

Review

Not peer-reviewed version

Machine Learning-Based Computer Vision for Depth Camera-Based Physiotherapy Movement Assessment: A Systematic Review

[Yafeng Zhou](#)*, [Fadilla 'Atyka Nor Rashid](#)*, [Marizuana Mat Daud](#), [Mohammad Kamrul Hasan](#),
Wangmei Chen

Posted Date: 12 December 2024

doi: 10.20944/preprints202412.1080.v1

Keywords: computer vision; depth camera; physiotherapy movement assessment; Machine learning; Systematic review



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Machine Learning-Based Computer Vision for Depth Camera-Based Physiotherapy Movement Assessment: A Systematic Review

Yafeng Zhou ^{1,2}, Fadilla 'Atyka Nor Rashid ^{1,*}, Marizuana Mat Daud ³, Mohammad Kamrul Hasan ⁴ and Wangmei Chen ⁴

¹ Center For Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia; Jalan Bangi, Bangi, 43600, Selangor, Malaysia

² Fotric Inc.; 2500 Xiupu Road, Pudong, 201315, Shanghai, China

³ Institute of Visual Informatics, Universiti Kebangsaan Malaysia; Jalan Bangi, Bangi, 43600, Selangor, Malaysia

⁴ Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia; Jalan Bangi, Bangi, 43600, Selangor, Malaysia

* Correspondence: fadilla@ukm.edu.my; Tel.: +60-13-736-5226 (F.A.N.R.)

Abstract: Background: Machine learning-based computer vision techniques using depth cameras have shown potential in physiotherapy movement assessment. However, a comprehensive understanding of their implementation, effectiveness, and limitations remains needed.

Methods: We conducted a systematic review following PRISMA guidelines, searching Web of Science, Scopus, PubMed, and Astrophysics Data System databases (2020-2024). From 371 initially identified publications, 18 met the inclusion criteria for detailed analysis.

Results: The analysis revealed three primary implementation scenarios: local (50%), clinical (33.4%), and remote (22.3%). Depth cameras, particularly the Kinect series (65.4%), dominated data collection methods. Data processing approaches primarily utilized RGB-D (55.6%) and skeletal data (27.8%), with algorithms split between traditional machine learning (44.4%) and deep learning (41.7%). Key challenges included limited real-world validation, insufficient dataset diversity, and algorithm generalization issues.

Conclusions: While machine learning-based computer vision systems demonstrated effectiveness in movement assessment tasks, further research is needed to address validation in clinical settings and improve algorithm generalization. This review provides a foundation for enhancing computer vision-based assessment tools in physiotherapy practice.

Keywords: computer vision; depth camera; physiotherapy movement assessment; Machine learning; Systematic review

1. Introduction

Physical rehabilitation plays a critical role in contemporary healthcare, helping individuals regain function after injury, surgery, or chronic disease. Recent advances in machine learning-based computer vision, from convolutional neural networks (CNN) [1] to vision transformers (ViT) [2], have achieved significant success in image classification, object detection, and other visual tasks [3]. These technological advances, particularly deep learning-based computer vision, have demonstrated substantial potential in physical rehabilitation [4–6] and physiotherapy movement assessment [7–9], offering opportunities to enhance traditional physical therapy practices.

The integration of depth cameras with machine learning algorithms has emerged as a promising approach for objective movement assessment in physiotherapy. These technologies enable contactless and accurate capture of patient movements, providing quantitative data for assessment and progress monitoring. Recent studies have highlighted how machine learning techniques can process this depth camera data to automate movement analysis and provide real-time feedback during rehabilitation exercises [4].

Artificial intelligence (AI) technologies, particularly deep learning (DL) methodologies, have revolutionized movement assessment in physiotherapy. As noted by Nogales *et al.*, "DL, the latest

breakthrough in AI, can provide more effective, personalized, and efficient care, leading to improved patient outcomes" [10]. Computer vision-based modeling and analysis of human movement have garnered significant academic attention [4], offering efficient and contactless solutions for patient movement data collection.

However, implementing machine learning-based computer vision systems for depth camera-based physiotherapy movement assessment presents various challenges. These include ensuring accurate movement detection, developing robust algorithms for real-time processing, and creating systems that can adapt to different patient conditions and rehabilitation needs. Understanding these challenges and current technological solutions is crucial for advancing this field.

This systematic review examines the application of machine learning-based computer vision in depth camera-based physiotherapy movement assessment. It analyzes current methodologies, technological implementations, and challenges in using depth cameras for human movement recognition and analysis in physiotherapy settings. The review aims to provide a comprehensive understanding of existing approaches while identifying areas for future research and development.

1.1. Domain Characteristics

This paper focuses on machine learning-based computer vision approaches utilizing depth cameras for physiotherapy movement assessment. Depth cameras have revolutionized physical therapy by enabling accurate 3D pose capture, real-time feedback mechanisms, personalized therapy monitoring, and objective evaluation of treatment outcomes. The integration of machine learning, particularly deep learning algorithms, has proven crucial for processing depth camera data effectively. These advanced algorithms enable efficient feature extraction from sensor data, precise movement analysis, and objective evaluation, significantly enhancing treatment effectiveness.

We systematically analyzed the literature on depth camera-based physiotherapy movement assessment, examining the methodologies, datasets, machine learning architectures, computer vision techniques, and implementation strategies employed. This comprehensive analysis provides insights into current approaches and guidance for future developments in machine learning-based movement assessment in physiotherapy. The combination of machine learning algorithms with depth camera technology has substantially advanced the field, particularly in automated movement analysis and assessment capabilities.

1.2. Scope of This Review

This systematic review examined relevant articles from 2020 to 2024, focusing on the integration of machine learning-based computer vision technology in physical rehabilitation. We specifically analyzed research implementing depth cameras for physiotherapy movement assessment. The review excluded papers on virtual rehabilitation and serious games based on visual sensors, despite their use of related technologies such as human skeleton keypoint detection and target tracking, to maintain focus on clinical movement assessment applications.

Recent studies have made significant contributions to this field. Several researchers have provided crucial insights into machine learning applications in healthcare AI [10–14]. Notable work includes [Burhani and Naqvi](#)'s research on specific applications like distal radius fracture assessment [15], while [Debnath et al.](#) contributed valuable insights into computer vision applications in rehabilitation [4].

This review specifically analyzes how researchers have implemented machine learning algorithms for human motion feature extraction and recognition using depth camera data. We paid particular attention to studies that utilized depth cameras as primary data capture sensors, examining their integration with machine learning approaches for movement assessment. This focused scope allows for detailed analysis of current technical implementations and their effectiveness in physiotherapy settings.

1.3. Outline

As illustrated in Figure 1, this comprehensive review is structured into six main sections. Section 1 establishes the foundation of the study by introducing the domain characteristics of machine learning-based computer vision in physiotherapy, defining the scope and objectives of the review, and presenting the structural organization of this article. Section 2 details the methodological framework, beginning with the formulation of research questions, followed by comprehensive search strategies, systematic data collection procedures, and the rigorous application of the PRISMA process. Section 3 presents a systematic analysis of the results through eight research questions (RQs), examining: sensing technologies (RQ1), available datasets (RQ2), data processing methods (RQ3), algorithmic approaches (RQ4), feature extraction techniques (RQ5), application scenarios (RQ6), movement assessment targets (RQ7), and problem statements (RQ8). Section 4 provides a comprehensive discussion that synthesizes key findings, contextualizes results within existing literature, addresses open challenges, and proposes future research directions. The paper concludes with Section 5, summarizing the main contributions and implications of the study, followed by Section 6, which details the funding sources supporting this research.

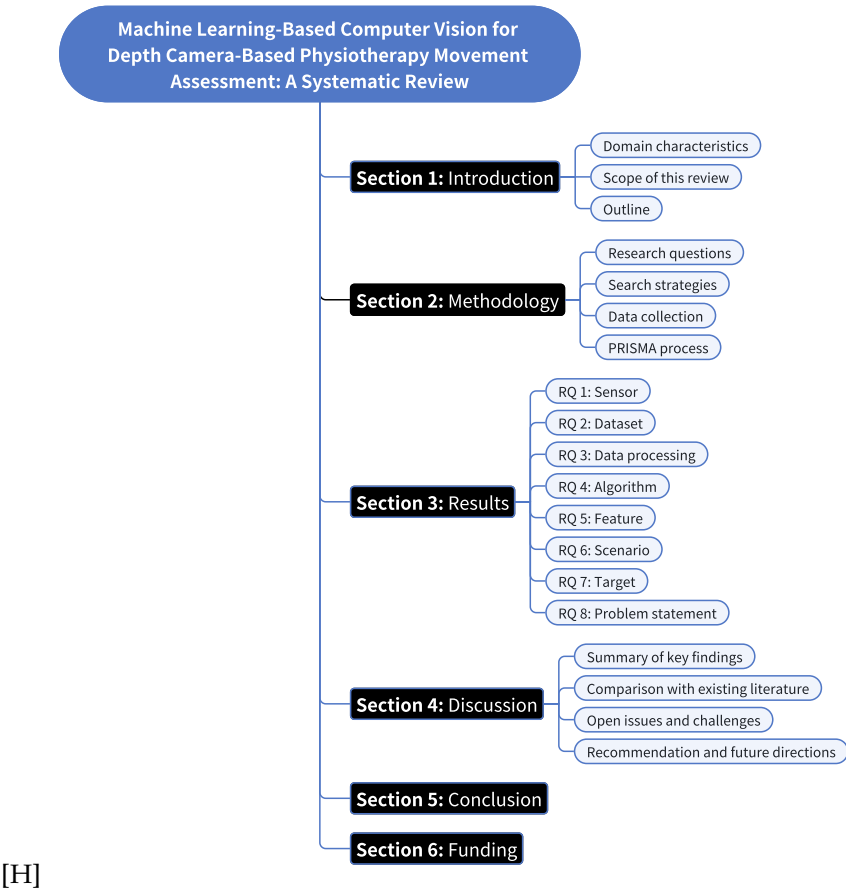


Figure 1. The structure of this article

2. Methodology

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [16,17] to ensure transparency and reproducibility. The review protocol included the following key components: research question formulation, literature search strategy development, inclusion and exclusion criteria establishment, data extraction process, and quality assessment methodology.

The researchers selected PRISMA over other methodological frameworks such as QUOROM [18], Cochrane Handbook [19], and JBIRM [20] due to its comprehensive reporting structure and wide acceptance in both clinical and technical research domains. This choice aligned with current journal publishing standards and facilitated systematic data collection and analysis.

The research process, as illustrated in Figure 3, comprised systematic search procedures, screening protocols, and data synthesis methods. These steps ensured a rigorous and replicable review process focused on machine learning-based computer vision applications in physiotherapy movement assessment using depth cameras.

2.1. Research Questions

This review explores systematic advances in recent research on computer vision for camera-depth sensor-based physiotherapy movement assignments. The primary objective is to summarize and analyze the relevant scientific literature, examine the characteristics and contributions of the work covered, and critically assess its limitations. Through this process, the paper aims to extract valuable insights and trends to guide future research. With this in mind, this review focuses on the following research questions to clearly articulate its purpose and usefulness (See Table 1).

Table 1. Research Questions and Motivations for Computer Vision in Camera-Depth Sensor-Based Physiotherapy Movement Assessments

RQ 1: What types of depth camera sensors are used in physiotherapy movement assessment, and what are their characteristics and applications?

Motivation 1: Depth camera technologies (e.g., time-of-flight, structured light, and stereo vision) are essential to physiotherapy movement assessment. Each depth camera type has unique advantages, and choosing the right camera for a particular treatment scenario is critical. A comprehensive analysis of the depth cameras used in this study and their actual efficacy in physical therapy can provide an essential reference for practical applications to improve the evaluation system’s stability and performance, optimize the data acquisition quality, and enhance evaluation accuracy.

RQ 2: How do the dataset type, construction methods, and feature selection impact algorithm performance and the effectiveness of applications in depth camera-based physiotherapy exercise assessments?

Motivation 2: The quality and applicability of datasets are critical factors in determining the effectiveness of computer vision techniques in physiotherapy movement assessment. To systematically analyze the effects of dataset types, construction methods, and dataset sizes and the effectiveness of their application as used in the current literature to evaluate existing models’ reliability and provide methodological guidance for future research to optimize dataset construction strategies.

RQ 3: What are the primary data processing methods used in depth-camera-based movement analysis for physical therapy, and how do they affect the assessment?

Motivation 3: Data processing directly affects the assessment’s quality and accuracy. Systematic study of current data processing techniques (e.g., skeletal data extraction, coordinate system transformation, data normalization) can help optimize the processing flow, improve data quality, and enhance model generalization to improve the accuracy and clinical relevance of the assessment and drive depth camera-based movement analysis in physical therapy toward higher accuracy and a more comprehensive range of applications.

RQ 4: Which algorithms perform best for depth camera-based physiotherapy movement assessment, and what are the advantages and limitations of different algorithms for various physical therapy tasks?

Motivation 4: Algorithm selection is crucial for depth camera-based physiotherapy assessment, impacting accuracy, efficiency, and clinical utility. Comparing traditional machine learning, deep learning, and specialized algorithms across various physiotherapy tasks, elucidating their strengths and limitations. Optimizing algorithm choice aims to enhance assessment techniques, improve patient outcomes, and establish a robust foundation for intelligent, personalized physiotherapy interventions.

RQ 5: What are the main innovative features of current research in physical therapy exercise assessment with depth camera sensors? How do these features improve rehabilitation outcomes and accessibility?

Motivation 5: Depth camera sensor technology presents multiple innovative features in physiotherapy movement assessment, and systematic analysis of these features is essential to capture advances in the field and guide future research. By exploring how these innovations can work together to promote advances in physical therapy, comprehensive and effective rehabilitation systems could be developed, improving diagnostic accuracy and personalized rehabilitation outcomes while optimizing home rehabilitation and remote monitoring solutions.

RQ 6: What are the main application scenarios for depth cameras in physical therapy, and how do these scenarios affect the implementation strategies and rehabilitation outcomes?

Motivation 6: Depth camera technology is used in different scenarios, such as remote, clinical, and home. An in-depth understanding of the specific needs and challenges of these scenarios can help optimize the implementation strategy, improve the efficiency of rehabilitation resource allocation, and promote the development of physical therapy in the direction of more intelligent, personalized, and universal, thus improving the overall rehabilitation effect.

RQ 7: What are the most common body parts targeted by camera depth sensor-based physiotherapy, and where is this technology most frequently implemented?

Motivation 7: In physical rehabilitation, movement assessment on different body parts is critical to patient recovery. Clarifying the application of depth camera technology in assessing various body parts can help develop more comprehensive and accurate assessment methods, improve the relevance and effectiveness of rehabilitation treatment, and provide a reliable basis for developing individualized rehabilitation plans.

RQ 8: What are the main challenges and limitations of computer vision and depth sensor technology in physiotherapy movement assessment?

Motivation 8: Although depth camera technology has shown great potential in physical therapy, it still faces many challenges. A thorough analysis of these challenges will not only help to understand the limitations of the current technology, but also point to future research that will lead to more accurate and reliable physiotherapy movement assessment systems and improved rehabilitation outcomes.

2.2. Search Strategies

The literature search is essential in literature review studies to provide adequate literature support and relevant information. This paper aims to present a literature search strategy on "Applying Computer Vision for Camera Depth Sensor-Based Physiotherapy Movements Assessment". Web of



In the literature search, the authors used different search strategies, and the search process followed a combination of computer vision and physiotherapy movements to identify and retrieve keywords. When identifying keywords, the researchers set the logical operator between the fields of computer vision and physiotherapy movements to AND to ensure that the retrieved literature covered both fields. In addition, the authors also combined the logical AND between physical therapy and movement or rehabilitation in the physical therapy domain, and the final search strategy architecture is shown in Figure 2 and in Table 2.

1 <https://www.webofscience.com/>
2 <https://www.scopus.com/>
3 <https://pubmed.ncbi.nlm.nih.gov/>
4 <https://ui.adsabs.harvard.edu/>

Table 2. Overview of Electronic Search Strategy: Including Databases, Search Queries.

Database	Query
WOS, PubMed	("Computer Vision" OR "Depth Camera*" OR "Depth Sensor*" OR "Kinect" OR "RGBD" OR "RGB-D") AND ("Physiotherapy" OR "Physiotherapies" OR "Physical Therapy" OR "Physical Therapies") AND ("Habilitation" OR "Rehabilitation" OR "Movement" OR "Exercise*" OR "Action" OR "Recognition")
Scopus	TITLE-ABS-KEY (("Computer Vision" OR "Depth Camera*" OR "Depth Sensor*" OR "Kinect" OR "RGBD" OR "RGB-D") AND ("Physiotherapy" OR "Physiotherapies" OR "Physical Therapy" OR "Physical Therapies") AND ("Habilitation" OR "Rehabilitation" OR "Movement" OR "Exercise*" OR "Action" OR "Recognition"))
Astrophysics	(title:"Computer Vision" OR title:"Depth Camera*" OR title:"Depth Sensor*" OR title:"Kinect" OR title:"RGBD" OR title:"RGB-D" OR keyword:"Computer Vision" OR keyword:"Depth Camera*" OR keyword:"Depth Sensor*" OR keyword:"Kinect" OR keyword:"RGBD" OR keyword:"RGB-D" OR abstract:"Computer Vision" OR abstract:"Depth Camera*" OR abstract:"Depth Sensor*" OR abstract:"Kinect" OR abstract:"RGBD" OR abstract:"RGB-D") AND (title:"Physiotherapy" OR title:"Physiotherapies" OR title:"Physical Therapy" OR title:"Physical Therapies" OR keyword:"Physiotherapy" OR keyword:"Physiotherapies" OR keyword:"Physical Therapy" OR keyword:"Physical Therapies" OR abstract:"Physiotherapy" OR abstract:"Physiotherapies" OR abstract:"Physical Therapy" OR abstract:"Physical Therapies") AND (title:"Habilitation" OR title:"Rehabilitation" OR title:"Movement" OR title:"Exercise*" OR title:"Action" OR title:"Recognition" OR keyword:"Habilitation" OR keyword:"Rehabilitation" OR keyword:"Movement" OR keyword:"Exercise*" OR keyword:"Action" OR keyword:"Recognition" OR abstract:"Habilitation" OR abstract:"Rehabilitation" OR abstract:"Movement" OR abstract:"Exercise*" OR abstract:"Action" OR abstract:"Recognition") AND (pubdate:[2020-01-01 TO 2024-12-31])

Combining the above strategies, a comprehensive search string was constructed by combining keywords in computer vision and physiotherapy movement. By applying this search strategy across retrieval sources, the relevant literature covering the combined application of physiotherapy movement and computer vision based on deep-camera sensors was obtained. Notably, the deletion and addition of relevant papers were carried out to obtain the final literature used for the literature review research.

2.3. Data Collection

A systematic literature review was conducted in databases such as WOS, Scopus, PubMed, and Astrophysics Data System using keywords related to physiotherapy, computer vision, depth camera sensors, and deep learning. Studies not focused on computer vision techniques, especially those that didn’t use depth camera sensors for physiotherapy movement assessment, were excluded. Critical information, such as camera types, dataset characteristics, learning architectures, and implementation strategies, was extracted.

2.4. PRISMA Process

This systematic review was conducted following PRISMA guidelines by a multidisciplinary team comprising three primary reviewers: a computer vision algorithm engineer (Zhou), a senior research expert in physiotherapy movement intelligence (Fadilla), and an expert in computer vision, deep learning and extensive experience in medical imaging and healthcare (Marizwana). This three-reviewer structure ensured comprehensive evaluation through majority voting and maintained methodological rigor.

The review process began with keyword definition in two domains. For computer vision, we initially identified: “Machine Learning”, “Deep Learning”, “Neural Networks”, “Computer Vision”, “Depth Camera*”, “Depth Sensor*”, “Kinect”, “RGBD”, and “RGB-D”. These terms were systematically

combined with physiotherapy movement keywords using Boolean operators. Our preliminary Web of Science search yielded 211 articles. To enhance specificity, we refined our search to focus on machine learning-based computer vision applications, resulting in 97 papers.

The screening process followed a structured three-phase approach:

1. **Initial Screening:** Each reviewer independently assessed titles, abstracts, and keywords against our inclusion criteria focusing on machine learning-based computer vision for depth camera-based physiotherapy movement assessment. Zhou evaluated technical aspects of computer vision implementations, Fadilla assessed physiotherapy relevance, and Marizwana provided oversight on methodology and deep learning applications.
2. **Quality Assessment:** Papers passing initial screening underwent rigorous quality assessment using standardized tools:
 - Risk of bias assessment using the QUADAS-2 tool for diagnostic accuracy studies
 - Methodological quality evaluation using the PEDro scale for clinical trials
 - Reporting completeness assessment using the PRISMA checklist
3. **Final Selection:** Disagreements were resolved through majority voting among the three reviewers. Each paper required approval from at least two reviewers for inclusion.

Selection criteria were established to ensure methodological rigor and relevance:

1. **Technical Relevance:** Studies must incorporate machine learning-based computer vision and depth sensor applications in physiotherapy contexts
2. **Methodological Quality:** Research must demonstrate robust experimental design with clear methodology, including:
 - Detailed description of machine learning architectures and parameters
 - Comprehensive evaluation metrics
 - Appropriate statistical analysis
3. **Results Validation:** Studies must include thorough validation of results with:
 - Clear performance metrics
 - Statistical significance testing
 - Comparison with baseline methods
4. **Innovation:** Research must demonstrate original contributions to the field
5. **Reporting Quality:** Papers must meet PRISMA reporting guidelines

We implemented systematic data extraction and synthesis methods to analyze the selected studies. This included assessment of:

1. Machine learning architectures and implementations
2. Depth camera specifications and configurations
3. Physiotherapy movement assessment protocols
4. Performance metrics and validation methods
5. Clinical relevance and practical applicability

3. Results

A diagram explaining PRISMA can be seen in Figure 3. We found 371 papers through our search on different databases: 97 papers on WOS, 138 papers on Scopus, 130 papers on PubMed, and 6 papers on Astrophysics Data System. After removing duplicates and papers without DOIs or written in languages other than English, 238 papers were left for the first screening. 23 articles were excluded at this stage, and 169 papers were excluded as they did not meet the criteria of "Applying Deep Learning-based Computer Vision for Camera Depth Sensor-Based Physiotherapy Movements Assessment." In the second screening, 46 articles were analyzed, of which 18 papers were finally selected for this systematic review as shown in Table 3.

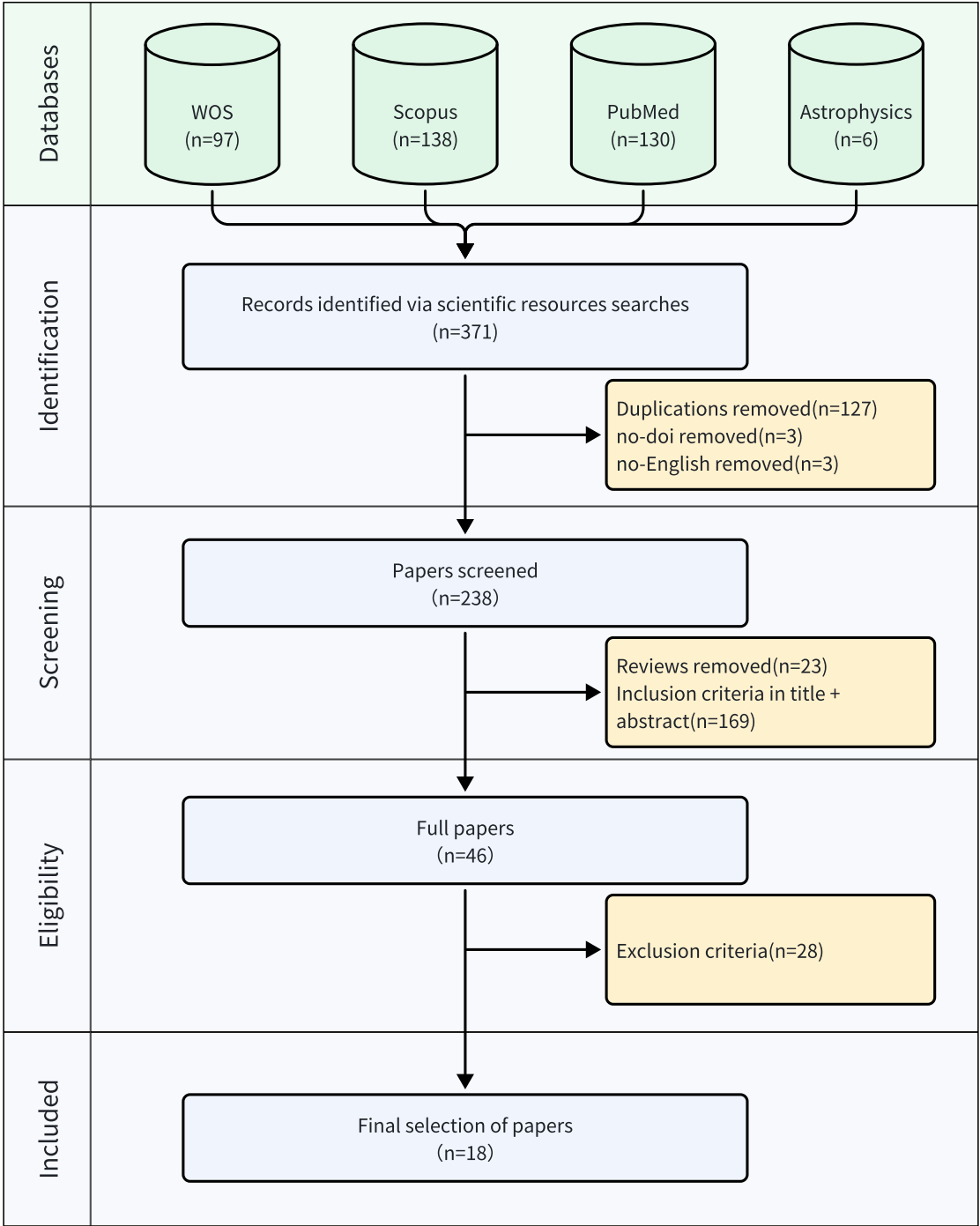


Figure 3. PRISMA diagram of the systematic review

Table 3. The final selection of articles for the review

ID	Author	Year	Country	Database	Summary
1	Lim et al.	2024	Canada	WOS, PubMed	Feasibility of depth cameras & pressure pads as alternatives to force plates.
2	Wagner et al.	2023	Poland	PubMed	Depth-sensor gait methods compared.
3	Raza et al.	2023	Pakistan, Saudi Arabia	WOS, Scopus	AI for pose estimation in physiotherapy exercises.
4	Maskeliūnas et al.	2023	Lithuania	WOS, Scopus	BiomacVR for posture & movement analysis in rehabilitation.
5	Lim et al.	2023	China	Scopus, PubMed	Adaptive Cobot system for assistive rehab training.
6	Khan et al.	2023	USA	WOS	Quantum neural network for post-stroke exercise assessment.
7	Bijalwan et al.	2023	India	WOS, Scopus	Automated system for upper limb exercise detection using an RGB-Depth camera.
8	Keller et al.	2022	USA	PubMed	Unsupervised ML for low back pain exercise strategies.
9	Zhao et al.	2021	USA, China	WOS, Scopus	Home TKR rehab system development.
10	Trinidad-Fernández et al.	2021	Spain, Belgium	PubMed	RGB-D camera validates motion capture in spondyloarthritis.
11	Hustinawaty et al.	2021	Indonesia	Scopus	Kinect SDK for a study of straight leg lift exercise.
12	Girase et al.	2021	USA	PubMed	Key factors identified for spine, hip, and knee assessment from sit-to-stand.
13	Çubukçu et al.	2021	Turkey	WOS, Scopus	Kinect-based mentor for shoulder injury telerehab.
14	Wei et al.	2020	USA	WOS, Scopus	Sensors and DL for automated balance assessment.

Table 3 (continued)

ID	Author	Year	Country	Database	Summary
15	Uccheddu <i>et al.</i>	2021	Italy	WOS, Scopus	Hybrid approach for 3D pose estimation proposed.
16	Trinidad-Fernández <i>et al.</i>	2020	Belgium, Spain, Australia	PubMed	RGB-D camera kinematic assessment results.
17	Saratean <i>et al.</i>	2020	Romania	WOS, Scopus	Kinect-based physical therapy guidance system.
18	García <i>et al.</i>	2020	Brazil	WOS, Scopus	RGB-D camera analysis of compensatory trunk movements.

3.1. RQ 1: Sensor

Camera sensors are crucial to acquiring visual data for movement analysis and physical therapy. The type of sensor chosen directly affects the quality of the captured data, the performance of the deep learning model, and the design of the evaluation protocol. In this review, depth cameras accounted for 65.4% (See Figure 4). Standard depth camera sensors are shown below:

1. **Kinect series:** RGB-D camera introduced by Microsoft, which obtains accurate depth and image data through infrared ranging and color image acquisition technologies, and the main models are Kinect V1, Kinect V2, and so on.
2. **RealSense series:** Intel’s RGB-D camera product line, using visual-inertial ranging technology, can obtain high-quality depth and motion data.
3. **Other RGB-D cameras:** Besides the mainstream products mentioned above, third-party vendors, such as Xtion Pro, provide some RGB-D camera devices.
4. **Ordinary RGB cameras:** Only capture color image data and must combine with other depth estimation algorithms for data processing and analysis.

After reviewing 18 related literature (See Table 8), the Kinect series is the most widely used sensor, and 12 literature [8,21,22,25–29,31–33,37] adopt Kinect V2 as the primary Kinect camera for data acquisition. The Kinect camera is recognized as the mainstream choice in this field because of its high accuracy and reliability. Intel RealSense series has also gained some applications; two papers [24,35] used the RealSense L515, D435i, and D415 models, respectively. RealSense cameras are technologically advanced, have excellent performance, and are expected to be used more widely in the future.

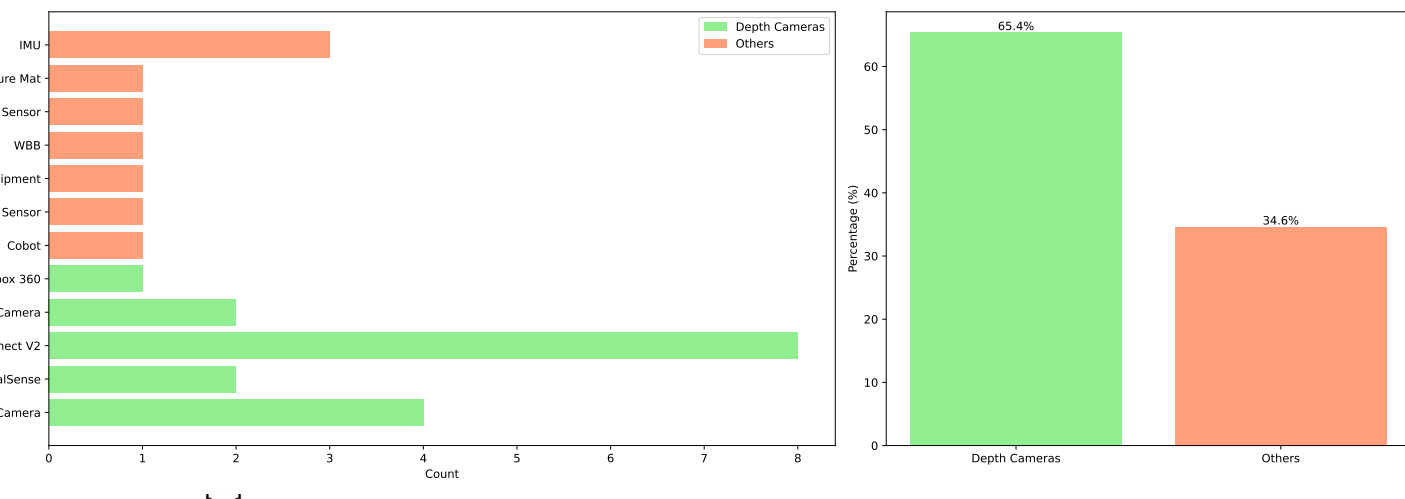


Figure 4. Sensor selection: depth-camera vs. other sensors.

In addition, three papers [23,30,36] used ordinary RGB cameras and other RGB-D cameras (e.g. Xtion Pro) to collect data. This approach has low hardware requirements but requires the development of appropriate algorithms to process and analyze the data.

Kinect series sensors are generally widely used due to their reliable performance and mature applications. However, emerging RGB-D cameras such as RealSense also show good prospects for development. Depending on the specific research objectives and application scenarios, choosing the right sensor is crucial to obtaining high-quality data. The advantages and disadvantages of different sensors must be weighed to utilize depth cameras in physical therapy fully.

3.2. RQ 2: Dataset

In computer vision-based physiotherapy movement assessment research, datasets are a key driver for its development. High-quality and diverse raw data and labeled information are the basis for developing excellent deep-learning models. At the same time, these datasets support the ability of the models to generalize across different application scenarios, facilitating the establishment of standardized physiotherapy movement assessment methods. In the reviewed literature, researchers used a variety of data types (See Figure 5), as shown in the following:

1. **RGB-D image/video data:** Many studies [8,22,24,25,27,30,31,34,35] have used color image and depth information captured by RGB-D cameras, usually image sequences or videos. This raw data can directly reflect the motion process and provide the basic input for subsequent motion detection and analysis.
2. **Joint and Skeletal Data:** Several studies [23,26,32,33,37] have utilized the joint positions and skeletal information extracted by depth cameras to construct skeletal datasets. This structured data directly represents the key features of human movement and is used to model and analyze joint motion trajectories.
3. **Combined datasets:** There are also some studies [29,30,36] that combine RGB-D data and auxiliary data captured by other sensors (e.g., IMU) to capture the motion process from different perspectives, to obtain more comprehensive and accurate motion information.

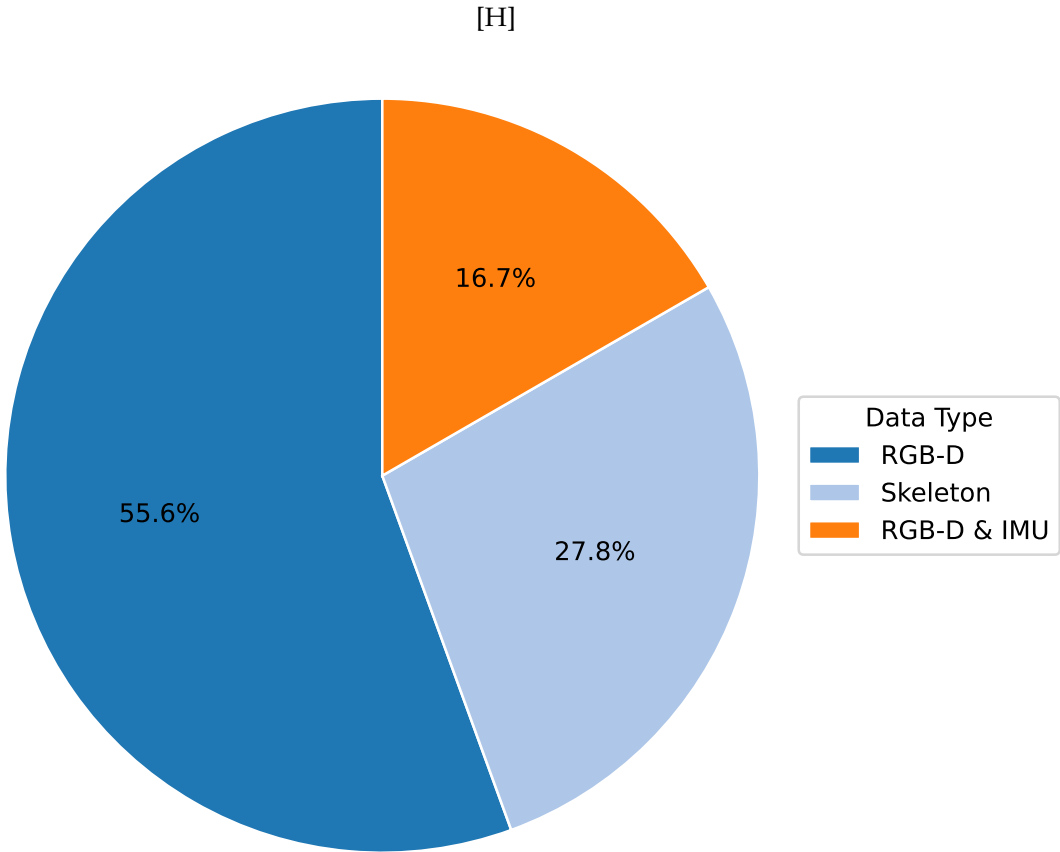


Figure 5. Pie chart of percentage distribution of data types.

Regarding dataset construction (See Table 4), the main approaches taken by existing studies are public datasets, self-constructed datasets, and the fusion of multiple datasets. Most studies [8,21,22,24,25,28,30–34,36] collected the data themselves. The size of the datasets ranged from tens to hundreds of participants, with some small to medium datasets and some larger datasets, such as Girase *et al.* containing 411 participants. In data collection and construction, researchers generally consider data from people of different ages, genders, and health conditions. This diversity enhances the robustness of the physical therapy movement dataset and improves the ability to generalize the model.

In addition, different studies have used existing publicly available datasets to accelerate model development by utilizing existing labeled data. For example, Raza *et al.* used "Multi-Class Exercise Poses for Human Skeleton"⁵, and Khan *et al.* used UI-PRMD [5]. In contrast, Bijalwan *et al.* used a fusion of multiple data sets by combining publicly available datasets (UTD-MHAD [38], mHealth [39], OU-ISIR [40], HAPT [41]) with the self-collected datasets for the construction of combinations.

In general, self-constructed datasets dominate computer vision-based physical therapy exercise and assistance research. Researchers have constructed or utilized multiple datasets according to specific needs, covering different groups and sizes of participants and containing multimodal data from different sensor sources. These datasets provide a solid data foundation for deep learning model training and physiotherapy exercise analysis. Proper utilization and expansion of high-quality datasets will provide strong support for developing this field.

⁵ <https://www.kaggle.com/datasets/dp5995/gym-exercise-mediapipe-33-landmarks>

Table 4. Summary of Data Types and Datasets in Physiotherapy Movement Assessment

ID	Data Type	Dataset
1	Joint displacement data series	10 non-disabled participants: 7 males, 3 females
2	RGB-D Images	5 subjects: 2 males, 3 females
3	Skeleton Data	Multi-Class Exercise Poses for Human Skeleton
4	RGB-D Videos	16 healthy subjects, 10 post-stroke patients
5	RGB-D Image	5 healthy subjects
6	Joint-Skeletal	UI-PRMD
7	RGB-D	UTD-MHAD, mHealth, OU-ISIR, HAPT
8	RGB-D	111 participants: back pain 43, control 26, surgery 4
9	RGB-D & IMU	/
10	RGB-D Videos & IMU	17 subjects: 54.35 (\pm 11.75) years
11	RGB-D	10 human objects
12	RGB-D Time Series	3 patient groups and one control group: 78 control, 130 LBP, 90 hip, and 113 knee
13	Skeleton Data	29 shoulder damaged volunteers: 18 males, 11 females
14	RGB-D Image	41 subjects: 26 males, 15 females; 21 healthy subjects and 20 patients with PD
15	RGB-D Videos	/
16	RGB-D & IMU	30 subjects: 18 65 years with non-specific lumbar pain
17	Skeleton Data	/
18	RGB-D	14 volunteers: 9 range of movement capture tests, 5 trunk compensation tests

3.3. RQ 3: Data Processing

Data processing plays a key role in depth camera-based physiotherapy movement analysis studies and directly affects the quality and precision of subsequent analyses. As Table 5 shows, through a systematic review of 18 works in the literature, the author can summarize several major types of data processing methods and techniques. First, skeletal data extraction and processing are the basis of most studies, e.g., [Girase et al.](#), [Çubukçu et al.](#), [Uccheddu et al.](#) use Kinect SDK 2.0 and OpenPose library to extract human skeletal data from raw depth images, respectively, which provides structured input for subsequent analysis. Second, to ensure the consistency and comparability of the data, some studies such as [Wagner et al.](#) and [Khan et al.](#) performed coordinate system transformation and alignment operations, which helped to eliminate errors caused by different devices or shooting angles.

Data standardization and normalization is another common processing step, such as the min-max normalization method used in [27], which helps to eliminate the effects of different scales and allows various features to be compared on the same magnitude. To improve data quality, some studies have used filtering and noise removal techniques, such as the Kalman filter and low-pass Butterworth filter used in [28], which effectively remove noise and improve signal quality. Feature extraction and selection also play an important role in machine learning, as evidenced by several articles in the literature [23,31]. These techniques help to reduce data dimensionality and improve the efficiency and generalization of the model.

For studies involving multimodal data, data synchronization and fusion become key issues. [Zhao et al.](#) and [Trinidad-Fernández et al.](#) explore the fusion of accelerometer and gyroscope measurements and visual synchronization via timestamps, respectively. In addition, to increase the diversity of training samples and improve the robustness of the model, [Wei et al.](#) employs data augmentation

techniques, which are particularly useful in deep learning model training to alleviate the problem of insufficient data effectively. Some studies such as [26] also mention the processing of feature transformations, including operations such as dimensionality reduction and feature combination, which help to extract more meaningful feature representations.

These diverse data processing methods cover the whole process from raw data acquisition to feature extraction, improving data quality and providing more reliable and effective inputs for subsequent algorithmic models. However, it should be noted that different studies used different combinations of processing methods according to their specific objectives and data characteristics. This diversity reflects the complexity and importance of data processing in physiotherapy exercise analysis.

Researchers may need to explore further advanced data processing techniques, such as automated feature engineering and more sophisticated multimodal data fusion methods, to cope with the increasingly complex demands of exercise analysis. Meanwhile, maximizing the extraction of useful information while maintaining the authenticity of the data is also a direction worthy of in-depth research. As technology advances, innovations in data processing methods will continue to drive depth camera-based physiotherapy movement analysis research toward higher accuracy and broader application areas.

Table 5. Summary of Algorithm and Processing in Physiotherapy Movement Assessment

ID	Algorithm	Processing
1	/	Joint Displacement Data.
2	Savitzky-Golay Filter	Transform the coordinate system using the KD, CH, and FV data processing methods.
3	RF, LR, GRU, LSTM, LogRF	MediaPipe Pose Marker, Feature Selection, and Hyperparameter Tuning.
4	ANN, DNN, CNN, CPM	Human skeletal movement was observed using visible information.
5	Imitation Learning for Adaptive Learning	/
6	High-Quality Neural Network	Align the length and center, and perform characteristic transformation.
7	HDL of CNN, RNN, CNN-GRU, and CNN-LSTM	Apply Min-Max normalization.
8	PCA, NLPCA, LR, Kaiser and Scree Plot Rules, Pattern Matching Statistics	Use Kalman filter, sequential second-order, and low pass Butterworth filtering.
9	/	Fuse accelerometer and gyroscope measurements.
10	Statistical analysis	Synchronize the dataset with the timestamp and visualize it using OpenNI2, NiTE2, and MRPT.
11	Detecting and Tracking	Calibration, skeletalization process, and feature extraction.

Table 5 (continued)

ID	Algorithm	Processing
12	SVM, RF, MLPs, CCNN, Semi-Supervised Learning, Unscented Kalman Filter	Estimate joint center positions using the standard Kinect 2 Body Tracking library.
13	Statistical Analysis	Use the Kinect SDK 2.0.
14	CNN, RF Classifier	Perform data augmentation.
15	OpenPose	Process video frames from the RGB sensor with the Open-Pose library.
16	/	Synchronize and use OpenNI2 and NiTE2 to create a virtual skeleton representation.
17	Effort-Based Method	Parameterization /
18	PrimeSense	Use the Kinect SDK 2.0.

3.4. RQ 4: Algorithm

In the research field of deep camera-based physiotherapy movement assessment, the selection and design of algorithms are crucial and directly affect the accuracy and efficiency of movement recognition, assessment, and analysis. Based on a systematic review of existing literature, algorithms can be broadly categorized into three groups: traditional machine learning algorithms, deep learning algorithms, and dedicated algorithms for specific tasks (See Figure 6). Each of these algorithms excels in different application scenarios, collectively contributing to the field’s rapid development (See Table 5).

Traditional machine-learning algorithms have been widely used in several studies. For instance, Random Forest (RF) [42], Logistic Regression (LR) [43], Support Vector Machine (SVM) [44], and Principal Component Analysis (PCA) [45] are favored for their interpretability and computational efficiency. Raza et al. utilize RF, LR, and LSTM algorithms for human posture estimation. At the same time, PCA, nonlinear PCA (NLPCA), and LR are combined to identify movement strategies in patients with low back pain [28]. These methods perform well with structured skeletal data and provide reliable tools for clinical assessment, especially in cases with well-defined features and small data sizes.

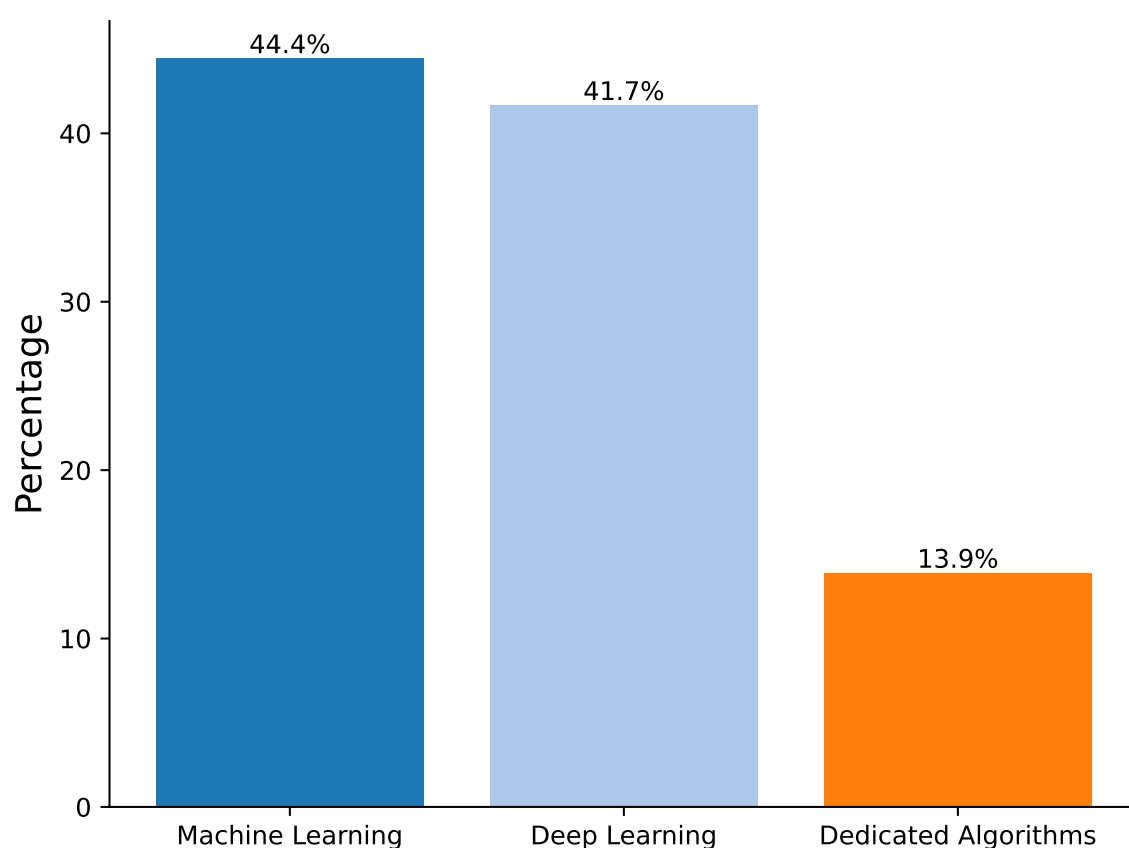


Figure 6. Diagram of bars for the algorithm in current literature.

With the rapid development of deep learning techniques, more studies are using deep neural network models. Architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory Networks (LSTMs) show significant advantages when dealing with complex time-series motion data. For example, a combination of Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Convolutional Pose Machine (CPM) achieves accurate analysis of human posture and motion [24], while a Hybrid Deep Learning (HDL) model combining CNN, RNN, and CNN-GRU effectively improves the detection and recognition accuracy of upper limb rehabilitation movements [27]. Another study uses a combination of CNN and RF classifiers to estimate the human body's center of gravity (CoM) [34]. These deep learning methods automatically extract deep features from raw data, significantly reducing the workload of manual feature engineering and improving model generalization, which is particularly suitable for processing large-scale and high-dimensional visual data.

In addition, some researchers have developed specialized algorithms for specific physiological therapy tasks. For example, Imitation Learning is used to achieve adaptive learning for multifunctional upper limb rehabilitation [25], while the Effort-Based Parameterization Method (EBPM) provides a theoretical basis for a home rehabilitation guidance system [37]. Though narrow in application, these specific task-oriented algorithms often provide precise and efficient solutions, reflecting the researchers' deep understanding and innovative thinking about actual clinical needs.

It should be noted that some studies have adopted the algorithm fusion strategy to utilize the advantages of different algorithms fully. For example, combining SVM, RF, multilayer perceptrons (MLPs), and cascaded convolutional neural networks (CCNNs) with semi-supervised learning and a traceless Kalman filter identifies the key factors of pathological movements [32]. Another approach integrates algorithms such as KD, CH, and FV with the Zebris FDM platform for accurately estimating

gait parameters [22]. This fusion strategy improves the performance and stability of the model, providing new ideas for solving complex physiological treatment problems.

Data preprocessing and postprocessing play equally important roles in the application of algorithms. Techniques such as coordinate system transformation [22], feature selection and hyperparameter tuning [23], min-max normalization [27], Kalman filtering and Butterworth low-pass filtering [28] improve the quality of the data and performance of the algorithms. The OpenPose library, for instance, is used to process video frames from RGB sensors for joint estimation [35]. These processing techniques complement the core algorithm, forming a complete analysis flow.

In general, movement analysis based on depth research cameras for physiological treatments has shown a trend of diversification, specialization, and convergence in algorithm selection and design. Researchers flexibly utilize existing machine learning and deep learning algorithms and develop innovative solutions based on specific application scenarios. This diversified algorithmic application strategy has effectively promoted technological advancements and improvements in clinical practice.

However, several noteworthy issues and research directions remain. First, although deep learning algorithms perform well, their "black-box" nature may affect the interpretability and credibility of models in clinical practice. Balancing model performance and interpretability requires further exploration. Second, while specialized algorithms for specific tasks are effective, their generalizability has yet to be verified. Designing algorithms that are specialized and flexible enough to adapt to different physiological treatment needs is a challenging research direction.

Furthermore, considering the special characteristics of physiological treatment data (e.g., data privacy, high labeling cost), improving algorithm performance under limited data conditions through techniques such as transfer learning and few-shot learning is crucial. With the development of edge computing and IoT technology, designing lightweight and efficient algorithms to meet the demand for real-time processing and feedback is also worthy of in-depth research.

With continuous progress in AI technology, particularly federated learning, interpretable AI, and adaptive learning, more efficient, accurate and safe algorithms will emerge in the future. These advancements will further promote depth camera-based physiotherapy movement analysis technology, providing patients with more personalized, intelligent, and effective rehabilitation programs, ultimately achieving the goal of precision medicine and intelligent rehabilitation.

3.5. RQ 5: Feature

As shown in Table 6, the summary of current research and innovations in physiotherapy and rehabilitation highlights the integration of advanced technologies and methods. Incorporating computer vision technology, depth cameras, and other sensors has been pivotal in enhancing the effectiveness and accuracy of rehabilitation treatments. These technologies offer cost-effective training feedback, improve gait assessment, increase the accuracy of human posture estimation, enhance patient engagement and the precision of posture and movement analysis, enable personalized adaptive learning, accelerate post-stroke movement assessment, and support clinical decision-making. Additionally, these technologies are used to develop home rehabilitation protocols and remote physiotherapy guidance systems, promoting continuous patient care and recovery. This systematic review categorizes and discusses the features of using depth camera sensors in physiotherapy movement assessment, focusing on the following aspects:

1. **Integration of Depth Sensors with Other Technologies:** Research on integrating depth cameras with other sensors, such as pressure mats, is crucial in physiotherapy movement assessment. For example, [Lim et al.](#) demonstrated efficient balance training feedback by combining depth cameras and pressure mats.
2. **Gait and Posture Assessment:** Gait and posture assessment are critical in physiotherapy. Several studies explore the application of depth sensors in these areas. For instance, [Wagner et al.](#) utilized depth sensors to improve gait assessment accuracy, enhancing diagnosis and treatment. [Hustinawaty et al.](#) applied Kinect SDK skeletonization to assess the straight leg raise accurately,

- aiding in lumbar condition diagnosis. [Girase et al.](#) used machine learning to automatically detect and classify pathological movements from sit-to-stand transitions.
- 3. Upper and Lower Limb Rehabilitation:** Upper and lower limb rehabilitation is a key research direction in physiotherapy. [Lim et al.](#) explored a personalized adaptive learning system for upper limb rehabilitation, improving patient outcomes. [Uccheddu et al.](#) developed a method using RGB-D sensors for precise joint angle estimation in-home rehabilitation. [Saratean et al.](#) implemented a Kinect-based remote physiotherapy guidance system, promoting continuous care.
 - 4. Applications of Machine Learning and Deep Learning:** Machine learning and deep learning are widely applied in physiotherapy assessments. [Raza et al.](#) improved human pose estimation using the LogRF and random forest algorithms. [Khan et al.](#) introduced a hybrid quantum neural network to enhance the speed and accuracy of post-stroke movement assessments. [Keller et al.](#) utilized unsupervised learning to analyze motion capture data, identifying movement strategies in low back pain patients. [Bijalwan et al.](#) combined deep learning models to enhance spatiotemporal feature modeling in stroke rehabilitation.
 - 5. Home Rehabilitation and Remote Monitoring:** Home rehabilitation and remote monitoring are current research hotspots. [Zhao et al.](#) proposed a home rehabilitation protocol for post-knee replacement using convenient technology. [Çubukçu et al.](#) developed a system for dynamic monitoring and correction of shoulder movements using Kinect. [Garcia et al.](#) used RGB-D cameras to analyze compensatory trunk movements, improving upper limb rehabilitation strategies.
 - 6. Clinical Applications and Validation:** Several studies validate the effectiveness of depth sensors in clinical settings. [Trinidad-Fernández et al.](#) accurately assessed movement limitations caused by spinal arthritis using RGB-D cameras, supporting better clinical decision-making. [Trinidad-Fernández et al.](#) demonstrated the reliability and speed of kinematic assessments using RGB-D cameras in clinical environments.

Table 6. Summary of the Research Feature

ID	Feature
1	Demonstrates integration of depth cameras and pressure mats as cost-effective, accessible feedback mechanisms for balance training.
2	Advances gait assessment techniques using depth sensor data, improving diagnostic and treatment accuracy.
3	Uses the LogRF method and random forest algorithms for improved human pose estimation in physiotherapy.
4	Employs virtual reality to increase precision and engagement in posture and movement analysis during rehabilitation.
5	Features a personalized adaptive learning system for upper-limb rehabilitation, enhancing patient-specific outcomes.
6	Introduces a hybrid quantum neural network to enhance speed and accuracy in post-stroke exercise assessments.

Table 6 (continued)

ID	Feature
7	Combines deep learning models to enhance modeling of spatio-temporal features in stroke rehabilitation.
8	Utilizes unsupervised machine learning to analyze movement capture data, identifying movement strategies in low back pain patients.
9	Proposes a protocol for home-based rehabilitation post-knee replacement using accessible technology.
10	Uses an RGB-D camera to accurately assess movement limitations in spondyloarthritis, supporting better clinical decisions.
11	Applies Kinect SDK skeletonization to accurately assess the straight leg raise, aiding lumbar condition diagnoses.
12	Automates detection and classification of pathologies from sit-to-stand movements using machine learning.
13	Develop a system using Kinect to monitor and correct shoulder exercises dynamically.
14	Integrates sensors and deep learning to enable real-time balance evaluations, enhancing therapy effectiveness.
15	Develop a hybrid method using RGB-D sensors for accurate joint angle estimation in-home rehabilitation.
16	Validates the use of RGB-D cameras for reliable and responsive kinematic assessments in clinical settings.
17	Implements a Kinect-based system for remote physiotherapy coaching, facilitating continuous care.
18	Analyze compensatory trunk movements with RGB-D cameras to refine upper limb rehabilitation strategies.

In summary, computer vision-based physiotherapy movement assessment using depth camera sensors have shown diverse applications and significant advancements. These technologies not only demonstrate great potential in enhancing diagnostic accuracy, personalized rehabilitation, and patient engagement but also pave the way for more effective and accessible physiotherapy solutions. Studies indicate that depth sensors are widely applied in gait and posture assessment and upper and lower limb rehabilitation, and, when combined with machine learning and deep learning technologies, have achieved breakthroughs in home rehabilitation and remote monitoring. These studies cover balance training, virtual reality integration, and home rehabilitation, providing real-time, accurate, and personalized feedback mechanisms that improve treatment outcomes and patient participation. Future

research should focus on integrating these features into comprehensive systems to enhance further diagnostic accuracy, movement assessment speed, and home rehabilitation efficacy, promoting more efficient and convenient rehabilitation practices. In general, applying depth cameras and advanced algorithms brings innovative solutions to physical therapy, significantly improving the efficiency and coverage of rehabilitation training.

3.6. RQ 6: Scenario

The movement analysis based on depth cameras for physical therapy shows significant value and potential in three major scenarios (See Table 7): remote, clinical, and local (See Figure 7). In remote scenarios, the technology breaks through geographical limitations and enables patients to receive real-time rehabilitation guidance and assessment at home, improving the accessibility and continuity of rehabilitation services; in clinical scenarios, it provides medical professionals with accurate exercise data and objective assessment tools, which help to formulate personalized and efficient treatment plans; and in local scenarios, such as at home or in community-based rehabilitation centers, the technology supports autonomous training and daily monitoring and enhances patients’ self-management ability. The integration of these three scenarios optimizes the allocation of rehabilitation resources and realizes an all-round multilevel rehabilitation care system, significantly improving the overall effect of physical therapy and patient experience. With the advancement of technology and in-depth clinical practice, this multi-scenario application mode is reshaping the traditional rehabilitation concept and promoting the development of physical therapy in the direction of intelligence, personalization, and popularization.

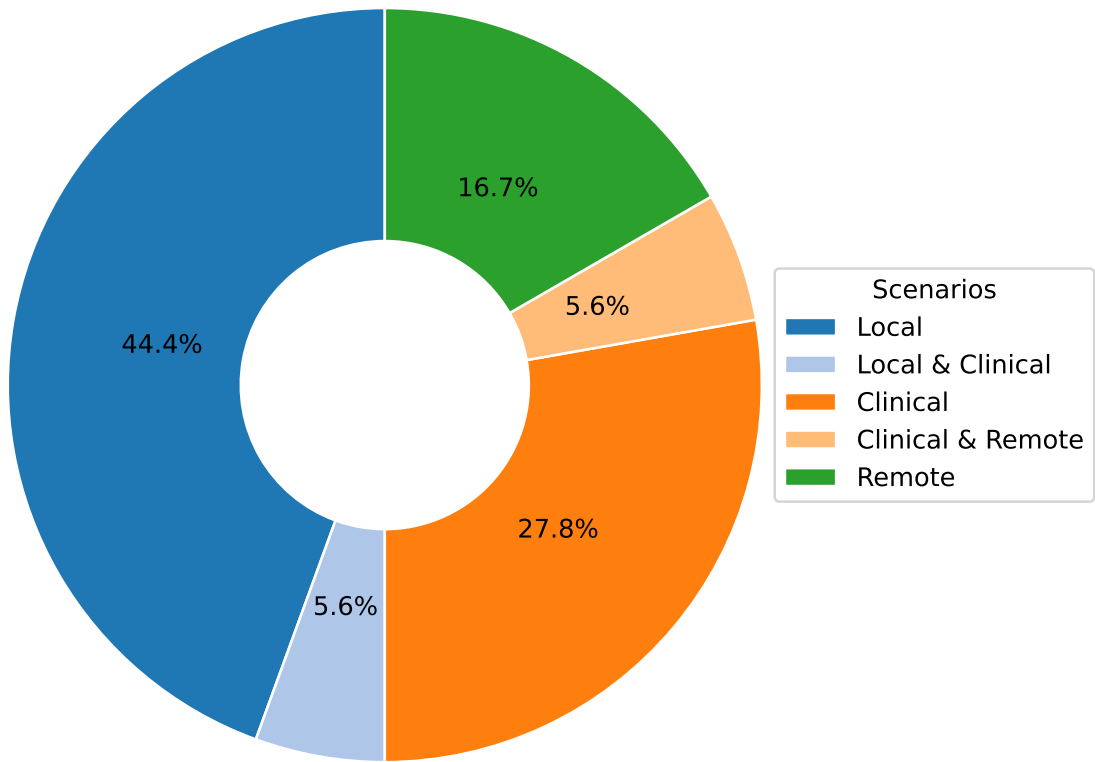


Figure 7. Pie chart depicting research application scenarios for physiotherapy movement assessment.

In a remote scenario, the main objective is to provide patients with a convenient home rehabilitation program. With the development of telemedicine technology, remote physical therapy movement assessment based on depth cameras has become a reality. For example, [Maskeliūnas et al.](#) describe

BiomacVR, a virtual reality (VR)-based rehabilitation system that combines a VR physical training monitoring environment and upper limb rehabilitation technology for precise interaction and improves patient engagement, which is applied to a real-time physical therapy sports wellness system for telerehabilitation. Authors [Lim et al.](#) [25] propose an adaptive learning system based on imitation learning for multi-purpose upper extremity rehabilitation that allows patients to perform rehabilitation at home. Authors [Çubukçu et al.](#) [33] examine the development of a Kinect 2 sensor-based telerehabilitation system that observes and evaluates exercise in patients with shoulder impairments through a web application used for communication between the patient and the therapist and a console application that helps the patient perform the exercise correctly. [Saratean et al.](#) proposes an approach based on effort parameterization for monitoring a home rehabilitation system to ensure correctness and adherence to rehabilitation exercises.

In clinical scenarios, physiotherapy movement assessment research focuses on accurately analyzing and assessing patients’ motion status to provide key data support for clinical diagnosis and treatment. For example, [Wagner et al.](#) realize the accurate analysis of patients’ gait through the estimation of gait parameters, which effectively assists clinical diagnosis and the formulation of treatment plans. Due to the wide application of artificial intelligence technology in this field, for example, [Maskeliūnas et al.](#) [24] utilize neural network algorithms to observe human skeletal motion through visible information, which can accurately analyze patient posture and movement patterns. In addition, [Bijalwan et al.](#) [27] apply deep learning techniques for the detection and recognition of detecting and recognizing upper limb rehabilitation exercises, which helps clinicians assess the progress of the rehabilitation of patients. In disease-specific studies, such as [28], machine learning methods are applied to identify exercise strategies for patients with low back pain, providing a scientific basis for developing clinical treatment programs. [Trinidad-Fernández et al.](#), which validates and analyzes patients’ trunk movement limitations by synchronizing and visualizing datasets, helps clinicians gain insights into patients’ movement abilities and limitations. Authors [Hustinawaty et al.](#) [31] employ advanced detection and tracking techniques, combining calibration, skeletonization process, and feature extraction, to achieve monitoring and analysis of key movements in the rehabilitation process, providing detailed movement data support for clinical decision-making. [Girase et al.](#) [32] applie semi-supervised learning algorithms to estimate the joint center position through the standard Kinect 2 body tracking library, successfully identifying and classifying critical factors of pathological movements and providing more accurate data support for rehabilitation treatment.

Table 7. Summary of Scenarios in Physiotherapy Movement Assessment

ID	Scenario	Objective
1	Local	Evaluate balance training effectiveness with depth cameras and pressure mats.
2	Local	Enhance gait analysis accuracy using new spatiotemporal methods and depth sensors.
3	Local	Improve exercise correction in physiotherapy with innovative pose estimation.
4	Remote, Clinical	Develop a VR system for precise human posture and motion analysis to boost rehabilitation engagement.
5	Remote	Build a personalized adaptive learning system with collaborative robots for upper-limb rehab.
6	Local	Improve post-stroke exercise assessments with a hybrid quantum neural network.
7	Clinical	Enhance upper extremity rehab post-stroke by modeling spatio-temporal features with deep learning.

Table 7 (continued)

ID	Scenario	Objective
8	Clinical	Discover low back pain strategies using unsupervised learning on motion data.
9	Local	Establish a comprehensive home protocol for post-knee replacement recovery.
10	Clinical	Validate RGB-D cameras for precise trunk movement analysis in spondyloarthritis.
11	Clinical, Local	Analyze straight leg raises accurately using Kinect SDK's skeletonization.
12	Clinical	Automate diagnosis of spinal, hip, and knee pathologies from sit-to-stand movements.
13	Remote	Develop a Kinect-based system to monitor and correct shoulder rehab exercises.
14	Local	Develop an on-demand balance evaluation tool integrating sensors with deep learning.
15	Local	Merge 2D and 3D RGB-D data for precise joint angle estimation in-home rehab.
16	Local	Validate the reliability and responsiveness of kinematic assessments with RGB-D cameras.
17	Remote	Implement a Kinect-based remote physiotherapy coaching system to ensure exercise adherence.
18	Clinical	Analyze compensatory trunk movements in upper limb rehab using RGB-D cameras.

In local scenarios, physiotherapy exercise and assessment research focus on using advanced algorithms and data processing techniques to automate the evaluation of patient rehabilitation training and posture recognition and improve rehabilitation effects and patient compliance. For example, [Lim et al.](#) provides visual feedback to patients through the acquisition and processing of joint displacement data in real-time to help them perform effective balance training at home. [Raza et al.](#) applies AI algorithms combined with MediaPipe pose labeling, feature selection, and hyper-parameter tuning to achieve a high-precision estimation of human posture, which provides important data support for rehabilitation training. [Khan et al.](#) [26] realize automated evaluation of exercises through high-quality neural network alignment of length and center and feature transformation, which significantly improves the efficiency and effectiveness of rehabilitation training. In rehabilitation after specific surgeries, [Zhao et al.](#) [29] incorporate accelerometer and gyroscope measurement techniques to perform home rehabilitation training evaluation after total knee replacement. This makes it possible to monitor patient rehabilitation progress in the home environment. [Wei et al.](#) [34] accurately estimate the patient's center of mass position through data augmentation techniques, providing a scientific basis for balance training and assessment. [Uccheddu et al.](#) [35] achieve an accurate estimation of joint position by processing video frame data from RGB sensors, providing strong support for motion analysis. In addition, [Trinidad-Fernández et al.](#) [36] create virtual skeletal representations to assess patients' ability to perform functional tasks, further extending the scope and depth of local rehabilitation assessment. These studies fully demonstrate that patient rehabilitation training can be effectively monitored and evaluated in local scenarios with advanced algorithms and data processing techniques, improving the rehabilitation effect and significantly enhancing patient compliance.

3.7. RQ 7: Target

In physiotherapy movement assessment, the selection of an appropriate target for the study is critical. This selection is directly related to the type of visual data to be captured and its depth, which affects the design and application of deep learning models. By carefully selecting targets for the human body, researchers can ensure that the acquired motion data is pertinent and complete, providing high-quality input for movement recognition and assessment.

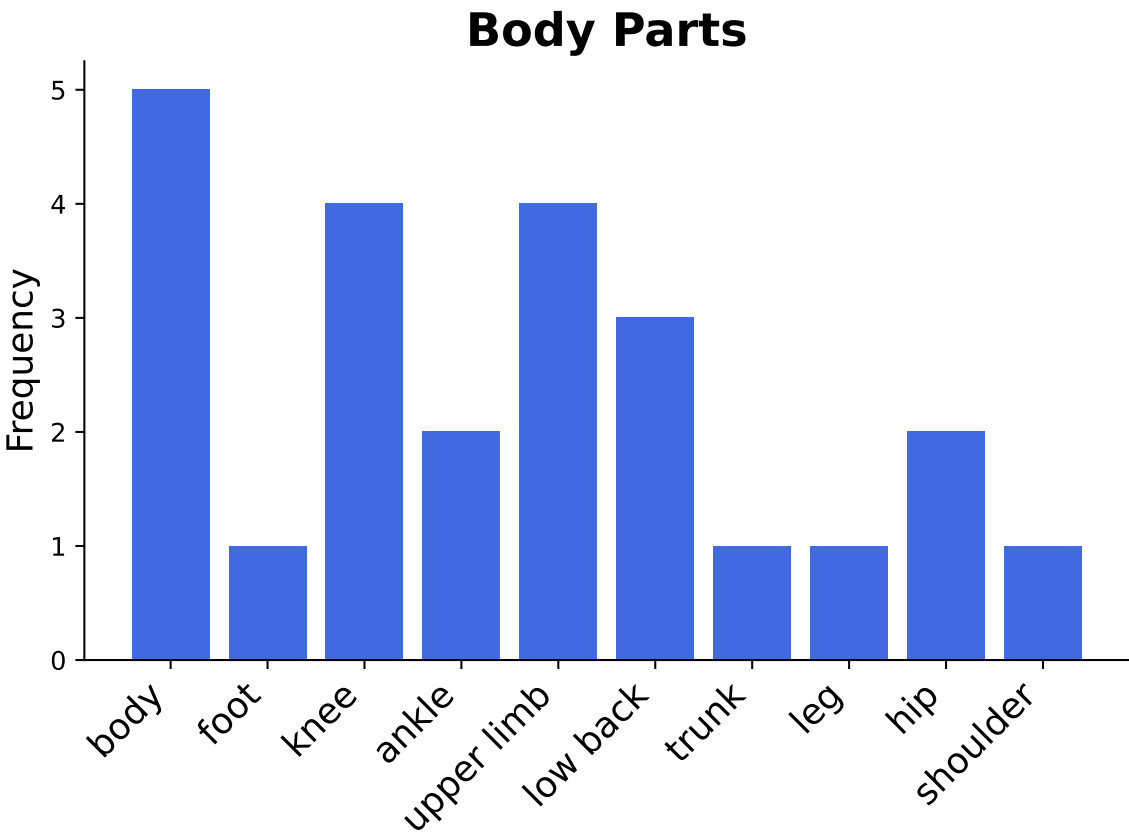


Figure 8. Diagram of bars for the studied body parts.

As shown in Figure 8, the existing literature shows that researchers generally focus on the main body parts, such as the entire body, the upper limb, and the lower limb, as well as specific joint parts, such as the knee, ankle, and shoulder (See Table 8). For example, of the 18 articles, five [21,23,26,34,37] provided in-depth studies on the recognition and evaluation of full body motion recognition and evaluation, which focus on how to capture body motion data using a depth camera and analyze it using deep learning models.

The upper limb is also one of the key targets, with a total of four articles [8,24,25,27] focusing on recognizing and evaluating movements of the upper limb. These studies focused on capturing motion data at the arm and elbow joints to guide upper limb rehabilitation.

The lower limbs have received similarly extensive attention, including the foot, ankle, knee, and hip. Six publications [22,29,31,32,35,36] have also been extensively analyzed aimed at providing assessment and guidance for lower limb rehabilitation. Of particular note, motor analysis of the regions of the lower back and trunk, whose motor status is critical to evaluating general physical mobility, was examined explicitly in 3 publications [28,30,32].

In summary, the existing literature covers movement recognition and assessment of the full body, upper and lower limbs, and key joints. Reasonable selection of the study target is decisive in designing efficient data acquisition schemes, building accurate deep learning models, and deriving results from

targeted movement analysis. Ensuring the reliability of the study results and the integrity of the acquired motion data, thus fully utilizing the potential of depth cameras and deep learning techniques in the scope of physiological therapy.

Table 8. Summary of Target and Sensor in Physiotherapy Movement Assessment

ID	Target	Sensor
1	body	Kinect V2, Pressure Mat
2	foot, knee, ankle	Kinect V2
3	body	Ordinary Camera
4	upper limb	Intel RealSense L515/D435i, HTC Vive VR Equipment
5	upper limb	Kinect Camera, Cobot, Force/Torque Sensor
6	body	Kinect
7	upper limb	Kinect V2
8	low-back	Kinect V2
9	knee	Kinect V2, IMU(Shimmer)
10	trunk	RGB-D Camera, IMU(MP67B)
11	leg	Kinect
12	low-back, hip, knee	Kinect V2
13	shoulder	Kinect V2
14	body	Kinect, WBB
15	hip, knee, ankle	Intel RealSense D415
16	low-back	RGB-D Camera (Xtion Pro), IMU(MP67B)
17	body	Kinect for Xbox 360
18	upper limb	Kinect V2

3.8. RQ 8: Problem Statement

As shown in Table 9, most studies indicate that using computer vision and depth sensor technology for patient movement analysis and rehabilitation training has become a significant development direction in modern physiotherapy and assessment. However, the problem statements from a systematic review reveal numerous challenges in applying current technologies and methods. To better understand the bottlenecks in current research and future development directions, this paper categorizes and discusses these problems as follows:

- 1. Equipment and Feedback Mechanism Issues:** Several studies [21,34,37] highlight that current balance training and motion analysis equipment is often expensive and bulky, limiting its use in resource-constrained clinical environments. There needs to be more effective on-demand balance assessment tools in physiotherapy, further restricting treatment flexibility and real-time feedback capabilities.
- 2. Data Utilization and Diagnostic Accuracy Issues:** Traditional gait analysis and kinematic assessment methods need to effectively utilize depth sensor data, leading to decreased diagnostic accuracy [22,36]. For example, conventional methods cannot accurately capture the straight leg

- raise motion [31], complicating lumbar assessments. Furthermore, home rehabilitation methods inaccurately estimate joint angle ranges [35], which affects patient treatment outcomes.
3. **Accuracy Issues in Posture Estimation and Motion Analysis:** Posture estimation in physiotherapy often lacks accuracy [23,24], impacting the correction of exercises and rehabilitation effectiveness. Traditional motion analysis tools lack precision and interactivity, failing to meet the demands of efficient rehabilitation.
 4. **Personalization and Dynamic Adaptability Issues:** Standard upper limb rehabilitation devices and post-stroke assessment systems lack dynamic adaptation to patient progress [25,26], limiting their effectiveness. Existing shoulder rehabilitation systems lack precise and interactive exercise monitoring [33], making personalized treatment difficult. Inadequate motion data analysis limits low back pain treatment strategies [28].
 5. **Spatiotemporal Feature Modeling Issues:** Insufficient spatiotemporal feature modeling in stroke rehabilitation [27,29] affects the effectiveness of rehabilitation exercises. Moreover, current methods also fall short in analyzing compensatory trunk movements [8], further impacting upper limb rehabilitation outcomes.
 6. **Pathological Diagnosis and Assessment Tool Issues:** Current automated diagnostic tools for spine, hip, and knee pathologies [32], as well as tools to assess movement limitations in ankylosing spondylitis [30], are inadequate. These issues indicate that the existing automated diagnostic and assessment tools still need to meet clinical needs and require further development and optimization.

Table 9. Summary of the Research Problem Statement

ID	Problem Statement
1	Current feedback mechanisms in balance training are restricted by their reliance on expensive, bulky equipment.
2	Traditional gait analysis needs to harness depth sensor data, impacting diagnostic accuracy effectively.
3	Current pose estimation in physiotherapy often lacks precision, leading to ineffective exercise correction.
4	Conventional motion analysis tools lack the precision and interactivity required for effective rehabilitation.
5	Standard upper-limb rehabilitation devices do not adapt to patient progress, limiting their effectiveness.
6	Existing post-stroke assessments lack precision and speed, necessitating advanced computational solutions.
7	Spatio-temporal feature modeling in stroke rehabilitation is inadequate, hindering exercise effectiveness.
8	Personalized treatment strategies in low back pain are limited by poor analysis of movement data.
9	Spatio-temporal feature modeling in stroke rehabilitation is inadequate, hindering exercise effectiveness.
10	Tools for assessing movement limitations in spondyloarthritis need to be improved.
11	Current methods need to capture the straight leg raise, complicating lumbar assessments accurately.
12	Automated diagnostic tools for spine, hip, and knee pathologies need to be improved.
13	Existing shoulder rehab systems lack precise, interactive monitoring of exercises.

Table 9 (continued)

ID	Problem Statement
14	Current physical therapy lacks practical on-demand balance evaluation tools.
15	Home rehab methods inaccurately estimate joint angular ranges, affecting treatment outcomes.
16	Kinematic assessments lack the reliability and responsiveness required for effective clinical decisions.
17	Current physical therapy lacks effective on-demand balance evaluation tools.
18	Methods to analyze compensatory trunk movements in upper limb rehab are ineffective.

In summary, current computer vision and depth sensor-based physiotherapy movement assessment technologies face numerous challenges regarding equipment, data utilization, posture estimation, personalization adaptability, spatiotemporal feature modeling, and pathological diagnosis tools. Addressing these issues requires multidisciplinary collaboration and technological innovation. Future research should focus on improving the portability and affordability of equipment, optimizing data analysis methods, enhancing the dynamic adaptability of systems, and developing high-precision diagnostic and assessment tools.

4. Discussion

4.1. Summary of Key Findings

This systematic literature review provides an in-depth exploration of the application of depth camera-based computer vision techniques in physiotherapy movement assessment, revealing the following key findings:

- Diversification of application scenarios:** The study shows that depth camera technology presents a diversified application trend in physical therapy, which is mainly distributed in three major scenarios: remote (16.7%), clinical (27.8%), and local (50%). This diversified application mode optimizes the allocation of rehabilitation resources and significantly promotes the development of physical therapy in the direction of intelligence, personalization, and popularization. It is particularly noteworthy that in remote and home rehabilitation, depth camera technology effectively breaks through geographical limitations, dramatically improves the accessibility and continuity of rehabilitation services, and provides patients with more flexible and convenient treatment options.
- Dominance of sensor technology:** Regarding sensor selection, depth cameras dominate (65.4%) in physical therapy applications. The Kinect family of sensors is widely adopted due to its excellent reliability and technological maturity. This trend fully reflects the unique advantages of depth sensors in providing markerless, non-contact human movement information, which provides a solid technological foundation to drive the development of personalized rehabilitation.
- Diversity of data types and processing techniques:** Regarding data types and processing, the study shows that RGB-D data (55.6%) and skeletal data (27.8%) are the most dominant data types. The data processing techniques show a diverse trend, covering a wide range of aspects from skeletal data extraction to coordinate system transformation to feature selection and noise reduction. This diversity highlights the critical importance of data quality for subsequent analysis and reflects the researchers’ unremitting efforts to improve the precision and efficiency of data processing.

4. **Balance and trend of algorithm selection:** Regarding algorithm selection, traditional machine learning algorithms (44.4%) and deep learning algorithms (41.7%) show a relatively balanced proportion of usage. This balance reflects the researchers' efforts to seek an optimal balance between the pursuit of model interpretability and performance. Meanwhile, deep learning algorithms show increasingly significant advantages in processing complex time series motion data, further indicating that future research will increasingly adopt it.
5. **Distribution of research focus:** The research focuses mainly on rehabilitation assessment and movement analysis of the upper limb, lower back, knee, and whole body. This point reflects the importance of these body parts in daily activities and quality of life and highlights the great potential of deep camera technology in comprehensively assessing human movement. This finding provides valuable guidance for future research directions, suggesting that depth camera technology is expected to play a vital role in a broader range of body part rehabilitation assessments.

In summary, the findings of the literature review not only systematically elucidate the current state of the application of depth camera technology in physical therapy but also provide a clear direction for future research. These findings highlight the critical role of technological innovation in driving personalization, precision, and accessibility in physical therapy and the need for interdisciplinary collaboration in advancing the field.

4.2. Comparison with Existing Literature

This systematic literature review resonates with the existing literature and provides essential extensions and additions through a multidimensional analysis. Regarding the range of technical applications, this study extends the work of [Debnath et al.](#) emphasizing computer vision's potential in evaluating human movement. In contrast, this study quantifies the trend of depth camera applications through systematic analysis and provides a specific distribution of application scenarios and evaluation of their effects. This in-depth analysis provides a more precise direction guide for future research.

This study presents a more comprehensive and dynamic picture of technology evolution than [Rashid et al.](#) [46]. Although Kinect still dominates, our study reveals that emerging depth camera technologies (e.g., Intel RealSense) are gaining more and more attention. This finding reflects the trend of technology diversification and provides researchers and clinical practitioners with a broader perspective on technology selection.

Regarding AI applications, this study complements the work of [Sardari et al.](#) and [Sumner et al.](#) [14, 47]. While these studies focused on using AI in skeletal data analysis and holistic physical rehabilitation, the present study provides a more comprehensive view, covering the entire process from data collection to algorithm selection. Our findings highlight the applicability of AI techniques in different assessment tasks, providing essential insights into the optimization and selection of AI systems.

In terms of data processing, this study extends the work of [Abdullah and Al-Kazzaz](#) [48]. We focus on common RGB-D and skeletal data and also delve into various data processing techniques, such as coordinate system transformations and feature selection. This comprehensive analysis provides researchers with a more decadent choice of data processing strategies, which helps improve the accuracy and efficiency of data analysis.

In terms of the assessment system framework, compared to the AI-assisted physical therapy assessment framework proposed by [Ekambaram and Ponnusamy](#) [49], this study provides a more systematic and comprehensive assessment system. We focused on emotion detection and movement recognition. We explored how to simulate the clinical assessment process, laying a theoretical foundation for building a better AI-assisted rehabilitation system.

In addition, this study complements the research of [Momin et al.](#) [50] on applying depth sensors in in-home activity monitoring for older people. Our study covers home, clinical, and remote rehabilitation scenarios, providing a more comprehensive analysis of application scenarios and helping to promote the flexible application of depth camera technology in different rehabilitation settings.

Overall, this study significantly contributes to specific application scenarios and effectiveness assessment of depth camera technology, diversity and applicability analysis of data processing methods, application and selection considerations of AI algorithms in different assessment tasks, and proposing a more systematic and comprehensive review framework for assisted rehabilitation assessment. These findings fill the gaps in the existing literature and provide a solid theoretical foundation for interdisciplinary collaboration and technological innovation in physical therapy. The systematic analysis of this study provides a clear direction for future research, especially in improving the applicability of the technology, data processing efficiency, and algorithmic accuracy, which is expected to promote the development of physical therapy technology in the direction of being more innovative, more accurate, and more personalized.

4.3. Open Issues and Challenges

As shown in Table 10, there are several limitations in the research on using computer vision with depth sensors for physiotherapy movement assessment. Most studies have been conducted on healthy participants or small datasets, lacking validation on patients with various pathological conditions. Algorithms and systems are often tested in laboratory settings, which do not reflect real-world complexity. The high computational resources required for deep learning models and virtual reality systems, along with patient discomfort, limit their widespread application. Sensor resolution and reliability issues, inaccuracies in Kinect SDK capture, and sensitivity to environmental interference and specific clothing also impact assessment results. Many studies do not consider the sustainability of long-term home rehabilitation plans and the complexity of dynamic movements. For instances:

- 1. **Data Samples:** Several studies are limited to testing on healthy participants or small demographic groups, failing to include patients with spinal cord injuries or varying cognitive impairments, which restricts the generalizability of the findings [21,28].
- 2. **Experimental Environment:** Many studies are conducted in controlled laboratory settings, which do not adequately reflect real-world diversity and complexity. Some algorithms perform well in controlled environments but may significantly underperform in cluttered or complex real-world settings [22,23,34].
- 3. **Technical Solutions:** Depth sensors like Kinect have limited resolution and field of view, potentially failing to capture fine joint movements. The accuracy of these sensors can be affected by wearable interference and environmental factors [30,31,35–37].
- 4. **Algorithm Application:** Many current algorithms are validated only on static exercises, not sufficiently exploring the complexity of dynamic movements. Although deep learning and machine learning models show some effectiveness, they require significant computational resources and are limited by the size of datasets, which can affect model generalization [27,33].
- 5. **User Experience:** Virtual reality systems may cause discomfort or dizziness in some patients, limiting their broader application. Studies often fail to consider the feasibility of long-term adherence to home rehabilitation programs, impacting practical effectiveness [24,29].

Table 10. Summary of the Research Issues and Challenges

ID	Ref	Limitation
1	[21]	The study only tested non-disabled participants, not spinal cord injury patients.
2	[22]	Gait analyses were conducted under controlled lab conditions, which may not reflect real-world variability.

Table 10 (continued)

ID	Ref	Limitation
3	[23]	Pose estimation algorithms tested primarily in well-controlled environments, may not perform as well in cluttered spaces.
4	[24]	VR systems may cause discomfort or dizziness in some patients, limiting their widespread usability.
5	[25]	System’s adaptability is not tested on patients with varying degrees of cognitive impairments.
6	[26]	Validation is restricted to small datasets which may not generalize to broader populations.
7	[27]	Deep learning models require extensive computational resources, limiting deployment in low-resource settings.
8	[28]	Machine learning models derived from a limited demographic, potentially affecting the universality of findings.
9	[29]	Study did not account for long-term adherence to home-based programs.
10	[30]	Camera’s depth resolution is insufficient to capture fine-grained joint movements accurately.
11	[31]	Kinect SDK’s accuracy in capturing leg movements is not verified against gold-standard clinical assessment tools.
12	[32]	Diagnostic accuracy is dependent on the precise execution of sit-to-stand movements, which varies widely among patients.
13	[33]	Limited to static exercises; dynamic movements’ complexity not fully explored.
14	[34]	Balance evaluation algorithms not validated in diverse real-world environments.
15	[35]	Estimations may not be accurate for patients with severe joint deformities or those wearing certain types of clothing that interfere with sensor accuracy.
16	[36]	Kinematic data’s reliability is compromised by occasional sensor inaccuracies and environmental interferences.
17	[37]	Kinect sensor’s limited field of view can restrict the range of exercises that can be monitored.
18	[8]	Analysis does not account for simultaneous lower limb movements, which can influence trunk motion.

In summary, while depth cameras and computer vision technologies offer new possibilities for physiotherapy, the aforementioned limitations need to be addressed in future research. This requires technological innovation, interdisciplinary collaboration, and careful consideration of cost-effectiveness and user experience. In addition, these limitations highlight the need for future research to focus

on diverse patient populations, real-world environment testing, sensor improvement, and long-term rehabilitation evaluation. By addressing these issues, future research can provide more comprehensive, reliable, and applicable tools for physiotherapy movement assessment, benefiting clinical practice and patient rehabilitation. Therefore, future research should focus on the following directions:

1. **Expanding Sample Diversity:** Future studies should include a more diverse range of participants, particularly patients with spinal cord injuries and varying cognitive impairments, to improve the generalizability and application value of the results.
2. **Real-World Environment Validation:** Validate the performance of algorithms and systems in real-world environments, such as homes and community settings, to ensure applicability.
3. **Improving Sensor Technology:** Enhance the resolution and field of view of depth sensors, design appropriate sensor application schemes, optimize motion capture performance, and develop new sensors and improved algorithms to reduce environmental and wearable interference.
4. **Developing Generalizable Algorithms:** Create algorithms that generalize well to dynamic movements, using large-scale, diverse datasets to improve model generalization.
5. **Enhancing User Experience:** Improve virtual reality systems, use non-contact methods where possible to reduce patient discomfort, increase acceptability, and design systems and protocols that promote long-term home rehabilitation adherence.

4.4. Recommendation and Future Directions

With the continuous progress of computer vision technology, its application in camera depth sensor-based physiotherapy movement assessment, although there are many shortcomings and limitations, also shows many promising research and development directions. Based on the previous analysis, future research should focus on the following aspects:

First, in-depth research is needed to investigate how to effectively apply depth camera technology in physical therapy rehabilitation exercises, especially the data acquisition that does not require professional setup and is contactless, and in this way, improve the algorithms to capture and analyze movement data in complex dynamic environments more accurately.

Secondly, developing lightweight and efficient algorithms for resource-constrained environments is also an important direction for future research. In terms of interdisciplinary collaboration, it is recommended that collaboration in computer science, physical therapy, neuroscience, and biomedical engineering be strengthened. Perspectives and methodologies from different fields can provide more comprehensive solutions that can drive innovation and application of physical therapy technologies.

In addition, existing machine learning and deep learning models need to be validated on more extensive and diverse datasets to improve their generalizability and robustness and explore new theoretical frameworks and algorithms to explain and optimize complex phenomena in movement capture and rehabilitation training. For example, such as the limited application of self-supervised learning algorithms to the learning extraction of physical therapy movement feature data, and the design of personalized physical therapy rehabilitation programs through downstream tasks.

Future research should explore the application of deep sensors in home rehabilitation, telemedicine, and community health centers, which will significantly enhance the accessibility and continuity of rehabilitation services, especially the development of rehabilitation devices and programs for home environments that can provide personalized and real-time feedback to improve patient engagement and rehabilitation outcomes, especially in today's era of increasing population aging for the elderly to develop physical therapy Smart applications for exercise and assessment.

In summary, future research efforts should build on existing research by focusing on the full-scale application of deep cameras, the construction of adequate datasets, the application of self-supervised learning algorithms in physical therapy sports, as well as the exploration of new methods, the promotion of interdisciplinary collaborations, and the validation of theoretical hypotheses, with an eye on practical applications and emerging trends. By addressing these future directions, depth

camera-based computer vision for physiotherapy movement assessment will continue to evolve, ultimately leading to more effective, convenient, and personalized patient care.

5. Conclusion

This systematic review examined machine learning-based computer vision techniques for depth camera-based physiotherapy movement assessment through analysis of 18 high-quality papers. The findings demonstrated both the potential and limitations of current approaches in this field.

The analysis revealed that depth cameras, particularly the Kinect family of sensors, dominated physiotherapy applications due to their ability to provide markerless, non-contact motion capture. The review identified three main data types used in the studies: RGB-D data (55.6%), skeletal data (27.8%), and multimodal data combining RGB-D with IMU measurements. Most studies employed the Kinect SDK 2.0 or OpenPose library for skeletal data extraction, supplemented by data alignment, normalization, and transformation techniques.

Regarding algorithmic approaches, the field showed a balanced distribution between traditional machine learning methods (44.4%) and deep learning models (41.7%), with specialized task-specific algorithms accounting for the remaining 13.9%. Convolutional Neural Networks emerged as the predominant deep learning architecture, while Random Forest and Support Vector Machines were the most commonly used traditional approaches.

The applications primarily focused on local environments (50%), followed by clinical settings (33.4%), with remote rehabilitation representing a smaller portion (22.3%). The research concentrated on assessing movements of the upper limbs, lower back, knees, and whole-body motion, with implementations varying across rehabilitation, assessment, and movement analysis applications.

However, several limitations were identified in the current research. These included insufficient validation on diverse patient populations, limited testing in real-world environments, small dataset sizes, and challenges in handling dynamic movements. Future research should address these limitations by expanding patient diversity, conducting real-world validation studies, developing more robust algorithms for dynamic movement analysis, and creating larger, more comprehensive datasets.

The findings suggest that while machine learning-based computer vision shows promise for objective and efficient physiotherapy movement assessment, significant work remains to bridge the gap between laboratory success and clinical implementation. Advances in sensor technology, algorithm development, and validation studies will be crucial for realizing the full potential of these systems in clinical practice.

6. Funding

This work was supported by the GGPM 2023-049 grant and funding from the Faculty of Information Science & Technology (FTSM) at Universiti Kebangsaan Malaysia (UKM).

Abbreviations

The following abbreviations are used in this manuscript:

KD	Knee Distance
CH	Centre Height
FV	Foot Velocity
RF	Random Forest
LR	Logistic Regression
GRU	Gated Recurrent Unit
LSTM	Long Short-Term Memory
LogRF	Logistic regression Recursive Feature elimination
ANN	Artificial Neural Network
DNN	Deep Neural Network
CNN	Convolutional Neural Network
CPM	Convolutional Pose Machines
HDL	Hybrid Deep Learning

RNN	Recurrent Neural Network
PCA	Principal Component Analysis
NLPCA	Non-Linear Principal Component Analysis
SVM	Support Vector Machine
MLPs	Multi-Layer Perceptrons
CCNN	Causal dilated Convolutional Neural Network
MRPT	Mobile Robot Programming Toolkit [51,52]

References

1. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., Vol. 25.
2. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; Uszkoreit, J.; Houlsby, N. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale, 2021, [[arxiv:cs/2010.11929](https://arxiv.org/abs/2010.11929)].
3. Chai, J.; Zeng, H.; Li, A.; Ngai, E.W. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications* **2021**, 6, 100134.
4. Debnath, B.; O'Brien, M.; Yamaguchi, M.; Behera, A. A Review of Computer Vision-Based Approaches for Physical Rehabilitation and Assessment. *Multimedia Systems* **2022**, 28, 209–239. doi:10.1007/s00530-021-00815-4.
5. Liao, Y.; Vakanski, A.; Xian, M. A Deep Learning Framework for Assessing Physical Rehabilitation Exercises. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2020**, 28, 468–477. doi:10.1109/TNSRE.2020.2966249.
6. Francisco, J.A.; Rodrigues, P.S. Computer vision based on a modular neural network for automatic assessment of physical therapy rehabilitation activities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2022**, 31, 2174–2183.
7. Patel, S.; Park, H.; Bonato, P.; Chan, L.; Rodgers, M. A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation* **2012**, 9, 1–17.
8. Garcia, A.T.; Kelbouscas, A.L.D.S.; Guimarnes, L.L.; Silva, S.A.V.E.; Oliveira, V.M. Use of RGB-D Camera for Analysis of Compensatory Trunk Movements in Upper Limbs Rehabilitation. 2020 IEEE 18th International Conference on Industrial Informatics (INDIN); IEEE: Warwick, United Kingdom, 2020; pp. 243–248. doi:10.1109/INDIN45582.2020.9442112.
9. Rashid, F.A.N.; Suriani, N.S.; Mohd, M.N.; Tomari, M.R.; Zakaria, W.N.W.; Nazari, A. Deep Convolutional Network Approach in Spike Train Analysis of Physiotherapy Movements. In *Advances in Electronics Engineering*; Zakaria, Z.; Ahmad, R., Eds.; Springer Singapore; Vol. 619, pp. 159–170. doi:10.1007/978-981-15-1289-6_15.
10. Nogales, A.; Rodríguez-Aragón, M.; García-Tejedor, Á. A Systematic Review of the Application of Deep Learning Techniques in the Physiotherapeutic Therapy of Musculoskeletal Pathologies. *Computers in Biology and Medicine* **2024**, p. 108082. doi:10.1016/j.compbimed.2024.108082.
11. Tack, C. Artificial intelligence and machine learning | applications in musculoskeletal physiotherapy. *Musculoskeletal Science and Practice* **2019**, 39, 164–169.
12. Ravali, R.S.; Vijayakumar, T.M.; Lakshmi, K.S.; Mavaluru, D.; Reddy, L.V.; Retnadhas, M.; Thomas, T. A systematic review of artificial intelligence for pediatric physiotherapy practice: past, present, and future. *Neuroscience Informatics* **2022**, 2, 100045.
13. Naik, S.; Rathod, D.; Agarwala, P.; Phadke, S.; Tilak, P. A Literature review of artificial intelligence in physiotherapy practice **2022**.
14. Sardari, S.; Sharifzadeh, S.; Daneshkhah, A.; Nakisa, B.; Loke, S.W.; Palade, V.; Duncan, M.J. Artificial Intelligence for Skeleton-Based Physical Rehabilitation Action Evaluation: A Systematic Review. *Computers in Biology and Medicine* **2023**, 158, 106835. doi:10.1016/j.compbimed.2023.106835.
15. Burhani, T.; Naqvi, W.M. Impact of Artificial Intelligence in the Physiotherapy Rehabilitation of Distal Radial Fracture Patients: A Review. *Journal of Pharmaceutical Research International* **2021**, 33, 1982–1988.
16. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; the PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine* **2009**, 151, 264–269. doi:10.7326/0003-4819-151-4-200908180-00135.
17. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; Chou, R.; Glanville, J.; Grimshaw, J.M.; Hróbjartsson, A.; Lalu, M.M.; Li, T.

- Loder, E.W.; Mayo-Wilson, E.; McDonald, S.; McGuinness, L.A.; Stewart, L.A.; Thomas, J.; Tricco, A.C.; Welch, V.A.; Whiting, P.; Moher, D. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. doi:10.1136/bmj.n71.
18. Moher, D.; Cook, D.J.; Eastwood, S.; Olkin, I.; Rennie, D.; Stroup, D.F. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* **1999**, *354*, 1896–1900.
19. Higgins, J.P.; Green, S.; others. Cochrane handbook for systematic reviews of interventions **2008**.
20. Institute, J.B.; others. Joanna Briggs Institute reviewers' manual: 2014 edition. *Australia: The Joanna Briggs Institute* **2014**, pp. 88–91.
21. Lim, D.; Pei, W.; Lee, J.W.; Musselman, K.E.; Masani, K. Feasibility of Using a Depth Camera or Pressure Mat for Visual Feedback Balance Training with Functional Electrical Stimulation. *BioMedical Engineering OnLine* **2024**, *23*, 19. doi:10.1186/s12938-023-01191-y.
22. Wagner, J.; Szymański, M.; Błażkiewicz, M.; Kaczmarczyk, K. Methods for Spatiotemporal Analysis of Human Gait Based on Data from Depth Sensors. *Sensors* **2023**, *23*, 1218. doi:10.3390/s23031218.
23. Raza, A.; Qadri, A.M.; Akhtar, I.; Samee, N.A.; Alabdulhafith, M. LogRF: An Approach to Human Pose Estimation Using Skeleton Landmarks for Physiotherapy Fitness Exercise Correction. *IEEE Access* **2023**, *11*, 107930–107939. doi:10.1109/ACCESS.2023.3320144.
24. Maskeliūnas, R.; Damaševičius, R.; Blažauskas, T.; Canbulut, C.; Adomavičienė, A.; Griškevičius, J. BiomacVR: A Virtual Reality-Based System for Precise Human Posture and Motion Analysis in Rehabilitation Exercises Using Depth Sensors. *Electronics* **2023**, *12*, 339. doi:10.3390/electronics12020339.
25. Lim, J.H.; He, K.; Yi, Z.; Hou, C.; Zhang, C.; Sui, Y.; Li, L. Adaptive Learning Based Upper-Limb Rehabilitation Training System with Collaborative Robot. 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2023, pp. 1–5, [arxiv:cs/2305.10642]. doi:10.1109/EMBC40787.2023.10340313.
26. Khan, M.A.A.H.; Murikipudi, M.; Azmee, A.A. Post-Stroke Exercise Assessment Using Hybrid Quantum Neural Network. 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC); IEEE: Torino, Italy, 2023; pp. 539–548. doi:10.1109/COMPSAC57700.2023.00078.
27. Bijalwan, V.; Semwal, V.B.; Singh, G.; Mandal, T.K. HDL-PSR: Modelling Spatio-Temporal Features Using Hybrid Deep Learning Approach for Post-Stroke Rehabilitation. *Neural Processing Letters* **2023**, *55*, 279–298. doi:10.1007/s11063-022-10744-6.
28. Keller, A.V.; Torres-Espin, A.; Peterson, T.A.; Booker, J.; O'Neill, C.; Lotz, J.C.; Bailey, J.F.; Ferguson, A.R.; Matthew, R.P. Unsupervised Machine Learning on Motion Capture Data Uncovers Movement Strategies in Low Back Pain. *Frontiers in Bioengineering and Biotechnology* **2022**, *10*, 868684. doi:10.3389/fbioe.2022.868684.
29. Zhao, W.; Yang, S.; Luo, X. Towards Rehabilitation at Home after Total Knee Replacement. *Tsinghua Science and Technology* **2021**, *26*, 791–799. doi:10.26599/TST.2020.9010034.
30. Trinidad-Fernández, M.; Cuesta-Vargas, A.; Vaes, P.; Beckwée, D.; Moreno, F.Á.; González-Jiménez, J.; Fernández-Nebro, A.; Manrique-Arija, S.; Ureña-Garnica, I.; González-Sánchez, M. Human Motion Capture for Movement Limitation Analysis Using an RGB-D Camera in Spondyloarthritis: A Validation Study. *Medical & Biological Engineering & Computing* **2021**, *59*, 2127–2137. doi:10.1007/s11517-021-02406-x.
31. Hustinawaty, H.; Rumambi, T.; Hermita, M. Skeletonization of the Straight Leg Raise Movement Using the Kinect SDK. *International Journal of Advanced Computer Science and Applications* **2021**, *12*.
32. Girase, H.; Nyayapati, P.; Booker, J.; Lotz, J.C.; Bailey, J.F.; Matthew, R.P. Automated Assessment and Classification of Spine, Hip, and Knee Pathologies from Sit-to-Stand Movements Collected in Clinical Practice. *Journal of Biomechanics* **2021**, *128*, 110786. doi:10.1016/j.jbiomech.2021.110786.
33. Çubukçu, B.; Yüzgeç, U.; Zileli, A.; Zileli, R. Kinect-Based Integrated Physiotherapy Mentor Application for Shoulder Damage. *Future Generation Computer Systems* **2021**, *122*, 105–116. doi:10.1016/j.future.2021.04.003.
34. Wei, W.; Mcelroy, C.; Dey, S. Using Sensors and Deep Learning to Enable On-Demand Balance Evaluation for Effective Physical Therapy. *IEEE Access* **2020**, *8*, 99889–99899. doi:10.1109/ACCESS.2020.2997341.
35. Uccheddu, F.; Governi, L.; Furferi, R.; Carfagni, M. Home Physiotherapy Rehabilitation Based on RGB-D Sensors: A Hybrid Approach to the Joints Angular Range of Motion Estimation. *International Journal on Interactive Design and Manufacturing (IJIDeM)* **2021**, *15*, 99–102. doi:10.1007/s12008-020-00728-y.
36. Trinidad-Fernández, M.; Beckwée, D.; Cuesta-Vargas, A.; González-Sánchez, M.; Moreno, F.A.; González-Jiménez, J.; Joos, E.; Vaes, P. Validation, Reliability, and Responsiveness Outcomes of Kinematic Assessment

- with an RGB-D Camera to Analyze Movement in Subacute and Chronic Low Back Pain. *Sensors* **2020**, *20*, 689. doi:10.3390/s20030689.
37. Saratean, T.; Antal, M.; Pop, C.; Cioara, T.; Anghel, I.; Salomie, I. A Physiotherapy Coaching System Based on Kinect Sensor. 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP); IEEE: Cluj-Napoca, Romania, 2020; pp. 535–540. doi:10.1109/ICCP51029.2020.9266178.
 38. Chen, C.; Jafari, R.; Kehtarnavaz, N. UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 168–172. doi:10.1109/ICIP.2015.7350781.
 39. Banos, O.; Garcia, R.; Saez, A. MHEALTH. UCI Machine Learning Repository, 2014. doi:10.24432/C5TW22.
 40. Makihara, Y.; Mannami, H.; Tsuji, A.; Hossain, M.A.; Sugiura, K.; Mori, A.; Yagi, Y. The OU-ISIR Gait Database Comprising the Treadmill Dataset. *IPSI Transactions on Computer Vision and Applications* **2012**, *4*, 53–62. doi:10.2197/ipsjtcva.4.53.
 41. Reyes-Ortiz, J.; Anguita, D.; Oneto, L.; Parra, X. Smartphone-Based Recognition of Human Activities and Postural Transitions. UCI Machine Learning Repository, 2015. doi:10.24432/C54G7M.
 42. Breiman, L. Random forests. *Machine learning* **2001**, *45*, 5–32. doi:https://doi.org/10.1023/A:1010933404324.
 43. Cox, D.R. The regression analysis of binary sequences. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **1958**, *20*, 215–232. doi:https://doi.org/10.1111/j.2517-6161.1958.tb00292.x.
 44. Cortes, C.; Vapnik, V. Support-vector networks. *Machine learning* **1995**, *20*, 273–297.
 45. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemometrics and intelligent laboratory systems* **1987**, *2*, 37–52. doi:https://doi.org/10.1016/0169-7439(87)80084-9.
 46. Rashid, F.A.N.; Suriani, N.S.; Nazari, A. Kinect-Based Physiotherapy and Assessment: A Comprehensive Review. *Indonesian Journal of Electrical Engineering and Computer Science* **2018**, *11*, 1176–1187. doi:10.11591/ijeecs.v11.i3.pp1176-1187.
 47. Sumner, J.; Lim, H.W.; Chong, L.S.; Bundele, A.; Mukhopadhyay, A.; Kayambu, G. Artificial Intelligence in Physical Rehabilitation: A Systematic Review. *Artificial Intelligence in Medicine* **2023**, *146*, 102693. doi:10.1016/j.artmed.2023.102693.
 48. Abdullah, N.Y.; Al-Kazzaz, S.A. Evaluation of Physiotherapy Exercise by Motion Capturing Based on Artificial Intelligence: A Review. *Al-Rafidain Engineering Journal (AREJ)* **2023**, *28*, 237–251.
 49. Ekambaram, D.; Ponnusamy, V. AI-assisted Physical Therapy for Post-injury Rehabilitation: Current State of the Art. *IEIE Transactions on Smart Processing & Computing* **2023**, *12*, 234–242. doi:10.5573/IEIESPC.2023.12.3.234.
 50. Momin, M.S.; Sufian, A.; Barman, D.; Dutta, P.; Dong, M.; Leo, M. In-Home Older Adults' Activity Pattern Monitoring Using Depth Sensors: A Review. *Sensors*, 2022, Vol. 22, pp. 9067–9067. doi:10.3390/s22239067.
 51. Blanco, J. *Development of Scientific Applications with the Mobile Robot Programming Toolkit—The MRPT Reference Book*; 2010. [cited 2017 Nov 11].
 52. MRPT. MRPT—Empowering C++ Development in Robotics, 2017. [cited 2017 Nov 11].

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.