Informatics Bulletin, Faculty of Computers and Artificial Intelligence, Helwan University
Published Online *Vol 6 Issue 2, July 2024*
(https://fcihib.journals.ekb.eg)

Artificial Intelligence Techniques in Enhancing Home-Based Rehabilitation: A Survey

Moamen Zaher[*] [†] , Amr Ghoneim [‡] , Laila Abdelhamid[§], and Ayman Atia [†] [¶]
[*]Software Engineering, Faculty of Computers and Artificial Intelligence, Helwan University, Giza, Egypt
[†] October University for Modern Sciences and Arts (MSA), Giza, Egypt
[‡] Computer Science Dept. , Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt
[§]Information System Dept. , Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt
[¶]HCI-Lab, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt
{e-mail: moibrahem@msa.edu.eg, amr.ghoneim@fci.helwan.edu.eg,
laila.abdelhamid@fci.helwan.edu.eg, aezzat@msa.edu.eg}

***Abstract*—Globally, an estimated 2.4 billion people live with health conditions that may benefit from rehabilitation, yet there is a significant shortage of skilled rehabilitation practitioners, particularly in low- and middle-income countries, with only 10 per 1 million population according to World Health Organization(WHO). The global demand for rehabilitation services, exacerbated by the COVID-19 pandemic, underscores the need for innovative solutions to improve accessibility and efficiency. Instead of increasing the number of physiotherapists, This research focuses on enhancing physiotherapist productivity by monitoring more patients simultaneously through home-based rehabilitation. This study investigates the integration of Human-Computer Interaction (HCI), computer vision, and sensor technologies to transform physical therapy. Key challenges include ensuring model generalizability, various data acquisition sensors, and overcoming barriers to real-world implementation. A comprehensive framework is proposed for home-based rehabilitation, utilizing HCI, computer vision, and sensor technologies to automate exercise assessment and classification. This framework aims to enable personalized rehabilitation programs and alleviate the strain on healthcare systems.**

***Index Terms*—Rehabilitation, Kinect, RGB, Skeleton-based, Machine Learning, Transfer Learning, Deep Learning, Fusion, Ensemble Learning.**

I. INTRODUCTION

In recent years, the convergence of Human-Computer Interaction (HCI), computer vision, and sensor technologies has emerged as a transformative force within the realm of physical therapy [1]–[4]. This interdisciplinary fusion is driven by the escalating global need for innovative rehabilitation services, underscored by the World Health Organization's report that nearly a third of the world's population could benefit from such interventions [5]. The advent of the COVID-19 pandemic has further accentuated this demand, highlighting the critical role of accessible, personalized, and efficient rehabilitation in supporting the world's aging population and those recovering from illnesses or injuries.

Rehabilitation exercises are pivotal in enhancing physical functionality and well-being, enabling individuals to return to their daily routines [6]. Traditional approaches to rehabilitation, however, grapple with challenges in accessibility, personalization, and patient engagement, exacerbated by a global shortage of skilled practitioners, especially in resource-limited settings [7], [8]. These challenges underscore the pressing necessity for innovative, technology-driven solutions to augment the effectiveness and reach of rehabilitation services. This survey delves into the utilization of computer vision and sensor technologies to revolutionize physical therapy. By harnessing the power of advanced deep learning techniques and innovative sensor applications, this research explores automated frameworks for the assessment and classification of rehabilitation exercises. Employing cutting-edge technologies, such as Kinect cameras and sophisticated algorithms like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks [9]–[13], the study aims to facilitate the creation of customized treatment plans and real-time monitoring of patient progress. These technological advancements promise to elevate patient engagement, improve clinical outcomes, and alleviate the logistical and financial burdens associated with traditional rehabilitation methods [14]–[17].

Furthermore, the integration of these digital technologies into home-based rehabilitation represents a paradigm shift towards more personalized, accessible, and efficient healthcare solutions [18]–[20]. This transition is not only in line with the Sustainable Development Goals, particularly those focusing on health and well-being, but also paves the way for enhancing global health outcomes and accessibility to crucial rehabilitation services [21], [22]. The development of novel frameworks and models proposed in this research holds the potential for integration into mobile applications, allowing patients to perform rehabilitative exercises from their homes. This approach addresses both the risk of physical visits during the pandemic and the acute shortage of physiotherapists worldwide.

Through a comprehensive examination of the intersection between HCI, computer vision, sensor technologies, and physical therapy, this thesis aims to forge new pathways in rehabilitating care. By advancing the classification, assessment, and personalization of physical therapy practices through these technologies, the research contributes to a broader vision of achieving more effective, personalized, and accessible rehabilitation services globally.

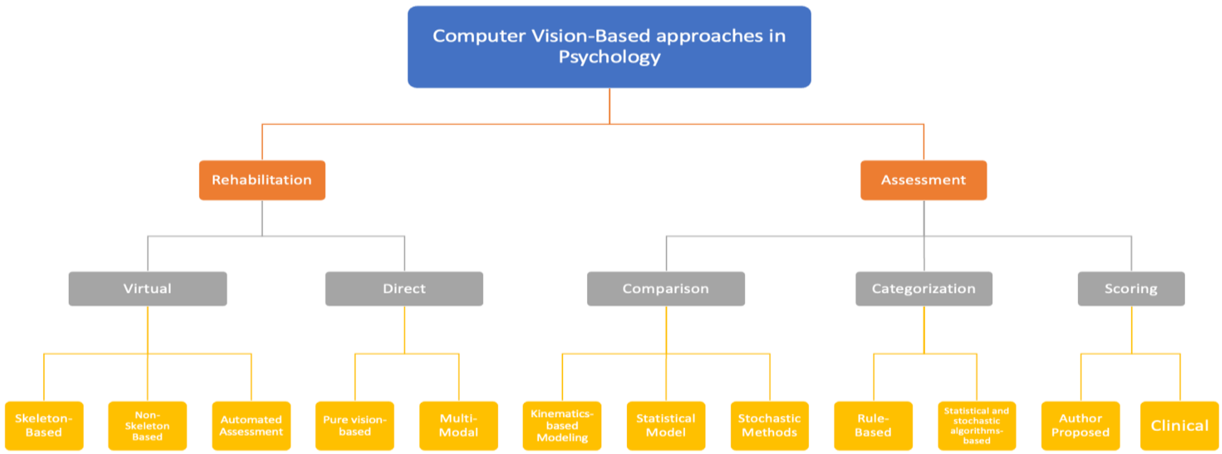

Fig. 1. Computer Vision Approaches in Physical Therapy

This manuscript offers a detailed exploration of computer vision's role in psychology across eight sections, covering data collection methods, key datasets, feature extraction, and various machine learning approaches. Section II outlines computer vision's applications in psychology. Section III examines data acquisition techniques and the sensors used. Section IV reviews essential benchmark datasets. Section V discusses feature extraction methods. Section VI presents machine learning strategies, including classical approaches, deep learning, transfer learning, ensemble learning, and transformers. Section VII identifies challenges and future research directions. Section VIII discusses key insights, and Section IX concludes the survey, weaving together the diverse topics explored.

II. Physical Rehabilitation Application

Liao et. al [23] conducted a survey on vision based approaches in psychology rehabilitation. Various computer vision approaches were applied to contribute in physical rehabilitation and the assessment of the quality of the movement.

Figure 1 provides a structured overview of computer vision-based approaches in the field of psychology, with a particular focus on two main application areas: rehabilitation and assessment. In rehabilitation, the framework is divided into virtual and direct methods. Virtual rehabilitation encompasses skeleton-based and non-skeleton-based approaches, as well as automated assessment tools that leverage computer vision to evaluate patients' movements and progress. Direct rehabilitation methods, on the other hand, utilize pure vision-based techniques or multi-modal approaches, which might combine visual data with other sensor input for a comprehensive analysis.

The assessment domain is further bifurcated into comparison, categorization, and scoring methodologies. Comparison involves kinematics-based modeling, statistical models, and stochastic methods to evaluate and contrast different movements or postures. Categorization employs rule-based systems as well as statistical and stochastic algorithms to classify psychological states or behaviors. Lastly, the scoring segment illustrates the use of author-proposed algorithms alongside clinical standards to quantify the outcomes of psychological assessments.

Figure 1 serves as a guide to the spectrum of computer vision applications in psychology, highlighting how different methodologies can be applied to the analysis, treatment, and evaluation of psychological and rehabilitative processes. Each branch and sub-branch represents a specific set of techniques, reflecting the diverse and intricate ways in which computer vision can contribute to advancements in psychological practices.

This survey focus on rehabilitation approaches. Rehabilitation approaches can be further divided into virtual and direct rehabilitation.

1) Virtual Rehabilitation: In virtual rehabilitation, a patient's performance in a virtual world is assessed rather than directly assessing a patient's physical performance. This includes an avatar performing tasks in a virtual world and the use of serious games for rehabilitation. Here, subjects are required to perform activities in a virtual world through real world movements as shown in Figure 2.

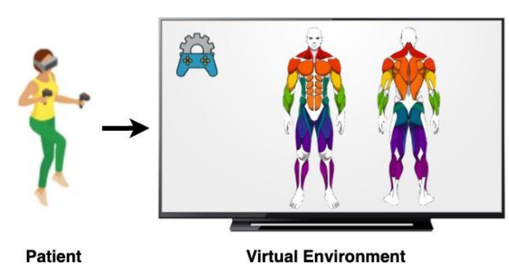

Fig. 2. Virtual Rehabilitation: A patient engages in therapeutic exercises within a virtual reality environment, guided by an avatar in a gamified setting, showcasing the seamless integration of physical movement and digital technology.

2) Direct Rehabilitation: In Direct Rehabilitation systems, users are guided by a web-based interface to perform rehabilitation exercises, while their movements are directly tracked through vision-based sensor. In this case, physical performance of patient is measured instead of their avatar's performance or their ability to complete tasks in a virtual world. Patient assessment may be inbuilt or may require clinicians as shown in Figure 3.

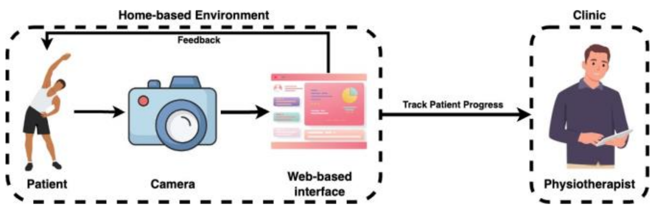

Fig. 3. Direct Rehabilitation: A patient performs rehabilitative exercises under the guidance of a computer vision system, with real-time movement analysis projected on a digital screen, while a physiotherapist monitors the progress on a tablet.

Figure 4 illustrates various stages of a general AI-driven digital Rehab program. digital Rehab programs typically include a clinical assessment and clinician meetings with patients virtually or in person, and then the prescription of individualized digital Rehab programs.

This framework consist of multiple phases, starting with patient performing exercises at home in persistence of a sensor. The sensor type will be discussed in data acquisition phase (Section III), The data are then processed (Section V represents the preprocessing phase followed by feature extraction). Then, the extracted features are fed into Artificial Intelligence models which perform tasks as classification whether the user performed which exercise and if he/she performed it correctly or not. or can even evaluate the quality of the movement and generate progress reports which are then sent back to the clinicians to monitor patient status.

III. Data Acquisition

Data Acquisition is considered the first phase in human activity recognition [24]. The section on data acquisition in human activity recognition is crucial for understanding how information about human behavior is collected. Different sensor modalities and technologies are employed to capture relevant data. Below is a comprehensive overview of data acquisition methods, including sensor-based and camera-based approaches. In the realm of human activity recognition, a wide array of sensor modalities and technologies are deployed to capture the multifaceted aspects of human movements and behaviors, each offering unique advantages and challenges. A comprehensive understanding of these data acquisition methods is essential for designing systems that are both accurate and practical for real-world applications. The section is further divided into two sub-sections: Sensor-based and Camera-based.

A. Sensor-Based

Numerous types of sensors have been utilized for acquiring movement data. The sub-section is further divided into 3 categories based on the type of the sensor: Inertial Sensors, Physiological Sensors and Location-based Sensors.

1) Inertial Sensors: Inertial Measurement Units (IMUs) represent a cornerstone in the landscape of motion detection and analysis, with accelerometers, gyroscopes, and magnetometers serving as their critical components. These sensors, either used individually or combined, offer a nuanced understanding of human movement by capturing various aspects of physical dynamics. The synergy of data they provide forms the backbone of numerous applications, from fitness tracking and sports analytics to rehabilitation and gesture recognition.

- Accelerometers: Measure acceleration and provide information about changes in velocity. Commonly used for recognizing various physical activities, such as walking, running, or gestures, by capturing changes in linear acceleration. [25]–[28].
- Gyroscopes: Measure angular velocity, aiding in determining the orientation and rotation of body parts. Commonly used for tracking rotational movements, like the twist of a wrist or the turn of a head, complementing accelerometer data [27]–[29].
- Magnetometers: Measure the strength and direction of magnetic fields, useful for compass-like orientation. However, It's usually used alongside accelerometer and gyroscope for Human Activity Recognition (HAR) tasks [27], [28]

2) Physiological Sensors:

- Electrocardiography (ECG or EKG): Measures electrical activity of the heart, providing insights into the user's cardiac response [30], [31].
- Electromyography (EMG): Records muscle activity and can be used to identify specific gestures or motions [32], [33].
- Galvanic Skin Response (GSR): Measures changes in skin conductance due to emotional or physiological arousal [34], [35].

3) Location-based Sensors:

- Global Positioning System (GPS) : Provides geo-spatial coordinates, enabling the tracking of outdoor activities and movement patterns [36], [37].

B. Camera-Based Data Acquisition

Also known as Skeleton-Based Cameras, this category involves the use of different cameras specifically designed to detect and track the human skeleton. The information derived from these cameras, capturing the skeletal structure, is subsequently employed for activity or exercise recognition. This section explores various types of cameras that fall under the Skeleton-Based category, highlighting their role in extracting essential skeletal data for applications such as human activity recognition.

1) RGB Cameras: Also known as Color Cameras, RGB cameras capture visual information in the form of color images, which are useful for recognizing activities based on visual cues. Additional Step is needed for extracting body joints when using RGB cameras. Nevertheless, additional processing is essential to extract skeletal joints. Two viable approaches for using RGB cameras as a primary sensor are OpenPose [38] and BlazePose [39]. OpenPose and BlazePose

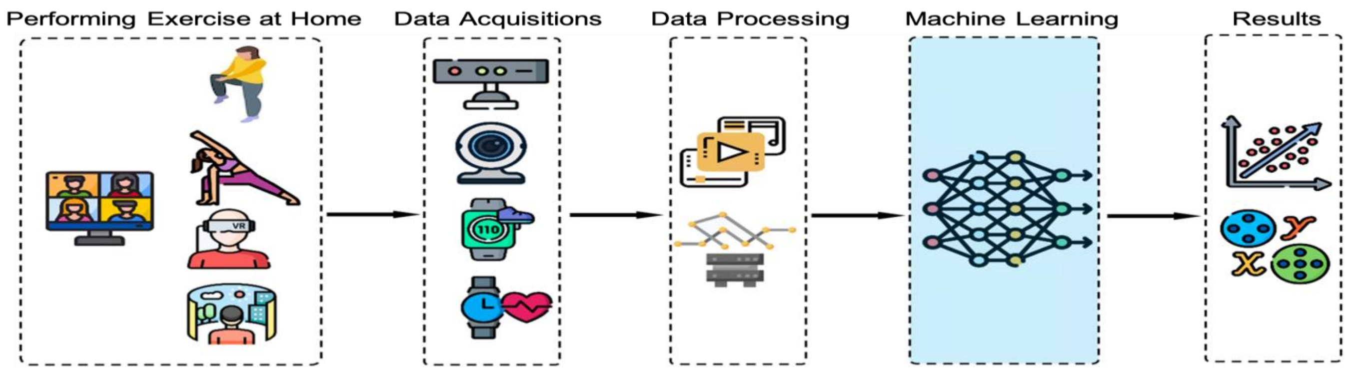

Fig. 4. A conceptual diagram depicting various stages of AI-driven digital Rehab platforms.

are widely adopted solutions for extracting skeletal joints from the human body. BlazePose offers a significantly faster runtime than OpenPose [40], making it suitable for diverse applications such as movement pre-screening and activity classification. BlazePose, functioning as a lightweight and efficient (CNN) model, excels in real-time pose estimation by predicting the 3D pose of an individual from a single image or video frame. The model employs CNNs to extract features from input images, followed by regression layers to predict the location of body key points. On the other hand, MediaPipe [41], a framework based on the BlazePose model, is better suited for constructing real-time machine-learning pipelines. RGB cameras have utilized for previous researches in the real of Human Activity Recognition (HAR) [42]–[44] specially in exercise classification [45]

2) Depth Cameras: Depth cameras can be implemented through various technologies. Two common types are:

- Time-of-Flight (ToF) Cameras:
  Also known as depth cameras, ToF cameras measure the time it takes for light to travel from the camera to the subject, providing depth information. Examples include the *Microsoft Kinect*, *Intel RealSense D series*, and *Azure Kinect* [46]–[48].
- Infrared Cameras
  Infrared cameras detect heat radiation, enabling the capture of thermal images. This is particularly useful for tracking human presence and activity in low-light conditions. Examples include the *FLIR ONE* attachment for smartphones and specialized thermal imaging cameras like the *FLIR Axiom Series* [49], [50].

The utilization of camera-based systems proves particularly advantageous in the realm of physical rehabilitation exercises. This is attributed to the heightened accuracy facilitated by the precise capture of orientation and spatio-temporal coordinates of various body joints. This detailed information plays a critical role in the effective monitoring and classification of physical exercises within the rehabilitation context.

The adoption of RGB cameras in physical rehabilitation presents a set of limitations that are pivotal to consider. Real-time human pose estimation, a critical component of direct rehabilitation applications, can be effectively facilitated by solutions like MediaPipe. However, MediaPipe is limited by its ability to detect only a single human pose at a time, posing a significant constraint in environments where multiple individuals need to be tracked simultaneously. Alternative frameworks, such as YOLO or BlazePose, extend this capability to multiple human figures, yet they often encounter trade-offs in terms of processing speed. The reduced frame rate inherent to these models can impair their real-time applicability, a non-negotiable requirement for feedback-sensitive rehabilitative scenarios.

Transitioning to a Kinect camera may circumvent these issues, offering robust multi-person tracking at higher frame rates, thereby enabling more seamless real-time interaction. Nevertheless, the implementation of a Kinect system introduces additional economic considerations due to the cost associated with the hardware. For rehabilitation centers and patients alike, this investment may represent a significant financial burden. This highlights the broader dilemma in the field of digital rehabilitation: the balance between technological capability and accessibility. Further research is necessary to optimize the trade-offs between cost, performance, and real-time processing capabilities to ensure that the benefits of advanced HCI technologies in rehabilitation can be realized universally.

When comparing the use of cameras and inertial sensors for human activity recognition, both modalities have distinct advantages and limitations. Cameras, particularly RGB cameras, provide rich visual information that can capture detailed features of human actions and interactions with the environment, making them effective in recognizing complex activities and gestures. However, cameras require a clear line of sight and sufficient lighting conditions, which can limit their effectiveness in certain environments or during nighttime, and privacy concerns may arise due to their intrusive nature, especially in private or sensitive settings. On the other hand, inertial sensors, such as accelerometers and gyroscopes, are wearable and unobtrusive, making them suitable for continuous monitoring of human activities in various environments, particularly in scenarios where camera-based systems are impractical or invasive, such as in sports training or healthcare applications. However, inertial sensors have limitations in capturing detailed visual information, and their accuracy can be affected by sensor placement and orientation, limiting their ability to

TABLE I
COMPARISON OF DATASETS FOR PHYSICAL REHABILITATION

Dataset Name	Exercise	Subjects	Number of Activities	Details	Sensor/Data
SPHERE-Staircase2014 [53]	Walking-up stairs	12	2	48 sequences, normal and abnormal gait	Kinect/Open NI skeleton
SPHERE-Walking2015 [54]	Walking	10	2	40 sequences, normal and abnormal gait	Kinect/Kinect SDK, OpenNI SDK skeleton
SPHERE-SitStand2015 [55]	Sit to stand	10	2	109 sequences, restricted knee, hip, freezing	Kinect/Kinect SDK, OpenNI SDK skeleton
TRSP [56]	Stroke, compensatory movement	19	4	4 compensatory movements, frame-by-frame label	Kinect, Haptic robot/Kinect SDK skeleton
Parkinson's pose estimation [57]	PD, LID, UPDRS assessment tasks	9	3	526 sequences, PD, LID patients, 4 UPDRS assessment tasks	RGB Camera/CPM [58] skeleton
UI-PRMD [52]	General rehabilitation	10	10	10 exercises, 10 repetitions	Kinect Vicon/Kinect SDK skeleton
KIMORE Dataset [59]	Stroke, PD, back pain exercises	78	5	5 exercises, 5 repetitions	Kinect/RGB, depth, skeleton
AHA-3D Dataset [60]	Senior lower body fitness	21	4	4 exercises	Kinect/RGB, depth, skeleton
UTD-MHAD [61]	Multiple activities of daily living (ADLs)	8	27	27 different actions, 4 repetitions	Depth Camera (Kinect), Inertial Sensors (Accelerometers, gyroscopes)
UAV-Human [62]	General activities	119	155		RGB /Azure DK Depth / 3D joints
HARTH [63]	General activities	22	12	Two Accelerometers	Accelerometer
PAMP2 [64]	Walking, Cycling, Playing soccer, and more	9	18		Accelerometer, Gyroscope, Temperature, Heart Rate Monitor
HHAR [65]	Biking, Sitting, Standing, Walking, Stair Up and Stair down	9	6		Accelerometer and gyroscopes

recognize complex activities that involve subtle movements or interactions with the environment. In summary, while cameras excel in capturing detailed visual information for human activity recognition, they may be limited by environmental factors and privacy concerns, whereas inertial sensors offer a wearable and unobtrusive solution but may lack the visual context and detailed information provided by cameras. The choice between the two modalities depends on the specific requirements of the application, balancing between accuracy, privacy, and practicality.

IV. BENCH-MARKING DATASETS

In this section, we delve into the methodologies employed for data acquisition and the structural characteristics of selected datasets crucial for human activity recognition in the realm of physical rehabilitation. Each dataset is scrutinized with regard to its distinctive features, strengths, weaknesses, limitations, and prevalent applications.

Here, most authors have used their own small datasets and thus, it is difficult to ascertain their generalisability. Owing to availability of skeleton positions, kinematic parameters have been used for performing statistical comparisons like ANOVA analysis. Small datasets are not sufficient for the application of Deep Learning (DL) algorithms but other algorithms such as HMM, DTW could have been used for comparing temporal sequences. Joint angle comparison is good for posture recognition. However, time sequence comparison algorithms are essential for comparing joint angle and/or joint position trajectories.

In recent times, Generative Adversarial Networks (GANs) have been used to generate synthetic data including, but not limited to, human faces and human poses. Li et al. [51] used the UI-PMRD dataset [52] to generate a synthetic dataset of incorrect human activities. Four different GANs models were trained, which included two Deep Convolutional GANs (DCGAN), a Wasserstein GAN and a Recurrent GAN. A 1D Convolutional Neural Network (CNN) was trained as discriminator with the GANs and a soft-metric based on absolute differences was used for evaluating the performance of GANs. Modelling or replicating kinematic data through GAN is a major contribution of this article, although it aims to classify physical movements.

The table presented in Table I provides a concise summary of the most frequently employed datasets within the field of physical rehabilitation exercises.

- SPHERE-Staircase2014 Focused on stair climbing, this dataset captures gait patterns during ascent using RGB and depth sensors [53]. Comprising 48 sequences from 12 subjects, the dataset offers valuable insights into rehabilitation scenarios. The data is formatted as sequences of frames, where each frame includes color and depth information.
- SPHERE-Walking2015 Specializing in walking activities, this dataset employs Kinect and OpenNI SDK for RGB and depth data acquisition [54]. With 40 sequences from 10 subjects, it contributes to the comprehension of diverse walking patterns. Data is structured as sequential frames, integrating color and depth information.
- SPHERE-SitStand2015 Targeting sit-to-stand movements, this dataset employs Kinect and OpenNI SDK [55]. It features 109 sequences involving 10 individuals with specific conditions. The dataset's structure encompasses sequential frames, incorporating color and depth information relevant to rehabilitation scenarios.
- TRSP Tailored for stroke and compensatory movement analysis, TRSP integrates Kinect and Haptic robot sensors [56]. The dataset includes 10 sequences with 10 healthy subjects and 10 stroke patients, structured as sequential frames capturing RGB and depth information.
- Parkinson's Pose Estimation Centered on Parkinson Disease assessments, this dataset utilizes RGB cameras for pose estimation [57]. Involving Parkinson Disease and Levodopa-induced Dyskinesia patients in 526 sequences for UPDRS assessment tasks, the dataset is characterized by sequential frames capturing essential pose-related data.
- UI-PRMD address general rehabilitation exercises, UI-PRMD employs Kinect Vicon sensors [52]. With 10 subjects performing 10 exercises for 10 repetitions, the dataset's sequential frames encompass RGB and skeletal information, offering versatility in rehabilitation scenarios.
- KIMORE Dataset Catering to stroke, Parkinson Disease, and back pain exercises, KIMORE uses Kinect sensors for RGB, depth, and skeletal data [59]. The dataset includes 78 subjects performing exercises structured as sequential frames, providing diverse insights into targeted impairments.
- AHA-3D Dataset Focused on senior lower body fitness, AHA-3D employs Kinect for RGB, depth, and skeletal data [60]. With 21 subjects performing exercises, the dataset's sequential frames capture nuanced movements, contributing to the understanding of elderly fitness.
- UTD-MHAD Encompassing multiple ADLs, UTD-MHAD utilizes depth cameras (Kinect) and inertial sen-

sors (accelerometers, gyroscopes) [61]. The dataset involves eight subjects performing 27 actions, capturing diverse activities and sensor modalities in sequential frames. The relatively small subject pool and nuanced movement capture present notable challenges for analysis. challenges for analysis.

These datasets collectively contribute to the understanding of rehabilitation strategies, each offering specific insights and considerations. Researchers should carefully select datasets aligned with their study objectives, weighing factors such as dataset size, diversity, and the targeted rehabilitation scenarios.

V. Feature Extraction

Feature extraction is a pivotal step in human activity recognition, involving the conversion of raw sensor or camera data into a representation that captures essential patterns for analysis. Various methodologies exist for feature extraction, each with distinct advantages and applications.

The utilization of numeric data, whether in one-dimensional (1D) or two-dimensional (2D) formats, is fundamental to the feature extraction process in human activity recognition. In the context of 1D data, such as temporal sequences from sensors or joint angles, numerical representations are derived through techniques encompassing statistical measures and signal processing. These 1D numeric features effectively capture temporal patterns and variations critical for comprehending sequential activities.

Conversely, in the realm of image-based data, typically represented in a 2D format like video frames or depth maps, feature extraction employs methods like convolutional neural networks (CNNs). These CNNs autonomously acquire hierarchical features and spatial relationships within the visual input, proving highly effective in capturing spatial dependencies and discerning patterns in human activities within the visual domain. It is noteworthy that 2D data serves a dual purpose, being utilized as numeric data for recurrent neural networks (RNNs), and it can also be transformed into images for tasks related to image classification. The integration of both 1D and 2D numeric representations in feature extraction ensures a comprehensive understanding of the diverse aspects intrinsic to human activity data.

In this section, we delve into the diverse approaches employed for feature extraction, data representation methods, spanning conventional statistical measures, signal processing methods, and domain-specific descriptors. Subsequently, we explore the transformative impact of deep learning, particularly through Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), as well as attention-based feature extraction methods facilitated by transformer architectures. Understanding the nuances of these methodologies is crucial for researchers and practitioners seeking to optimize feature extraction strategies for robust and context-aware human activity recognition.

Table II provides a Comprehensive review of different feature extraction techniques for Time-Series Data specially in the domain of Human Activity Recognition (HAR).

A. 1D Approaches

Traditional or hand-crafted features are manually designed representations that encapsulate specific characteristics of the input data usually used to convert 2D data to 1D. These features are often tailored to capture domain-specific knowledge and are carefully selected based on the requirements of the recognition task. Examples of hand-crafted features include statistical measures, signal processing techniques, and domain-specific descriptors. While these methods provide interpretability, they may lack adaptability to diverse datasets and may require expert domain knowledge for effective design.

1) Statistical Measures: Numerous statistical methods are used to represent 2D data into 1D vector. Table II lists some of the most popular statistical methods and their equations [24].

B. 2D Approaches

1) Signal Processing: A spectrum of mathematical methods is employed for data representation, each tailored to reveal specific nuances within datasets. The Fourier Transform decomposes signals into sinusoidal functions, elucidating their inherent frequency components. Short-Time Fourier Transform (STFT) [66] refines this analysis by incorporating temporal considerations. Discrete Wavelet Transform (DWT) [67] excels in extracting both frequency and temporal information, particularly suited for non-stationary signals. Principal Component Analysis (PCA) [68] serves as potent dimensionality reduction techniques, unveiling fundamental structures within data. T-distributed Stochastic Neighbor Embedding (t-SNE) [69] crafts lower-dimensional representations for exploratory data analysis. Non-negative Matrix Factorization (NMF) [70] identifies parts-based features, while Kernel Principal Component Analysis (KPCA) [71] extends PCA to nonlinear relationships. Mel-Frequency Cepstral Coefficients (MFCC) [72] and the Gabor Transform [73] are specialized techniques adept at capturing frequency content in sound signals and localized components in signals, respectively. Continuous Wavelet Transform (CWT) [74] stands out for its simultaneous analysis of signals in both time and frequency domains, offering insights into the time-varying frequency components of a signal. Collectively, these methods comprise a comprehensive toolkit for researchers and practitioners aiming to comprehend diverse data structures across various domains.

- Fourier Transform: It is widely used to transform time-domain signals into their frequency components, helping in the analysis of the frequency characteristics of the signals. The FT is particularly useful for stationary signals [75]–[80].
- Wavelet Transform (WT): This technique provides time and frequency information simultaneously, making it suitable for analyzing non-stationary signals. It has been widely used in HAR to extract features from raw data due to its ability to capture both high-frequency and low-frequency components of a signal [81]–[87].
- Principal Component Analysis (PCA): PCA is a statistical technique used to reduce the dimensionality of the dataset while retaining most of the variability in the data. It is

TABLE II: COMPREHENSIVE FEATURE EXTRACTION TECHNIQUES FOR TIME-SERIES DATA USED IN HUMAN ACTIVITY RECOGNITION (HAR)

Method	**Sub-Method**	**Equation/Description**
Statistical Features	Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$
	Median	$median(S_1, S_2, ..S_n)$
	Minimum	$min(S_1, S_2, ..S_n)$
	Maximum	$max(S_1, S_2, ..S_n)$
	Coefficients of Variation	σ/μ
	Percentile	$percentile(S,p) = (1-f)S_k + f_{k+1}$
	Peak-to-Peak Amplitude	$max(S) - \min(S)$
	Interquartile Range (IQR)	$Q_3 - Q_1$, difference between the third and first quartiles, percentile ($S_{,75}$) – percentile ($S_{,25}$)
	Median Crossings	Number of times the signal crosses its median
	Skewness	$\frac{1}{n\sigma^3}\sum_{i=1}^{n}(S_i - \mu_s)^3$
	Kurtosis	$\frac{1}{n\sigma^4}\sum_{i=1}^{n}(S_i - \mu_s)^4$
	Signal Power	$\sum_{i=1}^{n} S_i^2$
	Root Mean Square (RMS)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} S_i^2}$
	Peak Intensity	*The number of signal peakks within a certain period of time*
	Pearson's Correlation Coefficient	$\frac{cov(X,Y)}{\sigma_x \sigma_y}$ *where cov is covariance*
	Inter-axis Cross-correlation	$\frac{\sum_{i=1}^{n}(a_i - \mu_a)(b_i - \mu_b)}{\sqrt{\sum_{i=1}^{n}(a_i - \mu_a)^2 \sum_{i=1}^{n}(b_i - \mu_b)^2}}$
Time-domain Features	Autocorrelation	$R(k) = \frac{1}{(n-k)\sigma^2}\sum_{i=1}^{n}(S_i - \mu)(S_{i+k} - \mu)\sqrt{k}$
	Trapezoidal Numerical Integration	*Approximation of the integral of a function using trapezoids*
	Signal Magnitude Area	$\frac{1}{n}\sum_{i=1}^{n}(\lvert x_i\rvert + \lvert y_i\rvert + \lvert z_i\rvert)$
	Signal Vector Magnitude	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i^2 + y_i^2 + z_i^2)}$
Frequency-domain Features	Power Spectral Density	$\frac{1}{n-1}\sum_{i=1}^{n}\left(S_i \cos\left(\frac{2\pi f_i}{n}\right)\right)^2 + \left(S_i \sin\left(\frac{2\pi f_i}{n}\right)\right)^2$
	Fourier Transform	$\int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} dt$
Time-frequency Features	Continuous Wavelet Transform (CWT)	$\frac{1}{\sqrt{\lvert s\rvert}}\int_{-\infty}^{\infty} \mathcal{X}(t)\psi * \left(\frac{t-\tau}{s}\right) dt$
Time-domain Features	Zero-Crossing Rate	$\frac{1}{T-1}\sum_{i=1}^{T-1}\lvert sgn(x_i - 1)\rvert$
Entropy-based Features	Approximate Entropy	$\phi^m(r) - \phi^{m-1}(r)$
	Sample Entropy	$-ln\frac{A}{B}$
	Spectral Entropy	$-\sum_i P(f_i) ln P(f_i)$
Others	Pitch Angle	$arctan\left(\frac{x_1}{\sqrt{y_1^2 + z_1^2}}\right)$
	Roll Angle	$arctan\left(\frac{y_1}{\sqrt{x_1^2 + z_1^2}}\right)$

often used in HAR to reduce the complexity of the data before classification [88]–[90].

- Sensor Fusion and Feature Fusion: Combining data from multiple sensors or the features extracted from these data using techniques like concatenation or more complex fusion methods can enhance the recognition performance by providing a more comprehensive view of the activities [91]–[95].

2) Domain-Specific Descriptors: For skeletal data, joint angles or distances between joints can be considered as domain-specific descriptors.

C. Deep Learning Feature Extraction

In recent years, deep learning has emerged as a powerful paradigm for automatic feature learning. Deep neural networks, particularly convolutional neural networks (CNNs) for sensor data and recurrent neural networks (RNNs) for temporal sequences, have demonstrated the ability to automatically extract hierarchical and abstract features from raw input. This end-to-end learning approach eliminates the need for manual feature engineering and can adapt to different data distributions. However, deep learning models often require large labeled datasets for training and may be computationally intensive.

1) Convolutional Neural Networks (CNNs): : For image-based data such as video frames, CNNs can automatically learn hierarchical features, recognizing patterns and objects relevant to human activities.

2) Recurrent Neural Networks (RNNs): : Applied to sequential data, RNNs capture temporal dependencies. In the context of human activity recognition, they might capture the order and timing of movements.

VI. LEARNING TECHNIQUES

In the realm of recognizing human movements during physical rehabilitation exercises, our exploration of learning methods has evolved over time. This section takes a step-by-step look at these methods, each representing a different stage in how we understand and categorize actions accurately. Starting with traditional machine learning, where features were manually crafted, we move through the transformative phases of deep learning, transfer learning, and transformer-based learning. Eventually, we explore ensemble learning, where combining various models can achieve better results. This journey not only traces the historical progression but also highlights the adaptive strategies developed to tackle challenges in the nuanced field of physical rehabilitation.

A. Machine Learning

Machine Learning (ML) represents a transformative approach in the field of computer science, leveraging algorithms and statistical models to enable machines to improve at tasks through experience. At its core, ML employs data-driven techniques to automate predictive analysis and decision-making processes, bypassing the need for explicit programming for every individual task. This approach encompasses a broad spectrum of algorithms, including supervised learning, where models are trained on labeled datasets; unsupervised learning, which discovers hidden patterns in data without pre-assigned labels; and reinforcement learning, where an agent learns to make decisions by performing actions and receiving feedback in a dynamic environment.

The versatility of ML methodologies has found applications across numerous domains such as healthcare, where it aids in diagnosing diseases and personalizing treatment plans; finance, for predicting market trends and managing risks; and autonomous vehicles, by enabling them to perceive their surroundings and make safe navigation decisions. This paradigm shift towards data-centric computing has not only accelerated technological advancements but also posed new challenges and ethical considerations, particularly concerning data privacy, model interpretability, and the potential for automated decision-making systems to perpetuate biases. Nonetheless, the continuous evolution of ML techniques, coupled with growing computational power and data availability, promises to further expand its capabilities and societal impact.

Utilizing features taken from Human Activity Recognition (HAR) datasets, a variety of classification algorithms, including logistic regression (LR) [96]–[99], support vector machine (SVM) [100]–[107], K-nearest neighbors (KNN) [108]–[112], random forest (RF) [113]–[116], Extra tree [45], XGBoost [117]–[120], and classifier stacking [121], [122], have been applied to activity classification. In Fang's study [123], LR, SVM, and KNN, employing manual feature extraction, demonstrated notable performance, with KNN achieving acceptable accuracies of in recognizing seven daily living activities (ADL).

In the trajectory of action classification for physical rehabilitation, the sequential exploration of learning paradigms unfolds a compelling narrative. Traditionally, machine learning techniques have laid the groundwork, employing handcrafted features to interpret sensor data intricacies. Taylor et. al investigated various machine learning models for human activity recognition including KNN, Random Forrest and Support Vector Machine [112].

B. Deep Learning

However, the paradigm shift to deep learning marks a revolutionary leap. Neural network architectures, such as Convolutional Neural Networks (CNNs) [9]–[11] designed for image-based or 1-dimensional data, and Recurrent Neural Networks (RNNs) utilized for sequential data [124], [125], inherently capture hierarchical features and temporal dependencies. This eliminates the necessity for explicit feature engineering. Notable recurrent models include Long-Short Term Memory (LSTM) [12], [13], [126], Bidirectional LSTM (BiLSTM) [127]–[129], and the combined architecture of CNN-LSTM [130], [131]. These advancements collectively contribute to the automated extraction of relevant features and dependencies in diverse data modalities.

C. Transfer Learning

Transitioning seamlessly, transfer learning emerges as a pragmatic strategy when confronted with limited labeled data. This approach leverages pre-trained models, originating from extensive datasets in general action recognition or computer vision, and fine-tunes them for specific rehabilitation tasks [132]. Capitalizing on broader domain knowledge, transfer learning enhances performance in scenarios characterized by constrained labeled data. [133]–[135] Various Transfer Learning techniques were applied to the domain of human activity recognition, encompassing notable architectures such as VGG [136], ResNet, MobileNet, Xception, Inception, and DenseNet [74]. These diverse approaches illuminate the versatility and efficacy of transfer learning in enhancing the understanding of human activities across different models.

D. Transformer

Building on this foundation, transformer-based learning, inspired by transformer architectures, introduces attention mechanisms that selectively focus on pertinent segments of input data [137]. This innovative approach has exhibited considerable success in capturing dependencies within sequences, rendering it especially well-suited for unraveling the intricacies of rehabilitation movements. Noteworthy models encompass Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT), along with Vision Transformer and Graph Transformer [138]. Additionally, ST Dynamic Graph Attention [139] and ST-GCN [140] exemplify the versatility and efficacy of this methodology in modeling complex dynamics within the realm of rehabilitation movements.

E. Ensemble Learning

Also Known as Late Fusion, Ensemble Models in machine learning represent a robust approach that combines the predictions from multiple learning algorithms to make more accurate predictions than any individual model could. This methodology is founded on the principle that a group of weak learners can come together to form a strong learner, thereby enhancing the performance of models on complex datasets. Ensemble techniques such as Bagging (Bootstrap Aggregating), Boosting, and Stacking are among the most popular methods employed to aggregate the predictions of several base estimators built with a given learning algorithm, aiming to improve stability, reduce variance, and increase prediction accuracy.

The strength of ensemble models lies in their versatility and efficacy across a wide range of applications, from winning

TABLE III
LIMITATIONS OF DIFFERENT LEARNING TECHNIQUES

Learning Technique	Advantages	Challenges	Train Time	Testing Time	Accuracy	Complexity	Model Footprint
Machine Learning	Interpretable models, require less data	Limited in handling complex patterns, requires additional feature engineering processing	Moderate	Low	Moderate	Low to Medium	Low to Medium
Deep Learning	Excellent at learning complex patterns, Can achieve high accuracy	Requires large amounts of data	High	Moderate	High	High	High
Transfer Learning	Leverages pre-trained models, Faster training times, Can be effective with limited data	Performance depends on similarity of tasks	Low to Moderate	Low to Moderate	Moderate	Low to Medium	Low to Medium
Attention Models	Focuses on relevant parts of input sequences	computationally intensive	High	High	High	High	High
Ensemble Learning	Combines multiple models for improved accuracy, Reduces variance and over-fitting	Can be complex to implement and interpret, choosing the appropriate voting strategy is challenging	These factors varies depending on the models chosen for ensemble				

Kaggle competitions to critical uses in banking for credit scoring and fraud detection, in healthcare for disease prediction and patient diagnosis, and in e-commerce for recommendation systems. For instance, Random Forests, an ensemble of decision trees, is renowned for its superior classification performance, achieved by averaging the predictions of numerous decision trees trained on different parts of the same training set. Similarly, Gradient Boosting Machines (GBMs) iteratively correct the mistakes of weak learners to improve accuracy. Stacking models, on the other hand, learn to combine the predictions of several other models, thereby leveraging their distinct strengths.

Despite their advantages, ensemble models also pose challenges, such as increased computational complexity and the risk of overfitting, particularly with models like boosting that iteratively focus on hard to classify instances. Moreover, the interpretability of ensemble models can be lower compared to that of single models due to their complex nature. Nonetheless, with careful tuning and the appropriate selection of base models, ensemble methods continue to be a powerful tool in the machine learning toolkit, offering unmatched accuracy and robustness across a diverse array of tasks and domains.

Ensemble learning, explored comprehensively by Zhang et al. [141], integrates data fusion, modeling, and mining in a unified framework. Effective ensemble methods carefully combine members to enhance performance, avoiding haphazard fusion pitfalls. In classification tasks, these methods are categorised into data-level, feature-level, decision-level, and model-level approaches. Chenguang et al. [142] utilize ensemble learning for hand function assessment, while Chihiro et al. [143] apply it for predicting functional outcomes after spinal cord injury. Additionally, Wenchuan et al. [144] leverage ensemble learning for personalized remote training in Parkinson's disease patients. Yu et al. [145] proposed an ensemble-based framework called EGCN which demonstrated robust performance on both UI-PRMD and KIMORE datasets in the realm of skeleton-based rehabilitation.

Table III presents a comparative analysis of various learning techniques, namely machine learning, deep learning, transfer learning, attention models, and ensemble learning, across several key factors. These factors encompass the advantages and challenges associated with each technique, along with considerations such as training time, testing time, model footprint (which is particularly crucial for deployment on resource-constrained devices), and complexity, which includes aspects like model architecture and hyper-parameter tuning.

The insights presented in this table are derived from seminal works in the field [132], [146]–[149]. In summary, each learning technique exhibits distinct advantages and limitations. Classical machine learning, for instance, offers interpretability and can operate with less data but may struggle with intricate patterns. In contrast, deep learning excels at discerning complex patterns but demands substantial data and computational resources. Transfer learning capitalizes on pre-existing models for new tasks, yet its effectiveness hinges on the similarity between tasks. Attention models are adept at identifying pertinent segments within input sequences, albeit at a potentially higher computational cost. Ensemble learning, which combines multiple models to enhance accuracy, presents challenges in terms of implementation and interpretation complexity. The selection of a suitable technique hinges on the specific requirements and constraints inherent to the given task.

VII. CHALLENGES AND FUTURE DIRECTIONS

The intersection of Human-Computer Interaction (HCI) and physical rehabilitation is a fertile ground for innovation, yet it is not without its challenges and issues. Addressing these will be vital for the continued advancement and integration of these technologies into mainstream healthcare. This section highlights the primary challenges in the domain and outlines potential avenues for future research:

Lack of large-scale annotated datasets presents a challenge for developing robust multi-person detection and tracking systems, especially for communal rehabilitation settings. Current systems like MediaPipe are limited to single-person detection, highlighting the need for scalable multi-person tracking solutions that maintain real-time performance. Frame rate optimization is another concern, as solutions capable of multi-person tracking often sacrifice frame rates, impacting real-time applicability. Enhancing processing speed without compromising accuracy remains an ongoing challenge. Additionally, there is a cost vs. accessibility trade-off, where technologies like the Kinect camera offer sophisticated tracking capabilities but add to the cost barrier. Future research should focus on developing cost-effective solutions that do not sacrifice functionality. Data privacy and security are paramount, requiring robust security protocols to protect patient information. Sensor fusion and data integration from diverse sources pose technical hurdles, requiring seamless blending of data streams while ensuring reliability and interpretability. Algorithm performance is critical, as it affects classification results, real-time usability, and compatibility with low-power devices. Algorithms must balance accuracy with speed to provide real-time feedback in rehabilitation exercises. User-centered design and usability are essential for adoption, requiring technologies to be user-friendly and tailored to the needs of patients and therapists. Finally, clinical validation and interdisciplinary collaboration are crucial, with long-term studies needed to validate efficacy and safety, aligning with regulatory frameworks for clinical adoption. Collaboration across disciplines is necessary to address these multifaceted challenges in physical rehabilitation technology. Table IV summarize the challenges in this domain

TABLE IV
CHALLENGES

Limitation	Reference
Lack of large-scale annotated datasets	[23], [150]–[152]
Multi-Person Detection and Tracking	[40], [45]
Frame Rate Optimization	[153]–[156]
Cost vs. Accessibility Trade-off	[45], [157]–[161]
Algorithmic trade-off	[45], [157]
Data Privacy and Security	[162]–[166]
Sensor Fusion and Data Integration	[167]–[171]
User-Centered Design and Usability	[172]–[174]
Clinical Validation and Regulatory Compliance	[175]–[178]

In addressing these issues, researchers have the opportunity to significantly impact the field of physical rehabilitation. The development of sophisticated, accessible, and user-centric HCI solutions has the potential to enhance patient outcomes and reshape the landscape of rehabilitative care. It is through the lens of these challenges that future research can find its direction, ensuring that advances in the domain are both technologically sound and aligned with the holistic needs of patients.

VIII. Discussion

This survey provides a comprehensive overview of current HCI techniques in physical rehabilitation, emphasizing virtual and direct methodologies. These technologies offer potential to revolutionize physical therapy by overcoming traditional limitations and meeting the demand for personalized, remote services.

In data acquisition, sensor-based methods offer high precision but can be intrusive, while camera-based approaches are non-intrusive but raise privacy and environmental challenges. Benchmarking datasets are crucial for development but need to better represent diverse patient populations.

Feature extraction techniques, including 1D statistical methods, 2D signal approaches, and deep learning methods like CNNs and RNNs, play a critical role. Deep learning shows promise but requires substantial resources and large datasets.

Various learning techniques, such as machine learning, ensemble methods, deep learning, transfer learning, and transformer models, offer unique advantages. However, challenges remain in ensuring model generalizability, real-world validation, and addressing ethical considerations like data privacy.

Future research should focus on interdisciplinary collaboration to address these challenges and advance HCI in rehabilitation for improved patient care worldwide.

IX. Conclusion

In conclusion, the integration of Human-Computer Interaction (HCI) in physical rehabilitation offers a promising avenue to augment traditional therapeutic methods, making rehabilitation more accessible, engaging, and efficient. This survey has highlighted significant advancements in virtual and direct rehabilitation, the nuances of data acquisition methods, the importance of benchmarking datasets, and the evolution of feature extraction and learning techniques. While challenges such as multi-person detection, real-time processing, cost considerations, and ethical concerns persist, they also chart the course for future research. The potential of HCI to revolutionize physical rehabilitation is immense, contingent upon continuous innovation and interdisciplinary collaboration. As we move forward, the focus must remain on creating inclusive, patient-centric solutions that not only embrace technological advancements but also uphold the highest standards of care and ethics. The journey of HCI in rehabilitation is just beginning, and its trajectory promises to reshape the future of physical therapy and patient care.

References

[1] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, and J. S. Suri, "Human activity recognition in artificial intelligence framework: A narrative review," *Artificial Intelligence Review*, vol. 55, pp. 4755–4808, August 2022.

[2] X. Zhou, W. Liang, I. Kevin, K. Wang, H. Wang, L. T. Yang, and Q. Jin, "Deep-learning-enhanced human activity recognition for internet of healthcare things," *IEEE Internet of Things Journal*, vol. 7, pp. 6429–6438, April 2020.

[3] H. F. Nweke, Y. W. Teh, G. Mujtaba, and M. A. Al-Garadi, "Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions," *Information Fusion*, vol. 46, pp. 147–170, March 2019.

[4] S. Qiu, H. Zhao, N. Jiang, Z. Wang, L. Liu, Y. An, H. Zhao, X. Miao, R. Liu, and G. Fortino, "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," *Information Fusion*, vol. 80, pp. 241–265, April 2022.

[5] W. H. Organization, "Rehabilitation," 2021.

[6] K. Witkiewitz, K. S. Montes, F. J. Schwebel, and J. A. Tucker, "What is recovery?," *Alcohol Research: Current Reviews*, vol. 40, no. 3, 2020.
[7] B. Debnath, M. O'brien, M. Yamaguchi, and A. Behera, "A review of computer vision-based approaches for physical rehabilitation and assessment," *Multimedia Systems*, vol. 28, no. 1, pp. 209–239, 2022.
[8] A. A. Khan, A. Laghari, and S. A. Awan, "Machine learning in computer vision: a review," *EAI Endorsed Transactions on Scalable Information Systems*, vol. 8, pp. e4–, April 2021.
[9] J. Yang, M. N. Nguyen, Y. Guo, and Y. Li, "Deep learning for human activity recognition with smartphones," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 12, pp. 1107–1118, 2015.
[10] Z. Li, H. Zhu, Y. Liu, and L. Wang, "A multi-modal convolutional neural network for human activity recognition using smartphone sensors and wearable sensors," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 8, pp. 8994–9004, 2022.
[11] K. He, X. Zhang, S. Ren, and J. Sun, "Convolutional neural networks for human activity recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2005–2015, 2016.
[12] Y. Chen, K. Zhong, J. Zhang, Q. Sun, and X. Zhao, "Lstm networks for mobile human activity recognition," in *2016 International conference on artificial intelligence: technologies and applications*, pp. 50–53, Atlantis Press, January 2016.
[13] S. Mekruksavanich and A. Jitpattanakul, "Lstm networks using smartphone data for sensor-based human activity recognition in smart homes," *Sensors*, vol. 21, p. 1636, February 2021.
[14] D. Li, Z. Yang, P. Kang, and X. Xie, "Home-based compared with hospital-based rehabilitation program for patients undergoing total knee arthroplasty for osteoarthritis: a systematic review and meta-analysis of randomized controlled trials," *American journal of physical medicine & rehabilitation*, vol. 96, pp. 440–447, Jun 2017.
[15] G. Saposnik, M. Levin, G. PT., *et al.*, "Designing personalized exercise interventions for individuals with disabilities: A human-centered approach," *Disability and Rehabilitation*, vol. 39, no. 16, pp. 1626–1634, 2017.
[16] J. Koivisto and A. Malik, "Gamification for older adults: a systematic literature review," *The Gerontologist*, vol. 61, pp. e360–e372, Oct 2021.
[17] Y. Chen, K. Abel, J. Janecek, Y. Chen, K. Zheng, and S. Cramer, "Home-based technologies for stroke rehabilitation: A systematic review," *International journal of medical informatics*, vol. 123, pp. 11–22, Mar 2019.
[18] B. Cunha, R. Ferreira, and A. S. Sousa, "Home-based rehabilitation of the shoulder using auxiliary systems and artificial intelligence: an overview," *Sensors*, vol. 23, no. 16, p. 7100, 2023.
[19] A. Arntz, F. Weber, M. Handgraaf, K. La¨lla¨, K. Korniloff, K.-P. Murtonen, J. Chichaeva, A. Kidritsch, M. Heller, E. Sakellari, *et al.*, "Technologies in home-based digital rehabilitation: scoping review," *JMIR rehabilitation and assistive technologies*, vol. 10, p. e43615, 2023.
[20] H. Loellgen, P. Zupet, N. Bachl, and A. Debruyne, "Physical activity, exercise prescription for health and home-based rehabilitation," *Sustainability*, vol. 12, no. 24, p. 10230, 2020.
[21] E. Fascio, J. A. Vitale, P. Sirtori, G. Peretti, G. Banfi, and L. Mangiavini, "Early virtual-reality-based home rehabilitation after total hip arthroplasty: a randomized controlled trial," *Journal of Clinical Medicine*, vol. 11, no. 7, p. 1766, 2022.
[22] F. Correia, A. Nogueira, I. Magalha˜es, J. Guimara˜es, M. Moreira, I. Barradas, L. Teixeira, J. Tulha, R. Seabra, J. Lains, and V. Bento, "Home-based rehabilitation with a novel digital biofeedback system versus conventional in-person rehabilitation after total knee replacement: a feasibility study," *Sci Rep*, vol. 8, no. 1, p. 11299, 2018.
[23] Y. Liao, A. Vakanski, M. Xian, D. Paul, and R. Baker, "A review of computational approaches for evaluation of rehabilitation exercises," *Computers in biology and medicine*, vol. 119, p. 103687, 2020.
[24] S. O. Slim, A. Atia, M. M. Elfattah, and M.-S. M. Mostafa, "Survey on human activity recognition based on acceleration data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 3, 2019.
[25] E. Fridriksdottir and A. G. Bonomi, "Accelerometer-based human activity recognition for patient monitoring using a deep neural network," *Sensors*, vol. 20, no. 22, p. 6424, 2020.
[26] M. M. Baig, S. Afifi, H. GholamHosseini, and F. Mirza, "A systematic review of wearable sensors and iot-based monitoring applications for older adults–a focus on ageing population and independent living," *Journal of medical systems*, vol. 43, pp. 1–11, 2019.
[27] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey," *IEEE access*, vol. 8, pp. 210816–210836, 2020.
[28] M. Straczkiewicz, P. James, and J.-P. Onnela, "A systematic review of smartphone-based human activity recognition methods for health research," *NPJ Digital Medicine*, vol. 4, no. 1, p. 148, 2021.
[29] A. Barna, A. K. M. Masum, M. E. Hossain, E. H. Bahadur, and M. S. Alam, "A study on human activity recognition using gyroscope, accelerometer, temperature and humidity data," in *2019 international conference on electrical, computer and communication engineering (ecce)*, pp. 1–6, IEEE, 2019.
[30] D. Gholamiangonabadi and K. Grolinger, "Personalized models for human activity recognition with wearable sensors: deep neural networks and signal processing," *Applied Intelligence*, vol. 53, no. 5, pp. 6041–6061, 2023.
[31] S. Farrokhi, W. Dargie, and C. Poellabauer, "Human activity recognition based on wireless electrocardiogram and inertial sensors," *IEEE Sensors Journal*, 2024.
[32] W. Jiang and Z. Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks," in *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 1307–1310, 2015.
[33] K. Nurhanim, I. Elamvazuthi, L. Izhar, G. Capi, and S. Su, "Emg signals classification on human activity recognition using machine learning algorithm," in *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*, pp. 369–373, IEEE, 2021.
[34] T. Paul, C. Bhattacharyya, P. Sen, R. Prasad, S. Shaw, and S. Das, "Human emotion recognition using gsr and eeg," *International Journal of Scientific and Research Publication*, vol. 10, no. 5, pp. 394–400, 2020.
[35] C. Mundell, J. P. Vielma, and T. Zaman, "Predicting performance under stressful conditions using galvanic skin response," *arXiv preprint arXiv:1606.01836*, 2016.
[36] I. da Penha Natal, L. Correia, A. C. Garcia, and L. Fernandes, "Efficient out-of-home activity recognition by complementing gps data with semantic information," *First Monday*, 2019.
[37] V. Dentamaro, V. Gattulli, D. Impedovo, and F. Manca, "Human activity recognition with smartphone-integrated sensors: A survey," *Expert Systems with Applications*, p. 123143, 2024.
[38] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, pp. 172–186, Jan. 2021.
[39] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "BlazePose: On-device real-time body pose tracking," June 2020.
[40] S. Mroz, N. Baddour, C. McGuirk, P. Juneau, A. Tu, K. Cheung, and E. Lemaire, "Comparing the quality of human pose estimation with blazepose or openpose," in *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, pp. 1–4, 2021.
[41] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann, "MediaPipe: A framework for building perception pipelines," June 2019.
[42] P. Wang, W. Li, P. Ogunbona, J. Wan, and S. Escalera, "Rgb-d-based human motion recognition with deep learning: A survey," *Computer vision and image understanding*, vol. 171, pp. 118–139, 2018.
[43] A. Franco, A. Magnani, and D. Maio, "A multimodal approach for human activity recognition based on skeleton and rgb data," *Pattern Recognition Letters*, vol. 131, pp. 293–299, 2020.
[44] P. Khaire and P. Kumar, "Deep learning and rgb-d based human action, human–human and human–object interaction recognition: A survey," *Journal of Visual Communication and Image Representation*, vol. 86, p. 103531, 2022.
[45] M. Zaher, A. Samir, A. Ghoneim, L. Abdelhamid, and A. Atia, "A framework for assessing physical rehabilitation exercises," in *2023 Intelligent Methods, Systems, and Applications (IMSA)*, pp. 526–532, 2023.
[46] A. L. Paredes, Q. Song, and M. H. Conde, "Performance evaluation of state-of-the-art high-resolution time-of-flight cameras," *IEEE Sensors Journal*, 2023.
[47] D. K. Vishwakarma and K. Jain, "Three-dimensional human activity recognition by forming a movement polygon using posture skeletal data from depth sensor," *ETRI Journal*, vol. 44, no. 2, pp. 286–299, 2022.
[48] S. Almasi, H. Ahmadi, F. Asadi, L. Shahmoradi, G. Arji, M. Alizadeh, H. Kolivand, *et al.*, "Kinect-based rehabilitation systems for stroke patients: A scoping review," *BioMed research international*, vol. 2022, 2022.
[49] M. Mangone, E. Marinelli, G. Santilli, N. Finanore, F. Agostini, V. Santilli, A. Bernetti, M. Paoloni, S. Zaami, *et al.*, "Gait analysis

advancements: rehabilitation value and new perspectives from forensic application," *European Review for Medical and Pharmacological Sciences*, vol. 27, no. 1, pp. 3–12, 2023.

[50] A. Lubkowska and W. Pluta, "Infrared thermography as a non-invasive tool in musculoskeletal disease rehabilitation—the control variables in applicability—a systematic review," *Applied Sciences*, vol. 12, no. 9, p. 4302, 2022.

[51] L. Li and A. Vakanski, "Generative adversarial networks for generation and classification of physical rehabilitation movement episodes," *International journal of machine learning and computing*, vol. 8, no. 5, p. 428, 2018.

[52] A. Vakanski, H.-P. Jun, D. Paul, and R. Baker, "A data set of human body movements for physical rehabilitation exercises," *Data (Basel)*, vol. 3, Mar. 2018.

[53] A. Paiement, L. Tao, S. Hannuna, M. Camplani, D. Damen, and M. Mirmehdi, "Online quality assessment of human movement from skeleton data," in *British Machine Vision Conference*, pp. 153–166, BMVA press, 2014.

[54] L. Tao, A. Paiement, D. Damen, M. Mirmehdi, S. Hannuna, M. Camplani, T. Burghardt, and I. Craddock, "A comparative study of pose representation and dynamics modelling for online motion quality assessment," *Computer vision and image understanding*, vol. 148, pp. 136–152, 2016.

[55] L. Tao, A. Paiement, D. Damen, M. Mirmehdi, S. Hannuna, M. Camplani, T. Burghardt, and I. Craddock, "A comparative study of pose representation and dynamics modelling for online motion quality assessment," *Computer vision and image understanding*, vol. 148, pp. 136–152, 2016.

[56] E. Dolatabadi, Y. X. Zhi, B. Ye, M. Coahran, G. Lupinacci, A. Mihailidis, R. Wang, and B. Taati, "The toronto rehab stroke pose dataset to detect compensation during stroke rehabilitation therapy," in *Proceedings of the 11th EAI international conference on pervasive computing technologies for healthcare*, pp. 375–381, 2017.

[57] M. H. Li, T. A. Mestre, S. H. Fox, and B. Taati, "Vision-based assessment of parkinsonism and levodopa-induced dyskinesia with pose estimation," *Journal of neuroengineering and rehabilitation*, vol. 15, no. 1, pp. 1–13, 2018.

[58] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 4724–4732, 2016.

[59] M. Capecci, M. G. Ceravolo, F. Ferracuti, S. Iarlori, A. Monteriu, L. Romeo, and F. Verdini, "The kimore dataset: Kinematic assessment of movement and clinical scores for remote monitoring of physical rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 7, pp. 1436–1448, 2019.

[60] J. Antunes, A. Bernardino, A. Smailagic, and D. P. Siewiorek, "Aha-3d: A labelled dataset for senior fitness exercise recognition and segmentation from 3d skeletal data.," in *BMVC*, p. 332, 2018.

[61] 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 *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 168–172, IEEE, 2015.

[62] T. Li, J. Liu, W. Zhang, Y. Ni, W. Wang, and Z. Li, "Uav-human: A large benchmark for human behavior understanding with unmanned aerial vehicles," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16266–16275, 2021.

[63] A. Logacjov, K. Bach, A. Kongsvold, H. B. Ba˚rdstu, and P. J. Mork, "Harth: a human activity recognition dataset for machine learning," *Sensors*, vol. 21, no. 23, p. 7853, 2021.

[64] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *2012 16th international symposium on wearable computers*, pp. 108–109, IEEE, 2012.

[65] B. Blunck, K. Prentow, and Anind, "Heterogeneity Activity Recognition." UCI Machine Learning Repository, 2015. DOI: https://doi.org/10.24432/C5689X.

[66] C. Ding, Y. Jia, G. Cui, C. Chen, X. Zhong, and Y. Guo, "Continuous human activity recognition through parallelism lstm with multi-frequency spectrograms," *Remote Sensing*, vol. 13, no. 21, p. 4264, 2021.

[67] A. Dahou, M. A. Al-qaness, M. Abd Elaziz, and A. M. Helmi, "Mlcnnwav: Multi-level convolutional neural network with wavelet transformations for sensor-based human activity recognition," *IEEE Internet of Things Journal*, 2023.

[68] S. Balli, E. A. Sag˘bas¸, and M. Peker, "Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm," *Measurement and Control*, vol. 52, no. 1-2, pp. 37–45, 2019.

[69] F. H. M. Oliveira, A. R. Machado, A. O. Andrade, *et al.*, "On the use of t-distributed stochastic neighbor embedding for data visualization and classification of individuals with parkinson's disease," *Computational and mathematical methods in medicine*, vol. 2018, 2018.

[70] A. Eweiwi, M. S. Cheema, and C. Bauckhage, "Discriminative joint non-negative matrix factorization for human action classification," in *German Conference on Pattern Recognition*, pp. 61–70, Springer, 2013.

[71] D. Wu, H. Zhang, C. Niu, J. Ren, and W. Zhao, "Inertial sensor based human activity recognition via reduced kernel pca," in *Advances in Body Area Networks I: Post-Conference Proceedings of BodyNets 2017*, pp. 447–456, Springer, 2019.

[72] F. Cruciani, A. Vafeiadis, C. Nugent, I. Cleland, P. McCullagh, K. Votis, D. Giakoumis, D. Tzovaras, L. Chen, and R. Hamzaoui, "Comparing cnn and human crafted features for human activity recognition," in *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 960–967, IEEE, 2019.

[73] D. Vishwakarma, P. Rawat, and R. Kapoor, "Human activity recognition using gabor wavelet transform and ridgelet transform," *Procedia Computer Science*, vol. 57, pp. 630–636, 2015.

[74] O. Pavliuk, M. Mishchuk, and C. Strauss, "Transfer learning approach for human activity recognition based on continuous wavelet transform," *Algorithms*, vol. 16, no. 2, p. 77, 2023.

[75] A. Nedorubova, A. Kadyrova, and A. Khlyupin, "Human activity recognition using continuous wavelet transform and convolutional neural networks," *arXiv preprint arXiv:2106.12666*, 2021.

[76] A. Khtun and S. G. S. Hossain, "A fourier domain feature approach for human activity recognition & fall detection," in *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 40–45, IEEE, 2023.

[77] Z. Li, Y. Liu, B. Liu, J. Le Kernec, and S. Yang, "A holistic human activity recognition optimisation using ai techniques," *IET Radar, Sonar & Navigation*, vol. 18, no. 2, pp. 256–265, 2024.

[78] Z. Qin, Y. Zhang, S. Meng, Z. Qin, and K.-K. R. Choo, "Imaging and fusing time series for wearable sensor-based human activity recognition," *Information Fusion*, vol. 53, pp. 80–87, 2020.

[79] I. Dirgova´ Lupta´kova´, M. Kubovcˇ´ık, and J. Posp´ıchal, "Wearable sensor-based human activity recognition with transformer model," *Sensors*, vol. 22, no. 5, p. 1911, 2022.

[80] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper, "Sensing technology for human activity recognition: A comprehensive survey," *Ieee Access*, vol. 8, pp. 83791–83820, 2020.

[81] C. Ieracitano, N. Mammone, A. Hussain, and F. C. Morabito, "A novel multi-modal machine learning based approach for automatic classification of eeg recordings in dementia," *Neural Networks*, vol. 123, pp. 176–190, 2020.

[82] P. Jadhav and S. Mukhopadhyay, "Automated sleep stage scoring using time-frequency spectra convolution neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022.

[83] L. Jalal and A. Peer, "Emotion recognition from physiological signals using continuous wavelet transform and deep learning," in *International conference on human-computer interaction*, pp. 88–99, Springer, 2022.

[84] A. Sarkar, S. S. Hossain, and R. Sarkar, "Human activity recognition from sensor data using spatial attention-aided cnn with genetic algorithm," *Neural Computing and Applications*, vol. 35, no. 7, pp. 5165–5191, 2023.

[85] T. H. Vuong, T. Doan, and A. Takasu, "Deep wavelet convolutional neural networks for multimodal human activity recognition using wearable inertial sensors," *Sensors*, vol. 23, no. 24, p. 9721, 2023.

[86] F. Abuhoureyah, Y. C. Wong, A. S. M. Isira, and J. H. Chuah, "Addressing location dependency in human activity recognition using channel state information via 3d-cwt approach," in *2023 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 212–217, IEEE, 2023.

[87] F. S. Butt, L. La Blunda, M. F. Wagner, J. Scha¨fer, I. Medina-Bulo, and D. Go´mez-Ullate, "Fall detection from electrocardiogram (ecg) signals and classification by deep transfer learning," *Information*, vol. 12, no. 2, p. 63, 2021.

[88] Y. Zhao, H. Zhou, S. Lu, Y. Liu, X. An, and Q. Liu, "Human activity recognition based on non-contact radar data and improved pca method," *Applied Sciences*, vol. 12, no. 14, p. 7124, 2022.

[89] I. A. Showmik, T. F. Sanam, and H. Imtiaz, "Human activity recognition from wi-fi csi data using principal component-based wavelet cnn," *Digital Signal Processing*, vol. 138, p. 104056, 2023.

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

[91] A. Franco, A. Magnani, and D. Maio, "A multimodal approach for human activity recognition based on skeleton and rgb data," *Pattern Recognition Letters*, vol. 131, pp. 293–299, 2020.

[92] H. Zhou, Y. Zhao, Y. Liu, S. Lu, X. An, and Q. Liu, "Multi-sensor data fusion and cnn-lstm model for human activity recognition system," *Sensors*, vol. 23, no. 10, p. 4750, 2023.

[93] L. Cao, S. Liang, Z. Zhao, D. Wang, C. Fu, and K. Du, "Human activity recognition method based on fmcw radar sensor with multi-domain feature attention fusion network," *Sensors*, vol. 23, no. 11, p. 5100, 2023.

[94] W. Tao, H. Chen, M. Moniruzzaman, M. C. Leu, Z. Yi, and R. Qin, "Attention-based sensor fusion for human activity recognition using imu signals," *arXiv preprint arXiv:2112.11224*, 2021.

[95] Y. Wang, K. Yu, Y. Wang, and H. Xue, "Multi-view fusion transformer for sensor-based human activity recognition," *arXiv preprint arXiv:2202.12949*, 2022.

[96] O. M. Giggins, K. T. Sweeney, and B. Caulfield, "Rehabilitation exercise assessment using inertial sensors: a cross-sectional analytical study," *Journal of neuroengineering and rehabilitation*, vol. 11, pp. 1–10, 2014.

[97] Q. Zhang, Z. Zhang, X. Huang, C. Zhou, J. Xu, *et al.*, "Application of logistic regression and decision tree models in the prediction of activities of daily living in patients with stroke," *Neural Plasticity*, vol. 2022, 2022.

[98] W. J. Gordon-Evans, A. L. Johnson, K. E. Knap, and D. J. Griffon, "The effect of body condition on postoperative recovery of dachshunds with intervertebral disc disease treated with postoperative physical rehabilitation," *Veterinary surgery*, vol. 48, no. 2, pp. 159–163, 2019.

[99] L. X. Chen, J. J. Mao, S. Fernandes, M. L. Galantino, W. Guo, P. LaRiccia, V. L. Teal, M. A. Bowman, H. R. Schumacher, and J. T. Farrar, "Integrating acupuncture with exercise-based physical therapy for knee osteoarthritis: a randomized controlled trial," *JCR: Journal of Clinical Rheumatology*, vol. 19, no. 6, pp. 308–316, 2013.

[100] T. Hamaguchi, T. Saito, M. Suzuki, T. Ishioka, Y. Tomisawa, N. Nakaya, and M. Abo, "Support vector machine-based classifier for the assessment of finger movement of stroke patients undergoing rehabilitation," *Journal of Medical and Biological Engineering*, vol. 40, pp. 91–100, 2020.

[101] A. E. Minarno, W. A. Kusuma, and H. Wibowo, "Performance comparisson activity recognition using logistic regression and support vector machine," in *2020 3rd International conference on intelligent autonomous systems (ICoIAS)*, pp. 19–24, IEEE, 2020.

[102] A. A. Mohamed, M. I. Awad, S. A. Maged, and M. M. Gaber, "Automated upper limb motor functions assessment system using one-class support vector machine," in *2021 16th International Conference on Computer Engineering and Systems (ICCES)*, pp. 1–6, IEEE, 2021.

[103] S.-C. Yeh, M.-C. Huang, P.-C. Wang, T.-Y. Fang, M.-C. Su, P.-Y. Tsai, and A. Rizzo, "Machine learning-based assessment tool for imbalance and vestibular dysfunction with virtual reality rehabilitation system," *Computer methods and programs in biomedicine*, vol. 116, no. 3, pp. 311–318, 2014.

[104] Y. Bouteraa, I. B. Abdallah, and K. Boukthir, "A new wrist–forearm rehabilitation protocol integrating human biomechanics and svm-based machine learning for muscle fatigue estimation," *Bioengineering*, vol. 10, no. 2, p. 219, 2023.

[105] Y. Long, Z.-J. Du, W.-D. Wang, G.-Y. Zhao, G.-Q. Xu, L. He, X.-W. Mao, and W. Dong, "Pso-svm-based online locomotion mode identification for rehabilitation robotic exoskeletons," *Sensors*, vol. 16, no. 9, p. 1408, 2016.

[106] S. Cai, Y. Chen, S. Huang, Y. Wu, H. Zheng, X. Li, and L. Xie, "Svm-based classification of semg signals for upper-limb self-rehabilitation training," *Frontiers in neurorobotics*, vol. 13, p. 31, 2019.

[107] N. Alhammad and H. Al-Dossari, "Recognizing physical activities for spinal cord injury rehabilitation using wearable sensors," *Sensors*, vol. 21, no. 16, p. 5479, 2021.

[108] M. Zhu, W. Chen, J. P. Hirdes, and P. Stolee, "The k-nearest neighbor algorithm predicted rehabilitation potential better than current clinical assessment protocol," *Journal of clinical epidemiology*, vol. 60, no. 10, pp. 1015–1021, 2007.

[109] K.-H. Chen, P.-C. Chen, K.-C. Liu, and C.-T. Chan, "Wearable sensor-based rehabilitation exercise assessment for knee osteoarthritis," *Sensors*, vol. 15, no. 2, pp. 4193–4211, 2015.

[110] Q. Guo and S. S. Khan, "Exercise-specific feature extraction approach for assessing physical rehabilitation," in *4th IJCAI Workshop on AI for Aging, Rehabilitation and Intelligent Assisted Living. IJCAI*, 2021.

[111] S. Noor, M. Hasham, Z. Habib, and A. B. Saragano, "Real-time rehabilitation and fitness system using depth sensor," in *2019 International Conference on Frontiers of Information Technology (FIT)*, pp. 215–2155, IEEE, 2019.

[112] Z. Liu, S. Li, J. Hao, J. Hu, and M. Pan, "An efficient and fast model reduced kernel knn for human activity recognition," *Journal of Advanced Transportation*, vol. 2021, pp. 1–9, 2021.

[113] D. S. Wood, K. Jensen, A. Crane, H. Lee, H. Dennis, J. Gladwell, A. Shurtz, D. T. Fullwood, M. K. Seeley, U. H. Mitchell, *et al.*, "Accurate prediction of knee angles during open-chain rehabilitation exercises using a wearable array of nanocomposite stretch sensors," *Sensors*, vol. 22, no. 7, p. 2499, 2022.

[114] A. Van, V. C. Gay, P. J. Kennedy, E. Barin, and P. Leijdekkers, "Understanding risk factors in cardiac rehabilitation patients with random forests and decision trees," in *Conferences in Research and Practice in Information Technology Series*, 2010.

[115] M. Javeed, A. Jalal, and K. Kim, "Wearable sensors based exertion recognition using statistical features and random forest for physical healthcare monitoring," in *2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*, pp. 512–517, IEEE, 2021.

[116] L. Bavan, K. Surmacz, D. Beard, S. Mellon, and J. Rees, "Adherence monitoring of rehabilitation exercise with inertial sensors: A clinical validation study," *Gait & Posture*, vol. 70, pp. 211–217, 2019.

[117] W. Zhang, X. Zhao, and Z. Li, "A comprehensive study of smartphone-based indoor activity recognition via xgboost," *IEEE Access*, vol. 7, pp. 80027–80042, 2019.

[118] K. Gusain, A. Gupta, and B. Popli, "Transition-aware human activity recognition using extreme gradient boosted decision trees," in *Advanced Computing and Communication Technologies: Proceedings of the 10th ICACCT, 2016*, pp. 41–49, Springer, 2018.

[119] K. Li, R. Habre, H. Deng, R. Urman, J. Morrison, F. D. Gilliland, J. L. Ambite, D. Stripelis, Y.-Y. Chiang, Y. Lin, *et al.*, "Applying multivariate segmentation methods to human activity recognition from wearable sensors' data," *JMIR mHealth and uHealth*, vol. 7, no. 2, p. e11201, 2019.

[120] M. H. Abid, A.-A. Nahid, M. R. Islam, and M. P. Mahmud, "Human activity recognition based on wavelet-based features along with feature prioritization," in *2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA)*, pp. 933–939, IEEE, 2021.

[121] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Computer Science*, vol. 34, pp. 450–457, 2014.

[122] F. Rustam, A. A. Reshi, I. Ashraf, A. Mehmood, S. Ullah, D. M. Khan, and G. S. Choi, "Sensor-based human activity recognition using deep stacked multilayered perceptron model," *IEEE Access*, vol. 8, pp. 218898–218910, 2020.

[123] L. Fang, S. Yishui, and C. Wei, "Up and down buses activity recognition using smartphone accelerometer," in *2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference*, pp. 761–765, IEEE, 2016.

[124] P. Kumar and S. Suresh, "Deep-har: an ensemble deep learning model for recognizing the simple, complex, and heterogeneous human activities," *Multimedia Tools and Applications*, pp. 1–28, 2023.

[125] Y. Liao, A. Vakanski, and M. Xian, "A deep learning framework for assessing physical rehabilitation exercises," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 468–477, 2020.

[126] J. D. Domingo, J. Go´mez-Garc´ıa-Bermejo, and E. Zalama, "Improving human activity recognition integrating lstm with different data sources: Features, object detection and skeleton tracking," *IEEE Access*, vol. 10, pp. 68213–68230, June 2022.

[127] J. Su, Z. Liao, Z. Sheng, A.-X. Liu, D. Singh, and H.-N. Lee, "Human activity recognition using self-powered sensors based on multilayer bi-directional long short-term memory networks," *IEEE Sensors Journal*, August 2022.

[128] J. He, X. Wu, Z. Cheng, Z. Yuan, and Y. Jiang, "Db-lstm: Densely-connected bi-directional lstm for human action recognition," *Neurocomputing*, vol. 444, pp. 319–331, July 2021.

[129] S. Arif and J. Wang, "Bidirectional lstm with saliency-aware 3d-cnn features for human action recognition," *Journal of Engineering Research*, vol. 9, September 2021.

[130] A. Nouriani, R. A. McGovern, and R. Rajamani, "Deep-learning-based human activity recognition using wearable sensors," *IFAC-PapersOnLine*, vol. 55, pp. 1–6, January 2022.

[131] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "Skeleton-based human activity recognition using convlstm and guided feature learning," *Soft Computing*, pp. 1–4, Oct 2022.

[132] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[133] L. Yuan, Z. He, Q. Wang, L. Xu, and X. Ma, "Spatial transformer network with transfer learning for small-scale fine-grained skeleton-based tai chi action recognition," in *IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1–6, IEEE, 2022.

[134] I. Boukhennoufa, X. Zhai, K. McDonald-Maier, V. Utti, and J. Jackson, "Improving the activity recognition using gmaf and transfer learning in post-stroke rehabilitation assessment," in *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, pp. 391–398, IEEE, 2021.

[135] Y. Qiu, J. Wang, Z. Jin, H. Chen, M. Zhang, and L. Guo, "Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training," *Biomedical Signal Processing and Control*, vol. 72, p. 103323, 2022.

[136] P. Khaire, P. Kumar, and J. Imran, "Combining cnn streams of rgb-d and skeletal data for human activity recognition," *Pattern Recognition Letters*, vol. 115, pp. 107–116, 2018.

[137] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[138] K. Re´by, I. Dulau, G. Dubrasquet, and M. Aimar, "Graph transformer for physical rehabilitation evaluation," in *2023 IEEE 17th International Conference on Automatic Face and Gesture Recognition (FG)*, pp. 1–8, IEEE, Jan 2023.

[139] M. Rahevar, A. Ganatra, T. Saba, A. Rehman, and S. A. Bahaj, "Spatial–temporal dynamic graph attention network for skeleton-based action recognition," *IEEE Access*, vol. 11, pp. 21546–21553, February 22 2023.

[140] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[141] X. Dong, Z. Yu, W. Cao, Y. Shi, and Q. Ma, "A survey on ensemble learning," *Frontiers of Computer Science*, vol. 14, pp. 241–258, 2020.

[142] H. Y. Chenguang Li, Long Cheng and F. Huang, "An automatic rehabilitation assessment system for hand function based on leap motion and ensemble learning," *Cybernetics and Systems*, vol. 52, no. 1, pp. 3–25, 2021.

[143] C. Kato, O. Uemura, Y. Sato, and T. Tsuji, "Functional outcome prediction after spinal cord injury using ensemble machine learning," *Archives of Physical Medicine and Rehabilitation*, 2023.

[144] W. Wei, C. McElroy, and S. Dey, "Towards on-demand virtual physical therapist: Machine learning-based patient action understanding, assessment and task recommendation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 9, pp. 1824–1835, 2019.

[145] A. Bruce, Y. Chen, and Z. Liu, "Egcn: An ensemble-based learning framework for exploring effective skeleton-based rehabilitation exercise assessment," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 511–517, 2022.

[146] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[147] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[148] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[149] R. Polikar, "Ensemble learning," *Ensemble machine learning: Methods and applications*, pp. 1–34, 2012.

[150] M. A. B. Siddiq, F. A. Rathore, D. Clegg, and J. J. Rasker, "Pulmonary rehabilitation in covid-19 patients: A scoping review of current practice and its application during the pandemic," *Turkish journal of physical medicine and rehabilitation*, vol. 66, no. 4, p. 480, 2020.

[151] T. S. Jesus, M. D. Landry, D. Brooks, and H. Hoenig, "Physical rehabilitation needs per condition type: results from the global burden of disease study 2017," *Archives of physical medicine and rehabilitation*, vol. 101, no. 6, pp. 960–968, 2020.

[152] W. Zou *et al.*, "Analysis and research on the rehabilitation effect of physical exercise on college students' mental depression based on multidimensional data mining," *Occupational Therapy International*, vol. 2022, 2022.

[153] R. R. Koli and T. I. Bagban, "Human action recognition using deep neural networks," in *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, pp. 376–380, IEEE, 2020.

[154] M. A. Al-Qaness, A. M. Helmi, A. Dahou, and M. A. Elaziz, "The applications of metaheuristics for human activity recognition and fall detection using wearable sensors: A comprehensive analysis," *Biosensors*, vol. 12, no. 10, p. 821, 2022.

[155] L. Xiao, Y. Cao, Y. Gai, E. Khezri, J. Liu, and M. Yang, "Recognizing sports activities from video frames using deformable convolution and adaptive multiscale features," *Journal of Cloud Computing*, vol. 12, no. 1, p. 167, 2023.

[156] F. Harjanto, Z. Wang, S. Lu, A. C. Tsoi, and D. D. Feng, "Investigating the impact of frame rate towards robust human action recognition," *Signal Processing*, vol. 124, pp. 220–232, 2016.

[157] M. Zaher, A. S. Ghoneim, L. Abdelhamid, and A. Atia, "Unlocking the potential of rnn and cnn models for accurate rehabilitation exercise classification on multi-datasets," *Multimedia Tools and Applications*, pp. 1–41, 2024.

[158] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," *IEEE transactions on cybernetics*, vol. 43, no. 5, pp. 1318–1334, 2013.

[159] A. Jalal, S. Kamal, and D. Kim, "A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments," *Sensors*, vol. 14, no. 7, pp. 11735–11759, 2014.

[160] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.

[161] P. Augustyniak, M. Smolen´, Z. Mikrut, and E. Kan´toch, "Seamless tracing of human behavior using complementary wearable and house-embedded sensors," *Sensors*, vol. 14, no. 5, pp. 7831–7856, 2014.

[162] I. Y. Jung, "A review of privacy-preserving human and human activity recognition," *International Journal on Smart Sensing and Intelligent Systems*, vol. 13, no. 1, pp. 1–13, 2020.

[163] S. K. Yadav, S. Sai, A. Gundewar, H. Rathore, K. Tiwari, H. M. Pandey, and M. Mathur, "Csitime: Privacy-preserving human activity recognition using wifi channel state information," *Neural Networks*, vol. 146, pp. 11–21, 2022.

[164] A. Jain, R. Akerkar, and A. Srivastava, "Privacy-preserving human activity recognition system for assisted living environments," *IEEE Transactions on Artificial Intelligence*, 2023.

[165] G. Diraco, G. Rescio, A. Caroppo, A. Manni, and A. Leone, "Human action recognition in smart living services and applications: context awareness, data availability, personalization, and privacy," *Sensors*, vol. 23, no. 13, p. 6040, 2023.

[166] S. Sakka, V. Liagkou, and C. Stylios, "Exploiting security issues in human activity recognition systems (harss)," *Information*, vol. 14, no. 6, p. 315, 2023.

[167] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing, "Multimodal fusion-adaboost based activity recognition for smart home on wifi platform," *IEEE Sensors Journal*, vol. 22, no. 5, pp. 4661–4674, 2022.

[168] B. Vidya and P. Sasikumar, "Wearable multi-sensor data fusion approach for human activity recognition using machine learning algorithms," *Sensors and Actuators A: Physical*, vol. 341, p. 113557, 2022.

[169] S. Hazra, A. A. Pratap, D. Tripathy, and A. Nandy, "Novel data fusion strategy for human gait analysis using multiple kinect sensors," *Biomedical Signal Processing and Control*, vol. 67, p. 102512, 2021.

[170] L. M. S. d. Nascimento, L. V. Bonfati, M. L. B. Freitas, J. J. A. Mendes Junior, H. V. Siqueira, and S. L. Stevan Jr, "Sensors and systems for physical rehabilitation and health monitoring—a review," *Sensors*, vol. 20, no. 15, p. 4063, 2020.

[171] Z. Zhao, J. Wang, S. Wang, R. Wang, Y. Lu, Y. Yuan, J. Chen, Y. Dai, Y. Liu, X. Wang, *et al.*, "Multimodal sensing in stroke motor rehabilitation," *Advanced Sensor Research*, vol. 2, no. 9, p. 2200055, 2023.

[172] M. R´ıos-Herna´ndez, J. M. Jacinto-Villegas, O. Portillo-Rodr´ıguez, and A. H. Vilchis-Gonza´lez, "User-centered design and evaluation of an upper limb rehabilitation system with a virtual environment," *Applied Sciences*, vol. 11, no. 20, p. 9500, 2021.

[173] L. van Velsen, G. Ludden, C. Gru¨nloh, *et al.*, "The limitations of user- and human-centered design in an ehealth context and how to move beyond them," *Journal of medical Internet research*, vol. 24, no. 10, p. e37341, 2022.

[174] I. Ten Klooster, J. Wentzel, F. Sieverink, G. Linssen, R. Wesselink, and L. van Gemert-Pijnen, “Personas for better targeted ehealth technologies: user-centered design approach,” *JMIR human factors*, vol. 9, no. 1, p. e24172, 2022.

[175] S. Huhn, M. Axt, H.-C. Gunga, M. A. Maggioni, S. Munga, D. Obor, A. Sie´, V. Boudo, A. Bunker, R. Sauerborn, *et al.*, “The impact of wearable technologies in health research: scoping review,” *JMIR mHealth and uHealth*, vol. 10, no. 1, p. e34384, 2022.

[176] R. De Fazio, V. Mattei, B. Al-Naami, M. De Vittorio, and P. Visconti, “Methodologies and wearable devices to monitor biophysical parameters related to sleep dysfunctions: an overview,” *Micromachines*, vol. 13, no. 8, p. 1335, 2022.

[177] J. C. Goldsack, A. Coravos, J. P. Bakker, B. Bent, A. V. Dowling, C. Fitzer-Attas, A. Godfrey, J. G. Godino, N. Gujar, E. Izmailova, *et al.*, “Verification, analytical validation, and clinical validation (v3): the foundation of determining fit-for-purpose for biometric monitoring technologies (biomets),” *npj digital Medicine*, vol. 3, no. 1, p. 55, 2020.

[178] J. B. Bro¨nneke, J. Mu¨ller, K. Mouratis, J. Hagen, and A. D. Stern, “Regulatory, legal, and market aspects of smart wearables for cardiac monitoring,” *Sensors*, vol. 21, no. 14, p. 4937, 2021.