Proceedings of Machine Learning Research 126:1–26, 2021 Machine Learning for Healthcare

Person Identification And Tinetti Score Prediction Using Balance Parameters : A Machine Learning Approach To Determine Fall Risk

Firstname Lastname NAME@EMAIL.EDU
Department of ML and Health Research
University
City, State, Country

Firstname Lastname NAME@EMAIL.EDU
Department of ML and Health Research
University
City, State, Country

Editor: Editor's name

Abstract

This paper presents a Machine Learning approach using sensor data from a Smart Floor aimed at addressing a substantial health problem among the elderly population namely falls, a major cause of accidental home deaths. Studies show approximately one-third of community-dwelling people over 65 years of age will experience one or more falls each year. Balance and walking patterns are useful indicators to determine the risk of a fall in an individual and are highly influenced by several parameters and conditions. A deterioration in the balance and walking stability of an individual can occur because of natural processes related to aging or as a result of various underlying health conditions, fatigue, changes in muscle tone, or impaired balance. The Tinetti test is widely used to assess the gait and balance in elder adults to determine the perception of balance and stability during daily activities and fear of falling. It is considered a good indicator of the fall risk of an individual. In this research, we aimed to provide a new way for fall risk reduction and early detection of the onset of chronic health conditions by creating a Machine Learning model for predicting Tinetti scores based on foot pressure data arising in common everyday activities. The goal is for this to help improve eldercare through constant monitoring and by reducing the white-coat syndrome that inhibits clinical examinations.

This paper mainly focuses on designing algorithms to extract balance parameters from the quiet standing instances arising during normal everyday life and to build individualized models capable of differentiating normal or abnormal patterns for an individual. A variety of time and frequency domain parameters are build based on the center of pressure (CoP) values obtained from time-series data from a pressure monitoring floor sensor. A classification model is build using a support vector machine for distinguishing 30 individuals based solely on these balance parameters. Further, using these parameters, a regression model is built to predict the balance and the Tinetti score of an individual which is used to predict the fall risk. This novel approach for Tinetti score prediction by just using balance analysis could be used in isolation or in combination with other assessments to provide information to the individual or care providers in order to assess health changes and to prevent falls before they happen.

1. Introduction

Falls are one of the most common problems and the leading causes of fatal and non-fatal injuries for elderly individuals around the world [40]. Falls not only threaten an individual's safety and independence but also generate enormous economic and personal cost. A fall is defined as an event which results in a person coming to rest inadvertently on the ground or floor or at another lower level. It is the second leading cause of accidental or unintentional injury or death after road traffic injuries. The rates were higher in hospitalized patients and nursing home residents. There are a number of factors that are involved in fall risks, such as weak muscles, loss of consciousness, foot problems, memory loss, confusion or difficulties with thinking or problem solving, vision and hearing problems, or taking medication that makes you dizzy or drowsy. All these factors may contribute to poor balance, causing unsteadiness on your feet, and can ultimately result in a fall. Various underlying health conditions like Parkinson's disease, Alzheimer's disease (AD) and Frontotemporal dementia (FTD), sensory abnormalities, cardiovascular diseases, and musculoskeletal disorders also contribute to poor postural stability and might result in falls [50].

Gait and balance are some of the major factors that indicate and determine a person's postural stability. A human's gait refers to an individual style of walking whereas balance is an ability to maintain the line of gravity, that is a vertical line from the center of mass of a body, within the base of support with minimal postural sway. Changes in many chronic conditions as well as short term health events and muscular control limitations often directly reflect in changes of gait and balance patterns, making them a potential early indicator of health changes. Also, disorders in gait and balance are among the most common causes of falls in older adults. They are usually multi-factorial in origin and require a comprehensive assessment to determine contributing factors and targeted interventions [40]. Balance, postural steadiness, or static posturography characterizes the performance of the postural control system in a static position and environment during quiet standing [21]. There are environmental factors that can also affect balance such as light conditions, floor surface changes, as well as short-term influences with effects on fall risk, such as alcohol, drugs, and ear infection. Of particular interest and importance are are balance impairments associated with aging. Age-related decline in the ability of the above systems to receive and integrate sensory information contributes to poor balance in older adults [4]. As a result, the elderly are at an increased risk of falls. In fact, one in three adults aged 65 and over will fall each year [5].

The Tinetti test is a traditional approach which was published by Mary Tinetti (Yale University) to assess the gait and balance in older adults and to assess the perception of balance and stability during activities of daily living and fear of falling. It is a good indicator of the fall risk of an individual and widely used in assessing balance and motor control in clinical settings. The Tinetti test has a gait score and a balance score. It uses a 3-point ordinal scale of 0, 1 and 2 in each of its contributing sections. Sections in each score categories are added and overall gait is scored over 12 and balance is scored over 16, totaling 28 as the maximum for the complete Tinetti test. The lower the score on the Tinetti test, the higher the risk of falling.

While a number of systems exist to detect falls once they happened, allowing them to contact help and avoid extended periods unconscious or injured on the floor, addressing the problem of fall prediction and prevention is more challenging and no good systems for this are in use today. One of the main reasons for this is that fall prediction requires that the person's condition and postural stability needs to be monitored continually to identify and potentially prevent a fall in future. The causes for falls do not generally appear in one day unless the fall is due to some environmental factors or unexpected accidents. Balance deterioration usually occurs over a period of time, implying that, if detected early, fall risks can potentially be mitigated, implying that falling does not have to be an inevitable result of aging. If we are able to monitor the balance characteristics of the individual over a certain period of time to analyze and capture the abnormality in balance and gait patterns in the early stage, then it will help in planning preventive strategies to avoid falls.

Although there is significant amount of research done in this area to prevent the risk of a fall[46], existing approaches have a number of limitations that make them hard to apply in practice. Firstly, most of the existing approaches for gait and balance analysis make use of devices worn on the body[45], video surveillance cameras, electrodes mounted on the skin, needles pierced into the muscle, or kinematic systems to obtain data for detection of abnormality in gait and balance. As a consequence, many of the devices currently in use require a significant amount of effort and involvement from the user and rely heavily on the willingness of the user at any specific time to use the device. Secondly, in most of these systems, human in the loop is required to carry out the clinical analysis to determine neurological or musculoskeletal disorders using the measurements obtained from these approaches or other clinical trials.

Therefore, we see a need to find a way to continuously track the person's postural stability without inhibiting the user's convenience and without the need for active human intervention to provide measurements and monitor the condition of the person. Hence, the main motivation of the research project is to prevent falls among elder individuals by providing a smart care environment which is an intelligent, sensor-driven living environment for the elderly that would be deployable to individuals' homes or independent living facilities and that can monitor relevant gait and balance characteristics unobtrusively during arbitrary in-home activities. The research presented in this paper is a part of this smart care system where the person's gait and balance characteristics are studied on a sensor-based smart floor. This paper in particular focuses on the balance characteristics of an individual and on creating a machine learning-based, automated system that helps to early detect abnormal patterns in the balance that could suggest early signs of a physical or cognitive issue.

Generalizable Insights about Machine Learning in the Context of Healthcare

This paper discusses various algorithms designed to automatically extract balance parameters from quiet standing instances arising in everyday activities in the home. The parameter values are obtained from time-series data from a pressure monitoring floor sensor. It demonstrates how Machine Learning classification models can be leveraged to distinguishing 30 individuals based solely on these balance parameters. Furthermore, it shows how another Machine Learning-based regression model can be built to differentiate normal or abnormal patterns for an individual and to predict the balance and the Tinetti score of an individual unobtrusively from everyday activity data, which, in turn, can be used to predict fall risk and initiate fall mitigation activities. This represents a novel approach for Tinetti score prediction by just using automatically (and potentially continuously) assessed balance parameters.

2. Related Work

There is a significant amount of work to determine and extract balance parameters to evaluate the postural stability of an individual for assessing the risk of a fall.

The Research paper "Measures on Postural Steadiness" [21] aims to characterize the dynamics of the postural control system associated with maintaining balance during quiet standing and discusses various balance parameters that are extracted from the Center of Pressure(CoP). The objective of this study was to evaluate the relative sensitivity of center of pressure (CoP)-based measures to changes in postural steadiness related to age. A variety of time and frequency domain measures of postural steadiness were compared between a group of twenty healthy young adults (21-35 years) and a group of twenty healthy elderly adults (66-70 years) under both eyes-open and eyes-closed conditions. The CoP coordinate time series, in Anterior-Posterior (AP) and Medial-lateral(ML) direction, are commonly used to compute measures of postural steadiness in this paper. There are a small number of other papers that performed experimentation using similar balance features but using different approaches and with different goals.

The Research paper 'A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population' [24] conducted a study of postural balance and risk of falling in an ambulatory and independent elderly population. These balance tests were performed on 100 volunteers (aged 62-96) and falling was monitored prospectively over a one-year period. The balance testing comprised measurements of: (a) spontaneous postural sway, (b) induced anterior-posterior sway, (c) induced medial-lateral sway, (d) anticipatory adjustments preceding volitional arm movements, (e) timed one-leg stance, and (f) performance on a clinical balance assessment scale. Small pseudorandom platform motions were used to perturb balance in the induced-sway tests. Using force plates, the spontaneous- and induced-sway responses were quantified in terms of the amplitude, speed, and mean frequency of the center-of-pressure displacement; input-output models were also used to parameterize the induced-sway performance.

Another paper, "The contribution of postural balance analysis in older adult fallers" [27] also uses similar balance parameters. The main objective of this research was the identification of postural characteristics of older adults at risk of falling using both static and dynamic postural balance assessments. The research claims that center of pressure (CoP) path length, CoP velocity and sway in medial lateral-and anterior-posterior direction are the variables that distinguish older adult fallers from non-fallers.

In our research we have used a number of similar balance parameters but with a different approach at extracting them from sensor data as well as with a different goal, namely to automatically form individualized models from pressure floor data during common day activities using Machine Learning and to use them to predict Tinetti scores in order to automatically monitor balance characteristics and with it potential changes in health and fall risk. A detailed explanation of features used and mathematical algorithms formulated to extract these features from raw time series data for this research is provided in Section 3.2

In addition to balance feature characterizations, there is also a significant amount of research on person identification based on gait parameters using a variety of wearable sensors, computer vision, or foot pressure information [4]. Most of the approaches for person identification from gait characteristics are based on video data [4]. For instance , C. Ben-Abdelkader et al [51] perform person identification from spatio-temporal features extracted from a video. The features they extracted from the walk were estimated stride length and cadence. With data of 17 individuals the person was verified with an error rate of 11 and correctly identified with a probability of 40% .

In another Research, Liang Wang et al [52] proposes a technique for person identification based on spatial-temporal silhouette analysis by background subtraction on a video. Principal component analysis was applied to reduce the dimensionality of time-varying distance signals. Supervised learning was performed on a lower dimensional Eigenspace to recognize individuals.

But when it comes to the person identification task using balance parameters, there are no significant research experiments in the current literature.

When it comes to fall prediction, there are various approaches to assess the risk of a fall in elderly individuals, including the Tinetti Score.

The research paper 'Evaluation of balance in fallers and non-fallers elderly' [30]. This study was designed to identify balance impairments associated with falling with a purpose to evaluate the balance between fallers and non-fallers amongst the elderly. This study reports comparative results of Computerized Dynamic Posturography (CDP) and Berg Balance Scale (BBS) tests carried out on either group among which Group I consist of 15 elderly subjects who are reported to have experienced two or more unexpected falls during the past 12 months and Group II which includes elderly people that are non-fallers. A simple predictive model was reported using logistic regressions that combined the Berg Balance Scale (BBS) scores with a selfreported history of imbalance to predict the risk of falls.

Over the past two decades, many clinical balance examinations have been developed for evaluating human balance ability. For example, the Berg Balance Scale (BBS), the Timed Up and GO Test (TUGT) and the Short Physical Performance Battery (SPPB). Recently, the above mentioned examines were further used to probe into the relationship between balance ability and cognitive function.The Timed Up and Go test is a fast and reliable diagnostic tool. Persons who have difficulty or demonstrate unsteadiness performing the Timed Up and Go test require further assessment, usually with a physical therapist, to help elucidate gait impairments and related functional limitations. The most effective strategy for falls prevention involves a multi-factorial evaluation followed by targeted interventions for identified contributing factors. Evidence on the effectiveness of interventions for gait and balance disorders is limited because of the lack of standardized outcome measures determining gait and balance abilities[53].

These approaches mostly require human in the loop for the analysis and manual experimentation and computation of the score. As a result, these techniques are applicable mainly in more clinical settings and can assess changes in risk only at intervals representing successive doctor's visits with corresponding assessments. Currently, there are not many significant experiments where this process of Tinetti score prediction or the prediction of any of the related scores is automated using Machine learning based techniques and moved into a setting where assessment can be performed during common everyday activities rather than by performing a pre-determined, scripted sequence of movements.

The approach we presented in this paper and the solution differ from the existing approaches in multiple ways: Firstly, to extract balance parameters, our approach uses floor-mounted pressure sensors which are designed to collect data unobtrusively, over long periods of time, and without interfering with gait or inconveniencing the user.

Secondly, this paper tries to provide a novel approach for person identification based on the balance parameters extracted from CoP coordinates. The experimentation is performed on the smart floor, but instead of gait, this research mainly focuses on corresponding standing segments individuals to extract the features corresponding to balance stability.

Finally, in this paper, we tried to provide a novel approach for Tinetti score predictions by using machine learning based models predicting these scores based on balanced parameters extracted for Individuals. These scores will help to identify the individuals with high risk of fall in future.

3. Methods

In this paper we introduce a system that utilizes data from a low cost, pressure-sensitive smart floor during everyday activities to perform person identification as well as to predict Tinetti balance and overall scores solely based on data corresponding to easy to identify activity episodes where the person is standing still. For this, we utilize a sensor grid that is embedded in the floor and collects data at a rate of 25Hz with a spatial resolution of 1 square foot. The data from the previously calibrated sensor floor is then processed by removing outlier noise data and by extracting the overall center of pressure (CoP) and the corresponding total pressure at that point to provide a time sequence of CoP and pressure values. This time series is then analyzed in the frequency domain to extract time series segments that correspond to situations where the person is standing still. These sequences form the basis for the subsequent feature engineering which provides the input to the two Machine Learning components that are aimed at person identification and Tinetti and balance score prediction, respectively. Figure 1 shows the components of the presented work as well as the information flow between them.

This diagram shows the basic components of the used methodology. Smart Floor and Preprocessing to extract data segments that correspond to the person standing here use prior work [1][2][3] and will be described for completeness at a high level in Section 3.1. Section 3.2 then introduces the approach to feature engineering taken in this work in order to provide a useful feature set for analyzing balance characteristics from CoP data. Based on this data, Section 3.3 presents the proposed methods for person identification to demonstrate the discerning ability of balance characteristics across individuals and to be able to distinguish between different individuals in order to be able to make individualized models and fall risk predictions. Section 3.4, finally, describes the proposed approach to learn to predict balance and Tinetti scores from the balance data to detect changes in individuals and allow basic assessment of fall risk level.

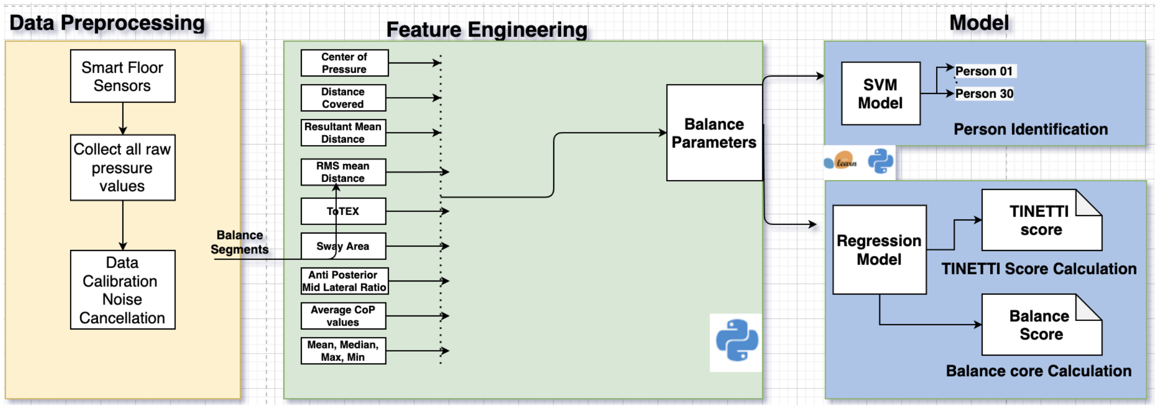


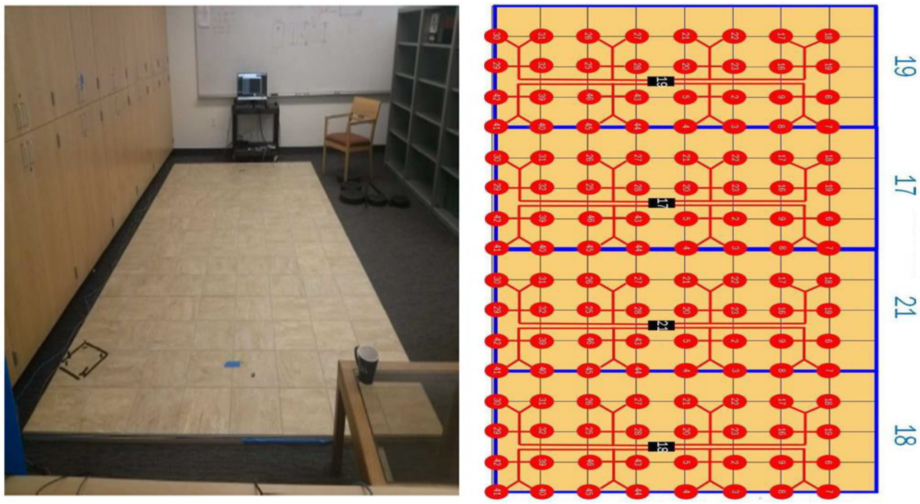
Figure 1: High Level Diagram of Methodology Components for Person Identification and Fall Risk Prediction from Pressure Floor Data

3.1. Smart Floor Pressure Data and Activity Segmentation

For data collection a pressure-sensitive smart floor is used with a series of pressure monitoring sensors placed underneath the floor tiles to record pressure data. The Pressure exerted by a subject while performing activities like standing and walking are collected at a rate of 25Hz. Data is transmitted continuously from the floor containing 128 sensors placed under 128 tiles to a nearby computer. The size of each tile is 30 cm x 30 cm. The laid-out tiles form a grid of 8 x 16 tiles. Figure 2 shows the Pressure Sensitive Smart Floor set up in the laboratory with the Layout of sensors beneath the floor.

The data obtained from the floor was obtained with approximately pre-calibrated sensors to ensure that sensor values in different floor regions are providing pressure information in terms of a uniform measurement unit. There are a total of 128 sensors on the experimental floor. Each sensor can output a value from 0-1023. Calibration of the slope and intercept was performed using a set of standard weights that were placed on the sensor locations. A linear least square fit was then applied to obtain the calibration parameters for each sensor.

After data calibration, the pressure being exerted by the subject while standing or walking on the floor is obtained along with the weight of the tile. The weight of the tile is subtracted from the data after finding the mode for each sensor. This allows us to extract the pressure exerted by the subject on the floor. The data comprises the location coordinates x and y and the associated pressure value. The pressure value is determined by averaging the pressure over the region of activated sensors refrerred to as center of pressure(CoP). The preprocessed data which is CoP coordinates in the X and Y axis and average CoP value, has a mixture of standing, slow walking, walking and various segments. The segments are converted from the time domain to the frequency domain using multidimensional Fourier Transform to obtain the frequency spectrum along each dimension. Then similar segments were grouped together by using unsupervised Agglomerative Hierarchical clustering using spectral coherence as a similarity metric on data segments of 30 subjects. Standing segments are extracted alone from this segmented data for this research purpose.



(i) (ii)

Figure 2: (i)Pressure Sensitive Smart Floor in the Laboratory Used for Experimentation (ii) Layout of sensors underneath the floor

3.2. Balance Feature Engineering from Generic Standing Situation CoP Data

Following are the Features or Balance Parameters and the corresponding Mathematical Algorithms used in this research for building Person Identification and Tinetti score prediction Machine Learning Models. Few of these parameters are computed by taking references from previous work [1][2][3][21].

Center of Pressure Speed (CoP Speed) : The CoP speed is the rate of change of the CoP over time. It is directly proportional to the coordinates of the CoP. Hence, if the displacement of the CoP from its initial position is high, there is an increase in the CoP speed and vice versa when there is a low displacement [3]. Let T_i be the time of the i^{th} time step in a data segment with N data points, and (x_i, y_i) be the corresponding CoP coordinate. Using these, the CoP speed can be extracted using the algorithm shown below.

for $t = 0$ to N
obtain the cop coordinate (x_{t+1}, y_{t+1}) at Time T_{t+1}
obtain the Cop coordinate (x_t, y_t) at time T_t
Distance: $D_t = \sqrt{(x_{t+1} - y_{t+1})^2 + (x_t - y_t)^2}$
Compute change in time $(T_t) = dt = (T_{t+1} - T_t)$
end for
$CoP_{\text{speed}} = \sum_{t=0}^{N} \frac{D_t}{T_t}$

Center of Pressure Distance (Total Sway path): CoP Distance is the total length of the path covered by a subject over the period of time. It is directly proportional to the coordinates of the CoP. This also represents the total sway path. Let T_i be the time of the

i^{th} time step in a data segment with N data points, and (x_i, y_i) be the corresponding CoP coordinate.

for $t = 0$ to N
Obtain the cop coordinatee (x_{t+1}, y_{t+1}) at Time T_{t+1}
obtain the cop coordinate (x_t, y_t) at Time T_t
Distance: $D_t = \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}$
end for
CoP Distance $= \sum_{t=0}^{N} D_t$

Average Center of Pressure: Average Center of Pressure (CoP) is the average value of CoP for each user on the floor. With the center of pressure, a subject's balance as a measure of postural sway while a person is standing can be measured. Considering the center of pressure's x coordinate, CoP(x,), it can be calculated using the formula:

CoP$(x) = \frac{\sum_i x_i \cdot F_i}{\sum_i F_i}$

where Fi is the pressure at a sensor location xi relative to a reference point in the x-direction. Computation of the y position of the CoP is performed in the same way. And the average is computed by considering the number of tiles where is pressure value is above a threshold.

Mean Resultant Distance (MRDIST): MRDIST is the mean of the Resultant Distance(RD) and represents the average distance from the mean CoP. Below are the steps to compute MRDIST, MRDIST in AP and ML direction respectively. The parameters described in this section are the most commonly used measures of postural steadiness [32].

Mean Distance-AP (MDISTAP) and Mean Distance-ML(MDISTML) are the mean absolute value of the AP and ML time series respectively and represents the average AP distance from the mean CoP. Here N = 0 to 125 corresponding to 5 secs of a data segment.

$$MDISTAP = AP_{\text{mean}} = \frac{1}{N}\sum_{i=0}^{N-1} X[i]$$
$$MDISTML = ML_{\text{mean}} = \frac{1}{N}\sum_{i=0}^{N-1} Y[i]$$

The Resultant Distance (RD) time series is the vector distance from the mean CoP to each pair of points in the AP and ML time series.

for $t = 0$ *to* $N - 1$
$AP[i] = X[i] - AP_{\text{mean}}$
$ML[i] = Y[i] - ML_{\text{mean}}$
$RD[i] = \sqrt{AP[i]^2 + ML[i]^2}$
end for

Finally by using the above computed values The MRDIST, MRDIST in AP and ML direction are calculated as below. Figure 3 and Figure 4 shows the corresponding body sway in Anterior-Posterior (AP) and Medial-Lateral (ML) direction for one of the subjects.

$$MRDIST = \frac{1}{N}\sum_{i=1}^{N} RD[i]$$
$$MDIST_{AP} = \frac{1}{N}\sum_{i=1}^{N} |AP[i]|$$
$$MDIST_{ML} = \frac{1}{N}\sum_{i=1}^{N} |ML[i]|$$

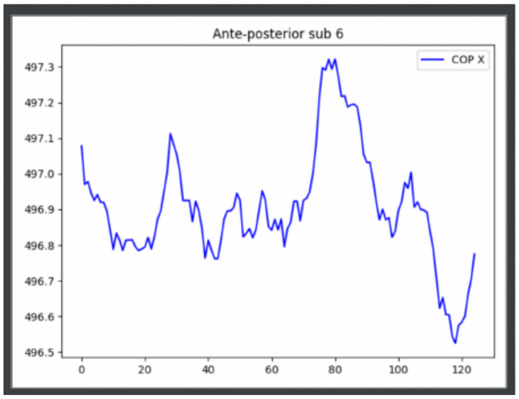


Figure 3: Anterior-posterior body sway for subject 6

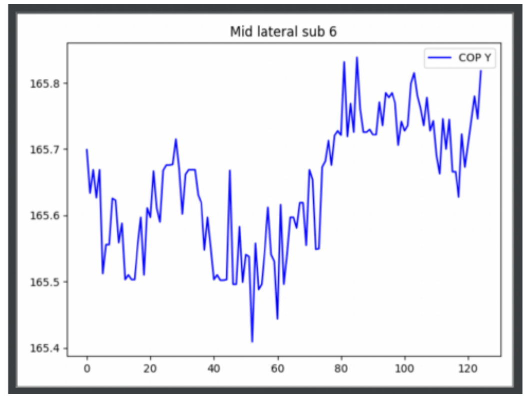


Figure 4: Medial-Lateral body sway for subject 6

RMS Distance: The RMS distance (RMSDIST) is the root mean squared distance from the mean CoP. Below Algorithm shows computation for RMS Resultant distance, RMS distance in AP and ML.

$$RMSDIST_{RD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(RD[i])^2}$$
$$RMSDIST_{AP} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(AP[i])^2} \quad RMSDIST_{ML} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(ML[i])^2}$$

Total Excursions (TOTEX): The total excursions (TOTEX) is the total length of the CoP path and is approximated by the sum of the distances between consecutive points on the CoP path. The total excursions-AP (TOTEXAP) the total length of the CoP path in the AP direction, and is approximated by the sum of the distances between consecutive points in the AP time series. Below algorithm shows computation for TOTEX, TOTEX in AP and ML.

$$\text{TOTEX} = \sum_{i=0}^{N-1}\sqrt{(AP[i+1]-AP[i])^2+(ML[i+1]-ML[i])^2}$$

$$\text{TOTEX}_{AP} = \sum_{i=0}^{N-1}|AP[i+1]-AP[i]|$$
$$\text{TOTEX}_{ML} = \sum_{i=0}^{N-1}|ML[i+1]-ML[i]|$$

Sway Area: The sway area describes the enclosed area covered by the C0P as it oscillates within the base of support. Multiple studies suggest that high Sway Area could be related to a distorted balance condition [27]. There are a number of methods explained in previous research that use sway area measures [34]. In this experimentation two methods are used to calculate the Sway Area.

Area of Stabilogram (AREA-SW): Sway Area (AREA-SW) estimates the area enclosed by the CoP path per unit of time. This measure is approximated by summing the area of the triangles formed by two consecutive points on the CoP path and the mean CoP [13]. Sway area is dependent on the distance from the mean CoP and the distance traveled by the CoP and can be conceptualized as proportional to the product of mean distance and mean velocity[27]. Figure 5 and Figure 6 shows the sway path in 3D and 2D respectively for one of the subjects.

Sway Area $= \frac{1}{2T}\sum_{i=0}^{N-1} |(AP[i+1]ML[i]) - (AP[i]ML[i+1])|$

Gauss-Green: A numerical approximation of the Gauss-Green formula is used to calculate the sway area [27].

Guass Sway Area=
$\frac{1}{4}\sum_{i=1}^{N-1} \mid ((AP[i+1]+AP[i])(ML[i+1]-ML[i]))-((AP[i+1]-AP[i])(ML(i+1)+ML(i))) \mid$

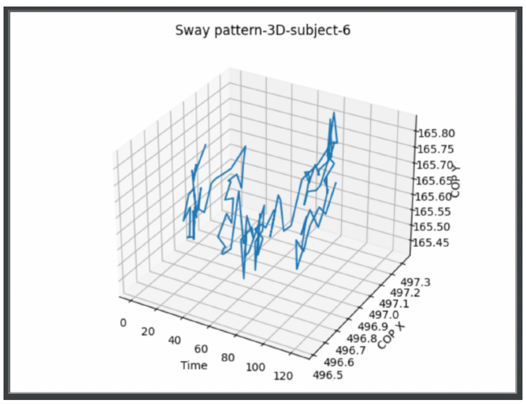


Figure 5: Sway Pattern for subject 6 in 3D space

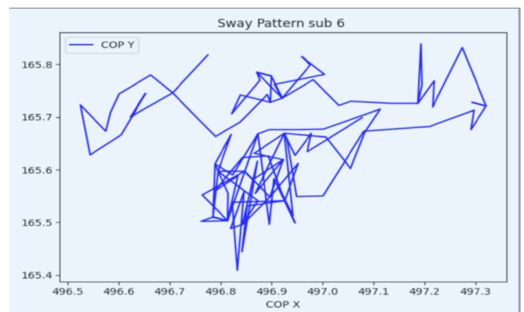


Figure 6: Sway Pattern for subject 6 in 2D space

AP-ML Ratio: The AP-ML ratio is the mean of the ratio of AP to ML. Here i=0 to N-1 .

$$\text{AP-ML Ratio } = \frac{AP[i]}{ML[i]}$$

$$\text{Mean Ratio } = \frac{1}{N}\sum_{i=1}^{N-1} \text{ AP-ML Ratio } [i]$$

3.3. Person Identification

In this research, we use a novel approach for person identification solely based on balance parameters. Person identification is basically a classification technique used to classify the different subjects based on the balance parameters (features) derived from the CoP based data. For this, the data corresponding to standing instances of 30 subjects are considered and meaningful features or balance parameters are extracted for each individual as explained in Section 3.2.

The main motivation for this component and the corresponding experimentation is to analyze how well the balance parameters that are extracted represent the balance characteristics of the subject and how well it helps to distinguish one subject from the other. While we are ultimately interested in predicting fall risks and detecting anomalies that might indicate a health change or deterioration, person identification can be seen as a test example where we can determine whether balance features for a different person can be identified as anomalies for an individual. These features will subsequently also be used to build a model for balance score prediction which is discussed in Section 3.4. Given that we have multiple classes (30 Subjects) to classify, we have used a multi-class classifier and will here use various One vs Rest Supervised Classification models for Person Identification due to their high robustness especially in the context of limited amounts of data for each class. Hyperparameter tuning is performed to achieve the best performance accuracy. The experimental results are compared in Section 5.2. Machine Learning Algorithms used for experimentation are:

1. One vs Rest Support Vector Classifiers
2. Logistic Regression

3.4. Tinetti Score Prediction

To facilitate automatic and continuous fall risk assessment, his research also introduces a novel approach for the Tinetti score prediction using a Machine Learning based model. For this, this work proposes a regression process wherein the input data is the balance features extracted as discussed in Section 3.2, and the output is the predicted balance score values and Tinetti score values. For balance score prediction, the regression model is trained on the balance score values for 30 subjects which are obtained from the experimentation process performed on the smart floor in our laboratory as described in Section 4. Similarly, the model for Tinetti score prediction is trained on Tinetti scores obtained in a similar fashion.The experimental results are discussed in Section 5.3. The Machine Learning models used for experimentation are:

1. Linear Regression
2. Lasso Regression
3. Ridge Regression
4. Support Vector Regression

4. Cohort

To build the machine learning prediction models and to evaluate the feasibility and capability of the methodology proposed here, a human subject study was performed on the laboratory's Smart Floor shown in the previous section in Figure 2. In this pilot study, 30 individuals performed a scripted sequence of activities aimed at representing common everyday activities, including standing up, walking, opening a door, standing still, and moving an item from a table on one side of the room to a location on the other side. In addition to recording the sensor data through the Smart Floor, a Tinetti test was performed for each individual and balance, mobility, and overall Tinetti scores were recorded by three independent assessors to ensure consistency of the recorded scores. This data was later used to train and evaluate the Machine Learning models.

4.1. Cohort Selection

The experimentation is performed on 30 subjects, 11 men and 19 women on a sensors-based pressure sensitive smart floor. The participants were able-bodied individuals between 18 and 72 years old, with 3 below 20, 7 between 20 and 30, 4 between 30 and 40, 3 between 40 and 50, 5 between 50 and 60, 6 between 60 and 70, and 2 above 70. All the participants were volunteers and acknowledged and signed a consent form before taking part in the study. Participants performed multiple scripted activities that were designed to generate representative balancing and walking gait data for the subject. In this, one activity was specifically designed to obtain static balancing data. As part of these activities, a Tinetti balance and gait assessment score was also elicited and a corresponding form was filled out by 3 trained investigators to assign appropriate gait and balance scores to each participant [3]. These scores are used in this research to compare the Balance and Tinetti scores obtained from Machine Learning models. The table in Figure 7 shows the obtained Tinetti and balance scores for the individuals. Due the composition of the cohort in this pilot study, which only contained relatively able-bodied individuals of different ages, the range in both scores is somewhat limited, making score prediction more sensitive. In a future study we intend to obtain data from additional individuals with age-related limitations to expand the prediction to the full range of Tinetti scores.

Subject	4	5	6	7	8	9	10	11	12	13	14	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Balance Score	15	15	15	15	15	15	15	14	13	16	15	13	13	12	16	14	15	12	15	14	14	13	13	13	14	14	14	14	14	16
Tinetti Score	27	28	27	27	27	27	27	26	25	28	27	25	25	24	28	24	27	24	27	26	26	25	25	25	26	26	26	26	26	28

Figure 7: Tinetti Balance and Overall Scores for the 30 Participants in the Pilot Study

4.2. Data Extraction

Data used for this experimentation is time series data collected using the pressure-sensitive smart floor. The data obtained from the sensors is calibrated to ensure that all sensors read pressure in a uniform way. Center of pressure (CoP) coordinates and total pressure values are then computed and form the basis for the data used for subsequent identification and segmentation of standing segments and for balance-relevant feature extraction. Using CoP coordinates and average CoP value, an unsupervised Hierarchical Agglomerative clustering algorithm is used to segment the entire data corresponding to each subject into various clusters at different hierarchy levels. The data segmented into 3 types of clusters, namely Walking, Standing and Other. All results of calibration, segmentation, and clustering are available from the previous experiments[1][2][3][4]. The research presented in this paper is built on top of the standing segment data obtained from the previous research on the same pressure sensitive smart floor. The following discussion represents the core part of the current research.

Data Transformation and Datasets:
In this paper, as we focus mainly on the standing segment data, we have extracted the standing segments for 30 subjects (subject 4 to subject 34, excluding subject 15 for which insufficient standing data was available). The original standing segment data has CoP coordinates in X and Y direction and average CoP value and has been divided into segments with 25 records, corresponding to 1 sec of data. For extraction of balance parameters, however, 1 second of data is too short to be able to extract all relevant balance parameters and thus overlapping 5 seconds data segments were constructed. These longer segments are necessary since balance features such as sway area tend to correspond to lower frequency events, thus requiring longer segments to extract and filter. The resulting new standing segments have 5 seconds of data each with 125 CoP and pressure value records. The new segments are formed by combining 5 consecutive 1 second segments and contain a 4 second overlap with the next balance segment in order to ensure sufficiently many balance data segments. For instance, subject 1 has 20 consecutive original segments, each representing 1 sec of data (representing the 20 seconds of standing data in the experiment), results in balance data segments formed in the following way:

- New segment 1 = old segments 1 to 5
- New segment 2 = old segments 2 to 6
- ...

Similarly, all data is re-segmented for all the 30 subjects. This new dataset from here onwards will be referred to as Raw CoP data which is one of the datasets that we used for features extraction.

Data Transformation Using Principal Components:
Now the dataset that we have consists of segments of 5 seconds of data with COP coordinates in X and Y directions and overall pressure values. One of the major attributes or parameters to analyze in the context of assessments of postural balance is sway. To be able to do Ante-Posterior (AP) and Mid-Lateral (ML) sway analysis, it is necessary to identify the orientation of the subject from the raw balance data. In our scenario, we have the X and Y axis of CoP coordinates, but we do not have the information which direction (axis) the standing subject is facing. Hence to overcome this problem we used Principal Component Analysis (PCA) to transform the X and Y axis into Major and Minor axis. In this process, we applied PCA on the newly segmented 5 seconds of data, and the resulting eigenvector with maximum eigenvalue is considered the Major axis and the other value the minor axis. The main rationale here is that we would generally expect sway in the AP direction to be stronger than in the ML direction as individuals are standing on both legs, thus allowing us to associate the major axis of the PCA converted data with AP and the minor axis with ML. The resulting data set will be referred to as the PCA Transformed dataset.

Frequency Domain Transformation:
The raw CoP data or the PCA Transformed data may not be feasible to be used as direct input data for Deep Learning or Machine Learning algorithms for feature extraction or classification purposes. The main reason for this lies in the limited size of the available data

set for each individual with a very limited number of data segments for each individual (around 20 data sequences per individual) and the enormous complexity of time potential CoP time series of length 125. The Cop X, Y coordinates or PCA Transformed X, Y coordinates representing the location parameters are constantly varying and themselves carry limited meaning. We, therefore, transform the data from the time domain to the frequency domain, where each data point represents the amplitude of the decomposed data sequence at a specific frequency. We used Fast Fourier Transformation for this purpose and the PCA Transformed data and the Raw CoP data are transformed to the corresponding frequency domain representation. The datasets so formed are referred to as the Frequency Domain dataset. Also, for easy reference, X, Y coordinates in Raw CoP data and Major and Minor axis in PCA Transformed Data will be referred to as AP and ML coordinate, respectively.

Now that we have these different raw spatial and frequency domain data sets, we extract the set of balance-relevant features used to capture the postural stability and used as the input in our person identification and Tinetti score prediction components.

4.3. Feature Choices

There are a variety of factors that determine balance during quiet standing, but not every parameter can be extracted from the pressure-based data that we have. Also, a number of parameters, for instance height, weight, age of subjects also contribute to the balance of a person, but we did not use these factors for our experimentation as these values are specific for each subject in our pool, and this would make the model overfit these values. We rather wanted to create a more generalized Machine Learning model for Person Identification and Tinetti Score prediction. Hence, considering all these factors and taking reference from the existing work[21], the core features listed below are used for this research. Mathematical Algorithms formulated to compute these features are discussed in Section 3.2.

1. Center of Pressure Speed
2. Center of Pressure Distance
3. Average Center of Pressure
4. Mean Resultant Distance
5. RMS Distance
6. Sway Area
7. AP ML Ratio
8. Total Excursions

Apart from this, a few supporting features, namely Maximum, Median and Mean values of the above features are also computed. However, it turned out that these supporting features did not contribute much to improve the models performance.

All the above discussed parameters or features of balance for each individual have been extracted from the AP and ML (CoP X, Y coordinates for raw data, or major and minor axis in PCA Transformed data) coordinates and CoP value obtained from the 3 datasets. Hence parameters are built and analyzed on frequency domain data and PCA Transformed data as well as on Raw CoP data in the time domain to determine the balance characteristics of individuals.

5. Results on Real Data

All the experiments are performed on balance parameters extracted using the algorithms discussed in Sections 3.2 and 4. The following 4 categories of datasets are used and results are for each type of data are compared and analyzed:

1. Features extracted from PCA Transformed Time series data
2. Features extracted from PCA Transformed Frequency domain data
3. Features combined from 1 and 2
4. Features extracted from Raw CoP data

Input data to both Person Identification and Tinetti Score prediction are feature vectors extracted from the above mentioned 4 datasets representing the same underlying movement sequences. Output data are subject 4 to subject 34 modeled as output classes 0-29 in the case of the classification task for the Person Identification problem. The classification model is trained using the subjects data corresponding to these classes. In case of the regression task for balance score prediction, the output data is the balance score and the Tinetti score of the corresponding individual. This model is trained on the balance and Tinetti score separately and output results are compared with the experimental test results which were manually determined by trained experimenters as provided in Figure 7. Both the Person Identification and the Tinetti and Balance score prediction experiments are conducted using all 30 subjects, as well as on a smaller set comprised of the first 10 subject, to observe how the model behaves with different variance in the datasets.The training and test data are shuffled and split according to a 80 - 20 ratio using a seed to reproduce the results.

In the following, the test data referred to is the data not seen in training in both Person Identification as well as in the Balance and Tinetti score prediction problems. In the case of the Balance and Tinetti score prediction experiments, two different test conditions are used; the first corresponds to the above-mentioned test set containing data not used for training from individuals that were present in the training set. The second condition, referred to as Unseen data refers to test data consisting of 20% of the subjects that were not included in the training. This case is only applicable for Balance and Tinetti score experiments since Person Identification cannot be performed on individuals that were not part of the training set.

5.1. Evaluation Approach/Study Design

To Evaluate the performance of the experiments , we considered Accuracy as the evaluation metric for classification based Person Identification Models. For the Regression based Balance Score and Tinetti Score prediction model we used Root Mean Square (RMS) Error as the evaluation metric.

As the approach used in this paper is novel for both Person Identification and Tinneti Score Prediction in that it performs these tasks based on balance features extracted from standing activities, we do not have any baseline from prior work to compare the results with. Instead we are comparing results from a range of different Machine Learning approaches on multiple feature representations for balance parameters in the spatial and frequency domain to determine the best feature and learning algorithm combination for the two problems.

5.2. Results on Person Identification

In Person Identification, the classifier used for this task are One vs Rest classifiers to address the significant number of classes in this task. Two main learning algorithms, namely Logistic Regression and Support Vector Machines (SVM) were applied on different feature representations and the results were compared.

The first training algorithm used was Logistic Regression. Though the Logistic regression model performed well on the training dataset, reaching a training accuracy of 90.86% when using combined Frequency and Time feature data, it was prone to over-fitting and performed poorly on the test dataset with the highest accuracy achieved on test date being 29.88%, again on the combined Frequency and Time feature data set. Experiments were performed by tuning the regularization parameter but still the result did not improve with the best results for each feature representation, as well as the algorithm settings shown in the table in Figure 8.

Algorithm	**Experiments**	**Accuracy**	**Time Based**	**Frequency**	**Raw**	**Frequency and Time**
SVC One-Vs-Rest	EXP1 C=100, kernel='poly', gamma=2, degree=7	Train	86.53	76.84	**80.99**	90.86
		Test	44.28	47.23	**70.47**	44.64
	EXP2 C=1000, kernel='rbf', gamma=2, degree=3	Train	70.87	65.4	67.52	84.4
		Test	47.6	50.18	**60.14**	47.23
Logistic Regression One-Vs-Rest	EXP1 :multiclass='ovr', penalty='l1', solver='saga',	Train	86.53	76.84	80.99	90.86
		Test	28.04	24.72	17.34	29.88

Figure 8: One vs Rest Logistic Regression and Support Vector Machine Performance Comparison for Various Feature Sets

Another training algorithm used in this experiment is the One vs Rest Support Vector Machine Classifier (SVC) with various kernel functions and hyperparameters. These hyperparameters are selected after performing grid search over a wide range of degree, regularization and gamma values. To further counteract overfitting, we also performed Principal Component Analysis (PCA) on the feature space to reduce the dimensions of input features and analyzed the model performance. However, there was no improvement in model accuracy using the reduced feature set but instead accuracy slightly dropped. As a result, we are not including the results for this reduced feature set here. The table in Figure 8 shows the results for the One vs Rest SVM classifier with a polynomial kernel of degree 7 and with a Radial Basis Function kernel. The listed parameters where obtained using a preliminary grid search over different parameter values and kernel degrees.

While these results still show some overfitting, especially in the case of the combined Frequency and Time features, it is significantly less pronounced than in the case of One vs Rest Logistic Regression and, in the case of the raw feature set is limited. Moreover, test accuracies for both kernel types are significantly higher that for the Logistic Regression model with the best performance and the least overfitting produced using SVM with polynomial kernels, which was consequently chosen as our classification model.

Performing additional, more fine grained grid search to further optimize the model parameters, a final One vs Rest SVM model with C=100, gamma=3, kernel=poly, degree=7, provides the highest accuracy of 74.17% on the Raw CoP dataset Test Data for the 30 class classification problem involved in person identification. To investigate the effect of the larger number of classes on accuracy, we compare the results of Raw CoP test data across the number of subjects by applying it to a reduced data set involving only 10 subjects. Doing this we can see that the percentage of correctly identified individuals increases to 81.98% for the first 10 individuals from the 74.17% for all 30 individuals. This illustrates the expected effect of an increase in the number subjects causing an increase in the probability of wrong prediction. It is important to note here that in both cases the classification accuracy is dramatically higher than the prior likelihood which in the case of 30 subjects would be 1:29 and 1:9 in the case of 10 subjects.

Figure 9 shows the results on the best performing SVM model for the different feature set options and for both 30 and 10 subjects.

Datasets	30 Subjects			10 Subjects		
	Train	Test	Train Test (avg)	Train	Test	Train Test (avg)
Time Series	93.45	45.02	69	98.19	68.47	83
Frequency Domain	83.03	45.39	64	92.29	65.77	79
Raw COP	89.21	**74.17**	82	98.19	**81.98**	90
Frequency+Time	95.85	46.49	71	97.73	60.36	79

Figure 9: Best Performing Support Vector Classification Model for Person Identification on 30 and 10 Individuals Using Different Feature Sets

5.3. Results on Tinetti Score and Balance Score Prediction

Tinetti score and balance score prediction experimentation is performed using various regression models, namely Linear Regression(LR), Lasso, Ridge Regression, and Support Vector Regression(SVR). For all models and feature choices, the regression model performance was almost the same on Training and Test data, indicating that no overfitting seemed to occur in the Balance and Tinetti prediction task.

Two types of prediction tasks were performed here, where in the first the model was trained using all individuals with a subset of the their balance data with tests being performed using test data from the same individuals. This corresponds to the situation where we want to determine and track the Tinetti score for a known individual. In the second task we investigate whether and to what degree Balance and Tinetti prediction models can transfer between individuals. Here the model was trained on the data of a subset of the subjects and used to predict the scores for completely Unseen data representing subjects whose data was not included in training. The experimentation's were again performed on the 4 feature types and on 30 subjects and on the first 10 subjects to investigate the effect of diversity of individuals on the results.

Prediction for Known Subjects:
In the first setting, data of all subjects is used for training with the test data being seprated beforehand and not seen during training. Figure 10 shows tables with the results for each of the different regression approaches and feature representation choices.

Algorithm	Experiments	Scores	RMSE	Time Based	Frequency	Raw	Frequency and Time
Support Vector Regression	**EXP1** C=100 degree=3 kernel='rbf'	Balance	Train	0.797	0.821	0.826	0.799
			Test	0.832	0.863	**0.791**	0.878
		TINETTI	Train	0.948	0.967	1.010	0.950
			Test	1.045	1.026	0.977	1.071
	EXP2 C=100 degree=6 kernel='rbf' gamma=2	Balance	Train	0.768	0.805	0.820	0.739
			Test	0.834	0.883	0.793	0.860
		TINETTI	Train	0.910	0.949	1.002	0.875
			Test	1.051	1.042	0.980	1.053
	EXP3 Unseen C=100 degree=6 kernel='rbf' gamma=2	Balance	Train	0.814	0.895	0.848	0.793
			Test	0.936	0.810	0.845	0.867
		TINETTI	Train	0.990	1.052	1.064	0.954
			Test	1.094	0.982	0.862	1.032

SVR Model

Algorithm	Experiments	Scores	RMSE	Time Based	Frequency	Raw	Frequency and Time
Lasso	**EXP1** alpha=2	Balance	Train	0.898	0.898	0.898	0.898
			Test	0.863	0.863	0.863	0.863
		TINETTI	Train	1.071	1.071	1.071	1.071
			Test	1.051	1.051	1.051	1.051
	EXP2 alpha=10	Balance	Train	0.898	0.898	0.898	0.896
			Test	0.863	0.863	0.863	0863
		TINETTI	Train	1.071	1.071	1.071	1.071
			Test	1.051	1.051	1.051	1.051
	EXP3 Unseen alpha=10	Balance	Train	0.944	0.944	0.944	0.944
			Test	0.638	0.638	0.638	0.638
		TINETTI	Train	1.150	1.150	1.150	1.150
			Test	0.635	0.635	0.635	0.635

Lasso Model

Algorithm	Experiments	Scores	RMSE	Time Based	Frequency	Raw	Frequency and Time
Linear Regression	**EXP1** Polynomial Degree=1	Balance	Train	0.829	0.844	0.823	0.809
			Test	0.833	0.850	0.801	0.858
		TINETTI	Train	1.005	1.012	0.985	0.979
			Test	1.026	1.017	0.986	1.049
	EXP2 Polynomial Degree=2	Balance	Train	0.715	0.693	0.610	0.00006
			Test	1.025	2.387	2.287	2.70
		TINETTI	Train	1.005	1.012	0.985	0.979
			Test	1.026	1.017	0.986	1.049
	EXP3 Unseen Polynomial Degree=1	Balance	Train	0.863	0.886	0.848	0.851
			Test	0.702	0.656	0.752	0.684
		TINETTI	Train	1.069	1.075	1.039	1.046
			Test	0.758	0.719	0.780	0.772

Linear Regression Model

Algorithm	Experiments	Scores	RMSE	Time Based	Frequency	Raw	Frequency and Time
Ridge	**EXP1** alpha=2	Balance	Train	0.869	0.864	0.881	0.862
			Test	0.836	0.833	0.035	0.832
		TINETTI	Train	1.038	1.028	1.052	1.029
			Test	1.010	1.001	1.016	1.004
	EXP2 alpha=10	Balance	Train	0.876	0.869	0.886	0.867
			Test	0.839	0.837	0.844	0.832
		TINETTI	Train	1.038	1.028	1.052	1.029
			Test	1.010	1.001	1.016	1.004
	EXP3_Unseen alpha=10	Balance	Train	0.919	0.911	0.923	0.909
			Test	0.622	0.621	0.655	0.625
		TINETTI	Train	1.106	1.094	1.117	1.096
			Test	0.654	0.652	0.680	0.651

Ridge Regression Model

Figure 10: Root Mean Square (RMS) Error for Different Machine Learning Algorithms and Feature Sets for Balance and Tinetti Score Prediction With 30 Individuals. Both Seen and Unseen Scenarios are Shown.

The best performing model here was the Support Vector Regression (SVR) model with Radial Basis Function kernels with an average The Root Mean Square (RMS) error for the balance score of 0.791 and for the Tinetti score of 0.977 when using C=100 and degree=3. This represents an average error of predicting the Balance and Tinetti Scores of significantly less than 1 point and thus demonstrates that capabilities of the system.

Prediction for Unseen Subjects:
Experiments were repeated where the test data was from individuals not present in the training set, referred to here as Unseen. For the predictions on Unseen data, the SVR model again performed best, yielding an average RMS error of 0.845 for the balance and 0.862 for the Tinetti score as shown in the bottom rows of the tables in Figure 10. Again, predictions of the score can be made using this model with an average error significantly below 1 point (the minimum resolution of the original test) even for individuals that have not been seen as part of training. This again suggests the potential of the presented methodology to detect changes in balance and thus predict changing fall risks.

Reduced Training Data Set:
Repeating the experiments using the data of only 10 individuals yielded somewhat expected results in that it allowed for more accurate prediction of scores for known individuals but decreased prediction accuracy for unknown individuals. Tables in Figure 11 show the experimental results using only the first 10 subjects on SVR and Ridge regression. As expected, when applied to known individuals, the model is performing better with RMS error for the balance score of 0.527 and for the Tinetti score of 0.578. This is because there is very low variation in the expected output values. That is also the reason that the model is performing bad on the Unseen data with RMS erro increasing to 1.225 and 1.267 for Balance and Tinetti prediction, respectively. Training with this few individuals leads to significant overfitting and customization to these individuals and thus other individuals scores can no longer be predicted well.

Algorithm	**Experiments**	**Scores**	**RMSE**	**Time Based**	**Frequency**	**Raw**	**Frequency and Time**
Support Vector Regression	**EXP1** C=100 degree=3 kernel='rbf'	Balance	Train	0.461	0.522	0.437	0.464
			Test	0.462	0.572	0.440	0.551
		TINETTI	Train	0.516	0.586	0.556	0.520
			Test	0.535	0.592	0.551	0.610
	EXP2_Unseen C=100 degree=6 kernel='rbf' gamma=2	Balance	Train	0.081	0.093	0.112	0.080
			Test	1.237	1.316	1.278	1.225
		TINETTI	Train	0.234	0.286	0.390	0.221
			Test	1.299	1.274	1.295	1.267

SVR Model

Algorithm	**Experiments**	**Scores**	**RMSE**	**Time Based**	**Frequency**	**Raw**	**Frequency and Time**
Ridge	**EXP1** alpha=2	Balance	Train	0.566	0.619	0.659	0.580
			Test	0.559	0.596	0.644	0.575
		TINETTI	Train	0.654	0.686	0.783	0.649
			Test	0.649	0.652	0.722	0.656
	EXP2_Unseen alpha=10	Balance	Train	0.256	0.317	0.390	0.280
			Test	1.248	1.313	1.334	1.259
		TINETTI	Train	0.437	0.479	0.597	0.426
			Test	1.272	1.311	1.383	1.282

Ridge Regression Model

Figure 11: Root Mean Square (RMS) Error for Different Machine Learning Algorithms and Feature Sets for Balance and Tinetti Score Prediction With 10 Individuals. Both Seen and Unseen Scenarios are Shown.

5.4. Upsampling the Data Set to Balance Training Data

There is some degree of imbalance in the data points corresponding to each subject which might lead to reduced prediction quality. To combat this problem, we did not want to use synthetic data. Using synthetic data might give biased results but introduces potentially incorrect data points that can yield misleading predictions. Here we wanted to analyze the results solely on real data in order to avoid incorrect data points. Hence, we tried to utilize a sampling technique on selected subjects to achieve more balanced data and thus better classification and prediction results. For person Identification, the left table in Figure 12 shows that the accuracy of the best performing SVM model on upsampled data for 11 subjects increased to an average of 92.86%.

Balanced 11 subjects SVM: c=100, Kernal=poly , gamma=3			
	Train Accuracy	Test Accuracy	Avg Accuracy
Balanced degree =3	77.99	67.5	73
Balanced degree=7	79.87	77.27	79
up sample degree 7	96.75	92.86	95

Balanced 11 Subjects		
	Train RMSE	Test RMSE
	Balance Score	
Lasso, Alpha=50	0.778	0.776
SVR C=1000, g=3, degree=3,kernel=poly	0.954	0.978
	Tinetti Score	
Lasso, Alpha=50	0.5	0.583
SVR C=1000, g=3, degree=3,kernel=poly	0.756	0.878

Figure 12: Person Identification (left) and Balance and Tineeti Score Prediction (right) on Upsampled, Balanced Data

For Balance and Tinetti score prediction, the right table in Figure 12 shows that the RMS error for the best performing Lasso model reduces to 0.778 for balance and 0.5 for Tinetti Score which, while slightly improved, is not significantly different from the results on the original 10 subject dataset.

6. Discussion

In this research we extracted parameters corresponding to balance from real time series pressure data obtained from a pressure sensitive Smart Floor. Leveraging this data we built baseline Machine Learning models for Person Identification and Balance and Tinetti score prediction based on balance parameters alone. The table in Figure 13 shows the final performance metrics achieved with this research for both Person Identification and Tinetti Score prediction.

	Person Identification			**Balance Score Prediction**			**Tinetti Score Predictions**		
Model	SVC One-Vs-Rest C=100, gamma=3, kernel=poly, degree=7			SVR model with C=100, Kernel='RBF', degree=3		Lasso , alpha=50	SVR model with C=100, Kernel='RBF', degree=3		Lasso , alpha=50
Subjects Included	30 Subjects	10 Subjects	11 Subjects Balanced Upsampled Dataset	30 Subjects	10 Subjects	11 Subjects Balanced Upsampled Dataset	30 Subjects	10 Subjects	11 Subjects Balanced Upsampled Dataset
Accuracy/ RMS Error on Test Data	**74.17**	**81.98**	**92.86**	**0.845**	**0.527**	**0.778**	**0.862**	**0.578**	**0.5**

Figure 13: Best Performing Models for Person Identification and Balance and Tinetti Score Prediction on Original Data for 30 and 10 Subjects and on Upsampled Data for 11 Subjects

These results demonstrate that Machine learning can be used in place of traditional approaches to assess balance characteristics and with this to assess fall risks. The resulting system can make these assessments continuously and unobtrusively without the need for human intervention while the subject is performing their normal activities in the home.

In the future we are planning to train these models with a wider range of individuals living in our live-in laboratory and to integrate these models in a broader system for monitoring the postural stability of the individuals that can provide early indications of health changes and fall risks and convey them to the inhabitant or a care provider.

Currently, in this research, we used only balance parameters to predict the balance scores and Tinetti scores. In the future gait parameters can also be extracted in a similar fashion and gait scores can be added to the input for the prediction model in order to further improve the system's ability to identify even minor systematic changes in a person's health status. This can further enhance fall prevention but also be potentially used to identify changes in chronic conditions ar issues arising as a consequence of a changed medication schedule.

Limitations The data that we used for this research introduces a number of limitations. Overall, the models in most of the experiments suffer from some degree of overfitting. This yields higher accuracy during training than during the actual application of the system. This effect is further enhanced by the fact that Person Identification has a significant number of classes and the data that we used has some degree of imbalance. While we conducted some experiments to balance the data using upsampling, more research into this would be useful. Moreover, a larger dataset for each individual would be beneficial to reduce the overfitting. However, in practical applications obtaining large datasets in the laboratory is difficult and we are thus aiming to collect such a significantly larger dataset from a live-in laboratory.

In case of Tinetti score prediction, we trained the model with the available subjects who have balance scores in the limited range of 12-16 and Tinetti scores in the range of 24-28. We do not have data samples for other possible values for Balance and Tinetti scores. However, the model performed still well on the used dataset. The model could be made more robust by training it with individuals' data that includes a wider range possible values of balance and Tinetti score. However, such data was not available from the pilot study. We aim to obtain more varied data from our live-in laboratory by attracting participants with a wider range of health conditions and mobility limitations.

References

[1] "References removed to preserve double-blind review. References will be provided in the final version of the paper"

[2] "References removed to preserve double-blind review. References will be provided in the final version of the paper"

[3] "References removed to preserve double-blind review. References will be provided in the final version of the paper"

[4] "References removed to preserve double-blind review. References will be provided in the final version of the paper"

[5] "References removed to preserve double-blind review. References will be provided in the final version of the paper"

[6] Pirker W, Katzenschlager R, "Gait disorders in adults and the elderly: A clinical guide. Wien Klin Wochenschr". 2017;129(3-4):81-95. doi:10.1007/s00508-016-1096-4

[7] Hall SJ, "Equilibrium and Human Movement". In: Hall SJ. eds. Basic Biomechanics, 8e New York, NY: McGraw-Hill;
[8] AMA Velayutham SG, Chandra SR, Bharath S, Shankar RG, "Quantitative Balance and Gait Measurement in Patients with Frontotemporal Dementia and Alzheimer Diseases: A Pilot Study. Indian J Psychol Med". 2017;39(2):176-182. doi:10.4103/0253-7176.203132
[9] International Journal of Nursing Sciences (Volume 4, Issue 4, 10 October 2017). Factors associated with balance function in patients with knee osteoarthritis: An integrative review
[10] Yoon, Jung; Yoon, Tae; Shin, Bo & Na, Eun. (2012), "Factors Affecting Test Results and Standardized Method in Quiet Standing Balance Evaluation". Annals of rehabilitation medicine. 36. 112-8.10.5535/arm.2012.36.1.112.
[11] The Center of Gravity and Stability: https://accessphysiotherapy.mhmedical.com/content.aspx?bookid=446&sectionid=41564591
[12] Balance. (2020, June 11). Physiopedia, .Retrieved 03:57, August 14, 2020 from https://www.physiopedia.com/index.php?title=Balance&oldid=240814.
[13] Martinez, Matthew Thomas, De Leon, Phillip L, and Keeley, David, "Anomaly Detection for Prediciting Falls Risk using Smartphone Gait Data", United States: N. p., 2016. Web.
[14] Mancini M, Salarian A, Carlson-Kuhta P, et al, " ISway: a sensitive, valid and reliable measure of postural control". J Neuroeng Rehabil. 2012;9:59. Published 2012 Aug 22. doi:10.1186/1743-0003-9-59
[15] Martina Mancini, John G. Nutt, Fay B, "Horak,Balance Dysfunction in Parkinson's Disease,Academic Press", 2020,Pages xi-xiii, ISBN 9780128138748,https://doi.org/10.1016/B978-0-12-813874-8.11001-3.
[16] Schoneburg B, Mancini M, Horak F, Nutt JG, "Framework for understanding balance dysfunction in Parkinson's disease". Mov Disord. 2013 Sep;28(11) 1474-1482. doi:10.1002/mds.25613. PMID: 23925954; PMCID:PMC4001822.
[17] Martina Mancini,Laurie King ,Fay Horak , "ISWAY: An Objective Test of Balance Control". DOI:https://doi.org/10.1016/j.apmr.2012.08.171
[18] Dr. Fay Horak and Dr. Martina Mancini - APDM https://www.apdm.com/webinar-with-dr-fay-horak-and-dr-martina-mancini/
[19] Marchetti GF, Whitney SL, Redfern MS, Furman JM, "Factors associated with balance confidence in older adults with health conditions affecting the balance and vestibular system". Arch Phys Med Rehabil. 2011;92(11):1884- 1891. doi:10.1016/j.apmr.2011.06.015
[20] Hylton B. Menz, Meg E. Morris, Stephen R. Lord, "Foot and Ankle Characteristics Associated with Impaired Balance and Functional Ability in Older People", The Journals of Gerontology: Series A, Volume 60, Issue 12, December 2005, Pages 1546–1552, https://doi.org/10.1093/gerona/60.12.1546
[21] T. E. Prieto, J. B. Myklebust, R. G. Hoffmann, E. G. Lovett and B. M. Myklebust, "Measures of postural steadiness: differences between healthy young and elderly adults," in IEEE Transactions on Biomedical Engineering, vol. 43, no. 9, pp. 956-966, Sept. 1996, doi: 10.1109/10.532130.
[22] Jorstad EC, Hauer K, Becker C; Lamb SE; ProFaNE Group, "Measuring the psychological outcomes of falling: a systematic review". J Am Geriatr Soc. 2005;53(3):501–510.85
[23] Juras, Grzegorz; Kajetan, Słomka; Fredyk, Artur; Sobota, Grzegorz & Bacik, Bogdan. (2008). "Evaluation of the Limits of Stability (LOS)". Balance Test. Journal of Human

Kinetics. 19. 39-52. 10.2478/v10078-008- 0003-0.
[24] Maki BE, Holliday PJ, Topper AK, "A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population". J Gerontol. 1994;49(2):M72-M84. doi:10.1093/geronj/49.2.m72
[25] Greene B, R, Doheny E, P, Walsh C, Cunningham C, Crosby L, Kenny R, A, "Evaluation of Falls Risk in Community-Dwelling Older Adults Using Body-Worn Sensors". Gerontology 2012; 58:472-480. doi: 10.1159/000337259
[26] Nissim M, Livny A, Barmatz C, et al. Effects of aquatic physical intervention on fall risk, working memory and hazard-perception as pedestrians in older people: a pilot trial. BMC Geriatr. 2020;20(1):74. Published 2020 Feb 19. doi:10.1186/s12877-020-1477-4
[27] Pizzigalli L, Micheletti Cremasco M, Mulasso A, Rainoldi A, "The contribution of postural balance analysis in older adult fallers: A narrative review." J Bodyw Mov Ther. 2016;20(2):409-417. doi:10.1016/j.jbmt.2015.12.008
[28] Baloh RW, Fife TD, Zwerling L, et al, "Comparison of static and dynamic posturography in young and older normal people. J Am Geriatr Soc". 1994;42(4):405-412. doi:10.1111/j.1532-5415. 1994.tb07489.x
[29] Piirtola M, Era P, "Force platform measurements as predictors of falls among older people - a review". Gerontology. 2006;52(1):1-16. doi:10.1159/000089820
[30] Banu Mujdeci, Songul Aksoy, Ahmet Atas, "Evaluation of balance in fallers and non-fallers elderly", Brazilian Journal of Otorhinolaryngology Volume 78, Issue 5,2012, Pages 104-109, ISSN 1808- 8694,https://doi.org/10.5935/1808-8694.20120016.
[31] Puszczalowska-Lizis E., Bujas P., Omorczyk J., Jandzis S., Zak M, "Feet deformities are correlated with impaired balance and postural stability in seniors over 75". (2017) PLoS ONE, 12 (9) , art. no. e0183227
[32] Hertel J, Olmsted-Kramer LC, Challis JH, "Time-to-boundary measures of postural control during single leg quiet standing. J Appl Biomech". 2006;22(1):67-73. doi: 10.1123/jab.22.1.67
[33] Ruhe A, Fejer R, Walker B, "Center of pressure excursion as a measure of balance performance in patients with non-specific low back pain compared to healthy controls: a systematic review of the literature". Eur Spine J. 2011;20(3):358-368. doi:10.1007/s00586-010-1543-2
[34] Thomas Wollseifen, i3, Wiesbaden, Germany, "Different Methods of Calculating Body Sway Area 2011".
[35] GyuTae Kim1, Mohammed Ferdjallah2, Gerald F. Harris, "Fast Computational Analysis of Sway Area Using Center of Pressure Data in Normal Children and Children with Cerebral Palsy ", ISSN: 1937-9080
[36] Lemay, J., Gagnon, D.H., Nadeau, S. et al, "Center-of-pressure total trajectory length is a complementary measure to maximum excursion to better differentiate multidirectional standing limits of stability between individuals with incomplete spinal cord injury and able-bodied individuals". J NeuroEngineering Rehabil 11, 8 (2014). https://doi.org/10.1186/1743-0003-11-8
[37] Sulaiman A. Alshammari, Abdullatif M. Alhassan, Matar A. Aldawsari, Faisal O. Bazuhair, Fahad K. Alotaibi, Ahmed A. Aldakhil, Faroq W. Abdulfattah, "Falls among elderly and its relation with their health problems and surrounding environmental factors in Riyadh". J Family Community Med. 2018 Jan-Apr; 25(1): 29–34. doi: 10.4103/jfcm.JFCM_48_17

[38] Thomas Wollseifen, InVentiv Health Clinical, Eltville, Germany , “Determination of Body Sway Area by Fourier Analysis of its Contour “, PhUSE 2013
[39] Alshammari SA, Alhassan AM, Aldawsari MA, et al, “ Falls among elderly and its relation with their health problems and surrounding environmental factors in Riyadh”. J Family Community Med. 2018;25(1):29- 34. doi:10.4103/jfcm.JFCM_48_17
[40] BROOKE SALZMAN, MD, “Gait and Balance disorders on older adults”. Thomas Jefferson University, Philadelphia, Pennsylvaniahttps://www.aafp.org/afp/2010/0701/p61.html
[41] Tomohisa Yamamoto, Charles E Smith, Yasuyuki Suzuki, Ken Kiyono, Takao Tanahashi, Saburo Sakoda, Pietro Morasso, Taishin Nomura Physiol. “Universal and individual characteristics of postural sway during quiet standing in healthy young adults”. Rep. 2015 Mar; 3(3): e12329. Published online 2015 Mar 16. doi: 10.14814/phy2.12329
[42] Christopher K. Rhea, “A new measure of the CoP trajectory in postural sway: Dynamics of heading change”. http://www.cog.brown.edu/research/ven_lab/Publications/Rhea_CoP_HeadingChange_MEP14.pdf (December 2013)
[43] Lamyaa Sadouk, “CNN Approaches for Time Series Classification”. Submitted: May 7th 2018Reviewed: August 27th 2018Published: November 5th 2018.DOI: 10.5772/intechopen.81170
[44] Tinetti ME, Williams TF, Mayewski R, “Fall Risk Index for elderly patients based on number of chronic disabilities”. Am J Med 1986:80:429-434
[45] Hsu, Yu-Liang; Chung, Pau-Choo; Wang, Wei-Hsin; Pai, Ming- Chyi; Wang, Chun-Yao; Lin, Chien-Wen; Wu, Hao-Li & Wang, Jeen- Shing. (2014). “Gait and Balance Analysis for Patients With Alzheimer’s Disease Using an Inertial-Sensor-Based Wearable Instrument”. IEEE journal of biomedical and health informatics. 18. 1822-1830. 10.1109/JBHI.2014.2325413.
[46] Camille J. Shanahan, Frederique M. C. Boonstra, L. Eduardo Cofré Lizama, Myrte Strik, Bradford A. Moffat, Fary Khan, Trevor J. Kilpatrick, Anneke van der Walt, Mary P. Galea, Scott C. Kolbe. “Technologies for Advanced Gait and Balance Assessments in People with Multiple Sclerosis”.Front Neurol. 2017; 8: 708. Published online 2018 Feb 2. doi: 10.3389/fneur.2017.00708
[47] B. J. Benda, P. O. Riley and D. E. Krebs, ”Biomechanical relationship between center of gravity and center of pressure during standing,” in IEEE Transactions on Rehabilitation Engineering, vol. 2, no. 1, pp. 3-10, March 1994, doi: 10.1109/86.296348.
[48] Brian a. Blanksby, Graeme a. Wood, and Leonard freedman. Human kinesiology. Department of human movement & recreation studies (ba.b., ga.w); department of anatomy and human biology (bab., l.f.), Nedlands, western Australia, 6009, Australia
[49] Thomas Muehlbauer, Albert Gollhofer, Urs Granacher , “Associations Between Measures of Balance and Lower-Extremity Muscle Strength/Power in Healthy Individuals Across the Lifespan: A Systematic Review and Meta- Analysis”. Sports Med. 2015; 45(12): 1671–1692. Published online 2015 Sep 28. Doi: 10.1007/s40279-015-0390-z
[50] Selva Ganapathy Velayutham, Sadanandavalli Retnaswami Chandra, Srikala Bharath, Ravi Girikamatha Shankar. “Quantitative Balance and Gait Measurement in Patients with Frontotemporal Dementia and Alzheimer Diseases: A Pilot Study”. Indian J Psychol Med. 2017 Mar-Apr; 39(2): 176– 182. doi: 10.4103/0253-7176.203132
[51]C. BenAbdelkader, R. Cutler and L. Davis, ”Stride and cadence as a biometric in automatic person identification and verification,” Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 2002, pp. 372-377.

doi: 10.1109/AFGR.2002.1004182
[52]Liang Wang, Tieniu Tan, Huazhong Ning and Weiming Hu, ”Silhouette analysis-based gait recognition for human identification,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pp. 1505-1518, Dec. 2003. doi: 10.1109/TPAMI.2003.1251144