

# Video-Based Detection of Freezing of Gait in Daily Clinical Practice in Patients With Parkinsonism

Yuki Kondo<sup>ID</sup>, Kyota Bando<sup>ID</sup>, Ippei Suzuki, Yuta Miyazaki<sup>ID</sup>, Daisuke Nishida<sup>ID</sup>, Takatoshi Hara, Hideki Kadone<sup>ID</sup>, and Kenji Suzuki<sup>ID</sup>, *Senior Member, IEEE*

**Abstract**—Freezing of gait (FoG) is a prevalent symptom among individuals with Parkinson’s disease and related disorders. FoG detection from videos has been developed recently; however, the process requires using videos filmed within a controlled environment. We attempted to establish an automatic FoG detection method from videos taken in uncontrolled environments such as in daily clinical practices. Motion features of 16 patients were extracted from timed-up-and-go test in 109 video data points, through object tracking and three-dimension pose estimation. These motion features were utilized to form the FoG detection model, which combined rule-based and machine learning-based models. The rule-based model distinguished the frames in which the patient was walking from those when the patient has stopped, using the pelvic position coordinates; the machine learning-based model distinguished between FoG and stop using a combined one-dimensional convolutional neural network and long short-term memory (1dCNN-LSTM). The model achieved a high intraclass correlation coefficient of 0.75–0.94 with a manually-annotated duration of FoG and %FoG. This method is novel as it combines object tracking, 3D pose

estimation, and expert-guided feature selection in the pre-processing and modeling phases, enabling FoG detection even from videos captured in uncontrolled environments.

**Index Terms**—Freezing of gait, machine learning, Parkinson’s disease, pose estimation, video.

## I. INTRODUCTION

**F**REEZING of gait (FoG) is a characteristic symptom of Parkinson’s disease (PD) and related disorders, defined as “episodic absence or pronounced reduction in the forward progression of the feet during ambulation despite the intention to walk” [1]. The reported prevalence of FoG is 50.6%, which gradually increases with the progression of the disease [2]. Moreover, FoG frequently leads to falls [3], [4], fall-related injuries, and loss of independence [5], [6], [7].

Accurate assessment of FoG is essential for initiating or modifying treatment and evaluating its efficiency. Establishing the effectiveness of any intervention requires a precise assessment of FoG [8], [9]. Various clinical methods are used to assess the severity of FoG, including qualitative tools such as the Movement Disorder Society-sponsored revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) [10], FoG questionnaire (FoG-Q) [11], and New FoG Questionnaire (NFoG-Q) [12]. In addition, more sophisticated quantitative methods such as video-recorded [13], [14], [15], three-dimensional (3D) motion capture systems [16], force plates [17], [18], Inertial Measurement Units (IMU) [19], [20], [21], [22], [23], [24], [25], [26], [27], and electroencephalography caps [28], [29] have been employed to detect FoG episodes [30].

Scoring FoG based on video-recorded walking tasks is being increasingly acknowledged as the gold standard for FoG assessment [13], [14], [15]. However, identifying FoG events requires trained experts and is time-consuming. Therefore, there is a growing interest in developing automatic video-based FoG detection methods using machine learning. Automatic FoG detection has previously been attempted using red–green–blue (RGB) cameras for 2D keypoint recognition [31], [32], [33], [34], allowing image capturing with a monocular camera, without the use of external scales or markers to estimate joint coordinates. Compared to specialized equipment—such as depth sensors, IMU, and 3D motion

Manuscript received 30 August 2023; revised 16 January 2024 and 7 May 2024; accepted 9 June 2024. Date of publication 12 June 2024; date of current version 24 June 2024. This work was supported in part by JSPS KAKENHI under Grant 23H00485. (Corresponding author: Yuki Kondo.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted in part by the Ethics Committee of the University of Tsukuba, Japan, under Approval No. 2023R739; and in part by the National Center of Neurology and Psychiatry under Approval No. A2021-062, and performed in line with the Declaration of Helsinki.

Yuki Kondo is with the Department of Intelligent Interaction Technologies, Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba 305-8577, Japan, and also with the Department of Physical Rehabilitation, National Center Hospital, National Center of Neurology and Psychiatry, Tokyo 187-8551, Japan (e-mail: kondo.ncnp@gmail.com).

Kyota Bando, Ippei Suzuki, Yuta Miyazaki, Daisuke Nishida, and Takatoshi Hara are with the Department of Physical Rehabilitation, National Center Hospital, National Center of Neurology and Psychiatry, Tokyo 187-8551, Japan.

Hideki Kadone is with the Center for Innovative Medicine and Engineering and the Center for Cybernics Research, University of Tsukuba, Tsukuba 305-8577, Japan.

Kenji Suzuki is with the Institute of Systems and Information Engineering and the Center for Cybernics Research, University of Tsukuba, Tsukuba 305-8573, Japan.

This article has supplementary downloadable material available at <https://doi.org/10.1109/TNSRE.2024.3413055>, provided by the authors. Digital Object Identifier 10.1109/TNSRE.2024.3413055

capture systems—affordable devices like smartphones enable recording of different FoG phenomena by researchers, patients, and their families, in the form of easily available videos. Furthermore, unlike wearable devices which rely on estimates and cannot be verified after the event has occurred, video-based methods allow for the confirmation of FoG episodes even after screening. Because FoG rarely occurs in a laboratory setting [35], [36], [37], it is essential to develop a method that is capable of detecting FoG at home or other settings. FoG detection from videos is therefore a promising approach for addressing the ailment.

This study aimed to establish a method for automatically detecting FoG in videos captured during daily clinical practices. In addition, we validated the ability of this method to quantify FoG, and also evaluated its reliability.

Although there have been recent advancements in automatic FoG detection, several challenges still persist that limit the accurate quantification of FoG and its automatic detection in diverse daily clinical and home environments. First, the environments in which videos can be captured for FoG detection are limited. Most studies rely on videos captured in controlled environments, such as laboratories, where camera positions are fixed. Uncontrolled environments include different field-of-view, variation in light condition, disturbance by other therapists or patients, or video captured without indicating the start and end of gait. However, because individuals with PD and related disorders often experience FoG in their daily lives [35], [36], [37], whether at home or in outdoor environments, it is challenging to capture videos under consistent conditions. Therefore, establishing methods that are less affected by variations in camera positions is essential. Second, in previous studies, videos typically featured only the participant [33], or the participant with no more than two assistants [31], [32], [34]. However, videos from daily clinical practices often capture multiple individuals, including patients, their families, and therapists who prevent falls. Therefore, algorithms capable of selecting an analysis target among multiple individuals in a video are required.

Considering the aforementioned aspects, we chose and applied an object-tracking technique on videos captured in daily clinical practices to select the target patient, followed by 3D pose estimation to extract motion features. An algorithm was developed to detect the duration of FoG based on the temporal characteristics of these motion features. In contrast to quantifying FoG in a controlled environment, it was more convenient to quantify it using camera videos, as it allowed unrestricted movement for patients and could help detect FoG at home; in addition, this approach allowed detection of FoG from videos capturing other individuals as well. Furthermore, the ability to directly quantify FoG from videos may reduce the workload, which would be both time-saving and patient-friendly, and therefore convenient for daily use.

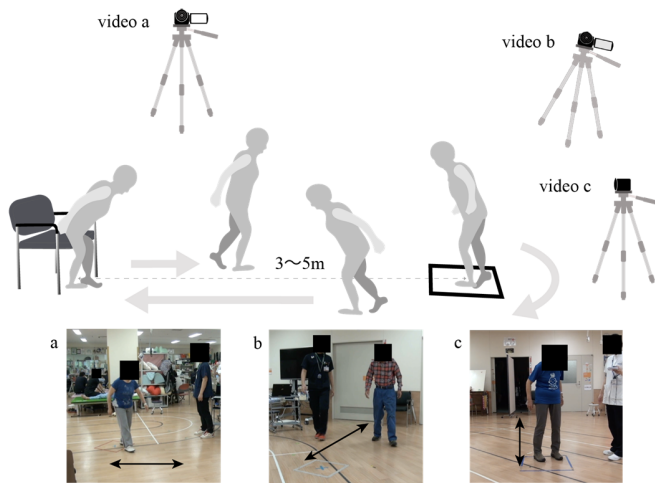
The novelty of our method primarily resides in the preprocessing and modeling phases. By focusing on these stages, we successfully developed a method capable of automatically detecting FoG from videos captured in uncontrolled environments. Our approach markedly differs from previous studies like those by Hu et al. [31], [32], Li et al. [33], Shalin et al. [38], and Sun et al. [34], especially in terms of video capture

methodology. Unlike these prior works, which often relied on videos shot from limited or fixed camera angles, our method can effectively utilize videos captured from a variety of angles. In the preprocessing phase, rather than training the model directly on the raw video data, we included an initial step that involved isolating the subject from videos showing multiple individuals, using object tracking techniques. Following this, we applied 3D pose estimation to the acquired 3D skeletal data. When using videos captured from the same position, similar features can be obtained from 2D poses with sufficient accuracy. However, for videos captured from different positions and angles, the features obtained vary depending on the viewpoint, thereby limiting the effectiveness of 2D pose extraction methods. Thus, we considered that 3D estimation would allow us to extract poses from any viewpoint, resulting in more accurate FoG detection. In the modeling phase, these data were then utilized to construct our model, integrating features that were critically observed by experts during the assessment of FoG. In the Forward Progression Identification Model, we employed a simple rule-based approach, using the pelvic marker to determine the presence of any forward movement. Segments identified as stops by this model were then passed to the FoG Classification Model, which employed machine learning to differentiate between FoG and stops. The features used for this model were carefully selected based on variables that experts focus on when evaluating FoG. By combining these two models—the Forward Progression Identification Model and the FoG Classification Model—we completed our novel FoG Detection Model. Our study demonstrates that even without incorporating multi-modal learning, a standard CNN-LSTM algorithm can achieve high accuracy. This underscores the effectiveness of our preprocessing and data handling techniques. By aligning the model's focus with the critical features identified by medical experts, we significantly enhanced the system's predictive performance.

## II. MATERIALS AND METHODS

### A. Dataset

We collected the video data recorded during clinical practice at the National Center of Neurology and Psychiatry (NCNP), between April 2012 and March 2021. The recorded video data included the following: (1) clinical findings of PD and related disorders, (2) performance in the Timed Up and Go (TUG) test to provoke FoG [15], and (3) occurrence of FoG. In addition, our dataset includes videos that captured FoG episodes, regardless of the medication state (ON or OFF). We focused on identifying the observable FoG episodes captured in the videos. The TUG test is commonly used in the rehabilitation assessment of dynamic balance and involves getting up from a chair, walking at an easy or moderate pace, turning and walking back to the chair, and sitting down [39], [40]. Shine et al. [15] found that the TUG test is a reliable method for provoking FoG in a clinical setting because all the conditions associated with FoG, such as the start of a walk, turning, or moving in front of a target, are included. Videos were taken in daily clinical practice settings, resulting in variations in patient distances and camera angles in each video (Fig. 1). The extracted videos were captured using a



**Fig. 1.** Circumstances under which the video was shot videos taken from various distances and camera angles during the tug test were used. Video **a** was taken from directly beside the tug test; **b** was taken from an oblique angle; **c** was taken from directly in front of the tug test.

commercially-available monocular RGB camera (Panasonic HC-V720M-T) with a resolution of  $720 \times 480$  pixels and a frame rate of 30 fps.

This study was approved by the Institutional Review Board of the University of Tsukuba, Japan (approval number 2023R739) and NCNP, Japan (approval number A2021-062), and conducted according to the guidelines of the Ethics Committee of University of Tsukuba, the NCNP, and the Declaration of Helsinki.

The datasets were divided into training and validation datasets, and test datasets, as follows: videos taken at an angle oblique to the TUG test (Fig. 1b) were used as the training and validation data, and those taken from other angles (Fig. 1a and 1c) were used as the test data. The video data comprised videos captured in the same room but under uncontrolled environment, which implies that camera positions as well as the arrangements of objects (such as computers) and the positions of other people varied in the videos (Supplementary Figure 1). Owing to the lack of consistency in camera positioning, videos not captured directly beside or in front of the TUG test were classified as “obliquely captured videos” (Fig. 1b).

### B. Video Annotation

For annotating videos, we used a free template by Gilat which can be implemented in open-source software to score FoG based on videos [41]. The ELAN (version 6.0) software was used, which is a tool for creating complex annotations on video and audio recordings. The FoG scoring template used in ELAN contains predefined annotations frequently used to score FoG in research and clinical practices. This template and the user guide for measuring the percentage of time spent in FoG during the walking task (%FoG) can be downloaded freely.

The duration of FoG and %FoG were evaluated based on the videos. Using ELAN, the duration of FoG was scored by

three specialists in rehabilitation for movement disorders. The evaluators did not offer any insights regarding the assessment of the video recordings. A total of 109 videos for 16 patients were presented to each evaluator in the same sequence. The three evaluators interpreted the definition of FoG by Nutt et al. [1]—which suggested episodic absence or pronounced reduction in the forward progression of the feet during ambulation despite the intention to walk—and designated FoG scores through discussion. The FoG start time was defined as “the moment when the foot of the participant is suddenly no longer producing an effective step forward and displays FoG-related features, despite the participant’s intention to continue walking.” The end time was defined as “the moment of initial toe-off after FoG when the participant is again able to perform at least two effective alternating steps with both legs showing no FoG-related features” [1], [41], [42]. In addition, the start of FoG during a turn was defined as “the moment of toe-off of the first step that touches down in the area where the turn should be performed, with the foot pointing towards the direction of the turn.” The end during a turn was defined as “the moment of heel-strike of the first step that leaves the area where the turn was performed, with the foot pointing towards the end target of the gait task” [41]. The %FoG was calculated using  $T_{FoG}$  and  $T_v$ , where  $T_{FoG}$  represents the duration of FoG observed, and  $T_v$  denotes the total duration of the TUG test. The following equation (1) was used [41]:

$$\%FOG = \frac{T_{FoG} \times 100}{T_v} \quad (1)$$

Based on the error range reported by Kondo et al. [42], we employed a protocol of consensus discussion among the evaluators. Specifically, a group discussion was initiated if any video, as rated by the three evaluators, exhibited an error in FoG duration as  $>1.6$  s or 5.5% in %FoG. The discussions were aimed at reaching consensus and creating a common interpretation for each video.

### C. Automatic FoG Detection

The two steps in developing the model were preprocessing (pose estimation) and modeling. A flowchart of the automatic FoG detection is shown in Fig. 2.

#### D. Preprocessing (Pose Estimation)

1) *Object Tracking*: In clinical practice settings, video recordings often capture multiple individuals, including medical staff, families of patients, and other accompanying patients, who prevent falls. Therefore, before obtaining skeletal information, extracting the target individual for analysis from the captured video is essential. We used the “LightTrack” model as an object-tracking solution for our method to address this complication (Fig. 3). Object tracking is a technique that involves assigning numbers to bound boxes obtained from object detection to track the target individual. LightTrack is a lightweight and efficient object-tracking model built based on the neural architecture search approach, enabling recent efficient execution with limited resources [43].

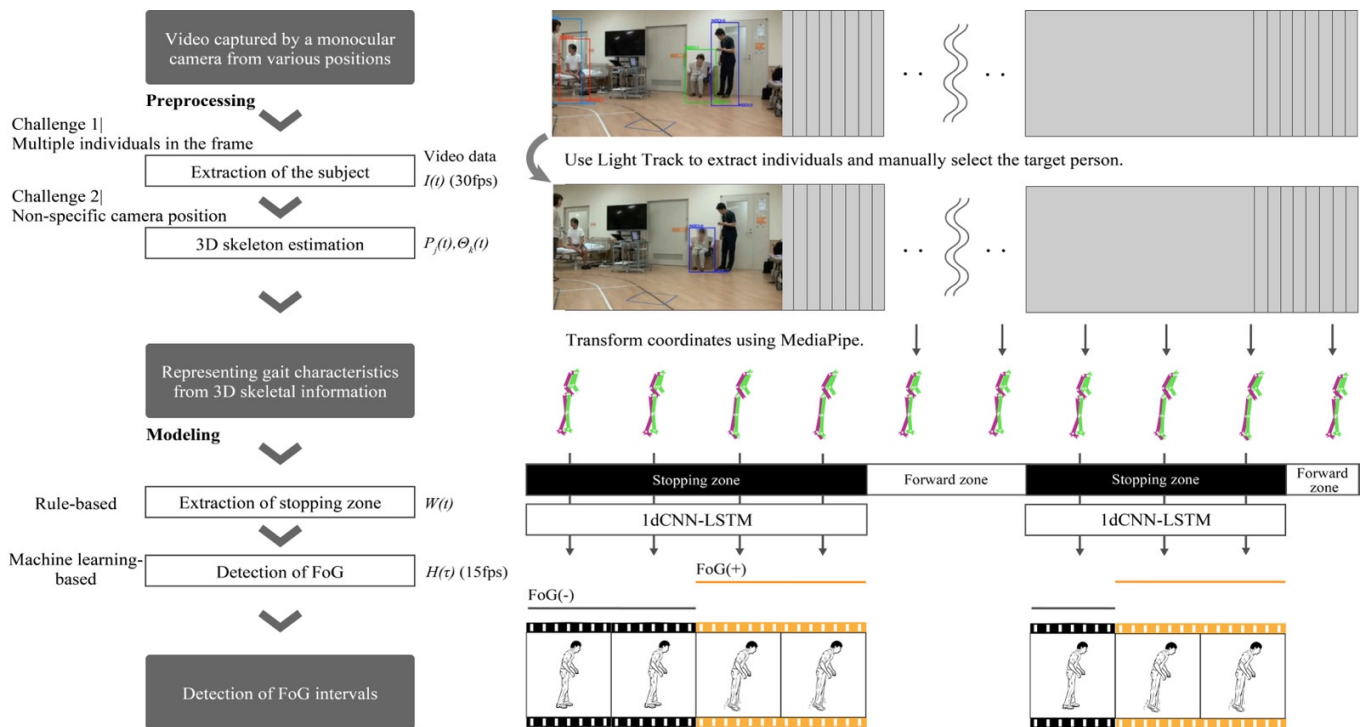


Fig. 2. Summary of the automatic FoG detection method after selecting the target subjects using object tracking, coordinate data were extracted using a 3D pose estimation model. An algorithm was developed to determine the duration of FoG based on the temporal characteristics of the obtained motion features.

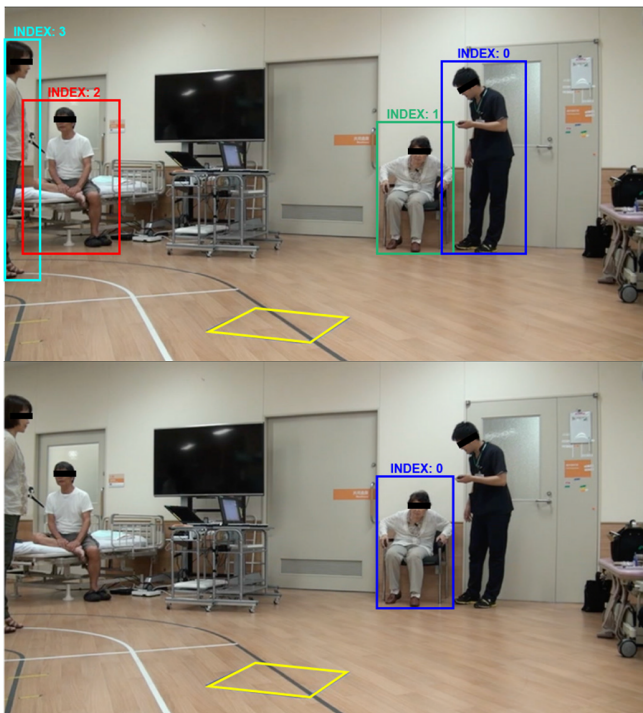


Fig. 3. Examples of videos taken in clinical practice and object tracking. The top image depicts the setting before the bounding box was specified and the bottom image shows the setting after the bounding box was specified.

2) 3D Pose Estimation: In videos with varying camera positions, accurately identifying the pose of the patient using 2D pose estimation is challenging. For example, in videos that

capture a person walking from the background to the foreground, determining whether the person is moving forward is complicated. We believed that this could be solved using a 3D pose-estimation model. In our study, we employed markerless motion capture technology, artificial intelligence-based motion analysis techniques, and the MediaPipe model, which converts 2D video information into 3D skeletal information. This allowed us to overcome the aforementioned challenges and analyze human motion in 3D. MediaPipe, a data stream processing machine learning application development framework developed and open-sourced by Google, is a learning development framework with built-in fast machine learning reasoning and processing, which can achieve end-to-end acceleration on common hardware. Hence, data collection, model training and making predictions (the ‘end-to-end’ part) can be done more quickly (‘acceleration’) even on regular consumer-grade computers or devices (‘common hardware’), without requiring high-end, specialized equipment. MediaPipe has been used successfully for detecting and analyzing characteristic tremors associated with Parkinson’s disease [38]. In this framework, the center of the screen was the origin, and the obtained coordinate data are shown in Fig. 4.

We visually confirmed that both object tracking and 3D pose estimations were performed correctly in our study. However, we also observed that occlusion could potentially reduce the accuracy of pose estimation. To address this issue, we preemptively identified and removed videos having occluded subjects from our analysis.

We initially hypothesized that monocular cameras might have limitations in accurately capturing depth information. However, the proposed method did not rely on precise distance

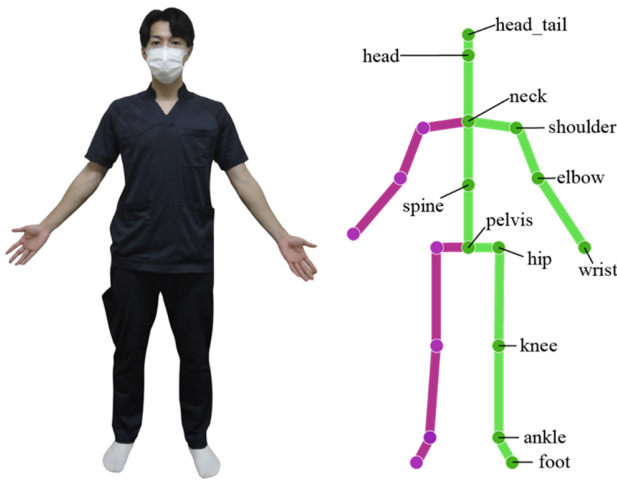


Fig. 4. Coordinates data obtained with MediaPipe.

measurements. We rather focused on determining whether the patient was moving forward, as we believe this could enable the detection of FoG.

### E. Modeling

1) *FoG Detection Model*: The segments identified as walking stops in Model I were extracted. A model for automatic FoG detection was created by subtracting from Model I the segment recognized as a walking stop in Model II that machine learning-based model and not subtracting from Model I the segment recognized as FoG in Model II. Using the strategy of first calculating the displacement of the pelvis and detecting the FoG, we developed the FoG Detection Model. We considered that it could reduce errors by not using multiple machine learning models.

2) *[Model I] Forward Progression Identification Model*: We calculated the displacement from the starting position using (2) for the pelvic position coordinates obtained using the 3D pose estimation. Outliers were removed using a median filter, and the data were smoothed. We defined a walking stop as a segment where the displacement  $d_i$  at the  $i$ -th segment was less than 3% of the total distance traveled in the TUG test within 1 s (30 frames) and developed a model to calculate the duration of walking stops by summing up the number of frames.

$$d_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \quad (2)$$

3) *[Model II] FoG Classification Model*: In the training data, 30 consecutive frames identified as walking stops in Model I were manually classified by author by referring to the ground truth mentioned earlier in video annotation as FoG and walking stop instance. To prevent overfitting, we randomly used three instances of 30-frame segments per video dataset during training and did not train all walking stops. For example, if there were eight instances of a 30-frame FoG within a single video data, it meant selecting any three instances for training data of the FoG instances. Consequently, the training data were divided into 214 FoG instances and 197 non-FoG instances (walking stops).

TABLE I  
LIST OF VARIABLES USED IN TRAINING

Feature group	Number of variables	Variables
Coordinate of lower limb	30	right & left hip xyz (1-6) right & left knee xyz (7-12) right & left ankle xyz (13-18) right & left foot xyz (19-24) right & left heel xyz (25-30)
Lower limb angle	6	right & left hip joint angle (31-32) right & left knee joint angle (32-34) right & left ankle joint angle (35-36)
Left-right difference in lower limb angle	3	hip joint angle (37) knee joint angle (38) ankle joint angle (39)

We trained a binary classification model for time-series data using a combined one-dimensional convolutional neural network (1dCNN) and a long short-term memory (LSTM) architecture. This model is known to be effective for time-series analysis and has shown high accuracy in the analysis of electroencephalographic (EEG) signals [44]. Video data can essentially be interpreted as a type of time-series data. A video is a series of frames (images) that progress over time. Therefore, models that are effective for time-series analysis, such as the 1d-CNN combined with LSTM, can likely be used for analyzing video data. The 1d CNN-LSTM architecture for Model II is shown in Fig. 5. The variables corresponding to the segments identified as walking stops in Model I, as listed in Table I, were used for training in Model II.

We constructed a 1D CNN-LSTM model for our study using different hyperparameters. The input shape for the model included three different input sequences, each consisting of 15 frames captured at a rate of 30 frames per second. We strategically selected 15 frames for the input sequences, equivalent to a total of half a second at a frame rate of 30 frames per second. This helped balance computational efficiency and the capacity of the model to capture temporal patterns associated with FoG events. While shorter video sequences may lack sufficient temporal context for FoG detection, longer sequences pose a risk of diluting crucial FoG information across an extensive number of frames. The CNN layers comprised 32 filters with a kernel size of 7 employing the Swish activation function. Average pooling was applied to compress the temporal dimension by half. The LSTM layer had 32 hidden units with a dropout rate of 0.2 to mitigate overfitting. In addition, dropout regularization was employed to enhance the robustness of the model. The output layer involves using a sigmoid activation function to produce binary classification results. The model was trained for 128 epochs using a batch size of 8. To prevent overfitting and improve the generalization performance, we employed early stopping. We identified the hyperparameters of our model by referencing those used in previous 1dCNN-LSTM models as well as via an iterative process of trial and error.

To assess the generalization performance of our 1dCNN-LSTM model, we employed 5-fold cross-validation [45].

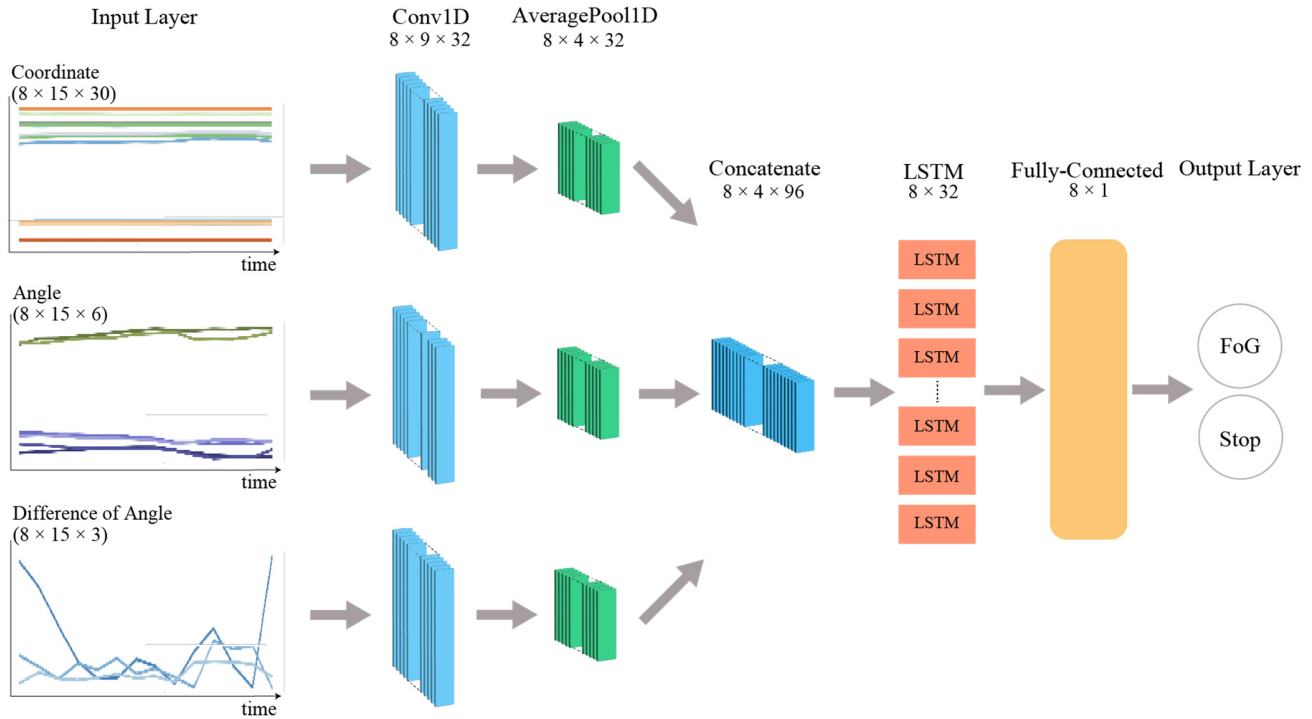


Fig. 5. 1dCNN-LSTM architecture in Model II.

Specifically, we randomly split the entire dataset into five equal-sized folds. We used each fold as the validation set and trained the model on the remaining 5-1 folds. We repeated this process five times, and each fold was used once as the validation set. After each run, the validation accuracy, precision, recall/sensitivity, specificity and F-measure score were obtained.

To determine the optimal threshold for distinguishing between FoG and stop, we applied the trained model to the training data and generated the receiver operating characteristic (ROC) curve and Precision–Recall (PR) curve with thresholds ranging from 0 to 1.0, in increments of 0.1.

#### F. Statistical Analysis

To confirm whether the data obtained in Model I and FoG detection model in each validation and test dataset agreed with the annotation data obtained by the expert, the intraclass correlation coefficients (ICCs) (2, 1) between the data obtained in each model and the expert were calculated. The ICCs  $< 0.5$  showed a poor agreement, those between 0.5 and 0.75 showed a moderate agreement, those between 0.75 and 0.9 showed a good agreement, and those greater than 0.90 showed an excellent agreement [46]. All statistical analyses were performed using Python 3.9.

Bland–Altman plots were used to assess the characteristics of measurement errors, focusing on the limit of agreement (LoA) between the ground truth and the FoG detection model. This analysis aided in better understanding the magnitude of discrepancies between the two sets of values obtained [47]. The 95% LoA was determined as the “average difference  $\pm 1.96 \times$  standard deviation”.

### III. RESULTS

#### A. Details of the Dataset

A total of 109 video data points were collected from 16 patients, nine of whom had PD and seven had progressive supranuclear palsy. There were six men and ten women with a mean  $\pm$  standard deviation age of  $69.6 \pm 8.4$  years and a disease duration of  $12.8 \pm 9.0$  years. Five patients were in stage 3, and eleven were in stage 4 of the disease, according to the Hoehn and Yahr clinical staging. The number of videos per participant varied, ranging from as few as one video to as many as 21 videos.

The 109 videos comprised 74 videos of patients performing a 180-degree turn, and 35 of those performing a 540-degree turn. There were 95 videos as training and validation data and 14 as test data. It should be noted here that among the 14 test data videos, only 2 videos (1 data point each from 2 different patients) included footage of patients who were also part of the training data, albeit filmed from different angles (Supplementary Table I).

#### B. Video Annotation Model II Evaluation of Validation Data

The results of the 5-fold cross-validation are presented in Table II. On average, Model II, which differentiates between FoG and walking stops, demonstrated an accuracy of 93.2% accuracy, 97.9% precision, 88.8% recall/ sensitivity, 97.9% specificity, and 93.1% F-value across the cross-validation results. ROC curve and PR curve of Model II were drawn (Fig. 6). Based on the results, the threshold for distinguishing between FoG and stop was set at 0.5.

TABLE II  
K-FOLD (5-FOLD) CROSS VALIDATION OF MODEL II

	accuracy	precision	recall/ sensitivity	specificity	F-value
1	96.3%	97.6%	95.2%	97.5%	96.4%
2	94.1%	97.5%	90.7%	97.4%	94.0%
3	91.7%	97.4%	86.0%	97.4%	91.4%
4	90.6%	97.3%	83.7%	97.4%	90.0%
5	94.0%	100%	88.1%	100%	93.7%

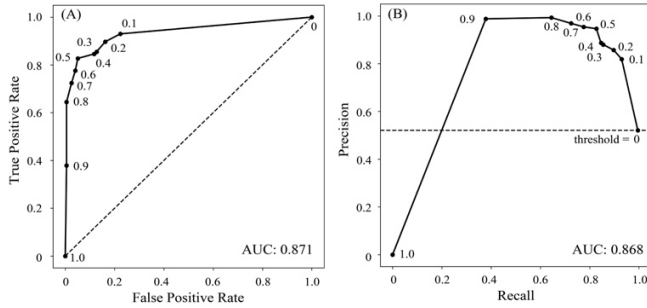


Fig. 6. ROC curve (A) and PR curve (B) of the trained 1dCNN-LSTM model.

### C. Example of Sample

An example of the FoG durations obtained using these three models is shown in Fig. 7. Model I involved extracting the walking stops, and the extracted regions were identified as FoG or walking stops using Model II. The FoG detection model is a complete model used in calculating the duration of FoG by subtracting the parts identified as walking stops in Model II from the total number of walking stops extracted in Model I.

### D. ICCs of Model I and FoG Detection Mode

The durations of FoG, FoG%, and respective ICCs (2,1) for Model I and the FoG detection model for the validation and test data are shown in Table III. The Bland–Altman analyses of the duration of FoG and %FoG are shown in Fig. 8, with separate plots for each dataset (training and test) and individual subjects denoted.

## IV. DISCUSSION

This study aimed to establish a method for automatically detecting and quantifying FoG from videos taken in daily clinical practices. The FoG detection model was constructed by combining the rule-based model and the machine learning-based model, ensuring high reliability of our method.

A key advantage of our method is its capability to utilize videos captured in uncontrolled clinical settings. Unlike current systems that require special equipment, our method allows for the use of commonly available devices like smartphones to capture videos, making it accessible to medical professionals, patients, and their families. While 3D motion analysis offers high accuracy in detecting FoG [48], [49], its applicability is limited outside laboratory settings. Given the challenges of eliciting FoG in laboratory settings [35], [36], [37], we believe that the ability to quantify FoG episodes at home, where FoG mostly occurs, using readily-available equipment and from any camera position, has high potential. For example, therapists or

TABLE III  
GRAND TRUTH AND INTER-RATER RELIABILITY OF EACH MODEL

	duration of FoG (s)	%FoG (%)	duration of FoG	
			ICCs (2,1) [95% CI]	%FoG ICCs (2,1) [95% CI]
validation data (95 video data)				
GRAND TRUTH	16.9±25.5	37.9±22.8	-	-
Model I	18.4±24.0	48.3±16.8	0.87 [0.61-0.94]	0.66 [0.27-0.83]
FoG Detection Model	15.4±23.9	35.1±19.6	0.94 [0.91-0.96]	0.84 [0.77-0.89]
test data (14 video data)				
GRAND TRUTH	11.8±15.6	32.4±20.6	-	-
Model I	13.0±12.6	42.0±12.4	0.83 [0.37-0.95]	0.52 [0.01-0.82]
FoG Detection Model	9.0±11.5	27.9±14.3	0.91 [0.74-0.97]	0.75 [0.59-0.91]

family members can likely record the walking movement of the patient using smartphones or similar devices. In addition, because the recording is done by trusted individuals for a set period of time, rather than continuously by permanently-installed cameras, privacy concerns can be addressed. The method proposed in this study can assist clinicians in accurately detecting the duration of FoG, which is often used as an indicator of treatment effectiveness. The strength of the duration of FoG and %FoG is that it is an objective outcome of the ratio measurement level that directly reflects the severity of FoG at the time of testing, as opposed to subjective scales such as FoG questionnaires [41]. Moreover, automatic FoG detection can improve the efficiency of FoG assessment, reducing the reliance on trained experts, and minimizing subjective judgments of evaluators. Therefore, we believe that our automatic FoG detection method is the first step toward detecting true endpoints of occurring FoG in daily clinical and home environments.

Based on videos, the FoG detection model showed good reliability regarding FoG duration and %FoG in the FoG scoring. This was indicated by the high ICC obtained from the validation (ICC = 0.94 for the duration of FoG and 0.84 for %FoG) and test data (ICC = 0.91 for the duration of FoG and 0.75 for %FoG). ICC values greater than 0.75 are considered good, according to Koo and Li [46]. While Sun et al. [34] also attempted the calculation of FoG duration by counting FoG frame events from videos, the absence of reported error values in their study precludes a direct comparison with our findings. Previous studies employing graph-based neural networks [31], [32], [33] and Transformer neural networks [34] have used raw video data as input; however, by adopting the preprocessing steps established in this study, it may be possible to refine the feature extraction process. Furthermore, features obtained from 2D poses may lack consistency across different angles, potentially affecting the performance of FoG detection.

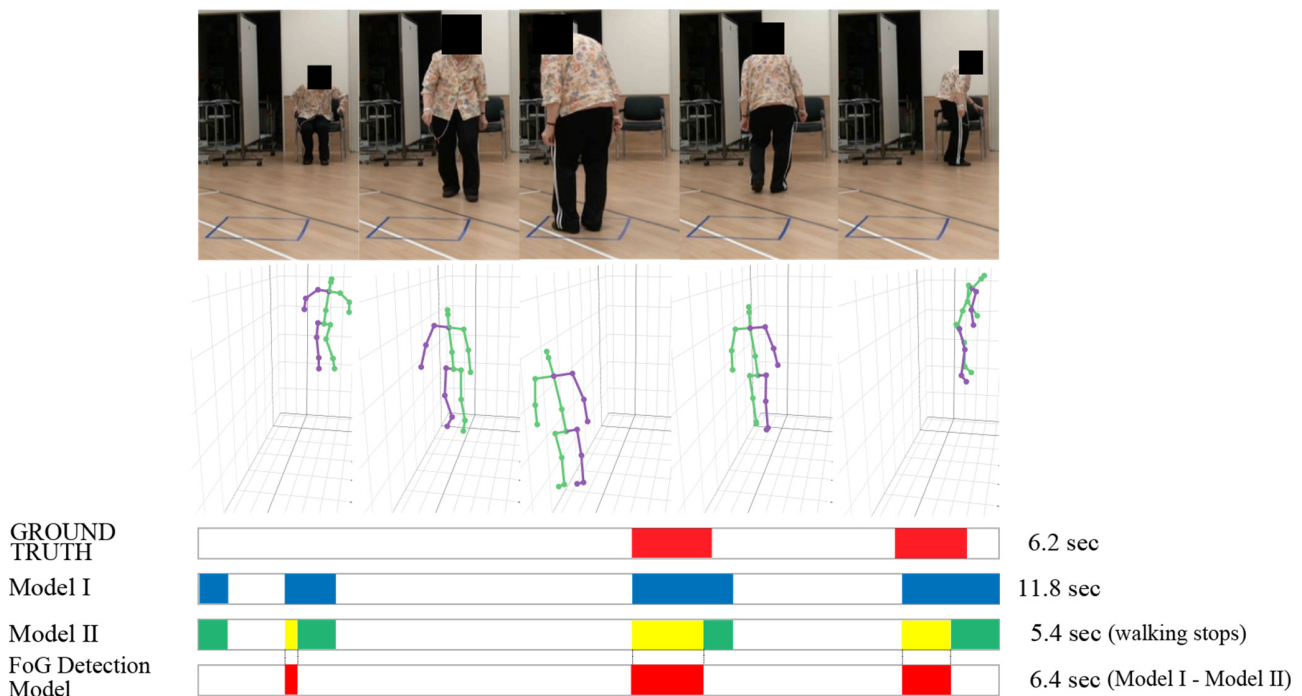


Fig. 7. Example sample The areas identified as walking stops in Model I were indicated by blue bars. The green bar identified as walking stops and the yellow bar identified as FoG by Model II; Model I minus Model II was detected as FoG and was shown by a red bar.

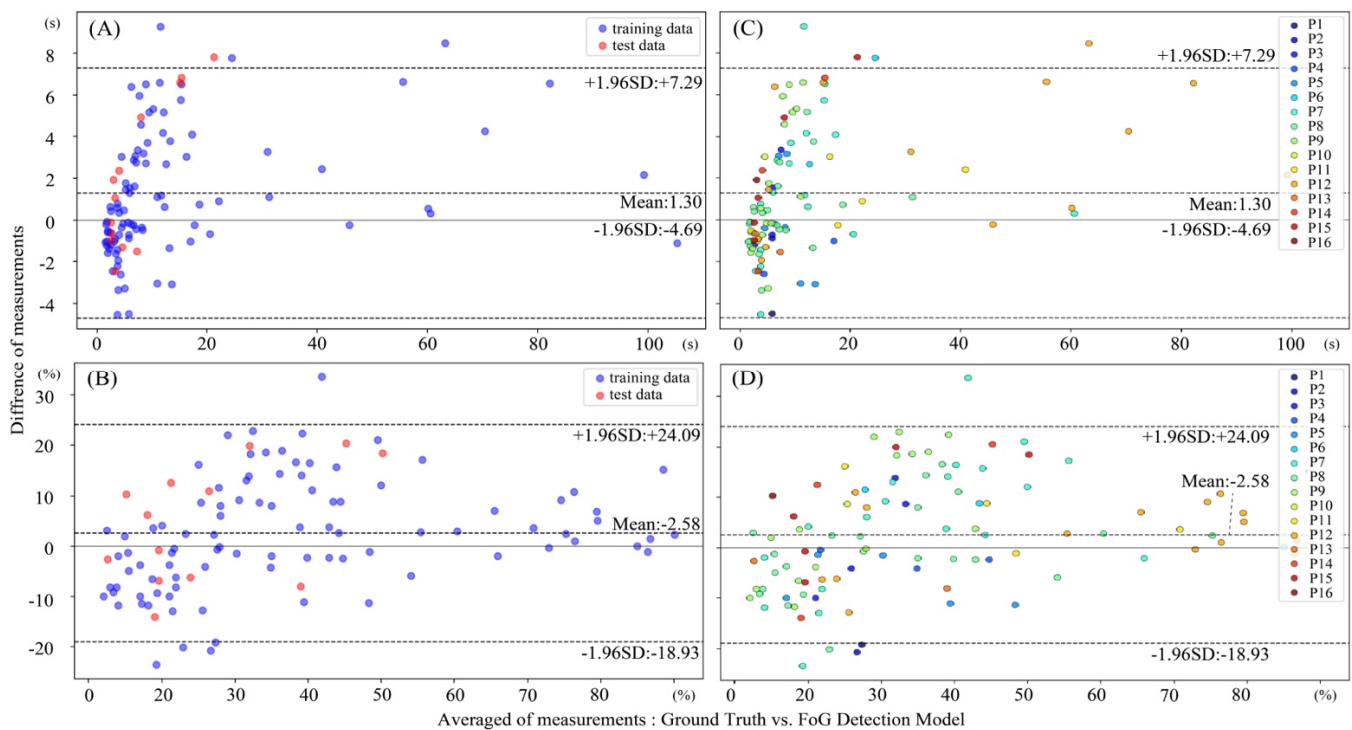


Fig. 8. Bland–Altman plots Panels (A) and (B) show the results for each dataset, with (A) depicting the duration of FoG and (B) illustrating %FoG. Panels (C) and (D) showcase the same analyses for individual subjects, with each subject denoted by a unique color. Panel (C) represents the duration of FoG and panel (D) represents %FoG.

In contrast, our method, which utilizes 3D pose estimation, can extract pose information that is invariant to camera viewpoints, allowing for consistent feature extraction regardless of the camera angle.

The difference in ICC values between the duration of FoG and %FoG must be considered after understanding that they represent different perspectives of the FoG phenomenon and might be influenced by distinct factors. FoG focused on the



absolute time during which FoG occurred, irrespective of the total task duration, essentially providing an isolated view of FoG events. In contrast, %FoG offers a relative measure of FoG, interpreting it as a part of the total task duration. Hence, it can give a more comprehensive overview of the degree by which FoG affects the task performance. In our experimental setup, the total task duration was constant and did not involve any automated identification by the algorithm. Thus, the discrepancy in ICC values between FoG duration and %FoG is unlikely to be a consequence of the variation in task duration.

The Bland–Altman plots for the duration of FoG initially indicate larger errors for medium duration FoG episodes (5 to 30s); however, when FoG was expressed as a percentage of the total task, these biased errors appear to be less pronounced. Furthermore, the minimum/smallest detectable change, which can provide clinicians with useful and easy-to-understand criteria to assess change (improvement or decline) in individual performance, was reported as 6.9 s for the duration of FoG and 18.3% for %FoG [42]. This implies that a potential variability of 6.9 s and 18.3% occurs even in evaluations conducted by experts. Consequently, the differences between the mean values of the ground truth and those of the FoG detection models calculated in this study fall within these ranges, attesting to the accuracy of our model.

The test data comprised videos captured from varying angles and distances that were distinct from the validation data. Our model demonstrated a robust generalization performance unaffected by camera positions, as evidenced by achieving ICC values  $>0.75$  for FoG duration and %FoG. The narrow confidence interval further attests to the consistent outcomes, reinforcing the potential for generalization of our findings. In addition, the lack of detection bias across datasets and subjects, as shown in the Bland–Altman plots, further supports the generalizability of our approach. Though the test data in the present study comprises videos taken in the same room, they differ not only with respect to the camera positions but also the arrangements of objects such as computers and beds, and the number and positions of people captured in the videos. Concerns were initially raised regarding the accuracy of the depth information when using a monocular camera, which could potentially affect the FoG detection accuracy. However, without precise distance information, accurate detection remained achievable by distinguishing forward movements. Thus, the proposed FoG detection model is generalizable and suitable for practical clinical applications.

Model II, which is a model to distinguish whether the portion where there is no forward movement in Model I is FoG or a walking stop, was highly accurate for predictions, having 93.2%, 98.0%, 88.7%, 97.9%, and 93.1% accuracy, precision, recall/sensitivity, specificity, and F-value, respectively. These results indicate that Model II is efficient for accurately predicting positive instances while effectively balancing false and true positives. Previous research employing video-based detection for classifying FoG, such as the studies conducted by Hu et al. [31], [32] and Li et al. [33], the method for classifying FoG reported an accuracy range of 81.6–82.5%. However, the present study achieving a higher accuracy of

93.2% compared to the previous studies. In comparison, the present study reported sensitivity and specificity of 88.7% and 97.9%, respectively, thereby demonstrating the substantial progress made in FoG detection. By employing a relatively simple method for classifying instances as FoG or walking stops rather than differentiating between FoG and normal walking, we achieved a highly accurate Model II. This model could help us discern whether an event is FoG or a stop using 15 frames, which translates to 0.5 s. Hence, FoG episodes lasting less than 0.5 s cannot be detected by our system. However, it is challenging even for experts to determine whether an episode is FoG based on video footage that lasts  $< 0.5$  s.

In addition to the significant advancements that our current FoG detection methodology ensues, we recognize the importance of identifying early signs of FoG within the gait cycle. The ability to detect these preliminary signs will have profound implications for both diagnosis and intervention strategies. Our technology of utilizing video analysis holds the potential to identify these early indicators. This feature could be pivotal in developing real-time monitoring systems, which could potentially alert patients or caregivers to impending FoG episodes, enabling timely interventions such as cueing strategies to prevent the onset of FoG. It is well-recognized that preserving optimal gait can effectively prevent FoG [50]. The prospect of integrating this technology into therapeutic interventions, especially in home settings, is particularly promising. It could offer a proactive approach for managing FoG, enhancing the quality of life for patients. As we move forward, expanding our research to encompass the detection of early signs of FoG will be our key focus, aiming to contribute further to the field of PD management.

This study has some limitations. First, the accuracy of our automatic FoG detection method at home remains unknown. Although the accuracy of different camera positions is guaranteed, system performance may vary in real-world environments with inconsistent conditions. Nevertheless, this study serves as a foundation for future research on the application of video-based FoG detection methods in uncontrolled environments. Second, our study was limited to the TUG test, which implies that the accuracy of applying this system for assessing other gait-related characteristics remains unknown. Therefore, further evaluation of system performance in diverse clinical and home environments is required. Third, it is unknown whether the camera distance to the patient influences accuracy. The videos in this study were captured in a routine clinical setting, and the precise camera distance to the patient was not measured. Fourth, accurate FoG detection is impossible in the presence of occlusion. Occlusions of the patient by other people or objects in the environment, such as furniture and obstacles, may hinder accurate FoG detection. Fifth, the sample size was small. However, our study included a substantial number of data points (109 data points) collected in a clinical setting. Our dataset meets the COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) guidelines, which recommends that a sample of 100 data points adequate [51]. Furthermore, previous studies on FoG detection have been conducted with as few as seven

cases [49]. Considering that our study primarily aims to propose a novel method, this sample size can be considered reasonable within the context of methodological development. Sixth, our system is not currently capable of real-time FoG detection, as the necessity for such immediate identification is relatively limited. However, if real-time detection were achievable, it could be utilized for interventions such as providing external cues upon detecting early signs of FoG. Seventh, our study lacks an explicit evaluation protocol considering subject independence. Regardless, the overlap between subjects in the training and testing sets was limited in our study, and our method could detect FoG with a certain level of accuracy even from diverse video data collected during daily clinical practice, suggesting its potential contribution to advances in the engineering field. Finally, two manual processes were performed for the analysis. The first was the manual trimming of the recorded video in the situation to be analyzed, and the second was selecting the analysis target from the bounding box extracted by object tracking, which requires manual work. These processes are time-consuming and may limit practicing the system in clinical settings.

## V. CONCLUSION

The automatic FoG detection method established in this study exhibited high reliability, with an accuracy comparable with that of expert identification. This is the first system capable of automatically detecting FoG in an uncontrolled environment, regardless of differences in camera positions or the presence of individuals other than the patient in the frame. This automatic FoG detection method can improve the clinical evaluation and management of patients with PD.

## ACKNOWLEDGMENT

The authors would like to thank Editage ([www.editage.com](http://www.editage.com)) for English language editing. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## REFERENCES

- [1] J. G. Nutt, B. R. Bloem, N. Giladi, M. Hallett, F. B. Horak, and A. Nieuwboer, "Freezing of gait: Moving forward on a mysterious clinical phenomenon," *Lancet Neurol.*, vol. 10, no. 8, pp. 734–744, Aug. 2011, doi: [10.1016/S1474-4422\(11\)70143-0](https://doi.org/10.1016/S1474-4422(11)70143-0).
- [2] W.-S. Zhang, C. Gao, Y.-Y. Tan, and S.-D. Chen, "Prevalence of freezing of gait in Parkinson's disease: A systematic review and meta-analysis," *J. Neurol.*, vol. 268, no. 11, pp. 4138–4150, Nov. 2021, doi: [10.1007/s00415-021-10685-5](https://doi.org/10.1007/s00415-021-10685-5).
- [3] B. R. Bloem, J. M. Hausdorff, J. E. Visser, and N. Giladi, "Falls and freezing of gait in Parkinson's disease: A review of two interconnected, episodic phenomena," *Movement Disorders*, vol. 19, no. 8, pp. 871–884, 2004.
- [4] Y. Okuma, A. L. Silva de Lima, J. Fukae, B. R. Bloem, and A. H. Snijders, "A prospective study of falls in relation to freezing of gait and response fluctuations in Parkinson's disease," *Parkinsonism Rel. Disorders*, vol. 46, pp. 30–35, Jan. 2018.
- [5] M. Michałowska, U. Fiszer, A. Krygowska-Wajs, and K. Owczarek, "Falls in Parkinson's disease. Causes and impact on patients' quality of life," *Funct. Neurol.*, vol. 20, no. 4, pp. 163–168, 2005.
- [6] S. Rahman, H. J. Griffin, N. P. Quinn, and M. Jahanshahi, "On the nature of fear of falling in Parkinson's disease," *Behavioural Neurol.*, vol. 24, no. 3, pp. 219–228, 2011, doi: [10.3233/ben-2011-0330](https://doi.org/10.3233/ben-2011-0330).
- [7] N. E. Allen, A. K. Schwarzel, and C. G. Canning, "Recurrent falls in Parkinson's disease: A systematic review," *Parkinson's Disease*, vol. 2013, pp. 1–16, Jan. 2013.
- [8] M. Mancini, B. R. Bloem, F. B. Horak, S. J. G. Lewis, A. Nieuwboer, and J. Nonnekes, "Clinical and methodological challenges for assessing freezing of gait: Future perspectives," *Movement Disorders*, vol. 34, no. 6, pp. 783–790, Jun. 2019, doi: [10.1002/mds.27709](https://doi.org/10.1002/mds.27709).
- [9] C. K. Cui and S. J. G. Lewis, "Future therapeutic strategies for freezing of gait in Parkinson's disease," *Frontiers Human Neurosci.*, vol. 15, Nov. 2021, Art. no. 741918, doi: [10.3389/fnhum.2021.741918](https://doi.org/10.3389/fnhum.2021.741918).
- [10] 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*, vol. 23, no. 15, pp. 2129–2170, Nov. 2008, doi: [10.1002/mds.22340](https://doi.org/10.1002/mds.22340).
- [11] N. Giladi, H. Shabtai, E. S. Simon, S. Biran, J. Tal, and A. D. Korczyn, "Construction of freezing of gait questionnaire for patients with parkinsonism," *Parkinsonism Rel. Disorders*, vol. 6, no. 3, pp. 165–170, Jul. 2000.
- [12] A. Nieuwboer et al., "Reliability of the new freezing of gait questionnaire: Agreement between patients with Parkinson's disease and their carers," *Gait Posture*, vol. 30, no. 4, pp. 459–463, Nov. 2009.
- [13] T. R. Morris et al., "A comparison of clinical and objective measures of freezing of gait in Parkinson's disease," *Parkinsonism Rel. Disorders*, vol. 18, no. 5, pp. 572–577, Jun. 2012.
- [14] T. R. Morris et al., "Clinical assessment of freezing of gait in Parkinson's disease from computer-generated animation," *Gait Posture*, vol. 38, no. 2, pp. 326–329, Jun. 2013.
- [15] J. M. Shine et al., "Assessing the utility of freezing of gait questionnaires in Parkinson's disease," *Parkinsonism Rel. Disorders*, vol. 18, no. 1, pp. 25–29, Jan. 2012.
- [16] A. Delval et al., "Objective detection of subtle freezing of gait episodes in Parkinson's disease," *Movement Disorders*, vol. 25, no. 11, pp. 1684–1693, Aug. 2010, doi: [10.1002/mds.23159](https://doi.org/10.1002/mds.23159).
- [17] M. Plotnik, N. Giladi, Y. Balash, C. Peretz, and J. M. Hausdorff, "Is freezing of gait in Parkinson's disease related to asymmetric motor function?" *Ann. Neurol.*, vol. 57, no. 5, pp. 656–663, May 2005, doi: [10.1002/ana.20452](https://doi.org/10.1002/ana.20452).
- [18] M. B. Popovic, M. Djuric-Jovicic, S. Radovanovic, I. Petrovic, and V. Kostic, "A simple method to assess freezing of gait in Parkinson's disease patients," *Brazilian J. Med. Biol. Res.*, vol. 43, no. 9, pp. 883–889, Sep. 2010, doi: [10.1590/s0100-879x2010007500077](https://doi.org/10.1590/s0100-879x2010007500077).
- [19] S. T. Moore, H. G. MacDougall, and W. G. Ondo, "Ambulatory monitoring of freezing of gait in Parkinson's disease," *J. Neurosci. Methods*, vol. 167, no. 2, pp. 340–348, Jan. 2008, doi: [10.1016/j.jneumeth.2007.08.023](https://doi.org/10.1016/j.jneumeth.2007.08.023).
- [20] S. T. Moore et al., "Autonomous identification of freezing of gait in Parkinson's disease from lower-body segmental accelerometry," *J. NeuroEng. Rehabil.*, vol. 10, no. 1, p. 19, 2013, doi: [10.1186/1743-0003-10-19](https://doi.org/10.1186/1743-0003-10-19).
- [21] E. E. Tripoliti et al., "Automatic detection of freezing of gait events in patients with Parkinson's disease," *Comput. Methods Programs Biomed.*, vol. 110, no. 1, pp. 12–26, Apr. 2013, doi: [10.1016/j.cmpb.2012.10.016](https://doi.org/10.1016/j.cmpb.2012.10.016).
- [22] H. Zach et al., "Identifying freezing of gait in Parkinson's disease during freezing provoking tasks using waist-mounted accelerometry," *Parkinsonism Rel. Disorders*, vol. 21, no. 11, pp. 1362–1366, Nov. 2015, doi: [10.1016/j.parkrel.2015.09.051](https://doi.org/10.1016/j.parkrel.2015.09.051).
- [23] S. Rezvani and T. Lockhart, "Towards real-time detection of freezing of gait using wavelet transform on wireless accelerometer data," *Sensors*, vol. 16, no. 4, p. 475, Apr. 2016, doi: [10.3390/s16040475](https://doi.org/10.3390/s16040475).
- [24] C. Ahlrichs et al., "Detecting freezing of gait with a tri-axial accelerometer in Parkinson's disease patients," *Med. Biol. Eng. Comput.*, vol. 54, no. 1, pp. 223–233, Jan. 2016, doi: [10.1007/s11517-015-1395-3](https://doi.org/10.1007/s11517-015-1395-3).
- [25] T. T. Pham et al., "Freezing of gait detection in Parkinson's disease: A subject-independent detector using anomaly scores," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 11, pp. 2719–2728, Nov. 2017, doi: [10.1109/TBME.2017.2665438](https://doi.org/10.1109/TBME.2017.2665438).
- [26] L. Borzi, I. Mazzetta, A. Zampogna, A. Suppa, G. Olmo, and F. Irrera, "Prediction of freezing of gait in Parkinson's disease using wearables and machine learning," *Sensors*, vol. 21, no. 2, p. 614, Jan. 2021, doi: [10.3390/s21020614](https://doi.org/10.3390/s21020614).
- [27] T. Ashfaq Mostafa, S. Soltaninejad, T. L. McIsaac, and I. Cheng, "A comparative study of time frequency representation techniques for freeze of gait detection and prediction," *Sensors*, vol. 21, no. 19, p. 6446, Sep. 2021, doi: [10.3390/s21196446](https://doi.org/10.3390/s21196446).

- [28] A. M. A. Handojoseno, J. M. Shine, T. N. Nguyen, Y. Tran, S. J. G. Lewis, and H. T. Nguyen, "Analysis and prediction of the freezing of gait using EEG brain dynamics," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 5, pp. 887–896, Sep. 2015, doi: [10.1109/TNSRE.2014.2381254](https://doi.org/10.1109/TNSRE.2014.2381254).
- [29] A. M. A. Handojoseno et al., "Prediction of freezing of gait in patients with Parkinson's disease using EEG signals," *Stud. Health Technol. Inf.*, vol. 246, pp. 124–131, Jan. 2018.
- [30] S. K. Bansal, B. Basumatary, R. Bansal, and A. K. Sahani, "Techniques for the detection and management of freezing of gait in Parkinson's disease—A systematic review and future perspectives," *MethodsX*, vol. 10, Nov. 2023, Art. no. 102106, doi: [10.1016/j.mex.2023.102106](https://doi.org/10.1016/j.mex.2023.102106).
- [31] K. Hu et al., "Graph sequence recurrent neural network for vision-based freezing of gait detection," *IEEE Trans. Image Process.*, vol. 29, pp. 1890–1901, 2020.
- [32] K. Hu et al., "Vision-based freezing of gait detection with anatomic directed graph representation," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 1215–1225, Apr. 2020.
- [33] W. Li et al., "Recognition of freezing of gait in Parkinson's disease based on machine vision," *Frontiers Aging Neurosci.*, vol. 14, Jul. 2022, Art. no. 921081, doi: [10.3389/fnagi.2022.921081](https://doi.org/10.3389/fnagi.2022.921081).
- [34] R. Sun et al., "Higher order polynomial transformer for fine-grained freezing of gait detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Apr. 12, 2023, doi: [10.1109/TNNLS.2023.3264647](https://doi.org/10.1109/TNNLS.2023.3264647).
- [35] A. H. Snijders, M. J. Nijkrake, M. Bakker, M. Munneke, C. Wind, and B. R. Bloem, "Clinimetrics of freezing of gait," *Movement Disorders*, vol. 23, no. S2, pp. S468–S474, Jul. 2008, doi: [10.1002/mds.22144](https://doi.org/10.1002/mds.22144).
- [36] M. Sawada, K. Wada-Isoe, R. Hanajima, and K. Nakashima, "Clinical features of freezing of gait in Parkinson's disease patients," *Brain Behav.*, vol. 9, no. 4, Apr. 2019, Art. no. e01244, doi: [10.1002/brb3.1244](https://doi.org/10.1002/brb3.1244).
- [37] J. Nonnekes, A. H. Snijders, J. G. Nutt, G. Deuschl, N. Giladi, and B. R. Bloem, "Freezing of gait: A practical approach to management," *Lancet Neurol.*, vol. 14, no. 7, pp. 768–778, Jul. 2015.
- [38] G. Shalin, S. Pardoel, E. D. Lemaire, J. Nantel, and J. Kofman, "Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural networks," *J. NeuroEng. Rehabil.*, vol. 18, no. 1, p. 167, Nov. 2021, doi: [10.1186/s12984-021-00958-5](https://doi.org/10.1186/s12984-021-00958-5).
- [39] D. Podsiadlo and S. Richardson, "The timed 'up & go': A test of basic functional mobility for frail elderly persons," *J. Amer. Geriatrics Soc.*, vol. 39, no. 2, pp. 142–148, Feb. 1991, doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x).
- [40] S. Morris, M. E. Morris, and R. Ianssek, "Reliability of measurements obtained with the timed 'Up & Go' test in people with Parkinson disease," *Phys. Therapy*, vol. 81, no. 2, pp. 810–818, Feb. 2001, doi: [10.1093/ptj/81.2.810](https://doi.org/10.1093/ptj/81.2.810).
- [41] M. Gilat, "How to annotate freezing of gait from video: A standardized method using open-source software," *J. Parkinson's Disease*, vol. 9, no. 4, pp. 821–824, Oct. 2019, doi: [10.3233/jpd-191700](https://doi.org/10.3233/jpd-191700).
- [42] Y. Kondo et al., "Measurement accuracy of freezing of gait scoring based on videos," *Frontiers Human Neurosci.*, vol. 16, May 2022, Art. no. 828355, doi: [10.3389/fnhum.2022.828355](https://doi.org/10.3389/fnhum.2022.828355).
- [43] Y. Bin, P. Houwen, W. Kan, W. Dong, F. Jianlong, and L. Huchuan, "LightTrack: Finding lightweight neural networks for object tracking via one-shot architecture search," 2021, *arXiv:2104.14545*.
- [44] G. Xu, T. Ren, Y. Chen, and W. Che, "A one-dimensional CNN-LSTM model for epileptic seizure recognition using EEG signal analysis," *Frontiers Neurosci.*, vol. 14, Dec. 2020, Art. no. 578126, doi: [10.3389/fnins.2020.578126](https://doi.org/10.3389/fnins.2020.578126).
- [45] S. Arlot and M. Lerasle, "Why V=5 is enough in V-fold cross-validation," HAL-Inria, 2014. [Online]. Available: <https://hal.inria.fr/hal-00743931v2>
- [46] 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, Jun. 2016, doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012).
- [47] J. Martin Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986.
- [48] H. Kwon et al., "An explainable spatial-temporal graphical convolutional network to score freezing of gait in parkinsonian patients," *Sensors*, vol. 23, no. 4, p. 1766, Feb. 2023, doi: [10.3390/s23041766](https://doi.org/10.3390/s23041766).
- [49] B. Filtjens, P. Ginis, A. Nieuwboer, P. Slaets, and B. Vanrumste, "Automated freezing of gait assessment with marker-based motion capture and multi-stage spatial-temporal graph convolutional neural networks," *J. NeuroEng. Rehabil.*, vol. 19, no. 1, p. 48, Dec. 2022, doi: [10.1186/s12984-022-01025-3](https://doi.org/10.1186/s12984-022-01025-3).
- [50] P. Ginis, E. Nackaerts, A. Nieuwboer, and E. Heremans, "Cueing for people with Parkinson's disease with freezing of gait: A narrative review of the state-of-the-art and novel perspectives," *Ann. Phys. Rehabil. Med.*, vol. 61, no. 6, pp. 407–413, Nov. 2018, doi: [10.1016/j.rehab.2017.08.002](https://doi.org/10.1016/j.rehab.2017.08.002).
- [51] C. A. C. Prinsen et al., "COSMIN guideline for systematic reviews of patient-reported outcome measures," *Qual. Life Res.*, vol. 27, no. 5, pp. 1147–1157, May 2018, doi: [10.1007/s11136-018-1798-3](https://doi.org/10.1007/s11136-018-1798-3).