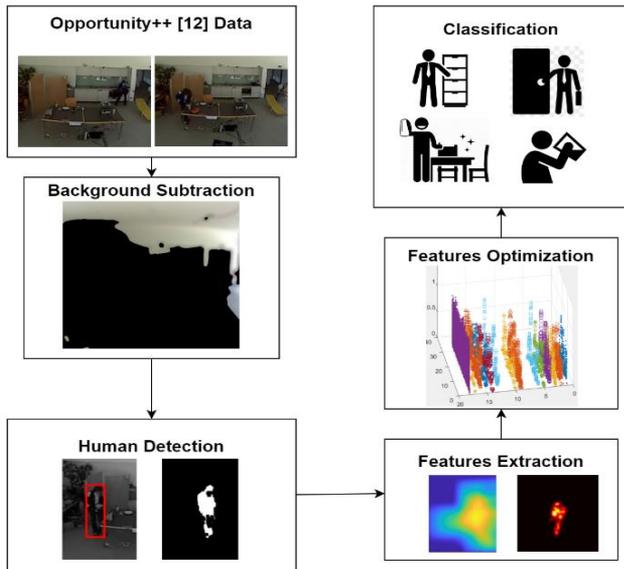
Figure 1. Architecture flow diagram for intelligent fine-grained daily living locomotion prediction system

This section describes the proposed implemented system in detail. The problems of misclassification and correct human skeleton points detection have been taken into consideration here and a model based on visual sensors has been proposed. First, the videos have been acquired from a well-known dataset named Opportunity++ [12]. Next, the data acquired has been pre-processed through filtration [25] and background subtraction processes to get more precise raw data. Then, a human skeleton has been exhibited by finding relevant skeleton points on the human body. Further, the features have been extracted using energy and saliency map features extraction techniques. These two feature extraction methods are based on skeleton points detected. Moreover, these extracted features are then reduced by using quadratic discriminant analysis methodology to avoid dimensionality issue. Finally, CNN has been utilized for human locomotion prediction via ADL classification. Figure 1 has illustrated the idea of the implemented system through an architecture flow diagram.

3.1 Data pre-processing

For data pre-processing, we have processed the image sequences from videos of the Opportunity++ [12] dataset. Then, we filtered those image sequences by reducing noise as below.

3.1.1 Weiner filter

The image sequences are more sensitive toward motion blur noise and can cause more computational costs for the implemented system [26]. Therefore, we proposed to reduce the noise in image sequences through a Wiener filter [27]. This filter calculates the mathematical relationships as in Eq. (1).

$$w(x, y) = \frac{H^*(x, y)}{|H(x, y)|^2 + K(x, y)} \tag{1}$$

where, $w(x, y)$ is the actual data, $K(x, y)$ represents the filter values, and $H^*(x, y)$ gives the large area [28]. Figure 2 indicates the pre-processed data sample along with its original image. From the figure, we can see that filtered image has less noise as compared to the original image.

Figure 2. Pre-processed image sequence for (a) Original image (b) Filtered Image over Opportunity++ dataset

3.2 Background subtraction and human detection

Data pre-processing is being followed by the background subtraction from images applied over the filtered image sequences [29]. For this purpose, we have proposed to select a base background image from the dataset, converted that image into a binary image [30], and used it as background image to subtract the background from image sequences [31]. Then, we have detected the movable parts from these sequences of images [32] and used its binary image along with the background subtraction to get human movable parts from the image sequences [33]. Figure 3 presents a sample image with its movable parts detected for further processing. It is evident from the figure that a complete human can be extracted from the moveable parts detected by this methodology.

Figure 3. Human detection through moveable parts in a sample frame sequence from Opportunity++ dataset

3.3 Skeleton modeling

Figure 4. Human Skeleton model containing five detected skeleton points and their connections

Furthermore, we have extracted the centroids from blobs [34] in order to detect the human body's five skeleton points [35]. The five points include head, shoulder, elbow, hand, and

foot [36]. Figure 4 shows the detected skeleton model from a random image sequence, where the yellow line shows the head point connected to the shoulder skeleton point, the green line represents the shoulder connected to the elbow skeleton point, the red line gives the elbow connected to the hand skeleton point, and the cyan colored line offers the skeleton points connection between elbow and foot. The red dots represent all the skeleton points detected in the proposed methodology. This shows a human skeleton model extracted from the five skeleton points detected.

3.4 Features extraction

Locomotion prediction for intelligent ADL is an important task that consists of another vital step known as features extraction. The features are extracted from the skeleton points and help in identifying the salient features from the frame sequences. This paper has proposed to extract two different feature types comprising of energy and saliency map features.

3.4.1 Energy features

Energy features have been extracted as the human body parts move from one frame to another and have been captured in the form of a heat map [37]. The more the human parts move, the more yellowish colors are formed on the map, whereas the lesser motion of a human will generate more reddish or blackish color over the map [38]. The map has been represented in the shape of a matrix and the values of matrix range from 0 to 6000 that is calculated as in Eq. (2) and Eq. (3).

$$H_k = \frac{p_k}{\acute{p}_{n\epsilon k}} \tag{2}$$

$$TM(x) = \sum_{i=0}^{k} lnR(i) \tag{3}$$

where, x is the 1D vector consisting of extracted vector, i is the index value, and R refers to the RGB values [39]. Figure 5 displays the heat map extracted from an image sequence.
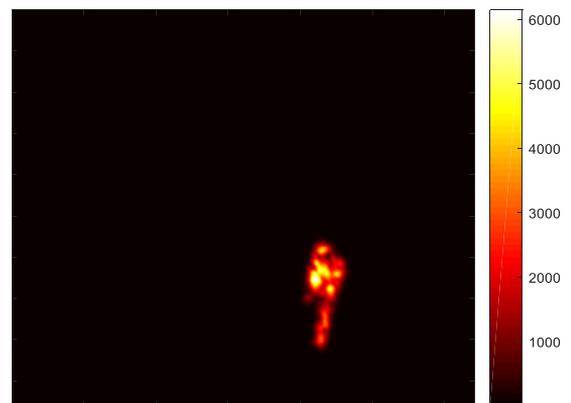
Figure 5. Energy features extraction results

3.4.2 Saliency maps

The next extracted features are through salient locations of a frame because the entire frame cannot be processed at the same time [40]. The fixated region is analyzed and then attention is redirected to other salient regions using saccades movements that are more to be focused upon [41]. Saliency

maps are a successful biologically plausible technique for modeling visual attention and hence being used in this study [42]. The generalized Gaussian distribution is also utilized to model each of these and calculated as in Eq. (4):

$$P(f_i) = \frac{\theta_i}{2\sigma_i \gamma(\theta_i^{-1})} \exp\left(-\left|\frac{f_i}{\sigma_i}\right|^{\theta_i}\right) \qquad (4)$$

where, $\theta_i > 0$ is the shape parameter, $\sigma_i > 0$ describes the scale parameter, and γ is the gamma function [43]. Figure 6 illustrates the saliency map extracted from an image sequence of Opportunity++ dataset.
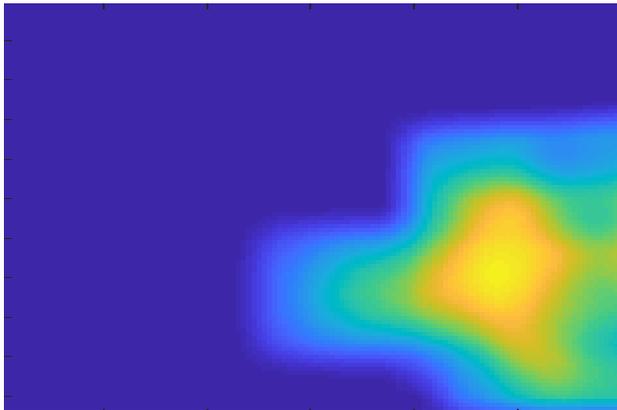
Figure 6. Saliency map extracted for an image sequence over Opportunity++ dataset

3.5 Features optimization

After extracting the stochastic features, dimensions of the vector increased enormously. In order to make the features vector as connected to the goals as possible, this study proposed to apply quadratic discriminant analysis over the extracted features vector of energy and saliency map features. Other techniques for features optimization such as, sequential forward selection, linear discriminant analysis, genetic algorithm, and factorial discriminant analysis have been applied over this system as well. Due to the non-linear nature of data in this system, we have found quadratic discriminant analysis (QDA) to give the superlative outcomes.

3.5.1 Quadratic discriminant analysis
The extracted features from both the techniques explained in the previous subsection are present in a non-linear way [44]. Hence, quadratic discriminant analysis is supportive in features optimization of such non-linear vectors when compared to other linear methods [45]. We cannot assume the locomotion dispersion, therefore QDA is a good option in our proposed scenario [40]. The covariance matrix M_i is calculated for each ADL as $i \in \{1, \ldots \ldots, I\}$. QDA can be calculated as in Eq. (5):

$$\sigma_i(x) = -\frac{1}{2}\log\left|\sum_i\right| - \frac{1}{2}(x - \varphi_i)^T \sum_i^{-1}(x - \varphi_i) + \log \pi_i \qquad (5)$$

where, x represents the extracted features vector and π_i gives I activity priors [46]. Figure 7 explains the optimization of features using QDA from fine-grained activities of the Opportunity++ dataset [12].
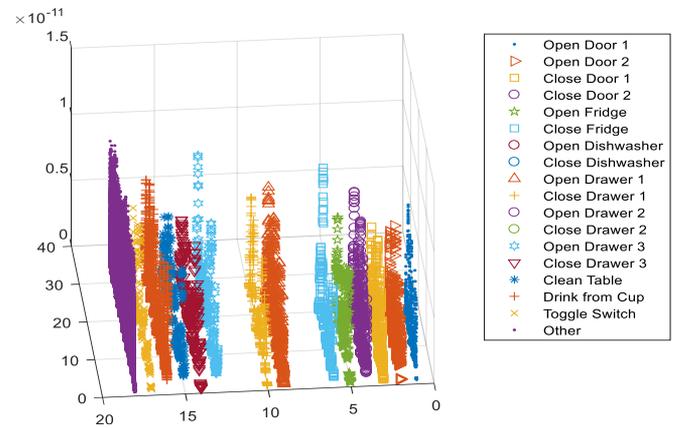
Figure 7. Optimized features extracted via QDA over the Opportunity++ dataset

3.6 Classification through CNN

CNN has been utilized in learning algorithms in order to recognize classes, objects, actions etc. It is a neural network based architecture that learns from given input data. A variety of methodologies has been utilizing the CNN as a classification step [47]. There are one-dimensional and two-dimensional CNN implementations [48]. However, multi-dimensional CNN can also be used to classify fine-grained activities [49]. It can solve complex tasks in less time than conventional artificial neural networks [50, 51].
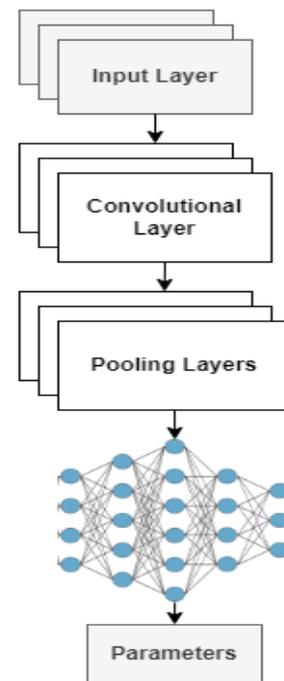
Figure 8. CNN architecture diagram

CNN consists of different layers providing input to each other, such as convolutional, stride, max pooling, and fully connected layers [52]. The proposed CNN architecture consists of input layer, convolutional layer1, pooling layer1, convolutional layer2, pooling layer2, two fully connected layers, and finally a softmax layer. This architecture was implemented for two times and was compared to other architectures of CNN as well. But it provided the best possible results, hence we proposed this architecture of CNN for ADL

recognition. This study also applies the activation function named ReLU (rectified linear unit). The ReLU function trains by mapping the negative values to zero and by keeping positive values only. The learning rate was set to 0.0005 and the maximum epoch was also set to 200. Figure 8 displays the proposed method's structure applied over the Opportunity++ dataset. The input data from the selected dataset has been given in the form of reduced energy and saliency map features from QDA. Then, convolutional layer was applied where a set of convolutional filters has been used to activate the important features from video sequences. Further, a max-pooling layer is applied to downsample output and multiple parameters used for network learning. Finally, a fully connected artificial neural network has been used to recognize the ADL. Whole set of parameters and activation size have been displayed in Table 3.

Table 3. CNN-based parameters

Layer	Activation Shape	Activation Size	#Parameters
Input Layer	(32,32,3)	3072	0
CONV1(f=4,s=1)	(28,28,8)	6272	392
POOL1	(14,14,8)	1568	0
CONV2(f=4,s=1)	(10,10,16)	1600	2064
POOL2	(5,5,16)	400	0
FC3	(120,1)	120	48120
FC4	(84,1)	84	10164
Softmax	(17,1)	17	1445

4. EXPERIMENTAL OUTCOMES

In this section, a detailed review of experiments performed along with their outcomes has been given. The experiments have been performed on the MATLAB tool. The hardware system used was Intel Core i7-8th Gen with 2.40GHz processing power, 24GB RAM having x64 based Windows 10. We used a publicly available dataset for validation of the proposed system. First, we assessed the performance of ADL detection through the Opportunity++ dataset [12]. Next, we evaluated the performance using precision, recall, and F1-score. Then, the last part compared the proposed system with other state-of-the-art models. This section is split into two subsections: dataset description and performance metrics and results.

4.1 Datasets description

A publicly available dataset called Opportunity++ [12] having 25 hours of video data from 12 subjects performing coarse-grained and finer-grained types of activities inside a room has been used for experimental examination of the implemented system. This system focused on 17 finer-grained level activities such as *open door*, *close door*, *open fridge*, *close fridge*, *open dishwasher*, *close dishwasher*, *open drawer*, *close drawer*, *clean table*, *drink from cup*, and *toggle switch*. There were 6 different runs for each subject having five ADL runs followed by a drill run scenario [53]. The runs included different ADL situations like the *groom*, *relax*, *preparing food*, *eating food*, *cleaning up*, and *break*. The videos were recorded using 640×480 pixels resolution and 10 frames per second speed. The videos were anonymized to hide the subjects' identity [54, 55]. Figure 9 gives a few image sequence samples from the dataset.

Figure 9. Sample image sequences from Opportunity++ dataset [12]

The variety of activities on both levels including fine-grained level and high-level have suggested that this system can be applicable to a number of daily life routine actions. The diversity in captured data using multiple sensors has also recommended the robustness of this system for efficient results. These factors support the long-establishment of our proposed system and it can also be generalized to other datasets and the real-world human locomotion related situations.

4.2 Performance metrics and results

We utilized multiple performance evaluation metrics to assess the performance of the implemented system. The performance of ADL detection was evaluated with five evaluation metrics including accuracy, precision, recall, F1-score, and confusion matrix calculated as in Eqs. (6)-(8), and Eq. (9) as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{F1} - \text{score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right) \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

where, *TP* means the true positive values, *TN* gives the true negative values, *FP* is the false positive values, and *FN* is the false negative values. Accuracy in Eq. (6) depicts how accurate the implemented system and precision is in Eq. (7) mentions the confidence in the implemented system's results. In Eq. (8), the F1-score states the system's accuracy over the Opportunity++ dataset. The recall in Eq. (9) presents the fraction of relevancy among the finer-grained level activities.

4.2.1 Experiment 1: Locomotion ADL recognition over opportunity++ dataset

CNN has been used to test the implemented system over the Opportunity++ dataset [56]. We chose the k-fold cross-validation technique to avoid overfitting issues [57]. The results are given in the form of a confusion matrix [58] in Table 4. A total of 17 fine-grained level activities have been classified through CNN [59]. The mean accuracy achieved over Opportunity++ video data is 82.94%. Table 5 contributes towards the precision, recall, and F1-score calculated for each fine-grained locomotion ADL in the system. The system also showed a low computational cost of 0.59ms, which demonstrates that the proposed locomotion ADL recognition

system is quite efficient. A receiver operating characterictic (ROC) curve has been shown in Figure 10 to describe the performance of proposed system in terms of false positive rate and true positive rate using Eq. (10) and Eq. (11) as:

$$FPR = \frac{FP}{FP + TP} \qquad (10)$$

$$TPR = \frac{TP}{TP + FP} \qquad (11)$$

where, *TP* shows the actual true positives and *FP* gives the actual false positives.

4.2.2 Experiment 2: Locomotion ADL detection comparison with state-of-the-art methods

The proposed system has been compared with other well-known methods in this experiment. Table 6 compares with other state-of-the-art systems [60, 61] based on locomotion prediction models. Different models were proposed in the past for ADL recognition via locomotion prediction [62].

The authors [63] proposed an ADL recognition system based on videos of actions performed by elderly people and a CNN-based network. Donaj and Maučec [64] suggested an hidden Markov models based ADL recognition strategy by modifying the Viterbi algorithm. A four modules based method for locomotion prediction has been presented in the study [65], namely, processing, extraction, optimization, and recognition modules. A multi-layer perceptron based approach

has been proposed using IoT-based data from Opportunity++ dataset [66]. Next, they pre-processed and extracted a bag of features, which is further optimized. The authors [67] have extracted ten skeleton points to further extract cues from the data from Opportunity++. They have utilized deep belief network to identify the actions. Micro activities have been identified by Sridharan et al. [68] by location-aware algorithm. A transfer learning methodology-based system has been proposed in the study [69], which used random forest classifier to recognize ADL.
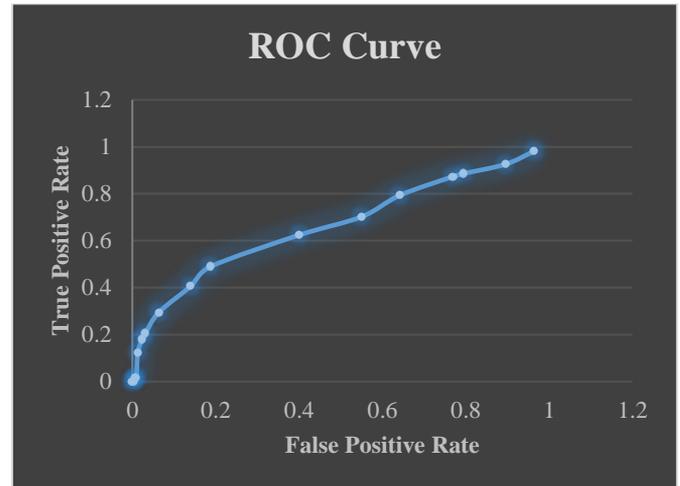
Figure 10. ROC curve for Opportunity++ dataset

Table 4. Confusion matrix for fine-grained locomotion ADL prediction accuracy over Opportunity++ dataset

Opportunity++ Dataset	OD1	OD2	CD1	CD2	OF	CF	ODW	CDW	ODR1	CDR1	ODR2	CDR2	ODR3	CDR3	CT	DC	TS
OD1*	**9**	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
OD2	0	**8**	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
CD1	1	0	**9**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CD2	0	0	0	**7**	1	1	1	0	0	0	0	0	0	0	0	0	0
OF	0	0	0	1	**8**	0	0	1	0	0	0	0	0	0	0	0	0
CF	0	0	0	1	0	**7**	1	1	0	0	0	0	0	0	0	0	0
ODW	0	0	0	0	0	1	**8**	1	0	0	0	0	0	0	0	0	0
CDW	0	0	0	0	0	1	1	**8**	0	0	0	0	0	0	0	0	0
ODR1	0	0	0	0	0	0	0	0	**9**	0	0	0	0	0	0	1	0
CDR1	0	1	0	0	0	0	0	0	0	**9**	0	0	0	0	0	0	0
ODR2	0	0	0	0	0	0	1	1	0	0	**8**	0	0	0	0	0	0
CDR2	0	0	0	0	0	0	0	0	0	1	1	**7**	0	0	0	1	0
ODR3	0	1	0	0	0	0	0	0	0	0	0	0	**9**	0	0	0	0
CDR3	0	0	0	0	0	0	0	0	0	0	0	0	0	**10**	0	0	0
CT	1	0	0	0	0	0	0	0	0	0	0	0	1	1	**7**	0	0
DC	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**9**	0
TS	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	**9**
Mean Accuracy = 82.94%																	

*OD1=Open Door 1, OD2=Open Door 2, CD1=Close Door 1, CD2=Close Door 2, OF=Open Fridge, CF=Close Fridge, ODW=Open Dishwasher, CDW=Close Dishwasher, ODR1=Open Drawer 1, CDR1=Close Drawer 1, ODR2=Open Drawer 2, CDR2=Close Drawer 2, ODR3=Open Drawer 3, CDR3=Close Drawer 3, CT=Clean Table, DC=Drink from Cup, TS = Toggle Switch.

Table 5. Precision, recall, and F1-score for fine-grained locomotion ADL prediction over Opportunity++ dataset

Performance Measures	OD1	OD2	CD1	CD2	OF	CF	ODW	CDW	ODR1	CDR1	ODR2	CDR2	ODR3	CDR3	CT	DC	TS	Mean
Precision	0.80	0.80	0.90	0.70	0.80	0.70	0.80	0.80	0.90	0.90	0.80	0.70	0.90	1.00	0.70	0.90	0.90	**0.82**
Recall	0.80	0.80	0.82	0.70	0.89	0.70	0.67	0.67	1.00	0.90	0.89	0.87	0.82	0.91	0.87	0.82	1.00	**0.83**
F1-Score	0.80	0.80	0.86	0.70	0.84	0.70	0.73	0.73	0.95	0.90	0.84	0.77	0.86	0.95	0.77	0.86	0.95	**0.82**

A novel parts-based model has been proposed in the study [70] to recognize interaction between people and fully convolutional network has also been utilized to classify them. These compared systems couldn't perform very well due to no

filtration technique applied to video sequences and no hand-crafted features being extracted. Therefore, these models [63-70] did not perform very well and hence achieved less accuracy rates. Our implemented system has attained better

prediction in ADL recognition due to the introduced ADL recognition system of background subtraction, human detection, features extraction, optimization, and classification.

Table 6. Comparison of locomotion prediction accuracies with other state-of-the-art models

Models	Average Accuracy (%)
Gabrielli et al. [63]	64.55
Donaj and Maučec [64]	70.95
Javeed et al. [65]	74.26
Azmat et al. [66]	74.70
Akhter et al. [67]	74.70
Sridharan et al. [68]	76.67
Myagmar et al. [69]	78.90
Ghadi et al. [70]	82.55
Implemented System	**82.94**

4.2.3 Experiment 3: Locomotion ADL detection comparison with other neural networks

The proposed system has been compared with other well-known neural networks in this experiment. Table 7 compares the CNN based precision (prec) and recall (rec) results to artificial neural network (ANN) and Adaboost algorithms. It is evident from the comparison that CNN performed much better than other compared classifiers as CNN achieved precision of 82.0% and recall of 83.0%.

Table 7. Comparison of locomotion prediction with other neural networks

Opportunity++Dataset	ANN		Adaboost		CNN	
	Prec	Rec	Prec	Rec	Prec	Rec
OD1	0.72	0.70	0.77	0.74	**0.80**	**0.80**
OD2	0.78	0.78	0.68	0.66	**0.80**	**0.80**
CD1	0.80	0.82	0.59	0.60	**0.90**	**0.82**
CD2	0.67	0.66	0.74	0.71	**0.70**	**0.70**
OF	0.89	0.87	0.85	0.89	**0.80**	**0.89**
CF	0.91	0.90	0.77	0.75	**0.70**	**0.70**
ODW	0.77	0.66	0.71	0.72	**0.80**	**0.67**
CDW	0.67	0.70	0.67	0.67	**0.80**	**0.67**
ODR1	0.78	0.72	0.69	0.74	**0.90**	**1.00**
CDR1	0.72	0.70	0.70	0.73	**0.90**	**0.90**
ODR2	0.65	0.61	0.84	0.80	**0.80**	**0.89**
CDR2	0.69	0.67	0.66	0.64	**0.70**	**0.87**
ODR3	0.73	0.76	0.76	0.78	**0.90**	**0.82**
CDR3	0.80	0.79	0.80	0.77	**1.00**	**0.91**
CT	0.70	0.67	0.74	0.76	**0.70**	**0.87**
DC	0.67	0.61	0.90	0.90	**0.90**	**0.82**
TS	0.89	0.80	0.80	0.78	**0.90**	**1.00**
Mean	**0.75**	**0.73**	**0.74**	**0.74**	**0.82**	**0.83**

5. DISCUSSION

There are certain challenges associated with the extraction of skeleton model from an image. The skeleton points may not be visible in few sample images that can cause limitations to model skeleton. There is also a possibility of mix up for two or more skeleton points when it comes to variable number of subjects in the same frame and different lighting conditions. Figure 11 presents such limitations for Opportunity++ dataset.

Other potential limitations of this system are related to the environment site. The selected dataset is based on indoor environment and thus we can say that the system's implementation is limited to indoor settings, such as hospitals, smart homes, industrial indoor areas etc. This could affect the

locomotion prediction in general environment. The visual data alone is not enough to predict human locomotion in general. We would require other sensors including motion and ambient sensors for the application of this system in outdoor environments. There can be another issue in the implementation of this method for real-world patient monitoring systems due to its implementation complexities including time complexity and cost factor.

Another possible limitation is the background subtraction using a background frame because it can lead to inaccurate detection of skeleton points when the background is dynamic. Therefore, we can apply this system for videos with static backgrounds.

Figure 11. Examples of limitations related to ADL over Opportunity++ dataset

6. CONCLUSION

This paper has implemented an intelligent daily living locomotion prediction system that will be supporting human daily routine activities. It consists of data acquisition from state-of-the-art Opportunity++ dataset in the form of videos, which are further pre-processed using the Weiner filter and the background has been subtracted through detecting moveable body parts. Then, a human skeleton model has been extracted via five body points, which are further utilized to extract the stochastic features such as the energy heat map and saliency map. Moreover, the large feature vectors are reduced using the quadratic discriminant analysis technique followed by the classification step through CNN has supported the implemented system in achieving an 82.94% mean accuracy rate for fine-grained ADL recognition for locomotion prediction over the Opportunity++ dataset.

The system requires more related features extraction and a combination of multiple types of sensors. Hence, in the future, we will improve the system by applying robust algorithms for filtration and optimization along with multi-sensory devices. The system has used only one dataset named Opportunity++ that is limited to indoor data. There is a need to experiment the proposed system over outdoor data as well. We will also be focusing on the presented dataset limitation in upcoming studies.

REFERENCES

[1] Figueiredo, J., Carvalho, S.P., Goncalve, D., Moreno, J.C., Santos, C.P. (2020). Daily locomotion recognition and prediction: A kinematic data-based machine learning approach. IEEE Access, 8: 33250-33262. https://doi.org/10.1109/ACCESS.2020.2971552

[2] Akhter, I., Jalal, A., Kim, K. (2021). Adaptive pose estimation for gait event detection using context-aware model and hierarchical optimization. Journal of Electrical Engineering & Technology, 16: 2721-2729. https://doi.org/10.1007/s42835-021-00756-y

[3] Javeed, M., Gochoo, M., Jalal, A., Kim, K. (2021). HF-SPHR: Hybrid features for sustainable physical healthcare pattern recognition using deep belief networks. Sustainability, 13(4): 1699. https://doi.org/10.3390/su13041699

[4] Javeed, M., Jalal, A., Kim, K. (2021). Wearable sensors based exertion recognition using statistical features and random forest for physical healthcare monitoring. In 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, pp. 512-517. https://doi.org/10.1109/IBCAST51254.2021.9393014

[5] Jalal, A., Kim, Y., Kim, D. (2014). Ridge body parts features for human pose estimation and recognition from RGB-D video data. In Fifth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Hefei, China, pp. 1-6. https://doi.org/10.1109/ICCCNT.2014.6963015

[6] Javeed, M., Shorfuzzaman, M., Alsufyani, N., Chelloug, S.A., Jalal, A., Park, J. (2022). Physical human locomotion prediction using manifold regularization. PeerJ Computer Science, 8: e1105. https://doi.org/10.7717/peerj-cs.1105

[7] Javeed, M., Mudawi, N.A., Alabduallah, B.I., Jalal, A., Kim, W. (2023). A multimodal IoT-based locomotion classification system using features engineering and recursive neural network. Sensors, 23(10): 4716. https://doi.org/10.3390/s23104716

[8] Su, B., Liu, Y.X., Gutierrez-Farewik, E.M. (2021). Locomotion mode transition prediction based on gait-event identification using wearable sensors and multilayer perceptrons. Sensors, 21(22): 7473. https://doi.org/10.3390/s21227473

[9] Batool, M., Ghadi, Y.Y., Alsuhibany, S.A., Al Shloul, T., Jalal, A., Park, J. (2022). Self-Care assessment for daily living using machine learning mechanism. Computers, Materials & Continua, 72(1): 1747-1764. https://doi.org/10.32604/cmc.2022.025112

[10] Jalal, A., Kim, J.T., Kim, T.S. (2012). Development of a life logging system via depth imaging-based human activity recognition for smart homes. In Proceedings of the International Symposium on Sustainable Healthy Buildings, Seoul, Korea, 19: 91-96.

[11] Azmat, U., Ghadi, Y.Y., Shloul, T.A., Alsuhibany, S.A., Jalal, A., Park, J. (2022). Smartphone sensor-based human locomotion surveillance system using multilayer perceptron. Applied Sciences, 12(5): 2550. https://doi.org/10.3390/app12052550

[12] Ciliberto, M., Fortes Rey, V., Calatroni, A., Lukowicz, P., Roggen, D. (2021). Opportunity++: A multimodal dataset for video-and wearable, object and ambient sensors-based human activity recognition. Frontiers in Computer Science, 3: 792065. https://doi.org/10.3389/fcomp.2021.792065

[13] Jonic, S., Jankovic, T., Gajic, V., Popvic, D. (1999). Three machine learning techniques for automatic determination of rules to control locomotion. IEEE Transactions on Biomedical Engineering, 46(3): 300-310. https://doi.org/10.1109/10.748983

[14] Gochoo, M., Tahir, S.B.U.D., Jalal, A., Kim, K. (2021). Monitoring real-time personal locomotion behaviors over smart indoor-outdoor environments via body-worn sensors. IEEE Access, 9: 70556-70570. https://doi.org/10.1109/ACCESS.2021.3078513

[15] Azmat, U., Jalal, A. (2021). Smartphone inertial sensors for human locomotion activity recognition based on template matching and codebook generation. In 2021 International Conference on Communication Technologies (ComTech), Rawalpindi, Pakistan, pp. 109-114. https://doi.org/10.1109/ComTech52583.2021.9616681

[16] Zhou, H., Yang, D., Li, Z., Zhou, D., Gao, J., Guan, J. (2021). Locomotion mode recognition for walking on three terrains based on sEMG of lower limb and back muscles. Sensors, 21(9): 2933. https://doi.org/10.3390/s21092933

[17] Bejinariu, S.I., Luca, R., Onu, I., Petroiu, G., Costin, H. (2021). Image processing for the rehabilitation assessment of locomotion injuries and post stroke disabilities. In 2021 International Conference on E-Health and Bioengineering (EHB), Iasi, Romania, pp. 1-4. https://doi.org/10.1109/EHB52898.2021.9657714

[18] Pan, S., Berges, M., Rodakowski, J., Zhang, P., Noh, H.Y. (2020). Fine-grained activity of daily living (ADL) recognition through heterogeneous sensing systems with complementary spatiotemporal characteristics. Frontiers in Built Environment, 6: 560497. https://doi.org/10.3389/fbuil.2020.560497

[19] Yan, T., Sun, Y., Liu, T., Cheung, C.H., Meng, M.Q.H. (2018). A locomotion recognition system using depth images. In 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, pp. 6766-6772. https://doi.org/10.1109/ICRA.2018.8460514

[20] Moriya, K., Nakagawa, E., Fujimoto, M., Suwa, H., Arakawa, Y., Kimura, A., Miki, S., Yasumoto, K. (2017). Daily living activity recognition with ECHONET lite appliances and motion sensors. In 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, HI, USA, pp. 437-442. https://doi.org/10.1109/PERCOMW.2017.7917603

[21] Khalid, N., Gochoo, M., Jalal, A., Kim, K. (2021). Modeling two-person segmentation and locomotion for stereoscopic action identification: A sustainable video surveillance system. Sustainability, 13(2): 970. https://doi.org/10.3390/su13020970

[22] Zhang, X., Xu, Z., Liao, H. (2022). Human motion tracking and 3D motion track detection technology based on visual information features and machine learning. Neural Computing and Applications, 34(15): 12439-12451. https://doi.org/10.1007/s00521-021-06703-2

[23] Wang, L., Tan, T., Ning, H., Hu, W. (2003). Silhouette analysis-based gait recognition for human identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12): 1505-1518. https://doi.org/10.1109/TPAMI.2003.1251144

[24] Guo, C., Zuo, X., Wang, S., Liu, X., Zou, S., Gong, M., Cheng, L. (2022). Action2video: Generating videos of human 3d actions. International Journal of Computer Vision, 130(2): 285-315. https://doi.org/10.1007/s11263-021-01550-z

[25] Ghadi, Y.Y., Akhter, I., Aljuaid, H., Gochoo, M., Alsuhibany, S.A., Jalal, A., Park, J. (2022). Extrinsic behavior prediction of pedestrians via maximum entropy Markov model and graph-based features mining. Applied Sciences, 12(12): 5985. https://doi.org/10.3390/app12125985

[26] Chen, J., Benesty, J., Huang, Y., Doclo, S. (2006). New insights into the noise reduction Wiener filter. IEEE Transactions on Audio, Speech, and Language Processing, 14(4): 1218-1234. https://doi.org/10.1109/TSA.2005.860851

[27] Dogariu, L.M., Ciochină, S., Paleologu, C., Benesty, J., Oprea, C. (2020). An iterative Wiener filter for the identification of multilinear forms. In 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, pp. 193-197. https://doi.org/10.1109/TSP49548.2020.9163453

[28] Rafique, A.A., Jalal, A., Ahmed, A. (2019). Scene understanding and recognition: Statistical segmented model using geometrical features and Gaussian naïve bayes. In IEEE conference on International Conference on Applied and Engineering Mathematics, Taxila, Pakistan, Vol. 57. https://doi.org/10.1109/ICAEM.2019.8853721

[29] Jalal, A., Batool, M., ud din Tahir, S.B. (2021). Markerless sensors for physical health monitoring system using ECG and GMM feature extraction. In 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, pp. 340-345. https://doi.org/10.1109/IBCAST51254.2021.9393243

[30] Rafique, A.A., Jalal, A., Kim, K. (2020). Statistical multi-objects segmentation for indoor/outdoor scene detection and classification via depth images. In 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, pp. 271-276. https://doi.org/10.1109/IBCAST47879.2020.9044576

[31] Mahmood, M., Jalal, A., Sidduqi, M.A. (2018). Robust spatio-temporal features for human interaction recognition via artificial neural network. In 2018 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, pp. 218-223. https://doi.org/10.1109/FIT.2018.00045

[32] Ghadi, Y., Akhter, I., Alarfaj, M., Jalal, A., Kim, K. (2021). Syntactic model-based human body 3D reconstruction and event classification via association based features mining and deep learning. PeerJ Computer Science, 7: e764. https://doi.org/10.7717/peerj-cs.764

[33] Javeed, M., Abdelhaq, M., Algarni, A., Jalal, A. (2023). Biosensor-based multimodal deep human locomotion decoding via internet of healthcare things. Micromachines, 14(12): 2204. https://doi.org/10.3390/mi14122204

[34] Akhter, I., Jalal, A., Kim, K. (2021). Pose estimation and detection for event recognition using Sense-Aware features and Adaboost classifier. In 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, pp. 500-505. https://doi.org/10.1109/IBCAST51254.2021.9393293

[35] Khalid, N., Ghadi, Y.Y., Gochoo, M., Jalal, A., Kim, K. (2021). Semantic recognition of human-object interactions via Gaussian-based elliptical modeling and pixel-level labeling. IEEE Access, 9: 111249-111266. https://doi.org/10.1109/ACCESS.2021.3101716

[36] Wibowo, G.H., Sigit, R., Barakbah, A. (2016). Feature extraction of character image using shape energy. In 2016 International Electronics Symposium (IES), Denpasar, Indonesia, pp. 471-475. https://doi.org/10.1109/ELECSYM.2016.7861052

[37] Gochoo, M., Akhter, I., Jalal, A., Kim, K. (2021). Stochastic remote sensing event classification over adaptive posture estimation via multifused data and deep belief network. Remote Sensing, 13(5): 912. https://doi.org/10.3390/rs13050912

[38] Ghadi, Y.Y., Khalid, N., Alsuhibany, S.A., Al Shloul, T., Jalal, A., Park, J. (2022). An intelligent healthcare monitoring framework for daily assistant living. Computers, Materials & Continua, 72(2): 2597-2615. https://doi.org/10.32604/cmc.2022.024422

[39] Pervaiz, M., Jalal, A., Kim, K. (2021). Hybrid algorithm for multi people counting and tracking for smart surveillance. In 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, pp. 530-535. https://doi.org/10.1109/IBCAST51254.2021.9393171

[40] Waheed, M., Jalal, A., Alarfaj, M., Ghadi, Y.Y., Al Shloul, T., Kamal, S., Kim, D.S. (2021). An LSTM-based approach for understanding human interactions using hybrid feature descriptors over depth sensors. IEEE Access, 9: 167434-167446. https://doi.org/10.1109/ACCESS.2021.3130613

[41] Damaševičius, R., Maskeliūnas, R., Woźniak, M., Polap, D., Sidekerskienė, T., Gabryel, M. (2017). Detection of saliency map as image feature outliers using random projections based method. In 2017 13th International Computer Engineering Conference (ICENCO), Cairo, Egypt, pp. 85-90. https://doi.org/10.1109/ICENCO.2017.8289768

[42] Kanan, C., Cottrell, G. (2010). Robust classification of objects, faces, and flowers using natural image statistics. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, pp. 2472-2479. https://doi.org/10.1109/CVPR.2010.5539947

[43] Elkhalil, K., Kammoun, A., Couillet, R., Al-Naffouri, T.Y., Alouini, M.S. (2017). Asymptotic performance of regularized quadratic discriminant analysis based classifiers. In 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), Tokyo, Japan, pp. 1-6. https://doi.org/10.1109/MLSP.2017.8168172

[44] Qin, A., Hu, Q., Lv, Y., Zhang, Q. (2018). Concurrent fault diagnosis based on Bayesian discriminating analysis and time series analysis with dimensionless parameters. IEEE Sensors Journal, 19(6): 2254-2265. https://doi.org/10.1109/JSEN.2018.2885377

[45] Quaid, M.A.K., Jalal, A. (2020). Wearable sensors based human behavioral pattern recognition using statistical features and reweighted genetic algorithm. Multimedia Tools and Applications, 79(9): 6061-6083. https://doi.org/10.1007/s11042-019-08463-7

[46] Ghadi, Y.Y., Javeed, M., Alarfaj, M., Al Shloul, T., Alsuhibany, S.A., Jalal, A., Kamal, S., Kim, D.S. (2022). MS-DLD: Multi-sensors based daily locomotion detection via kinematic-static energy and body-specific HMMs. IEEE Access, 10: 23964-23979. https://doi.org/10.1109/ACCESS.2022.3154775

[47] Ye, Q., Huang, P., Zhang, Z., Zheng, Y., Fu, L., Yang, W. (2021). Multiview learning with robust double-sided twin SVM. IEEE Transactions on Cybernetics, 52(12): 12745-12758. https://doi.org/10.1109/TCYB.2021.3088519

[48] Jalal, A., Batool, M., Kim, K. (2020). Stochastic recognition of physical activity and healthcare using tri-axial inertial wearable sensors. Applied Sciences, 10(20): 7122. https://doi.org/10.3390/app10207122

[49] Al Shloul, T., Javeed, M., Gochoo, M., Alsuhibany, S.A., Ghadi, Y.Y., Jalal, A., Park, J. (2023). Student's health exercise recognition tool for E-learning education. Intelligent Automation & Soft Computing, 35(1): 149-161. http://doi.org/10.32604/iasc.2023.026051

[50] Al-Qatf, M., Lasheng, Y., Al-Habib, M., Al-Sabahi, K. (2018). Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access, 6: 52843-52856. https://doi.org/10.1109/ACCESS.2018.2869577

[51] Batool, M., Jalal, A., Kim, K. (2019). Sensors technologies for human activity analysis based on SVM optimized by PSO algorithm. In 2019 International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, pp. 145-150. https://doi.org/10.1109/ICAEM.2019.8853770

[52] Shuvo, M.M.H., Ahmed, N., Nouduri, K., Palaniappan, K. (2020). A hybrid approach for human activity recognition with support vector machine and 1D convolutional neural network. In 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington DC, DC, USA, pp. 1-5. https://doi.org/10.1109/AIPR50011.2020.9425332

[53] Lukowicz, P., Pirkl, G., Bannach, D., Wagner, F., Calatroni, A., Förster, K., Holleczek, T., Rossi, M., Roggen, D., Troester, G., Doppler, J., Holzmann, C., Riener, A., Ferscha, A., Chavarriaga, R. (2010). Recording a complex, multi modal activity data set for context recognition. In 23th International Conference on Architecture of Computing Systems 2010, Hannover, Germany, VDE. pp. 1-6.

[54] Jalal, A., Quaid, M.A.K., Tahir, S.B.U.D., Kim, K. (2020). A study of accelerometer and gyroscope measurements in physical life-log activities detection systems. Sensors, 20(22): 6670. https://doi.org/10.3390/s20226670

[55] Sagha, H., Digumarti, S.T., Millán, J.D.R., Chavarriaga, R., Calatroni, A., Roggen, D., Tröster, G. (2011). Benchmarking classification techniques using the Opportunity human activity dataset. In 2011 IEEE International Conference on Systems, Man, and Cybernetics, Anchorage, AK, USA, pp. 36-40. https://doi.org/10.1109/ICSMC.2011.6083628.

[56] Roggen, D., Calatroni, A., Rossi, M., Holleczek, T., Förster, K., Tröster, G., Lukowicz, P., Bannach, D., Pirkl, G., Ferscha, A., Doppler, J., Holzmann, C., Kurz, M., Holl, G., Chavarriaga, R., Sagha, H., Bayati, H., Creatura, M., Millàn, J.D.R. (2010). Collecting complex activity datasets in highly rich networked sensor environments. In 2010 Seventh International Conference on Networked Sensing Systems (INSS), Kassel, Germany, pp. 233-240. https://doi.org/10.1109/INSS.2010.5573462

[57] Javeed, M., Jalal, A. (2021). Body-worn hybrid-sensors based motion patterns detection via bag-of-features and fuzzy logic optimization. In 2021 International Conference on Innovative Computing (ICIC), Lahore, Pakistan, pp. 1-7. https://doi.org/10.1109/ICIC53490.2021.9692924

[58] Mahmood, M., Jalal, A., Kim, K. (2020). WHITE STAG model: Wise human interaction tracking and estimation (WHITE) using spatio-temporal and angular-geometric (STAG) descriptors. Multimedia Tools and Applications, 79(11): 6919-6950. https://doi.org/10.1007/s11042-019-08527-8

[59] Jalal, A., Mahmood, M. (2019). Students' behavior mining in e-learning environment using cognitive processes with information technologies. Education and Information Technologies, 24: 2797-2821. https://doi.org/10.1007/s10639-019-09892-5

[60] Jalal, A., Nadeem, A., Bobasu, S. (2019). Human body parts estimation and detection for physical sports movements. In 2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), Islamabad, Pakistan, pp. 104-109. https://doi.org/10.1109/C-CODE.2019.8680993

[61] Nadeem, A., Jalal, A., Kim, K. (2021). Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy Markov model. Multimedia Tools and Applications, 80: 21465-21498. https://doi.org/10.1007/s11042-021-10687-5

[62] ud din Tahir, S.B., Jalal, A., Batool, M. (2020). Wearable sensors for activity analysis using SMO-based random forest over smart home and sports datasets. In 2020 3rd International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, pp. 1-6. https://doi.org/10.1109/ICACS47775.2020.9055944

[63] Gabrielli, M., Leo, P., Renzi, F., Bergamaschi, S. (2019). Action recognition to estimate Activities of Daily Living (ADL) of elderly people. In 2019 IEEE 23rd International Symposium on Consumer Technologies (ISCT), Ancona, Italy, pp. 261-264. https://doi.org/10.1109/ISCE.2019.8900995

[64] Donaj, G., Maučec, M.S. (2019). Extension of HMM-based ADL recognition with Markov chains of activities and activity transition cost. IEEE Access, 7: 130650-130662. https://doi.org/10.1109/ACCESS.2019.2937350

[65] Javeed, M., Al Mudawi, N., Alazeb, A., Alotaibi, S.S., Almujally, N.A., Jalal, A. (2023). Deep ontology-based human locomotor activity recognition system via multisensory devices. IEEE Access, 11: 105466-105478. https://doi.org/10.1109/ACCESS.2023.3317893

[66] Azmat, U., Jalal, A., Javeed, M. (2023). Multi-sensors fused IoT-based home surveillance via Bag of visual and motion features. In 2023 International Conference on Communication, Computing and Digital Systems (C-CODE), Islamabad, Pakistan, pp. 1-6.

https://doi.org/10.1109/C-CODE58145.2023.10139889

[67] Akhter, I., Javeed, M., Jalal, A. (2023). Deep skeleton modeling and hybrid hand-crafted cues over physical exercises. In 2023 International Conference on Communication, Computing and Digital Systems (C-CODE), Islamabad, Pakistan, pp. 1-6. https://doi.org/10.1109/C-CODE58145.2023.10139863

[68] Sridharan, M., Bigham, J., Campbell, P.M., Phillips, C., Bodanese, E. (2019). Inferring micro-activities using wearable sensing for ADL recognition of home-care patients. IEEE Journal of Biomedical and Health Informatics, 24(3): 747-759. https://doi.org/10.1109/JBHI.2019.2918718

[69] Myagmar, B., Li, J., Kimura, S. (2018). A novel supervised heterogenuos feature transfer learning scheme for ADL recognition. In 2018 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Ishigaki, Japan, pp. 135-140. https://doi.org/10.1109/ISPACS.2018.8923251

[70] Ghadi, Y.Y., Waheed, M., Al Shloul, T., Alsuhibany, S.A., Jalal, A., Park, J. (2022). Automated parts-based model for recognizing human-object interactions from aerial imagery with fully convolutional network. Remote Sensing, 14(6): 1492. https://doi.org/10.3390/rs14061492