

Sway Risk Analysis Based on Age Group Classification

Hafsa Ismail^{1,2}, Ibrahim Radwan³, Hanna Suominen¹⁻⁴, Gordon Waddington¹ and Roland Goecke¹

Abstract—Human postural sway is the horizontal movement, which is generated to control a person’s balance while walking or standing. Changes in postural sway patterns are a reflection of changes in brain signals and in physical health affected by a person’s aging process. This increases the likelihood of falling. In this paper, we propose a method to estimate the age groups of elderly people based on the extracted sway measurements. These measurements are computed based on learning the postural sway signal from video footage only. Then, we use the person’s real age and the estimated age group to define a risk of a fall for the elderly people becoming more likely. This may lead to an early intervention and help them to prevent serious falls or injuries by putting countermeasures in place, e.g. physical exercise regimes. The conducted experiments show the reliability of the proposed method to produce a risk analysis tool for elderly based on their estimated sway signals.

I. INTRODUCTION

Postural sway is the slight horizontal movements of the human body’s centre of pressure (CoP) to maintain the balance within the body’s base of support [1]. It is a combination of antero-posterior (front-to-back) and medial-lateral (side-to-side) sway as illustrated in Fig. 1. Clinical studies show that medial-lateral sway is more related to maintaining body balance, thus more affected by changes in it.

Maintaining body balance is achieved by the coordination of input from the different sensory systems, *vestibular, somatosensory, and visual systems* [2]. Each of these systems gathers different information related to different reference points. Directional information that relates to head position and other organs that regulate the equilibrium is sensed by the vestibular system. The spatial position and movements relative to the support surface and the position of different body parts relative to each other are sensed by the somatosensory system, and spatial location relative to other objects is sensed by the visual system.

Maintaining balance stability is related to the brain’s ability to integrate the information from the different sensory systems and the muscle motor process of the different body parts to modify these processes as a response to the environmental factors, such as illumination, flooring, medication, alcohol/drug usage, and/or ear infection.

Aging affects the efficiency of a person’s sensory systems and the brain-to-muscle integration to respond to environmental changes. Thus, maintaining body balance becomes harder while aging and the likelihood of a fall increases.

To measure postural sway, different metrics have been identified, such as sway area, speed, frequency, and total

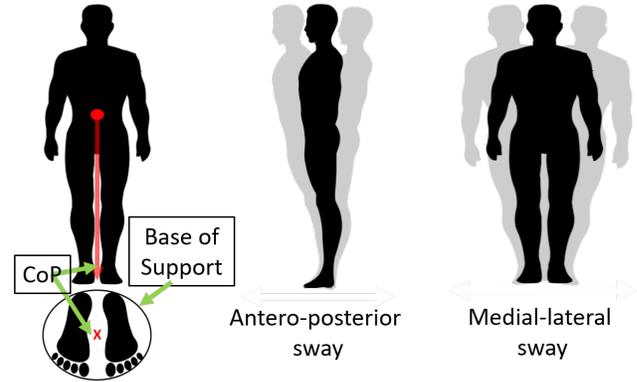


Fig. 1: Postural sway, the horizontal CoP movements to maintain the balance within the body’s base of support.

path length [4]. These metrics are used to analyse body movements and to detect abnormality when exist. Measuring such metrics requires special equipment, like force plate that captures the CoP movements or wearable sensors, which need to be attached to the body parts to capture their movements. In either cases, a special expensive installation and supervision are needed or, in the case of the wearable sensor, it needs to be attached to the body all the time to capture the required data.

In recent years, studying and analysing human movements has received more interest from the computer vision point-of-view for the purpose of human recognition and identification, as well as for fall detection.

Due to a lack of publicly available datasets for studying the fall risk for elderly people, this paper, firstly, describes the collected dataset that contains video as well as force plate balance data for two groups of people: athletes and elderly. Secondly, a computer vision and machine learning method is proposed to transfer the postural sway from the pixel space of the video data to the millimeter force-plate likewise signals. Then, the sway measurements are extracted from the transferred signals.

In this paper, we answer the question of the possibility of detecting increases in fall risk early, especially for elderly people, based on analysing sway features from the predicted postural sway signals using only video data.

The next section overviews the related work, followed by presenting the collected dataset and the research methodology. Then, the experiments and results are discussed. Finally, the conclusions and suggested future work are presented.

¹ University of Canberra, Bruce, Australia

² DATA61/CSIRO, Acton, Australia

³ Australian National University, Acton, Australia

⁴ University of Turku, Turku, Finland

II. RELATED WORK

Determining and analysing human movements, gait and sway, have been initiated from clinical and medical studies to determine non-normal movement patterns from normal ones. Human movement analysis is used in the clinical routine to identify pathological movements using a simplified model of the human body structure. In clinical re-habitation, human movement analysis is used in diagnosis, monitoring and implementation of methodologies that mitigate the effect of some pathologies associated with the movements. Human movement analysis in an aged population has and still receives wide attention in clinical studies to reduce the risk of a fall. Falls may cause serious injuries, especially in elderly people. The most frequent cause of falling in the elderly is the incorrect body weight shifting (balance control) that is associated with walking or standing still.

To assess human balance control, different functional balance tests have been developed. Examples of these test include the Romberg Test [6], Berg Balance test [6], Performance-Oriented Mobility Assessment (POMA) [6], Standing test for Imbalance and Disequilibrium (SIDE) [7], and the Balance Error Scoring System (BESS) [8]. Each of these balance test requires performing some stances and/or actions while being observed by a clinically trained examiner, who looks for deviations from the proper stance. With the technological improvements, some devices can be used for more accurate assessment of the body movement data. These devices can be wearable devices, such as sensors, non-wearable devices, such as force plates [3], or even video cameras to replay and assess the test more carefully.

The need for a laboratory or clinical environment and specialist supervisor with special equipment to assess the balance and diagnose any problems is an expensive and unreliable option when talking about continuous monitoring of the balance. On the other hand, using one (or more) wearable sensor for continuous monitoring of the human movements that may spot a balance disorder/problem requires wearing such sensor(s) all the time, which is not a reliable option, especially when talking about elder people.

The computer vision interest in human movements began with the human recognition and identification [9] for security and surveillance purposes. Also, for health purposes, such as fall detection, received a lot of attention [10] to improve the quality of human life but focused on the fall itself rather than preventing it in the first place.

In this study, using computer vision methods and techniques for measuring and analysing human movements, we present a cheaper and more reliable alternative for sway monitoring and analysis.

The next sections present the BESS balance test and the collected dataset in detail as well as the proposed methodology to extract the sway measurements from video recordings and to analyse these measurements toward early risk detection based on the sway. Experiments, results, conclusion and future work are presented afterwards.

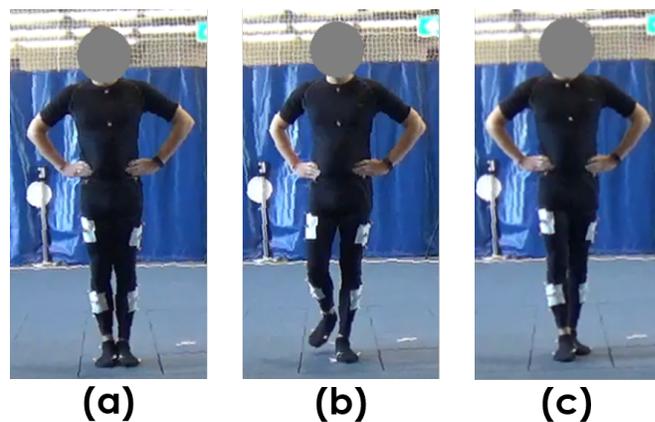


Fig. 2: Balance Error Scoring System (BESS) stances. (a) Double legs stance, (b) Single leg stance, and (c) Tandem stance

III. DATASET

Existing datasets that contain gait and balance test have been usually collected in clinical data settings and, hence, are confidential and not publicly available. Moreover, videos or motion capture data that are needed to assess the model accuracy were not included. On the other hand, datasets that have been collected for computer vision research usually do not include the balance test data that we require.

Two ethic approvals and research permissions were furnished and approved to collect our dataset parts: 1) the ground truth, and 2) the elderly data. The ground truth part is used to establish and assess the accuracy of the proposed model(s) to apply on the elderly dataset part. The two parts of the dataset are discussed in more details in following subsections, however firstly, we explain the balance error scoring system that is used to evaluate the postural stability.

1) *Balance Error Scoring System (BESS)*: The BESS balance test is used in our study. It is simple, affordable, accurate, and commonly used as an objective method to generally assess the static postural stability. It is also used particularly in sports settings to assess the effects of mild to moderate head injury on postural stability [8]. A typical BESS test assesses three stances: 1) double-leg stance by standing with touching feet side-by-side, 2) single-leg stance by standing on the non-dominant foot, and 3) tandem stance by standing heel to toe with the non-dominant foot at the back. BESS stances are shown in Fig. 2. The test is performed on two different surfaces (firm surface, ground/floor, and unstable surface, medium density foam, which is a more challenging balance task). Each stance is required to be taken for 20-30 seconds on each surface with closed eyes and hands on hips. A deviation is noted when a participant opens eyes, uses hand for balance, falls, lifts the forefoot or heel off the testing surface, or remains out of the proper testing position for more than five seconds.

In our dataset, the BESS test is performed on only the firm surface. Each test is taken on a Kissler force plate to precisely

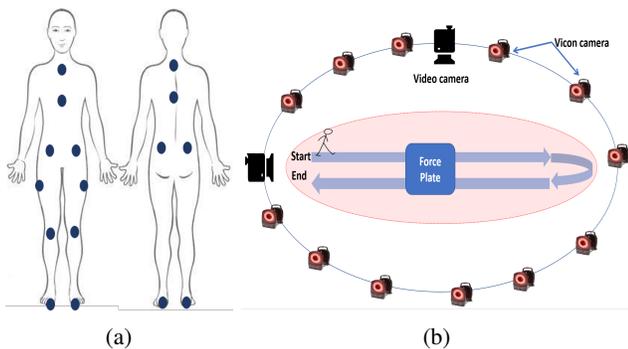


Fig. 3: (a) Selected Vicon markers and their positions on the participant's body, and (b) Ground truth recording lab set-up

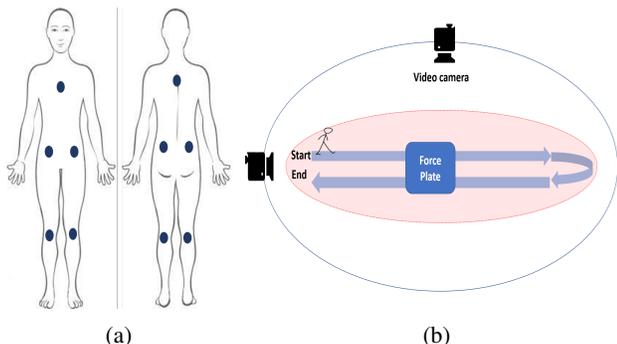


Fig. 4: (a) Locations of the distinctive colored markers on the participant's body, and (b) Elderly recording lab set-up

capture the actual Centre of Pressure (CoP) displacements, while recording the test with video cameras. As we aim to measure and analyse the postural sway only from video, the medial-lateral sway is measured and analysed.

2) *Ground Truth*: Seventeen athletic adults, six women and eleven men with an average age of thirty years, participated to create the ground truth part of the dataset. They were asked to walk normally for about ten metres, then making a U-turn at the end of the recording area, before also taking the BESS test. In the ground truth part, twelve motion capture cameras (Vicon) were used to collect body movements from sixteen markers that were placed on the participant's body as shown in Fig. 3a. An AMTI force plate was used to measure participant body's CoP movements when performing the balance test or when stepping on it while walking. Vicon system and the force plate were systemically synchronised. Two RGB cameras were used to capture the participant's movements from frontal and side views. Fig. 3b illustrates the ground truth recording lab set-up. The two views were aligned manually using a video editor application. Video and Vicon data were synchronised based on the 'start' and 'end' recording commands.

3) *Elderly dataset*: Seventy eight elderly people over fifty years old, sixty women and eighteen men, participated to collect the elderly part of the dataset. In the Elderly dataset, the main data are the video recordings for the participants walking and performing the BESS balance test. Two video

cameras are placed to capture the frontal and the side views. Also, a force plate is used to capture the CoP when performing the BESS test and when stepping on it while walking. The elderly recording lab setup is shown in Fig. 4b. Distinctive colour points are placed on the participant's upper torso, hips and knees, as shown in Fig. 4a, to facilitate body parts tracking over video frames.

IV. METHODOLOGY

Our research methodology starts with data pre-processing, which prepares the video and force plate data to be fed in the next step. A Recurrent neural network (RNN) predicts the video sway signal, measured in pixels, in the force plate measurement space, measured in mm [16]. For each subject, the sway measurements, *i.e.* maximum sway and sway frequencies, are calculated. Using the ground truth part of the dataset, a Gaussian mixture model (GMM) is used to build age group clusters based on the calculated sway measurements. Another GMM is used to classify the test subject into one of the age group clusters. Comparing that subject's classification with his/her real age can be used to identify early risks. The next subsections discuss in details each part of the methodology.

A. Data Pre-processing

Preparing the data for the training and testing includes extracting the tracking features and signal sampling as following:

Feature Tracking: While the upper body movements have more impact on postural sway as in [5], a Kanade-Lukas-Tomasi (KLT) tracker [11] is used to track the marker, which was placed on the person's upper torso over the video frames, while performing the BESS test. The tracking results in a signal, which represents the trajectory of the upper body sway over time of the recording.

Signal Resampling: The force plate captures the CoP displacements at a high frequency (1,000 Hz), which results in a noisy signal. On the other hand, the video was recorded at 25fps. Since the signal of interest (the postural sway) is characterised by low frequencies, the Butterworth filter [12] is applied to cut off the high frequencies of the force plate signal to reduce the noise and to make it suitable for training and testing.

B. Video-based Postural Sway Estimation

Given the tracking features of an input video, $X = \{x_1, x_2, \dots, x_n\}$, where x_i are the extracted features from frame i , and a target force plate signal $y = \{y_1, y_2, \dots, y_n\}$, the input and output signals are split into chunks of size n . Learning the conditional distribution $p(y|X)$ is done by encoding the source signal into fixed-length hidden states followed by decoding these hidden states into the target space. We have built a model based on the gated-recurrent neural network (GRU), which has shown outstanding performance in domain-space translation and adaptation [13], [14], [15].

To predict a force plate likewise signal from the tracked point over the video frames, a GRU encoder-decoder is

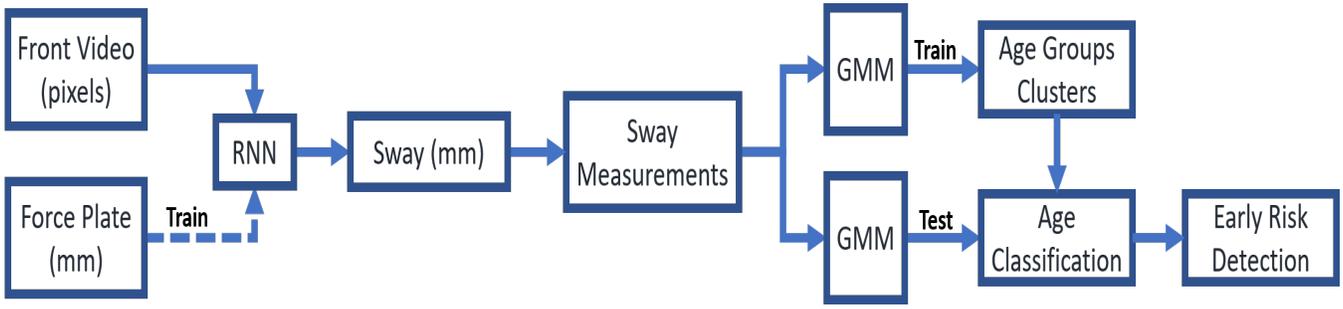


Fig. 5: The proposed methodology trains an RNN model to estimate upper body sway from video tracked points and force plate signals. Then, it generates age group clusters by training a Gaussian mixture model using sway measurements. In testing, the trained RNN model is used to predict the upper body sway from video, then the trained GMM is used to classify the sway measurements for the testing case into one of the age group clusters. Lastly, an early risk analysis is produced from the age classification.

proposed. the input sequence is transferred into a state (thought) vector $h = \{h_1, h_2, \dots, h_n\}$ by the encoder. The hidden states are updated iteratively by,

$$h_t = f(h_{t-1}, x_t) \quad (1)$$

where f is a non-linearity function.

The thought vector transforms the input signals into the space of the target signals. At each step, the GRU units decide how much information should be passed to the decoder. In the proposed GRU framework (Fig. 6), f is used as one gated recurrent unit (GRU [13]) to loop through n time steps and to transfer the input to an output-like space. A non-linear function q is applied by the encoder on the hidden states to calculate the thought vector v as follows:

$$v = q(h_1, h_2, \dots, h_n). \quad (2)$$

The cost function is applied on the output of each time step, *i.e.* frame, to estimate the output via the decoder process as follows:

$$y_t = g(h_t, y_{t-1}, v) \quad (3)$$

At each time step, the previous output state is considered as well as the thought vectors as input to a fully-connected 'Dense' layer. Then, a non-linear activation function g is applied on the output of every frame to produce the estimated output at this time step.

C. Sway Measurements

Postural sway is usually determined by measuring the bodys CoP displacements that are captured by the force plate. The area enclosed by the movements of the CoP in the $X-Y$ plane is known as the sway area, which is used as a basic measurement of the postural sway. Total path length, sway speed and frequency are other parameters that are commonly used to describe the amount of postural sway from force plate data. In this paper, the proposed method in [16] is applied on the frontal video recordings to estimate the medial-lateral sway of the body movements. Medial-lateral sway is the main predictor for the increased postural sway that reflects the body balance while standing or walking.

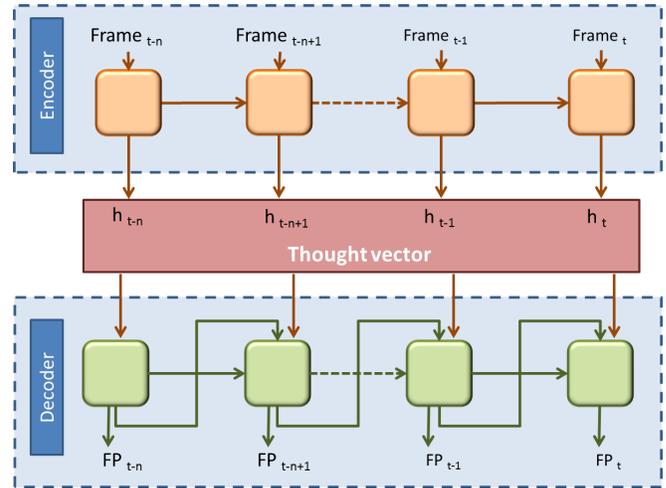


Fig. 6: The proposed sequence-to-sequence GRU framework. Top row: a batch of size n of the input sequence up to frame t . Bottom row: corresponding target sequence to be predicted. Middle row: the thought vector, which encodes the current state of the hidden units to produce the output

For measuring postural sway from video recordings, two measurements are extracted and used in the experiments of this paper: maximum sway and sway frequency as explained next.

1) *Maximum Sway*: When one direction of the medial-lateral (side-to-side) sway is noted as d , the maximum sway, M , can be determined by the difference between the farthest reached points in both directions d and $-d$.

2) *Sway Frequency*: Within the specified ranges, a sway frequency f_r represents the number of times the direction of the bodys medial-lateral sway changes within one of these ranges $r = (0.25, 0.50, 0.75, 1.00)$ of the maximum say M in the BESS time period. Fig. 7 shows the ranges, which are considered in our experiments. High maximum sway and sway frequencies close to the maximum sway range reflect the persons instability and indicate higher possibility to fall.

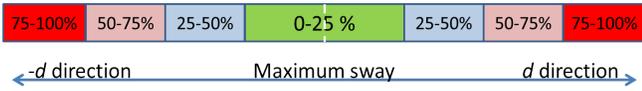


Fig. 7: Proposed sway ranges used to calculate sway frequencies

D. Age Group Modeling using GMM

Given the sway measurements $x = (M, f_{0.25}, f_{0.50}, f_{0.75}, f_{1.00})$ for each postural sway signal, two Gaussian mixture models (GMMs) are built to model the distributions of the input data, X . The GMM is a generative approach that can be used to estimate the likelihood of a sway signal pertinence to a specific age group. A GMM is parameterised by the weights of the components, $W = (w_1, w_2, \dots, w_K)$, where K is the number of components, and mean, μ and co-variance, Σ . Each component has a mean μ^k and co-variance Σ^k . These parameters are learned using the common Expectation-Maximisation (EM) algorithm.

We used the GMM in a semi-supervised manner, where the ground truth signals, which are belonging to the athlete participants are modelled with the participant age labels as they are considered to be under one group of age. The rest of the data for the elderly people are fitted in an unsupervised manner.

The probability of a sway signal belongs to a specific age group (C_k) is given as following:

$$p(C_k|x) = \frac{w_k \mathcal{N}(x|\mu_k, \Sigma_k)}{\sum_{i=1}^K w_i \mathcal{N}(x|\mu_i, \Sigma_i)} \quad (4)$$

where

$$\mathcal{N}(x|\mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^K |\Sigma_k|}} \exp\left(-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)\right) \quad (5)$$

and $\sum_{k=1}^K w_k = 1$. Based on the sway measurements, we classify the subject using the maximum probability of Eq. 4 to classify his/her sway signal to belong to either the athlete group or to one of the elderly age groups. Then, we compare the estimated age group of the subjects with their real age to check whether the subjects are classified in the same age group as they are or less, which is considered as being in a *low risk* group, or in a higher age group, which is considered as a *high risk* group. This will be discussed in more details in the experiments section.

V. EXPERIMENTAL RESULT

A. Sway measurements

Using the ground truth part of the dataset, the proposed GRU-RNN model is trained on force plate (as the target) signal and the corresponding tracked signal from the frontal view camera (as the source signal) to transfer it to the force plate measurement (*i.e.* in mm). Then, for each subject in the elderly part of the dataset, the trained GRU-RNN model is used to predict the sway signal for this subject. The proposed RNN model is composed of a GRU unit with 100 hidden

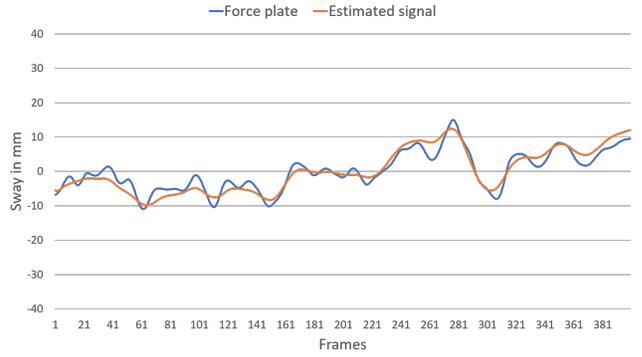


Fig. 8: Example of estimated sway signal from video and the corresponding force plate signal.

units, followed by a 'timedistributed' fully connected layer with 100 units. The Stochastic Gradient Descent (SGD) with a learning rate of 0.1 and momentum of 0.9 is used to train the model.

In Fig. 8, we show example of the estimated sway signals from video and their corresponding target force-plate instances. This qualitatively shows the robustness of the GRU based proposed method to estimate a force-plate likewise signals, only from video footage.

Sway measurements, maximum sway M and sway frequencies in the different ranges f_r , are calculated for each subject in the two parts of the dataset, the ground truth and the elderly.

B. Risk detection and analysis

The sway measurements are fed into Gaussian Mixture Models to model the distributions of the sway signals for the athlete group and the elderly groups. For the athlete participants, we consider them as in one group (healthy) and the elderly are considered in three age groups (low risk, normal and high risk). The first GMM for the healthy group has one mixture, where the second GMM for the elderly has three components.

The GMMs for the elderly people are built in unsupervised manner using the Expectation Minimization (EM) approach. The age groups of elderly are sorted based on their distance from the distribution of the ground truth GMM. We use the normalised Euclidean distance between the mean of the ground truth GMM component and the mean of each component of the elderly mixtures.

The testing was performed on the elderly people only by leave-one-subject-out for testing and merging the others with the ground truth signals for sake of training.

We use the well-known leave-one-out scheme for training and evaluating the proposed models, which guarantees testing the models on unseen data for sake of generalisation.

Comparing the real age group of the elderly subject with the age group cluster, which this subject has been classified into, this subject can be identified as high risk, increased risk, normal, reduced risk, or low risk based on the distance to the athlete age group.

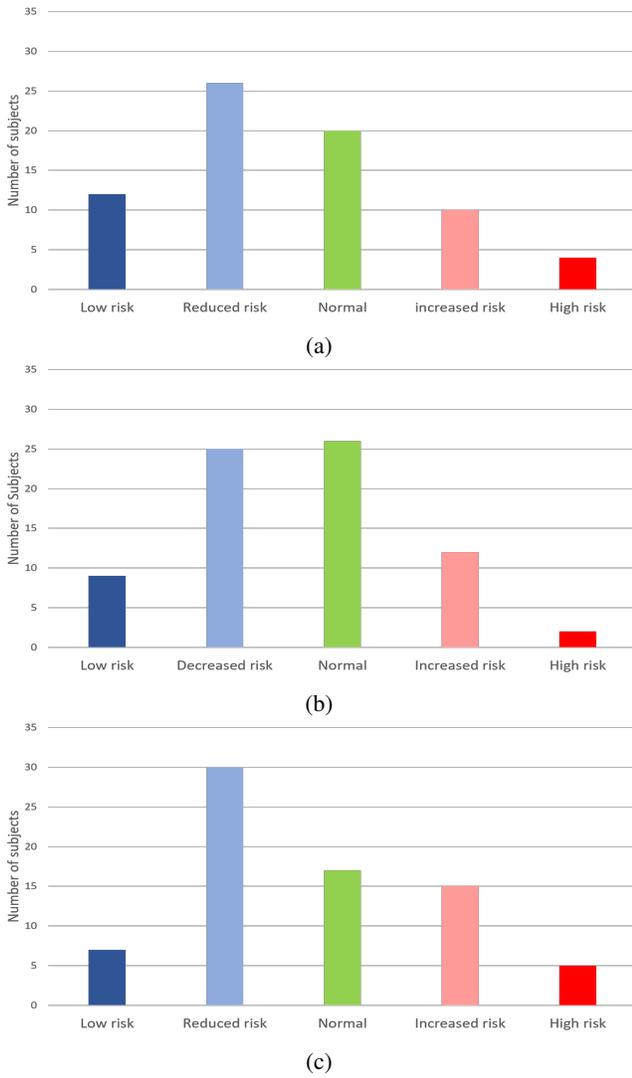


Fig. 9: Elderly subjects classification over the different risk groups for the (a) **Double leg** stance, (b) **Single leg** stance, and (c) **Tandem** stance based on the sway signal predicted from the video and the real age of the subject

Fig. 9a illustrates the risk analysis applied on the double legs stance of the BESS test. In the double legs stance, most of the subjects classified as normal to low risk depends on their sway. In double stance, maintaining balance is easier than the single leg stance, as shown in Fig. 9b, or tandem stance as shown in Fig. 9c, where a higher number of subjects classified into the increased and high risk groups.

Fig. 10 shows the predicted sway signals from the proposed GRU-RNN model for a selected subjects from each risk group, low risk, normal, and high risk. For each stance of the BESS test, double legs stance (Fig. 10a), single leg stance (Fig. 10b), and tandem (Fig. 10c), the differences in the maximum sway and the sway frequencies are obvious between the risk groups, especially in single leg and tandem stances.

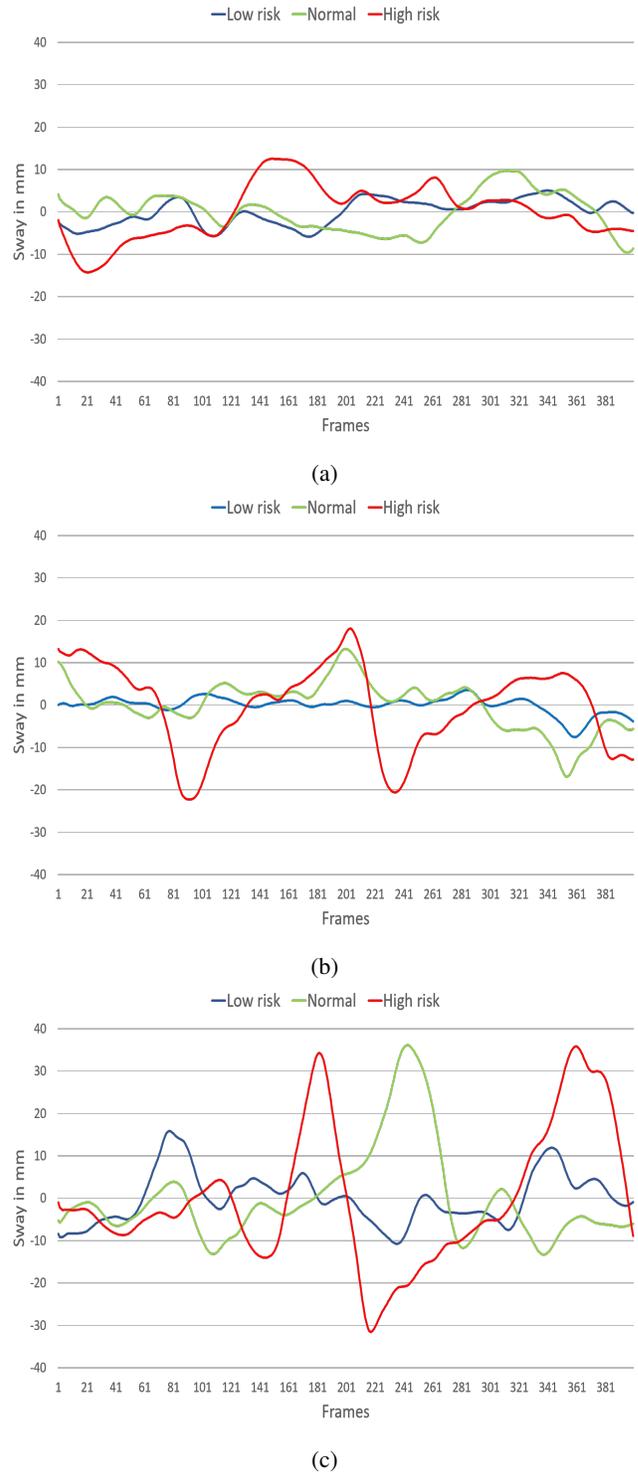


Fig. 10: Sway signals for selected subjects from the low risk, normal, and high risk in the (a) **Double leg** stance, (b) **Single leg** stance, and (c) **Tandem** stance

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we present an algorithm to estimate the age groups of elder people based on the extracted sway measurements. These measurements are derived via learning the postural sway signal from video data. Then, the person's real age and the estimated age group are used to analyse the risk of fall for the elder people. The experiments show the effectiveness of the proposed method to detect increased fall risks for elderly based on their estimated sway signals. This study present the first automated approach to detect the fall risk from video data based on postural sway measurements. This work can be extended by tracking more body joints to be more representative for the body sway. Also, repeated phases of data collection are required to study the changes of the sway measurements over time.

REFERENCES

- [1] A. Shumway-Cook, D. Anson, and S. Haller, Postural sway biofeedback: its effect on reestablishing stance stability in hemiplegic patients, *Archives of physical medicine and rehabilitation*, vol. 69, no. 6, pp. 395400, 1988.
- [2] P. A. Gribble and J. Hertel, Effect of lower-extremity muscle fatigue on postural control, *Archives of physical medicine and rehabilitation*, vol. 85, no. 4, pp. 589592, 2004
- [3] A. M. Herran and B. Garcia-Zapirain and A Mendez-Zorrilla, "Gait analysis methods: an overview of wearable and Non-wearable systems, highlighting clinical applications", *Sensors*, vol. 14(2), pp. 33623394, 2014
- [4] T. Wollseifen, Different methods of calculating body sway area, *Pharmaceutical Programming*, vol. 4, no. 1-2, pp. 91106, 2011
- [5] H. Ismail and I. Radwan and H. Suominen and G. Waddington and R. Goecke, "Human Postural Sway Estimation from Noisy Observations", In 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), pp. 454-461, 2017.
- [6] S. B. OSullivan, T. J. Schmitz, and G. Fulk, "Physical rehabilitation", FA Davis, 2013.
- [7] T. Teranishi, I. Kondo, S. Sonoda, Y. Wada, H. Miyasaka, G. Tanino, W. Narita, H. Sakurai, M. Okada, and E. Saitoh, Validity study of the standing test for imbalance and disequilibrium (side): Is the amount of body sway in adopted postures consistent with item order?, *Gait & posture*, vol. 34, no. 3, pp. 295299, 2011.
- [8] J. T. Finnoff, V. J. Peterson, J. H. Hollman, and J. Smith, Intrarater and interrater reliability of the balance error scoring system (bess), *Physical Medicine and Rehabilitation (PM&R)*, vol. 1, no. 1, pp. 5054. 2009.
- [9] J. Wang and M. She and S. Nahavandi and Abbas Kouzani. "A review of vision-based gait recognition methods for human identification", *International Conference on Digital Image Computing, Techniques and Applications (DICTA)*, pp. 320327, 2010
- [10] Mubashir. M and Shao. L, and Seed. L." A survey on fall detection: Principles and approaches", *Neurocomputing*, 100:144152, 2013
- [11] B. D. Lucas, T. Kanade, et al., An iterative image registration technique with an application to stereo vision, in *IJCAI*, vol. 81, pp. 674679, 1981.
- [12] S. Butterworth, On the theory of filter amplifiers, *Wireless Engineer*, vol. 7, no. 6, pp. 536541, 1930.
- [13] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, Learning phrase representations using rnn encoder-decoder for statistical machine translation, *arXiv preprint rXiv:1406.1078*, 2014.
- [14] I. Sutskever, O. Vinyals, and Q. V. Le, Sequence to sequence learning with neural networks, in *Advances in neural information processing systems*, pp. 31043112, 2014.
- [15] D. Bahdanau, K. Cho, and Y. Bengio, Neural machine translation by jointly learning to align and translate, *arXiv preprint arXiv:1409.0473*, 2014.
- [16] H. Ismail, I. Radwan, H. Suominen, G. Waddington, and R. Goecke, Human Postural Sway Estimation from Noisy Observations, 12th IEEE International Conference on Automatic Face & Gesture Recognition FG 2017, pp.454-461