Pathological Gait Analysis With an Open-Source Cloud-Enabled Platform Empowered by Semi-Supervised Learning-PathoOpenGait

Ming-Yang Ho , Ming-Che Kuo , Ciao-Sin Chen , Ruey-Meei Wu , Ching-Chi Chuang, Chi-Sheng Shih , *Senior Member, IEEE,* and Yufeng Jane Tseng

Abstract—We present PathoOpenGait, a cloud-based platform for comprehensive gait analysis. Gait assessment is crucial in neurodegenerative diseases such as Parkinson's and multiple system atrophy, yet current techniques are neither affordable nor efficient. PathoOpenGait utilizes 2D and 3D data from a binocular 3D camera for monitoring and analyzing gait parameters. Our algorithms, including a semi-supervised learning-boosted neural network model for turn time estimation and deterministic algorithms to estimate gait parameters, were rigorously validated on annotated gait records, demonstrating high precision and consistency. We further demonstrate PathoOpenGait's applicability in clinical settings by analyzing gait trials from Parkinson's patients and healthy controls. PathoOpenGait is the first open-source, cloud-based system for gait analysis, providing a user-friendly tool for continuous patient care and monitoring. It offers a cost-effective and accessible solution for both clinicians and patients, revolutionizing the field of gait assessment. PathoOpenGait is available at https://pathoopengait.cmdm.tw.

Index Terms—Gait analysis, semi-supervised learning, deep learning, cloud-based platform, neurodegenerative diseases.

I. INTRODUCTION

THE assessment of gait is integral to the diagnosis and management of neurodegenerative diseases, such as Parkinson's disease (PD) and multiple system atrophy (MSA) [1], [2], [3]. They often result in movement impairments marked by muscle rigidity, bradykinesia, resting limb tremors, and postural instability [4], considerably lowering the quality of life [5]. Monitoring gait parameters such as stride length (SL), stride width (SW), stride time (ST), velocity, and cadence is vital not only for tracking symptom progression but also for diagnosing neurodegenerative diseases [1], [6], [7]. Furthermore, assessing turn time is valuable in detecting freezing of gait, a common symptom in PD [8].

Nevertheless, current medical protocols rely on nonstandardized and subjective visual assessments [9], [10]. Given the sporadic and inconsistent nature of gait disruptions, which are affected by various internal and external elements [11], medical practitioners often lack the necessary time and resources to thoroughly examine an individual's entire gait pattern. This presents an urgent demand for a systematic tool to support early detection and prolonged monitoring of gait parameters within healthcare systems [12].

Numerous technological innovations have been utilized in gait analysis. Carpet-based systems, such as GaitRite [13], [14], and 3D optical camera-array-based systems, like Vicon [15], [16], are regarded as gold standards in the technological field [13], [17]. However, their high costs and extensive space requirements limit their applicability in clinical and domestic settings [17]. Alternatively, systems based on wearable devices might be relatively more affordable [12], [18], [19], [20]. However, the additional equipment an individual must wear can be disruptive, and the setup is often inconvenient [21]. Moreover, the location
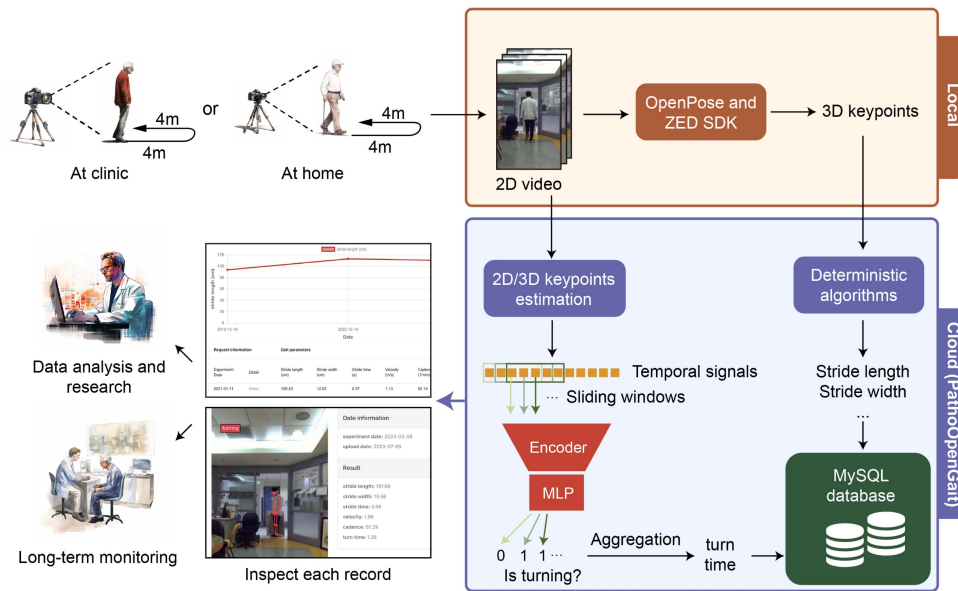
Fig. 1. **PathoOpenGait's Comprehensive Framework.** PathoOpenGait is transforming gait analysis by offering a user-friendly, cloud-based service, facilitating patients and doctors in recording and uploading gait trial data seamlessly from any location. The platform utilizes 2D videos and precomputed 3D keypoints to estimate turn time and additional gait parameters. The turn time estimation algorithm encompasses a neural network-based model and a sliding window-based method. Once the inference process is completed, users can visualize the results on the interactive dashboard.

where wearable devices are attached can influence estimation performance [22]. While devices attached to the feet might provide more effective signals for gait analysis than those attached to the knees, the former can be more uncomfortable for users than the latter [20].

The recent evolution of computer vision technology, especially human pose estimation [23], [24], has enabled the development of monocular camera-based systems [25]. Studies conducted by Kidziński Ł. et al. [26] and Stenum J. et al. [27] utilized sagittal-plane videos and proposed 2D camera-based solutions, but these required either multiple cameras to broaden the viewing field or manual adjustments of a 2D camera by technicians, making them impractical. To navigate the limitations present in these sagittal-plane methodologies, 3D human pose estimation techniques have been employed to augment depth data in 2D frontal-plane videos [28], [29]. Nevertheless, the absence of accurate depth information adversely affects the estimation of depth-related gait parameters, such as SL and velocity. While attempts have been made to leverage infrared-based 3D cameras like Kinect [30], their capacity to capture accurate depth information remains insufficient, thereby diminishing performance in SL estimation. On the other hand, it's worth noting that current visual-based gait analysis methods have not yet incorporated a turn time estimation algorithm, limiting their utility in monitoring common symptoms associated with PD [8].

Another challenge arises in the realm of accessibility, underscoring the necessity for an open-source platform for gait analysis. Despite the strides made in developing gait analysis techniques, a notable deficit exists as most research studies neither offer a comprehensive platform for applying their proposed algorithms nor release their source code, impeding the application of these techniques in real clinical settings. Additionally, the closed-source nature of commercial platforms [31], [32] not only limits their adoption in clinical practices but also prevents patients from using them for extended monitoring at home due to their significant costs.

In this study, we developed an innovative cloud-based platform called PathoOpenGait (Fig. 1), designed to comprehensively monitor and analyze human gait using 2D and 3D data obtained from a validated binocular 3D camera [33]. Our study begins by utilizing 2D videos extracted from 3D footage to estimate the turn time (TT), employing a novel neural network-based methodology. To enhance the performance of our model, we introduce a semi-supervised training strategy [34]. Moreover, we employ a deterministic algorithm that relies on the captured 3D temporal keypoints to determine other essential gait parameters, including the SL, SW, ST, velocity, and cadence.

We conducted rigorous assessments at two distinct medical centers to verify our algorithms. Our evaluation encompassed 91 annotated gait records collected from 23 individuals, comprising a diverse group of participants, including healthy controls, patients diagnosed with PD, and patients with MSA. These individuals underwent an 8-meter walking test, and our results demonstrate remarkable precision and consistency between our system's estimations and the ground truth values.

Furthermore, a significant step forward is taken by showcasing the real-world applicability of PathoOpenGait in clinical settings. Using this platform, we collected and statistically analyzed gait trials from a substantial cohort of 109 PD patients (432 trials) and 122 healthy controls (476 trials). These findings highlight the capability of our system for extensive and ongoing gait monitoring.

Notably, our pioneering study invents a 2D video-based algorithm for quantifying turn time. Additionally, we introduce an open-source, cloud-based system, PathoOpenGait, for gait examination. This platform empowers patients to conveniently

record their gait from any location, thus providing an invaluable tool for continuous patient care and monitoring.

II. METHODS

A. Patient Recruitment Process

Patients diagnosed with Parkinson's disease (PD) or multiple system atrophy (MSA) were recruited from two prominent medical centers: the National Taiwan University Hospital (NTUH) and the National Taiwan University Cancer Center (NTUCC). These patients were diagnosed by neurologists specializing in movement disorders. To ensure the study's validity, all patients had been regularly followed in the clinic for at least three years and reported no subjective complaints about poor gait performance. Participants with severe cardiopulmonary disease or orthopedic impairments in the spine or limb joints were excluded from the study.

Several demographic and clinical variables were documented for each participant, including age, sex, height, Hoehn and Yahr (H&Y) stage, age of disease onset, disease duration, Mini-Mental State Examination (MMSE) scores, levodopa equivalent daily dose (LEDD), and specific characteristics of their gait presentations.

Healthy controls (HCs) between the ages of 20 and 100 years, without any neurodegenerative diseases or gait disorders, were also enrolled in the study. Individuals with conditions that could interfere with gait, such as vertebral fractures or severe metabolic diseases, were excluded from the control group.

B. Video Recording of the 8-M Walking Test (8MWT)

All participants included in the study were instructed to perform a series of walking tasks along a 4-meter line, both forward and backward, returning to the starting point after completing each walk. They were specifically instructed to walk comfortably yet quickly and safely. The entire walking sequence of 8 meters was recorded using Stereolab's ZED camera. This advanced stereo camera can capture high-resolution 3D video with exceptional spatial accuracy, with deviations of less than 0.5 cm [33]. The recorded video was captured at a resolution of 1080 p, with a frame rate of 30 frames per second.

Each data instance produced by the ZED camera is denoted as $\mathbf{X} = \{\mathbf{x}^v, \mathbf{x}^{pc}\} \sim \mathcal{X}$ and consists of a 2D video component represented by $\mathbf{x}^v \in R^{H \times W \times C \times T}$, where H represents the height, W represents the width, C denotes the RGB channels, and T signifies the duration of the video. Additionally, it contains a 3D point cloud component denoted by $\mathbf{x}^{pc} \in R^{H \times W \times Z \times T}$, where Z represents the depth dimension. The space defined by the ZED camera output data is represented by $\mathcal{X}$.

C. Establishing Ground Truth for Turn Time and Gait Parameters

The ground truth values for the turn time measurement involved manual segmentation based on the 2D video. Meanwhile, the ground truth values for the remaining gait parameters were established through the employment of a validated 3D optical system [35], which facilitated the recording of 3D trajectories, subsequently calibrated in alignment with real-world coordinates. Adhering to the established protocol [35], domain experts delineated the various gait phases using both the 2D video and trajectory data, thereby deriving the specific gait parameters. Additionally, these domain experts undertook a secondary refinement of the ground truth data to ensure the precision.

D. Data Splitting for Training and Testing

From the NTUH data, a maximum of 15 participants were randomly chosen from each of the HC, PD, and MSA groups. The data was split into training and testing sets at a 2:1 ratio. Only the testing set had all gait parameters annotated, while the training set was annotated solely for turn time. Data from the NTUCC was fully annotated and set aside as a testing set for cross-center validation.

E. Gait Parameter Estimation Algorithm Development

First, we employed the ZED camera's Software Development Kit (SDK) in combination with OpenPose [36] to extract the temporal positions of the ankle keypoint in terms of the horizontal (x) and depth (z) coordinates from the 3D point cloud data $\mathbf{x}^{pc}$. Specifically, we obtained the temporal signals denoted as $\mathbf{x}^{\text{left}}$, $\mathbf{x}^{\text{right}}$, $\mathbf{z}^{\text{left}}$, and $\mathbf{z}^{\text{right}}$, which represent the temporal variations in the horizontal and depth positions of the left and right ankles, respectively. These signals reside in the space R^T, where T denotes the video duration. Through our algorithm $\mathcal{F}$, we can estimate parameters such as stride length ($\hat{y}^{sl}$), stride width ($\hat{y}^{sw}$), stride time ($\hat{y}^{st}$), velocity ($\hat{y}^v$), and cadence ($\hat{y}^c$), all of which are within R. The relation is expressed as follows:

$$\{\hat{y}^j\} = \mathcal{F}\left(\mathbf{x}^{\text{left}}, \mathbf{x}^{\text{right}}, \mathbf{z}^{\text{left}}, \mathbf{z}^{\text{right}}\right), j \in \{sl, sw, st, v, c\} \tag{1}$$

Initially, all four temporal signals undergo a smoothing process by applying a running median ϕ_m with a window size of 5:

$$\widetilde{x}_i^k = \phi_m(x_{i-2}^k, \ldots, x_{i+2}^k), k \in \{left, right\}, i \in \{1, 2, \ldots, T\} \tag{2}$$

$$\widetilde{z}_i^k = \phi_m(z_{i-2}^k, \ldots, z_{i+2}^k), k \in \{left, right\}, i \in \{1, 2, \ldots, T\} \tag{3}$$

Subsequently, the starting and ending times of the turning phase are utilized as delimiters to segregate the forward and backward phases. Specifically, signals before the onset of the turning phase are classified as the forward phase, while those after the termination of the turning phase are classified as the backward phase.

For each phase, we compute the median depth as follows:

$$\widetilde{\mathbf{z}}^{mid} = \frac{\widetilde{\mathbf{z}}^{\text{left}} + \widetilde{\mathbf{z}}^{\text{right}}}{2} \tag{4}$$

We then employ $\widetilde{z}^{mid}$ to delineate the steps. This means that whenever $\widetilde{z}_i^{\text{left}}$ equals $\widetilde{z}_i^{mid}$ for a given $i \in \{1, 2, \ldots, T\}$, we consider these moments as the time points to partition the left steps, with an analogous process for the right steps.

$$\mathcal{T}^{\text{left}} = \{i : \widetilde{z}_i^{mid} \simeq \widetilde{z}_i^{\text{left}}, \forall i \in \{1, 2, \ldots, T\}\} \tag{5}$$

$$\mathcal{T}^{\text{right}} = \{i : \widetilde{z}_i^{mid} \simeq \widetilde{z}_i^{\text{right}}, \forall i \in \{1, 2, \ldots, T\}\} \tag{6}$$

It is assumed that N^{left} left steps are detected; that is, $|\mathcal{T}^{\text{left}}| = N^{\text{left}}$. At time points $t_i \in \mathcal{T}^{\text{left}} (i \in \{1, 2, \ldots, N^{\text{left}} - 1\})$, we can calculate the gait parameters as follows:

$$\hat{y}_i^{sl,\text{left}} = \left| \widetilde{z}_{t_{i+1}}^{\text{left}} - \widetilde{z}_{t_i}^{\text{left}} \right| \tag{7}$$

$$\hat{y}_i^{sw,\,left} = \left| \widetilde{x}_{t_i}^{\text{left}} - \widetilde{x}_{t_i}^{\text{right}} \right| \tag{8}$$

$$\hat{y}_i^{st,\text{left}} = t_{i+1} - t_i \tag{9}$$

$$\hat{y}_i^{v,\text{left}} = \frac{\hat{y}_i^{sl,\text{left}}}{\hat{y}_i^{st,\text{left}}} \tag{10}$$

$$\hat{y}_i^{c,\text{left}} = \frac{1}{\hat{y}_i^{st,\text{left}}} \tag{11}$$

These equations can also be applied for the right steps. Our final outputs are the mean values of the corresponding measures for both the left and right steps.

F. Handcrafted Turn Time Estimation Algorithm

An intuitive algorithm for estimating turn time can be devised from three temporal signals: $\widetilde{\mathbf{x}}^{\text{left}}$, $\widetilde{\mathbf{x}}^{\text{right}}$, and $\widetilde{\mathbf{z}}^{mid}$. Turning typically occurs at the trial's farthest position. Thus, we first determine the k_1^{th} percentile of $\widetilde{\mathbf{z}}^{mid}$ as $\widetilde{z}_{(k_1)}^{mid}$. Given the feet's width decreases during a turn, we compute the absolute difference between $\widetilde{\mathbf{x}}^{\text{left}}$ and $\widetilde{\mathbf{x}}^{\text{right}}$ as:

$$\widetilde{\mathbf{x}}^{diff} = \left| \widetilde{\mathbf{x}}^{\text{left}} - \widetilde{\mathbf{x}}^{\text{right}} \right| \tag{12}$$

Next, the k_2^{th} percentile of $\widetilde{\mathbf{x}}^{diff}$ is defined as $\widetilde{x}_{(k_2)}^{diff}$. The turning phase's timeframes are then identified by:

$$\mathcal{T}^{\text{turn}}$$
$$= \left\{ i \in \{1, 2, \ldots, T\} \mid \left(\widetilde{z}_i^{mid} > \widetilde{z}_{(k_1)}^{mid} \right) \wedge \left(\widetilde{x}_i^{diff} < \widetilde{x}_{(k_2)}^{diff} \right) \right\} \tag{13}$$

The turn time $\hat{y}^{tt}$ is obtained as $|\mathcal{T}^{\text{turn}}|$ via selecting k_1 and k_2 using annotated data pairs, e.g., the training data.

G. Neural Network-Based Turn Time Estimation Algorithm

Considering a 2D video $\mathbf{x}^v$, a model $\mathcal{M}$ can be constructed to estimate turn time $\hat{y}^{tt}$ as follows:

$$\hat{y}^{tt} = \mathcal{M}(\mathbf{x}^v), \mathbf{x}^v \in R^{H \times W \times C \times T}, \hat{y}^{tt} \in R \tag{14}$$

However, devising the algorithm above can be challenging due to the discrepancy in dimensionality between the sparse prediction and the input size. Consequently, we decompose $\mathcal{M}$ into several steps.

First, a 2D pose estimation model, $\mathcal{E}_{2D}$, is employed to extract 2D keypoints. We select Mask R-CNN in Detectron2 [37] as our pretrained model. We then utilize a pretrained 3D estimation model [38], $\mathcal{E}_{3D}$, to augment the keypoint space as follows:

$$\mathbf{x}^k = \mathcal{E}_{3D}(\mathcal{E}_{2D}(\mathbf{x}^v)), \mathbf{x}^k \in R^{17 \times 3 \times T} \tag{15}$$

The transformations performed by these models are defined as follows:

$$\mathcal{E}_{2D} : R^{H \times W \times C \times T} \to R^{17 \times 2 \times T} \tag{16}$$

$$\mathcal{E}_{3D} : R^{17 \times 2 \times T} \to R^{17 \times 3 \times T} \tag{17}$$

Subsequently, we design a model $\mathcal{M}' : R^{17 \times 3 \times T} \to R$ to estimate the turn time. Despite this simplification, the mapping relationship remains complex, with potential $2^{3264 \cdot T}$ relationships, considering that a 64-bit floating-point computation is involved.

To simplify the problem further, we transform it into a classification task. Specifically, for a given frame and its neighboring frames, we determine whether this frame falls within the turn time. Considering a time point i with 64 neighboring frames on each side, the input $\mathbf{x}_i^k$ now resides in $R^{17 \times 3 \times 129}$. That is,

$$\mathbf{x}_i^k = \mathbf{x}^k [\ldots] [\ldots] [i - 64, i - 63, \ldots, i + 64], i \in \{1, \ldots, T\} \tag{18}$$

Consequently, our new model $\mathcal{M}'' : R^{17 \times 3 \times 129} \to \{0, 1\}$ has only $2^{51 \times 129}$ relationships, which is a much simpler form for model design.

Since portraying the change in the keypoints during the turn time is challenging, we model $\mathcal{M}''$ as a neural network $\mathcal{D}_\theta$ that outputs logits. Namely, $\mathcal{D}_\theta : R^{17 \times 3 \times 129} \to R^2$. We structure our function set $\mathcal{H}$ as an encoder, consisting of a 3-layer 1D convolutional neural network (CNN), followed by a 3-layer multiple layer perceptron (MLP). For given labeled pairs (x_i^k, y_i^l), we aim to find $\mathcal{D}_\theta \in \mathcal{H}$ that fulfills the following objective:

$$\theta = \arg \min_\theta \mathcal{L}_s \tag{19}$$

$$= \arg \min_\theta \mathbb{E}_{i \sim T} \left[y_i^l \cdot \log \left(\mathcal{D}_\theta \left(\mathbf{x}_i^k \right) \right) \right.$$

$$\left. + \left(1 - y_i^l \right) \cdot \log \left(1 - \mathcal{D}_\theta \left(\mathbf{x}_i^k \right) \right) \right] \tag{20}$$

Once this step is complete, we can predict whether each time point corresponds to a turning point. We then apply dilation ϕ_{dilate} and erosion ϕ_{erode} operations with a kernel size of 10 to consolidate the results and filter outliers. Finally, the turn time is determined by the maximum number of consecutive frames identified as "1".

$$\hat{y}_i^l = \arg \max \left(\mathcal{D}_\theta \left(\mathbf{x}_i^k \right) \right), \forall i \in \{1, 2, \ldots, T\} \tag{21}$$

$$\hat{\mathbf{y}}^l = \left[\hat{y}_1^l, \hat{y}_2^l, \ldots, \hat{y}_T^l \right] \tag{22}$$

$$\widehat{\mathbf{y}'}^l = \phi_{dilate} \left(\phi_{erode} \left(\hat{\mathbf{y}}^l \right) \right) \tag{23}$$

$$\hat{y}^{tt} = \max \left\{ len(S) \mid S \text{ is consecutive } 1 \text{ in } \widehat{\mathbf{y}'}^l \right\} \tag{24}$$

H. Semi-Supervised Learning for Neural Network-Based Turn Time Estimation

Implementing the above optimization requires manual, framewise labeling, which is time-consuming and labor-intensive. We adopt a semi-supervised learning strategy to use unlabeled videos or 3D temporal keypoints to enhance the model's performance. Given that most existing algorithms are tailored for image-based datasets, we must adapt these for our 1D signal

data. We do this by modifying the FixMatch [39] method and designing appropriate weak and strong augmentations.

For the labeled data pairs $\left(\mathbf{x}_i^k, y_i^l\right)$ and unlabeled data $\left(\mathbf{x}_i^{UK}\right)$, we calculate the cross-entropy loss for paired data directly. For the unpaired data, we devise a weak-augmentation function $a_{weak}^{\sigma^2} \sim A_{weak}$, which introduces Gaussian noise with a small standard deviation to the original data:

$$
\begin{aligned}
\widetilde{\mathbf{x}}_i^k &= a_{\text{weak}}^{\sigma^2}\left(\mathbf{x}_i^k\right) \\
&= \mathbf{x}_i^k + \mathcal{N}\left(0, \sigma^2\right), \quad \sigma^2 \sim \mathcal{U}(0.00005, 0.00015) \quad (25)
\end{aligned}
$$

Here, $\mathcal{U}$ signifies a uniform distribution.

For the strong-augmentation function $a_{\text{strong}}^{\sigma^2, \text{mask}} \sim A_{\text{strong}}$, we introduce more Gaussian noise and mask certain data points:

$$
\widetilde{\mathbf{x}}_i^k = a_{\text{strong}}^{\sigma^2, \text{mask}}\left(\mathbf{x}_i^k\right) = \left(\mathbf{x}_i^k + \mathcal{N}\left(0, \sigma^2\right)\right) \cdot \mathbf{m} \quad (26)
$$

$$
\sigma^2 \sim \mathcal{U}(0.0005, \ 0.0015) \quad (27)
$$

$$
\mathbf{m} \sim \{0, 1\}^{17 \times 3 \times 129} \quad (28)
$$

Subsequently, we compute the pseudolabel y_i^{pl} and the unsupervised loss for unlabeled data:

$$
\mathbf{y}_i^{pl} = \mathcal{D}_\theta\left(A_{\text{weak}}\left(\mathbf{x}_i^{UK}\right)\right) \quad (29)
$$

$$
\hat{\mathbf{y}}_i^{pl} = \arg\max\left(\mathbf{y}_i^{pl}\right) \quad (30)
$$

$$
\begin{aligned}
\mathcal{L}_u = \mathop{\mathbb{E}}_{i \sim T} &\left[\mathbb{1} \cdot \left(\max\left(\mathbf{y}_i^{pl}\right) \ge 1\right) \right. \\
&\cdot \left(\hat{\mathbf{y}}_i^{pl} \cdot \log\left(\mathcal{D}_\theta\left(A_{\text{strong}}\left(\mathbf{x}_i^{UK}\right)\right)\right)\right. \\
&\left.\left. + \left(1 - \hat{\mathbf{y}}_i^{pl}\right) \cdot \log\left(1 - \mathcal{D}_\theta\left(A_{\text{strong}}\left(\mathbf{x}_i^{UK}\right)\right)\right)\right) \right] \quad (31)
\end{aligned}
$$

Finally, we employ equal weights for both losses to optimize the model $\mathcal{D}_\theta$.

$$
\mathcal{L} = \mathcal{L}_s + \mathcal{L}_u \quad (32)
$$

This approach balances the contributions of labeled and unlabeled data to the learning process, enhancing the model's performance.

I. Settings for Training the Turn Time Estimation Model

The training process of $\mathcal{D}_\theta$ was conducted using the stochastic gradient descent (SGD) optimizer, initialized with a learning rate of 0.01 and a momentum value of 0.9. A cosine annealing scheduler was employed to dynamically adjust the learning rate to enhance the learning process. The scheduler utilized a rate of $\frac{7}{16}\pi$. The model was trained over 12,000 iterations, with a batch size of 256. A data instance was randomly sampled from the entire dataset during each iteration. The model was implemented using PyTorch 1.13.1 with Python 3.9, while training was performed on an NVIDIA RTX 3090 Ti GPU. The entire training process was executed within an Ubuntu 20.04 Docker image environment.

J. Evaluation Metrics

We employed several metrics: mean absolute error (MAE), mean squared error (MSE), Pearson correlation coefficient (Corr.), intraclass correlation coefficient (3, 1) (ICC$_{3, 1}$), and relative error rate (Err.) to assess our algorithms. MAE and MSE are useful for identifying outliers, given their distinct error penalties. Corr., recognized for detecting linear variable relationships, provides valuable cross-study comparability. ICC$_{3, 1}$, essential for evaluating new clinical devices [40], harmonizes both correlation and agreement between measurements. An ICC$_{3, 1}$ value over 0.9 indicates "excellent" reliability, 0.75 to 0.9 is "good", 0.5 to 0.75 is "moderate", and below 0.5 is "poor". We also presented the relative error rate with its 95% confidence interval to showcase error dispersion.

K. PathoOpenGait System Design

To facilitate the recording, uploading, access, and analysis of gait trials by patients and doctors at any time and from any location, we developed PathoOpenGait as a cloud-based system. To ensure the security and privacy of patient data, we incorporated an authentication module. This module allows different types of users, such as doctors and patients, to have specific functionalities while preventing potential data breaches.

We introduced a flexible abstract interface to enhance the system's flexibility. This interface enables the easy integration of upgraded models, addition of new algorithms, and incorporation of different types of data inputs as per user requirements.

The system's front end was implemented using ReactJS, while the back end was developed using the Python Flask framework. The MySQL database was utilized to store the relevant data, and Python Celery was employed as the scheduler. An Nginx load balancer was introduced to optimize performance at the system's entry point. Furthermore, the entire system was containerized using Docker, allowing for convenient deployment on any cloud platform.

L. Demonstration of PathoOpenGait Within a Clinical Context

To effectively showcase the practical application of PathoOpenGait in a clinical setting, we conducted further recruitment of additional patients diagnosed with PD and HCs from the NTUH. The recruited participants underwent an 8-meter walking test, and the collected gait data were subsequently uploaded to the PathoOpenGait system for analysis.

Statistical analysis was performed to compare the gait parameters of the HC and PD groups and among the different subgroups within the PD population. The Wilcoxon signed-rank test was adjusted using Holm's method to compare parameters of the HC and PD groups. In contrast, the Kruskal–Wallis test was utilized to compare parameters across multiple groups.

Moreover, we examined the relationships between gait parameters and variables such as MMSE scores, LEDD, and disease duration in the PD group using Spearman's rank correlation. To ensure the validity of the comparison, only HCs who were age-

	Medical Center 1			Medical Center 2	
	HC	PD	MSA	HC	PD
# of participants	5	5	4	6	3
# of records	20	20	16	23	12
Age (year)	49.4 ± 21.2	66.0 ± 9.5	60.8 ± 10.1	58.8 ± 17.5	72.3 ± 8.7
Age range (year)	[20, 83]	[52, 78]	[49, 76]	[21, 72]	[63, 84]
Height (cm)	165.2 ± 8.3	161.8 ± 5.7	160.0 ± 1.9	160.2 ± 5.0	162.7 ± 10.2
Height range (cm)	[157, 178]	[155, 169]	[157,162]	[155, 169]	[150, 175]
Male (%)	40 (2)	60 (3)	50 (2)	16.7 (5)	66.6 (2)

and height-matched to the patient group were included in the statistical analysis.

III. RESULTS

A. Dataset Compilation for Algorithm Development and Verification

The testing dataset comprised 91 records obtained from 23 individuals associated with two distinct medical centers (Table I). This dataset contains two subsets: the first subset was collected from the NTUH, and it consists of 56 records derived from 14 participants. There were 20 records within this subset from 5 healthy controls (HCs), 20 from 5 patients diagnosed with PD, and 16 from 4 patients diagnosed with MSA. The second subset encompassing nine individuals who underwent 35 gait trials was obtained from the NTUCC. Within this subset were 23 records from 6 HCs and 12 records from 3 patients with PD.

An auxiliary training set from the NTUH was created to optimize the turn time prediction model. This set comprised 111 records obtained from 30 participants, including 38 records from 10 HCs, 39 from 10 patients with PD, and 34 from 10 patients with MSA. Furthermore, an unlabeled dataset containing 444 records from 114 individuals, with no additional information, was included for semi-supervised training. After performing data preprocessing, the final number of instances available for training comprised approximately 30,000 labeled data points and 120,000 unlabeled data points, each in $R^{17 \times 3 \times 129}$.

All the datasets used in this study were mutually exclusive to ensure the integrity of the verification results and prevent any information leakage. This approach ensured that the verification process was conducted with authenticity.

B. Algorithm Validation

Compared to the ground truth, our algorithm's validation results are presented in Table II and Fig. 2(a-l) to provide an insightful analysis of the algorithm's performance across the two exclusive datasets. Notably, the majority of the $\text{ICC}_{3, 1}$ values fall within the "good" to "excellent" reliability range, with error rates below 10%, underscoring the effectiveness of our approach. Yet, a closer examination reveals slight inferiority in SW and turn time performance.

In particular, a 12% turn time error in the first subset, coupled with a "moderate" $\text{ICC}_{3, 1}$ in the second, could stem from challenges annotators face in pinpointing exact action boundaries (Fig. 2(m)–(n)), introducing minor inconsistencies. On another

note, the second subset's SW error rate is twice that of the first, likely due to floor color. The first center's green floor contrasts with shoes, whereas the second's white floor blends, potentially affecting keypoint stability and causing SW estimation errors. In response to this finding, we plan to enhance our system's robustness to environmental variables in future work.

Further analysis of the Pearson correlation coefficient reveals that many parameters exhibit coefficients exceeding 0.90, indicative of strong correlations with the ground truth. However, a few parameters demonstrate coefficients ranging from 0.7 to 0.8. This discrepancy can potentially be attributed to the limited size of our dataset, as Pearson correlation coefficients tend to increase with larger datasets. Therefore, we anticipate even higher correlation coefficients with a more extensive dataset.

C. Ablation Study of the Semi-Supervised Learning Strategy

The ablation study results of the semi-supervised strategy are presented in Table III. Remarkably, the utilization of unlabeled data demonstrates a notable enhancement in the performance in terms of all evaluated metrics. This improvement is evident across both datasets, particularly in the second, signaling a bolstered adaptation of our algorithm and substantiating the efficacy of our approach in exploiting unlabeled data.

D. Cloud-Based PathoOpenGait System

To align with the clinical context, our PathoOpenGait system implements a user categorization framework, which classifies registered users into four distinct categories: administrator, manager, general user, and guest (Fig. 3). The administrator sets up the service and has full control over account management. Managers, usually doctors, monitor patients' gait parameters, create patient accounts, and upload data for analysis. Patients, categorized as general users, can upload gait trial records and view their results. However, they need their doctor's assistance to register an account.

To ensure privacy, doctors within the manager category do not have access to the information of other doctors' patients. Moreover, even administrators are restricted from viewing individual data to ensure confidentiality.

The system allows the addition of detailed patient statuses during uploads. Patient profiles capture crucial clinical data, such as diagnosis, MMSE score, and LEDD. A download feature lets doctors retrieve gait parameters alongside patient information for deeper analysis.

As the dataset grows and more data become available, it is reasonable to expect improved model performance through fine-tuning. We applied an abstracted Python interface called the Analyzer to facilitate this process, and it enables users to easily implement and update algorithms or incorporate different data formats. Without substantial code alterations, new models or data types are easily integrated, giving doctors choices in gait parameter analysis models.

Efficiency is key. We have adopted a first-in-first-out scheduling, with user-friendly status updates on data uploads. Load balancing ensures smooth service, and for security, every
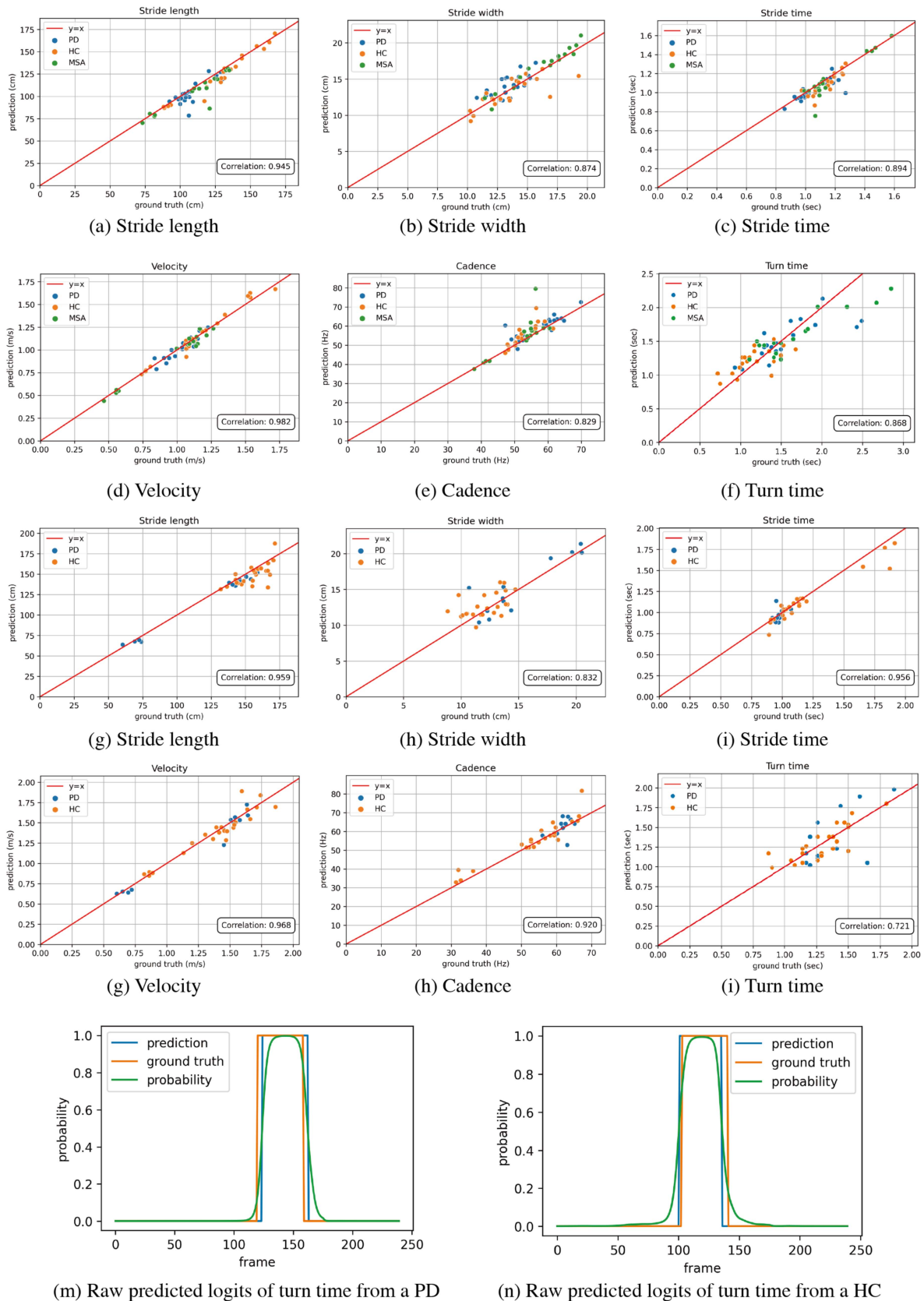
Fig. 2. **Comparison of the Ground Truth and Gait Parameter Estimations.** Subplots (a)–(f) correspond to the dataset collected at National Taiwan University Hospital, depicting the stride length, stride width, stride time, velocity, cadence, and turn time, respectively. Subplots (g)–(l) represent the National Taiwan University Cancer Center dataset, showing the same gait parameters. Subplots (m)–(n) demonstrate the raw predicted logits from our turn-time estimation model for one case of PD and one HC, respectively. Notably, the probability gradually increases and decreases within the boundaries, which makes it challenging for annotators to make precise judgments.

TABLE II

VALIDATION RESULTS OF THE GAIT PARAMETER ESTIMATION ALGORITHMS. THE TABLE PRESENTS THE VALIDATION RESULTS OF THE ALGORITHMS IN ESTIMATING VARIOUS GAIT PARAMETERS, INCLUDING STRIDE LENGTH, STRIDE WIDTH, STRIDE TIME, VELOCITY, CADENCE, AND TURN TIME. MOST OF THE ESTIMATED GAIT PARAMETERS DEMONSTRATE A "GOOD" OR "EXCELLENT" $ICC_{3,1}$ VALUE WITH A LOW ERROR RATE AND A STRONG CORRELATION WITH THE GROUND TRUTH. HOWEVER, ONE DATASET'S TURN TIME AND STRIDE WIDTH EXHIBIT A SLIGHTLY HIGHER ERROR RATE OF 10-12%, AND ONE TURN TIME SHOWS A "MODERATE" $ICC_{3,1}$ VALUE

	Medical Center 1: National Taiwan University Hospital					Medical Center 2: National Taiwan University Cancer Center				
	MAE	**MSE**	**Corr.**	**$ICC_{3,1}$**	**Err. (%) [95%CI]**	**MAE**	**MSE**	**Corr.**	**$ICC_{3,1}$**	**Err. (%) [95%CI]**
Stride length	5.746 cm	74.747 cm^2	0.945	0.945	5.0 [3.7; 6.2]	7.072 cm	93.132 cm^2	0.959	0.959	4.8 [3.7; 6.0]
Stride width	0.975 cm	1.609 cm^2	0.874	0.873	6.9 [5.6; 8.1]	1.354 cm	2.964 cm^2	0.832	0.831	11.3 [8.3; 14.4]
Stride time	0.052 s	0.006 s^2	0.894	0.890	4.7 [3.5; 5.9]	0.056 s	0.008 s^2	0.956	0.947	4.8 [3.3; 6.2]
Velocity	0.037 m/s	0.002 (m/s)^2	0.982	0.981	3.5 [2.9; 4.1]	0.064 m/s	0.008 (m/s)^2	0.968	0.968	4.6 [3.4; 5.8]
Cadence	2.842 (1/min)	22.417 (1/min)^2	0.829	0.815	5.2 [3.7; 6.8]	2.732 (1/min)	16.610 (1/min)^2	0.920	0.919	5.0 [3.4; 6.6]
Turn time	0.184 s	0.058 s^2	0.868	0.805	12.6 [10.7; 14.5]	0.141 s	0.034 s^2	0.721	0.707	10.7 [8.2; 13.2]

MAE: mean absolute error; MSE: mean square error; Corr.: Pearson correlation coefficient; $ICC_{3,1}$: intraclass correlation coefficient (3, 1); Err.: error rate; CI: confidence interval

TABLE III

PERFORMANCE COMPARISON OF TURN TIME ESTIMATION ALGORITHMS. THE NEURAL NETWORK-BASED ALGORITHM EXHIBITS SUPERIOR PERFORMANCE COMPARED TO THE HANDCRAFTED ALGORITHM. MOREOVER, UTILIZING A SEMI-SUPERVISED LEARNING STRATEGY FURTHER BOOSTS PERFORMANCE AND ENHANCES THE MODEL'S ADAPTABILITY ACROSS DATASETS

	Medical Center 1: National Taiwan University Hospital					Medical Center 2: National Taiwan University Cancer Center				
	MAE (s)↓	**MSE (s^2)↓**	**Corr.↑**	**$ICC_{3,1}$↑**	**Err. (%) [95%CI]↓**	**MAE (s)↓**	**MSE (s^2)↓**	**Corr.↑**	**$ICC_{3,1}$↑**	**Err. (%) [95%CI]↓**
Handcrafted algorithm*	0.663	0.687	0.455	0.445	39.7 [33.8; 45.6]	0.780	0.807	0.270	0.242	48.8 [41.0; 56.6]
NN-based algorithm	0.199	0.061	0.840	**0.822**	14.3 [12.0; 16.6]	0.195	0.076	0.574	0.530	14.6 [10.6; 18.6]
NN-based algorithm (with semi-supervised)	**0.184**	**0.058**	**0.868**	0.805	**12.6 [10.7; 14.5]**	**0.141**	**0.034**	**0.721**	**0.707**	**10.7 [8.2; 13.2]**

*: achieved by optionally selecting threshold values ($k_1 = 75$ and $k_2 = 10$); NN: neural network; MAE: mean absolute error; MSE: mean square error;
Corr.: Pearson correlation coefficient; $ICC_{3,1}$: intraclass correlation coefficient (3, 1); Err.: error rate; CI: confidence interval; ↓: the lower the better; ↑: the higher the better
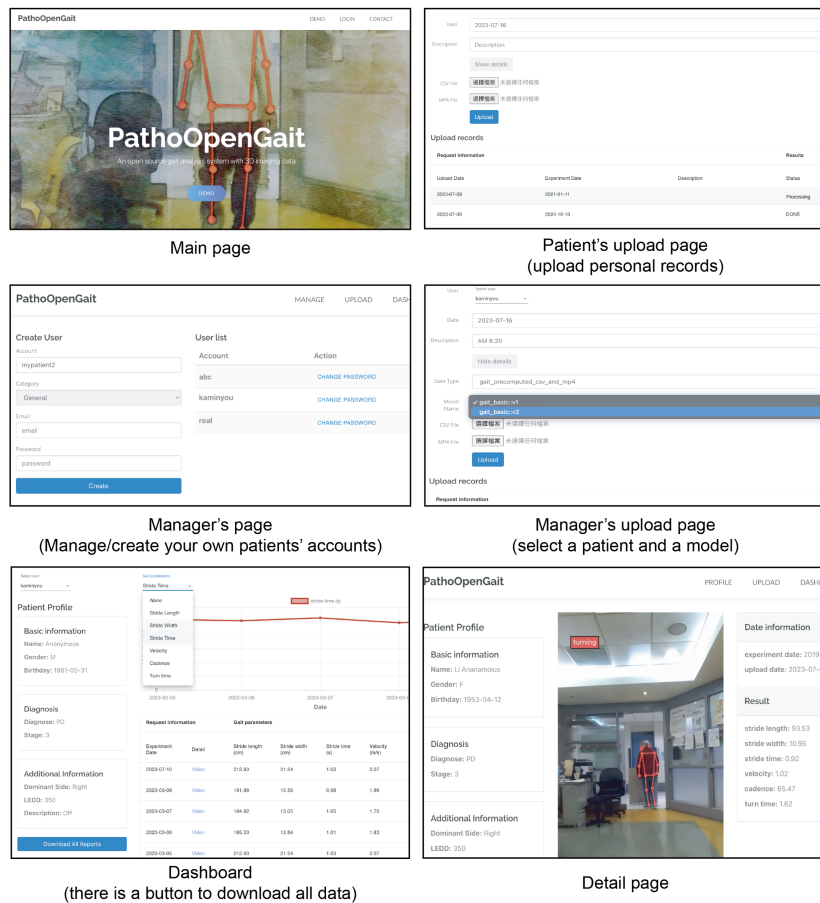The best-performing method in each experiment is highlighted in bold and underlined.

Main page

Patient's upload page
(upload personal records)

Manager's page
(Manage/create your own patients' accounts)

Manager's upload page
(select a patient and a model)

Dashboard
(there is a button to download all data)

Detail page

Fig. 3.   **User Categories and Functionality in the PathoOpenGait System.** PathoOpenGait system classifies users into four categories: guests, general users, managers, and administrators. The administrator holds complete control over all accounts within the system. On the other hand, the manager, typically a doctor, possesses the authority to create accounts solely for their patients. Managers can upload videos for their patients, view the inference results on the dashboard, and access a download button to obtain all estimated gait parameters for further analysis. General users, typically patients, can upload their data, modify their profiles, and utilize the dashboard to monitor the progression of their disease. Guests, however, are limited to viewing the demo page without access to additional functionalities.

<div align="center">

TABLE IV

CHARACTERISTICS OF THE PARTICIPANTS. THE P VALUES FOR COMPARING THE AGE, HEIGHT, SEX, AND THE DISTRIBUTION OF GAIT SYMPTOMS AMONG GROUPS WERE ALL > 0.05

</div>

	Total	H&Y stage 1	H&Y stage 2	H&Y stage 3	HC
# of participants	105	30	45	30	41
Age (year)	70.6 ± 7.8	67.5 ± 7.5	70.7 ± 6.9	73.5 ± 8.5	68.1 ± 9.3
Height (cm)	158.5 ± 17.6	160.2 ± 7.3	161.0 ± 8.2	153.2 ± 30.1	158.8 ± 9.3
Male (%)	53 (50)	14 (47)	23 (51)	16 (53)	14 (34)
Disease duration (year)	6.3 ± 4.3	4.6 ± 3.6	6.2 ± 4.1	8.2 ± 4.5	n.a.
LEDD (mg)	567.8 ± 383.0	376.9 ± 300.0	603.2 ± 354.8	704.8 ± 432.3	n.a.
MMSE	24.8 ± -4.2	25.3 ± 3.0	25.6 ± 3.4	22.3 ± 6.1	n.a.
Predominant side (L/R)	61/44	19/11	23/22	19/11	n.a.
UPDRS (III)	12.3 ± 8.2	7.2 ± 4.3	11.7 ± 7.0	18.2 ± 9.4	n.a.
UPDRS-tremor	1.4 ± 2.2	1.3 ± 2.0	1.1 ± 2.3	1.9 ± 2.4	n.a.
UPDRS-PGID	3.4 ± 3.0	2.0 ± 2.6	3.2 ± 2.8	4.9 ± 2.9	n.a.
Gait symptoms					
Shuffling gait (%)	44 (42)	13 (43)	15 (33)	16 (53)	n.a.
Freezing of gait (%)	9 (9)	2 (7)	4 (9)	3 (10)	n.a.
Difficulty in turning (%)	12 (11)	4 (13)	5 (11)	3 (10)	n.a.
Difficulty in initiation (%)	9 (9)	2 (7)	4 (9)	3 (10)	n.a.
Gait parameters					
Stride length (cm)	96.3 ± 28.4	109.1 ± 27.4	98.5 ± 24.1	80.1 ± 28.4	125.3 ± 25.0
Stride width (cm)	14.0 ± 2.8	13.4 ± 2.9	14.3 ± 2.9	14.1 ± 2.4	12.7 ± 2.1
Stride time (s)	1.1 ± 0.2	1.1 ± 0.1	1.0 ± 0.1	1.1 ± 0.2	1.0 ± 0.1
Velocity (cm/s)	93.1 ± 31.9	105.0 ± 31.3	96.7 ± 28.2	75.8 ± 31.3	123.3 ± 29.4
Cadence (1/min)	57.2 ± 7.1	57.1 ± 6.7	58.3 ± 6.4	55.4 ± 8.4	58.7 ± 5.1
Turn time (s)	1.5 ± 0.4	1.5 ± 0.3	1.6 ± 0.4	1.7 ± 0.5	1.3 ± 0.3

H&Y, Hoehn, and Yahr; LEDD, levodopa equivalent daily dose; UPDRS, United Parkinson's Disease Rating Scale; n.a.: not applicable
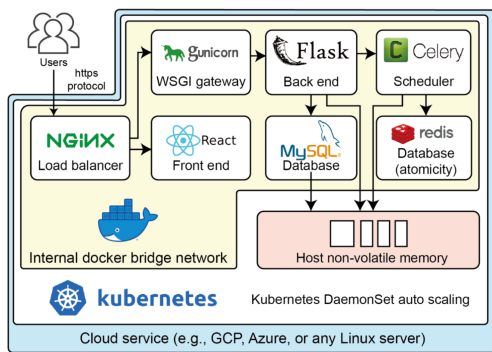

Fig. 4. **PathoOpenGait System Design.** The system incorporates an entry point facilitated by the Nginx load balancer module. The front end is developed using ReactJS, while the back end is implemented using the Python Flask framework. Python Celery serves as the scheduler, efficiently dispatching inference jobs with a queuing algorithm achieved through the atomicity of Redis. Users' data and inference results are stored in MySQL, while the videos are stored in the nonvolatile memory of the host. All components are interconnected through an internal Docker bridge network, containerized, and easily managed by Kubernetes for seamless autoscaling. The system can be deployed on any cloud service or Linux server, providing implementation flexibility and versatility.

component resides within Docker containers, safeguarding internal operations while exposing only the front-end endpoint.

Setting up our PathoOpenGait system is effortless with the Docker-compose command. It can be seamlessly deployed in various Linux environments or on commercial cloud platforms such as Google Cloud Platform and Amazon Web Services, providing flexibility and scalability (Fig. 4).

E. Demonstration of the PathoOpenGait System

We used the PathoOpenGait system to collect data from 122 patients diagnosed with PD and 124 HCs. A total of 363 gait records for PD patients and 152 records for HCs were then used for the statistical analysis. To ensure comparability, a data cleaning process was performed to align age and sex information. Detailed characteristics of the participants can be found in Table IV.

Initially, we compared the gait parameters of the PD patients and HCs, both overall and across different stages of PD. Our findings revealed that compared to HCs PD patients exhibited reduced SL ($p < 0.001$), along with slight increases in SW ($p < 0.001$) and TT ($p < 0.001$). To further investigate the minor differences in ST and cadence, we conducted separate analyses on patients in stages 1 to 3 of the disease. As anticipated, SL ($p < 0.001$) and velocity ($p < 0.001$) consistently decreased as the disease progressed among PD patients. However, we only observed slight increases in ST ($p < 0.01$) and cadence ($p < 0.01$) in stage 3.

Furthermore, we performed analyses on subgroups of PD patients. Individuals who exhibited poor initiation, shuffling, and freezing showed reduced SL and V (all $p < 0.05$). Those with difficulties turning demonstrated decreased ST and increased cadence (all $p < 0.05$). In contrast, those with freezing patterns exhibited increased TT ($p < 0.001$).

In addition, we explored the relationship between gait parameters and other variables, such as MMSE scores, LEDD, and disease duration. We found that increased TT alone may be associated with lower MMSE scores and longer disease duration ($p < 0.001$). Decreased SL and V also indicated lower MMSE scores ($p < 0.001$). These strong correlations were observed in both stage 1 and stage 2 when conducting a stratified analysis based on the disease stage. A higher LEDD was associated with increased SW ($p < 0.05$), while a decline in ST or an increase in cadence indicated a lower LEDD in stages 1 ($p < 0.05$) and 2 ($p < 0.05$). Finally, an increase in SL was associated with longer disease duration in stage 3 ($p < 0.05$).

	Stride length			Stride time			Velocity			Cadence		
	Corr.↑	$ICC_{3,1}$↑	Err. (%)↓	Corr.↑	$ICC_{3,1}$↑	Err. (%)↓	Corr.↑	$ICC_{3,1}$↑	Err. (%)↓	Corr.↑	$ICC_{3,1}$↑	Err. (%)↓
Wearables (on waist) [19]	0.840	n.a.	12.3	0.860	n.a.	3.7	0.870	n.a.	26.6	0.910	n.a.	3.0
Wearables (on condyles)* [20]	0.890	0.610	n.a.	-	-	-	0.980	0.930	n.a.	0.990	0.990	n.a.
Wearables (on feet)* [20]	0.990	0.990	n.a.	n.a.	n.a.	n.a.	0.990	0.990	n.a.	0.990	0.990	n.a.
2D camera (sagittal view) [26]	-	-	-	-	-	-	0.730	n.a.	13.0	0.790	n.a.	8.0
2D camera (frontal view) [28]	0.704	n.a.	n.a.	0.948	n.a.	n.a.	0.877	n.a.	n.a.	0.984	n.a.	n.a.
3D infrared camera* [30]	0.840	0.670	9.0	0.930	0.940	1.0	-	-	-	-	-	-
3D binocular camera (ours)†	0.963	0.963	4.9	0.928	0.926	4.7	0.978	0.978	3.9	0.876	0.872	5.2

*: involved healthy participants only.; †: involved testing datasets from both medical centers; Corr.: Pearson correlation coefficient; $ICC_{3,1}$: intraclass correlation coefficient (3, 1); Err.: error rate; n.a.: not applicable due to the metric not being measured; -: not applicable due to the gait parameter not being measured; ↓: the lower the better; ↑: the higher the better

IV. DISCUSSIONS

Over the years, various techniques have been developed for estimating gait parameters and conducting gait analysis [3], [41], [42], [43]. These techniques range from traditional carpet-based methods [14] to more advanced approaches utilizing 3D camera arrays [16], wearable devices [12], and computer vision-based methods in either 2D or 3D environments [26], [29]. However, even with advancements in these techniques, most studies haven't provided platforms or source code, limiting clinical application. In this study, we developed and demonstrated PathoOpenGait, a portable and cloud-based platform for gait analysis that supports cross-center tracking and long-term monitoring. Patients can conveniently upload their gait records from the comfort of their own homes. To our knowledge, PathoOpenGait is currently the only open-source and flexible platform available for gait analysis.

With the introduction of our platform, understanding the functioning and optimal use of our platform is vital. Our algorithms leverage components from pre-trained model weights, such as the Mask R-CNN, followed by a 3D keypoint augmentation. In scenarios with dark, complex, or occluded scenes, and non-contrasting floors that lead to inaccurate 2D keypoint extraction, these errors can propagate to the 3D augmentation, adversely affecting the final estimation of gait parameters. Therefore, recording in well-lit, clear settings is essential for accurate results. Moreover, to minimize errors, such as becoming undetectable in the video when moving beyond a 4-meter trial distance, patients must strictly adhere to the trial instructions. To quantify potential errors, our system evaluates the presence and reliability of detected keypoints. If our system detects keypoints with confidence scores below 0.5, as recommended by the official OpenPose and Detectron2 documentation [36], [44], or identifies null values, it will prompt users to re-record.

We present a comparison of gait parameter estimation performance with various methods in Table V. SW, reported by only one infrared camera study [30] ($ICC_{3,1} = 0.710$; Corr. = 0.910), and turn time, noted by a single study using wearables on the lumbar spine [45] ($ICC_{3,1} = 0.690$ to 0.800), are omitted. Wearable devices, when affixed to the waist or condyles, tend to exhibit suboptimal performance, particularly in depth-related gait parameters like SL ($ICC_{3,1} = 0.610$) and velocity (Err. = 26.6%). 2D cameras utilizing frontal view videos may outperform those with sagittal view; however, the relevant

3D estimation algorithms often fail to precisely capture depth information, compromising SL performance (Corr. = 0.704). Additionally, 3D infrared cameras demonstrate limited capacity in capturing depth information, resulting in an inadequate estimation of SL ($ICC_{3,1} = 0.670$). Conversely, our binocular camera system, coupled with our proposed algorithms, is capable of estimating all gait parameters with $ICC_{3,1}$ values ranging from "good" to "excellent" reliability (SW of 0.858 and turn time of 0.791), as observed in a combined analysis of two datasets.

Furthermore, we developed the world's first turn time estimation algorithm that employs 2D video as its input. While its "moderate" to "good" $ICC_{3,1}$ values and 10% to 12% error rates may seem inferior compared to other gait parameters, our algorithm is still able to pinpoint the freezing of gait symptom in the PD group ($p < 0.001$). The slight inferiority is a consequence of the challenges in determining specific moments within a turn as part of the turning phase (Fig. 2(m)–(n)). Moreover, individuals' diverse range of turning patterns, such as stopping before turning or turning while walking, makes it infeasible to mathematically model these patterns using handcrafted deterministic algorithms (Table III). In such cases, neural networks offer an effective approach, as they can automatically learn and capture these complex patterns during optimization [46].

However, utilizing neural networks for turn-time estimation necessitates framewise annotations, which can be laborious and time-consuming. In this study, we demonstrated the effectiveness of semi-supervised learning in the context of 1D signals, showcasing its potential for reducing the labeling burden in gait analysis research. This is a notable contribution, particularly considering that most studies applying semi-supervised learning focus on benchmarking in image applications [39], [47], [48].

Additionally, we showcased the suitability and efficiency of PathoOpenGait for doctors to conduct gait analysis and easily compare features across different groups. One of the most appealing features of PathoOpenGait is its flexibility. Making PathoOpenGait an open-source platform and providing abstract interfaces for developers to add plugins allows the entire framework to be easily modified for various applications involving data uploading and inference. This eliminates the need for reimplementing the authentication module and scheduler, which are typically labor-intensive tasks, thus enabling seamless transfer to other applications with minimal effort. Our codebase is ready for deployment in a hospital research setting. For clinical use beyond studies, users must heed local regulations.

Despite its merits, our study does have certain limitations. Although a $449 3D binocular camera is more affordable than most mobile phones and can be accessible to many users, it requires an additional device purchase. However, it is still more cost-effective than carpets or 3D array cameras, and PathoOpenGait is freely available. To mitigate this limitation, our system retains the capability to estimate turn time using just 2D videos in mp4 format from a 2D camera. We also offer guidelines for adapting 3D trajectories from various cameras to our system. However, users should be cautious, as performance may differ with cameras beyond our original design. Our future goal is to improve current 3D estimation algorithms and develop a system that accurately works with solely 2D videos.

Another limitation involves the linguistic accessibility of PathoOpenGait's current interface, which may present difficulties for users who are not adept in English. We encourage users and developers worldwide to contribute by making pull requests to the PathoOpenGait GitHub repository and helping translate the interface into their respective languages.

Finally, for our PathoOpenGait to transition from a study context to a clinical application, thorough software verification is crucial [49]. This includes security checks, QA standards, automated tests, and extensive documentation. As regulatory standards vary by country [50], achieving universal clinical readiness is complex. We've employed static analysis for security flaws in our code and plan to add dynamic analysis. Additionally, we intend to strengthen our code with unit and integration tests, supported by a detailed QA report.

In conclusion, we introduce PathoOpenGait, as the first open-source, cloud-based platform for gait analysis, facilitating clinical research and long-term monitoring of patients in any location. We demonstrate the efficacy of semi-supervised learning in enhancing model performance through the case of turn time estimation. The flexibility and modifiability of PathoOpenGait empower developers to further improve model performance or adapt the platform for other applications that support the advancement of digital medicine.

ACKNOWLEDGMENT

REFERENCES

[1] L. Di Biase et al., "Gait analysis in Parkinson's disease: An overview of the most accurate markers for diagnosis and symptoms monitoring," *Sensors*, vol. 20, no. 12, 2020, Art. no. 3529.

[2] V. Sidoroff et al., "Characterization of gait variability in multiple system atrophy and Parkinson's disease," *J. Neurol.*, vol. 268, pp. 1770–1779, 2021.

[3] G. Cicirelli, D. Impedovo, V. Dentamaro, R. Marani, G. Pirlo, and T. R. D'Orazio, "Human gait analysis in neurodegenerative diseases: A review," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 1, pp. 229–242, Jan. 2022.

[4] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly: A clinical guide," *Wiener Klinische Wochenschrift*, vol. 129, no. 3-4, pp. 81–95, 2017.

[5] S. Perez-Lloret et al., "Prevalence, determinants, and effect on quality of life of freezing of gait in Parkinson disease," *JAMA Neurol.*, vol. 71, no. 7, pp. 884–890, 2014.

[6] R. Bouca-Machado et al., "Gait kinematic parameters in Parkinson's disease: A systematic review," *J. Parkinson's Dis.*, vol. 10, no. 3, pp. 843–853, 2020.

[7] M. R. Daliri, "Automatic diagnosis of neuro-degenerative diseases using gait dynamics," *Measurement*, vol. 45, no. 7, pp. 1729–1734, 2012.

[8] C. Gao, J. Liu, Y. Tan, and S. Chen, "Freezing of gait in Parkinson's disease: Pathophysiology, risk factors and treatments," *Transl. Neurodegener.*, vol. 9, pp. 1–22, 2020.

[9] C. G. Goetz et al., "Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disorders: Official J. Movement Disorder Soc.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[10] A. L. Behrman, K. E. Light, and G. M. Miller, "Sensitivity of the tinetti gait assessment for detecting change in individuals with Parkinson's disease," *Clin. Rehabil.*, vol. 16, no. 4, pp. 399–405, 2002.

[11] A. Mirelman et al., "Gait impairments in Parkinson's disease," *Lancet Neurol.*, vol. 18, no. 7, pp. 697–708, 2019.

[12] W. Tao, T. Liu, R. Zheng, and H. Feng, "Gait analysis using wearable sensors," *Sensors*, vol. 12, no. 2, pp. 2255–2283, 2012.

[13] K. E. Webster, J. E. Wittwer, and J. A. Feller, "Validity of the GAITRite walkway system for the measurement of averaged and individual step parameters of gait," *Gait Posture*, vol. 22, no. 4, pp. 317–321, 2005.

[14] B. Bilney, M. Morris, and K. Webster, "Concurrent related validity of the GAITRite walkway system for quantification of the spatial and temporal parameters of gait," *Gait Posture*, vol. 17, no. 1, pp. 68–74, 2003.

[15] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, "Comparative abilities of microsoft kinect and vicon 3D motion capture for gait analysis," *J. Med. Eng. Technol.*, vol. 38, no. 5, pp. 274–280, 2014.

[16] E. Mirek, M. Rudzińska, and A. Szczudlik, "The assessment of gait disorders in patients with Parkinson's disease using the three-dimensional motion analysis system vicon," *Neurologia I Neurochirurgia Polska*, vol. 41, no. 2, pp. 128–133, 2007.

[17] B. Carse, B. Meadows, R. Bowers, and P. Rowe, "Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3D motion analysis system," *Physiotherapy*, vol. 99, no. 4, pp. 347–351, 2013.

[18] L. C. Benson, C. A. Clermont, E. Bošnjak, and R. Ferber, "The use of wearable devices for walking and running gait analysis outside of the lab: A systematic review," *Gait Posture*, vol. 63, pp. 124–138, 2018.

[19] M. Zago et al., "Gait evaluation using inertial measurement units in subjects with Parkinson's disease," *J. Electromyogr. Kinesiol.*, vol. 42, pp. 44–48, 2018.

[20] K. Homan, K. Yamamoto, K. Kadoya, N. Ishida, and N. Iwasaki, "Comprehensive validation of a wearable foot sensor system for estimating spatiotemporal gait parameters by simultaneous three-dimensional optical motion analysis," BMC Sports Science, *Med. Rehabil.*, vol. 14, no. 1, pp. 1–11, 2022.

[21] M. Moro, G. Marchesi, F. Hesse, F. Odone, and M. Casadio, "Markerless vs. marker-based gait analysis: A proof of concept study," *Sensors*, vol. 22, no. 5, 2022, Art. no. 2011.

[22] Y. Celik, S. Stuart, W. L. Woo, and A. Godfrey, "Wearable inertial gait algorithms: Impact of wear location and environment in healthy and parkinson's populations," *Sensors*, vol. 21, no. 19, 2021, Art. no. 6476.

[23] J. Wang et al., "Deep 3D human pose estimation: A review," *Comput. Vis. Image Understanding*, vol. 210, 2021, Art. no. 103225.

[24] W. Liu, Q. Bao, Y. Sun, and T. Mei, "Recent advances of monocular 2D and 3D human pose estimation: A deep learning perspective," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–41, 2022.

[25] D. Sethi, S. Bharti, and C. Prakash, "A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work," *Artif. Intell. Med.*, vol. 129, 2022, Art. no. 102314.

[26] Ł. Kidziński, B. Yang, J. L. Hicks, A. Rajagopal, S. L. Delp, and M. H. Schwartz, "Deep neural networks enable quantitative movement analysis using single-camera videos," *Nature Commun.*, vol. 11, no. 1, 2020, Art. no. 4054.

[27] J. Stenum, C. Rossi, and R. T. Roemmich, "Two-dimensional video-based analysis of human gait using pose estimation," *PLoS Comput. Biol.*, vol. 17, no. 4, 2021, Art. no. e1008935.

[28] R. J. Cotton, E. McClerklin, A. Cimorelli, A. Patel, and T. Karakostas, "Transforming gait: Video-based spatiotemporal gait analysis," in *Proc. IEEE 44th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2022, pp. 115–120.

[29] A. Azhand, S. Rabe, S. Müller, I. Sattler, and A. Heimann-Steinert, "Algorithm based on one monocular video delivers highly valid and reliable gait parameters," *Sci. Rep.*, vol. 11, no. 1, 2021, Art. no. 14065.

[30] M. Eltoukhy, J. Oh, C. Kuenze, and J. Signorile, "Improved kinect-based spatiotemporal and kinematic treadmill gait assessment," *Gait Posture*, vol. 51, pp. 77–83, 2017.

[31] A. L. McDonough, M. Batavia, F. C. Chen, S. Kwon, and J. Ziai, "The validity and reliability of the GAITRite system's measurements: A preliminary evaluation," *Arch. Phys. Med. Rehabil.*, vol. 82, no. 3, pp. 419–425, 2001.

[32] M. Kelly, P. Jones, R. Wuebbles, V. Lugade, D. Cipriani, and N. G. Murray, "A novel smartphone application is reliable for repeat administration and comparable to the Tekscan Strideway for spatiotemporal gait," *Measurement*, vol. 192, 2022, Art. no. 110882.

[33] L. E. Ortiz, E. V. Cabrera, and L. M. Gonçalves, "Depth data error modeling of the zed 3D vision sensor from stereolabs," *ELCVIA: Electron. Lett. Comput. Vis. Image Anal.*, vol. 17, no. 1, pp. 1–15, 2018.

[34] J. E. Van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Mach. Learn.*, vol. 109, no. 2, pp. 373–440, 2020.

[35] M. Zago, M. Luzzago, T. Marangoni, M. De Cecco, M. Tarabini, and M. Galli, "3D tracking of human motion using visual skeletonization and stereoscopic vision," *Front. Bioeng. Biotechnol.*, vol. 8, 2020, Art. no. 181.

[36] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 7291–7299.

[37] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2961–2969.

[38] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7753–7762.

[39] K. Sohn et al., "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, vol. 33, pp. 596–608.

[40] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *J. Chiropractic Med.*, vol. 15, no. 2, pp. 155–163, 2016.

[41] P. Khera and N. Kumar, "Role of machine learning in gait analysis: A review," *J. Med. Eng. Technol.*, vol. 44, no. 8, pp. 441–467, 2020.

[42] Y. Celik, S. Stuart, W. L. Woo, and A. Godfrey, "Gait analysis in neurological populations: Progression in the use of wearables," *Med. Eng. Phys.*, vol. 87, pp. 9–29, 2021.

[43] A. Michelini, A. Eshraghi, and J. Andrysek, "Two-dimensional video gait analysis: A systematic review of reliability, validity, and best practice considerations," *Prosthetics Orthotics Int.*, vol. 44, no. 4, pp. 245–262, 2020.

[44] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron 2," 2019. [Online], Available: https://github.com/facebookresearch/detectron2

[45] M. M. Koop, S. J. Ozinga, A. B. Rosenfeldt, and J. L. Alberts, "Quantifying turning behavior and gait in Parkinson's disease using mobile technology," *IBRO Rep.*, vol. 5, pp. 10–16, 2018.

[46] A. Cochocki and R. Unbehauen, *Neural Networks for Optimization and Signal Processing.* Hoboken, NJ, USA: Wiley, 1993.

[47] Z. Cai et al., "Semi-supervised vision transformers at scale," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, vol. 35, pp. 25697–25710.

[48] B. Zhang et al., "Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling," in *Proc. Adv. Neural Inf. Process. Syst.*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., 2021, pp. 18408–18419. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2021/file/995693c15f439e3d189b06e89d145dd5-Paper.pdf

[49] US Department of Health and Human Services. Food and Drug Administration, "General principles of software validation: Final guidance for industry and FDA staff," 2002. Accessed: Oct. 19, 2023. [Online]. Available: https://www.fda.gov/media/73141/download

[50] D. B. Kramer, Y. T. Tan, C. Sato, and A. S. Kesselheim, "Ensuring medical device effectiveness and safety: A cross-national comparison of approaches to regulation," *Food Drug Law J.*, vol. 69, no. 1, 2014, Art. no. 1.