Northumbria Research Link

Article

Internet-of-Things-Enabled Markerless Running Gait Assessment from a Single Smartphone Camera

Fraser Young [1], Rachel Mason [2], Rosie Morris [2], Samuel Stuart [2] and Alan Godfrey [1,*]

[1] Department of Computer and Information Sciences, Northumbria University, Newcastle-upon-Tyne NE1 8ST, UK

[2] Department of Health and Life Sciences, Northumbria University, Newcastle-upon-Tyne NE1 8ST, UK

[*] Correspondence: alan.godfrey@northumbria.ac.uk

Abstract: Running gait assessment is essential for the development of technical optimization strategies as well as to inform injury prevention and rehabilitation. Currently, running gait assessment relies on (i) visual assessment, exhibiting subjectivity and limited reliability, or (ii) use of instrumented approaches, which often carry high costs and can be intrusive due to the attachment of equipment to the body. Here, the use of an IoT-enabled markerless computer vision smartphone application based upon Google's pose estimation model BlazePose was evaluated for running gait assessment for use in low-resource settings. That human pose estimation architecture was used to extract contact time, swing time, step time, knee flexion angle, and foot strike location from a large cohort of runners. The gold-standard Vicon 3D motion capture system was used as a reference. The proposed approach performs robustly, demonstrating good (ICC(2,1) > 0.75) to excellent (ICC(2,1) > 0.90) agreement in all running gait outcomes. Additionally, temporal outcomes exhibit low mean error (0.01–0.014 s) in left foot outcomes. However, there are some discrepancies in right foot outcomes, due to occlusion. This study demonstrates that the proposed low-cost and markerless system provides accurate running gait assessment outcomes. The approach may help routine running gait assessment in low-resource environments.

Keywords: gait analysis; computer vision; deep learning; signal analysis; pose estimation; BlazePose; smartphone application

1. Introduction

Running for sport and exercise has continually grown in popularity due to its obvious health benefits and low (cost, skill, resource) barrier to entry [1]. As participation in running has increased, so too has the assessment and analysis of running gait to, e.g., avoid injury or improve efficiency. Running gait assessment generally encapsulates the study of lower-extremity kinematics [2]. Particularly, running gait assessment is paramount in providing injury prevention and rehabilitation mechanisms through quantifying the factors pertaining to the unique patterns of limb movement and co-ordination [3]. For example, runners exhibiting a rear-foot strike have been found to be at almost twice the risk of a repetitive strain injury in comparison to those with a fore or mid-foot strike [4]. As such, identifying and adapting small changes such as strike location to the running stride can minimize injury risk. Additionally, identifying shortcomings and adapting the running stride can lead to performance optimizations and overall better technique and speed. For example, adopting a forefoot strike location can minimize ground contact time, contributing to an increased race speed [5], and improve metabolic economy [6].

Traditionally, running gait assessment has relied upon manual, visual/video-based observation of treadmill running by a trained individual (e.g., bio-mechanist, sports therapist). However, there is some dispute as to the reliability of such approaches between assessors, often being reliant upon the observed plane and/or assessor experience [7].

Consequently, running gait assessment has moved towards instrumented approaches such as wearable technology including inertial measurement units (IMU), force/pressure plate analysis [8–10], as well as three-dimensional (3D) motion tracking [11,12] in an effort to provide reliable, reproducible outcomes. Despite research-grade wearable technology's utility in providing a wide range of gait outcomes, they are currently limited in use due to the cost [13], tethering of peripheral technologies, and reliance upon bespoke environments with expert assistance [9]. For example, wearable sensors or 3D-motion reflective markers often rely upon precise anatomical placement for optimal use [14,15]. Additionally, such technologies are naturally intrusive, often requiring adhesion to the skin or clothes by, e.g., medical tape [16], which may cause discomfort and thus, affect a natural running cycle; limiting the usability of such approaches for habitual use. Consequently, investigation into the use of markerless assessment systems for running gait assessment is warranted.

Markerless systems typically encapsulate human pose estimation methods from 2D or 3D video streams. For example, OpenPose [17] utilizes deep learning-based part affinity fields to identify anatomical locations from a 2D video stream, providing a digital skeleton of an individual performing a range of tasks. Alternatively, Microsoft's Kinect [18] utilizes depth sensors and infrared cameras to provide a 3D mapping of a human skeleton in near real-time. Although 3D approaches such as Kinect have demonstrated validity in certain spatiotemporal gait outcomes [19], their innate high cost (e.g., specific hardware rather than a generic camera) in comparison with 2D-based deep learning approaches limit their use much like wearable and marker-based tracking technologies. Accordingly, a 2D-based approach is generally preferred (and is the focus of this study).

To date, 2D markerless pose estimation has primarily seen research interest within normal walking/gait assessment. For example, markerless pose estimation has been investigated for use within clinical-based studies [20–22], where computer vision-based gait assessment could be used as a primary biomarker to assess disease onset and well-being. In the referenced studies, both spatial [20–22] and temporal [21] outcomes are evaluated with generally positive results. In contrast, there are significantly fewer developments for running gait with 2D pose estimation, although the use of OpenPose has demonstrated an ability to examine cadence [23]. Yet, within running gait, there are a significantly higher number of relevant outcomes to inform, e.g., performance optimization such as contact time, swing time, and knee flexion angle [24,25]. Accordingly, this warrants further exploration of the validity of markerless pose estimation within running gait. Furthermore, any contemporary approach to analyzing running gait should ensure flexibility to assess in any environment, beyond bespoke facilities to evoke natural running patterns [26]. Accordingly, reliance upon the OpenPose infrastructure, despite its reliability, requires significant computational cost (e.g., multi-GPU, high-powered computer) and so limits use within habitual or low-resource settings. In contrast, routine technologies (e.g., smartphones) may be more suitable. As such, low-powered 2D approaches must be investigated to enable running gait assessment beyond complex, research-grade environments.

BlazePose is a low-powered markerless pose estimation technique and has recently demonstrated the ability to run on inexpensive hardware (Pixel 2 smartphone) [27]. The approach has been shown to provide a tradeoff relationship with OpenPose, sacrificing some anatomical accuracy for a significant reduction in computational cost in clinical environments [28]. Consequently, the approach could augment assessments within low-resource settings through deployment on low-powered hardware such as a smartphone.

Considering approx. 90% of adults own a smartphone [29], and the technology is generally ubiquitous in everyday life. Accordingly, smartphones are providing a mechanism for remote healthcare, data capture, and transmission within the Internet of Things (IoT) [30] due to their relatively low computational power, connectivity and storage capabilities. Equally, a smartphone can capture data, while offloading complex computations to an external device (e.g., server). Consequently, smartphones could be used to augment

complex assessments that were previously reliant on expert in-person visits. Applied to running gait assessment, a smartphone was conceptualized as the data capture mechanism (via video) and then transmitted to a local edge device to enable a more accessible running gait assessment. As such, here, a single-camera approach to running gait assessment utilizing the low-powered BlazePose 2D pose estimation framework is proposed. The study aims to assess the validity of the low-powered approach, running on a custom smartphone application and low-cost server in comparison to a reference-standard 3D motion tracking system. Through validating the low-resource approach, the work aims to contribute to moving the field of running gait assessment out of bespoke facilities (e.g., lab) and into low-resource (e.g., everyday/habitual) environments by providing a simplified mechanism for running gait assessment (smartphone camera) in comparison to existing gold-standard approaches, that could conceivably be used, e.g., in the home, or at the gym.

2. Materials and Methods

2.1. Participants

Thirty-one healthy, experienced runners (34.5 ± 9.7 years; 1.75 ± 0.30 m/m; 76.2 ± 4.1 kg/kg; 20 male:11 female) were recruited from running clubs throughout the Northeast of England. Participants were screened for previous running-related injuries (RRI) and their ability to perform unsupervised, short treadmill-based running bouts. No participants reported any gait-affecting injuries or conditions that would adversely affect their ability to participate in the study and all had previous treadmill running experience.

Ethical approval for the study was granted by Northumbria University's Research Ethics Committee (reference: 21603). Prior to testing, participants were given informed consent and provided verbal and written consent following a short briefing. Upon consenting, participants were provided with a standardized, neutral cushioning running shoe (Saucony Guide Runner) to wear during testing in order to remove bias from gait-affecting cushioning within, e.g., support cushioning running shoes [31].

2.2. Video Capture

Participants were video recorded from the side and rear angles during treadmill running sessions utilizing two iPhone 13 smartphones, capturing at 240 frames per second (FPS) to provide slow-motion video streams for the application of 2D pose estimation. The smartphones were placed in separate static mounts approx. six feet from the left side and rear of the treadmill, standardizing the video capture sessions. Before running, participants were filmed performing a short static test, where they stood up straight with their hands to the side to establish a baseline reading. Once calibrated, participants performed up to five 1 min runs on a treadmill between 8 km/h (km/h) and 14 km/h, selected based upon a pace comparable to their most recent outdoor 5 km pace.

2.3. Reference System and Data Labelling

To provide a ground truth to benchmark the system performance against, a 14-camera 3D motion tracking system (Vicon Vertex, UK, www.vicon.com, accessed on 28 November 2022) was used, providing a high-resolution 3D skeletal mapping of participants during treadmill running. Participants were fitted with 16 neo-reflective markers, located at the calcaneal tuberosity (heel), lateral malleoli (ankle), base of the second metatarsal (front-foot/toe), lateral mid-shank, lateral knee joint line, mid-lateral thigh, anterior superior iliac spine, and posterior superior iliac spine. The Vicon 3D motion tracking system was configured to poll at 200 Hertz (Hz) to provide a suitable rate of detail for intricate gait assessment.

To ensure robust outcomes were obtained from the Vicon system, the built-in gait analysis suites were utilized, which have been used to provide a gold-standard set of gait outcomes for comparison with other sensing technologies such as infrared cameras [32]

and wearable technology [33]. The following gait outcomes were obtained from the Vicon system.

- Initial contact (IC)—the point at which the foot first contacts the ground.
- Final contact (FC)—the point at which the foot first leaves the ground.
- Contact time (CT)—the total time elapsed between IC and FC (i.e., time foot spent in contact with the ground.
- Swing time (ST)—the time elapsed between an FC event and a proceeding IC (i.e., time foot spent off the ground).
- Step time (StT)—the time elapsed between two IC events.
- Cadence—the number of steps taken per minute of running.
- Knee flexion angle—the angle between lateral mid-shank, lateral knee joint line, and mid-lateral thigh throughout a gait cycle.
- Foot strike location—the angle of the foot during contact with the ground during IC.

2.4. Proposed Low-Cost Approach

Typically, OpenPose [17], has been used for 2D computer vision-based running gait assessment but requires a complex GPU configuration due to reliance on convolution pose machines [34] and part affinity fields [35], which may create a barrier of entry to typical users through high associated computational costs. Furthermore, through utilizing part affinity fields, OpenPose performs multi-person pose detection, which creates additional complexities that are unnecessary within running gait assessment (for a single individual). Alternatively, Blazepose provides 33 anatomical key points for single-person videos, which include hips, knees, ankles, heels, and toes. That provides a suitable level of anatomical detail for understanding running gait biomechanics [36], while retaining a substantially lower computational cost by naively relying on identification of the head and its relative position to the body [27]. By utilizing the naïve approach, the BlazePose architecture could in turn augment gait assessment on low-cost hardware such as a smartphone or within an edge computing context, especially when considered against, e.g., OpenPose.

2.4.1. Proposed Infrastructure

To maximize the utility of the proposed 2D approach, a smartphone application was developed, interfacing with a Raspberry Pi server where pose estimation and analysis take place. By utilizing a smartphone in combination with IoT-based edge processing, the system could in turn be adapted for full-scale development/release, lowering the barriers of entry to running gait assessment. An overview of the infrastructure can be found in Figure 1, detailing the flow from a custom smartphone application to the Raspberry Pi-based server, where a cloud processing unit extracts gait outcomes for display on the smartphone application.
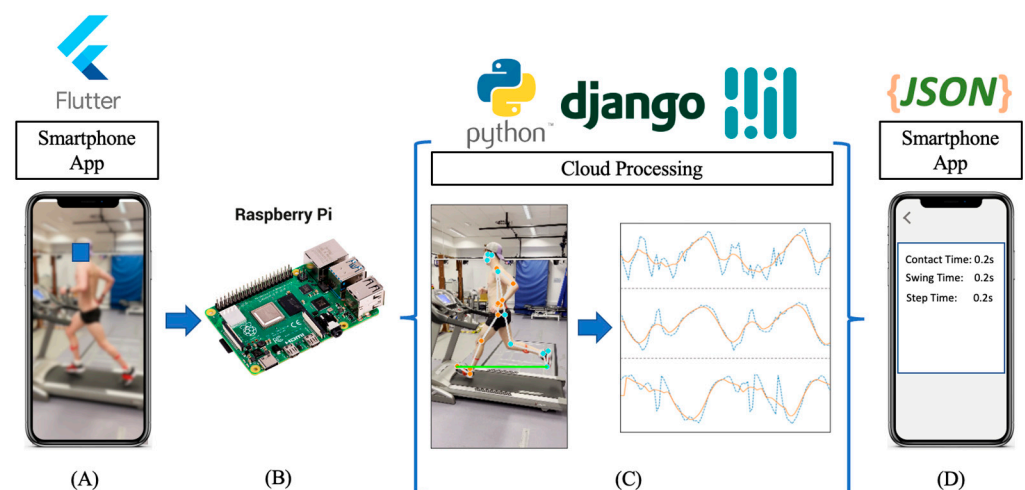
Figure 1. The infrastructure of the smartphone application and its cloud interface. Here, (**A**) the smartphone application captures and uploads side-view videos, transmitting to (**B**) a cloud processing unit running on a Raspberry Pi 4 where a Python Django implementation of Blazepose estimates and (**C**) stores the location of anatomical locations/keypoints, wherein signals are analyzed for gait outcomes. (**D**) Outcomes are then returned to the user via JSON for interpretation in the smartphone application.

2.4.2. Smartphone Application and Cloud Infrastructure

Here, a cross-platform (iOS, Android) smartphone application was developed with the Flutter 2.0 software development kit (SDK) wherein users can capture and upload 1 min (left) side-view videos of treadmill running. Selected videos are compressed using H.264 compression to optimize data transmission to a low-powered cloud infrastructure, wherein analysis takes place.

Despite Blazepose's ability to perform on-device pose estimation, the approach still requires reasonably modern hardware (e.g., iOS version 12+). As such, a low-powered, low-cost cloud infrastructure was developed utilizing a Raspberry Pi 4 (GBP ≈ £35) and Python's Django library to maximize the compatibility with a larger number of smartphone devices (Figure 1). The cloud infrastructure first receives videos from the smartphone application and temporarily stores them. Then, the video is run through a BlazePose instance within a Django server, wherein the X and Y locations of relevant anatomical key points of each frame are stored in a .CSV file. The generated .CSV file is then analyzed by the running gait feature extraction layer; wherein extracted outcomes are rendered as a JSON object for transmission to the smartphone application for interpretation, i.e., running assessment. Following a successful analysis, the video is deleted from the cloud storage to maintain user privacy.

2.5. *Feature Extraction*

The feature extraction layer of the proposed system was performed using libraries that are commonplace within Python's data science and computer vision capabilities (Figure 1). Particularly, OpenCV [37] handles the video stream playback and metadata, where SciPy, Pandas, and NumPy [38,39] were used in combination for data manipulation, signal analysis, statistics, and other mathematical operations detailed throughout the paper. The following section details how running gait features are extracted from the X and Y locations of anatomical keypoints utilizing signal analysis.

2.5.1. Data Preprocessing

Occasionally, keypoints identified by BlazePose can behave erratically and may be subject to noise that is not indicative of the real position of the anatomical location. Consequently, applying a preprocessing layer to the signals is paramount for accurate analysis. Here, a 5-step moving average is applied to each signal stream, optimally selected by manual observation of the signal, ensuring noise is reduced while minimizing loss of signal fidelity.

2.5.2. Gait Mechanics: Identifying Key Features

Initial contact and final contact are key biomarkers within gait as they inform a vast range of temporal and angle gait features. As such, quantifying IC and FC before other features will augment the rest of the gait feature extractions. Within a running stride, IC and FC typically proceed with a maximum extension of the leg [40], wherein the angle between the hip, knee, and ankle is at its maximum value (tending towards 180°), referred to here as the leg extension angle. To calculate the leg extension angle, locations are converted to vectors for vector scalar product analysis such that:

$$\overrightarrow{KH} = K - H \ and \ \overrightarrow{KA} = K - A \tag{1}$$

where *K* = the location of the knee, *H* = the location of the hip, and *A* = the location of the ankle extracted by the pose estimation layer. $\overrightarrow{KH}$ refers to the vector between knee and hip, whereas $\overrightarrow{KA}$ refers to the vector between knee and ankle, with *K* being the intersecting point of two vectors. Proceeding vector conversion, the scalar product of the two intersecting vectors can be represented as:

$$\overrightarrow{KH}.\overrightarrow{KA} = \left|\overrightarrow{KA}\right|\left|\overrightarrow{KH}\right|cos\,(\theta) \tag{2}$$

Here, e.g., $\left|\overrightarrow{KH}\right|$ refers to the magnitude of the vector between the knee and hip such that:

$$\left|\overrightarrow{KH}\right| = \sqrt{(K_x - K_y)^2 + (H_x - H_y)^2} \tag{3}$$

and θ is the angle between the intersecting vectors. Consequently, the following expression can be solved for θ such that:

$$\theta = cos^{-1}\left(\frac{\overrightarrow{KH}.\overrightarrow{KA}}{\left|\overrightarrow{KA}\right|\left|\overrightarrow{KH}\right|}\right) \tag{4}$$

With the angle between the hip, knee, and ankle calculated, IC and FC are calculated in the following manner:

1. A dynamic threshold is set at the 90th percentile of the maximum leg extension angle within the signal.
2. A zero-crossing gradient maxima peak detection algorithm such as those found within other gait assessment applications, e.g., [41] detects two peaks above the dynamic threshold, IC and FC.
3. IC and FC are then distinguished apart by observing the minima of the signal prior to the identified peak. Before an IC, the signal will dip significantly lower than the signal prior to FC due to minimum leg extension (i.e., lowest angle between hip, knee, and ankle; highest flexion) during the swing phase of gait, opposed to a maximum extension during contact [40] (Figure 2).
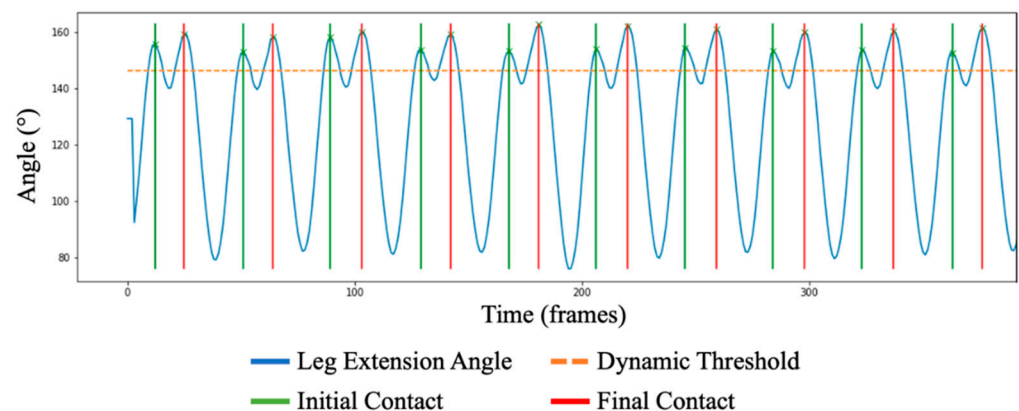

Figure 2. The identification of IC and FC, quantified by observing peaks within the leg extension angle data stream. Here, the dynamic threshold has been set at 90% of the maximum observed angle, wherein peaks are detected above the threshold. As seen, there is a significant angle reduction before an IC, with a smaller reduction before FC; providing distinguishable features between the two.

2.5.3. Temporal Outcomes and Cadence

Upon the identification of IC and FC; CT, ST and StT can be calculated such that:

$$\text{CT} = (FC_x - IC_x)/FPS \tag{5}$$

$$\text{ST} = (IC_{x+1} - FC_x)/FPS \tag{6}$$

$$StT = (IC_{x+1} - IC_x)/FPS \qquad (7)$$

where x refers to the currently observed stride (e.g., $x+1$ refers to the next stride). Values are then normalized from frames (*FPS*) into seconds for user-interpretable outcomes (Figure 3). Proceeding this, cadence is then quantified by first calculating the average number of steps per second $\times$ 60 s.
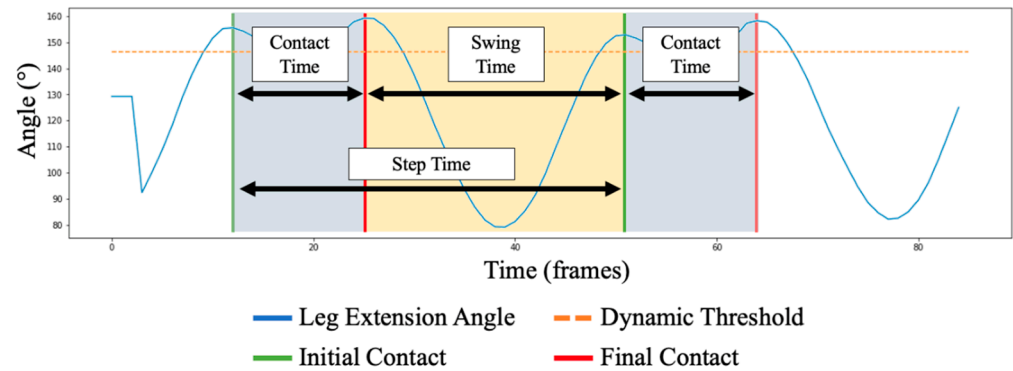

Figure 3. The following figure shows a magnified view of Figure 2. Here, an illustration of the calculation of temporal running gait outcomes is presented. CT is shown as the time between an IC and proceeding FC, ST is shown as the time between FC and a proceeding IC; whereas StT is the time between two respective IC events.

2.5.4. Knee Flexion Angle

Utilizing the leg extension angle calculation detailed as part of Section 2.5.2, the knee flexion angle can be calculated as a natural byproduct. Here, the knee flexion angle is recorded during an identified IC event, during an identified FC event, as well as the minimum value within a single step (i.e., within two ICs). Knee flexion angle is also stored continually throughout strides for observational comparison between estimated and ground truth angles.

2.5.5. Foot Strike Angle and Location

As previously described, foot strike location refers to the location of the foot that makes contact with the ground during IC, often referred to by rearfoot, midfoot, and forefoot striking patterns [42]. As such, calculating the angle of the foot is necessary. The angle of the foot is calculated between the toe, heel, and a point placed directly vertical from the heel (vH), where the angle is calculated using the vector scalar product as above (Section 2.5.2), with hip, knee, and ankle replaced with toe, heel, and vH, respectively.

During the calibration period in data capture, wherein participants stood still for a short time before running, a baseline reading of the foot angle is taken to account for any offset experienced during video capture (e.g., horizontal recording may not be perfectly aligned with treadmill, therefore it cannot be assumed that the resting foot lies at 90°). After calibration, the foot strike angle is then recorded during an identified IC event. An average angle is taken throughout all identified strikes and compared against the baseline reading. Foot strike location is then described as: >5° from baseline = forefoot; <−5° from baseline = rearfoot; and between −5° and 5° from baseline = midfoot strike patterns.

2.6. Statistical Analysis

The statistical analysis examining the performance of the proposed approach was performed in SPSS v27. Shapiro–Wilk tests indicated a parametric distribution for CT, ST, knee flexion angle, and foot strike angle; and non-parametric distribution for StT. As such, a $\log_{10}$ transformation was applied to non-parametric data to allow for Pearson's correlation scores to be used throughout the outcomes. Additionally, to examine the reliability

between 3D motion tracking and pose estimation approaches, intra-class correlation ($ICC_{(2,1)}$) was deployed, examining the absolute agreement between the measures. $ICC_{(2,1)}$ performance was defined in line with the accepted performance guideline such that the result can be described as poor (<0.500), moderate (0.500–0.750), good (0.750–0.900), and excellent (>0.900) [43]. Additionally, mean error, Bland–Altman, and box plot analysis were used to provide a visual assessment of the performance of the system in comparison to ground truth 3D motion tracking data.

3. Results

No data dropout or loss occurred throughout the data capture process, resulting in the full range (n = 31) participants being utilized in the results. In total, 148 running bouts were captured (between 4 and 5 running bouts per participant), observing 9327 unique strides during the proposed study. (Here, accuracies for all runs are presented, see Supplementary Materials for results breakdown across all running speeds.)

3.1. Temporal Outcomes

For all running speeds, intraclass correlation demonstrates good ($ICC_{(2,1)} \geq 0.751$) agreement across CT, ST, and SwT features. This indicates results obtained from 2D pose estimation are strongly indicative of measures obtained from the gold standard 3D motion tracking approach. Additionally, the features observe relatively low mean error (left foot 0.011–0.014 s; right foot 0.014–0.033 s) from 3D motion tracking-produced outcomes (Table 1). Figures 4 and 5 containing Bland–Altman and box plots (respectively) illustrate the performance of each measure. Accuracy of pose estimation outcomes is almost unanimously higher within the left foot in opposition to the right, with right foot swing time and step time measures tending to overestimate outcomes.
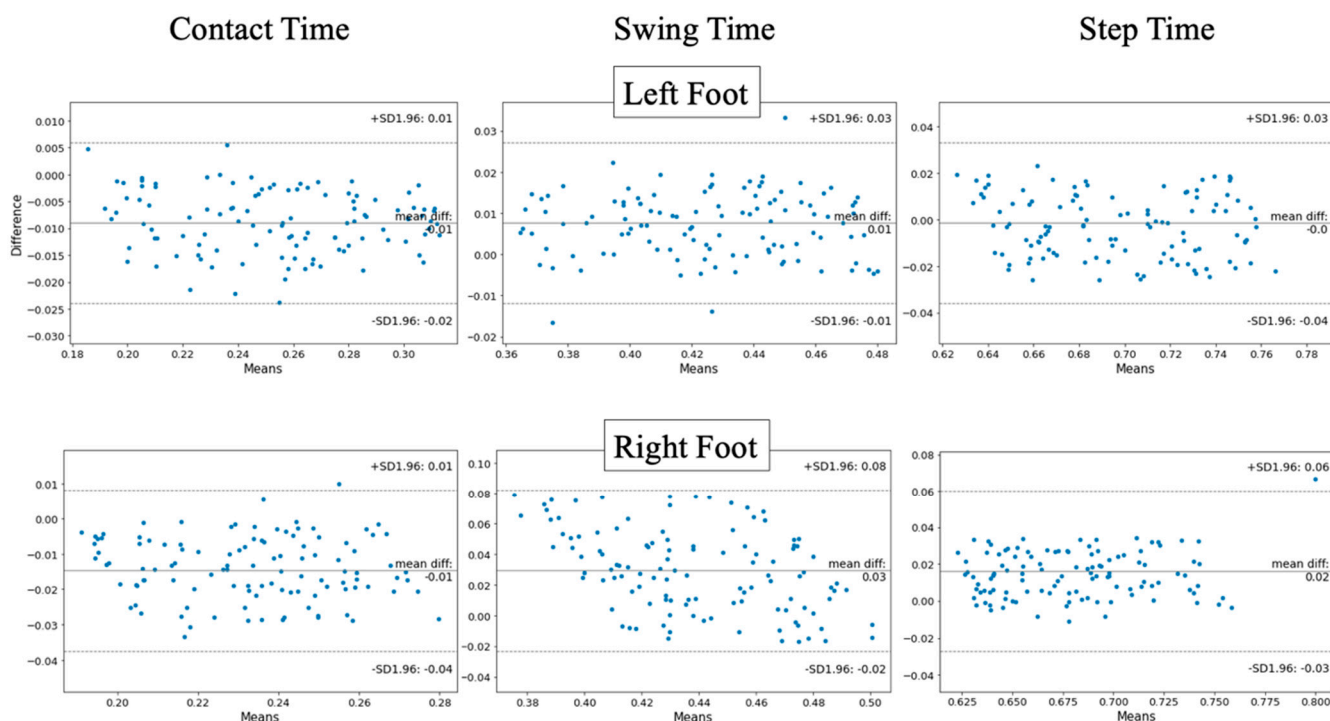

Figure 4. Bland–Altman plots illustrating the mean error between BlazePose estimation and reference standard 3D motion tracking data streams for contact time, swing time, and step time between left and right feet.
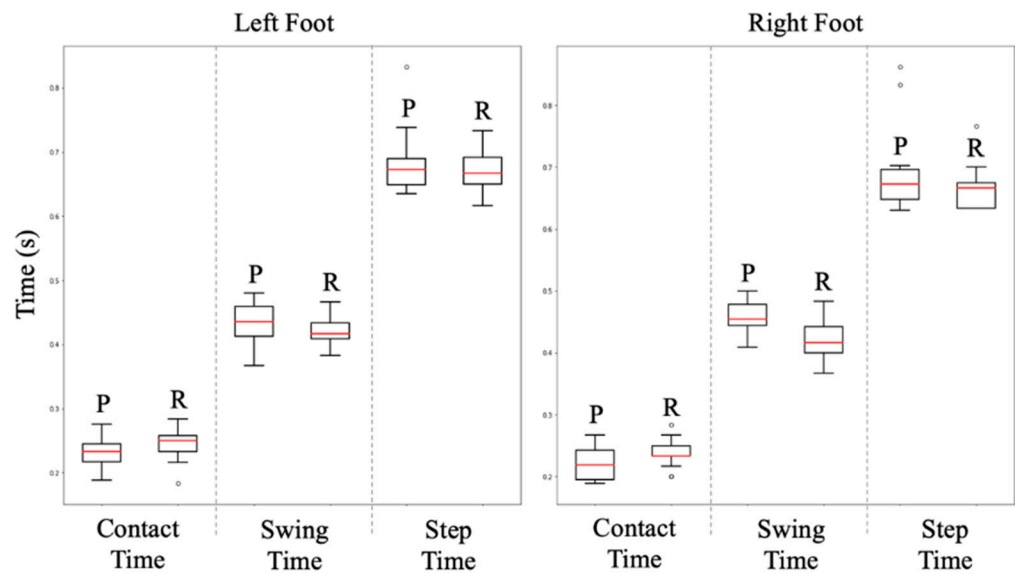
Figure 5. Box plots illustrating the temporal difference between BlazePose predicted (**P**) and reference (**R**) data streams for contact time, swing time, and step time at all speeds, observing the differences between left and right foot outcomes.

Table 1. Merged intraclass correlations (ICC$_{(2,1)}$), Pearson's correlation, mean and mean error scores for the range of outcomes for all running speeds obtained from the proposed pose estimation (predicted) approach in comparison with 3D motion tracking reference data. In addition, min/max predicted and reference knee flexion angle and foot strike location are presented.

	Left Foot				
Outcome	**Mean Predicted**	**Mean Reference**	**Mean Error**	**ICC$_{(2,1)}$**	***r***
Contact Time (s)	0.232	0.243	0.011	0.862	0.858
Swing Time (s)	0.434	0.420	0.014	0.837	0.883
Step Time (s)	0.682	0.671	0.010	0.811	0.845
	Min Predicted	**Max Predicted**	**Min Reference**	**Max Reference**	**ICC$_{(2,1)}$**	***r***
Knee Flexion (°)	19.9	98.9	1.4	103.1	0.961	0.975
Foot Strike Loc. (°)	−86.2	5.6	−96.7	20.3	0.981	0.980
	Right Foot				
Outcome	**Mean Predicted**	**Mean Reference**	**Mean Error**	**ICC$_{(2,1)}$**	***r***
Contact Time (s)	0.220	0.239	0.014	0.861	0.854
Swing Time (s)	0.457	0.423	0.033	0.821	0.781
Step Time (s)	0.689	0.665	0.024	0.751	0.769
	Min Predicted	**Max Predicted**	**Min Reference**	**Max Reference**	**ICC$_{(2,1)}$**	***r***
Knee Flexion (°)	30.4	91.1	12.5	97.1	0.979	0.982
Foot Strike Loc. (°)	−124.0	−51.0	−114.4	3.3	0.844	0.911

3.2. Cadence

The proposed approach to cadence identification demonstrates excellent intraclass correlation scores (ICC$_{(2,1)}$ = 0.981) across the range of tested running bouts. With a mean error rate of 1.2 steps, pose estimation can accurately estimate cadence in comparison to reference tracking data.

3.3. Knee Flexion Angle

Here, the performance score is derived from the average measure experienced across an accumulation of all strides. $ICC_{(2,1)}$ scores indicate excellent (left leg $ICC_{(2,1)}$ = 0.961; right leg $ICC_{(2,1)}$ = 0.979) correlation between data streams obtained between pose estimation and 3D motion tracking outcomes, with no significant deviation between left and right legs. However, when observing the average obtained waveforms (Figure 6) there is a widening distance between the measures surrounding IC and FC events, wherein the pose estimation approach overestimates the knee angle at the point of maximum extension.
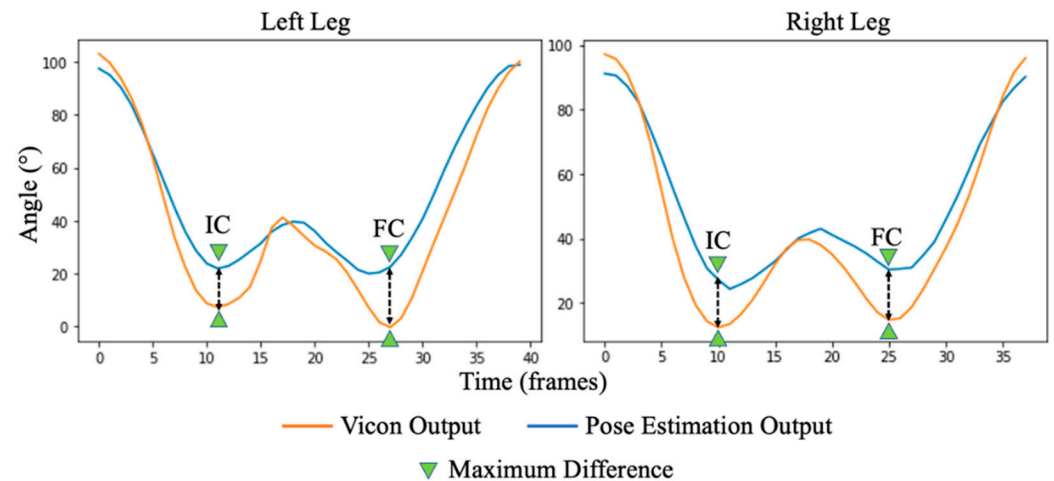

Figure 6. An illustration of the knee flexion angle of an average stride within the dataset between left and right legs. Here, it can be seen that pose estimation (orange) can closely match 3D motion tracking (blue) above ~40°; however, a maximum difference is exacerbated below ~40°, indicating pose estimation may overestimate knee flexion angles during maximum extension events (IC/FC).

3.4. Foot Angle and Foot Strike Location

Assessment of foot angle and strike location is separated to assess the performance of the pose estimation approach throughout both (i) the entire stride and (ii) at impact (i.e., providing foot strike location), providing a rearfoot, midfoot, or forefoot strike classification.

3.4.1. Foot Angle

Throughout the stride, outcomes demonstrate excellent ($ICC_{(2,1)}$ = 0.981) and good ($ICC_{(2,1)}$ = 0.844) correlation between pose estimation and 3D motion capture outcomes for left and right feet, respectively. Observing waveforms obtained from pose estimation and 3D motion capture systems and min/max values (Figure 7 and Table 1), the pose estimation approach consistently underestimates foot angle in both left and right feet. In both, and particularly within the right foot, a significant underestimation of foot angle is experienced as the angle tends towards positive. However, the foot angle is significantly closer surrounding the location of a typical IC event, where foot angle primarily affects gait.
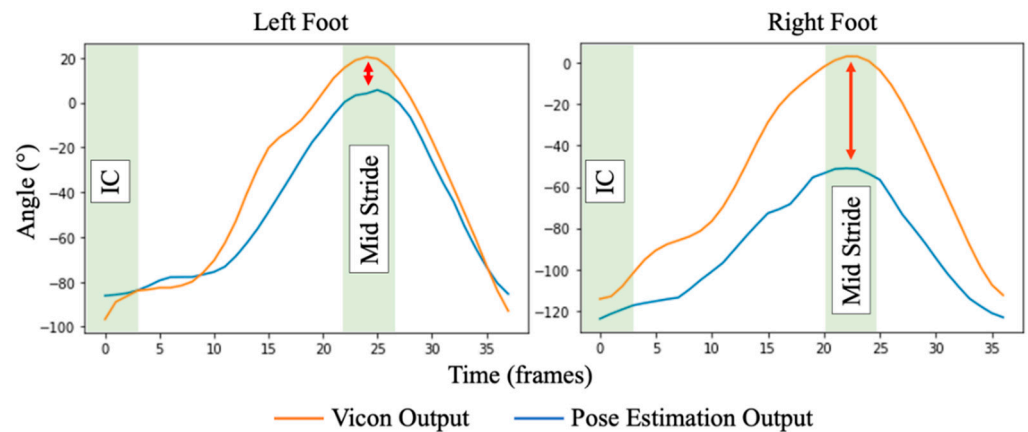
Figure 7. An illustration of the foot angle of an average stride within the dataset between left and right feet. Here, there is a significant deviation between left and right feet. In the left foot, signals are closely associated with surrounding initial contact events and begin to slightly deviate during the mid-stride. Alternatively, the right foot demonstrates a further deviance between pose estimation and 3D motion tracking signals, exacerbated during the mid-stride.

3.4.2. Foot Strike Location

Classifications were measured based on each participant's baseline resting foot angle, measured within video stream calibration. Accuracy of the foot strike location (i.e., forefoot, rearfoot and midfoot) demonstrated 95.2% and 89.8% in left and right feet, respectively. This indicates that despite deviance in foot angle throughout the stride, foot strike location can be accurately assessed utilizing BlazePose estimation.

4. Discussion

Understanding running gait is crucial in performance optimization and providing mechanisms for injury prevention and rehabilitation [4]. Providing low-cost mechanisms to measure running gait beyond bespoke settings is important to enabling accessibility and assessments for a wider range of individuals. This study has demonstrated the viability of a lightweight pose estimation, smartphone camera-based approach to provide running gait assessment.

4.1. Pose Estimation Performance

The single-camera BlazePose approach demonstrates high performance in both angle and temporal running gait outcomes in comparison to a Vicon 3D motion tracking reference system. Particularly, $ICC_{(2,1)}$ scores throughout the range of extracted outcomes indicate good to excellent agreement with reference streams; whereas temporal outcomes maintain a low mean error.

4.1.1. Temporal Outcomes

In comparison to similar studies within the field utilizing computer vision techniques, the proposed approach performs comparably within temporal gait outcomes, despite operating from a smartphone camera. For example, Garcia-Pinillos et al. [44] utilized the Microsoft Kinect system to extract contact and swing time, compared against both the OptoGait and high-speed video reference streams. In the study, the Kinect-based approach can extract temporal running gait features with an ICC of 0.712 and 0.894 in comparison to OptoGait; and 0.667 and 0.838 in comparison to high-speed video streams for contact and swing time, respectively. In comparison to the proposed approach exhibiting ICC scores of 0.862/0.861 (left/right) and 0.837/0.821 (left/right) for contact and swing time, 2D pose estimation can extract temporal gait outcomes comparably to approaches using dedicated hardware such as the Kinect.

There is a small effect on the performance of the temporal feature extraction layer between slow (8 km/h) and fast (14 km/h+) running speeds, with a slightly observably lower agreement and higher mean error experienced in higher speeds in comparison to those in lower speeds (Supplementary Materials). The higher mean error in faster speeds could be attributed to the video frame rate provided by the smartphone camera (60 FPS) not providing a suitable resolution to distinguish the naturally shorter temporal outcomes associated with faster running [45]. In fact, it has been shown that producing reliable gait outcomes at higher speeds (14.4 km/h+) will require up to 100FPS video capture streams [46]. Consequently, further investigation into the effects of frame rate (e.g., 120/240 FPS) on the pose estimation approach may be necessary to move into, e.g., elite running speeds, where the likes of contact time could be significantly shorter than those observed within this study.

Although the approach generally exhibits a low mean error from reference streams, there is an observably lower agreement in comparison to studies utilizing wearable technology such as IMU sensors. For example, a foot-mounted IMU has demonstrated an ability to extract contact and swing time with excellent (>0.9) agreement with reference streams [47]. The reduction in performance could be dependent on lighting. Particularly, fluctuations of light within a scene can greatly affect the pixel values, effectively creating noise within an image [48], which may hinder the performance of pose estimation frameworks. Additionally, wearable sensors can be configured to capture at a faster rate (e.g., 100 Hz–200 Hz [49]) which would provide considerably higher resolution for temporal gait outcomes, which can be relevant within 1/100 thousands of a second.

4.1.2. Knee Flexion and Foot Strike Angle

Observing waveforms and agreement measures of knee flexion and foot angle, the pose estimation approach closely follows the 3D motion tracking reference stream. Within foot angle, the pose estimation approach tends to underestimate the angle, whereas within knee flexion angle the inverse is true, with the pose estimation approach overestimating the angle during maximum extension periods. This is likely due to the camera placement relative to the treadmill. For example, in Figure 8, it is evident that the treadmill is not always straight within the sagittal plane, i.e., at 90° to the treadmill. As such, the angles obtained will naturally deviate in comparison to 3D motion tracking, which uses multiple infrared cameras to provide a more accurate angle estimation.
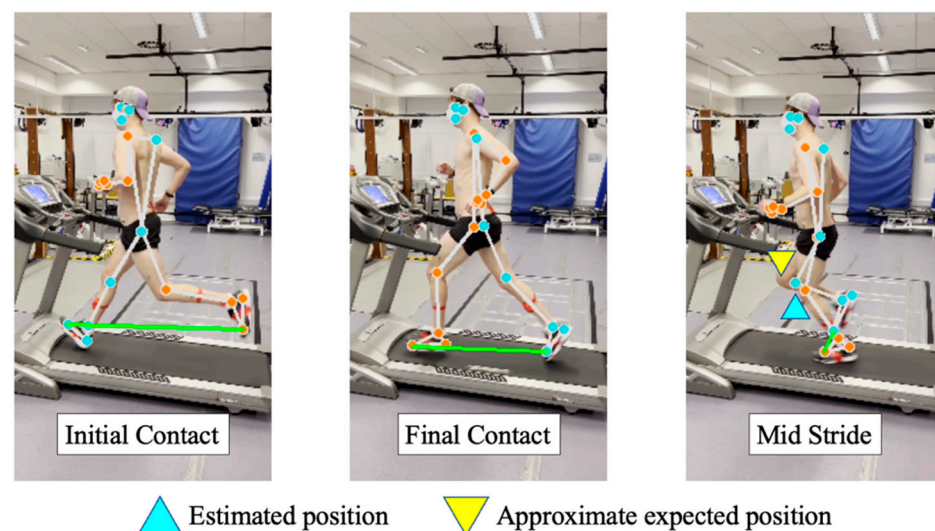

Figure 8. An example of pose estimation output at three different stages of the gait cycle, (i) initial contact, (ii) final contact, and (iii) mid-stride. As shown, during non-occluded segments of the gait cycle (IC/FC), estimated anatomical locations closely resemble their expected positions. However, observing mid-stride (during occlusion between left/right legs) one can begin to see the estimated

anatomical locations deviate from expected positions on the right leg (e.g., knee), providing insight behind lower accuracy experienced by right leg gait outcomes.

Prevalently within foot angle measures, there is a noticeable reduction in performance experienced between left and right measures (Table 1 and Figure 7). Upon investigation, there were bouts of noise experienced from the pose estimation output where the left leg (facing towards the camera) occludes the right leg (always facing away from the camera), an issue commonly found within pose estimation accuracy and validity measures [50]. Particularly, this provides an explanation for the significant deviance in the right foot angle surrounding the mid-stride, with the left and right legs intersecting throughout the mid-stride (Figure 8). Some studies have begun to address occlusion within the wider pose estimation domain with varying rates of accuracy [51–53]. However, the approaches generally rely on expensive computations (requiring complex hardware), which may limit the routine deployment of the system. As such, within the constraints of the study, higher accuracy may be achieved pragmatically by capturing videos from both sides (i.e., right leg at front) and running separate inferences, rather than the inclusion of a potentially expensive occlusion dampening approach.

Despite deviance between left and right feet, the pose estimation approach performs comparably with IMU-based approaches for foot strike location. For example, Young et al. [54] validated an IoT-enabled foot-mounted IMU polling at 60 Hz for the extraction of foot strike location during running gait at 94.3% accuracy. These results are comparable to those obtained within the proposed study, with 95.2% and 89.8% accuracy obtained through analyzing pose estimation data streams. The foot strike location accuracy is largely unaffected by deviation in the angle between estimated and reference streams due to simply observing the angular change surrounding impact. For example, rather than measuring the angle of the foot during impact (which may be inaccurate compared to reference streams), it is more pragmatic to assess the angular change between heel and toe, e.g., if the toe is moving down in comparison to the heel during impact, it is assumed a heel strike has occurred.

5. Limitations and Future Work

Currently, there is some discrepancy between left- and right-footed outcomes due to stationary camera location (left side only). This resulted in occlusion experienced when the left leg (front) obstructs the right leg (back). Future work will assess the approach utilizing combined videos from both sides (left and right foot forward) or a moving camera to examine whether a multi or panning video-based approach would improve the occluded discrepancies.

The data capture sessions took place in a well-lit, tailored laboratory which isn't necessarily indicative of habitual environments that may have, e.g., dimmer lighting or lower-quality smartphone camera equipment. As such, evaluating the systems performance in a range of low-resource environments (including outside) in comparison to, e.g., IMU-based outcomes, will help to validate the approach as a habitual running gait assessment tool.

Finally, spatial outcomes could be incorporated within the proposed system to provide greater utility within running gait assessment. For example, stride length has been shown to affect oxygen uptake [55] and shock attenuation [56] during running, and as such, can be optimized for running performance and injury mechanisms. Measuring stride length will require testing the system on an instrumented treadmill and during overground runs to assess the step/stride and total distance. Consequently, future work will aim to expand the functionality of the approach through use in overground settings, while expanding the suite of gait outcomes.

6. Conclusions

The proposed work demonstrates the efficacy of a low-powered pose estimation framework (BlazePose) to provide running gait outcomes within an IoT context. By operating from a smartphone interfacing with an edge device (Raspberry Pi-based server), the approach could contribute towards moving running gait assessment out of bespoke environments and into everyday/habitual settings. The approach exhibits good to excellent (ICC$_{(2,1)}$ 0.751–0.981) agreement for temporal outcomes, cadence, and knee angles compared with reference standard 3D motion tracking. Foot strike angle accuracy was %. Furthermore, the approach demonstrates low mean error with <0.014 s and <0.033 s for left and right streams, respectively. The approach may help routine running gait assessment in a low-resource environment.

References

1. Shipway, R.; Holloway, I. Running free: Embracing a healthy lifestyle through distance running. *Perspect. Public Health* **2010**, *130*, 270–276.
2. Dugan, S.A.; Bhat, K.P. Biomechanics and analysis of running gait. *Phys. Med. Rehabil. Clin.* **2005**, *16*, 603–621.
3. Agresta, C. Running Gait Assessment. In *Clinical Care of the Runner*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 61–73.
4. Daoud, A.I.; Geissler, G.J.; Wang, F.; Saretsky, J.; Daoud, Y.A.; Lieberman, D.E. Foot strike and injury rates in endurance runners: A retrospective study. *Med. Sci. Sport. Exerc.* **2012**, *44*, 1325–1334.
5. Hayes, P.; Caplan, N. Foot strike patterns and ground contact times during high-calibre middle-distance races. *J. Sport. Sci.* **2012**, *30*, 1275–1283.
6. ARDIGO', L.; Lafortuna, C.; Minetti, A.; Mognoni, P.; Saibene, F. Metabolic and mechanical aspects of foot landing type, forefoot and rearfoot strike, in human running. *Acta Physiol. Scand.* **1995**, *155*, 17–22.
7. Reinking, M.F.; Dugan, L.; Ripple, N.; Schleper, K.; Scholz, H.; Spadino, J.; Stahl, C.; McPoil, T.G. Reliability of two-dimensional video-based running gait analysis. *Int. J. Sport. Phys. Ther.* **2018**, *13*, 453.
8. Higginson, B.K. Methods of running gait analysis. *Curr. Sport. Med. Rep.* **2009**, *8*, 136–141.
9. Benson, L.C.; Räisänen, A.M.; Clermont, C.A.; Ferber, R. Is This the Real Life, or Is This Just Laboratory? A Scoping Review of IMU-Based Running Gait Analysis. *Sensors* **2022**, *22*, 1722.
10. Bailey, G.; Harle, R. Assessment of foot kinematics during steady state running using a foot-mounted IMU. *Procedia Eng.* **2014**, *72*, 32–37.
11. Zrenner, M.; Gradl, S.; Jensen, U.; Ullrich, M.; Eskofier, B.M. Comparison of different algorithms for calculating velocity and stride length in running using inertial measurement units. *Sensors* **2018**, *18*, 4194.
12. Albert, J.A.; Owolabi, V.; Gebel, A.; Brahms, C.M.; Granacher, U.; Arnrich, B. Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: A pilot study. *Sensors* **2020**, *20*, 5104.

13. Ye, M.; Yang, C.; Stankovic, V.; Stankovic, L.; Cheng, S. Gait phase classification for in-home gait assessment. In Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, China, 10–14 July 2017; pp. 1524–1529.

14. Anwary, A.R.; Yu, H.; Vassallo, M. Optimal foot location for placing wearable IMU sensors and automatic feature extraction for gait analysis. *IEEE Sens. J.* **2018**, *18*, 2555–2567.

15. Tan, T.; Strout, Z.A.; Shull, P.B. Accurate impact loading rate estimation during running via a subject-independent convolutional neural network model and optimal IMU placement. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 1215–1222.

16. Young, F.; Mason, R.; Wall, C.; Morris, R.; Stuart, S.; Godfrey, A. Examination of a Foot Mounted IMU-based Methodology for Running Gait Assessment. *Front. Sport. Act. Living* **2022**, *4*, 956889.

17. Cao, Z.; Simon, T.; Wei, S.-E.; Sheikh, Y. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7291–7299.

18. Preis, J.; Kessel, M.; Werner, M.; Linnhoff-Popien, C. Gait recognition with kinect. In Proceedings of the 1st International Workshop on Kinect in Pervasive Computing, New Castle, UK, 18–22 June 2012; pp. 1–4.

19. Springer, S.; Yogev Seligmann, G. Validity of the kinect for gait assessment: A focused review. *Sensors* **2016**, *16*, 194.

20. D'Antonio, E.; Taborri, J.; Palermo, E.; Rossi, S.; Patane, F. A markerless system for gait analysis based on OpenPose library. In Proceedings of the 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Dubrovnik, Croatia, 25–28 May 2020; pp. 1–6.

21. Stenum, J.; Rossi, C.; Roemmich, R.T. Two-dimensional video-based analysis of human gait using pose estimation. *PLoS Comput. Biol.* **2021**, *17*, e1008935.

22. Viswakumar, A.; Rajagopalan, V.; Ray, T.; Parimi, C. Human gait analysis using OpenPose. In Proceedings of the 2019 fifth international conference on image information processing (ICIIP), Shimla, India, 15–17 November 2019; pp. 310–314.

23. Tay, C.Z.; Lim, K.H.; Phang, J.T.S. Markerless gait estimation and tracking for postural assessment. *Multimed. Tools Appl.* **2022**, *81*, 12777–12794.

24. McCann, D.J.; Higginson, B.K. Training to maximize economy of motion in running gait. *Curr. Sport. Med. Rep.* **2008**, *7*, 158–162.

25. Moore, I.S.; Jones, A.M.; Dixon, S.J. Mechanisms for improved running economy in beginner runners. *Med. Sci. Sport. Exerc.* **2012**, *44*, 1756–1763.

26. Mason, R.; Pearson, L.T.; Barry, G.; Young, F.; Lennon, O.; Godfrey, A.; Stuart, S. Wearables for Running Gait Analysis: A Systematic Review. *Sport. Med.* **2022**, *53*, 241–268.

27. Bazarevsky, V.; Grishchenko, I.; Raveendran, K.; Zhu, T.; Zhang, F.; Grundmann, M. Blazepose: On-device real-time body pose tracking. *arXiv* **2020**, arXiv:2006.10204.

28. Mroz, S.; Baddour, N.; McGuirk, C.; Juneau, P.; Tu, A.; Cheung, K.; Lemaire, E. Comparing the Quality of Human Pose Estimation with BlazePose or OpenPose. In Proceedings of the 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), Paris, France, 8–10 December 2021; pp. 1–4.

29. Deloitte. Digital Consumer Trends: The UK Cut. Availabe online: https://www2.deloitte.com/uk/en/pages/technology-media-and-telecommunications/articles/digital-consumer-trends.html (accessed on 8 December 2022).

30. Gupta, A.; Chakraborty, C.; Gupta, B. Medical information processing using smartphone under IoT framework. In *Energy Conservation for IoT Devices*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 283–308.

31. Roca-Dols, A.; Losa-Iglesias, M.E.; Sánchez-Gómez, R.; Becerro-de-Bengoa-Vallejo, R.; López-López, D.; Rodríguez-Sanz, D.; Martínez-Jiménez, E.M.; Calvo-Lobo, C. Effect of the cushioning running shoes in ground contact time of phases of gait. *J. Mech. Behav. Biomed. Mater.* **2018**, *88*, 196–200.

32. Pfister, A.; West, A.M.; Bronner, S.; Noah, J.A. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *J. Med. Eng. Technol.* **2014**, *38*, 274–280.

33. Simoes, M.A. *Feasibility of Wearable Sensors to Determine Gait Parameters*; University of South Florida: Tampa, FL, USA, 2011.

34. Wei, S.-E.; Ramakrishna, V.; Kanade, T.; Sheikh, Y. Convolutional pose machines. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4724–4732.

35. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 172–186.

36. Nicola, T.L.; Jewison, D.J. The anatomy and biomechanics of running. *Clin. Sport. Med.* **2012**, *31*, 187–201.

37. Bradski, G.; Kaehler, A. OpenCV. *Dr. Dobb's J. Softw. Tools* **2000**, *3*, 120.

38. Bressert, E. *SciPy and NumPy: An Overview for Developers*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2012.

39. McKinney, W. pandas: A foundational Python library for data analysis and statistics. *Python High Perform. Sci. Comput.* **2011**, *14*, 1–9.

40. Dicharry, J. Kinematics and kinetics of gait: From lab to clinic. *Clin. Sport. Med.* **2010**, *29*, 347–364.

41. Trojaniello, D.; Ravaschio, A.; Hausdorff, J.M.; Cereatti, A. Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: Application to elderly, post-stroke, Parkinson's disease and Huntington's disease subjects. *Gait Posture* **2015**, *42*, 310–316.

42. Larson, P.; Higgins, E.; Kaminski, J.; Decker, T.; Preble, J.; Lyons, D.; McIntyre, K.; Normile, A. Foot strike patterns of recreational and sub-elite runners in a long-distance road race. *J. Sport. Sci.* **2011**, *29*, 1665–1673.

43. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163.

44. García-Pinillos, F.; Jaén-Carrillo, D.; Hermoso, V.S.; Román, P.L.; Delgado, P.; Martinez, C.; Carton, A.; Seruendo, L.R. Agreement Between Spatiotemporal Gait Parameters Measured by a Markerless Motion Capture System and Two Reference Systems—A Treadmill-Based Photoelectric Cell and High-Speed Video Analyses: Comparative Study. *JMIR Mhealth Uhealth* **2020**, *8*, e19498.
45. Di Michele, R.; Merni, F. The concurrent effects of strike pattern and ground-contact time on running economy. *J. Sci. Med. Sport* **2014**, *17*, 414–418.
46. Padulo, J.; Chamari, K.; Ardigò, L.P. Walking and running on treadmill: The standard criteria for kinematics studies. *Muscles Ligaments Tendons J.* **2014**, *4*, 159.
47. Young, F.; Stuart, S.; Morris, R.; Downs, C.; Coleman, M.; Godfrey, A. Validation of an inertial-based contact and swing time algorithm for running analysis from a foot mounted IoT enabled wearable. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Mexico City, Mexico, 1–5 November 2021; pp. 6818–6821.
48. Shirmohammadi, S.; Ferrero, A. Camera as the instrument: The rising trend of vision based measurement. *IEEE Instrum. Meas. Mag.* **2014**, *17*, 41–47.
49. Powell, D.; Stuart, S.; Godfrey, A. Investigating the use of an open source wearable as a tool to assess sports related concussion (SRC). *Physiotherapy* **2021**, *113*, e141–e142.
50. Sárándi, I.; Linder, T.; Arras, K.O.; Leibe, B. How robust is 3D human pose estimation to occlusion? *arXiv* **2018**, arXiv:1808.09316.
51. Tsai, Y.-S.; Hsu, L.-H.; Hsieh, Y.-Z.; Lin, S.-S. The real-time depth estimation for an occluded person based on a single image and OpenPose method. *Mathematics* **2020**, *8*, 1333.
52. Angelini, F.; Fu, Z.; Long, Y.; Shao, L.; Naqvi, S.M. 2D pose-based real-time human action recognition with occlusion-handling. *IEEE Trans. Multimed.* **2019**, *22*, 1433–1446.
53. Cheng, Y.; Yang, B.; Wang, B.; Tan, R.T. 3d human pose estimation using spatio-temporal networks with explicit occlusion training. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 10631–10638.
54. Young, F.; Coulby, G.; Watson, I.; Downs, C.; Stuart, S.; Godfrey, A. Just find it: The Mymo approach to recommend running shoes. *IEEE Access* **2020**, *8*, 109791–109800.
55. Cavanagh, P.R.; Williams, K.R. The effect of stride length variation on oxygen uptake during distance running. *Med. Sci. Sport. Exerc.* **1982**, *14*, 30–35.
56. Mercer, J.A.; Vance, J.; Hreljac, A.; Hamill, J. Relationship between shock attenuation and stride length during running at different velocities. *Eur. J. Appl. Physiol.* **2002**, *87*, 403–408.