IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 25 January 2024, accepted 25 February 2024, date of publication 1 March 2024, date of current version 9 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3372593

RESEARCH ARTICLE

Effective Gait Abnormality Detection in Parkinson's Patients for Multi-Sensors Surveillance System

ABDULWAHAB ALAZEB[1], MOUAZMA BATOOL[2], NAIF AL MUDAWI[1], MOHAMMED S. ALSHEHRI[1], SULTAN ALMAKDI[1], NOUF ABDULLAH ALMUJALLY[3], AND ASAAD ALGARNI[4]

[1]Department of Computer Science, College of Computer Science and Information System, Najran University, Najran 55461, Saudi Arabia
[2]Department of Computer Science, Air University, Islamabad 44000, Pakistan
[3]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[4]Department of Computer Science, Faculty of Computing and Information Technology, Northern Border University, Rafha 91911, Saudi Arabia

Corresponding author: Naif Al Mudawi (naalmudawi@nu.edu.sa)

The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Research Group Funding program grant code (NU/DRP/SERC/12/26).

ABSTRACT Parkinson's disease affects bodily functions and there is a growing need for advanced solutions to offer therapeutic advice to patients. A framework using artificial intelligence and machine learning techniques has been proposed to address this. The proposed system employs a combination of RGB, inertial, and depth sensors data. The inertial signals have been filtered using a notch filter to obtain the optimal wearable sensor data by examining the upper and lower cutoff frequencies. Multiple features have been calculated, including mel frequency cepstral coefficients (MFCC), statistical features and Gaussian mixture model (GMM) features. On the other hand, silhouettes have been extracted from RGB and depth images, and four crucial parameters have been employed to gauge the level of accuracy with which patients with neurological disorders performed their activities, including the angle formation between the hands and lower/upper half of the body and center of the body, and the angle between the hands of the silhouette. The resulting features have been fused and classified using principle component analysis and a reweighted genetic algorithm. Evaluation using cross-validation on mRI (multi-modal 3d human pose estimation dataset using mmwave, RGB-D, and inertial sensors) and MHEALTH (mobile health) datasets showed a recognition accuracy rate of 97.29% and a 97.94%. The study highlights the need for more datasets to address challenges in rehabilitation using human activity recognition with multi-modal sensors.

INDEX TERMS Depth, human activity recognition, inertial sensors, multi-modal sensors, neurological disorder, RGB.

I. INTRODUCTION

Human Neurological disorders impact the central or peripheral nervous system, affecting various parts of the human body like muscles, brain, cranial nerves, spinal cord, peripheral nerves, and neuro muscular junction [1]. These disorders lead to Alzheimer's disease, epilepsy, dementia and so on [2]. These disorders are also the main cause of cerebrovascular diseases, such as stroke, headaches (migraines, Parkinson's disease, sclerosis, so on), malnutrition related neurological disorders, and infections of the nervous system, traumatic nervous system injuries, and brain tumors [3], [4], [5], [6], [7]. These imbalances, interfere with the voluntary movements of humans [8]. While voluntary movement may appear simple, it is a complex process involving several nerve structures making reflexive and intentional decisions [9]. This mechanism controls movement through nerve impulses aimed at the musculoskeletal system. Neurological disorders is the

The associate editor coordinating the review of this manuscript and approving it for publication was Xinyu Du.

primary contributor to disability and ranks as the second most prevalent cause of mortality on a global scale [10], [11], [12].

Artificial Intelligence (AI) has played significant progress in neurological rehabilitation, with a particular focus on supporting patients grappling with neurological disorders [13]. Personalized rehabilitation programs that cater to each patient's distinctive requirements are one of the pivotal areas where these technologies are being utilized [14], [15], [16]. These technologies are scrutinizing the patient's movement patterns, evaluating their motor capabilities, and monitoring their advancement over time but with limited capabilities till now [17]. Subsequently, based on this analysis, customized exercise routines can be generated to target specific regions of the body or brain. To illustrate, in the instance of patients who have suffered a stroke, AI and machine learning algorithms can examine their movement patterns and identify which areas of the brain have been impacted [18]. This enables the algorithms to devise exercises focusing on those specific areas, facilitating recovery. Additionally, these algorithms can supervise the patient's progress and adjust the exercise routines as necessary to ensure they are both secure and effective [19]. By implementing AI and machine learning algorithms in this manner, patients can receive more individualized and productive rehabilitation programs, ultimately culminating in improved outcomes and a speedier recovery [20].

Physical rehabilitation exercise aims to improve health-related physical fitness and prevent illness [21]. Three key areas are considered: body composition, musculoskeletal fitness, and cardiorespiratory capacity. Body composition looks at the distribution of body tissues, while musculoskeletal fitness relates to muscle strength and flexibility. Cardiorespiratory capacity refers to the ability of the body's systems to supply oxygen during prolonged activity [22]. Different exercises can achieve specific fitness goals, e.g., resistance for muscles, aerobics for cardiorespiratory health, and intensity for anaerobic capabilities. A personalized exercise plan prescribed by a specialized healthcare professional is crucial in implementing physical exercise in rehabilitation, considering potential complications and individual factors, and intensity is the most important factor in physical prescribing exercise. Low/moderate intensity exercises were used, now American Heart Association recommends high-intensity exercises for increasing V O2MAX, but the prescription is complex [23]. AI-based exercise regulation methods are needed to ensure safe rehabilitation and prevent excessive fatigue.

The present study has outlined a novel machine learning-based methodology for physical rehabilitation of patients with neurological disorder, incorporating diverse features for inertial, RGB, and depth images. In this regard, the mRI dataset has been employed for remote health monitoring and rehabilitation training. The outcomes of this research highlight the efficacy of employing a combination of multi-modal approaches, highlighting its potential to optimize physical rehabilitation outcomes through machine learning. The primary significance of this research can be condensed into the following key points:

- Based on our estimation, there is no precedent available research on the recognition of activities in patients with neurological disorders using multi-modal sensors physical rehabilitation
- The principal component analysis (PCA) algorithm has been employed to enhance
- the data for complex human activity patterns, providing both contextual information and behavior classification.
- Furthermore, the long short-term memory-recurrent neural networks (LSTM-RNN)
- classifier has been utilized to categorize the mRI and MHEALTH benchmark datasets, achieving notably superior results compared to other cutting-edge methods.

The subsequent sections of the paper are structured as follows: Section II presents a comprehensive overview of the most recent intelligent frameworks utilized in human activity recognition (HAR). Section III introduces the fusion architecture proposed in this study and concisely explains the evaluation protocols employed. Section IV presents the experimental findings, provides details about the dataset used, and compares the results with the current cutting-edge approaches. Lastly, Section V formulate the research and outlines prospective directions for a future investigation.

II. RELATED WORKS

With the progress made in machine learning and deep learning methodologies, researchers have introduced a multitude of efficient and accurate frameworks to enhance healthcare services for patients. These frameworks leverage human activity recognition (HAR) as a means to accomplish their objectives. This study emphasizes the latest and most promising machine learning and deep learning-based frameworks in this field.

Betancourt et al. [24] developed a platform for human activity recognition, which incorporated self-attention networks and utilized smartphone-collected data. Their methodology involved extracting essential features from time-series data by intelligently allocating attention to significant input attributes. The platform's performance was evaluated using two publicly available datasets, attaining an average accuracy of 97.1% on the UCIHAR dataset. Nevertheless, it is important to acknowledge that the proposed architecture has limitations, as it is restricted to wearable device data and does not adequately account for the learning period. Nazari et al. [25] carried out comparative analysis, assessing various machine learning (ML) methodologies for HAR utilizing the data of knee angle. The study examined raw data and manually derived features in training the ML models. The experimental findings demonstrated that the Gradient Boosting algorithm exhibited remarkable performance of 94% by employing 11-fold cross-validation on trained. On the other hand, the manually extracted features resulted in a

decline in recognition accuracy. Tan et al. [26] employed fusion architecture of convolutional neural network and long short-term memory (CNN-LSTM) to acquire features from RGB images. The twelve discrete actions on HAR activities have been recognized on the resultant extracted patterns. The developed system achieved an impressive accuracy of 97.02%. Singh et al. [27] introduced was a deep learning architecture designed for human activity recognition (HAR), which integrated self-attention mechanisms to identify crucial information within time-series data. By including a self-attention layer, the model acquired weights that captured the inherent connections among input time points of raw sensory data, facilitating accurate activity decoding. The experiments encompassed six distinct public datasets, with the MHEALTH dataset attaining the highest accuracy of 94.86%.

Moreover, Wan et al. [28] utilized five distinct machine learning algortihms including: CNN, SVM, LSTM, MLP, and BiLSTM for the recognition of miscellaneous activities using inertial sensors data. In this paper, the proposed framework has intially utilized CNN for the extraction of high-level patterns and for performing feature classification of the input data. The extracted features have been then later classified using SVM, LSTM, MLP, and BiLSTM classifier. On UCI-HAR dataset, the paper has revealed exceptional precision of 93.21%. Dua et al. [29] incorporated gated recurrent unit (GRU) and CNN together to autonomously extract features for the recognition of human activities. CNN has been employed to isolate distinctive features from the input data, while, GRU finally classify the data. The model has been validated on three individual datasets and gain a phenomenal accuracy of around 97%. An et al. [30] introduced mRI, an innovative dataset for estimating the poses of 3D human movements across multiple modalities. This dataset encompassed over 160,000 synchronized frames and was constructed by capturing 20 subjects engaged in rehabilitation exercises. Unlike prior datasets that mainly emphasized home based health monitoring and single modality, mRI integrated mmWave, RGB-D, and inertial sensor modalities. The researchers conducted comprehensive evaluations on the mRI dataset, analyzing the unique strengths offered by each modality. By making this dataset publicly available, their intention was to stimulate research in pose estimation, action comprehension, and multi-modal learning, and foster the application of home-based health monitoring.

III. PROPOSED MODEL

Multi-modal methodology for neurological disorder rehabilitation using RGB, depth, and inertial sensors for home-based health monitoring has been developed. To implement this methodology, we have developed an application in Python on intel i5-8250 CPU, 64-bit operating system, 1.8GHz processor, and 16GB RAM. For benchmark purposes, we chose the mRI (multi-modal 3d human pose estimation dataset with mmwave, RGB-D, and inertial sensors) dataset, as it is the only multi-modal dataset available for physical rehabilitation

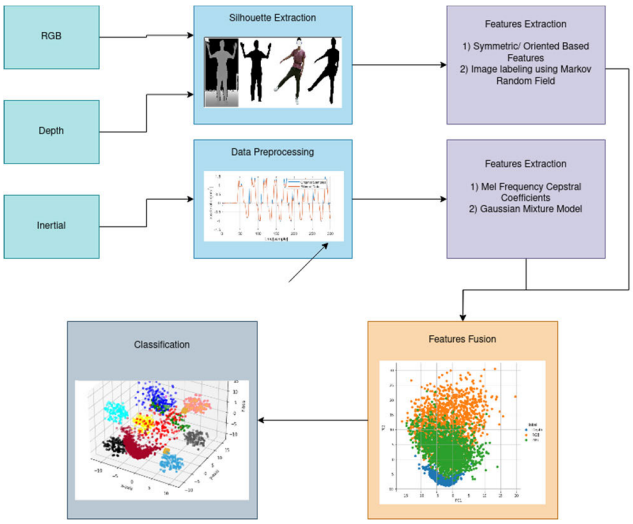

FIGURE 1. Graphical illustration of the suggested model using multi-modal sensors.

of patients for home-based health monitoring. Our approach for data acquisition and processing only requires data of RGB-D and inertial sensors, which are available in mRI datasets. The system's structure encompasses several stages, including preprocessing, feature extraction, fusion, and classification, utilizing PCA and a reweighted genetic algorithm. We tested and trained the mRI dataset using LOSO method. Our proposed framework is illustrated in Figure 1.

A. DATA PROCESSING

Each sensor underwent independent data preprocessing. In order to extract the silhouette from RGBD images, several operations such as replacement, scaling, median/canny edge filtering, and depth mapping on image to RGB images were carried out during the preprocessing step. Additionally, Kalman filtering was utilized to pre-process inertial sensor data [31], [32], [33]. Further elaboration of each preprocessing technique is provided in the subsequent descriptions.

1) RGB-D IMAGE'S SILHOUETTE EXTRACTION USING DEPTH SILHOUETTES

Background subtraction (BGS) serves as the primary and crucial stage for human detection in image, and it is utilized further for feature extraction purposes. Several BGS techniques have been suggested previously, such as self-organizing maps, statistical methods, temporal, and feature based methods [34]. However, BGS methods usually rely on color spaces, such as RGB, YUV, and HSV that include limitations of color camouflage and lighting changes. The preprocessing step involves the use of both substitution & scaling operations. The depth sensing kinect camera encounters limitations in acquiring comprehensive information about depth pixels due the presence of light scattering barriers [22]. To address this, substitution operation has been used to interpolate the missing depth pixel by searching for its

FIGURE 2. Block diagram of the proposed system based on RGB-D data for action recognition.

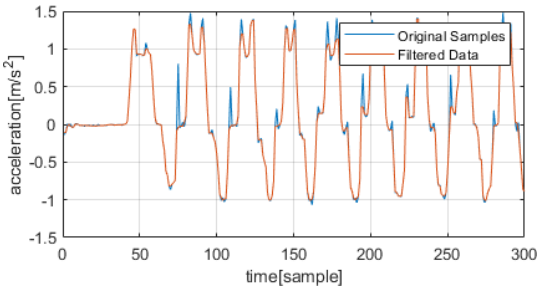

FIGURE 3. The inertial sensors data undergoes Kalman filtration. The red signal represents the filtered data, while the blue represents unfiltered/raw data.

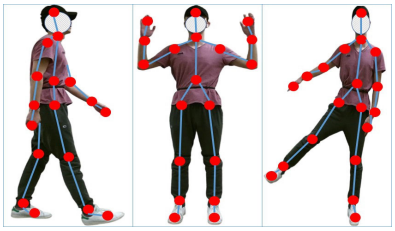

FIGURE 4. Visual representation of nineteen key points identified on a silhouette image.

left and right neighboring pixels and replacing the larger pixel with neighboring depth pixels. Meanwhile, scaling involves linearly scaling the depth image within the range of 0 to 255.

$$V^{\sim} = \frac{V - V_{\min}}{V_{max} - V_{\min}} \omega^{t} + \nu \min \tag{1}$$

The original values are represented by V, while V$^{\sim}$ represents the rescaled values. The maximum, minimum values of the pixels before scaling are denoted as Vmax, and Vmin, respectively [35]. A minimum value of 0 has been established as the lower bound for the depth image, while The disparity between the upper and lower bounds has been determined by ωt, set to 25. Finally, an affine transformation has been used to map the depth silhouette to RGB images Silhouettes were derived from RGB frames through a pre-processing step, and the outcomes of this process are displayed in Fig 2.

2) INERTIAL SENSOR'S PREPROCESSING USING KALMAN FILTER

This step involves preprocessing of inertial sensors data using a Kalman filter, as depicted in Figure 3, to eliminate inconsistencies in the data. The Kalman filter functions as an optimal estimator that extracts relevant parameters from inaccurate observations [36], [37], [38]. This filter process the signal until the maximum likelihood of the signal has been obtained and thus reduces the mean square error up to its maximum best possible level. Its determine as follows:

$$K_s = T_s\,(I_s - 1) + MV_c + P_n \tag{2}$$

where, Ts transition-based model intended for use with the preceding signal Is 1. Moreover, M specify the control input model, which is subsequently applied to the control vector determine by Vc and process noise depicted as Pn of the signal. From covariance and zero mean, the filtered noise is determined, and represented by Pn in the subsequent equation.

B. FEATURE EXTRACTION

1) RGB-D SENSORS BASED FEATURES EXTRACTION

The symmetric based feature extraction methodology underwent two phases [39], [40]. Firstly, a skeleton model was developed to identify the key points of the human skeleton. Secondly, oriented-based features were extracted to determine the angles against upper body, lower body, and middle body points with respect to hands. Further details are provided below.

Symmetric/ Oriented Based Features:

Before advancing to oriented-based feature extraction, it is necessary to perform skeleton modeling as an initial stage for identifying crucial points of the human body.

Skeleton Modelling:

Algorithm 1 depicts the comprehensive explanation of human key point detection model and is also illustrated 4. The model identified thirteen crucial points on the human body, which have been grouped into the upper/lower/middle skeleton points fragmentation. The detection process has been intiated by detecting links among the ankle, feet, hips, knees, elbows, hand points, wrists, shoulders, neck, and head. Every body part play a significant role during the completion of the task. The torso, central point of human body has been identified by calculating the outer shape of the human silhouette. Moreover, the positioning of the ankle is established by identifying the midpoint that lies one-fourth within the region spanning from the foot to the knee 4 In order to approximate the location of the wrist point, The distance encompassing the hands and elbow points has been divided into equal quarters as 4 to estimate the wrist point.

2) ORIENTED-BASED FEATURES:

An angle is always formed between the hands and the body during movement, with a specific measurement taken into account for the upper, middle, and lower body points to maintain symmetry. The upper body points comprises of shoulders Sh1, Sh2, neck Nk, and head Hd. While, the center Mp and lower points of the body is depicted by Hp1, Hp2. Lower

Algorithm 1 Recognition of Essential Keypoints on Human Silhouette

Input: HumanSilhouette(Hsil)
Output: Nineteensalientpointsastorso, hips, shoulders, ankle, elbows, knees, wrists, feet, head, hands, neck
while ((Silhouettedetection)&&(leftandrightSiderecognition)! = NULL) do
  for i = 1 to M
    Ph = Head T racking(Hsil)
    Pn = Upper body localization(Hsil)
    Pt = T orso localization(height, width)
    Pl = Lower body segmentation(Hsil)
    Pkn = Midsection positioning(Pt, Pl)
    Php = Midsection positioning(Ph, Pn)
    Pep = Midsection positioning(Php, Ph)
    Pwp = Midsection positioning(Php, Pep)
    Phip = Pt
    Panp = Midesction positioning(Pkn, Pl)/4
  End
End While

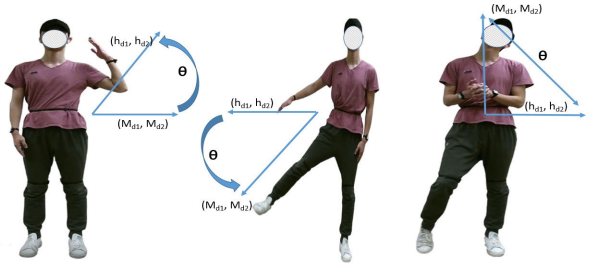

FIGURE 5. Angles demonstration between the hands and three regions of the body, including the upper, lower, and middle sections on body silhouette.

body points consist of knees Kn1, Kn2, and feet Ft1, Ft2. These angles have been evaluated every 0.5 seconds over a sequence of ten consecutive frames, to assess if they satisfy the specific factors used to determine the appropriateness of a particular action at a given moment. The variation in angle configuration between the hands and upper half of the body contributes to the diversity and uniqueness of these six activities: stretching of right/left/both upper limb, squat, right front lunge, and left front lunge. Moreover, the variation the angle formation between hands and center part of the body contribute to the formation of right side lunge, left limb extension, left side lunge, and right limb extension. The critical factor for determining three activities: stretching in free forms, relaxing, and walking; within the chosen mRI benchmark datasets is the configuration of angles formed by the hands in relation to the lower body points. The angle detection process is depicted in Figure 5, where the formation of angle A between the coordinates of hands hd1 and hd2 in relation to the upper, lower, and middle body points Mp1 and Mp2 at time t is represented as follows:

$$\boldsymbol{A}\,(\boldsymbol{t}) = \boldsymbol{tan}\frac{\boldsymbol{h_{d1}} - \boldsymbol{M_{p1}}}{\boldsymbol{h_{d2}} - \boldsymbol{M_{p2}}} \tag{3}$$

3) IMAGE LABELLING USING MARKOV RANDOM FIELD

Markov Random Field (MRFs) is a type of graphical model that assume a set of random variables has the Markov

FIGURE 6. Results of image labeling using Markov random field on right upper limb extension, left front lunge, and left limb extension.

property. The main objective of using MRFs is to capture an image's spatia coherence and context by modeling the relationships between neighboring pixels [41]. The first step in this process is to represent the image as a graph, where the set of indexes are represented by D = 1, . . . , d, the set of pixels are represented as Px = px1, . . . , pxd, the set of labels is represented as L = l1, . . . , ld, and the set of neighbors is represented as Ng = ng1, . . . , ngd. The graph is constructed by connecting neighboring pixels or nodes in Px using the set of neighbors Ng [54], [55]. The edges in the graph represent the relationships between the neighboring pixels. Therefore, by grouping together nodes that are connected by local relationships, a graph can be represented a:

$$N_{gi} = i.j \in D\left[f\left(px_i, px_j\right)\right]^2 \leq r, \quad i \neq j \tag{4}$$

In this context, the Euclidean distance between i and j is denoted by f. The neighbors of i are defined as the collection of locations within a radius of r2, where r represents the order of the neighboring system.

MRF use an energy function to measure the compatibility between neighboring pixels' label assignments, which models the probability distribution over all label assignments in the image. A pairwise potential function is commonly used to penalize label assignments inconsistent with the local neighborhood structure. The optimization problem of finding the best label assignments is solved by minimizing the energy function using Gibbs sampling a

$$\boldsymbol{pr}\,(\boldsymbol{f}) = \boldsymbol{j} - \boldsymbol{1})\boldsymbol{Xe}^{\wedge} - (\boldsymbol{1}/\boldsymbol{c})\boldsymbol{E}(\boldsymbol{f}) \tag{5}$$

where, equation uses the clique potential to calculate the joint probability, with J as the normalizing constant or partition function, and C as a constant equal to 1. The resulting label assignments are obtained by minimizing the energy function, and the inference process is completed once the results are interpreted. Figure 6 depicted the results of image labelling via MRF.

4) INERTIAL SENSORS BASED FEATURES EXTRACTION

Mel Frequency Cepstral Coefficient

MFCCs are frequently employed in speech processing to analyze low frequency speech resolution. To accomplish this, the speech spectra are computed and matched to a Mel scale, which determines the energy in frequency bands [42]. The energy distribution in lower frequencies is similar in both speech and inertial signals, making MFCCs useful for feature extraction in recognizing everyday life activities. To compute

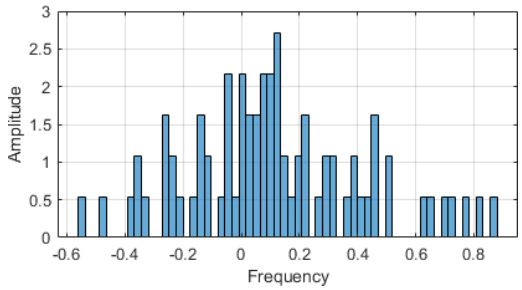

FIGURE 7. Mel frequency cepstral coefficients (MFCC) results of IMU sensors data over mRI dataset's activities.

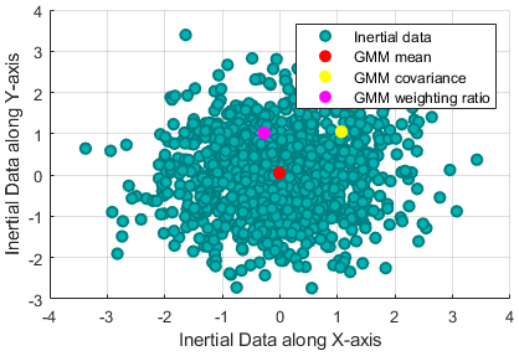

FIGURE 8. Mel frequency cepstral coefficients (MFCC) results of IMU sensors data over mRI dataset's activities.

MFCCs, the signal undergoes a process that includes performing a Fourier transform to acquire the power spectrum, followed by mapping it onto the Mel scale using overlapping windows. Then, after determining the power at Mel frequencies, the logarithm of the power is obtained and subsequently processed through the discrete cosine transform. The obtained MFCCs depict the amplitude of the spectrum, as illustrated in Figure 7.

$$MF_s = \sum_{i=0}^{n} f_i \cos[\frac{P_i}{n(i+0.5)l}] \tag{6}$$

$$S_j = \begin{cases} s_j & l = I_i \\ 0 & otherwise \end{cases} \tag{7}$$

The coefficient of a signal has been obtained by computing each critical band of the MFCCs signals individually using l, where s consists of n samples in the current frame. i denotes the current sample of the signal, while fi refers to the total number of sample signals within each frame [43]. The discrete cosine transform contributes to the reduced correlation observed among the MFCCs features. When the observed signal corresponds to the desired signal, the covariance matrix values [44] serve as the MFCCs coefficients for the signal. In contrast, if the observed signal does not match the desired signal, the signal values have been set to zero. Here, li denotes the index associated with the observed signal.

5) GAUSSIAN MIXTURE MODEL

Gaussian Mixture Model (GMM) is employed to model the clusters N in the dataset to estimate the covariance, mean, and weight vectors [45]. The vector is obtained by determining the maximum likelihood of a particular signal. To identify the weight vector with the highest probability, an iterative maximization method has been utilized. Later on, to quantify the variation between two samples within the current signals, the GMM covariance is computed. The aforementioned calculations yield the following parameters:

$$G\,(mean_x) = (\mathrm{x1} * \mathrm{x2} * \ldots * \mathrm{xn})^{\frac{1}{n}} \tag{8}$$

$$\boldsymbol{Gwv}\,(\boldsymbol{x}) = \frac{\left(\boldsymbol{w1} * (\boldsymbol{x1} - \boldsymbol{\mu})^{\boldsymbol{2}} + .. + \boldsymbol{wn} * (\boldsymbol{xn} - \boldsymbol{\mu})^{\boldsymbol{2}}\right)}{(\boldsymbol{w1} + \boldsymbol{w2} + \ldots \boldsymbol{wn})} \tag{9}$$

In the equation shown in Figure 8, a fram's current and previous samples are represented by xi and xi − 1, respectively, where i is the index number. The means of the accelerometer and gyroscope signals are denoted by μy and μz, respectively. The total number of samples is represented by n. The GMM coefficients, including the mean Gmean(x), weight vector Gwv(x), and covariance Gcov(x), are calculated on the signal using these variables.

C. FEATURES FUSION

Fusion of sensors data as shown in Figure 9, is required to combine RGB, depth, and inertial sensor's data which is of different format into a standard representational format [46]. Principal component analysis (PCA) is an evolving technique that systematically combines data of multiple sensors to make inferences about a particular problem. PCA analyzes the relationship between samples within a given frame by utilizing orthogonal information from multi-modal signals. It achieves this by combining the fram's samples in a manner that preserves variables with the least average square distance while eliminating others [47]. To start the process, the data's center is computed by subtracting the mean M from each attribute ai, placing the origin at the data's center, as demonstrated in the following equation.

$$P_{ca} = (a_1, a_2, a_3, ..a_i) - M \tag{10}$$

The transpose of the covariance matrix is computed, which includes information on the variance and covariance of the data. The eigenvalues Λ and eigenvectors $A^{\rightarrow}$ indicate the magnitude (variance) and direction (covariance) of the data, respectively. After selecting the most suitable eigenvalue and its corresponding lambda, the formula shows that the resulting eigenvector have given the best fit results over multi-modal signals data.

$$\Sigma A^{\rightarrow} = \Lambda A^{\rightarrow} \tag{11}$$

D. CLASSIFICATION

In this paper, the long short-term memory-recurrent neural networks (LSTM-RNN) has been implemented to classify

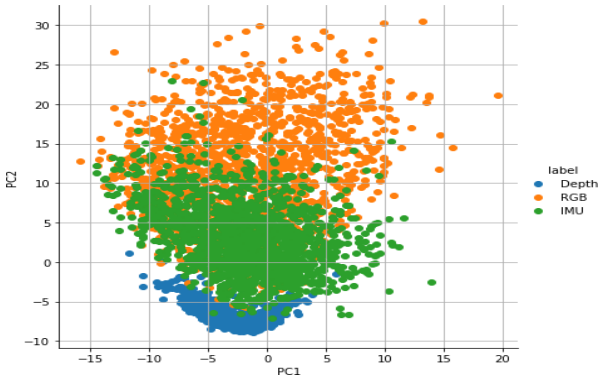

FIGURE 9. Principal component analysis (PCA) results of IMU sensors data, RGB, and depth over mRI dataset's activities.

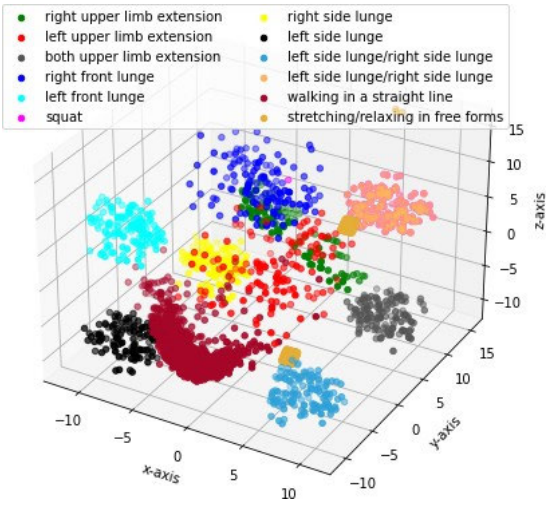

FIGURE 10. Classification results of IMU sensors data, RGB, and depth over mRI dataset's activities.

multi-modal sensors data [33]. Traditional neural networks, due to their absence of memory components, face limitations in linking prior knowledge to the present task and making meaningful inferences about past events. On the other hand, the distinctive characteristic of recurrent neural networks (RNNs) is their ability to process sequential data effectively by allowing information to propagate through recurrent connections in the network topology [48]. The ability to learn long-term dependencies, primarily attributed to the adoption of LSTM (Long Short-Term Memory), is a key factor. Hence, in our research, LSTM-RNNs has been employed to predict probable behaviors by analyzing sequential sensor data.

For our research, a simple LSTM model consisting of single input layer, output layer, and hidden layer have been employed with having 10, 9, and 42 neurons, respectively. During the training phase, the learning rate has been configured as 0.005, while the batch size has been set to 1600. The outcomes of the LSTM-RNN classifiers has been depicted in Figure 10.

FIGURE 11. The twelve rehabilitation exercises covered in mRI dataset include right upper limb extension, left upper limb extension, both upper limb extension, right front lunge, left front lunge, squat, right side lunge, left side lunge, right limb extension, left limb extension, walking in straight line, and stretching/ relaxing.

IV. EXPERIMENTAL SETUP AND RESULTS

This section details the dataset, experimental results, recognition accuracy, and a comparison of our method with current state-of-the-art human activity recognition systems for patients with neurological disorders.

A. DATASET DESCRIPTION

1) THE mRI DATASET

A multi-modal 3D human pose estimation dataset with mmWave, RGB-D, and inertial Sensors (mRI) dataset [30] is a distinctive commercial dataset that integrates data from Inertial sensors, mmWave radar, and RGB-D to assist patients with neurological disorders during their rehabilitation. The dataset encompasses over 160k synchronized frames from three sensing modalities, enabling comprehensive 3D human pose estimation. The data was gathered from 20 human subjects who performed twelve clinically recommended rehabilitation movements including right/left/both upper limb extensions, right/left front lunge, squat, right/left side lunge, right/left limb extension, stretching/ relaxing, and straight line walkin; that involve the essential upper and lower body muscles for human mobility as shown in Figure 11. The data collection process involved utilizing a commercially available low-power and cost-effective mmWave radar, along with six inertial measurement units (IMUs) and two depth cameras. These exercises play a vital role in the rehabilitation of patients who are recovering from central nervous system disorders such as Parkinso's disease (PD) and cerebrovascular diseases like stroke.

2) THE MHEALTH DATASET

The dataset [49] comprises accelerometer, magnetometer, gyroscope and ECG sensors. The ten volunteers performed 12 different physical activities including: Jumping forth and back, running, jogging, cycling, knees bending, frontal elevation of arms, waist bend forward, climbing stairs, walking, lying down, sitting/ relaxing, and standing still. The sensors were placed on the subjec's chest, right wrist and left ankle and5] attached using elastic straps. A sampling rate of 50 Hz is employed to record all sensing modalities, which is deemed suitable for capturing human activity. A video camera was employed to record each session. The dataset encompasses a wide range of activities of daily living, incorporating the

TABLE 1. The evaluation of classification performance of the mRI dataset via confusion matrix.

Activity	RUL	LUL	BUL	RFL	LFL	SQT	RSL	LSL	RLE	LLE	SAR	WSL
RUL	97.22	0	0.33	0	0	1.09	0	1.10	0	0	0.26	0
LUL	1.03	95.08	0	0	2.10	0.80	0	0	0.92	0	0	0.07
BUL	0	0	98.65	0	0.30	0	0.05	0	0	0	1.00	0
RFL	0.50	0.50	0	97.44	0	0	1.50	0	0	0	0.06	0
LFL	0.30	0.05	0.5	0.5	94.65	0.5	1.5	0.5	1.00	0	0	0.5
SQT	0	0.5	0	0.05	0	98.45	0	0.5	0	0.5	0	0
RSL	0.6	0.5	0.07	0	0.5	0.5	95.33	0	1.5	0	0.5	0.5
LSL	0	0	0.3	0	0	0	0	99.77	0	0	0	0
RLE	0.7	0	0	0.06	0	0.5	0	0	98.24	0	0.5	0
LLE	0	0	0.8	0	0.02	0	0.5	0.5	0	97.18	0.5	0.5
SAR	0	0	0.9	0	1.01	0	0.5	0	0	0	97.09	0.5
WSL	0	0	0.6	0	0	0.02	0	0.5	0.5	0	0	98.38
Mean Accuracy = 97.29%												

RUL=right upper limb extension, LUL=left upper limb extension, BUL=both upper limb extension, RFL=right front lunge, LFL=left front lunge, SQT=squat, RSL=right side lunge, LSL=left side lunge, RLE=right limb extension, LLE=left limb extension, SAR=stretching and relaxing, WSL=walking in straight line.

FIGURE 12. The 27 sample activities performed in HWU-USP dataset.

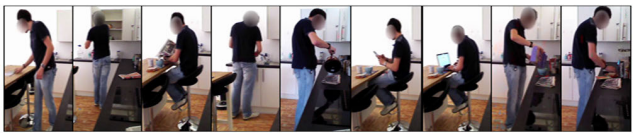

FIGURE 13. The 9 sample activities performed in HWU-USP dataset.

involvement of various body parts (e.g., arms and knees), diverse levels of exertion (e.g., cycling and relaxation), and variations in movement speed or dynamism (e.g., running and stationary positions). Hence, it exhibits generalizability to typical daily activities. The activities were gathered in a non-laboratory setting, allowing participants the freedom to execute them without specific constraints, except for the expectation that they would exert their best effort during the execution.

3) THE UTD-MHAD DATASET

The UTD-MHAD dataset has been implemented by Chen et al. [56] at University of Texas Dallas. The 8 participants including 4 males and 4 females have performed 27 different actions in an indoor environment. Each action is performed four times by each participant with 861 total sequences. The wearable and kinect sensors have been used to record RGB, skeleton, inertial, and depth data respectively. The list of activities include swipe left and right using right arm, draw clock-wise and counter clock-wise circle using right hand, sitting, standing, waving using right hand, clapping, bowling, boxing, pushing, catching, throwing using right arm, jogging, squating, walking, throw or pickup using right hand, forward lunge, knock door, curl arms, swing or serve tennis, swing baseball, shoot basketball, draw triangle, cross arms, and draw X using right hand. The sample images of these activities have been included in Figure 12.

4) THE HWU-USP DATASET

The HWU-USP dataset has been implemented by Ranieri et al. [57] at university of Sao Paulo. The 16 participants have performed 9 activities in a smart home environment. The RGBD data has been collected using TIAGO robot and inertial data have been collected by placing the sensors on wrist and waist. The list of activities include reading newspaper, setting table, cleaning kitchen, using phone/laptop, making cereals/ sandwich/ tea, and dishwashing. The sample images have been depicted in Figure 13.

The system under consideration was tested using the leave one subject out (LOSO) cross validation approach, with separate training and testing datasets. In order to differentiate between different postures and movements, the classification of human activity has been conducted on both the mRI and MHEALTH datasets. The evaluation of the system has been conducted using recall, precision, and F-measure as:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{12}$$

$$Recall = \frac{True\ Positive}{True\ Positive + True\ Negative} \tag{13}$$

$$F1score = \frac{2\,(Precision \times Recall)}{(Precision + Recall)} \tag{14}$$

The confusion matrix of twelve different activities in the mRI and MHEALTH datasets are in Table 1 and 2. According to Table 3, 4, 5, and 6 the system achieved an average accuracy of 97% on F-measure over 4 benchmark datasets, which

TABLE 2. The evaluation of classification performance of the MHEALTH dataset via confusion matrix.

Activity	STS	STR	LYD	WLK	CLS	WBF	FEA	KBC	CYL	JOG	RUN	JFB
STS	98.62	0.30	0.5	0.08	0.5	0.08	0	0	0	0	0	0
STR	0	97.90	0	0	0.1	0	0.5	0	0.5	0	0.5	0.5
LYD	0.4	0	98.57	0	0.03	0	0.5	0	0.5	0	0	0
WLK	0	0.5	0.5	98.83	0	0.1	0	0.07	0	0	0	0
CLS	0.5	0.5	0	0	97.55	0.5	0.5	0	0.4	0.05	0	0
WBF	0	0	0.5	0	0	98.45	0	0.5	0.5	0	0.05	0
FEA	0.6	0.07	0.5	0.5	0	0.5	96.33	0.5	0.5	0	0.5	0
KBC	0.01	0	0.4	0.4	0.5	0.5	0	97.19	0.5	0	0	0.5
CYL	0	0.01	0.5	0	0.5	0	0.5	0	98.49	0	0	0
JOG	0.06	0	0.5	0	0.5	0	0.5	0.5	0	97.44	0	0.5
RUN	0.01	0	0.5	0	0.5	0	0.5	0	0.5	0	97.99	0
JFB	0	0	0.08	0	0.5	0	0.5	0	0	0	0	98.92
Mean Accuracy = 97.94%												

STS=standing still, STR=sitting/relaxing, LYD=lying down, WLK=walking, CLS=climbing stairs, WBF=waist bend forward, FEA=frontal elevation of arms, KBC=knees bending/crouching, CYL=cycling, JOG=jogging, RUN=running, JFB=jump front back

TABLE 3. The determination of classification performance of the mRI dataset via precision, recall, and F1 score.

Class	Precision	Recall	F1 score	Class	Precision	Recall	F1score
RUL	0.972	0.970	0.970	RSL	0.953	0.959	0.955
LUL	0.950	0.959	0.954	LSL	0.997	0.992	0.994
BUL	0.986	0.985	0.985	RLE	0.982	0.987	0.984
RFL	0.974	0.975	0.974	LLE	0.971	0.974	0.972
LFL	0.946	0.949	0.946	SAR	0.970	0.976	0.972
SQT	0.984	0.987	0.985	WSL	0.983	0.987	0.984
Mean Accuracy = 97.29%							

TABLE 4. The determination of classification performance of the MHEALTH dataset via precision, recall, and F1 score.

Class	Precision	Recall	F1 score	Class	Precision	Recall	F1score
STS	0.986	0.987	0.986	FEA	0.963	0.964	0.963
STR	0.979	0.978	0.978	KBC	0.971	0.970	0.970
LYD	0.985	0.986	0.985	CYL	0.984	0.986	0.986
WLK	0.988	0.987	0.987	JOG	0.974	0.975	0.974
CLS	0.975	0.974	0.974	RUN	0.979	0.978	0.978
WBF	0.984	0.986	0.984	JFB	0.989	0.990	0.989
Mean Accuracy = 97.94%							

TABLE 5. The determination of classification performance of the UTD-MHAD dataset via precision, recall, and F1 score.

Class	Precision	Recall	F1 score	Class	Precision	Recall	F1score
A1	0.960	0.961	0.960	A15	0.962	0.963	0.962
A2	0.964	0.965	0.964	A16	0.964	0.965	0.964
A3	0.969	0.968	0.968	A17	0.971	0.970	0.970
A4	0.970	0.971	0.970	A18	0.963	0.962	0.962
A5	0.973	0.972	0.972	A19	0.966	0.965	0.965
A6	0.969	0.968	0.968	A20	0.974	0.973	0.973
A7	0.966	0.967	0.966	A21	0.968	0.969	0.968
A8	0.979	0.978	0.978	A22	0.966	0.967	0.966
A9	0.970	0.971	0.970	A23	0.978	0.977	0.977
A10	0.965	0.966	0.965	A24	0.972	0.971	0.971
A11	0.972	0.971	0.971	A25	0.969	0.968	0.968
A12	0.969	0.968	0.968	A26	0.970	0.971	0.970
A13	0.970	0.971	0.970	A27	0.969	0.968	0.968
A14	0.966	0.965	0.965				
Mean Accuracy = 96.86%							

combines recall and precision. Table 7 compares the proposed approach with existing framework techniques on both mRI and MHEALTH datasets respectively. Finally, Table 8 evaluated the proposed approach in relation to other cutting-edge domains, revealing that our proposed strategy outperforms the state-of-the-art methods.

TABLE 6. The determination of classification performance of the HWU-USP dataset via precision, recall, and F1 score.

Class	Precision	Recall	F1 score	Class	Precision	Recall	F1score
STT	0.975	0.976	0.975	ULP	0.977	0.976	0.976
CTK	0.962	0.963	0.962	MCL	0.969	0.968	0.968
NPR	0.966	0.965	0.965	SMG	0.981	0.980	0.980
DWG	0.979	0.978	0.978	MTE	0.984	0.985	0.984
UPE	0.983	0.982	0.982				
Mean Accuracy = 97.51%							

TABLE 7. The comparison of proposed methodology against mRI and MHEALTH datasets againsts state-of-the-art methods.

METHODS	MRI	MHEALTH
2D TO 3D KEYPOINTS TRANSFORMATION [30]	94.3%	-
BAGGED TREE ENSEMBLE APPROACH [50]	-	95.2%
MULTILAYER PERCEPTRON [51]	-	91.70%
FEDERATED LEARNING NEURAL NETWORK [52]	-	90.2%
EXTREME GRADIENT BOOSTING [53]	-	89.97%
PROPOSED METHOD	**97.29%**	**97.94%**

V. DISCUSSION

The paper presents a successful framework for human activity recognition (HAR) that can aid individuals with neurological conditions during their rehabilitation. The performance of the model highly depends on the quality of the input data and the sufficient amount of data for training the model. The model with existing benchmark datasets has demonstrated an average accuracy of 97% across four designated benchmark datasets. Notably, the lowest accuracy of 94% is observed in the case of left limb extension within the MRI dataset. Recall, precision, and F1-score are computed using equations (11)-(13). Misleading outcomes are attributed to imbalances and the insufficient amount of data in the existing datasets caused suboptimal performance in analogous activities. However, the model could adapt the unseen data over time which ensures its long-term adaptability. Additionally, the proposed model is contrasted with established machine learning and deep learning techniques, as detailed in Tables 5 and 6.

The inertial data has been divided into 6-second windows and 10-second frames for both RGB and Depth images. The Kalman filter has produced an output of 600 samples. Simultaneously, the silhouette extraction process for vision sensors data has yielded 240 binary images. During the feature extraction process, 600 sample data points, each comprising 12 coefficients per frame, resulted in a total of 4,560 data samples, with an additional 1365 data samples extracted. In the subsequent feature fusion process, the PCA output generated 1784 sample data. Finally, the LSTM-RNN classifier provided classification results for 1605 sample data points as shown in Figure 14.

The computational complexity of LSTM models per weight and time step with the RNN model is O(1) and the computational complexity per time step is O(N).

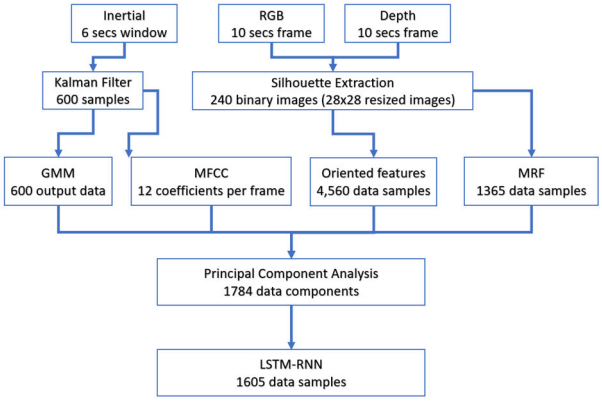

FIGURE 14. Overview of model decision making process.

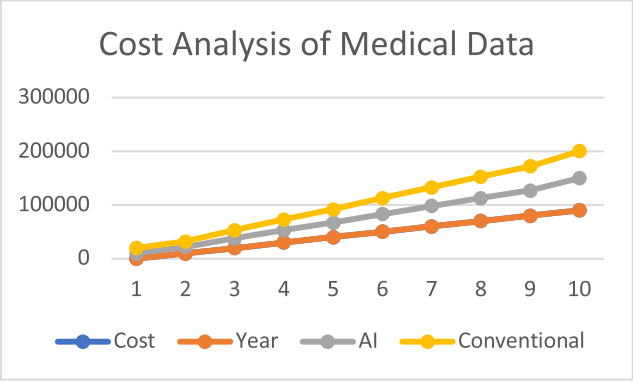

FIGURE 15. Cost-benefit analysis of conventional and AI model in medical data.

Additionally, average learning time is dominated by number of memory cell n_c and number of outputs n_0 as $n_c x(4xn_c + n_0)$ factors. However, the model quite becomes computationally expensive when the number of output units becomes too large to store temporal contextual information. Given its effective performance on benchmark datasets, the model demonstrates a notable capability in recognizing complex tasks when compared to traditional models.

The cost analysis of medical data for assessing the economic viability of the system in real-world deployment primarily focuses on the past 10 years, as explored by researchers. During this examination, data from 20 patients across 20 hospitals were initially considered, leading to a progression to 65 patients per day across a total of 38 hospitals over 10 years. The accompanying Figure 15 illustrates the

TABLE 8. The evaluation of the proposed approach in relation to other cutting-edge domains.

Authors	Domain	Methodology	Accuracy
Betancourt et al. [24]	Time Series	Self-attention features	97.1%
Nazari et al. [25]	Machine Learning	Gradient Boosting algorithm Convolutional neural network (CNN)	94%
Tan et al. [26]	Fusion Architecture	+ Long short-term memory *LSTM*	97.02%
Singh et al. [27]	Deep Learning Model	Model learned weights	94.86%
Wan et al. [28]	Neural network	Convolutional neural network (CNN)	92.71%
Dua et al. [29]	Deep Learning Model	+ Gated recurrent unit (GRU)	97.21%
Proposed Methodology	**Machine Learning**	**Neural N etworks (LSTM-RNN)**	**97.94%**

cost savings achieved by the conventional method in comparison to machine learning and deep learning models. Notably, the cost savings demonstrate an upward trend corresponding to the advancements in AI development. Building upon the insights gained from the preceding analysis, it can be inferred that our proposed model is poised to generate cost savings in medical data, emphasizing its economic feasibility.

VI. CONCLUSION AND FUTURE DIRECTION

The framework employs multi-modal sensors and incorporates oriented-based features for RGB-D sensor, and MFCC, GMM for inertial sensor-based features. These features are used to select the most relevant data for analysis. The RGB, depth, and inertial sensors have been fused using Principal Component Analysis (PCA), and the classification of different movements is achieved using LSTM-RNN. Our proposed system has demonstrated average accuracy of 97% over the benchmark datasets, indicating its potential for real-world application in recognizing patients with neurological disorders during their rehabilitation.

This study has the potential to reduce the time, costs, and errors linked to diagnosing patient health, enabling more effective interventions compared to traditional methods. Additionally, ongoing patient monitoring could detect early warning signs, providing healthcare providers with alerts before the situation becomes critical. In future, this research will investigate the smooth integration of machine learning models into Electronic Health Record (EHR) systems, aiming to provide real-time decision support for clinicians. This integration would expect to improve the model's applicability and its influence on patient care.

REFERENCES

[1] M. Qi, S. Cui, X. Chang, Y. Xu, H. Meng, Y. Wang, and T. Yin, "Multi-region nonuniform brightness correction algorithm based on L-channel gamma transform," *Secur. Commun. Netw.*, vol. 2022, pp. 1–9, Apr. 2022, doi: 10.1155/2022/2675950.

[2] T. Li, T. Xia, H. Wang, Z. Tu, S. Tarkoma, Z. Han, and P. Hui, "Smartphone app usage analysis: Datasets, methods, and applications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 937–966, 2nd Quart., 2022, doi: 10.1109/COMST.2022.3163176.

[3] J. Zhang, Q. Shen, Y. Ma, L. Liu, W. Jia, L. Chen, and J. Xie, "Calcium homeostasis in Parkinson's disease: From pathology to treatment," *Neurosci. Bull.*, vol. 38, no. 10, pp. 1267–1270, Oct. 2022, doi: 10.1007/s12264-022-00899-6.

[4] H. Liu, Y. Xu, and F. Chen, "Sketch2Photo: Synthesizing photo-realistic images from sketches via global contexts," *Eng. Appl. Artif. Intell.*, vol. 117, Jan. 2023, Art. no. 105608.

[5] Y. Shan et al., "Evidence of a large current of transcranial alternating current stimulation directly to deep brain regions," *Mol. Psychiatry*, Jul. 2023, doi: 10.1038/s41380-023-02150-8.

[6] S. Lu, J. Yang, B. Yang, Z. Yin, M. Liu, L. Yin, and W. Zheng, "Analysis and design of surgical instrument localization algorithm," *Comput. Model. Eng. Sci.*, vol. 137, no. 1, pp. 669–685, 2023, doi: 10.32604/cmes.2023.027417.

[7] Y. Miao, X. Wang, S. Wang, and R. Li, "Adaptive switching control based on dynamic zero-moment point for versatile hip exoskeleton under hybrid locomotion," *IEEE Trans. Ind. Electron.*, vol. 70, no. 11, pp. 11443–11452, Nov. 2023, doi: 10.1109/TIE.2022.3229343.

[8] Z. Hu, L. Ren, G. Wei, Z. Qian, W. Liang, W. Chen, X. Lu, L. Ren, and K. Wang, "Energy flow and functional behavior of individual muscles at different speeds during human walking," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 294–303, 2023, doi: 10.1109/tnsre.2022.3221986.

[9] R. Zhang, L. Li, Q. Zhang, J. Zhang, L. Xu, B. Zhang, and B. Wang, "Differential feature awareness network within antagonistic learning for infrared-visible object detection," *IEEE Trans. Circuits Syst. Video Technol.*, 2023, doi: 10.1109/TCSVT.2023.3289142.

[10] M. K. A. Jannat, M. S. Islam, S.-H. Yang, and H. Liu, "Efficient Wi-Fi-based human activity recognition using adaptive antenna elimination," *IEEE Access*, vol. 11, pp. 105440–105454, 2023, doi: 10.1109/access.2023.3320069.

[11] L. Cai, S. Yan, C. Ouyang, T. Zhang, J. Zhu, L. Chen, X. Ma, and H. Liu, "Muscle synergies in joystick manipulation," *Frontiers Physiol.*, vol. 14, Oct. 2023, doi: 10.3389/fphys.2023.1282295.

[12] J. Li, J. Li, C. Wang, F. J. Verbeek, T. Schultz, and H. Liu, "Outlier detection using iterative adaptive mini-minimum spanning tree generation with applications on medical data," *Frontiers Physiol.*, vol. 14, Oct. 2023, doi: 10.3389/fphys.2023.1233341.

[13] Y. Zou, M. Zhong, S. Li, Z. Qing, X. Xing, G. Gong, R. Yan, W. Qin, J. Shen, H. Zhang, Y. Jiang, Z. Wang, and C. Zhou, "Flexible wearable strain sensors based on laser-induced graphene for monitoring human physiological signals," *Polymers*, vol. 15, no. 17, p. 3553, Aug. 2023, doi: 10.3390/polym15173553.

[14] J. Li, J. Luo, L. Liu, H. Fu, and L. Tang, "The genetic association between apolipoprotein E gene polymorphism and Parkinson disease: A meta-analysis of 47 studies," *Medicine*, vol. 97, no. 43, 2018, Art. no. e12884.

[15] N. Wang, J. Chen, W. Chen, Z. Shi, H. Yang, P. Liu, X. Wei, X. Dong, C. Wang, L. Mao, and X. Li, "The effectiveness of case management for cancer patients: An umbrella review," *BMC Health Services Res.*, vol. 22, no. 1, p. 1247, Oct. 2022, doi: 10.1186/s12913-022-08610-1.

[16] S. Hu, W. Chen, H. Hu, W. Huang, J. Chen, and J. Hu, "Coaching to develop leadership for healthcare managers: A mixed-method systematic review protocol," *Systematic Rev.*, vol. 11, no. 1, p. 67, Dec. 2022, doi: 10.1186/s13643-022-01946-z.

[17] M. A. Välimäki, T. Lantta, K. Hipp, J. Varpula, G. Liu, Y. Tang, W. Chen, S. Hu, and X. Li, "Measured and perceived impacts of evidence-based leadership in nursing: A mixed-methods systematic review protocol," *BMJ Open*, vol. 11, no. 10, Oct. 2021, Art. no. e055356, doi: 10.1136/bmjopen-2021-055356.

[18] X. Shen, S.-C. Du, Y.-N. Sun, P. Z. H. Sun, R. Law, and E. Q. Wu, "Advance scheduling for chronic care under online or offline revisit uncertainty," *IEEE Trans. Autom. Sci. Eng.*, 2004, doi: 10.1109/TASE.2023.3310116.

[19] R. Fu, M. Han, F. Wang, and P. Shi, "Intentions recognition of EEG signals with high arousal degree for complex task," *J. Med. Syst.*, vol. 44, no. 6, p. 110, Jun. 2020, doi: 10.1007/s10916-020-01571-0.

[20] J. J. Liszka-Hackzell and D. P. Martin, "Categorization and analysis of pain and activity in patients with low back pain using a neural network technique," *J. Med. Syst.*, vol. 26, no. 4, pp. 337–347, 2002, doi: 10.1023/a:1015820804859.

[21] A. A. Rafique, A. Jalal, and A. Ahmed, "Scene understanding and recognition: Statistical segmented model using geometrical features and Gaussian Naïve Bayes," in *Proc. Int. Conf. Appl. Eng. Math. (ICAEM)*, Taxila, Pakistan, Aug. 2019, pp. 225–230.

[22] L. Gao, A. K. Bourke, and J. Nelson, "Activity recognition using dynamic multiple sensor fusion in body sensor networks," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 1077–1080.

[23] S.-Y. Chang, C.-F. Lai, H.-C.-J. Chao, J. H. Park, and Y.-M. Huang, "An environmental-adaptive fall detection system on mobile device," *J. Med. Syst.*, vol. 35, no. 5, pp. 1299–1312, Oct. 2011.

[24] C. Betancourt, W.-H. Chen, and C.-W. Kuan, "Self-attention networks for human activity recognition using wearable devices," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2020, pp. 1194–1199.

[25] F. Nazari, D. Nahavandi, N. Mohajer, and A. Khosravi, "Human activity recognition from knee angle using machine learning techniques," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2021, pp. 295–300.

[26] T.-H. Tan, C.-J. Huang, M. Gochoo, and Y.-F. Chen, "Activity recognition based on FR-CNN and attention-based LSTM network," in *Proc. 30th Wireless Opt. Commun. Conf. (WOCC)*, Oct. 2021, pp. 146–149.

[27] S. P. Singh, M. K. Sharma, A. Lay-Ekuakille, D. Gangwar, and S. Gupta, "Deep ConvLSTM with self-attention for human activity decoding using wearable sensors," *IEEE Sensors J.*, vol. 21, no. 6, pp. 8575–8582, Mar. 2021.

[28] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mobile Netw. Appl.*, vol. 25, no. 2, pp. 743–755, Apr. 2020.

[29] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input CNN-GRU based human activity recognition using wearable sensors," *Computing*, vol. 103, no. 7, pp. 1461–1478, Jul. 2021.

[30] S. An, Y. Li, and U. Ogras, "MRI: Multi-modal 3D human pose estimation dataset using mmWave, RGB-D, and inertial sensors," 2022, *arXiv:2210.08394*.

[31] S. Velliangiri and J. Premalatha, "A novel forgery detection in image frames of the videos using enhanced convolutional neural network in face images," *Comput. Model. Eng. Sci.*, vol. 125, no. 2, pp. 625–645, 2020.

[32] J. Chen, J. Li, and Y. Li, "Predicting human mobility via long short-term patterns," *Comput. Model. Eng. Sci.*, vol. 124, no. 3, pp. 847–864, 2020.

[33] S. Lahmiri, "Gait nonlinear patterns related to Parkinson's disease and age," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 7, pp. 2545–2551, Jul. 2019.

[34] A. Jazib, W. Tariq, and M. Mahmood, "Sentiment analysis using ensemble classifier for entrepreneurs based on Twitter analytics," in *Proc. 19th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Islamabad, Pakistan, Aug. 2022, pp. 207–212.

[35] 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, pp. 70556–70570, 2021.

[36] 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, pp. 16793–16807, 2017.

[37] 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, pp. 111249–111266, 2021.

[38] J. Lu, Y. Wu, M. Hu, Y. Xiong, Y. Zhou, Z. Zhao, and L. Shang, "Breast tumor computer-aided detection system based on magnetic resonance imaging using convolutional neural network," *Comput. Model. Eng. Sci.*, vol. 130, no. 1, pp. 365–377, 2022.

[39] J. Shi, L. Ye, Z. Li, and D. Zhan, "Unsupervised binary protocol clustering based on maximum sequential patterns," *Comput. Model. Eng. Sci.*, vol. 130, no. 1, pp. 483–498, 2022.

[40] D. Cui, D. Li, and S. Zhou, "Design of multi-coupled laminates with extension-twisting coupling for application in adaptive structures," *Comput. Model. Eng. Sci.*, vol. 130, no. 1, pp. 415–441, 2022.

[41] A. Jalal, M. A. K. Quaid, S. B. U. D. Tahir, and K. Kim, "A study of accelerometer and gyroscope measurements in physical life-log activities detection systems," *Sensors*, vol. 20, no. 22, p. 6670, Nov. 2020.

[42] R. Abu-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 SomeError analysis," *Comput. Model. Eng. Sci.*, vol. 130, no. 1, pp. 299–313, 2022.

[43] Z. Kh. Abdul and A. K. Al-Talabani, "Mel frequency cepstral coefficient and its applications: A review," *IEEE Access*, vol. 10, pp. 122136–122158, 2022, doi: 10.1109/ACCESS.2022.3223444.

[44] M. S. Nordin, A. L. Asnawi, N. A. Zainal, R. F. Olanrewaju, A. Z. Jusoh, S. N. Ibrahim, and N. F. M. Azmin, "Stress detection based on TEO and MFCC speech features using convolutional neural networks (CNN)," in *Proc. IEEE Int. Conf. Comput. (ICOCO)*, Kota Kinabalu, Malaysia, Nov. 2022, pp. 84–89, doi: 10.1109/ICOCO56118.2022.10031771.

[45] H.-B. Zhang, Y.-X. Zhang, B. Zhong, Q. Lei, L. Yang, J.-X. Du, and D.-S. Chen, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, no. 5, p. 1005, Feb. 2019.

[46] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–40, May 2022.

[47] Y. Y. Ghadi, N. Khalid, S. A. Alsuhibany, T. Al Shloul, A. Jalal, and J. Park, "An intelligent HealthCare monitoring framework for daily assistant living," *Comput., Mater. Continua*, vol. 72, no. 2, pp. 2597–2615, 2022.

[48] A. Prati, C. Shan, and K. I.-K. Wang, "Sensors, vision and networks: From video surveillance to activity recognition and health monitoring," *J. Ambient Intell. Smart Environ.*, vol. 11, no. 1, pp. 5–22, 2019.

[49] B. Oresti, G. Rafael, and S. Alejandro, "MHEALTH dataset," UCI Mach. Learn. Repository, Tech. Rep., 2014, doi: 10.24432/C5TW22.

[50] S. Khatun and B. I. Morshed, "Fully-automated human activity recognition with transition awareness from wearable sensor data for mHealth," in *Proc. IEEE Int. Conf. Electro/Inf. Technol. (EIT)*, Rochester, MI, USA, May 2018, pp. 0934–0938.

[51] M. A. Kutlay and S. Gagula-Palalic, "Application of machine learning in healthcare: Analysis on MHEALTH dataset," *Southeast Eur. J. Soft Comput.*, vol. 4, no. 2, pp. 46–51, Mar. 2016.

[52] J. C. Liu, J. Goetz, S. Sen, and A. Tewari, "Learning from others without sacrificing privacy: Simulation comparing centralized and federated machine learning on mobile health data," *JMIR mHealth uHealth*, vol. 9, no. 3, Mar. 2021, Art. no. e23728.

[53] J. O'Halloran and E. A. Curry, "Comparison of deep learning models in human activity recognition and behavioural prediction on the MHEALTH dataset," in *Proc. AICS*, 2019, pp. 212–223.

[54] J. Wang, X. Nie, Y. Xia, Y. Wu, and S.-C. Zhu, "Cross-view action modeling, learning, and recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014.

[55] M. Waheed, A. Jalal, M. Alarfaj, Y. Y. Ghadi, T. A. Shloul, S. Kamal, and D.-S. Kim, "An LSTM-based approach for understanding human interactions using hybrid feature descriptors over depth sensors," *IEEE Access*, vol. 9, pp. 167434–167446, 2021.

[56] C. Chen, R. Jafari, and N. Kehtarnavaz, "UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 168–172.

[57] C. M. Ranieri, S. MacLeod, M. Dragone, P. A. Vargas, and R. A. F. Romero, "Activity recognition for ambient assisted living with videos, inertial units and ambient sensors," *Sensors*, vol. 21, no. 3, p. 768, Jan. 2021, doi: 10.3390/s21030768.

ABDULWAHAB ALAZEB received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2007, the M.S. degree in computer science from the Department of Computer Science, University of Colorado Denver, USA, in 2014, and the Ph.D. degree in cybersecurity from the University of Arkansas, USA, in 2021. He is currently an Assistant Professor with the Department of Computer Science and Information System, Najran University. He received a Graduate Certificate in cybersecurity from the University of Arkansas. His research interests include cybersecurity, cloud and edge computing security, machine learning, and the Internet of Things.

MOUAZMA BATOOL is currently pursuing the Ph.D. degree with Air University, Pakistan. Her research interests include wearable and optical sensors, signal acquisition, the IoT, and life-log generation.

NAIF AL MUDAWI received the master's degree in computer science from Australian La Trobe University, in 2011, and the Ph.D. degree from the Collage of Engineering and Informatics, University of Sussex, Brighton, U.K., in 2018. He is an Assistant Professor with the Department of Computer Science and Information System, Najran University. He has many published research and scientific articles in many prestigious journals in various disciplines of computer science. He was a member of the Australian Computer Science Committee.

MOHAMMED S. ALSHEHRI received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2010, the M.S. degree in computer science from the University of Colorado Denver, Denver, CO, USA, in 2014, and the Ph.D. degree in computer science, with a concentration on information security from the University of Arkansas, Fayetteville, AR, USA, in 2021. His research interests include cybersecurity, computer networks, blockchain, machine learning, and deep learning. He received a Graduate Certificate in cybersecurity from the University of Arkansas, in 2020.

SULTAN ALMAKDI received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2010, the M.S. degree in computer science from the University of Colorado Denver, Denver, USA, in 2014, and the Ph.D. degree in computer science from the University of Arkansas, Fayetteville, USA, in 2020. He is currently an Assistant Professor with the Department of Computer Science and Information Systems, Najran University, Saudi Arabia. His research interests include cloud security, fog security, edge computing security, the IoT security, and computer security. He received a Graduate Certificate in cybersecurity from the University of Arkansas, in 2020.

NOUF ABDULLAH ALMUJALLY received the Ph.D. degree in computer science from the University of Warwick, U.K. She is currently an Assistant Professor of computer science with the Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), Riyadh, Saudi Arabia. Her research interests include human–computer interaction (HCI), artificial intelligence (AI), machine learning, deep learning, and computer based applications.

ASAAD ALGARNI received the Ph.D. degree in software engineering from North Dakota State University, USA. He is an Assistant Professor with the Department of Computer Sciences, College of Computing and Information Technology, Northern Border University, Saudi Arabia. His research interests revolve around software engineering, computer vision applications, and machine learning.