

**FEASIBILITY OF MARKERLESS MOTION CAPTURE IN CLINICAL GAIT  
ANALYSIS IN CHILDREN WITH CEREBRAL PALSY**

Afet Mustafaoglu

Biology of Physical Activity Master's Thesis in  
Biomechanics  
Faculty of Sport and Health Sciences  
University of Jyväskylä  
Autumn 2021  
Supervisors: Neil Cronin, Harri Piitulainen,  
Francesco Cenni

## ABSTRACT

Mustafaoglu, A. 2023. Feasibility of markerless motion capture in clinical gait analysis in children with cerebral palsy. Faculty of Sport and Health Sciences, University of Jyväskylä, Master's thesis, 75 pp., 6 appendices.

The main purpose of this study is to explore the feasibility of deep learning based markerless motion capture in clinical settings. This study compares the markerless motion capture (Blazepose) to the gold standard marker-based system (Vicon) by assessing the hip, knee, and ankle joint flexion angles of cerebral palsy (CP) patients and their typically developed (TD) peers.

The participant group included six CP patients and six TD individuals. The participants walked on an 8-meter gait path at a self-selected pace. 11 Vicon cameras (200 Hz) and 3 GoPro (60 Hz) cameras were used for the marker-based and markerless system setup. Both systems were synchronized and recorded simultaneously. The keypoint trajectories from Blazepose were obtained by feeding the images collected with GoPros as an input to the algorithm. Further analysis included calibration, 3D reconstruction, and data filtering in Matlab. Skeletal modeling and joint angle calculations were conducted in OpenSim for both systems to eliminate the methodological difference. SPM1D Matlab package was used for statistical analysis.

Significant differences were observed in ankle and hip joint angles between the Blazepose and Vicon systems at specific gait cycle phases in both the CP and TD groups. The ankle angle showed significant differences in the CP group at 0.7–1.3% ( $p < 0.016$ ) of the gait cycle and in the TD group at 38–46% ( $p < 0.016$ ). For hip flexion, significant differences in the CP group were noted at 0.81–1.81% ( $p < 0.016$ ), 13–40% ( $p < 0.016$ ), and 89–93% ( $p < 0.016$ ) of the gait cycle, while in the TD group, a significant difference was observed at 75–84% of the gait cycle ( $p < 0.016$ ). No significant difference was observed for the knee angle in both groups.

The results of this study highlight the potential use and certain limitations of markerless motion capture systems like Blazepose in clinical settings. While showing promise in certain aspects of joint angle tracking, the study emphasizes the need for more refined datasets and advanced algorithms to enhance the accuracy and reliability of such systems, especially for clinical applications involving CP patients.

Key words: deep learning, markerless motion capture, clinical gait analysis, cerebral palsy, kinematics

## **ABBREVIATIONS**

AI	artificial intelligence
CNN	convolutional neural network
CP	cerebral palsy
DoF	degrees of freedom
GMFCS	gross motor function classification system
HPE	human pose estimation
TD	typically developing
2D	two-dimensional
3D	three-dimensional

# CONTENTS

## ABSTRACT

1	INTRODUCTION .....	1
2	CEREBRAL PALSY.....	3
2.1	Epidemiology and classification.....	4
2.1.1	Topographic description for cerebral palsy .....	6
2.1.2	Classification of tone and movement abnormality in cerebral palsy.....	7
2.1.3	The gross motor classification system .....	8
2.2	Human gait .....	10
2.2.1	Gait in cerebral palsy .....	11
2.2.2	Gait patterns in cerebral palsy .....	12
2.2.3	Importance of gait analysis in cerebral palsy .....	14
3	MACHINE LEARNING AND DEEP LEARNING.....	17
3.1	How does machine learning work? .....	18
3.2	Types of machine learning .....	19
3.3	Artificial neural networks.....	21
3.3.1	Structure of neural networks.....	21
3.3.2	Training neural networks.....	22
3.4	Deep neural networks .....	23
4	MOTION ANALYSIS .....	26
4.1	Marker-based motion analysis.....	26
4.2	Markerless motion capture .....	29
4.2.1	Earlier markerless motion capture methods .....	30
4.2.2	Human pose estimation .....	31
5	PURPOSE OF THE STUDY .....	38
6	METHODS.....	40

6.1	Participants .....	41
6.2	Experimental setup and equipments .....	41
6.3	Experimental protocol .....	43
6.4	Data analysis for markerless data .....	43
6.4.1	Calibration .....	45
6.4.2	3D Reconstruction .....	46
6.4.3	Filtering the keypoint trajectories.....	47
6.5	Skeletal modelling and joint angle calculation.....	49
6.6	Statistical analysis .....	50
7	RESULTS.....	51
8	DISCUSSION.....	50
8.1	Ankle joint assesment.....	40
8.2	Knee joint assesment .....	41
8.3	Hip joint assesment.....	42
8.4	Accuracy concerns.....	43
8.5	Technical and methodological considerations.....	48
8.6	Usibility and potential of markerless motion capture.....	49
9	CONCLUSION .....	50
	REFERENCES .....	53
	APPENDICES	

## 1. INTRODUCTION

The term "gait" refers to a series of limb movements that results in locomotion. Gait analysis can be used to monitor a patient's recovery after a therapy, as well as diagnose illnesses caused by aging, injuries, or neurological disorders in clinical applications (Albuquerque et al. 2021). Gait analysis is an important part of clinical evaluations for determining functional outcomes following neurological or musculoskeletal disorder like cerebral palsy. Gait deficits are common in a range of health conditions, including stroke, Parkinson's disease, cerebral palsy, and spinal cord injury (Lonini et al. 2022). In terms of the kinematic and kinetic parameters of patients with lower limb impairment differs from that of healthy people and it is crucial to measure these properties (Takeda et al. 2021).

Marker-based motion tracking has become the most commonly used method for determining accurate kinematic gait parameters and became the gold standard over the years (Latorre et al. 2019). Typically, this application involves placing markers on specific anatomical landmarks and using several motion cameras to extract kinematic characteristics from the observed motion (Albuquerque et al. 2021). These systems allow for precise spatiotemporal and kinematic analysis however, their high cost and stationariness, have limited their application to controlled surroundings like research labs and specialist personnel, as well as setup and calibration processes that can be time-consuming and impracticable. In addition, these systems require extensive financial resources to build such facilities. (Albuquerque et al. 2021; Latorre et al. 2019) Although the marker-based system is the gold standard, the obligation to wear markers may affect normal movement patterns, and the marker placement is subject to inter-session and inter-tester variability. Moreover, marker placements do not always correlate to the exact anatomical joint centers they represent, and soft tissue artefact might result in further measurement inaccuracies. (Needham et al. 2021) Another limitation of this method is that in clinical settings patients with neuromuscular disorders like cerebral palsy can get tired very quickly and feel uncomfortable with the markers which can affect the normal gait pattern.

As a result, over the last few years, new approaches based on alternative technologies have been investigated. Recent improvements in computer vision and deep learning, as well as the availability of annotated data sets of people and body landmarks, have made it possible to detect body landmarks (keypoints) automatically from single videos without the need for reflective markers. As a result, various open-source algorithms for 2D and 3D pose estimation are now

available, for both clinical and sport applications. Deep learning based markerless systems require one or more cameras, and they mainly utilize two types of cameras, either depth or RGB cameras. Such methods refer to processing two-dimensional (2D) images, captured by a single (or multiple) depth or RGB camera, to locate a set of anatomical landmarks (or joint centers) in the three-dimensional (3D) capture volume and then with simple geometry to calculate the joint angles (i.e., the angle between two vectors that each represent a body segment) (Cronin 2021; Lonini et al. 2022; Vafadar et al. 2021).

Markerless motion capture approaches based on artificial intelligence, particularly deep learning, have made substantial development in recent years. (Vafadar et al. 2021). These systems are cost-effective and easy-to-use alternatives, especially for clinical gait analysis (Cronin 2021). However, the accuracy of these systems has not been studied enough, especially in clinical settings that require high accuracy. Thus, the main purpose of this study is to explore the feasibility of deep learning based markerless motion capture in clinical settings. This study compares the markerless motion capture (Blazepose) to the gold standard marker-based system (Vicon) by assessing hip, knee and ankle joint flexion angles of cerebral palsy (CP) patients and their typically developed (TD) peers.

## 2. Cerebral palsy

The term "cerebral palsy" refers to a group of conditions with varying degrees of severity that share some developmental characteristics rather than a specific disease entity (Graham et al. 2016). The official definition is as follows and was established by an international panel in April 2006 by consultants and by feedback and recommendations from several reviewers' responses to drafts made available to the global community:

*“Cerebral palsy (CP) describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems.”* (Rosenbaum et al. 2007, p.9)

CP is a neurodevelopmental condition that starts in early childhood and lasts throughout one's lifetime (Rosenbaum et al. 2007). It is characterized by abnormal movement, posture, and muscle tone, which restrict the affected person's activity (Stavsky et al. 2017). CP is likely caused by a lesion in the fetal brain (Graham & Selber 2003), which is linked to non-progressive abnormalities, or a complication of premature birth (Rosenbaum et al. 2007). The clinical expression of CP may change over time, even though CP itself is a non-progressive condition (Stavsky et al. 2017). It is therefore preferable to think of it as permanent, though not constant (Morgan & McGinley 2018). Although it is important to add that, typically, a newborn with CP is born without any musculoskeletal anomalies or deformities, the development of scoliosis, dislocation of the hip, and contractures happens during the child's rapid growth (Graham & Selber 2003).

Since CP is an umbrella term and categorized by the movement disorder and its topographical distribution (Graham & Selber 2003), this section will describe the epidemiology, various classification systems for CP, as well as the abnormal gait patterns caused by the initial brain lesion and their effects on the functional ability of CP patients.



## 2.1 Epidemiology and classification

The most prevalent motor disability in children is CP and has historically been shown to range between 1.5 and 2.5 per 1,000 live births, with the majority of these population-based registries being in Europe (Cans 2007). However, recent studies in the United States (Christensen et al. 2014), Taiwan (Chang et al. 2015), and Egypt (Shehata et al. 2014) have discovered prevalence rates in adults between the ages of 4 and 48 exceeding 3 per 1,000 live births (Graham et al. 2016). Children born prematurely before 28 weeks of gestation have a much greater prevalence, ranging from 40 to 100 per 1000 live births (Cans 2007). Although gestational age is a more reliable indicator, low birthweight is frequently used in surveillance studies since it commonly denotes preterm birth (Figure 1) (Wimalasundera & Stevenson 2016).

The majority of patients shared two characteristics in the initial clinical descriptions of children with cerebral palsy: premature birth and challenging labor with neonatal asphyxia, which is caused by a lack of oxygen and blood flow to the brain during birth (Graham et al. 2016; Little J. 1861). Even though neonatal asphyxia is considered as the most common cause, it is anticipated that prenatal factors account for about 75% of CP, with perinatal hypoxia (deficiency of oxygen) accounting for less than 10%. Less than 20% of instances involve postnatal infection or damage, and 30% of patients have no known underlying cause. (Table 1) (Morgan & McGinley 2018)

About 90% of the time, CP is caused by damage to the healthy brain rather than abnormal development in the fetal brain (Graham et al. 2016; Morgan & McGinley 2018). However, the exact cause of CP is still not fully understood (Sewell et al., 2014). It is important to note that the timing and type of the injury to the brain can cause different co-morbidities, such as epilepsy, dysphagia, cognitive and communication impairments, and can affect the type of musculoskeletal deformities of varying severity (Graham et al. 2016; Wimalasundera & Stevenson 2016). Thus, CP is classified by topographical distribution, motor type, and functional level to provide appropriate intervention and prognostications (Wimalasundera & Stevenson 2016).

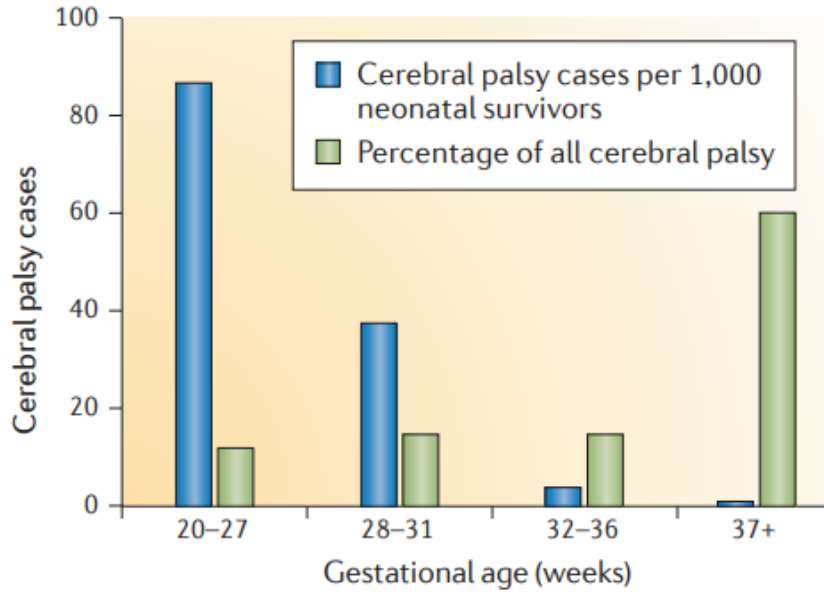


FIGURE 1. Correlation between gestational age and cerebral palsy prevalence. It is important to add that despite the fact that there is a high correlation between prematurity and the likelihood of developing the condition, most people with cerebral palsy are born at term (Graham et al. 2016; Reid et al. 2016).

TABLE 1. Possible causes and risk factors for CP. Prenatal refers to the period before birth. Perinatal includes the time from the 20th week of gestation through the first week after birth. Postnatal care starts from the moment of birth and extends into the early weeks and months of the individual. Table modified from Peterson & Walton (2016)

<b>Prenatal (75%)</b>	<b>Perinatal (8%)</b>	<b>Postnatal (17%)</b>
Prematurity	Hypoxia	Infection
Inflammation	Hypoglycaemia	Injury
Infection (ToRCH)	Jaundice	
Ingested toxins	Infection	
Genetic syndromes		
Congenital brain malformation		
Metabolic conditions		

### 2.1.1 Topographic description for cerebral palsy

The distribution of limb involvement was classified as bilateral or unilateral by the European Surveillance of Cerebral Palsy Group (SCPE) (Surveillance of Cerebral Palsy in Europe, 2000). Bilateral further divides into diplegia, triplegia, and quadriplegia. In diplegia, the lower limbs are significantly more impaired than the upper limbs, which often only exhibit fine motor dysfunction. In triplegia, bilateral lower limb involvement and unilateral upper limb involvement are the typical patterns. On the same side where the upper limb is involved, the lower limb is almost always more impacted. All four limbs, as well as the trunk, are affected by quadriplegia. Tetraplegia and “whole-body involvement” are other terms used to describe quadriplegia. (Figure 2) (Graham et al. 2016; Wimalasundera & Stevenson 2016)

Unilateral further divides into monoplegia and hemiplegia. Only one limb is affected in monoplegia, and typically it is the lower limb. Whereas one side of the body is affected in hemiplegia, and often the upper limb is more affected than the lower limb. (Figure 2) (Graham et al. 2016; Wimalasundera & Stevenson 2016)

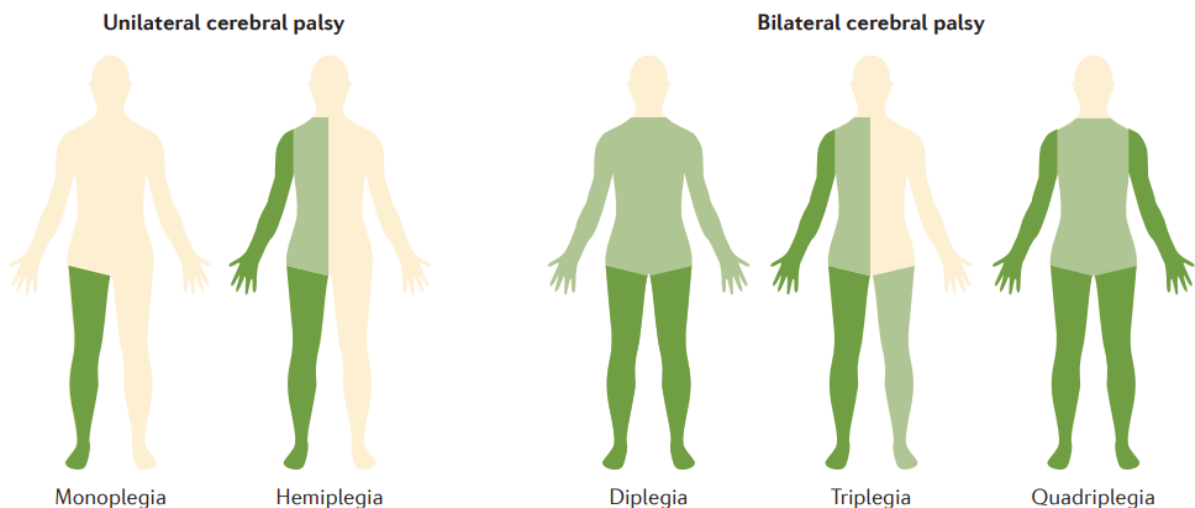


FIGURE 2. Topographical distribution in cerebral palsy (Graham et al., 2016).

### **2.1.2 Classifications of tone and movement abnormality in cerebral palsy**

The three main types of CP are spastic, dyskinetic, and ataxic, with each describing distinct movement and/or muscle tone dysfunctions. Spastic CP accounts for between 85% and 90% of all cases of CP (1/3 unilateral, 2/3 bilateral), making it the most prevalent type (Wimalasundera & Stevenson 2016). Spasticity essentially refers to an overactivation of the stretch reflex with respect to velocity in response to a passive muscular stretch. The increased resistance to stretching (hyper-resistance) in CP is a result of this irregular spike in muscle activity. (Graham et al. 2016; Lance 1980) At rest, background muscle tone is electrically silent, whereas in spasticity, excessive background muscle tone may be evident (Peacock 2009). Additionally, spasticity induces large-scale movements as opposed to delicate and individual ones, while voluntary motions become slow and demanding (Gulati & Sondhi 2018).

Individuals with dyskinetic CP frequently exhibit multiple forms of involuntary movement. These movements can lead to stiffness in the limbs, especially during attempted movements or in response to emotional stimuli. The tendon reflexes in these patients may vary, ranging from normal to challenging to elicit. (Gulati & Sondhi 2018; Sewell et al. 2014; Wimalasundera & Stevenson 2016)

Ataxic CP is a relatively uncommon condition (4% of CP). Decreased muscle tone (generalized hypotonia) with loss of muscle coordination is commonly observed in this type of CP. Developmental milestones related to motor skills and language are often delayed in these cases. However, it's important to note that ataxia typically shows improvement over time. (Gulati & Sondhi 2018; Sewell et al. 2014; Wimalasundera & Stevenson 2016)

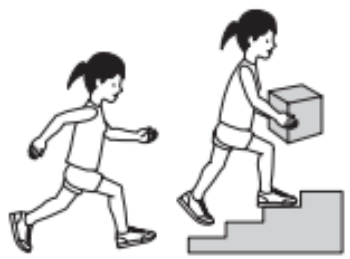
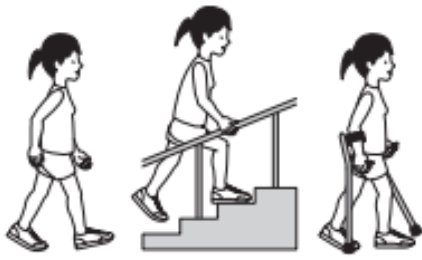



### **2.1.3 The gross motor classification system**

To comprehend the clinical picture of cerebral palsy, it is essential to assess characteristics at both the impairment and function levels. This includes factors like the limb distribution of the clinical "syndromes" (number of limbs with impaired motor control) and the type and severity of the motor disorder. (Gorter et al. 2004) Functional classification systems, which focus on children's abilities and participation in society, are considered the ideal method for classifying children with cerebral palsy. Unlike physical classification systems (topographic description and classification of tone and movement abnormalities) , they provide information on activity status and level of participation in life situations. (Gorter et al. 2004; Palisano et al. 1997; Sewell et al. 2014)

The Gross Motor Function Classification System (GMFCS) is an age-dependent system (2-4 years, 4-6 years, 6-12 years, and 12-18 years) that categorizes children into five levels according to their mobility abilities. This classification reflects their overall gross motor skills and the extent of motor impairment they experience (Figure 3). (Palisano et al. 1997; Sewell et al. 2014) Level 1 indicates minimal disability, while Level 5 indicates complete reliance on equipment or caregivers to maintain posture (Figure 3). These levels are not influenced by the type or distribution of CP motor impairments. They solely reflect the primary method by which the individual mobilizes and do not assess the quality of movement. (Wimalasundera & Stevenson 2016) The significance of the system lies in its ability to predict walking function, assess the required orthotic and equipment support for improved mobility, and evaluate the possibility of needing orthopedic surgery (Sewell et al. 2014).

The gross motor function measure assesses physical activity by examining various activities like rolling, walking, running, and jumping. In GMFCS level 1, approximately 90% of potential motor development occurs around the age of 5 years. For level 2, it's around 4.5 years; level 3 is about 3.75 years; level 4 is at 3.5 years; and level 5 reaches around 2.75 years. After this period, motor ability tends to plateau, especially in more severe GMFCS levels, making this information crucial for providing a prognosis regarding walking. An example of this is a 4-year-old child with cerebral palsy at GMFCS level 4, who is not expected to attain independent walking in the future. (Wimalasundera & Stevenson 2016)

# GMFCS E & R between 12<sup>th</sup> and 18<sup>th</sup> birthday: Descriptors and illustrations

	<p><b>GMFCS Level I</b></p> <p>Youth walk at home, school, outdoors and in the community. Youth are able to climb curbs and stairs without physical assistance or a railing. They perform gross motor skills such as running and jumping but speed, balance and coordination are limited.</p>
	<p><b>GMFCS Level II</b></p> <p>Youth walk in most settings but environmental factors and personal choice influence mobility choices. At school or work they may require a hand held mobility device for safety and climb stairs holding onto a railing. Outdoors and in the community youth may use wheeled mobility when traveling long distances.</p>
	<p><b>GMFCS Level III</b></p> <p>Youth are capable of walking using a hand-held mobility device. Youth may climb stairs holding onto a railing with supervision or assistance. At school they may self-propel a manual wheelchair or use powered mobility. Outdoors and in the community youth are transported in a wheelchair or use powered mobility.</p>
	<p><b>GMFCS Level IV</b></p> <p>Youth use wheeled mobility in most settings. Physical assistance of 1-2 people is required for transfers. Indoors, youth may walk short distances with physical assistance, use wheeled mobility or a body support walker when positioned. They may operate a powered chair, otherwise are transported in a manual wheelchair.</p>
	<p><b>GMFCS Level V</b></p> <p>Youth are transported in a manual wheelchair in all settings. Youth are limited in their ability to maintain antigravity head and trunk postures and control leg and arm movements. Self-mobility is severely limited, even with the use of assistive technology.</p>

GMFCS descriptors: Palisano et al. (1997) Dev Med Child Neurol 39:214-23  
CanChild: www.canchild.ca

Illustrations copyright © Kerr Graham, Bill Reid and Adrienne Harvey,  
The Royal Children's Hospital, Melbourne

FIGURE 3. Descriptors and illustrations of the Gross Motor Classification System (GMFCS) for ages 12-18 years (CanChild n.d.).

## 2.2 Human gait

Each step involves a constantly shifting alignment between the body and the supporting foot during stance and a selective advancement of the limb segments during swing in order to perform the fundamental activities necessary for walking. As a result, a variety of motion patterns are performed by the hip, knee, and ankle. The basic divisions of the gait cycle are stance and swing, which further divide into eight functional patterns (Figure 4). The functional significance of the various motions happening at the particular joints can be directly determined by breaking down the walking patterns into phases. Additionally, it offers a way to connect the simultaneous movements of the separate joints into patterns of whole-limb movement. It is a crucial strategy for understanding the functional impacts of impairment. (Perry 1992)

In the typical walking process, the foot-off phase happens at roughly 60% of the gait cycle. As a result, the stance phase constitutes approximately 60% of the gait cycle, while the swing phase comprises 40%. At about 10% and 50% of the cycle, opposite foot-off and opposite initial contact happen, respectively. Thus, there are two instances of "double support" during walking, each lasting around 10% of the cycle. The first double-support period occurs immediately after initial contact, while the second occurs just before foot-off. During the loading response, which is a deceleration phase, the body absorbs the shock of impact. Following the initial double-support period, there is a period of single support, accounting for approximately 40% of the cycle. A second period of double support known as "pre-swing" occurs in the late stance phase. It starts at around 50% of the gait cycle and lasts until foot-off on the stance side. Therefore, the loading response and the pre-swing on the opposite side share the same time frame and represent essentially the same event. (Figure 4) (Gage & Schwartz 2009b) Following sections of this chapter describes gait in cerebral palsy, the main gait patterns that are observed in CP, and the importance of clinical gait analysis in CP.

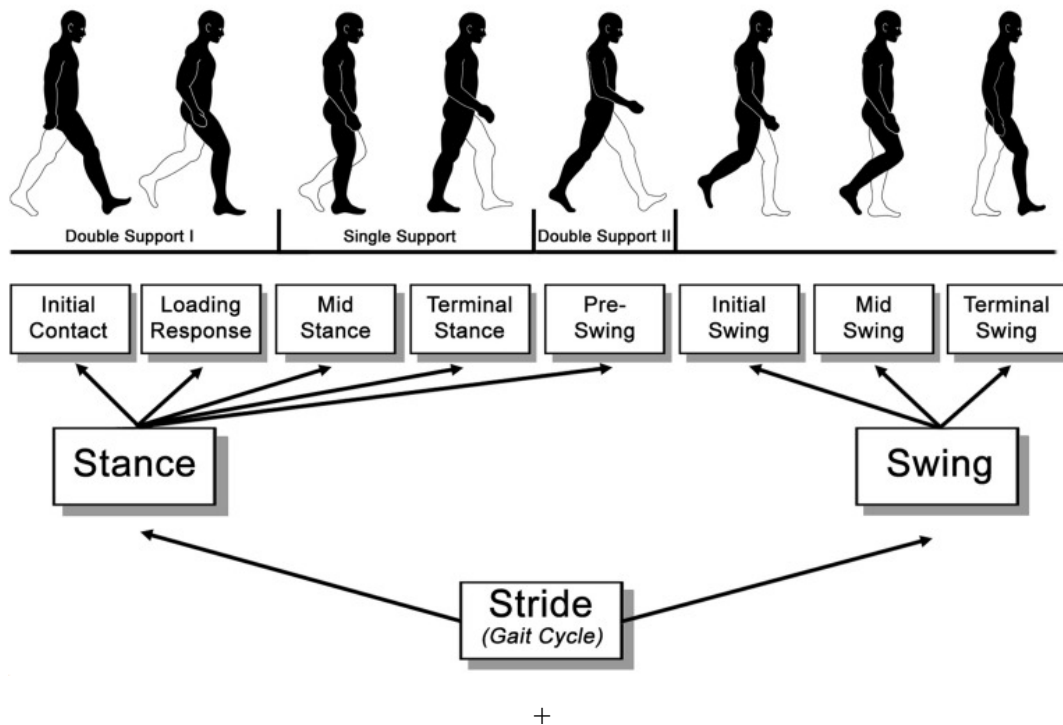


FIGURE 4. Eight functional patterns of the gait cycle (Stöckel et al. 2015).

### 2.2.1 Gait in cerebral palsy

Cerebral palsy arises from a primary injury in the central nervous system, yet its clinical symptoms become evident in the peripheral neuromuscular system, particularly in skeletal muscles. A common complication of cerebral palsy is the occurrence of muscle contractures, which result from increased passive muscle force and lead to restricted joint movement. In children with cerebral palsy, muscle contractures are frequent, and their muscles tend to be shorter, smaller, and have reduced diameter muscle fibers (Rose et al. 1994), indicating a decrease in overall strength. (Graham et al. 2016) The elongated sarcomeres observed in individuals with cerebral palsy are believed to be a contributing factor to reduced force generation and functional weakness (Morgan & McGinley 2018). Musculoskeletal abnormalities result in disadvantageous biomechanical conditions and less energy-efficient gait patterns. Thus, it is important to identify pathological gait in CP since the focus of treatment is on managing symptoms, optimizing developmental potential, and reducing musculoskeletal abnormalities. (Wimalasundera & Stevenson 2016)



### **2.2.2 Gait patterns in cerebral palsy**

Pathological gait often lacks five crucial attributes found in normal walking. These include stability during stance, sufficient foot clearance in the swinging motion, appropriate pre-positioning of the foot in the swing phase, an adequate step length, and energy conservation. (Gage 1993; Gage & Schwartz 2009b)

Patients with CP often exhibit distinct gait patterns. However, clinicians can recognize and identify certain patterns. From day to day and stride to stride, spastic motor patterns are generally quite consistent. However, over the long term, aging and intervention frequently lead to changes. (J. Rodda & Graham 2001) These patterns have been developed for patients with unilateral or bilateral spastic CP. The primary emphasis of these patterns lies in the sagittal plane, describing walking patterns that involve multiple joints. In order to suggest management strategies to address the underlying issues, these patterns attempted to identify and group frequent kinematic abnormalities across numerous joints. (Sangeux & Stephane 2015)

Winters et al. (1987) proposed the first classification for gait patterns in unilateral spastic CP. The system has four types, each representing greater gait disruption. Drop foot in late swing followed by absent first rocker in early stance are the main disturbances observed in type 1 patients. Overactive plantarflexors, a weak tibialis anterior muscle, or poor selective motor control may all contribute to the related deficit. In type 2, patients exhibit drop foot and decreased dorsiflexion in stance. Compared to type 1, a contracture of the plantarflexors may cause further disability. Type 3 patients exhibit the same deficits of types 1 and 2, with additional spasticity or contracture of the hamstring or rectus muscles. They may also exhibit decreased or delayed knee flexion in swing. In addition to type 1-3 deficits, type 4 patients exhibit impairments at the hip with reduced extension and the pelvis with increased anterior tilt. Rodda & Graham (2001) further added patients with hyperextension at the knee and transverse plane deviations at the hip since Winters et al. have defined these impairments in the sagittal plane (Figure 5). (Sangeux & Stephane 2015)

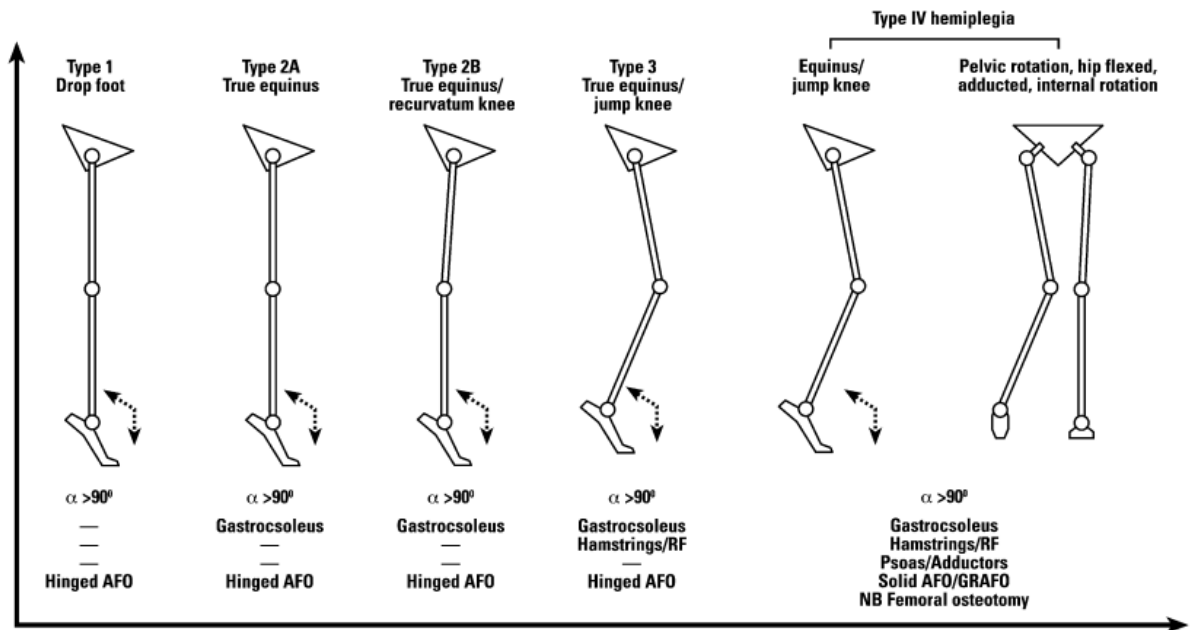


FIGURE 5. Gait patterns in spastic hemiplegia. (AFO: ankle foot orthosis, GRAFO: ground reaction foot orthosis) (J. Rodda & Graham 2001)

Rodda & Graham (2001 & 2004), have based their classification for bilateral spastic CP on previous works by Miller et al. (1995) and Sutherland & Davids (1993). Five groups were listed under the system: mild gait, real equinus, jump gait, apparent equinus, and crouch gait. In mid-stance, patients with true equinus exhibit excessive plantarflexion. In cases of jump gait, there is excessive plantarflexion and knee flexion whereas in cases of apparent equinus, patients display normal ankle kinematics but knee flexion, in mid-stance. The patients with crouch gait exhibit excessive dorsiflexion and knee flexion in mid-stance. It is important to note that the impairment in two limbs may exhibit different levels of involvement. Thus, the authors proposed an asymmetric group. Lastly, patients with mild gait do not exhibit any significant impairment in the saggital plane; however, they may exhibit impairments in other planes. (Figure 6) (Sangeux & Stephane 2015)

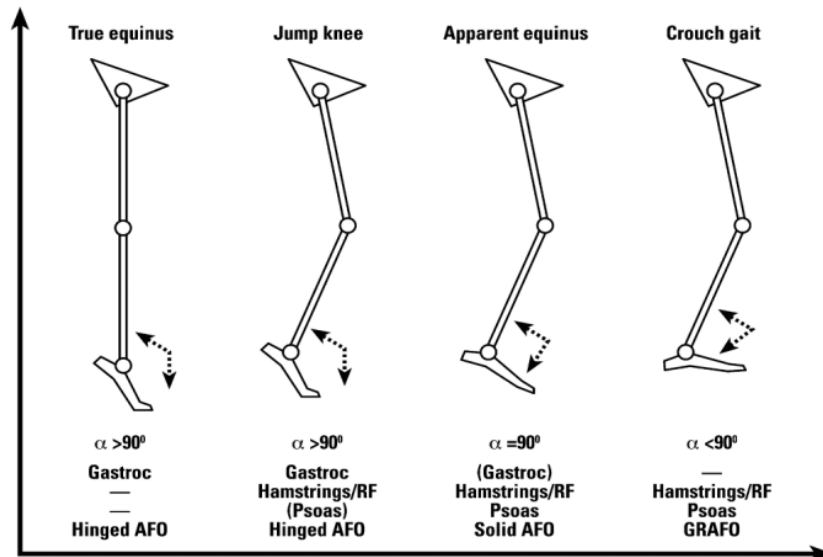


FIGURE 6. Gait patterns in spastic diplegia. (AFO: ankle foot orthosis, GRAFO: ground reaction foot orthosis) (J. Rodda & Graham 2001)

Both of these classification systems were developed based on clinical observations and were aimed at identifying common kinematic deviations observed in real-world clinical scenarios (Sangeux & Stephane 2015).

### 2.2.3 Importance of gait analysis in cerebral palsy

Gait analysis typically involves the following elements in a motion analysis lab: quantitative 3-dimensional motion measurement (kinematics), measurements of moments and power production occurring in the major lower extremity articulations (kinetics), the on-off signals of specific muscles and/or muscle groups (dynamic EMG), and dynamic foot pressure during gait. Therefore, gait analysis offers an accurate "snapshot in time" of a person's walking or running gait. (Dugan & Shilt 2020; Gage & Stout 2009)

As previously described, patients with CP exhibit neuromotor abnormalities, including primary (neurological) and secondary (growth) deficits. Primary impairments caused by cerebral palsy are increased muscle tone, impaired selective motor control, and impaired balance mechanisms. Secondary impairments, which are caused by primary impairments, encompass muscle shortening, muscle weakness, and decreased joint range of motion. (Baker et al. 2016; Fosdahl et al. 2019; Gage & Schwartz 2009) The dynamic nature of CP requires ongoing evaluation of

the patient's clinical condition, treatment objectives, and planning for further treatment modalities. The common evaluation approach entails a clinical examination of the child, assessing aspects such as joint mobility, muscle tone, spasticity, muscle strength, and selective control level. However, there is a need for a precise dynamic assessment tool like gait analysis that can quantitatively measure the patient's motor functions in comparison to their typically developing peers and track changes in motor patterns resulting from specific treatments. (Baker et al. 2016) According to Baker et al. (2016), there are four potential reasons for performing a clinical gait analysis: (1) diagnose disease entities; (2) assess the severity, extent, or nature of a disease or injury; (3) monitor progress in the presence or absence of intervention; and (4) predict the outcome of intervention or non-intervention.

Gait analysis is crucial to the treatment's ability to be tailored. The decision to switch from reversible to irreversible therapies may be supported by longitudinal follow-up of a patient using gait analysis as well as decisions about the best time to implement each treatment option. Data from gait analysis and clinical examinations can be used to determine and specify the specific treatment objectives for children with CP. These treatments include botulinum neurotoxin treatment, orthopaedic surgery, physical therapy, and orthotic management (AFO-ankle foot orthosis). (Baker et al. 2016) Using the wrong treatment approach can lead to negative outcomes, including a decline in functionality, unnecessary surgeries, diminished quality of life, and increased financial burden (Dugan & Shilt 2020). Gait analysis is also beneficial for tracking the progression of disorders like muscular dystrophy over time by comparing patients' multiple assessments. It also plays a crucial role in assessing the impact of orthotic devices or medications on an individual's gait. (Gage & Stout 2009)

According to the study by Lofterød et al. (2007), gait analysis proved to be a valuable source of additional information that significantly influenced preoperative surgical planning. As a result of their study, 70% (42 out of 60 patients) of the patient's treatment plan changed after gait analysis. Out of 60 patients who have been recommended surgical treatment, 49 were recommended for a surgical procedure, while 11 were recommended for a nonoperative treatment after the gait analysis. Moreover, Wren et al. (2013) showed that for 84% of patients, surgical modifications were made after gait analysis recommendations.

Abnormal cerebral control, muscle contractures, or lever-arm dysfunction force individuals to incorporate compensatory abnormalities into their gait as a way to cope with the challenges

arising from their condition. Gait analysis is also important to differentiate the coping mechanisms caused by these musculoskeletal impairments(Gage & Stout 2009)

Addressing gait complexities through gait analysis is a demanding task, and without it, finding solutions becomes extremely challenging. Thus, interpreting motion analysis requires a holistic approach, considering the patient's medical history and other assessments for a comprehensive understanding. Relying solely on clinical evaluations is inadequate to fully comprehend the complexities of gait disorders associated with cerebral palsy. (Dugan & Shilt 2020; Gage 1993; Gage & Stout 2009)

### 3. MACHINE LEARNING AND DEEP LEARNING

Artificial intelligence (AI) is a wide-ranging field that utilizes technologies to construct machines and computers with the capacity to imitate human cognitive abilities, including understanding and responding to language, analyzing data, and making recommendations. While often perceived as a standalone entity, artificial intelligence comprises a set of technologies integrated into systems to enable reasoning, learning, and effective problem-solving for complex issues. AI systems require the capacity to learn features by identifying patterns from unprocessed data. This capability is referred to as machine learning, which makes machine learning a subfield of AI. (Figure 7) (Galbusera et al. 2019; Goodfellow et al. 2016b; Mintz & Brodie 2019)

Machine learning refers to programming computers with the aim of optimizing performance criteria using example data or prior experience. Models are structured with defined parameters and learning entails employing computer programs to refine these parameters based on training data or past experiences. These models can serve the purpose of predicting future outcomes, extracting features from data, or even combining both functions. Statistics plays a crucial role in building mathematical models for machine learning, as it revolves around drawing inferences from sample data. In the realm of computer science, there are two key aspects: firstly, the need for efficient algorithms in training to handle optimization and data management, and secondly, the requirement for streamlined representation and algorithms for efficient inference once learning is complete. (Alpaydin 2009; Goodfellow et al. 2016b)

Deep learning, a branch of machine learning, mimics human brain functions by assessing various data sets concurrently. It iteratively processes this data through multiple evaluations called layers (often referred to as deep neural networks), with each subsequent assessment relying on the results of the preceding one. Every evaluation takes place in a distinct layer, building on the output of the layer before it. (Mintz & Brodie 2019) Its design is inspired by the human brain's neural architecture, enabling it to automatically learn features and representations from vast amounts of data (LeCun et al. 2015).

This chapter further describes the basic principles and concepts of machine learning and deep learning, as well as the methods and common machine learning algorithms.

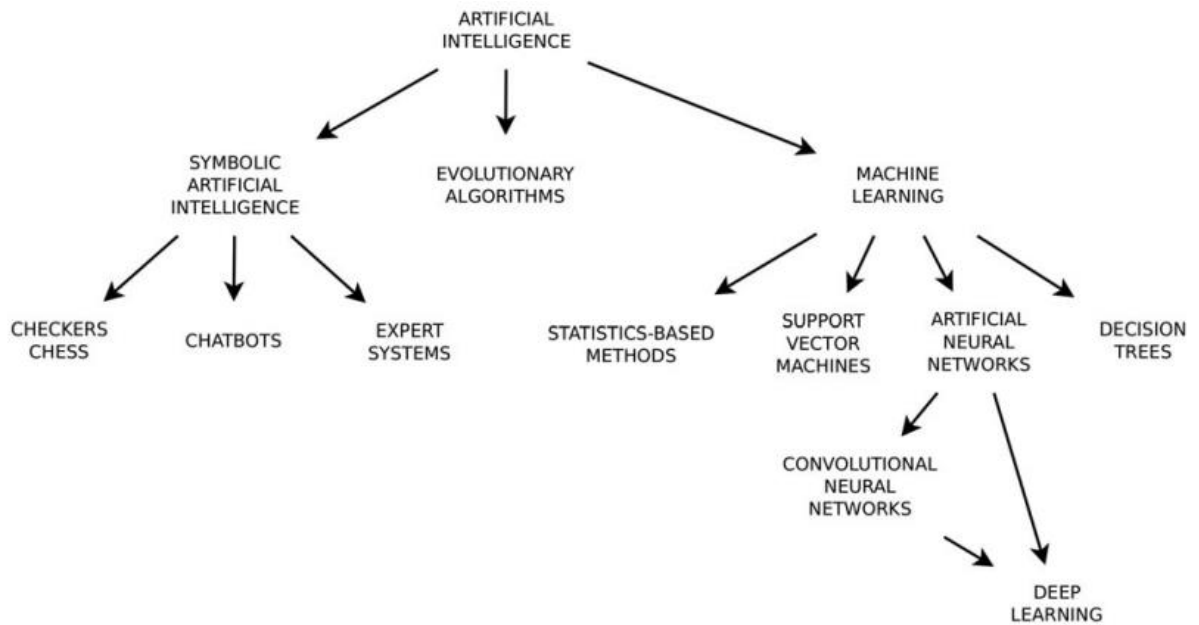


FIGURE 7. Schematic overview of the main branches of artificial intelligence and some of the machine learning methods (Galbusera et al. 2019).

### 3.1 How does machine learning work?

Machine learning basically operates on three fundamental elements that enable it to identify patterns and attributes, process input data, and produce the appropriate output to adjust the model parameters. These elements encompass the decision process, error function, and optimization process. (Jung 2023)

The primary focus in the decision-making process is making the best choices based on the relevant probabilities from the data (Bishop 2006a). The process frequently starts with the problem definition, followed by the choice of a suitable model, data training, and predictions or classifications based on the trained model. It is important to understand that the decision-making process is a workflow rather than a single entity. (Figure 8) (Bishop 2006a; Goodfellow et al. 2016b; Jung 2023)

The error function (often called the loss function or cost function) measures how closely the model's predictions match the true values. It gives an indication of the model's performance. This is often the discrepancy between the predicted values and the true values, and there are

various functions we can choose for the task that is required. During training, it is important to get this value as low as possible.(Goodfellow et al. 2016b; Jung 2023)

The optimization process in machine learning refers to the systematic procedure of refining model parameters to minimize the loss function. Finding parameter values that provide the least amount of error is the ultimate goal of optimization, which makes sure that the model's predictions are reliable and consistent. There are several optimization algorithms and strategies used in machine learning, such as gradient based optimization methods. The nature of the problem, the model's structure, the qualities of the data, and the type of machine learning task all influence the optimization strategy that is chosen. (Goodfellow et al. 2016b; Jung 2023; Shalev-Shwartz & Ben-David 2014b)

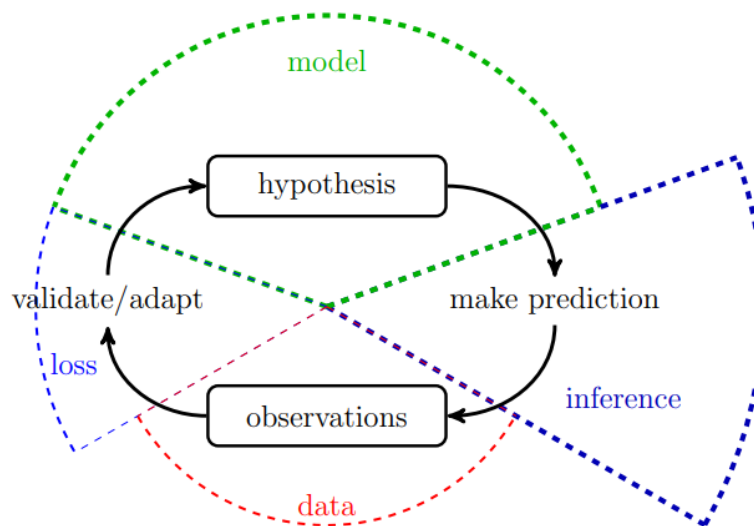


FIGURE 8. Basic framework of the decision-making process. Model, data and loss are the three components that provide a basic framework for the decision-making process. (Jung 2023)

### 3.2 Types of machine learning

Learning types differ based on the nature of the task. This chapter will cover the most relevant learning types utilized in human pose estimation models. These are supervised learning, unsupervised learning, semi-supervised learning, and transfer learning.



The algorithm is trained on a labelled dataset in supervised learning. A data point with a known value is called labelled data point. Such labelled data points can be sourced from human specialists who provide these data points with their respective label values. (Jung 2023) This implies that the appropriate output for each input is known. To enable the model to generate precise predictions on previously unobserved data, the objective is to learn a mapping from inputs to outputs. It is crucial to give accurate labels to the model to minimize prediction errors. (Alpaydin 2009; Galbusera et al. 2019) The most common supervised learning applications are classification and regression. A data point is classified by putting it in one of the predetermined categories or classes depending on its characteristics. The procedure requires a dataset that has been labelled, which means that each example has been matched with the appropriate category or label. Modelling the relationship between a dependent variable and one or more independent variables is what regression entails. (Bishop 2006; Galbusera et al. 2019)

In unsupervised learning, the algorithm is trained on an unlabelled dataset, and therefore the outcome is unknown. Without explicit instruction, the objective is to find patterns, structures, or representations in the data. (Galbusera et al. 2019; Jung 2023) The most common applications in unsupervised learning are clustering and density estimation. In clustering, the aim is to find similar groups that exhibit similar patterns in the data. Whereas in density estimation, the aim is to find the structure or shape of the data distribution. (Alpaydin 2009; Bishop 2006b; Galbusera et al. 2019)

Semi-supervised learning uses both labelled and unlabelled data for training. Along with unlabelled data, the algorithm receives some guiding details, although not for every instance. Here, the data can be split into two segments: points with known labels and points where the labels remain unidentified. (Chapelle et al. 2006) The objective is to learn a pattern so that examples belonging to the same class have representations that are similar (Goodfellow et al. 2016b).

Transfer learning is the process of retraining a model that has been trained to perform a certain task, such as classification, as part of a network to perform another task or pose estimation while retaining or "transferring" the previously learned features to the modified model. Transfer learning enables using smaller datasets since the modified model has already been trained with a large dataset. When compared to training from scratch, transfer learning typically increases the robustness of the model and the convergence speed (the rate of the algorithm's ability to

find a solution that is close enough to the best possible outcome) of model training. (Cronin 2021; Mathis et al. 2020)

### **3.3 Artificial neural networks**

Neural networks are modelled after the structure and operation of the human brain, particularly how its billions of interconnected neurons receive and send information. The "artificial neuron," a computational equivalent of biological neurons, is the fundamental unit of a neural network. Similar to how our brains learn from experiences, the network adapts the strength (or weights) of connections between these neurons, improving its predictions through iterative feedback. This design gives computers the ability to spot patterns, make choices, and learn from data. (Alpaydin 2010; Basheer & Hajmeer 2000; Bishop 2006; Galbusera et al. 2019; Krogh 2008; Shalev-Shwartz & Ben-David 2014a)

#### **3.3.1 Structure of neural networks**

In order to do complicated computations, numerous neurons are connected by communication links (weights) in neural networks (Shalev-Shwartz & Ben-David 2014a). When the data is provided to the system, it undergoes various processes and transformations through layers, eventually leading to a specific output or result. Typically, there's an input layer, which is where the data enters the system, one or more hidden layers, where the data is processed, and an output layer, which produces the final output. (Figure 9A) (Basheer & Hajmeer 2000; Galbusera et al. 2019; Shalev-Shwartz & Ben-David 2014a) This data flow process is called forward propagation (Goodfellow et al. 2016a).

The process starts with each artificial neuron taking its input and multiplying it by a weight. Weights adjust during training to improve the accuracy of the network's output. Then all these weighted inputs are summed together. Often, a bias (or threshold in simpler models) is added to the weighted sum (Basheer & Hajmeer 2000). Bias influences the value that is passed to the activation function, thereby affecting the activation of a neuron (Alpaydin 2010). After summing the weighted inputs and adding bias, the result is passed through an activation function. This function can be linear (a straight-line relationship) or nonlinear (a curved

relationship) (Galbusera et al. 2019). The purpose of this function is to introduce non-linearity to the model, allowing it to learn from errors and make complex decisions. (Figure 9B) (Basheer & Hajmeer 2000; Shalev-Shwartz & Ben-David 2014a) .

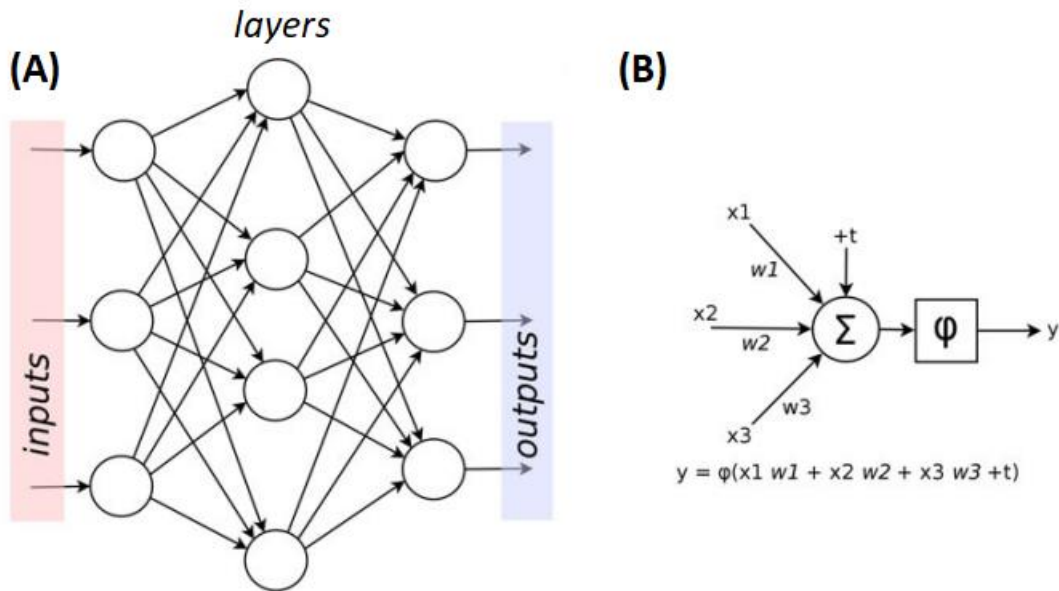


FIGURE 9. An artificial neural network (A) and an artificial neuron unit (B) are shown schematically. In each unit, the inputs (“x1,3”) are multiplied by weights (“w1,3”), a bias term is added (“+t”), and the result is passed through a linear or nonlinear activation function (“φ”). Modified from Galbusera et al. (2019).

### 3.3.2 Training neural networks

Upon completion of the forward propagation phase, the neural network generates an output, which is then compared to the desired target or true value to determine the prediction error (Goodfellow et al. 2016a). This error quantification provides a measure of how well the neural network's current weights and biases are performing (Aggarwal 2018). To adjust these weights and biases for better future predictions, the error is propagated backward through the network in a process termed "backward propagation" or simply "backpropagation" (Goodfellow et al. 2016a; Rumelhart et al. 1986).

Starting from the output layer, the error is used to determine how much each neuron in the following layers contributed to the overall error (Goodfellow et al. 2016a). The gradients of the

loss function with respect to the weights and biases are calculated using the chain rule of calculus (Aggarwal 2018). These gradients effectively capture the direction and magnitude by which adjustments should be made to the network's weights to reduce the overall error (Goodfellow et al. 2016a; Krogh 2008; Rumelhart et al. 1986).

These gradients are then used by an optimization algorithm, such as gradient descent, to iteratively update and modify the network's weights and biases. The aim is to reduce the prediction error of the network throughout future iterations of training (Bishop 2006; Krogh 2008). This continuous process of forward and backward propagation, followed by weight adjustments, forms the foundation of training most neural networks, refining them for enhanced performance on unseen data. (Bishop 2006; Goodfellow et al. 2016a; Krogh 2008)

### **3.4 Deep neural networks**

Deep neural networks are a subset of artificial neural networks and are primarily utilized in deep learning applications (Figure 7) (Galbusera et al. 2019). Deep neural networks add complexity and capacity by adding numerous layers between the input and output, enabling the hierarchical representation of data (Aggarwal 2018; Galbusera et al. 2019). These networks' depth, which represents the number of layers, enables them to extract complex characteristics at different levels of abstraction (Galbusera et al. 2019; Goodfellow et al. 2016a), making them especially effective for tasks like image and speech recognition (LeCun et al. 2015). Deep networks can have dozens or even hundreds of hidden layers, whereas standard neural networks could only have a few (Figure 10, Figure 9A) (Aggarwal 2018).

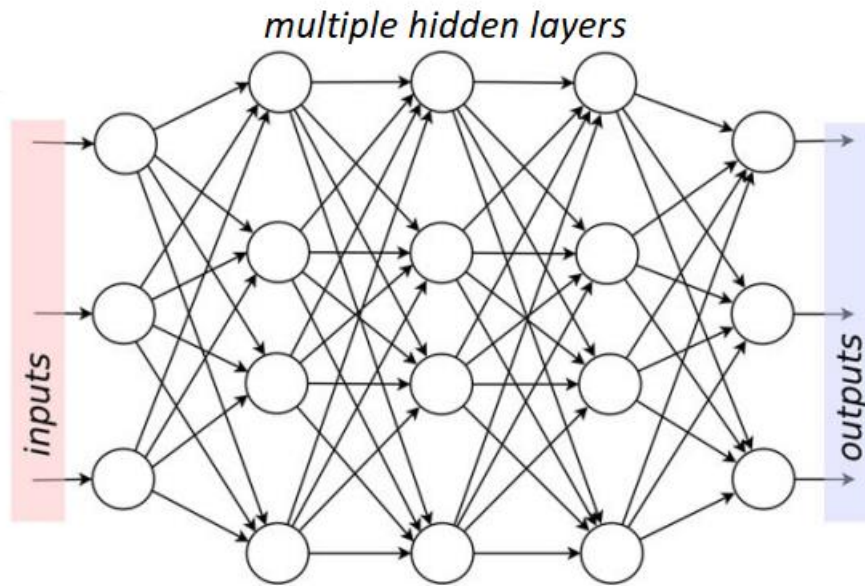


FIGURE 10. A deep neural network with multiple hidden layers. Modified from Galbusera et al. (2019).

Certain specialized architectures within the field of deep neural networks are very useful for tasks such as human pose estimation. One of the most well-known structures is a convolutional neural network (CNN, also called ConvNet), which is well-known for its capacity to analyze spatial patterns in images (LeCun et al. 2015). Convolutional layers in CNNs use filters that systematically scan and analyze overlapping portions of the input image as they "slide" or "convolve" over it. Simple edges to complex textures and features are all detected by this technique. (Figure 11) (Galbusera et al. 2019) The processed information is subsequently down-sampled using pooling layers in order to simplify the data and emphasize the most important aspects. This decreases the computational load while keeping important patterns (Galbusera et al. 2019; Scherer et al. 2010). Given that human poses in images are essentially spatial, CNNs and their variations have become a key component in many pose estimation methods (Bazarevsky et al. 2020; Cao et al. 2018; Fang et al. 2022; Mathis et al. 2018).

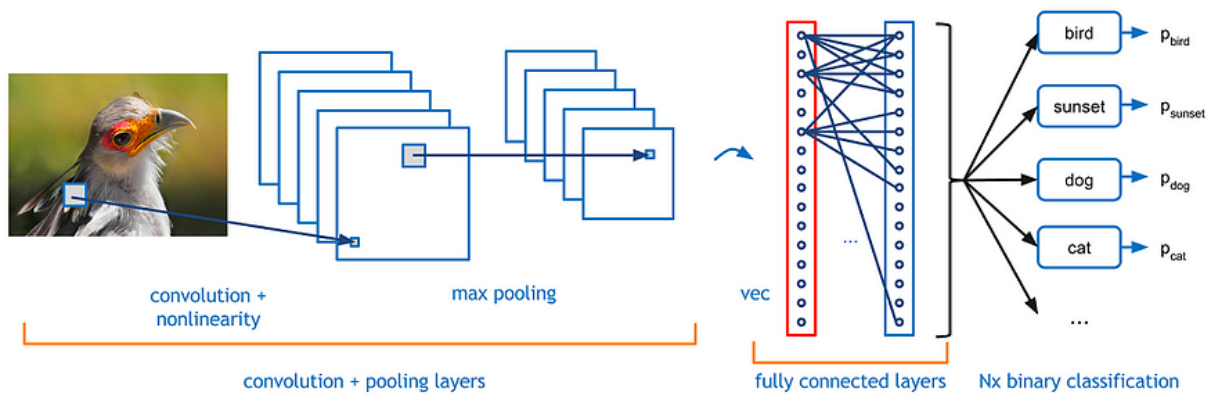


FIGURE 11. Demonstration of convolutional neural network. The filter convolves over the input image. The processed image is down sampled by the pooling layers and the fully connected layers take the features and use them to determine the final task specific output (in this case a classification problem). (Shahid 2019)

## **4. MOTION ANALYSIS**

Motion analysis is a scientific discipline dedicated to the study of movement in its various forms (Winter 2009). Its origins can be traced back to the pioneering work of Eadweard Muybridge in 1878, who used sequential photographs to capture the running motion of horses, which makes this study a significant milestone in biomechanics. Since then, motion analysis has evolved and provides valuable tools to understand and quantify motion. (Colyer et al. 2018; Richards et al. 2018)

Motion analysis has evolved to incorporate a wide range of tools and techniques, from simple devices like instrumented walkmats to force platforms, Inertial Measurement Units (IMUs), optical motion capture systems, and the recent emergence of markerless approaches. Each phase of development has expanded our capabilities in capturing, analyzing, and interpreting motion data. (Colyer et al. 2018)

With the introduction of markerless motion analysis, which is powered by depth sensor cameras and sophisticated artificial intelligence algorithms, motion analysis is currently set for further advancement. This invention has the potential to make motion analysis more approachable and non-invasive in a variety of fields. (Colyer et al. 2018; Wade et al. 2022)

This chapter will further delve into marker-based measurement systems, markerless measurement systems, their working principles, and their limitations.

### **4.1 Marker-based motion analysis**

Instrumented walkways and force plates have enabled the precise assessment of spatiotemporal features derived from ground reaction forces during walking (Hansen et al. 2002; Menz et al. 2004). However, monitoring the locations of several joints throughout trials is necessary for determining kinematic characteristics. To achieve this, various technical solutions have been employed to estimate sensor positions attached to specific anatomical landmarks. (Colyer et al. 2018) Among these solutions, optical motion tracking has emerged as the predominant method for accurately determining kinematic gait parameters, with the use of markers still being considered the standard practice. (Latorre et al. 2019; Robertsen et al. 2013)

These systems allow us to gain insights into various aspects of movement, such as joint kinematics (e.g., joint angles and joint range of motion), joint kinetics (e.g., forces and moments acting on joints), and muscle activity patterns (Robertsen et al. 2013; Wade et al. 2022). In order to determine the location and orientation of the body segments (6 degrees of freedom) in three dimensions, marker-based systems compute the translation of the limb in the sagittal frontal and transverse planes and the rotation in flexion/extension, abduction/adduction, and rotation about the longitudinal axis (Figure 12). This information is valuable for understanding factors like clinical gait analysis, sports performance, injury prevention, rehabilitation, and ergonomics (Colyer et al. 2018; Wade et al. 2022).

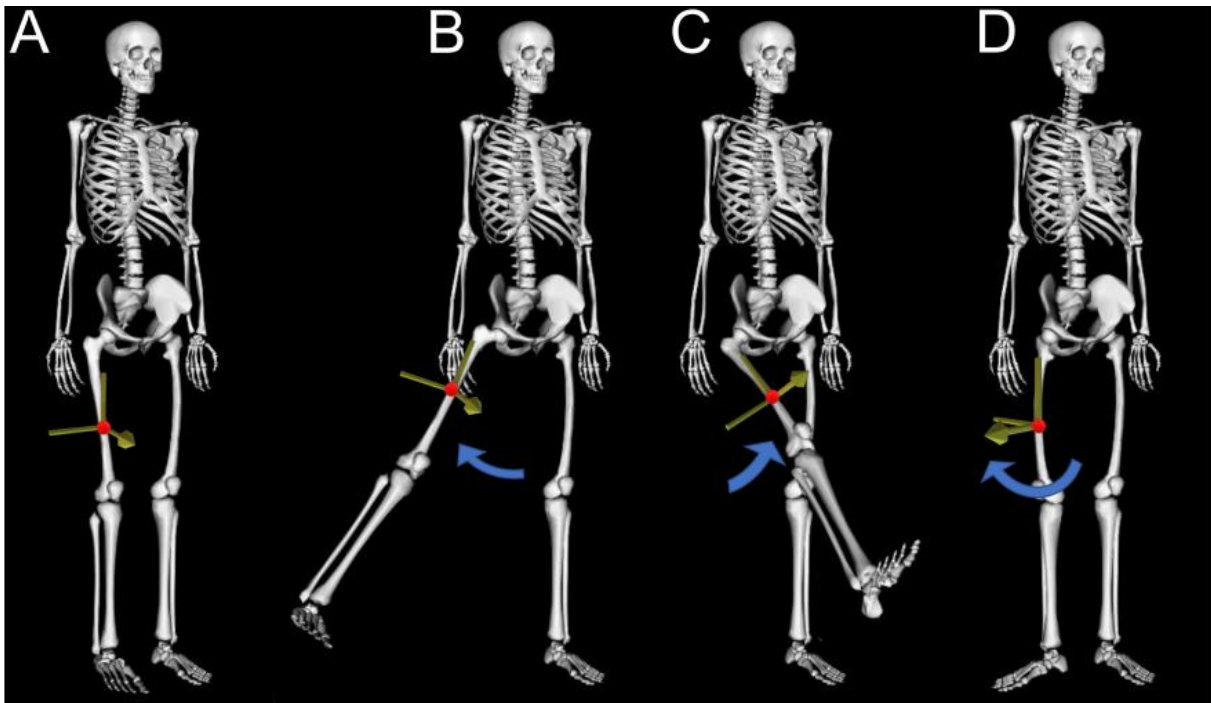


FIGURE 12. 6 degrees of freedom (6 DoF) demonstrated in thigh segment. Segment center of mass marked with red dot and rotation in three planes marked with blue arrow. (A) The reference standing posture, (B) adduction/abduction, (C) flexion/extension, (D) rotation about the longitudinal axis. (Wade et al. 2022)

Marker-based motion capture systems are preferred, especially in clinical gait analysis, because they offer high accuracy and reliability in capturing the detailed movement of specific body parts, despite their limitations. However, it is worth noting that the true gold standard for movement analysis is considered to be bi-planar video radiography, which involves using X-ray imaging to track skeletal movements (Kessler et al. 2019; Miranda et al. 2011). However,



even this method is not without its challenges and errors. Nonetheless, this approach exhibits translational errors in joint center positioning of 0.3 mm and rotational errors of  $0.44^\circ$ , as reported by Miranda et al. (2011). Furthermore, its high expenses, limited capture volume (suitable for single-joint analysis only), and radiation exposure make it impractical for use in clinical or sports-related applications. Therefore, marker-based systems are now the current standard method due to these limitations. (Wade et al. 2022)

Typically, a marker-based approach involves attaching markers to specific anatomical landmarks and utilizing multiple motion cameras to extract kinematic characteristics from the observed motion (Albuquerque et al. 2021). These markers are typically small, reflective, or non-reflective objects that can be attached to the objects or individuals whose movements need to be tracked. Markers can come in various forms, such as spherical reflective balls (commonly used), LED lights, or passive markers (Richards et al. 2018; Robertsen et al. 2013). Multiple cameras are strategically placed around the area where the motion capture is taking place. These infrared cameras are equipped with sensors to capture images of the markers from various angles (Colyer et al. 2018; Moro et al. 2022). The motion capture system is calibrated before any data is collected in order to determine the spatial relationship between the cameras and build a coordinate system for the data that is taken. Calibration is crucial to ensuring accurate measurements. The cameras capture the marker locations from different angles, allowing for three-dimensional reconstruction of the subject's movement. Specialized software for the utilized systems is then used to analyze and process the marker positions to calculate joint angles, segment velocities, accelerations, and other biomechanical properties. (Richards et al. 2018; Robertsen et al. 2013)

Marker-based motion capture systems offer high accuracy and precision in capturing movement data, but they also come with several limitations. One primary limitation is the precise marker placement. The markers on the skin don't always align with the actual three-dimensional positions of joints. There are numerous techniques for using markers positioned on anatomical landmarks to identify segment pose and/or find joint centers. However, these methods can give different results for motions that are not purely in the vertical or horizontal planes during repeated gait cycles. (Colyer et al. 2018; Ferrari et al. 2008) Additionally, there are inevitable variations in marker placement from day to day and between different technicians, which compromises the consistency of marker-based motion capture systems, especially when

measuring movements in the transverse plane.(Colyer et al. 2018; Growney et al. 1997; Tsushima et al. 2003)

Another main source of error in marker-based motion capture is the false assumption that underlying bone locations are well-reflected by surface markers. This misperception results in a phenomenon called soft tissue artifact, in which the markers alter in a way that is inconsistent with bone movement due to the movements of muscles, fat, and skin. (Camomilla et al. 2017; Peters et al. 2010) According to studies, this difference can cause a 5-7 mm inaccuracy for markers placed on softer surfaces, like the shank, and a 3-5 mm error for markers placed on tougher surfaces, such as the bony landmarks on the foot (Kessler et al. 2019). Although some of these errors can be reduced by data filtering techniques, the underlying problem of soft tissue movement cannot be resolved without directly affixing markers to the bones (intra-cortical bone pins), which is invasive (Wade et al. 2022).

Another limitation is marker occlusion, where markers can be temporarily hidden from the view of cameras, leading to gaps in the data that may require additional manual correction or interpolation (Richards et al. 2018). Furthermore, the process of attaching markers to the subject's body can be time-consuming, and the presence of markers can affect natural movement, potentially altering the results. High costs associated with equipment, calibration processes, and data processing are also another critical limitation. Additionally, lighting conditions and subject cooperation further contribute to potential inaccuracies. Moreover, these systems rely on a clear line of sight between markers and cameras, making tracking accuracy susceptible to obstructions or subjects moving out of view, which forces these measurements to be in a controlled environment which ultimately making these systems impractical to use in outdoor settings. (Colyer et al. 2018; Wade et al. 2022)

## **4.2 Markerless motion capture**

Markerless motion capture technologies are emerging as a solution to many of the challenges posed by traditional marker-based methods. Utilizing advanced algorithms in computer vision and image processing, these systems can track human movement without the need for physical markers, mitigating issues of marker placement, interference, and occlusion. They offer the advantage of tracking movements directly from video footage, making them less susceptible to line-of-sight issues and reducing equipment and setup complexity. Markerless systems have

gained popularity in recent years, particularly in applications where marker-based systems face challenges, such as capturing natural movements or movements in non-laboratory settings (Cronin et al. 2019; Pagnon et al. 2021; Palucci Vieira et al. 2022). While they have their own set of challenges, including the need for robust algorithms and computational resources, markerless systems represent a promising alternative for motion capture and biomechanical research.

#### **4.2.1. Earlier markerless motion capture methods**

Before the advent of sophisticated human pose estimation models, several simpler and more manual techniques were employed in markerless motion capture systems to analyze human movement.

One of the early methods was silhouette analysis. Silhouette analysis focused on the outline of a person as captured by a single camera. The system tracked the changes in this outline to infer movement. However, this method faced challenges, particularly in terms of depth perception and differentiating between body parts when they overlapped or were close to one another. (Colyer et al. 2018; Corazza et al. 2010)

A visual hull is a more comprehensive 3D representation that is produced by combining silhouettes from various camera angles. A volumetric model of the subject was created by intersecting the various silhouette views from several directions. (Laurentini 1994) Although this provided a better sense of depth and form, the visual hull was still an approximation. It was unable to capture the nuances of individual body parts or to discern between closely interacting limbs, often resulting in a somewhat rough and imprecise representation of the human body. While it was possible to achieve more detailed and accurate 3D reconstructions, it required more advanced and intensive computational processes. (Colyer et al. 2018; Corazza et al. 2010) Nevertheless, the visual hull technique has shown precise outcomes for automated markerless capture of body movement in a study conducted by Corazza et al. (2010), Mundermann et al. (2005) and in another study by Liu et al. (2013) facilitating kinematic analysis of multiple individuals.

Manual digitization is another early technique in markerless motion capture where operators manually mark body landmarks in video frames to analyze movement. This labor-intensive

process involves frame-by-frame analysis and is dependent on the operator's anatomical knowledge and consistency. While it provided a foundation for modern motion analysis, its time-consuming nature and subjectivity have led to its replacement by automated systems. (Colyer et al. 2018) Despite this, manual digitization remains a useful tool for motion capture technologies- especially for sports applications that typically occur outdoors- and is still used to validate markerless motion capture (Palucci Vieira et al. 2022; Takeda et al. 2021).

The arrival of depth sensing cameras marked a significant step forward. These cameras used the time-of-flight (ToF) principle or structured light to capture the depth information of a scene. These ToF cameras work by emitting a light pulse and measuring the time it takes for the light to reflect back from objects in the scene. This method provides a detailed in-depth map, allowing for an intricate 3D portrayal of the environment and the subjects within it. (Colyer et al. 2018) Systems like the Microsoft Kinect (no longer commercially available) popularized this technology for both recreational and research purposes (van Schaik & Dominici 2020). Nevertheless, even with their advanced capabilities, depth sensors had a limited field of view and were sensitive to environmental factors like lighting and background variability which forced these systems to be used in controlled settings (Clark et al. 2013; Wade et al. 2022).

These methods' limitations highlighted the need for a more sophisticated approach, particularly in accurately capturing complex motions in unconstrained environments. This led to the rise of deep learning-based human pose estimation (HPE) models.

#### **4.2.2. Human pose estimation**

Human pose estimation is a revolutionary advancement in markerless motion capture technology. Compared to the earlier methods, HPE models utilize convolutional neural networks (CNNs) and other machine learning architectures to accurately identify and track human anatomical keypoints from image and video data. There have been many open-source models such as OpenPose (Cao et al. 2018), AlphaPose (Fang et al. 2022), DeepLabCut (Mathis et al. 2018), BlazePose (Bazarevsky et al. 2020) and commercial models such as Theia (Theia Markerless Inc.), Captury (Captury), and Simi (Simi Reality Motion Systems GmbH) introduced over the years. These models are trained on large, diverse datasets of annotated images that encompass a wide range of human shapes, sizes, environments, postures, and activities (e.g., COCO Keypoint Detection Challenge, MPII Human Pose Dataset, and VGG

Pose Dataset). (Colyer et al. 2018; Mathis et al. 2020) They are increasingly being adopted for motion analysis in diverse settings, from sports biomechanics (Cronin et al. 2019; Evans et al. 2018; Pagnon et al. 2021; Palucci Vieira et al. 2022) to clinical assessments (Kidziński et al. 2020; Lonini et al. 2022; Sato et al. 2019; Takeda et al. 2021; Vafadar et al. 2021) , offering a less intrusive and more flexible approach compared to traditional methods.

HPE models involve complex processes to detect body keypoints in images or video sequences and use this information to infer body posture. Keypoint detection, pose inference, learning, and model architecture are the main components that comprise the overall structure of an HPE system.

HPE models detect keypoints in the human body, such as the elbows, wrists, knees, hips, shoulders, and ankles (the number of key points varies from model to model). This is typically achieved through CNNs, which have become the standard in these models due to their effectiveness in handling large image data. The input image passes through convolutional layers to extract the features and patterns associated with the anatomical landmarks. These patterns are then mapped onto a set of heatmaps, each corresponding to a specific keypoint. In these heatmaps, the pixel values represent the confidence that a particular keypoint is present at that location. Generally the higher the confidence value, the higher the accuracy of the tracked coordinates of the keypoint location. (Mathis et al. 2020)

Once the keypoints are identified, the HPE system forms the skeletal structure by connecting these points. This stage is crucial as it involves associating the correct keypoints to form the limbs and joints of the skeleton. Different models approach keypoint detection in individuals that are detected in various ways. For instance, OpenPose (Cao et al. 2018) and DeepLabCut (Mathis et al. 2018) utilize a bottom-up approach to multi-person pose estimation, detecting every instance of a particular body part while also connecting each body part by using a set of 2D vector fields (Part Affinity Fields) that encode the location and orientation of limbs in the image. (Figure 13A) Detectron (Facebook AI Research), AlphaPose (Fang et al. 2022), and BlazePose (Bazarevsky et al. 2020) takes a top-down method to multi-person pose estimation, first detecting individual people inside an image (bounding box detector), then using pose estimation to detect critical points of each detected person (Figure 13B) (Mehdizadeh et al. 2021; Needham et al. 2021). However, the top-down approach fails if the person has not been detected (Mathis et al., 2020). This step is key to assembling the detected keypoints into a

coherent and accurate representation of the human pose, providing the basis for further analysis or application.

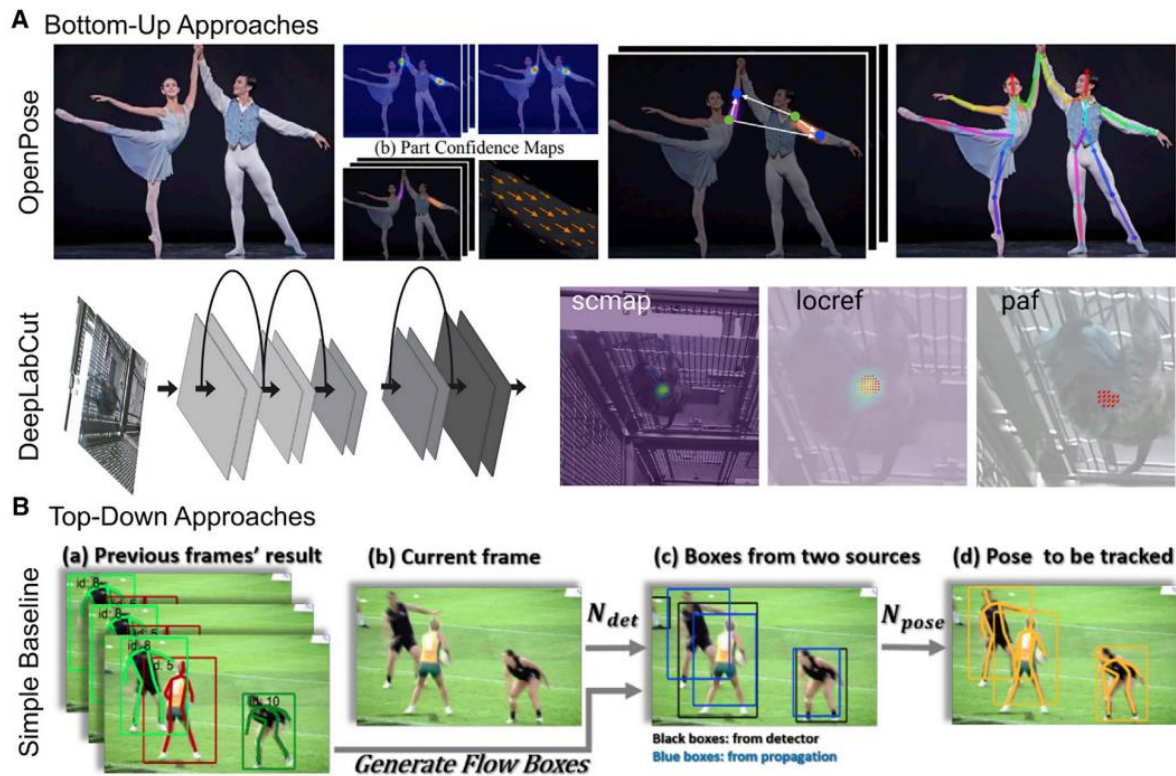


FIGURE 13. Pose estimation approaches. A) Bottom-up approaches utilized by OpenPose and DeepLabCut. In bottom-up approach, each body part and limbs in the image is first identified by the part confidence maps, and then each keypoint is put together to form the posture by using the part affinity fields. B) Top-down approaches first detect the individuals with the bounding box detectors then predict the posture. (Mathis et al. 2020)

Learning in HPE is another crucial step. With a supervised learning approach, the network can detect learned features in new images (Figure 14). Transfer learning is also often used, where a model pre-trained on a large dataset is fine-tuned with specific data to achieve higher accuracy. As mentioned before, these models are trained on large, annotated datasets where each image includes the labeled location of keypoints. During training, the model adjusts its parameters to minimize the difference between its predictions and the provided annotations. However, this brings the main limitation of these models being prone to subjectivity because of their dependence on the labeller's anatomical knowledge. Most of these large datasets were labelled by annotators without anatomical knowledge and for other purposes than biomechanics, which

caused some of the models to incorrectly identify certain keypoints. However, even with an annotator with biomechanical and anatomical knowledge, subjectivity remains a problem since the ground truth is often not known. One solution suggested by Cronin (2021) is that the data should be labelled by multiple people, and after that, agreement between their estimations should be verified using a predetermined reliability criterion. Another critical point is the variability of the datasets. This means that the datasets should include images from different angles, scales, postures, lighting, clothing conditions, and so on. Extending the training set through specific changes (e.g., rotating or scaling image size) is known as data augmentation. Models become more accurate and invariant to rotations, scale changes, and distortions based on the selected corruptions (with less training data). Data augmentation ultimately aids in enhancing resistance to noise, such as motion blur and image artifacts. (Cronin 2021; Mathis et al. 2020)

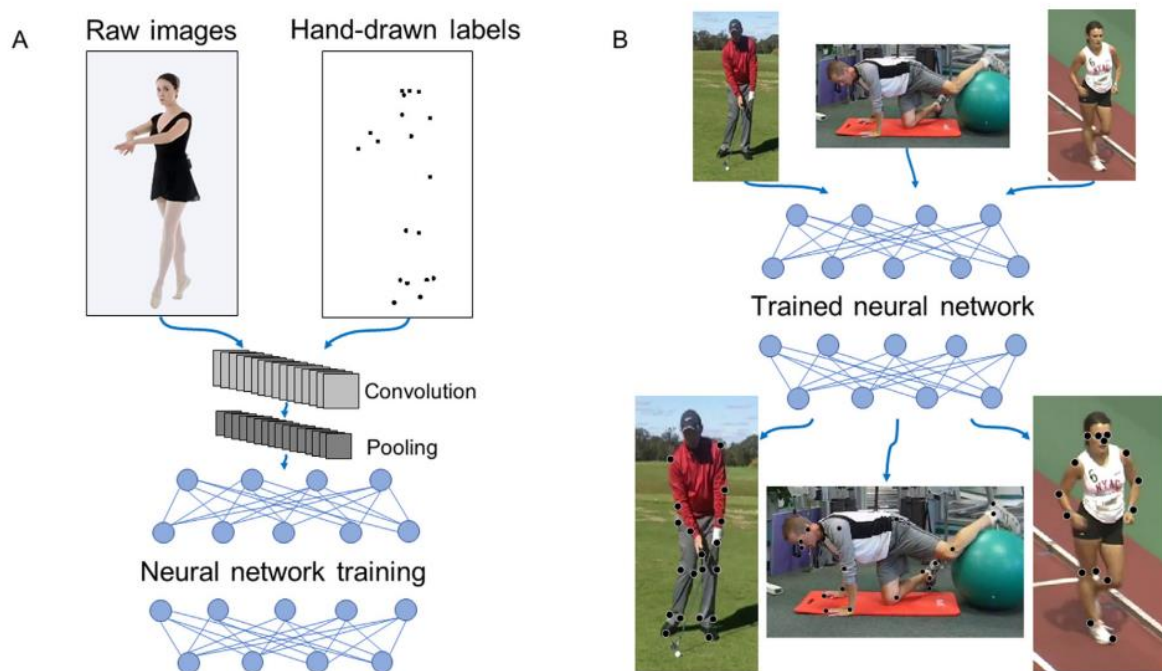


FIGURE 14. A) Training the human pose estimation model with supervised learning approach by utilizing CNN and feeding the labelled images to the network. B) After training the model is able to detect the learned anatomical keypoints in new images. (Cronin 2021)

These key components enable HPE systems to function and improve. The keypoint detection provide the coordinates of the keypoints, the skeletal structure is added by pose inference, learning procedures optimize the model's parameters for precise prediction, and complicated

model networks and architectures allow for the effective processing of the visual data. Models are becoming more robust to detect the various types of inputs of body types, sizes, and occlusions as HPE technology develops, expanding their use to increasingly complex and dynamic real-world situations.

As mentioned before HPE models utilize artificial intelligence to handle the complex visual data that a typical video camera collects, while marker-based motion capture primarily relies on hardware (markers and infrared cameras) to extract the anatomic keypoints and the poses. The main difference between these approaches is that while marker-based systems use three keypoints to calculate 6 degrees of freedom (6DoF), the majority of the existing pose estimation models have been trained to extract two points on each segment (the proximal and distal joint center locations). While information about the sagittal and coronal planes can be obtained from two keypoints, rotation about the segment's longitudinal axis requires the third keypoint. For this reason, most of the introduced HPE models are limited to 5DoF. Although there are certain methods to overcome this limitation, it requires additional processing with musculoskeletal modelling or training the model with a relabelled data (Needham et al. 2021). (Wade et al. 2022)

Although HPE models have the potential to change biomechanics by allowing researchers to collect data outside of the laboratory, there are still concerns about the accuracy and convenience of use of these methods (Drazan et al. 2021). Human pose estimation models must be validated before being used in research settings. Numerous studies have examined the validity and reliability of the markerless approach in comparison to the marker-based system in settings with one or more cameras and a range of movement tasks and environments.

The most studied open-source HPE model in the literature was OpenPose (Åberg et al. 2021; Kidziński et al. 2020; Nakano et al. 2020; Ota et al. 2020; Pagnon et al. 2021; Palucci Vieira et al. 2022; Sato et al. 2019; Slembrouck et al. 2020; Takeda et al. 2021; Zago et al. 2020). Several studies showed differences of 3D average joint center location between 10-50 mm in comparison of markerless and marker-based systems (Nakano et al. 2020; Slembrouck et al. 2020; Zago et al. 2020). Temporospatial differences showed 15 mm of difference compared to marker-based systems (Zago et al. 2020). Nakano et al. (2020) observed better results in comparison to marker-based systems from slower movements in average joint center differences of 10-30 mm in walking, where in fast movements like jumping and throwing they



observed a difference of 20-40 mm. Other studies also reported that slow video frames may increase the error rate in faster movements as well (Slembrouck et al. 2020; Zago et al. 2020).

Three studies have compared DeepLabCut to marker-based systems in 2D sagittal view with their own manually labelled datasets in countermovement jump (Drazan et al. 2021), underwater running (Cronin et al. 2019), and walking of stroke patients (Moro et al. 2020). They have reported joint center location differences of 10-20 mm compared to marker-based systems (Drazan et al. 2021, Moro et al. 2020) and manual labelling (Cronin et al. 2019). However, they found no significant difference for temporospatial measures and joint angles during walking, countermovement jump, and underwater running. (Cronin et al. 2019; Drazan et al. 2021; Moro et al. 2020) Cronin et al. (2019) also reported that 300-400 images were sufficient to train the model to track joint center locations with a similar accuracy to the human labeller (mean difference < 3 pixels, around 1 cm). These studies showed very promising results for 2D applications from sports to clinical settings.

Needham et al. (2021) performed a comparison study between OpenPose, DeepLabCut (standard version) and AlphaPose in walking, running, and jumping with using marker-based system as the reference. AlphaPose and OpenPose showed the smallest 3D lower body joint center location differences of 16-34 mm in walking, 23-48 mm in running and 14-36 mm in jumping compared to the marker-based system. They also have reported that the largest differences were at the hip joint for all three models. According to their paper, this is probably the result of inaccurately labelled open-access datasets, with the hip joint being the worst because it is extremely difficult to correctly identify without physical palpation which ultimately leads to the main limitation of these open-source models. Another study (Stenum et al. 2021) reported similar issue for the hip joint. They have interpreted this as a result of different tracking and modelling logics of the two systems. For marker-based systems marker placement relies on physical palpation of bony landmarks, whereas Open Pose relies on visibly labelled generalized key points (e.g., "ankle," "knee" vs. lateral malleolus, femoral epi condyle).

Some of the studies also have tested these open-source HPE models in clinical settings. Moro et al. (2020) was able to detect significant differences in affected and unaffected sides of stroke patients by utilizing DeepLabCut. Moreover, head and trunk segment angles in young children with cerebral palsy were evaluated in another study in 2D, allowing clinical procedures to be automated to assess head and spine posture (Cunningham et al. 2019). Baldewijns et al. (2016)

evaluated walking speed using a camera discreetly captured in patients' homes, illustrating how markerless approaches could allow for ongoing patient monitoring during their daily lives. Martinez et al. (2018) examined walking cadence and automated the calculation of an anomaly score for patients with Parkinson's disease using OpenPose with the aim of giving clinicians an objective, broad picture of the patients' disease progression. Shin et al. 2021 examined the temporospatial measures of patients with Parkinson's disease in frontal view. They were able to identify slight gait abnormalities that the clinicians were not able to detect and showed high correlations between subjective clinical gait tests.

Finally, the most important example for the clinical application is the study conducted by Kidziński et al. (2020). Kidziński et al. (2020) analyzed around 1800 images of cerebral palsy patients' gait over the years. They have extracted 2D joint centers from these images with OpenPose and used the trajectories as an input for a neural network that predicted clinically relevant parameters (e.g., walking speed, cadence, joint angles). One limitation of this study was that they were not able to compare these outcomes with marker-based systems due to the data collection methods. Nevertheless, this study showed the great potential of the markerless systems in clinical settings by providing a gait report from the outcome measures that was generated from a deep learning based markerless method. Moreover, the studies by Kidziński et al. (2020) and Shin et al. (2021) show the importance of a dataset specifically created for clinical purposes which ultimately might overcome the limitations regarding the datasets that have not been specifically designed for biomechanical purposes.

## 5 PURPOSE OF THE STUDY

Compared to marker-based techniques, markerless motion capture offers the opportunity to do movement analysis noninvasively without using markers, with less time spent gathering data and enabling the measurements to be conducted outside the laboratory. The accuracy of joint center locations and joint angles is still lacking, even though markerless temporospatial measures seem to be somewhat similar to marker-based motion capture. Temporospatial measures may be accurate enough for practical uses, but joint center localization and angle identification still require a great deal of work, and more studies need to be conducted (Wade et al. 2022).

As mentioned before, some studies have tested deep learning-based markerless methods in clinical settings. However, most of these studies (Baldewijns et al. 2016; Cunningham et al. 2019; Martinez et al. 2018; Moro et al. 2020; Shin et al. 2021) did not assess joint angle measures. Only Kidziński et al. (2020) was able to assess this parameter though they were not able to validate their method. Thus, we have decided to test the feasibility of deep learning based markerless motion capture in clinical settings by assessing the joint angles of CP patients and their TD peers. The reason we chose this particular population was that CP patients exhibit neuromotor abnormalities, and gait analysis is crucial for precisely assessing and tailoring the treatment accordingly. Often, marker-based motion capture is utilized in these settings; however, it can be very time-consuming and tiring for the patients. Markerless motion capture can overcome these issues and bring flexibility to clinicians since not every institution has these expensive marker-based systems. Moreover, most of the previous studies utilized OpenPose (Cao et al. 2018), DeepLabCut (Mathis et al. 2018) and AlphaPose (Fang et al. 2022). In addition, previous studies (Bittner et al. 2022; Mroz et al. 2021) indicated that the BlazePose model is lightweight and performs faster during runtime which offers more promise for real-time use. For this reason, we have decided to utilize BlazePose (Bazarevsky et al. 2020) in our study and examine the accuracy of this model.

Thus, the main purpose of this study is to explore the feasibility of deep learning-based markerless motion capture in clinical settings by assessing the hip, knee, and ankle joint flexion angles of CP patients and their TD peers. The research questions of this study are;

- 1) How accurately are joint angles tracked in CP patients compared to marker-based systems?

Hypothesis 1: There is only one study in the literature that tested markerless motion capture in CP patients (Kidziński et al. 2020). However, this study utilized OpenPose to extract the keypoint trajectories and included another neural network to predict the kinematic outcomes, which is not comparable to our study. We expect BlazePose to perform less accurately with this population since it is very challenging to measure CP gait parameters even with marker-based systems. Additionally, as these models have been trained on datasets that have not been annotated for biomechanical purposes, it may be challenging to extract clinically meaningful and accurate information (Seethapathi et al. 2019).

- 2) How accurately are joint angles tracked in TD participants compared to marker-based systems?

Hypothesis 2: Previous studies that examined OpenPose reported significant differences in joint center localization compared to the marker-based systems (Nakano et al. 2020; Slembrouck et al. 2020; Zago et al. 2020). However, studies that assessed the joint angles with DeepLabCut reported no significant differences (Cronin et al. 2019; Drazan et al. 2021; Moro et al. 2020, 2022). It is important to note that each one of these models in the literature may perform differently (Needham et al. 2021) depending on factors such as the movement being studied. In addition, Moro et al. (2022) highlighted the fact that the tracking limitations of these models can be overcome with additional processing. For this reason, we expect to find qualitatively similar results with BlazePose as reported earlier with other models.

## 6 METHODS

In this study, we compared the performance of BlazePose (Bazarevsky et al. 2020) to Vicon motion capture system (Oxford Metrics, Inc., Oxford, UK) in CP patients and TD peers. We used GoPro (GoPro Inc., San Mateo, CA, US) cameras to collect the markerless data and synchronized with Vicon camera system for simultaneous acquisition. We further extracted the keypoint trajectories by feeding the images collected with GoPros as an input to the BlazePose algorithm in Python 3.11 (Python Software Foundation). BlazePose provides 33 keypoints annotated on the subject's body in 3D coordinates as an output for each camera view (Figure 15). While the x and y coordinates are directly from 2D image plane, the z-coordinate is an estimate of the depth, inferred by the model from the 2D image (Grishchenko et al., 2022). For this reason, we applied 3D reconstruction techniques to calculate more accurate 3D coordinates (see chapter 6.4.1). To obtain the keypoints faster, a CSC (Finnish IT Center for Science Ltd.) Puhti supercomputer was used. We did further data processing in Matlab R2020b (Mathworks Inc., US) and calculation of joint angles in OpenSim 4.3 (Simbios, US). We also performed the joint angle calculation for Vicon data in OpenSim to eliminate the methodological difference. We used the SPM1D (Pataky, 2012) tool for the statistical analysis.

This chapter will further explain the participant information, the equipment that was used, the experimental protocol, the data analysis steps, and finally the statistical analysis.

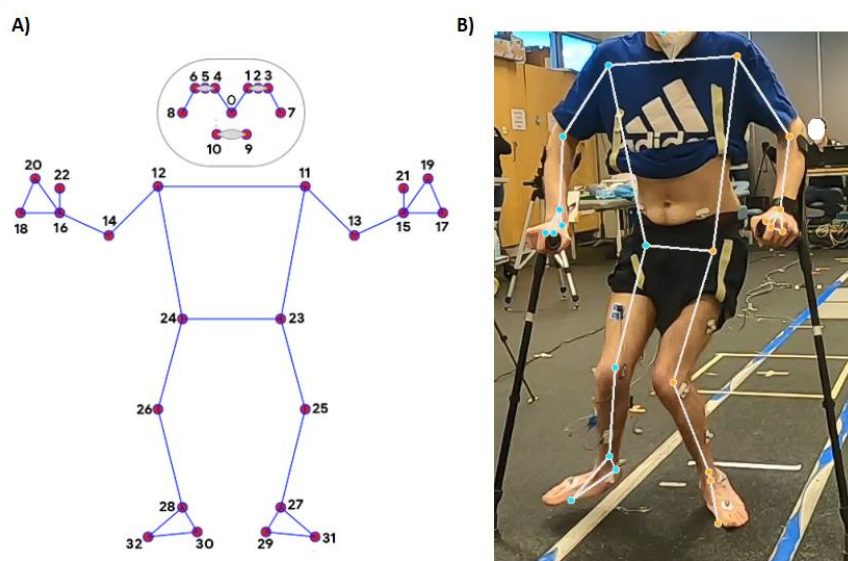


FIGURE 15. A) 33 BlazePose keypoint topology (Bazarevsky et al. 2020). B) BlazePose output with detected keypoints on the CP participant from the frontal view.

## 6.1 Participants

For this study, twelve CP patients (nine males, three females,  $17.2\pm 4.2$  years; Gross Motor Function Classification System: level I = 8, level III = 4) and twelve TD peers (nine males, three females,  $16.8\pm 4.7$  years) volunteered. However, due to the extensive amount of time required for the data processing, we were only able to include the data of six CP patients (five males, one female,  $17.5\pm 5.5$  years; GMCFS: level I = 3, level III = 6) and six TD peers.

Participants were recruited as part of the Integrating sensorimotor disorders with ankle hyper-resistance in cerebral palsy project (Cenni et al. 2023) (MSCA-IF, I-SENS). The acquisitions took place at the Neuromuscular Research Center, University of Jyväskylä. The Central Finland Hospital District's Research Ethics Committee approved the study (Dnro 8U/2017, amended 2021), and participants and the legal guardians of those under the age of 18 provided written informed consent.

## 6.2 Experimental setup and equipments

We used 11 Vicon motion capture cameras (Vero v2.2) for the marker-based setup (sampling frequency 200 Hz); 20 reflective spherical markers were placed on the participants body following the Vicon lower-limb plug-in-gait model. Ethernet cables were used to link the cameras to a PoE switch box (D-Link, DGS-1026MP, 26-port), and another Ethernet cable was used to connect the switch box to a laptop. Vicon Nexus 2.10.2 data capture software was used for the acquisition.

For the markerless setup, we used 3 GoPro (Hero 8, GoPro Inc., San Mateo, CA, US) cameras (sampling frequency 60 Hz, standard linear settings, with 1920x1080 pixel resolution). We mounted the GoPro cameras on tripods at hip level. The first camera was placed in frontal view, angled 45° from the corner of the gait path, and the second camera was placed in posterior view, angled 45° from the other corner of the gait path, to visualize the largest possible area. (Figure 16) The positions of the cameras were marked on the floor to ensure the camera locations were the same for each measurement. The GoPro Smart Remote was used to start and stop the recording of the cameras simultaneously.

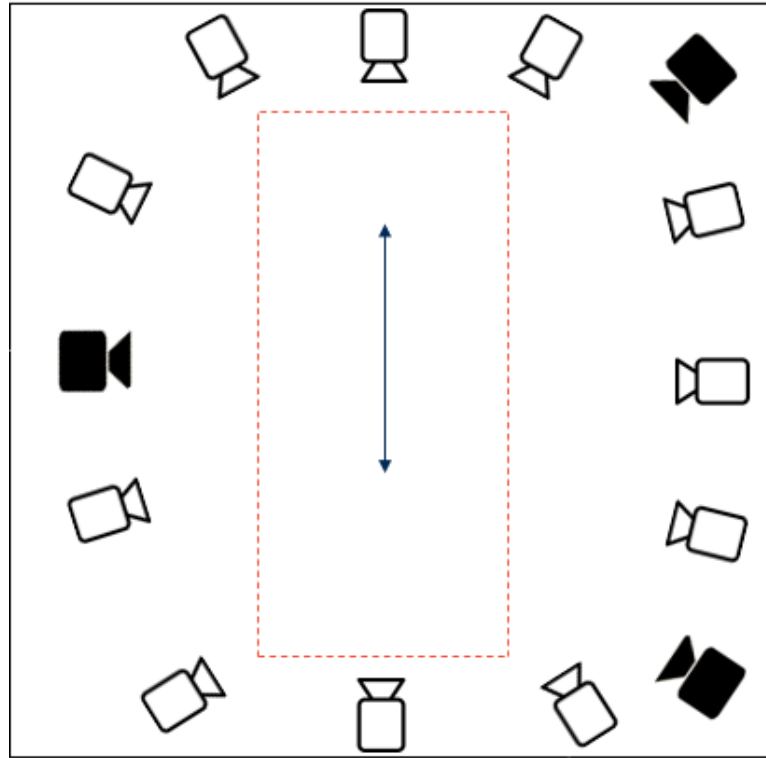


FIGURE 16. Experimental setup. The white cameras represent the Vicon camera system. The black cameras represent the GoPro cameras.

LED boxes were used to synchronize the GoPro and Vicon systems. The LED boxes were placed in front of each camera and connected to each other with BNC connectors. To trigger the light, we connected the LEDs to a voltage level shifter and from there to the Vicon Control Box (Lock+, for synchronizing third party devices). This setup allowed us to mark the trial beginning by lighting the LEDs when the acquisition started and turned the LED lights off when the acquisition stopped simultaneously with the Vicon system.

Before each measurement session, the Vicon camera system was calibrated. Before the calibration, zoom and focus were adjusted. Any undesired marker detections that occurred from reflections or other cameras were masked in Nexus. A T-shaped wand with five markers attached was used to calibrate the system by waving the wand in the capture volume to get at least 1000 frames in each camera. The world error reported in Nexus was checked at the end of the calibration to ensure the quality of the capture. If the error value exceeded 1, the calibration was repeated. The origin was set in Nexus by placing the wand on the side of the force plate in the gait path vertically. The volume calculation for the markerless system was done separately; volume calculation and calibration steps for the markerless method are outlined in 6.4.1.

### **6.3 Experimental protocol**

Anthropometric measures (height, body mass, ankle and knee width, leg length) were taken upon the arrival of the participants. These measures were required for the biomechanical model of the participant in Nexus. Then, 20 reflective spherical markers were attached to the patient's body following the lower-limb plug-in-gait model, including additional markers on the head of the fibula, tibial tuberosity, and medial malleolus to define a local shank reference frame (Cenni et al. 2023). Medial malleolus and tibial tuberosity markers were removed after the static acquisition. The participants were asked to wear shorts, and all the clothing was secured with tape to eliminate any marker occlusion due to the clothing. The participants were barefoot, and the markers were attached to the skin and secured with additional tape to ensure the markers did not detach during the acquisition. However, the markers on the anterior-superior iliac spine were attached on top of the participant's clothing.

After the preparation of the markers, the participants walked along an 8-meter gait path at a self-selected speed to familiarize themselves with the environment and to ensure they walked as naturally as possible during the acquisitions. After the familiarization, the measurements started with two static acquisitions by standing still on the force platform to create the biomechanical model of the participant. After static acquisition, a minimum of six trials were collected. The participant started to walk from the beginning of the gait path to the end and stopped. The following trial started from the end of the path and stopped at the beginning of the path (i.e., data were collected in both directions). Since both systems were synchronized, the acquisition happened simultaneously.

### **6.4 Data analysis for markerless data**

Analysis of the markerless data consisted of eight stages in total, including the statistical analysis (Figure 16). It started with capture volume calculation and digitizing the object coordinates from the volume capture (acquisition environment). After extracting the 2D keypoint coordinates from the BlazePose algorithm by feeding in the GoPro images, we synchronized all the coordinates from the three camera views. We filtered the trajectories, labeled the anatomical landmarks, and exported them into OpenSim. We then created the skeletal model in OpenSim and ran the inverse kinematics (IK) for the joint angle calculation. Another filter was applied to the joint angles to smooth out the noise. After that, we overlaid



the curves and time normalized both markerless and Vicon data. Finally, we performed the statistical analysis with the SPM1D Matlab package. All these steps were repeated for each subject. (Figure 17)

This chapter further delves into the details of the calibration, 3D reconstruction, and filtering keypoint trajectories for markerless data. Then, in the following chapters, the skeletal modeling, joint angle calculation, overlaying of the curves, and time normalization stages are explained for both Vicon and markerless data.



FIGURE 17. The overview of the data analysis workflow.

### 6.4.1 Calibration

We performed a lab volume calculation to transfer the capture volume (laboratory environment) dimensions into the virtual environment (in other words, 3D world coordinates to 2D image coordinates) and to set the origin. This method is known as the direct linear transformation (DLT) (Chen et al. 1994; Hatze 1988). For this calculation, we first marked 16 fixed points in the capture environment that we knew would not change with each measurement. We then measured the coordinates of these marked points with a tacheometer (a mapping device) (Topcon R500, Leica, Germany) (Figure 18). After obtaining the coordinates of the capture volume, we digitized the points in Matlab for each camera view by manually marking the known points in the images. To compute the DLT coefficients (which are used to transform 2D image coordinates into 3D world coordinates, allowing mapping of the camera's view space to the real-world coordinate system), we then backprojected the same points (measured with a tacheometer) to overlap the manually marked points (Figure 19). Calibration was successful if the points overlapped. This step was done for each subject to eliminate any calibration errors.



FIGURE 18. The coordinates of the marked points in the capture volume were measured with a tacheometer for the laboratory volume calculation.

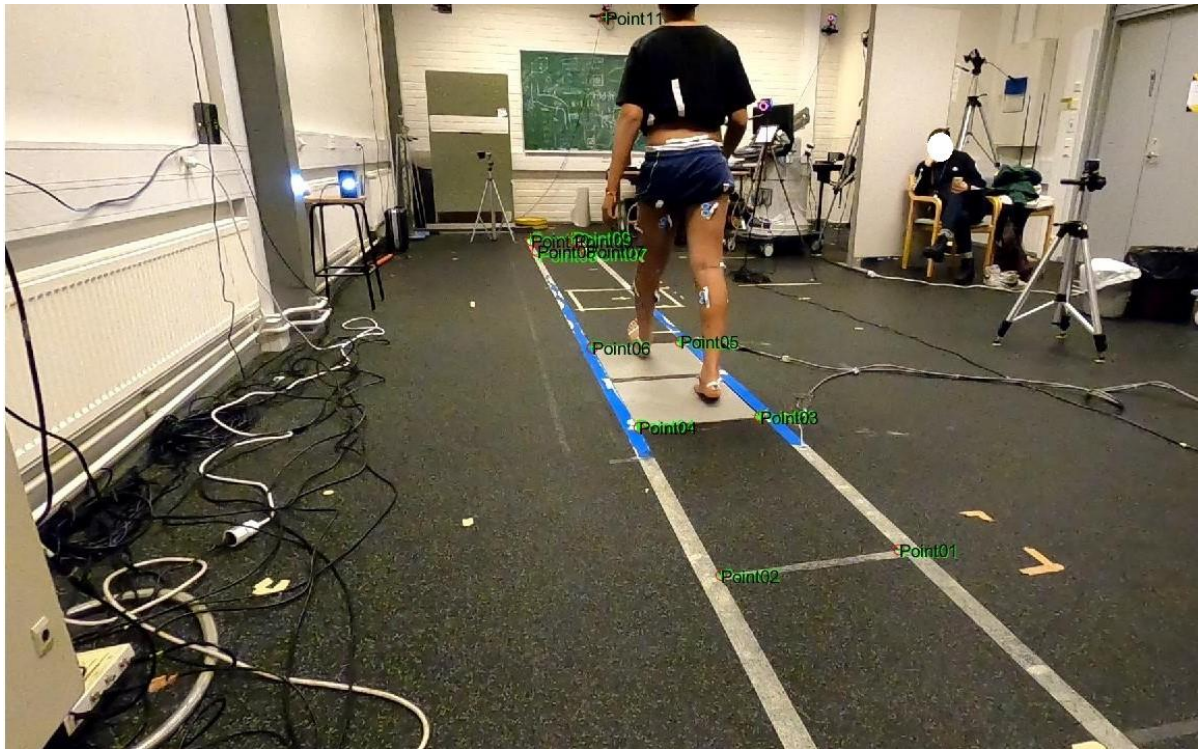


FIGURE 19. The backprojection of the points to compute the DLT coefficients from view of camera1. The calibration is successful if the points are matched. The green shade is the backprojected points and the points that are manually clicked are in black.

#### 6.4.2 3D Reconstruction

This step started by manually determining the frames where the LED lights went on and off in each camera view. We then used the frame information and the keypoint coordinates obtained from BlazePose for each camera and aligned these 2D image coordinates across cameras, adjusted them for synchronization, and then computed 3D coordinates using the DLT coefficients in Matlab. The Matlab code finds the overlapping time frames where data from all cameras is available and synchronizes these frames according to the provided synchronization information. It also aligns the DLT coefficients with the respective cameras. For each synchronized time point, the code combines the 2D coordinates from all three cameras and uses the DLT coefficients to calculate the 3D coordinates. This step involves selecting the closest indices for each camera based on the target timestamp and computing the 3D coordinates if the points are visible from at least two cameras. The final data is saved to an Excel file, and a visualization of the 3D points is created and saved as a video within a loop to process all trials (Figure 20).

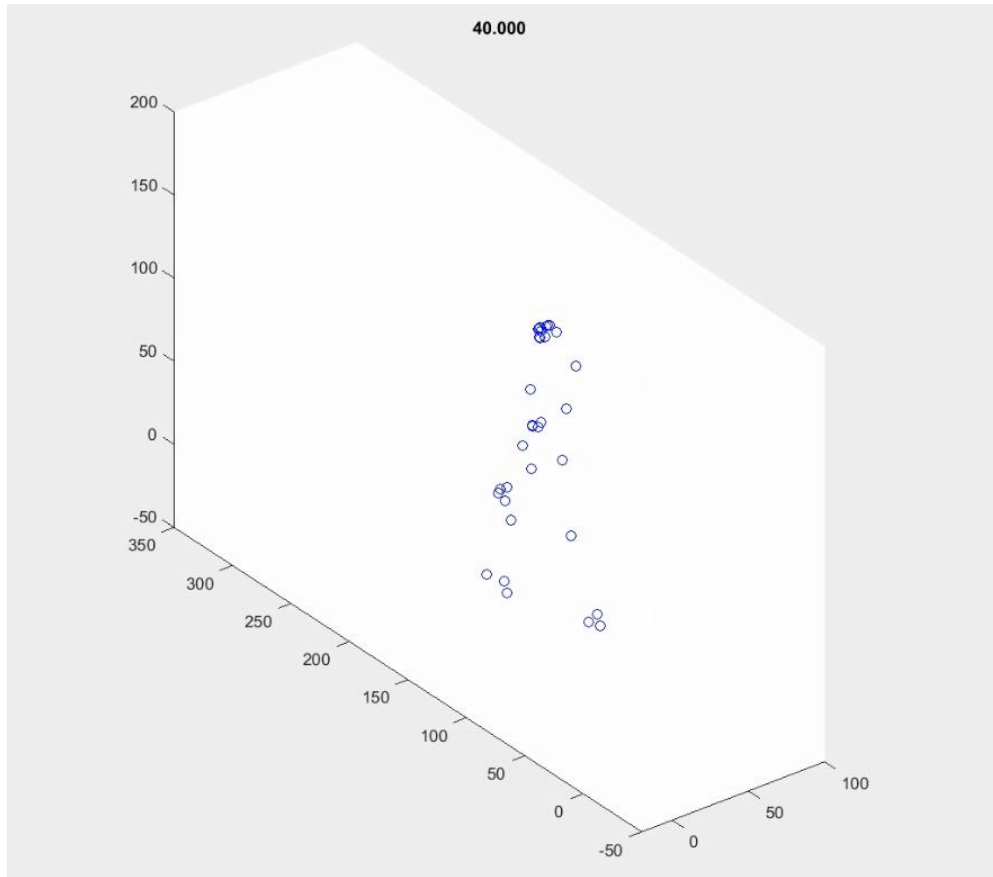


FIGURE 20. The visualization of the 3D reconstruction.

### 6.4.3 Filtering the keypoint trajectories

We made a filter selection to choose the most suitable filter for our data to smooth out the noise in keypoint trajectories without losing information from the data. Due to the high noise level of the keypoint trajectories in our data, we first chose a moving average filter with a window size of 8. The moving average filter is well-known for being simple to use and efficient at cutting down on random noise (Crenna et al. 2021). But even with lower window sizes, it was clear from application that this strategy severely smoothed the data, resulting in a substantial loss of crucial signal features. We then switched to a median filter due to its robustness in handling impulsive noise (Ohki et al. 1995), which is a common issue in motion capture data. We chose a window size of 5 to balance noise reduction with signal detail preservation. However, we still observed a loss in critical signal details, which finally led us to apply the Butterworth filter, known for its smooth frequency response and minimal signal distortion and commonly used for biomechanical signals (Crenna et al. 2021). We applied a 4th-order lowpass Butterworth filter

with a cutoff frequency of 5 (estimated with the residual analysis) to the keypoint trajectories. This filter was the best fit for the characteristics of our data. (See figure 21 for the comparison.)

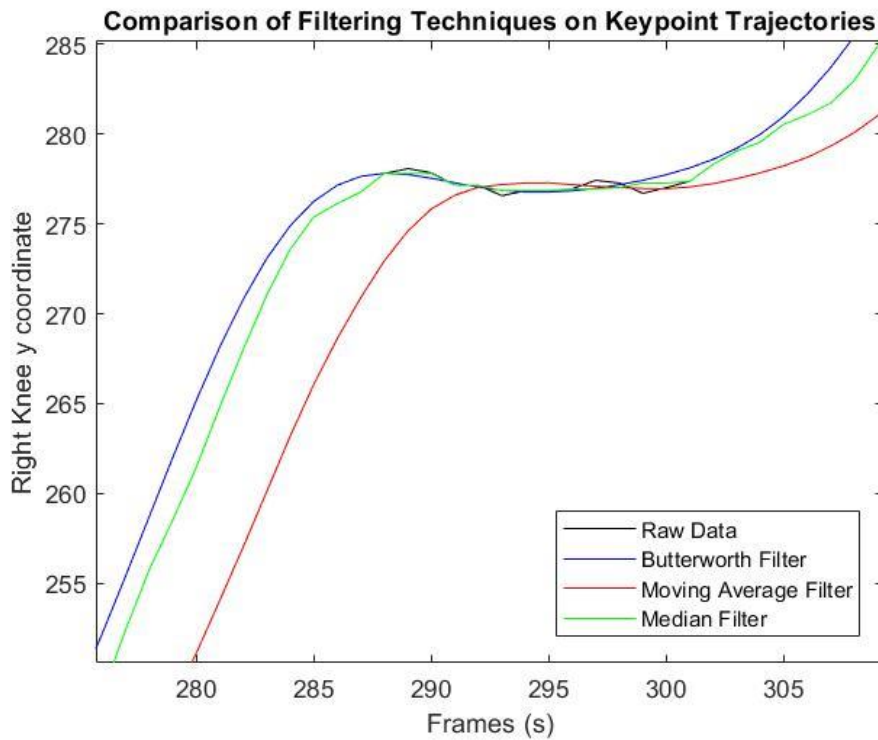


FIGURE 21. An example of all three filters applied on raw right knee y coordinate. Butterworth filter followed the original data without losing information.

## 6.5 Skeletal modelling and joint angle calculation

Marker trajectories from Vicon were first processed in Nexus. The markers were labelled, and subject-specific static calibration was performed using the plug-in-gait model for the 3D reconstruction. Gaps (caused by marker occlusion) in the marker trajectories were filled with the "rigid body fill" function in Nexus. This method uses the known positions of other markers that are part of the same segment to estimate the position of the missing marker. The marker trajectories were then exported as C3D files.

For the BlazePose data, additional virtual markers (for knee and ankle) were created in Matlab to prevent the femur from flipping, stabilize the limb representation, and improve kinematic analysis. We applied the cross product to pairs of vectors representing different body parts (e.g., the right foot orientation vector and the upward direction vector of the right tibia), resulting in new vectors perpendicular to the plane formed by the initial pair (e.g., the lateral-medial direction of the tibia). These new vectors are then normalized to a unit vector to ensure they only represent direction. Finally, we labeled the BlazePose keypoints and converted both BlazePose (CSV to TRC) and Vicon (C3D to TRC) data files in Matlab to make them compatible in OpenSim.

We utilized the Gait2392 gait model (Anderson & Pandy 1999; Delp et al. 1990; Yamaguchi & Zajac 1989) to calculate IK in OpenSim to obtain joint angles. The reason the joint angle calculation was performed in OpenSim for both Vicon and BlazePose data was to eliminate the methodological difference since the plug-in-gait model and the Gait2392 model in OpenSim may have different approaches to biomechanical modeling, including how they define and calculate joint angles. The model was scaled for each subject. The scale factors were calculated based on the marker pairs that were manually defined for each segment. The weights for markers were set to 1 (the default setting) to specify the marker matching accuracy of the model. After scaling and adjusting the weights, the IK was calculated, and the results were exported as STO files.

Joint angles were filtered with a 4th-order low-pass Butterworth filter with a cutoff frequency of 6 (estimated with residual analysis). For the comparison of Vicon and BlazePose, the data was trimmed to obtain the relevant gait cycles of hip, knee, and ankle joint angles. Then both data were time normalized to 101 data points (one gait cycle) due to the sampling frequency

difference (Vicon: 200 Hz, GoPro: 60 Hz). Time normalization to 101 data points per gait cycle is essential because it enables the Vicon and BlazePose data to be directly compared, regardless of the sampling frequencies used in the original data sets. We ensure that every data point corresponds to the same relative period within the gait cycle by dividing each gait cycle into an equal number of data points. This allows for a meaningful comparison of joint angle patterns across time. After the time normalization, the gait cycles were averaged (5 trials per subject), and the standard deviation of the mean curves was taken.

## 6.6 Statistical analysis

One-dimensional statistical parametric mapping (SPM1D) was used to compare hip, knee, and ankle joint angle curves from Vicon and BlazePose. SPM1D was developed by Todd Pataky (2010) and was inspired by statistical parametric mapping (SPM) (Friston et al. 1994), which is widely used for the analysis of brain imaging data. The concepts and methods of SPM were modified by Todd Pataky for use with one-dimensional data, allowing for comparable advanced statistical analyses in the field of biomechanics. SPM analyses were implemented with the open-source SPM1D M.0.4.11 software package in Matlab 2020b.

The normality of the data was assessed with the Shapiro-Wilk test. A non-parametric two-tailed paired t-test ( $\alpha = 0.05$ ) was then performed to assess the point-wise curve differences. For every supra-threshold cluster in the data, a p-value was assigned. Significant p-values highlight the time points in the data where the observed effects exceed the critical significance level, i.e. where there are differences between the two methods being compared. For non-parametric tests, SPM1D utilizes a permutation technique that repeatedly rearranges the labels of the data and re-analyses them to determine whether the differences observed are significant. The total number of permutations determines the minimum possible p-value that can be detected (Pataky et al. 2015):

$$P_{min} = \frac{1}{N}$$

## 7. RESULTS

The range of joint angle values of CP and TD groups for each joint and method are shown in table 2. Mean ankle angle traces for each group and both methods are shown in Figure 22 (A and B). SPM1D analysis from the comparison of time-normalized and averaged joint angle data of BlazePose and Vicon for the CP group (Figure 22C) showed that there was one supra-threshold cluster observed (significant difference) in ankle angle at 0.7–1.3% ( $p < 0.016$ ) (~loading response) of the gait cycle. For the TD group (Figure 22D), there was one supra-threshold cluster observed at 38–46% ( $p < 0.016$ ) (~terminal stance) of the gait cycle. Individual mean ankle angle traces are shown in appendix 1 for CP and appendix 2 for TD groups.

TABLE 2. Minimum/maximum flexion angles of CP and TD groups for BlazePose and Vicon.

	CP min/max flex (deg)		TD min/max flex (deg)	
	BlazePose	Vicon	BlazePose	Vicon
Ankle	0 – 35.7	-1.1 – 30.7	-1.1 – 38.6	-1.1 – 28.2
Knee	-0.7 – 49.4	-0.6 – 45.6	-2.7 – 54.1	-0.4 – 59.5
Hip	-1.8 – 46.5	-1.3 – 37.4	-1.4 – 54.4	-0.9 – 40.1



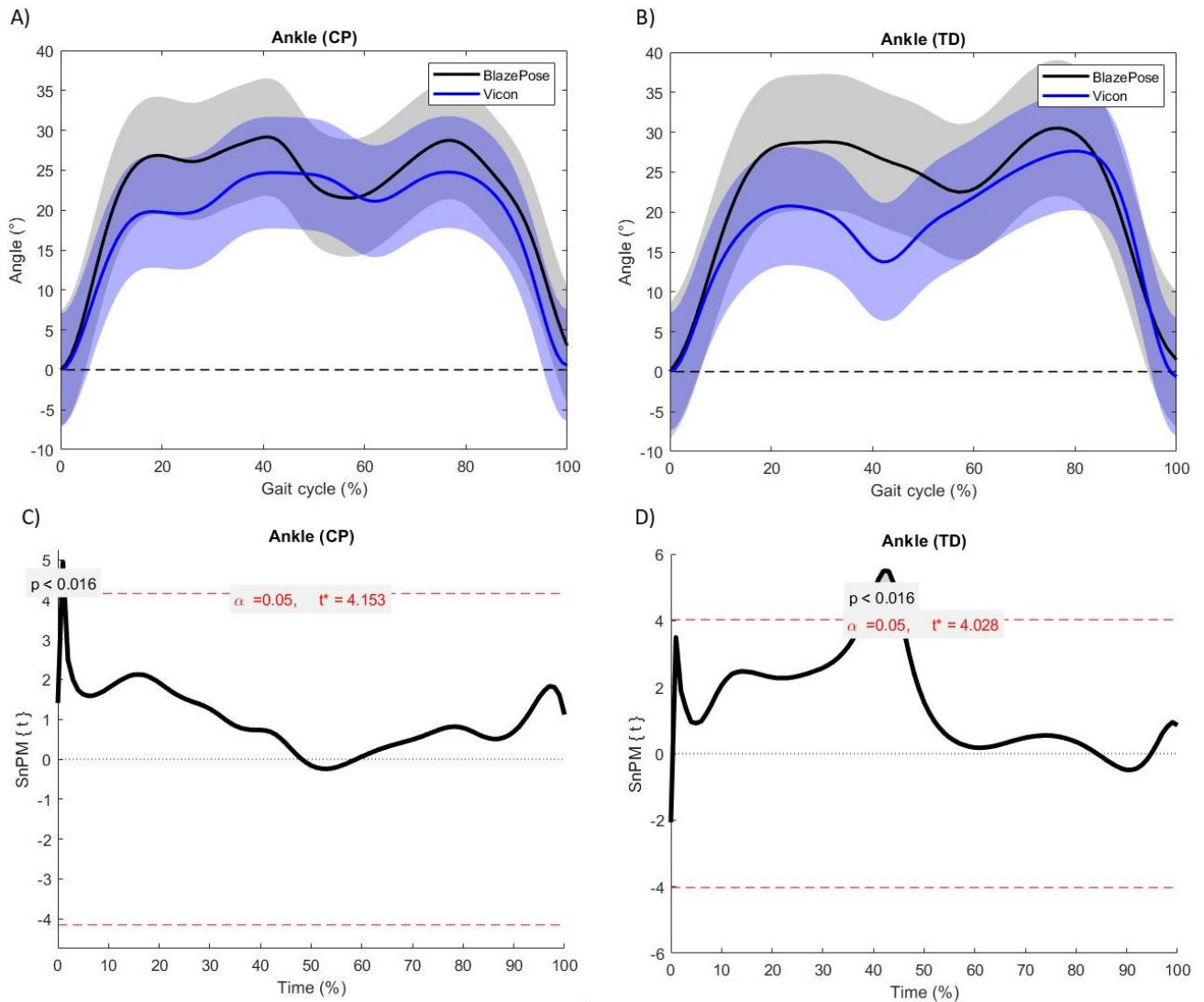


FIGURE 22. A and B: Comparison of mean ankle joint angle for each group. The thick line represents the mean joint angles, and the shaded area represents 1 SD. C and D: results obtained from SPM1D for CP and TD participants. The clusters are assigned, and p-values represent the significant difference between methods in the related data points.

No supra-threshold clusters were found in the knee joint angle comparison for both groups, indicating there was no significant difference observed between the methods (Figure 23). Individual mean knee angle traces are shown in appendix 3 for CP and appendix 4 for TD groups.

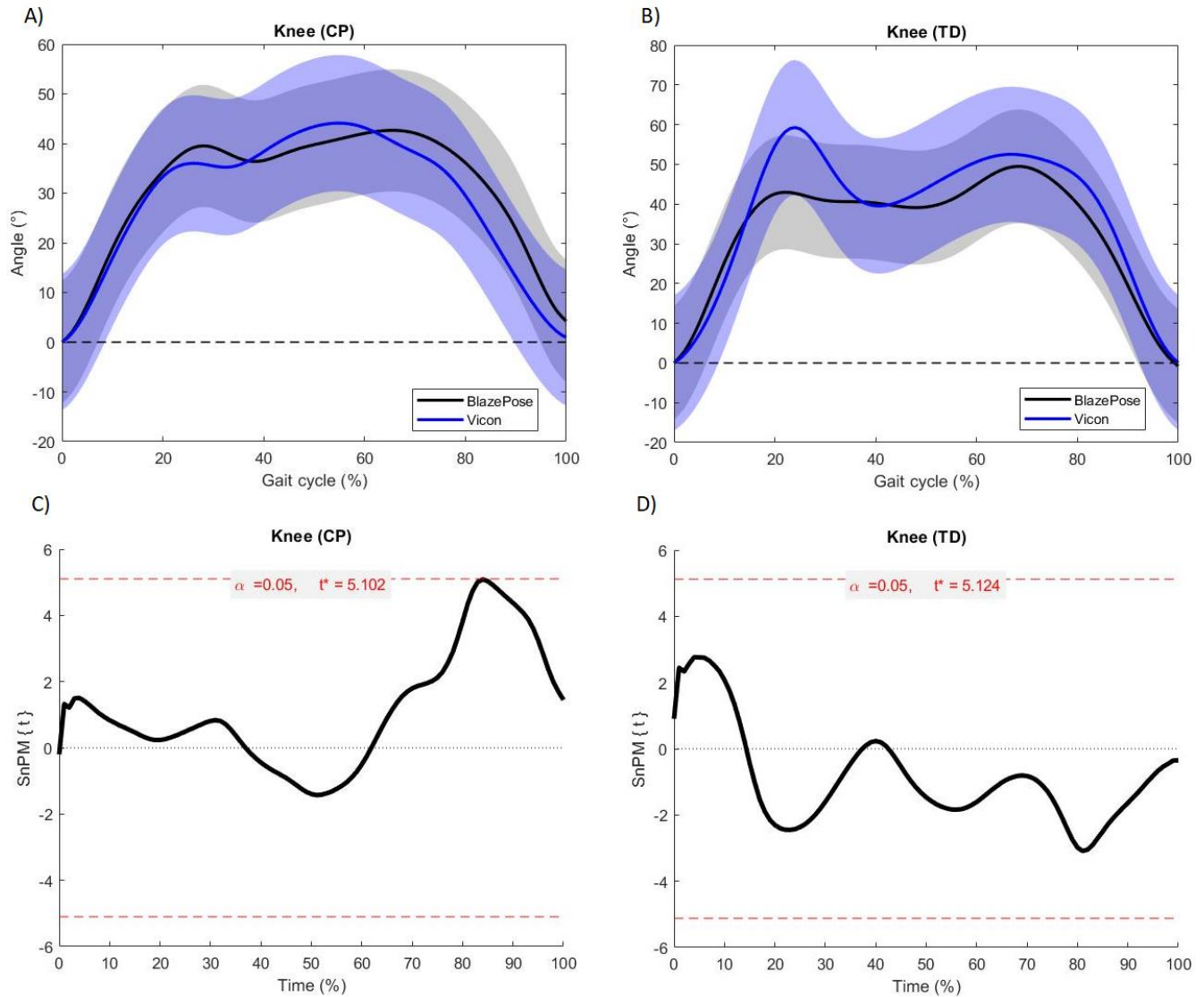


FIGURE 23. A and B: Comparison of mean knee joint angle for each group. The thick line represents the mean joint angles, and the shaded area represents 1 SD. C and D: results obtained from SPM1D for CP and TD participants. The clusters are assigned, and p-values represent the significant difference between methods in the related data points.

Mean hip flexion results are shown in Figure 24. For the hip flexion angle of the CP group, supra-threshold clusters were observed at 0.81–1.81% ( $p < 0.016$ ) (~loading response), 13–40% ( $p < 0.016$ ) (~from mid stance to terminal stance), and 89–93% ( $p < 0.016$ ) (~terminal swing) (Figure 24C). In the TD group, the only cluster that was observed occurred at 75–84% ( $p < 0.016$ ) (~mid swing) (Figure 24D). Individual mean hip angle traces are shown in appendix 5 for CP and appendix 6 for TD groups.

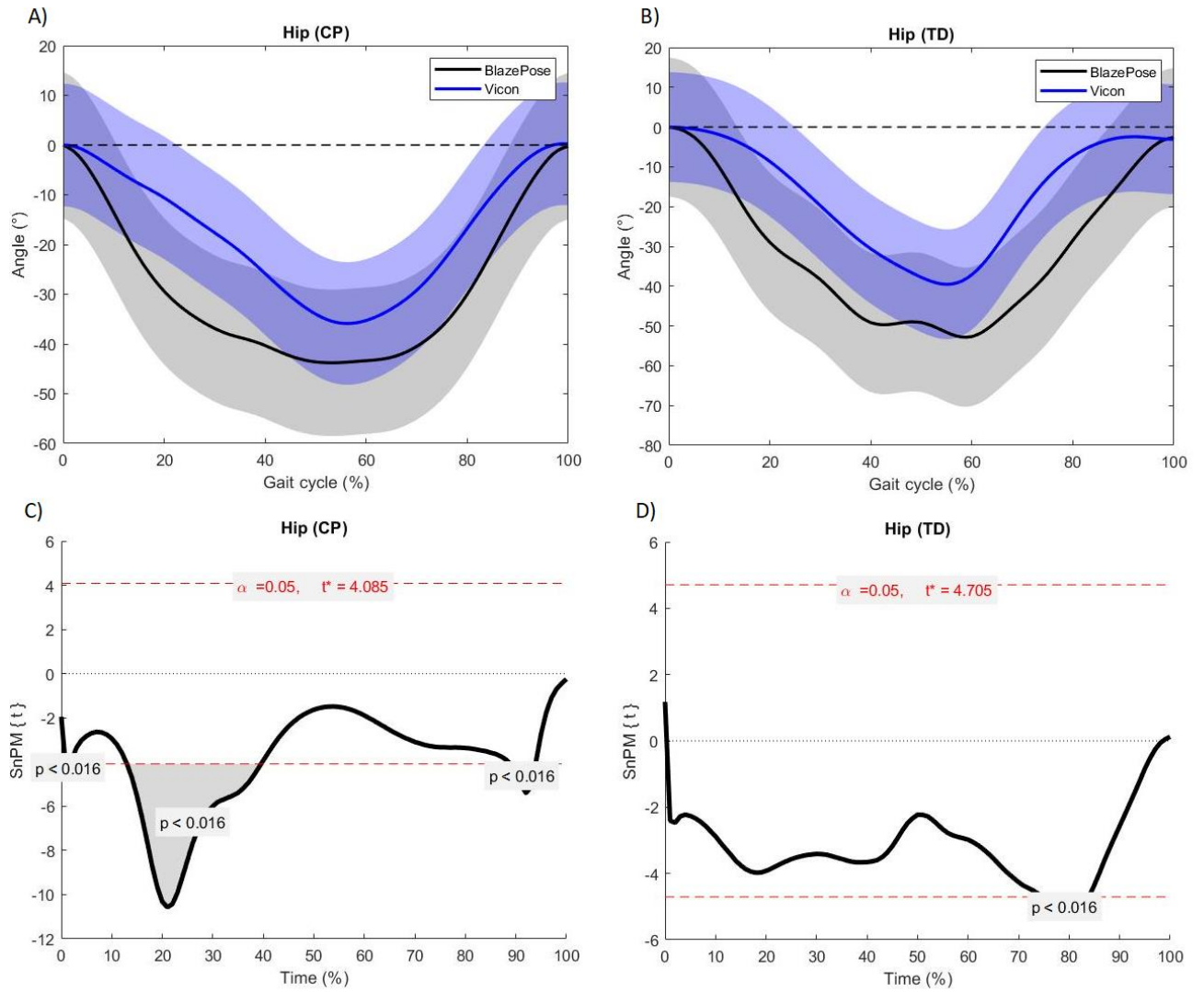


FIGURE 24. A and B: Comparison of mean knee joint angle for each group. The thick line represents the mean joint angles, and the shaded area represents 1 SD. C and D: results obtained from SPM1D for CP and TD participants. The clusters are assigned, and p-values represent the significant difference between methods in the related data points.

## **8. DISCUSSION**

In this study we compared markerless measurement system (Blazepose) to marker-based measurement system (Vicon) by assessing hip, knee and ankle joint angles in CP patients and TD peers. In general, Blazepose was able to follow similar patterns to Vicon however, it was not able to track as accurately as Vicon for both groups as expected in our hypothesis.

### **8.1 Ankle joint assessment**

Blazepose seemed to overestimate the tracking of the ankle for both groups. For the CP group, a significant difference occurred in the loading response phase of the gait cycle. This difference may be due to the altered gait mechanics commonly seen in CP patients. For instance, CP patients often have limited dorsiflexion range in their ankle joints, preventing the heel from making initial contact; instead, initial contact happens with forefoot or full-foot contact (Rodda et al. 2004), and Blazepose may not be able to detect this altered movement. This could be critical for individuals with CP, who may have increased stiffness or altered muscle activation patterns that require more sensitive measurement systems.

For the TD group, a significant difference was observed in the terminal stance phase of the gait cycle, where the heel rises, and the contralateral foot makes contact with the ground, which indicates Blazepose may not be able to detect finer movements occurring during normal gait. The reason this was not observed in the CP group could be due to the stiffness in the ankles of these patients, which likely prevents the occurrence of minor movements within the joint. A similar behavior was observed in another study when the participant walked with an ankle orthosis. The difference between OpenPose and marker-based system was less than walking normally because the OpenPose algorithm did not take toe flexion into consideration (Takeda et al. 2021).

### **8.2 Knee joint assessment**

Blazepose was able to track the knee best for both groups compared to Vicon, showing no significant differences between both systems. However, we observed a tendency to have a difference (not significant, Blazepose overestimates) from mid-stance to terminal-stance of the gait cycle in the TD group (Figures 23B, 23D), which shows Blazepose may not be able to track

finer movements in a healthy gait. The reason we did not see this in the CP group might be that the CP patients often have an inability to move or flex their joints through a full range and may have stiff joints (Wang et al. 2015). The stiff knee is most prevalent in patients with GMFCS level III (Rethlefsen et al. 2017) which three of our patient's GMFCS level was III. Moreover, for the CP group, we observed a tendency to have a difference (not significant, BlazePose underestimates) in swing phase of the gait cycle (Figures 23A, 23C). For the TD group, we did not observe this. The reason might be that CP patients may have issues with foot clearance due to an inability to fully flex their knee or hip during swing phase, which ultimately causes an altered gait with swinging their leg in a semi-circle (Wang et al. 2015) and BlazePose might be having difficulties tracking these altered movements. Nevertheless, BlazePose was generally able to track the knee joint similarly to Vicon showing promising results for this particular joint.

### **8.3 Hip joint assessment**

BlazePose particularly struggled to track the hip angles, especially for CP patients, aligning with previous studies indicating the open-source model's difficulties in accurate hip joint tracking (Needham et al. 2021; Stenum et al. 2021). For the CP group, BlazePose showed significant differences in loading response, mid-to-terminal stance, and terminal swing. Weakness in the hip abductors can lead to hip instability during loading for CP patients, which often causes hip instability. Moreover, body weight transfer and hip stabilization remain to be a challenge throughout the gait cycle, resulting in asymmetrical hip movements, which ultimately BlazePose was not able to track these gait characteristics of CP patients.

For the TD group, only at the mid-swing phase was a significant difference observed, which is a crucial part of the gait where the hip joint moves into flexion for propelling the leg forward. This indicates that BlazePose may not be able to detect rapid movements in the hip. For the rest of the gait cycle, BlazePose was able to track the hip motions similarly to Vicon.

### **8.4 Accuracy concerns**

In general, the curve patterns were similar between BlazePose and Vicon. However, there were differences observed in tracking in certain phases of the gait, especially for the CP group. Perhaps the biggest reason for this is due to the shortcomings of the current open-source models

since they were never designed for biomechanical purposes, resulting in inaccurate joint center locations and angles. Mainly because labelers without anatomical knowledge were recruited for the manual labeling of the training datasets. (Needham, Evans, Cosker, Wade, et al., 2021; Wade et al., 2022) In addition, not all images in open-source labeled datasets need to pass a second verification phase; thus, two people may interpret a joint center very differently, which could result in inconsistently classified images (Cronin 2021), which is not sufficient for specific populations and clinical applications.

It is evident that tailoring a dataset that is specific for biomechanical and clinical purposes increases the accuracy of models trained on that data significantly. Vafadar et al. (2021) created a new dataset (ENSAM), which included images of individuals with musculoskeletal conditions. The annotation data were obtained from the recorded data by the marker-based system and the EOS imaging system (bi-planar radiography). The anatomic landmarks were specified by orthopaedic surgeons. They first trained the pose estimation model (Iskakov et al. 2019) on a pre-existing dataset (Human3.6M), then retrained it (fine-tuned) on the ENSAM dataset. The results showed that the confidence intervals for the accuracy of the joint center estimations and the tracking errors were significantly reduced after fine-tuning on the ENSAM dataset compared to marker-based motion capture systems and medical imaging. In another study, Wren et al. (2023) compared the kinematics between commercially available Theia markerless (trained for biomechanical use) and marker-based systems in clinical patients, including patients with foot deformities and wearing orthotics. Theia markerless have been tested several times previously and the studies reported similar kinematic and temporospatial parameters on healthy subjects compared to the marker-based systems however, large differences in transverse plane (RMS difference: for thigh  $8.5^\circ$ , for shank  $13^\circ$ ) and ankle inversion/eversion angles (RMS difference:  $8^\circ$ ) were reported (Kanko et al. 2021a; Kanko et al. 2021b; Kanko et al. 2021c). Wren et al. (2023) reported similar results, and differences were attributed to segment definition rather than tracking accuracy. However, larger RMSDs were observed for patients with foot deformities, wearing orthoses, or using assistive devices, which indicates the need for datasets that include various pathological gait characteristics. While a dataset tailored for biomechanical use enhances tracking accuracy, a dataset that is diverse with clinical movements may be key to achieving outcomes that are relevant and applicable in a clinical setting.

Furthermore, as mentioned previously, in marker-based systems, marker placement relies on physical palpation on anatomical landmarks, whereas HPE models rely on visibly labeled generalized key points (e.g., lateral malleolus, femoral epi condyle vs. "ankle," "knee," "hip"), which results in tracking inaccuracies, especially for the hip joint since it is very challenging to identify it accurately without physical palpation (Needham et al., 2021). In addition, an insufficient number of keypoints leads to an increase in tracking errors and incorrect limb orientations because the major joints are treated as rigid segments, as reported in our study, such as the hip being rigid and the femur flipping. We were able to overcome the femur flipping issue by creating virtual markers for the knee and the ankle. However, we were not able to create virtual markers for the hip because the upper body did not have enough keypoints to perform the cross product, ultimately the reason for the large errors observed in hip joint angle. In addition, creating virtual markers added an extra step to the data processing, making it more time-consuming. Previous studies interpreted these large errors as a result of the different tracking and modeling approaches of the two systems. Even though we have modeled both data from BlazePose and Vicon in OpenSim to overcome the modeling differences, the tracking approach is hard to overcome with pre-trained open-source models without developing new models with new datasets and tracking approaches.

Thus, new open-source datasets must be designed specifically for biomechanical and clinical applications, include at least three keypoints for each segment for stabilizing limbs, and be annotated by labelers with anatomical knowledge by including a verification step. It is also important to note that for clinical datasets, consent and ethical approval should be considered (Wade et al. 2022).

## **8.5 Technical and methodological considerations**

Due to the relatively limited sample size of our study, there were statistical limitations when applying SPM1D. As mentioned before, SPM1D uses a permutation technique where the total number of permutations determines the minimum possible p-value that can be detected. Since we had six subjects, the total number of permutations was 64. Therefore, the minimum possible p-value that was detected and continuously seen at each significance level was 0.016 (see results), which might not be as generalizable because this limitation might have restricted the

ability to see more details in the data. Larger and more varied sample sizes should be used in future research to improve the result's statistical power and generalizability.

Moreover, due to the instable frame rate of the GoPro cameras, there was extra noise in the data caused by the synchronization errors. Even though all three GoPro cameras were synchronized and the frame rate was set to 60 Hz, from time to time the cameras showed instability in the frame rates, which ultimately affected the 3D joint coordinate calculation and caused twitches and noise in the keypoint trajectories. We were able to overcome the noise by filtering the data; however, there were still inaccuracies and twitches in the keypoint trajectories, which is hard to overcome since overfiltering can cause loss in the data.

Another consideration is the number of cameras. Three GoPro cameras were sufficient to calculate the 3D coordinates in our study. It is evident that multiple cameras increase the accuracy of the tracking. While it increases the cost and space because of additional hardware, to extract the biomechanically relevant 3D joint angles and avoid the self-occlusion errors reported by the previous studies (Serrancoli et al. 2020; Stenum et al. 2021b), it is essential to have multiple cameras. However, it is also important to note that more cameras also mean additional data processing, and in clinical settings where quick assessments are needed, this may not be efficient. Furthermore, camera alignment is very important to consider too. We have angled frontal and posterior cameras 45° from the corner side of the gait path (Figure 16) to aim for better capture volume. On the other hand, compared to a frontal or 45° angle, joints like the hip or shoulder may appear different from the side (Needham et al. 2021). While the reason for the large errors observed in hip joint angles was because the hip was treated as a rigid segment in BlazePose, camera placement might have contributed to the large errors observed in hip joint angles too. Thus, increasing the camera number and optimizing camera placement must be considered for better accuracy. Additionally, the accuracy of markerless systems will be impacted by resolution (both temporal and spatial), just like marker-based systems are (Colyer et al. 2018).

Calibration is another crucial step for achieving higher accuracy. While we were able to successfully calculate the capture volume and calibrate, there are important steps to consider. To start with, measuring the marked points accurately is essential for successful calibration because even one incorrect point can affect the calculation and ultimately the 3D coordinate calculations. While a mapping device is not crucial to measuring these marked and known



points, it reduces measurement errors and allows for better calibration and correct translation of 3D coordinates. However, it is also important to consider that it is uncommon to have these mapping devices in actual clinical settings. Instead, a measuring tape can be used for the calculation of known point coordinates in such cases.

While it was very quick and effortless to collect markerless data, one significant issue with these systems is the length of time it takes to obtain the keypoints and further process the data. Expensive computer infrastructure is typically required for faster computation. Using the supercomputer to obtain the keypoints significantly reduced the computational load however, further analysis took a long time, which can be problematic in clinical situations where quick analysis is frequently required. In addition, for the analysis, an operator with computer science knowledge is needed, which makes these systems less flexible. For users without a background in computer science or programming, this technology needs to be made more user-friendly. This emphasizes how crucial it is to provide quicker and more effective data processing techniques for these sophisticated motion capture methods. Some commercial systems, such as Theia3D, have made their software user-friendly by implementing their algorithm with software from known marker-based motion capture companies (e.g., Qualisys and Vicon). Nevertheless, it is important to note that the user cannot determine the keypoints that their method is obtaining with these commercial systems because they utilize a private dataset and posture estimation algorithm. (Wade et al. 2022)

Finally, it is also important to question whether it is logical to align markerless motion capture with marker-based motion capture since even marker-based approaches have their own errors, like soft tissue artifact as discussed previously. Perhaps, future studies must compare the markerless motion capture models to bi-planar video radiography to determine the true accuracy.

## **8.6 Usability and potential of markerless motion capture**

While there are still many limitations to overcome with markerless motion capture, it still showed great potential. While data processing can be time-consuming with these systems, the study by Kidziński et al. (2020) showed the potential to automatically process video data and provide quantitative results that a clinician might employ right away with an additional neural network. With this method, virtual clinic measurements could be possible. For example,

patients could film a video at home and email it to their physicians for further assesment. Additionally, it has the ability to evaluate data automatically and speed up the entire data processing process.

Blazepose can operate on conventional computers with moderate graphical processing units or even smaller devices like mobile phones, unlike previous pose estimation techniques that required significant processing power stored in high end computers (Cao et al. 2018). While better camera quality is still needed, it will become more practical to use the phone cameras and processors to enable small-sized, more accessible motion capture as HPE models advance. Moreover, automatic real-time markerless motion capture could be performed by researchers, clinicians, and coaches without requiring a significant setup expense, especially for CP patients, where this process can be very tiring. In addition, these systems can be used where temporospatial information (step length, stride length, stride time, walking speed and cadence) is needed since previous studies (Azhand et al. 2021; Kanko et al. 2021c; Moro et al. 2020; Shin et al. 2021) reported good agreement compared to marker-based systems and GAITRite pressure walkway. Lastly, markerless systems have the potential to gather patient data in home settings during everyday activities, eliminating the requirement for an operator or specialized clothing. These systems could be installed in shared spaces of nursing homes, allowing for the collection of data from multiple patients in a setting that captures their natural walking patterns (Wade et al. 2022).

## 9. CONCLUSION

This study tested the feasibility of markerless motion capture by comparing a human pose estimation model (Blazepose) to a marker-based motion capture system (Vicon) by assessing ankle, knee and hip joint angles of CP patients and their TD peers. Blazepose in general was able to follow similar patterns compared to Vicon. While Blazepose was able to track the knee joint angle, there were significant differences observed in ankle and hip joint angles between methods. Blazepose shows great potential, however, high accuracy and reliability are needed for clinical applications showing that markerless motion capture is not ready to be used in clinical settings.

Moreover, while it was very easy and effortless to collect data, extensive amount of time and programming knowledge is needed for data processing which is not efficient for clinical settings. Especially where quick assessments are needed. Before markerless systems can be used in actual clinical settings, first the algorithms must change. The algorithms must use better ground truth in their datasets, recruit labelers who possess anatomical knowledge and employ a verification process to avoid the annotation errors. Furthermore, for clinical applications datasets that are diverse with various musculoskeletal conditions and pathologic movements must be created with data augmentation techniques. Lastly, more studies with clinical populations with a larger sample size must be conducted for better generalization.

This study highlights the need for accuracy in clinical settings and suggests the direction for future research to enhance motion capture technology.

## REFERENCES

- Åberg, A. C., Olsson, F., Åhman, H. B., Tarassova, O., Arndt, A., Giedraitis, V., Berglund, L., & Halvorsen, K. (2021). Extraction of gait parameters from marker-free video recordings of Timed Up-and-Go tests: Validity, inter- and intra-rater reliability. *Gait and Posture*, 90, 489–495. <https://doi.org/10.1016/j.gaitpost.2021.08.004>
- Aggarwal, C. C. (2018). Training Deep Neural Networks. In *Neural Networks and Deep Learning*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-94463-0>
- Albuquerque, P., Verlekar, T. T., Correia, P. L., & Soares, L. D. (2021). A spatiotemporal deep learning approach for automatic pathological gait classification. *Sensors*, 21(18). <https://doi.org/10.3390/s21186202>
- Alpaydin, E. (2009). *Introduction to Machine Learning* (2nd ed.). The MIT Press.
- Alpaydin, E. (2010). Multilayer Perceptrons. In T. Dietterich (Ed.), *Introduction to Machine Learning* (2nd ed., pp. 233–275). The MIT Press.
- Anderson, F. C., & Pandy, M. G. (1999). A Dynamic Optimization Solution for Vertical Jumping in Three Dimensions. *Computer Methods in Biomechanics and Biomedical Engineering*, 2(3), 201–231. <https://doi.org/10.1080/10255849908907988>
- Azhand, A., Rabe, S., Müller, S., Sattler, I., & Heimann-Steinert, A. (2021). Algorithm based on one monocular video delivers highly valid and reliable gait parameters. *Scientific Reports*, 11(1), 14065. <https://doi.org/10.1038/s41598-021-93530-z>
- Baker, R., Esquenazi, A., Benedetti, M. G., & Desloovere, K. (2016). Gait analysis: clinical facts. *European Journal of Physical and Rehabilitation Medicine*, 52(4), 560–574.
- Baldewijns, G., Claes, V., Debar, G., Mertens, M., Devriendt, E., Milisen, K., Tournoy, J., Croonenborghs, T., & Vanrumste, B. (2016). Automated in-home gait transfer time analysis using video cameras. *Journal of Ambient Intelligence and Smart Environments*, 8(3), 273–286. <https://doi.org/10.3233/AIS-160379>
- Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), 3–31. [https://doi.org/10.1016/S0167-7012\(00\)00201-3](https://doi.org/10.1016/S0167-7012(00)00201-3)
- Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). *BlazePose: On-device Real-time Body Pose tracking*.
- Bishop, C. (2006). Neural Networks. In M. Jordan, J. Kleinberg, & B. Schölkopf (Eds.), *Pattern Recognition and Machine Learning* (pp. 225–284). Springer Science+Business Media, LLC.

- Bishop, C. M. (2006a). Decision Theory. In M. Jordan, J. Kleinberg, & B. Schölkopf (Eds.), *Pattern Recognition and Machine Learning* (pp. 38–46). Springer Science+Business Media, LLC.
- Bishop, C. M. (2006b). *Pattern Recognition and Machine Learning* (M. Jordan, J. Kleinberg, & B. Schölkopf, Eds.). Springer Science+Business Media, LLC.
- Bittner, M., Yang, W.-T., Zhang, X., Seth, A., van Gemert, J., & van der Helm, F. C. T. (2022). Towards Single Camera Human 3D-Kinematics. *Sensors*, 23(1), 341. <https://doi.org/10.3390/s23010341>
- Camomilla, V., Bonci, T., & Cappozzo, A. (2017). Soft tissue displacement over pelvic anatomical landmarks during 3-D hip movements. *Journal of Biomechanics*, 62, 14–20. <https://doi.org/10.1016/j.jbiomech.2017.01.013>
- CanChild. (n.d.). GMFCS - E&R, Gross Motor Function Classification System - Expanded & Revised.
- Cans, C. (2007). Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Developmental Medicine & Child Neurology*, 42(12), 816–824. <https://doi.org/10.1111/j.1469-8749.2000.tb00695.x>
- Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2018). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.
- Chang, M. J., Ma, H. I., & Lu, T. H. (2015). Estimating the prevalence of cerebral palsy in Taiwan: A comparison of different case definitions. *Research in Developmental Disabilities*, 36C, 207–212. <https://doi.org/10.1016/J.RIDD.2014.10.001>
- Chapelle, O., Scholkopf, B., & Zien, Eds. , A. (2006). *Semi-Supervised Learning* (T. Dietterich, Ed.). The MIT Press.
- Chen, L., Armstrong, C. W., & Raftopoulos, D. D. (1994). An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. *Journal of Biomechanics*, 27(4), 493–500. [https://doi.org/10.1016/0021-9290\(94\)90024-8](https://doi.org/10.1016/0021-9290(94)90024-8)
- Christensen, D., Van Naarden Braun, K., Doernberg, N. S., Maenner, M. J., Arneson, C. L., Durkin, M. S., Benedict, R. E., Kirby, R. S., Wingate, M. S., Fitzgerald, R., & Yeargin-Allsopp, M. (2014). Prevalence of cerebral palsy, co-occurring autism spectrum disorders, and motor functioning - Autism and Developmental Disabilities Monitoring Network, USA, 2008. *Developmental Medicine & Child Neurology*, 56(1), 59–65. <https://doi.org/10.1111/dmcn.12268>

- Clark, R. A., Bower, K. J., Mentiplay, B. F., Paterson, K., & Pua, Y.-H. (2013). Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. *Journal of Biomechanics*, 46(15), 2722–2725. <https://doi.org/10.1016/j.jbiomech.2013.08.011>
- Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports Medicine - Open*, 4(1), 24. <https://doi.org/10.1186/s40798-018-0139-y>
- Corazza, S., Mündermann, L., Gambaretto, E., Ferrigno, G., & Andriacchi, T. P. (2010). Markerless Motion Capture through Visual Hull, Articulated ICP and Subject Specific Model Generation. *International Journal of Computer Vision*, 87(1–2), 156–169. <https://doi.org/10.1007/s11263-009-0284-3>
- Crenna, F., Rossi, G. B., & Berardengo, M. (2021). Filtering Biomechanical Signals in Movement Analysis. *Sensors*, 21(13), 4580. <https://doi.org/10.3390/s21134580>
- Cronin, N. J. (2021). Using deep neural networks for kinematic analysis: Challenges and opportunities. In *Journal of Biomechanics* (Vol. 123). Elsevier Ltd. <https://doi.org/10.1016/j.jbiomech.2021.110460>
- Cronin, N. J., Rantalainen, T., Ahtiainen, J. P., Hynynen, E., & Waller, B. (2019). Markerless 2D kinematic analysis of underwater running: A deep learning approach. *Journal of Biomechanics*, 87, 75–82. <https://doi.org/10.1016/j.jbiomech.2019.02.021>
- Cunningham, R., Sánchez, M. B., Butler, P. B., Southgate, M. J., & Loram, I. D. (2019). Fully automated image-based estimation of postural point-features in children with cerebral palsy using deep learning. *Royal Society Open Science*, 6(11), 191011. <https://doi.org/10.1098/rsos.191011>
- Delp, S. L., Loan, J. P., Hoy, M. G., Zajac, F. E., Topp, E. L., & Rosen, J. M. (1990). An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering*, 37(8), 757–767. <https://doi.org/10.1109/10.102791>
- Drazan, J. F., Phillips, W. T., Seethapathi, N., Hullfish, T. J., & Baxter, J. R. (2021). Moving outside the lab: Markerless motion capture accurately quantifies sagittal plane kinematics during the vertical jump. *Journal of Biomechanics*, 125. <https://doi.org/10.1016/j.jbiomech.2021.110547>
- Dugan, E. L., & Shilt, J. S. (2020). The Role of Motion Analysis in Surgical Planning for Gait Abnormalities in Cerebral Palsy. *Physical Medicine and Rehabilitation Clinics of North America*, 31(1), 107–115. <https://doi.org/10.1016/j.pmr.2019.09.009>

- Evans, M., Colyer, S., Cosker, D., & Salo, A. (2018). Foot Contact Timings and Step Length for Sprint Training. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 1652–1660. <https://doi.org/10.1109/WACV.2018.00184>
- Fang, H.-S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., Li, Y.-L., & Lu, C. (2022). AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time.
- Ferrari, A., Benedetti, M. G., Pavan, E., Frigo, C., Bettinelli, D., Rabuffetti, M., Crenna, P., & Leardini, A. (2008). Quantitative comparison of five current protocols in gait analysis. *Gait & Posture*, 28(2), 207–216. <https://doi.org/10.1016/j.gaitpost.2007.11.009>
- Fosdahl, M. A., Jahnsen, R., Kvalheim, K., & Holm, I. (2019). Effect of a Combined Stretching and Strength Training Program on Gait Function in Children with Cerebral Palsy, GMFCS Level I & II: A Randomized Controlled Trial. *Medicina*, 55(6), 250. <https://doi.org/10.3390/medicina55060250>
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. -P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210. <https://doi.org/10.1002/hbm.460020402>
- Gage, J. R. (1993). Gait analysis: An essential tool in the treatment of cerebral palsy. *Clinical Orthopaedics and Related Research*, 288, 126–134.
- Gage, J. R., & Schwartz, M. H. (2009a). Consequences of brain injury on musculoskeletal development. In J. R. Gage, M. H. Schwartz, & S. Koop (Eds.), *Identification and Treatment of Gait Problems in Cerebral Palsy* (Vols. 180–181, pp. 107–130). Mac Keith Press.
- Gage, J. R., & Schwartz, M. H. (2009b). Normal Gait. In J. R. Gage, M. H. Schwartz, & S. E. Koop (Eds.), *Identification and Treatment of Gait Problems in Cerebral Palsy* (1st ed., Vols. 180–181, pp. 31–65). Mac Keith Press.
- Gage, J. R., & Stout, J. L. (2009). Gait Analysis: Kinematics, Kinetics, Electromyography, Oxygen Consumption and Pedobarography. In J. R. Gage, M. H. Schwartz, & S. E. Koop (Eds.), *Identification and Treatment of Gait Problems in Cerebral Palsy* (1st ed., Vols. 180–181, pp. 260–285). Mac Keith Press.
- Galbusera, F., Casaroli, G., & Bassani, T. (2019). Artificial intelligence and machine learning in spine research. *JOR Spine*, 2(1), e1044. <https://doi.org/10.1002/jsp2.1044>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016a). Deep Feedforward Networks. In *Deep Learning*. The MIT Press.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016b). *Deep Learning*. The MIT Press.

- Gorter, J. W., Rosenbaum, P. L., Hanna, S. E., Palisano, R. J., Bartlett, D. J., Russell, D. J., Walter, S. D., Raina, P., Galuppi, B. E., & Wood, E. (2004). Limb distribution, motor impairment, and functional classification of cerebral palsy. *Developmental Medicine and Child Neurology*, 46(7), 461–467. <https://doi.org/10.1017/s0012162204000763>
- Graham, H. K., Rosenbaum, P., Paneth, N., Dan, B., Lin, J. P., Damiano, Di. L., Becher, J. G., Gaebler-Spira, D., Colver, A., Reddihough, Di. S., Crompton, K. E., & Lieber, R. L. (2016). Cerebral palsy. *Nature Reviews. Disease Primers*, 2. <https://doi.org/10.1038/NRDP.2015.82>
- Graham, H. K., & Selber, P. (2003). MUSCULOSKELETAL ASPECTS OF CEREBRAL PALSY. *The Journal of Bone and Joint Surgery. British Volume*, 85-B(2), 157–166. <https://doi.org/10.1302/0301-620X.85B2.14066>
- Grishchenko, I., Bazarevsky, V., Zanfira, A., Bazavan, E. G., Zanfira, M., Yee, R., Raveendran, K., Zhdanovich, M., Grundmann, M., & Sminchisescu, C. (2022). BlazePose GHUM Holistic: Real-time 3D Human Landmarks and Pose Estimation.
- Growney, E., Meglan, D., Johnson, M., Cahalan, T., & An, K.-N. (1997). Repeated measures of adult normal walking using a video tracking system. *Gait & Posture*, 6(2), 147–162. [https://doi.org/10.1016/S0966-6362\(97\)01114-4](https://doi.org/10.1016/S0966-6362(97)01114-4)
- Gulati, S., & Sondhi, V. (2018). Cerebral Palsy: An Overview. *Indian Journal of Pediatrics*, 85(11), 1006–1016. <https://doi.org/10.1007/s12098-017-2475-1>
- Hansen, A. H., Childress, D. S., & Meier, M. R. (2002). A simple method for determination of gait events. *Journal of Biomechanics*, 35(1), 135–138. [https://doi.org/10.1016/s0021-9290\(01\)00174-9](https://doi.org/10.1016/s0021-9290(01)00174-9)
- Hatze, H. (1988). High-precision three-dimensional photogrammetric calibration and object space reconstruction using a modified DLT-approach. *Journal of Biomechanics*, 21(7), 533–538. [https://doi.org/10.1016/0021-9290\(88\)90216-3](https://doi.org/10.1016/0021-9290(88)90216-3)
- Iskakov, K., Burkov, E., Lempitsky, V., & Malkov, Y. (2019). Learnable Triangulation of Human Pose.
- Jung, A. (2023). *Machine Learning: The Basics*. Springer Science+Business Media, LLC.
- Kanko, R. M., Laende, E. K., Davis, E. M., Selbie, W. S., & Deluzio, K. J. (2021). Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *Journal of Biomechanics*, 127. <https://doi.org/10.1016/j.jbiomech.2021.110665>
- Kanko, R. M., Laende, E. K., Strutzenberger, G., Brown, M., Selbie, W. S., DePaul, V., Scott, S. H., & Deluzio, K. J. (2021). Assessment of spatiotemporal gait parameters using a deep



- learning algorithm-based markerless motion capture system. *Journal of Biomechanics*, 122, 110414. <https://doi.org/10.1016/j.jbiomech.2021.110414>
- Kanko, R. M., Laende, E., Selbie, W. S., & Deluzio, K. J. (2021). Inter-session repeatability of markerless motion capture gait kinematics. *Journal of Biomechanics*, 121, 110422. <https://doi.org/10.1016/j.jbiomech.2021.110422>
- Kessler, S. E., Rainbow, M. J., Lichtwark, G. A., Cresswell, A. G., D'Andrea, S. E., Konow, N., & Kelly, L. A. (2019). A Direct Comparison of Biplanar Videoradiography and Optical Motion Capture for Foot and Ankle Kinematics. *Frontiers in Bioengineering and Biotechnology*, 7, 199. <https://doi.org/10.3389/fbioe.2019.00199>
- Kidziński, Ł., Yang, B., Hicks, J. L., Rajagopal, A., Delp, S. L., & Schwartz, M. H. (2020). Deep neural networks enable quantitative movement analysis using single-camera videos. *Nature Communications*, 11(1). <https://doi.org/10.1038/s41467-020-17807-z>
- Krogh, A. (2008). What are artificial neural networks? *Nature Biotechnology*, 26(2), 195–197. <https://doi.org/10.1038/nbt1386>
- Lance, J. W. (1980). Spasticity: disordered motor control. In R. G. Feldman, R. R. Young, & W. P. Koella (Eds.), *Symposium synopsis* (pp. 485–500). Symposia Specialists.
- Latorre, J., Colomer, C., Alcañiz, M., & Llorens, R. (2019). Gait analysis with the Kinect v2: normative study with healthy individuals and comprehensive study of its sensitivity, validity, and reliability in individuals with stroke. *Journal of NeuroEngineering and Rehabilitation*, 16(1), 97. <https://doi.org/10.1186/s12984-019-0568-y>
- Laurentini, A. (1994). The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2), 150–162. <https://doi.org/10.1109/34.273735>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Little J., W. (1861). On The Influence of Abnormal Parturition, Difficult Labours, Premature Birth, and Asphyxia Neonatorum, on the Mental and Physical Condition of the Child, Especially in Relation to Deformities. *Transactions of the Obstetrical Society of London*, 3, 243–344.
- Liu, Y., Gall, J., Stoll, C., Qionghai Dai, Seidel, H.-P., & Theobalt, C. (2013). Markerless Motion Capture of Multiple Characters Using Multiview Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11), 2720–2735. <https://doi.org/10.1109/TPAMI.2013.47>

- Lofterød, B., Terjesen, T., Skaaret, I., Huse, A.-B., & Jahnsen, R. (2007). Preoperative gait analysis has a substantial effect on orthopedic decision making in children with cerebral palsy: Comparison between clinical evaluation and gait analysis in 60 patients. *Acta Orthopaedica*, 78(1), 74–80. <https://doi.org/10.1080/17453670610013448>
- Lonini, L., Moon, Y., Embry, K., Cotton, R. J., McKenzie, K., Jenz, S., & Jayaraman, A. (2022). Video-Based Pose Estimation for Gait Analysis in Stroke Survivors during Clinical Assessments: A Proof-of-Concept Study. *Digital Biomarkers*, 9–18. <https://doi.org/10.1159/000520732>
- Martinez, H. R., Garcia-Sarreón, A., Camara-Lemarroy, C., Salazar, F., & Guerrero-González, M. L. (2018). Accuracy of Markerless 3D Motion Capture Evaluation to Differentiate between On/Off Status in Parkinson’s Disease after Deep Brain Stimulation. *Parkinson’s Disease*, 2018, 1–7. <https://doi.org/10.1155/2018/5830364>
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Mathis, A., Schneider, S., Lauer, J., & Mathis, M. W. (2020). A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron*, 108(1), 44–65. <https://doi.org/10.1016/j.neuron.2020.09.017>
- Menz, H. B., Latt, M. D., Tiedemann, A., Mun San Kwan, M., & Lord, S. R. (2004). Reliability of the GAITRite® walkway system for the quantification of temporo-spatial parameters of gait in young and older people. *Gait & Posture*, 20(1), 20–25. [https://doi.org/10.1016/S0966-6362\(03\)00068-7](https://doi.org/10.1016/S0966-6362(03)00068-7)
- Miller, F., Dabney, K. W., & Rang, M. (1995). Complications in cerebral palsy treatment. In C. H. Epps & J. R. Bowen (Eds.), *Complications in Pediatric Orthopaedic Surgery* (pp. 477–544). J.B. Lippincott Company.
- Mintz, Y., & Brodie, R. (2019). Introduction to artificial intelligence in medicine. *Minimally Invasive Therapy & Allied Technologies*, 28(2), 73–81. <https://doi.org/10.1080/13645706.2019.1575882>
- Miranda, D. L., Schwartz, J. B., Loomis, A. C., Brainerd, E. L., Fleming, B. C., & Crisco, J. J. (2011). Static and dynamic error of a biplanar videoradiography system using marker-based and markerless tracking techniques. *Journal of Biomechanical Engineering*, 133(12), 121002. <https://doi.org/10.1115/1.4005471>

- Morgan, P., & McGinley, J. L. (2018). Cerebral palsy. *Handbook of Clinical Neurology*, 159, 323–336. <https://doi.org/10.1016/B978-0-444-63916-5.00020-3>
- Moro, M., Marchesi, G., Hesse, F., Odone, F., & Casadio, M. (2022). Markerless vs. Marker-Based Gait Analysis: A Proof of Concept Study. *Sensors (Basel, Switzerland)*, 22(5). <https://doi.org/10.3390/s22052011>
- Moro, M., Marchesi, G., Odone, F., & Casadio, M. (2020). Markerless gait analysis in stroke survivors based on computer vision and deep learning. *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2097–2104. <https://doi.org/10.1145/3341105.3373963>
- Mroz, S., Baddour, N., McGuirk, C., Juneau, P., Tu, A., Cheung, K., & Lemaire, E. (2021). Comparing the Quality of Human Pose Estimation with BlazePose or OpenPose. *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, 1–4. <https://doi.org/10.1109/BioSMART54244.2021.9677850>
- Mundermann, L., Corazza, S., Chaudhari, A. M., Alexander, E. J., & Andriacchi, T. P. (2005). *Most favorable camera configuration for a shape-from-silhouette markerless motion capture system for biomechanical analysis* (J.-A. Beraldin, S. F. El-Hakim, A. Gruen, & J. S. Walton, Eds.; p. 278). <https://doi.org/10.1117/12.587970>
- Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Iino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in Sports and Active Living*, 2. <https://doi.org/10.3389/fspor.2020.00050>
- Needham, L., Evans, M., Cosker, D. P., & Colyer, S. L. (2021). Can Markerless Pose Estimation Algorithms Estimate 3D Mass Centre Positions and Velocities during Linear Sprinting Activities? *Sensors*, 21(8), 2889. <https://doi.org/10.3390/s21082889>
- Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-00212-x>
- Ohki, M., Zervakis, M. E., & Venetsanopoulos, A. N. (1995). 3-D Digital Filters. In C. T. Leondes (Ed.), *Multidimensional Systems: Signal Processing and Modeling Techniques* (Vol. 77, pp. 49–88). Academic Press Inc. [https://doi.org/10.1016/S0090-5267\(05\)80038-6](https://doi.org/10.1016/S0090-5267(05)80038-6)
- Ota, M., Tateuchi, H., Hashiguchi, T., Kato, T., Ogino, Y., Yamagata, M., & Ichihashi, N. (2020). Verification of reliability and validity of motion analysis systems during bilateral

- squat using human pose tracking algorithm. *Gait and Posture*, 80, 62–67. <https://doi.org/10.1016/j.gaitpost.2020.05.027>
- Pagnon, D., Domalain, M., & Reveret, L. (2021). Pose2sim: An end-to-end workflow for 3D markerless sports kinematics—Part 1: Robustness. *Sensors*, 21(19). <https://doi.org/10.3390/s21196530>
- Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine and Child Neurology*, 39(4), 214–223. <https://doi.org/10.1111/j.1469-8749.1997.tb07414.x>
- Palucci Vieira, L. H., Santiago, P. R. P., Pinto, A., Aquino, R., Torres, R. da S., & Barbieri, F. A. (2022). Automatic Markerless Motion Detector Method against Traditional Digitisation for 3-Dimensional Movement Kinematic Analysis of Ball Kicking in Soccer Field Context. *International Journal of Environmental Research and Public Health*, 19(3), 1179. <https://doi.org/10.3390/ijerph19031179>
- Pataky, T. C. (2010). Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *Journal of Biomechanics*, 43(10), 1976–1982. <https://doi.org/10.1016/j.jbiomech.2010.03.008>
- Pataky, T. C. (2012). One-dimensional statistical parametric mapping in Python. *Computer Methods in Biomechanics and Biomedical Engineering*, 15(3), 295–301. <https://doi.org/10.1080/10255842.2010.527837>
- Pataky, T. C., Vanrenterghem, J., & Robinson, M. A. (2015). Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. *Journal of Biomechanics*, 48(7), 1277–1285. <https://doi.org/10.1016/j.jbiomech.2015.02.051>
- Peacock, W. J. (2009). The Pathophysiology of Spasticity. In Gage James R, Schwartz Micheal H., & Koop Steven E. (Eds.), *Identification and Treatment of Gait Problems in Cerebral Palsy* (1st ed., Vols. 180–181, pp. 65–178). Mac Keith Press.
- Perry, J. (1992). Phases of Gait. In *Gait Analysis: Normal and Pathological Function* (pp. 9–16). SLACK Incorporated.
- Peters, A., Galna, B., Sangeux, M., Morris, M., & Baker, R. (2010). Quantification of soft tissue artifact in lower limb human motion analysis: A systematic review. *Gait & Posture*, 31(1), 1–8. <https://doi.org/10.1016/j.gaitpost.2009.09.004>
- Peterson, N., & Walton, R. (2016). Ambulant cerebral palsy. *Orthopaedics and Trauma*, 30(6), 525–538. <https://doi.org/10.1016/j.mporth.2016.08.005>

- Reid, S. M., Meehan, E., McIntyre, S., Goldsmith, S., Badawi, N., Reddihough, D. S., & Australian Cerebral Palsy Register Group. (2016). Temporal trends in cerebral palsy by impairment severity and birth gestation. *Developmental Medicine and Child Neurology*, *58 Suppl 2*, 25–35. <https://doi.org/10.1111/dmcn.13001>
- Rethlefsen, S. A., Blumstein, G., Kay, R. M., Dorey, F., & Wren, T. A. L. (2017). Prevalence of specific gait abnormalities in children with cerebral palsy revisited: influence of age, prior surgery, and Gross Motor Function Classification System level. *Developmental Medicine & Child Neurology*, *59*(1), 79–88. <https://doi.org/10.1111/dmcn.13205>
- Richards, J., Thewlis, D., Sinclair, J., & Hobbs, S. J. (2018). Methods of Analysis of Movement. In *The Comprehensive Textbook of Clinical Biomechanics* (2nd ed., pp. 156–180). Elsevier.
- Robertsen, G. D., Hamill, J., Kamen, G., & Caldwell, E. G. (2013). Planar Kinematics . In *Research Methods in Biomechanics* (2nd ed.).
- Rodda, J., & Graham, H. K. (2001a). Classification of gait patterns in spastic hemiplegia and spastic diplegia: a basis for a management algorithm. *European Journal of Neurology*, *8 Suppl 5*, 98–108. <https://doi.org/10.1046/j.1468-1331.2001.00042.x>
- Rodda, J., & Graham, H. K. (2001b). Classification of gait patterns in spastic hemiplegia and spastic diplegia: a basis for a management algorithm. *European Journal of Neurology*, *8*(s5), 98–108. <https://doi.org/10.1046/j.1468-1331.2001.00042.x>
- Rodda, J. M., Graham, H. K., Carson, L., Galea, M. P., & Wolfe, R. (2004). Sagittal gait patterns in spastic diplegia. *The Journal of Bone and Joint Surgery. British Volume*, *86*(2), 251–258. <https://doi.org/10.1302/0301-620x.86b2.13878>
- Rose, J., Haskell, W. L., Gamble, J. G., Hamilton, R. L., Brown, D. A., & Rinsky, L. (1994). Muscle pathology and clinical measures of disability in children with cerebral palsy. *Journal of Orthopaedic Research: Official Publication of the Orthopaedic Research Society*, *12*(6), 758–768. <https://doi.org/10.1002/jor.1100120603>
- Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., Dan, B., & Jacobsson, B. (2007). *A report: the definition and classification of cerebral palsy April 2006*.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. <https://doi.org/10.1038/323533a0>
- Sangeux, M., & Stephane, A. (2015). Kinematic Deviations in Children with Cerebral Palsy. In F. Canavese & J. Deslandes (Eds.), *Orthopedic Management of Children with Cerebral Palsy* (pp. 242–253). Nova Science Publishers, Inc.

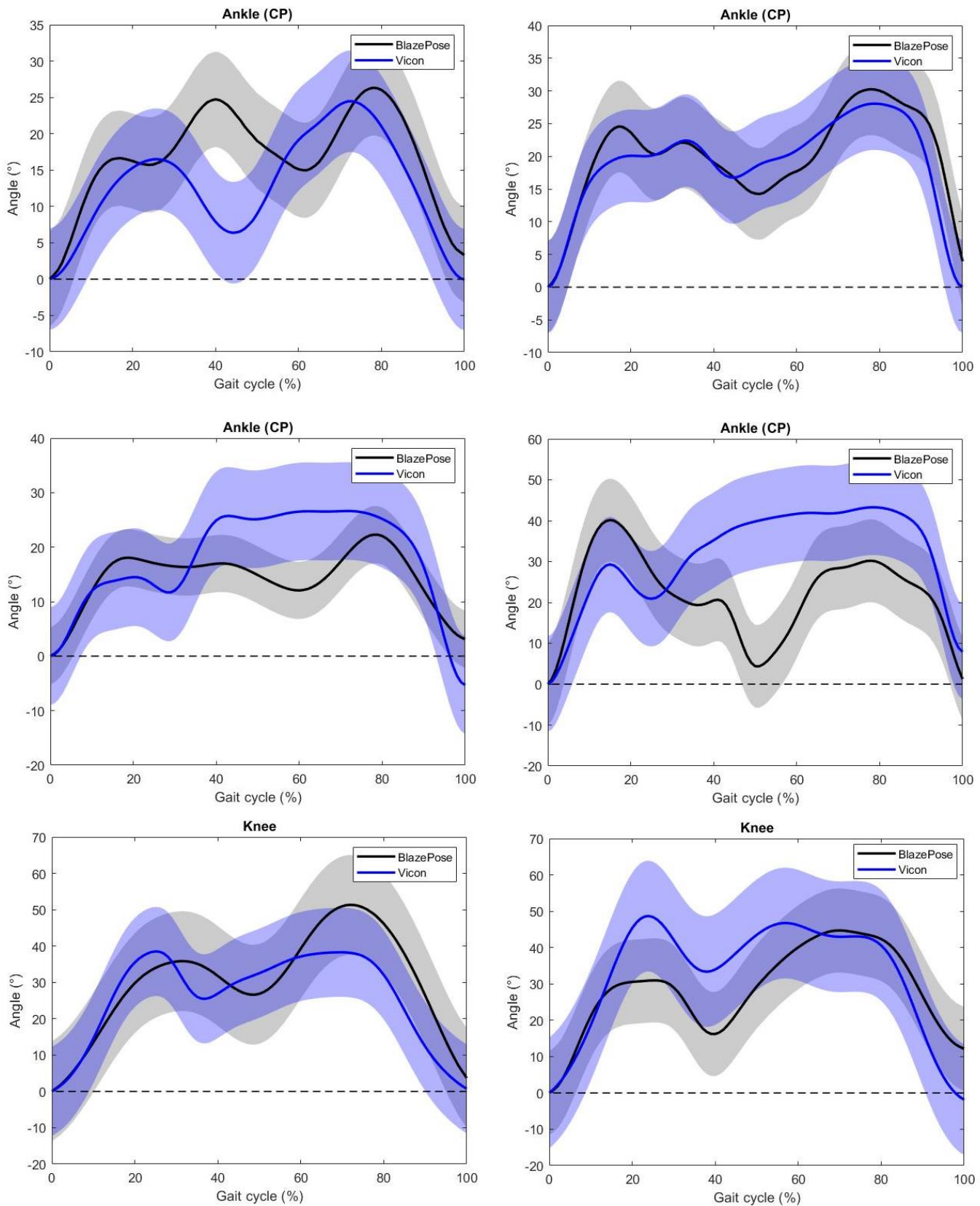
- Sato, K., Nagashima, Y., Mano, T., Iwata, A., & Toda, T. (2019). Quantifying normal and parkinsonian gait features from home movies: Practical application of a deep learning-based 2D pose estimator. *PLoS ONE*, *14*(11). <https://doi.org/10.1371/journal.pone.0223549>
- Scherer, D., Müller, A., & Behnke, S. (2010). *Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition* (pp. 92–101). [https://doi.org/10.1007/978-3-642-15825-4\\_10](https://doi.org/10.1007/978-3-642-15825-4_10)
- Seethapathi, N., Wang, S., Saluja, R., Blohm, G., & Kording, K. P. (2019). *Movement science needs different pose tracking algorithms*.
- Serranoli, G., Bogatikov, P., Huix, J. P., Barbera, A. F., Egea, A. J. S., Ribe, J. T., Kanaan-Izquierdo, S., & Susin, A. (2020). Marker-Less Monitoring Protocol to Analyze Biomechanical Joint Metrics During Pedaling. *IEEE Access*, *8*, 122782–122790. <https://doi.org/10.1109/ACCESS.2020.3006423>
- Sewell, M. D., Eastwood, D. M., & Wimalasundera, N. (2014). Managing common symptoms of cerebral palsy in children. *BMJ*, *349*(sep25 7), g5474–g5474. <https://doi.org/10.1136/bmj.g5474>
- Shahid, M. D. (2019, February 24). *Convolutional Neural Network: Learn Convolutional Neural Network from basic and its implementation in Keras*. Towards Data Science.
- Shalev-Shwartz, S., & Ben-David, S. (2014a). Neural Networks. In *Understanding Machine Learning: From Theory to Algorithms* (pp. 268–282). Cambridge University Press.
- Shalev-Shwartz, S., & Ben-David, S. (2014b). Stochastic Gradient Descent. In *Understanding Machine Learning: From Theory to Algorithms* (pp. 184–201). Cambridge University Press.
- Shehata, G., El-Tallawy, H., Farghaly, wafaa, Rageh, T., Badary, R., Sayed, M., Abd-Elwarth, A., Kandil, M., Metwally, N., & Abd El Hamed, M. (2014). Cerebral palsy in Al-Quseir City, Egypt: prevalence, subtypes, and risk factors. *Neuropsychiatric Disease and Treatment*, *12*(6), 1267. <https://doi.org/10.2147/NDT.S59599>
- Shin, J. H., Yu, R., Ong, J. N., Lee, C. Y., Jeon, S. H., Park, H., Kim, H.-J., Lee, J., & Jeon, B. (2021). Quantitative Gait Analysis Using a Pose-Estimation Algorithm with a Single 2D-Video of Parkinson's Disease Patients. *Journal of Parkinson's Disease*, *11*(3), 1271–1283. <https://doi.org/10.3233/JPD-212544>
- Slembrouck, M., Luong, H., Gerlo, J., Schütte, K., Van Cauwelaert, D., De Clercq, D., Vanwanseele, B., Veelaert, P., & Philips, W. (2020). *Multiview 3D Markerless Human*

- Pose Estimation from OpenPose Skeletons* (pp. 166–178). [https://doi.org/10.1007/978-3-030-40605-9\\_15](https://doi.org/10.1007/978-3-030-40605-9_15)
- Stavsky, M., Mor, O., Mastrolia, S. A., Greenbaum, S., Than, N. G., & Erez, O. (2017). Cerebral Palsy-Trends in Epidemiology and Recent Development in Prenatal Mechanisms of Disease, Treatment, and Prevention. *Frontiers in Pediatrics*, 5, 21. <https://doi.org/10.3389/fped.2017.00021>
- Stenum, J., Rossi, C., & Roemmich, R. T. (2021a). Two-dimensional video-based analysis of human gait using pose estimation. *PLoS Computational Biology*, 17(4). <https://doi.org/10.1371/journal.pcbi.1008935>
- Stenum, J., Rossi, C., & Roemmich, R. T. (2021b). Two-dimensional video-based analysis of human gait using pose estimation. *PLOS Computational Biology*, 17(4), e1008935. <https://doi.org/10.1371/journal.pcbi.1008935>
- Stöckel, T., Jacksteit, R., Behrens, M., Skripitz, R., Bader, R., & Mau-Moeller, A. (2015). The mental representation of the human gait in young and older adults. *Frontiers in Psychology*, 6, 943. <https://doi.org/10.3389/fpsyg.2015.00943>
- Surveillance of Cerebral Palsy in Europe. (2000). Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). *Developmental Medicine and Child Neurology*, 42(12), 816–824. <https://doi.org/10.1017/s0012162200001511>
- Sutherland, D. H., & Davids, J. R. (1993). Common gait abnormalities of the knee in cerebral palsy. *Clinical Orthopaedics and Related Research*, 288, 139–147.
- Takeda, I., Yamada, A., & Onodera, H. (2021). Artificial Intelligence-Assisted motion capture for medical applications: a comparative study between markerless and passive marker motion capture. *Computer Methods in Biomechanics and Biomedical Engineering*, 24(8), 864–873. <https://doi.org/10.1080/10255842.2020.1856372>
- Tsushima, H., Morris, M. E., & McGinley, J. (2003). Test-Retest Reliability and Inter-Tester Reliability of Kinematic Data from a Three-Dimensional Gait Analysis System. *Journal of the Japanese Physical Therapy Association*, 6(1), 9–17. <https://doi.org/10.1298/jjpta.6.9>
- Vafadar, S., Skalli, W., Bonnet-Lebrun, A., Khalifé, M., Renaudin, M., Hamza, A., & Gajny, L. (2021). A novel dataset and deep learning-based approach for marker-less motion capture during gait. *Gait and Posture*, 86, 70–76. <https://doi.org/10.1016/j.gaitpost.2021.03.003>

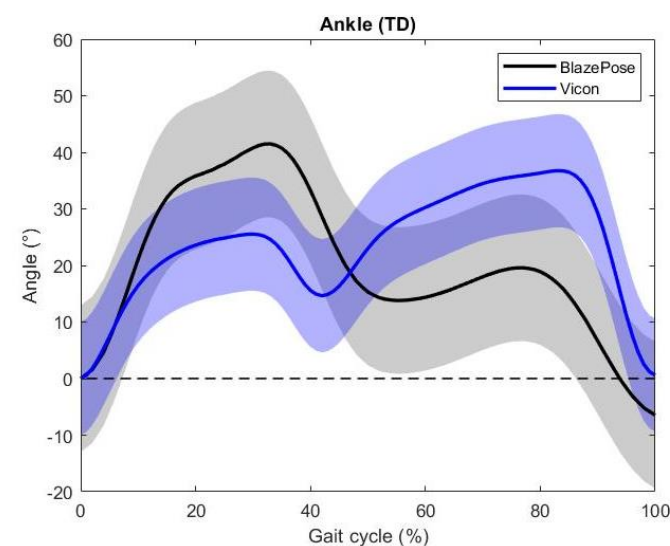
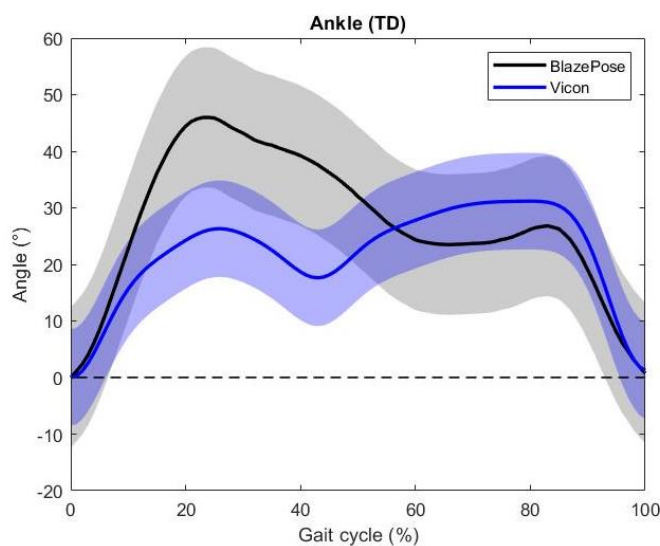
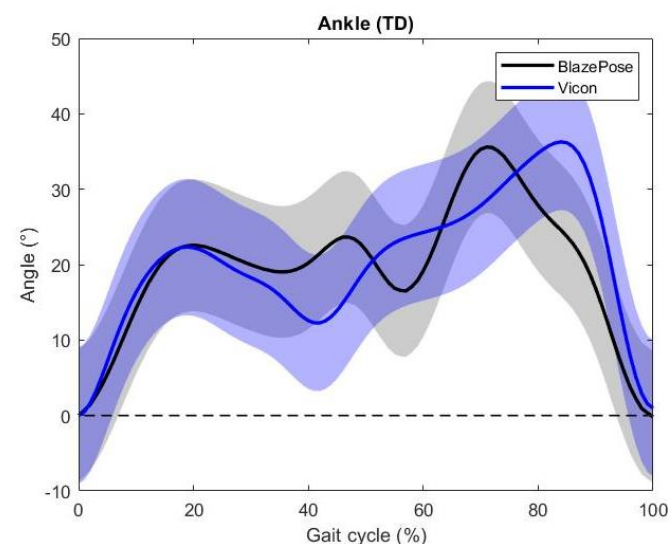
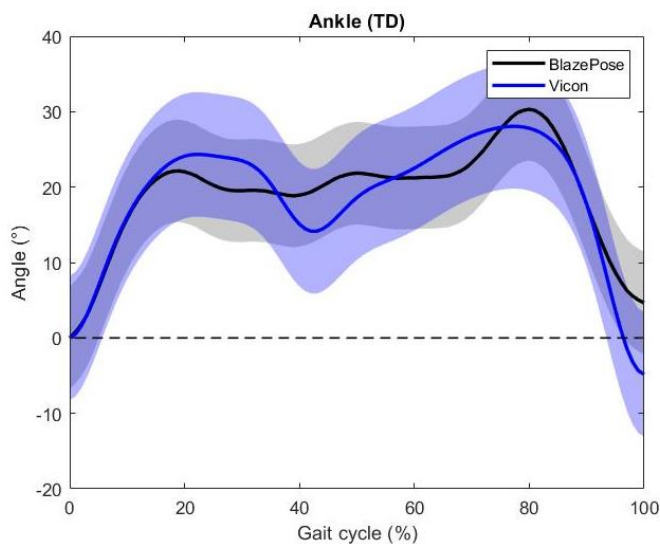
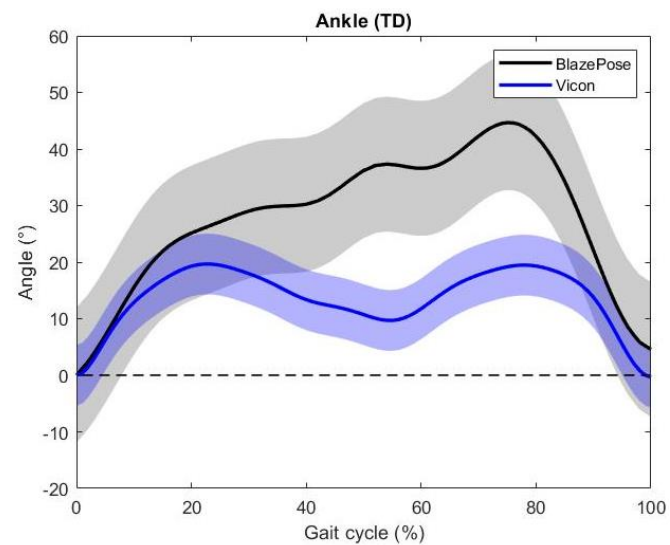
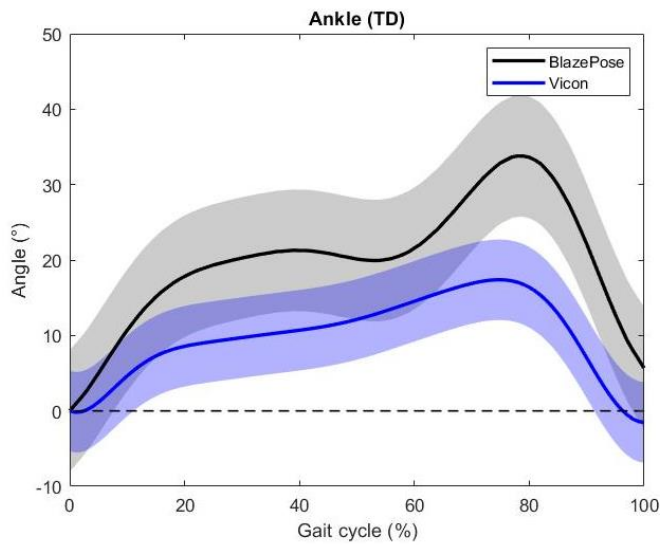
- van Schaik, J. E., & Dominici, N. (2020). Motion tracking in developmental research: Methods, considerations, and applications. In *Developmental motion tracking* (Vol. 254, pp. 89–111). Elsevier. <https://doi.org/10.1016/bs.pbr.2020.06.007>
- Wade, L., Needham, L., McGuigan, P., & Bilzon, J. (2022). Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. *PeerJ*, *10*, e12995. <https://doi.org/10.7717/peerj.12995>
- Wang, T.-M., Huang, H.-P., Li, J.-D., Hong, S.-W., Lo, W.-C., & Lu, T.-W. (2015). Leg and Joint Stiffness in Children with Spastic Diplegic Cerebral Palsy during Level Walking. *PLOS ONE*, *10*(12), e0143967. <https://doi.org/10.1371/journal.pone.0143967>
- Wimalasundera, N., & Stevenson, V. L. (2016). Cerebral palsy. *Practical Neurology*, *16*(3), 184–194. <https://doi.org/10.1136/practneurol-2015-001184>
- Winter, D. A. (2009). Biomechanics as an Interdiscipline. In *Biomechanics and Motor Control of Human Movement* (pp. 1–14). Wiley. <https://doi.org/10.1002/9780470549148>
- Winters, T. F., Gage, J. R., & Hicks, R. (1987). Gait patterns in spastic hemiplegia in children and young adults. *The Journal of Bone and Joint Surgery. American Volume*, *69*(3), 437–441.
- Wren, T. A. L., Elihu, K. J., Mansour, S., Rethlefsen, S. A., Ryan, D. D., Smith, M. L., & Kay, R. M. (2013). Differences in implementation of gait analysis recommendations based on affiliation with a gait laboratory. *Gait & Posture*, *37*(2), 206–209. <https://doi.org/10.1016/j.gaitpost.2012.07.008>
- Yamaguchi, G. T., & Zajac, F. E. (1989). A planar model of the knee joint to characterize the knee extensor mechanism. *Journal of Biomechanics*, *22*(1), 1–10. [https://doi.org/10.1016/0021-9290\(89\)90179-6](https://doi.org/10.1016/0021-9290(89)90179-6)
- Zago, M., Luzzago, M., Marangoni, T., De Cecco, M., Tarabini, M., & Galli, M. (2020). 3D Tracking of Human Motion Using Visual Skeletonization and Stereoscopic Vision. *Frontiers in Bioengineering and Biotechnology*, *8*. <https://doi.org/10.3389/fbioe.2020.00181>



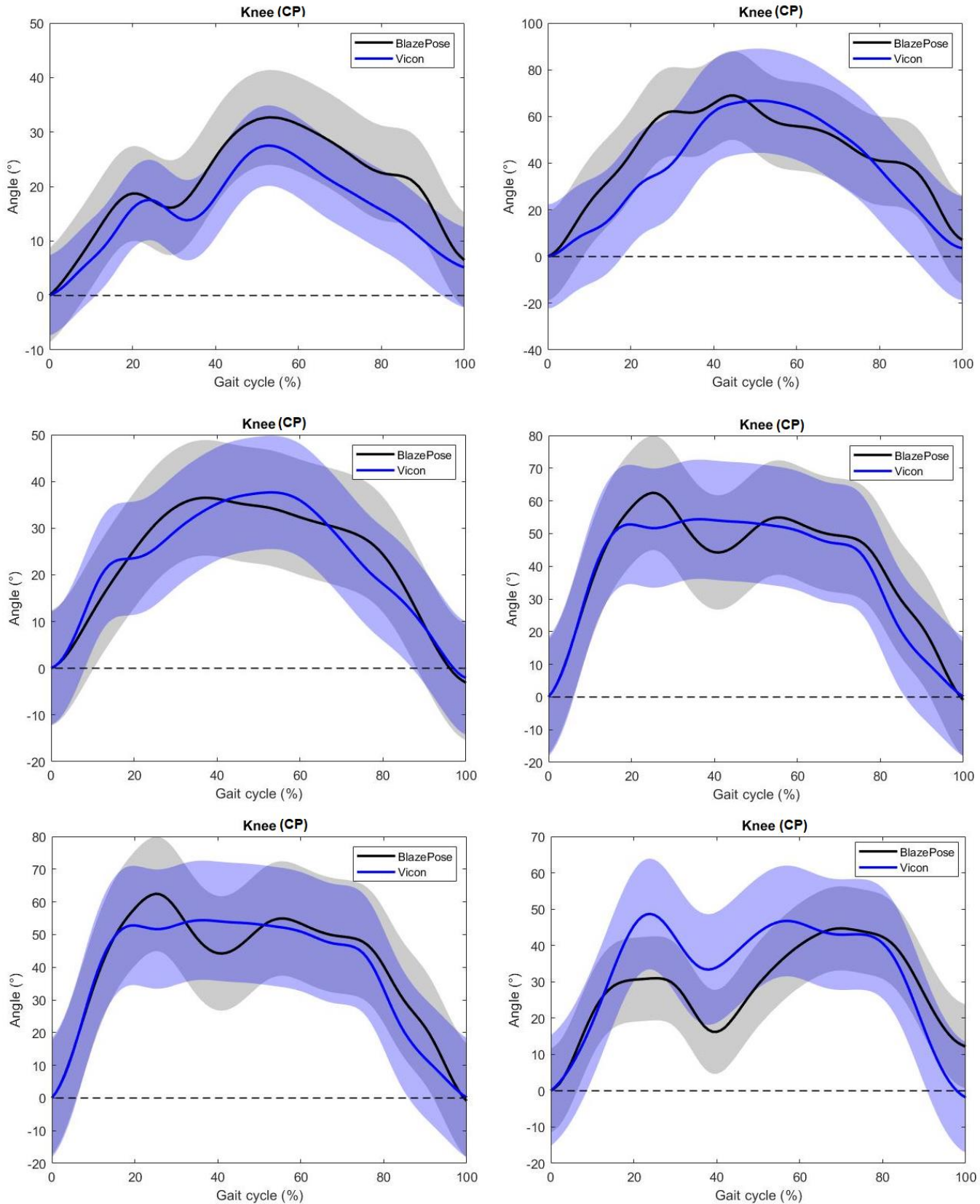
APPENDIX 1. Comparison of mean ankle joint angle (average of 5 trials per subject) for each CP participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.



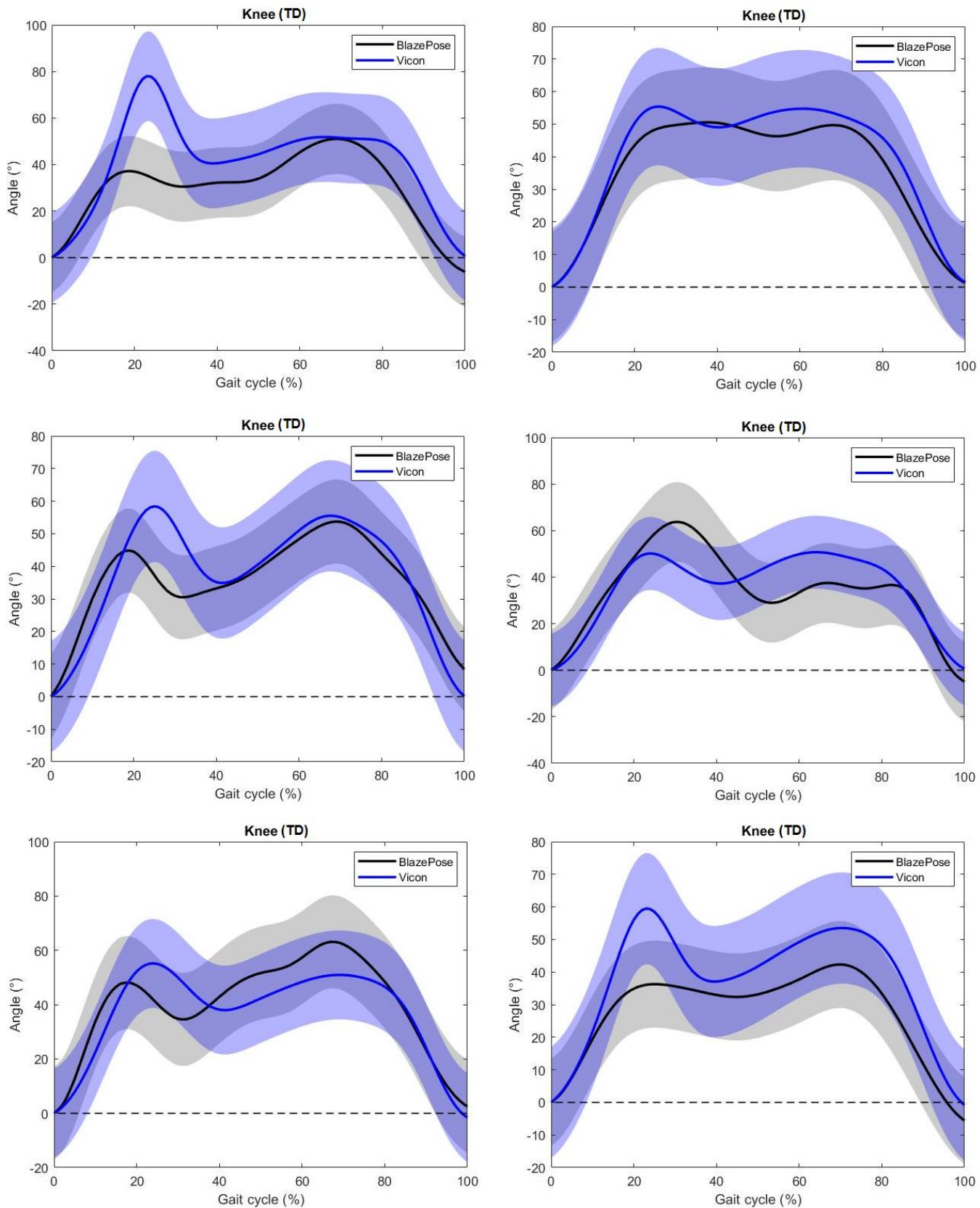
APPENDIX 2. Comparison of mean ankle joint angle (average of 5 trials per subject) for each TD participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.



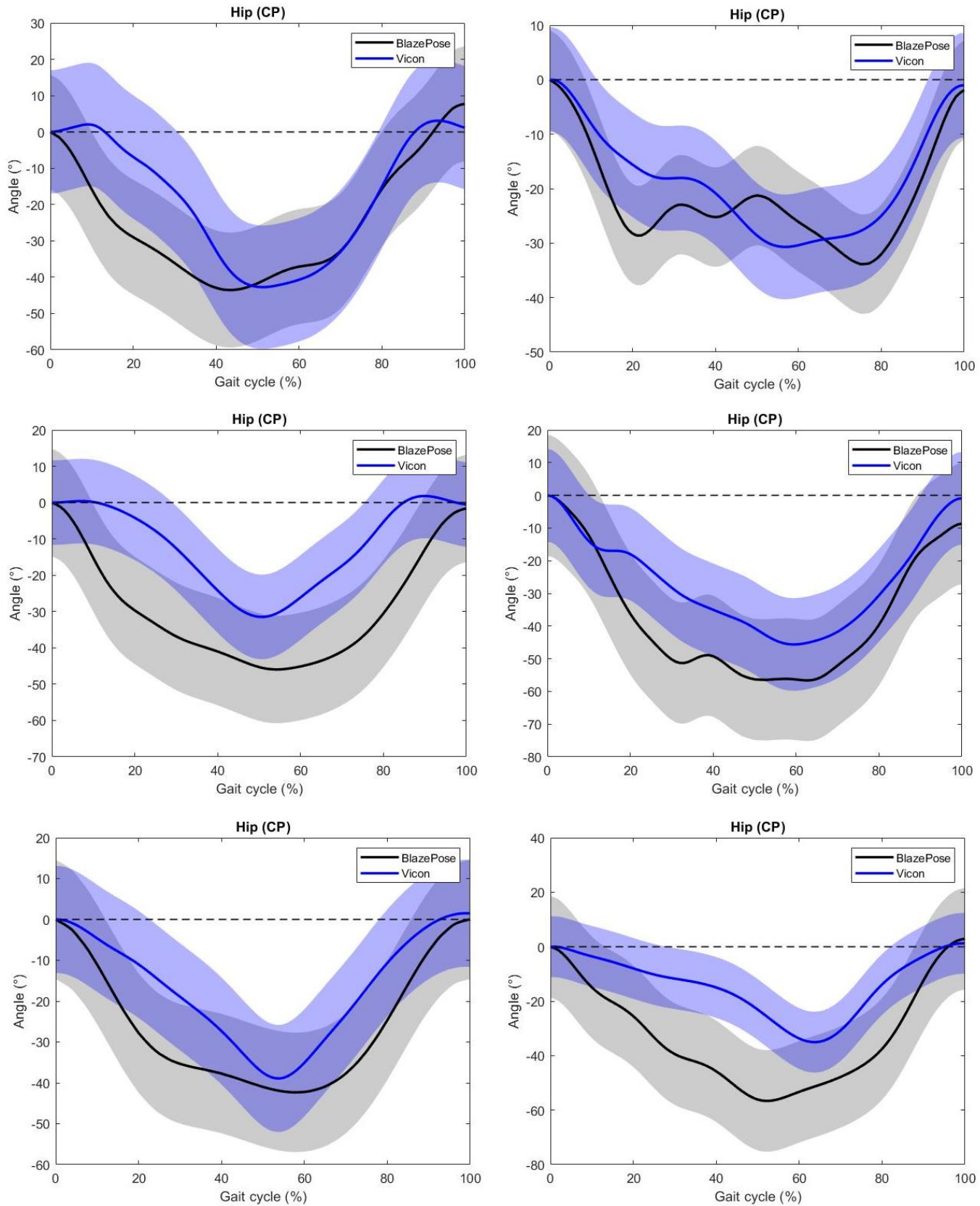
APPENDIX 3. Comparison of mean knee joint angle (average of 5 trials per subject) for each CP participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.



APPENDIX 4. Comparison of mean knee joint angle (average of 5 trials per subject) for each TD participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.



APPENDIX 5. Comparison of mean hip joint angle (average of 5 trials per subject) for each CP participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.



APPENDIX 6. Comparison of mean hip joint angle (average of 5 trials per subject) for each TD participant. The thick line represents the mean joint angles, and the shaded area represents 1 SD.

