*pediatric reports*

MDPI

*Systematic Review*

# Markerless Motion Capture for Human Movement Estimation Using Artificial Intelligence: A Systematic Review

**Georgina Domènech-Garcia [1,2,*], Xavier Marimon [3,4], Andoni Carrasco-Urribarren [1], Alejandro E. Portela [5] and Caritat Bagur-Calafat [1,*]**

1 Department of Physiotherapy, Universitat Internacional de Catalunya (UIC), 08195 Barcelona, Spain; acarrasco@uic.es
2 Fundació Aspace Catalunya, 08038 Barcelona, Spain
3 Department of Strength of Materials and Structural Engineering, Universitat Politècnica de Catalunya (UPC-12 Barcelona TECH), 08028 Barcelona, Spain; xavier.marimon@upc.edu
4 Institute for Research and Innovation in Health (IRIS), 08028 Barcelona, Spain
5 Bioengineering Institute of Technology, Universitat Internacional de Catalunya (UIC), 08195 Barcelona, Spain; aeportela@uic.es
* Correspondence: gdomenech@uic.es (G.D.-G.); cbagur@uic.es (C.B.-C.)

## Abstract

**Background**: Artificial intelligence (AI)-driven markerless motion capture (MMC) technologies are increasingly being integrated into pediatric healthcare to improve the assessment and management of movement disorders. These video-based systems enable non-invasive motion analysis without wearable sensors, facilitating more natural movement assessment in children, particularly those with neurological or developmental conditions. **Objectives**: We evaluated the clinical applicability of AI-based MMC tools in pediatric settings for diagnosis, monitoring of motor development, and rehabilitation. **Methods**: This systematic review was registered in PROSPERO (CRD42024511787) and conducted by two independent reviewers, with a third reviewer resolving disagreements. The literature published between 2018 and 2025 was systematically searched. Studies involving pediatric populations or clinically relevant pediatric applications of MMC were included. **Results**: Of 1521 identified studies, 52 were finally selected. The included studies evaluated populations across a wide age range. However, seven of the included articles were specifically focused on underage populations. Infant studies primarily analyzed whole-body movements, emphasizing the relevance of global motor patterns in early development. OpenPose and AlphaPose were the most frequently used frameworks in pediatric research because of their automatic full-body key point detection, whereas DeepLabCut was commonly selected for its customizable labeling capabilities. Theia3D emerged as a promising clinically applicable solution with high accuracy. Most studies evaluated kinematic parameters as objective markers of motor performance and development. However, methodological heterogeneity and limited pediatric-specific validation remain important limitations. **Conclusions**: AI-driven MMC technologies show considerable potential to support objective, accessible, and child-friendly movement assessment in pediatric clinical practice.

**Keywords:** artificial intelligence; motion capture; data accuracy; markerless

Check for updates

Academic Editor: Maurizio Aricò

Received: 9 May 2026
Revised: 16 June 2026
Accepted: 17 June 2026
Published: 23 June 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

Artificial intelligence (AI) has become a hot topic of discussion in recent times. As technology continues to advance, AI has become an area of interest not only for experts

but also for the public. However, there are still misconceptions about AI among clinicians. AI was conceived by McCarthy at the Dartmouth conference in 1956 as "the science and engineering of making intelligent machines" [1].

AI encompasses various subdisciplines, with machine learning being a prominent one [2]. It utilizes extensive datasets to discern interaction patterns among variables. Deep learning (DL), a subset of machine learning, emulates the neural operations of the human brain through multiple layers of artificial neuronal networks. This AI technique yields automated predictions from training datasets and finds compelling applications in image recognition, among other domains [3].

Among the diverse array of applications in everyday human activities, spanning financial services [4], automotive field [5,6], smart home technologies [6], and more, the incorporation of AI into healthcare stands out as holding tremendous potential. Research in this specific application predominantly focuses on addressing challenges related to cancer [7], nervous system disorders [8], and cardiovascular diseases [9], given their significant impact on disability and mortality rates. However, numerous articles have showcased promising advancements in utilizing AI for specific medical assessments [10–15] and training purposes [16].

Clinicians are looking for tools or methods to assess the progress of their patients easily and objectively. Nowadays, popular objective movement assessments systems are being developed in large laboratories equipped with extensive instrumentation [17]. Thanks to these AI techniques, low-cost alternatives are possible. Finding a robust and accurate way to measure human motion has always been the focus of researchers, and one of the great advantages of the use of DL-based methods is that they are extremely flexible and allow researchers to define what to track [18].

Motion estimation is a process used in computer vision and image processing to determine the motion of objects, features, or regions between consecutive frames in a sequence, such as a video or an image series. The integration of AI, particularly deep learning, has revolutionized motion estimation, offering promising advancements and future possibilities, like improved accuracy and robustness and end-to-end learning, as deep learning models can be trained end-to-end by directly learning to estimate motion from raw data without relying on handcrafted features [19].

For motion estimation, imaging is perhaps the most common and widely used because it allows non-invasive, high-resolution motion observations in a variety of settings [20,21]. Beyond estimating motion, AI can infer high-level semantics, like identifying specific actions, understanding a possible diagnostic, documenting disease progression, and evaluating treatment outcomes [22]. However, recording motion analysis in a special laboratory may not accurately reflect the actual movement patterns of individuals. These factors include the artificial nature of the setting, the influence of wearing specialized sensors or markers, and the potential for subjects to alter their behavior due to being observed, often referred to as the Hawthorne effect. Moreover, individuals may exhibit different behaviors in an unfamiliar laboratory or simulated environment during motion capture [23].

Recording motion analysis can be very expensive because it usually requires specialized laboratory equipment, like multiple high-speed cameras, reflective markers, and infrared (IR) cameras. These setups also need a controlled environment and trained professionals to operate and maintain them, which adds to the cost. A more affordable solution is to use an optic method that involves just a single video camera paired with a computer vision algorithm based on AI. This method can capture and analyze motion using standard video footage, without the need for expensive equipment or a controlled setting. Lam et al. 2023 [24] propose a solution to the issue above by advocating for the use of a smartphone camera in conjunction with an algorithm for analysis. They suggest that this combination

could offer a feasible approach for evaluating patients' daily movements through MMC, utilizing the capabilities of a smartphone and an advanced algorithm.

Therefore, in this review, we emphasize exploring the optimal MMC technology as a crucial assessment tool in human motion analysis. Specifically, we focus on the AI software and algorithms employed in human biomechanical analysis and examine their key features. Additionally, we investigate the accuracy, clinical utility, and ease of use of these technologies in various contexts.

## 2. Materials and Methods

### *2.1. Search Strategy*

This review was registered in the PROSPERO database (CRD42024511787). A comprehensive literature search was conducted across the following databases: Web of Science (WOS), PubMed/Medline, and IEEE Xplore. This systematic review was conducted in accordance with PRISMA guidelines (The PRISMA checklist is provided as Supplementary Materials (Table S2).

The equation search was the following: ("motion analysis" OR "movement analysis" OR "biomechanical assessment" OR "movement assessment" OR "motion tracking" OR "kinematic assessment" OR "motion capture" OR "motion estimation" OR "motion capture technology" OR "Computer Vision" OR "Video-based" OR "Pose Estimation") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND (markerless OR detectorless OR clusterless OR label-free). The final searching strategy was conducted up to 2 January 2026.

### *2.2. Inclusion and Exclusion Criteria*

Studies were included if they met the following criteria: (1) presented any markerless motion capture AI algorithm for diagnosis or evaluation; (2) were published between 2018 and 2025; (3) were randomized or non-randomized clinical trials, validation studies, or observational studies, whether pilot or full-scale; and (4) were published in English or Spanish.

Studies were excluded if they met any of the following criteria: (a) involved non-human subjects; (b) were not markerless; (c) did not report accuracy or evaluation metrics; or (d) were of a different study type. Specifically, criterion (a) referred to studies involving animals or cadavers. Criterion (b) included studies that used external devices, relied on bounding-box detection rather than joint motion tracking, or employed multiple cameras or depth cameras. Criterion (d) referred to studies investigating a different topic or whose design did not correspond to the predefined eligible study types.

### *2.3. Selection Process*

The studies collected by the searching strategy were screened by two independent authors (XMS, GDG). In cases of discrepancies, authors not involved in the initial screening acted as arbitrators.

A total of 1521 articles were identified, with 1491 from databases search and 30 from citation searching. Duplicate records were removed using Mendeley Desktop v. 1.19.8 Reference Manager (Elsevier, Amsterdam, The Netherlands). According to their suitability, eligible articles were examined by title, abstract, and full text. The reasons for excluding studies were recorded. The selection process is summarized in the PRISMA flowchart in Figure 1.

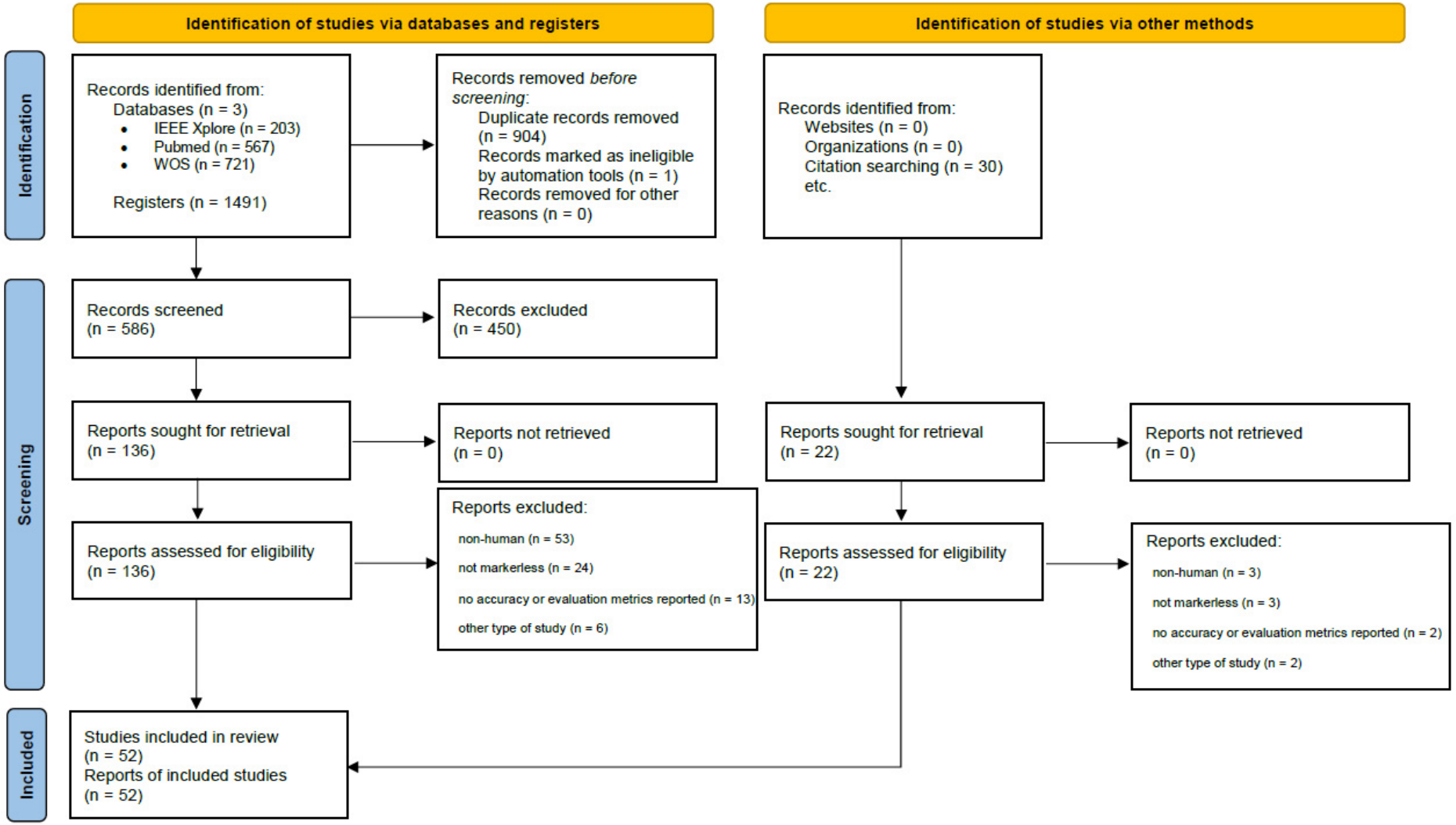



**Figure 1.** PRISMA 2020 flow diagram of the systematic review, which included searches of databases, registers, and other sources. We considered, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers) [25].

### *2.4. Data Extraction and Analysis*

Two reviewers were responsible for data extraction (XMS, GDG). A tabulated worksheet in Microsoft Excel (see Table S1 in Supplementary Materials) was used for the registration of relevant data from all studies included in this review. It includes the following research variables: (a) characteristics of the population; (b) characteristics of the video registration software and (c) AI motion tracking algorithm used, as well as (d) the accuracy results and (e) kinematic features; and (f) points of interest (PoI).

Of these, characteristics of the population refer to their age and health condition; characteristics of the video registration refer to its global dataset size, the AI motion tracking algorithm used, its training methodology, camera resolution, sampling frequency, the accuracy metrics reported, and the kinematic features, such as joint angles or velocity; and PoI refers to the analyzed body parts, such as upper or lower extremities, trunk, and whole body.

## 3. Results

A total of 1491 studies were found in the searches carried out in the WOS, PubMed/Medline, and IEE Xplore databases. In addition, 30 studies were found that could be included through other means, such as reviewing bibliographic references. A total of 904 duplicate titles and one other that could not be valued correctly were eliminated. Finally, after reading the title and abstract and applying the selection criteria of the present review, a total of 563 articles were eliminated, including 52 studies in Table S1 in Supplementary Materials.

Out of all the excluded articles, the majority focused on detecting laboratory characteristics like cancer diagnosis, which the authors deemed off topic. Among those that analyzed physical human motion tracking, various physical sensors were used, such as exoskeletons, robotic devices, or wearables, thus not meeting the criteria for being fully markerless. Others did not emphasize joint motion tracking but rather viewed participants as whole bodies (bounding box approach). Additionally, some studies employed multicamera setups and/or depth cameras, which were not addressed in this article because they are not accessible in standard rehabilitation clinics using everyday consumer electronics. Moreover, many novel MMC systems are still in early development and reported mainly in conference proceedings, with limited validation within the scientific community.

### *3.1. Demographic Characteristics*

The articles encompassed a broad age range, spanning from newborns to older individuals, with some studies focusing on specific age groups. Most of the studies (65%) included adult participants, whereas others included participants spanning multiple age groups, from children to adults [12,15,26–28]. Only a few studies were conducted exclusively with underage participants—infants [13,29–32] and children [33,34]. In the end, seven of these did not report the age range.

Nearly half of the articles (46%) examined motion tracking in healthy individuals. The remaining studies focused on populations with neurological conditions, including cerebral palsy [12,15,27,34,35], orthosis [36], Parkinson's Disease [37–39], Multiple Sclerosis [40], preterm [13,30], human immunodeficiency virus encephalopathy [33], and post-stroke [41–45]. Other studies investigated populations with traumatic or musculoskeletal conditions, such as osteoarthritis [46,47]. In the end, 15% of the articles did not provide information about the health status under consideration.

### *3.2. Points of Interest*

Considerable variability was observed in the regions of interest examined across studies. Most studies (41%) analyzed full-body kinematics, encompassing the entire body from the forehead to the great toe, and the lower extremities alone represented the second-most studied region (33%). Other studies focused specifically on the distal upper extremities [37,48,49] or both distal upper and lower extremities [38]. Some examined the trunk in combination with the upper extremities [27,50] or with the lower extremities [51–55]. In the end, 6% of the studies did not report their points of interest.

The following graph in Figure 2 illustrates a trend where studies focusing on younger populations tend to analyze the entire body, due to their general movement objectives [13,29–33], whereas research involving older populations increasingly focuses on specific parts of the body or points of interest (trunk and upper or lower extremity for gait analysis, for example).

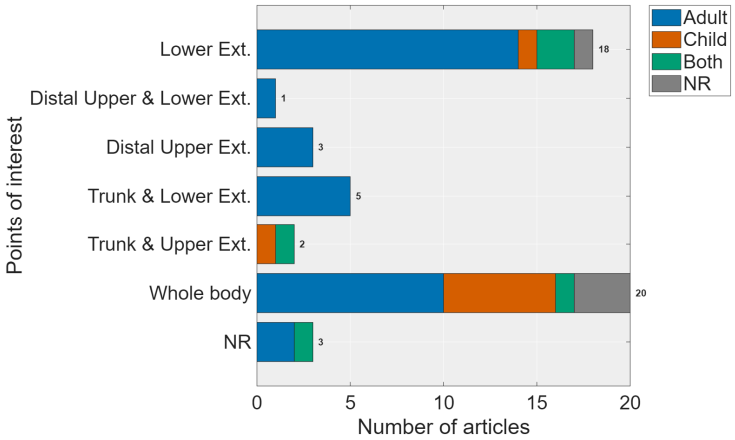



**Figure 2.** Number of articles describing different points of interest in each age group; Note: Ext.: extremity; NR: not reported.

### *3.3. Recording Specifications*

The range of camera resolutions employed in the studies varied, with the highest resolution set at 2704 × 1520 pixels [56] and the lowest at 520 × 520 pixels [39]. Most of the studies report 1920 × 1080 pixels while nearly half of them do not mention it. A diverse range of cameras was utilized in the research, including RGB cameras, smartphone cameras, and other types of cameras. Regarding the sampling frequency (Fs) from those that reported it, this was within a range of 25 Hz [15,26,27] up to 250 Hz [32], although 100 Hz was the most used (19%).

Although not all studies provided explicit details, there was variability in camera positioning, and it was adjusted according to the specific object/subject being recorded. While not all articles provided details on the dataset size of the collected videos, most authors separated them into training and validation datasets, with 60–90% of the collected data used for training and 10–20% of the remaining data used for validation. The smallest dataset size observed in the studies was six videos [57] and the highest 1140 videos [53].

### *3.4. Kinematic Feature Extraction*

The kinematic features analyzed in the research varied according to the specific aims of the studies. Most studies (63%) focused on joint angle variations, followed closely by joint maker trajectories (50%). Other derived measures were also investigated, including joint velocity [12,27,39,58,59], joint acceleration [39], segment length [51], joint positions [32,51], and joint distances to a middle point of a reference line [26], which are features calculated mainly for shoulders, elbows, and wrists. One article analyzed the interlimb correlation [30].

The authors suggest the following features groups, extracted from all the articles features included in this review:

- Temporal regularity (T.R.): These time domain features measure timing and coordination aspects of movement, indicating how synchronized different body parts or movements are (mean interpeak intervals, interpeak variability, rhythm regularity).
- Joint marker trajectories (J.M.T.): These measures relate to a joint's or body part's spatial positioning and movement (joint position, distance to middle point, distance to line, trajectory deviation).
- Joint angles (J.A.): These measures pertain to the body's segments of lengths and spatial relationships, which are directly related to joint angles (joint angle, segment length, mean amplitude, amplitude, variability, joint orientation).
- Interlimb correlation (I.C.): These measures refer to the relationship and coordination between the movements of different limbs (interlimb synchronization, interlimb complexity).
- Derivate measures (D.M.): These measures are more complex and derived from basic movement data, often reflecting higher-order characteristics of movement patterns (velocity, joint angular velocity, acceleration, curvature).

At the processing level, raw data frames are typically used. While most authors in this review work with raw data, some perform preprocessing through median filtering [51] and Gaussian filtering [30].

Statistical features of the data are also exploited to measure the interpeak variability of kinematic variables, such as mean amplitude, mean interpeak intervals, amplitude variability, and interpeak intervals variability [30].

Surprisingly, time and frequency domain features are not used in any of the publications examined in this systematic review. Only one publication used features based on the frequency domain [60]. A few articles were identified that used features based on non-linear measurements, chaos theory or information theory, and Sample Entropy measures [30]. Of interest are the features proposed by McCay et al. [60], where the kinematic parameters are encoded in 2D histograms, such as histograms of joint orientation and histograms of joint displacement. It is also interesting to note how Guayacan et al. proposed their results, where the curvature is encoded by a polar histogram scheme (dartboards) representing the magnitudes and directions of each feature [39].

### *3.5. Motion Estimation Algorithms*

The predominant algorithm employed for motion estimation was the Convolutional Neural Network (CNN), utilized in 70% of the studies, where the type of network used was mainly ResNet [38] and FCNet [61] architecture. Several studies did not specify the deep learning algorithm employed [47,52,54,62–65]. Other approaches included Random Forest [59], Transformer [29,55], and ST-GCN [13]. Two studies did not provide information regarding the AI techniques used [46,66,67]. None of the articles reported the application of a dimensionality reduction technique to the data.

Regarding software tools, DeepLabCut (DLC) was the most used software (26%) followed by OpenPose (19%) and Theia3D (15%). Additional software solutions included MediaPipe [48,49,68–70], DensePose [39], AlphaPose [30,71], BlazePose [50], and KinaTrax [42,72]. Seven studies either customized the algorithm without specifying the tool or did not report this information. Moreover, Python is the primary programming language of such algorithm tools.

Figure 3 presents which algorithm was used depending on health conditions being analyzed. It can be appreciated that DLC was the most used for its key points customization characteristics and OpenPose was the most used for its high range of accuracy. More-

over, Theia3D has gained prominence, providing end-to-end markerless motion capture pipelines that combine deep learning-based 2D key point detection with multi-view 3D reconstruction and biomechanical model fitting.

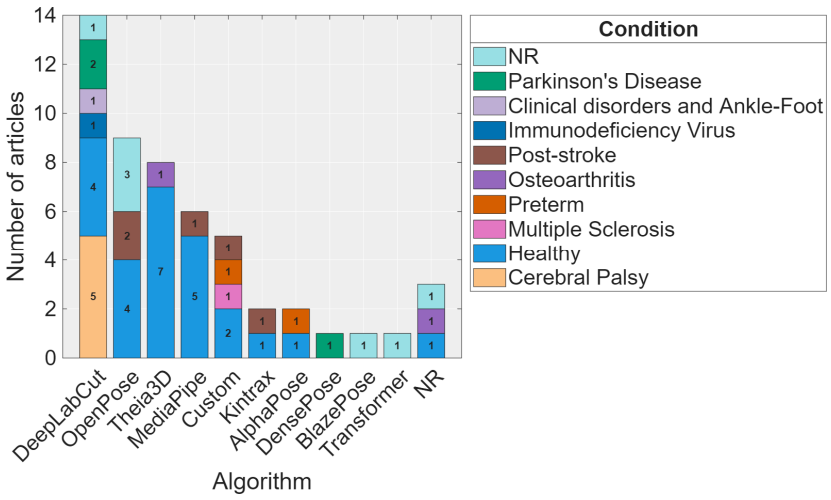



**Figure 3.** Algorithm distribution for each health condition. Note: number indicate the number of studies in which each algorithm was applied for each health condition.

Moreover, the temporal evolution in the use of each algorithm is presented in Figure 4. In 2024, there was an exponential increase in the number of newly introduced algorithm tools. Among them, Theia3D demonstrated the most substantial growth.

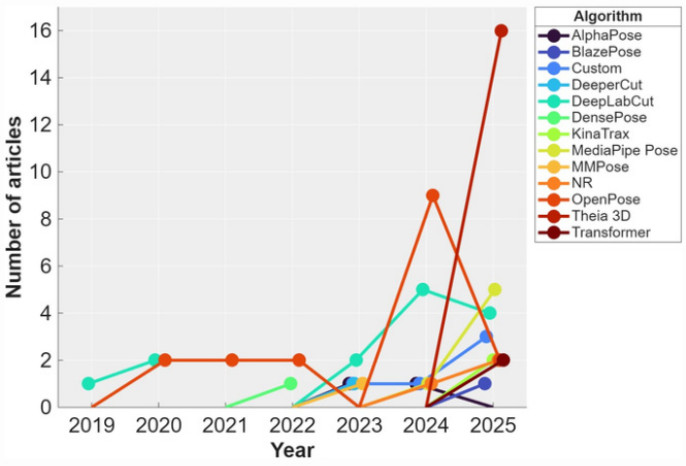



**Figure 4.** Evolution of algorithm usage.

This growth of algorithms usage is accompanied by their variability in accuracies, as represented in Figure 5. In the beginning, accuracies were reported in different ways: as accuracy percentages, r-squared (R2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and F-1 score. Accuracy was the most frequently used. Specifically, Shin HI. et al. reported their results with the r-square metric of 0.93 [30]. Another group of articles reported results in RMSE (2.93 pixels) [51] and others reported in MAE, ranging from 0.14 [26] up to 8.39 pixels [37].

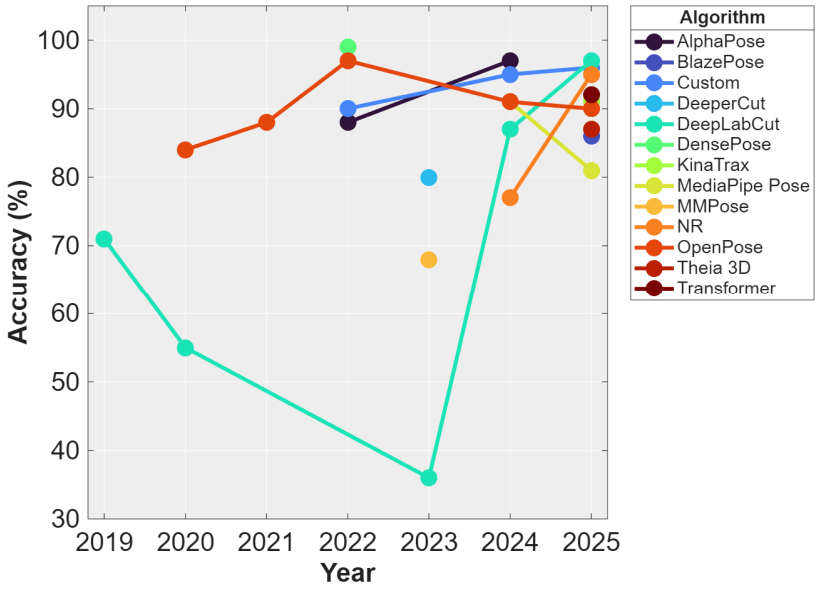



**Figure 5.** Evolution of algorithms accuracies. Note: Reported accuracy values should be interpreted with caution, as the underlying studies were not assessed for proper training methodology or healthy learning dynamics; thus, reported values may reflect overfitting or other methodological artifacts rather than true model generalizability.

### *3.6. Motion Estimation Accuracies and Their Training Method*

Most reviewed studies did not undertake active model training, localized fine-tuning, or architectural optimization. Instead, they evaluated established pre-trained frameworks or commercial platforms operating in inference-only modes (e.g., based on COCO or MPII datasets), or they provided no information regarding training procedures. Only a small subset of studies implemented active machine learning processes involving model training or adaptation using the acquired data. [13,15,32,36,51,55,56,59,67,69].

Moreover, none of the articles provided evidence of learning curves while some of them selected transparent hyperparameters, either fully or partially [12,13,15,26,34,39,46,50,55,56,63,67,68,73–76].

However, when considering only studies employing appropriate training methodologies and reporting accuracies in a common metric (percentage), DensePose still achieves a higher percentage (99,62%), followed by AlphaPose (88–95%), Theia3D (70–95%), and OpenPose (84–88%). DLC had a major range of accuracy levels, from 93% up to 22.7%. Finally, DeeperCut was 80% accurate.

These accuracies are held depending on the target age group, as demonstrated in Figure 6. Overall, most algorithm applications were predominantly evaluated in adult populations, where the highest concentration of studies and the widest range of reported accuracies were observed. In adults, several algorithms demonstrated high performance (DensePose, 99%) and achieved consistently strong results. Notably, DeepLabCut was one of the few tools consistently applied across multiple age categories (young, adults, and mixed samples), maintaining accuracies above 74% in all cases. In mixed-age samples, however, DLC showed a lower reported accuracy (65%).

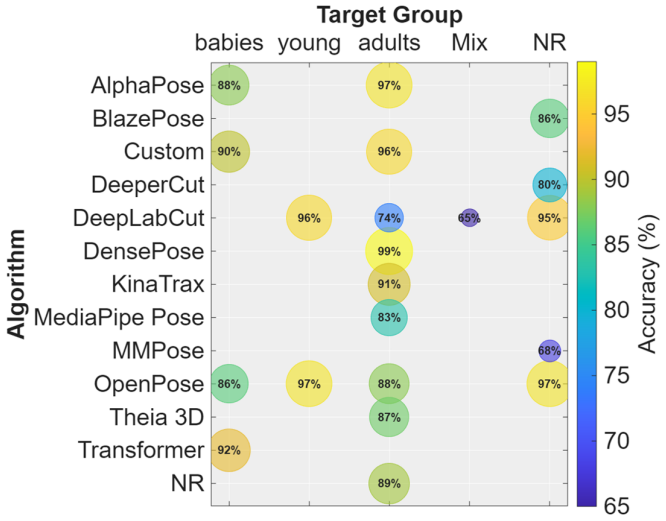



**Figure 6.** Accuracy distribution stratified by target age group. Note: Reported accuracy values should be interpreted with caution, as the underlying studies were not assessed for proper training methodology or healthy learning dynamics; thus, reported values may reflect overfitting or other methodological artifacts rather than true model generalizability.

In younger populations, OpenPose [32] and AlphaPose [30] (88%) achieved the highest reported accuracies in infant populations, whereas DeepLabCut was used in a children population, achieving a 96% of accuracy [33].

### *3.7. Validation*

A significant portion of the articles outlined various validation methods. The k-fold-cross validation method was widely utilized [13,39], using mostly k = 5 folds of cross-validation. Additionally, studies commonly assessed inter-rater agreement between two raters and compared manual versus automatic labeling. [26,32]. Other approaches involved comparisons with an external device [38]; alternative motion tracking systems, such as Vicon [64]; other markerless algorithms [29]; the application of various traditional machine learning classifiers [31]; or evaluation against a CNN [37].

## 4. Discussion

Recent advances in AI have significantly expanded its application in various scientific fields, one of which is markerless pose estimation. This trend has been partly driven by the availability of open-source software. Despite these advancements, the use, comprehension, and interpretation of these technologies are not universally accessible. Therefore, this paper focuses on evaluating the performance and accuracy of various algorithm tools used in markerless pose estimation.

The present review compares several algorithm tools specifically developed for markerless human pose estimation, highlighting their distinctive features and suitability according to different research objectives. It includes articles from 2019 to 2025, as AI markerless technology is relatively new, with notable improvements in quality and quantity starting in 2024.

Across the included studies, most analyses focused on either the whole body or the lower extremities. In pediatric populations, assessments predominantly targeted whole-body motion, whereas in adults, evaluations covered a broader range of body segments (upper limb, lower limb, trunk, or full body). This distribution may reflect the clinical focus in children, where general movement patterns and global motor development are typically assessed. In such cases, predefined full-body markerless models may offer greater efficiency

and practicality. Conversely, adult studies often investigate more specific biomechanical questions, requiring segmental or joint-specific analyses.

Regarding the populations studied, most participants were healthy individuals. Nevertheless, clinical populations were also represented, with cerebral palsy (CP) and post-stroke conditions being the most frequently investigated pathologies. This indicates a growing interest in validating markerless systems within neurorehabilitation and movement disorder contexts.

Considering the inclusion criteria—specifically studies employing purely markerless approaches without depth cameras—DeepLabCut (DLC) emerged as the most frequently used algorithm, followed by Theia3D. The versatility of DLC allows researchers to define and customize key points according to specific experimental needs, making it particularly suitable for tailored research applications. As one of the earliest and most widely adopted deep learning-based tools for markerless pose estimation, DLC has been extensively applied in biomechanics research over the past few years [77].

OpenPose supports full-body pose estimation, including facial landmarks, hand key points, and trunk key points, making it a versatile solution for comprehensive assessments [78]. Its continued use is largely attributed to its open-source framework and multi-person detection capabilities [79]. DensePose differs conceptually by providing dense 3D surface mapping rather than sparse anatomical key points, enabling more detailed surface-based analyses [39]. AlphaPose integrates symmetric heatmap aggregation and pose-guided proposals, achieving a favorable balance between precision and stability [80]. MediaPipe Pose, developed by Google, has demonstrated high sensitivity in clinical contexts and offers efficient real-time performance [81]. BlazePose and KinaTrax emerged more prominently in 2024 and are still under evaluation within the research community.

More recently, commercially available systems such as Theia3D have gained prominence, providing end-to-end markerless motion capture pipelines that combine deep learning-based 2D key point detection with multi-view 3D reconstruction and biomechanical model fitting [82].

When interpreting the reported accuracies of each algorithm, it is essential to consider the sample size and the characteristics of the dataset analyzed. Accuracy values may be influenced by the number of participants and the heterogeneity of the sample. In small cohorts, movement patterns may appear more homogeneous, potentially simplifying kinematic detection and increasing apparent accuracy. However, as sample size increases, inter-individual variability in anthropometrics, motor strategies, and movement quality also increases, which may introduce greater complexity and reduce performance consistency. In the reviewed studies, sample sizes ranged considerably from as few as five participants [48] to as many as 1176 individuals [55], which should be considered when comparing reported outcomes.

DeepLabCut (DLC) exhibited the widest variability in reported accuracies, particularly in studies including mixed-age populations, with values ranging from 17% to 99%. This broad range suggests potential sensitivity to heterogeneous datasets and diverse movement characteristics. Notably, DLC was the only algorithm consistently applied across such a wide age spectrum, reflecting researchers' confidence in its adaptability and customizable framework [12,15,26–28].

OpenPose maintained relatively stable and generally high accuracy values over time, ranging from 68% to 99% [45,57] and demonstrating robustness across different study designs and populations. Overall, most algorithm tools reported accuracy values above 75%, indicating that markerless pose estimation has reached a level of performance that is broadly acceptable for many biomechanical and clinical research applications, although variability persists depending on population characteristics and methodological factors.

The authors recommend comparing the evaluations of two raters and assessing the differences between manual and automatic labeling. This recommendation is based on the opinion of the clinician writing, emphasizing the ease of accessibility for clinical use and the parallels observed in other types of studies. Moreover, some algorithm tools still lack sufficient clinical and comprehensive validation; therefore, further studies are needed to establish clear criteria for their optimal use.

Comprehensive details regarding the number of full-length videos utilized, their durations, and the allocation between training and validation sets were often lacking in the reviewed studies. The predominance of datasets used for training underscores the necessity for more consistent reporting practices in research. Furthermore, the absence of a standardized accuracy reporting format has led to a wide range of evaluation metrics being employed, complicating direct comparisons across studies. However, all algorithms reviewed demonstrate excellent accuracy, with accuracy percentages > 80%, R2 values > 90%, and F1-scores > 0.9, indicating high effectiveness. Given the comprehensive review of the literature, if the authors were tasked with recommending an evaluation metric for markerless motion estimation using AI for clinicians, the suggestion would be to adopt percentage accuracy as the gold standard measure. This recommendation is based on the widespread use of percentage accuracy in various studies, indicating its widespread acceptance and applicability in the field. Moreover, the familiarity and simplicity of percentage accuracy make it easily interpretable for both researchers and clinical practitioners, enhancing the accessibility and utility of study.

The most used features are time domain measurements, i.e., joint-related kinematic features. Frequency domain or time-frequency measurements are absent. Non-linear measurements are very few and are based on entropy. In this review, it becomes evident that AI-based models, particularly CNNs, are highly effective at learning complex patterns in motion data. CNNs excel at processing and interpreting spatial information, making them adept at handling visual inputs and in recognizing intricate details despite variations in lighting or partial occlusions. These CNN models are less affected by noise and environmental changes compared to traditional methods. Their ability to learn and adapt from diverse and noisy data leads to more precise motion estimation. Consequently, the use of CNNs in motion analysis results in improved accuracy and robustness, enabling more reliable and consistent performance in various real and clinical conditions.

In the field of motion analysis, AI offers capabilities that go beyond simple motion estimation by enabling the extraction of high-level semantic information. This advanced functionality allows AI systems to not only track and quantify motion but also recognize and categorize specific actions and behaviors, which is particularly valuable for diagnostics, identifying abnormalities in motion patterns that may indicate the presence of certain health conditions. Overall, AI's ability to interpret movement data at a high level of abstraction enhances its utility in clinical practice, offering deeper insights that support accurate diagnosis, effective monitoring of disease progression, and optimal evaluation of treatment efficacy.

Particularly in pediatric populations, AlphaPose, OpenPose, and DeepLabCut are among the most widely used pose estimation frameworks due to their distinct methodological characteristics. AlphaPose and OpenPose provide predefined full-body key point models, whereas DeepLabCut offers a flexible, customizable framework for tracking user-defined anatomical landmarks. In babies, research typically focuses on full-body analysis, whereas in children and adolescents, studies more often concentrate on specific body segments or functional tasks, such as gait or grasping. Therefore, the selection of these tools generally depends on the study objectives, particularly on whether the focus is on whole-body motor coordination or on the movement of specific body segments.

Deep learning models offer a significant advantage in motion estimation by enabling end-to-end training directly from raw data. This means that these models learn to estimate motion without the need for manually designed features or preprocessing steps. Instead, the models automatically extract relevant features and patterns from the raw input data, such as images or video frames.

End-to-end learning facilitates real-time or live-motion estimation by processing raw data directly through a unified model, which enables us to rapidly interpret and analyze data, delivering motion estimates with minimal delay. The model's ability to efficiently handle diverse conditions and variations ensures accurate performance even in dynamic environments. By integrating all processing tasks into a single model, it reduces latency and provides timely results, making it ideal for live diagnostics.

Unfortunately, the use of AI among clinicians is not yet widely standardized due to its complexity and challenges in understandability. This complexity makes it difficult for healthcare professionals to integrate AI into their workflows effectively. The challenge in the future lies in making AI outputs clear and actionable for clinicians who may not have a deep technical background. As a result, the widespread use of AI in clinical settings remains uneven, with many healthcare providers struggling to fully leverage its capabilities.

Moreover, to effectively integrate AI into clinical practice, the authors recommend making clear which PoI is intended to analyze and the age and health condition of this population before selecting an algorithm for motion analysis. It is important that AI systems present results in a way that clinicians can easily understand. This means designing AI outputs to be straightforward and actionable, so that healthcare professionals can quickly interpret and use the information in their decision-making processes. In addition, having medical engineers or data scientists as part of the clinical team can address more complex aspects of data analysis. These professionals code and train complicated AI models and bridge the gap between advanced AI algorithms and practical clinical applications. Their expertise ensures that AI findings are translated into actionable medical insights, increasing the overall utility of AI in healthcare.

Lam W et al. (2023) conducted a comparable review using different inclusion criteria (e.g., depth-sensing Kinect cameras) [24]. They concluded that MMC technology holds promise both as an assessment tool and symptom detection, potentially supporting AI-based early disease screening. Our review builds on this work, confirming their conclusions and providing deeper, software-specific analysis [24].

A major limitation identified across the reviewed literature is the widespread lack of transparent reporting regarding the underlying learning dynamics of the deployed machine learning algorithms. In studies focused on model training or customization, final accuracy metrics cannot be considered inherently trustworthy indicators of generalizability without explicit evidence of proper training methodologies, such as data separation protocols, evaluation of learning curves, and robust hyperparameter selection. High accuracy outcomes can frequently mask methodological errors, including data leakage between training and testing sets or severe overfitting to specific laboratory settings. Furthermore, for those studies that evaluate pre-trained off-the-shelf systems, the reported validity metrics remain deeply dependent on the demographic and contextual characteristics of the original training repositories (e.g., COCO or MPII datasets). Consequently, readers and clinicians must interpret aggregate performance metrics with appropriate caution, as the available literature rarely provides a standardized auditing of a model's internal learning dynamics or resistance to overfitting artifacts.

Overall, the field is evolving rapidly, with a clear shift toward more automated, computationally efficient, and clinically adaptable deep learning-based markerless motion capture systems. High quality systematic reviews of this topic should be conducted

periodically and include a broader range of databases to ensure the literature remains up to date and its clinical application.

## 5. Conclusions

Among the algorithms analyzed, DeepLabCut was the most frequently utilized algorithm across the reviewed studies in both pediatric and adult populations. OpenPose often demonstrated higher and more stable accuracy values and was more frequently applied in pediatric populations. Moreover, Theia3D is also emerging as a prominent commercial solution, benefiting from its reported high accuracy and integrated end-to-end commercial workflow. However, comparisons of reported algorithm performance across studies should be interpreted with caution because of methodological heterogeneity, differences in datasets, outcome measures, and evaluation procedures, as well as the limitations discussed throughout the manuscript. These factors may influence reported accuracies and hinder direct comparisons between studies. Furthermore, the present review did not assess whether the included studies employed appropriate training methodologies or healthy learning dynamics, as this was beyond the scope of the analysis.

In line with these limitations, the comparison of accuracy metrics across studies remains challenging due to the heterogeneity of outcome measures. Different units (e.g., pixel error, millimeters, mean per-joint position error, percentage of correct key points, correlation coefficients) limit direct comparability between systems. Therefore, there is a clear need for methodological standardization. Based on the findings of this review, adopting percentage accuracy as a common reporting metric for AI-based markerless motion estimation may improve interpretability and facilitate comparisons across studies. Percentage-based metrics are widely understood, relatively intuitive, and potentially more accessible for both researchers and clinical practitioners.

In conclusion, AI-driven markerless pose estimation demonstrates strong potential and is evolving at a rapid pace. Nevertheless, a significant gap remains between technological development and clinical implementation and validation. Many clinicians may lack the technical background required to critically interpret AI-derived kinematic outputs. Integrating medical engineers or data scientists into clinical and research teams could facilitate data interpretation, improve methodological rigor, and accelerate translation into practice. Continued interdisciplinary collaboration will be essential to ensure that advances in markerless motion capture meaningfully benefit research and clinical care.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/pediatric18040083/s1, Table S1: Extraction table [10,12,13,15,20,26–29,32–34,36–45,47–57,59,60,62,64,65,67–75,80,83–86]; Table S2: PRISMA checklist.

**Author Contributions:** Conceptualization, C.B.-C., A.E.P., G.D.-G. and X.M.; methodology, C.B.-C., A.C.-U., G.D.-G. and X.M.; formal analysis, G.D.-G. and X.M.; investigation, G.D.-G. and X.M.; resources, G.D.-G. and X.M.; data curation, G.D.-G. and X.M.; writing—original draft preparation, G.D.-G. and X.M.; writing—review and editing, C.B.-C., G.D.-G. and X.M.; supervision, C.B.-C., A.E.P., X.M. and A.C.-U. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created.

**Acknowledgments:** We gratefully acknowledge the support of the libraries at Universitat Internacional de Catalunya and Universitat Politecnica de Catalunya. The resources provided by both institutions were essential to the completion of this research.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| AI | Artificial intelligence |
| MMC | Markerless motion capture |
| DL | Deep learning |
| Ext | Extremity |
| NR | Not reported |
| TR | Temporal regularity |
| JMT | Joint marker trajectory |
| JA | Joint angle |
| DM | Derivative measure |
| CNN | Convolutional neural network |
| R2 | Root square |
| RMSE | Root mean squared error |
| MAE | Mean absolute error |
| DLC | DeepLabCut |
| COCO | Common objects in context |
| MPII | Max Planck Institute for Informatics |

## References

1. McCarthy, J.; Minsky, M.L.; Rochester, N.; Shannon, C.E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Mag.* **2006**, *27*, 12.
2. Koprowski, R.; Foster, K.R. Machine learning and medicine: Book review and commentary. *Biomed. Eng. Online* **2018**, *17*, 17. [CrossRef]
3. Lee, E.-J.; Kim, Y.-H.; Kim, N.; Kang, D.-W. Deep into the Brain: Artificial Intelligence in Stroke Imaging. *J. Stroke* **2017**, *19*, 277–285. [CrossRef] [PubMed]
4. Bahoo, S.; Cucculelli, M.; Goga, X.; Mondolo, J. Artificial intelligence in Finance: A comprehensive review through bibliometric and content analysis. *SN Bus. Econ.* **2024**, *4*, 23. [CrossRef]
5. Soegoto, E.S.; Utami, R.D.; Hermawan, Y.A. Influence of artificial intelligence in automotive industry. *J. Phys. Conf. Ser.* **2019**, *1402*, 066081. [CrossRef]
6. Stolojescu-Crisan, C.; Crisan, C.; Butunoi, B.-P. An IoT-Based Smart Home Automation System. *Sensors* **2021**, *21*, 3784. [CrossRef] [PubMed]
7. Xi, G.; Wang, Q.; Zhan, H.; Kang, D.; Liu, Y.; Luo, T.; Xu, M.; Kong, Q.; Zheng, L.; Chen, G.; et al. Automated classification of breast cancer histologic grade using multiphoton microscopy and generative adversarial networks. *J. Phys. D Appl. Phys.* **2022**, *56*, 015401. [CrossRef]
8. Fanous, M.; Shi, C.; Caputo, M.P.; Rund, L.A.; Johnson, R.W.; Das, T.; Kuchan, M.J.; Sobh, N.; Popescu, G. Label-free screening of brain tissue myelin content using phase imaging with computational specificity (PICS). *APL Photonics* **2021**, *6*, 76103. [CrossRef] [PubMed]
9. Bhat, S.; Ohn, J.; Liebling, M. Motion-based structure separation for label-free high-speed 3-D cardiac microscopy. *IEEE Trans. Image Process.* **2012**, *21*, 3638–3647. [CrossRef] [PubMed]
10. Park, S.; Bae, B.; Kang, K.; Kim, H.; Nam, M.S.; Um, J.; Heo, Y.J. A Deep-Learning Approach for Identifying a Drunk Person Using Gait Recognition. *Appl. Sci.* **2023**, *13*, 1390. [CrossRef]
11. Haberkamp, L.D.; Garcia, M.C.; Bazett-Jones, D.M. Validity of an artificial intelligence, human pose estimation model for measuring single-leg squat kinematics. *J. Biomech.* **2022**, *144*, 111333. [CrossRef] [PubMed]
12. Huijsmans, H.; Haberfehlner, H.; Bonouvrié, L.A.; Van de Ven, S.S.; Harlaar, J.; Van der Krogt, M.M.; Buizer, A. Markerless motion tracking to assess upper limb dyskinesia in children and young adults with cerebral palsy. *Gait Posture* **2021**, *90*, 106–107. [CrossRef]
13. Groos, D.; Adde, L.; Aubert, S.; Boswell, L.; De Regnier, R.A.; Fjørtoft, T.; Gaebler-Spira, D.; Haukeland, A.; Loennecken, M.; Msall, M.; et al. Development and Validation of a Deep Learning Method to Predict Cerebral Palsy From Spontaneous Movements in Infants at High Risk. *JAMA Netw. Open* **2022**, *5*, e2221325. [CrossRef] [PubMed]

14. Balta, D.; Kuo, H.H.; Wang, J.; Porco, I.G.; Schladen, M.; Cereatti, A.; Lum, P.; Della Croce, U. Estimating infant upper extremities motion with an RGB-D camera and markerless deep neural network tracking: A validation study. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, UK, 11–15 July 2022; pp. 2548–2551. [CrossRef] [PubMed]
15. Haberfehlner, H.; Roth, Z.; Vanmechelen, I.; Buizer, A.I.; Vermeulen, R.J.; Koy, A.; Aerts, J.-M.; Hallez, H.; Monbaliu, E. A Novel Video-Based Methodology for Automated Classification of Dystonia and Choreoathetosis in Dyskinetic Cerebral Palsy During a Lower Extremity Task. *Neurorehabil. Neural Repair* **2024**, *38*, 479–492. [CrossRef] [PubMed]
16. Knippenberg, E.; Verbrugghe, J.; Lamers, I.; Palmaers, S.; Timmermans, A.; Spooren, A. Markerless motion capture systems as training device in neurological rehabilitation: A systematic review of their use, application, target population and efficacy. *J. Neuroeng. Rehabil.* **2017**, *14*, 61. [CrossRef]
17. Nagymáté, G.; Kiss, R.M. Application of OptiTrack motion capture systems in human movement analysis: A systematic literature review. *Recent Innov. Mechatron.* **2018**, *5*, 1–9. [CrossRef]
18. Mathis, A.; Schneider, S.; Lauer, J.; Mathis, M.W. A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron* **2020**, *108*, 44–65. [CrossRef] [PubMed]
19. Mishra, A.K.; Kohli, N. An intelligent optimization algorithm with a deep learning-enabled block-based motion estimation model. *Expert Syst.* **2022**, *39*, e13074. [CrossRef]
20. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **1973**, *14*, 201–211. [CrossRef]
21. O'Connell, A.F.; Nichols, J.D.; Karanth, K.U. *Camera Traps in Animal Ecology: Methods and Analyses*; Springer: New York, NY, USA, 2011; pp. 1–271. [CrossRef]
22. Sambati, L.; Baldelli, L.; Calandra Buonaura, G.; Capellari, S.; Giannini, G.; Scaglione, C.L.M.; Armaroli, M.; Zoni, E.; Cortelli, P.; Martinelli, P. Observing movement disorders: Best practice proposal in the use of video recording in clinical practice. *Neurol. Sci.* **2019**, *40*, 333–338. [CrossRef] [PubMed]
23. Tronick, E.; Als, H.; Brazelton, T.B. Early Development of Neonatal and Infant Behavior. In *Human Growth*; Springer: Boston, MA, USA, 1979; pp. 305–328. [CrossRef]
24. Lam, W.W.T.; Tang, Y.M.; Fong, K.N.K. A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. *J. Neuroeng. Rehabil.* **2023**, *20*, 57. [CrossRef]
25. 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.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
26. Haberfehlner, H.; van de Ven, S.S.; van der Burg, S.A.; Huber, F.; Georgievska, S.; Aleo, I.; Harlaar, J.; Bonouvrié, L.A.; van der Krogt, M.M.; Buizer, A.I. Towards automated video-based assessment of dystonia in dyskinetic cerebral palsy: A novel approach using markerless motion tracking and machine learning. *Front. Robot. AI* **2023**, *10*, 1108114. [CrossRef] [PubMed]
27. Vanmechelen, I.; Van Wonterghem, E.; Aerts, J.M.; Hallez, H.; Desloovere, K.; Van de Walle, P.; Buizer, A.I.; Monbaliu, E.; Haberfehlner, H. Markerless motion analysis to assess reaching-sideways in individuals with dyskinetic cerebral palsy: A validity study. *J. Biomech.* **2024**, *173*, 112233. [CrossRef] [PubMed]
28. Verhoeven, M.; Zandvoort, C.S.; Dominici, N. From marker to markerless: Validating DeepLabCut for 2D sagittal plane gait analysis in adults and newly walking toddlers. *J. Biomech.* **2025**, *186*, 112708. [CrossRef] [PubMed]
29. Gama, F.; Mísař, M.; Navara, L.; Popescu, S.T.; Hoffmann, M. Automatic infant 2D pose estimation from videos: Comparing seven deep neural network methods. *Behav. Res. Methods* **2025**, *57*, 280. [CrossRef] [PubMed]
30. Shin, H.I.; Shin, H.I.; Bang, M.S.; Kim, D.K.; Shin, S.H.; Kim, E.K.; Kim, Y.-J.; Lee, E.S.; Park, S.G.; Ji, H.M.; et al. Deep learning-based quantitative analyses of spontaneous movements and their association with early neurological development in preterm infants. *Sci. Rep.* **2022**, *12*, 3138. [CrossRef]
31. McCay, K.D.; Ho, E.S.L.; Shum, H.P.H.; Fehringer, G.; Marcroft, C.; Embleton, N.D. Abnormal Infant Movements Classification with Deep Learning on Pose-Based Features. *IEEE Access* **2020**, *8*, 51582–51592. [CrossRef]
32. Reich, S.; Zhang, D.; Kulvicius, T.; Bölte, S.; Nielsen-Saines, K.; Pokorny, F.B.; Peharz, R.; Poustka, L.; Wörgötter, F.; Einspieler, C.; et al. Novel AI driven approach to classify infant motor functions. *Sci. Rep.* **2021**, *11*, 9888. [CrossRef] [PubMed]
33. Eken, M.M.; Meyns, P.; Lamberts, R.P.; Langerak, N.G. Markerless Upper Body Movement Tracking During Gait in Children with HIV Encephalopathy: A Pilot Study. *Appl. Sci.* **2025**, *15*, 4546. [CrossRef]
34. Homlong, E.G.; Quadir, H.A.; Kumar, R.P.; Elle, O.J.; Wiig, O. Addressing Occlusions and Pose Challenges in Clinical Gait Analysis: A Robust 3D-to-2D Motion Pipeline. In Proceedings of the 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Copenhagen, Denmark, 14–18 July 2025; pp. 1–7. [CrossRef] [PubMed]

35. Haberfehlner, H.; Ven SSvan de Burg Svan der Aleo, I.; Bonouvrié, L.A.; Harlaar, J.; Buizer, A.I.; van der Krogt, M.M. Using DeepLabCut for tracking body landmarks in videos of children with dyskinetic cerebral palsy: A working methodology. *medRxiv* **2022**, medRxiv:2022.03.30.22272088. [CrossRef]
36. van der Waard, M.W.P.; de Jong, L.A.F.; Keijsers, N.L.W. Motion tracking with automated pose estimator can enhance ankle-foot-orthoses alignment. *Clin. Biomech.* **2025**, *121*, 106375. [CrossRef] [PubMed]
37. Williams, S.; Zhao, Z.; Hafeez, A.; Wong, D.C.; Relton, S.D.; Fang, H.; Alty, J.E. The discerning eye of computer vision: Can it measure Parkinson's finger tap bradykinesia? *J. Neurol. Sci.* **2020**, *416*, 117003. [CrossRef] [PubMed]
38. Shin, J.H.; Ong, J.N.; Kim, R.; Park Smin Choi, J.; Kim, H.J.; Jeon, B. Objective measurement of limb bradykinesia using a marker-less tracking algorithm with 2D-video in PD patients. *Park. Relat. Disord.* **2020**, *81*, 129–135. [CrossRef] [PubMed]
39. Guayacán, L.C.; Manzanera, A.; Martínez, F. Quantification of Parkinsonian Kinematic Patterns in Body-Segment Regions During Locomotion. *J. Med. Biol. Eng.* **2022**, *42*, 204–215. [CrossRef]
40. Moro, M.; Marchesi, G.; Hesse, F.; Odone, F.; Casadio, M. Markerless vs. Marker-Based Gait Analysis: A Proof-of-Concept Study. *Sensors* **2022**, *22*, 2011. [CrossRef] [PubMed]
41. Hu, Z.; Zhang, Y.; Xing, Y.; Zhao, Y.; Cao, D.; Lv, C. Toward Human-Centered Automated Driving: A Novel Spatiotemporal Vision Transformer-Enabled Head Tracker. *IEEE Veh. Technol. Mag.* **2022**, *17*, 57–64. [CrossRef]
42. Alammari, B.J.; Schoenwether, B.; Ripic, Z.; Kirk-Sanchez, N.; Eltoukhy, M.; Bishop, L. Validity of AI-Driven Markerless Motion Capture for Spatiotemporal Gait Analysis in Stroke Survivors. *Sensors* **2025**, *25*, 5315. [CrossRef] [PubMed]
43. Wagh, V.; Scott, M.W.; Andrushko, J.W.; Jones, C.B.; Larssen, B.C.; Boyd, L.A.; Kraeutner, S.N. Using MediaPipe to track upper-limb reaching movements after stroke: A proof-of-principle study. *J. Neuroeng. Rehabil.* **2025**, *22*, 268. [CrossRef] [PubMed]
44. Peng, Y.; Wang, W.; Zeng, Y.; Chen, Z.; Li, H.; Li, G. Analysis of Pelvis and Lower Limb Coordination in Stroke Patients Using Smartphone-Based Motion Capture. *IEEE Trans. Biomed. Eng.* **2025**, *72*, 2425–2436. [CrossRef] [PubMed]
45. Yang, Z.; Stenum, J.; Varghese, R.; Roemmich, R.T. A graphical user interface for editing keypoints from human pose estimation algorithms. *medRxiv* **2025**. [CrossRef] [PubMed]
46. Hu, B.; Wang, J.; Xu, W.; Li, T.; Nie, Y.; Li, K. Dual-Camera Markerless Motion Capture System for Precise Lower-Limb Kinematic Analysis in Osteoarthritis. *Ann. Biomed. Eng.* **2025**, *53*, 2949–2965. [CrossRef] [PubMed]
47. Haddas, R.; Morriss, N.; Schillinger, E.; Minto, J.; Castle, P.; Greif, D.N.; Ramirez, G.; Barber, P.; Nicandri, G.; Mannava, S.; et al. Comparison of Markerless and Conventional Marker-Based Shoulder Kinematics Models During Activities of Daily Living in Patients with Glenohumeral Osteoarthritis. *J. Orthop. Res.* **2025**, *43*, 1907–1923. [CrossRef] [PubMed]
48. Dinh, H.G.; Zhou, J.Y.; Benmira, A.; Kenney, D.E.; Ladd, A.L. Proof of Concept and Validation of Single-Camera AI-Assisted Live Thumb Motion Capture. *Sensors* **2025**, *25*, 4633. [CrossRef] [PubMed]
49. Wagh, V.; Scott, M.W.; Kraeutner, S.N. Quantifying Similarities Between MediaPipe and a Known Standard to Address Issues in Tracking 2D Upper Limb Trajectories: Proof of Concept Study. *JMIR Form. Res.* **2024**, *8*, e56682. [CrossRef] [PubMed]
50. Cunha, B.; Macaes, J.; Amorim, I. Smartphone-Based Markerless Motion Capture for Accessible Rehabilitation: A Computer Vision Study. *Sensors* **2025**, *25*, 5428. [CrossRef] [PubMed]
51. Cronin, N.J.; Rantalainen, T.; Ahtiainen, J.P.; Hynynen, E.; Waller, B. Markerless 2D kinematic analysis of underwater running: A deep learning approach. *J. Biomech.* **2019**, *87*, 75–82. [CrossRef] [PubMed]
52. Nishikawa, N.; Watanabe, S.; Yamamoto, K. Comparison of kinematics between markerless and marker-based motion capture systems for change of direction maneuvers. *J. Biomech.* **2025**, *192*, 112965. [CrossRef] [PubMed]
53. Yoma, M.; Herrington, L.; Starbuck, C.; Llurda, L.; Jones, R. Between-Day Reliability of Kinematic Variables Using Markerless Motion Capture for Single-Leg Squat and Single-Leg Landing Tasks. *Int. J. Sports Phys. Ther.* **2025**, *20*, 1160–1175. [CrossRef] [PubMed]
54. Cameron-Whytock, H.; Divall, H.; Lewis, M.; Apps, C. Marker based and markerless motion capture for equestrian rider kinematic analysis: A comparative study. *J. Biomech.* **2025**, *186*, 112728. [CrossRef] [PubMed]
55. Falisse, A.; Uhlrich, S.D.; Chaudhari, A.S.; Hicks, J.L.; Delp, S.L. Marker Data Enhancement for Markerless Motion Capture. *IEEE Trans. Biomed. Eng.* **2025**, *72*, 2013–2022. [CrossRef] [PubMed]
56. Carriere, J.; Oliver, M.L.; Hamilton-Wright, A.; Young, C.; Gordon, K.D. Can Machine Learning Enhance Computer Vision-Predicted Wrist Kinematics Determined from a Low-Cost Motion Capture System? *Appl. Sci.* **2025**, *15*, 3552. [CrossRef]
57. Liu, L.; Dai, Y.; Liu, Z. Real-time pose estimation and motion tracking for motion performance using deep learning models. *J. Intell. Syst.* **2024**, *33*, 20230288. [CrossRef]
58. Song, Y.F.; Zhang, Z.; Shan, C.; Wang, L. Constructing Stronger and Faster Baselines for Skeleton-Based Action Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 1474–1488. [CrossRef] [PubMed]

59. Zhang, L.; Sidarta, A.; Wu, T.-L.; Jatesiktat, P.; Wang, H.; Li, L.; Kwong, P.W.-H.; Long, A.; Long, X.; Ang, W.T. Towards Clinical Application of Enhanced Timed Up and Go with Markerless Motion Capture and Machine Learning for Balance and Gait Assessment. *IEEE J. Biomed. Health Inform.* **2025**, 1–9. [CrossRef] [PubMed]
60. Sakkos, D.; Mccay, K.D.; Marcroft, C.; Embleton, N.D.; Chattopadhyay, S.; Ho, E.S.L. Identification of Abnormal Movements in Infants: A Deep Neural Network for Body Part-Based Prediction of Cerebral Palsy. *IEEE Access* **2021**, *9*, 94281–94292. [CrossRef]
61. Moro, M.; Marchesi, G.; Cellerino, M.; Boffa, G.; Odone, F.; Inglese, M.; Casadio, M. Markerless Video-Based Gait Analysis in People with Multiple Sclerosis. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2025**, *33*, 2743–2749. [CrossRef] [PubMed]
62. Thomas, C.; Nolte, K.; Schmidt, M.; Jaitner, T. Comparison of Marker-Based and Markerless Motion Capture Systems for Measuring Throwing Kinematics. *Biomechanics* **2025**, *5*, 100. [CrossRef]
63. Li, Z.; Shin, S.; Phan, V.; Meinders, E.; Halilaj, E. Impact of Multi-View Fusion and Biomechanical Modeling on Markerless Motion Tracking. *IEEE Trans. Biomed. Eng.* **2025**, *73*, 2054–2062. [CrossRef] [PubMed]
64. Yang, C.; Wei, L.; Huang, X.; Tu, L.; Xu, Y.; Li, X.; Hu, Z. Comparison of lower limb kinematic and kinetic estimation during athlete jumping between markerless and marker-based motion capture systems. *Sci. Rep.* **2025**, *15*, 18552. [CrossRef] [PubMed]
65. Kearney, J.; Greaves, H.; Robinson, M.A.; Barton, G.J.; O'Brien, T.D.; Pinzone, O.; Wright, D.M.; Gibbon, K.; Foster, R.J. Comparison of Theia3D and the conventional gait model in typically developing children and adults in a clinical gait laboratory. *J. Biomech.* **2025**, *193*, 112995. [CrossRef] [PubMed]
66. Park, J.; Kim, Y.; Kim, S.; Park, K. Markerless Kinematic Data in the Frontal Plane Contributions to Movement Quality in the Single-Leg Squat Test: A Comparison and Decision Tree Approach. *J. Sport Rehabil.* **2024**, *34*, 126–133. [CrossRef] [PubMed]
67. Cho, S.W.; Cho, S.J.; Park, E.; Park, N.; Han, S.; Rhee, Y.; Hong, N. Video-estimated peak jump power using deep learning is associated with sarcopenia and low physical performance in adults. *Osteoporos. Int.* **2025**, *36*, 1193–1201. [CrossRef] [PubMed]
68. Kim, T.; Jung, M.-C.; Mo, S.-M. Suggestion for Camera Location in Monocular Markerless 3D Motion Capture System: Focused on Accuracy Comparison with Marker-Based System for Upper Limb Joints. *IEEE Access* **2025**, *13*, 47605–47616. [CrossRef]
69. Tahara, A.K.; Chinaglia, A.G.; Monteiro, R.L.M.; Bedo, B.L.S.; Cesar, G.M.; Santiago, P.R.P. Predicting Walkway Spatiotemporal Parameters Using a Markerless, Pixel-Based Machine Learning Approach. *Braz. J. Mot. Behav.* **2025**, *19*, e462. [CrossRef]
70. Asaeda, M.; Onishi, T.; Ito, H.; Miyahara, S.; Mikami, Y. Reliability and validity of knee valgus angle calculation at single-leg drop landing by posture estimation using machine learning. *Heliyon* **2024**, *10*, e36338. [CrossRef] [PubMed]
71. Ceriola, L.; Taborri, J.; Donati, M.; Rossi, S.; Patanè, F.; Mileti, I. Comparative Analysis of Markerless Motion Capture Systems for Measuring Human Kinematics. *IEEE Sens. J.* **2024**, *24*, 28135–28144. [CrossRef]
72. Schoenwether, B.; Ripic, Z.; Nienhuis, M.; Signorile, J.F.; Best, T.M.; Eltoukhy, M. Reliability of artificial intelligence-driven markerless motion capture in gait analyses of healthy adults. *PLoS ONE* **2025**, *20*, e0316119. [CrossRef]
73. Boldo, M.; Di, R.; Martini, E.; Nardon, M.; Bertucco, M.; Bombieri, N. On the reliability of single-camera markerless systems for overground gait monitoring. *Comput. Biol. Med.* **2024**, *171*, 108101. [CrossRef] [PubMed]
74. Galasso, S.; Carissimo, C.; Cerro, G.; Molinara, M.; Ferrigno, L.; Salvatore Calabrò, R.; de Nunzio, A.M. A Novel Measurement Procedure for Error Correction in Single Camera Gait Analysis. *IEEE Access* **2024**, *12*, 158052–158064. [CrossRef]
75. Jatesiktat, P.; Lim, G.M.; Lim WSen Ang, W.T. Anatomical-Marker-Driven 3D Markerless Human Motion Capture. *IEEE J. Biomed. Health Inform.* **2025**, *29*, 6186–6199. [CrossRef] [PubMed]
76. Han, H.; Chang, J. Volleyball Motion Analysis Model Based on GCN and Cross-View 3D Posture Tracking. *Int. J. Adv. Comput. Sci. Appl.* **2024**, *15*, 804–815. [CrossRef]
77. Mathis, A.; Mamidanna, P.; Cury, K.M.; Abe, T.; Murthy, V.N.; Mathis, M.W.; Bethge, M. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **2018**, *21*, 1281–1289. [CrossRef] [PubMed]
78. Sahin, I.; Modi, A.; Kokkoni, E. Evaluation of OpenPose for Quantifying Infant Reaching Motion. *Arch. Phys. Med. Rehabil.* **2021**, *102*, e86. [CrossRef]
79. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.E.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 172–186. [CrossRef] [PubMed]
80. Fang, H.S.; Li, J.; Tang, H.; Xu, C.; Zhu, H.; Xiu, Y.; Li, Y.-L.; Lu, C. AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 7157–7173. [CrossRef] [PubMed]
81. Ferraris, C.; Amprimo, G.; Cerfoglio, S.; Vismara, L.; Cimolin, V. A Deep Dive into MediaPipe Pose for Postural Assessment: A Comparative Investigation. *IEEE Access* **2025**, *13*, 211055–211074. [CrossRef]
82. Varcin, F.; Boocock, M.G. The accuracy, validity and reliability of Theia3D markerless motion capture for studying the biomechanics of human movement: A systematic review. *Artif. Intell. Med.* **2026**, *173*, 103332. [CrossRef] [PubMed]
83. Cronin, N.J.; Walker, J.; Tucker, C.B.; Nicholson, G.; Cooke, M.; Merlino, S.; Bissas, A. Feasibility of OpenPose markerless motion analysis in a real athletics competition. *Front. Sports Act. Living* **2024**, *5*, 1298003. [CrossRef] [PubMed]
84. Milone, D.; Longo, F.; Merlino, G.; De Marchis, C.; Risitano, G.; D'Agati, L. MocapMe: DeepLabCut-Enhanced Neural Network for Enhanced Markerless Stability in Sit-to-Stand Motion Capture. *Sensors* **2024**, *24*, 3022. [CrossRef] [PubMed]

85. Mundt, M.; Colyer, S.; Wade, L.; Needham, L.; Evans, M.; Millett, E.; Alderson, J. Automating Video-Based Two-Dimensional Motion Analysis in Sport? Implications for Gait Event Detection, Pose Estimation, and Performance Parameter Analysis. *Scand. J. Med. Sci. Sports* **2024**, *34*, e14693. [CrossRef] [PubMed]
86. Panconi, G.; Grasso, S.; Guarducci, S.; Mucchi, L.; Minciacchi, D.; Bravi, R. DeepLabCut custom-trained model and the refinement function for gait analysis. *Sci. Rep.* **2025**, *15*, 2364. [CrossRef] [PubMed]

**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.