

Review

# Applications of Pose Estimation in Human Health and Performance across the Lifespan

Jan Stenum <sup>1,2</sup>, Kendra M. Cherry-Allen <sup>2</sup>, Connor O. Pyles <sup>3</sup> , Rachel D. Reetzke <sup>4,5</sup>, Michael F. Vignos <sup>3</sup> and Ryan T. Roemmich <sup>1,2,\*</sup> 

- <sup>1</sup> Center for Movement Studies, Kennedy Krieger Institute, Baltimore, MD 21205, USA; jstenum1@jhmi.edu  
<sup>2</sup> Department of Physical Medicine and Rehabilitation, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA; kcherry6@jhu.edu  
<sup>3</sup> Johns Hopkins Applied Physics Laboratory, Laurel, MD 20723, USA; Connor.Pyles@jhuapl.edu (C.O.P.); Mike.Vignos@jhuapl.edu (M.F.V.)  
<sup>4</sup> Center for Autism and Related Disorders, Kennedy Krieger Institute, Baltimore, MD 21211, USA; Reetzke@kennedykrieger.org  
<sup>5</sup> Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA  
\* Correspondence: rroemmi1@jhmi.edu

**Abstract:** The emergence of pose estimation algorithms represents a potential paradigm shift in the study and assessment of human movement. Human pose estimation algorithms leverage advances in computer vision to track human movement automatically from simple videos recorded using common household devices with relatively low-cost cameras (e.g., smartphones, tablets, laptop computers). In our view, these technologies offer clear and exciting potential to make measurement of human movement substantially more accessible; for example, a clinician could perform a quantitative motor assessment directly in a patient's home, a researcher without access to expensive motion capture equipment could analyze movement kinematics using a smartphone video, and a coach could evaluate player performance with video recordings directly from the field. In this review, we combine expertise and perspectives from physical therapy, speech-language pathology, movement science, and engineering to provide insight into applications of pose estimation in human health and performance. We focus specifically on applications in areas of human development, performance optimization, injury prevention, and motor assessment of persons with neurologic damage or disease. We review relevant literature, share interdisciplinary viewpoints on future applications of these technologies to improve human health and performance, and discuss perceived limitations.

**Keywords:** pose estimation; movement tracking; computer vision; artificial intelligence; markerless motion capture; assessment; kinematics; development; machine learning



**Citation:** Stenum, J.; Cherry-Allen, K.M.; Pyles, C.O.; Reetzke, R.D.; Vignos, M.F.; Roemmich, R.T. Applications of Pose Estimation in Human Health and Performance across the Lifespan. *Sensors* **2021**, *21*, 7315. <https://doi.org/10.3390/s21217315>

Academic Editor: Angelo Maria Sabatini

Received: 24 September 2021  
Accepted: 31 October 2021  
Published: 3 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



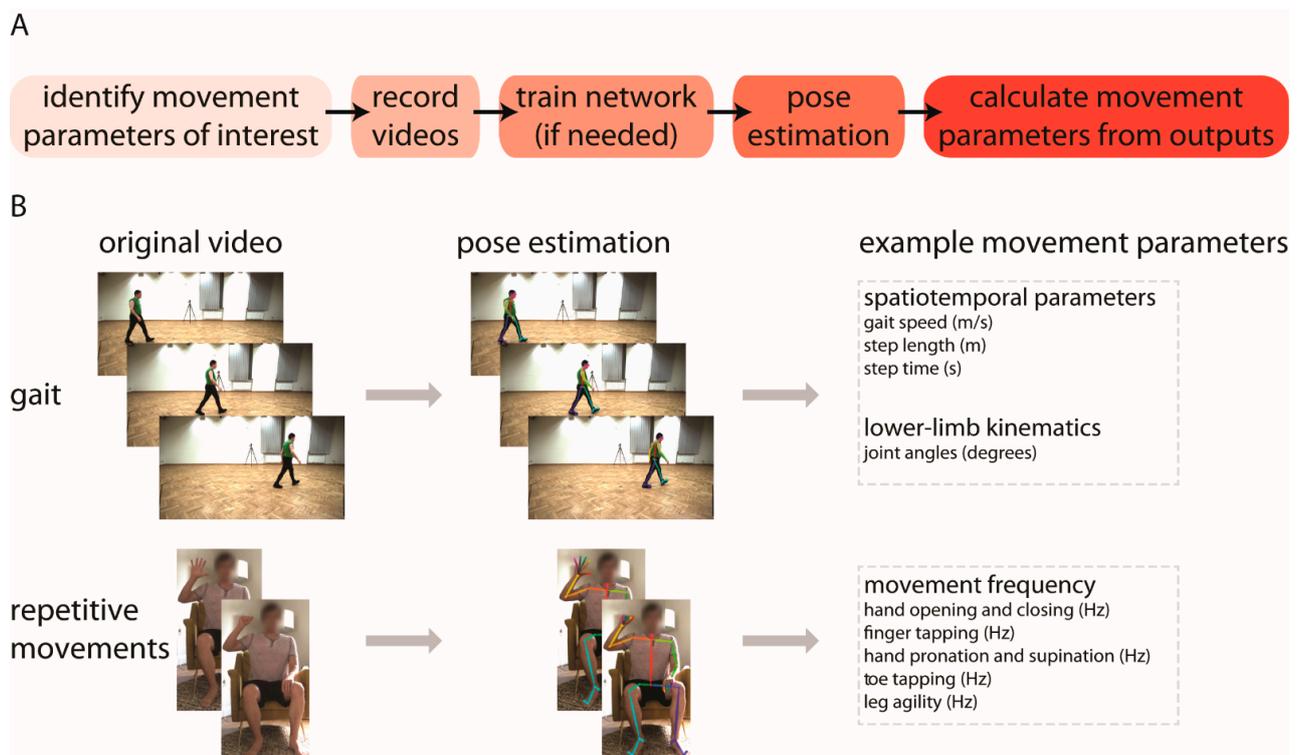
**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Humans have long been interested in quantitative measurement of our movements [1,2]. This is evident in many aspects of life: an Olympic judge scrutinizes and scores a figure skater's performance; a physical therapist measures a patient's walking speed to assess mobility; a running coach inspects and adjusts a distance runner's foot-strike pattern to prevent injury. We also interpret the movements of others to communicate (e.g., sign language) or make inferences about emotional state (i.e., "reading body language"; [3–5]).

In this review, we focus on applications of human pose estimation, an emerging technology for quantitative measurement of human movement kinematics [6–13]. Pose estimation algorithms use computer vision to identify key landmarks on the body (e.g., fingertip, elbow, knee) from simple digital videos that can be recorded using common household devices (example workflow and applications are shown in Figure 1A,B, respectively). This simplicity offers exciting potential for measuring whole-body kinematics in

nearly any setting, with minimal costs of money, time, and effort. We also see significant opportunities for the ongoing maturation and validation of these approaches to offer robust supplements or alternatives to subjective visual motor assessments and to improve accessibility to measurement of movement kinematics by removing long-standing barriers. The ability to capture quantitative, whole-body kinematics using a household device could substantially reduce reliance on traditional methods that are inaccessible or data-limited, such as expensive research-grade motion capture systems or wearable devices.



**Figure 1.** (A) Basic workflow for using pose estimation to measure movement kinematics from video; (B) Example applications of using pose estimation to quantify spatiotemporal and kinematic gait parameters (top) and frequencies of repetitive upper and lower extremity movements (bottom). These applications are described in greater detail in [14,15]. The gait images shown in (B) are taken from the GPJATK dataset [16].

We focus specifically on applications of human pose estimation for improving human health and performance. We note that pose estimation algorithms are used for many other applications (e.g., intelligent video surveillance [17], activity recognition [18], sign language translation [19]), and prior reviews have discussed technical aspects of various algorithms and their perceived advantages and disadvantages [20–22]. Here, we focus less on the technical aspects of pose estimation and instead discuss applications of these algorithms, both in terms of current applications and those that we perceive may be possible in the future. We cover areas of application across the human lifespan, including human development, human performance optimization, musculoskeletal injury prevention, and motor assessment of persons with neurologic damage or disease.

We also integrate the clinical perspective on pose estimation applications. Much prior work on human pose estimation (including our own) has suggested promise for clinical application. However, in our view, the clinician’s (i.e., end user) viewpoint on these potential applications has not received adequate consideration or representation, and applications of pose estimation have not been contextualized within current models of clinical care. We aim to address these issues by providing an interdisciplinary perspective that integrates views from physical therapy, speech-language pathology, movement science, and engineering.

## 2. What Is Pose Estimation?

Markerless human pose estimation relies on recent advances in computer vision to automatically track anatomical landmarks—so-called keypoints—of the human body from digital videos. Examples of possible tracked keypoints include the ankle, knee, hip, wrist, elbow, shoulder, foot (e.g., heel, big toe, and small toe), hand (e.g., tip and three joints of every finger), and face (e.g., ears, eyes, nose, and mouth). Current state-of-the-art algorithms used to track human poses have been trained on large datasets of digital images and/or videos of human movement in which keypoints have been manually annotated [23,24]. The trained algorithms can then track new, unlabeled videos of humans. This enables automated, video-based human movement tracking, with the greatest accuracy achieved for movements similar to those in the training dataset.

The primary output from pose estimation is a series of two-dimensional pixel coordinates of the tracked keypoints, as they appear projected onto the image sensor of the camera. From the two-dimensional pixel coordinates, different approaches of analyzing and processing data have been reported, and fall into three broad categories. First, some studies use the output to represent planar two-dimensional kinematics of human movement, from which specific metrics of interest can be calculated [15,25–28]. An example of an instance in which this approach may be appropriate is capturing a video of the sagittal view of human locomotion and subsequently calculating sagittal gait kinematics (e.g., lower limb joint angles). Second, it is possible to reconstruct three-dimensional kinematics of human movement if capturing videos from multiple viewpoints using at least two cameras [29–31]. This approach offers significant advantages over a single camera view, in part because occlusions occur and out-of-plane motions are not well-captured by a single camera; however, this approach also has potential drawbacks associated with setup and computational complexity. Last, it is also possible to use the pose estimation output as an input for further processing by neural networks designed to predict specific metrics of interest [32–34]. Subsequent processing by neural networks may be appropriate when predicting a scalar value such as peak knee flexion during walking or clinical ratings, but this approach may be less accurate when predicting frame-by-frame time-series data. This inaccuracy is commonly due to the fact that most algorithms do not aim to minimize frame-to-frame variation when performing pose estimation with video data.

These diverse approaches to data analysis of pose estimation of human movement make it possible to obtain many parameters associated with movement. For example, pose estimation has been used to study human locomotion [15,34,35] and provide kinematic measures such as lower limb joint angles; spatiotemporal measures such as gait speed, step length, and step time; and clinical ratings such as the Gait Deviation Index in patients with cerebral palsy or MDS–UPDRS gait scores for persons with Parkinson’s disease. Other studies have used pose estimation to assess neuromotor risk and development in human infants [36,37]. These areas of application are introduced briefly here, but will be covered in greater detail in later sections of this manuscript.

## 3. What Tools Are Available?

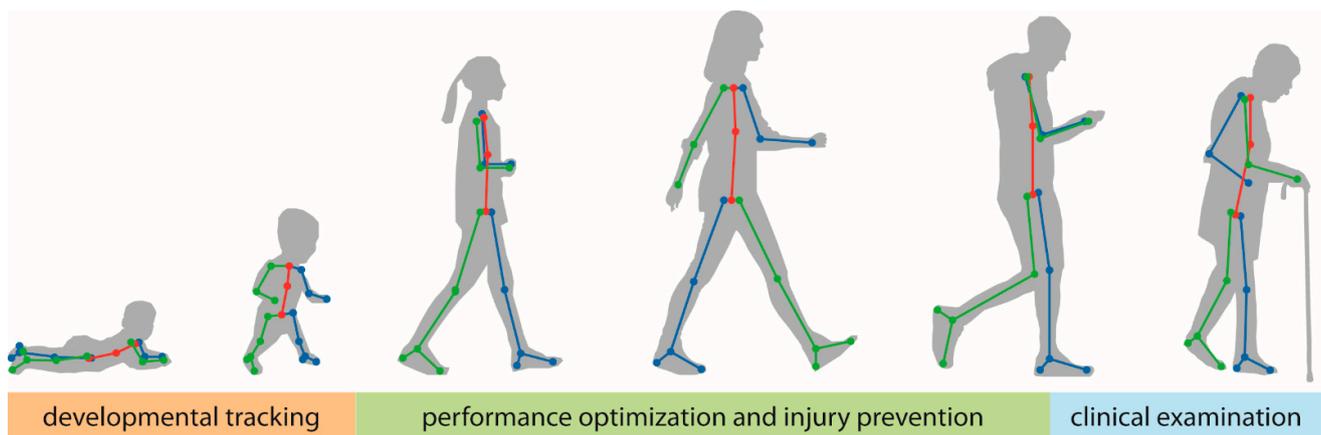
Several different algorithms for pose estimation have been published over the past decade (e.g., OpenPose [13], DeepLabCut [12], DeepPose [10], DeeperCut [8], AlphaPose [38], ArtTrack [7]). Using these algorithms, it is possible to take advantage of pre-trained networks that are freely available, or train new networks customized for various research or clinical needs. For example, a commonly used pretrained network is the human pretrained demo of OpenPose that includes keypoints of the body, feet, hands, and face [13,39] and has been used in several recent studies for quantitative analysis of human movement [15,26,29,31,34,40].

The computations needed for training a new network and tracking new videos often require intensive computing capabilities. Therefore, the computing power of a graphics processing unit (GPU) may be necessary in order for processing times to reach acceptable limits (many algorithms provide documentation with hardware recommendations,

as in [11]). If a user does not have their own GPU, some computing environments (e.g., Google Colaboratory) provide GPU access for faster processing; however, these may not be suitable for applications involving protected health information because the processing occurs externally. Processing without a GPU is slower but may be sufficient depending on the user's time constraints and processing needs (e.g., length of videos, number of people tracked, number of keypoints tracked). Furthermore, it is also possible to use pose estimation for real-time movement tracking (as is available with OpenPose, for example [39]). This capability may be particularly useful to some users, as it could be implemented to provide real-time biofeedback for various applications. Beyond these increasingly popular deep learning approaches, other approaches also use optimization [41–43] and filtering [44,45] techniques to perform pose estimation.

#### 4. How Can These Tools Be Used to Improve Human Health and Performance?

In the following subsections, we will focus on three specific areas of application across the human lifespan: (1) human development, (2) performance optimization and injury prevention, and (3) motor assessment of persons with neurologic damage or disease (Figure 2). Certainly, many additional areas of application exist beyond the scope of this review. We focus on these applications due to the emerging nature of the relevant literature and the expertise of the authors. We expect that many of the principles discussed below are likely to generalize to other applications and/or populations of interest.



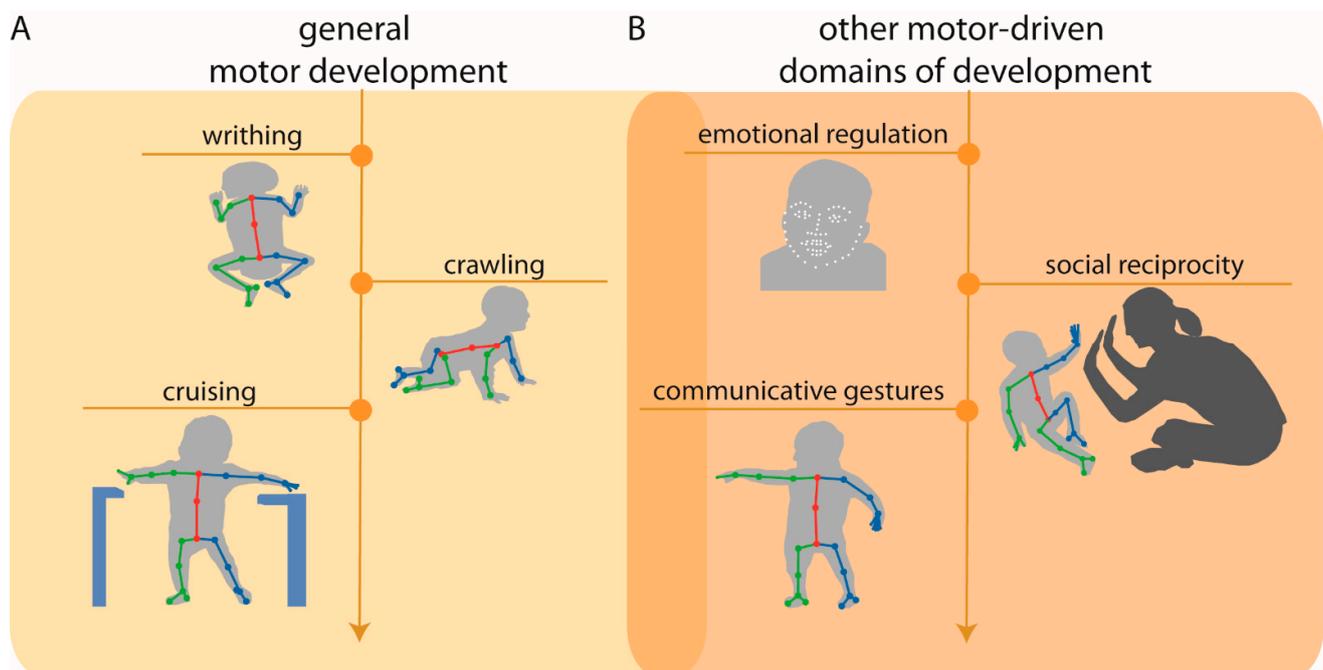
**Figure 2.** In this manuscript, we focus on three general areas of applications of pose estimation in human health and performance across the lifespan: tracking of motor and non-motor development in young children (orange), performance optimization and injury prevention in athletes and other populations that are primarily young or middle-aged adults (green), and clinical examinations of persons with neurologic damage or disease who are primarily older adults (blue).

##### 4.1. Tracking General Motor Development

Developmental scientists study the emergence of specific behaviors from infancy to adolescence in many different settings, including the laboratory, home environment, clinic, and classroom. Accordingly, video recordings are an integral component of most, if not all, developmental research programs. Video-based approaches have been used to study multiple domains of development, including gross and fine motor development as well as social, language, and play development [46–49]. One major limitation of current video-based approaches is the time-intensive but necessary process of manually coding child behaviors of interest by clinicians and researchers. Pose estimation technologies offer a much-needed opportunity to accelerate video coding to capture specific behaviors of interest in such developmental investigations. Due to the extensive manual video coding that has been done in the field over decades, there are large existing video databases that have already undergone human coding/reliability checks and can provide a valuable source of ground truth data for training and validation of machine learning models of de-

velopment (e.g., [50]). Such approaches could further help decrease reliance on assessment tools that require the expertise and time of trained clinicians for interpretation and, in turn, offer cost-effective and scalable alternatives to more subjective measures of typical and atypical development.

Although in the early stages of application, pose estimation approaches are beginning to be applied to the study of general motor development [36,51] (Figure 3A). For example, pose estimation has been used to detect normal writhing movements (i.e., typical spontaneous movements produced by newborns) vs. abnormal movements from video recordings of newborns in their first days of life [51]. Preliminary findings are promising and suggest that normal vs. abnormal writhing movements can be automatically classified with 80% accuracy, a percentage comparable to expert human classification.



**Figure 3.** (A) Example applications of pose estimation to quantify early motor developmental milestones (left), including writhing movements (e.g., [51]), crawling, and cruising (e.g., [36]); and (B) other motor-driven domains of development, including emotional regulation, social reciprocity, and communicative gestures. Overlap between (A,B) denotes that these areas of development are intimately linked with one another. Arrow indicates that application of pose estimation is not restricted to these examples and can be applied to quantify later motor and motor-driven developmental milestones.

As infants progress in their gross motor development, the onsets of crawling and walking—gross motor advances that allow infants to explore and learn from their environment—have been found to be intimately linked with growth in other developmental domains [52,53]. Indeed, findings from developmental science literature suggest that delays in the onset of walking may result in limited opportunities for exploration and input from caregivers and family members, leading to subsequent delays in language and social communication development [48,54,55]. As a result, it is critical to improve the early detection of delays in locomotor development in order to intervene prior to any cascading effects on other domains of development.

Researchers have begun to implement pose estimation as a useful tool for quantitative tracking of infant locomotor development. For example, Ossmy and Adolph [36] used a combination of pose estimation, machine learning, and time-series analyses to examine the role of experience in infant acquisition of interlimb coordination based on video recordings of the infants “cruising” (i.e., side-stepping with support of the upper extremities)—which is the transitional behavior between crawling and walking—at 11 months of age. More specifically, the authors used pose estimation to track frame-by-frame body movements

and subsequently calculated the distance between the limbs (i.e., the distance between the hands and the distance between the feet) for each tracked video frame to extract the coordination pattern for cruising. The results of this study provided insight into the mechanisms by which infants learn to optimally cruise and, as a result, may hold implications for future work aiming to investigate early detection and intervention for delays in locomotor development.

#### 4.2. Clinical Use in Pediatric Populations

Early detection of atypical development is critical for the diagnosis of congenital movement-based disorders (e.g., cerebral palsy) and neurodevelopmental disorders (e.g., autism spectrum disorder) to ensure timely access to early intervention services to improve motor outcomes (e.g., coordination, postural support) and other domains of development (e.g., social, language). Advances in pose estimation approaches and the emergence of novel machine learning-based models offer exciting potential for the assessment of movement-based predictors of clinical disorders. For example, pose estimation is beginning to be applied, not only to measure predictors of later motor-based disorders, but also predictors of other motor-driven domains of development (social communication; Figure 3B). In this subsection, we provide examples of these advances.

Cerebral palsy (CP) is the most common movement disorder in childhood, caused by abnormal neural development or injury that impairs the ability to control movement and posture [56]. Diagnosis of CP using conventional assessments typically occurs between age 12 and 24 months; however, using a combination of standardized assessments and neonatal magnetic resonance imaging (MRI), CP can be accurately predicted before 6 months corrected age [57]. Yet, there remain significant drawbacks to this approach: standardized assessments are based on subjective human observation that requires substantial training and clinical expertise, and neonatal MRI is expensive and often inaccessible in low-resource areas [58].

Recent research efforts have attempted to address these shortcomings by aiming to use video recordings to implement low-cost, automatic, objective alternatives for the detection of CP risk. Such investigations have succeeded in predicting CP based on automatic movement assessment from infant video recordings with performance comparable to standardized CP risk measures [59–61]. For example, in a multi-site cohort investigation, an automated, objective, movement assessment of infant video recordings was compared to standard risk assessment measures (i.e., the General Movement Assessment and neonatal neuroimaging) at 9–15 weeks corrected age to predict CP status and motor function at approximately 3.7 years of age. The results of this investigation found that the automated, video-based approach exhibited sensitivity and specificity comparable to standard measures used to predict CP [61].

There are also clear applications for pose estimation to potentially improve the early identification of neurodevelopmental disorders, such as the early detection of autism spectrum disorder (ASD). Although parents often report first concerns about ASD when their child is between 12 to 14 months of age [62,63] and reliable ASD diagnosis is possible by age 2, the majority of children with ASD remain undiagnosed until 4 years of age [64]. Shortages of ASD expert clinicians and limited capacities at autism tertiary diagnostic centers contribute to the long wait times for families [65]. Families living in rural and low-resource communities are often required to travel long distances to receive diagnostic services, placing them at an even greater disadvantage in accessing services. Indeed, a recent report indicates that approximately 84% (2635/3142) of U.S. counties do not have the necessary ASD diagnostic resources [66]. Given these barriers to a timely diagnosis, a significant portion of children with ASD are missing a critical window for early intervention services, as evidence shows that intervention before the age of 2 significantly improves behavioral and developmental outcomes for children with ASD [67–69]. The detrimental impact of diagnostic delays has resulted in federal prioritization of early identification of ASD and an urgency to develop accessible and accurate early screening methods [64].

Leveraging advances in machine learning, efforts have been made to develop scalable, video-based ASD screeners to improve access to diagnostic and early intervention services. For example, Crippa et al. developed an algorithm to examine the predictive value of motor behavioral biomarker measures in ASD to discriminate preschool children with ASD from children with typical development using a simple upper-limb reach-to-drop task [70]. The resulting model showed an accuracy rate of 96.7%, suggesting that video-based approaches combined with machine learning can be a useful method of classification and discrimination in the diagnostic process [70].

The emerging evidence supporting the application of automated, video-based assessments to monitor general gross motor development and promote early detection of both motor-based and neurodevelopmental disorders is promising. In order to establish the clinical utility of pose estimation, future work is needed to examine the feasibility and acceptability of clinician use of such techniques.

#### *4.3. Human Performance Optimization, Injury Prevention, and Safety*

Numerous applications of pose estimation exist within optimization of human performance and safety, with these applications spanning injury risk assessment, rehabilitation, and enhancing human performance. This application space commonly consists of some type of instructor, such as a coach, trainer, or clinician, attempting to assess an individual's movement patterns to determine whether the individual is at an increased risk for injury, is moving differently from a healthy, uninjured individual, or is moving with some level of inefficiency that can be modified to improve performance. Within injury assessment, common applications of pose estimation have been to evaluate an individual's risk for specific musculoskeletal injuries and to perform a post-hoc analysis following the occurrence of an injury. For example, two-dimensional pose estimation techniques have been applied to develop proof-of-concept screening technologies that detect abnormal gait patterns during walking and running [71–75], fall detection [76–78], abnormal movements that are indicative of injury risk in manual labor work environments [79–81], and risk of sports-related injury, such as anterior cruciate ligament rupture [82–84]. Post-hoc analysis following an injury has primarily been targeted towards sports performance applications and focused on understanding mechanisms of injury, with the ultimate goal of developing techniques to mitigate injury risk [85,86].

Applications of pose estimation to rehabilitation following injury or surgery typically focus on using these techniques to monitor an individual's return to normal movement patterns and to guide the motion of rehabilitation technology that is designed to interface with a patient. Pose estimation techniques have been used to measure a patient's range of motion and movement during functional exercises and assess their progression towards a healthy range of motion [87–89]. In particular, there has been an emphasis on the use of pose estimation to monitor rehabilitation progress outside of the clinic, such as in home or on an athletic field [90–93]. Additionally, many technologies have been designed to actively interface with an individual to either support their movement during rehabilitation or to help provide a mechanical stimulus to enhance rehabilitation. These technologies are commonly referred to as rehabilitation robotics, and techniques have been developed that leverage pose estimation to inform the movement of these systems [94–97].

The use of pose estimation for enhancing human performance remains a challenging application, given the large range of joint articulation, out of plane motion, and fast movements that can be difficult to capture with the relatively slow sampling rates of common video recording devices and risk of occlusion that occurs in these applications [98,99]. However, a number of proof-of-concept systems have been developed to inform pose of an athlete during training, particularly for sports in which success for the athlete is directly linked to pose (e.g., gymnastics and skiing) [100–102]. Development of new pose estimation techniques for human performance applications have focused on achieving high accuracy with 'in the wild' pose estimations, given the importance of performing these measurements outside of the lab in these applications [11,103,104]. While this previous

research has demonstrated applications that may be made possible with pose estimation, very few of these proof-of-concept technologies have made the transition to regular use in a clinical, athletic, or other relevant environments. This likely derives from the fact that many unique requirements arise when attempting to apply these techniques to human performance applications outside of the laboratory.

For pose estimation to influence the broader human performance community, including non-clinical populations, research must drive towards robust ‘in the wild’ pose estimation encompassing a range of environments and populations. To this end, we will define desirable components of an ideal dataset for pose estimation algorithm development, training, and validation. Future studies should focus on capturing and making available these datasets to expand the application space of pose estimation or define functional limitations of the current hardware or software technology.

Many injury and performance evaluations are based on highly dynamic motion analysis [85,86,105], requiring that any pose estimation validation datasets should include accurate ground truth measurements of human joint kinematics for as many degrees of freedom as feasible. Ideally, this will include kinematics of complex joints, such as the ankle, wrists, intervertebral joints, and scapular motion—all of which play a key role in many injuries and are not estimated in most existing pose estimation techniques. Linear kinematics of the various body components should also be reported on, especially in relation to conditions that result from impact injuries (e.g., traumatic brain injury, chronic traumatic encephalopathy) [106]. Optical motion tracking is currently the gold standard for such ground truth measurements, but further accuracy (and cost) improvements are desirable due to artifacts arising from relative marker motion with respect to the underlying bony anatomy [107]. Therefore, researchers should aim to account for these artifacts within the pose estimation process.

Validation datasets should be captured outside of laboratory environments and include complexities such as partial occlusion (self-occlusion, inter-subject occlusion, environmental occlusion), various illuminations, loose-fitting clothing, and multiple camera standoffs or viewing angles. Recent examples of pose estimation outside of the lab are primarily based on monocular RGB images [108–111]. However, these techniques are generally less accurate—especially in three dimensions—when compared to laboratory pose estimation. The fusion of other pose estimation modalities, including inertial measurement units and infrared imaging, with single or multi-view RGB images is a promising direction for improved pose estimation [112], and should be included in validation datasets, such as those provided by Malleon et al. [113].

As new pose estimation algorithms are developed for human performance applications, special consideration should be given to the evaluation metrics reported. Motion type classification is of limited usefulness for in-depth biomechanical analysis and, instead, joint kinematic errors should be reported for each degree of freedom. Furthermore, estimation accuracies should be reported under varying conditions, including differences between lab-based and outdoor estimations. Finally, the computational cost per frame of pose estimation should be reported to understand applicability to real-time, highly dynamic application spaces [113].

#### *4.4. Clinical Motor Assessment in Adult Neurologic Conditions*

Clinical assessments and the resulting outcome measures are critical to motor rehabilitation in adults with neurologic conditions. These clinical assessments are typically administered to capture either a patient’s status at a specific point in time or to track their motor function longitudinally. When administered at a single time point, assessments are used to classify the severity of an individual’s deficits. When administered longitudinally, assessments are commonly used to track disease progression/regression, measure recovery, or evaluate the effectiveness of an intervention.

The International Classification of Functioning, Disability and Health (ICF) is a common, widely accepted framework developed by the World Health Organization for describ-

ing health and disability at individual and population levels [114]. It provides standard language and has a wide range of uses across different sectors by identifying three primary levels of human functioning:

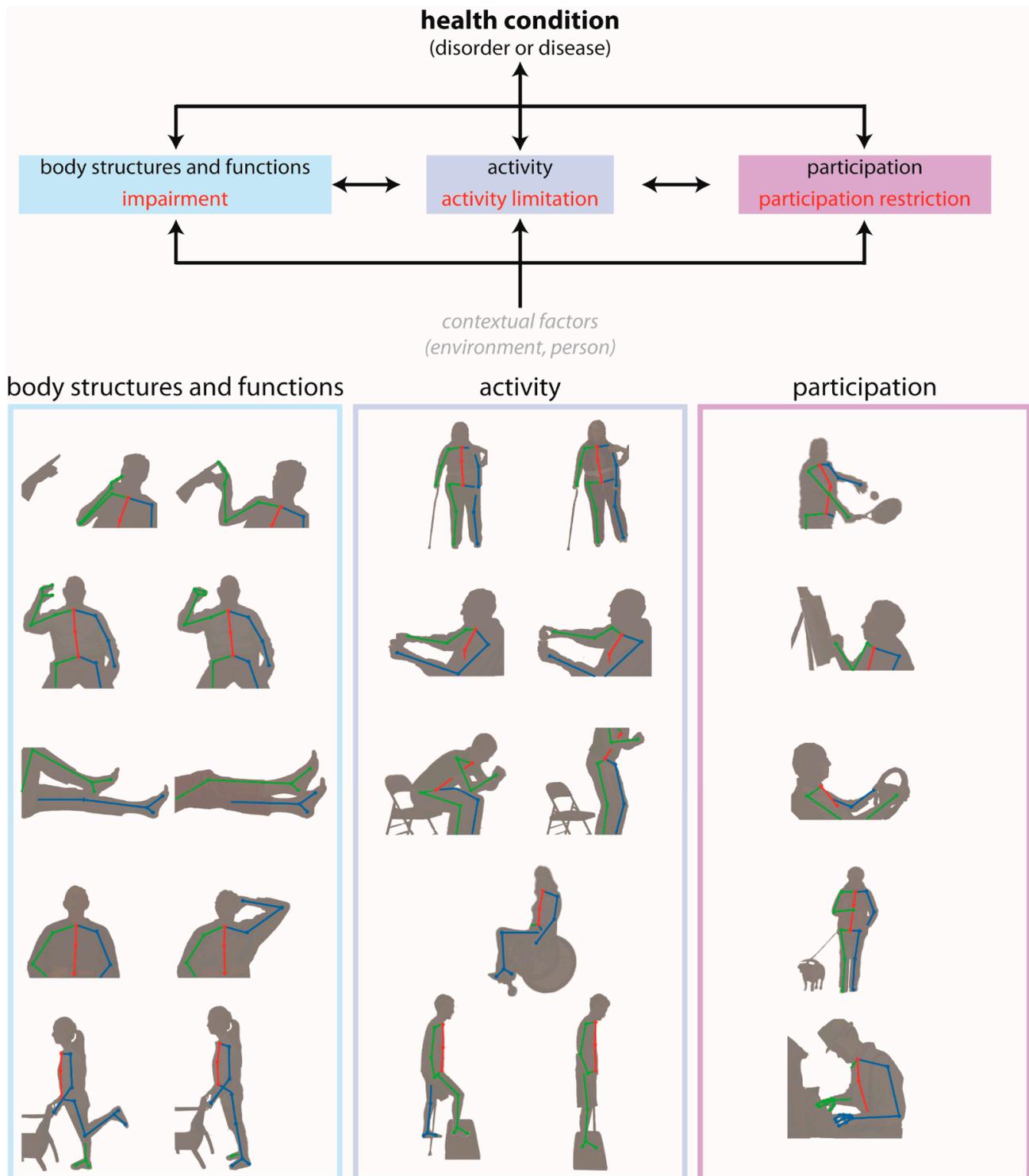
1. *Body structures and functions* are anatomical parts of the body and physiological functions of the body systems, respectively. The term impairment refers to problems in *body structure or function*.
2. *Activity* is the execution of a task or action by an individual. The term activity limitation describes difficulties with completion of an *activity*.
3. *Participation* is involvement in a life situation. Participation restrictions are problems that an individual encounters during *participation* in real-world situations.

To provide a concrete example of how this framework is used, consider a person who has experienced a stroke. This person might experience changes in all three levels of human functioning: the impairment of left-sided hemiparesis (*body structures and functions* level), the activity limitation of difficulty walking (*activity* level), and the participation restriction of inability to attend their desired religious activities (*participation* level). One can quickly observe that, while the three levels may be related to one another, there are independent needs for quantitative measurement within each level. In other words, there are needs for quantitative measurement of the hemiparesis, daily walking activity, and the inability to attend religious activities in this particular example.

Clinical outcome measures for each level of the ICF are administered as a part of routine clinical practice. Current measures of impairment involve a skilled clinician observing a patient as they perform a series of movements designed to expose deficits in *body structure and function*. For instance, one item on the Fugl–Meyer Assessment—a widely used quantitative measure of motor impairment after stroke—involves asking the patient to move their hand from the contralateral knee to ipsilateral ear while individual elements (e.g., shoulder retraction, shoulder elevation, elbow flexion, forearm supination) of this movement are scored subjectively from 0 to 2 [115]. Measures of activity limitations involve the patient performing one or more tasks that simulate *activities* encountered in daily life. An example of an ecologically valid task is the water pouring item of the Action Research Arm Test—an extensively used activity level measure for people with stroke [116]—where the person pours water from one glass to another. Lastly are measures of participation restrictions, which are often self-reported measures of the person’s perceptions of their movement abilities and resulting impact on their quality of life (e.g., the Stroke Impact Scale [117], a self-report questionnaire that evaluates disability and health-related quality of life after stroke) and daily *participation*. The data gathered from existing outcome measures are valuable for their use in diagnosing movement disorders, establish rehabilitation goals, and track changes in patient status.

Pose estimation tools have the potential to address two important challenges that exist within current clinical assessments spanning all three levels of the ICF (Figure 4). First, they can increase the accuracy, precision, and frequency with which movement kinematics are measured and assessed. Presently, *body structure/function* and *activity* level assessments primarily rely on visual observation of movement or task performance, and many are scored on ordinal scales that require a clinic visit or other similarly time-consuming interaction for both patients and their providers. Pose estimation offers the potential to provide precise, quantitative, and continuous data about single joint or whole-body movements through short video recordings that could be recorded in virtually any setting with much higher frequency. This opportunity to obtain frequent, quantitative motor assessments could significantly enhance the abilities of clinicians to detect and track impairments and activity limitations in their patients longitudinally. Second, current assessments of participation restrictions are almost exclusively self-reported. The self-report format has been necessary due to the difficulty of measuring movement kinematics in the home, but many self-report measures lack reliability and often do not correlate with clinically-administered motor assessments. There is clear potential for the propagation of telerehabilitation and pose estimation tools to make a significant impact in this area by providing

significantly improved accessibility for clinicians and researchers to obtain quantitative data about how people move and participate in their home and community environments.



**Figure 4.** Depiction of potential applications of pose estimation for movement tracking during clinical assessments across the domains of the International Classification of Functioning, Disability and Health (ICF) model. For instance, finger–nose coordination testing the body structures and functions domain (**left**), walking assessment in the activity domain (**middle**), and playing tennis in the participation domain (**right**).

The uses of pose estimation in clinical populations are expanding, but ultimately remain in the beginning stages. At the *body structure/function* level, early work has involved detecting hallmark motor signs in persons with Parkinson’s disease (PD). For instance, dyskinesia is an involuntary movement of the head, arm, leg, or entire body. Dyskinesia is commonly seen in persons with PD, often as a side effect of long-term levodopa treatment. A number of recent studies have used pose estimation to assess dyskinesias in persons with PD and found similar or superior performance with standard clinical assessments [118–120]. Bradykinesia, or slowness of movement, is another cardinal motor sign of PD. Liu et al. report that their computer vision-based method was 89.7% accurate in quantifying bradykinesia severity in people with PD as they performed repetitive movements including finger tapping, hand clasping, and alternating hand pronation/supination movements [121].

There are also a number of studies that have begun to use pose estimation to measure *activity-level* behaviors. Gait assessment, in particular, has been an early clinical target for these evolving tools. Video-based tools have been used to successfully capture gait parameters such as step lengths, step width, step time, stride length, gait velocity, and cadence in people with stroke [122], PD [25,123] or dementia [124]. Beyond gait, the timed up and go is a widely accepted assessment of functional mobility in patients with a range of neurological disorders or disease. Li et al. recently validated and used a video-based activity classification to automatize timed-up-and-go sub-task segmentation (sit-to-stand, walk, turn, walk-back, sit-back) in people with PD [125].

Future work should focus on further validation of pose estimation with gold standard kinematic tools and interpretability alongside standard clinical assessments. Additional patient populations with a wide range of different movement patterns should be included in these investigations in order to develop algorithms that are broadly applicable. The potential of video-based analysis and pose estimation to quantitatively measure participation-level data in the home and the community should also be a top priority. Precise data captured in the real world not only will provide clinicians with important data from which they can make clinical decisions, but this may also facilitate early diagnosis of movement disorders and the ability to track movement patterns throughout a disease course.

We summarize many of the applications discussed in Section 4 in Table 1 below.

**Table 1.** Summary of example applications of pose estimation in human health and performance across the lifespan.

Domain	Behavior/Movement Pattern Tracked	References
Motor and non-motor development	Infant cruising (early locomotion)	[36]
	Infant play/general movement	[37]
	Infant writhing	[51]
Human performance optimization, injury prevention, and safety	Healthy repetitive movements	[14]
	Healthy gait	[15,26,29–31,35,40]
	Sign language	[19]
	Healthy running	[27,35]
	Bilateral squat	[28]
	Healthy gait/jumping/throwing	[29]
	Lifting	[79,84]
	Various unsafe working behaviors	[80,81]
	ACL injury risk	[82,85,86]
	Handcart pushing and pulling	[83]
	Ergonomic postural assessment	[87]
	Remotely-delivered rehabilitation	[88,91–93]
Healthy finger movements	[90]	
Rehabilitation robotics	[94–97]	
Athletic training	[100,101]	
Swimming	[102]	
Clinical motor assessment	Gait in Parkinson’s disease	[25,33,123]
	Knee kinetics in osteoarthritis	[32]
	Gait in cerebral palsy	[34]
	Simulated abnormal gait	[72,74]
	Gait in older adults	[73]
	Fall detection	[76–78]
	Dyskinesias in Parkinson’s disease	[118–120]
	Gait in older adults with dementia	[124]
Timed up-and-go in Parkinson’s disease	[125]	

## 5. What Are the Limitations of Pose Estimation?

While many of our perspectives on the limitations of human pose estimation algorithms with regard to applications in human health and performance are embedded within the sections above, we considered that it may be helpful to include a condensed summary section here. As mentioned previously, technical limitations have been discussed extensively in prior reviews [20,21]. Here, we list perceived limitations in two general areas: *application limitations* and *barriers to implementation*. We consider application limitations to be those associated with obtaining high quality, usable data from video recordings via pose estimation (some are also discussed in [21]) and barriers to implementation to be limitations associated with the uptake and implementation of pose estimation approaches for common use among clinicians and researchers (with an emphasis on implementation in clinical settings).

### 5.1. Application Limitations

- **Occlusions:** these occur when one or more of the anatomical locations desired to be tracked are not visible. This may be due to occlusion by other body segments, by other people in the frame, or by inanimate objects (e.g., assistive devices—canes, walkers, crutches, orthoses, robotics; clinical objects—beds, hospital gowns, medical devices; sporting equipment—helmets, balls, bats, sticks).
- **Limited training data:** networks that are trained on sets of images that lack diversity (e.g., clothing, poses, illuminations, viewpoints, unusual postures associated with clinical conditions) may not perform well in applications where the videos are quite different from those included in the training set. Applications of current techniques that require a training dataset may require creation of a new training dataset if movements/images of a patient population are substantially different from those included in the existing training dataset (e.g., abnormal hand postures after stroke). This is particularly important given that most training datasets are biased toward healthy movement patterns.
- **Capture errors:** pose estimation algorithms may identify and track unwanted human or human-like figures in the field of view (e.g., people in the background, images on posters or artwork).
- **Positional errors:** tracking may be difficult when conditions introduce uncertainty into the positions of anatomical locations within the image (e.g., wearing a dress, hospital gown, athletic uniform or padding). This may also occur when attempting to track a movement from a suboptimal viewpoint (e.g., measuring knee flexion from a frontal view).
- **Limitations of recording devices:** use of devices with low sampling rates (e.g., the sampling rate of common video recording devices is often approximately 30 Hz) may be unable to capture accurate movement kinematics of movements that occur at high speeds or high frequencies. The aperture and shutter speed of recording devices can also impact image quality and introduce blurring, which can impact the quality of the tracking achieved through pose estimation.

Examples of application limitations are depicted in Figure 5.

### 5.2. Barriers to Implementation

- **User-friendliness:** we currently lack plug-and-play options for pose estimation. While we certainly understand and acknowledge the many reasons for this, pose estimation is unlikely to be used widely in clinical settings in particular until user-friendliness improves. We outline several relevant components to user-friendliness below:
  - **Set up time:** in our experience, many users want point-and-click capability. They want to be able to carry a recording device in their pocket, use it to record a quick video of their patient or research participant when needed, and ultimately obtain meaningful information about movement kinematics. Alternatively, they want a reserved space where a recording device could be permanently mounted and

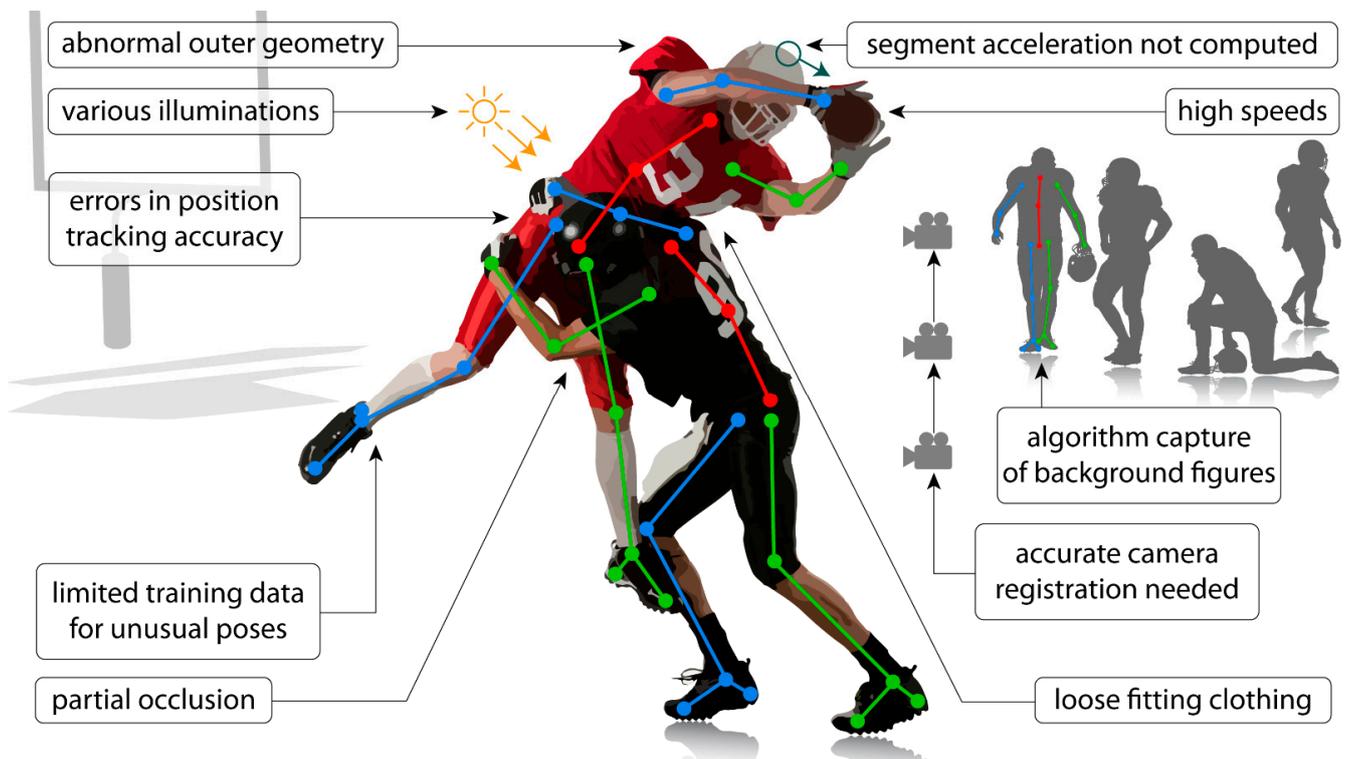
easily started and stopped (e.g., a tablet mounted to a wall). Any configuration that requires multi-camera calibration or prolonged set up time is unlikely to be adopted for widespread clinical use.

- Delayed results: many users want results in near real-time. There is a need for fast, automated approaches that immediately process the pose estimation outputs, calculate relevant movement parameters, and return interpretable data.
  - Programming and training requirements: some existing pose estimation options are very easy to download, install, and use for users with basic technical expertise. However, even these can remain prohibitively daunting for clinicians and researchers without technical backgrounds. Technologies that require any amount of programming or significant training are unlikely to reach widespread use in clinical settings.
- Outcome measure challenges: in some cases, users want to use movement data to improve clinical or performance-related decision-making, but it is not immediately clear what parameters of the movement will lead to improved outcomes (e.g., a user may express interest in measuring “walking” but is not sure which specific gait parameters are most relevant to their research study or clinical intervention). Therefore, there is a desire to collect kinematic data, but how these data should be used is not well-defined. Similarly, in the case of clinical assessments, there needs to be a clear link to relevant clinical and translational outcomes—the users should have input as to what output metrics are important.
  - Limited hardware infrastructure: as described above, some applications of pose estimation for human movement tracking require significant computational power. Some clinical and research settings are unlikely to have access to the hardware (e.g., GPUs) needed to execute their desired applications in a timely manner.
  - Technology challenges: many technologies that promise potential for clinical or human performance impact are made available before they are fully developed. This can lead to buggy software and frequent updating, which harms trust and credibility among users. This can, in turn, exacerbate the hesitancy in adopting new technologies present in some clinical and research communities, especially in artificial intelligence technologies (such as pose estimation) that are purported to supplement or even replace expert human assessment.
  - Lack of validation and feasibility data: there is a need for large-scale studies to validate pose estimation outputs against ground truth measures in a wide range of different populations. This may be accomplished in a variety of ways, including (but not limited to) comparisons with three-dimensional motion capture, wearable devices with proven accuracy, expert clinical ratings and/or assessments, or even possibly other pose estimation algorithms. The error (relative to the ground truth measurement) that is deemed acceptable is likely to depend on the use case and the metrics being used. In our experience, users who study very specific movements of joints or other anatomical landmarks (e.g., biomechanics or motor control researchers) are likely to seek greater accuracy than, for example, a clinician who may wish to incorporate a video-based assessment of walking speed as part of a larger clinical examination. It may be desirable to begin to develop field-specific accuracy standards for some applications.

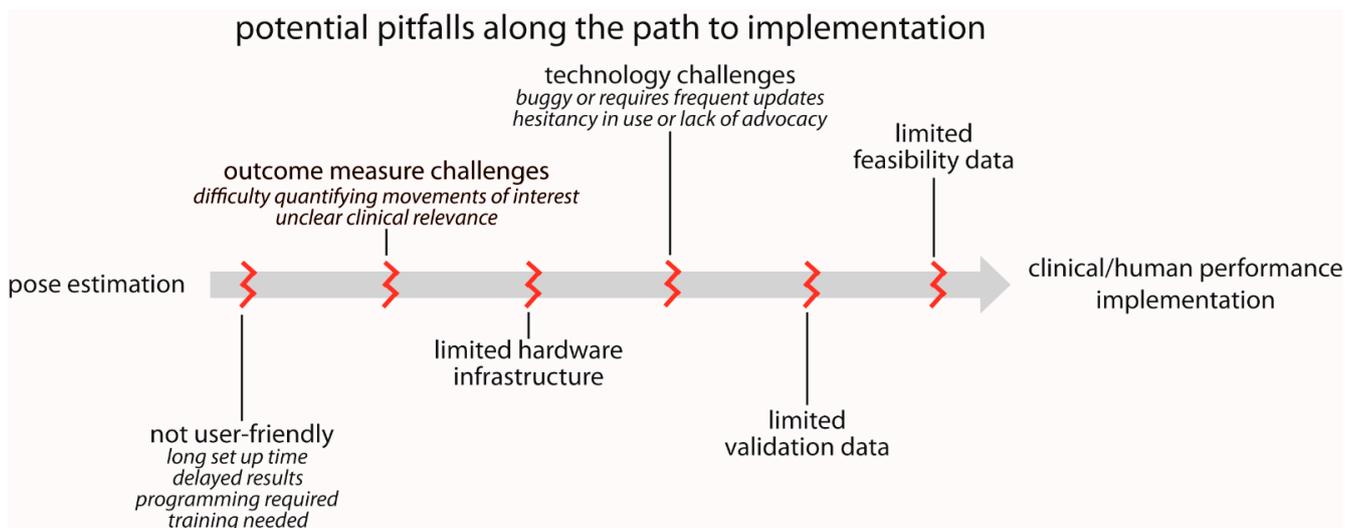
There is also a need for testing of sensitivity, specificity, feasibility, and reliability. When a new clinical outcome measure is developed, a first step should be to establish criterion-validity or construct validity between the pose estimated measures and age-concurrent, clinician-coded, gold-standard clinical measures. Next, using receiver operating characteristic (ROC) analysis, sensitivity and specificity should be compared to assess the ability of the new pose estimated measure in predicting dichotomous outcomes (e.g., motor impaired vs. motor unimpaired). Area under the curve (AUC) should further be computed as a measure of the ability to distinguish between groups. Finally, it is important to evaluate the feasibility and acceptability of the new pose estimation protocol. One way to assess

feasibility is to assess the number of completed and submitted usable videos by patients (i.e., the total number of videos submitted divided by the number expected, multiplied by 100). One way to assess acceptability is through satisfaction questionnaires/surveys. For example, after video submission, patients, families of patients (if patients are children), and clinicians can complete a brief satisfaction questionnaire/survey regarding their experience using the pose estimation protocol.

These potential pitfalls along the path to implementation are shown in Figure 6.



**Figure 5.** Common application limitations with current pose estimation algorithms and challenges with using these algorithms outside of the laboratory. These applications commonly require three-dimensional kinematics of multiple people moving at relatively high speeds to be tracked in environments with background figures (e.g., irrelevant people and objects shaped similarly to people). This leads to challenges with segment occlusion, unintentional capture of background figures, and registration of multiple cameras. Additionally, using current algorithms for scenarios different than the training dataset (e.g., different movements, different types of clothing or equipment being worn, different lighting) may lead to reduced accuracy in the predicted kinematics or, potentially, failure of the algorithm. Finally, most algorithms do not predict kinematic metrics that are required for some applications (e.g., head acceleration to assess concussion risk), and limitations with using current algorithms on time-series data make it challenging to accurately derive these metrics.



**Figure 6.** Common pitfalls that must be avoided on the path to widespread implementation of pose estimation applications for human health and performance. These are covered in greater detail in the “**What are limitations of pose estimation?**” section of the manuscript.

## 6. Conclusions

The emergence and continued development of human pose estimation approaches offer exciting potential for making quantitative assessments of human movement kinematics significantly more accessible. Pose estimation algorithms directly address an important and widespread need for low cost, easy to use, accessible technologies that enable human movement tracking in virtually any environment, including the home, clinic, classroom, playing field, and other ‘in the wild’ settings. Applications in health and human performance have begun to emerge in the literature, but we perceive that these technologies are still in their relative infancy with regard to the potential for research and clinical implementation. Many limitations persist, and it is important that users are aware of these and adjust expectations accordingly. However, we anticipate that applications of pose estimation in human health and performance will continue to expand in coming years, and these technologies will provide powerful tools for capturing meaningful aspects of human movement that have been difficult to capture with conventional techniques.

**Author Contributions:** Conceptualization: J.S., K.M.C.-A., C.O.P., R.D.R., M.F.V. and R.T.R.; writing—original draft preparation: J.S., K.M.C.-A., C.O.P., R.D.R., M.F.V. and R.T.R.; writing—review and editing: J.S., K.M.C.-A., C.O.P., R.D.R., M.F.V. and R.T.R.; funding acquisition: C.O.P., R.D.R., M.F.V. and R.T.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by a Kennedy Krieger Institute Goldstein Innovation Grant to RDR, NIH grant R21 AG059184 to RTR.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We would like to thank the Janney Program within the Johns Hopkins University Applied Physics Laboratory for providing partial funding for this work, which nurtures a culture of discovery, embraces risk, and welcomes being at the center of a vibrant innovation ecosystem.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. All authors have read and agreed to the published version of the manuscript.

## References

1. Mündermann, L.; Corazza, S.; Andriacchi, T.P. The Evolution of Methods for the Capture of Human Movement Leading to Markerless Motion Capture for Biomechanical Applications. *J. NeuroEng. Rehabil.* **2006**, *3*, 6. [[CrossRef](#)]
2. Baker, R. The History of Gait Analysis before the Advent of Modern Computers. *Gait Posture* **2007**, *26*, 23–28. [[CrossRef](#)] [[PubMed](#)]
3. Roether, C.L.; Omlor, L.; Christensen, A.; Giese, M.A. Critical Features for the Perception of Emotion from Gait. *J. Vis.* **2009**, *9*, 1–32. [[CrossRef](#)] [[PubMed](#)]
4. Michalak, J.; Troje, N.F.; Fischer, J.; Vollmar, P.; Heidenreich, T.; Schulte, D. Embodiment of Sadness and Depression-Gait Patterns Associated with Dysphoric Mood. *Psychosom. Med.* **2009**, *71*, 580–587. [[CrossRef](#)] [[PubMed](#)]
5. Kendon, A. Movement Coordination in Social Interaction: Some Examples Described. *Acta Psychol.* **1970**, *32*, 101–125. [[CrossRef](#)]
6. Martinez, G.H.; Raaj, Y.; Idrees, H.; Xiang, D.; Joo, H.; Simon, T.; Sheikh, Y. Single-Network Whole-Body Pose Estimation. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 6982–6991. [[CrossRef](#)]
7. Insafutdinov, E.; Andriluka, M.; Pishchulin, L.; Tang, S.; Levinkov, E.; Andres, B.; Schiele, B. ArtTrack: Articulated Multi-Person Tracking in the Wild. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6457–6465. [[CrossRef](#)]
8. Insafutdinov, E.; Pishchulin, L.; Andres, B.; Andriluka, M.; Schiele, B. Deepercut: A Deeper, Stronger, and Faster Multi-Person Pose Estimation Model. In *Computer Vision—ECCV 2016; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer International Publishing: Cham, Switzerland, 2016; pp. 34–50. [[CrossRef](#)]
9. Pishchulin, L.; Insafutdinov, E.; Tang, S.; Andres, B.; Andriluka, M.; Gehler, P.; Schiele, B. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4929–4937. [[CrossRef](#)]
10. Toshev, A.; Szegedy, C. DeepPose: Human Pose Estimation via Deep Neural Networks. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1653–1660. [[CrossRef](#)]
11. Nath, T.; Mathis, A.; Chen, A.C.; Patel, A.; Bethge, M.; Mathis, M.W. Using DeepLabCut for 3D Markerless Pose Estimation across Species and Behaviors. *Nat. Protoc.* **2019**, *14*, 2152–2176. [[CrossRef](#)] [[PubMed](#)]
12. 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](#)]
13. Cao, Z.; Simon, T.; Wei, S.E.; Sheikh, Y. Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1302–1310. [[CrossRef](#)]
14. Cornman, H.L.; Stenum, J.; Roemmich, R.T. Video-Based Quantification of Human Movement Frequency Using Pose Estimation. *bioRxiv* **2021**. [[CrossRef](#)]
15. Stenum, J.; Rossi, C.; Roemmich, R.T. Two-Dimensional Video-Based Analysis of Human Gait Using Pose Estimation. *PLoS Comput. Biol.* **2021**, *17*, e1008935. [[CrossRef](#)]
16. Kwolek, B.; Michalczyk, A.; Krzeszowski, T.; Switonski, A.; Josinski, H.; Wojciechowski, K. Calibrated and Synchronized Multi-View Video and Motion Capture Dataset for Evaluation of Gait Recognition. *Multimed. Tools Appl.* **2019**, *78*, 32437–32465. [[CrossRef](#)]
17. Wang, L.; Tan, T.; Ning, H.; Hu, W. Silhouette Analysis-Based Gait Recognition for Human Identification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1505–1518. [[CrossRef](#)]
18. Holte, M.B.; Cuong, T.; Trivedi, M.M.; Moeslund, T.B. Human Pose Estimation and Activity Recognition from Multi-View Videos: Comparative Explorations of Recent Developments. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 538–552. [[CrossRef](#)]
19. Isaacs, J.; Foo, S. Hand Pose Estimation for American Sign Language Recognition. In Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory, Atlanta, GA, USA, 16 March 2004; pp. 132–136. [[CrossRef](#)]
20. Cronin, N.J. Using Deep Neural Networks for Kinematic Analysis: Challenges and Opportunities. *J. Biomech.* **2021**, *123*, 110460. [[CrossRef](#)] [[PubMed](#)]
21. Seethapathi, N.; Wang, S.; Saluja, R.; Blohm, G.; Kording, K.P. Movement Science Needs Different Pose Tracking Algorithms. *arXiv* **2019**, arXiv:1907.10226.
22. Arac, A. Machine Learning for 3D Kinematic Analysis of Movements in Neurorehabilitation. *Curr. Neurol. Neurosci. Rep.* **2020**, *20*, 29. [[CrossRef](#)]
23. Andriluka, M.; Pishchulin, L.; Gehler, P.; Schiele, B. 2D Human Pose Estimation: New Benchmark and State of the Art Analysis. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3683–3693. [[CrossRef](#)]
24. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In *Computer Vision—ECCV 2014; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer International Publishing: Cham, Switzerland, 2014; pp. 740–755. [[CrossRef](#)]

25. Sato, K.; Nagashima, Y.; Mano, T.; Iwata, A.; Toda, T. Quantifying Normal and Parkinsonian Gait Features from Home Movies: Practical Application of a Deep Learning–Based 2D Pose Estimator. *PLoS ONE* **2019**, *14*, e0223549. [[CrossRef](#)]
26. Chambers, C.; Kong, G.; Wei, K.; Kording, K. Pose Estimates from Online Videos Show That Side-by-Side Walkers Synchronize Movement under Naturalistic Conditions. *PLoS ONE* **2019**, *14*, e0217861. [[CrossRef](#)]
27. 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](#)]
28. Ota, M.; Tateuchi, H.; Hashiguchi, T.; Kato, T.; Ogino, Y.; Yamagata, M.; Ichihashi, N. Verification of Reliability and Validity of Motion Analysis Systems during Bilateral Squat Using Human Pose Tracking Algorithm. *Gait Posture* **2020**, *80*, 62–67. [[CrossRef](#)]
29. Nakano, N.; Sakura, T.; Ueda, K.; Omura, L.; Kimura, A.; Iino, Y.; Fukashiro, S.; Yoshioka, S. Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Front. Sports Act. Living* **2020**, *2*, 50. [[CrossRef](#)]
30. Zago, M.; Luzzago, M.; Marangoni, T.; De Cecco, M.; Tarabini, M.; Galli, M. 3D Tracking of Human Motion Using Visual Skeletonization and Stereoscopic Vision. *Front. Bioeng. Biotechnol.* **2020**, *8*, 181. [[CrossRef](#)] [[PubMed](#)]
31. D’Antonio, E.; Taborri, J.; Palermo, E.; Rossi, S.; Patanè, F. A Markerless System for Gait Analysis Based on OpenPose Library. In Proceedings of the 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Dubrovnik, Croatia, 25–28 May 2020; pp. 1–6. [[CrossRef](#)]
32. Boswell, M.A.; Uhlrich, S.D.; Kidziński, Ł.; Thomas, K.; Kolesar, J.A.; Gold, G.E.; Beaupre, G.S.; Delp, S.L. A Neural Network to Predict the Knee Adduction Moment in Patients with Osteoarthritis Using Anatomical Landmarks Obtainable from 2D Video Analysis. *Osteoarthr. Cartil.* **2021**, *29*, 346–356. [[CrossRef](#)] [[PubMed](#)]
33. Lu, M.; Poston, K.; Pfefferbaum, A.; Sullivan, E.V.; Fei-Fei, L.; Pohl, K.M.; Niebles, J.C.; Adeli, E. Vision-Based Estimation of MDS-UPDRS Gait Scores for Assessing Parkinson’s Disease Motor Severity. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020*; Springer International Publishing: Cham, Switzerland, 2020; Volume 12263, pp. 637–647. [[CrossRef](#)]
34. Kidziński, Ł.; Yang, B.; Hicks, J.L.; Rajagopal, A.; Delp, S.L.; Schwartz, M.H. Deep Neural Networks Enable Quantitative Movement Analysis Using Single-Camera Videos. *Nat. Commun.* **2020**, *11*, 4054. [[CrossRef](#)] [[PubMed](#)]
35. Ota, M.; Tateuchi, H.; Hashiguchi, T.; Ichihashi, N. Verification of Validity of Gait Analysis Systems during Treadmill Walking and Running Using Human Pose Tracking Algorithm. *Gait Posture* **2021**, *85*, 290–297. [[CrossRef](#)]
36. Ossmy, O.; Adolph, K.E. Real-Time Assembly of Coordination Patterns in Human Infants. *Curr. Biol.* **2020**, *30*, 4553–4562. [[CrossRef](#)] [[PubMed](#)]
37. Chambers, C.; Seethapathi, N.; Saluja, R.; Loeb, H.; Pierce, S.R.; Bogen, D.K.; Prosser, L.; Johnson, M.J.; Kording, K.P. Computer Vision to Automatically Assess Infant Neuromotor Risk. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 2431–2442. [[CrossRef](#)]
38. Fang, H.; Xie, S.; Lu, C. RMPE: Regional Multi-Person Pose Estimation. *arXiv* **2018**, arXiv:1612.001737.
39. Cao, Z.; Hidalgo Martinez, G.; Simon, T.; Wei, S.-E.; Sheikh, Y.A. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 172–186. [[CrossRef](#)]
40. Viswakumar, A.; Rajagopalan, V.; Ray, T.; Parimi, C. Human Gait Analysis Using OpenPose. In Proceedings of the IEEE International Conference Image Information Processing, Shimla, India, 15–17 November 2019; pp. 310–314. [[CrossRef](#)]
41. Ye, Q.; Yuan, S.; Kim, T.-K. Spatial Attention Deep Net with Partial PSO for Hierarchical Hybrid Hand Pose Estimation. *arXiv* **2016**, arXiv:1604.03334.
42. Ivekovic, S.; Trucco, E. Human Body Pose Estimation with PSO. In Proceedings of the 2006 IEEE International Conference on Evolutionary Computation, Vancouver, BC, Canada, 16–21 July 2006; pp. 1256–1263. [[CrossRef](#)]
43. Lee, K.-Z.; Liu, T.-W.; Ho, S.-Y. Model-Based Pose Estimation of Human Motion Using Orthogonal Simulated Annealing. In *Intelligent Data Engineering and Automated Learning*; Springer: Berlin/Heidelberg, Germany, 2003. [[CrossRef](#)]
44. Halvorsen, K.; Söderström, T.; Stokes, V.; Lanshammar, H. Using an Extended Kalman Filter for Rigid Body Pose Estimation. *J. Biomech. Eng.* **2004**, *127*, 475–483. [[CrossRef](#)] [[PubMed](#)]
45. Janabi-Sharifi, F.; Marey, M. A Kalman-Filter-Based Method for Pose Estimation in Visual Servoing. *IEEE Trans. Robot.* **2010**, *26*, 939–947. [[CrossRef](#)]
46. Fanning, P.A.J.; Sparaci, L.; Dissanayake, C.; Hocking, D.R.; Vivanti, G. Functional Play in Young Children with Autism and Williams Syndrome: A Cross-Syndrome Comparison. *Child Neuropsychol. J. Norm. Abnorm. Dev. Child. Adolesc.* **2021**, *27*, 125–149. [[CrossRef](#)]
47. Kretch, K.S.; Franchak, J.M.; Adolph, K.E. Crawling and Walking Infants See the World Differently. *Child Dev.* **2014**, *85*, 1503–1518. [[CrossRef](#)] [[PubMed](#)]
48. LeBarton, E.S.; Iverson, J.M. Fine Motor Skill Predicts Expressive Language in Infant Siblings of Children with Autism. *Dev. Sci.* **2013**, *16*, 815–827. [[CrossRef](#)]
49. Masek, L.R.; Paterson, S.J.; Golinkoff, R.M.; Bakeman, R.; Adamson, L.B.; Owen, M.T.; Pace, A.; Hirsh-Pasek, K. Beyond Talk: Contributions of Quantity and Quality of Communication to Language Success across Socioeconomic Strata. *Infancy Off. J. Int. Soc. Infant Stud.* **2021**, *26*, 123–147. [[CrossRef](#)]
50. Le, H.; Hoch, J.E.; Ossmy, O.; Adolph, K.E.; Fern, X.; Fern, A. Modeling Infant Free Play Using Hidden Markov Models. In Proceedings of the 2021 IEEE International Conference on Development and Learning (ICDL), Beijing, China, 23–26 August 2021; pp. 1–6. [[CrossRef](#)]

51. Doroniewicz, I.; Ledwoń, D.J.; Affanasowicz, A.; Kieszczyńska, K.; Latos, D.; Matyja, M.; Mitas, A.W.; Myśliwiec, A. Writhing Movement Detection in Newborns on the Second and Third Day of Life Using Pose-Based Feature Machine Learning Classification. *Sensors* **2020**, *20*, 5986. [[CrossRef](#)] [[PubMed](#)]
52. Iverson, J.M.; Shic, F.; Wall, C.A.; Chawarska, K.; Curtin, S.; Estes, A.; Gardner, J.M.; Hutman, T.; Landa, R.J.; Levin, A.R. Early Motor Abilities in Infants at Heightened versus Low Risk for ASD: A Baby Siblings Research Consortium (BSRC) Study. *J. Abnorm. Psychol.* **2019**, *128*, 69. [[CrossRef](#)]
53. Iverson, J.M. Developing Language in a Developing Body: The Relationship between Motor Development and Language Development. *J. Child Lang.* **2010**, *37*, 229–261. [[CrossRef](#)]
54. Alcock, K.J.; Krawczyk, K. Individual Differences in Language Development: Relationship with Motor Skill at 21 Months. *Dev. Sci.* **2010**, *13*, 677–691. [[CrossRef](#)]
55. Adolph, K.E.; Hoch, J.E. Motor Development: Embodied, Embedded, Enculturated, and Enabling. *Annu. Rev. Psychol.* **2019**, *70*, 141–164. [[CrossRef](#)] [[PubMed](#)]
56. Rosenbaum, P.; Paneth, N.; Leviton, A.; Goldstein, M.; Bax, M.; Damiano, D.; Dan, B.; Jacobsson, B. A Report: The Definition and Classification of Cerebral Palsy April 2006. *Dev. Med. Child Neurol. Suppl.* **2007**, *109*, 8–14. [[PubMed](#)]
57. Novak, I.; Morgan, C.; Adde, L.; Blackman, J.; Boyd, R.N.; Brunstrom-Hernandez, J.; Cioni, G.; Damiano, D.; Darrach, J.; Eliasson, A.-C. Early, Accurate Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment. *JAMA Pediatr.* **2017**, *171*, 897–907. [[CrossRef](#)] [[PubMed](#)]
58. Geethanath, S.; Vaughan, J.T.J. Accessible Magnetic Resonance Imaging: A Review. *J. Magn. Reson. Imaging JMRI* **2019**, *49*, e65–e77. [[CrossRef](#)] [[PubMed](#)]
59. Adde, L.; Helbostad, J.L.; Jensenius, A.R.; Taraldsen, G.; Grunewaldt, K.H.; Støen, R. Early Prediction of Cerebral Palsy by Computer-based Video Analysis of General Movements: A Feasibility Study. *Dev. Med. Child Neurol.* **2010**, *52*, 773–778. [[CrossRef](#)]
60. Rahmati, H.; Aamo, O.M.; Stavdahl, Ø.; Dragon, R.; Adde, L. Video-Based Early Cerebral Palsy Prediction Using Motion Segmentation. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2014**, 3779–3783. [[CrossRef](#)]
61. Ihlen, E.A.; Støen, R.; Boswell, L.; de Regnier, R.-A.; Fjørtoft, T.; Gaebler-Spira, D.; Ladori, C.; Loenneken, M.C.; Msall, M.E.; Möinichen, U.I. Machine Learning of Infant Spontaneous Movements for the Early Prediction of Cerebral Palsy: A Multi-Site Cohort Study. *J. Clin. Med.* **2020**, *9*, 5. [[CrossRef](#)] [[PubMed](#)]
62. Chawarska, K.; Klin, A.; Paul, R.; Volkmar, F. Autism Spectrum Disorder in the Second Year: Stability and Change in Syndrome Expression. *J. Child Psychol. Psychiatry* **2007**, *48*, 128–138. [[CrossRef](#)] [[PubMed](#)]
63. Landa, R.; Garrett-Mayer, E. Development in Infants with Autism Spectrum Disorders: A Prospective Study. *J. Child Psychol. Psychiatry* **2006**, *47*, 629–638. [[CrossRef](#)]
64. Maenner, M.J.; Shaw, K.A.; Baio, J. Prevalence of Autism Spectrum Disorder among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *MMWR Surveill. Summ.* **2020**, *69*, 1–12. [[CrossRef](#)]
65. Gordon-Lipkin, E.; Foster, J.; Peacock, G. Whittling down the Wait Time: Exploring Models to Minimize the Delay from Initial Concern to Diagnosis and Treatment of Autism Spectrum Disorder. *Pediatr. Clin.* **2016**, *63*, 851–859. [[CrossRef](#)]
66. Ning, M.; Daniels, J.; Schwartz, J.; Dunlap, K.; Washington, P.; Kalantarian, H.; Du, M.; Wall, D.P. Identification and Quantification of Gaps in Access to Autism Resources in the United States: An Infodemiological Study. *J. Med. Internet Res.* **2019**, *21*, e13094. [[CrossRef](#)] [[PubMed](#)]
67. Brian, J.A.; Smith, I.M.; Zwaigenbaum, L.; Bryson, S.E. Cross-site Randomized Control Trial of the Social ABCs Caregiver-mediated Intervention for Toddlers with Autism Spectrum Disorder. *Autism Res.* **2017**, *10*, 1700–1711. [[CrossRef](#)]
68. Dawson, G.; Rogers, S.; Munson, J.; Smith, M.; Winter, J.; Greenson, J.; Donaldson, A.; Varley, J. Randomized, Controlled Trial of an Intervention for Toddlers with Autism: The Early Start Denver Model. *Pediatrics* **2010**, *125*, e17–e23. [[CrossRef](#)] [[PubMed](#)]
69. Landa, R.J.; Holman, K.C.; O’Neill, A.H.; Stuart, E.A. Intervention Targeting Development of Socially Synchronous Engagement in Toddlers with Autism Spectrum Disorder: A Randomized Controlled Trial. *J. Child Psychol. Psychiatry* **2011**, *52*, 13–21. [[CrossRef](#)]
70. Crippa, A.; Salvatore, C.; Perego, P.; Forti, S.; Nobile, M.; Molteni, M.; Castiglioni, I. Use of Machine Learning to Identify Children with Autism and Their Motor Abnormalities. *J. Autism Dev. Disord.* **2015**, *45*, 2146–2156. [[CrossRef](#)]
71. Karatsidis, A.; Richards, R.E.; Konrath, J.M.; van den Noort, J.C.; Schepers, H.M.; Bellusci, G.; Harlaar, J.; Veltink, P.H. Validation of Wearable Visual Feedback for Retraining Foot Progression Angle Using Inertial Sensors and an Augmented Reality Headset. *J. NeuroEng. Rehabil.* **2018**, *15*, 78. [[CrossRef](#)] [[PubMed](#)]
72. Guo, Y.; Deligianni, F.; Gu, X.; Yang, G.-Z. 3-D Canonical Pose Estimation and Abnormal Gait Recognition with a Single RGB-D Camera. *IEEE Robot. Autom. Lett.* **2019**, *4*, 3617–3624. [[CrossRef](#)]
73. Kondragunta, J.; Hirtz, G. Gait Parameter Estimation of Elderly People Using 3D Human Pose Estimation in Early Detection of Dementia. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2020**, 5798–5801. [[CrossRef](#)]
74. Chaaoui, A.A.; Padilla-López, J.R.; Flórez-Revuelta, F. Abnormal Gait Detection with RGB-D Devices Using Joint Motion History Features. In Proceedings of the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 4–8 May 2015; pp. 1–6. [[CrossRef](#)]
75. Li, G.; Liu, T.; Yi, J. Wearable Sensor System for Detecting Gait Parameters of Abnormal Gaits: A Feasibility Study. *IEEE Sens. J.* **2018**, *18*, 4234–4241. [[CrossRef](#)]

76. Chen, Y.; Du, R.; Luo, K.; Xiao, Y. Fall Detection System Based on Real-Time Pose Estimation and SVM. In Proceedings of the 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Nanchang, China, 26–28 March 2021; pp. 990–993. [[CrossRef](#)]
77. Bian, Z.; Hou, J.; Chau, L.; Magnenat-Thalmann, N. Fall Detection Based on Body Part Tracking Using a Depth Camera. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 430–439. [[CrossRef](#)]
78. Huang, Z.; Liu, Y.; Fang, Y.; Horn, B.K.P. Video-Based Fall Detection for Seniors with Human Pose Estimation. In Proceedings of the 2018 4th International Conference on Universal Village (UV), Boston, MA, USA, 21–24 October 2018; pp. 1–4. [[CrossRef](#)]
79. Mehrizi, R.; Peng, X.; Tang, Z.; Xu, X.; Metaxas, D.; Li, K. Toward Marker-Free 3D Pose Estimation in Lifting: A Deep Multi-View Solution. *arXiv* **2018**, arXiv:1802.01741.
80. Han, S.; Lee, S. A Vision-Based Motion Capture and Recognition Framework for Behavior-Based Safety Management. *Autom. Constr.* **2013**, *35*, 131–141. [[CrossRef](#)]
81. Han, S.; Achar, M.; Lee, S.; Peña-Mora, F. Empirical Assessment of a RGB-D Sensor on Motion Capture and Action Recognition for Construction Worker Monitoring. *Vis. Eng.* **2013**, *1*, 6. [[CrossRef](#)]
82. Blanchard, N.; Skinner, K.; Kemp, A.; Scheirer, W.; Flynn, P. “Keep Me In, Coach!”: A Computer Vision Perspective on Assessing ACL Injury Risk in Female Athletes. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 7–11 January 2019; pp. 1366–1374. [[CrossRef](#)]
83. Vukicevic, A.M.; Macuzic, I.; Mijailovic, N.; Peulic, A.; Radovic, M. Assessment of the Handcart Pushing and Pulling Safety by Using Deep Learning 3D Pose Estimation and IoT Force Sensors. *Expert Syst. Appl.* **2021**, *183*, 115371. [[CrossRef](#)]
84. Mehrizi, R.; Peng, X.; Metaxas, D.N.; Xu, X.; Zhang, S.; Li, K. Predicting 3-D Lower Back Joint Load in Lifting: A Deep Pose Estimation Approach. *IEEE Trans. Hum. Mach. Syst.* **2019**, *49*, 85–94. [[CrossRef](#)]
85. Krosshaug, T.; Nakamae, A.; Boden, B.P.; Engebretsen, L.; Smith, G.; Slauterbeck, J.R.; Hewett, T.E.; Bahr, R. Mechanisms of Anterior Cruciate Ligament Injury in Basketball: Video Analysis of 39 Cases. *Am. J. Sports Med.* **2007**, *35*, 359–367. [[CrossRef](#)]
86. Olsen, O.-E.; Myklebust, G.; Engebretsen, L.; Bahr, R. Injury Mechanisms for Anterior Cruciate Ligament Injuries in Team Handball: A Systematic Video Analysis. *Am. J. Sports Med.* **2004**, *32*, 1002–1012. [[CrossRef](#)] [[PubMed](#)]
87. Kim, W.; Sung, J.; Saakes, D.; Huang, C.; Xiong, S. Ergonomic Postural Assessment Using a New Open-Source Human Pose Estimation Technology (OpenPose). *Int. J. Ind. Ergon.* **2021**, *84*, 103164. [[CrossRef](#)]
88. Li, Y.; Wang, C.; Cao, Y.; Liu, B.; Tan, J.; Luo, Y. Human Pose Estimation Based In-Home Lower Body Rehabilitation System. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8. [[CrossRef](#)]
89. Cordella, F.; Di Corato, F.; Zollo, L.; Siciliano, B. A Robust Hand Pose Estimation Algorithm for Hand Rehabilitation. In *New Trends in Image Analysis and Processing—ICIAP 2013*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8158, pp. 1–10. [[CrossRef](#)]
90. Zhu, Y.; Lu, W.; Gan, W.; Hou, W. A Contactless Method to Measure Real-Time Finger Motion Using Depth-Based Pose Estimation. *Comput. Biol. Med.* **2021**, *131*, 104282. [[CrossRef](#)]
91. Milosevic, B.; Leardini, A.; Farella, E. Kinect and Wearable Inertial Sensors for Motor Rehabilitation Programs at Home: State of the Art and an Experimental Comparison. *Biomed. Eng. Online* **2020**, *19*, 25. [[CrossRef](#)]
92. Tao, Y.; Hu, H.; Zhou, H. Integration of Vision and Inertial Sensors for 3D Arm Motion Tracking in Home-Based Rehabilitation. *Int. J. Robot. Res.* **2007**, *26*, 607–624. [[CrossRef](#)]
93. Ranasinghe, I.; Dantu, R.; Albert, M.V.; Watts, S.; Ocana, R. Cyber-Physiotherapy: Rehabilitation to Training. In Proceedings of the 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM), Bordeaux, France, 17–21 May 2021; pp. 1054–1057.
94. Tao, T.; Yang, X.; Xu, J.; Wang, W.; Zhang, S.; Li, M.; Xu, G. Trajectory Planning of Upper Limb Rehabilitation Robot Based on Human Pose Estimation. In Proceedings of the 2020 17th International Conference on Ubiquitous Robots (UR), Kyoto, Japan, 22–26 June 2020; pp. 333–338. [[CrossRef](#)]
95. Palermo, M.; Moccia, S.; Migliorelli, L.; Frontoni, E.; Santos, C.P. Real-Time Human Pose Estimation on a Smart Walker Using Convolutional Neural Networks. *Expert Syst. Appl.* **2021**, *184*, 115498. [[CrossRef](#)]
96. Airò Farulla, G.; Pianu, D.; Cempini, M.; Cortese, M.; Russo, L.O.; Indaco, M.; Nerino, R.; Chimienti, A.; Oddo, C.M.; Vitiello, N. Vision-Based Pose Estimation for Robot-Mediated Hand Telerehabilitation. *Sensors* **2016**, *16*, 208. [[CrossRef](#)]
97. Sarsfield, J.; Brown, D.; Sherkat, N.; Langensiepen, C.; Lewis, J.; Taheri, M.; McCollin, C.; Barnett, C.; Selwood, L.; Standen, P.; et al. Clinical Assessment of Depth Sensor Based Pose Estimation Algorithms for Technology Supervised Rehabilitation Applications. *Int. J. Med. Inf.* **2019**, *121*, 30–38. [[CrossRef](#)] [[PubMed](#)]
98. Xu, W.; Chatterjee, A.; Zollhoefer, M.; Rhodin, H.; Mehta, D.; Seidel, H.-P.; Theobalt, C. MonoPerfCap: Human Performance Capture from Monocular Video. *arXiv* **2018**, arXiv:1708.02136.
99. Habermann, M.; Xu, W.; Zollhoefer, M.; Pons-Moll, G.; Theobalt, C. LiveCap: Real-Time Human Performance Capture from Monocular Video. *arXiv* **2019**, arXiv:1810.02648.
100. Wang, J.; Qiu, K.; Peng, H.; Fu, J.; Zhu, J. AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training Assistance. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 374–382. [[CrossRef](#)]

101. Einfalt, M.; Dampeyrou, C.; Zecha, D.; Lienhart, R. Frame-Level Event Detection in Athletics Videos with Pose-Based Convolutional Sequence Networks. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 42–50. [[CrossRef](#)]
102. Einfalt, M.; Zecha, D.; Lienhart, R. Activity-Conditioned Continuous Human Pose Estimation for Performance Analysis of Athletes Using the Example of Swimming. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 446–455. [[CrossRef](#)]
103. Güler, R.A.; Neverova, N.; Kokkinos, I. DensePose: Dense Human Pose Estimation in the Wild. *arXiv* **2018**, arXiv:1802.00434.
104. Patacchiola, M.; Cangelosi, A. Head Pose Estimation in the Wild Using Convolutional Neural Networks and Adaptive Gradient Methods. *Pattern Recognit.* **2017**, *71*, 132–143. [[CrossRef](#)]
105. Fong, C.-M.; Blackburn, J.T.; Norcross, M.F.; McGrath, M.; Padua, D.A. Ankle-Dorsiflexion Range of Motion and Landing Biomechanics. *J. Athl. Train.* **2011**, *46*, 5–10. [[CrossRef](#)]
106. Caccese, J.B.; Buckley, T.A.; Tierney, R.T.; Rose, W.C.; Glutting, J.J.; Kaminski, T.W. Sex and Age Differences in Head Acceleration during Purposeful Soccer Heading. *Res. Sports Med.* **2018**, *26*, 64–74. [[CrossRef](#)] [[PubMed](#)]
107. Cerveri, P.; Pedotti, A.; Ferrigno, G. Kinematical Models to Reduce the Effect of Skin Artifacts on Marker-Based Human Motion Estimation. *J. Biomech.* **2005**, *38*, 2228–2236. [[CrossRef](#)]
108. Joo, H.; Neverova, N.; Vedaldi, A. Exemplar Fine-Tuning for 3D Human Model Fitting Towards In-the-Wild 3D Human Pose Estimation. *arXiv* **2020**, arXiv:200403686.
109. Habibie, I.; Xu, W.; Mehta, D.; Pons-Moll, G.; Theobalt, C. In the Wild Human Pose Estimation Using Explicit 2D Features and Intermediate 3D Representations. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 10897–10906. [[CrossRef](#)]
110. Mehta, D.; Rhodin, H.; Casas, D.; Fua, P.; Sotnychenko, O.; Xu, W.; Theobalt, C. Monocular 3D Human Pose Estimation in The Wild Using Improved CNN Supervision. *arXiv* **2017**, arXiv:1611.09813.
111. Mehta, D.; Sridhar, S.; Sotnychenko, O.; Rhodin, H.; Shafiei, M.; Seidel, H.-P.; Xu, W.; Casas, D.; Theobalt, C. VNect: Real-Time 3D Human Pose Estimation with a Single RGB Camera. *ACM Trans. Graph.* **2017**, *36*, 1–14. [[CrossRef](#)]
112. Gilbert, A.; Trumble, M.; Malleson, C.; Hilton, A.; Collomosse, J. Fusing Visual and Inertial Sensors with Semantics for 3D Human Pose Estimation. *Int. J. Comput. Vis.* **2019**, *127*, 381–397. [[CrossRef](#)]
113. Malleson, C.; Gilbert, A.; Trumble, M.; Collomosse, J.; Hilton, A.; Volino, M. Real-Time Full-Body Motion Capture from Video and IMUs. In Proceedings of the 2017 International Conference on 3D Vision (3DV), Qingdao, China, 10–12 October 2017; pp. 449–457. [[CrossRef](#)]
114. WHO. *International Classification of Functioning, Disability and Health: ICF 2001, Title of Beta 2, Full Version: International Classification of Functioning and Disability: ICDH-2 (WHO Document no. WHO/HSC/ACE/99.2)*; WHO: Geneva, Switzerland, 2001.
115. Fugl-Meyer, A.R.; Jääskö, L.; Leyman, I.; Olsson, S.; Steglind, S. The Post-Stroke Hemiplegic Patient. 1. a Method for Evaluation of Physical Performance. *Scand. J. Rehabil. Med.* **1975**, *7*, 13–31. [[PubMed](#)]
116. Yozbatiran, N.; Der-Yeghiaian, L.; Cramer, S.C. A Standardized Approach to Performing the Action Research Arm Test. *Neurorehabil. Neural Repair* **2008**, *22*, 78–90. [[CrossRef](#)] [[PubMed](#)]
117. Duncan, P.W.; Wallace, D.; Lai, S.M.; Johnson, D.; Embretson, S.; Laster, L.J. The Stroke Impact Scale Version 2.0. *Stroke* **1999**, *30*, 2131–2140. [[CrossRef](#)]
118. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Automated Assessment of Levodopa-Induced Dyskinesia: Evaluating the Responsiveness of Video-Based Features. *Parkinsonism Relat. Disord.* **2018**, *53*, 42–45. [[CrossRef](#)]
119. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Automated Vision-Based Analysis of Levodopa-Induced Dyskinesia with Deep Learning. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea, 11–15 July 2017. [[CrossRef](#)] [[PubMed](#)]
120. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Vision-Based Assessment of Parkinsonism and Levodopa-Induced Dyskinesia with Pose Estimation. *J. Neuroeng. Rehabil.* **2018**, *15*, 97. [[CrossRef](#)] [[PubMed](#)]
121. Liu, Y.; Chen, J.; Hu, C.; Ma, Y.; Ge, D.; Miao, S.; Xue, Y.; Li, L. Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 1952–1961. [[CrossRef](#)] [[PubMed](#)]
122. Aung, N.; Bovonsunthonchai, S.; Hiengkaew, V.; Tretriluxana, J.; Rojasavastera, R.; Pheung-Phrarattanatrai, A. Concurrent Validity and Intratester Reliability of the Video-Based System for Measuring Gait Poststroke. *Physiother. Res. Int. J. Res. Clin. Phys. Ther.* **2020**, *25*, e1803. [[CrossRef](#)]
123. Shin, J.H.; Yu, R.; Ong, J.N.; Lee, C.Y.; Jeon, S.H.; Park, H.; Kim, H.-J.; Lee, J.; Jeon, B. Quantitative Gait Analysis Using a Pose-Estimation Algorithm with a Single 2D-Video of Parkinson’s Disease Patients. *J. Park. Dis.* **2021**, *11*, 1271–1283. [[CrossRef](#)]
124. Ng, K.-D.; Mehdizadeh, S.; Iaboni, A.; Mansfield, A.; Flint, A.; Taati, B. Measuring Gait Variables Using Computer Vision to Assess Mobility and Fall Risk in Older Adults With Dementia. *IEEE J. Transl. Eng. Health Med.* **2020**, *8*, 2100609. [[CrossRef](#)]
125. Li, T.; Chen, J.; Hu, C.; Ma, Y.; Wu, Z.; Wan, W.; Huang, Y.; Jia, F.; Gong, C.; Wan, S.; et al. Automatic Timed Up-and-Go Sub-Task Segmentation for Parkinson’s Disease Patients Using Video-Based Activity Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 2189–2199. [[CrossRef](#)]