Multichannel Sensor Analysis for Fall Detection: Evaluating Accelerometer, Gyroscope, and Magnetometer Data in Real-World Settings

Aanika Desai, Mieko Chun

23 August 2025

Abstract

Falls are a leading health risk among older adults and individuals with conditions such as Ataxia, PSP (Progressive Supranuclear Palsy), and Dementia. Accurate fall detection can reduce the time for at risk individuals to receive medical care and mitigate fall-related injuries. Many studies use simulated data, which limits generalizability as models trained on simulated data are not representative of real-world falls. While recent work has shown that time-series models trained on accelerometer data achieve strong results, there is a lack of information on how additional measurements such as gyroscope and magnetometer affect model performance. This study evaluates whether multi-channel sensor data improves fall detection performance. Using the FARSEEING dataset, we analyzed accelerometer, gyroscope, and magnetometer fall signals. We trained tabular models with and without feature extraction, as well as time-series models, and compared performance across acceleration-only and multi-channel inputs. Accelerometer-only models appeared to outperform multi-channel approaches, particularly in the time-series setting. The best-performing model, QUANT, created with acceleration-only data, achieved an F1 score of 0.99 ± 0.02. This finding demonstrates that accelerometer data is sufficient for accurate fall detection using time-series models. Our results further suggest that accelerometer-based systems may be the most practical option for future fall data collection.

1 Introduction

Falls are typically caused by a sudden loss of balance and represent one of the most serious health risks for older adults. According to one study, falls account for approximately 50% of all injuries that lead to hospitalization [source]. They are the second leading cause of unintentional injury-related deaths worldwide [source], and their frequency also places additional stress on both individuals and healthcare systems. The risk of falling is higher in older individuals and those with health conditions such as ataxia, progressive supranuclear palsy (PSP), dementia, and Parkinson's disease. For these reasons, accurate and reliable fall detection is a developing area of research, offering the potential to provide assistance during falls and reduce the physical and mental effects.

Traditional fall detection methods face several challenges. Because real-world falls are relatively rare and difficult to capture in controlled settings, many studies rely on simulated falls to train machine learning models. Simulated falls are typically performed by young, healthy volunteers, and as a result, they do not accurately depict the intricacies of real-world falls [1]. Therefore, models trained exclusively on simulated data often cannot be generalized to real individuals. Prior research has also explored the use of tabular models with feature extraction, with reasonable success. However, recent works suggest that time-series models may provide a more effective way to detect fall events. In particular, Aderinola et al. demonstrated that time-series models achieve high performance when applied to real-world fall data, establishing a foundation for our work [1].

Despite these advances some limitations still remain. Many existing studies utilize simulated data and while acceleration is primarily used to create models, there is not a field-wide accepted measurement established. Various wearable sensors are capable of capturing data in addition to the traditional acceleration data, including gyroscope and magnetometer signals. Combining these measures could improve model performance by providing a more complete picture of body movement. Another challenge is minimizing false alarms and missed falls. Missing a fall can result in severe health consequences, while excessive false alarms can negatively impact day-to-day life. Addressing these challenges requires the development of robust machine learning models with consideration of their impact on individuals that suffer from falls.

In this work, we investigate whether incorporating gyroscope and magnetometer channels, in addition to accelerometer data, improves fall detection. We focus on participants whose data contains all three measurements to compare the performance of acceleration-only models to models that use all three channels. Signal data was collected from wearable sensors placed on the lower back, near the L5 vertebra, or thigh, both of which capture motion from the body's center of mass. Data was recorded at 100 Hz which accurately captures the body's movements by recording 100 measurements per second. Accelerometers capture linear acceleration, gyroscopes measure angular velocity, and magnetometers track orientation relative to the Earth's magnetic field. All together, these measures create a comprehensive depiction of body movement. Using the FARSEEING dataset, we conducted various experiments: (1) applying tabular models directly to raw data, (2) testing feature extraction methods with tabular models, and (3) evaluating time-series classification models. Each of these approaches was tested both with acceleration-only data and with the all three channels.

While our initial hypothesis was that incorporating gyroscope and magnetometer channels would improve performance, we found that acceleration-only models achieved the best results. In particular, the QUANT time-series model trained on accelerometer data reached an F1 score of 0.99 ± 0.02, an AUC of 1.0, a false alarm rate of 0.0, and a miss rate of 0.03. This strong result is consistent with past findings [1]. Additionally, these results show that time-series methods can reliably detect falls using only accelerometer data, reducing the need for additional sensor channels or complex feature engineering. This finding has practical implications for future data work, suggesting that simple sensor data may be sufficient for robust fall detection.

2 Literature Review

2.1 *Accurate and Efficient Real-World Fall Detection Using Time Series Techniques* [1]

In *Accurate and Efficient Real-World Fall Detection Using Time Series Techniques* [1], Aderinola et al. explore how to accurately predict falling, a common issue faced by older adults. The growing older population and in turn number of incidents of falls pose real-world consequences to healthcare systems as falls often require emergency services and medical attention. Detecting falls in a timely manner can reduce strain on the healthcare industry and alleviate fears that older adults may have of falling. The researchers' goal was to build a robust model to detect falls that exhibits lower false positive and false negative rates than existing models. Additionally, they aimed to explore the impact of using real-world fall data instead of simulated data. While simulated fall data is more available than real-world data, the researchers questioned its usability to predict actual falls. To approach this issue, the researchers analyzed one dataset using real-world data, *FARSEEING*, and two datasets using simulated data, *FallAllD* and *SisFall*. To identify incidences of falls, the researchers looked at features that measured each subject's acceleration in three directions. To account for differences in measurement across the three datasets, the researchers' combined the three directional features into a single feature, the univariate magnitude. They then segmented data into incidents of falls and ADLs (activities of daily living) for training and testing by setting an acceleration threshold to identify times of interest. This approach allowed for the creation of models with efficient runtime that were also easily interpretable. After analyzing baseline tabular models, the researchers developed sophisticated time-series models and analyzed performance primarily based on F_1 score which combines precision and recall measures. Their analyses showed that the time series models consistently outperformed the tabular models across each of the three datasets. Their analyses also showed that models trained on simulated data showed a drop in performance when tested on real world data. The results demonstrate the importance of training models to predict falling on data that includes real-world data.

2.2 *Review of fall detection techniques: A data availability perspective [2]*

Falls are generally caused by the sudden loss of balance, and according to one study, accounts for 50% of all injuries that lead to hospitalization. However, the paper highlights the idea that despite the high need for fall detection, there is not sufficient data to address this issue. Although there are many strategies for fall detection depending on sensors, methods, and features used, they are based on having sufficient data on falls. The authors review past literature, and find that the majority of fall detection is currently using simulated falls, which is not as effective or generalizable to real life falls. To address this, they created a taxonomy to help determine what strategies can be used in fall detection based on how much fall training data is available. Category I occurs when sufficient fall data exists, typiclly due to simulated falls, and algorithms can be trained on the fall data itself. In this scenario, viable solutions include supervised learning and the use of thresholds. When fall data is insufficient, several strategies are proposed based on the literature reviewed: over- or under-sampling to balance the training set, applying semi-supervised learning, using cost-sensitive classification, or using OCC techniques. This second classification is more realistic and can also be highly effective. In addition, models can also predict falls without actual fall data. Instead, falls can be defined as abnormalities from normal behavior by some models. This strategy works best when accurate normal activity data is collected, specifically with a combination of vision-based sensors and wearable sensors. This review advocates for a shift toward fall detection methods that rely on using normal activities or those that are realistic to use for insufficient fall data.

3 Materials and Methods

3.1 Dataset

We analyzed data from the FARSEEING dataset which consists entirely of real-world fall data. For the purpose of analyzing the impact of magnetometer and gyroscope data, we only kept signals that contained accelerometer, gyroscope, and magnetometer measurements. This left us with 116 signals from 38 participants.

Subject ID	Falls	Sex	Age	Setting	Condition	Height (cm)	Weight (kg)
93807530	3	F	57	Community Dwelling	Ataxia	167	70
00002186	5	M	70	Community Dwelling	PSP	167	82
42542799	2	M	86	Assisted Living	Dementia	NA	NA
87486959	3	M	51	Community Dwelling	Ataxia	169	77
08332163	2	F	40	Community Dwelling	Ataxia	169	77

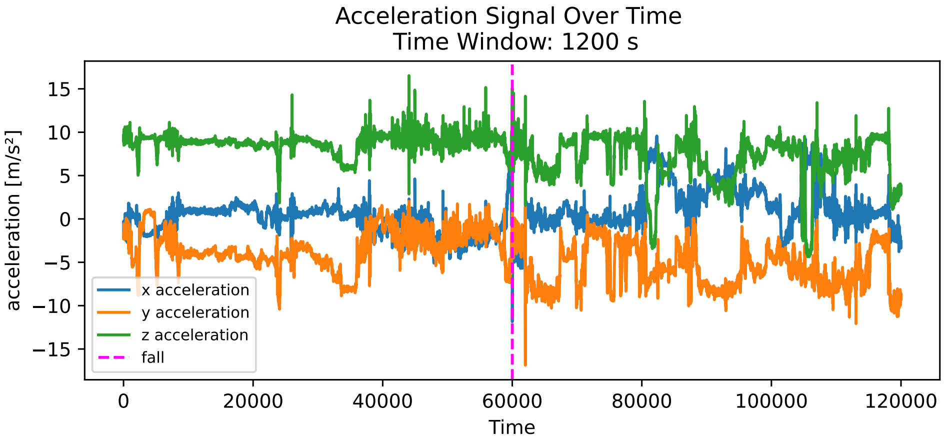
Table 1: Demographic Information for Five Subjects

Table 1 details the personal characteristics for five example subjects in the dataset, including their physical attributes, medical conditions, and the setting in which their falls were recorded. Of the 38 participants, there were 20 women and 18 men. The overall mean age was 69.7 ± 14.4 years.

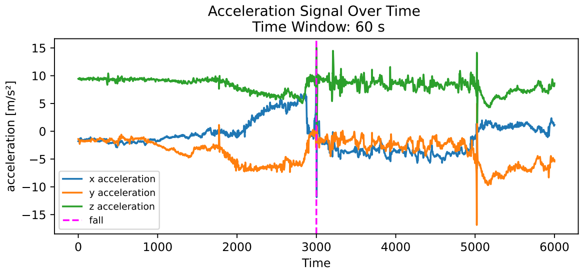
Subject ID	Device	Sensor Location	Sensor Type	Units	Sample Rate (Hz)
93807530	Samsung Galaxy S3	L5	acc, gyro, mag	m/s², °/s, µT	100
00002186	Samsung Galaxy SII	L5	acc, gyro, mag	m/s², °/s, µT	100
42542799	uSense	L5	acc, gyro, mag	m/s², °/s, µT	100
87486959	Samsung Galaxy S3	L5	acc, gyro, mag	m/s², °/s, µT	100
08332163	Samsung Galaxy S3	L5	acc, gyro, mag	m/s², °/s, µT	100

Table 2: Sensor Information for Five Subjects

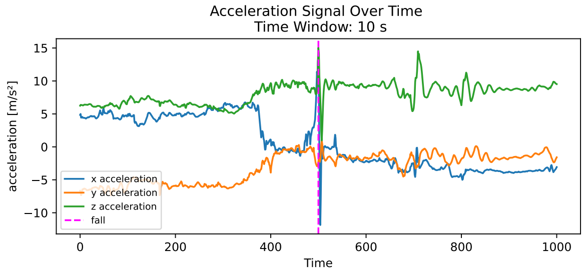
Table 2 contains information about the sensors used to record each subject's fall data. Sensor types present in the dataset include accelerometers, gyroscopes, and magnetometers. For all of the subjects, the sensors were positioned at the L5 location (the fifth lumbar vertebra) or on the thigh. In fall studies, wearable sensors are often placed at L5, which is close to the body's center of mass. This allows the sensors to efficiently capture the body's movements. The sample rate was 100 Hz for all subjects, which means that during each fall, sensor data was collected 100 times per second. It is important to have a sample rate that is high enough to ensure that the sensors capture a comprehensive picture of the body's movements during a fall.



(a) 1200 second Fall Acceleration Signal



(b) 60 second Fall Acceleration Signal



(c) 10 second Fall Acceleration Signal

Figure 1: Fall 93807530-02 signal breakdown: full signal (top) 60 and 10 second signals (bottom).

Figure 1 shows three directional accelerometer signals recorded during an example fall in three different time windows. In each plot, the fall is located at the center. When the fall occurs, there are noticeable changes in the direction of the three accelerometer measures. This indicates a large change in movement and corresponds to a fall. During this fall, the participant recovered immediately and there was no lying period present. This is seen in plots (b) and (c) of Figure 1 that show fluctuations in acceleration immediately after the fall corresponding to the participant getting up. Comparatively, if the three measures were stagnant after a fall, this would correspond to the presence of a lying period. Falls that include lying periods indicate the fall contributing to a potential injury.

3.2 Data Processing

3.2.1 Transformation

The original dataset contains recordings from accelerometers, gyroscopes, and magnetometers in three directions, x, y, and z. To account for differences in body placement and orientation, we transformed the three directional measures for each sensor into a single measure by taking its magnitude.

3.2.2 Segmentation

For the purpose of classification, we aimed to segment the signals into falls and ADLs (Activities of Daily Living). This would allow for binary classification of the two categories. To achieve this we utilized time markers to distinguish between the phases of each fall. t_0 marks the beginning of the pre-fall phase, t_1 the impact phase, t_2 the recovery phase, and t_3 the end of the fall. All together each fall segmentation contains 7 seconds of monitoring with the pre-fall phase lasting 1 second, the impact phase lasting 1 second, and the recovery phase lasting 5 seconds, as seen in Figure 2.

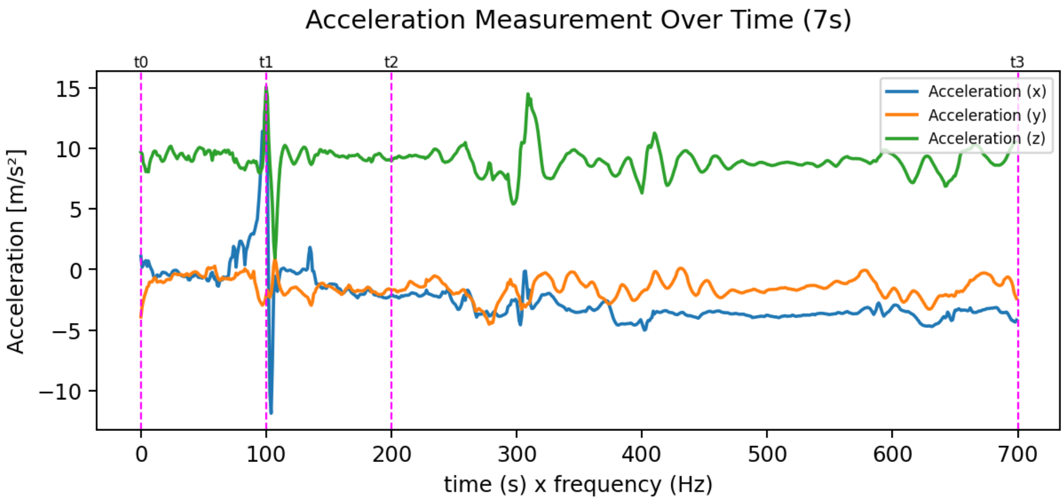


Figure 2: Segmented Fall Signal with Time Markers (Sample 93807530-02)

To segment ADLs from each signal, we discarded measures taken after t_0 that correspond to measures taken during and after the fall. These measures would not be valid because participants would likely behave cautiously after falling, and the measurements would not reflect normal activity. With the remaining data, we applied a sliding-window technique to analyze pariticipant behavior before each fall. Starting with the beginning of the signal, each window was offset from the previous by one second and had a length of seven seconds, the same length as each fall sample. For each window, we analyzed the middle one second, which would represent the impact phase, and looked at the magnitude of acceleration. If the maximum acceleration during the one second window exceeded the threshold, we would keep the window as a valid ADL. We selected an acceleration threshold of 1.4 g or 13.7 m/s^2 [1]. The use of the threshold was to identify significant movements that resembled falls to train the model to differentiate between ADLs and falls.

After segmenting the data, we selected up to five ADLs from each signal. If there were fewer than five valid ADLs, we selected all valid ADLs from the signal. In the end, we had 116 falls, 480 ADLs and a total of 596 data points to train our models.

3.2.3 Normalization

As an additional pre-processing step, we applied standardization using sklearn's StandardScaler to our data. After baseline experiments without normalization, using MinMaxScaler, and StandardScaler, we determined that standardization was the best approach for the classification experiments we aimed to conduct.

4 Experiments & Results

4.1 Tabular Models without Feature Extraction

Using commonly used tabular methods, we trained models both with acceleration-only data and with all sensor channels. Each sample consisted of 700 features representing the measurements recorded at each of 700 timepoints obtained from segmentation, along with a binary label indicating either fall or ADL. The acceleration-only models therefore contained 700 features, while the all-channel models contained $700 \times 3 = 2100$ features. This larger feature space led to a slower fit time for the all-channel models.

The models were evaluated using five-fold cross-validation with sklearn's StratifiedKFold, which was chosen to address class imbalance, since ADL samples greatly outnumber fall samples.

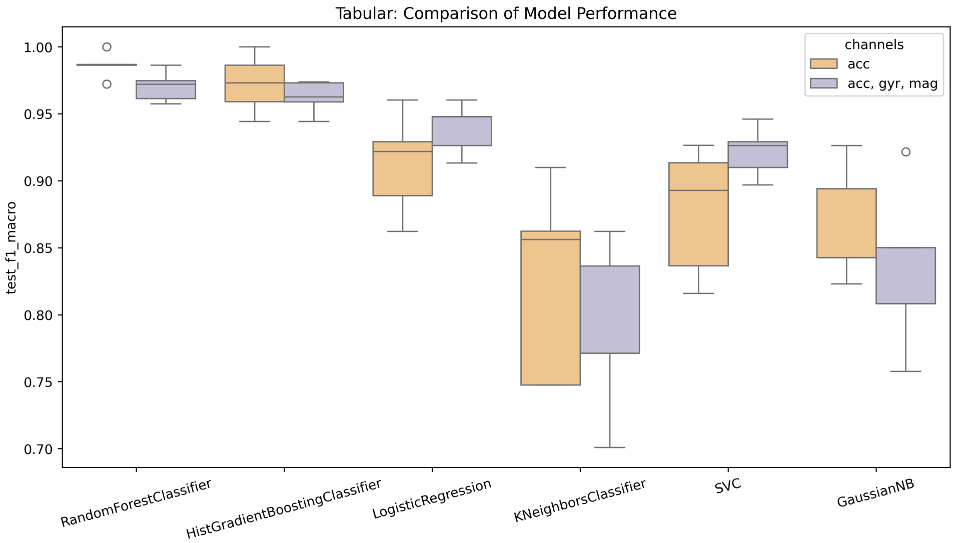


Figure 3: Comparison of F1 Scores Across Tabular Models

Our results were consistent with those reported in prior research [1]. As expected, the RandomForest-Classifier performed strongly, with the acceleration-only model achieving an F1 score of 0.99 ± 0.01 and the all-channel model achieving an F1 score of 0.97 ± 0.01, seen in Table 3. These results confirm that we successfully replicated earlier findings and can use this framework for future experiments.

Across models, the miss rate was consistently higher than the false alarm rate, but both remained very low. As seen in Table 3, the RandomForest acceleration model had a miss rate of 0.03 ± 0.00 with essentially no false alarms, while some weaker models like Logistic Regression, had higher miss rates but still very low false alarm rates. This trend was observed in both the acceleration-only and all-channel models, confirming that the main source of error is missed falls rather than incorrectly labeling ADLs as falls. Interestingly, we observed that both the acceleration-only and all-channel models achieved nearly identical performance.

Channels Used	Model Name	Metrics					
		F1 Score	AUC	Precision	Recall	Miss Rate	False Alarm Rate
Acceleration	RandomForest	0.99 ± 0.01	1.00 ± 0.00	0.99 ± 0.01	0.98 ± 0.02	0.03 ± 0.00	0.00 ± 0.00
	HistGradientBoosting	0.97 ± 0.02	1.00 ± 0.00	0.98 ± 0.02	0.96 ± 0.03	0.07 ± 0.00	0.00 ± 0.00
	LogisticRegression	0.91 ± 0.04	0.97 ± 0.04	0.93 ± 0.02	0.90 ± 0.06	0.18 ± 0.00	0.02 ± 0.00
	KNeighbors	0.83 ± 0.07	0.92 ± 0.05	0.95 ± 0.02	0.78 ± 0.08	0.45 ± 0.00	0.00 ± 0.00
	SVC	0.88 ± 0.05	0.99 ± 0.01	0.96 ± 0.01	0.83 ± 0.06	0.34 ± 0.00	0.00 ± 0.00
	GaussianNB	0.87 ± 0.04	0.93 ± 0.04	0.84 ± 0.04	0.91 ± 0.04	0.09 ± 0.00	0.10 ± 0.00
All Channels	RandomForest	0.97 ± 0.01	0.99 ± 0.00	0.97 ± 0.02	0.97 ± 0.02	0.05 ± 0.00	0.01 ± 0.00
	HistGradientBoosting	0.96 ± 0.01	0.99 ± 0.00	0.96 ± 0.01	0.97 ± 0.02	0.05 ± 0.00	0.02 ± 0.00
	LogisticRegression	0.94 ± 0.02	0.99 ± 0.00	0.95 ± 0.01	0.93 ± 0.04	0.12 ± 0.00	0.02 ± 0.00
	KNeighbors	0.79 ± 0.06	0.92 ± 0.04	0.92 ± 0.05	0.74 ± 0.06	0.52 ± 0.00	0.00 ± 0.00
	SVC	0.92 ± 0.02	0.99 ± 0.00	0.96 ± 0.02	0.89 ± 0.03	0.21 ± 0.00	0.01 ± 0.00
	GaussianNB	0.83 ± 0.06	0.89 ± 0.04	0.81 ± 0.07	0.88 ± 0.05	0.11 ± 0.00	0.13 ± 0.00

Table 3: Comparison of Metrics Across Tabular Models

4.2 Tabular Models with Feature Extraction

In addition to evaluating models on raw tabular time-series data, we investigated the impact of applying feature extraction for dimensionality reduction. Each raw signal consisted of 700 time points. From each signal, we computed four descriptive statistics, mean, standard deviation, minimum, and maximum, for each magnitude measurement. This process yielded 12 features when using acceleration, gyroscope, and magnetometer channels, and 4 features when using only acceleration. The reasoning behind this approach was that feature extraction may reduce runtime, though with a potential loss in performance. We aimed to analyze if the resulting performance was acceptable given the better runtime.

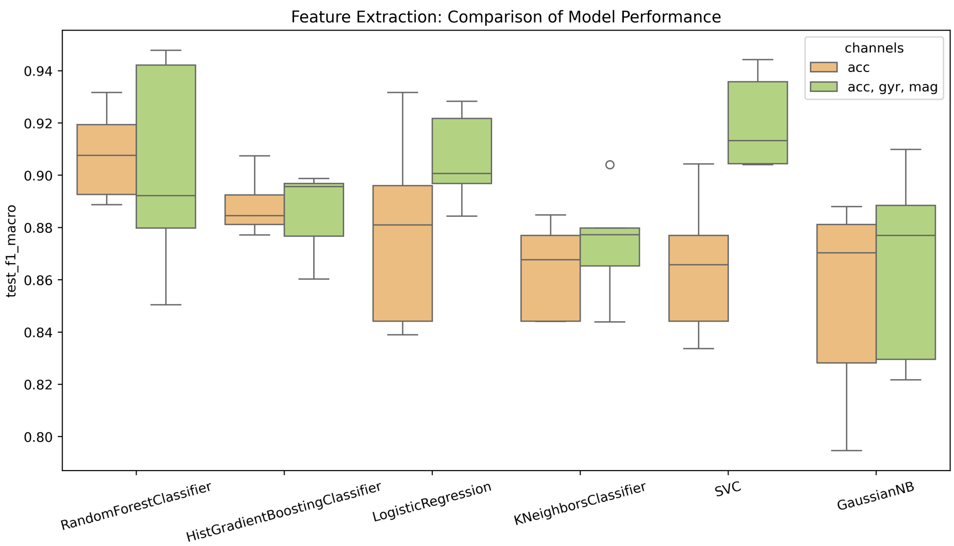


Figure 4: Comparison of F1 Scores Across Tabular Models Using Feature Extraction

Table 4 summarizes the performance of the classifiers for both methods. When using only acceleration, Random Forest achieved the highest performance, achieving an F1 score of 0.91 ± 0.02, with a miss rate of 0.17 ± 0.00 and a false alarm rate of 0.03 ± 0.00. For the all-channel models, Support Vector

Classification (SVC) performed best, with an F1 score of 0.92 ± 0.02, a miss rate of 0.12 ± 0.00, and a false alarm rate of 0.03 ± 0.00.

Although feature extraction led to a slight reduction in classification performance compared to tabular models trained on raw data, the drop in performance was not substantial. The smaller dimensionality provided lower runtime complexity, suggesting benefits when scaling to larger datasets. Additionally, performance was comparable across acceleration-only and all-channel settings, though the miss rate was consistently higher when using only acceleration.

Channels Used	Model Name	Metrics					
		F1 Score	AUC	Precision	Recall	Miss Rate	False Alarm Rate
Acceleration	RandomForest	0.91 ± 0.02	0.97 ± 0.02	0.92 ± 0.01	0.90 ± 0.02	0.17 ± 0.00	0.03 ± 0.00
	HistGradientBoosting	0.89 ± 0.01	0.96 ± 0.02	0.90 ± 0.01	0.88 ± 0.02	0.21 ± 0.00	0.04 ± 0.00
	LogisticRegression	0.88 ± 0.04	0.97 ± 0.01	0.89 ± 0.03	0.87 ± 0.04	0.21 ± 0.00	0.04 ± 0.00
	KNeighbors	0.86 ± 0.02	0.94 ± 0.03	0.88 ± 0.02	0.85 ± 0.03	0.25 ± 0.00	0.04 ± 0.00
	SVC	0.86 ± 0.03	0.94 ± 0.03	0.87 ± 0.03	0.86 ± 0.03	0.24 ± 0.00	0.05 ± 0.00
	GaussianNB	0.85 ± 0.04	0.95 ± 0.02	0.83 ± 0.04	0.88 ± 0.04	0.15 ± 0.00	0.09 ± 0.00
All Channels	RandomForest	0.90 ± 0.04	0.97 ± 0.01	0.91 ± 0.06	0.90 ± 0.04	0.16 ± 0.00	0.04 ± 0.00
	HistGradientBoosting	0.89 ± 0.02	0.97 ± 0.01	0.88 ± 0.04	0.90 ± 0.03	0.16 ± 0.00	0.05 ± 0.00
	LogisticRegression	0.91 ± 0.02	0.98 ± 0.00	0.92 ± 0.02	0.90 ± 0.05	0.16 ± 0.00	0.03 ± 0.00
	KNeighbors	0.87 ± 0.02	0.95 ± 0.02	0.89 ± 0.03	0.87 ± 0.03	0.22 ± 0.00	0.04 ± 0.00
	SVC	0.92 ± 0.02	0.97 ± 0.00	0.92 ± 0.03	0.92 ± 0.03	0.12 ± 0.00	0.03 ± 0.00
	GaussianNB	0.87 ± 0.04	0.96 ± 0.01	0.86 ± 0.07	0.90 ± 0.02	0.12 ± 0.00	0.09 ± 0.00

Table 4: Comparison of Metrics Across Feature Extraction Models

4.3 Time Series Models

To conclude our experiments, we investigated the use of time-series classification models. The same data were used as in the tabular setting, where each signal was represented by 700 time points per channel. In the acceleration-only setting, each sample contained 700 features, creating a 700x1x1 input while in the all-channel setting (acceleration, gyroscope, and magnetometer) was inputted as a matrix of 700x1x3. We implemented a set of time-series models that have been shown to perform well in prior work [1].

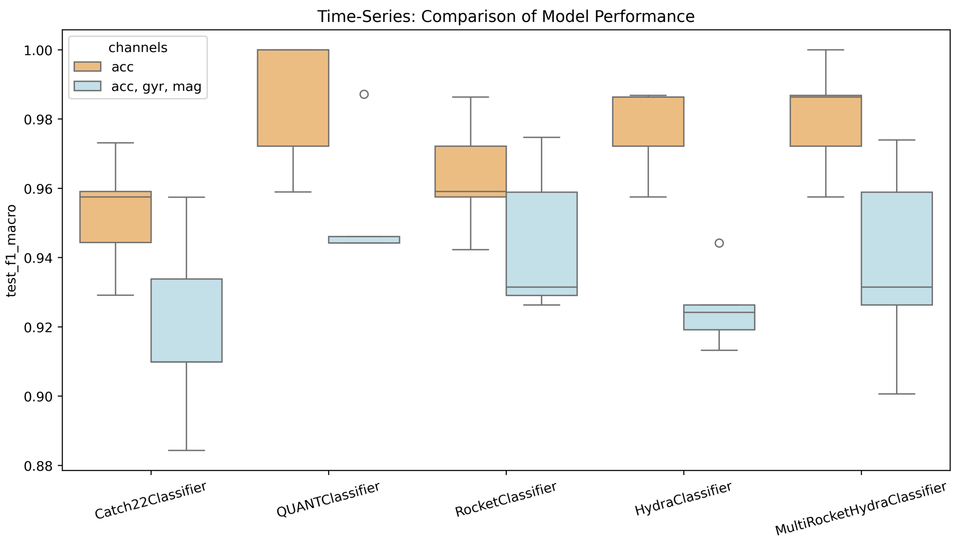


Figure 5: Comparison of F1 Scores Across Time Series Models

Table 5 summarizes the performance of the time-series classifiers. We found that models trained on acceleration-only data consistently outperformed those trained on all channels. In both cases, QUANT achieved the strongest performance. For acceleration-only classification, QUANT achieved an F1 score of 0.99 ± 0.02, with a miss rate of 0.03 ± 0.00 and a false alarm rate of 0.00 ± 0.00. In contrast, classification using all channels yielded lower results, with an F1 score of 0.96 ± 0.03, a miss rate of 0.09 ± 0.00, and a false alarm rate of 0.01 ± 0.00.

These results demonstrate that acceleration alone is sufficient for high-performing fall detection, and that the inclusion of additional channels does not improve performance. In fact, miss rates and false alarm rates were consistently lower in the acceleration-only setting, suggesting that additional channels may introduce noise rather than contribute useful information. Compared with tabular models, time-series classifiers achieved similar F1 scores but reduced miss and false alarm rates. We conclude that time-series classification with acceleration-only data represents the most viable modeling approach for accurate fall detection.

Channels Used	**Model Name**	**Metrics**					
		F1 Score	**AUC**	**Precision**	**Recall**	**Miss Rate**	**False Alarm Rate**
Acceleration	Catch22	0.95 ± 0.02	0.97 ± 0.02	0.97 ± 0.01	0.94 ± 0.02	0.12 ± 0.00	0.01 ± 0.00
	Hydra	0.98 ± 0.01	0.97 ± 0.02	0.99 ± 0.00	0.97 ± 0.02	0.07 ± 0.00	0.00 ± 0.00
	MultiRocketHydra	0.98 ± 0.02	0.97 ± 0.02	0.99 ± 0.01	0.97 ± 0.02	0.06 ± 0.00	0.00 ± 0.00
	QUANT	0.99 ± 0.02	1.00 ± 0.00	0.99 ± 0.01	0.98 ± 0.03	0.03 ± 0.00	0.00 ± 0.00
	Rocket	0.96 ± 0.02	0.94 ± 0.02	0.99 ± 0.01	0.94 ± 0.02	0.11 ± 0.00	0.00 ± 0.00
All Channels	Catch22	0.89 ± 0.04	0.94 ± 0.03	0.94 ± 0.05	0.86 ± 0.03	0.26 ± 0.00	0.01 ± 0.00
	Hydra	0.93 ± 0.03	0.91 ± 0.04	0.95 ± 0.04	0.91 ± 0.04	0.16 ± 0.00	0.01 ± 0.00
	MultiRocketHydra	0.93 ± 0.02	0.91 ± 0.03	0.95 ± 0.02	0.91 ± 0.03	0.16 ± 0.00	0.01 ± 0.00
	QUANT	0.96 ± 0.03	0.99 ± 0.00	0.97 ± 0.02	0.95 ± 0.04	0.09 ± 0.00	0.01 ± 0.00
	Rocket	0.94 ± 0.03	0.93 ± 0.04	0.94 ± 0.03	0.93 ± 0.04	0.11 ± 0.00	0.02 ± 0.00

Table 5: Comparison of Metrics Across Time-Series Models

5 Conclusion

Falls pose a serious health risk to populations worldwide, particularly older adults and individuals with high risk conditions. Proactive fall detection systems offer a way to reduce health risks by providing timely alerts to individuals and healthcare providers. In this study, we aimed to improve the fall detection framework by exploring the impact of different sensor measurements and classification methods. We evaluated whether incorporating gyroscope and magnetometer data, in addition to accelerometer signals, increases detection performance. Our goal was to address a gap in the literature: understanding how sensor measurements can be best utilized to develop robust models that minimize false alarm and miss rates.

We evaluated our models using the RATE framework: reality, accuracy, timeliness, and explainability [1]. We meet the requirement of reality by training exclusively on real-world fall data. Our time-series models achieved high accuracy, with the QUANT model trained on accelerometer data alone reaching an F1 score of 0.99 ± 0.02. In terms of timeliness, our models detect falls in a timely manner that is promising for real-world use. Finally, the use of simple measures, acceleration, gyroscope, and magnetometer magnitudes ensures explainability. Together, these results demonstrate that time-series models trained solely on accelerometer data are effective, and highlight the potential for the use of simplistic data in fall detection.

This work also provides new directions for further research. Expanding the amount of real-world fall data would allow more opportunity to certify the usability of fall detection systems. Our findings suggest that focusing on accelerometer data may simplify data collection without sacrificing performance. Also, incorporating additional sensors beyond the methods we tested could be explored. Another area to investigate in future work is feature extraction methods that generalize across multiple datasets. Finally, we want to continue quantifying and minimizing false alarm and miss rates to ensure that models can be smoothly implemented for detecting real falls.

Our results suggest that simple and interpretable machine learning models, particularly time-series models trained on accelerometer data, offer a path toward reliable fall detection.

References

[1] Timilehin B. Aderinola, Luca Palmerini, Ilaria D'Ascanio, Lorenzo Chiari, Jochen Klenk, Clemens Becker, Brian Caulfield, and Georgiana Ifrim. Accurate and efficient real-world fall detection using time series techniques. In Vincent Lemaire, Georgiana Ifrim, Anthony Bagnall, Thomas Guyet, Simon Malinowski, Patrick Schäfer, and Romain Tavenard, editors, *Advanced Analytics and Learning on Temporal Data*, pages 52–79, Cham, 2025. Springer Nature Switzerland.

[2] Shehroz S. Khan and Jesse Hoey. Review of fall detection techniques: A data availability perspective. *Medical Engineering Physics*, 39:12–22, 2017.