Artificial neural network to classify cognitive impairment using gait and clinical variables

Zhou, Yuhan; van Campen, Jos; Hortobágyi, Tibor; Lamoth, Claudine JC

Published in:
Intelligence-Based Medicine

DOI:
10.1016/j.ibmed.2022.100076

Publication date:
2022

Intelligence-Based Medicine 6 (2022) 100076

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

Artificial neural network to classify cognitive impairment using gait and clinical variables

Yuhan Zhou [a], Jos van Campen [b], Tibor Hortobágyi [a,c,d,e], Claudine JC. Lamoth [a,*]

[a] University of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, 9713, AV, Groningen, the Netherlands
[b] Department of Geriatric Medicine, OLVG Hospital, 1091, AC, Amsterdam, the Netherlands
[c] Hungarian University of Sports Science, Department of Kinesiology, Budapest, Hungary
[d] Department of Sport Biology, Institute of Sport Sciences and Physical Education, University of Pécs, Hungary
[e] Somogy County Kaposi Mór Teaching Hospital, Kaposvár, Hungary

ARTICLE INFO

Keywords:
Neural network
Dynamic gait variables
Clinical variables
Cognitive impairments
Geriatrics

ABSTRACT

Combining gait and clinical variables could increase the accuracy of identifying cognitive impairment (CI) in geriatric patients. We aimed to classify geriatric patients with and without CI based on clinical variables, gait, or a combination of clinical and gait variables, using two machine learning methods, Random Forest (RF) and Artificial Neural Network (ANN). The most accurate classification model examined how interactions between clinical and gait variables would improve classification accuracy and determine the contributions of key variables. Based on Minimal Mental State Examination (MMSE) scores, 131 geriatric patients were divided into a cognitive impaired and a cognitively healthy (CH) group. From 3D accelerometer data collected during 3 min of walking at a habitual speed, we computed 23 dynamic gait variables. In conclusion, an ANN model incorporating the interaction between clinical and gait variables classified geriatric patients with an accuracy of 96%, an area of the receiver operating characteristic curve of 0.95, and a model validation score of 0.97 (F1) based on their clinical status. Machine learning analyses of gait and clinical variables can inform geriatricians about the diagnosis of geriatric patients' cognitive status.

1. Introduction

Physical and cognitive function declines with natural aging. Age-related brain pathologies lead to declines in memory, executive and visuospatial functions, and processing and the prevalence of cognitive impairment (CI) increases with age [1]. Mini-mental state examination (MMSE) [2], the 7-min screen (7MS) [3], and other tests can identify CI. Brain imaging (Computerized Tomography or Magnetic Resonance Imaging scanning) and analyses of cerebrospinal fluid can increase diagnosis accuracy but are costly and invasive.

In addition to cognitive screens, walking ability is also an indicator of current and future health [4]. Indeed, gait impairments predict mortality [5], fall risk [6] and future cognitive decline in geriatric patients [7]. Gait control is mainly mediated by the frontal subcortical circuits, brain structures that also control executive and attentional function [8]. There is thus a strong inter-relationship between gait and cognition [9]. Although gait speed predicts age-related cognition decline [10], gait speed by itself is unlikely to be sensitive enough to detect subtle changes in cognition in the presence of comorbidities [11]. Gait variables other than gait speed comprise information about subtle modifications in gait caused by age-related structural and functional changes in central nervous system [12,13]. Gait is controlled by brain areas that also underlie specific cognitive functions. Models that classify patients with and without CI based on gait could therefore help in recognizing patients in the early stages of CI [7]. Indeed, age-related decreases in gait regularity and predictability were associated with cognitive decline [14]. Variability of step length and stance time also correlated with executive function [15] and a diagnosis of mild cognitive impairment [16]. Therefore, features of gait other than speed, which may lack specificity, could identify CI in geriatric patients.

In addition to gait, questionnaires and functional tests target yet other elements of geriatric disease towards the identification patients with CI. The Drug burden index (DBI) [17], Timed-up-and-Go (TUG) [18], and neuropsychological tests [19], Instrumental Activities of Daily

* Corresponding author. University of Groningen, University Medical Center Groningen, Department of Human Movement Sciences, Antonius Deusinglaan 1, 9713, AV, Groningen, the Netherlands.
E-mail address: c.j.c.lamoth@umcg.nl (C.JC. Lamoth).

https://doi.org/10.1016/j.ibmed.2022.100076
Received 22 April 2021; Received in revised form 8 August 2022; Accepted 14 September 2022
Available online 20 September 2022

Living, and the Geriatric Depression Scale are additional instruments geriatricians use to identify geriatric patients with CI. The accuracy of such assessments is up to 81% [20].

Although gait function as a predictor of CI has been previously examined [21,22], the results are inconsistent and the relationship between specific gait variables and CI is especially poorly understood [22]. Partial least square discriminant analysis (PLS-DA) is a classification method used to analyze datasets with high dimensions and a linear data structure. While PLS-DA classified healthy older adults versus geriatric patients 96% accurately based on dynamic gait variables [13], the classification accuracy of geriatric patients with or without CI using the same dynamic gait variables reached only 38% [13]. This poor classification accuracy may be related to comorbidities and polypharmacy and that PLS-DA is not sensitive to the non-linearities underlying clinical and gait variables [23]. Still, the predominant approach has been the use of clinical and gait variables in isolation rather than in combination to classify geriatric patients with CI. Hence, to understand differences in gait performance between geriatric patients with and without CI, clinical variables should be integrated with data that quantifies the dynamics of gait [24]. To analyze these non-linearly correlated and high dimensional clinical/gait data, an alternative method is needed.

One approach is to use machine learning that can capture the non-linear relationships between clinical and gait variables and cognition. While machine learning classification methods such as Support Vector Machine and Deep Learning have been in use for some time [25], it is difficult to interpret the model outputs because the contribution of each input variable (weights) to the classification remains unknown. Tree-based methods such as decision trees have their own limitations. The reproducibility of the Decision Tree model is highly sensitive: small changes in the data substantially affect the tree structure. Space and time complexity of the decision tree model is relatively higher, leading to longer model training time. A single decision tree is often a weak learner; hence a rich decision tree (known as Random Forest) is required for accurate prediction. Random Forest is a more powerful model as it relies on a single decision tree and creates an ensemble model out of hundreds or thousands of trees to reduce the variance. Thus, ensemble tree-based methods have the advantage over single tree-based methods of having the ability to produce more accurate and stable results. In the present study, we used two relatively interpretable machine learning algorithms for classifying geriatric patients with and without CI, i.e., Random Forest (RF) and Artificial Neural Network (ANN). RF is not sensitive to sample size because of the bootstrap aggregating: every subject can be repeatedly classified in each RF decision tree [26]. However, RF establishes each decision tree independently and disregards the interrelations between decision trees, which can reduce classification performance [27]. Artificial Neural Networks (ANN) do consider the non-linear interrelations between variables by including activation functions such as the hyperbolic tangent but ANN is sensitive to sample size [28]. Although the exact computational process remains hidden in the layers of the ANN, the model can output variable weight coefficients similar to RF. Both ANN and RF can automatically weigh variables and adjust the algorithm to account for their relationships without any prior knowledge to select clinically relevant variables. This automatic adjustment is crucial for clinical data sets where the complicated relationship between variables is unknown.

Therefore, accurate machine learning classification models based on the relationship between dynamic gait variables on the one hand and clinical variables on the other hand, could support clinicians in the classification of geriatric patients with or without CI. The first aim was to compare the classification performance of two machine learning models, i.e., RF and ANN, to classify geriatric patients with and without CI based on: 1) clinical variables, 2) dynamic gait variables, and 3) a combination of clinical and gait variables, the procedures were shown on Fig. 1. The second aim was to identify the variables that contribute most to the classification of geriatric patients with and without CI. As geriatric patients have many comorbidities, we hypothesized that the combination of gait and clinical variables rather than each of these variable sets alone would most accurately classify these two groups of geriatric patients.

2. Methods and models

2.1. Patient characteristics

Accelerometer data, recorded during 3 min of walking, were extracted from an existing database of geriatric patients who visited a diagnostic geriatric day clinic in a teaching hospital in Amsterdam between 2009 and 2018 [13,29,30]. 131 patients' current cognitive status

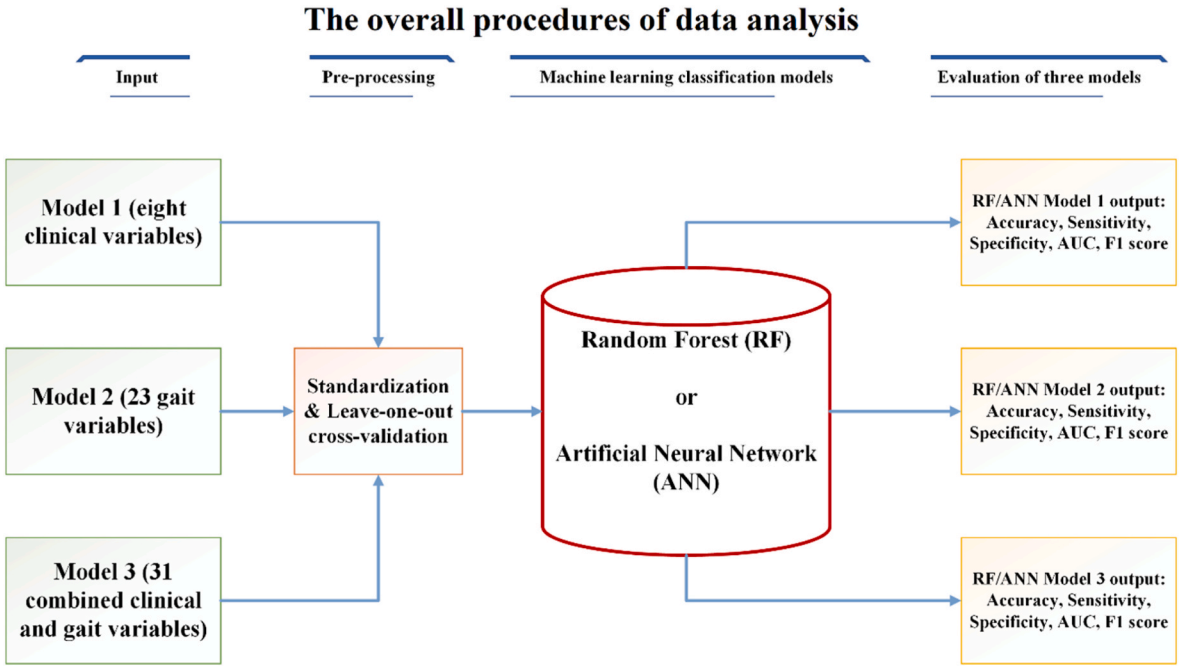

Fig. 1. The overall procedures of data analysis.

was evaluated by the MMSE score; geriatric patients were assigned to a group with the CI (MMSE <24; mean = 19.89 ± 2.66); and a CH group (MMSE ≥24; mean = 26.93 ± 1.88) [31]. The CH group had a mean age of 79.57 ± 5.53 years old and 51/86 were female, and the CI group was on average 80.36 ± 6.52 years old and 28/45 were female. General practitioners referred patients to the day clinic for combined cognitive/somatic complaints. A geriatrician conducted an extensive physical, psychological, and cognitive examination. Exclusion criteria for the study were: (1) inability to walk for 3 min without a walking aid, (2) neurological disorders such as Parkinson's disease and stroke, (3) pain and severe mobility disability caused by orthopedic conditions, and (4) inability to speak and understand the Dutch language. The hospital's Medical Ethics Committee approved the protocol and informed consent documents signed by each patient. The investigation has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. Each patient completed a uniform walking test that included walking in a hallway for 3 min at a self-selected walking speed without a walking aid. Trunk accelerations in 3D were measured by either a stand-alone accelerometer unit, the DynaPort hybrid unit (McRoberts BV, The Hague, the Netherlands) or a built-in tri-axial acceleration sensor, iPod Touch G4 (iOS 6; Apple Inc.) [32].

2.2. Data description

2.2.1. Clinical variables

All patients underwent an extensive clinical assessment including questionnaires and functional tests. The Charlson Comorbidity Index (CCI) was used to measure comorbidity [33]; mood was assessed by the Geriatric depression scale (GDS); and frailty (Frail) was indexed by the Fried's Frailty Scale, including weight loss, exhaustion, low physical activity, slowness and weakness [34]. Functional tests included: Timed-Up-and-Go test (TUG) [35] and maximal grip strength of the dominant hand assessed with a Jamar hand-held dynamometer (3 trials averaged) (HandGrip) [36]. The number of medications was determined from historical medical records (NumMed) and the cumulative exposure to anticholinergic and sedative medications was quantified with the Drug Burden Index (DBI) [17]. Finally, Body mass index (BMI) was determined based on body height and weight. The median and interquartile range of each clinical variable between CH and CI groups are shown in Appendix Fig. A. Table 1 summarizes patients' characteristics and scores obtained for the eight clinical variables in the CH and CI group.

Table 1
Patient characteristics, including the eight clinical variables.

	CH	CI
	Mean ± SD	Mean ± SD
MMSE	26.93 ± 1.88	19.89 ± 2.66
Age, y	79.57 ± 5.53	80.36 ± 6.52
Height, m	1.66 ± 0.09	1.66 ± 0.09
BMI, kg·m^2	26.19 ± 4.29	25.65 ± 4.24
HandGrip, kg	26.17 ± 7.63	25.56 ± 7.78
TUG, s	13.37 ± 4.31	12.69 ± 7.00
CCI	1.58 ± 1.56	1.73 ± 1.18
GDS	4.51 ± 3.51	4.09 ± 2.43
Frail	1.19 ± 1.12	1.11 ± 1.25
DBI	0.58 ± 0.71	0.55 ± 0.75
NumMed	5.69 ± 3.80	4.69 ± 3.48

In terms of abbreviations in the table, CH, cognitive health; CI, cognitive impairment; SD, standard deviation; MMSE, Mini-mental state examination; BMI, Body mass index; HandGrip, grip strength of hand; TUG, Timed-Up-and-Go test; CCI, Charlson Comorbidity Index; GDS, Geriatric depression scale; Frail, frailty criteria; DBI, Drug Burden Index; NumMed, the number of medications used.

2.2.2. Dynamic gait variables

Dynamic gait variables were calculated by using custom-made software in MATLAB (version 2014b; The MathWorks Inc.) from anterior-posterior (AP), medio-lateral (ML) and vertical (V) trunk accelerations. Twenty-three dynamic gait variables represent how gait evolves over time; these gait variables were calculated related to pace, predictability, regularity, symmetry, variability, stability, synchronization and smoothness (for details of the variable calculation, see Refs. [13, 37]). Gait speed (Gaitspeed) was calculated by dividing walking distance (m) by time (s). The Root Mean Square (RMS) acceleration is a measure for the variability of the amplitude of accelerations. Multiscale Entropy (MsEn) quantifies the predictability at different time scales, reflecting signal complexity with 0 denoting a wholly predictable and non-complex signal. Gait step or stride regularity (Step/StriReg) was calculated by the unbiased auto-correlation function of the acceleration signal in AP and V directions. The signal was phase shifted with a window approximating average step and stride time. Perfectly regular steps or strides are represented by a value of 1. The difference between step and stride regularity reflects gait symmetry (Symm), a value of 0 representing a perfectly symmetrical gait. Frequency variability (FreqVar) reflects the relative fluctuations in step frequency. The maximal Lyapunov exponent (mLyap) was calculated by the Wolff algorithm to represent the local stability of trunk acceleration patterns. The higher values show greater sensitivity to local perturbations. The Cross-sample Entropy (CrEn) quantifies the degree of synchronization between AP and ML, AP and V, and ML and V accelerations. A value of 0 reflects perfect synchronization between acceleration signals. The Index of Harmonicity (IH) represents the gait smoothness. IH values are ranged from 0 to 1, and a value of 1 reflects a perfectly smooth gait. The median and interquartile range of each gait variable between CH and CI groups are shown in Appendix Fig. A.

2.3. Machine learning approaches for classification

2.3.1. Random Forest (RF)

The RF model builds various decision trees and merges them based on standardized gait variables and clinical variables, to obtain the optimal classification performance and provide the weight coefficient of each variable of the classification model. The majority of voting was used in RF to make a decision. From the training set $\{(x_i, y_i)\}_{i=1}^{n}$ (x_i represents the training data and y_i represents its label, n represents the number of samples in the training set), a set of m decision trees were built with individual weight functions W_j with the individual clinical parameter or gait outcome as each tree leaf j, the predicted label of CH or CI group is $\hat{y}$ (0 or 1) of the new testing set x' with gait and/or clinical variables [27]:

$$\begin{aligned}\hat{y} &= \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} W_j(x_i, x')y_i \\ &= \sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m} W_j(x_i, x')\right)y_i\end{aligned} \tag{1}$$

Leave-one-out cross-validation (LOOCV) splits the dataset into a training set (n =238) and a testing set. The number of trees m = 128 are optimal for the RF classification.

2.3.2. Artificial Neural Network (ANN)

The ANN model consists of input of the standardized clinical and gait variables, the output of predicted CH and CI patients. The one hidden layer with five neurons computes each variable and their interrelations by the activation function "Rectified Linear Unit (ReLU). This activation function is proper for non-linear gait data structures [28], weight initialization scheme is "He Initialization" [38].

For the computational process in ANN, a neuron j (j runs over from 1 to 8 in model 1, from 1 to 23 in model 2, from 1 to 31 in model 3) re-

ceives an input of clinical variables and dynamic gait variables as $p_j(t)$. The activation function f computes the new activation variable at the next iteration time unit $t+1$ from an activation $a_j(t)$ that represents a neuron's state. θ_j is a fixed parameter for a neuron in the model. The new activation variable is computed as:

$$a_j(t+1) = f\left(a_j(t), p_j(t), \theta_j\right) \tag{2}$$

and then the function f_{out} outputs the prediction of CH or CI group for each subject by the activation function:

$$O_j(t) = f_{out}\left(a_j(t)\right) \tag{3}$$

The classification model is based on LOOCV, $k-1$ sets of training data and one set of testing data. This process was repeated k times (k = the total number of subjects in this dataset, 131). Note that the testing dataset was unique and different from the training dataset.

To examine the classification based on clinical variables and/or gait variables, three different models were computed: Model 1 included only clinical variables (N =8), Model 2 included only gait dynamic variables (N =23), Model 3 included both clinical and gait variables (N =31).

In general, the weight coefficient of ANN was calculated by the Stochastic Gradient Descent method [39], and adjusted by activation function ReLU, to sum the weighted inputs from each incoming synapse (connection) and pass the result to all neurons (variables) in the next layer. This process is repeated iteratively until the error derivatives drop below an acceptable threshold. The ANN weight coefficient in the present study was ranged from 0 to 100.

Depending on the present study results, two conditions were used to determine if a variable contributed significantly to the classification. First, the weight of the variable had exceeded one half of the upper limit (100). Second, these significant variables must have much higher weights than the rest of the variables.

2.4. Evaluation of classification

The accuracy, sensitivity, and specificity were calculated based on the confusion matrix to evaluate the three models' performances to identify CI patients. The results were averaged from LOOCV. Because the number of patients in CH and CI group was unbalanced, the receiver operating characteristic (ROC), the area under the ROC curve (AUC) and F1 score (see equation (6)) provides an overall validation and evaluation of the classification. To compute the F1 score, precision and recall should be calculated in advance based on a classification confusion matrix. In equation (4), precision is the number of correct positive results divided by the number of all positive results returned by the classifier. In equation (5), recall is the number of correct positive results divided by the number of all samples that should have been identified as positive. The baseline of AUC and F1 score is 0.5, and the perfect machine learning classification model has the AUC/F1 score = 1.

$$precision = \frac{true\ positive}{true\ positive + false\ positive} \tag{4}$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \tag{5}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{6}$$

2.5. Procedures of the study

The overall data analysis is illustrated in the flow chart in Fig. 1.

3. Results

3.1. Classification results of RF

In the RF classification, model 1 based on the eight clinical variables (Table 1) obtained a classification accuracy of 64%, with the high sensitivity of 85% and the low specificity of 24%. Also, RF model 2 with 23 dynamic gait variables and model 3 which combined clinical variables and dynamic gait variables (31 variables) had insufficient classification accuracy. Model 2 and 3 obtained a moderate accuracy of 60% and 63%, respectively, while the sensitivity was high of 90% and 92% but specificity was very low of 2% and 9%.

The high sensitivity and low specificity in three RF models imply that over 85% of CH patients were correctly classified as CH but less than 10% of CI patients were successfully classified as being differed from CH patients (Table 2). The classification performances of the three models were evaluated and validated by AUC (the corresponding ROC curves were shown in Fig. 2d) and F1 score and obtained the values near the baseline of 0.5. Because the classification performances in all three RF models were poor, we did not further consider this model.

3.2. Classification results of ANN

The ANN model 1 with the eight clinical variables produced classification accuracy, sensitivity, and specificity of 79%, 84%, and 71%, respectively. This means that 72 out of 86 CH patients were classified correctly to the CH group, while 32 out of 45 CI patients were classified as being different from CH patients (Fig. 2a, Table 3). ANN model 2 obtained an accuracy of 91%, with the sensitivity of 98%, and the specificity of 78%. Thus, 82 out of 84 CH patients were classified to CH group, while 35 out of 45 CI patients were discriminated from CH patients (Fig. 2b, Table 3). Classification of CH and CI patients by ANN models 1 and 2 was AUC = 0.77 and AUC = 0.87, respectively. The F1 score to measure the test's accuracy of model 1 was 0.84 and model 2 was 0.93.

Model 3 is based on the integrated eight clinical variables and the 23 dynamic gait variables; the classification accuracy was increased to 96%, with the higher sensitivity of 99% and the higher specificity of 91% than these values in model 1 and model 2 (Fig. 2c and Table 3). Model 3 obtained an AUC of 0.95 and an F1 score of 0.97 (Table 3). The ROC curves for models 1, 2 and 3, are shown in Fig. 2d.

3.3. The weighted variables of clinical and gait variables in the ANN models

The contributions of variables to the ANN models are quantified by the weight coefficients of ANN (see Fig. 3).

For model 1 (clinical variables), the CI vs. CH group was characterised by a weaker handgrip, lower TUG and fewer frailty criteria (Fig. 3a and Appendix Fig. A, Table 1). For the CI group in model 2 with the dynamic gait variables, the synchronization (CrEn) in ML-V direction, the smoothness of gait (IH) in AP and V directions, the pace (RMS) in AP, ML and V directions and the symmetry (Symm) in AP direction

Table 2
Results of Random Forest (RF) classification for geriatric patients with and without cognitive impairment (CI).

	Model 1	Model 2	Model 3
Accuracy (%)	64	60	63
Sensitivity (%)	85	90	92
Specificity (%)	24	2	9
AUC	0.56	0.50	0.45
F1 score	0.61	0.44	0.55

Model 1 with eight clinical variables, model 2 with 23 dynamic gait variables, and model 3 with the aggregated dataset of eight clinical and 23 dynamic gait variables. AUC, the area under the receiver operating characteristic curve.

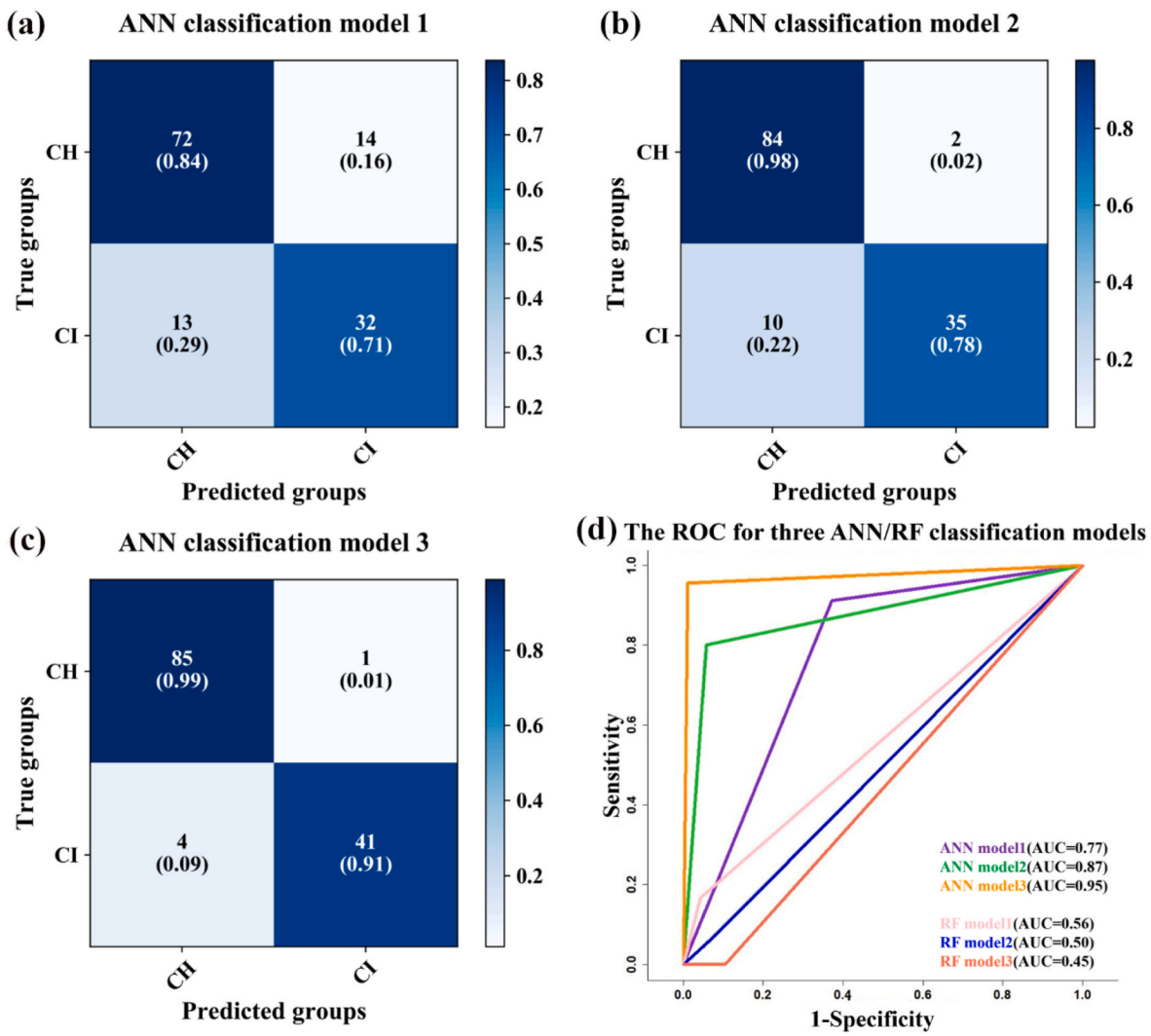

Fig. 2. Confusion matrix of ANN classification and ROC curve for ANN/RF. Panels **(a), (b)** and **(c)** show the classification confusion matrix. The x-axis represents the patients in the predicted groups and the y-axis shows the patients in the original groups. The dark blue means more patients were assigned to this group. The numbers of patients and their percentages in the original group are shown in the squares and braces. Figure **(d)** shows the ROC curves for ANN and RF classification, based on ANN model 1 (purple) with clinical variables, ANN model 2 (green) with dynamic gait variables, and ANN model 3 (yellow) with both clinical variables and dynamic gait variables. Pink, blue, and red lines, respectively denote the RF models 1 to 3. For abbreviations, CH, cognitive health; CI, cognitive impairment; ANN, Artificial Neural Network; RF, Random Forest; ROC, the receiver operating characteristic curve; AUC, area under the ROC curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Results of ANN classification of cognitive healthy (CH) and cognitive impaired (CI) patients.

	Model 1	Model 2	Model 3
Accuracy (%)	79	91	96
Sensitivity (%)	84	98	99
Specificity (%)	71	78	91
AUC	0.77	0.87	0.95
F1 score	0.84	0.93	0.97

Model 1 with eight clinical variables, model 2 with 23 dynamic gait variables, and model 3 with the aggregated dataset of eight clinical and 23 dynamic gait variables. AUC, the area under the receiver operating characteristic curve.

obtained higher weights than other gait variables. These variables contributed more to the classification of CH and CI patients (Fig. 3b and Appendix Fig. A).

In model 3 in Fig. 3c and Appendix Fig. A, and Table 1, the CI group versus CH group had a lower BMI, higher CCI, lower TUG, decreased gait pace, less predictability and stability in ML direction, higher frequency variability (FreqVar) and less step regularity in AP direction. These variables had the most considerable contribution to the classification model that discriminated CI patients from CH patients.

4. Discussion

The data from the present study supported the hypothesis that a combination of gait and clinical variables as an input to ANN machine learning models accurately classified geriatric patients with and without CI. We discuss the clinical and computational relevance of these data.

From the three RF/ANN classification models based on clinical variables (model 1), dynamic gait variables (model 2), and the 31 clinical and gait variables combined (model 3), ANN model 3 had the most accurate classification performance. Compared with cognitively intact patients, those with CI had more comorbidities, poorer mobility, a less predictable and stable and more variable gait (Fig. 3, Appendix Fig. A). The combination of clinical dynamic gait variables in the model produced high classification power (high weight in ANN model 3). The input variables to the machine learning models included eight clinical

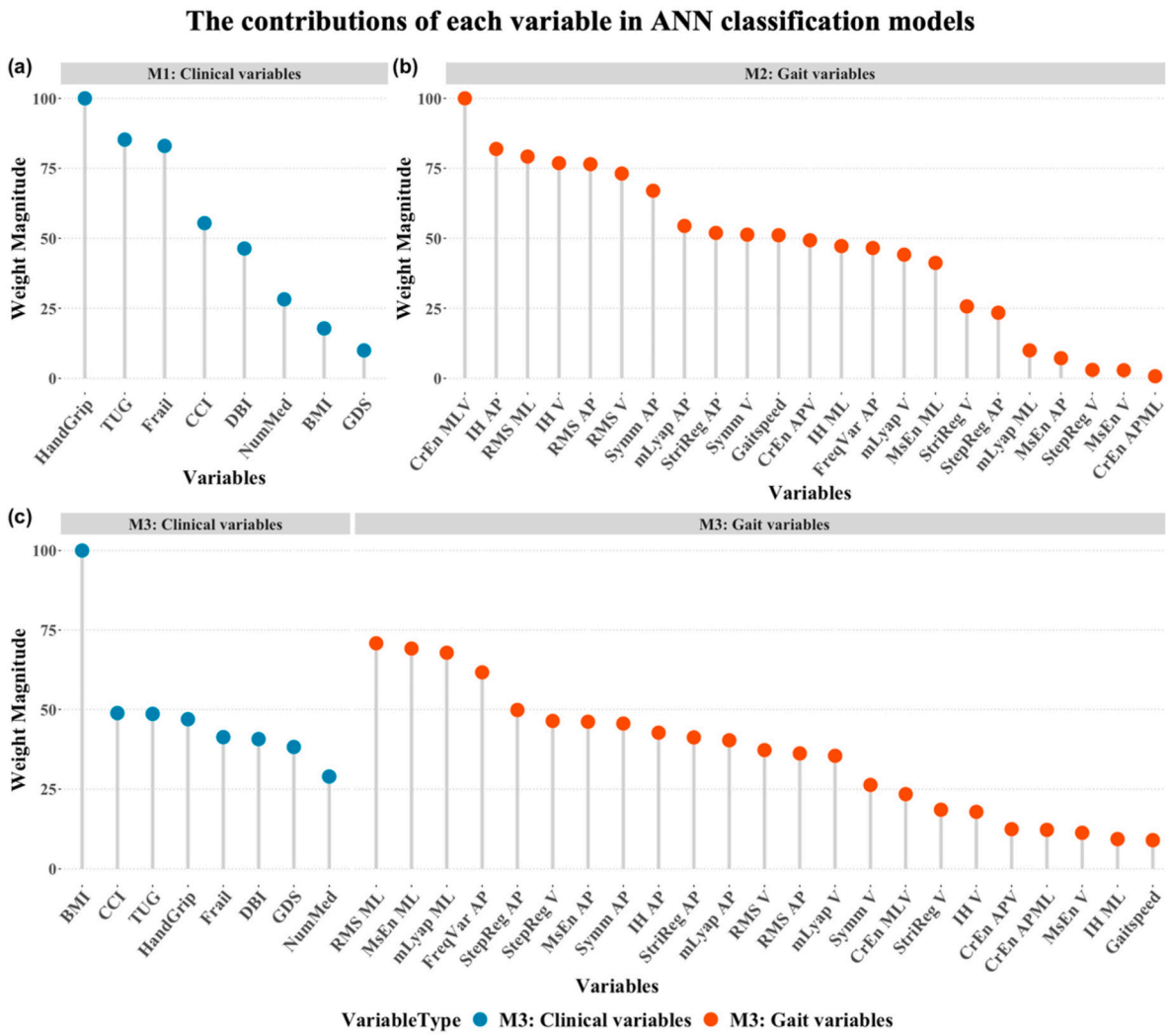

Fig. 3. The contributions of each variable in ANN classification models. (a) weights of eight clinical variables in model 1, **(b)** weights of 23 dynamic gait variables in model 2, and **(c)** weight of 31 variables which combined eight clinical variables with the 23 gait variables in model 3. Each variable was z-score-standardized prior to entering the model. For abbreviations, BMI, Body mass index; HandGrip, grip strength of hand; TUG, Timed-Up-and-Go test; CCI, Charlson Comorbidity Index; GDS, Geriatric depression scale; Frail, frailty criteria; DBI, Drug Burden Index; NumMed, the number of medications used; RMS, root mean square; MsEn, Multiscale Entropy; Step/StriReg, gait step or stride regularity; Symm, gait symmetry; FreqVar, Frequency variability; mLyap, maximal Lyapunov exponent; CrEn, Cross-sample Entropy; IH, Index of Harmonicity; AP, anterior-posterior direction; ML, medio-lateral direction; V, vertical direction.

variables to represent patients' comorbidities in terms of psychological state, physical function, frailty, and medical conditions. The 23 dynamic gait variables as inputs to machine learning models captured many features of gait over a long time period (minutes). The differences between geriatric patients with and without CI in these clinical and gait variables are indicative of CI and therefore are clinically relevant for a cognition-based classification of these geriatric patients. The reason for including so many clinical and gait variables is to capture the diverse effects of comorbidities on clinical and gait dysfunctions. Such an approach reduces the bias of selecting variables as inputs to machine learning analyses, which can handle high dimensions and non-linear associations among input variables in a relatively small sample size.

Using dynamic gait variables as input in a previous study, ANN classified young-middle age adults, healthy older adults, and geriatric patients with a classification accuracy of 90% and an AUC of 0.86 using dynamic gait variables [23]. As in that study, we also noticed that the classification performances of three ANN machine learning models (accuracy = 79%–96%) outperformed the RF classification performances (accuracy = 60%–64%). The RF model's low classification accuracy might be related to the random generation of multiple decision trees for gait and clinical variables without considering the interrelationships between these variables or trees [27]. In contrast to RF, ANN takes into account the non-linear, high dimensional interactions between the clinical variables and dynamic gait variables, based on the weight coefficients and activation functions (e.g., ReLU) [28,40].

Concerning the three ANN models, the classification based on the eight clinical variables (model 1) produced moderate classification accuracy with low sensitivity and specificity, implying that clinical variables alone are not sufficient to classify the two groups. When the classification was based on 23 dynamic gait variables, the classification performance had acceptable accuracy but with moderate specificity. In other words, model 2 incorrectly classified a high number of CI patients as CH patients so that dynamic gait variables alone did not accurately classify the two groups. However, model 3 was based on all 31 clinical and dynamic gait variables, produced an accurate classification, as nearly all CH patients were correctly assigned to CH group and most CI patients were distinguished correctly from CH. The most precise classification performance of ANN model 3 confirmed the hypothesis that interactions between clinical variables and dynamic gait variables underlie the classification of geriatric patients with and without CI.

The variables with the highest weights in ANN models 1 and 2

differed from those in ANN model 3, underscoring the idea that the interaction among input variables was key in geriatric patient classification in the present study (Fig. 3). With regard to the highly weighted variables in the ANN model 3, dynamic gait variables related to pace, predictability, stability, variability and regularity were identified to be sensitive indicators of CI. Although gait speed has previously emerged as a predictor of CI [41,42], we found that gait speed had a low weight and did not make a contribution to predicting CI (Fig. 3c). One reason could be that gait speed incorporates many features of gait; hence it is auto-correlated with most gait variables. For example, gait speed strongly correlated with gait regularity [43]. Moreover, unlike dynamical gait variables, gait speed is the most often used and reported mobility outcome, which increases the likelihood of accidental discovery [41,42]. The classification performance of ANN model 3 indicates that the inclusion of multiple gait variables as inputs underlie the high sensitivity and specificity of patient classification. This finding is in line with the successful classification of age groups based on gait dynamics, where stability and regularity had much higher weights than gait speed [23].

With respect to the direction of the acceleration, the gait of CI compared with CH patients is more unpredictable and unstable in the ML direction (Fig. 3c, Appendix Fig. A). Accelerations in the ML direction are key determinants of dynamic balance during walking [44]. Reduced gait predictability and stability are associated with an increase in fall risk in geriatric patients [45]. Furthermore, compared with CH patients, CI patients' less regular and more variable gait in AP direction might be related to the decline of executive function, a specific cognitive function that regulates planning and organisation relative to gait progression. The reduction in executive function may lead to incorrect control of limb movement, resulting in the inability to adjust and adapt gait made up of irregular steps [46].

The accuracy and efficiency of CI diagnosis in an early phase requires more extensive, expensive and burdensome neuropsychological testing and other diagnostics such as Magnetic Resonance Imaging scanning and analyses of cerebrospinal fluid [47]. Moreover, an accurate diagnosis of CI from small and perhaps inconsistent changes in cognition requires professional training [48]. Therefore, gait measurements combined with clinical parameters may be an inexpensive and accessible method to identify a cognitive decline in geriatric patients in an early stage [24]. Although many studies have measured gait or clinical variables alone in patients with CI and fall risk, the interactions of these variables cannot be identified by traditional statistical analysis.

ANN has been applied for clinical aims using in different patient populations. For example, in stroke patients ANN methods have been use to: a) screen patients who might suffer a stroke in the future [49,50]; b) identify patients who are at risk for a transient ischemic attack [51]; c) predict motor function [52], and d) to detect gait events [53]. The inertial measurement unit or IMU-based screening can be accuracy up to 99% [1,2]. Compared with RF, ANN also proved to be more accurate to classify age groups based on gait analysis (73% vs. 90% accuracy) [23]. ANN models have thus ability to accurately classify different populations based on non-linearly correlated, high dimensional human motion data. Unlike other machine learning methods, ANNs can be trained to detect the complex relationships between model inputs and model outputs. Such characteristics of ANN, after proper training, allow researchers to classify individual patients according to age, sex, and pathology. However, the use of ANN still requires caution and expertise, for instance, gender classification of Malaysian children based on gait analysis was only successful in a large sample size and the model had to be adjusted to produce an accurate classification without overfitting [54]. Additionally, researchers must carefully check and interpret the complex interaction between variables and 'translate' the results into a language clinicians can understand and use for diagnostic purpose or monitoring interventions.

Hence, a limitation for the clinical application of machine learning algorithms, even of those that output variable weight coefficients, is that the relationship between input variables, whether linear or non-linear, remains hidden. This limits the clinical interpretability of machine learning-generated outputs of gait classification. The current data were used to classify geriatric patients' cognitive status retrospectively and future studies will need to examine if the prospective classification of CI would be similarly accurate. Geriatric patients with CI in the present study had an irregular, unpredictable, and unstable gait, suggesting a patient-specific gait pattern but it is not clear if such gait properties would be the hallmarks of prospective identification of patients with CI [11]. Another limitation is that we classified patients based on MMSE without consideration of the dementia type or mild cognitive impairment. It remains unknown how different cognitive domains or dementia types are associated with gait and clinical variables.

5. Conclusion

In conclusion, an ANN model incorporating the interaction between clinical and gait variables classified geriatric patients with high accuracy, sensitivity, and specificity. This form of machine learning analyses of gait and clinical variables can inform geriatricians in the diagnosis of geriatric patients' cognitive status. There is also evidence suggesting that CI is frequently associated with falls and the prevalence of falls increases with the degree of CI [55]. Future studies may shed light on the question which patients with CI are more prone to a fall.

Data availability

The statistical description of the dataset (131 participants and 31 variables) is shown in Supplementary Information. Due to Institute Review Board related matters, the data can be available from the principal investigator Claudine Lamoth (c.j.c.Lamoth@umcg.nl) upon a reasonable request.

Funding

This work was supported by Keep Control project, funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721577.

Author contributions

Y.Z. and C.J.C.L. contributed to study concept and design. J.V.C. contributed to data acquisition. Y.Z. conducted the data analyses. All the authors contributed to the interpretation of the results. Y.Z. contributed to the drafting of the manuscript and figures. Y.Z., C.J.C.L., T.H., and J. V.C. critically revised the manuscript. All the authors made a significant contribution to the research and the development of the manuscript and approved the final version for publication.

Declaration of competing interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmed.2022.100076.

References

[1] Qiu C, Fratiglioni L. Aging without dementia is achievable: current evidence from epidemiological research. J Alzheim Dis 2018;62:933–42. https://doi.org/10.3233/JAD-171037.

[2] Folstein MF, Folstein SE, McHugh PR. Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98. https://doi.org/10.1016/0022-3956(75)90026-6.

[3] Meulen EFJ, Schmand B, Van Campen JP, De Koning SJ, Ponds RW, Scheltens P, et al. The seven minute screen: a neurocognitive screening test highly sensitive to various types of dementia. J Neurol Neurosurg Psychiatr 2004;75:700–5. https://doi.org/10.1136/jnnp.2003.021055.

[4] Neufeld S, Machacova K, Mossey J, Luborsky M. Walking ability and its relationship to self-rated health in later life. Clin Gerontol 2013;36:17–32. https://doi.org/10.1080/07317115.2012.731477.

[5] Rodríguez-Molinero A, Herrero-Larrea A, Miñarro A, Narvaiza L, Gálvez-Barrón C, Gonzalo León N, et al. The spatial parameters of gait and their association with falls, functional decline and death in older adults: a prospective study. Sci Rep 2019;9:1–9. https://doi.org/10.1038/s41598-019-45113-2.

[6] Zhou Y, Zia R, Rehman U, Hansen C, Maetzler W, Din S Del, et al. Classification of neurological patients to identify fallers based on spatial-temporal gait characteristics measured by a wearable device. Sensors 2020;20:4098. https://doi.org/10.3390/s20154098. 2020;20:4098.

[7] Zhang W, Low L-F, Schwenk M, Mills N, Gwynn JD, Clemson L. Review of gait, cognition, and fall risks with implications for fall prevention in older adults with dementia. Dement Geriatr Cognit Disord 2019;48:17–29. https://doi.org/10.1159/000504340.

[8] Beauchet O, Launay CP, Annweiler C, Allali G. Hippocampal volume, early cognitive decline and gait variability: which association? Exp Gerontol 2015;61:98–104. https://doi.org/10.1016/j.exger.2014.11.002.

[9] Verlinden VJA, Van Der Geest JN, Hofman A, Ikram MA. Cognition and gait show a distinct pattern of association in the general population. Alzheimer's Dementia 2014;10:328–35. https://doi.org/10.1016/j.jalz.2013.03.009.

[10] Peel NM, Alapatt LJ, Jones LV, Hubbard RE. The association between gait speed and cognitive status in community-dwelling older people: a systematic review and meta-analysis. J Gerontol Series A Biol Sci Med Sci 2019;74:943–8. https://doi.org/10.1093/gerona/gly140.

[11] Kikkert LHJ, Vuillerme N, van Campen JP, Hortobágyi T, Lamoth CJ. Walking ability to predict future cognitive decline in old adults: a scoping review. Ageing Res Rev 2016;27:1–14. https://doi.org/10.1016/j.arr.2016.02.001.

[12] Savica R, Wennberg AMV, Hagen C, Edwards K, Roberts RO, Hollman JH, et al. Comparison of gait parameters for predicting cognitive decline: the mayo clinic study of aging. J Alzheim Dis 2017;55:559–67. https://doi.org/10.3233/JAD-160697.

[13] Kikkert LHJCJ, Vuillerme N, Van Campen JP, Appels BA, Hortobágyi T, Lamoth CJC. Gait characteristics and their discriminative power in geriatric patients with and without cognitive impairment. J NeuroEng Rehabil 2017;14:1–10. https://doi.org/10.1186/s12984-017-0297-z.

[14] Kikkert LHJ, Vuillerme N, Van Campen JP, Appels BA, Hortobágyi T, Lamoth CJC. The relationship between gait dynamics and future cognitive decline: a prospective pilot study in geriatric patients. Int Psychogeriatr 2018;30:1301–9. https://doi.org/10.1017/S1041610217002770.

[15] Verlinden VJA, van der Geest JN, Hofman A, Ikram MA. Cognition and gait show a distinct pattern of association in the general population. Alzheimer's Dementia 2014;10:328–35. https://doi.org/10.1016/j.jalz.2013.03.009.

[16] Beauchet O, Allali G, Launay C, Herrmann FR, Annweiler C. Gait variability at fast-pace walking speed: a biomarker of mild cognitive impairment? J Nutr Health Aging 2013;17:235–9. https://doi.org/10.1007/s12603-012-0394-4.

[17] Wouters H, Hilmer SN, Gnjidic D, Van Campen JP, Teichert M, Van Der Meer HG, et al. Long-term exposure to anticholinergic and sedative medications and cognitive and physical function in later life. J Gerontol: Series A 2019;75:357–65. https://doi.org/10.1093/gerona/glz019.

[18] Mumic de Melo L, Hotta Ansai J, Giusti Rossi P, Carvalho Vale FA, Cristhine de Medeiros Takahashi A, Pires de Andrade L. Performance of an adapted version of the timed up-and-go test in people with cognitive impairments. J Mot Behav 2019;51:647–54. https://doi.org/10.1080/00222895.2018.1552917.

[19] Melikyan ZA, Corrada MM, Dick MB, Whittle C, Paganini-Hill A, Kawas CH. Neuropsychological test norms in cognitively intact oldest-old. J Int Neuropsychol Soc 2019;25:530–45. https://doi.org/10.1017/S1355617719000122.

[20] Weakley A, Williams JA, Schmitter-Edgecombe M, Cook DJ. Neuropsychological test selection for cognitive impairment classification: a machine learning approach. J Clin Exp Neuropsychol 2015;37:899–916. https://doi.org/10.1080/13803395.2015.1067290.

[21] Zhang W, Low L-F, Schwenk M, Mills N, Gwynn JD, Clemson L. Review of gait, cognition, and fall risks with implications for fall prevention in older adults with dementia. Dement Geriatr Cognit Disord 2019;48:17–29. https://doi.org/10.1159/000504340.

[22] Morris R, Lord S, Bunce J, Burn D, Rochester L. Gait and cognition: mapping the global and discrete relationships in ageing and neurodegenerative disease. Neurosci Biobehav Rev 2016;64:326–45. https://doi.org/10.1016/j.neubiorev.2016.02.012.

[23] Zhou Y, Romijnders R, Hansen C, Campen J van, Maetzler W, Hortobágyi T, et al. The detection of age groups by dynamic gait outcomes using machine learning approaches. Sci Rep 2020;10:1–12. https://doi.org/10.1038/s41598-020-61423-2.

[24] Jin L, Lv W, Han G, Ni L, sun D, Hu X, et al. Gait characteristics and clinical relevance of hereditary spinocerebellar ataxia on deep learning. Artif Intell Med 2020;103:101794. https://doi.org/10.1016/j.artmed.2020.101794.

[25] Rehman RZU, Zhou Y, Del Din S, Alcock L, Hansen C, Guan Y, et al. Gait analysis with wearables can accurately classify fallers from non-fallers: a step toward better management of neurological disorders. Sensors 2020;20:6992. https://doi.org/10.3390/s20236992.

[26] Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One 2019;14:e0224365. https://doi.org/10.1371/journal.pone.0224365.

[27] Lakshmanaprabu SK, Shankar K, Ilayaraja M, Nasir AW, Vijayakumar V, Chilamkurti N. Random forest for big data classification in the internet of things using optimal features. Int J Machine Learn Cybernet 2019;10:2609–18. https://doi.org/10.1007/s13042-018-00916-z.

[28] Bircanoglu C, Arica N. A comparison of activation functions in artificial neural networks. 2018 26th Signal Processing and Communications Applications Conference. In: SIU. IEEE; 2018. p. 1–4. https://doi.org/10.1109/SIU.2018.8404724.

[29] Lamoth CJ, van Deudekom FJ, van Campen JP, Appels BA, de Vries OJ, Pijnappels M. Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people. J NeuroEng Rehabil 2011;8:2. https://doi.org/10.1186/1743-0003-8-2.

[30] de Groot MH, van Campen JPCM, Kosse NM, de Vries OJ, Beijnen JH, Lamoth CJC, et al. The association of medication-use and frailty-related factors with gait performance in older patients. PLoS One 2016;11:e0149888. https://doi.org/10.1371/journal.pone.0149888.

[31] Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc 1992;40:922–35. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x.

[32] Kosse NM, Caljouw S, Vervoort D, Vuillerme N, Lamoth CJC. Validity and reliability of gait and postural control analysis using the tri-axial accelerometer of the iPod Touch. Ann Biomed Eng 2015;43:1935–46. https://doi.org/10.1007/s10439-014-1232-0.

[33] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis 1987;40:373–83. https://doi.org/10.1016/0021-9681(87)90171-8.

[34] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol: Series A 2001;56:M146–57. https://doi.org/10.1093/gerona/56.3.M146.

[35] Ibrahim A, Singh DKA, Shahar S. 'Timed up and Go' test: age, gender and cognitive impairment stratified normative values of older adults. PLoS One 2017;12:e0185641. https://doi.org/10.1371/journal.pone.0185641.

[36] Sallinen J, Stenholm S, Rantanen T, Heliövaara M, Sainio P, Koskinen S. Hand-grip strength cut points to screen older persons at risk for mobility limitation. J Am Geriatr Soc 2010;58:1721–6. https://doi.org/10.1111/j.1532-5415.2010.03035.x.

[37] Kosse NM, Vuillerme N, Hortobágyi T, Lamoth CJ. Multiple gait parameters derived from iPod accelerometry predict age-related gait changes. Gait Posture 2016;46:112–7. https://doi.org/10.1016/J.GAITPOST.2016.02.022.

[38] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. Proc IEEE Int Conf Comput Vis 2015:1026–34.

[39] Robbins H, Monro S. A stochastic approximation method. Ann Math Stat 1951;22:400–7. https://doi.org/10.1214/AOMS/1177729586.

[40] Pasini A. Artificial neural networks for small dataset analysis. J Thorac Dis 2015;7:953–60. https://doi.org/10.3978/j.issn.2072-1439.2015.04.61.

[41] Garcia-Pinillos F, Cozar-Barba M, Munoz-Jimenez M, Soto-Hermoso V, Latorre-Roman P. Gait speed in older people: an easy test for detecting cognitive impairment, functional independence, and health state. Psychogeriatrics 2016;16:165–71. https://doi.org/10.1111/psyg.12133.

[42] Hoogendijk EO, JJM Rijnhart, Skoog J, Robitaille A, van den Hout A, Ferrucci L, et al. Gait speed as predictor of transition into cognitive impairment: findings from three longitudinal studies on aging. Exp Gerontol 2020;129:110783. https://doi.org/10.1016/j.exger.2019.110783.

[43] Rabuffetti M, Scalera G, Ferrarin M. Effects of gait strategy and speed on regularity of locomotion assessed in healthy subjects using a multi-sensor method. Sensors 2019;19:513. https://doi.org/10.3390/s19030513.

[44] Buurke TJW, Lamoth CJC, van der Woude LHV, Hof AL, den Otter R. Bilateral temporal control determines mediolateral margins of stability in symmetric and asymmetric human walking. Sci Rep 2019;9:1–10. https://doi.org/10.1038/s41598-019-49033-z.

[45] Riva F, Toebes MJPJP, Pijnappels M, Stagni R, van Dieën JHH. Estimating fall risk with inertial sensors using gait stability measures that do not require step detection. Gait Posture 2013;38:170–4. https://doi.org/10.1016/J.GAITPOST.2013.05.002.

[46] Yogev-Seligmann G, Hausdorff JM, Giladi N. The role of executive function and attention in gait, vol. 23; 2008. https://doi.org/10.1002/mds.21720.

[47] Laske C, Sohrabi HR, Frost SM, López-De-Ipiña K, Garrard P, Buscema M, et al. Innovative diagnostic tools for early detection of Alzheimer's disease. Alzheimer's Dementia 2015;11:561–78. https://doi.org/10.1016/j.jalz.2014.06.004.

[48] Boron JB, Turiano NA, Willis SL, Schaie KW. Effects of cognitive training on change in accuracy in inductive reasoning ability. J Gerontol B Psychol Sci Soc Sci 2007;62. https://doi.org/10.1093/geronb/62.3.P179.

[49] A V, G N, T G, H N, H R, B-R J, et al. Novel screening tool for stroke using artificial neural network. Stroke 2017;48:1678–81. https://doi.org/10.1161/STROKEAHA.117.017033.

[50] Varrecchia T, Castiglia SF, Ranavolo A, Conte C, Tatarelli A, Coppola G, et al. An artificial neural network approach to detect presence and severity of Parkinson's disease via gait parameters. PLoS One 2021;16:e0244396. https://doi.org/10.1371/JOURNAL.PONE.0244396.

[51] Chan KL, Leng X, Zhang W, Dong W, Qiu Q, Yang J, et al. Early identification of high-risk TIA or minor stroke using artificial neural network. Front Neurol 2019;10:171. https://doi.org/10.3389/FNEUR.2019.00171.

[52] Thakkar HK, Liao W, Wu C, Hsieh Y-W, Lee T-H. Predicting clinically significant motor function improvement after contemporary task-oriented interventions using machine learning approaches. J NeuroEng Rehabil 2020;17. https://doi.org/10.1186/S12984-020-00758-3.

[53] A M. Gait event detection using a multilayer neural network. Gait Posture 2009;29:542–5. https://doi.org/10.1016/J.GAITPOST.2008.12.003.
[54] Zakaria NK, Jailani R, Tahir NM. Application of ann in gait features of children for gender classification. Procedia Comput Sci 2015;76:235–42.
[55] Montero-Odasso M, Speechley M. Falls in cognitively impaired older adults: implications for risk assessment and prevention. J Am Geriatr Soc 2018;66:367–75. https://doi.org/10.1111/jgs.15219.