

Improving the Ambient Intelligence Living Using Deep Learning Classifier

Yazeed Yasin Ghadi¹, Mouazma Batool², Munkhjargal Gochoo³, Suliman A. Alsubibany⁴,
Tamara al Shloul⁵, Ahmad Jalal² and Jeongmin Park^{6,*}

¹Department of Computer Science and Software Engineering, Al Ain University, Al Ain, 15551, UAE

²Department of Computer Science, Air University, Islamabad, 44000, Pakistan

³Department of Computer Science and Software Engineering, United Arab Emirates University, Al Ain, 15551, UAE

⁴Department of Computer Science, College of Computer, Qassim University, Buraydah, 51452, Saudi Arabia

⁵Department of Humanities and Social Science, Al Ain University, Al Ain, 15551, UAE

⁶Department of Computer Engineering, Tech University of Korea, Gyeonggi-do, 15073, Korea

*Corresponding Author: Jeongmin Park. Email: jmpark@tukorea.ac.kr

Received: 17 January 2022; Accepted: 30 March 2022

Abstract: Over the last decade, there is a surge of attention in establishing ambient assisted living (AAL) solutions to assist individuals live independently. With a social and economic perspective, the demographic shift toward an elderly population has brought new challenges to today's society. AAL can offer a variety of solutions for increasing people's quality of life, allowing them to live healthier and more independently for longer. In this paper, we have proposed a novel AAL solution using a hybrid bidirectional long-term and short-term memory networks (BiLSTM) and convolutional neural network (CNN) classifier. We first pre-processed the signal data, then used time-frequency features such as signal energy, signal variance, signal frequency, empirical mode, and empirical mode decomposition. The convolutional neural network-bidirectional long-term and short-term memory (CNN-biLSTM) classifier with dimensional reduction isomap algorithm was then used to select ideal features. We assessed the performance of our proposed system on the publicly accessible human gait database (HuGaDB) benchmark dataset and achieved an accuracy rates of 93.95 percent, respectively. Experiments reveal that hybrid method gives more accuracy than single classifier in AAL model. The suggested system can assist persons with impairments, assisting carers and medical personnel.

Keywords: Ambient assisted living; convolutional neural network; dimensionality reduction; frequency-time features; wearable technology

1 Introduction

Ambient assisted living (AAL) is an emerging technology for solving the challenges of elderly people. AAL helps people to stay connected socially and live autonomously into old age by integrating smart technology like sensors, smart interfaces, actuators, and artificial intelligence [1]. AAL has



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

evolved rapidly in daily-life activities like providing assistive technologies to disabled people and providing easiness in acceptability, accessibility, and usability of advanced technologies [2]. Moreover, emerging AAL technology has decision-making capability to anticipate and respond intelligently to the upcoming needs of elder people [3].

As elder people have many age-related diseases, and therefore the need of health assistance systems is increasing annually. The most generic method of monitoring these people is physical observation which is expensive and impractical in real life. AAL technologies like video surveillance systems, care-providing robots, and human-computer interaction recognition systems are generally targeting elderly people but these technologies can also assist physically and mentally impaired people [4]. Moreover, people suffering from obesity, diabetes, and people conscious of their fitness can also take advantage of such AAL systems. Hence, the AAL-based real-time monitoring is affecting people of every age ranging from people with disabilities to people linked with sports activities, which is the major branch of the human activity recognition (HAR) domain [5].

The key reason of using wearable sensors, i.e., accelerometers and gyroscopes in AAL systems is the ease of recording a person's daily activities [6]. Continuous activity records help the caretaker keep a check on the physical and mental condition of the patient, resulting in improving the person's mental and physical condition. The continuous feedback of the patient's activities makes it easier for the caretaker to diagnose the medical condition of the patient [7]. Moreover, AAL is also providing a system for independent living to elderly people and patients with other mental pathologies. Furthermore, HAR systems can be designed to interact with users to change their behaviors and lifestyles towards more active and healthier ones. Recently, numerous intelligent systems based on wearable sensors have been designed for diagnosis of medical conditions, like the parkinson disease, cardiac rehabilitation, physical therapy, detecting abnormal activities, physical activity of adolescents, detecting fall of elderly people, and analyzing sleep patterns of children, and adolescents [8].

This paper focuses on the classification of activities recorded by the movements and positioning of inertial sensors. The proposed system includes five main steps: First, signal acquisition and preprocessing is done using Chebyshev, Kalman, and dynamic data reconciliation (DDR) filters. Second, signal normalization is applied to the filtered signal. Third, time-frequency features are extracted from the signal. In addition to that, dimensionality reduction is done using projecting matrix and convolutional neural network-bidirectional long-term and short-term memory (CNN-biLSTM) classifier. For AAL environment system, the HuGaDB benchmark dataset composed of diverse patterns of activities has been used to assess the proposed model. The main contributions of this paper are as follows:

- We proposed multifeature extraction algorithms for a variety of signal patterns that include both time domain and frequency domain features.
- We used the Isomap algorithm to optimize the data for complicated human activity patterns, which offered contextual information as well as classifying behaviours.
- In addition, the CNN-biLSTM classifier was utilized to classify the HuGaDB public benchmark dataset, and the results were significantly better than other state-of-the-art approaches.

The rest of this paper has been structured as follows: Section 2 gives the detailed description of related work in the field of time-frequency and deep learning model. Section 3 covers the details of the proposed AAL architecture. Section 4 consists of experimental results. Finally, Section 5 comprises of the conclusion and future perspectives of our proposed model.

2 Related Work

The machine learning validation technologies for AAL system, established so-far could be broadly classified into two types as described below.

2.1 Activity Recognition Using Time-Frequency Features

Many researchers have extensively worked on the well-known features of human activity recognition systems in assorted domains. Shi et al. [9] employed features corresponding to time-frequency domain for brain activity recognition using the wavelet packet decomposition (WPD) tool. The WPD works by decomposing frequencies of low and high components within the signal to get the specific band of particular components. They have achieved an accuracy of 90.89%. The approach offers a high time-frequency resolution and noise suppression capability, but it requires a lot of computation, so it can only be used for small samples. Wang et al. [10] proposed an action detection method for an artificial knee system. They gathered data from inertial sensors and then fractional Fourier transform (FRFT) is applied to the data and then eight time-frequency features were extracted. In this paper, the standard deviation, interquartile range, variance, mean, and peaks were selected to form the feature vector and gain good recognition accuracy in this regard. The real-time performance of HAR system might be affected when a large amount of data is processed at once. Rosati et al. [11] performed feature extraction in time-domain for human activity recognition (HAR) classification. Later on, the resultant features are further reduced using genetic algorithm. Support vector machine (SVM) technology is used for accurate classification of data. It is easy to implement genetic algorithm with SVM classifier. However, because of the uncertainty principle's restriction, the approach is unable to achieve ideal temporal and frequency resolution at the same time. Debs et al. [12] linearly encoded acceleration signal in the frequency and time domain by applying a linear transformation on the inertial data. Acceleration signals are then anonymized by filtering them in time-frequency domain. Finally, they have obtained a recognition rate of 85%. The particular approach has poor noise suppression and time-frequency resolution. Jourdan et al. [13] designed a novel mechanism of preprocessing the raw signal with median and butterworth filter on a mobile phone. The preprocessed data is then sent to a server for feature extraction and activity classification. The frequency domain features include discrete Fourier transform (DFT), energy, and entropy. While, time domain features include mean, standard deviation, signal magnitude, and signal-pair correlation. The random forest classifier efficiently classified the data with a mean accuracy of 87%. Leonardis et al. [14] split the data into 5 s frames and extracted 342 features in the time, frequency, and time-frequency domains by using a correlation-based and genetic algorithm feature selection process. The author claim that the proposed technique successfully identified 90% of the actions. Chinimilli et al. [15] proposed a real-time human locomotion detection system using amplitude and omega features on six periodic activities, including level walking, stair ascent, stair descent, uphill, downhill, jogging, and running. Two successive peaks in the thigh angle signal has been further fed to k-nearest neighbors for the recognition of activities in an outdoor environment.

2.2 Activity Recognition Using CNN and LSTM

Deep learning techniques have been widely deployed in extensive research areas to automatically detect/classify data samples. Mohammadian et al. [16] designed a convolutional autoencoder to denoise the signal for abnormal human movements of accelerometer signals. The authors mentioned that they will implement the mobile application in real-time in the future. Nukala et al. [17] presented custom-designed approach for a gait analysis system using an artificial neural network. This system claims to efficiently classify falls and daily activities. The triaxial accelerometer and gyroscope sensors

are placed on human body at T4 and belt position. The implemented model is then implemented by transmitting data in between wireless gait analysis sensor (WGAS) and personal computer (PC). The model achieved a reliable accuracy for assisted living environments. Kwon et al. [18] performed gesture pattern classification using convolution neural network (CNN). The proposed model is tested in an indoor environment. The model efficiently classified the 10 gestures with a recognition accuracy of 90%. However, these systems has been tested only in an indoor activities and with constraint environment. Hence, these methods are not suitable enough to incorporate influence of the sensor errors in an unrestricted environment.

With the use of a neural network to classify complex human behaviors, many researchers have adopted long short-term memory (LSTM) to model their human activity recognition systems. These systems are usually based on unidirectional or bidirectional LSTMs, which utilize the past and future information to predict the current information for the recognition model. Jian et al. [19] anticipated hand gesture recognition system using LSTM based dynamic probability on the arabic numerals gesture (ANG) dataset. The data is firstly segmented into sub periods and then LSTM based classification algorithm is applied to each sub period to recognize hand gestures in a real-time environment. The non-maximum suppression is applied on the classified sub period to further eliminate any invalid substance from the resultant data. The proposed model has achieved substantial accuracy over the ANG dataset. Cui et al. [20] utilized unidirectional and bidirectional network architectures for the prediction of traffic state. The model verification is done on two real-world datasets. Mekruksavanich et al. [21] presented a smart home-based solution using LSTM in time domain series. Bayesian optimization is further applied on the resultant vector of LSTM to tune the hyperparameters of LSTM resultant vectors. The model is trained over UCI-HAR publicly available smartphone dataset and 10-fold cross-validation is applied for the validation of the proposed model. The main drawback of these systems are that only single LSTM classifier might not be sufficient enough to overcome noise suppression, to estimate continuous variables of activities, and to learn orientation of the sensor within a consecutive time frames.

Some researchers have also worked on hybrid approach: combination of neural networks and LSTM. Pienaar et al. [22] presented long short-term memory and recurrent neural network (LSTM-RNN) mechanism for human activity recognition (HAR). While Xia et al. [23] utilized long short-term memory and convolutional neural network (LSTM-CNN) based architecture for recognition of human activities. These systems have efficiently classify human activities within a controlled and uncontrolled environment. Hence, these systems inspired us to develop a novel hybrid approach for human activity recognition.

In general bidirectional networks, bidirectional long short-term memory (biLSTM) is quite better than unidirectional network LSTM, which preserves the information of the model from past and future. Moreover, deep neural network automatically extracts complex human behaviors from the raw datasets. In this study, we have used time-frequency feature extraction alongwith a hybrid approach of CNN and biLSTM to build a model which can recognize the complex AAL human behaviors in in a spontaneous way. The complete description of our proposed model is explained in Section 3.

3 Material and Methods

The model has been designed to evaluate the improvement in the ambient assisted living (AAL) system by using contextual information of inertial sensors. The aim was to classify complex activities within a real-time environment. In Section 2, we didn't find a system efficient and accurate enough for real-time decision making system. Hence, a system has been designed for a complex sequence of

actions, presented in the below subsections. The architecture of the proposed model has been elicited in Fig. 1. The proposed model has been divided into five phases: preprocessing, signal normalization, feature extraction, dimensionality reduction using projecting matrix, and a convolutional neural network-bidirectional long-term and short-term (CNN-biLSTM) classifier. Initially, the system takes 3-axis accelerometer and gyroscope values together as input to the system. The data has been taken from human gait database (HuGaDB) dataset that recorded the signal stream of six inertial sensors, attached on the left and right thighs, feet, and shins. This dataset is unique in the sense of its granularity, with actions composed of a single action performed for a long time and short movements of different activities performed for a few seconds. The inertial sensors data has been rectified by applying 3 different filters, i.e., lookahead filter, Kalman filter, and dynamic data reconciliation (DDR) filter. These signals has been further normalized to hinder any complex values in the feature normalization step. Finally, optimization and classification has been applied on the resultant data.

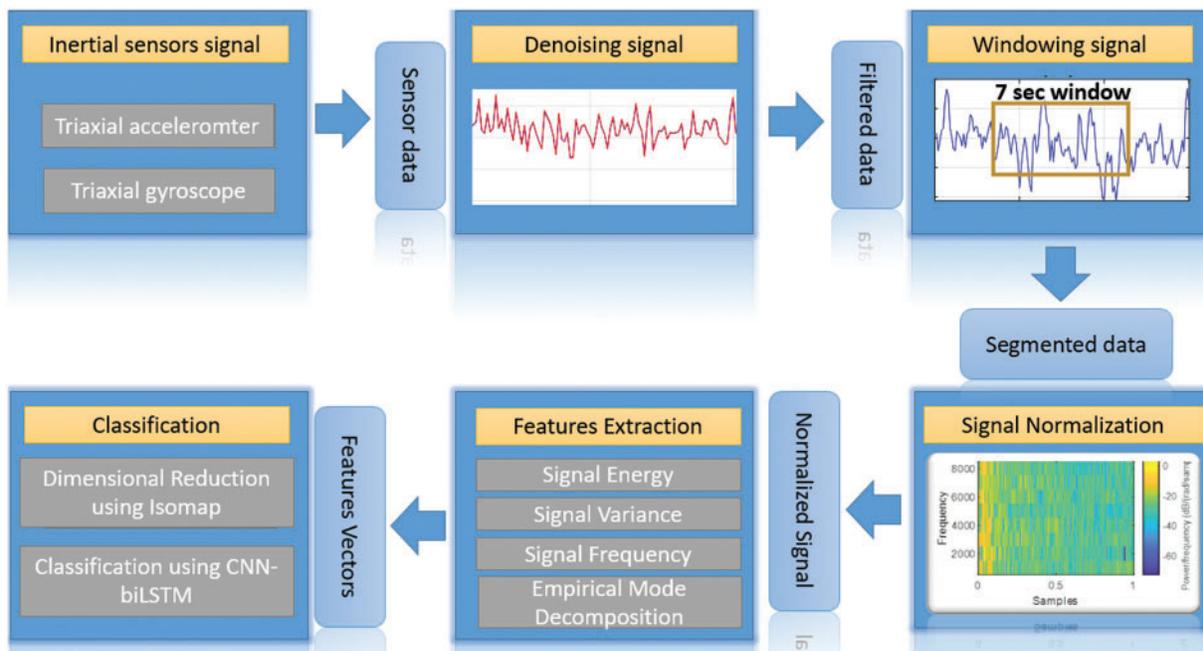


Figure 1: The classification framework of the proposed ambient assisted living (AAL) system

3.1 Signal Acquisition and Data Preprocessing

The inertial sensors data is highly sensitive to random noise which badly effects the feature extraction process [24]. Hence, three different filtration techniques, i.e., Chebyshev, Kalman, and dynamic data reconciliation (DDR) filter have been applied to eliminate the noise associated with the inertial data. The Chebyshev filter [25] smoothens the sensors data, eliminates power line interloping, and enhances the accuracy of inertial signal (See Fig. 2). The Kalman filter [26] is a finite, linear, and discrete time-varying filter that reduces the mean-square error (See Fig. 3). The DDR filter works on the data reconciliation principle which integrates the sample information corresponding to time instant t and time window H . The samples given to filter should be long enough to capture the dynamic behaviour of activities. In this system, one minute long data has been fed to the DDR [27] filter to get the desired results (See Fig. 4). The DDR filter, gave a better response than Chebyshev and Kalman

filter on human gait database (HuGaDB) dataset. Therefore, the DDR filter has been chosen for the next process.

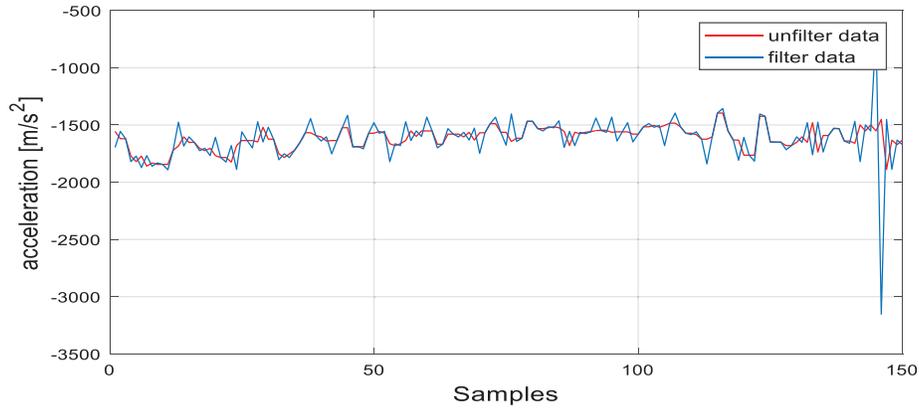


Figure 2: The filter response of chebyshev filter

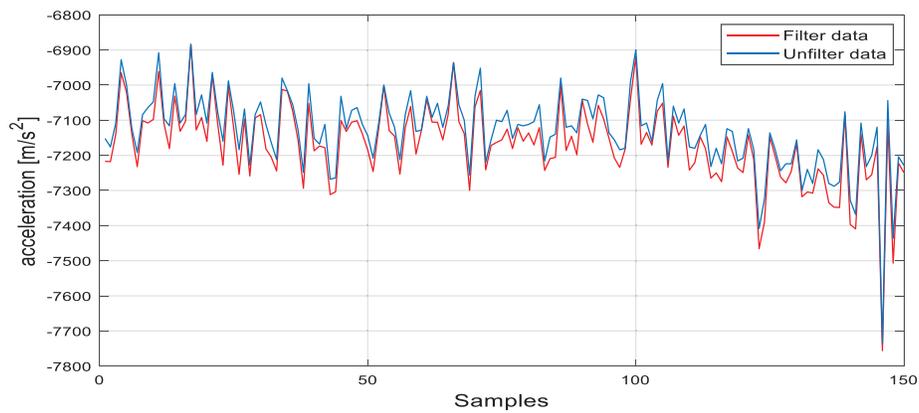


Figure 3: The filter response of kalman filter

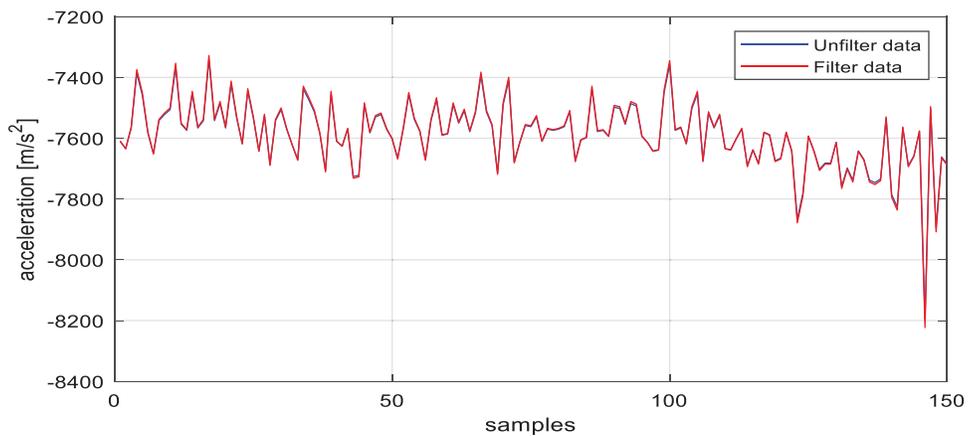


Figure 4: The filter response of dynamic data reconciliation (DDR) filter

3.2 Window Selection

The inertial signal cannot be fed to the system at once, therefore an appropriate window segment is needed to efficiently extract the contextual information of AAL activity behavior. The different window blocks have been tested by the researchers for the complete analysis of human behavior. Most researchers have used the block of 4-5 s for efficient analysis of human activity recognition [28]. The HuGaDB dataset contains some activities having repetitive motion patterns for a long duration of time like walking, sitting, standing, and bicycling. These activities can be accurately predicted due to high logical consistency with the previous frame but multiple activities performed during a short interval of time become more complex. These complex human motions need data to be long enough for accurate prediction of the subject's behavior. Out of many experiments, the window of 7 s kept balance in accurate prediction and recognition delay. As a result, 7 s window size was selected for maximum results. To maintain consistency in human motions, the contextual information of the previous frame is passed on to the next frame within a consecutive frame [29]. In our assessment 94% of the start and stop breakpoints of a given activity are accurately predicted.

3.3 Signal Normalization

The existence of sudden fluctuations in the signal impair the feature extraction and classification process and performance. There's a risk that variables with greater magnitude will be given more weight. As a result, before feeding the signal to the feature extraction process, it is critical to normalize the signal to ensure that time-domain features are not biased towards one feature. The normalization has been performed on the signal to rescale the data between 0 and 1. The signal normalization performs better and converges faster on a smaller scale [30]. The resultant rescaled coefficients has been scaled so they are less sensitive for the feature extraction process. The formula for normalization is given in Eq. (1)

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where the minimum value x_{min} has been subtracted from each entry x and the resultant value has been divided by the range. The range is the difference of minimum x_{min} and maximum x_{max} value. The min-max scaling has been done on the signal and is re-scaled between 0 and 1, as depicted in Fig. 5.

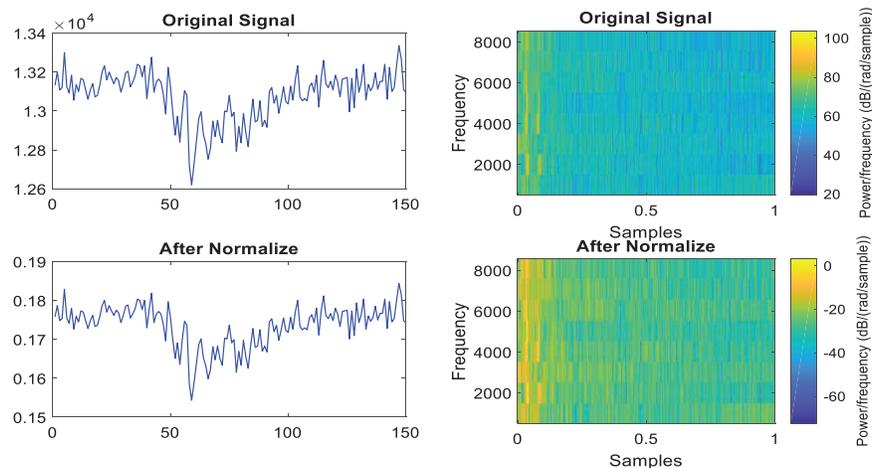


Figure 5: The scalable output of signal normalization

3.4 Feature Extraction

This is the most critical phase of a machine learning model, in which the meaningful attributes need to be represented for the accurate recognition of the performed activities. In the proposed system, time-frequency features simplify the analysis of inertial data. The stream of normalized data has been fed as input to this phase [31]. The detailed explanation of each feature has been described below.

3.4.1 Signal Energy

The energy of the signal has been obtained by squaring the normalized signal and summing the signal from initial window segment to the N number of sliding window signals [32]. The signal energy has been defined as Eq. (2).

$$S_e = \frac{1}{N} \sum_{n=0}^{N-1} S_{ig}[n]^2 \quad (2)$$

where, $S_{ig}[n]$, is the signal of the inertial sensor. S_e Represents the energy of the signal. The energy signal is a useful feature in detecting a transition within an activity motion and efficiently detecting a noticeable change in the energy of a signal obtained through the inertial sensor. The variation in the signal energy is illustrated in Fig. 6.

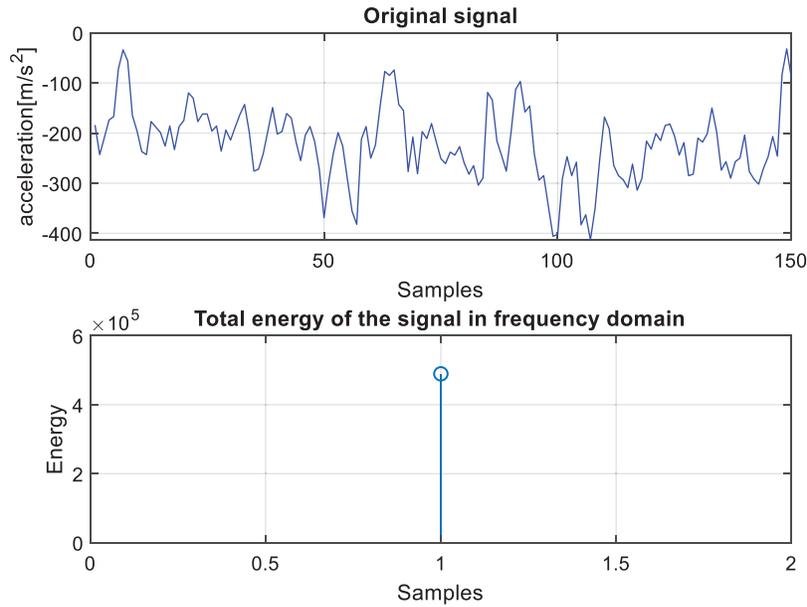


Figure 6: Inertial energy of the signal

3.4.2 Signal Variance

The signal variance is defined as the average of the squared difference [33]. The signal variance is defined as;

$$\sigma_s^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} \left(\|S_{ig}[n]\| - \frac{1}{N} \sum_{n=0}^{N-1} \|S_{ig}[n]\| \right) \quad (3)$$

where, $S_{ig}[n]$, is the inertial sensor's signal. σ_s^2 represents the variance of the signal. The signal variance is a useful feature to assist the discrimination between low and high-intensity motion.

3.4.3 Frequency Analysis

The frequency of a signal varies with different movements. The frequency analysis features help in capturing periodic movements related to the subject's activity [34]. The resultant peak in the signal gives useful information of the type of activity performed by the subject. The data of the inertial signal will remain constant if the person is standing still. The frequency analysis of the signal has been observed using fast Fourier transform (FFT) and reported in Fig. 7.

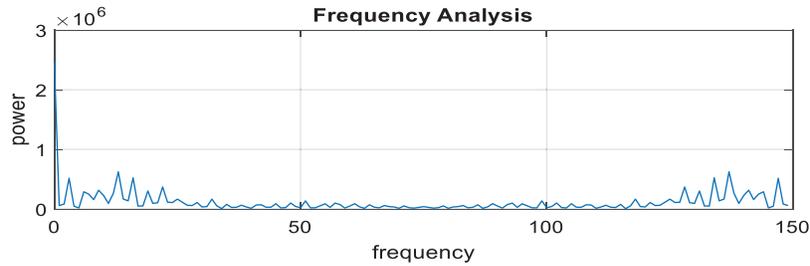


Figure 7: Frequency analysis of the signal

3.4.4 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) also known as intrinsic mode functions (IMFs) is commonly used for processing real-time signals from sensors which breaks every signal into many zero mean oscillatory modules [35]. Algorithm 1 represents the logical representation of EMD.

Algorithm 1: Empirical Mode Decomposition

Input: Inertial Sensors Signal ($S_{ig}[n]$)

Output: EMD_Vector_Array (E_v)

$E_v[n] \leftarrow []$

do

$E_{xtrema}[N] \leftarrow \text{calculate_extrema}(\text{GetLocalMaximaMinima}(S_{ig}[n]))$

$E_{nvelope}[N] \leftarrow \text{calculate_envelope}(E_{xtrema}[N])$

$A_{verage}[N] \leftarrow \text{calculate_average}(E_{nvelope}[N])$

$R_{esult}[N] \leftarrow S_{ig}[n] - A_{verage}[N]$

While ($R_{esult}[N] \neq \text{Stopping_Criteria}$)

$E_v[n] \leftarrow R_{esult}[N]$

return EMD_Vector_Array ($E_v[n]$)

3.5 Dimensionality Reduction Using Isomap

As a result of dimensionality reduction, the data consumes less training time and therefore the overall performance is increased. It efficiently transforms non-linear data to linearly separable form [36]. To further enhance the accuracy of classification in the proposed system, the dimensionality reduction method has been used to reduce the high dimensional feature matrix to low dimensional using Isomap. Isomap algorithm calculates the pairwise distances of the input feature vector. Next, a

k-nearest neighbors (KNN) classifier has been used to calculate the nearest neighbors of each feature's vector. Once the neighbors are found, the KNN builds a weighted graph where vectors are connected only if they are the nearest neighbors of each other [37]. Next, the Dijkstra algorithm has been used to calculate the minimal path between vector points in a graph called geodesic distance. Finally, multidimensional scaling (MDS) has been used to compute the vector points in low dimension such that the distance between the points has been completely preserved [38]. As shown in Fig. 8, the graph has achieved the spike at $x = 0.76$. This means that the values are improved at its maximum point of 0.76, indicating that we've achieved the global minimum and is thus achieved the maximum possible dimensional reduction.

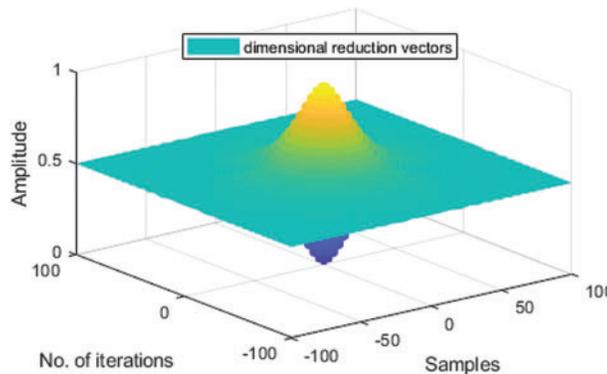


Figure 8: The results of Isomap, dimension reduction algorithm

3.6 CNN-biLSTM Classifier

In this paper, two deep neural networks bidirectional long-term and short-term memory networks (biLSTM) and convolutional neural network (CNN) have been merged to form a CNN-biLSTM classifier for the logical classification of extracted features [39]. The CNN model has three layers. The first layer of the CNN segments the inertial sensor's data by sliding the window of a signal. The next two layers effectively derived features from short and fixed lengths of the segments within the segments which are not highly relevant. The bidirectional long short-term memory (biLSTM) is similar to long short-term memory (LSTM) except for the recurrent block.

The one dimensional-convolutional neural network and long short-term memory (1D-CNN-biLSTM) learns the data in bidirectional layers [40]. First, it feeds the data in the forward direction, i.e., from beginning to end and then it feeds the data in the backward direction, i.e., from the end to beginning. Thus bidirectional LSTM acquires information from both past and future states as shown in Fig. 9.

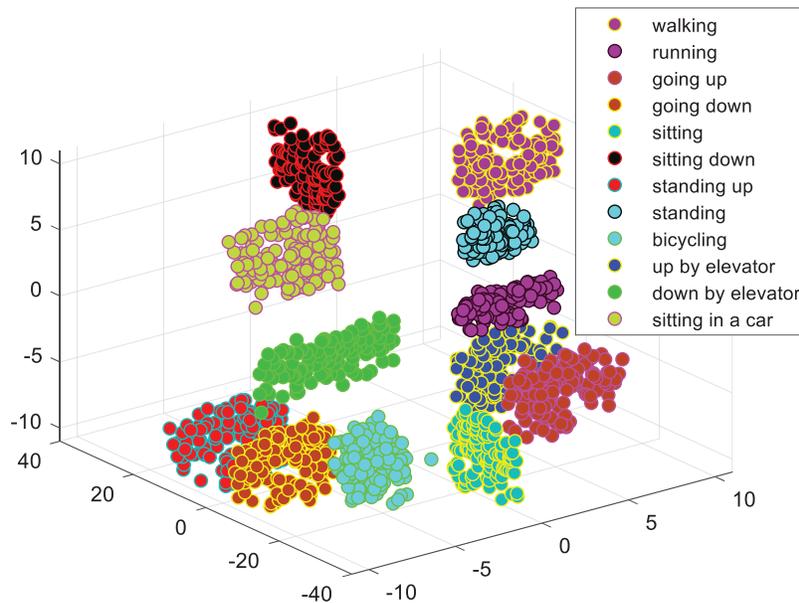


Figure 9: The logical classification of human gait database (HuGaDB) dataset via convolutional neural network-bidirectional long-term and short-term memory (CNN-biLSTM) classifier

4 Experimental Setting and Results

In this section, the experiments have been performed on a Desktop PC equipped with Intel(R) Process(R) CPU G3260, 8GB RAM, 3.30 GHz processing power, 64-bit operating system, and Windows 10, using Anaconda Spyder Python IDE. The leave-one-subject-out (LOSO) cross-validation method has been used to evaluate the performance of our proposed model on the human gait database (HuGaDB) benchmark dataset. The HuGaDB dataset has been already described in Section 3.1. Moreover, the proposed AAL system has been validated through precision, recall, and F-measure parameters. Finally, a comparison with other state-of-the-art methods has been done to prove that our model outperforms them in terms of accuracy.

4.1 The HuGaDB Datasets Description

The human gait database dataset (HuGaDB) [24] contains a continuous recording of data obtained from six inertial sensors. In total, 12 activities were performed by 18 participants which include running, going up and down, walking, sitting, standing up, sitting down, standing, bicycling, sitting in a car, up and down by elevator. The sensors were located on shins, feet, and left and right thighs. This dataset is unique in the sense that the participant performed each activity for a long time or a combination of activities for a short interval of time, and hence 10 h of total data is recorded.

4.2 Performance Parameters and Evaluations

The proposed methodology has been validated with recognition accuracy, precision, recall, and F1-score. Details of parameters for individual experiments have been deliberated as follows

4.2.1 Experiment I

The leave-one-subject-out (LOSO) cross-validation is an authentic model validation technique. It has been used to assess the performance of the proposed model on the HuGaDB dataset. The LOSO validation technique efficiently certified the performance of our AAL scheme with a mean accuracy of 93.95% as shown in [Tab. 1](#).

Table 1: Confusion matrix certified the accuracy of the human gait database (HuGaDB) dataset

Classes	SC	UE	DE	ST	SU	SI	SD	BC	GU	DS	RN	WK
SC	91.7	0	0	0	2.08	3.05	1.20	0	0	1.97	0	0
UE	0	92.09	4.60	0	0	0	0	0	0.91	2.4	0	0
DE	0	3.6	93.05	0	2.3	0	0	0	0	1.05	0	0
ST	0	0	0	92.70	2.2	2.1	3.0	0	0	0	0	0
SU	0	0	0	3.0	94.0	1.50	1.50	0	0	0	0	0
SI	3.05	0	0	0	2.6	94.35	0	0	0	0	0	0
SD	0	0	2.0	0	0	0	95.09	0	1.91	0	1.00	0
BC	0	0	0	1.54	0	0	0	93.06	0	0	2.4	3.0
GU	0	1.00	0	0.50	0	0.50	0	0	96.8	0	1.20	0
DS	0	0	0	0	2.00	0.90	0	3.00	0	94.02	0	0.08
RN	1.30	0	1.10	0	2.00	0	0	0	0	0	95.6	0
WK	0	1.00	0	0	0	0	0.31	0	2.00	1.00	0	95.0

Mean Accuracy = 93.95%

Notes: SC= sitting in a car; UE= up by elevator; DE=down by elevator; ST= standing; SU= standing up; SI= sitting; SD= sitting down; BC= bicycling; GU= going up; DS= down the stairs; RN= running; WK= walking. The diagonal values represent the prediction rate of positive classes.

Utilizing the confusion matrix, the going up (GU) activity has achieved the highest accuracy of 96.8%. While vectors of some activities are confused with each other like some vectors of sitting down and standing up are confused. Similarly, the activities of going up and down the stairs are also confused due to the repetition of similar movements. However, the overall accuracy of 93.95% proves the robustness of our model.

4.2.2 Experiment II

In this section, precision, recall, and F-measure parameters validate the performance of the proposed model as shown in [Tab. 2](#).

It is observed from [Tab. 2](#) that the bicycling, running, and walking activities achieved the maximum percentages of true positives while others have the least precision rates due to more false-positive results.

Table 2: Validation of proposed ambient assisted living (AAL) system with precision, recall and F1-measure over the benchmark dataset

Dataset	Interactions	Precision	Recall	F-measure
HuGaDB	Sitting in a car	0.92	0.94	0.93
	Up by elevator	0.93	0.94	0.94
	Down by elevator	0.92	0.91	0.91
	Standing	0.91	0.92	0.92
	Standing up	0.90	0.92	0.91
	Sitting	0.91	0.92	0.92
	Sitting down	0.90	0.93	0.93
	Bicycling	0.98	0.98	0.98
	Going up	0.93	0.94	0.94
	Down the stairs	0.89	0.91	0.90
	Running	0.98	0.98	0.98
Walking	0.98	0.98	0.98	
Average		92.9	93.9	93.6

4.2.3 Experiment III

In this section, comparison of the proposed model has been done against other state-of-the-art models. [Tab. 3](#) depicts the optimal performance of our proposed methodology over the HuGaDB dataset and [Tab. 4](#) depicts optimal performance of the proposed system with other state-of-the-art domains.

Table 3: The optimal accuracy of the proposed technique over state-of-the-art method

Dataset	Authors	Methodology	Recognition accuracy (%)
HuGaDB	Lahmiri et al. [41]	Complexity measures	85.91
	Gochoo et al. [42]	Adagrad	89.90
	Seiffert et al. [43]	Kinematic features with kernel sliding perceptron	91.76
	Gochoo et al. [42]	Stochastic gradient descent	92.50
	Proposed Methodology	Time-frequency features + CNN-biLSTM classifier	93.95

Table 4: The comparison of the proposed method with other state-of-the-art domains

Authors	Domain	Methodology	Recognition accuracy (%)
Jian et al. [19]	Long short-term memory	Long short-term memory-based dynamic probability (DP-LSTMs)	83.00
Debs et al. [12]	Time-frequency features	Linear transformation	85.00
Mohammadian et al. [16]	Neural network	Convolutional autoencoder	85.00
Jourdan et al. [13]	Time-frequency features	Discrete Fourier transform (DFT), energy, and entropy	87.00
Wang et al. [10]	Time-frequency features	Fractional Fourier transform (FRFT)	89.39
Kwon et al. [18]	Neural network	Convolution neural network	90.00
Shi et al. [9]	Time-frequency features	Wavelet packet decomposition	90.89
Xia et al. [23]	Hybrid approach	Long short-term memory and convolutional neural network LSTM-CNN	92.63
Proposed method	Hybrid approach	Time-frequency features + CNN-biLSTM classifier	93.95

5 Discussion

A total of 21,600 windows were processed (12 activities \times 18 participants \times 100 windows). The LOSO validation scheme was used to create the training and testing sets. The dynamic data reconciliation (DDR) filter captures the nonlinear response of activities in the first step. The signals were then split using a 7-second sliding window. Within a successive frame, the contextual information from the previous frame was passed on to the next frame. This option offers a nice balance of real-time data processing and immediate activity detection. The min-max signal scaling has been performed to the signal to dynamically adjust the data between 0 and 1 in order to offer fair weights and reduce rapid oscillations in the signal. Following that, the Isomap dimension reduction algorithm identified the optimal time-frequency features. The features has been reduced from 342 to 120 using the Isomap technique. Finally, the CNN-biLSTM classifier has been trained and validated using 120 features. The Section 4 reveal that hybrid method gives higher accuracy of 93.95% than state-of-the-art single classifier in AAL model.

6 Conclusion

Our proposed AAL recognition system attained high scores in Experiment II. Initially, the data of the triaxial accelerometer and gyroscope dataset is denoised with Chebyshev, Kalman, and dynamic data reconciliation (DDR) filters. The DDR filter has achieved a higher accuracy in the filtration process. The DDR filter output has been further processed with 7 s sliding windows that efficiently identify variation across 12 different activities. Furthermore, time-frequency features have been an addition to the performance of the model. Next, the obtained features have been processed through the Isomap dimensional reduction algorithm. Finally, the CNN-biLSTM classifier uses meaningful features to identify different activities. Finally, comparison has been done against other state-of-the-art methods to prove the robustness of our proposed AAL model.

In future work, the CNN-biLSTM model will be further developed in the future using other hyper parameters like as regularization, learning rate, batch size, and others. Furthermore, we will create our own AAL-based dataset with more difficult tasks to tackle significant problems in the AAL domain.

Funding Statement: This research was supported by a grant (2021R1F1A1063634) of the Basic Science Research Program through the National Research Foundation (NRF) funded by the Ministry of Education, Republic of Korea.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] H. B. Zhang, Y. X. Zhang, B. Zhong, Q. Lei and L. Yang, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, no. 5, pp. 1–20, 2019.
- [2] K. Chen, D. Zhang, L. Yao, B. Guo and Z. Yu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Computing Surveys*, vol. 54, no. 4, pp. 1–40, 2021.
- [3] A. Prati, C. Shan and K. Wang, "Sensors, vision and networks: From video surveillance to activity recognition and health monitoring," *Journal of Ambient Intelligence and Smart Environments*, vol. 11, no. 1, pp. 5–22, 2019.
- [4] A. Jalal, S. Kamal and D. Kim, "Facial expression recognition using 1D transform features and hidden markov model," *Journal of Electrical Engineering & Technology*, vol. 12, no. 4, pp. 1657–1662, 2017.
- [5] A. Jalal, S. Kamal and D. Kim, "A depth video-based human detection and activity recognition using multi-features and embedded hidden markov models for health care monitoring systems," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, no. 4, pp. 54–62, 2017.
- [6] F. Farooq, A. Jalal and L. Zheng, "Facial expression recognition using hybrid features and self-organizing maps," in *Proc. IEEE Int. Conf. on Multimedia and Expo*, Hong Kong, China, pp. 409–414, 2017.
- [7] A. Jalal, S. Kamal and D. S. Kim, "Detecting complex 3D human motions with body model low-rank representation for real-time smart activity monitoring system," *KSII Transactions on Internet and Information Systems*, vol. 12, no. 3, pp. 1189–1204, 2018.
- [8] M. Mahmood, A. Jalal and H. A. Evans, "Facial expression recognition in image sequences using 1D transform and gabor wavelet transform," in *IEEE Conf. on Int. Conf. on Applied and Engineering Mathematics*, Taxila, Pakistan, pp. 1–6, 2018.
- [9] Y. Shi, F. Li and T. Liu, "Beyette FR, song W. dynamic time-frequency feature extraction for brain activity recognition," in *Annu Int. Conf. IEEE Eng. Med. Biol. Soc.*, Honolulu, HI, USA, pp. 3104–3107, 2018.
- [10] T. Wang, N. Liu, Z. Su and C. Li, "A new time-frequency feature extraction method for action detection on artificial knee by fractional Fourier transform," *Micromachines (Basel)*, vol. 10, no. 5, pp. 333–349, 2019.
- [11] S. Rosati, G. Balestra and M. Knaflitz, "Comparison of different sets of features for human activity recognition by wearable sensors," *Sensors*, vol. 18, no. 4189, pp. 598–605, 2018.

- [12] N. Debs, T. Jourdan, A. Moukadem, A. Boutet and C. Frindel, "Motion sensor data anonymization by time-frequency filtering," in *28th European Signal Processing Conf. (EUSIPCO 2020)*, Amsterdam, Netherlands, pp. 50–58, 2020.
- [13] T. Jourdan, A. Boutet and C. Frindel, "Toward privacy in IoT mobile devices for activity recognition," in *15th Int. Conf. on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, EAI, pp. 155–165, 2020.
- [14] G. Leonardis, S. Rosati, G. Balestra, V. Agostini, E. Panero *et al.*, "Human activity recognition by wearable sensors: Comparison of different classifiers for real-time applications," in *Proc. of the Int. Conf. on Medical Measurements and Applications*, Rome, Italy, pp. 1–6, 2020.
- [15] P. T. Chinimilli, S. Redkar and T. Sugar, "A Two-dimensional feature space-based approach for human locomotion recognition," *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4271–4282, 2019.
- [16] R. N. Mohammadian, T. Laarhoven, C. Furlanello and E. Marchiori, "Novelty detection using deep normative modeling for IMU-based abnormal movement monitoring in Parkinson's disease and autism spectrum disorders," *Sensors*, vol. 18, no. 8, pp. 3316–3320, 2018.
- [17] B. T. Nukala, N. Shibuya, A. I. Rodriguez, J. Tsay, T. Q. Nguyen *et al.*, "A real-time robust fall detection system using a wireless gait analysis sensor and an artificial neural network," in *Proc. of the Int. Conf. on Healthcare Innovation; Seattle*, pp. 219–222, WA, USA, 2014.
- [18] M. C. Kwon, G. Park and S. Choi, "Smartwatch user interface implementation using CNN-based gesture pattern recognition," *Sensors*, vol. 18, no. 9, pp. 1–12, 2018.
- [19] C. Jian, J. Li and M. Zhang, "LSTM-based dynamic probability continuous hand gesture trajectory recognition," *IET Image Process*, vol. 13, no. 12, pp. 2314–2320, 2019.
- [20] Z. Cui, R. Ke, Z. Pu and Y. Wang, "Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting networkwide traffic state with missing values," *Transportation Research Part C: Emerging Technologies*, vol. 118, no. 4, pp. 968–978, 2020.
- [21] S. Mekruksavanich and A. Jitpattanakul, "LSTM networks using smartphone data for sensor-based human activity recognition in smart homes," *Sensors (Basel)*, vol. 21, no. 5, pp. 1–25, 2021.
- [22] S. W. Pienaar and R. Malekian, "Human activity recognition using LSTM-RNN deep neural network architecture," in *IEEE 2nd Wireless Africa Conf. (WAC)*, Pretoria, South Africa, pp. 1–5, 2019.
- [23] K. Xia, J. Huang and H. Wang, "LSTM-CNN architecture for human activity recognition," in *IEEE Access*, vol. 8, no. 6, pp. 56855–56866, 2020.
- [24] R. Chereshevnev and A. K.-Farkas, "HuGaDB: Human gait database for activity recognition from wearable inertial sensor networks," *arXiv: Computers and Society*, vol. 18, no. 1, pp. 191–203, 2018.
- [25] T. Tan, M. Gochoo, S. Huang, Y. Liu, S. Liu *et al.*, "Multi-resident activity recognition in a smart home using RGB activity image and DCNN," *IEEE Sensors Journal*, vol. 18, no. 23, pp. 9718–9727, 2018.
- [26] M. Mahmood, A. Jalal and K. Kim, "WHITE STAG model: Wise human interaction tracking and estimation (WHITE) using spatio-temporal and angular-geometric (STAG) descriptors," *Multimed Tools Appl*, vol. 79, no. 10, pp. 6919–6950, 2020.
- [27] M. Gabrielli, P. Leo, F. Renzi and S. Bergamaschi, "Action recognition to estimate activities of daily living (ADL) of elderly people," in *IEEE 23rd Int. Symp. on Consumer Technologies (ISCT)*, Ancona, Italy, pp. 261–264, 2019.
- [28] C. Siriwardhana, D. Madhuranga, R. Madushan and K. Gunasekera, "Classification of activities of daily living based on depth sequences and audio," in *14th Conf. on Industrial and Information Systems (ICIIS)*, Kandy, Sri Lanka, pp. 278–283, 2019.
- [29] A. Jalal, S. Kamal and K. Kim, "Human depth sensors-based activity recognition using spatiotemporal features and hidden markov model for smart environments," *Journal of Computer Networks and Communications*, vol. 2016, pp. 2090–7141, 2016.
- [30] D. Madhuranga, R. Madushan and C. Siriwardane, "Real-time multimodal ADL recognition using convolution neural networks," *The Visual Computer*, vol. 37, no. 12, pp. 1263–1276, 2021.

- [31] N. Khalid, Y. Y. Ghadi, M. Gochoo, A. Jalal and K. Kim, "Semantic recognition of human-object interactions via Gaussian-based elliptical modeling and pixel-level labeling," *IEEE Access*, vol. 9, no. 5, pp. 111249–111266, 2021.
- [32] J. Lu, Y. Wu, M. Hu, Y. Xiong and Y. Zhou, "Breast tumor computer-aided detection system based on magnetic resonance imaging using convolutional neural network," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 365–377, 2022.
- [33] J. Shi, L. Ye, Z. Li and D. Zhan, "Unsupervised binary protocol clustering based on maximum sequential patterns," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 483–498, 2022.
- [34] D. Cui, D. Li and S. Zhou, "Design of multi-coupled laminates with extension-twisting coupling for application in adaptive structures," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 415–441, 2022.
- [35] R. A.-Gdairi, S. Hasan, S. Al-Omari, M. Al-Smadi and S. Momani, "Attractive multistep reproducing kernel approach for solving stiffness differential systems of ordinary differential equations and some error analysis," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 299–313, 2022.
- [36] O. Ricou and M. Bercovier, "A dimensional reduction of the stokes problem," *Computer Modeling in Engineering & Sciences*, vol. 3, no. 1, pp. 87–102, 2002.
- [37] M. Batool, A. Jalal and K. Kim, "Sensors technologies for human activity analysis based on SVM optimized by PSO algorithm," in *IEEE ICAEM Conf.*, Taxila, Pakistan, pp. 145–150, 2019.
- [38] A. A. Rafique, A. Jalal and A. Ahmed, "Scene understanding and recognition: Statistical segmented model using geometrical features and Gaussian naïve Bayes," in *IEEE Conf. on Int. Conf. on Applied and Engineering Mathematics*, Taxila, Pakistan, pp. 225–230, 2019.
- [39] S. Velliangiri and J. Premalatha, "A novel forgery detection in image frames of the videos using enhanced convolutional neural network in face images," *Computer Modeling in Engineering & Sciences*, vol. 125, no. 2, pp. 625–645, 2020.
- [40] J. Chen, J. Li and Y. Li, "Predicting human mobility via long short-term patterns," *Computer Modeling in Engineering & Sciences*, vol. 124, no. 3, pp. 847–864, 2020.
- [41] S. Lahmiri, "Gait nonlinear patterns related to Parkinson's disease and age," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 7, pp. 2545–2551, 2019.
- [42] M. Gochoo, S. B. U. D. Tahir, A. Jalal and K. Kim, "Monitoring real-time personal locomotion behaviors over smart indoor-outdoor environments via body-worn sensors," *IEEE Access*, vol. 9, no. 12, pp. 70556–70570, 2021.
- [43] M. Seiffert, F. Holstein, R. Schlosser and J. Schiller, "Next generation cooperative wearables: Generalized activity assessment computed fully distributed within a wireless body area network," *IEEE Access*, vol. 5, no. 13, pp. 16793–16807, 2017.