Nyon-data, a Fall Detection Dataset from a Hinged Board Apparatus

Rogério Pais Dionísio[1,2][0000-0002-6810-2447], Ana Rafaela Rosa[1][0000-0002-9472-1954] and Cassandra Sofia dos Santos Jesus[1][0000-0003-2649-9375]

[1] Polytechnic Institute of Castelo Branco, 6000-084 Castelo Branco, Portugal
[2] DiSAC – Digital Services, Applications and Contents, 6000-767 Castelo Branco, Portugal
rdionisio@ipcb.pt

Abstract. This study aims to create a dataset that collects the data from a 3-axis acceleration sensor fixed on an apparatus that mimics a human fall event, The resulting dataset includes data from different angle positions and heights, corresponding to joints of the lower limbs of the human body (ankle, knee, and hip). We use the dataset with a threshold-based fall detection algorithm. The result from the Receiver Operating Curve shows a good behavior with a mean Area Under the Curve of 0.77 and allow to compute a best threshold value with False Positive Rate of 14.8 % and True Positive rate of 89.1 %.

Keywords: ROC, Dataset, Fall Detection Algorithm, 3-axis Accelerometer

1 Introduction

During the COVID-19 pandemic, elderly individuals have experienced heightened levels of isolation, leading to detrimental effects on their mental and physical well-being. This topic has become particularly important, especially in more developed countries such as in Europe, where the highest rates of aging are recorded globally. In 2021, Portugal ranked fourth on the list of the oldest countries in Europe, only surpassed by Italy, Greece, and Finland [1], with an aging index of 182.1 [2]. Aging is a natural process that happens continuously and inevitably. It involves many biological, psychological, and social changes, some of which can result in losses or gains [3]. These changes can lead to a decrease in function, which makes older people more vulnerable. This vulnerability increases the likelihood of developing health problems and, as a result, increases the risk of falling [4].

Falls are one of the leading causes of death and illness among elderly people worldwide. In the population over 65 years of age, falls can cause numerous hospitalizations, institutionalizations, and premature deaths, which contribute to the loss of functional capacity, increased consumption of health resources and, above all, a loss of quality of life [5, 7]. In Portugal, the prevalence of fall-related injuries rises to 76% in the 65-74 age group, reaching 90% in the group of people aged 75 and over [8].

Because Castelo Branco district boasts one of the highest aging rates in the interior of Portugal (242.9) [9] and given that the National Police reported the presence of 1840 elderly individuals living in isolation within the district [10], the project ZELAR@CB was proposed in 2019 to monitor elder falls using wearable devices. But with the lockdown measures taken during the pandemic, all contacts with elder people were canceled. The restrictions imposed to curb the spread of the virus, such as social distancing and lockdown measures, have significantly impacted scientific projects involving elders in various ways.

Many scientific projects, as ZELAR@CB, require face-to-face interactions. Social distancing measures and restrictions on physical contact have limited the ability to conduct in-person data collection and interactions with older participants. Moreover, the pandemic has made it difficult to recruit and retain voluntary participants due to concerns about their vulnerability to the virus. Thus, researchers have had to adapt their methods, exploring innovative approaches to continue their scientific projects involving older adults.

1.1 Objectives and structure of the paper

This study aims to create a dataset that collects the data from a 3-axis acceleration sensor fixed on an apparatus that mimics a human fall event and test it with a threshold-based fall detection algorithm. The resulting dataset includes data from different angle positions and heights, corresponding to joints of the lower limbs of the human body (ankle, knee, and hip).

The present work is divided into 5 sections: the introduction section presents a framework and rationale of the paper's theme and the objectives of this work; then the experimental setup used in the production of the Nyon dataset is described, followed by the description of the dataset. Finally, the results and respective discussion will be presented, and we finalize with the main conclusions.

2 Experimental setup

This section contains the methods used to collect the data. The materials used will be described, along with a description of how they were used during the data collection, in terms of hardware and the software developed for the data gathering procedure. Then, a description of the experimental setup is provided. Lastly, the experimental procedure's details are given.

2.1 Datalogging Electronic System

The purpose of the datalogging electronic system is to collect, process, and store data from a 3-axis digital accelerometer. The system we develop is made of off-the-shelf low-cost devices found in the laboratory, and consists of an accelerometer device, a data acquisition unit that converts analog signals into digital format, a processing unit, and a storage component such as a memory card. The wiring circuit was produced using Fritzing and is depicted in Fig. 1. In details, the bill of material is:

- An Arduino UNO or equivalent board;
- A 3-Axis digital accelerometer (ADXL345);
- A micro-SD card reader breakout board (Adafruit);
- A push-button to control the start and completion of the data collection process;
- A green LED to signal when the SD-card is recording;
- A red LED to signal that the program stopped successfully;
- Several resistors (220 Ω) for current limitation of the LED and the push-button;
- An electronic breadboard and conductive wires.

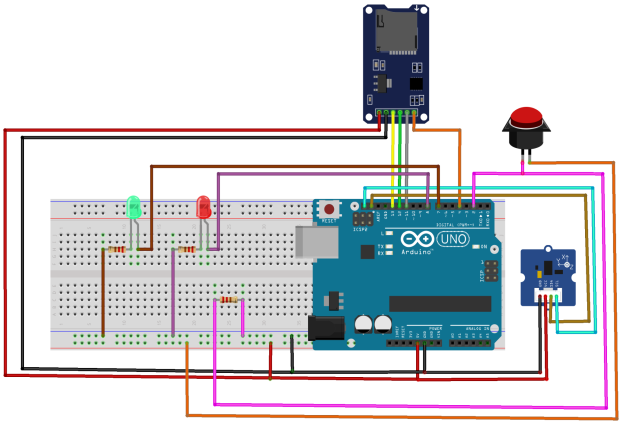

Fig. 1. Datalogger electronic system to measure and store acceleration measurements.

The program for the Arduino UNO simplifies the datalogging procedure, using a push-button to start and stop the data recording process, and two LED to inform the user about the state of the recording process. The pseudocode is described in Fig. 2.

```
Green LED OFF and Red LED ON
int measure = 1
int inic_timestamp = micro()
int elapsed_time
forever
  if(button pressed && measure == 1)
      measure = 0
      Green LED ON and Red LED OFF
      inic_timestamp = micro()
      read(AccX,AccY,AccZ)
      elapsed_time = micro() - inic_timestamp
      record_SD(elapsed_time,AccX,AccY,AccZ)
  end if
  if(button pressed && measure == 0)
      measure = 1;
      Green LED OFF and Red LED ON;
  end if
```

Fig. 2. Pseudo-code program of the datalogger system

2.2 Hinged Board

The sensor was taped in the surface of a hinged board, as seen in Fig. 3, at different height relative to the base, to emulate the position of the ankle, knee, and hip of a human body. The values were estimated from anthropometric proportions between the height of an adult human body (H) and the relative height of the ankle (0,039H), knee (0,285H) and hip (0,53H). Considering the Portuguese population, the mean heigh for men (172,9 cm) and women (163 cm) were averaged (168 cm), and the results are: Ankle height: 7 cm; Knee height: 48 cm; Hip height: 89 cm.

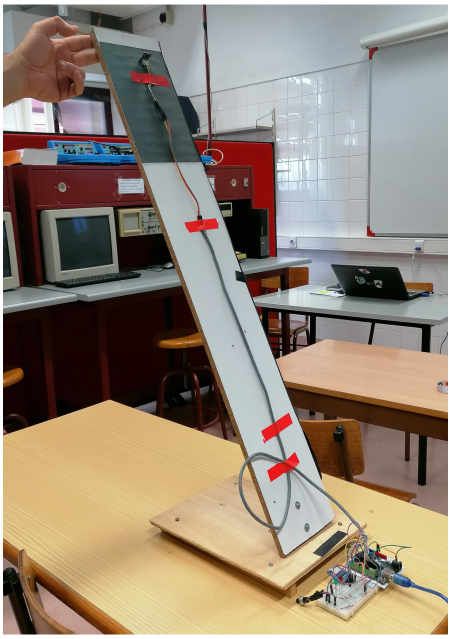

Fig. 3. Experimental apparatus to collect acceleration measurements in different scenarios (relative angle position and elevation of the accelerometer from the base).

3 Dataset Description

This section contains the information regarding the data included in the dataset. First the different positions and angles of the hinged board considered for the data gathering process, and finally a description of the organization of the data, from the different files to the format of the data in each file. The dataset is available on GitHub (https://github.com/rdionisio1403/Nyon/) under a Creative Commons (CC0) license.

3.1 Dataset Format Description

This dataset's files use the CSV format as it is widely used and can easily be formatted by researchers if necessary. The dataset is contained in one folder. The folder

Nyon_data contains 9 files, and the name of the files is a combination the height of the sensor (Ankle, Knee, or Hip) and the falling angle (30º, 45º or 90º). As an example, the file "knee_30.csv" contains acceleration measurements, with the sensor placed on the hinged board at the Knee height (48 cm) and with starting angle of 30º. Each of the file's content is organized as presented in Table 1.

Table 1. Common structure of a dataset file.

Time (s)	Accel. X axis (g)	Accel. Y axis (g)	Accel. X axis (g)
0	X Value 1	Y Value 1	Z Value 1
0,02	X Value 2	Y Value 2	Z Value 2
0,04	X Value 3	Y Value 3	Z Value 3
…	…	…	…

Each file is composed of four columns. The first column gives the time in second, the second, third and fourth columns give the measured acceleration in X, Y and Z axes, respectively, in g units (one g is equal to 9.8 m/s^2). Each of the files contains 3 sequential emulated fall events, separated by a period without any movement, which ultimately results in 27 recorded fall events. The sampling interval is 20 ms and was defined from the hardware and software implementation of the datalogger described in Section 2.1. With these files, an image with the timeline acceleration on three axes can be generated, as seen in Fig. 4.

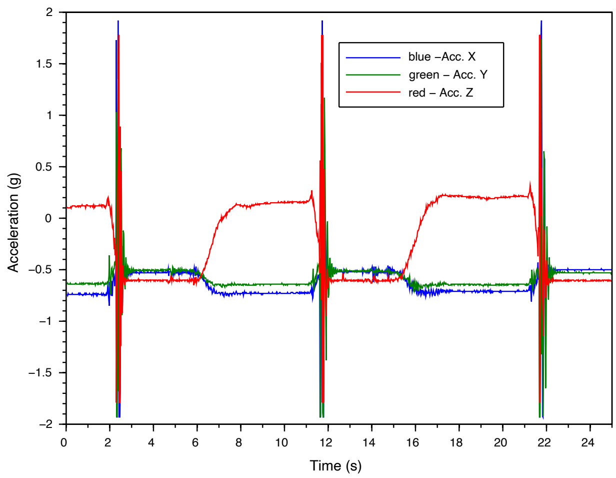

Fig. 4. Acceleration measurements in three axes, during three fall events with the hinged board starting at an angle of 45º, with the sensor at the ankle height (ankle_45.csv).

4 Results and Discussion

This section provides an overview of the evaluation process conducted for the dataset and its reliability to implement fall detection algorithms. The evaluation encompasses three key sub-sections: Fall detection algorithms, Receiver Operation Curve (ROC) computation, and Discussion. Each sub-section investigates the specific aspects of the evaluation process. Scylab-2023.1.0 is the scientific tool we use to evaluate the dataset.

4.1 Fall Detection Algorithm

Threshold-based algorithms are commonly used in fall detection wearable devices to detect falls based on predefined thresholds of motion sensor data. These algorithms rely on the assumption that the motion patterns associated with falls are different from normal activities of daily living (ADLs) and can be detected by setting appropriate thresholds on sensor data.

Accelerometer-based threshold-based algorithms are widely used in fall detection wearable devices [12] and are easy to implement with low computational complexity. These algorithms analyze the acceleration data from the accelerometer sensor, which measures the change in velocity over time, to detect falls. Acceleration data is typically analyzed in three axes: X-axis (horizontal), Y-axis (vertical), and Z-axis (perpendicular to the other two axes). The Differential Sum Vector Magnitude (A_{DSVM}) [13], is often used as a feature for fall detection.

$$A_{DSVM}(n) = \sqrt{(x_n - x_{n-1})^2 + (y_n - y_{n-1})^2 + (z_n - z_{n-1})^2} \quad (1)$$

where x, y and z are the three independent variables from the acceleration axes and n denotes the sample number.

Fig. 5 shows the computed A_{DSVM} of one dataset file (ankle_90_degree.csv), represented with a red line. Three fall events are clearly visible. The blue line sequence, represent the true positive fall event that was timed during the dataset recording. The results show a good agreement between the estimate and the true fall events.

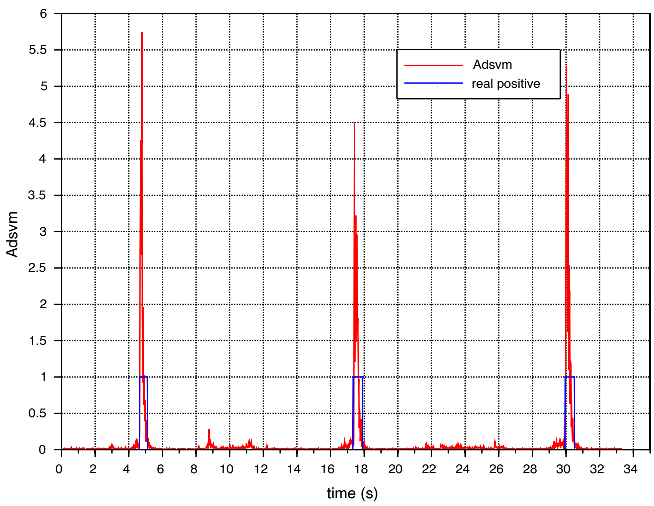

Fig. 5. A_{dsvm} computed from the dataset file ankle_90.csv.

4.2 ROC computation

Receiver Operating Characteristic (ROC) curve is a graphical plot that displays the trade-off between sensitivity and specificity for different threshold values in a binary classification problem, such as fall detection. ROC curves are generated by plotting the True positive Rate (TPR) on the y-axis against the False Positive Rate (FPR) on the x-axis, at different threshold values,

$$TPR = \frac{TP}{TP+FN} \tag{2}$$

$$FPR = \frac{FP}{FP+TN} \tag{3}$$

where TP stands for True Positive, FN is False Negative, FP is False Positive, and TN mean True Negative. The algorithm to compute TP, TN, FP and FN is presented in Fig. 6. `y_real` and `y_pred` are binary vectors (0 or 1) representing the real positive values (as seen in the blue line in Fig. 5) and the predicted positive values, respectively. The variable `adsvm` is a vector of the computed A_{DSVM}. `maxThreshold` is a constant value above the maximum value from the `adsvm` vector, for the resulting ROC curve to be fully represented. For the present case, the value 8 was used.

```
TP = FP = TN = FN = 0
L = length(adsvm)
indx = 1
for th = 0 to maxThreshold
   for k = 1 to L
    if adsvm(k) > th
       y_pred(k) = 1
    else
       y_pred(k) = 0
    end if
   end for
   for k = 1 to L
    if (y_real(k) + y_pred(k)) == 2
       TP = TP + 1
    end if
    if (y_pred(k) = 1) and (y_real(k) ≠ y_pred(k))
       FP = FP + 1
    end if
    if (y_real(k) + y_pred(k)) == 0
       TN = TN + 1
    end if
    if (y_pred(k) == 0) and (y_real(k) ≠ y_pred(k))
        FN = FN + 1
    end if
   end for
   TPR(indx) = TP / (TP + FN)
   FPR(indx) = FP / (FP + TN)
   indx++
end for
```

Fig. 6. Pseudo-code for the ROC parameters computation.

Threshold computation using ROC analysis involves finding the optimal threshold value that maximizes the sensitivity and specificity trade-off. This can be done by selecting the threshold value that corresponds to the point on the ROC curve closest to the top-left corner, which represents the perfect balance between sensitivity and specificity (i.e., the highest sensitivity with the lowest false positive rate), as represented by the code snippet from Fig. 7,

```
ampl = TPR-FPR;
best_ampl = find(ampl==max(ampl));
best_th = vect_th(best_ampl);
```

Fig. 7. Pseudo-code of the threshold computation.

The white dot visible in Fig. 8 is the optimal threshold (0.35), when TPR = 90.5 % and FPR = 18.6 %.

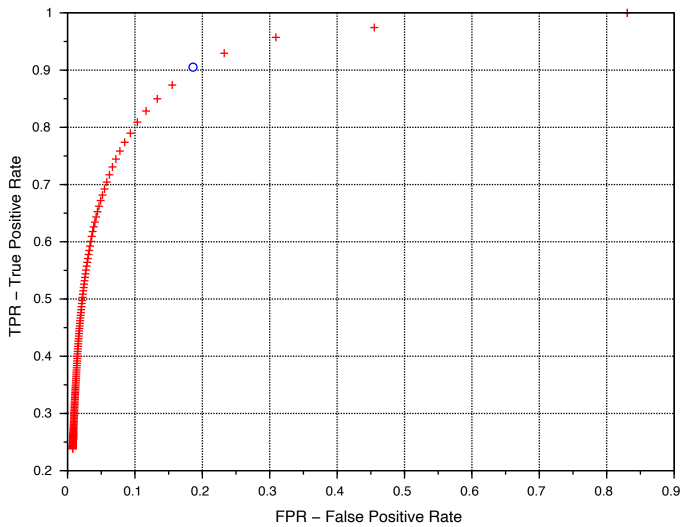

Fig. 8. Receiver operating curve of the fall detection device (hankle_90.csv).

4.3 Discussion

ROC analysis can provide valuable insights into the performance of threshold-based algorithms in fall detection devices, since it allows for the assessment of the algorithm's FPR and TPR at different threshold values, helping to determine the optimal threshold value for a particular population or use case. First, we assess the Area Under the Curve (AUC) for each ROC, as a measure of the performance of a binary classification model. The results from Table 2 shows high values of the AUC, with a mean value of 77.2 %, that indicates a good-performing model for all scenarios.

Using the algorithm from Fig. 7, we compute the best threshold from each dataset file, and presented the value together with the resulting FPR and TPR. The results from

Table 2 show that the best A_{dsvm} threshold value for the specific fall detection algorithm varies between 0.20 and 0.40, depending on the height and falling angle of the fall sensor. The mean value (0.23) can be used to test the algorithm for different fall events, although this value is not optimum for all cases. However, the mean FPR and TPR, respectively 14.8 % and 89.1 %, are good indicator as it concerns to the correctness and effectiveness of the algorithm.

Table 2. Classification model results for each dataset

Dataset file	AUC	A_{dsvm} threshold	FPR	TPR
hankle_30.csv	83.1 %	0.35	13.6 %	92.5 %
hankle_45.csv	79.0 %	0.30	14.4 %	90.0 %
hankle_90.csv	76.0 %	0.20	18.6 %	90.5 %
knee_30.csv	72.4 %	0.35	11.7 %	87.7 %
knee_45.csv	62.8 %	0.25	14.0 %	78.5 %
knee_90.csv	72.0 %	0.40	10.0 %	99.8 %
hip_30.csv	87.0 %	0.30	15.0 %	93.7 %
hip_45.csv	82.3 %	0.40	15.2 %	81.0 %
hip_90.csv	80.3 %	0.25	20.3 %	88.2 %
Mean value	**77.2 %**	**0.23**	**14.8 %**	**89.1 %**
Std Deviation	**7.3 %**	**0.07**	**3.16 %**	**6.43 %**

5 Conclusions

The process of aging is a natural and ongoing phenomenon that cannot be avoided. It involves a combination of biological, psychological, and social changes, which can result in both positive and negative outcomes. As people grow older, they may experience a decline in physical and cognitive abilities, making them more susceptible to various health issues. This increased vulnerability significantly raises the likelihood of developing diseases and, consequently, the risk of experiencing falls. Thus, the monitoring of vital parameters can improve the safety, mobility, communication, and monitoring of the elderly [14]. Some authors argue that a prompt response after a fall can significantly reduce the number of inherent consequences [15].

The analysis from the dataset files reveals that the best threshold value for the proposed fall detection algorithm (A_{dsvm}) ranges between 0.20 and 0.40, depending on the height and falling angle of the fall sensor. However, the optimal threshold value may vary depending on the specific population, activity patterns, and environmental conditions, which may require further customization and validation in real-world settings. The dataset available online, other fall detection algorithms can be evaluated and compared.

References

1. EUROSTAT. (2021). Archive: Estrutura populacional e envelhecimento. EUROSTAT. https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Population_structure_and_ageing/pt&oldid=510113 , last accessed 2021/12/20.
2. PORDATA. (2021). Indicadores de envelhecimento segundo os Censos. PORDATA. https://www.pordata.pt/Portugal/Indicadores+de+envelhecimento+segundo+os+Censos++-525 , last accessed 2022/06/20.
3. Moreira, M. J. (2020). Como envelhecem os portugueses: envelhecimento, saúde, idadismo. (1ª ed.). Fundação Francisco Manuel dos Santos.
4. Silva, D., Brasileiro, M., & Souza, D. (2018). Relação entre envelhecimento da população e o risco de quedas: revisão integrativa. Revista Científica de Enfermagem RECIEN, 8(23), 28-38. Disponível em https://www.recien.com.br/index.php/Recien/article/view/258
5. Agudelo-Botero, M., Giraldo-Rodríguez, L., Murillo-González, J., Mino-León, D., & CruzArenas, E. (2018). Factors associated with occasional and recurrent falls in Mexican community-dwelling older people. PLoS ONE 13(2), 1-12. doi: https://doi.org/10.1371/journal.pone.0192926
6. World Health Organization. (2018). Falls. Disponível em https://www.who.int/newsroom/fact-sheets/detail/falls
7. Sherrington, C., Fairhall, N., Wallbank, G., Tiedemann, A., Michaleff, Z., Howard, K., Clemson, L., Hopewell, S., & Lamb, S. (2019). Exercise for preventing fall in older people living in the community: an abridged Cochrane systematic review. Br J Sports Med 2020, 54, 885-891. http://dx.doi.org/10.1136/bjsports-2019-101512
8. Coimbra, V., Marques, E., & Chaves, C. (2019). Prevalência de Quedas em Idosos Residentes numa Comunidade Rural. Millenium, 2(3), 109-116. https://doi.org/10.29352/mill0203e.09.00218
9. PORDATA. (2023). Índice de envelhecimento - Onde há mais e menos idosos por 100 jovens?. https://www.pordata.pt/municipios/indice+de+envelhecimento-458
10. GNR. (2022). Operação “Censos Sénior 2022”- Balanço. Disponível em https://www.gnr.pt/noticias.aspx?linha=20474
11. Luna-Perejón, F., Muñoz-Saavedra, L., Civit-Masot, J., Civit, A. & Domínguez- Morales, M. (2021). AnkFall – Falls, Falling Risks and Daily-Life Activities Dataset with an Ankle-Placed Accelerometer and Training Using Recurrent Neural Networks. Sensors, 21(5). doi: https://doi.org/10.3390/s21051889
12. Kim, N. H., & Yu, Y. S. (2013). Fall recognition algorithm using gravity-weighted 3-axis accelerometer data. *Journal of the Institute of Electronics and Information Engineers*, *50*(6), 254-259.
13. Igual, R., Medrano, C., & Plaza, I. (2013). Challenges, issues and trends in fall detection systems. *Biomedical engineering online*, *12*(1), 1-24.
14. dos Santos Jesus, C. S., Rosa, A. R., & Dionísio, R. P. (2023, March). Nyon: A Ubiquitous Fall Detection Device for Elders. In *Innovations in Smart Cities Applications Volume 6: The Proceedings of the 7th International Conference on Smart City Applications* (pp. 175-184). Cham: Springer International Publishing.
15. GONÇALVES, Fábio [et al.] (2021) - IoT system to monitor the well-being of senior citizens who self-isolate during the pandemic. In European Conference on Networks and Communications and 6G, Porto, 8-11 de junho - EuCNC & 6G Summit.