Applied Computer Systems
ISSN 2255-8691 (online)
ISSN 2255-8683 (print)
2025, vol. 30, no. 1, pp. 157–165
https://doi.org/10.2478/acss-2025-0017
https://content.sciendo.com

RIGA TECHNICAL UNIVERSITY

A Comprehensive Vision-Based Gait Data Collection Framework with a Systematic Multi-Camera Placement Strategy

Pınar Güner Şahan[1*], Ilgar Akkaya[2], İbrahim Şahan[3], Suhap Şahin[4]
[1–4]*Department of Computer Engineering, Kocaeli University, Kocaeli, Turkey*

Abstract **– Systematic collection of gait data directly affects the reliability of gait analysis. Therefore, the first step that should be given importance in gait analysis is to ensure that the data collected is of high quality and reliable. Optimising and standardising the data collection process is a critical requirement that increases the success of the analysis. This study proposes a systematic multi-camera placement strategy and data processing process to collect gait data in real time with RGB cameras for gait recognition and analysis applications. The proposed method provides an end-to-end framework from the physical setup of the data collection environment to the data processing steps. The camera placement strategy aims to maximise the visibility of all body parts by capturing the participant's body from different directions during walking. The results of three different methods used for silhouette extraction from the acquired videos were compared. Furthermore, a video-based approach was used to calculate participants' walking speeds. Theoretical and practical information provided regarding the data collection and analysis process is detailed to guide future studies. In this respect, the paper is aimed to be a guiding study for researchers working in the related field.**

Keywords **– Computer vision, data collection, gait analysis, multi-camera systems, reliability of data.**

I. Introduction

Gait can be defined as the walking way of a person, and research associated with this topic is called gait analysis. Every person has their own gait with different features; therefore, gait can contain discriminative information about people [1]. Hence, gait data has been used extensively in several application fields such as person recognition and identification [2]–[6], healthcare [7]–[11], biomechanics and sport science [12]–[14] and psychology [15]–[17].

Gait data is usually obtained sensor-based using floor sensors and wearable sensors such as accelerometers, gyroscopes, and light sensors, or vision-based using cameras [18]. Sensor-based gait analysis systems need a complex device configuration to collect data, while vision-based systems allow data to be collected without any physical interaction using cameras [19]. Improving the performance of these systems largely depends on the number of cameras and their positioning to collect data from different angles throughout the participant's gait. Considering gait analysis and gait recognition studies, appearance changes are quite common in the real world. Therefore, it is important to include gait videos obtained from different perspectives in the datasets to improve system performance [20]. Using data captured by multiple cameras and providing robustness to different viewpoints offers the possibility of evaluating a person's 3D pose and dynamic features [21].

Most of the existing vision-based gait datasets [22]–[31] have been collected for clinical purposes or gait recognition and each has different features such as the number of subjects, number of cameras, environmental conditions, data type and variations in the appearance of the individual. Although RGB methods such as video are commonly used to record gait sequences, most of these datasets do not include the original RGB data due to challenges related to data storage and privacy concerns. Therefore, datasets usually comprise gait sequences represented in silhouette format, extracted from RGB video data [31]. This isolates individual movement patterns in the images, minimising variables related to lighting conditions or clothing, allowing the algorithms to focus solely on gait features.

In practice, collecting gait data is extremely challenging, because the processes of selecting participants for data collection, obtaining the consent of participants, providing the data collection environment and system installation, and recording and processing gait videos can be time-consuming and costly [32], [33]. Therefore, the data collection environment must be well planned in the beginning and must be set up in accordance with the purpose of the study. A well-designed data collection environment minimises errors, enabling costs to be managed effectively and reliable data to be obtained. In this context, it is critical to pay attention to factors such as camera placement and calibration and improving lighting conditions to increase data accuracy and save time and cost.

This study presents a systematic approach to gait data collection by detailing how a data acquisition setup can be established within a specific physical environment. During

* Corresponding author's e-mail: pinar.guner@kocaeli.edu.tr
Article received 2025-10-02; accepted 2025-11-07

implementation, a customized setup was designed based on the spatial characteristics of the selected environment and the specifications of the equipment used (e.g., number and properties of cameras). The gait videos recorded in this environment were analysed, and silhouettes were extracted to obtain a processable data format. Additionally, walking speeds were computed from the recorded videos and the results were evaluated in terms of the system's functionality. Note that this study emphasises the system for vision-based gait data collection rather than the database itself.

This study provides a comprehensive framework for collecting video-based gait data, and its main contributions are as follows:

- A systematic camera setup is proposed to capture participants from different directions. This layout, with six cameras, maximises body visibility across all gait phases and supports high-quality silhouette extraction.
- The study presents a full data acquisition pipeline, including setup design tailored to environmental constraints, RGB video recording, distortion correction, silhouette generation, and walking speed computation from video sequences.
- Three different methods are evaluated on the collected dataset. This allowed us to evaluate segmentation performance in a real gait data collection scenario.
- A video analysis-based method used for estimating walking speed by tracking the temporal change in bounding box position within predefined spatial boundaries, supporting real-time motion assessment without additional sensors.

The rest of the paper is organised as follows. Section II reviews related works focusing on gait data collection strategies. Section III extensively describes camera placement details and data collection protocol. Also, it explains the video preprocessing, silhouette extraction and walking speed calculation steps. Section IV gives a summary of the dataset and reports the research findings. Finally, Section V contains an overall assessment and presents our conclusions.

II. Related Work

Despite the extensive literature on gait recognition and gait analysis, the determination of strategies for collecting such data has received limited attention. The effect of different camera angles on the accuracy of gait analysis has been examined in some studies. Baldinger et al. examined the validity of joint angle estimation from different camera angles, including front left, back left, front right, and back right [34]. The findings of the study revealed that knee and hip joints were measured more consistently when sagittal and frontal angles were used, and these measurements were made with fewer errors when alternative angles such as back left were used. Han et al. stated in their review that camera placement was a critical issue in gait analysis studies [35]. They pointed out the lack of comparative studies evaluating the effectiveness of different camera placement combinations and emphasised the need for systematic research on the effects of camera placement on the quality of gait data collection. There are some studies that emphasise the importance of camera placement for collecting reliable gait data [36], [37], as well as a system proposed for automatic gait data collection [33]. Riberto et al. [37] stated that the location where the cameras was positioned depended on some factors such as the type of the study, the size of the room where the study was conducted, the number of cameras available and the characteristics of the cameras. Takemura et al. [26] collected gait data for biometric identification based on the gait collection framework introduced by Makihara et al. [33]. In the study, a multi-camera setup is used to collect gait data from different angles. In this way, gait features are analysed from different perspectives. However, this study has some limitations in terms of the positioning of the cameras. The camera setup given in the paper allows images to be taken from 14 viewing angles in the range of 0°–90°, 180°–270° (intervals of 15-degree). When the image was recorded in this way, it was observed that a portion of the subject's left posterior and right anterior region was not visible, as no camera was placed at any angle between 90°–180° with respect to the centre of the walking path. Therefore, gait features from this perspective are not recorded. The inadequate positioning of the cameras, preventing coverage of all perspectives of the gait may limit the scope of the dataset. These studies provide important insights regarding camera setup techniques for gait data collection.

III. Methods

A. Data Collection Setup and Acquisition

Within the scope of this study, gait data collection was performed in a foyer area of approximately 60 square metres provided by Kocaeli University, Faculty of Medicine (see in Fig. 1). In the data collection area, artificial lighting was used in addition to ambient daylight and stable lighting conditions were provided. The area was cleared of distractions and cameras were positioned to provide optimal viewing angles. The setup of the gait data collection area and the methods used are described in detail below.

Fig. 1. Data acquisition environment.

In this study, attention has been paid to strategic placement of cameras for a complete visualization of the person's gait movements and to record the gait from different angles. Six network cameras recording at 1920 × 1080 resolution and 25 fps are placed at 30-degree intervals along half of a circle whose centre coincides with the centre of the walking path to cover a wide field of view. An illustration of the camera placement is given in Fig. 2. Each camera in the figure is numbered from Cam1 to Cam6, and both ends of the walking path are labelled as points A and B. θ is the view angle that represents the position of the camera placements in relation to the centre of the walking path. This angle geometrically describes how the participant is seen by the cameras as they move along the walking path. Cam1 to Cam6 capture gait images from 12 view angles ranging 0°–150° while the person is walking from point A to point B, and ranging 180°–330° while the person is walking from point B to point A. The camera views for gait from point B to point A are shown virtually from Cam 1' to Cam 6'.

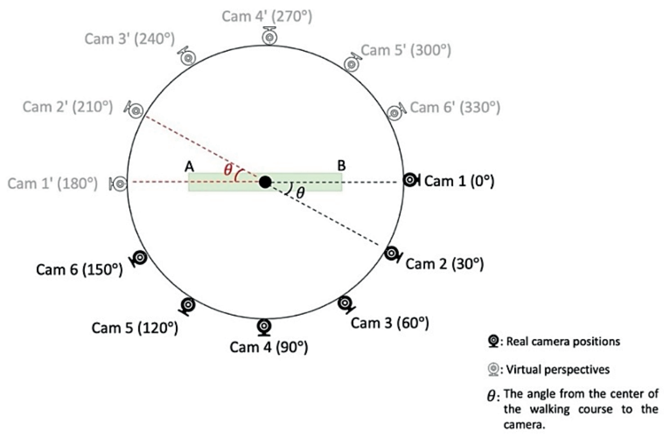

Fig. 2. Illustration of the camera placement.

The cameras were mounted on tripods 2 m above the ground, and each positioned approximately 5.4 m from the centre of the walking path. The participant walks from point A to point B and from point B to point A (once forward and once backward) on a straight 6.5 m long walking path. As a result, a total of 12 (2 gait sequences, forward and backward, and 6 camera) gait sequences are captured per participant. A gait sequence contains a varying number of images depending on the walking speed of the person and the frames per second captured by the camera. The cameras are placed at equal angular intervals following a semicircle with a radius of 5.4 m, allowing each camera to symmetrically observe both the participant's gait from point A to point B and from point B to point A.

In brief, six cameras were physically placed on one side of the walking path. By having the participant walk forward and backward along the walking path, it was possible for each camera to capture images of the person from different angles. This strategy ensures that the front, back, right, left and diagonal (front right, back right, front left, back left) directions are covered in a single walking session as shown in Fig. 3. As illustrated in the figure, the diagonal-view cameras Cam 2 and Cam 3 capture the subject from front-right and back-left directions, while cameras Cam 5 and Cam 6 observe back right and front left perspectives. Although these cameras target similar orientations, the differences in their angular placement result in distinct visual perspectives. This difference affects the body orientation and limb visibility throughout the gait cycle. Hence, each camera contributes visual information, enhancing the completeness and quality of the collected gait data.

In the study, gait images were collected from 30 willing participants to test the data collection setup. Participants walked in both directions with their shoes on a 6.5 m long (approximately 10–12 steps) walking path. To make accurate measurements in the following stages, we need to consider the speeding up and slowing down processes that may occur at the beginning and end of the walk [38]. To maintain a constant speed during the measurement, participants were allowed to start their walk one or two steps before the walking path and end it one or two steps after the path. This enabled the participants to reach a constant speed. In this way, it was aimed to obtain more reliable results by excluding the beginning and end parts of the walk. While participants walk, a total of 12 gait sequences are obtained from six cameras placed in the data collection environment. Six of these gait sequences contain the gait from point A to point B, and the other six contain the gait from point B to point A.

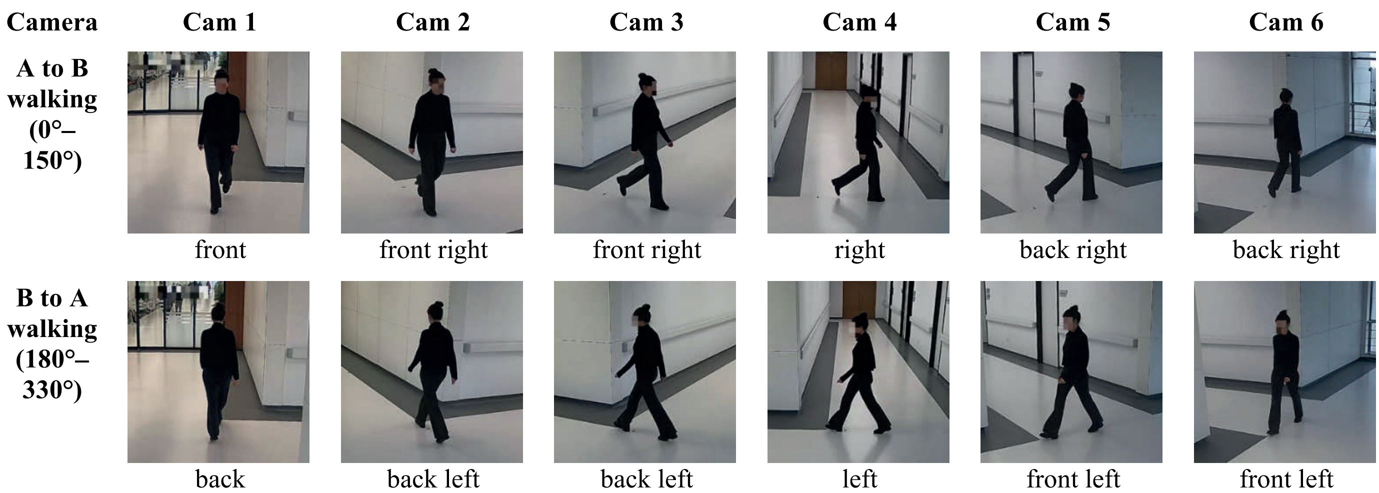

Fig. 3. Illustration of the camera placement.

B. Preprocessing and Silhouette Extraction

Real-Time Streaming Protocol (RTSP) [39], which provides real-time data transmission, is used to ensure simultaneous video recording from the cameras. In this way, video streams are recorded without interruption. The low latency and flexible structure provided by RTSP play an important role in the sensitive data collection processes required in the study. Data received simultaneously from multiple camera sources can be processed in a synchronized manner to produce high quality video recordings. This process ensures data integrity. In order to focus on the details of the gait dynamics, the focal points of the cameras are set to the centre point of the walking path. This is done by automatically marking the centre of the camera's field of view in the preview program written to monitor camera snapshots (see Fig. 4). In this way, the camera is directed to the desired focus. This process is carried out by physically moving each camera and aligning the point marked in the preview program with the centre of the walking path.

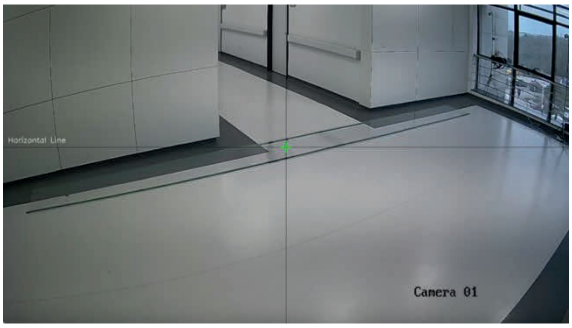

Fig. 4. Adjusted focus point of the camera.

Gait data of 1920 × 1080 pixels at 25 fps was collected in this way; some deformations occurred due to a lens distortion problem in the cameras. To solve this problem, CCD Camera Calibration method based on OpenCV was used [40]. This method provides a highly accurate calibration based on the pinhole camera model. Camera calibration is performed on a specific pattern such as a chessboard plane. In this paper, in order to detect all lens distortions, pictures were taken with the camera at different angles and positions from a chessboard plane image. The corner points of the pattern were identified in the captured images and these points were associated with real world coordinates. With this information, lens parameters such as focal length and lens distortion were calculated and used to correct the distortions in the videos.

For tasks such as gait recognition, precise human silhouette extraction from videos is required. The use of silhouette extraction is quite common in computer vision fields such as motion analysis, video surveillance, and human or object detection [41]. Silhouette extraction can be defined as the process of separating the target object from the background. A silhouette is usually in the form of a binary image where the object is represented in white and the background in black [42]. In this paper, distortion corrected videos are divided into frames, and deep learning-based silhouette extraction is performed on each frame. The pipeline is shown in Fig. 5.

For this purpose, first, person detection was performed using the latest version of YOLO (You Only Look Once), which is a popular deep learning model used for real-time object detection [43]. Considering object detection as a regression problem, YOLO divides the image into grids and estimates the bounding boxes of possible objects and the class probabilities of these objects for each grid. Among these predictions, objects marked with the person label are identified with their respective bounding boxes, and person detection is performed. YOLO11 offers improvements over previous versions, especially in terms of performance and flexibility [44]. After person detection, the images are resized to a fixed size to create normalized silhouettes. For this purpose, a new image of size 600 × 600 is created by taking the bounding box as the centre. This fixed size is chosen to ensure that the person stays inside the box when moving. Second, segmentation is performed to extract the silhouettes of the detected people from the image. Segmentation is the process of dividing an image into objects or regions called segments. This process is carried out by assigning a label to each pixel in the image that represents the class to which that pixel belongs, and in such a way that pixels in the same class have similar properties [45]. In this study, three different segmentation methods (SAM 2 [46], YOLO11 [44], DeepLabv3 [47]) were applied to generate silhouettes from video frames, and their performances were analysed.

SAM 2 [46], an improved version of the Segment Anything Model (SAM) [48] developed by Meta AI, generates segmentation masks of the corresponding object for a given frame. It conditions the segmentation prediction based on memories of previous predictions stored in its memory. It is known that this model is more accurate and six times faster than SAM in image segmentation [48]. The result of the person detection performed in the first step is given as input to the segmentation model. This limits the segmentation area of the model. Instead of working on the entire image, the model processes only a specific region.

Fig. 5. Silhouette extraction pipeline.

This method improves accuracy and reduces the time required for image segmentation.

In this study, YOLO11 [44] was used not only for bounding box detection of the walking person but also extended with a post-processing step to extract the silhouette by masking the detected region. While offering low latency in real-time applications, advanced attention mechanisms ensure that accurate masks are produced even in complex scenes. DeepLabv3 [47] is widely used in human segmentation tasks due to its strong boundary refinement capabilities. Its semantic approach makes it effective when the walking subject is well separated from the background.

After segmentation, each frame was converted into a binary silhouette image, where the subject's region was highlighted and the background suppressed to facilitate further analysis.

C. Walking Speed Computation

In this study, an important metric, walking speed, was calculated to analyse the gait dynamics of the participants by processing the data collected. This metric can be addressed in areas such as monitoring rehabilitation processes, evaluating mobility and neurological conditions, as well as sports performance analysis and stress level detection.

At this point, a video analysis-based method was preferred to measure walking speed. For this purpose, the starting and ending points of the walking path are represented by two separate bounding boxes. In order to track the change of the person's position over time, the previously mentioned YOLO11 [44] model is used for person detection. With this model, the bounding box coordinates of the walking person are obtained for each frame throughout the video. The walking speed detection from video performed within the scope of this article can be formulated as follows. The first time the person enters the bounding box at the start point, t_s, and the first time the person enters the bounding box at the end point, t_e, are determined in video frames according to (1) and (2), respectively. In these equations, B_s represents the bounding box coordinates at the start point, B_e represents the bounding box coordinates at the end point, and $B_p(t)$ represents the bounding box coordinates of the person at time t.

$$t_s = \min\left\{ t : B_p(t) \in B_s \right\}. \tag{1}$$

$$t_e = \min\left\{ t : B_p(t) \in B_e \right\}. \tag{2}$$

The time that the person walks on the walking path is indicated by T (walking time) and is calculated according to the frames per second (fps) information of the camera as in (3).

$$T = \frac{t_e - t_s}{fps}. \tag{3}$$

Finally, given the distance between the bounding boxes of the start and end points, i.e., the length of the walking path is denoted by d, the walking speed (v) of the person can be calculated in m/s by (4).

$$v = \frac{d}{T}. \tag{4}$$

Figure 6 shows the video-based speed computation process of a participant. In the figure, the start and the end areas of the walking path are indicated by bounding boxes, and the times the participant entered these areas were tracked frame by frame.

IV. Results

This section presents the analysis results of gait videos obtained using the proposed camera placement strategy. The findings enable evaluations regarding the applicability of the data collection process and the processability of the images obtained.

A. Collected Data Overview

As of this publication, gait images (1920 × 1080 pixels at 25 fps) were collected from 30 participants in order to test the camera setup. Participants consist of 10 females and 20 males aged between 19 and 29. For participant statistics, see Fig. 7.

Participants walked on a fixed 6.5 m long walking path to the endpoint and returned to the starting point. During the data collection process, six cameras were used, which were placed in the data collection environment as described in Section II-A. During each participant's walk, a total of 12 gait sequences in two different directions were recorded (from six cameras in two different directions). A table summarising key dataset statistics is shown in Table I.

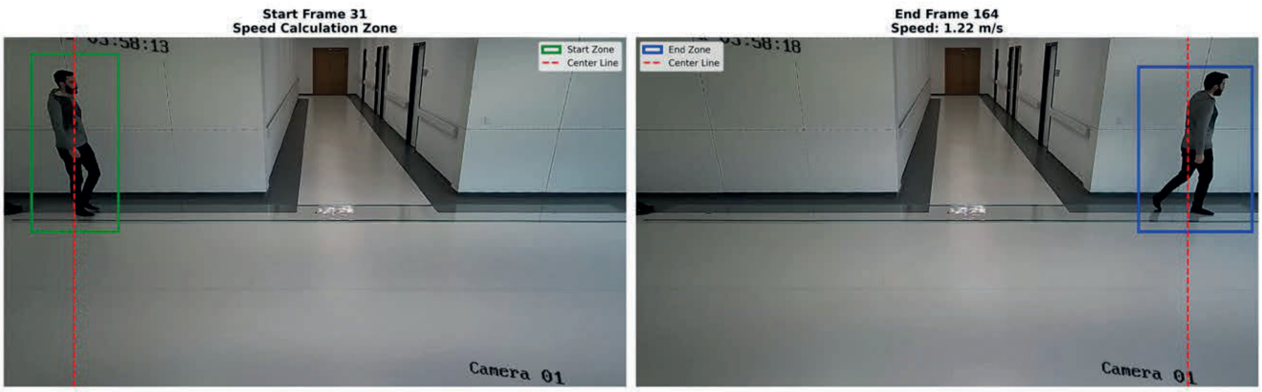

Fig. 6. Walking speed computation process.

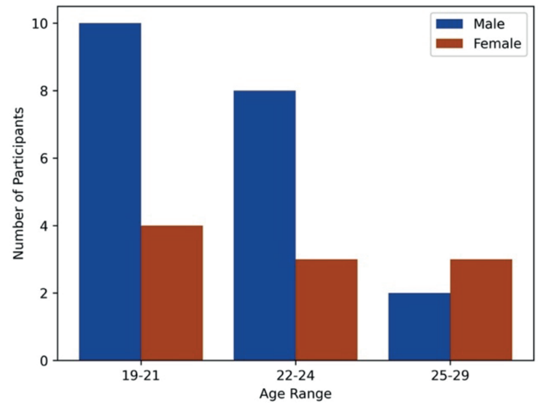

Fig. 7. Distribution of participants by age ranges and genders.

TABLE I
DATASET SUMMARY

Metric	Value
Number of participants	30
Number of cameras	6
Total gait sequences	360
Average frames per gait sequence	133
Frame rate	25
Video resolution	1920 × 1080

This approach allowed for a comprehensive assessment of gait in different aspects, as shown in Fig. 8. In the figure, the images in the blue frame show the person walking from point A to point B; the images in the orange frame show the same person returning from point B to point A. When all these images are combined in sequence and in succession, a 360-degree sequence of images of the person is created. These procedures were completed for 30 participants, and a total of 360 gait sequences were obtained.

Table II summarises the directional viewpoint coverage of popular gait datasets in comparison with the proposed multi-camera setup. The participant's body is viewed from the front, back, right, left and diagonal (front right, front left, back right, back left) directions along the walking path, enabling a detailed observation of posture and limb movements from multiple perspectives.

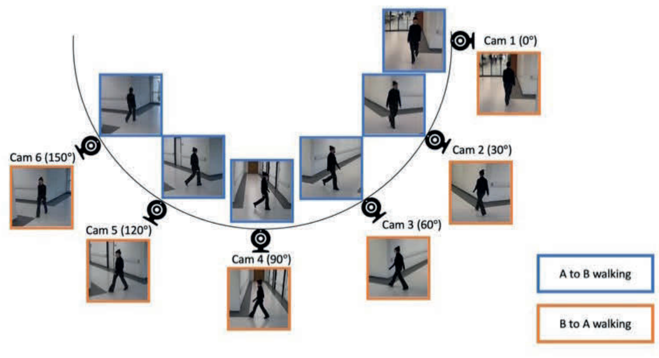

Fig. 8. Illustration of camera placement and data collection.

B. Data Quality and Usability Evaluation

The collected images were processed, and silhouettes were extracted as gait features using the techniques described in Section II-B. In the second part of the silhouette extraction, to find the most appropriate method, in addition to the SAM 2 [46], YOLO11 [44] and DeepLabv3 [47] models were also evaluated for segmentation within the scope of the study. Figure 9 shows silhouette examples of these models. For each model, the first two images show silhouette examples of images taken from a distance, and the last two images show silhouette examples of close-up images of the person. In this figure, it is seen that SAM 2 is the best-performing model in terms of accurate background separation and detailed representation of person boundaries. Particularly, this performance advantage becomes more pronounced when the person is distant.

One of the most important factors directly affecting silhouette quality is the camera placement strategy employed during data collection. The multi-camera placement, which allows for the participant's body to be viewed from the front, back, right, left, and diagonal directions along the walking path, not only contributed to segmentation algorithms producing clearer human silhouettes but also enabled the observation of posture and limb movements from different angles. Notable evidence of this effect is that body parts appear more evenly and balanced in images of the same scene taken from different angles.

TABLE II
COMPARISON OF VIEWPOINT COVERAGE IN POPULAR GAIT DATASETS AND THE PROPOSED SETUP

Camera setup	Number of cameras	Front	Back	Right	Left	Front-Right	Front-Left	Back-Right	Back-Left
Yu et al. (2006) [24]	11	✓	✓	×	✓	×	✓	×	✓
Takemura et al. (2018) [26]	7	✓	✓	✓	✓	×	✓	✓	×
Song et al. (2023) [20]	8	✓	✓	✓	✓	✓	×	×	✓
Proposed method	6	✓	✓	✓	✓	✓	✓	✓	✓

Fig. 9. Silhouette examples of different segmentation models.

When viewed from different angles, the visibility of each limb increases. In addition, since the overlap between limbs is reduced in images obtained from diagonal directions, contour lines become more distinct, allowing for clearer and unobstructed silhouettes. It has been observed that dynamics such as body sway and stride length during gait are better reflected in images taken from diagonal angles. This observation is supported by the improvement in contour sharpness and shape integrity observed in the example silhouette outputs presented in Fig. 10. The sample images were selected from images obtained from various angles (0°, 30°, 60°, 90°, 120°, 150°) while the participants were walking in a single direction (A to B).

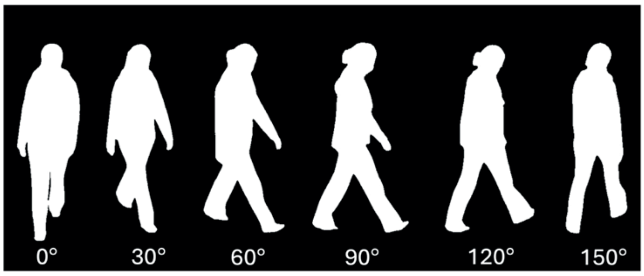

Fig. 10. Different viewpoint silhouettes from the same gait instance (for a gait sequence).

To evaluate the dynamic usability of the dataset, the walking speeds of the participants were computed and analysed. For this purpose, first, the walking times of the participants on the 6.5 m walking path were calculated as described in Section II-C. To verify the calculated walking time of a participant, the start and end times of the video recordings were adjusted to cover the participant's walk from the beginning to the end of the walking path. The video durations were compared with the calculated walking times and a high level of consistency was observed between them. This result demonstrated the reliability of the video analysis-based speed computation method used in this study.

As a result of the analysis, it has been determined that the average walking speed of the participants is 1.24 m/s, and the standard deviation is 0.12 m/s (Table III). This result indicates that walking speeds are concentrated close to the average. Previous studies indicate that the comfortable walking speed for healthy individuals is generally between 1.20 m/s and 1.40 m/s [49], [50]. The results obtained in this study have been found to be consistent with this range. The data obtained reveal that walking speed in healthy individuals shows a regular distribution and that these measurements can be used as a reference in clinical or physical assessments.

TABLE III

DESCRIPTIVE STATISTICS OF PARTICIPANT WALKING SPEEDS

Walking speed parameter	Measured value (m/s)
Average	1.24
Standard deviation	0.12
Minimum	0.92
Maximum	1.42

The camera placement strategy proposed in this study enables the acquisition of gait images from multiple angles, providing significant gains in terms of silhouette production quality and the quantitative analysis of gait data. The results demonstrate the high usability of both the proposed approach for gait data collection and the processing methods used.

V. LIMITATIONS

In this study, gait data were collected under stable lighting and obstacle-free conditions. These controlled conditions ensured consistency across trials and repeatability of measurements. While this approach is appropriate for demonstrating that the data collection framework design performs as intended, it does not reflect the diversity of the real world, such as variable lighting conditions and obstructed environments. Future studies testing the proposed approach under a wider range of environmental conditions by addressing these limitations will provide a more comprehensive evaluation.

VI. CONCLUSION

This paper provides detailed descriptions of the setup required for the efficient collection of gait data, which can be used in different fields such as health, sports, biometric identification, and the processing and analysing of the data obtained. Setting up the data collection environment correctly ensures that data is collected completely and accurately, increasing both the reliability of the processed data and the accuracy of the results obtained from the analysis. Therefore, the accuracy of camera positioning plays a critical role in the quality and reliability of the collected data in studies that aim to create a dataset of video images. In addition, the performance of the algorithms used to convert the data into silhouette format is critical for accurate and fast analysis. Precise measurement of metrics such as walking speed offers a wide range of potential applications in areas such as health and sports. In this study, a real-time multi-camera system has been used to obtain images from different angles and increase data diversity. The proposed framework enhances data reliability by balancing limb visibility across perspectives, providing a replicable methodology for future gait research. The placement of the cameras has been carried out carefully to increase the data quality. In addition, it is aimed to create a reliable basis for studies to be performed in this field by converting the data into silhouette format and calculating walking speed. The results of the study contribute to improving quality in gait data collection and emphasise the importance of systematic approaches in creating reliable datasets. The paper aims to ensure the reproducibility of the data collection process by addressing the dimensions of the special area provided for data collection, the features of the materials used, the camera placement details and processing details of the collected data. In this respect, our study is a guide for researchers who aim to create a unique data set on gait. Future research should extend this framework by including diverse environmental conditions and advanced data processing techniques.

Ethics Statement

Ethical approval was obtained from the Kocaeli University Ethics Committee for Natural and Engineering Sciences (050.99-732736), and all participants gave informed consent to take part in the study. The images used in the figures belong to the authors; no participant images were included in the article.

Acknowledgement

The authors would like to thank Research Assistant Tuğrul Hakan Gençtürk for his assistance in solving a technical issue related to the synchronization of multi-camera data.

References

[1] C. Prakash, R. Kumar, and N. Mittal, "Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges," *Artificial Intelligence Review*, vol. 49, no. 1, pp. 1–40, Sep. 2018. https://doi.org/10.1007/s10462-016-9514-6

[2] C. Fan *et al*., "GaitPart: Temporal part-based model for gait recognition," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 14213–14221. https://doi.org/10.1109/CVPR42600.2020.01423

[3] X. Huang *et al*., "Context-sensitive temporal feature learning for gait recognition," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, Oct. 2021, pp. 12889–12898. https://doi.org/10.1109/ICCV48922.2021.01267

[4] H. Li *et al*., "GaitSlice: A gait recognition model based on spatio-temporal slice features," *Pattern Recognition*, vol. 124, Apr. 2022, Art. no. 108453. https://doi.org/10.1016/j.patcog.2021.108453

[5] X. Huang, X. Wang, B. He, S. He, W. Liu, and B. Feng, "STAR: Spatio-temporal augmented relation network for gait recognition," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 5, no. 1, pp. 115–125, Jan. 2023. https://doi.org/10.1109/TBIOM.2022.3211843

[6] C. Fan, J. Liang, C. Shen, S. Hou, Y. Huang, and S. Yu, "OpenGait: Revisiting gait recognition toward better practicality," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, Jun. 2023, pp. 9707–9716. https://doi.org/10.1109/CVPR52729.2023.00936

[7] T. T. Verlekar, P. Lobato Correia, and L. D. Soares, "Using transfer learning for classification of gait pathologies," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Madrid, Spain, Dec. 2018, pp. 2376–2381. https://doi.org/10.1109/BIBM.2018.8621302

[8] R. Kaur, S. Menon, X. Zhang, R. Sowers, and M. E. Hernandez, "Exploring characteristic features in gait patterns for predicting multiple sclerosis," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, Jul. 2019, pp. 4217–4220. https://doi.org/10.1109/EMBC.2019.8857604

[9] R. Kaur, R. W. Motl, R. Sowers, and M. E. Hernandez, "A vision-based framework for predicting multiple sclerosis and Parkinson's disease gait dysfunctions – A deep learning approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 190–201, Jan. 2023. https://doi.org/10.1109/JBHI.2022.3208077

[10] V. W. S. Tan, W. X. Ooi, Y. F. Chan, T. Connie, and M. K. O. Goh, "Vision-based gait analysis for neurodegenerative disorders detection," *Journal of Informatics and Web Engineering*, vol. 3, no. 1, pp. 136–154, Feb. 2024. https://doi.org/10.33093/jiwe.2024.3.1.9

[11] J. Stenum, M. M. Hsu, A. Y. Pantelyat, and R. T. Roemmich, "Clinical gait analysis using video-based pose estimation: multiple perspectives, clinical populations, and measuring change," *PLOS Digital Health*, vol. 3, no. 3, 2024, Art. no. e0000467. https://doi.org/10.1371/journal.pdig.0000467

[12] X. Li, H. Xu, and J. T. Cheung, "Gait-force model and inertial measurement unit-based measurements: A new approach for gait analysis and balance monitoring," *Journal of Exercise Science & Fitness*, vol. 14, no. 2, pp. 60–66, Dec. 2016. https://doi.org/10.1016/j.jesf.2016.07.002

[13] H. Zhang, Y. Guo, and D. Zanotto, "Accurate ambulatory gait analysis in walking and running using machine learning models," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 1, pp. 191–202, Jan. 2020. https://doi.org/10.1109/TNSRE.2019.2958679

[14] L. Xiang, Y. Gu, Q. Mei, A. Wang, V. Shim, and J. Fernandez, "Automatic classification of barefoot and shod populations based on the foot metrics and plantar pressure patterns," *Frontiers in Bioengineering and Biotechnology*, vol. 10, Mar. 2022, Art. no. 843204. https://doi.org/10.3389/fbioe.2022.843204

[15] D. Adolph, W. Tschacher, H. Niemeyer, and J. Michalak, "Gait patterns and mood in everyday life: A comparison between depressed patients and non-depressed controls," *Cognitive Therapy and Research*, vol. 45, pp. 1128–1140, Feb. 2021. https://doi.org/10.1007/s10608-021-10215-7

[16] Y. Wang, J. Wang, X. Liu, and T. Zhu, "Detecting depression through gait data: examining the contribution of gait features in recognizing depression," *Frontiers in Psychiatry*, vol. 12, May 2021, Art. no. 661213. https://doi.org/10.3389/fpsyt.2021.661213

[17] Y. Wen, B. Li, X. Liu, D. Chen, S. Gao, and T. Zhu, "Using gait videos to automatically assess anxiety," *Frontiers in Public Health*, vol. 11, Mar. 2023, Art. no. 1082139. https://doi.org/10.3389/fpubh.2023.1082139

[18] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-based gait recognition: A survey," *IEEE Access*, vol. 6, pp. 70497–70527, Nov. 2018. https://doi.org/10.1109/ACCESS.2018.2879896

[19] M. H. Khan, M. S. Farid, and M. Grzegorzek, "Vision-based approaches towards person identification using gait," *Computer Science Review*, vol. 42, Nov. 2021, Art. no. 100432. https://doi.org/10.1016/j.cosrev.2021.100432

[20] C. Song, Y. Huang, W. Wang, and L. Wang, "CASIA-E: A large comprehensive dataset for gait recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 3, pp. 2801–2815, 2022. https://doi.org/10.1109/TPAMI.2022.3183288

[21] G. Zhao, G. Liu, H. Li, and M. Pietikainen, "3D gait recognition using multiple cameras," in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Southampton, UK, 2006, pp. 529–534. https://doi.org/10.1109/FGR.2006.2

[22] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *Applications and Science in Soft Computing*. Berlin, Germany: Springer, 2004, pp. 339–346. https://doi.org/10.1007/978-3-540-45240-9_46

[23] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, Feb. 2005. https://doi.org/10.1109/TPAMI.2005.39

[24] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *18th International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, China, Aug. 2006, pp. 441–444. https://doi.org/10.1109/ICPR.2006.67

[25] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, "The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition," *IEEE Transactions on Information*

Forensics and Security, vol. 7, no. 5, pp. 1511–1521, Oct. 2012. https://doi.org/10.1109/TIFS.2012.2204253

[26] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition," *IPSJ Transactions on Computer Vision and Applications*, vol. 10, no. 4, pp. 1–14, Feb. 2018. https://doi.org/10.1186/s41074-018-0039-6

[27] Z. Mu, F. M. Castro, M. J. Marín-Jiménez, N. Guil, Y.-R. Li, and S. Yu, "ReSGait: The real-scene gait dataset," in *2021 IEEE International Joint Conference on Biometrics (IJCB)*, Shenzhen, China, Aug. 2021, pp. 1–8. https://doi.org/10.1109/IJCB52358.2021.9484347

[28] M. Nieto-Hidalgo, F. J. Ferrández-Pastor, R. J. Valdivieso-Sarabia, J. Mora-Pascual, and J. M. García-Chamizo, "A vision based proposal for classification of normal and abnormal gait using RGB camera," *Journal of Biomedical Informatics*, vol. 63, pp. 82–89, Oct. 2016. https://doi.org/10.1016/j.jbi.2016.08.003

[29] D. Goyal, K. Rao Jerripothula, and A. Mittal, "Detection of gait abnormalities caused by neurological disorders," in *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, Finland, Sep. 2020, pp. 1–6. https://doi.org/10.1109/MMSP48831.2020.9287163

[30] A. Dadashzadeh, A. Whone, M. Rolinski, and M. Mirmehdi, "Exploring motion boundaries in an end-to-end network for vision-based Parkinson's severity assessment," *arXiv preprint arXiv:2012.09890*, Dec. 2020. https://doi.org/10.48550/arXiv.2012.09890

[31] R. Ranjan, D. Ahmedt-Aristizabal, M. A. Armin, and J. Kim, "Computer vision for clinical gait analysis: A gait abnormality video dataset," *arXiv preprint arXiv:2407.04190*, Jul. 2024. https://doi.org/10.48550/arXiv.2407.04190

[32] H. Dou, W. Zhang, P. Zhang, Y. Zhao, S. Li, Z. Qin, F. Wu, L. Dong, and X. Li, "VersatileGait: A large-scale synthetic gait dataset with fine-grained attributes and complicated scenarios," *arXiv preprint arXiv:2101.01394*, Jan. 2021. https://doi.org/10.48550/arXiv.2101.01394

[33] Y. Makihara *et al*., "Gait collector: An automatic gait data collection system in conjunction with an experience-based long-run exhibition," in *2016 International Conference on Biometrics (ICB)*, Halmstad, Sweden, Jun. 2016, pp. 1–8. https://doi.org/10.1109/ICB.2016.7550090

[34] M. Baldinger, L. M. Reimer, and V. Senner, "Influence of the camera viewing angle on OpenPose validity in motion analysis," *Sensors*, vol. 25, no. 3, Jan. 2025, Art. no. 799. https://doi.org/10.3390/s25030799

[35] X. Han, D. Guffanti, and A. Brunete, "A comprehensive review of vision-based sensor systems for human gait analysis," *Sensors*, vol. 25, no. 2, Jan. 2025, Art. no. 498. https://doi.org/10.3390/s25020498

[36] R. Raman, P. K. Sa, S. Bakshi, and B. Majhi, "Towards optimized placement of cameras for gait pattern recognition," *Procedia Technology*, vol. 6, pp. 1019–1025, 2012. https://doi.org/10.1016/j.protcy.2012.10.124

[37] M. Riberto, R. F. Liporaci, O. F. Vieira, and J. B. Volpon, "Setting up a human motion analysis laboratory: Camera positioning for kinematic recording of gait," *International Journal of Physical Medicine & Rehabilitation*, vol. 1, no. 4, pp. 1–3, 2013. https://doi.org/10.4172/2329-9096.1000131

[38] M. W. Whittle, "Gait analysis," in *The Soft Tissues*, G. R. McLatchie and C. M. E. Lennox, Eds. Oxford, U.K.: Butterworth-Heinemann, 1993, pp. 187–199. https://doi.org/10.1016/B978-0-7506-0170-2.50017-0

[39] H. Schulzrinne, A. Rao, and R. Lanphier, "Real time streaming protocol (RTSP)," *RFC 2326*, Apr. 1998. https://doi.org/10.17487/rfc2326

[40] S.-j. Sun and X.-d. Yang, "Method of CCD camera calibration based on OpenCV," in *2015 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration*, Wuhan, China, Dec. 2015, pp. 34–37. https://doi.org/10.1109/ICIICII.2015.96

[41] A. Dutta, R. Lal, D. S. Raychaudhuri, C.-K. Ta and A. K. Roy-Chowdhury, "POISE: Pose guided human silhouette extraction under occlusions," in *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, Jan. 2024, pp. 6141–6151. https://doi.org/10.1109/WACV57701.2024.00604

[42] K. H. Karstensen, "Silhouette extraction using graphics processing units," M.S. thesis, 2012.

[43] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91

[44] G. Jocher and J. Qiu, "Ultralytics YOLO11," GitHub repository, 2024. [Online]. Available: https://github.com/ultralytics

[45] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Hoboken, NJ, USA: Pearson, 2018.

[46] N. Ravi *et al*., "Sam 2: Segment anything in images and videos," *arXiv preprint arXiv:2408.00714*, Aug. 2024. https://doi.org/10.48550/arXiv.2408.00714

[47] L. C. Chen, "Rethinking Atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017. https://arxiv.org/pdf/1706.05587

[48] A. Kirillov et al., "Segment anything," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, Oct. 2023, pp. 3992–4003. https://doi.org/10.1109/ICCV51070.2023.00371

[49] S. Fritz and M. Lusardi, "Walking speed: The sixth vital sign," *Journal of Geriatric Physical Therapy*, vol. 32, no. 2, pp. 2–5, 2009. https://doi.org/10.1519/00139143-200932020-00002

[50] A. W. Andrews, S. A. Chinworth, M. Bourassa, M. Garvin, D. Benton, and S. Tanner, "Update on distance and velocity requirements for community ambulation," *Journal of Geriatric Physical Therapy*, vol. 33, no. 3, pp. 128–134, 2010.

Pınar Güner Şahan received her B.S. degree in Computer Engineering from Erciyes University, Turkey, in 2018, and her M.S. degree in Electrical and Computer Engineering from Abdullah Gül University, Turkey, in 2021. She is currently pursuing a PhD at the Department of Computer Engineering at Kocaeli University, Turkey, with a research focus on gait analysis using deep learning techniques. She also serves as a Lecturer at the Computer Application and Research Centre of the same university. Her research interests include computer vision, artificial intelligence, and bioinformatics.
E-mail: pinar.guner@kocaeli.edu.tr
ORCID iD: https://orcid.org/0000-0001-5979-0375

Ilgar Akkaya received his B.S. degree in Computer Engineering from Alanya Alaaddin Keykubat University, Turkey, in 2024. He is currently pursuing an M.S. degree at the Department of Computer Engineering at Kocaeli University, focusing on medical image segmentation using deep learning techniques. His research interests include computer vision and artificial intelligence.
E-mail: 245112058@kocaeli.edu.tr
ORCID iD: https://orcid.org/0009-0006-6242-9154

İbrahim Şahan received his B.S. and M.S. degrees in Computer Engineering from Erciyes University, Turkey, in 2018 and 2022, respectively. He is currently pursuing a PhD at the Department of Computer Engineering at Kocaeli University, Turkey. He also serves as a Research Assistant at the same department. His research interests include computer vision, artificial intelligence, and evolutionary computation.
E-mail: ibrahim.sahan@kocaeli.edu.tr
ORCID iD: https://orcid.org/0000-0001-6968-1770

Suhap Şahin received his PhD degree in Electronics and Communication Engineering from Kocaeli University, Turkey, in 2010. He is currently a Professor at the Department of Computer Engineering at Kocaeli University. His research interests include computer vision, pattern recognition, artificial intelligence, and computer hardware.
E-mail: suhapsahin@kocaeli.edu.tr
ORCID iD: https://orcid.org/0000-0003-1340-8972