Article

Gait Impairment Analysis Using Silhouette Sinogram Signals and Assisted Knowledge Learning

Mohammed A. Al-masni [1], Eman N. Marzban [2], Abobakr Khalil Al-Shamiri [3], Mugahed A. Al-antari [1], Maali Ibrahim Alabdulhafith [4], Noha F. Mahmoud [5], Nagwan Abdel Samee [4,*] and Yasser M. Kadah [6]

[1] Department of Artificial Intelligence and Data Science, College of Software & Convergence Technology, Sejong University, Seoul 05006, Republic of Korea; m.almasani@sejong.ac.kr (M.A.A.-m.); en.mualshz@sejong.ac.kr (M.A.A.-a.)
[2] Biomedical Engineering Department, Cairo University, Giza 12613, Egypt; eman.marzban@eng1.cu.edu.eg
[3] School of Computer Science, University of Southampton Malaysia, Iskandar Puteri 79100, Johor, Malaysia; a.k.n.al-shamiri@soton.ac.uk
[4] Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia; mialabdulhafith@pnu.edu.sa
[5] Rehabilitation Sciences Department, Health and Rehabilitation Sciences College, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia; nfmahmoud@pnu.edu.sa
[6] Electrical and Computer Engineering Department, King Abdulaziz University, Jeddah 22254, Saudi Arabia; ykadah@kau.edu.sa
*   Correspondence: nmabdelsamee@pnu.edu.sa

Abstract: The analysis of body motion is a valuable tool in the assessment and diagnosis of gait impairments, particularly those related to neurological disorders. In this study, we propose a novel automated system leveraging artificial intelligence for efficiently analyzing gait impairment from video-recorded images. The proposed methodology encompasses three key aspects. First, we generate a novel one-dimensional representation of each silhouette image, termed a silhouette sinogram, by computing the distance and angle between the centroid and each detected boundary points. This process enables us to effectively utilize relative variations in motion at different angles to detect gait patterns. Second, a one-dimensional convolutional neural network (1D CNN) model is developed and trained by incorporating the consecutive silhouette sinogram signals of silhouette frames to capture spatiotemporal information via assisted knowledge learning. This process allows the network to capture a broader context and temporal dependencies within the gait cycle, enabling a more accurate diagnosis of gait abnormalities. This study conducts training and an evaluation utilizing the publicly accessible INIT GAIT database. Finally, two evaluation schemes are employed: one leveraging individual silhouette frames and the other operating at the subject level, utilizing a majority voting technique. The outcomes of the proposed method showed superior enhancements in gait impairment recognition, with overall F1-scores of 100%, 90.62%, and 77.32% when evaluated based on sinogram signals, and 100%, 100%, and 83.33% when evaluated based on the subject level, for cases involving two, four, and six gait abnormalities, respectively. In conclusion, by comparing the observed locomotor function to a conventional gait pattern often seen in healthy individuals, the recommended approach allows for a quantitative and non-invasive evaluation of locomotion.

Keywords: gait disorders; gait analysis; deep learning; assisted knowledge learning; silhouette images

1. Introduction

Gait is the way an individual walks. It is the locomotion attained through the movement of limbs influenced by the integrated functioning of the musculoskeletal and nervous systems [1]. Any change in this distinctive sequential walking pattern can provide important clues about the state and development of a wide range of underlying health problems [2]. One reliable way to detect declining health symptoms caused by aging,

physical dysfunction, or neurodegenerative conditions is by observing gait. Such disease and disorders include, among others, Parkinson's disease, multiple sclerosis (MS), and strokes [3–6]. Clinicians analyze the gait patterns of patients using various types of data, such as temporal–spatial, kinematic, kinetic, or electromyography data. By doing so, they can accurately classify these gait patterns through clinically significant groups and develop suitable interventions accordingly [7,8].

When gait pattern and characteristics can be reliably identified in the clinical setting and then tracked and evaluated over time, it opens the door to more effective personalized therapy and more accurate predictive outcome evaluation [2]. It was reported by [9] that understanding clinical rehabilitation and machine learning is crucial, as shown by the evaluation of future research. There has been a recent paradigm shift towards personalized medicine, which considers the uniqueness of each patient. The use of instrumented gait analysis is being extensively employed in the assessment of patients, serving as a valuable tool in clinical decision-making processes and the evaluation of various treatment outcomes [10].

Clinical gait analysis is a multi-dimensional procedure that entails the comprehensive examination and analysis of data derived from various sources [11]. Neurologists and rehabilitation therapists are responsible for conducting various diagnostic examinations to evaluate and manually quantify gait disorders. The development of an automated system capable of identifying individuals with impaired mobility based on their gait patterns holds great promise for enhancing health care and various other applications related to gait analysis [12–16]. Physical examinations, the collection of patient history, and three-dimensional motion analysis collectively provide valuable data pertaining to the patient's gait impairments. Identifying patterns and trends within this extensive dataset can pose challenges. Nevertheless, the interpretation of clinical gait analysis results and subsequent treatment decisions are predominantly based on subjective pattern recognition by clinicians. Hence, the development of a method to accurately classify gait data could be beneficial in identifying cohorts of individuals sharing widespread underlying clinical conditions.

The utilization of automated recognition techniques for human gait patterns enables several valuable applications. Firstly, it facilitates a quantitative and non-invasive diagnosis of locomotion by comparing the observed locomotor function to a standard gait pattern associated with healthy individuals [17,18]. Secondly, it facilitates the identification of personalized gait training tasks by automatically adapting assistance based on the individual's recognized motor function [15]. Thirdly, it aids in the planning of potential treatment strategies tailored to the specific needs of the user, considering their previously identified gait impairment through automated recognition approaches [19]. Lastly, it enables the measurement and description of the progress made in gait treatment by assessing the individual's gait patterns during initial and follow-up sessions [7,20]. Moreover, these automated systems utilized in clinical gait analysis represent an impartial method for the extensive manipulation of gait data, and they offer enhanced efficiency and cost effectiveness compared to the traditional procedures commonly employed by clinicians [21].

The existing research on gait analysis can be broken down into two groups: model-based and appearance-based techniques. In order to obtain reliable pose parameters, model-based approaches often use a motion capture device or wearable sensors. However, model-based techniques require controlled laboratory conditions and subject cooperation; subjects must wear specialized equipment or gadgets in order to obtain accurate measurements of the trajectories of their joints and other moving body parts. Additionally, they typically endure high computation costs and poor pose estimation quality. On the other hand, appearance-based solutions employ captured videos directly; this eliminates the requirement for individuals to wear any specialized equipment. The method relies on analyzing a person's gait in order to determine things like their gender and verify their identity; the gait energy image (GEI) [22] is one such feature that has the potential to be employed for gait classification due to its strong authentication performance for individuals based on gait. In addition to its other benefits, the GEI does not necessitate high-quality

silhouette extraction and is hence highly resilient against the noise that is invariably present in extracted silhouette images. Due to its definition as the mean of successive silhouettes over a certain walking period, the GEI is well suited to encoding the shape of people and thus could prove useful for verification purposes. Despite the remarkable performance exhibited by the GEI feature in gait classification, it relies on prior knowledge of gait cycle sequences and requires the use of a two-dimensional classification model. To solve this problem, the proposed study seeks to investigate the effectiveness of alternative image features, namely silhouette sinograms, which bear resemblance to the GEI feature, in detecting gait impairments. To utilize these features, a 1D classification model that does not rely on gait cycle information is required. The following are the novel contributions achieved through this study.

1. Novel silhouette sinogram signals are generated for each silhouette image by computing the angle and distance between the centroid and all boundary points, facilitating the extraction of prominent representations of relative variations in pathological gait patterns.
2. We develop a 1D CNN model that can integrate multiple silhouette sinogram signals and train them together to enhance the network's understanding of gait abnormalities through assisted knowledge learning.
3. We analyze the influence of varying k-assisted silhouette sinogram signals, used as prior knowledge and fed into the network (k = 1, 10, 20, and 30). This work also conducts comprehensive experiments encompassing three distinct sizes of gait abnormalities, categorized into n = 2, 4, and 6 classes.
4. This work employs two evaluation methodologies: one based on individual silhouette frames and another operating at the subject level using a majority voting approach.

Quantitative gait analysis has been around for over 30 years, although it is still mostly used in labs and not much in clinical practice. This is mainly because clinical teams have varying levels of training, experience, and preferences, and because typical gait labs include expensive and heavy equipment, complicated protocols, and data management and analysis. Despite evidence of the decreased efficacy of quantifying gait, observational gait and qualitative measures are still employed extensively in clinics [2]. We believe that this gait representation method has not been proposed in the literature before. It offers the opportunity to utilize low-cost camera-based gait data collection methods to obtain detailed information about the movement of different parts of the body, similar to what sensor-based methods offer. Furthermore, the computational complexity is very modest, allowing for real-time implementation.

The rest of this paper is outlined as follows: Section 2 introduces related works and investigations for gait impairment analysis. Section 3 describes our proposed methodology, including the generation of silhouette sinogram signals and the developed 1D CNN model. Section 4 presents the experimental results and comparisons. Sections 5 and 6 discuss the findings and draw conclusions, respectively.

2. Related Works

The application of movement analysis has demonstrated its utility in comprehending the underlying mechanisms of various pediatric diseases and enhancing their subsequent monitoring and evaluation [23]. The utilization of a particular intervention in the management of cerebral palsy has experienced significant prevalence in recent years. According to Hebda-Boon et al. [24], the implementation of this intervention within hospital settings has led to a decrease in the number of surgical treatment recommendations. Additionally, it has been employed to enhance rehabilitation in diverse pediatric conditions by emphasizing the osteoarticular hazards caused by excessive trunk tilt motions [25]. Furthermore, the integration of motion analysis with X-ray imaging has been employed to enhance our comprehension of the correlation between alterations in skeletal structure and corresponding modifications in individuals with cerebral palsy [23] as well as those with diverse postural abnormalities [26].

We can categorize the current body of literature on gait analysis into two distinct categories: model-based approaches and appearance-based approaches. Typically, model-based approaches employ either a motion capture system [27–30] or wearable sensors [27,31–35] to accurately capture pose parameters. In their study, Hallemans et al. [36] employed a three-dimensional motion capture system to assess head orientation, stride length, and trunk flexion. Their objective was to investigate the potential impact of impaired vision on the dynamic stability of walking. Several model-based studies conducted between 2015 and 2020 have used camera-recorded movies to estimate pose parameters. These studies achieved this by employing fitting and tracking techniques on various body components. In their study, Khan et al. [37] utilized a skeletal model to analyze binary human silhouettes. This analysis allowed them to determine the posture lean and stride cycles of individuals, with the aim of detecting Parkinsonian gait. Costin et al. [38] have proposed a novel automated image processing technique for the assessment of human gait in patients undergoing rehabilitation. The present study assessed the atypical gait patterns exhibited by patients as a result of lower limb injuries. They used the MoveNet Thunder model to determine the positions of joints and key points in a human's body based on video data. The magnitudes as well as variations in the angles resulting from the symmetrical articulations of the human body are assessed in different frames, following the identification of specific body parts. The purpose of this evaluation is to analyze and monitor these angles throughout the rehabilitation process. The concept of a rehabilitation degree was established and quantified as the mean percentage of the pertinent angles of the body segments relative to the domains of normality for those angles. Based on the encouraging outcomes obtained from the utilization of these two model-based methodologies, it is affirmed that the identification of disordered gait necessitates the consideration of crucial factors such as the walking posture, temporal cue, and stability property. Indeed, model-based methodologies have the capability to accurately quantify the trajectories of mobile joints and bodily elements. However, their implementation necessitates controlled laboratory conditions and the active participation of the subjects, who are required to don specialized attire or apparatus. Moreover, it is common for these systems to exhibit subpar accuracy in estimating poses and require significant computational resources. Therefore, these methods are not suitable for the scheme we aim to accomplish.

In contrast, appearance-based methodologies rely on the direct utilization of recorded videos, thereby eliminating the necessity for subjects to use any specialized equipment. Chen et al. [13,14] employed binary silhouettes derived from color images as a means of differentiating individuals diagnosed with Parkinson's disease from those without the condition. Their study used binary images directly for classification purposes, making the results significantly dependent on the caliber of each silhouette. Previous studies [22,39,40] have attempted to analyze gait observations in order to identify certain characteristics that can be utilized for the purposes of gender recognition and human authentication. One frequently employed feature in gait analysis is the gait energy image, which has gained recognition for its ability to effectively authenticate individuals based on their gait patterns. One additional benefit of the GEI method is its ability to handle noise present in extracted silhouette images without the need for high-quality silhouette extraction. Given that GEIs are characterized as the means of consecutive silhouettes within a walking interval, it possesses a strong capability to accurately represent the physical form of individuals. Consequently, it is reasonable to assert that the GEI is a highly efficient method for the identification of pathological gait impairments from 2D camera-based and video-recorded images.

Gait recognition is a contemporary biometric technique that has gained prominence for its potential applications in security. It is characterized by its unobtrusive learning method. Bakchy et al. [41] employed the GEI feature for the purpose of gait recognition in security-related applications. The GEI effectively retains both the dynamic and static characteristics inherent in a gait sequence. The static information pertaining to the human body encompasses its physical appearance and shape, while the dynamic information encompasses fluctuations in frequency and phase. Nevertheless, the temporal aspect that

normalizes each silhouette within the GEI is not taken into account. In response to this issue, Luo et al. [42] suggested the cumulative frame difference energy image, which can reflect the time properties. Subsequently, gait recognition was achieved by employing the nearest neighbor classifier, which relied on the Euclidean distance metric. The experimental findings indicate that the proposed algorithm exhibits superior performance compared to the hybrid GEI and 2D-PCA methods, while also satisfying the real-time constraints. The GEI feature has demonstrated remarkable efficacy in gait classification, prompting us to improve the accuracy of current classification models by harnessing the capabilities of convolutional neural networks. One potential approach to achieve this objective involves employing sinogram images as an alternative to GEIs.

In the work of Chen et al. [13], they used an LDA as a feature extractor, followed by a minimum distance classifier (MDC) to differentiate patients with Parkinson's from normal subjects using video sequences as the input, and reached an overall accuracy of 95.49%. Their dataset comprised seven healthy subject and seven PD patients having video sequences of around 2.57 ± 0.5 min. In 2018, Verlekar et al. [43] used SVM to classify four gait patterns using the INIT. The involved features were body-related features such as the torso orientation and the shift in the center of gravity, and feet-related features such as the step length and the ratio between the foot flat period to the stance period. The parameters involved in their system involved the entire video sequence for both the training and the testing phases, and they reached an accuracy of 98.8%. In 2021, Albuquerque et al. [44] created their own simulated dataset named GAIT-IT and trained their CNN on it. They used GEIs and achieved an accuracy of 93.4% and 89.8% when testing on GAIT-IT and GAIT-IST [45], respectively. They included five gait types for classification purposes, namely, normal, scissor, spastic, steppage, and propulsive, which differ from the pathologies included in the INIT dataset. The same team, in their published work [46], achieved an accuracy of 96.5% and 91.4% when testing on GAIT-IT [44] and GAIT-IST [45], respectively. In this work, they selected nine key frames per gait cycle, then used a pretrained CNN for spatial feature extraction and a bidirectional LSTM for temporal feature extraction. Finally, the classification was performed by two fully connected neural network layers. In the studies of [13,43,44,46], the accuracy was the sole evaluation metric computed.

The current state of the art for predicting gait impairment in medical applications lacks advanced deep learning methods such as Federated Learning (FL). FL has emerged as a promising technique in the field of machine learning, particularly within the realm of mobile health (mHealth) [47]. Unlike conventional machine learning methods, FL enables collaborative model training across multiple data sources without requiring direct access to individual datasets. This approach is especially beneficial in the context of mHealth, where medical data are often fragmented across isolated silos, and privacy concerns pose significant barriers to data sharing. As a result, FL holds great potential for advancing research and innovation in the field of mHealth, offering a privacy-preserving solution for leveraging distributed data sources effectively. This approach, explored by Zhang et al. [48,49] for tasks like predicting the remaining useful life, holds promise for gait impairment prediction as well.

3. Materials and Methods

3.1. Dataset

Gait difficulties refer to any deviation from the standard walking or gait pattern. Neurological etiologies exhibit a higher prevalence compared to non-neurological etiologies. There are a multitude of diseases that impact on both the central and peripheral nervous systems, resulting in subsequent alterations to an individual's gait. Gait disturbances may arise as a result of subtle alterations in the interactions between these two systems. The examination should encompass the following aspects: Posture, the act of being in an upright position, with one's weight supported by one's feet, refers to the position and alignment of the body, particularly in relation to the spine and limbs. It encompasses

various aspects, such as narrow or wide stances. Additionally, the examination involves the evaluation of the walking process and the measurement of the step length. The gait cycle consists of two primary phases: stance, during which a specific foot is in contact with the ground, and swing, during which the same foot is no longer in contact with the ground and is propelling forward. In the context of typical human locomotion, it is anticipated that the stance and swing phases will account for approximately 62% and 38% of a complete gait cycle; accordingly, Figure 1 depicts the distribution of this phenomenon, as observed from the perspective of the right limb, throughout an entire gait cycle. In this research, a rehabilitation specialist examined several video-based indicators meant to distinguish between gait phases and instances.

Gait phases	Phase of the stance (62%)				Phase of the swing (38%)		
Contralateral Instances	Foot strike	Foot off	Swing phase	Foot strike	Stance phase	Foot strike	Foot off
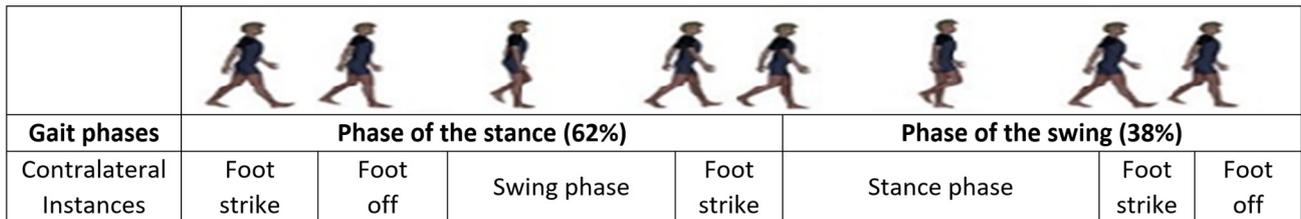
Figure 1. An example for a representation of a series of standardized binary silhouettes.

In this investigation, we made use of an already existing data collection called the INIT Gait database [5,50]. This dataset is made up of sequences of high-quality binary silhouettes that were extracted from RGB movies that were recorded at the specialized studio LABCOM at University Jaume I. LABCOM is a part of the audiovisual facilities at the university. A group of ten individuals, consisting of nine males and one female, were recruited as healthy volunteers to participate in a study including the simulation of several aberrant gait styles on a green chroma background. The utilization of a consistent background aided the process of converting the frames into binary form and extracting silhouettes of superior quality. This, in turn, decreased the level of uncertainty when assessing the precision of features. A total of seven different gait types were mimicked, involving modifications to both limb movement and overall body posture. The individuals in question draw inspiration from problematic gait patterns that are commonly associated with specific neurological disorders, such as Parkinson's disease. Additionally, a novel eighth style characterized by natural and uncontrived movement was incorporated. Each participant was recorded twice for each gait pattern, and all sequences were obtained from a lateral perspective, enabling a more detailed description of limb action and body position. Hence, a total of 160 sequences were captured, including a collective count of 41,104 silhouette images. Each of these images has a frame size of 400 × 800 pixels. The following table, Table 1, provides a summary of the gait styles contained in the INIT Gait database.

Table 1. Description of INIT Gait database walking patterns.

Gait Style	Description
Normal (NM)	It depicts the normal gait pattern of a healthy person, which is also referred to as neutral or regular appearance in the database.
Abnormal right leg (RL)	A pattern of walking in which the right leg takes steps that are approximately one half as short as the ones taken by the left leg.
Abnormal left leg (LL)	A pattern of walking in which the left leg takes steps that are approximately one half as short as the steps taken by the right leg.
Abnormal full body (FB)	A damaged gait pattern in which the complete body shows a multitude of abnormal gait symptoms, including the following: participants walk slowly, bending their knees, and taking extremely tiny steps, barely lifting their feet off the ground (shuffling gait).

Table 1. *Cont.*

Gait Style	Description
Abnormal half right arm (RA-0.5)	A walking pattern in which the right arm swings about one half as far as the left arm does.
Abnormal half left arm (LA-0.5)	A walking pattern in which the left arm swings about one half as far as the right arm does.
Abnormal right arm (RA)	A walking pattern in which the right arm does not swing at all while walking.
Abnormal left arm (LA)	A walking pattern in which the left arm does not swing at all while walking.

3.2. Proposed Framework

This study proposes a new pipeline for gait impairment analysis, consisting of three main stages. Firstly, silhouette sinogram signals are generated utilizing a novel angular centroid distance algorithm. Subsequently, a one-dimensional convolutional neural network (1D CNN) model is developed and trained by integrating shared information through assisted knowledge learning. Finally, two evaluation schemes are employed: one based on individual silhouette frames and another based on the subject level, utilizing a majority voting approach. The diagram scheme of the proposed framework is illustrated in Figure 2.
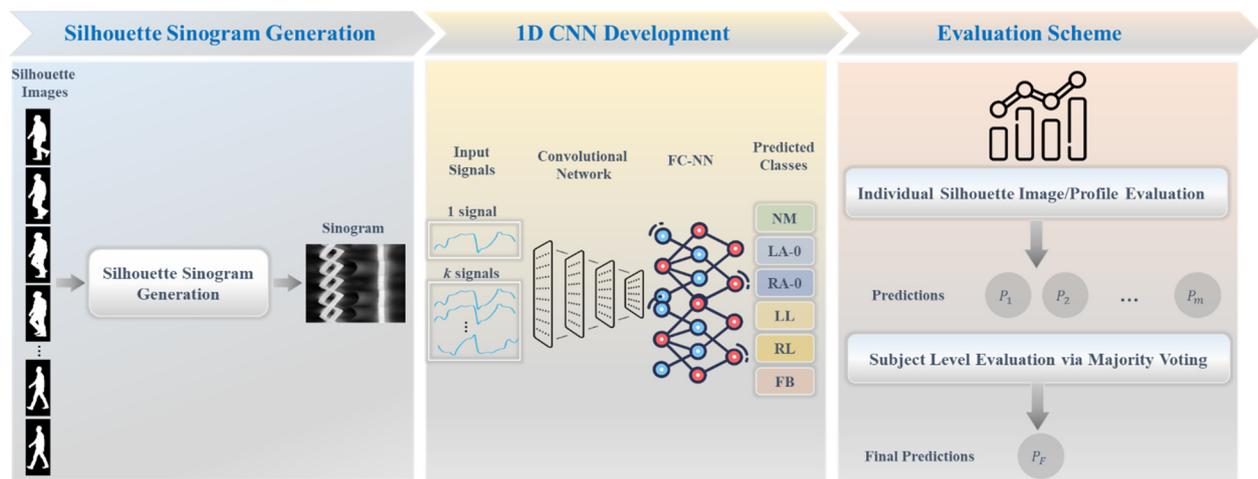

Figure 2. Overview diagram of the proposed framework, including silhouette sinogram generation from silhouette images, development of 1D CNN, and two evaluation schemes.

3.2.1. Generation of Silhouette Sinogram Signals

(1)   Image Cropping

The first step is to segment and crop all binary silhouettes in each subject's gait cycle so that they fit within the same image size. This is achieved by projecting the original silhouettes in the horizontal and vertical directions, then searching for nonzero elements within these projections to find the start and end points as the minimum and maximum of nonzero element positions. Given an initial binary silhouette image n in the sequence $S_n(x, y)$, the procedure can be expressed mathematically as follows:

$$S_{nx}(y) = \sum_x S(x, y) \text{ and } S_{ny}(x) = \sum_y S(x, y). \tag{1}$$

Then, the cropped image $S_n^c(x, y)$ can be obtained as the subset of rows and columns where the projections in the x and y directions are both nonzero, such that

$$S_n^c(x, y) = \{S_n(x, y)\}_{S_{nx}(y) > 0 \text{ and } S_{ny}(x) > 0}. \tag{2}$$

The cropped image sizes N_{nx} and N_{ny} are obtained as the cardinality of all nonzero elements in the projection vectors $S_{ny}(x)$ and $S_{nx}(y)$, which can be calculated as the 0-norm of each:

$$N_{nx} = \left\|S_{ny}(x)\right\|_0, \ N_{ny} = \left\|S_{nx}(y)\right\|_0. \tag{3}$$

Each binary silhouette is cropped in both the horizontal and vertical directions to contain only the points from the start and end point locations. This results in a sequence of binary silhouette images of variable horizontal and vertical sizes. In order to make all of them have the same size, all images are zero-padded equally from both directions to the maximum vertical and horizontal sizes of all nonuniformly sized cropped images to obtain uniformly sized cropped binary silhouette images. The common binary silhouette sizes in a particular sequence are obtained as follows:

$$N_x = \max_n\{N_{nx}\}, \ N_y = \max_n\{N_{ny}\}. \tag{4}$$

The result of zero-padding all binary silhouette images in the sequence to the common sizes N_x and N_y is termed $S_n^{cz}(x,y)$ for the image n in the sequence. This will ensure consistency in the subsequent movement assessment steps for the different silhouettes in the gait sequence.

(2)   Centroid and Boundary Calculations

For each gait sequence, the centroid location (c_{nx}, c_{ny}) of each of the uniformly sized cropped binary silhouette images $S_n^{cz}(x,y)$ is calculated as the location of the center of mass of each assuming uniform mass density everywhere in the silhouette region. This can be expressed mathematically as

$$c_{nx} = \frac{\sum_y \sum_x x \, S_n^{cz}(x,y)}{\sum_y \sum_x S_n^{cz}(x,y)}, \ c_{ny} = \frac{\sum_y \sum_x y \, S_n^{cz}(x,y)}{\sum_y \sum_x S_n^{cz}(x,y)}. \tag{5}$$

Then, uniformly sized cropped binary silhouette images undergo a boundary tracing operation to obtain an ordered sequence of points that represent the silhouette boundary. This is more sophisticated than edge detection in that the boundary is given as an ordered sequence of boundary image points rather than a binary contour image, particularly with complex silhouette region shapes. This can be performed using a number of existing approaches such as pixel-following- or vertex-following-based methods, where the basic assumptions of having input binary images that are zero-padded with no boundary points at the edge of the image matrix. In this study, the boundary tracing procedure was implemented based on the Moore neighbor tracing algorithm modified by Jacob's stopping criteria [51]. This starts with a boundary point (preferably the uppermost and leftmost one) and proceeds to examine its eight neighbors one by one in the same direction (for example, clockwise) starting from a neighbor that has a zero value (lies in the background) and selecting the first neighbor that has a one value (lies within the silhouette) to be the next point in the boundary and sequence. This process is repeated until one of Jacob's stopping criteria is met whereby the sequence of boundary points reaches the starting point. The result of this boundary detection operation is a list of boundary points' horizontal and vertical locations $\left\{\vec{b}_{nx}, \vec{b}_{ny}\right\}$ for each cropped binary silhouette image $S_n^{cz}(x,y)$. It should be noted that the size of such boundary sequences will vary across different images within the same sequence.

(3)   Silhouette Sinogram Calculations

In this part, the detected boundaries $\left\{\vec{b}_{nx}, \vec{b}_{ny}\right\}$ of uniformly sized cropped binary silhouettes $S_n^{cz}(x,y)$ along with their centroids (c_{nx}, c_{ny}) are used to generate a novel one-dimensional representation of each silhouette, which we will term a silhouette sinogram.

This is achieved by computing the vectors of distances and angles $\left\{ \overrightarrow{D}_n, \overrightarrow{A}_n \right\}$ between the centroid and each of the silhouette boundary points, as shown in Figure 3. The mathematical representation of these calculations for the point *m* in the boundary sequence are as follows:

$$D_{nm} = \left\| \begin{bmatrix} c_{nx} \\ c_{ny} \end{bmatrix} - \begin{bmatrix} b_{nmx} \\ b_{nmy} \end{bmatrix} \right\|_2, \quad A_{nm} = \tan^{-1}\left(\frac{b_{nmy} - c_{ny}}{b_{nmx} - c\_nx} \right). \tag{6}$$
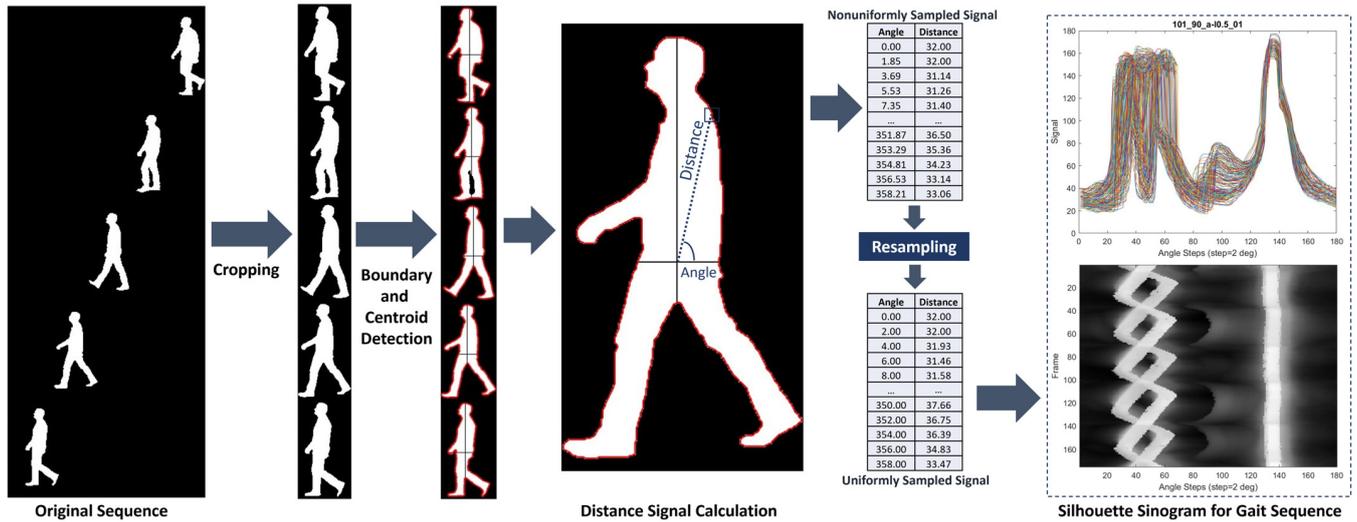

Figure 3. Schematic diagram of the generation process of the silhouette sinogram.

It should be noted that this representation is a generalization of the so-called signature of the 2D boundary [51], where the angles included in this representation are nonuniformly sampled in this work. We observe that this distance is, in essence, a projection of the silhouette in the angle that the line between each boundary point and the centroid makes with the horizontal direction. Such a one-dimensional representation of the two-dimensional silhouette in terms of projections bears similarity to the sinograms encountered in computed tomography. Hence, such a distance versus angle representation will be termed the silhouette sinogram in reference to such similarity.

Since the result of the above step is a set of nonuniformly sampled angles and their corresponding distances, issues related to the multiplicity of boundary points at the same angle as well as the comparison of the distance profiles across different silhouettes in the sequence may arise. Given that the same angle can have multiple distances, such as those corresponding to the proximal and distal boundary points of an extended arm, the algorithm should be able to resolve such multiplicity occurrences when detected to have only one distance value for each angle. This is achieved by sorting the set of angles and their corresponding distances by the angle values, the detection of multiplicity for each angle, and selecting the maximum distance for each angle from the different distance values. This ensures that the maximum extension information is kept given its greater relevance to the classification. The above steps can be expressed mathematically as follows:

$$\{A'_{nm}\} = \operatorname*{Sort}_{m}(\{A'_{nm}\}). \tag{7}$$

$$D'_{nm} = \max(\{D_{nm}\} \ \forall \ D_{nm} \text{ with similar } A_{nm}). \tag{8}$$

Subsequently, the set of angles and their corresponding distances now contain consistent entries that are repeated whenever multiplicity is encountered. Therefore, this set should be reduced to include only distinctive angle/distance pairs $\left\{ \vec{D'}_n,\ \vec{A'}_n \right\}$, such that

$$\vec{D'}_n = \left\{ A'_{nm} \right\}_{\neq}\ ,\quad \vec{A'}_n = \left\{ D'_{nm} \right\}_{\neq}. \tag{9}$$

Here, the common mathematical notation of $\{\,.\,\}_{\neq}$ is used to express the selection of unique elements within the enclosed set.

Furthermore, given that the resultant angle values are not uniformly sampled and that such sampling differs from one silhouette to another in the same sequence, resampling of the distance versus angle data is performed. The outcome of this resampling allows the final silhouette sinogram to be obtained as the distance versus angle data uniformly sampled with a specified resolution ΔA from 0 to 360 degrees. This can be achieved using 1D interpolation using linear or cubic interpolation methods to obtain the resampled sinogram signal $\vec{D^r}_n$. Here, linear interpolation was used. This is mathematically expressed as follows:

$$\vec{D^r}_n = \underset{\vec{A^r}_n = \{0,\ \Delta A,\ 2\Delta A,\ \dots\}}{\text{Resample}} \left(\left\{ \vec{D'}_n,\ \vec{A'}_n \right\} \right). \tag{10}$$

In this study, a 2-degree resolution was found to be robust in reducing small variations in close points while being adequate for maintaining the details of sinogram representation. Furthermore, the selection ensures that the resampling procedure will not encounter an aliasing problem given that the original density of the sampling will be more than or equal to the desired uniform sampling.

Once the silhouette sinograms are obtained from the gait sequence, we have a set of signals with each representing the distance versus angle profile for a silhouette at a certain point in time. The silhouette sinograms can be displayed as overlapping thin line plots or as an image where the distance is coded as intensity. A diagram showing the process of obtaining the silhouette sinograms is shown in Figure 3. It should be noted that the silhouette sinograms show marked differences between the different gait classifications as shown in Figure 4 and hence have significant potential for use in computer-aided diagnosis.
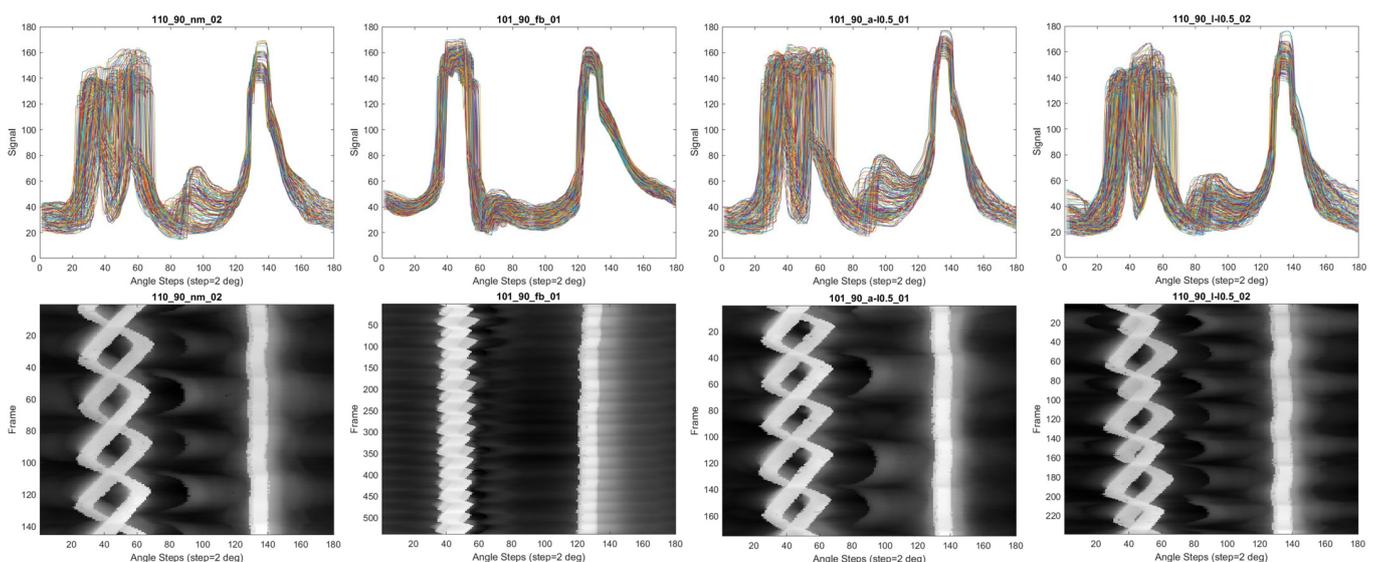

Figure 4. Silhouette sinogram signals for different gait patterns. Examples of normal (NM), full-body severe gait impairment (FB), motionlessness in the left arm, and left leg are shown from left-to-right, respectively.

It can be observed that the variations in such silhouette sinograms along the sequence of images in the gait cycle provide data that are similar to those of the more expensive sensor-based gait data acquisition setups. It should also be noted that silhouette sinograms can be normalized such that their variations across subjects of different height or body shape can be minimized whereas relative variations in motion at different angles are reliably used to detect gait classification.

3.2.2. Deep Learning Network Architecture

The main objective of the proposed deep learning network is to effectively differentiate between various types of gait abnormalities. We have adapted and introduced a 1D CNN architecture, inspired by the success of 2D CNNs in various image classification tasks in both natural and medical image domains. This architecture aims to learn and extract distinctive features of pathological gaits associated with each specific gait pattern. This adaptation allows for a more accurate assessment and evaluation of gait impairment, ultimately enhancing the analysis of gait-related abnormalities.

The aim behind the design of a 1D CNN structure is to seamlessly integrate it with the input data, denoted as x_i, which was generated in the aforementioned section as 1D silhouette sinogram signal information for each individual ith silhouette image. The proposed network architecture involves an encoder comprising four convolutional resolution blocks of varying sizes and eventually attached to fully connected neural networks (FC-NNs). Within each block, we employ two consecutive 1D convolutional layers, accompanied by batch normalization and rectified linear unit (ReLU) activation. In place of traditional subsampling pooling layers between these blocks, we employ stride convolutions with a value of 2. This approach promotes the further learning of pooling operations while enhancing the overall stability of the network [52]. Specifically, our convolutional filters consist of feature maps with dimensions of 32, 64, 128, and 256, kernel sizes of 7, 5, 5, and 5, and stride values of 1 for the four convolutional blocks, respectively. To further regularize the network training and avoid overfitting, a dropout layer with a probability of 0.3 is incorporated at the top of each block. Note that these convolutional encoder layers are responsible for automatically extracting and learning gait features directly from the input silhouette sinogram signals. The structural overview of the proposed 1D CNN is visually depicted in Figure 5.
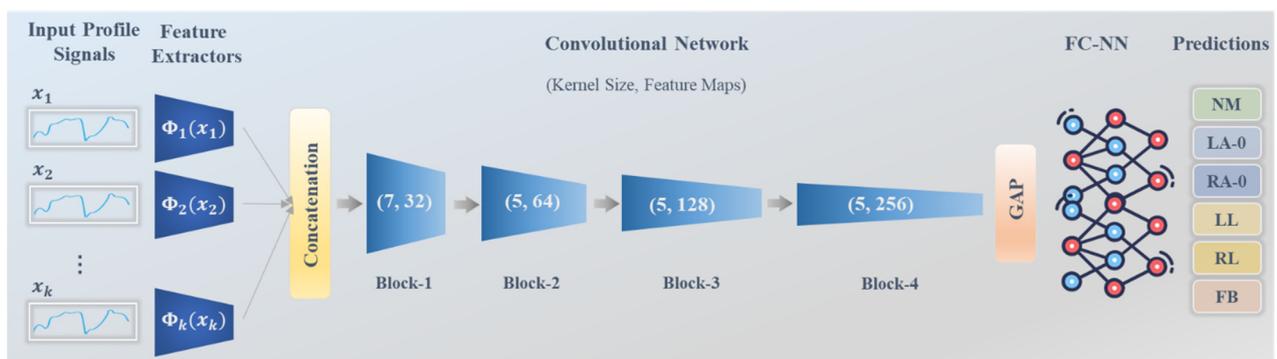

Figure 5. Schematic diagram of the proposed 1D CNN that involves multiple inputs serving as collaborative assisted knowledge learning.

Afterwards, these representation patterns are fed into a global average pooling (GAP) layer, which computes a single piece of spatial information from each feature map. This pooling operation aids in preventing overfitting by eliminating the need to optimize additional parameters, resulting in a more robust and generalized model. The extracted prominent features are subsequently passed into three dense FC-NN layers with 64, 16, and n neurons, respectively. Here, the value of n corresponds to the number of classes representing the different gait impairment abnormalities under investigation. The final

dense layer is connected to the Softmax activation function, facilitating the classification process. It is noteworthy that this paper investigates three distinct sizes of abnormalities, which are categorized into n = 2, 4, and 6 classes. In other words, three separate networks are trained and tested, differing solely in the number of classes specified in the last dense FC-NN layer.

One significant contribution of the proposed network architecture lies in its capability to acquire and learn additional assisted knowledge from the consecutive frames or silhouettes of each subject's data. This aspect is explained in detail in the following subsection, highlighting the network's capacity to leverage temporal information for enhanced learning and analysis.

3.2.3. Assisted Knowledge Learning

The proposed network is initiated by incorporating the subsequent silhouette sinogram signals of silhouette frames along with the current x_i input. This integration was achieved through the re-design of the 1D CNN, enabling the passage of multiple inputs into the network instead of a single signal. The objective of this process is to emulate the provision of essential knowledge about the gait cycle, thereby enhancing the network's ability to discern gait abnormalities more effectively. Specifically, multiple k-assisted silhouette sinogram signals are utilized as prior knowledge, derived from the same subject exhibiting a similar pattern. This assistance aids the network in making more accurate diagnostic decisions regarding the main x_i input. It is of note that this process represents a significant improvement compared to using only a single sinogram signal x_i.

The process of collaboratively assisted knowledge learning is accomplished through the sharing of intelligence from subsequent x_i sinogram signals. In regard to implementation, this process involves performing the following mathematic computations:

$$X = Concat\langle \Phi(x_1), \Phi(x_2), \ldots, \Phi(x_k)\rangle, \tag{11}$$

where $\Phi(x_i)$ represents the feature extractor for each consecutive sinogram signal, utilizing a separate convolutional block with a feature map of 32 and a kernel size of 7. $Concat\langle \cdot \rangle$ refers to the concatenation process of all the generated features derived from the shared information across k signals. Consequently, this collective knowledge, denoted as X, is then utilized as an input for the designed 1D CNN model. In this study, we investigate multiple values of k, specifically, k = 10, 20, or 30.

By incorporating this shared information into the model, we aim to leverage the accumulated insights from multiple silhouette sinogram signals to enhance the network's understanding of gait abnormalities. This collaborative assisted knowledge learning approach allows the network to capture a broader context and temporal dependencies within the gait cycle, enabling the more accurate diagnosis and classification of abnormalities. Through our investigation of various values for k, we seek to determine the optimal amount of shared information that yields the best performance in distinguishing gait abnormalities. This process is applicable to all input sinogram signals within both the training and testing datasets. Thus, the network is equipped with a single essential sinogram signal to evaluate, accompanied by k − 1 supportive knowledge derived from the subsequent silhouette frames. This design results in a multiple input single output model, wherein at any given time, the network generates a single decision pertaining to the essential input signal. By adopting this approach, we ensure that the network's classification decisions are based on a comprehensive understanding of the gait cycle, leveraging both the individual silhouette sinogram signals and the collective information provided by the subsequent frames. Figure 5 illustrates the concept of assisted knowledge learning.

3.2.4. Technical Specifications

This study employs a supervised learning approach to train the proposed 1D CNN. This methodology involves tuning and optimizing the network parameters through forward and backward propagation during training epochs. The optimization process compares the

early predictions of the network with reference true labels. The implementation details of the proposed work are outlined as follows. A learning rate of 0.003 is initially set, and then exponentially decreased by a factor of 10 until the 50th epoch. To ensure the inclusion of all frames or silhouette images of one gait cycle, a batch size of 50 is utilized. The networks are optimized using the Adam optimizer with the categorical cross-entropy loss function (L_{CE}). The L_{CE} is defined as

$$L_{CE} = -\sum_{i=1}^{i=N} y_i \cdot log(\hat{y}_i), \tag{12}$$

where y_i and $\hat{y}_i$ are the true label and predicted results from the proposed network, respectively. The summation over N represents the integration of cross-entropy loss for all classes.

The system implementation for this work was carried out on a PC equipped with a Cuda-enabled GPU, specifically the NVIDIA GeForce RTX 3080, and had 64 GB of RAM. The implementation of this work was performed using Python programming language version 3.7.10, along with the Tensorflow framework and the Keras library.

3.2.5. Evaluation Measures

Quantitative assessments were conducted on an unseen test set, which consisted of 20% of the entire INIT Gait dataset. The testing set comprised two subjects, each with two different experiments encompassing various gait pattern abnormalities. Therefore, the total number of experiments in the testing set amounted to four. This paper employed two distinct evaluation schemes. The first scheme involved comparing each predicted sinogram signal, which corresponds to an individual silhouette image, against its corresponding ground-truth label. The second evaluation scheme operated at the subject level, using a majority voting method to consider all silhouette sinogram signals (or silhouette images). Both evaluations involved conducting experimental studies with varying numbers of target classes, specifically, n = 2, 4, or 6. We trained the network as a binary classification system in the two-class problem, which distinguished between normal gait patterns (NM) and severe gait impairment affecting the full body (FB). For the four-class problem, two additional gait patterns were included, namely, half motion of the left leg (LL) and half motion of the right leg (RL). The six-class problem included the four aforementioned gait impairments, along with motionlessness in the upper limbs of the left arm (LA-0) and right arm (RA-0).

We quantitatively evaluated all these experiments using metrics like sensitivity, specificity, accuracy, and F1-score. Additionally, we provided a confusion matrix to visually depict the ratio of correctly predicted classes to false predictions, thereby providing an intuitive representation of the classification performance.

4. Results and Experiments

4.1. Gait Analysis Results

In this work, a 1D-CNN model was implemented for the classification purposes of gait pathologies. The test set comprised two subjects, each of which had two sequences for each of the five anomalies as well as the normal unimpaired sequences (NM). The total number of frames was 818, 1292, 1258, 843, 842, and 2616, corresponding to NM, LL, RL, LA-0, RA-0, and FB, respectively.

In this work, three classification problems were addressed: the binary classification of NM versus FB (n = 2); the classification of NM, FB, LL, and RL (n = 4); and the classification of NM, FB, LL, RL, LA-0, and RA-0 (n = 6). In addition, the impact of a different number of signals or images fed to the network was studied (k = 1, 10, 20, and 30).

For the first classification problem (n = 2 and k = 1), NM vs. FB, one image from the sequence was sufficient to achieve perfect classification results: 100% accuracy for all of the 818 NM and 2616 FB frames. Accordingly, there was neither a need to explore the impact of larger values of k, nor a need to evaluate based on a subject's complete sequence. Figure 6

displays the confusion matrices of all the other problems using both evaluation schemes. To analyze the proportion of correct classifications vs. misclassification with respect to each class, their proportions were calculated and are shown in Figures 7 and 8 for n = 4 and n = 6, respectively.
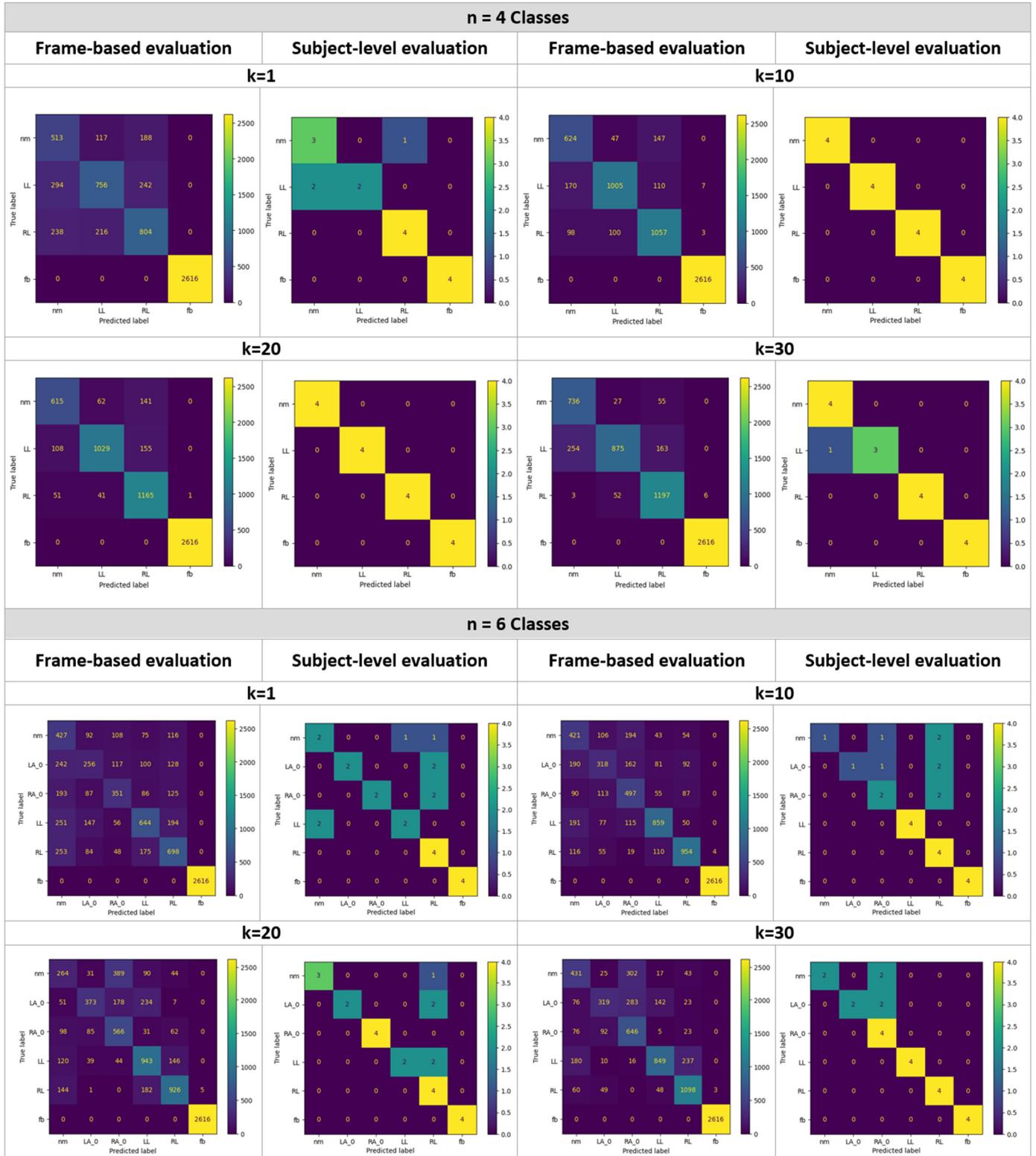

Figure 6. Confusion matrices of classifying n classes using k frames, where n = 4 (**top**) and n = 6 (**bottom**), and k = 1, 10, 20, and 30.
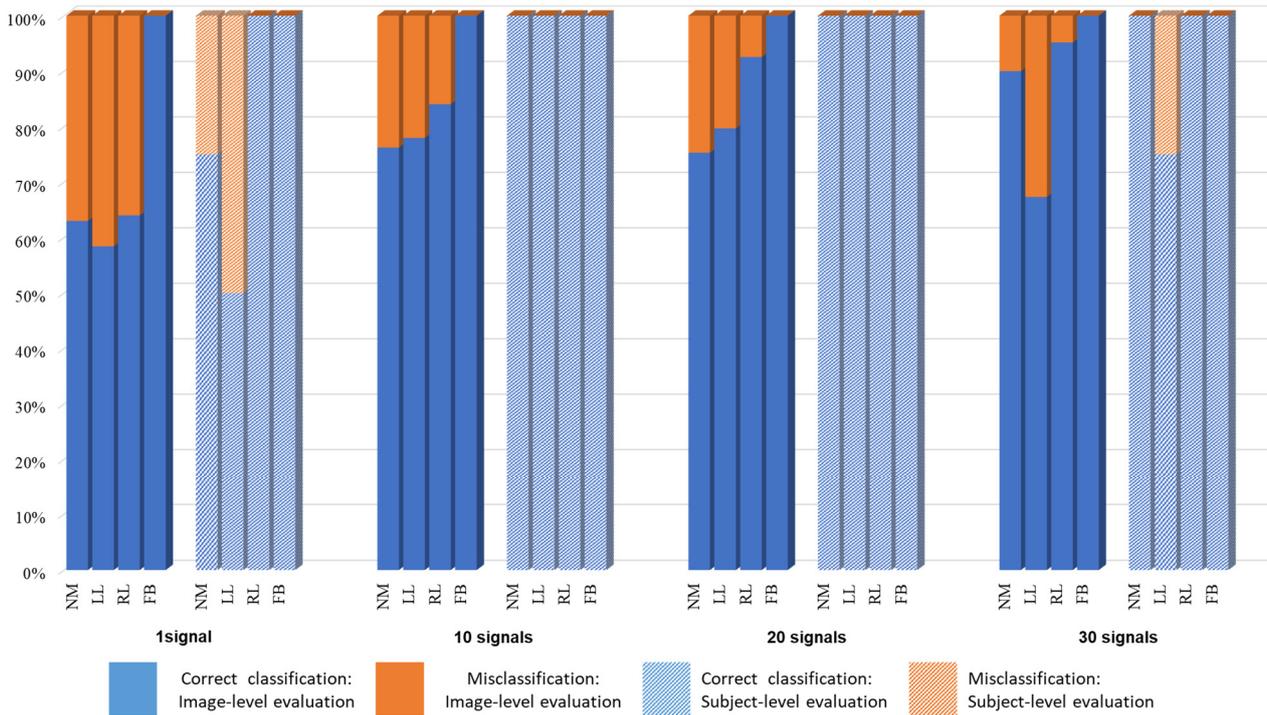
**Figure 7.** Correct vs. false classification for n = 4.
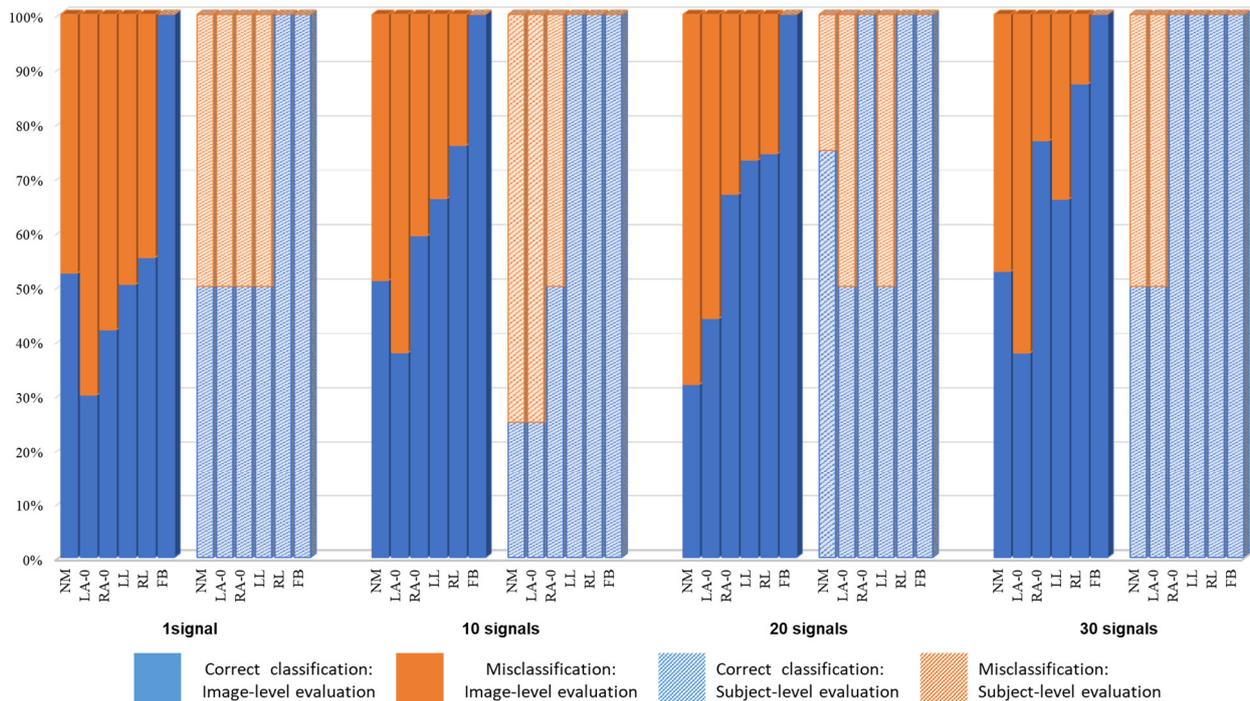
Figure 8. Correct vs. false classification for n = 6.

As can be seen, for all the values of n and k, no misclassifications occurred with FB deformation, where a subject suffered from severe gait cycle impairment featuring short steps, feet subtly touching the ground, and bending knees, chest, and head [5].

Table 2 shows the evaluation metrics for the aforementioned classification problems, weighted by the total number of frames per class. The best results, those having the highest F1-score, are highlighted; hence, testing the network with 20 and 30 images for the four-class and the six-class problems, respectively, is recommended to attain the best

performance. The accuracy is computed as the percentage of the correct classifications divided by the entire population, as per Kelleher et al. [53]; it is also equivalent to the weighted sensitivity, and they recommended this measure of accuracy over the weighted accuracy for the imbalanced datasets.

Table 2. Classification results of two evaluation schemes (signal-based and subject-based) with different number of prior frames (k = 1, 10, 20, 30) and two types of gait patterns (n = 4 and 6).

	Evaluation Based on a Signal/Image							
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	78.36	88.60	90.66	90.64	65.09	73.91	74.17	77.70
Specificity	95.15	97.34	97.78	97.91	94.76	96.24	96.01	96.73
Accuracy	78.36	88.60	90.66	90.64	65.09	73.87	74.17	77.70
F1-score	78.52	88.65	90.62	90.43	65.32	74.05	73.88	77.32
	Evaluation Based on Majority Voting of a Sequence of a Subject							
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	81.25	100	100	93.75	66.67	66.67	77.5	83.33
Specificity	93.75	100	100	97.92	93.33	93.33	95.24	96.67
Accuracy	81.25	100	100	93.75	66.67	66.67	79.17	83.33
F1-score	80.56	100	100	93.65	67	64.52	76.66	83.33

Table 2 depicts portions of the sequences fed to the network along with their predictions. The results are also displayed in Appendix A, where all the evaluation metrics were computed per class individually. Table 3 displays the class metrics for the suggested values of k—achieving the highest weighted F1-score—that were 20 and 30 for n = 4 and n = 6, respectively. The highest F1-score was computed for FB, followed by RL, LL, NM, RA, and lastly, LA. On the other hand, the highest accuracy was computed for FB, followed by NM, LL, RL, LA, and lastly, RA. These values were computed when the class in the header of the table was regarded as the positive class and all the remaining classes were the negative class.

Table 3. Evaluation metrics per class for the suggested *k* values.

	Evaluation Based on a Signal/Image									
	NM		LL		RL		FB		LA	RA
Classes	n = 4	n = 6	n = 4	n = 6	n = 4	n = 6	n = 4	n = 6	n = 6	n = 6
Frames/signals	k = 20	k = 30	k = 20	k = 30	k = 20	k = 30	k = 20	k = 30	k = 30	k = 30
Sensitivity	75.18	52.69	79.64	65.71	92.61	87.28	100	100	37.84	76.72
Specificity	96.92	94.28	97.8	96.68	93.74	94.91	99.97	99.94	97.42	91.2
Accuracy	93.95	89.84	93.88	91.46	93.5	93.66	99.98	99.96	90.87	89.61
F1-score	77.26	52.53	84.9	72.16	85.69	81.88	99.98	99.94	47.68	61.85
	Evaluation Based on Majority Voting of a Sequence of a Subject									
Frames/signals	k = 20	k = 30	k = 20	k = 30	k = 20	k = 30	k = 20	k = 30	k = 30	k = 30
Sensitivity	100	50	100	100	100	100	100	100	50	100
Specificity	100	100	100	100	100	100	100	100	100	80
Accuracy	100	91.67	100	100	100	100	100	100	91.67	83.33
F1-score	100	66.67	100	100	100	100	100	100	66.67	66.67

As mentioned above, all the FB frames were successfully classified for all the values of n and k; a sample of a sequence is shown in Table 4. In addition, the samples for an NM subject as well as an LL subject are shown in Table 4 with different values of n and k.

Table 4. Samples of the input signals and their corresponding predictions.

Correctly classified FB using n = 2, 4, and 6 for any k value. Blue corresponds to predicting FB.	
A part of an NM sequence; white, red, and green corresponds to NM, LL, and RL predictions, respectively. n = 4, k = 1.	
Same part as the row above; white and green corresponds to NM and RL predictions, respectively. n = 4, k = 10, 20.	
Same part as the row above; white and green corresponds to NM and RL predictions, respectively. n = 4, k = 30, and for n = 2.	
A part of an LL sequence; white, red, and green corresponds to NM, LL, and RL predictions, respectively. n = 4, k = 1.	
Same part as the row above; red corresponds to LL predictions. n = 4, k = 10, 20, and 30.	
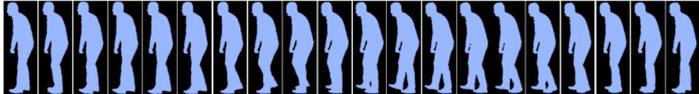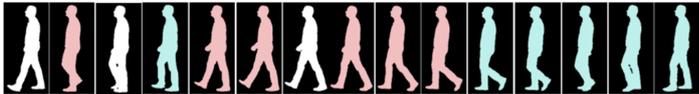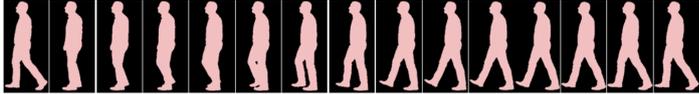
Figure 9 shows the detailed results of the four-class problem on the test set. In this figure, the x-axis denotes the subject ID along with the sequence number; e.g., NM1_Seq1 denotes the first sequence for the first subject with normal gait, and LL1_Seq2 denotes the second sequence of the first subject representing problems with the left leg, and so on. All the subjects' sequences are displayed except the four FB sequences since they yielded a 100% accuracy (for both n = 4 and n = 6). This bar chart portrays the proportions of labels for each sequence with respect to the number of frames. For example, regarding the entire sequence of frames for one of the two experiments of the first subject, when k = 1, frames predicted as NM, LL, and RL comprised 26%, 34%, and 40%, respectively; when k = 10, frames predicted as NM, LL, and RL comprised 71%, 11%, and 18%, respectively; when k = 20, frames predicted as NM, LL, and RL comprised 79%, 7%, and 14%, respectively; and, lastly, when k = 30, frames predicted as NM, LL, and RL comprised 93%, 3%, and 4%, respectively. For all the values of k and n, no misclassification with FB was encountered. All of these values are visualized as the first four bars and so on for the rest of the subjects. In a similar fashion, Figure 10 displays the detailed results of the six-class problem on the test set.
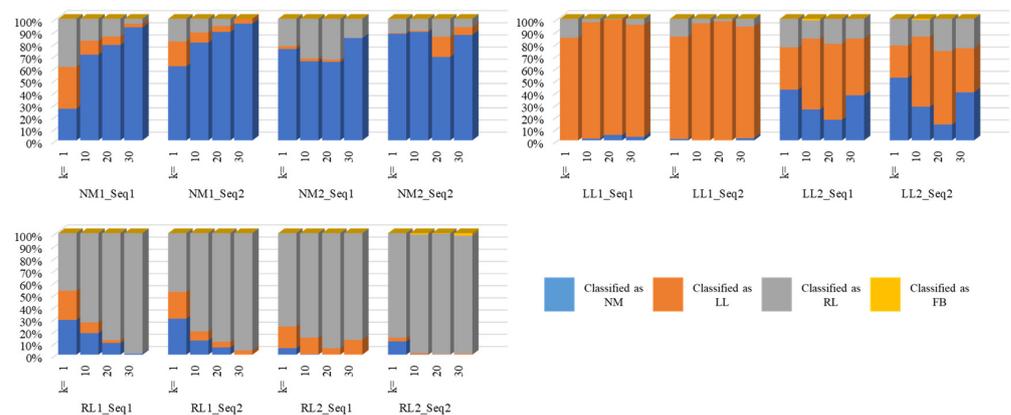

**Figure 9.** Four-class prediction labels for the test sets.
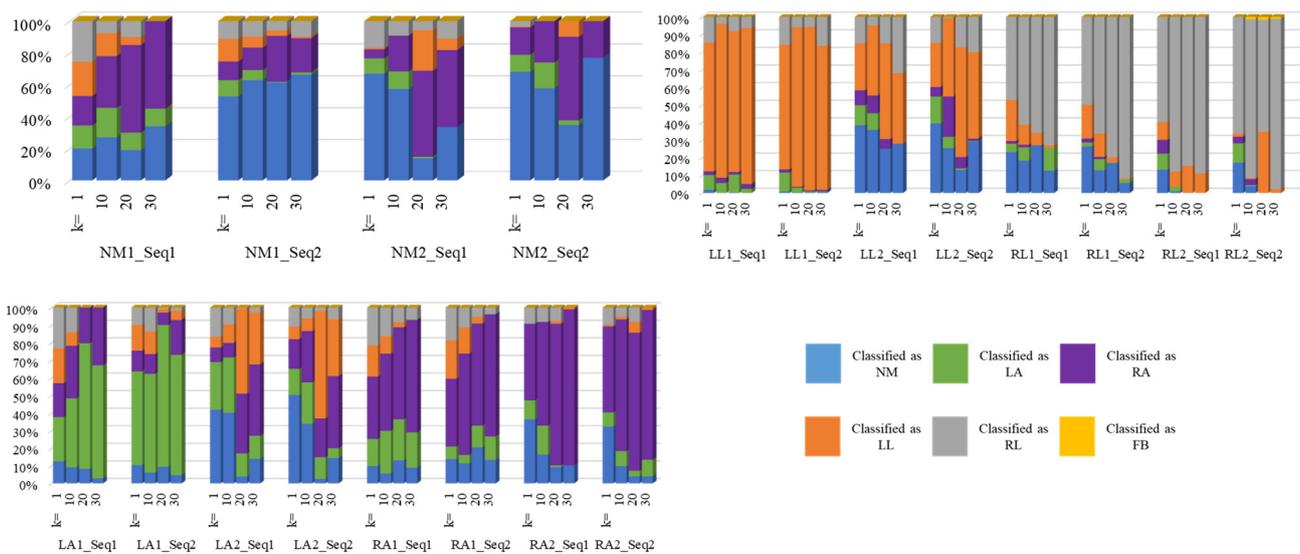
Figure 10. Six-class prediction labels for the test set.

4.2. Comparative Experiments

This study utilizes a deep learning method to identify abnormal gait patterns associated with neurological conditions like Parkinson's disease, multiple sclerosis, and strokes. Our focus is on gait classification, and as shown in Table 5, several prior works have employed similar methodologies. For instance, the vision-based system in [54] aimed to assess Parkinsonian gait severity. However, their classifiers, trained on data with limited availability, despite achieving good correlation with clinician labels, faced challenges due to the small datasets. Similarly, Cho et al. [13] were limited by a dataset comprising only seven Parkinson's patients and seven healthy controls.

Table 5. Comparison of used dataset, methodology, and achieved accuracy.

Study	Data	Approach	Accuracy
Chavez et al. [54]	Private dataset (Parkinsonian gait)	Classical machine learning methods (RFC and XGBoost)	50% (Normal vs. abnormal gait)
Cho et al. [13]	Private dataset (Parkinsonian gait)	Classical machine learning methods (hybrid PCA and LDA)	77% (Normal vs. abnormal gait)
Proposed study	Public dataset (INIT database [50])	Deep learning approach (1D CNN)	90% (4 classes of gait) 77% (6 classes of gait)

RFC: random forest classifier, XGBoost: extreme gradient boosting, PCA: principal component analysis, LDA: linear discriminant analysis.

Our study offers several advancements. First, we propose a novel gait representation method that allows us to leverage publicly available camera-based gait data from the INIT database. This approach captures detailed body movement information, akin to sensor-based methods, but with potentially broader accessibility and cost-effectiveness. Additionally, we employ an advanced deep learning method for multi-class gait classification, making it suitable for real-time implementation—a feature often absent in existing approaches. Overall, our work contributes a fresh perspective to gait analysis, addressing limitations in the current literature through innovative solutions.

In our study, we utilized a 1D CNN structure for its ability to effectively capture temporal information alongside main body movement features from silhouette binary images. Unlike 2D CNNs, which excel at learning contextual representations from 2D images, such as identifying key objects, our task presents challenges in identifying moving parts within a single silhouette.

One common approach to addressing this challenge is the generation of gait energy images (GEIs) by averaging energy across multiple sequences representing a gait cycle. However, we hypothesize that this method may not fully capture the available gait pattern information. Hence, we introduce a new approach to generate temporal information, incorporating body boundary angles and distances to create sinogram signals. Our choice to employ a 1D CNN allows us to effectively learn features by incorporating temporal representations from adjacent sequences.

To provide comparative insights, we implemented 2D CNN and 2D Residual Network (ResNet50) architectures using GEIs and compared their performance against our proposed 1D CNN using sinogram signals, utilizing identical training and testing splits. The results, presented in Table 6, demonstrate the superiority of our proposed 1D CNN approach over the 2D networks (CNN and ResNet50) for our specific task. This comparative analysis illustrates the efficiency of deep learning networks in capturing robust and prominent representations from the input data. Our method, which generates 1D sinogram signals, proves its capability compared to gait energy maps by considering calculations of angles and distances between all body boundary points and the centroid. This approach focuses more on body part movements, thereby enhancing gait impairment analysis.

Table 6. Comparison of the proposed 1D CNN that uses the generated 1D sinogram signals against 2D CNN and ResNet50 models that use 2D gait energy images (GEIs).

Methods	Number of Classes (n = 4)				Number of Classes (n = 6)			
	Sensitivity	Specificity	Accuracy	F1-Score	Sensitivity	Specificity	Accuracy	F1-Score
2D CNN	82.50	94.83	82.50	82.35	70.94	89.90	70.94	70.25
2D ResNet50	86.25	91.38	86.25	86.24	67.08	85.86	67.52	67.08
Proposed 1D-CNN (k = 20)	90.66	97.78	90.66	90.62	74.17	96.01	74.17	73.88
Proposed 1D-CNN (k = 30)	90.64	97.91	90.64	90.43	77.70	96.73	77.70	77.32

5. Discussion

In this study, the following classification problems were tackled: the binary classification of NM versus FB, the classification of NM, FB, LL, and RL, and the classification of NM, FB, LL, RL, LA-0, and RA-0. The dataset comprised eight subjects for training and two subjects for testing for each anomaly; each of them had two sequences with a total number of frames of 818, 1292, 1258, 843, 842, and 2616 for NM, LL, RL, LA-0, RA-0, and FB, respectively. In addition, the effect of having different numbers of inputs was assessed: 1, 10, 20, and 30. Hence, testing the network with 20 and 30 images for the four-class and six-class problems, respectively, is recommended to attain the highest F1-score. It is worth noting that no misclassifications were encountered for FB when using image-level evaluation, and no misclassifications were encountered for RA-0, RL, LL, and FB when using subject-based evaluation. However, it is important to mention that the test set for subject-level evaluation was small: only two sequences for two subjects, which might be risky to generalize.

The use of 20 and 30 silhouette frames in our framework is considered sufficient to provide the necessary information for assisted knowledge learning. For instance, in [43], the number of silhouette frames within each gait cycle is estimated to be around 25 to 35 frames for normal motion styles. While this number may vary and increase for full-body gait patterns due to slower body movements, our experimental results indicate that abnormal full-body cases were the easiest class to classify compared to other gait patterns. Our study aimed to determine the optimal number of frames (k = 1, 10, 20, and 30) required to capture temporal information from adjacent sequences and identify the location of gait patterns. Increasing the number of frames would result in a heavier network architecture with more training parameters, as each frame is processed by a convolution layer to extract features before being incorporated with others. Although this may improve the performance to some extent, it could also lead to unfair comparisons with networks

using fewer computations. Additionally, including more frames may introduce redundancy by incorporating information from multiple gait cycles of the same subject.

In the work of Verlekar et al. [43], where the input was the entire video sequence for testing purposes, they attained an accuracy of 98.8% at the subject level. Contrary to this, in this work, only 20 frames were sufficient to attain all the values of the F1-score, specificity, and accuracy at 100% at the subject level for the same four classes in INIT. In addition, when testing at the image level, the F1-score, specificity, and accuracy had values of 90.62%, 97.78%, and 90.66%, respectively. Regarding the six-class problem, the F1-score, specificity, and accuracy were 83.33%, 96.67%, and 83.33%, respectively, at the subject level. On the other hand, at the image level, the F1-score, specificity, and accuracy were 77.32%, 96.73%, and 77.7%, respectively.

The accuracy of the NM classification was 95.49%, 100%, 99%, and 99% in [13,43,44,46], respectively. The number of classes involved was 2, 4, 5, and 5 in [13,43,44,46], respectively. In this study, the NM classification exhibited an F1-score, specificity, and accuracy of 100% at the subject level, and 77.26%, 96.92%, and 93.95%, respectively, at the image level.

To the best of our knowledge, neither classifying into six classes nor image-level testing have been performed before. Although Albuquerque et al. [46] made their classification based on nine key frames, they needed the entire video sequence to detect those key frames in the first place.

6. Conclusions

The implementation of automated recognition algorithms for analyzing human gait patterns offers numerous valuable opportunities. The proposed method enables a quantitative and non-invasive assessment of locomotion by comparing observed locomotor function with a standard gait pattern often observed in individuals in good health. This work presents a novel pipeline for the examination of gait disability, which encompasses the following stages: The generation of silhouette sinogram signals is achieved by the utilization of a unique method based on angular centroid distance. Following that, a 1D CNN model is constructed and trained by incorporating shared information across multiple silhouette signals via assisted knowledge learning. This study utilized two separate evaluation methodologies. The initial approach entailed assessing every anticipated sinogram signal, which corresponds to an individual silhouette image, in comparison to its matching ground-truth label. The second evaluation system considered all the silhouette sinogram signals at the individual subject level, employing a majority vote approach. This study focused on three classification tasks: distinguishing between NM and FB in a binary classification issue, classifying NM, FB, LL, and RL, and classifying NM, FB, LL, RL, LA-0, and RA-0. Furthermore, an investigation was conducted to examine the effects of varying the quantity of signals or images inputted into the network. Specifically, this study explored scenarios where k equaled 1, 10, 20, and 30. According to the attained findings, it was seen that there were no instances of misclassifications in relation to FB deformation across all the values of n and k. This was particularly evident in cases when subjects exhibited pronounced impairment in their gait cycle, characterized by the presence of shortened steps, subtle foot–ground contact, and flexion in the knees, chest, and head.

Quantitative gait analysis is largely used in research institutes and not in therapeutic settings. Traditional gait laboratories' high costs/cumbersome equipment, complex protocols, and data management/analysis, as well as clinical teams' variable training/experience and preferences are the main reasons. Clinics use observational gait and qualitative measures most often, even if quantifying gait is less beneficial. It appears that our gait representation method has never been published. Similar to sensor-based approaches, low-cost camera-based gait data collecting can offer exact body component movement data. Real-time implementation is possible due to low processing complexity.

Appendix A

NM								
Evaluation based on a signal/image								
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	62.71	76.28	75.18	89.98	52.2	51.47	32.27	52.69
Specificity	89.7	94.81	96.92	95.03	86.29	91.43	93.97	94.28
Accuracy	86.01	92.28	93.95	94.33	82.66	87.16	87.39	89.84
F1-score	55.07	72.98	77.26	81.28	39.1	46.11	35.32	52.53
Evaluation based on majority voting of a sequence of a subject								
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	75	100	100	100	50	25	75	50
Specificity	83.33	100	100	91.67	90	100	100	100
Accuracy	81.25	100	100	93.75	83.33	87.5	96	91.67
F1-score	66.67	100	100	88.89	50	40	85.71	66.67
LL								
Evaluation based on a signal/image								
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	58.51	77.79	79.64	67.72	49.85	66.49	72.99	65.71
Specificity	92.9	96.87	97.8	98.32	93.16	95.47	91.58	96.68
Accuracy	85.48	92.75	93.88	91.71	85.87	90.58	88.45	91.46
F1-score	63.5	82.24	84.9	77.92	54.3	70.41	68.04	72.16
Evaluation based on majority voting of a sequence of a subject								
Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	50	100	100	75	50	100	50	100
Specificity	100	100	100	100	95	100	100	100
Accuracy	87.5	100	100	93.75	87.5	100	92	100
F1-score	66.67	100	100	85.71	57.14	100	66.67	100

FB
Evaluation based on a signal/image

Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	100	100	100	100	100	100	100	100
Specificity	100	99.7	99.97	99.82	100	100	99.9	99.94
Accuracy	100	99.83	99.98	99.9	100	100	99.93	99.96
F1-score	100	99.81	99.98	99.89	100	100	99.9	99.94

Evaluation based on majority voting of a sequence of a subject

Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	100	100	100	100	100	100	100	100
Specificity	100	100	100	100	100	100	100	100
Accuracy	100	100	100	100	100	100	100	100
F1-score	100	100	100	100	100	100	100	100

RL
Evaluation based on a signal/image

Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	63.91	84.02	92.61	95.15	55.48	76.08	73.61	87.28
Specificity	90.9	94.56	93.74	95.39	91.22	95.59	95.96	94.91
Accuracy	85.23	92.35	93.5	95.34	85.36	92.39	92.29	93.66
F1-score	64.53	82.19	85.69	89.56	55.42	76.6	75.81	81.88

Evaluation based on majority voting of a sequence of a subject

Classes	n = 4				n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30	k = 1	k = 10	k = 20	k = 30
Sensitivity	100	100	100	100	100	100	100	100
Specificity	91.67	100	100	100	75	70	76.19	100
Accuracy	93.75	100	100	100	79.17	75	80	100
F1-score	88.89	100	100	100	61.54	57.14	61.54	100

LA
Evaluation based on a signal/image

Classes	**n = 6**			
Frames/signals	k = 1	k = 10	k = 20	k = 30
Sensitivity	30.37	37.72	44.25	37.84
Specificity	93.99	94.85	97.71	97.42
Accuracy	87	88.57	91.84	90.87
F1-score	33.93	42.06	54.37	47.68

Evaluation based on majority voting of a sequence of a subject

Classes	n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30
Sensitivity	50	25	40	50
Specificity	100	100	100	100
Accuracy	91.67	87.5	88	91.67
F1-score	66.67	40	57.14	66.67

RA Evaluation based on a signal/image				
Classes	n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30
Sensitivity	41.69	59.03	67.22	76.72
Specificity	95.18	92.82	91.05	91.2
Accuracy	89.31	89.11	88.43	89.61
F1-score	46.12	54.35	56.07	61.85
Evaluation based on majority voting of a sequence of a subject				
Classes	n = 6			
Frames/signals	k = 1	k = 10	k = 20	k = 30
Sensitivity	50	50	100	100
Specificity	100	90	95.24	80
Accuracy	91.67	83.33	96	83.33
F1-score	66.67	50	88.89	66.67

References

1. Saboor, A.; Kask, T.; Kuusik, A.; Alam, M.M.; Le Moullec, Y.; Niazi, I.K.; Zoha, A.; Ahmad, R. Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review. *IEEE Access* **2020**, *8*, 167830–167864. [CrossRef]
2. Hulleck, A.A.; Menoth Mohan, D.; Abdallah, N.; El Rich, M.; Khalaf, K. Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies. *Front. Med. Technol.* **2022**, *4*, 901331. [CrossRef] [PubMed]
3. Hase, K. A computer simulation study on the causal relationship between walking and physical malfunctions in older adults. *Anthropol. Sci.* **2008**, *116*, 95–104. [CrossRef]
4. Zhou, C.J.; Mitsugami, I.; Yagi, Y. Detection of Gait Impairment in the Elderly Using Patch-GEI. *IEEJ Trans. Electr. Electr.* **2015**, *10*, S69–S76. [CrossRef]
5. Ortells, J.; Herrero-Ezquerro, M.T.; Mollineda, R.A. Vision-based gait impairment analysis for aided diagnosis. *Med. Biol. Eng. Comput.* **2018**, *56*, 1553–1564. [CrossRef] [PubMed]
6. Mohan, D.M.; Khandoker, A.H.; Wasti, S.A.; Alali, S.I.I.I.; Jelinek, H.F.; Khalaf, K. Assessment Methods of Post-stroke Gait: A Scoping Review of Technology-Driven Approaches to Gait Characterization and Analysis. *Front. Neurol.* **2021**, *12*, 650024. [CrossRef] [PubMed]
7. Carriero, A.; Zavatsky, A.; Stebbins, J.; Theologis, T.; Shefelbine, S.J. Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture* **2009**, *29*, 71–75. [CrossRef] [PubMed]
8. Chang, F.M.; Rhodes, J.T.; Flynn, K.M.; Carollo, J.J. The Role of Gait Analysis in Treating Gait Abnormalities in Cerebral Palsy. *Orthop. Clin. N. Am.* **2010**, *41*, 489–506. [CrossRef] [PubMed]
9. Zhang, X.; Rong, X.; Luo, H. Optimizing lower limb rehabilitation: The intersection of machine learning and rehabilitative robotics. *Front. Rehabil. Sci.* **2024**, *5*, 1246773. [CrossRef]
10. States, R.A.; Krzak, J.J.; Salem, Y.; Godwin, E.M.; Bodkin, A.W.; McMulkin, M.L. Instrumented gait analysis for management of gait disorders in children with cerebral palsy: A scoping review. *Gait Posture* **2021**, *90*, 1–8. [CrossRef]
11. Rozumalski, A.; Schwartz, M.H. Crouch gait patterns defined using k-means cluster analysis are related to underlying clinical pathology. *Gait Posture* **2009**, *30*, 155–160. [CrossRef]
12. Begg, R.; Kamruzzaman, J. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. *J. Biomech.* **2005**, *38*, 401–408. [CrossRef]
13. Cho, C.W.; Chao, W.H.; Lin, S.H.; Chen, Y.Y. A vision-based analysis system for gait recognition in patients with Parkinson's disease. *Expert Syst. Appl.* **2009**, *36*, 7033–7039. [CrossRef]
14. Chen, Y.Y.; Cho, C.W.; Lin, S.H.; Lai, H.Y.; Lo, Y.C.; Chen, S.Y.; Chang, Y.J.; Huang, W.T.; Chen, C.H.; Jaw, F.S.; et al. A vision-based regression model to evaluate Parkinsonian gait from monocular image sequences. *Expert Syst. Appl.* **2012**, *39*, 520–526. [CrossRef]
15. Yang, M.J.; Zheng, H.R.; Wang, H.Y.; McClean, S.; Hall, J.; Harris, N. A machine learning approach to assessing gait patterns for Complex Regional Pain Syndrome. *Med. Eng. Phys.* **2012**, *34*, 740–746. [CrossRef]
16. Khan, T.; Grenholm, P.; Nyholm, D. Computer vision methods for parkinsonian gait analysis: A review on patents. *Recent Pat. Biomed. Eng.* **2013**, *6*, 97–108. [CrossRef]
17. Alaqtash, M.; Yu, H.Y.; Brower, R.; Abdelgawad, A.; Sarkodie-Gyan, T. Application of wearable sensors for human gait analysis using fuzzy computational algorithm. *Eng. Appl. Artif. Intel.* **2011**, *24*, 1018–1025. [CrossRef]
18. Nadeau, S.; Duclos, C.; Bouyer, L.; Richards, C.L. Guiding task-oriented gait training after stroke or spinal cord injury by means of a biomechanical gait analysis. *Prog. Brain Res.* **2011**, *192*, 161–180. [CrossRef] [PubMed]
19. Papageorgiou, E.; Nieuwenhuys, A.; Vandekerckhove, I.; Van Campenhout, A.; Ortibus, E.; Desloovere, K. Systematic review on gait classifications in children with cerebral palsy: An update. *Gait Posture* **2019**, *69*, 209–223. [CrossRef]

20. Darbandi, H.; Baniasad, M.; Baghdadi, S.; Khandan, A.; Vafaee, A.; Farahmand, F. Automatic classification of gait patterns in children with cerebral palsy using fuzzy clustering method. *Clin. Biomech.* **2020**, *73*, 189–194. [CrossRef]
21. Figueiredo, J.; Santos, C.P.; Moreno, J.C. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. *Med. Eng. Phys.* **2018**, *53*, 1–12. [CrossRef] [PubMed]
22. Han, J.; Bhanu, B. Individual recognition using Gait Energy Image. *IEEE Trans. Pattern Anal.* **2006**, *28*, 316–322. [CrossRef]
23. Bailly, R.; Lempereur, M.; Thepaut, M.; Pons, C.; Houx, L.; Brochard, S. Relationship between 3D lower limb bone morphology and 3D gait variables in children with uni and bilateral Cerebral Palsy. *Gait Posture* **2022**, *92*, 51–59. [CrossRef] [PubMed]
24. Hebda-Boon, A.; Tan, X.L.; Tillmann, R.; Shortland, A.P.; Firth, G.B.; Morrissey, D. The impact of instrumented gait analysis on decision-making in the interprofessional management of cerebral palsy: A scoping review. *Eur. J. Paediatr. Neurol.* **2023**, *42*, 60–70. [CrossRef] [PubMed]
25. Bonnet-Lebrun, A.; Linglart, A.; De Tienda, M.; Khac, V.N.; Ouchrif, Y.; Berkenou, J.; Pillet, H.; Assi, A.; Wicart, P.; Skalli, W. Combined gait analysis and radiologic examination in children with X-linked hypophosphatemia. *Clin. Biomech.* **2023**, *105*, 105974. [CrossRef]
26. Otayek, J.; Bizdikian, A.J.; Yared, F.; Saad, E.; Bakouny, Z.; Massaad, A.; Ghanimeh, J.; Labaki, C.; Skalli, W.; Ghanem, I.; et al. Influence of spino-pelvic and postural alignment parameters on gait kinematics. *Gait Posture* **2020**, *76*, 318–326. [CrossRef] [PubMed]
27. Mostayed, A.; Mazumder, M.M.G.; Kim, S.; Park, S.J. Abnormal gait detection using discrete Fourier transform. In Proceedings of the 2008 International Conference on Multimedia and Ubiquitous Engineering (MUE 2008), Busan, Republic of Korea, 24–26 April 2008; pp. 36–40.
28. Wang, F.; Stone, E.; Dai, W.Q.; Skubic, M.; Keller, J. Gait Analysis and Validation Using Voxel Data. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 6127–6130.
29. Chiu, S.L.; Chou, L.S. Variability in inter-joint coordination during walking of elderly adults and its association with clinical balance measures. *Clin. Biomech.* **2013**, *28*, 454–458. [CrossRef] [PubMed]
30. Shin, S.S.; An, D.H.; Yoo, W.G. The effect of navigating obstacles when walking with deficient (inappropriate, deprived, limited) foot sensory input on spatiotemporal gait parameters of elderly individuals with different visual acuities. *Eur. Geriatr. Med.* **2015**, *6*, 1–6. [CrossRef]
31. Salarian, A.; Russmann, H.; Vingerhoets, F.J.G.; Dehollain, C.; Blanc, Y.; Burkhard, P.R.; Aminian, K. Gait assessment in Parkinson's disease: Toward an ambulatory system for long-term monitoring. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1434–1443. [CrossRef]
32. Chang, H.-C.; Hsu, Y.-L.; Yang, S.-C.; Lin, J.-C.; Wu, Z.-H. A wearable inertial measurement system with complementary filter for gait analysis of patients with stroke or Parkinson's disease. *IEEE Access* **2016**, *4*, 8442–8453. [CrossRef]
33. Amin, J.; Ruthiraphong, P. Cloud-based Gait Analysis Using a Single IMU for Parkinson Disease. In Proceedings of the 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 19–22 May 2021; pp. 1129–1132.
34. Khan, A.Z.; Aamir, F.; Kafeel, A.; Khan, M.U.; Aziz, S. Freezing of gait detection in parkinson's disease from accelerometer readings. In Proceedings of the 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), Karachi, Pakistan, 15–17 July 2021; pp. 1–5.
35. Kovalenko, E.; Shcherbak, A.; Somov, A.; Bril, E.; Zimniakova, O.; Semenov, M.; Samoylov, A. Detecting the Parkinson's Disease Through the Simultaneous Analysis of Data From Wearable Sensors and Video. *IEEE Sens. J.* **2022**, *22*, 16430–16439. [CrossRef]
36. Hallemans, A.; Ortibus, E.; Meire, F.; Aerts, P. Low vision affects dynamic stability of gait. *Gait Posture* **2010**, *32*, 547–551. [CrossRef] [PubMed]
37. Khan, T.; Westin, J.; Dougherty, M. Motion cue analysis for parkinsonian gait recognition. *Open Biomed. Eng. J.* **2013**, *7*, 1–8. [CrossRef] [PubMed]
38. Costin, H.; Luca, R.; Onu, I.; Rotaru, F.; Bejinariu, S.-I. Medical Rehabilitation Evaluation using Image Processing of Patients' Gait. In Proceedings of the 2023 13th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 23–25 March 2023; pp. 1–5.
39. Yang, X.C.; Zhou, Y.; Zhang, T.H.; Shu, G.; Yang, J. Gait recognition based on dynamic region analysis. *Signal Process.* **2008**, *88*, 2350–2356. [CrossRef]
40. Hu, M.D.; Wang, Y.H. A New Approach for Gender Classification Based on Gait Analysis. In Proceedings of the 2009 Fifth International Conference on Image and Graphics (ICIG 2009), Xi'an, China, 20–23 September 2009; pp. 869–874.
41. Bakchy, S.C.; Islam, M.R.; Mahmud, M.R.; Imran, F. Human gait analysis using gait energy image. *arXiv* **2022**, arXiv:2203.09549.
42. Luo, J.; Zhang, J.; Zi, C.; Niu, Y.; Tian, H.; Xiu, C. Gait recognition using GEI and AFDEI. *Int. J. Opt.* **2015**, *2015*, 763908. [CrossRef]
43. Verlekar, T.T.; Soares, L.D.; Correia, P.L. Automatic Classification of Gait Impairments Using a Markerless 2D Video-Based System. *Sensors* **2018**, *18*, 2743. [CrossRef]
44. Albuquerque, P.; Machado, J.P.; Verlekar, T.T.; Correia, P.L.; Soares, L.D. Remote Gait Type Classification System Using Markerless 2D Video. *Diagnostics* **2021**, *11*, 1824. [CrossRef] [PubMed]
45. Loureiro, J.; Correia, P.L. Using a Skeleton Gait Energy Image for Pathological Gait Classification. In Proceedings of the 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), Buenos Aires, Argentina, 16–20 November 2020; pp. 503–507.

46. Albuquerque, P.; Verlekar, T.T.; Correia, P.L.; Soares, L.D. A Spatiotemporal Deep Learning Approach for Automatic Pathological Gait Classification. *Sensors* **2021**, *21*, 6202. [CrossRef]
47. Wang, T.; Du, Y.; Gong, Y.; Choo, K.-K.R.; Guo, Y. Applications of federated learning in mobile health: Scoping review. *J. Med. Internet Res.* **2023**, *25*, e43006. [CrossRef]
48. Zhang, J.; Tian, J.; Yan, P.; Wu, S.; Luo, H.; Yin, S. Multi-hop graph pooling adversarial network for cross-domain remaining useful life prediction: A distributed federated learning perspective. *Reliab. Eng. Syst. Saf.* **2024**, *244*, 109950. [CrossRef]
49. Zhang, J.; Tian, J.; Alcaide, A.M.; Leon, J.I.; Vazquez, S.; Franquelo, L.G.; Luo, H.; Yin, S. Lifetime extension approach based on levenberg-marquardt neural network and power routing of dc-dc converters. *IEEE Trans. Power Electron.* **2023**, *38*, 10280–10291. [CrossRef]
50. INIT Gait Database. Available online: https://www.vision.uji.es/gaitDB/ (accessed on 10 January 2023).
51. Gonzalez, R.C.; Woods, R.E.; Eddins, S.L. *Digital Image Processing Using MATLAB*; Pearson Prentice-Hall: London, UK, 2004.
52. Springenberg, J.T.; Dosovitskiy, A.; Brox, T.; Riedmiller, M. Striving for simplicity: The all convolutional net. *arXiv* **2014**, arXiv:1412.6806.
53. Kelleher, J.D.; Mac Namee, B.; D'Arcy, A. *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*; MIT Press: Cambridge, MA, USA, 2015; p. 262029448.
54. Chavez, J.M.; Tang, W. A vision-based system for stage classification of parkinsonian gait using machine learning and synthetic data. *Sensors* **2022**, *22*, 4463. [CrossRef]