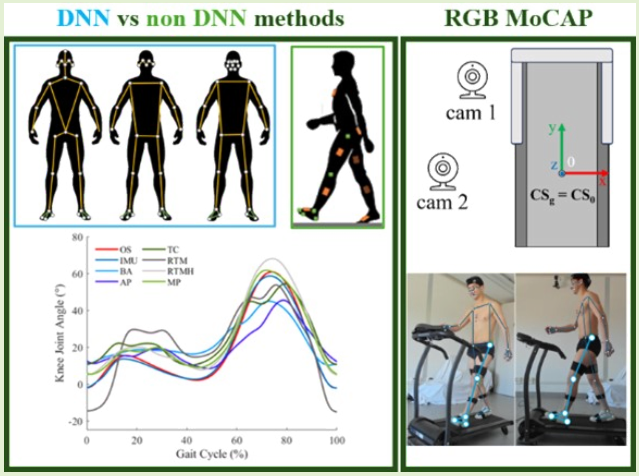
Comparative Analysis of Markerless Motion Capture Systems for Measuring Human Kinematics

Luca Ceriola, Juri Taborri, *Member, IEEE*, Marco Donati, Stefano Rossi, Fabrizio Patanè, and Ilaria Mileti

***Abstract*—To date, there are several measurement methods for evaluating human kinematics based on inertial sensors or vision systems. However, a comprehensive comparison has not been undertaken to determine which of these systems offers the most appropriate accuracy for clinical or sports evaluations. This study conducted a comparative analysis of different motion measurement systems: optoelectronic system (OS), inertial measurement units (IMUs), and vision-based methods, including deep neural network (DNN) and non-DNN approaches. Ten healthy subjects were involved, performing walking (W.) and running (R.) tests at various speeds (3.5, 5.0, and 7.0 km/h). The measurement of human kinematics was conducted by taking video images via two RGB cameras, together with an IMU-based system and an OS as the gold standard. Comparative analysis was conducted on a set of measurement methods, including IMU, a method based on blob analysis (BA), and DNN algorithms: Alphapose (AP), TC former (TC), RTMPose (RTM), and MediaPipe (MP). Data analysis involved triangulation and measurement of lower limb joint angles. Results showed that vision systems do not allow ankle joint measurement, and IMUs outperformed other methods in terms of RMSE and absolute error of range of motion (ε_{ROM}). RTM and MP exhibited results similar to IMUs, especially for the hip and knee joints, with the minimum absolute error reporting values of (3.1° ± 1.8°) and (3.5° ± 1.9°) for the hip joint and (4.0° ± 3.7°) and (4.8° ± 4.3°) for the knee joint, respectively.**



***Index Terms*— Comparative analysis, inertial measurement unit (IMU), kinematics measurements, markerless system, sensors.**

Manuscript received 21 May 2024; revised 12 July 2024; accepted 15 July 2024. Date of publication 26 July 2024; date of current version 1 September 2024. This work was supported by the Walking Over Balance-Board Learning Environment (WOBBLE) Project through Regione Lazio Legge Regionale 13/2008 Avviso pubblico "Intervento per il rafforzamento della ricerca e innovazione nel Lazio-incentivi per i dottorati innovativi per le imprese e per la PA." The associate editor coordinating the review of this article and approving it for publication was Dr. Sharmistha Bhadra. *(Corresponding author: Ilaria Mileti.)*

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

Luca Ceriola, Fabrizio Patanè, and Ilaria Mileti are with the Department of Engineering, University Niccolò Cusano, 00166 Rome, Italy (e-mail: luca.ceriola@unicusano.it; fabrizio.patane@unicusano.it; ilaria.mileti@unicusano.it).

Juri Taborri and Stefano Rossi are with the Department of Economics, Engineering, Society and Business Organization (DEIM), University of Tuscia, 01100 Viterbo, Italy (e-mail: juri.taborri@unitus.it; stefano.rossi@unitus.it).

Marco Donati is with Sensor Medica, Guidonia Montecelio, 00012 Rome, Italy (e-mail: marco.donati@sensormedica.com).

Digital Object Identifier 10.1109/JSEN.2024.3431873

Nomenclature

W.	Walking.
R.	Running.
DNN	Deep neural network.
BA	Blob analysis.
OS	Optoelectronic system.
IMU	Inertial measurement unit.
RMSE	Root mean square error.
TC	TC former method.
MP	MediaPipe method.
AP	Alphapose method.
RTMC	RTMPose method 1.
RTMH	RTMPose method 2.
ε_{ROM}	Absolute error of the range of motion.

I. Introduction

THE evaluation of human kinematics provides quantitative insights into motion behavior, which is fundamental for

clinical evaluation [1]. Identifying movement abnormalities and measuring their severity aid in the early detection of neuromuscular diseases, such as Parkinson's [2] and systemic sclerosis [3] or in the rehabilitation process of individuals with cerebral palsy [4]. The gold standard for the evaluation of human kinematics is the OS. It consists of a set of infrared cameras that track reflective markers attached to the subject's body segments according to specific biomechanical models, such as the plug-in gait [5]. Despite its high accuracy, this system has several drawbacks. It is costly, confines the assessments to indoor environments, and demands expertise for precise marker placement and postprocessing to assure the repeatability and reproducibility, and it is susceptible to soft tissue artifacts [6]. Currently, IMUs are validated tools [7] for estimating human kinematic parameters and can be a cost-effective alternative to OS. Following a specific setup and calibration, IMUs can compute 3-D joint angles of lower limbs [7], full-body kinematics [8], and perform human gait recognition [9]. Due to their wearability, they can be used for indoor or outdoor environments; however, the setup may be compromised in several ways. IMUs could suffer from drift, their orientation could be influenced by ferromagnetic disturbances when equipped with a magnetometer, and their use with noncooperative subjects is challenging. Other validated wearable systems that do not suffer from ferromagnetic disturbances, using FBG sensor [10] or piezoresistive sensors [11], [12], have been developed for joint angle estimation. A noninvasive alternative approach to estimate human kinematic parameters relies on feature detection and extraction using the image processing techniques. Among them, several research efforts have been devoted to the development and validation of marker-based tracking algorithms [13]. A thresholding technique, known as BA, discerns the 2-D coordinates of human joints by employing a color filtering approach, utilizing color markers with an RGB camera positioned in the subject's sagittal plane [14]. However, this classical processing technique was overtaken by the advent of DNN [15]. The potential of these markerless video-based systems is promising [16], as they are inexpensive, noninvasive, are not influenced by soft tissue artifacts, and require simple setups. However, they rely on two crucial aspects: training datasets and pose estimators. Training datasets consist of a series of images/videos labeled according to a certain topology of human annotation. To date, several databases exist in the literature. Some of them present a limited number of annotations, whereas others report a complete skeleton labeled with foot annotations necessary for a quantitative assessment of human kinematics [17], [18], [19]. It has to be noted that these datasets have some inherent pitfalls [20], such as 1) misalignments and offsets during the labeling process, which includes manual and semi-manual annotations [21]; 2) they are often task-specific and provide poor pose prediction when applied to other tasks [22]; and 3) many of them are not open source, including datasets focused on indoor gait [23]. The pose estimators are DNNs trained on the previously mentioned datasets, and the estimated pose depends on the labeled skeleton model. Openpose [24] is one of the most widely whole body pose estimators used in the literature for indoor gait analysis [25]. Other pose estimators have recently been gaining popularity, such as Alphapose (AP) [18], Blazepose (BZ) [17], TC [26], and RTMPose (RTM) [27], as they were often trained on newer and larger datasets [17], [19]. Open-source initiatives, such as OpenMMlab [28] and AP [18], help to use and test these pose estimators in a real-world scenario by providing user-friendly toolboxes and pretrained pose estimators. To date, several studies have compared some of these pose estimators, both for the assessment of 2-D and 3-D kinematics. Regarding the 2-D kinematic estimation, Menychtas et al. [29] performed a comparative analysis of three methods, two DNN methods (MP, OpenPose), and a non-DNN manual annotation tool (Kinovea) with respect to an OS system. The DNN methods did not show consistent trajectories for the ankle dorsiflexion angle. Moreover, hip and knee sagittal joint angles exhibited for both DNN and non-DNN methods a notable offset in comparison to the output angles from the OS. Furthermore, Van Hooren et al. [30] compared a custom-trained DNN method (DeepLabCut) and Openpose, with respect to an OS. They focused on lower limb joints in treadmill R. tests, particularly on the sagittal plane. After removing the offset between DNN and OS, the results of DeepLaabCut showed RMSEs for hip and knee joint angles of ($5.1° \pm 2.5°$) and ($7.9° \pm 3.4°$), while for Openpose of ($4.9° \pm 2.2°$) and ($6.5° \pm 2.4°$), respectively. Using Openpose, Wade et al. [31] evaluated the joint kinematics of the lower limbs of the visible and occluded lower joints in the sagittal plane, reporting the errors of the range of motion (ROM) of ($-3.6° \pm 4.6°$) for the hip and ($1.5° \pm 4.1°$) for the knee for the visible side. For the 3-D kinematic evaluation, after obtaining 2-D poses from one or more views, a 3-D gait analysis can be performed using various methods [32]. Among them, direct linear triangulation (DLT) is one of the most commonly used in the literature. Following this approach, D'Antonio et al. [33] calculated 3-D joint coordinates in three different stereo camera configurations using Openpose as a backbone, and 3-D sagittal angles extracted were compared with those of the inertial sensors, showing absolute errors of ($1.6° \pm 0.3°$) in an optimal setup condition. Instead, Vafadar et al. [34] trained and evaluated a DNN method [32] on a private custom gait dataset (ENSAM [23]). Their results showed a relative error in the ROM of hip, knee flexo-extension, and ankle dorsiflexion angles of ($4.5° \pm 8.2°$), ($-3.0° \pm 2.7°$), and ($1.1° \pm 4.7°$), respectively. Despite the considerable interest in the development of such measurement methods, the limitations of vision-based methods remain unclear. It is unclear whether such measurement systems can be used for clinical evaluation where methods with high accuracies and errors less than 5° are required, as in the case of inertial systems [7]. In addition, understanding whether better accuracies are achieved with the use of DNN-based methods for 3-D kinematic evaluation than with non-DNN-based methods, such as BA, is still an open question, never explored in literature. The objective of this study is to compare vision-based methods using different DNN-based algorithms and non-DNN, along with a wearable system based on inertial sensors already validated for clinical use. Our purpose was to investigate the accuracy of such methodologies for the evaluation of human kinematics in comparison with an OS, used as the gold standard.

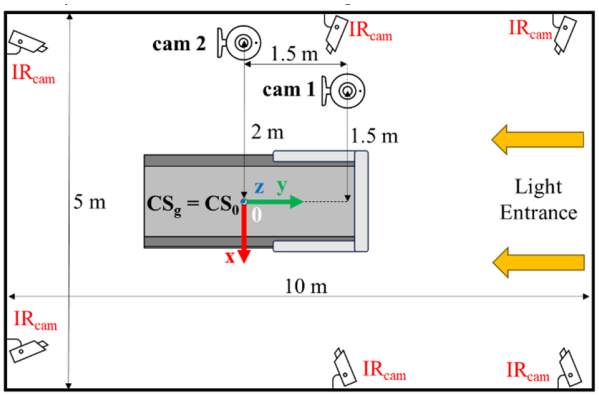


Fig. 1. Experimental setup with six infrared cameras (IR_{cam}), two RGB cameras (cam), and the treadmill.

II. Materials and Methods

A. Participants

Ten young and healthy male subjects (mean age ± standard deviation: 25 ± 4 years old and mean height ± standard deviation: 175 ± 7 cm) were involved in this study. None of the subjects wore orthopedic insoles, had any pre-existing neuromuscular or musculoskeletal conditions, or presented any intellectual, motor, or inner ear deficit. Subjects with joint pathologies, bones lesion, and orthopedic and/or neurological surgery history in the last three years were excluded. All participants presented normal vision with or without glasses and were able to walk and run independently. All participants provided their written consent to take part in the study and the protocol was in line with the guidelines outlined in the Declaration of Helsinki.

B. Experimental Design and Equipment

The experiments took place in the Laboratory of Mechanical Measurements and Experimental Biomechanics at the University of Tuscia, Viterbo, Italy. For the comparison of the measurement systems, a treadmill placed at the center of the laboratory was used, as shown in Fig. 1.

The gold standard was the OS Vicon Vero2.2, GPEM, Italy, equipped with six infrared cameras. As shown in Fig. 2(a), 39 reflective markers were attached to subjects according to the plug-in gait full-body model [5]. The 3-D trajectories of the reflective markers were recorded at 120 Hz. Video acquisition was carried out with two Logitech BRIO 4k Stream Edition webcams, positioned as shown in Fig. 1. Each camera was set with a resolution of 1920 × 1080 px and a frame rate of 60 Hz. All cameras pointed toward the origin of the laboratory reference frame CS_0. CS_0 was oriented as the reference frame CS_g of the OS, as depicted in Fig. 1. Synchronization between camera recordings was managed via OBS studio software. Six passive green markers were placed on the subject at the joint centers of the left limb according to [13], [14], and [35] and Table I and Fig. 2(b).

In addition, eight IMUs (MTws Xsens Technologies, Enschede, The Netherlands) were attached 1) to the sternum, between the two posterior iliac spines; 2) in midway between the pelvis and the center of the knee joint on the lateral side of the thighs; 3) to the lateral side of the lower legs aligned with the fibula above the lateral malleolus; and 4) above the midfoot with appropriate elastic bands, as shown in Fig. 2(a). The sampling rate of the IMUs was set to 60 Hz.

TABLE I
BA Markers Placement

Joint Centers	BA marker placement
Hip	Great Trochanter
Knee	Lateral epicondyle of the left knee
Ankle	Distal end of the left fibula
Heel	Left Calcaneus
1° met.	Over the first metatarsal head
5° met.	Left side of 5° metatarsal

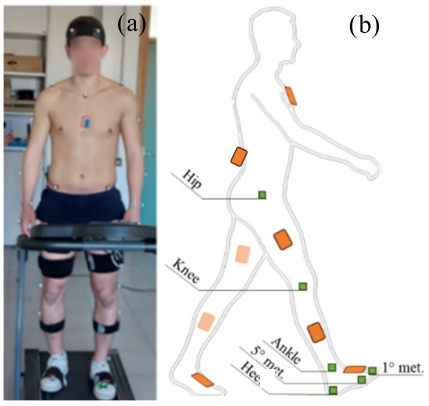


Fig. 2. (a) Complete sensors setup on the subject: 39 reflective markers, six passive green markers, and eight IMUs, including one on the chest. (b) Sagittal view of lower limb sensors setup: green squares represent the passive markers setup, instead orange rectangles represent the IMUs.

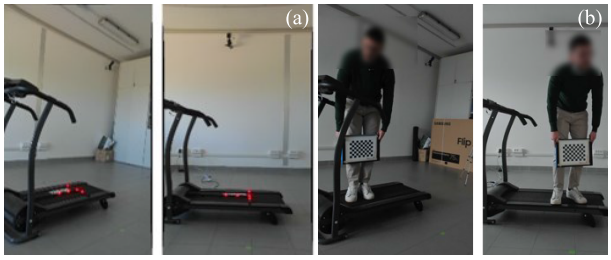


Fig. 3. Camera views during the calibration process of (a) OS origin and (b) RGB cameras.

C. Camera Calibration

For the origin calibration of the OS, the active wand was placed, as reported in Fig. 3(a). The same active wand placement was used to estimate CS_0 from each RGB camera. One stereo calibration was performed between the two cameras using the camera calibration tool of MATLAB (v.2022b, The Mathworks Inc., Natwick, USA): the Stereo Calibrator app (Computer Vision Toolbox 10.3) based on Zhang's [36] calibration algorithm. For the stereo calibration, a checkerboard (8 × 10) with squares 20 mm long [see Fig. 3(b)] was moved on the calibration volume. A total of 100 paired checkerboard images were captured and fed into the Stereo Calibrator app. Throughout the calibration process, stereo pairs with a reprojection error greater than 0.4 pixels (0.39 mm) were

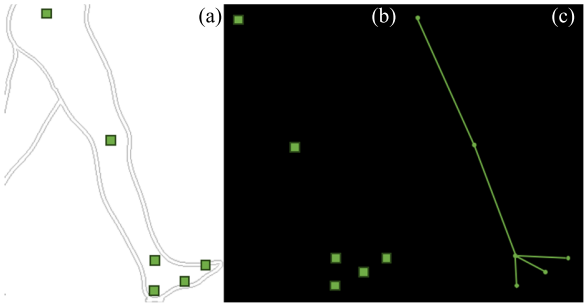


Fig. 4. (a) Input image, (b) green HSV color filter applied to the input frame, and (c) stick diagram based on blob detection.

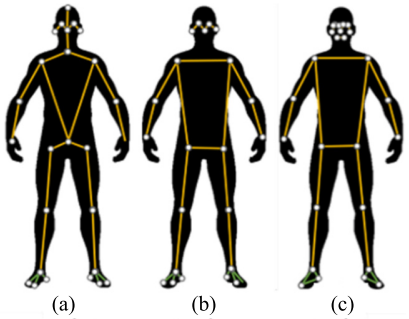


Fig. 5. Skeleton model composed by the 2-D keypoints for the video-based method DNN (a) Halpe full body, (b) COCO-whole body model, and (c) BZ model.

discarded. For the stereo calibration, the resulting average reprojection error was lower than 0.15 pixels (0.15 mm).

D. Experimental Protocol

Before the experimental session, each participant was asked to perform a static calibration and a functional calibration (FC) procedure for the body-to-IMUs alignment. The static calibration consisted in standing for 5 s at the center of the OS calibration volume, while maintaining a T pose. The FC consisted of a standing and sitting task, as reported in [7]. Participants were then asked to perform three different motor tasks for two minutes: 1) W. at 3.5 km/h (W3.5); 2) W. at 5.0 km/h (W5.0); and 3) R. at 7.0 km/h (R7.0). Two trials were performed for each motor condition. The order of tasks was randomized across subjects to prevent bias in results due to similar task sequences.

E. IMU-Based Method

For the analysis of the human joint angles, we used the biomechanical model described in [7]. The body-segment reference frames are defined as follows: the y-axis directed along the direction of progression, the z-axis vertically directed and pointing upward, and the yz-plane parallel to the sagittal plane. Joint angles were obtained by considering a Cardan-angle notation between consecutive body segments. Further details on the definition of joint axes and angles can be found in [7].

F. Video-Based Methods

Five approaches based on RGB cameras were used in this study. The first approach was marker-based and relied on a color threshold filter; along with BA, the other four were markerless and relied on DNN-based pose estimators. In particular, were tested: AP, TC, RTM, and MP. These pose estimators are based on a method called the "top–down" algorithm that uses a detector to locate the subject on a bounding box from which a pose estimator identifies the 2-D keypoints. As the choice of different detectors can affect the accuracy of pose estimation [19], a state-of-the-art pretrained detector YOLOX-X was used [37] for AP, TC, and RTM. In order to work properly, MPmethod needs, instead, to use its own detector.

1) Marker-Based Method: BA Method: A simple color threshold technique in the HSV color space was used to identify the six-green passive markers [see Fig. 4(a)]. All thresholds were set considering the room lighting and the scene background. After the HSV filter, the result was a monochrome image, with six different areas to be individuated on a black background [see Fig. 4(b)]. Then, the BA technique, validated in [14], was used to identify marker areas and determine their respective centers [see Fig. 4(c)]. For this method, the MATLAB app Color Threshold (Image Processing Toolbox 11.6) was used to create a mask color filter. BA has been carried out only for knee and ankle joints.

2) Markerless Method: AP Method: AP [18] offered many pretrained models and in this study, FastPose-DCN was used. ResNet-152 [38] is used as the "backbone," which acts as an encoder and feature extractor, and three dense upscaling convolution layers [39] are used to provide enhanced upscaling. The model was trained on the Halpe whole body dataset [18], which is composed of 50 000 images derived from the HICO-DET [40] dataset. The AP skeleton model is composed of 136 2-D keypoints: 20 were associated with the body, six with the feet, 42 with the hands, and 68 with the face. The HALPE 26 model annotation used in this study includes both body and foot annotations, as shown in Fig. 5(a).

3) Markerless Method: TC Method: TC [26] stands out as a token clustering transformer that employs a unique approach after the pose decoder. It divides the compressed image into a grid of tokens and sets the 2-D keypoints accordingly. Instead of utilizing deconvolution layers for upsampling, TC employs an attention transformer capable of combining tokens that are not pertinent to the task, such as the background, while preserving high resolution for tokens that contain high-level information. A pretrained TC model on the COCO-whole body dataset was used due to the OpenMMLab project that offers an open-source tool for pose estimation, MMPose. COCO-whole body dataset [26] is composed of 250k images that are labeled for each person with 133 2-D keypoints (17 for body, 6 for feet, 68 for face, and 42 for hands). Body and feet annotations were represented in Fig. 5(b).

4) Markerless Method: RTM Method: OpenMMLab also provides RTM [27]. With its open-source toolbox, such as MMpose and MMdeploy, RTM was deployed in the TensorRT framework to increase inference speed without losing accuracy. RTM predicted the 2-D keypoints using a SimCC [41]-based algorithm that treats keypoint localization as a dual classification task. This setup allows separate predictions for the x-axis and y-axis coordinates, effectively

determining the horizontal and vertical positions of the 2-D keypoints [47]. In this study, RTMPoseL, a pretrained model on both Halpe (RTMH) and COCO-Wholebody (RTMC) was selected.

5) Markerless Method: MP Method: Google offers MP as an open-source machine learning toolbox, and for pose estimation tasks, BZ is considered [17]. BZ follows a top–down approach (detector/tracker) for the first frame, while it switches to a bottom–up approach for subsequent frames, where only the tracker predicts the 2-D keypoints. BZ inference pipeline has also proposed a new person detector. Starting from a face detector, it encloses a person on a circular bounding box, like the Vitruvian man. BZ offers three pretrained models: lite, full, and heavy. The latter was chosen for its accuracy, although it requires more resource requirements. BZ is trained on a private Google dataset and consists of 85k images, 25k of which show fitness activity. The BZ skeleton model consists of 33 annotations [see Fig. 5(c)], with only four points dedicated to the feet (heels and first metatarsals). The training process is not public.

6) Data Analysis for Video-Based Methods: As the BA method did not estimate the centroids of the joint centers, but only the position of the markers on the skin, no triangulation was performed to compute the 3-D joint trajectories. The angles calculated by the BA method were therefore estimated using only 2-D images taken by the camera parallel to the sagittal plane. As regard the markerless methods, instead, they provide an estimation of the joint trajectories in two dimensions. The estimation of the 3-D coordinates was then carried out by means of stereo triangulation. MATLAB Camera Calibration Toolbox was used to evaluate the intrinsic and extrinsic parameters and to estimate the radial and tangential distortion parameters that minimized the reprojection errors. The accuracy of DLT can be affected by lens distortion [42]. All video footage was corrected by removing the lens distortion in postprocessing as reported in [43], in order to improve the accuracy of DLT [42]. Then, 2-D keypoints in pixels were extracted and triangulated using DLT. The resulting 3-D trajectories were finally and processed and expressed in terms of the global reference system. To extract lower limb joint angles from the skeleton models, thigh, shank, and foot coordinate systems were modeled. Their origins were set up on the pelvis, left knee, and left foot, respectively. Considering the global coordinate system CS_g, thigh, shank, and foot coordinate systems (CS_{th}, CS_{sh}, and CS_{ft}) were defined with the z-axis, as the line connecting the two joints of each body segment pointing upward. The x-axis was defined as the line parallel to x^g of the global reference system CS_g starting from their origins, whereas the y-axis has been defined as orthogonal to the third dimension, along the direction of progression [33]. Once computed CS_{th}, CS_{sh}, and CS_{ft}, the rotation matrices ${}^{pl}\boldsymbol{R}_{th}$, ${}^{th}\boldsymbol{R}_{sh}$, and ${}^{sh}\boldsymbol{R}_{ft}$ relative to the adjacent body segments were expressed as follows:

$$ {}^{pl}\boldsymbol{R}_{th} = \left({}^{g}\boldsymbol{R}_{pl}\right)^{T} {}^{g}\boldsymbol{R}_{th} \tag{1} $$

$$ {}^{th}\boldsymbol{R}_{sh} = \left({}^{g}\boldsymbol{R}_{th}\right)^{T} {}^{g}\boldsymbol{R}_{sh} \tag{2} $$

$$ {}^{sh}\boldsymbol{R}_{ft} = \left({}^{g}\boldsymbol{R}_{sh}\right)^{T} {}^{g}\boldsymbol{R}_{ft}. \tag{3} $$

The 3-D joint angles were computed from the previous matrices, considering the xyz Cardan sequence. A second-order Butterworth filter with a cutoff frequency of 10 Hz was used in the postprocessing for all methods [44]. All data were synchronized with the OS by means of an initial squat movement and through a cross correlation algorithm that used the sagittal knee angle as a reference. Using the 3-D trajectories of reflective heel markers from the OS, the heel strike events were identified by determining all minimum heel marker height values. All data were divided into gait cycles by considering two successive heel strike events. Then, joint angles related to each cycle were normalized at 100 samples.

G. Evaluation Parameters

In this study, the transversal and frontal planes were not investigated, because the analyzed video-based methods do not allow 3-D tracking of body segments, but only of joint centers. The joint angles of the lower limbs for all gait cycles were calculated discarding the first three and the last three gait cycles to avoid the acceleration and deceleration phases. Using the results of the OS as the ground truth, data from all other methods were compared. ROM, defined as the absolute difference between the maximum and the minimum of the joint angle, was calculated for each gait cycle. The RMSE was computed between the normalized joint angles for each n gait cycle, considering the OS as the reference system, along with ε_{ROM}. For the analysis of the RMSE, offset among tested methods and reference system was removed for both hip and knee joint angles. Offset was computed considering the mean value across each gait cycle

$$ \text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{\left(\hat{y}_i - y_i\right)^2}{n}} \tag{4} $$

$$ \varepsilon_{ROM} = |\text{ROM}_{OS} - \text{ROM}_x|. \tag{5} $$

All data were tested for normality with the Shapiro–Wilk test. Mean values across the gait cycle of RMSE and ε_{ROM} values were analyzed using two-way repeated measures ANOVA tests, with methods (seven levels) and tasks (three levels) as within-subject factors. When significant differences were found, a Bonferroni test was performed. The Greenhouse–Geisser correction was adopted when Mauchly's test was significant, and the assumption of sphericity was violated. Otherwise, the p-value of sphericity was considered. If the interaction effect methods $\times$ tasks was significant, the interactions were broken down, comparing methods and tasks separately. More specifically, the main effect methods was tested considering three one-way repeated measures ANOVA for each task, while the main effect task was tested considering seven one-way repeated measures ANOVA for each method. The significance level was set at 0.05 for all the tests.

III. Results

The analysis carried out on angles in the sagittal plane showed that algorithms based on vision techniques were not able to measure the ankle angle consistently. As illustrated in Fig. 6, the angles derived from video-based systems may not accurately represent the actual joint angles.

TABLE II
MEAN AND STANDARD DEVIATION OF THE ABSOLUTE RANGE-OF-MOTION ERROR (°) OF THE SAGITTAL PLANE ANGLE OF THE HIP JOINT AND KNEE JOINT FOR ALL METHODS AND ALL MOTOR TASKS. *** INDICATES $p < 0.01$ BETWEEN THE IMU METHOD AND THE VISION METHOD

		IMU	BA	AP	TC	RTMC	RTMH	MP
Hip	**W3.5**	2.8 ± 1.1	/	6.0 ± 2.9	5.2 ± 3.2	6.8 ± 3.5	3.8 ± 2.1	3.5 ± 1.9
	W5.0	3.8 ± 1.3	/	6.0 ±3.4	6.0 ± 4.1	7.4 ± 3.4	4.1 ± 3.7	5.1 ± 3.4
	R7.0	3.3 ± 1.3	/	5.9 ± 3.0	3.6 ± 2.5	6.3 ± 2.8	3.1 ± 1.8	3.8 ± 2.8
Knee	**W3.5**	3.7 ± 3.5	10.2 ± 4.4***	14.5 ± 6.7***	9.2 ± 3.5***	24.3 ± 8.9***	4.0 ± 3.7	5.6 ± 3.1***
	W5.0	3.3 ± 2.5	11.2 ± 4.3***	14.0 ± 4.4***	13.7 ± 10.5***	25.7 ± 9.0***	4.3 ± 3.5	5.1 ± 2.7***
	R7.0	3.4 ± 2.1	14.4 ± 6.8***	12.3 ± 4.2***	10.6 ± 4.5***	15.6 ± 7.1***	4.3 ± 2.6	4.8 ± 4.3***

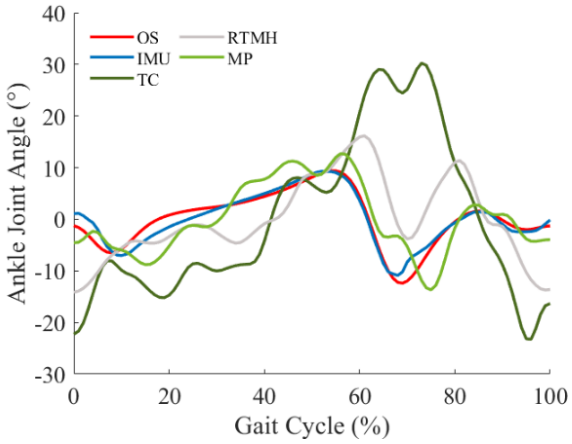


Fig. 6. Ankle joint angles measured by OS, IMU, MP, TC, and RTMH methods over a gait cycle in a W5.0 trial.

While it can be observed in the figure that the articular ranges for RTMH and MP are consistent with those of the reference system, their waveforms exhibit different trends. Therefore, it is inadvisable to employ these methodologies for the assessment of the kinematics of the ankle joint. Ankle joint kinematics could only be measured correctly by the IMU-based system, which reported the RMSE values of (2.1° ± 0.7°), (2.7° ± 2.0°), and (2.8° ± 1.3°) for W3.5, W5.0., and R7.0 tasks, respectively.

Fig. 7(a) and (b) illustrates the angles without offset removal in the sagittal plane of the hip and knee joint for the left limb, averaged over all subjects and trials for all methods employed in the execution of W5.0 task. Fig. 7(c) and (d) depicts the hip and knee angles in the sagittal plane, averaged over the subjects, without offset removal of the methods with less RMSE, i.e., IMU, MP, and RTMH methods, for W5.0 task. In these figures, a notable offset between the DNN methods and OS was observed, especially for the hip joint angle [see Fig. 7(a)]. A qualitative analysis revealed that the IMU-based method yielded better results in terms of waveform quality for both the hip and knee joints. Regarding vision systems, it has been observed that the BA method generated a flatter waveform than the reference method. Among other methods based on vision systems, the MP method and the RTMH method showed waveforms more similar to those of the OS reference system, as shown by the RMSE values in Fig. 7(e) and (f). It can also be observed that the RMSE values on the hip were lower than those on the knee for almost all methods and tasks. Table II presents the mean values and standard deviation of ε_{ROM} of the joint angles for all methods.

The lowest error value was observed for the IMU method for both the hip and knee joints. The highest error values were observed for the knee joint for the AP and RTMC methods, with errors of up to (14.5° ± 6.7°) for AP in the W3.5 task and (25.7° ± 9.0°) for RTMC in the W5.0 task. Although the BA vision method was not the one with the greatest inaccuracy, it reported considerable errors (14.4° ± 6.8°) for the knee joint during task R7.0. Video-based methods that exhibited lower absolute errors were the RTMH and the MP methods, with error values reaching (3.1° ± 1.8°) and (3.8° ± 2.8°) for the hip joint and and (4.0° ± 3.7°) and (4.8° ± 4.3°) for the knee joint, respectively.

A. Results of the Statistical Analysis

For the RMSE of the hip joint, the two-way repeated measures ANOVA reported a statistical difference for the main effect method ($p < 0.01$). No statistical differences were found for the task effect and the interaction effect. Considering the Bonferroni tests among different methods, the following pairs of methods resulted statistically different: IMU in comparison with AP ($p < 0.01$) and RTMC ($p < 0.01$) methods; AP in comparison with TC ($p < 0.01$), RTMH ($p < 0.01$), and MP ($p < 0.01$); TC in comparison with RTMC ($p < 0.01$); and RTMC in comparison RTMH ($p < 0.01$) and MP ($p < 0.01$).

For the RMSE of the knee joint, the interaction effect was statistically significant and the two-way repeated measures ANOVA was divided into three one-way repeated measures ANOVA for each task and seven one-way repeated measures ANOVA for each method. Considering separately the task, each one-way repeated measures ANOVA reports a statistical difference among main effect methods ($p < 0.01$). Considering the Bonferroni tests, statistical differences were found among the following pair of methods: IMU in comparison with BA for W3.5 ($p = 0.03$) and R7.0 ($p = 0.01$), IMU in comparison with AP for all the tasks ($p < 0.01$), IMU in comparison with TC for all the tasks ($p < 0.01$), IMU in comparison with RTMC for all the tasks ($p < 0.01$), IMU in comparison with MP for all the tasks ($p < 0.02$), BA in comparison with RTMH for the W3.5 ($p = 0.04$) and R7.0 ($p = 0.02$), AP in comparison with TC for the W3.5 ($p = 0.01$), TC in comparison with RTMH ($p < 0.02$), RTMC ($p < 0.02$), and MP ($p < 0.02$) for all the tasks, RTMC in comparison with RTMH ($p < 0.01$) and MP ($p < 0.01$) for all the tasks, and RTMH in comparison with MP for the W3.5 ($p = 0.01$). Considering separately the methods, one-way repeated measures ANOVA reports statistical differences among tasks for the following methods:

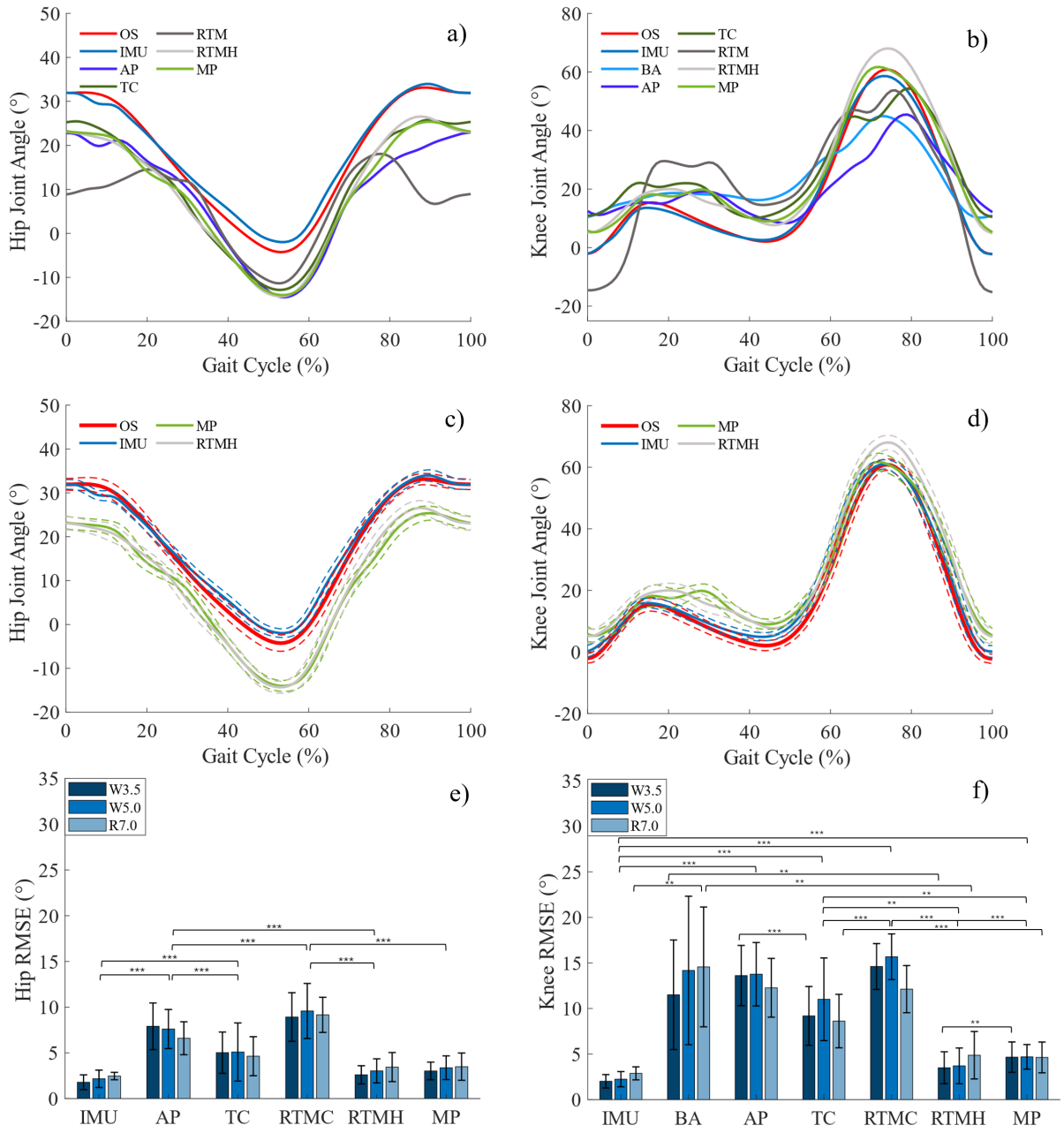


Fig. 7. (a) Hip and (b) knee angles in the sagittal plane averaged over the subjects for the W5.0 task without offset removal. (c) Hip and (d) knee angles in the sagittal plane with mean and standard deviation from subjects for the W5.0 task without offset removal. RMSE values for (e) hip and (f) knee all methods and tasks. *** indicates $p < 0.01$ and ** indicates $0.01 < p < 0.05$.

IMU ($p < 0.01$), TC ($p = 0.03$), RTMC ($p < 0.01$), and RTMH ($p = 0.02$).

For the ROM of the hip joint, the interaction effect was not statistically different, and the multicomparison between methods reported a statistical difference between TC and the RTMH ($p = 0.04$).

Similarly, for the RMSE of the knee, the interaction effect between methods and tasks of the ROM of the knee joint resulted statistically significant. Breaking down the ANOVA, the following statistical differences were found: IMU was found statistically different with all the methods ($p < 0.01$) for all the tasks, except with the RMTH, BA was found statistically different with IMU ($p < 0.01$) and RTMH ($p < 0.01$) for all tasks; AP, TC, and RTMC report a similar trend showing a statistical difference with IMU ($p \leq 0.01$) and RTMH ($p \leq 0.02$) for all tasks; RTMH was statistically different with all the methods ($p < 0.02$) and all task, expect for the IMU; MP was found statistically different with IMU ($p < 0.01$), RTMC ($p = 0.02$), and MP ($p < 0.02$) for all tasks. Considering separately the methods, one-way repeated measures ANOVA reports statistical differences among tasks only for the RTMC method ($p < 0.01$).

The statistical analysis showed that the IMU had a significantly lower RMSE and ε_{ROM} than the other methods for both the hip joint and the knee joint, except for RTMH. The RTMH and MP methods also exhibited statistically lower values than

most other methods for both RMSE and absolute error in both hip and knee joints, indicating a comparable level of measurement accuracy to that of IMUs.

IV. Discussion

This study aimed to compare several vision-based methods, both DNN and non-DNN based, along with an IMU system, against an OS considered as the gold standard. The experimental results demonstrated that while some video-based methods can accurately measure hip and knee joint angles, they faced challenges in accurately capturing ankle joint kinematics. The poor performance of these DNN methods in estimating sagittal ankle kinematics has been reported in several studies [29], [30], [31], [33]. Also, in the results reported by D'antonio et al. [33], it is observed that a DNN-based method (Openpose) failed to accurately track the ankle dorsiflexion angle during both W. and R. in healthy subjects. Using the same DNN method, Wade et al. [31] performed an Altman–Bland analysis on over ground W. and found limits of agreement for ankle dorsiflexion angles too high for clinical applications (±12°). IMU-based systems outperformed vision-based in terms of accuracy, particularly for ankle joint measurements. Regarding the hip and knee joint angle, among the vision-based methods, those utilizing deep learning algorithms, such as RTMH and MP, showed better performance compared to other video-based methods along with marker-based techniques, such as BA. In comparison with the results presented by Menychtas et al. [29], MP outperformed. In their study, MP showed significant errors for hip and knee joint angles with a Bland–Altman mean difference of −26.3° and −17.8° with limits of agreement of more than 15°. This disagreement could be due to the fact that the MP needs to process the whole person on the screen. In contrast to our previous work [45], which focused on a half lower body figure W. on a treadmill, MP showed better results in our current analysis, where the whole human figure was captured. In comparison with the DNN methods compared by Van Hooren et al. [30], only MP and RTMH got better results. RMSE values found in our study for R7.0 were, regarding the hip, (3.4° ± 1.6°) for RTMH and (3.5° ± 1.5°) for MP and (4.9° ± 2.6°) for RTMH and (4.7° ± 1.7°) for MP for the knee. The offset removed in the results could be due to a mismatch between the biomechanical model noted in the dataset, which has inherent pitfalls, and the biomechanical model used as a reference. In this study, the RMSE values assessed align with the findings reported in the literature. In comparison to the results obtained by Vadafar et al. [34], only the results on the hip joint were better with minimum values achieved by MP method of (3.5 ± 1.5)° in comparison with (−4.5° ± 8.2°). The AP, TC, and RTMC methods gave good results with an absolute error between 4° and 8°. However, errors below 4° were observed on both the knee and ankle, although this may be attributed to the fact that the DNN method was trained and tested on the same dataset, which limits its generalizability to other types of environments. Regarding the knee joint angle, our DNN method performs worse. Only RTMH and MP are comparable with respect to the literature results, achieving errors between 4° and 6° for each trial. One potential explanation for the greater accuracy of the MP and RTMH methods could be the backbone training method, as they have used heatmap-based methods that have been shown to increase the accuracy [27]. In our results, the non-DNN method performed worse than some DNN methods (TC, MP, and RTMH) and was not robust to uncontrolled lighting environments. For some subjects, the BA method was able to achieve acceptable errors of ROM, as previously reported [35]. However, considering the entire sample and a measurement condition with uncontrolled light, the evaluation of passive markers by applying color filters was found to be poor with errors even greater than 20°. These results can be attributed to the thresholding nature of the HSV color filters, which, in some trials, only detect a portion of the passive markers, depending on the luminosity of the scene and how the light is reflected upon them. Furthermore, the recorded scene and the change of even a small element of the surroundings could explain the different results obtained with the same method or the same dataset for video-based methods. In addition, it is important to note that different camera orientations [33] and positions [46] could affect video framing along with the accuracy of the results. In the assessment of the hip and knee flexo-extension angle, the MP and RTMH methods could offer a cost-effective alternative to IMUs and OSs in subjects who may be uncooperative. These methods may also be useful in the classification and identification of gait irregularity patterns in patients with cerebral palsy and spastic diplegia [47]. Furthermore, considering an appropriate setup for both cameras and scenes, they could be the core of a telemedicine homecare application for telerehabilitation [48].

V. Conclusion

This study compared vision-based methods, including DNN and non-DNN approaches, and an IMU system against the gold standard, an OS in measuring the kinematics during locomotion. Experimental results indicated that IMU system outperforms vision-based methods, especially for ankle joint measurements. Among vision-based methods, those employing DNN algorithms performed better, such as RTM and MP. For these methods, particularly concerning the hip and the knee joint, comparable measurement errors to those obtained by inertial systems are noted. Thus, it can be reported that in clinical or sports assessments where an inaccuracy not higher than 5° is required, these methods could serve as a viable alternative to IMU. In addition, it is not negligible the potentiality of the video-based analysis for the application in which the real-time evaluation is required, as, for example, the avatar tracking in the videogames or any entertainment scenario. This study highlighted the strengths and limitations of each system, guiding future research toward improving accuracy and applicability in human kinematic analysis. Further research is warranted to address the remaining challenges and improve the performance of vision-based systems, particularly in accurately capturing ankle joint kinematics and achieving higher measurement accuracy in comparison with IMU-based systems. Through this aim, the main future efforts must include both hardware and software updates. For example, considering the hardware, it is mandatory to understand

whether the use of additional cameras totally focused on the feet would help to perform a better kinematic analysis of the lower limbs. Moving to the software, the increment of training data variety should be considered helping the model to better generalize the results, as well to use transfer learning techniques to adapt pretrained models to new specific data due to a continuous fine-tuning of the algorithm parameters.

REFERENCES

[1] R. Baker, "Gait analysis methods in rehabilitation," *J. NeuroEng. Rehabil.*, vol. 3, no. 1, p. 4, Dec. 2006, doi: 10.1186/1743-0003-3-4.

[2] A. Zampogna et al., "Early balance impairment in Parkinson's disease: Evidence from robot-assisted axial rotations," *Clin. Neurophysiol.*, vol. 132, no. 10, pp. 2422–2430, Oct. 2021, doi: 10.1016/j.clinph.2021.06.023.

[3] A. Pacilli et al., "A wearable setup for auditory cued gait analysis in patients with Parkinson's disease," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2016, pp. 1–6, doi: 10.1109/MEMEA.2016.7533796.

[4] I. Mileti et al., "Accuracy evaluation and clinical application of an optimized solution for measuring spatio-temporal gait parameters," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2020, pp. 1–6, doi: 10.1109/MeMeA49120.2020.9137305.

[5] M. P. Kadaba, H. K. Ramakrishnan, and M. E. Wootten, "Measurement of lower extremity kinematics during level walking," *J. Orthopaedic Res.*, vol. 8, no. 3, pp. 383–392, May 1990, doi: 10.1002/jor.1100080310.

[6] V. Camomilla, R. Dumas, and A. Cappozzo, "Human movement analysis: The soft tissue artefact issue," *J. Biomechanics*, vol. 62, pp. 1–4, Sep. 2017, doi: 10.1016/j.jbiomech.2017.09.001.

[7] E. Palermo, S. Rossi, F. Marini, F. Patanè, and P. Cappa, "Experimental evaluation of accuracy and repeatability of a novel body-to-sensor calibration procedure for inertial sensor-based gait analysis," *Measurement*, vol. 52, pp. 145–155, Jun. 2014, doi: 10.1016/j.measurement.2014.03.004.

[8] X. Robert-Lachaine, H. Mecheri, C. Larue, and A. Plamondon, "Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis," *Med. Biol. Eng. Comput.*, vol. 55, no. 4, pp. 609–619, Apr. 2017, doi: 10.1007/s11517-016-1537-2.

[9] S. Glowinski, A. Blazejewski, T. Królikowski, and R. Knitter, "Gait recognition: A challenging task for MEMS signal identification," in *Sustainable Design and Manufacturing*, P. Ball, L. H. Huatuco, R. Howlett, and R. Setchi, Eds., Singapore: Springer, 2019, pp. 473–483.

[10] D. Lo Presti et al., "Fiber Bragg gratings for medical applications and future challenges: A review," *IEEE Access*, vol. 8, pp. 156863–156888, 2020, doi: 10.1109/ACCESS.2020.3019138.

[11] J. Di Tocco et al., "Wearable device based on a flexible conductive textile for knee joint movements monitoring," *IEEE Sensors J.*, vol. 21, no. 23, pp. 26655–26664, Dec. 2021, doi: 10.1109/JSEN.2021.3122585.

[12] J. Di Tocco, D. Lo Presti, A. Rainer, E. Schena, and C. Massaroni, "Silicone-textile composite resistive strain sensors for human motion-related parameters," *Sensors*, vol. 22, no. 10, p. 3954, May 2022, doi: 10.3390/s22103954.

[13] T. Zult, J. Allsop, J. Tabernero, and S. Pardhan, "A low-cost 2-D video system can accurately and reliably assess adaptive gait kinematics in healthy and low vision subjects," *Sci. Rep.*, vol. 9, no. 1, p. 18385, Dec. 2019, doi: 10.1038/s41598-019-54913-5.

[14] C. Prakash, K. Gupta, A. Mittal, R. Kumar, and V. Laxmi, "Passive marker based optical system for gait kinematics for lower extremity," *Proc. Comput. Sci.*, vol. 45, pp. 176–185, Jan. 2015, doi: 10.1016/j.procs.2015.03.116.

[15] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2D human pose estimation: New benchmark and state of the art analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3686–3693, doi: 10.1109/CVPR.2014.471.

[16] L. Wade, L. Needham, P. McGuigan, and J. Bilzon, "Applications and limitations of current markerless motion capture methods for clinical gait biomechanics," *PeerJ*, vol. 10, Feb. 2022, Art. no. e12995, doi: 10.7717/peerj.12995.

[17] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "BlazePose: On-device real-time body pose tracking," 2020, *arXiv:2006.10204*.

[18] H.-S. Fang et al., "AlphaPose: Whole-body regional multi-person pose estimation and tracking in real-time," 2022, *arXiv:2211.03375*.

[19] S. Jin et al., "Whole-body human pose estimation in the wild," 2020, *arXiv:2007.11858*.

[20] N. J. Cronin, "Using deep neural networks for kinematic analysis: Challenges and opportunities," *J. Biomechanics*, vol. 123, Jun. 23, 2021, Art. no. 110460, doi: 10.1016/j.jbiomech.2021.110460.

[21] Y. Zhu, N. Samet, and D. Picard, "H3WB: Human3.6M 3D WholeBody dataset and benchmark," 2022, *arXiv:2211.15692*.

[22] M. Andriluka et al., "PoseTrack: A benchmark for human pose estimation and tracking," 2017, *arXiv:1710.10000*.

[23] S. Vafadar et al., "A novel dataset and deep learning-based approach for marker-less motion capture during gait," *Gait Posture*, vol. 86, pp. 70–76, May 2021, doi: 10.1016/j.gaitpost.2021.03.003.

[24] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," 2018, *arXiv:1812.08008*.

[25] L. Kidzinski, B. Yang, J. L. Hicks, A. Rajagopal, S. L. Delp, and M. H. Schwartz, "Deep neural networks enable quantitative movement analysis using single-camera videos," *Nature Commun.*, vol. 11, no. 1, p. 4054, Aug. 2020, doi: 10.1038/s41467-020-17807-z.

[26] W. Zeng et al., "Not all tokens are equal: Human-centric visual analysis via token clustering transformer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11091–11101.

[27] T. Jiang et al., "RTMPose: Real-time multi-person pose estimation based on MMPose," 2023, *arXiv:2303.07399*.

[28] MMPose Contributors. (2020). *OpenMMLab Pose Estimation Toolbox and Benchmark*. [Online]. Available: https://github.com/open-mmlab/mmpose

[29] D. Menychtas et al., "Gait analysis comparison between manual marking, 2D pose estimation algorithms, and 3D marker-based system," *Frontiers Rehabil. Sci.*, vol. 4, Sep. 2023, Art. no. 1238134, doi: 10.3389/fresc.2023.1238134.

[30] B. Van Hooren, N. Pecasse, K. Meijer, and J. M. N. Essers, "The accuracy of markerless motion capture combined with computer vision techniques for measuring running kinematics," *Scandin. J. Med. Sci. Sports*, vol. 33, no. 6, pp. 966–978, Jun. 2023, doi: 10.1111/sms.14319.

[31] L. Wade et al., "Examination of 2D frontal and sagittal markerless motion capture: Implications for markerless applications," *PLoS ONE*, vol. 18, no. 11, Nov. 2023, Art. no. e0293917, doi: 10.1371/journal.pone.0293917.

[32] K. Iskakov, E. Burkov, V. Lempitsky, and Y. Malkov, "Learnable triangulation of human pose," 2019, *arXiv:1905.05754*.

[33] E. D'Antonio, J. Taborri, I. Mileti, S. Rossi, and F. Patané, "Validation of a 3D markerless system for gait analysis based on OpenPose and two RGB webcams," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17064–17075, Aug. 2021, doi: 10.1109/JSEN.2021.3081188.

[34] S. Vafadar, W. Skalli, A. Bonnet-Lebrun, A. Assi, and L. Gajny, "Assessment of a novel deep learning-based marker-less motion capture system for gait study," *Gait Posture*, vol. 94, pp. 138–143, May 2022, doi: 10.1016/j.gaitpost.2022.03.008.

[35] L. Ceriola, I. Mileti, J. Taborri, M. Donati, S. Rossi, and F. Patané, "Comparison of two video-based methods for knee joint angle measurement: A preliminary study," in *Proc. IEEE Int. Conf. Metrology eXtended Reality, Artif. Intell. Neural Eng. (MetroXRAINE)*, Oct. 2023, pp. 155–160, doi: 10.1109/MetroXRAINE58569.2023.10405673.

[36] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000, doi: 10.1109/34.888718.

[37] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," 2021, *arXiv:2107.08430*.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.

[39] P. Wang et al., "Understanding convolution for semantic segmentation," 2017, *arXiv:1702.08502*.

[40] Z. Hou, X. Peng, Y. Qiao, and D. Tao, "Visual compositional learning for human-object interaction detection," 2020, *arXiv:2007.12407*.

[41] Y. Li et al., "SimCC: A simple coordinate classification perspective for human pose estimation," 2021, *arXiv:2107.03332*.

[42] M. M. Rossi, A. P. Silvatti, F. A. S. Dias, and R. M. L. Barros, "Improved accuracy in 3D analysis using DLT after lens distortion correction," *Comput. Methods Biomechanics Biomed. Eng.*, vol. 18, no. 9, pp. 993–1002, Jul. 2015, doi: 10.1080/10255842.2013.866231.

[43] C. Yang et al., "Automation enhancement and accuracy investigation of a portable single-camera gait analysis system," *IET Sci., Meas. Technol.*, vol. 13, no. 4, pp. 563–571, Jun. 2019, doi: 10.1049/iet-smt.2018.5246.

[44] F. Crenna, G. B. Rossi, and M. Berardengo, "Filtering biomechanical signals in movement analysis," *Sensors*, vol. 21, no. 13, p. 4580, Jul. 2021, doi: 10.3390/s21134580.

[45] L. Ceriola, I. Mileti, M. Donati, and F. Patanè, "Comparison of video-based algorithms for 2D human kinematics estimation: A preliminary study," *J. Phys., Conf.*, vol. 2590, no. 1, Sep. 2023, Art. no. 012002, doi: 10.1088/1742-6596/2590/1/012002.

[46] J. Stenum, C. Rossi, and R. T. Roemmich, "Two-dimensional video-based analysis of human gait using pose estimation," *PLOS Comput. Biol.*, vol. 17, no. 4, Apr. 2021, Art. no. e1008935, doi: 10.1371/journal.pcbi.1008935.

[47] N. G. Langerak, N. Tam, J. du Toit, A. G. Fieggen, and R. P. Lamberts, "Gait pattern of adults with cerebral palsy and spastic diplegia more than 15 years after being treated with an interval surgery approach: Implications for low-resource settings," *Indian J. Orthopaedics*, vol. 53, no. 5, pp. 655–661, Oct. 2019, doi: 10.4103/ortho.ijortho_113_19.

[48] A. Nucita, G. Iannizzotto, M. Perina, A. Romano, and R. A. Fabio, "Telerehabilitation with computer vision-assisted markerless measures: A pilot study with rett syndrome patients," *Electronics*, vol. 12, no. 2, p. 435, Jan. 2023, doi: 10.3390/electronics12020435.

Luca Ceriola received the master's degree in mechanical engineering from Sapienza University, Rome, Italy, in 2022. He is currently pursuing the Ph.D. degree in industrial and civil engineering with Niccolò Cusano University, Rome, in collaboration with Sensor Medica srl. Luca Ceriola currently works on the project "Walking Over Balance-Board Learning Environment" (WOBBLE) Project, which aims to design and develop an all-round platform that interacts with Sensor Medica's Technology Environment to create innovative systems for human motion and posture assessment.

His research interests include the metrological validation of video-based algorithms based on deep learning architectures. His main interests include biomechanics, sensors, metrology, metrology for artificial intelligence, and human motion tracking.

Juri Taborri (Member, IEEE) received the master's (cum laude) degree in biomedical engineering from the Sapienza University, Rome, Italy, in 2014, and the Ph.D. degree in industrial and management engineering from the Sapienza University of Rome, Rome, in 2017.

He is currently an Assistant Professor in mechanical and thermal measurements with the University of Tuscia, Viterbo, Italy. His research interests include biomechanical measurements, validation and metrological characterization of innovative sensor systems, robotics, and machine-learning algorithms for human motion.

Marco Donati received the Computer Engineering degree from the University of Roma Tre, Rome, Italy, in 2003, and the Ph.D. degree in bioengineering from Alma Mater Studiorum, Bologna, Italy, in 2007.

Since 2007, he has been involved in high-resolution human motion analysis using stereophotogrammetric and inertial systems. He specializes in both the clinical field and in entertainment, utilizing motion capture techniques to animate 3-D characters. In 2011, he served as a Research Fellow at the University of Rome Foro Italico, Rome. In 2011, he founded the University Spin-Off Sensorize Srl, which focuses on sports gesture analysis through inertial sensors. In 2015, he founded Motustech Srl, which specializes in integrated systems for functional evaluation, where he served as the Technical Director in 2023. Since 2023, he has been responsible for software development at Sensor Medica.

Dr. Donati won the award for Best Methodological Contribution at Siamoc in 2006.

Stefano Rossi received the Mechanical Engineering (Hons.) degree from Sapienza University of Rome, Rome, Italy, in 2004, and the Ph.D. degree from Padua University, Padua, Italy, in 2008.

He is currently a Full Professor and the Vice-Director with the Department of Economics, Engineering, Society, and Business Organization, Tuscia University, Viterbo, Italy. He is also the Chair of the School of Engineering and the Head of the Industrial Measurements and Experimental Biomechanics Laboratory. His research interests include the design of innovative sensor systems for biomechanical applications, the design of new robotic systems for movement recovery, and the individuation of indices to objectively measure human motor performance.

Fabrizio Patanè received the Mechanical Engineering degree from La Sapienza University of Rome, Rome, Italy, in 2000, and the Ph.D. degree from Padua University, Padua, Italy, in 2004.

He is a Full Professor of Thermal and Mechanical Measurements with the Faculty of Engineering, Niccolò Cusano University, Rome. He is a Coordinator of both the Teaching and the Research Commissions and a member of the Doctorate School Teaching Staff in Industrial-civil engineering at Niccolò Cusano University. He is the author of more than 60 peer-reviewed scientific publications and three patents. His scientific skills regard mechanical and thermal measurements, environmental and inertial measurements, motion-analysis systems for biomedical applications, and robotic mechanisms for neurorehabilitation.

Ilaria Mileti received the master's degree in biomedical engineering from Sapienza University of Rome, Rome, Italy, in 2015, and the Ph.D. degree in industrial and management engineering from the Sapienza University of Rome, in 2019.

She is an Associate Professor in mechanical and thermal measurements at the Department of Engineering, University Niccolò Cusano, Rome, where she is a member of the Ph.D. program in Territory Innovation and Sustainability. Her research interests include measurements for experimental biomechanics, wearable systems, machine learning algorithms, and robotic devices.