IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 September 2023, accepted 19 October 2023, date of publication 23 October 2023, date of current version 2 November 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3327091*

RESEARCH ARTICLE

# A Fall Risk Assessment Model for Community-Dwelling Elderly Individuals Based on Gait Parameters

**KE ZHANG[1], WENBO LIU[1], JINGSHA ZHANG[2,3], ZENGYONG LI[2,3], AND JIZHONG LIU[1]**
[1]Nanchang City Key Laboratory of Integrated Medical and Industrial Technology, Nanchang University, Nanchang 330038, China
[2]Beijing Key Laboratory of Rehabilitation Technical Aids for Old-Age Disability, National Research Center for Rehabilitation Technical Aids, Beijing 100176, China
[3]Key Laboratory of Neuro-Functional Information and Rehabilitation Engineering of the Ministry of Civil Affairs, Beijing 100176, China

Corresponding authors: Jingsha Zhang (zhangjingsha105@163.com), Zengyong Li (lizengyong@mca.gov.cn), and Jizhong Liu (liujizhong@ncu.edu.cn)

This work was supported in part by the Fundamental Research Funds for Central Public Welfare Research Institutes of China under Grant 118009001000160001, in part by the Special Research Project on High Quality Development of Innovation and Entrepreneurship Education of the Chinese Society of Higher Education under Grant 21CXD01, and in part by the Key Research and Development Program of Jiangxi Province under Grant 20202BBGL73057.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Beijing Daxing District Integrated Chinese and Western Medicine Hospital.

**ABSTRACT** Assessing fall risk accurately is vital for the older adult population. However, existing fall risk assessments mainly depend on scales, which are inconvenient, subjective, and imprecise. The aim of this study was to explore a machine learning model based on gait parameters to evaluate the risk of falls in older adults living in the community over a one-year period. A total of 46 elderly subjects were recruited in this study. Information on demographics, disease history, and fall history was collected via questionnaire. Moreover, this study used a gait analysis system based on inertial measurement unit and Azure Kinect to acquire the spatiotemporal parameters of the subjects' gait. Based on the above data, various machine learning models, including k-nearest neighbor, support vector machine, gradient boosting decision tree, and voting classifier, were built to estimate the fall risk level of elderly individuals. K-nearest neighbor performed best among all the models with an accuracy of 0.80 on the individual test set, an F1 score of 0.67, and an area under the receiver operating characteristic curve of 0.83. Gait frequency was found to be the most significant feature associated with fall risk, followed by body mass index and gait cycle variability. The findings suggest that the k-nearest neighbor model can provide a quantitative and objective evaluation of fall risk for older adults living in the community and that the evaluation is more accurate when both gait parameters and disease history are taken into account.

**INDEX TERMS** Fall risk, machine learning, gait analysis, risk assessment.

## I. INTRODUCTION

Falls are one of the leading causes of injury and injury-related mortality among older adults living in a community [1], [2]. Annually, between 28% and 35% of community-dwelling older adults experience falls, and this percentage escalates with age [3]. Falls can entail many negative consequences, including serious injury, decreased mobility, and loss of independence; they can also induce anxiety, depression, fear of falling, and other negative emotions in older adults [4], [5], [6]. As society ages, falls in older adults are likely to exert more pressure on the health care system [7], [8].

Accurate fall risk assessment is essential to prevent the adverse consequences of falls in older adults. There have been numerous research reports on fall risk factors and fall risk assessments, and several fall risk assessment models have been developed. Clinical fall risk assessment for older adults typically comprises questionnaires or functional assessments

The associate editor coordinating the review of this manuscript and approving it for publication was Amjad Mehmood.

of posture and gait, such as the St. Thomas's Risk Assessment Tool, the Morse Fall Scale, and the Timed Up and Go Test [9], [10], [11], [12]. However, these methods have low sensitivity and specificity because falls are caused by multiple and complex factors [13], [14]. Furthermore, these methods are subjective and their applicability is often limited to a specific setting or population [14], [15]. It has been argued that no single fall risk assessment tool is significantly superior and none can accurately determine falls with high efficiency and confidence [16], [17].

Recently, researchers have used machine learning techniques to build fall risk assessment models with the help of big data and artificial intelligence technologies [18]. Mishra et al. [19] predicted 6-month fall risk in 92 older adults using geriatric assessments, GAITRite® measurements, and fall history. They found that the Support Vector Machine (SVM) model had the best performance, with an AUC of 0.80, sensitivity of 0.82, specificity of 0.72, and accuracy of 0.75. Hsu et al. [20] developed an eXtreme gradient boosting model to predict in-hospital falls, analyzing data from 639 participants. They reported that the model achieved an accuracy of 0.72 and an AUC of 0.70. Thapa et al. [21] used electronic health records (EHRs) from 2,785 patients in a proprietary database to predict 3-month fall risk among nursing home residents. The eXtreme gradient boosting model exhibited the highest performance, with an AUC of 0.846, specificity of 0.848, and sensitivity of 0.706. The above study constructed a fall risk assessment model for older adults using different predictors, such as demographic statistical characteristics, fall history, gait characteristics, and other information from EHRs. However, gait parameters were not included as predictors in most fall risk assessment studies for older adults. Only a few studies could obtain gait parameters and input them into the machine learning models. Moreover, some of the studies that included gait parameters only considered gait parameters as a risk factor [22], [23], [24].

Many factors increase the risk of falls in older adults, and the more risk factors present, the higher the fall risk [7]. Clinically, risk factors for falls can be classified as external and internal factors [25], [26]. External factors include slippery floors, uncomfortable shoes, poor lighting, etc. Internal factors are those risk factors associated with age-related changes and health that can affect systems related to balance and mobility. Internal factors include history of falls, gait and balance disorders, visual impairment, medication, chronic health conditions, etc. In addition, vestibular dysfunction is also considered a major intrinsic health risk that can lead to dizziness and imbalance in the elderly [27], [28]. Gait and balance impairment is a strong predictor of falls [29], as aging affects postural control and coordination. These changes are reflected in gait parameters, such as gait speed and stride length, which decrease with age [30], [31].

In recent decades, gait analysis techniques based on wearable sensors and vision-based sensors have been widely explored [32]. These sensors are portable, small, inexpensive, and most importantly, they have no special requirements for the experimental environment and can be applied in communities inhabited by elderly people. The acquisition of gait data in this study was based on an inertial measurement unit (IMU) and Azure Kinect, and we developed signal processing algorithms and feature extraction techniques to obtain a variety of gait characteristics, such as gait speed, stride length, gait cycle, and stance time.

A comprehensive fall risk assessment model should incorporate multiple potential risk factors and identify interactions between these risk factors, which may perform better in predicting fall risk in community-dwelling older adults. For these reasons, this study develops machine learning models to assess the fall risk of community-dwelling older adults by considering multiple risk factors including personal statistical characteristics, disease history, and gait parameters. We constructed several supervised learning algorithms, including k-nearest neighbor (KNN), SVM, and gradient boosting decision tree (GBDT). Their performance was compared to identify elderly fall risk assessment models with high sensitivity and specificity. In addition, the study was further based on feature ablation experiments to explore how changes in input features affect model performance, which is one of the key designs of this study. Finally, this study utilizes the SHapley Additive exPlanations (SHAP) [33] to analyze the model and the specific predictions made by the model.

The rest of this paper is organized as follows: In Section II, we describe the subjects, data collection and preprocessing in detail, as well as introduce the machine learning models and evaluation metrics. Section III reports the results of the models and SHAP analysis. Following that, Section IV discusses and interprets the results, and compares this study with existing literature. Finally, the conclusion summarizes the major findings and contributions of this paper.

## II. MATERIALS AND METHODS

### A. STUDY DESIGN AND PARTICIPANTS

Participants were recruited for this study from among community-dwelling older adults who met the following inclusion criteria: age $\geq$ 60 years; living in the community; able to walk independently; no severe cognitive impairment; and verbal communication skills. When participants were enrolled, they were asked about their demographics, medical history, fall history, and other information through questionnaires. At the same time, we acquired each participant's gait data using a gait analysis system based on an IMU and Azure Kinect. The IMU (AHR626, Jobrey Technology, Wuxi, China) was attached to both ankles of the participants using Velcro, the Azure Kinect was placed at a height of 1 metre above the ground, as shown in FIGURE 1. Considering the limitation of the Azure Kinect detection distance, participants started at 0.8 metres from the device and walked along a straight line at a normal speed to the 5-metre line wearing their usual footwear during the test. When the gait data were collected, all participants waited quietly at the starting

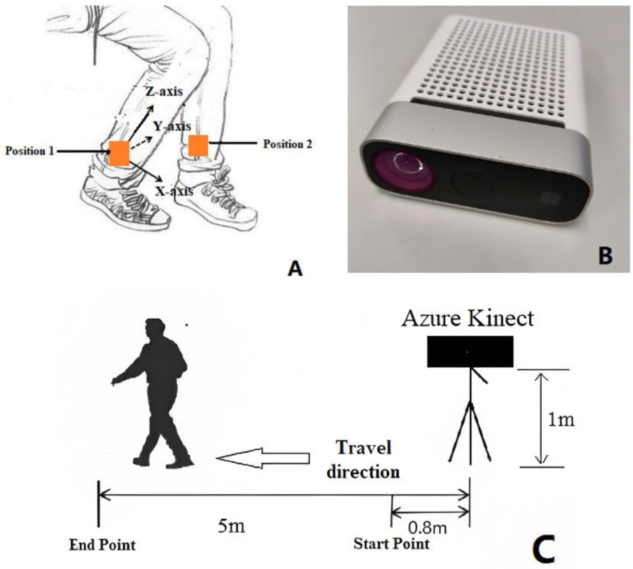


**FIGURE 1.** Gait analysis system. (A) IMU wearing position; (B) Azure Kinect; (C) Gait data acquisition process.

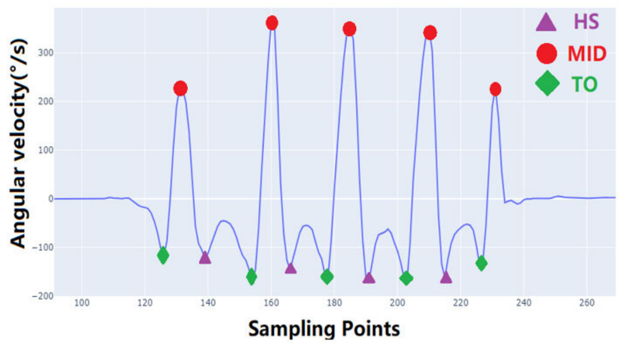


**FIGURE 2.** Gait event detection results. (MID: mid-swing; TO: toe-off; HS: heel-strike.)

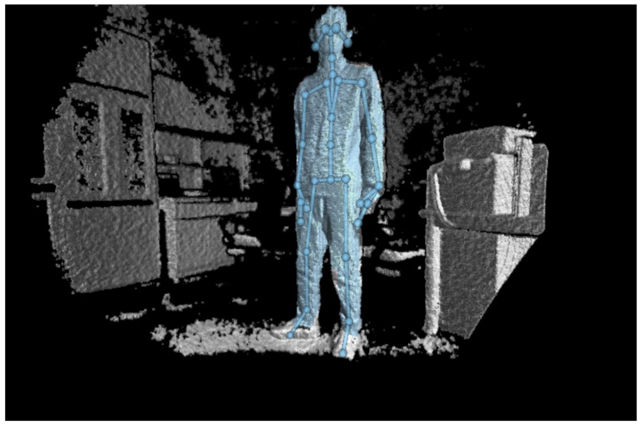

**FIGURE 3.** Azure Kinect Body Tracking Joints. (The skeleton consists of 32 joints, which are shown in spherical form.)

point for 3 seconds while the experimenter started the data collection and gave the participant a ''start'' signal to begin walking at their normal speed. When participants walked to the 5-metre finish line, they waited in place for 3 seconds, and then the experimenter ended the data collection. The IMU recorded the data at a sampling rate of 100 Hz, and the Azure Kinect acquired each participant's 3D skeletal information at 30 Hz.

Participants were divided into two groups based on their self-reported fall history within the year prior to data collection: high risk (at least one fall) and low risk (no falls). All participants signed a written informed consent from prior to the start of the study. All methods adhered to the ethical standards of the Declaration of Helsinki, and approved by the Ethics Committee of Beijing Daxing District Integrated Chinese and Western Medicine Hospital.

### B. EXTRACTION OF SPATIOTEMPORAL GAIT PARAMETERS

Gait spatiotemporal parameters are effective indicators for assessing fall risk. Many studies have been conducted using gait analysis to obtain these parameters, which were used as input features for the machine learning model in this study. In this study, gait data of elderly people were collected based on an IMU and Azure Kinect, and the spatiotemporal parameters of gait were extracted and analyzed. First, to avoid the interference of noise and anomalies in the collected IMU data on the detection of gait critical points, the Y-axis angular velocity values were filtered using a mean filter. Then, a thresholding method was used to perform gait event detection [34], [35], [36]. The gait events that were identified are heel-strike (HS), toe-off (TO), and mid-swing (MID). Mid-swing is the point of maximum Y-axis angular velocity within a gait cycle. HS is the local minimum of Y-axis angular velocity after the swing phase. Y-axis angular velocity reaches a local minimum when the toe leaves the ground, and then a maximum during the swing phase. TO is the local minimum point before the maximum (FIGURE 2). The gait time parameter can be calculated based on the detected critical points. Taking HS as the starting point of a gait cycle, then the gait cycle T is:

$$\mathrm{T}(n) = \mathrm{HS}(n+1) - \mathrm{HS}(n) \tag{1}$$

The swing phase is:

$$\mathrm{SP}(n) = \frac{\mathrm{HS}(n+1) - \mathrm{TO}(n)}{\mathrm{T}(n)} \tag{2}$$

Combining the gait time parameters of the right and left feet, the double support phase can be calculated, as follows.

Initial double support period:

$$\mathrm{IDSP}(n) = \mathrm{TO}_{\mathrm{left}}(n) - \mathrm{HS}_{\mathrm{right}}(n) \tag{3}$$

Final double support period:

$$\mathrm{FDSP}(n) = \mathrm{TO}_{\mathrm{right}}(n+1) - \mathrm{HS}_{\mathrm{left}}(n) \tag{4}$$

Double support phase:

$$\mathrm{DSP}(n) = \frac{\mathrm{IDSP}(n) + \mathrm{FDSP}(n)}{\mathrm{T}(n)} \tag{5}$$

**TABLE 1.** Inclusion characteristics.

| Feature Type | Content |
|---|---|
| Basic personal information | Sex, age, BMI, alcohol consumption, physical activity |
| Disease History | Hypertension, diabetes, dyslipidemia, cardiovascular disease, cerebrovascular disease, neurodegenerative diseases, cataract, osteoarthritis |
| Gait parameters | Gait cycle time, step frequency, swing phase, double support phase, stride length, gait speed, period_cv, stride_cv |

period_cv: gait cycle variability; stride_cv: stride variability.

Step frequency:

$$SF(n) = \frac{60}{T(n)} \times 2 \tag{6}$$

$n$ is the n-th gait cycle.

The Azure Kinect can capture 32 skeletal joint feature points of the human body without markers, and the 3D coordinate parameters of human joints are obtained by constructing a spatial right-angle coordinate system with the Azure Kinect camera location as the origin (FIGURE 3). Based on this, we developed algorithms to obtain the coordinate values of the ankle joints under the Azure Kinect reference coordinate system, from which the gait speed and stride length were calculated as the gait spatiotemporal parameters. The algorithm we developed is based on the Azure Kinect Sensor Software Development Kit (SDK) and the Body Tracking SDK. This algorithm can capture and generate human body tracking results, and obtain the coordinate values of the body tracking joints at each moment based on the tracking results.

The coefficient of variation (CV) of the stride length and cycle time was used to assess the gait variability. Stride variability and gait period variability are defined as stride_cv and period_cv, respectively. CV is the variability of a given gait parameter normalized to its mean value, expressed as a percentage. The analysis excluded the first and last steps of each participant's walk. We provided some gait parameters of partial subjects in the Supplementary Material.

### C. INCLUSION OF FEATURES AND DATA PREPROCESSING

Based on literature and expert knowledge, previous research reports have identified the following risk factors for falls: being female, advanced age, gait and balance impairments, sensory impairments (auditory and visual), cognitive impairments, and chronic health conditions (such as heart disease, diabetes, and arthritis) [7], [14], [25]. This study included three types of characteristics: basic personal information, disease history, and gait parameters, as shown in TABLE 1.

The dataset had an extremely unbalanced ratio of low-risk to high-risk samples, which could affect the performance of directly training the model. Chawla et al. [37] proposed an intelligent oversampling technique for unbalanced datasets called the Synthetic Minority Over-sampling Technique (SMOTE), which can well solve the classification overfitting problem caused by traditional oversampling techniques. In this study, the dataset was processed using the SMOTE algorithm to balance the samples. When inputting experimental data into a machine learning model consistently, inconsistent scaling can be an issue, which can affect the data comparability and introduce bias into the analysis results. For example, the feature "swing phase" has a value range of about 0.3 to 0.4, while the feature "step frequency" has a range of about 80 to 120. Larger values tend to carry greater significance in machine learning models, and inputting these two features without normalization could impact the model's classification performance. To avoid this problem, we used the min-max scaling method. This method normalizes continuous features by scaling their values between 0 and 1. We also used ordinal encoding for discrete features. For example, the feature "sex" was transformed into numerical values "0" and "1" to represent "male" and "female" respectively.

### D. MODEL DEVELOPMENT

In this study, three types of features were used as the machine learning model input features, and the dataset of elderly information collected experimentally was randomly sampled at a ratio of 8:2 to divide it into a training set and a test set. According to the fall history, we categorized the older adult into two groups based on their fall risk level: high risk and low risk. A binary machine learning model was constructed for classification. In this study, we selected various classifiers with different advantages and operation modes to ensure a comprehensive comparative study, including KNN, SVM, GBDT, and a voting classifier integrated the three classifiers.

KNN is based on the distance between different eigenvalues for classification [38], and the basic idea is that a sample will be classified in a category when the closest K samples in the feature space belong to a certain class. It has advantages such as simplicity, ease of understanding, no data assumptions, high accuracy, and insensitivity to outliers. SVM is a powerful and comprehensive machine learning model whose goal is to find the best separating hyperplane that maximizes the margin between instances of two different classes [39]. Using kernel functions, data can be mapped to higher dimensional spaces, so that the SVM can make nonlinear decisions. GBDT is an ensemble learning, boosting algorithm that uses CART decision trees as base classifiers. It trains each decision tree by continuously fitting the residuals based on the results of the previous decision tree. It reduces the residuals along the gradient direction and thus reduces the model error [40]. It can capture the non-linear relationships between features. It can also automatically handle feature selection and missing value problems when using decision trees as base classifiers. The voting classifier combines the predictions of each classifier. It uses the result with the most votes as the prediction category.

In this study, a 5-fold cross-validation grid search was used for the selection of hyperparameters to achieve the best model performance, and the reported results are the

**TABLE 2.** Participant characteristics.

| Variable | Fallers (n=10) | Nonfallers (n=36) | Overall (n=46) |
|---|---|---|---|
| Age, [years], (mean ± SD) | 73.10 ± 6.28 | 70.97 ± 4.74 | 71.43 ± 5.11 |
| Sex, [female], (%) | 9 (90) | 22 (61.11) | 31 (67.39) |
| BMI, (mean ± SD) | 25.90 ± 2.15 | 23.58 ± 2.85 | 24.09 ± 2.86 |
| Gait period, [s], (mean ± SD) | 1.29 ± 0.10 | 1.30 ± 0.19 | 1.30 ± 0.17 |
| Step frequency, [steps/min], (mean ± SD) | 93.90 ± 8.20 | 94.30 ± 13.72 | 94.21 ± 12.64 |
| Swing phase, [%], (mean ± SD) | 0.38 ± 0.03 | 0.38 ± 0.03 | 0.38 ± 0.03 |
| Double support phase, [%], (mean ± SD) | 0.27 ± 0.03 | 0.28 ± 0.05 | 0.28 ± 0.04 |
| Stride, [m], (mean ± SD) | 1.08 ± 0.17 | 1.11 ± 0.17 | 1.10 ± 0.16 |
| Pace, [m/s], (mean ± SD) | 0.69 ± 0.18 | 0.69 ± 0.16 | 0.69 ± 0.16 |
| Stride_cv, [%], (mean ± SD) | 5.15 ± 2.93 | 4.95 ± 2.87 | 5.00 ± 2.85 |
| Period_cv, [%], (mean ± SD) | 4.60 ± 2.27 | 4.51 ± 2.68 | 4.53 ± 2.58 |
| Hypertension, [n], (%) | 5 (50) | 17 (47.22) | 22 (47.83) |
| Diabetes, [n], (%) | 3(30) | 4 (11.11) | 7 (15.22) |
| Dyslipidemia, [n], (%) | 5 (50) | 16 (44.44) | 21 (45.65) |
| Cardiovascular disease, [n], (%) | 3 (30) | 13 (36.11) | 16 (34.78) |
| Neurodegenerative diseases/ Cerebrovascular disease, [n], (%) | 1 (10) | 2 (5.56) | 3 (6.52) |
| Cataract, [n], (%) | 6 (60) | 14 (38.89) | 20 (43.48) |
| Osteoarthritis, [n], (%) | 4 (40) | 11 (30.56) | 15 (32.61) |
| Alcohol consumption, [n], (%) | | | |
| No | 8 (80) | 24 (66.67) | 32 (69.57) |
| Drinking in small quantities | 2 (20) | 11 (30.56) | 13 (28.26) |
| Excessive drinking | 0 | 1 (2.78) | 1 (2.17) |
| Physical activity, [n], (%) | | | |
| Less than 2 hours | 3 (30) | 4 (11.11) | 7 (15.22) |
| 2-4 hours | 1 (10) | 5 (13.89) | 6 (13.04) |
| 4-7 hours | 6 (60) | 16 (44.44) | 22 (47.83) |
| More than 7 hours | 0 | 11 (30.56) | 11 (23.91) |

period_cv: gait cycle variability; stride_cv: stride variability.

average of five cross-validation performance measures. For KNN, the optimized hyperparameters include the number of neighbors (n_neighbors) and the method used to calculate the weights for the neighboring samples (weights). In this study, the n_neighbors parameter was selected in the range of 3,4,5,6,7, and the weights had two choices of ''uniform'' and ''distance''. For SVM, we explored different kernels, including linear, polynomial, and Gaussian RBF. We applied regularization techniques to the polynomial and RBF kernels, and selected the hyperparameters 'C' and 'gamma' within the range of 0.01, 0.1, 1, 10, 20, . . . , 100. For GBDT, the hyperparameters we optimized are ''n_estimators'', ''learning_rate'', ''max_depth'', ''min_samples_split'', and ''min_samples_leaf''. The range of values for n_estimators is 10, 20, . . . , 190, and 200. The range of values for learning_rate is 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, and 1. The values

**TABLE 3.** Performance results of the fall risk assessment model on the test set.

| | Accuracy | Sensitivity | Specificity | F-1 score |
|---|---|---|---|---|
| SVM | 0.60 | 0.33 | 0.71 | 0.33 |
| GBDT | **0.80** | **1.0** | 0.78 | 0.50 |
| KNN | **0.80** | 0.67 | **0.86** | **0.67** |
| Voting classifier | 0.70 | 0.50 | 0.75 | 0.40 |

**TABLE 4.** Fall risk assessment model performance results.

| | Accuracy | Sensitivity | Specificity | F-1 score |
|---|---|---|---|---|
| (a) | | | | |
| SVM | 0.60 | 0.33 | 0.71 | 0.33 |
| GBDT | 0.70 | 0.50 | 0.75 | 0.40 |
| KNN | **0.80** | **0.67** | **0.86** | **0.67** |
| Voting classifier | 0.60 | 0 | 0.67 | 0 |
| (b) | | | | |
| SVM | 0.60 | 0.33 | 0.71 | 0.33 |
| GBDT | 0.70 | 0 | 0.70 | 0 |
| KNN | **0.70** | **0.50** | **0.83** | **0.57** |
| Voting classifier | 0.50 | 0 | 0.63 | 0 |

(a) Performance of each model in the absence of gait parameters.
(b) Performance of each model in the absence of disease history.

for max_depth, min_samples_split, and min_samples_leaf range from 1 to 10. We discovered that the default GBDT performed the best. Finally, we combined the three best-performing models to form a voting classifier to investigate whether it improved the prediction accuracy.

The performance of each model was evaluated based on the accuracy, sensitivity, specificity, F1 score and AUC.

### E. ADDITIONAL ANALYSIS

In this section, we discuss feature ablation experiments that modified the input features of the machine learning model to evaluate their impact on the model performance. Two different combinations of features were tried in this study. First, we removed the gait parameters to explore whether the inclusion of the gait parameter had a positive impact on the performance of the model. Then, we removed the disease history category as a feature category.

This study also used SHAP analysis, an interpretable artificial intelligence technique, to interpret our model. SHAP analysis assessed the importance of features and understood which predictor variables contributed to increasing or decreasing the risk of individual predicted falls. It provided key clinical information for targeted interventions.

The above methods and analyses were implemented and executed based on the Python 3.9 environment and PyCharm Community Edition 2022 software. The main Python libraries used were "sklearn", "imblearn" and "shap".

## III. RESULTS

### A. PARTICIPANTS

In this study, data were collected from 46 community-dwelling older adults who met the inclusion criteria, of whom 10 (21.7%) were at high risk and 36 (78.3%) were at low risk. Nine of the 10 individuals at high risk were female. According to an 8:2 ratio, we assigned 36 participants to the training set and 10 participants to the independent test set. TABLE 2 provides a detailed description of the participant characteristics.

### B. MODEL PERFORMANCE

The performance of each model on the test set is shown in TABLE 3. Both KNN and GBDT achieved an accuracy of 0.80, higher than the 0.70 of SVM and voting classifiers, and both KNN and GBDT were able to accurately identify more than two-thirds of the high-risk individuals. The specificity of KNN was 0.86, and the F1 score was 0.67, both of which were higher than other models. The sensitivity of the DBDT model for high-risk individuals reached 1.0, but the F-1 score was only 0.50. The SVM performed poorly in all aspects, with a sensitivity of only 0.33. In addition, KNN outperformed the constructed voting classifier in all metrics. Overall, the KNN classifier performed best, with a sensitivity of 0.67, specificity of 0.86, and AUC of 0.83 (FIGURE 4(A)). The ROC curves of each model are shown in FIGURE 4(B).

### C. CHARACTERISTIC ABLATION EXPERIMENT

After modifying the model input features, it was found that either the lack of gait parameters or the history of disease reduced the performance of the model. First, the performance of each model decreased at different levels when the gait parameter was missing, and the best performance was still the KNN model. The AUC value of the KNN model without the gait parameter decreased slightly compared to the initial KNN model, which bad an AUC of 0.81. However, the performance of the GBDT and voting classifier, which performed well when all features are included, decreased significantly. The KNN model also performed better in terms of sensitivity and specificity when features such as disease history were missing, with a sensitivity of 0.50, a specificity of 0.83, and an F1 score of 0.57. Overall, we can conclude that the inclusion of gait parameters improves the performance of the model, and that the model performs best when both disease history and gait parameters are included. The specific evaluation indices of each model are shown in TABLE 4, and the ROC curves of each model are shown in FIGURE 4(C) and FIGURE 4(D).

### D. FEATURE IMPORTANCE AND INDIVIDUAL PREDICTION

This study used SHAP analysis to interpret the model and its predictions. FIGURE 5 shows the feature importance interpretation of the KNN model, and the top three features were the step frequency, BMI, and period_cv. The feature importance shows that two of the top three features were gait parameters, indicating that the gait parameters have a good contribution to the predicted values of our model. In addition, the FIGURE 5 also shows how the importance of each feature is distributed; for example, the lower the gait frequency is, the

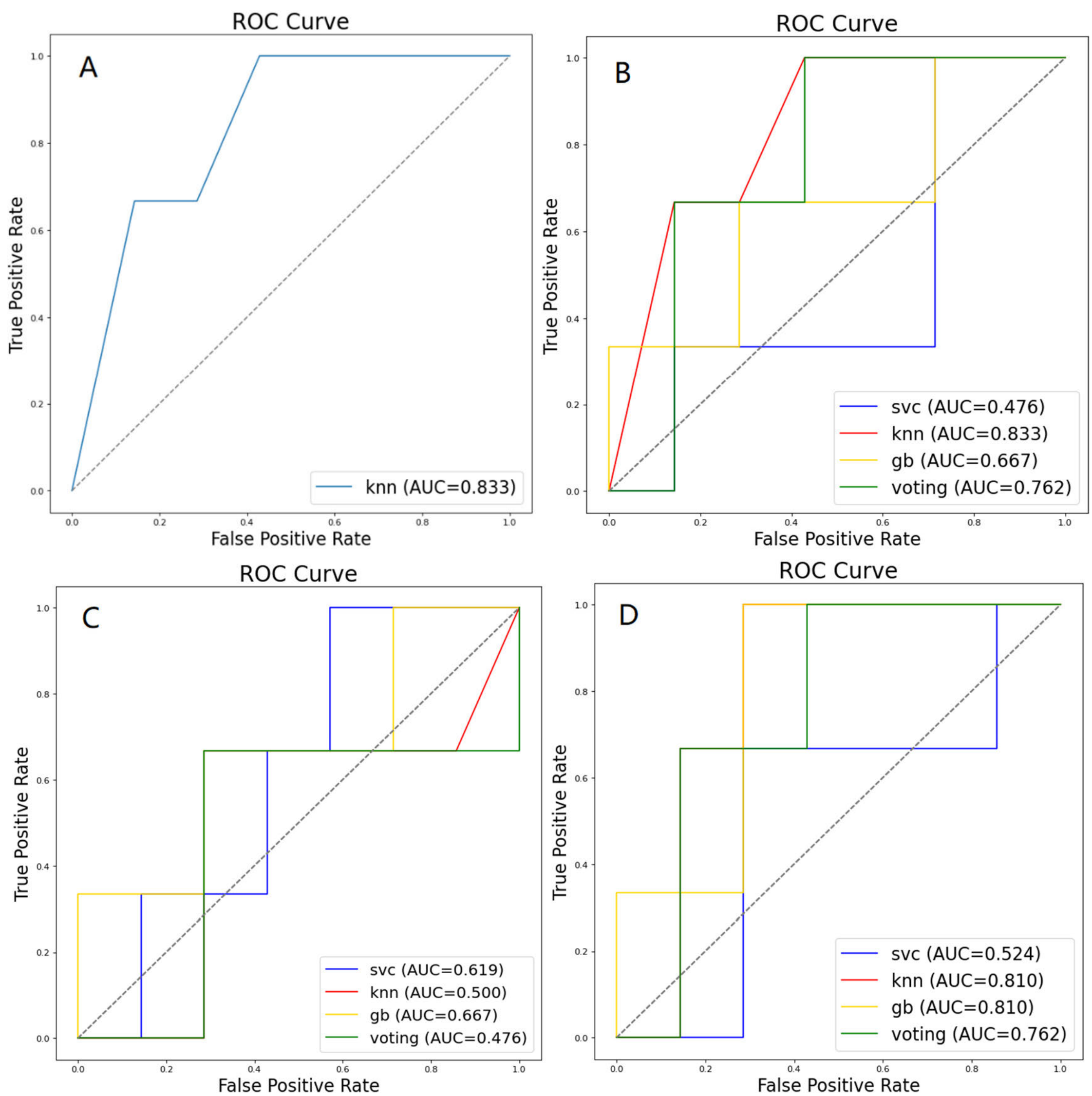


**FIGURE 4.** **(A) KNN receiver operating characteristic (ROC) curves. (B) ROC curves of KNN compared with other machine learning models. (C) Comparison of ROC curves of each model when the disease history feature is missing. (D) Comparison of ROC curves for each model when gait parameter features are missing.**

higher the model SHAP value and the higher the fall risk, and vice versa.

FIGURE 6 provides an individual interpretation of the predicted fall risk for two different older adults. FIGURE 6(A) shows that for Participant 1, having cardiovascular disease and higher age (77 years) increased the risk of falling, and lower BMI (20.66) and higher step frequency (100.10) decreased the risk of falling. FIGURE 6(B) shows that for Participant 2, a lower step frequency (78.55) increased the risk of falling, and a lower BMI (19.29) and the absence of cardiovascular disease reduced the risk of falling.

## IV. DISCUSSION

Fall prevention is an important aspect in the management of group safety in older adults. Studies have shown that the use of appropriate assessment tools to assess the risk of falls among older adults can provide targeted prevention and treatment, which is important for improving their quality of life and reducing the burden of care [41], [42]. Currently, most research is focused on three areas: fall assessment scales, wearable devices, and machine learning [43]. Multiple methods can be applied individually or in combination, but the most common method is still scale assessment. In recent years, combining wearable devices with machine learning to build fall risk assessment models has become a hot topic of research.

There are currently no appropriate and accurate fall risk assessment tools for older adults in a community setting. In this study, we extracted spatiotemporal gait parameters of community-dwelling elderly individuals based on IMU


**FIGURE 5.** KNN model feature importance.

and Azure Kinect. Subsequently, we developed a machine learning model using these gait parameters to assess the risk of falls among older persons within the next year. Of these models, the KNN model had the best performance with an accuracy of 0.80 and an AUC of 0.83. We found that incorporating gait parameters improved model performance when varying the input features of the models. The SHAP analysis identified the top three important features as the step frequency, BMI, and gait cycle variability, and provided additional insight into the results of the model output as to which features increase or decrease an individual's fall risk predictability.

In the models constructed in this study, we found a common problem that each model will have different degrees of overfitting, performing well on the training set but mediocre on the test set. Regardless of how we perform the hyperparameter transformation, the overfitting phenomenon will not be effectively improved. Therefore, we tentatively believe that the model overfitting is caused by the small amount of data. However, the KNN model still performs well, with an accuracy of 0.80, sensitivity of 0.67, and specificity of 0.86 on the test set. The KNN algorithm is related to only a very small number of neighboring samples in category decision-making, rather than relying on the discriminative class domain method to determine the class to which it belongs, which may be one of the reasons for the best performance of the KNN model [44]. Additionally, KNN does not require parameter estimation, which means that this model does not make any assumptions about the data. This means that the structure of the model built using KNN is determined by the data, which is more realistic [45]. Although SVM can handle high-dimensional data and non-linear problems, it has high computational complexity and is sensitive to the selection of parameters and kernel functions. On small datasets, SVM may overfit, resulting in poor generalization ability. GBDT is sensitive to outliers and has a long training time. For SVM and GBDT, when there is an imbalance of samples in the training data, the model may have difficulty learning the correct features and patterns. On small datasets, a voting classifier may be influenced by the base classifiers. If there is a large performance difference between the base classifiers or if there are classifiers with high error rates, the voting results may be inaccurate.

Compared to previous studies of falls in older adults in a community, the KNN model constructed in this study differs in its inclusion characteristics. Lockhart et al. [24] constructed the best classification model using both linear and nonlinear gait features, obtaining $81.6 \pm 0.7\%$ accuracy, $86.7 \pm 0.5\%$ sensitivity, and $80.3 \pm 0.2\%$ specificity in a blind test. The gait features used in the study were extracted from raw gyroscopic and acceleration data, which are difficult to understand compared to the gait spatiotemporal parameters incorporated in this study, reducing the interpretability of the model. Tunca et al. [23] argued that the domain knowledge inherent in the established spatiotemporal gait parameters is still valuable in helping the model to obtain high inference accuracy. The long and short-term memory neural network developed in the study based on a sequence of spatiotemporal gait parameters achieved an excellent classification accuracy of 92.1% on a separate test dataset. The two studies mentioned above only considered one category of gait characteristics and did not consider other factors related to falls, and increasing the incorporated features of the model may improve the model prediction performance.

With the identification of fall risk groups, it is important to identify important predictors from the feature set and take preventive measures. Through SHAP analysis, this study showed that gait frequency was the strongest predictor of fall risk assessment, followed by BMI and gait cycle variability. Hypertension, cataracts, and age also played an important role in fall risk assessment. Gait time-related characteristics play an important role in fall risk assessment, and gait variability refers to the fluctuation of a characteristic value from one step to another [46]. Both indicators of gait variability included in this study ranked high in feature importance, especially period variability. This is in line with the finding of Hausdorff et al. [47] that gait variability can be effective in assessing future falls. Among other gait temporal parameters, reduced gait speed and reduced swing phase share are important manifestations of falls in older persons [48]. However, the results of the analysis in this study suggest that their importance is minimal. In addition, we also noticed that women account for a larger proportion (90%) in the high-risk falling group, and research has shown that women are more prone to obesity [49]. Obesity

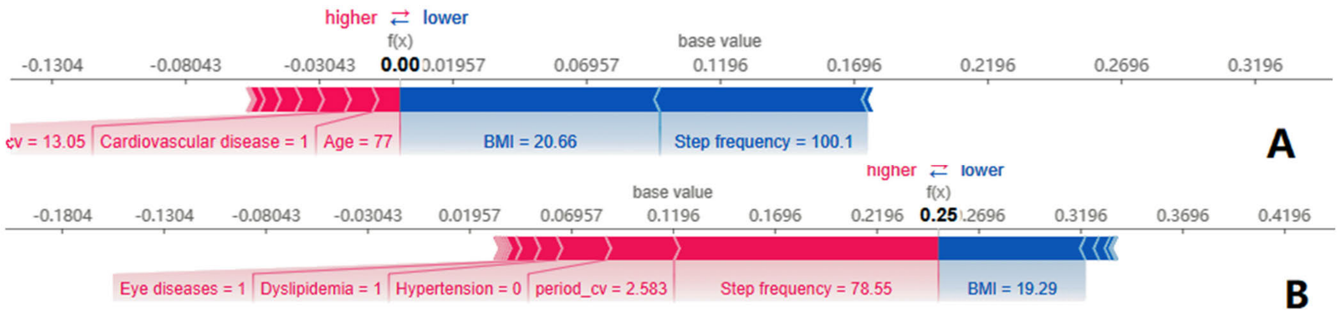


**FIGURE 6.** KNN model individual prediction explanation.

implies a high BMI and disrupted overall load balance, which further affects gait patterns. Moreover, SHAP analysis considers BMI as one of the important features. Therefore, the higher proportion of women in the high-risk group may lead to biased research results and average model performance. In the future, it is necessary to increase the sample size and reduce this influence.

Through feature ablation experiments, the study found that the accuracy of the best KNN model decreased from 0.8 to 0.7, the sensitivity decreased from 0.67 to 0.5, and the specificity decreased from 0.86 to 0.83 when the disease history class was removed from the incorporated features of the model; when the gait parameters class was removed, the accuracy, sensitivity, and specificity of the KNN model did not change, and the AUC decreased from 0.83 to 0.81. However, the performance of other models degraded significantly. These findings suggest that the combination of gait parameters and disease history results in the highest prediction accuracy compared to either feature alone. Moreover, the use of disease history improves the specificity as well as the sensitivity of the model, and the negative association between active disease and fall risk may be due to disease affecting systems related to balance and mobility [25].

There are several limitations in this study. First, machine learning-based classification models are prone to instability with small samples, and a larger dataset with more participants could improve the model generalization. Secondly, the dataset has highly unbalanced categories and a high rate of female falls in the dataset, which may lead to biased results. Third, due to the effective distance limitation of the Azure Kinect, the gait parameters measured in this study may not reflect the most realistic walking ability of the subjects. Moreover, they may change their walking habits due to strain during walking, which may also affect the accuracy of the model. Fourth, we believe dividing participants into fallers and non-fallers is a simple yet effective approach to reflect fall risk and health status among the elderly. However, we also recognize this approach may have some limitations. For example, one fall occurring within a year could be an accidental result and may not necessarily represent increased fall risk; also, there may be different fall patterns among fallers, such as single, recurrent or chronic falls, which may be associated with various risk factors, consequences and intervention needs. The future research directions will focus on developing deep learning models with more powerful algorithms to assess fall risk. It also involves increasing the volume of data and incorporating more risk factors related to falls. Additionally, we can explore the possibility of solely using IMU to obtain gait parameters, which would allow for the collection of data over longer distances. This study could help develop rapid and feasible fall risk assessment tests that can be performed with minimal risk in a community living setting.

## V. CONCLUSION

In this study, a fall risk assessment model for community-dwelling elderly individuals was conducted. We collected the gait data of community-dwelling elderly individuals using a wearable IMU and Azure Kinect. We used a fall risk assessment model based on gait parameters to assess their fall risk. The model feature importance analysis showed that the step frequency, BMI, and gait cycle variability were the most important evaluation metrics. The model performance was optimal when the model incorporated both gait parameters and disease history characteristics. Additionally, the SHAP analysis provided predictive interpretation that helped us suggest targeted interventions for older adults.

## ACKNOWLEDGMENT

The authors would like to thank Gongcheng Xu and Congcong Huo for their valuable comments on the writing of this article.

## REFERENCES

[1] J. E. Crews, C.-F. Chou, J. A. Stevens, and J. B. Saaddine, “Falls among persons aged ≥65 years with and without severe vision impairment—United States, 2014,” *Morbidity Mortality Weekly Rep.*, vol. 65, no. 17, pp. 433–437, May 2016.

[2] D. J. J. van Gulick, S. I. B. Perry, M. van der Leeden, J. G. M. van Beek, C. Lucas, and M. M. Stuiver, “A prediction model for falls in community-dwelling older adults in podiatry practices,” *Gerontology*, vol. 68, no. 11, pp. 1214–1223, 2022.

[3] E. R. Vieira, R. C. Palmer, and P. H. M. Chaves, “Prevention of falls in older people living in the community,” *BMJ*, vol. 353, p. i1419, Apr. 2016.

[4] T. Masud and R. O. Morris, “Epidemiology of falls,” *Age Ageing*, vol. 30, pp. 3–7, Nov. 2001.

[5] B. E. Maki, K. M. Sibley, S. B. Jaglal, M. Bayley, D. Brooks, G. R. Fernie, A. J. Flint, W. Gage, B. A. Liu, W. E. McIlroy, A. Mihailidis, S. D. Perry, M. R. Popovic, J. Pratt, and J. L. Zettel, "Reducing fall risk by improving balance control: Development, evaluation and knowledge-translation of new approaches," *J. Saf. Res.*, vol. 42, no. 6, pp. 473–485, Dec. 2011.
[6] M. Stenhagen, H. Ekström, E. Nordell, and S. Elmståhl, "Accidental falls, health-related quality of life and life satisfaction: A prospective study of the general elderly population," *Arch. Gerontology Geriatrics*, vol. 58, no. 1, pp. 95–100, Jan. 2014.
[7] K. S. F. Khow and R. Visvanathan, "Falls in the aging population," *Clinics Geriatric Med.*, vol. 33, no. 3, pp. 357–368, Aug. 2017.
[8] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, "Medical costs of fatal and nonfatal falls in older adults," *J. Amer. Geriatrics Soc.*, vol. 66, no. 4, pp. 693–698, 2018.
[9] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. A. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 8, pp. 843–851, Aug. 2002.
[10] D. Oliver, M. Britton, P. Seed, F. C. Martin, and A. H. Hopper, "Development and evaluation of evidence based risk assessment tool (STRATIFY) to predict which elderly inpatients will fall: Case-control and cohort studies," *BMJ*, vol. 315, no. 7115, pp. 1049–1053, Oct. 1997.
[11] T. E. Lockhart, R. Soangra, J. Zhang, and X. Wu, "Wavelet based automated postural event detection and activity classification with single IMU (TEMPO)," *Biomed. Sci. Instrum.*, vol. 49, pp. 224–233, Aug. 2013.
[12] 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.
[13] V. Milosevic, A. Linkens, B. Winkens, K. P. G. M. Hurkens, D. Wong, B. P. C. van Oijen, H. M. van der Kuy, and C. Mestres-Gonzalvo, "Fall incidents in nursing home residents: Development of a predictive clinical rule (FINDER)," *BMJ Open*, vol. 11, no. 5, May 2021, Art. no. e042941.
[14] N. Callis, "Falls prevention: Identification of predictive fall risk factors," *Appl. Nursing Res.*, vol. 29, pp. 53–58, Feb. 2016.
[15] R. Parsons, R. D. Blythe, S. M. Cramb, and S. M. McPhail, "Inpatient fall prediction models: A scoping review," *Gerontology*, vol. 69, no. 1, pp. 14–29, 2023.
[16] D. Beck Jepsen, K. Robinson, G. Ogliari, M. Montero-Odasso, N. Kamkar, J. Ryg, E. Freiberger, and T. Masud, "Predicting falls in older adults: An umbrella review of instruments assessing gait, balance, and functional mobility," *BMC Geriatrics*, vol. 22, no. 1, p. 615, Jul. 2022.
[17] S.-H. Park, "Tools for assessing fall risk in the elderly: A systematic review and meta-analysis," *Aging Clin. Experim. Res.*, vol. 30, no. 1, pp. 1–16, Jan. 2018.
[18] A. J. Young, A. Hare, M. Subramanian, J. L. Weaver, E. Kaufman, and C. Sims, "Using machine learning to make predictions in patients who fall," *J. Surgical Res.*, vol. 257, pp. 118–127, Jan. 2021.
[19] A. K. Mishra, M. Skubic, L. A. Despins, M. Popescu, J. Keller, M. Rantz, C. Abbott, M. Enayati, S. Shalini, and S. Miller, "Explainable fall risk prediction in older adults using gait and geriatric assessments," *Frontiers Digit. Health*, vol. 4, May 2022, Art. no. 869812.
[20] Y.-C. Hsu, H.-H. Weng, C.-Y. Kuo, T.-P. Chu, and Y.-H. Tsai, "Prediction of fall events during admission using extreme gradient boosting: A comparative validation study," *Sci. Rep.*, vol. 10, no. 1, p. 16777, Oct. 2020.
[21] R. Thapa, A. Garikipati, S. Shokouhi, M. Hurtado, G. Barnes, J. Hoffman, J. Calvert, L. Katzmann, Q. Mao, and R. Das, "Predicting falls in long-term care facilities: Machine learning study," *JMIR Aging*, vol. 5, no. 2, Apr. 2022, Art. no. e35373.
[22] A. N. Aicha, G. Englebienne, K. van Schooten, M. Pijnappels, and B. Kröse, "Deep learning to predict falls in older adults based on daily-life trunk accelerometry," *Sensors*, vol. 18, no. 5, p. 1654, May 2018.
[23] C. Tunca, G. Salur, and C. Ersoy, "Deep learning for fall risk assessment with inertial sensors: Utilizing domain knowledge in spatio-temporal gait parameters," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 1994–2005, Jul. 2020.
[24] T. E. Lockhart, R. Soangra, H. Yoon, T. Wu, C. W. Frames, R. Weaver, and K. A. Roberto, "Prediction of fall risk among community-dwelling older adults using a wearable system," *Sci. Rep.*, vol. 11, no. 1, p. 20976, Oct. 2021.
[25] C. Enderlin, J. Rooker, S. Ball, D. Hippensteel, J. Alderman, S. J. Fisher, N. McLeskey, and K. Jordan, "Summary of factors contributing to falls in older adults and nursing implications," *Geriatric Nursing*, vol. 36, no. 5, pp. 397–406, Sep. 2015.
[26] P. Pasquetti, L. Apicella, and G. Mangone, "Pathogenesis and treatment of falls in elderly," *Clin. Cases Mineral Bone Metabolism*, vol. 11, no. 3, p. 222, Sep. 2014.
[27] M. B. Liston, D.-E. Bamiou, F. Martin, A. Hopper, N. Koohi, L. Luxon, and M. Pavlou, "Peripheral vestibular dysfunction is prevalent in older adults experiencing multiple non-syncopal falls versus age-matched non-fallers: A pilot study," *Age Ageing*, vol. 43, no. 1, pp. 38–43, Jan. 2014.
[28] S. Fawzan, H. Kozou, F. Baki, and S. Asal, "Fall risk assessment and effect of vestibular rehabilitation in the elderly population," *Egyptian J. Otolaryngol.*, vol. 38, no. 1, p. 88, Dec. 2022.
[29] S. Deandrea, E. Lucenteforte, F. Bravi, R. Foschi, C. La Vecchia, and E. Negri, "Risk factors for falls in community-dwelling older people: 'A systematic review and meta-analysis,'" *Epidemiology*, vol. 21, no. 5, pp. 658–668, Sep. 2010.
[30] O. Beauchet, G. Allali, H. Sekhon, J. Verghese, S. Guilain, J.-P. Steinmetz, R. W. Kressig, J. M. Barden, T. Szturm, C. P. Launay, and S. Grenier, "Guidelines for assessment of gait and reference values for spatiotemporal gait parameters in older adults: The biomathics and Canadian gait consortiums initiative," *Frontiers Hum. Neurosci.*, vol. 11, p. 353, Aug. 2017.
[31] T. Hortobágyi, M. Lesinski, M. Gäbler, J. M. VanSwearingen, D. Malatesta, and U. Granacher, "Effects of three types of exercise interventions on healthy old adults' gait speed: A systematic review and meta-analysis," *Sports Med.*, vol. 45, no. 12, pp. 1627–1643, Dec. 2015.
[32] M. Chen, H. Wang, L. Yu, E. H. K. Yeung, J. Luo, K.-L. Tsui, and Y. Zhao, "A systematic review of wearable sensor-based technologies for fall risk assessment in older adults," *Sensors*, vol. 22, no. 18, p. 6752, Sep. 2022.
[33] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–10.
[34] A. Mannini and A. M. Sabatini, "Gait phase detection and discrimination between walking-jogging activities using hidden Markov models applied to foot motion data from a gyroscope," *Gait Posture*, vol. 36, no. 4, pp. 657–661, Sep. 2012.
[35] A. Salarian, H. Russmann, F. J. G. Vingerhoets, C. Dehollain, Y. Blanc, P. R. Burkhard, and K. Aminian, "Gait assessment in Parkinson's disease: Toward an ambulatory system for long-term monitoring," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 8, pp. 1434–1443, Aug. 2004.
[36] I. P. I. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, "A reliable gait phase detection system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 9, no. 2, pp. 113–125, Jun. 2001.
[37] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.
[38] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, pp. 207–244, Jun. 2009.
[39] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, Apr. 1995.
[40] J. H. Friedman, "Greedy function approximation: A gradient boosting machine.," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.
[41] K. Seaman, K. Ludlow, N. Wabe, L. Dodds, J. Siette, A. Nguyen, M. Jorgensen, S. R. Lord, J. C. T. Close, L. O'Toole, C. Lin, A. Eymael, and J. Westbrook, "The use of predictive fall models for older adults receiving aged care, using routinely collected electronic health record data: A systematic review," *BMC Geriatrics*, vol. 22, no. 1, p. 210, Mar. 2022.
[42] I. D. Cameron, G. R. Murray, L. D. Gillespie, M. C. Robertson, K. D. Hill, R. G. Cumming, and N. Kerse, "Interventions for preventing falls in older people in nursing care facilities and hospitals," *Cochrane Database Systematic Rev.*, no. 1, Jan. 2010, Art. no. CD005465.
[43] G. Zhao, L. Chen, and H. Ning, "Sensor-based fall risk assessment: A survey," *Healthcare*, vol. 9, no. 11, p. 1448, Oct. 2021.
[44] W. Zuo, D. Zhang, and K. Wang, "On kernel difference-weighted $k$-nearest neighbor classification," *Pattern Anal. Appl.*, vol. 11, nos. 3–4, pp. 247–257, Sep. 2008.
[45] Z. Zhang, "Introduction to machine learning: k-nearest neighbors," *Ann. Transl. Med.*, vol. 4, no. 11, p. 218, Jun. 2016.
[46] M. L. Callisaya, L. Blizzard, M. D. Schmidt, J. L. McGinley, and V. K. Srikanth, "Ageing and gait variability—A population-based study of older people," *Age Ageing*, vol. 39, no. 2, pp. 191–197, Mar. 2010.
[47] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg, "Gait variability and fall risk in community-living older adults: A 1-year prospective study," *Arch. Phys. Med. Rehabil.*, vol. 82, no. 8, pp. 1050–1056, Aug. 2001.

[48] M. M. Lusardi, S. Fritz, A. Middleton, L. Allison, M. Wingood, E. Phillips, M. Criss, S. Verma, J. Osborne, and K. K. Chui, "Determining risk of falls in community dwelling older adults: A systematic review and meta-analysis using posttest probability," *J. Geriatric Phys. Therapy*, vol. 40, no. 1, pp. 1–36, 2017.

[49] O. Pickett-Blakely, L. Uwakwe, and F. Rashid, "Obesity in women: The clinical impact on gastrointestinal and reproductive health and disease management," *Gastroenterol. Clinics*, vol. 45, no. 2, pp. 317–331, Jun. 2016.

**KE ZHANG** is currently pursuing the master's degree with the School of Advanced Manufacturing, Nanchang University. His current research interests include machine learning, smart rehabilitation, and wearable sensors.

**WENBO LIU** is currently pursuing the master's degree with Nanchang University. His current research interests include big data and artificial intelligence.

**JINGSHA ZHANG** is currently a Senior Engineer with the National Rehabilitation Assistive Devices Research Center. Her main research interests include intelligent medicine, medical image processing, and biomedical signal processing.

**ZENGYONG LI** received the Ph.D. degree. He is currently a Professor with the National Rehabilitation Assistive Devices Research Center. He has long been engaged in research in the fields of near-infrared brain function assessment, rehabilitation training model, and intelligent rehabilitation technology.

**JIZHONG LIU** is currently a Professor and a Doctoral Supervisor with the School of Advanced Manufacturing, Nanchang University. His current research interests include intelligent electromechanical systems and signal processing and machine learning.