Biomedical Signal Processing and Control 43 (2018) 320–328

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Prospective elderly fall prediction by older-adult fall-risk modeling with feature selection

Jennifer Howcroft[a], Edward D. Lemaire[b,c], Jonathan Kofman[a,*]

[a] Department of Systems Design Engineering, University of Waterloo, 200 University Ave West, Waterloo, ON N2L 3G1, Canada
[b] Centre for Rehabilitation Research and Development, Ottawa Hospital Research Institute, Ottawa, ON K1H 8M2, Canada
[c] Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada

ARTICLE INFO

Article history:
Received 9 November 2017
Received in revised form 14 February 2018
Accepted 17 March 2018

Keywords:
Fall-risk
Prediction
Prospective falls
Elderly
Wearable sensors
Feature selection
Accelerometer
Plantar pressure

ABSTRACT

This study implemented feature-selection methods for prospective elderly fall-risk prediction and compared modeling outcomes to retrospective fall occurrence classification. Seventy-five elderly adults without prior falls (75.2 (6.6) years; 47 non-fallers, 28 fallers based on 6 month prospective falls) walked 7.62 m while wearing F-Scan insoles and tri-axial accelerometers at the left and right shanks, pelvis, and head. Feature sets were reduced using Relief-F, correlation-based feature selection, and fast correlation based filter algorithms. Naïve Bayesian, multi-layer perceptron neural network, and support vector machine (SVM) classifiers were used for faller prediction. The top twenty feature selection-based models and top ten all-variable-based models determined by single 75:25 train:test stratified holdouts were further examined by training 10,000 models with randomized 75:25 train:test stratified holdouts. Model evaluation parameters were averaged across all 10,000 models and then ranked. Feature selection increased predictive accuracy by 9%, comparing models with and without feature selection. The increase in accuracy was greater than for retrospective faller classification using feature selection. Based on the 10,000 randomized holdouts, the highest ranked model had 65% accuracy and 59% sensitivity, achieved by a pressure-sensing-insole and left-shank-accelerometer based feature subset. The best performance without feature selection was 56% accuracy and 42% sensitivity. Feature selection improved model performance and should therefore be included as a model development step for elderly fall-risk prediction. Some model performance results, such as Relief-F providing the best feature selection and high single-sensor performance using the lower-back accelerometer, were consistent for prospective faller prediction and retrospective classification. However, other contributions to high model performance, such as dual-task assessment for single-sensor models and use of SVM, were specific to older adult fall-risk prediction based on prospective fall occurrence.

1. Introduction

The assessment of older-adult fall risk is clinically important given the relatively high risk of falls (approximately one third) in this population [1]. A wide variety of fall risk assessment tools have been used including questionnaires, checklists, subjective movement-based assessments such as the Timed Up & Go and the Berg Balance Scale, and quantitative movement-based assessments that rely on wearable-sensor-derived data [2,3]. This study focused on wearable-sensor-based assessment due to its potential to provide an objective, biomechanically-based fall risk assessment that could be applied at the point-of-care or home environment [4].

Within the field of wearable-sensor-based fall risk assessment, a wide range of sensors (accelerometer, gyroscopes, magnetometers, force sensors, pressure-sensing insoles), features, models, and participants have been investigated and resulted in models with a wide range of predictive accuracy (62–100%) [3–5]. The major-

Abbreviations: AP, anterior-posterior; ASUFSR, Arizona State University Feature Selection Repository; AV, all variables; CFS, correlation-based feature selection; CI, confidence intervals; CoP, center of pressure; CoV, coefficient of variation; DT, dual-task; FCBF, fast correlation based filter; FFTFQ, ast Fourier Transform first quartile; FS, feature selection; H, head accelerometer measures; I, pressure-sensing insole measures; I1, impulse, foot-strike to first peak; I2, impulse, first peak to minimum; I3, impulse, minimum to second peak; I4, impulse, second peak to foot-off; I5, impulse, foot-strike to minimum; I6, impulse, minimum to foot-off; I7, impulse, foot-strike to foot-off; LS, left shank accelerometer measures; MCC, Matthew's Correlation Coefficient; ML, medial-lateral; MLE, maximum Lyapunov exponent; NB, naïve Bayes; NB-L, linear naïve Bayesian; NN, neural network; NN-*a*, NN with *a*, the number of nodes in hidden layer; NPV, negative predictive value; P, pelvis accelerometer measures; PCA, principal component analysis; PPV, positive predictive value; REOH, ratio of even to odd harmonics; RS, right shank accelerometer measures; SD, standard deviation; SR, summed rank; ST, single-task gait; SVM, support vector machine; SVM-*b*, SVM with polynomial degree *b*.

* Corresponding author.
E-mail addresses: j2irwin@uwaterloo.ca (J. Howcroft), elemaire@ohri.ca (E.D. Lemaire), jkofman@uwaterloo.ca (J. Kofman).

https://doi.org/10.1016/j.bspc.2018.03.005
1746-8094/

ity of studies identified in [4] used retrospective fall data as their standard for evaluating predictive accuracy while only a small proportion (15%) used prospective fall data [4,5]. Retrospective fall data has two main limitations: inaccurate fall recall and gait pattern changes that occur between a fall and the assessment, due to injury or changes in gait patterns in an attempt to increase stability and reduce fear of falling [4,5]. Prospective fall occurrence should be considered the gold standard for evaluating fall risk assessment models [4,5]. Given that most studies to date have relied on retrospective fall occurrence to evaluate fall risk, it is important to determine the impact of this methodological decision on fall risk modeling results like optimal features and model parameters.

One modeling parameter that should be considered when creating a fall risk model is feature-space size reduction. A wide variety of variables can be derived from wearable sensors, with at least 130 inertial sensor-based variables already used to evaluate fall risk [4]. More signal features can lead to excessive computational cost and the "curse of dimensionality" [6,7], but this can be avoided if redundant and irrelevant features are removed as part of the analysis [6,7].

Most studies that applied feature-space size reduction techniques to older adult fall risk classification used retrospective fall history [8–12]. Only one study predicted older adult prospective fall occurrence in combination with feature-space size reduction techniques [13]. However, this study [13] was based on a six-minute walk test, which takes more time to administer at the point-of-care. In our earlier research with the easy to administer and shorter duration twenty-five foot (7.62 m) walk test [14], fall risk models were developed from prospective fall occurrence data; however, these were developed without feature selection. When feature selection was used on retrospective fall data [12], older adult fall-risk model classification performance improved over models without feature selection [15]. Since people who have fallen may have already modified their gait to avoid future falls, identifying fallers prospectively becomes a different and important classification task. No previous study based on a clinically efficient walk test has been performed with prospective data using feature selection to optimize prediction models. Therefore, research is needed to determine if feature selection methods can improve older adult prospective fall prediction based on a clinically efficient walk test.

This study investigated prospective elderly fall prediction based on the twenty-five foot walk test, using fall-risk modeling with feature selection. The modeling outcomes were compared to retrospective fall occurrence classification results [12], thereby determining how prospective and retrospective approaches to fall prediction differ when appropriate feature selection methods are implemented. We hypothesized that feature selection would improve prospective fall prediction, and that model features and outcomes would differ for prospective fall prediction compared to retrospective fall classification.

2. Methods

2.1. Participants

A convenience sample of 76 older adults, 65 years or older, with no falls six months prior to the study, were recruited from the community. Data from this participant group were also used in our previous research on prospective fall occurrence without feature selection [14], allowing a direct comparison between prospective faller prediction models with and without feature selection. For prospective analysis, participants were identified as fallers if they reported at least one fall (defined in [1]) during a six month follow-up period. Seventy-five participants completed the six month follow-up; 47 were classified as non-fallers and 28 fallers (Table 1). There were no significant differences between non-fallers and fallers in age, height, or weight as evaluated by the *t*-test. The mean number of falls during the six month follow-up for the 28 fallers was 1.3 (range: one to four falls). Participants with a cognitive disorder (self-reported) or those unable to walk for six minutes without an assistive device were excluded. The University of Waterloo Research Ethics Committee approved the study (ORE#: 19106) and all participants gave informed written consent.

2.2. Protocol

A complete description of the data collection protocol can be found in [14]. Briefly, participants wore accelerometers (X16-1C, Gulf Coast Data Concepts, Waveland, MS) attached to the posterior head, posterior pelvis, and lateral shanks just above the ankle. F-Scan pressure-sensing insoles (F-Scan 3000E, Tekscan, Boston, MA) were worn in their shoes. Data were collected at 50 Hz for accelerometers and 120 Hz for pressure-sensing insoles. Participants walked 7.62 m (25 ft) with a cognitive load (dual task: DT) and without a cognitive load (single task: ST), in separate trials, while plantar pressure and accelerometer data were collected. The cognitive task required participants to say words starting with A, F, or S [16]. Participant fall information was collected over a six-month follow-up period.

2.3. Data processing

For ST and DT trials, calculated plantar-pressure parameters were: number, length, and duration of both medial-lateral (ML) and posterior deviations per stance; anterior-posterior (AP) and ML stance phase center of pressure (CoP) path coefficients of variation (CoV); gait velocity; cadence; stride time; stance time; swing time; percent stance time; percent double support time; stride time symmetry index; stride time, stance time, and swing time CoV; and impulse parameters [15]. Accelerometer parameters were: cadence; stride time; and for each axis: maximum, mean, and standard deviation of acceleration; Fast Fourier Transform First Quartile (FFTFQ); ratio of even to odd harmonics (REOH); and maximum Lyapunov exponent (MLE) [15]. Temporal parameters (cadence and stride time) were calculated for each accelerometer and pressure-sensing insoles to ensure that models relying on only one sensor still included these temporal parameters. Accelerometer cadence and stride time calculations were based on detection of unique peaks or troughs that occurred consistently once per stride, with all detections confirmed manually [17–19]. Insole-based heel strike and foot off detections were based on a total force threshold of 5N plus the minimum, baseline force recorded during the walking session. This baseline force was due to contact between the shoe and foot, with the insole in-between, during all stages of gait, including swing. Thresholds between 0 and 10 N and thresholds that added participant-specific baseline forces to the threshold have been used in the literature [20–23].

2.4. Feature selection

The feature selection techniques used in this study were used in our previous work [12], allowing direct comparison to retrospective faller classification models. Three feature selection methods were used as detailed in [12]: correlation-based feature selection (CFS), fast correlation based filter (FCBF), and Relief-F. The number of features in the Relief-F feature subsets was determined using the runExperiment algorithm within the Arizona State University Feature Selection Repository (ASUFSR) [24].

Feature selection was performed as a pre-processing step for all 31 possible sensor combinations [14] using ASUFSR algorithms [24]. Identical feature subsets were inputs for single stratified hold-