*Research Article*

# On the Heterogeneity of Existing Repositories of Movements Intended for the Evaluation of Fall Detection Systems

**Eduardo Casilari** ![ORCID], **José A. Santoyo-Ramón** ![ORCID], **and José M. Cano-García** ![ORCID]

*Departamento de Tecnología Electrónica, Universidad de Málaga, ETSI Telecomunicación, 29071 Málaga, Spain*

Correspondence should be addressed to Eduardo Casilari; ecasilari@uma.es

Due to the serious impact of falls on the autonomy and health of older people, the investigation of wearable alerting systems for the automatic detection of falls has gained considerable scientific interest in the field of body telemonitoring with wireless sensors. Because of the difficulties of systematically validating these systems in a real application scenario, Fall Detection Systems (FDSs) are typically evaluated by studying their response to datasets containing inertial sensor measurements captured during the execution of labelled nonfall and fall movements. In this context, during the last decade, numerous publicly accessible databases have been released aiming at offering a common benchmarking tool for the validation of the new proposals on FDSs. This work offers a comparative and updated analysis of these existing repositories. For this purpose, the samples contained in the datasets are characterized by different statistics that model diverse aspects of the mobility of the human body in the time interval where the greatest change in the acceleration module is identified. By using one-way analysis of variance (ANOVA) on the series of these features, the comparison shows the significant differences detected between the datasets, even when comparing activities that require a similar degree of physical effort. This heterogeneity, which may result from the great variability of the sensors, experimental users, and testbeds employed to generate the datasets, is relevant because it casts doubt on the validity of the conclusions of many studies on FDSs, since most of the proposals in the literature are only evaluated using a single database.

## 1. Introduction

Falls, in particular falls among elderly, are a major social concern in current societies. The World Health Organization has reported that 646,000 persons die from falls each year worldwide, so they represent the second cause of unintentional injury deaths after car accidents [1]. In this respect, it has been shown that a rapid response after a fall can lower the risk of hospitalization by 26% and the death rate by 80% [2]. As a consequence, during the past decade, great research efforts have been devoted to the development of efficient and low-cost technologies for automatic Fall Detection Systems (FDSs).

Falls are generically and ambiguously defined as a loss of balance or accident that causes an individual to rest involuntarily on the ground or other lower level [3]. Most unintentional falls can be easily distinguished from other movements by human visual inspection. However, this task is not so evident when it is carried out by an automatic system. Accordingly, the problem of fall detection has been addressed through different approaches, which can be clustered into two great generic strategies: context-aware and wearable systems. Under the first strategy, an FDS can be deployed by placing video cameras and other ambient sensors, such as pressure sensors and microphones, in the vicinity of the user to be monitored. However, in most practical cases, the mobility of the patients can be tracked in a more adaptive and cost-effective way by employing lightweight sensors that can be directly transported on the clothes or as another garment or a piece of jewelry (e.g., as a pendant). The decreasing costs and widespread popularity of electronic wearables and especially those intended for sporting activities have fostered the adoption of this type of transportable solutions to investigate and implement FDSs.

Under a wearable FDS, a detection algorithm is permanently in charge of analyzing the signals captured by the sensors worn by the user to identify any anomalous mobility pattern that can be linked to the occurrence of a fall. As soon as a fall is presumed, an alerting message (phone call and SMS) to a remote monitoring point (medical premises and patients' relative) will be forwarded by the FDS. In the vast majority of wearable architectures, the detection decision is based on the measurements provided by an accelerometer and, in some cases, a gyroscope (integrated in the same Inertial Measurement Unit, IMU), which are attached to a certain part of the user's body.

The general goal of an FDS is to simultaneously minimize both the number of falls that remain unnoticed and the generation of false alarms, that is to say, conventional movements or Activities of Daily Living (ADLs) that are misinterpreted as falls. A crucial element in the investigation of a wearable FDS is the procedure by which the detection algorithm will be methodically evaluated to check its actual capacity to discriminate ADLs from falls.

In almost all works existing in the related literature, FDSs are tested against a set of labelled movements that include both ADLs and falls. In order to repeat the analysis by changing the detection techniques and the parameterization of the algorithms, the movements are previously prerecorded in files that contain the corresponding timestamp and measurements gathered by the inertial sensors. The quality and representativeness of the employed dataset of movements are a key aspect to assess the validity of the evaluation. In this regard, it has been estimated that it is necessary to record between 70,000 and 100,000 days to collect about 100 actual falls by continuously monitoring persons aged over 65 [4]. Owing to the obvious practical difficulties of monitoring actual falls experienced by elderly people, the general procedure followed by the literature to evaluate a fall detection algorithm is using datasets of activity traces that are intentionally created by experimental users. For this purpose, the participants in the experiments normally execute a series of predetermined movements while they transport the corresponding wearable sensors in one or several positions of their bodies. These movements typically incorporate different types of conventional ADLs (sitting, climbing stairs, picking up objects from the floor, etc.) and falls, which are mimicked taking into account different aspects, such as the direction (lateral and backwards) or the cause of the fall (slipping, stumbling, and tripping).

In almost all initial studies on FDSs, a group of volunteers were recruited to generate a specific dataset which was employed for the evaluation of the proposed architecture. These datasets were rarely released by the authors to enable their use by other researchers to validate new algorithms. To tackle this lack of a benchmarking framework, a nonnegligible number of datasets have recently been produced and made publicly available on the Web to cross compare FDSs with a common reference.

The use of normally young and healthy volunteers that emulate falling in a systematic way in a 'controlled' scenario, as surrogates for actual falls of older persons, is still a controversial issue in the field of FDSs. By tracking during six months two groups of persons totaling 16 older people, Kangas et al. conducted a study aiming at comparing the dynamics of real-life falls of older people with those simulated by middle-aged volunteers [5]. From the results, the authors concluded that the features of the acceleration data captured during accidentals falls follow a similar pattern to those measured from emulated falls, although some significant differences were detected (for example, in the timing of the different phases of the falls or in the acceleration magnitude measured during the impact against the floor). In a similar study [6], Klenk et al. compared the actual backward falls suffered by four elderly people to those mimicked by 18 young individuals. Results seem to indicate that the 'compensation' strategies to avoid the damages of the impact followed by the subject during the unintentional falls introduce relevant differences (e.g., jerkier movements with higher changes in the acceleration) with respect to the case of the emulated falls.

Besides, Bagalà et al. [7] have shown that the efficacy of certain algorithms successfully tested against datasets of emulated falls may notably decrease when they are evaluated with traces captured in a real scenario. In other works, such as that by Sucerquia et al., the ability of the proposed FDS to avoid false alarms is evaluated by monitoring elderly people that transport the wearable detection system during their daily routines. In these cases, the sensitivity of the detector cannot be computed unless a real fall occurs during the monitoring period. A similar strategy is described by Aziz et al. in [8]. These authors report that the number of false alarms of an FDS, which is based on a Support Vector Machine classifier, deteriorates when it is employed by a community of 19 older adults. In this scenario, 2 out of 10 actual falls suffered by the participants were not identified by the system.

In any case, these studies are based on the analysis of a very small number of real falls. The fact is that, to the best of our knowledge, the repository provided by the FARSEEING European project [9] is the only dataset that provides inertial measurements of real-world falls of elderly patients although again the number of samples that are publicly available, only 22, is quite limited. Thus, this work mainly focuses on those datasets grounded on emulated falls and ADLs (although in some cases, ADLs were captured not by an execution of predetermined activities on a laboratory but by monitoring the participants during their daily routines).

On the other hand, although the use of public and well-known datasets is gaining an increasing acceptance in the literature, most studies base their validation on the use of just one or, at most, two repositories. So, a question arises about the correctness of extrapolating the results obtained with a particular dataset when another repository is considered.

The goal of this study is to recap and compare the characteristics of the existing public repositories of inertial measurements intended for the assessment of FDSs.

The paper is organized as follows. Section 2 revises the available datasets, synopsizing their basic properties and the testbeds (employed sensors, characteristics of the experimental users, and typology of the movements) which were

deployed to generate the data. The section also describes the criteria to select the datasets to be compared. Section 3 presents the statistical features employed to characterize the mobility of the traces of the datasets, while Section 4 compares the datasets by showing the results of the analysis of variance (ANOVA) of these characteristics. The main conclusions are summarized in Section 5.

## 2. Revision and Selection of Public Datasets

As aforementioned, a key problem for the development of an automatic fall detection architecture is the need of trustworthy repositories that can be employed to thoroughly evaluate the accuracy of the detection decisions, i.e., the capacity of the system to correctly identify ADLs and falls by simultaneously avoiding false alarms and undetected falls.

Table 1 presents a comprehensive list of the authors, references, institutions, and year of publication of the existing datasets intended for the study of wearable systems. All these datasets comprise the measurements collected by the inertial sensors worn by the selected volunteers during their daily life or while performing a preconfigured set of movements in a controlled testbed. In this revision we do not include those available databases of inertial measurements (such as those presented in [10] or [11]) that are envisioned for other types of HAR (Human Activity Recognition) systems but do not incorporate falls among the represented activities .

In the case of Context-Aware Systems (CAS), different research groups have also published datasets containing the measurements captured by fixed video camera, motion and depth sensors (such as Kinect), and/or other ambient sensors (vibration detectors, pressure, infrared, and Doppler sensors, and near-field imaging systems), while a set of volunteers emulate falls and ADLs in a predefined testbed. Among these databases, we can mention the following: CIRL Fall Recognition [12], Le2i FDD [13], SDUFall [14], EDF&OCCU [15], eHomeSeniors [16], Multiple Camera Fall [17] KUL High-Quality Fall Simulation [18], UTA [19], FUKinect-Fall [20], or MEBIOMEC [21] datasets, as well as the infrared video clips described by Mastoraky and Makris in [22] or those sequences provided by Adhikari et al. in [23]. These datasets are out of the scope of this paper although we do consider those databases, such as UR Fall or UP Fall, which were conceived to test hybrid CAS-type and wearable FDSs, i.e., systems that make their detection decision from the joint analysis of video images (and/or magnitudes collected by environmental sensors) and measurements from inertial sensors transported by the users.

The number of samples, the considered typologies of the emulated ADLs and falls, and the duration of the traces (i.e., the duration of the recorded movements), as well as the basic characteristics of the participants (number, gender, weight, and age range) of each dataset, are enumerated in Table 2.

Table 2 illustrates the great heterogeneity of criteria used to define the experimental framework where the samples were captured, both with regard to the selection of the test subjects and the number and type of simulated movements. In some repositories, such as tFall, the ADLs were not emulated (scheduled and executed in a laboratory) but obtained by tracking the real-life movements of the subjects during a certain period of time. As expected, in most cases, the movements were exclusively carried out by volunteers under the age of 60. In the few testbeds in which older subjects participated, almost none of the older participants simulated any fall, so their samples are limited to examples of ADLs.

Table 3 summarizes, in turn, the type and basic properties (sampling rate and range) of the sensors employed to generate the repositories. The table also indicates the corporal position on which the inertial sensors were located or attached during the experiments. As it can be observed from the table, although there are cases where up to seven sensing positions have been considered, most datasets include just a single measuring point. In all cases, the sensor embeds, at least, an accelerometer and, less often, a gyroscope, a magnetometer, and/or an orientation sensor. In any case, the table shows the variability of the characteristics of the sensors (e.g., with sampling rates ranging from 10 to 200 Hz) and the body location considered to collect the measurements in the different testbeds again.

In the recent literature about FDSs, the use of some of these public datasets as benchmarking tools is becoming more and more common. However, in most studies, just one or, at most, two repositories are utilized to evaluate the effectiveness of the proposed detection algorithm. Khojasteh et al. [24] employed four datasets, although two of them (DaLiac [25] and Epilepsy [26] databases) do not encompass falls, which only allows assessing the capability of the system to avoid misinterpreting ADLs as falls. As a consequence, the conclusions of most works are mainly based on the results obtained when the proposed system is tested against a very particular set of samples.

Given the huge diversity of the experimental setups in which the datasets were generated, it is legitimate to question whether the conclusions achieved with a certain repository can be extrapolated to scenarios with a different typology of subjects, movements (simulated or not), or to a different parameterization of the inertial sensors.

In this context, Medrano et al. utilized three repositories (tFall, DLR, and MobiFall) in [27] to show that the effectiveness of an FDS based on a supervised machine learning strategy remarkably diminishes when the discrimination algorithms are tested against a database different from that utilized for training. In a more recent work [28], we concluded that even when the algorithm is trained and tested with traces of the same datasets and users, the quality metrics of the classification process may differ notably. In particular, we analyzed the performance of a deep learning classifier (a convolutional neural network) when it is individually trained and evaluated as a fall detector with 14 of the repositories presented in Table 1. Results clearly indicated that the performance dramatically varies depending on the dataset to which the detector is applied.

In the following sections, we thoroughly analyze the statistical properties of a representative number of these datasets to get a deeper understanding of the existing divergences between these repositories.

TABLE 1: Authors and references of the existing public datasets (*the check mark indicates those datasets employed in this study).

| Dataset | Ref. | Authors | Institution | City (country) | Year | * |
|---|---|---|---|---|---|---|
| DLR | [29] | Frank et al. | German Aerospace Center (DLR) | Munich (Germany) | 2010 | √ |
| LDPA | [30] | Kaluza et al. | Jožef Stefan Institute | Ljubljana (Slovenia) | 2010 | |
| MobiFall | [31] | Vavoulas et al. | BMI Lab (Technological Educational Institute of Crete) | Heraklion (Greece) | 2013 | |
| MobiAct | [32] | | | | 2016 | |
| EvAAL | [33] | Kozina et al. | Department of Intelligent Systems, Jozef Stefan Institute | Ljubljana (Slovenia) | 2013 | |
| TST fall detection | [34] | Gasparrini et al. | TST Group (Università Politecnica delle Marche) | Ancona (Italy) | 2014 | |
| tFall | [35] | Medrano et al. | EduQTech (University of Zaragoza) | Teruel (Spain) | 2014 | |
| UR fall detection | [36] | Kępski et al. | Interdisciplinary Center for Computational modelling (University of Rzeszow) | Krakow (Poland) | 2014 | |
| Erciyes University | [37] | Özdemir and Barshan | Department of Electrical and Electronics Engineering (Erciyes University) | Kayseri (Turkey) | 2014 | √ |
| Cogent Labs | [38] | Ojetola et al. | Labs (Coventry University) | Coventry (UK) | 2015 | |
| Gravity Project | [39] | Vilarinho et al. | SINTEF ICT | Trondheim (Norway) | 2015 | |
| Graz UT OL | [40] | Wetner et al. | Graz University of Technology | Graz (Austria) | 2015 | |
| UMAFall | [41] | Casilari et al. | Dpto. Tecnología Electrónica (University of Málaga) | Málaga (Spain) | 2016 | √ |
| FARSEEING | [42] | Klenk et al. | FARSEEING Consortium (SENSACTION-AAL European Commission Project) | Five hospital or scholar centers in Germany and one university in New Zealand | 2016 | |
| SisFall | [43] | Sucerquia et al. | SISTEMIC (University of Antioquia) | Antioquia (Colombia) | 2017 | √ |
| UniMiB SHAR | [44] | Micucci et al. | Department of Informatics, Systems and Communication (University of Milano) | Bicocca, Milan (Italy) | 2017 | |
| SMotion | [45] | Ahmed et al. | Department of Computer Science (University of Karachi) | Karachi (Pakistan) | 2017 | |
| IMUFD | [46] | Aziz et al. | Injury Prevention and Mobility Laboratory (Simon Fraser University) | Burnaby (BC, Canada) | 2017 | √ |
| CGU-BES | [47] | Wang et al. | Chang Gung University | Taoyuan (Taiwan) | 2018 | |
| CMDFALL | [48] | Tran et al. | International Research Institute MICA (Hanoi University of Science and Technology) | Hanoi (Vietnam) | 2018 | |
| DU-MD | [49] | Saha et al. | Department of Electrical and Electronic Engineering (University of Dhaka) | Dhaka (Bangladesh) | 2018 | |
| SmartFall and Smartwatch datasets | [50] | Mauldin et al. | Department of Computer Science, Texas State University | San Marcos (TX, USA) | 2018 | |
| UP-Fall | [51] | Martínez-Villaseñor et al. | Facultad de Ingeniería (Universidad Panamericana) | Mexico City (Mexico) | 2019 | √ |
| DOFDA | [52] | Cotechini et al. | Department of Information Engineering (Università Politecnica delle Marche) | Ancona (Italy) | 2019 | √ |

TABLE 2: Personal characteristics of the volunteers, typology distribution, number, and duration of the executed movements.

| Dataset | Number of subjects (females/males) | Age (years) | Weight (kg) | Height (cm) | Number of types of ADLs/falls | Number of samples (ADLs/falls) | Duration of the samples (s) |
|---|---|---|---|---|---|---|---|
| DLR | 19 (8/11) | [23–52] | n.i. | [160–183] | 15/1 | 1017 (961/56) | [0.27–864.33] |
| LDPA | 5 (n.i.) | n.i. | n.i. | n.i. | 10/1 | 100/75 | Up to 300 s |
| MobiFall | 24 (7/17) | [22–47] | [50–103] | [160–189] | 9/4 | 630 (342/288) | [0.27–864.33] |
| MobiAct | 57 (15/42) | [20–47] | | | 9/4 | 2526 (1879/647) | [4.89–300.01] |
| EvAAL | 1 (n.i.) | n.i. | [50–120] | [160–193] | 7/1 | 57 (55/2) | [0.162–30.172] |
| TST fall detection | 11 (n.i.) | [22–39] | n.i. | [162–197] | 4/4 | 264 (132/132) | [3.84–18.34] s |
| tFall | 10 (3/7) | [20–42] | [54–98] | [161–184] | n.i./8 | 10909 (9883/1026) | 6 s (all samples) |
| UR fall detection | 6 (0/6)[3] | n.i. (over 26) | n.i. | n.i. | 5/4 | 70 (40/30) | [2.11–13.57] |
| Erciyes University | 17 (7/10) | [19–27] | [47–92] | [157–184] | 16/20 | 3302(1476/1826) | [8.36–37.76] |

TABLE 2: Continued.

| Dataset | Number of subjects (females/males) | Age (years) | Weight (kg) | Height (cm) | Number of types of ADLs/falls | Number of samples (ADLs/falls) | Duration of the samples (s) |
|---|---|---|---|---|---|---|---|
| Cogent Labs | 42 (6/36) | [18–51] | [43–108] | [150–187] | 8/6 | 1968 (1520/448) | [0.53–55.73] |
| Gravity Project | 2 (n.i.)[4] | [26–32] | [63–80] | [170–185] | 7/12 | 117 (45/72) | [9.00–86.00] |
| Graz UT OL | 5 (n.i.) | n.i. | n.i. | n.i. | 10/4 | 2460 (2240/220) | [0.18–961.23] |
| UMAFall | 19 (8/11) | [18–68] | [50–97] | [156–193] | 12/3 | 746 (538/208) | 15 s (all samples) |
| FARSEEING | 15 (8/7) | [56–86] | [51–101] | [148–190] | 0/22 | 22 (0/22) | 1200 |
| SisFall | 38 (19/19) | [19–75] | [41.5–102] | [149–183] | 19/15 | 4505 (2707/1798) | [9.99–179.99] s |
| UniMiB SHAR | 30 (24/6) | [18–60] | [50–82] | [160–190] | 9/8 | 7013 (5314/1699) | 1 s (all samples) |
| SMotion | 120 (40/71 + 9 n.i.) | [17–79] | [35–95] | [125–186] | 3/1 | 309 (304/5) | [0.52734–27.1875] |
| IMUFD | 10 (n.i.) | n.i. | n.i. | n.i. | 8/7 | 600(390/210) | [15–20.01] |
| CGU-BES | 15 (4/11) | 21.8 ± 1.8 | 63.0 ± 10.1 kg | 167.7 ± 6.0 | 8/4 | 195 (135/60) | [11.49–16.73] |
| CMDFALL | 50 (20/30) | [21–40] | n.i. | n.i. | 12/8 | 1000 (600/400) | 450 s[1] |
| DU-MD | 10 (4/6) | [16–22] | [40–101] | [147–185] | 8/2 | 3299 (2309/990) | [2.85–11.55] |
| Smartfall | 7 (n.i.) | [21–55] | n.i. | n.i. | 4/4 | 181 (90/91) | [0.576–16.8] |
| Smartwatch | 7 (n.i.) | [20–35] | n.i. | n.i. | 7/4 | 2563 (2456/107) | [1–3.776] |
| UP-Fall | 17 (8/9) | [18–24] | [53–99] | [157–175] | 6/5 | 559(304/255) | [9.409–59.979] |
| DOFDA | 8 (2/6) | [22–29] | [60–94] | [173–187] | 5/13 | 432 (120/312) | 1.96–17.262 |

1. For the CMDFAL dataset, all the 20 programmed movements are executed in a continuous manner during 7.5 minutes. 2. n.i.: not indicated by the authors.

TABLE 3: Position and characteristics of the sensor used in the different datasets.

| Dataset | Number of sensing points | Captured signals in each sensing point | Positions of the sensing points | Type of device | Sampling rate (Hz) | Range |
|---|---|---|---|---|---|---|
| DLR | 1 | 3 (A, G, M) | Waist (belt) | 1 external IMU | 100 | ±5 g (A) ±1200°/s (G) ±75 µT (M) |
| LDPA | 4 | Position (x,y,z coordinates) | Right ankle, left ankle, waist (belt), and chest | 4 external IMUS (tags) | 10 | Tens of meters |
| MobiFall and MobiAct | 1 | 3 (A, G, O) | Thigh (trouser pocket) | 1 smartphone | 87 (A) 100 (G,O) | ±2 g (A) ±200°/s (G) ±360° (O) |
| EvAAL | 2 | 1 (A) | Chest and right thigh | 2 external IMUs | 50 | ±16 g (A) |
| TST fall detection | 2 | 1 (A) | Waist and wrist | 2 external IMUs | 100 | ±8 g (A) |
| Erciyes University | 6 | 3(A, G, M) | Chest, head, ankle, thigh, wrist, and waist | 6 external IMUs | 25 | ±16 g (A) ±1200°/s (G) ±150 µT (M) |
| tFall | 1 | 1 (A) | Alternatively: thigh (right or left pocket) and hand bag (left or right side) | 1 smartphone | 45 (±12) | ±2 g (A) |
| UR fall detection | 1 | 1 (A) | Waist (near the pelvis) | 1 external IMU | 256 | ±8 g (A) |
| Cogent Labs | 2 | 2 (A, G) | Chest and thigh | 2 external IMUs | 100 | ±8 g (A) ±2000°/s (G) |
| Gravity Project | 2 | 1 (A) | Thigh (smartphone in a pocket) Wrist (smartwatch) | 1 smartphone 1 smartwatch | 50 (SP) 157 (SW) | ±16 g (A) ±2 g (A) |

Table 3: Continued.

| Dataset | Number of sensing points | Captured signals in each sensing point | Positions of the sensing points | Type of device | Sampling rate (Hz) | Range |
|---|---|---|---|---|---|---|
| Graz UT OL | 1 | 2 (A, O) | Waist (belt bag) | 1 smartphone[3] | 5 | ±2 g (A) ±360° (O) |
| UMAFall | 5 | 3(A, G, M) | Ankle, chest, thigh, and waist | 1 smartphone[4] | 100 (SP) | ±16 g (A) ±256°/s (G) ±4800 μT (M) |
| | | | Wrist | 4 external IMUs | 20 (IMUs) | |
| FARSEEING | 1 | 2 (A,G) | Waist or thigh | 1 external IMU | 100 | ±6 g (A) ±100°/s (G) |
| SisFall | 1 | 3 (A, A, G) | Waist | 2 accelerometers and a gyroscope in a single mode | 200 | ±16 g (A1) ±8 g (A2) ±2000°/s (G) |
| UniMiB SHAR | 1 | 1 (A) | Thigh (left or right trouser pocket) | 1 smartphone | 50 | ±2 g (A) |
| SMotion | 1 | A, G | Waist | 1 external IMU | 51 | ±4 g (A2) ±500°/s (G) |
| IMUFD | 7 | 3(A, G, M) | Chest, head, left ankle, left thigh, right ankle, right thigh, and waist | 7 external IMUs | 128 | ±16 g (A) ±2000°/s (G) ±800 μT (M) |
| CGU-BES | 1 | 2 (A, G) | Chest | 1 sensing mote with a gyroscope and accelerometer | 200 | ±3.6 g (A) ±400°/s (G) |
| CMDFALL[1] | 2 | 1 (A) | Left wrist and left hip | 1 external IMU | 50 | ±16 g (A) |
| DU-MD | 1 | 1 (A) | Wrist | 1 external IMU | 33 | ±4 g (A) |
| SmartFall | 1 | 1 (A) | Wrist | 1 external IMU | 31.25 | ±16 g (A) |
| Smartwatch | 1 | 1 (A) | Wrist (left hand) | Smartwatch (MS band) | 31.25 | ±8 g (A) |
| UP-Fall | 5 | 2 (A, G) | Ankle, neck, and thigh (pocket) Waist and wrist | 5 external IMUs | 14 | ±8 g (A) ±2000°/s (G) |
| DOFDA | 1 | 4 (A, G, O, M) | Waist | 1 external IMU | 33 | ±16 g (A) ±2000°/s (G) ±800 μT (M) |

*Note.* A : accelerometer, G : gyroscope, O : orientation measurements, M : magnetometer, SP : smartphone. 1. TST, UR, CMDFALL, and UP datasets also include the measurements (RGB, depth, and skeleton information) of Kinect sensors or video cameras, not considered in this Table 2. n.i.: not indicated by the authors.

2.1. Election of the Compared Datasets. In order to compare the properties of the signals provided by different repositories on equal terms, we only select those datasets that contain inertial measurements captured on the same position. In particular, in a first analysis, we focus on those traces collected on the waist as several studies [53–57] have shown that this is one of the most adequate positions to place an inertial sensor aimed at characterizing the general dynamics of the body. This election benefits from the fact that the waist is near the center of mass of the human body in a standing posture. When compared to other placements such as a limb or the chest, the waist also provides better ergonomics as it may enable the user to transport the wearable sensor almost in a seamless way (e.g., attached to a belt).

To ensure that the analysis is performed with a minimum number of samples, we only take into account those datasets

with, at least, 300 samples. Consequently, we discard UR, FARSEEING, LDPA, and TST datasets, although they include traces captured with the sensor located on the waist. For a similar reason, we exclude the SMotion dataset [45], which is actually aimed at assessing fall risk and not fall detection systems, as it only contains 5 falls.

Finally, the Graz UT OL dataset is also discarded because of the small range of the employed accelerometer (±2g), which can prevent a proper representation of the acceleration peaks caused by falls (typically exceeding 4-5g).

## 3. Selection of the Characteristics for the Analysis

As in most works in the literature, the study will be based on the signals collected by the triaxial accelerometers ($A_X[i]$,

$A_Y[i]$, $A_Z[i]$ for the $i$-th measurement), which are provided by the datasets. Future studies should contemplate the analysis of the signals collected by the gyroscope and, secondarily, the magnetometer. Nevertheless, it is still under discussion that the information provided by the gyroscope may significantly improve the success rate of methods merely based on the accelerometry signals (see [58] for a revision of this issue).

During the free-fall period before the impact, a collapse typically prompts a sudden drop of the acceleration components, which is interrupted by a sharp peak of the acceleration magnitude (sometimes followed by several secondary peaks) produced by the collision against the floor [59]. Therefore, to define a common basis to compare the traces, which present a wide variety of lengths, we focus on the interval of every measurement sequence where the highest difference between the "valleys" (decays) and peaks of the acceleration components is detected. Once this analysis interval is extracted, the rest of the trace is ignored. For this purpose, we set up a sliding observation window of duration $t_W = 0.5$ s, consisting of $N_W$ samples:

$$N_W = t_W \cdot f_s, \tag{1}$$

where $fs$ indicates the sampling rate of the sensors.

To find the analysis interval within each trace, we follow the procedure presented in [60]. Thus, for each possible observation window within the sequences, we calculate the magnitude of the maximum variation of the acceleration components ($A_{w_{\mathrm{diff}}}[m]$) as

$$A_{w_{\mathrm{diff}}}[m] = \sqrt{\left(A_{X_{\max}}[m] - A_{X\min}[m]\right)^2 + \left(A_{Y_{\max}}[m] - A_{Y\min}[m]\right)^2 + \left(A_{Z_{\max}}[m] - A_{Z\min}[m]\right)^2}, \tag{2}$$

where $A_{X_{\max}}[m]$, $A_{Y_{\max}}[m]$, and $A_{Z_{\max}}[m]$ designate the maximum values of the components measured by the accelerometer in the $x$-, $y$- and $z$-axis, respectively, in the $m$-th sliding observation interval. Thus, for the $x$-axis, we have

$$A_{X_{\max}}[m] = \max\left(A_X[i]\right), \quad \forall i \in [m, m + N_W - 1]. \tag{3}$$

The analysis or observation interval will correspond to the subset of consecutive samples $[k_o, k_o + N_W - 1]$ where the maximum $A_{w_{\mathrm{diff}(\max)}}$ of $A_{w_{\mathrm{diff}}}[m]$ is located:

$$A_{w_{\mathrm{diff}(\max)}} = A_{w_{\mathrm{diff}}}[k_o] = \max\left(A_{w_{\mathrm{diff}}}[m]\right),$$
$$\forall m \in [1, N - N_W + 1]), \tag{4}$$

where $k_o$ is the index of the first sample of the analysis interval while $N$ denotes cardinality (number of samples for each axis) of the trace.

In order to compare the different datasets, we extract the acceleration components of the signals during the analysis interval to compute the following twelve statistical features for all the traces.

All these features have been regularly employed by the related literature on FDSs and human activity recognition systems (see, for example, the FDS described in [37, 43, 54, 61–72] or the comprehensive analyses presented by Vallabh in [73] or by Xi in [74]).

(1) The mean Signal Magnitude Vector ($\mu_{\mathbf{SMV}}$), which gives an idea of the average mobility experienced by the body during the analysis interval. This mean can be calculated as

$$\mu_{\mathrm{SMV}} = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \mathrm{SMV}[i], \tag{5}$$

where SMV[$i$] represents the Signal Magnitude Vector (SMV) of the acceleration for the $i$-th sample:

$$\mathrm{SMV}[i] = \sqrt{\left(A_X[i]\right)^2 + \left(A_Y[i]\right)^2 + \left(A_Z[i]\right)^2}. \tag{6}$$

(2) The standard deviation ($\sigma_{\mathbf{SMV}}$) of SMV[$i$], which describes the variability of the acceleration during the observation window:

$$\sigma_{\mathrm{SMV}} = \sqrt{\frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \left(\mathrm{SMV}[i] - \mu_{\mathrm{SMV}}\right)^2}. \tag{7}$$

(3) The mean absolute difference ($\mu_{\mathbf{SMV}_{\mathbf{diff}}}$) between two consecutive samples of the acceleration module, which is estimated as

$$\mu_{\mathrm{SMV}_{\mathrm{diff}}} = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} |\mathrm{SMV}[i+1] - \mathrm{SMV}[i]|. \tag{8}$$

This parameter is useful as it informs about the brusque fluctuations of the acceleration during a fall [75].

(4) The mean rotation angle ($\mu_\theta$) may help to detect the changes of the body orientation of the body caused by a fall [75]. This angle is computable as

$$\mu_\theta = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \left( \cos^{-1}\left[ \frac{A_X[i] \cdot A_X[i+1] + A_Y[i] \cdot A_Y[i+1] + A_Z[i] \cdot A_Z[i+1]}{\mathrm{SMV}[i] \cdot \mathrm{SMV}[i+1]} \right] \right). \qquad (9)$$

(5) The acceleration component in the direction perpendicular to the floor plane is strongly determined by the gravity. Thus, the tilt of the body provoked by the falls usually triggers a noteworthy alteration of the acceleration components that are parallel to the floor plane when the individual remains static in an upright posture. To characterize the alteration of the body position with respect to the standing position, we also compute the **mean magnitude ($\mu_{Ap}$)** of the vector formed by these two acceleration components:

$$\mu_{Ap} = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \sqrt{\left(A_{H1}[i]\right)^2 + \left(A_{H2}[i]\right)^2}, \quad (10)$$

where the pair $(A_{H1}[i], A_{H2}[i])$ of acceleration components may alternatively represent $(A_X[i], A_Y[i])$ $(A_X[i], A_Z[i])$ or $(A_Y[i], A_Z[i])$ depending on the placement and orientation of the accelerometer in each dataset.

(6) The aforementioned value of $\mathbf{A_{w_{diff(max)}}}$, which gives an insight of the range of the variability of the three acceleration components.

(7) The peak or maximum ($\mathbf{SMV_{max}}$) of the SMV, as a key element to describe the violence of the impact against the floor:

$$\mathrm{SMV_{max}} = \max(\mathrm{SMV}[i]), \quad \forall i \in [k_o, k_o + N_W - 1].$$
$$(11)$$

(8) The "valley" or minimum ($\mathbf{SMV_{min}}$) of the SMV to characterize the phase of free-fall:

$$\mathbf{SMV_{min}} = \min(\mathrm{SMV}[i]), \quad \forall i \in [k_o, k_o + N_W - 1].$$
$$(12)$$

(9) The skewness of SMV$[i]$ ($\gamma_{\mathbf{SMV}}$), which describes the symmetry of the distribution of the acceleration:

$$\gamma_{\mathrm{SMV}} = \frac{1}{\sigma_{\mathrm{SMV}}^3 \cdot N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \left(\mathrm{SMV}[i] - \mu_{\mathrm{SMV}}\right)^3.$$
$$(13)$$

(10) The Signal Magnitude Area (**SMA**) [43]. This parameter, which is an extended feature used to evaluate the physical activity, can be estimated as

$$\mathrm{SMA} = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \left( \left|A_X[i]\right| + \left|A_Y[i], \right| + \left|A_Z[i]\right| \right).$$
$$(14)$$

(11) Energy (**E**). Since falls are associated to rapid and energetic movements, we also consider the sum of the energy ($E$) estimated in the three axes during the observation interval [72]:

$$E = \frac{1}{N_W} \cdot \left( \sum_{i=0}^{N_W-1} \left|\mathrm{FFT}_X[i]\right|^2 + \sum_{i=0}^{N_W-1} \left|\mathrm{FFT}_Y[i]\right|^2 \right.$$
$$\left. + \sum_{i=0}^{N_W-1} \left|\mathrm{FFT}_Z[i]\right|^2 \right),$$
$$(15)$$

where $\mathrm{FFT}_X[i]$, $\mathrm{FFT}_Y[i]$, and $\mathrm{FFT}_Y[i]$, respectively, indicate the Discrete Fourier Transform of the acceleration components $A_X[i], A_Y[i]$, and $A_Z[i]$ in the analysis interval, straightforwardly computable (for the $x$-axis) as

$$\mathrm{FFT}_X[i] = \sum_{m=0}^{N_W-1} A_X[k_o+m] \cdot e^{-j2\pi(im/N_W)},$$
$$\text{for } i = 0, 1, \ldots, N_W - 1. \qquad (16)$$

(12) Mean of the autocorrelation function ($\mu_{\mathbf{R}}$) of the acceleration magnitude captured during the observation interval:

$$\mu_R = \frac{1}{N_W} \cdot \sum_{l=0}^{N_W-1} R[m], \qquad (17)$$

where $R[m]$ represents the $m$-th lag value in the series of the normalized autocorrelation coefficients of SMV$[i]$:

$$R[m] = \frac{1}{\sigma_{\mathrm{SMV}}^2 \cdot N_W} \cdot \left( \sum_{i=k_o}^{k_o+N_W-m-1} \left(\mathrm{SMV}[i] - \mu_{\mathrm{SMV}}\right) \right.$$
$$\left. \cdot \left(\mathrm{SMV}[i+M] - \mu_{\mathrm{SMV}}\right) \right), \quad \text{for } m = 0, 1, \ldots, N_W - 1.$$
$$(18)$$

This feature $\mu_R$ is taken into account as long as the acceleration during a conventional activity normally exhibits a certain degree of self-correlation that could be impacted by the unexpected movements caused by a fall.

## 4. Comparison and Discussion of the Datasets

For an initial comparison of the statistical features of the different datasets, we utilize boxplots (or box-and-whisker

plots), an extended and intuitive visual tool, to display the data distribution in a standardized manner.

Figures 1–12 show the boxplots of the twelve statistics when they are separately calculated for the ADLs and the fall movements of the seven datasets under study. In the graphs, for each dataset and type of activity (ADL/fall), the median of the corresponding statistic is denoted by the central line in each box while the 25th and 75th percentiles are indicated by the lower and upper limits of the box. The dotted lines or "whiskers" represent an interval over and under the box of 1.5 IQR (the height of the box or Interquartile range between the 25th and 75th percentiles). All the data outside these margins (box and whiskers) are considered to be outliers and marked as red crosses in the figures.

The graphs show the high inter- and intravariability of the statistics of the traces. As it refers to the intravariability, within each repository, the analysis identifies a wide IQR interval and a high number of outliers for almost all the characteristics, in particular for the ADLs. Similarly, when the boxplots of the different databases are compared, a huge heterogeneity is also present.

This intravariability among datasets is also noticeable (both for ADLs and falls) even in the case of a basic feature, such as the mean acceleration magnitude during the observation window (which is assumed to be linked to the period of greatest alteration in the body acceleration). For all the considered statistics and for both ADLs and falls, we can observe several pairs of datasets where the IQR intervals (which concentrate 50% of the samples) do not even overlap, i.e., the 25% quartile of the corresponding feature of a certain dataset exhibits a higher value than that of the 75% quartile for the same feature of a different dataset. In addition, the magnitude of the IQR interval strongly differs from one repository to another. In some cases, the estimated mean of certain statistics in one dataset is several times higher when compared to others. This is more visible for those characteristics associated with the loss of verticality: the mean rotation angle ($\mu_\theta$) and the mean magnitude of the acceleration components ($\mu_{Ap}$) perpendicular to the vertical plane while standing.

The statistical significance of these divergences among the repositories can be systematically confirmed by an ANOVA (Analysis of variance) test. Figures 13 and 14 depict the post hoc multiple comparison of the estimated means of the twelve features based on the results achieved by a one-way (or single-factor) ANOVA. In the bars of the figure, the circular marks indicate the mean whereas the corresponding comparison interval for a 95% confidence level is represented by the line extending out from the symbol. The group means are considered to be significantly different if the intervals determined by the lines are disjoint.

Each subgraph in these two figures shows, in red, those datasets that have a characteristic with a significantly different mean than that of the fall or ADL movements of another dataset (marked in blue), which is taken as a

reference by way of an example. As can be seen in the figure, there are very few cross comparisons, indicated in grey, in which the null hypothesis is not rejected as the differences between the means of the characteristics are not significantly relevant.

This inconsistency in the characterization of the different datasets is also appreciated if we consider other duration of the time observation window in which the maximum variation of the acceleration components is detected. Figures 15 and 16 present the analysis of variance when it is applied to the features computed for two different observation intervals (0.5 s and 1 s, respectively). For the sake of simplicity, the graphs only show the six first characteristics although a similar disparity can be found if the other six features were shown.

*4.1. Comparison of the Different Types of ADLs.* The differences analyzed in the previous section could be partly justified by the fact that the terms 'ADL' and 'falls' may hide a huge variety of different movements. This is particularly true for the groups labelled as ADLs, as they can encompass activities ranging from those that require almost no effort, such as standing, to those that are much more physically demanding (such as running). In spite of this evident heterogeneity, the authors of the datasets normally select the typology of the ADLs to be emulated by the volunteers without previously discussing the degree of mobility that the selected activities actually require.

In order to minimize the effects of this heterogeneity in the ADLs, we propose to individualize the previous ANOVA study taking into account the nature (physical effort) of the ADLs. For this purpose, as we also suggested in [76], we split the ADLs of each repository into three generic subcategories: basic ordinary movements (such as getting up, sitting, standing, and lying down), standard routines that entail some physical effort or a higher degree of mobility or leaning of the body (walking, climbing up and down stairs, picking an object from the floor, and tying shoe laces), and finally, sporting activities (running, jogging, jumping, and hopping).

By taking into account this taxonomy, Table 4 displays and catalogues the different types of ADLs and falls contained in the seven datasets under analysis. The table shows that each subcategory in each dataset is basically represented by the same three or four types of common movements. Thus, a certain homogeneity could be presumed. In two of the datasets (DOFDA and IMUFD), there are no sporting activities. As an extra type of 'nonfall' movements, the table also indicates which repository includes the emulation of near falls, that is to say, missteps, stumbles, trips, or any other type of accidental movements that involve a loss of balance but do not result in a fall.

The individualized ANOVA analyses of the series of the six statistical features of the datasets are depicted in Figures 17 and 18 (for basic movements), Figures 19 and 20
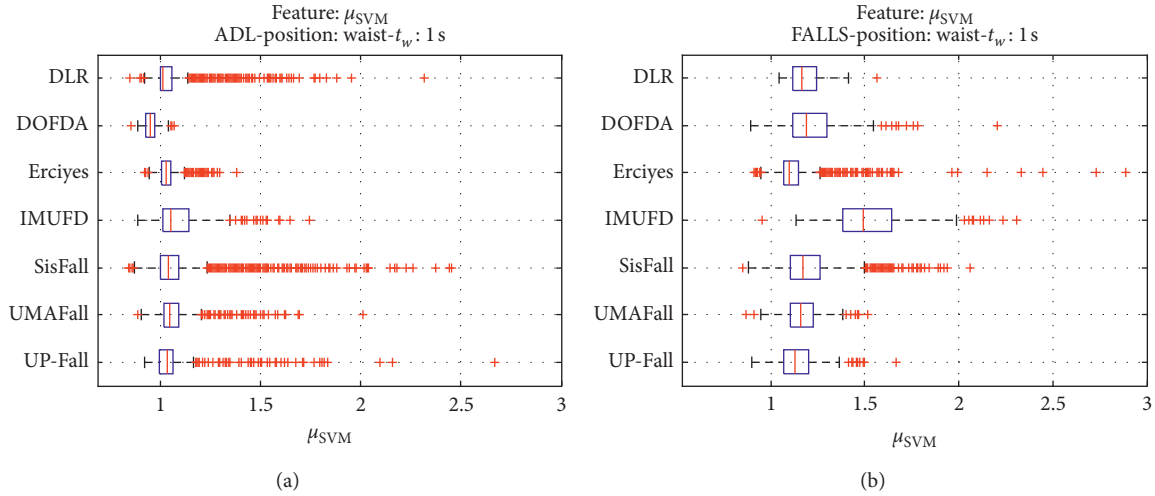
Feature: $\mu_{\text{SVM}}$
ADL-position: waist-$t_w$: 1 s

(a)

Feature: $\mu_{\text{SVM}}$
FALLS-position: waist-$t_w$: 1 s

(b)

FIGURE 1: Boxplots of the mean Signal Magnitude Vector ($\mu_{\text{SMV}}$) for the ADLs (left column) and falls. (a) ADLs. (b) Falls (right column) of all datasets.
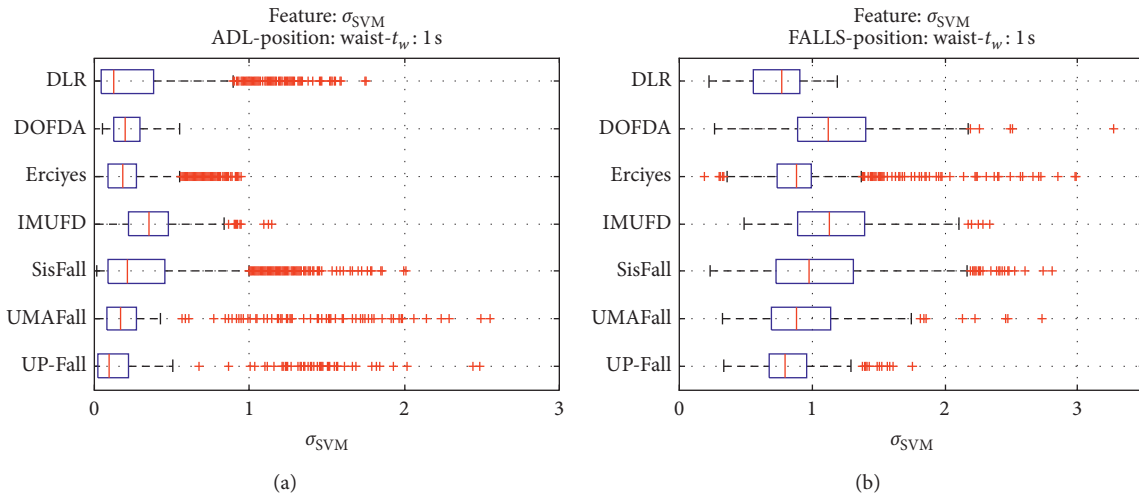


Feature: $\sigma_{\text{SVM}}$
ADL-position: waist-$t_w$: 1 s

(a)

Feature: $\sigma_{\text{SVM}}$
FALLS-position: waist-$t_w$: 1 s

(b)

FIGURE 2: Boxplots of the maximum variation of the standard deviation of the Signal Magnitude Vector ($\sigma_{\text{SMV}}$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

(for standard movements), and Figures 21 and 22 (for sporting movements).

Despite the categorization and clustering of the traces, the graphs again reveal the great variability of the datasets when they are compared to each other. For all three movement types and for all metrics, the mean of the six statistical features of each dataset is significantly different from that calculated for, at least, two other datasets. Figures evince that in a nonnegligible combination of cases (some of which are highlighted in blue in the graphs), the null hypothesis can be rejected for the comparison of a certain mean of a particular dataset with the mean of the same metric of the rest of datasets. For example, five out of the six contemplated features in the basic movements of the UMAFall repository present a mean value significantly different to those of all the other datasets. A similar behavior is detected in other repositories and types of movements (e.g., the sporting activities in the UP dataset).

A similar conclusion can be reached by analyzing the near-fall movements existing in two datasets (IMUFD and Erciyes). Figures 23 and 24 confirm that the six statistics with which these movements have been characterized present mean values that significantly differ for the two repositories.

4.2. Comparison for the Same Type of Movement: Walking. The disparity in the statistical characterization of the traces is confirmed even when the same type of movement is considered as the basis for comparing the datasets. Figures 25 and 26 depict the results obtained when the ANOVA is exclusively applied to those movement samples (measured on the waist) labelled as "walking". We select this ADL due to its importance in real-life scenarios of FDSs as it is the movement that normally precedes falls and because it is present in the seven datasets (DLR, DOFDA, Erciyes, IMUFD, SisFall, UMAFall, and UP-Fall) that employ a
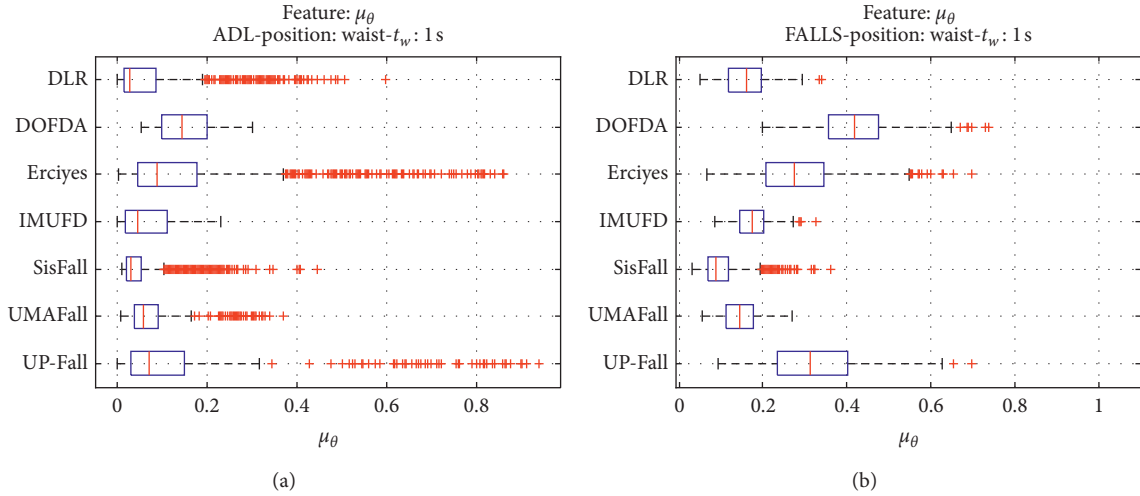
FIGURE 3: Boxplots of the mean rotation angle ($\mu_\theta$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
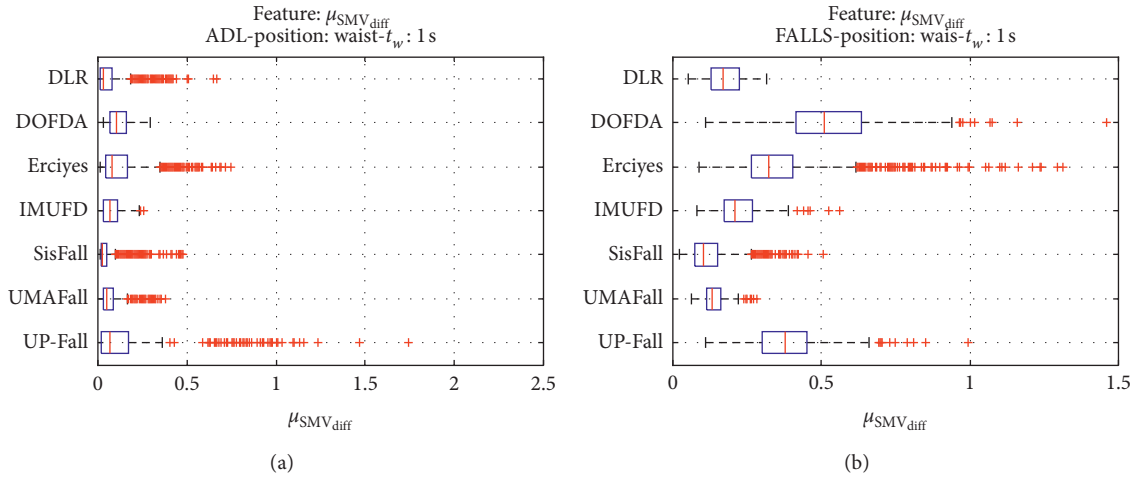


FIGURE 4: Boxplots of the mean absolute difference between consecutive samples ($\mu_{SMV_{diff}}$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
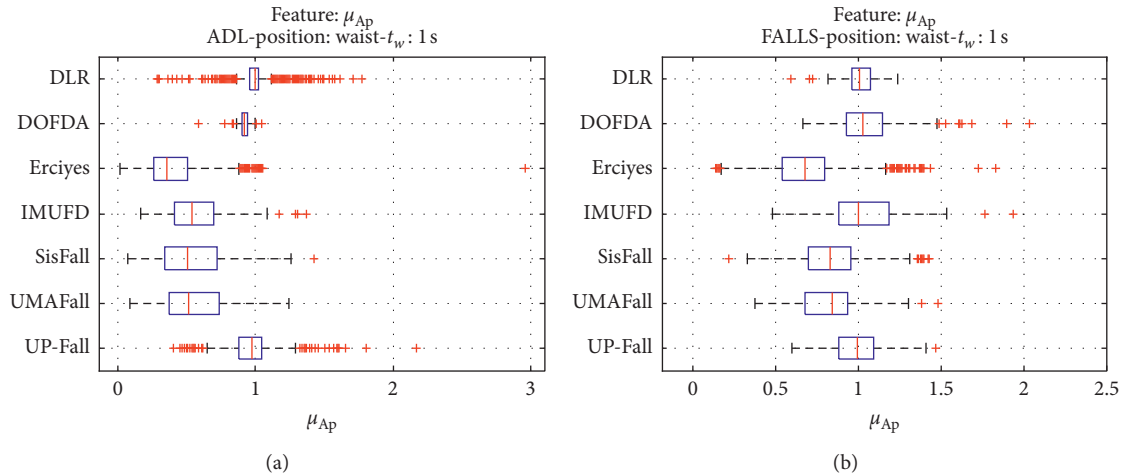


FIGURE 5: Boxplots of the mean module of the not perpendicular acceleration components ($\mu_{Ap}$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
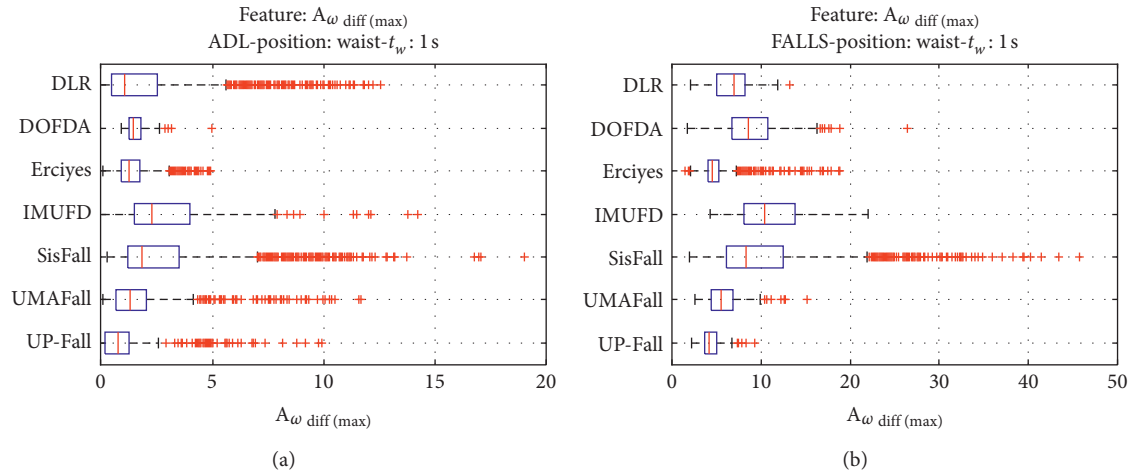
FIGURE 6: Boxplots of the maximum variation of the acceleration components ($A_{\omega_{\text{diff(max)}}}$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
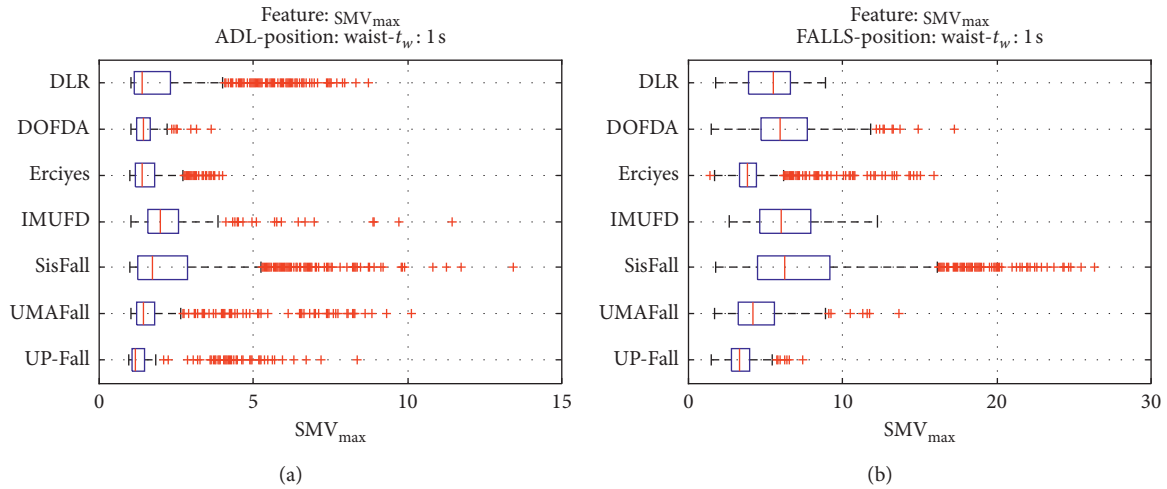


FIGURE 7: Boxplots of the maximum ($\text{SMV}_{\text{max}}$) of the SMV for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
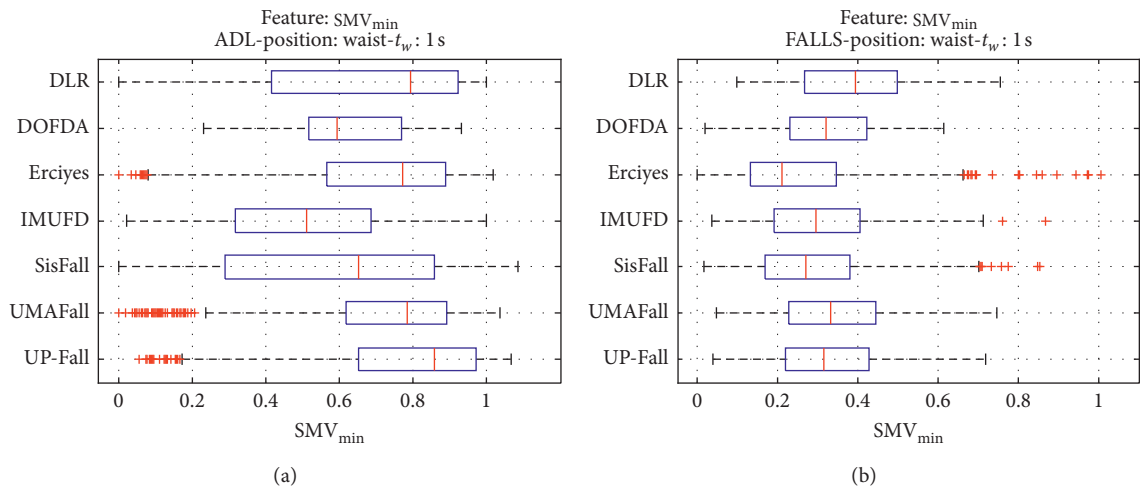


FIGURE 8: Boxplots of the minimum ($\text{SMV}_{\text{min}}$) of the SMV for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
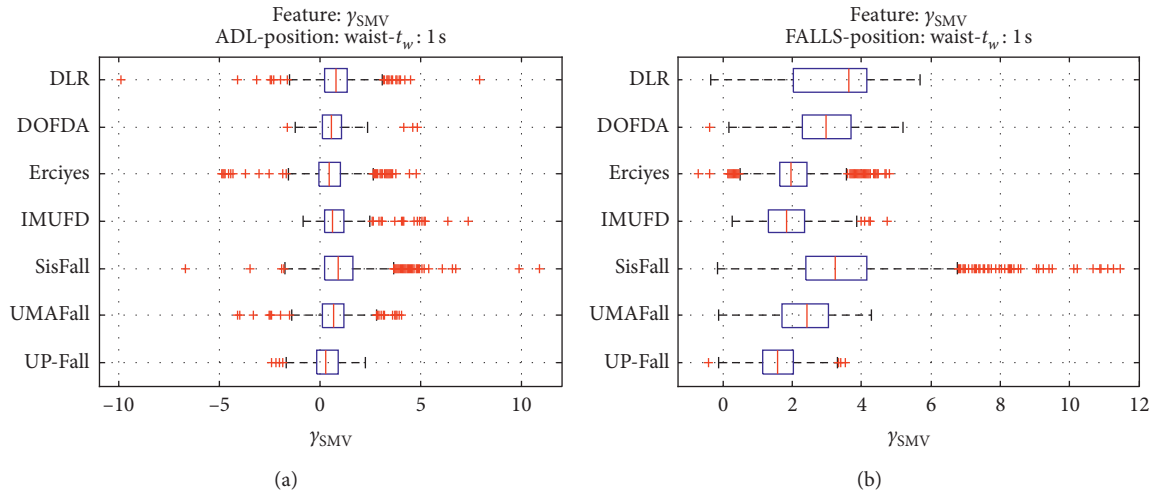
FIGURE 9: Boxplots of the skewness of *SMV* ($\gamma_{SMV}$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
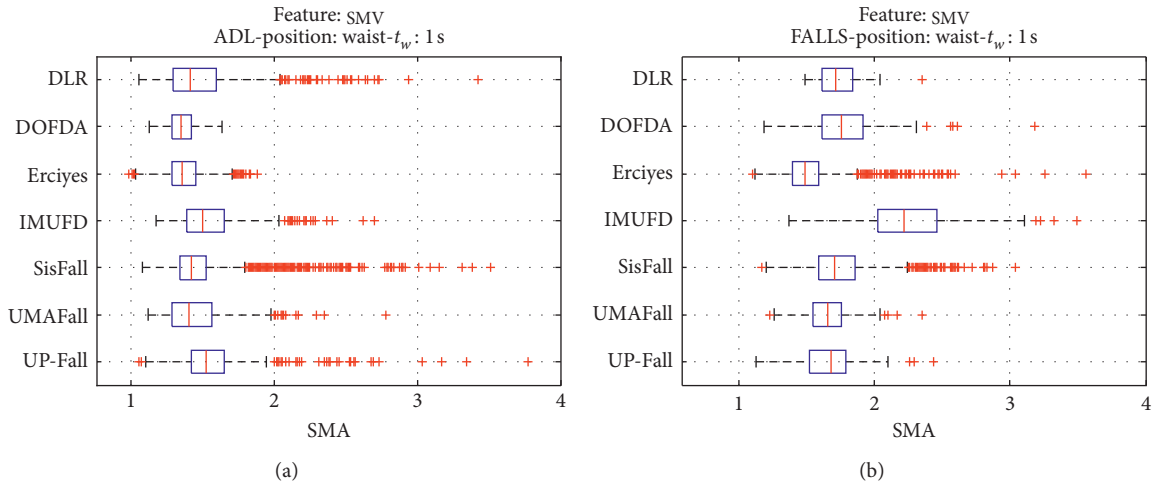


FIGURE 10: Boxplots of the Signal Magnitude Area (SMA) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
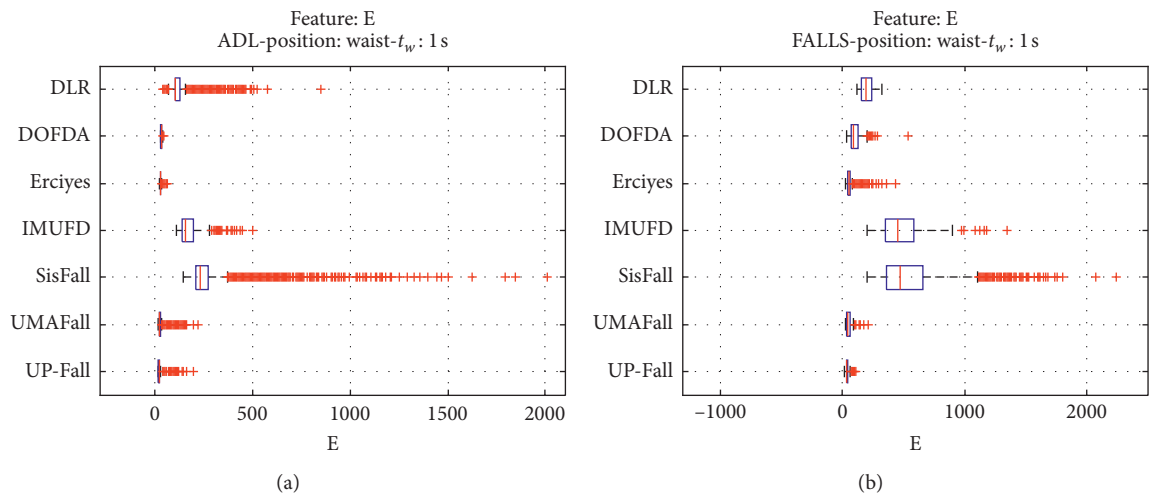


FIGURE 11: Boxplots of the sum of the energy (E) estimated in the three axes for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
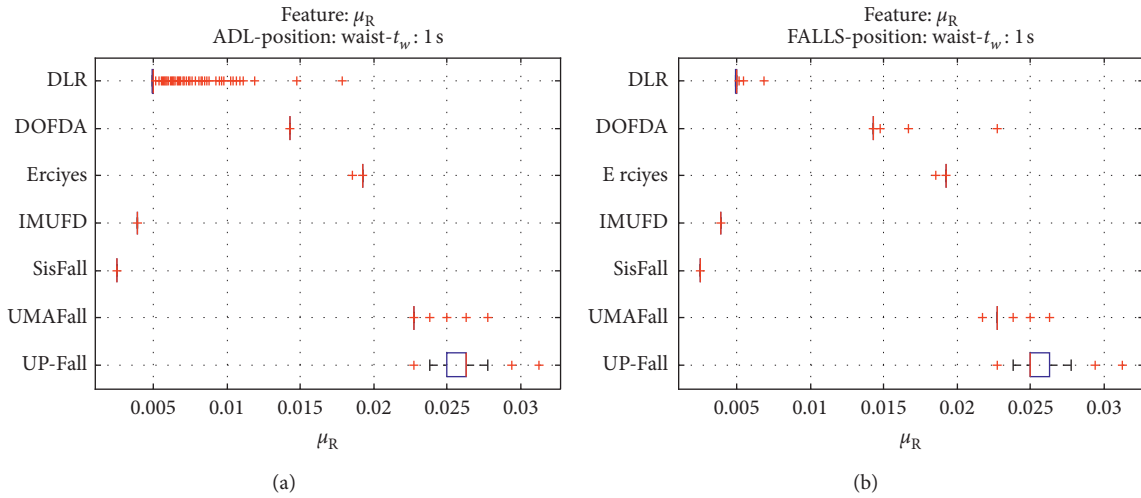
(a)



(b)

FIGURE 12: Boxplots of the mean of the autocorrelation function ($\mu_R$) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.
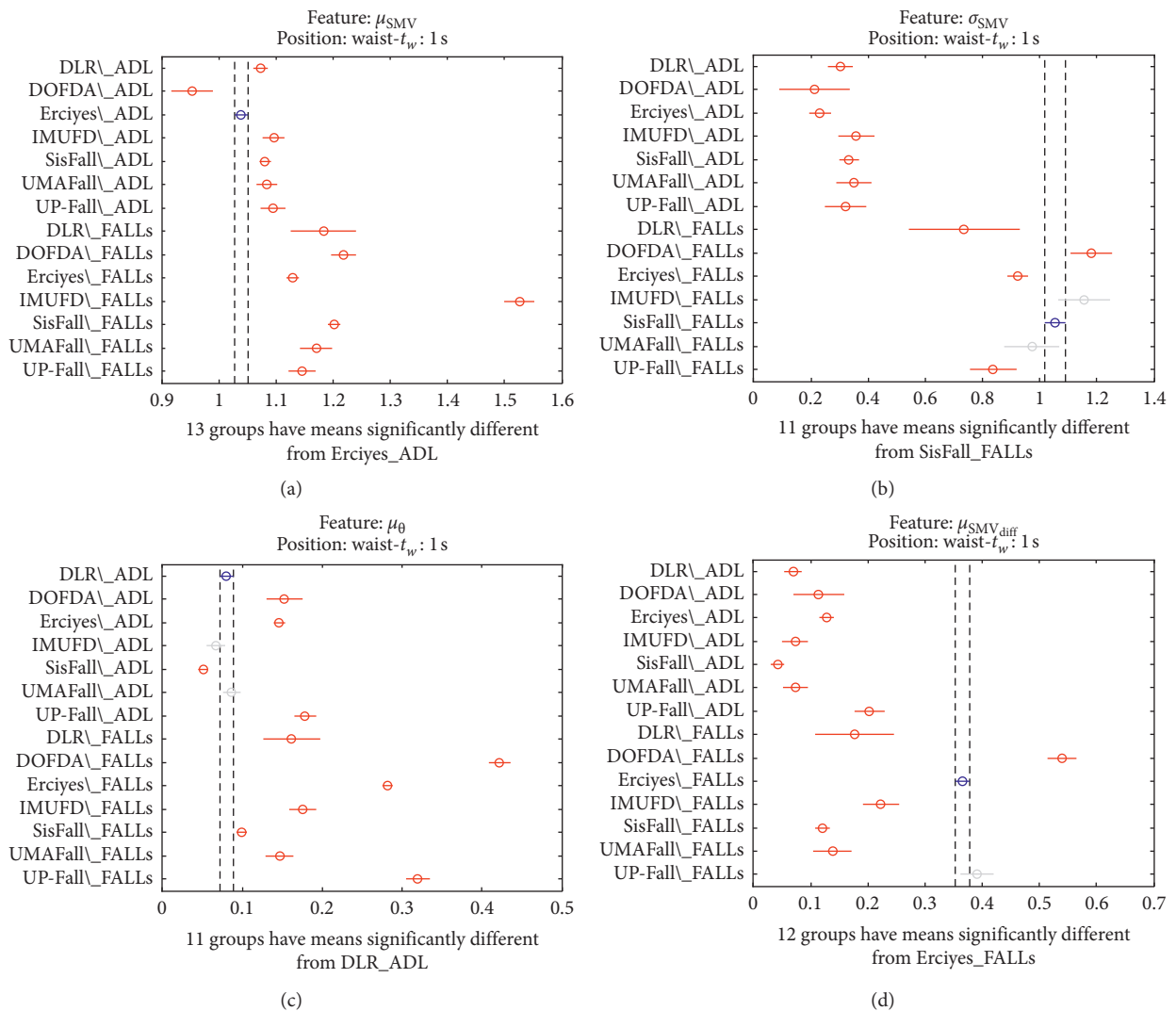


13 groups have means significantly different
from Erciyes_ADL

(a)



11 groups have means significantly different
from SisFall_FALLs

(b)



11 groups have means significantly different
from DLR_ADL

(c)



12 groups have means significantly different
from Erciyes_FALLs

(d)

FIGURE 13: Continued.

Feature: $\mu_{AP}$
Position: waist-$t_w$: 1 s

13 groups have means significantly different
from Erciyes_ADL

(e)

Feature: $A_{\omega\,\mathrm{diff\,(max)}}$
Position: waist-$t_w$: 1 s

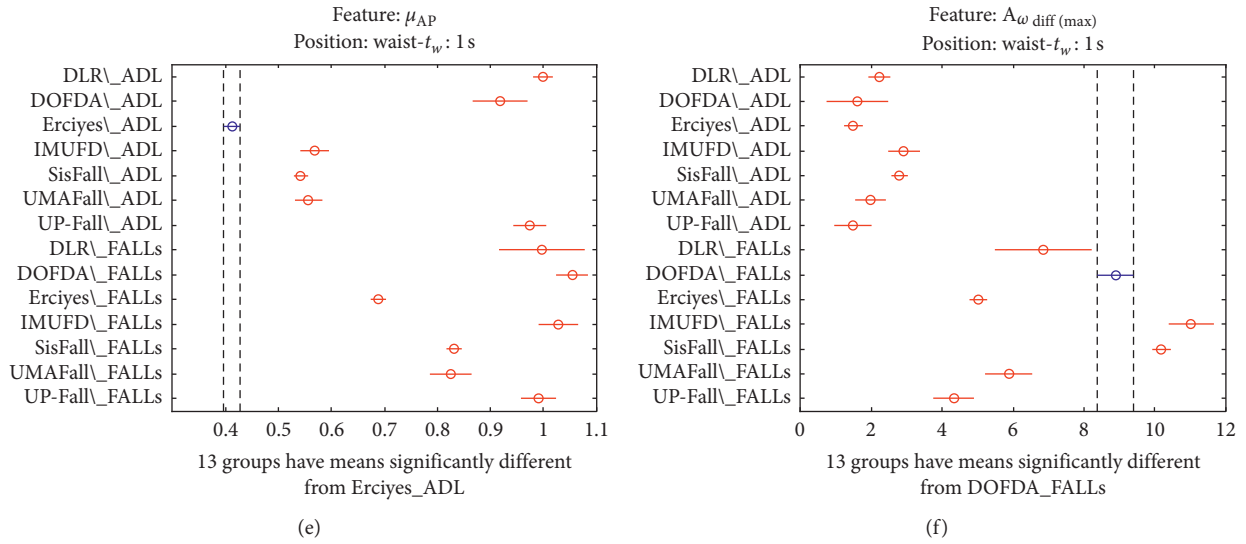13 groups have means significantly different
from DOFDA_FALLs

(f)

Figure 13: Multiple comparison test of the means of the following statistical characteristics of the datasets: (a) $\mu_{\mathbf{SMV}}$. (b) $\sigma_{\mathbf{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\mathbf{SMV}_{(\mathrm{diff})}}$. (e) $\mu_{\mathbf{Ap}}$. (f) $\mathbf{A}_{\omega_{\mathbf{diff(max)}}}$.

Feature: $\mathrm{SMV}_{\max}$
Positiont: waist-$t_w$: 1 s

13 groups have means significantly different
from SisFall_FALLs

(a)

Feature: $\mathrm{SMV}_{\min}$
Position: waist-$t_w$: 1 s

13 groups have means significantly different
from UP-Fall_ADL

(b)

Feature: $\gamma_{\mathrm{SMV}}$
Position: waist-$t_w$: 1 s

12 groups have means significantly different
from SisFall_FALLs

(c)

Feature: $\gamma_{\mathrm{SMV}}$
Position: waist-$t_w$: 1 s

11 groups have means significantly different
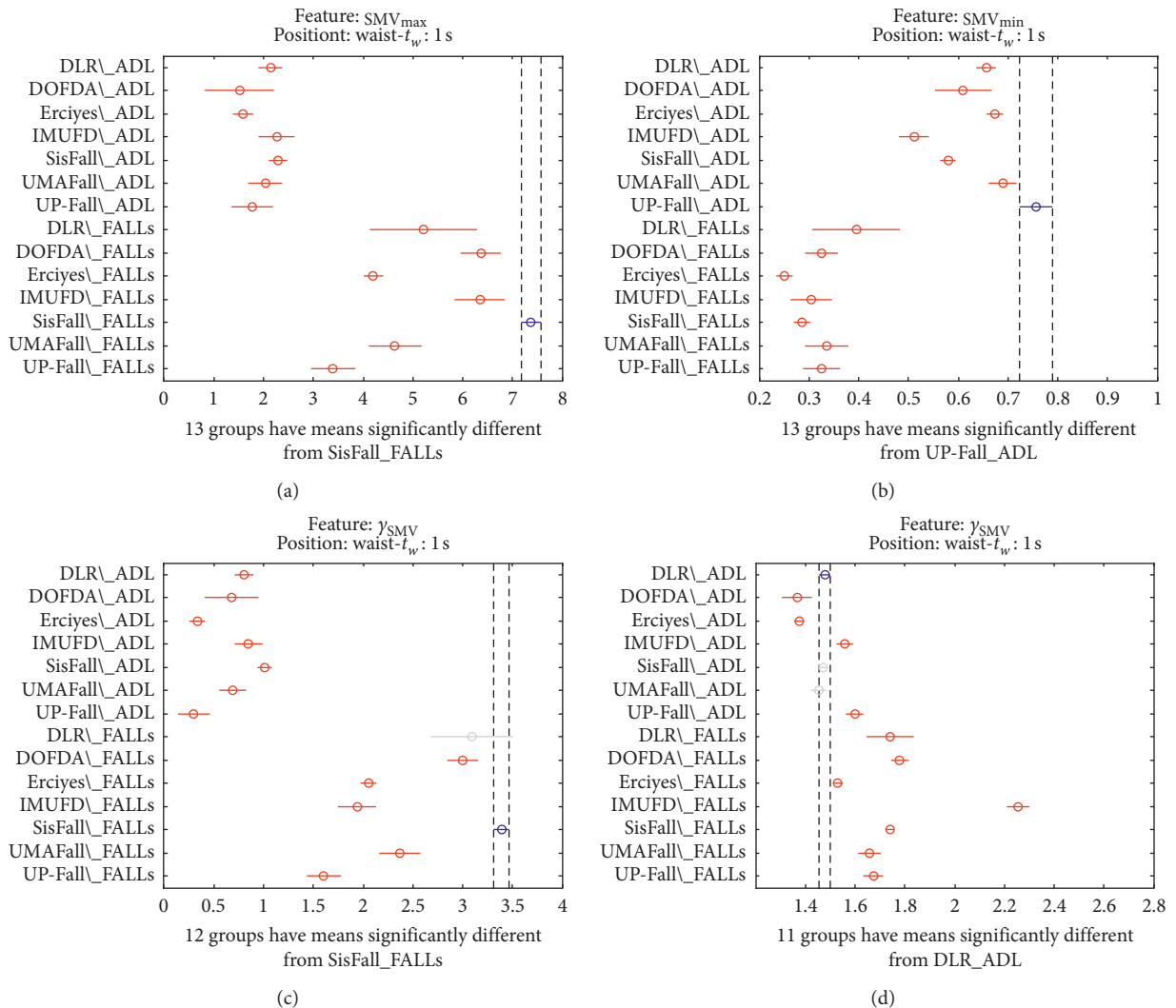from DLR_ADL

(d)
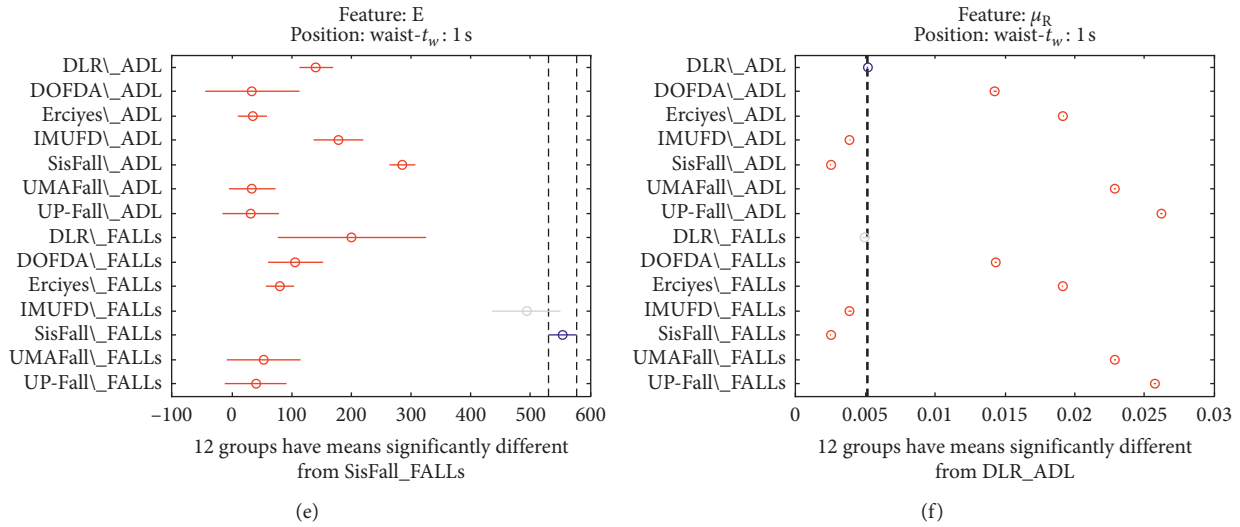
Figure 14: Continued.

(e)



(f)

Figure 14: Multiple comparison test of the means of the following statistical characteristics of the datasets: (a) $SMV_{max}$. (b) $SMV_{min}$. (c) $\gamma_{SMV}$. (d) SMA. (e) **E**. (f) $\mu_R$.
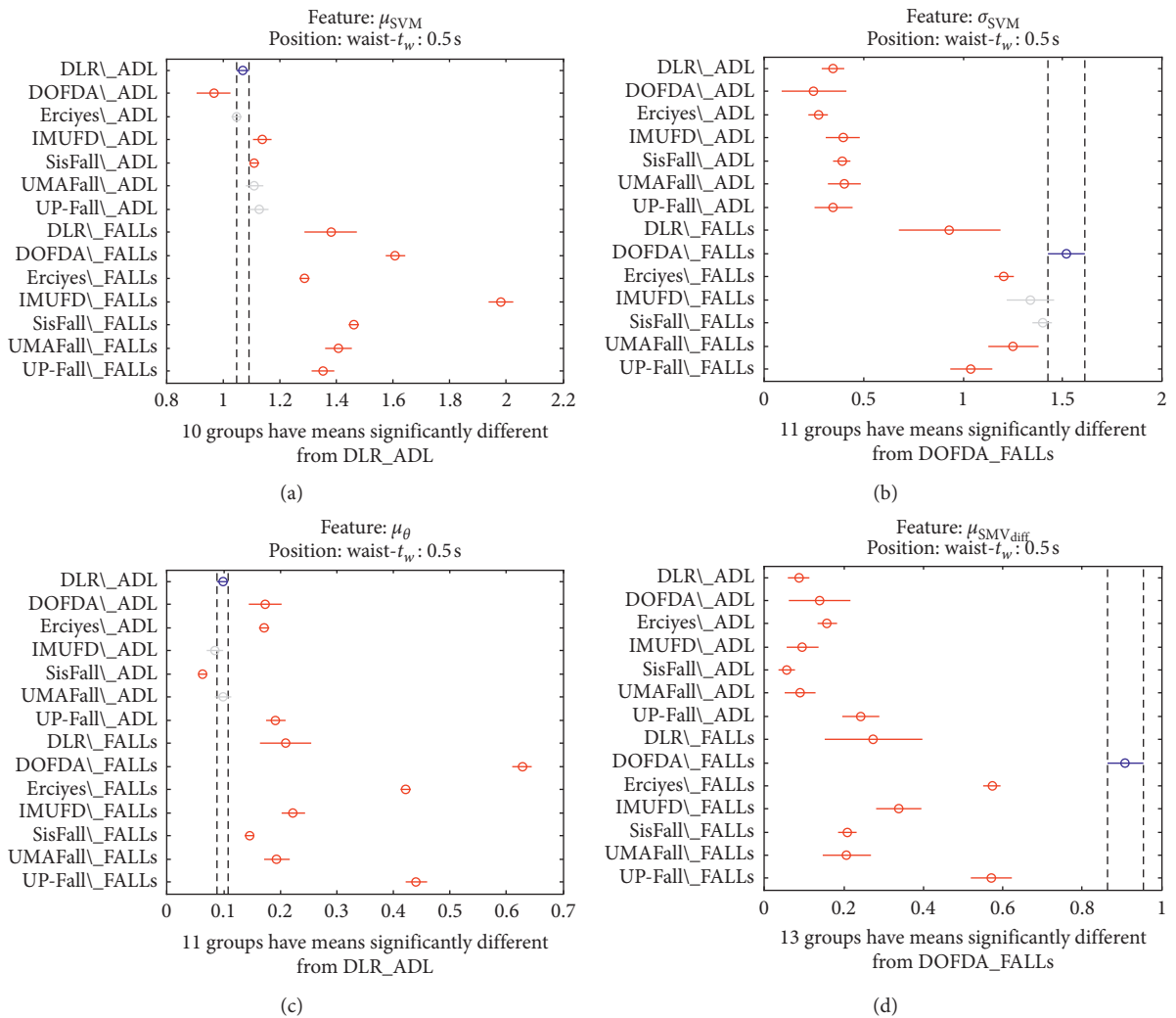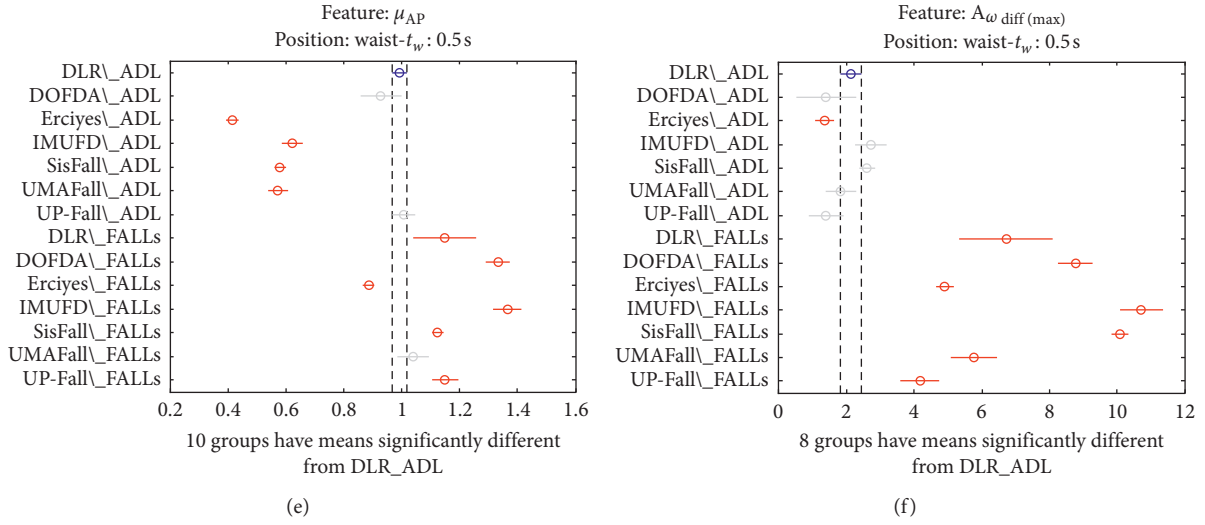


(a)



(b)



(c)



(d)

Figure 15: Continued.

(e)

(f)

FIGURE 15: Multiple comparison test of the means of the statistical characteristics of the datasets: (a) $\mu_{\text{SMV}}$. (b) $\sigma_{\text{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\text{SMV}_{(\text{diff})}}$. (e) $\mu_{\text{Ap}}$. (f) $A_{\omega_{\text{diff(max)}}}$. Observation window of 0.5 s.
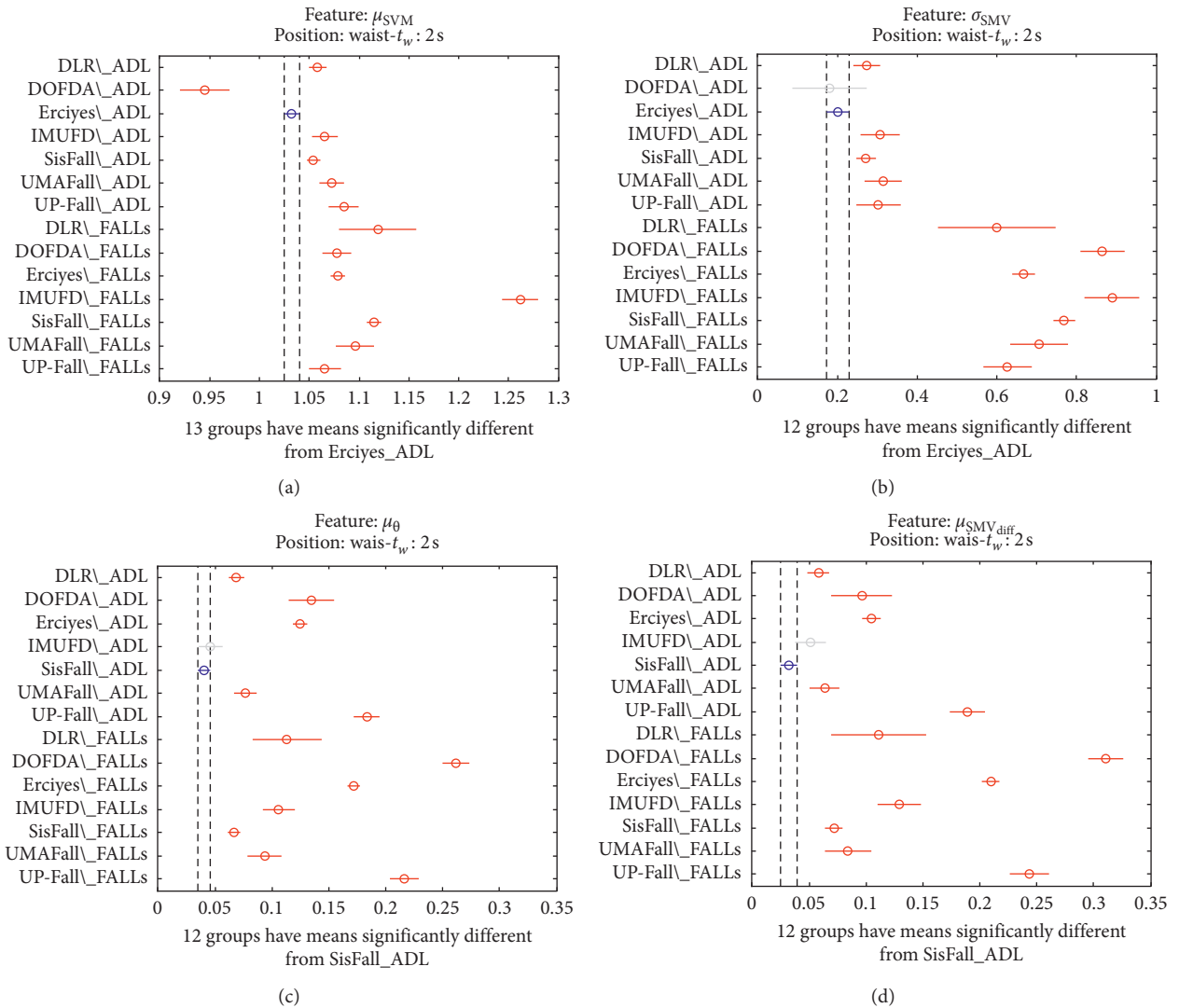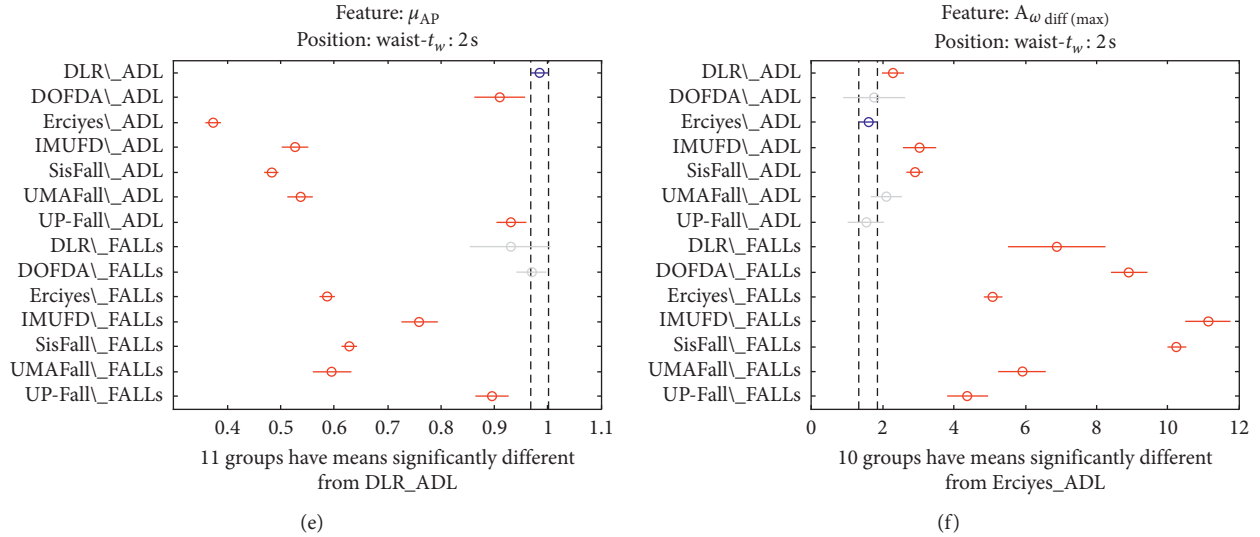


(a)

(b)

(c)

(d)

FIGURE 16: Continued.

(e)



(f)

FIGURE 16: Multiple comparison test of the means of the statistical characteristics of the datasets: (a) $\mu_{\text{SMV}}$. (b) $\sigma_{\text{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\text{SMV}_{(\text{diff})}}$. (e) $\mu_{\text{Ap}}$. (f) $A_{\omega_{\text{diff}(\text{max})}}$. Observation window of 2 s.

TABLE 4: Distribution of the activities in the traces among the different considered general types of movements.

| Dataset | DLR | DOFDA | Erciyes U. | IMUFD | SisFall | UMAFall | UP-fall |
|---|---|---|---|---|---|---|---|
| Number of types of ADLs/falls | 15/1 | 5/13 | 16/20 | 8/7 | 19/15 | 12/3 | 6/5 |
| BASIC MOVEMENTS | 4 | 2 | 8 | 4 | 11 | 7 | 3 |
| Standing | 1 | | | 1 | | | 1 |
| Rising/descending from(to) lying/kneeling | 1 | 1 | 1 | 1 | 2 | 1 | |
| Lying | 1 | | 1 | | 1 | | 1 |
| Descending to sitting/rising from sitting | 1 | 1 | 4 | 2 | 8 | 1 | 1 |
| Bending | | | 1 | | | 1 | |
| Hand movements (making a call and applauding) | | | | | | 4 | |
| Others | | | 1 | | | | |
| STANDARD MOVEMENTS | 4 | 3 | 5 | 4 | 4 | 3 | 2 |
| Walking | 1 | 1 | 2 | 1 | 2 | 1 | 1 |
| Going down | 1 | | | | | | |
| Climbing stairs (up and/or down) | 2 | 1 | | 2 | 2 | 2 | |
| Picking | | 1 | 1 | 1 | | | 1 |
| Others | | | 2 | | | | |
| SPORTING MOVEMENTS | 7 | | 1 | | 3 | 2 | 1 |
| Running/Jogging | 1 | | 1 | | 2 | 2 | |
| Jumping/Hopping | 4 | | | | 1 | | 1 |
| Others | 2 | | | | | | |
| NEAR FALLS | | | 2 | | 1 | | |
| Stumble | | | 1 | | 1 | | |
| Trip | | | 1 | | | | |
| FALLS | 1 | 13 | 20 | 7 | 15 | 3 | 5 |
| Backwards | | 4 | 4 | | 2 | 1 | 1 |
| Forward/Frontal | | 4 | 8 | | 2 | 1 | 2 |
| Lateral | | 4 | 4 | | 2 | 1 | 1 |
| Slipping | | | | 1 | 3 | | |
| Tripping/hitting/bumping | | | | 2 | 2 | | |
| Missteps | | | | 3 | | | |
| Syncope/Fainting/collapse | | | 1 | 2 | 1 | 4 | |
| Others | 1 | | 2 | | | | 1 |

sensor on the waist. As it can be appreciated from the figures, even for a common physical activity as walking, the characteristics show noteworthy discrepancies among the datasets. Figures show that there are only three

characteristics ($\sigma_{\text{SMV}}$, $A_{\omega_{\text{diff}(\text{max})}}$, and $\text{SMV}_{\text{max}}$) for which the null hypothesis cannot be rejected as long as no dataset exhibits a mean that can be considered significantly different from those computed for other databases. For some
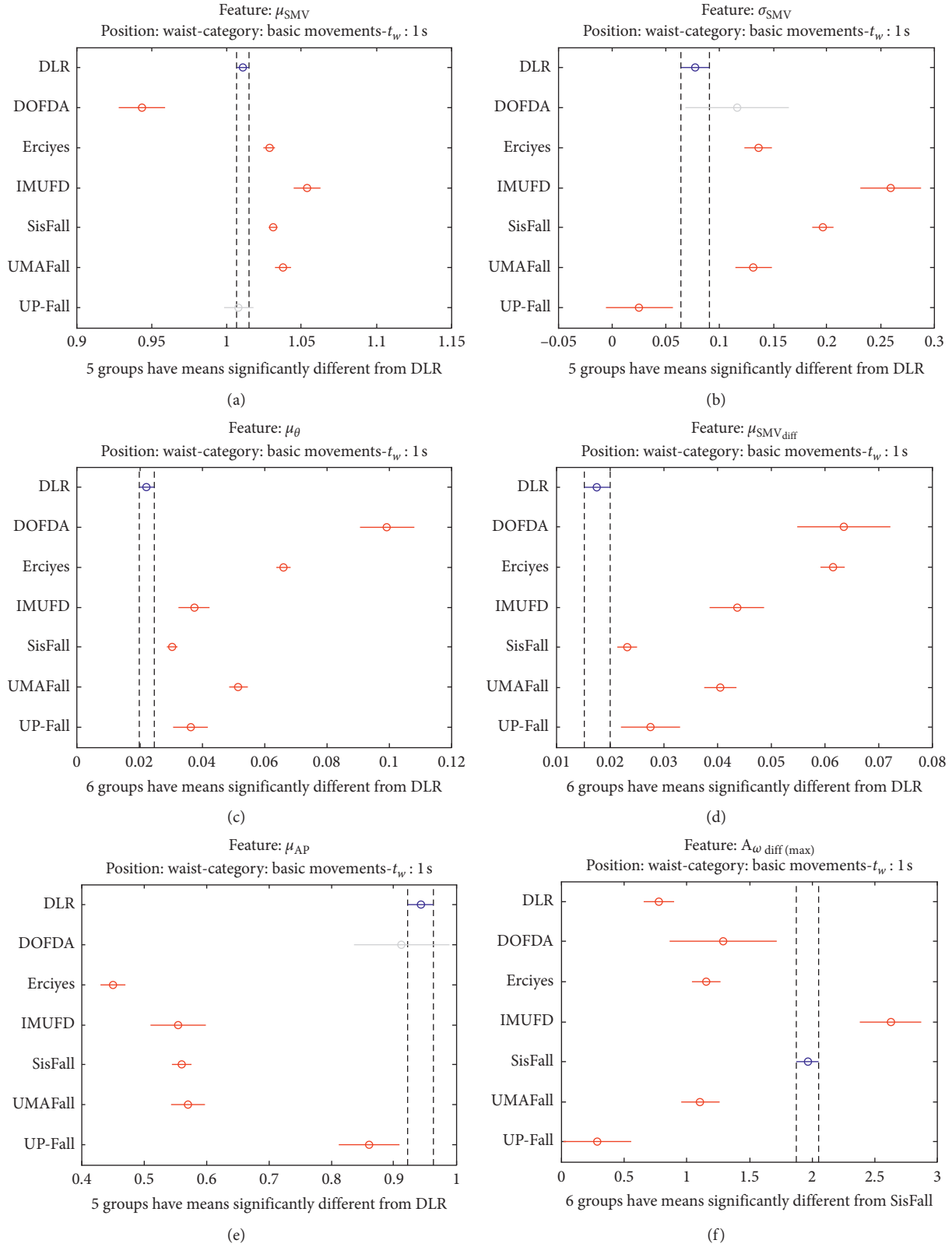
FIGURE 17: Multiple comparison test of the means of the statistical characteristics of the basic movements in the datasets: (a) $\mu_{\text{SMV}}$. (b) $\sigma_{\text{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\text{SMV}_{(\text{diff})}}$. (e) $\mu_{\text{Ap}}$. (f) $\text{A}_{\omega_{\text{diff}(\text{max})}}$. Basic movements.
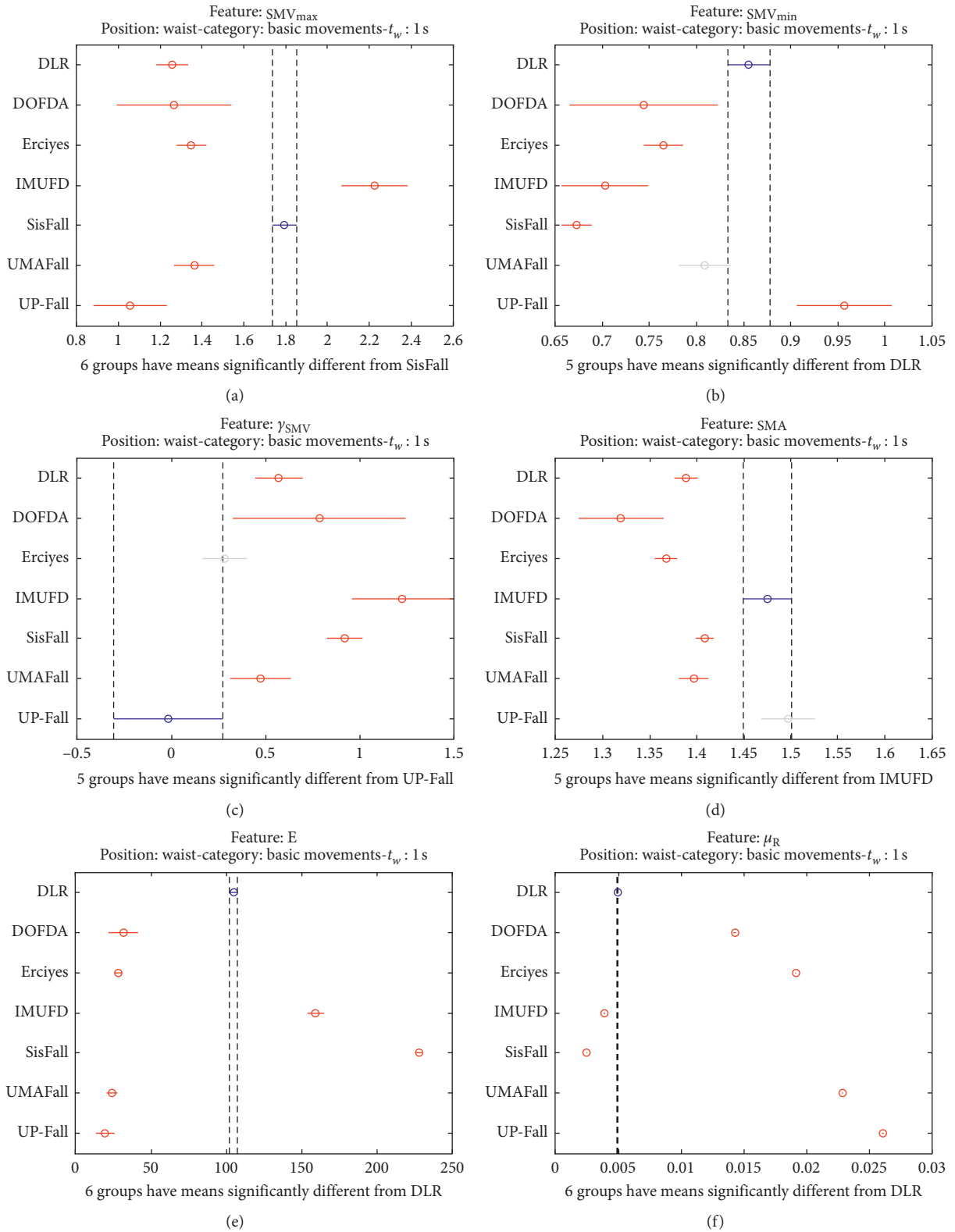
FIGURE 18: Multiple comparison test of the means of the statistical characteristics of the basic movements in the datasets: (a) $\mathbf{SMV_{max}}$. (b) $\mathbf{SMV_{min}}$. (c) $\gamma_{\mathbf{SMV}}$. (d) SMA. (e) $\mathbf{E}$. (f) $\mu_{\mathbf{R}}$. Basic movements.
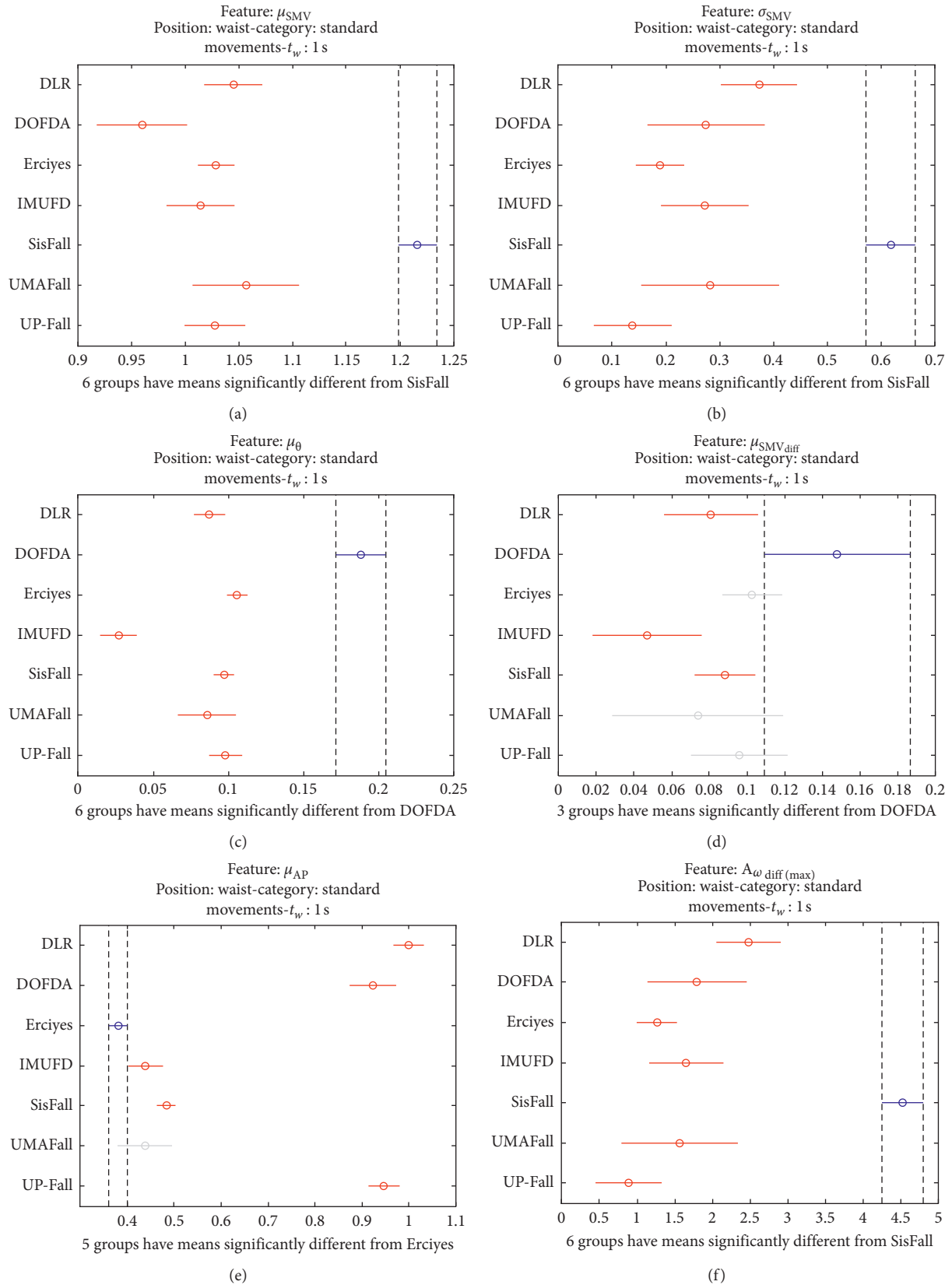
Figure 19: Multiple comparison test of the means of the statistical characteristics of the standard movements in the datasets: (a) $\mu_{SMV}$. (b) $\sigma_{SMV}$. (c) $\mu_{\theta}$. (d) $\mu_{SMV_{(diff)}}$. (e) $\mu_{Ap}$. (f) $A_{\omega_{diff(max)}}$. Standard movements.
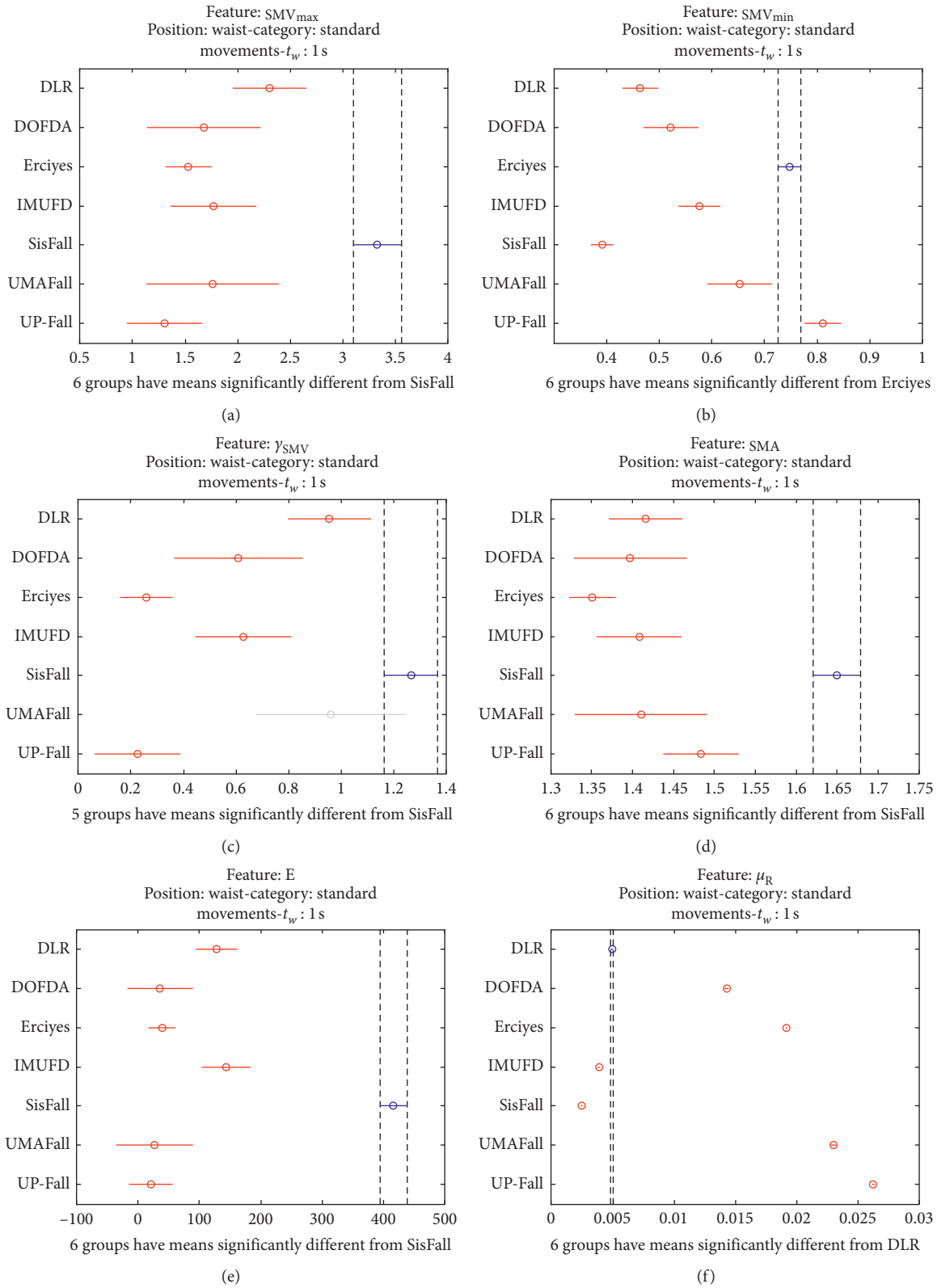
Figure 20: Multiple comparison test of the means of the statistical characteristics of the standard movements in the datasets: (a) $SMV_{max}$. (b) $SMV_{min}$. (c) $\gamma_{SMV}$. (d) SMA. (e) E. (f) $\mu_R$. Standard movements.

Figure 21: Multiple comparison test of the means of the statistical characteristics of the sporting movements in the datasets: (a) $\mu_{SMV}$. (b) $\sigma_{SMV}$. (c) $\mu_\theta$. (d) $\mu_{SMV_{(diff)}}$. (e) $\mu_{Ap}$. (f) $A_{\omega_{diff(max)}}$. Sporting movements.
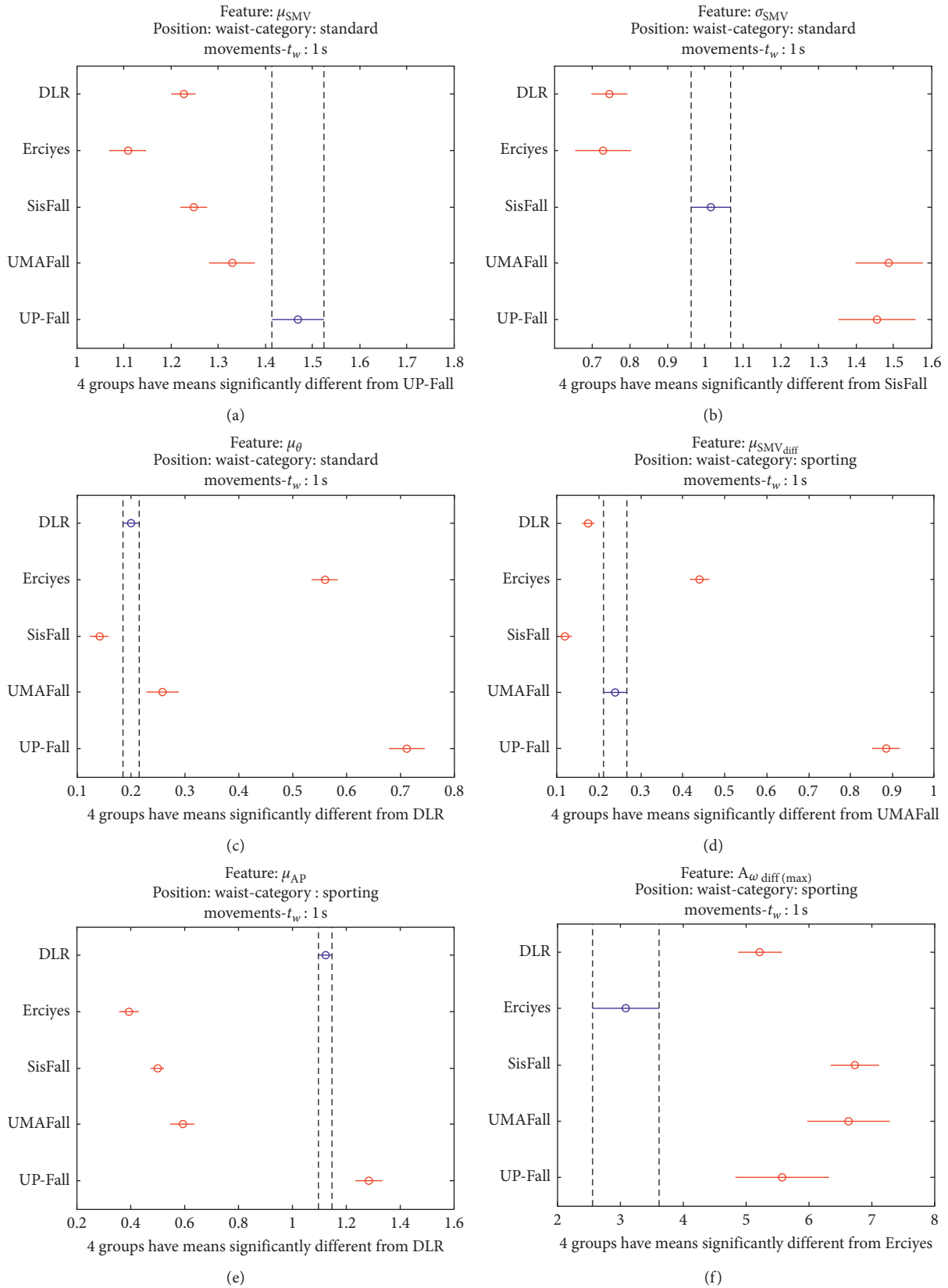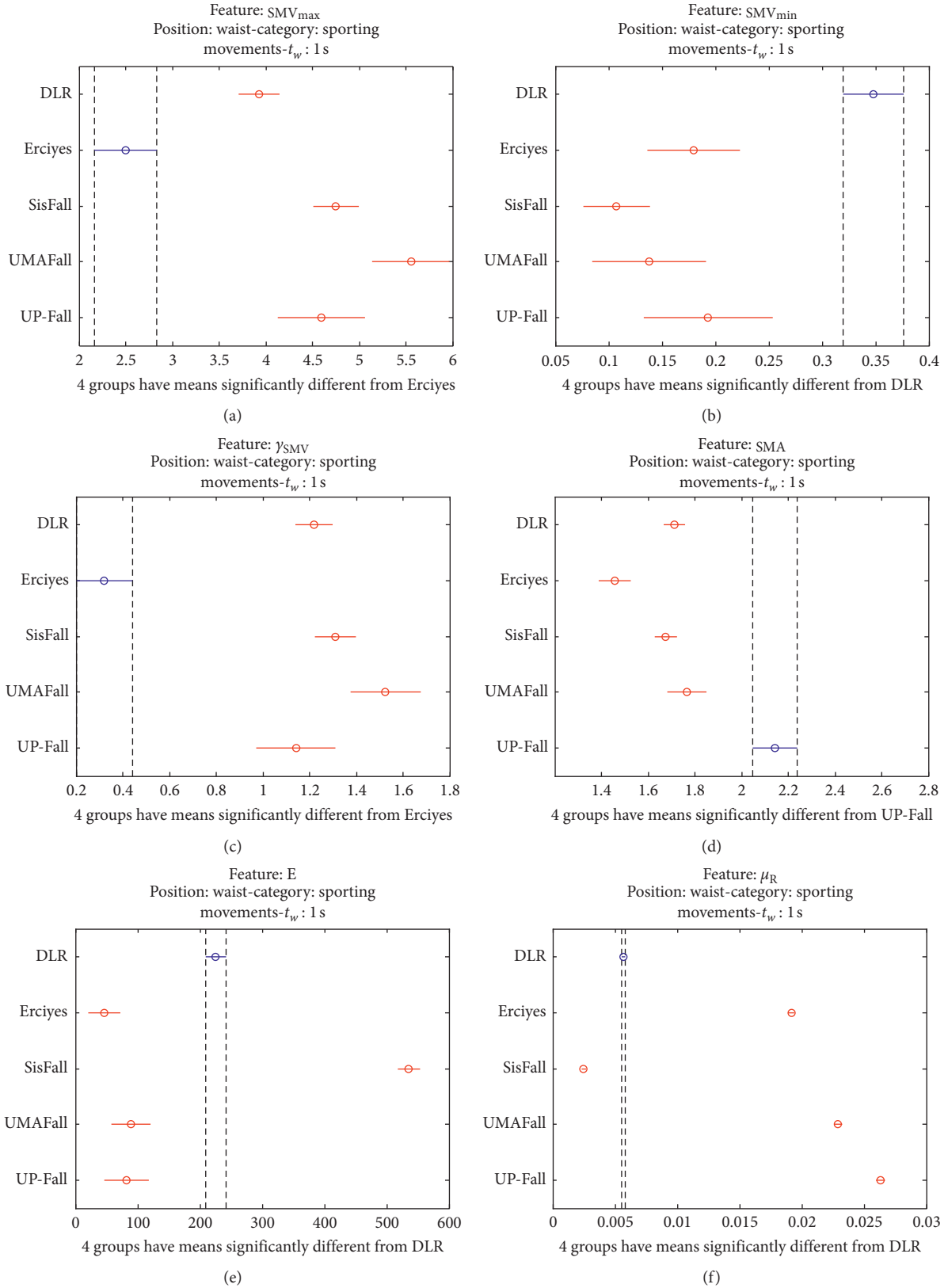
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 22: Multiple comparison test of the means of the statistical characteristics of the sporting movements in the datasets: (a) **SMV_max**. (b) **SMV_min**. (c) $\gamma_{SMV}$. (d) SMA. (e) **E**. (f) $\mu_R$. Sporting movements.

FIGURE 23: Multiple comparison test of the means of the statistical characteristics of the "near falls" in the two datasets that contain this type of movement: (a) $\mu_{SMV}$. (b) $\sigma_{SMV}$. (c) $\mu_\theta$. (d) $\mu_{SMV_{(diff)}}$. (e) $\mu_{Ap}$. (f) $A_{\omega_{diff(max)}}$.

Feature: $\text{SMV}_{\text{max}}$
Position: waist-category: near falls-$t_w$ : 1 s

The means of groups Erciyes and IMUFD
are significantly different

(a)

Feature: $\text{SMV}_{\text{min}}$
Position: waist-category: near falls-$t_w$ : 1 s

The means of groups Erciyes and IMUFD
are significantly different

(b)

Feature: $\gamma_{\text{SMV}}$
Position: waist-category: near falls-$t_w$ : 1 s

No groups have means significantly different
from Erciyes

(c)

Feature: $\text{SMA}$
Position: waist-category: near falls-$t_w$ : 1 s

The means of groups Erciyes and IMUFD
are significantly different

(d)

Feature: E
Position: waist-category : near falls-$t_w$ : 1 s

The means of groups Erciyes and IMUFD
are significantly different

(e)

Feature: $\mu_{\text{R}}$
Position: waist-category: near falls-$t_w$ : 1 s

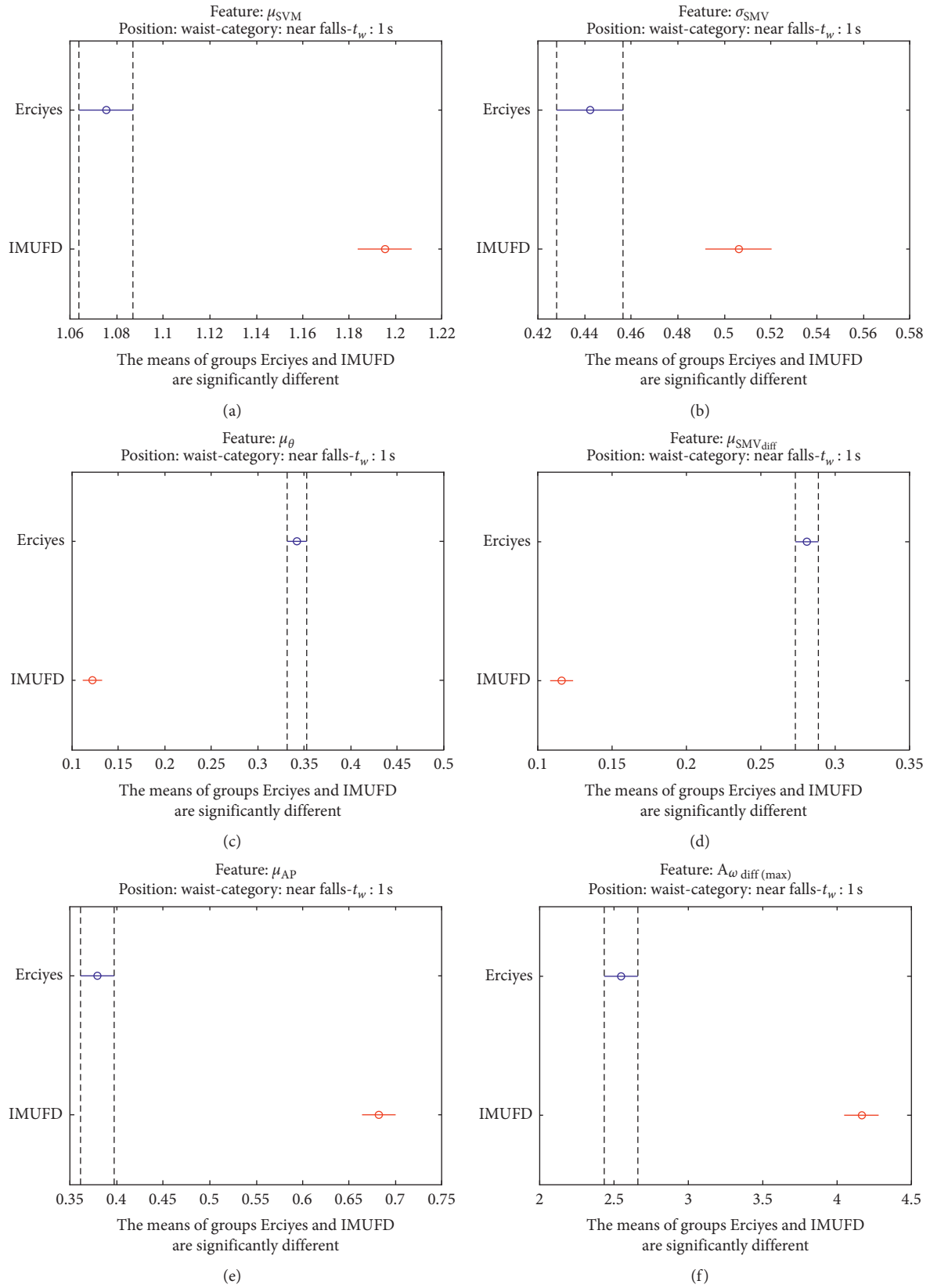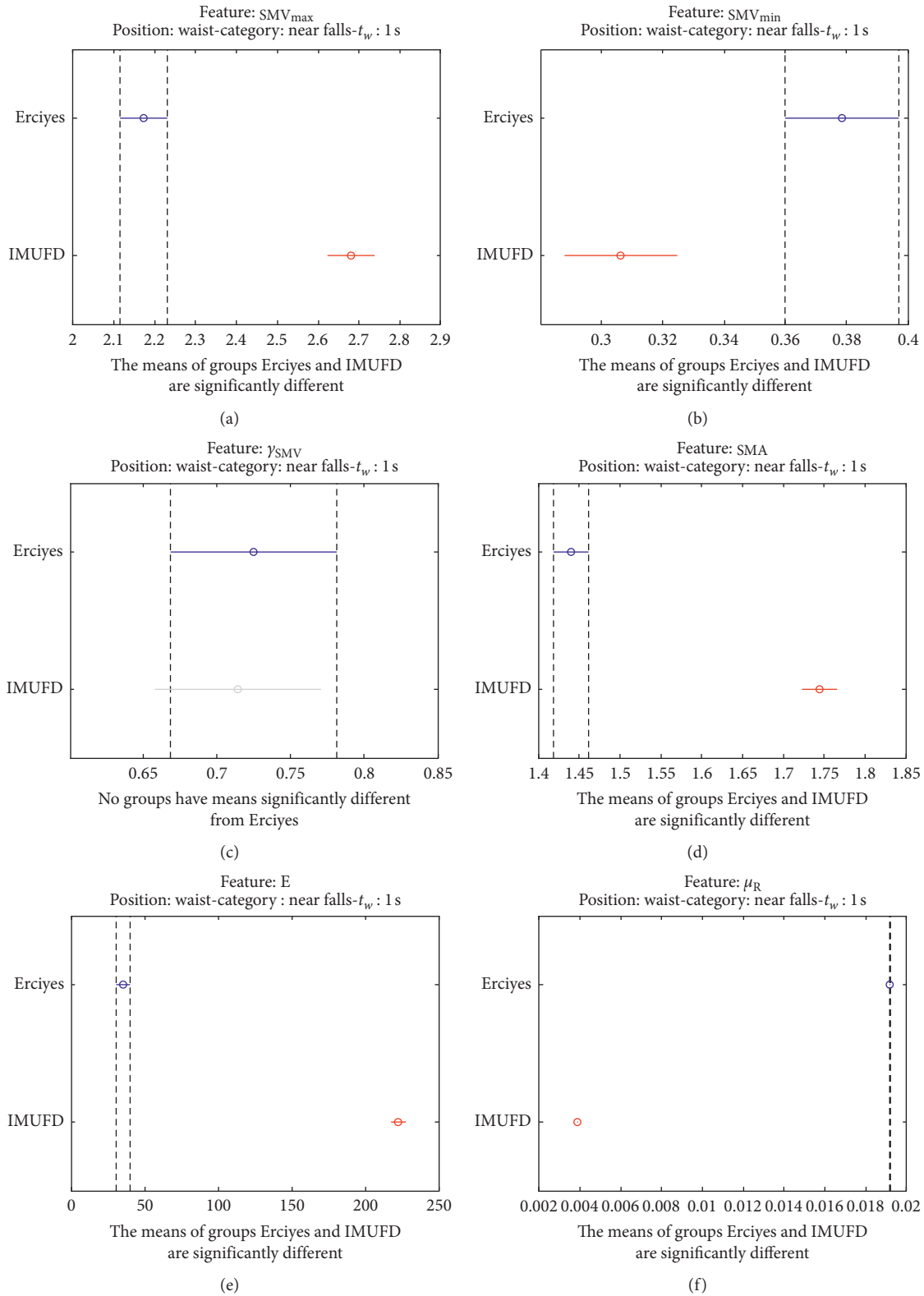The means of groups Erciyes and IMUFD
are significantly different

(f)

FIGURE 24: Multiple comparison test of the means of the statistical characteristics of the "near falls" in the two datasets that contain this type of movement: (a) $\textbf{SMV}_{\textbf{max}}$. (b) $\textbf{SMV}_{\textbf{min}}$. (c) $\gamma_{\textbf{SMV}}$. (d) SMA. (e) $\textbf{E}$. (f) $\mu_{\textbf{R}}$. Near falls.
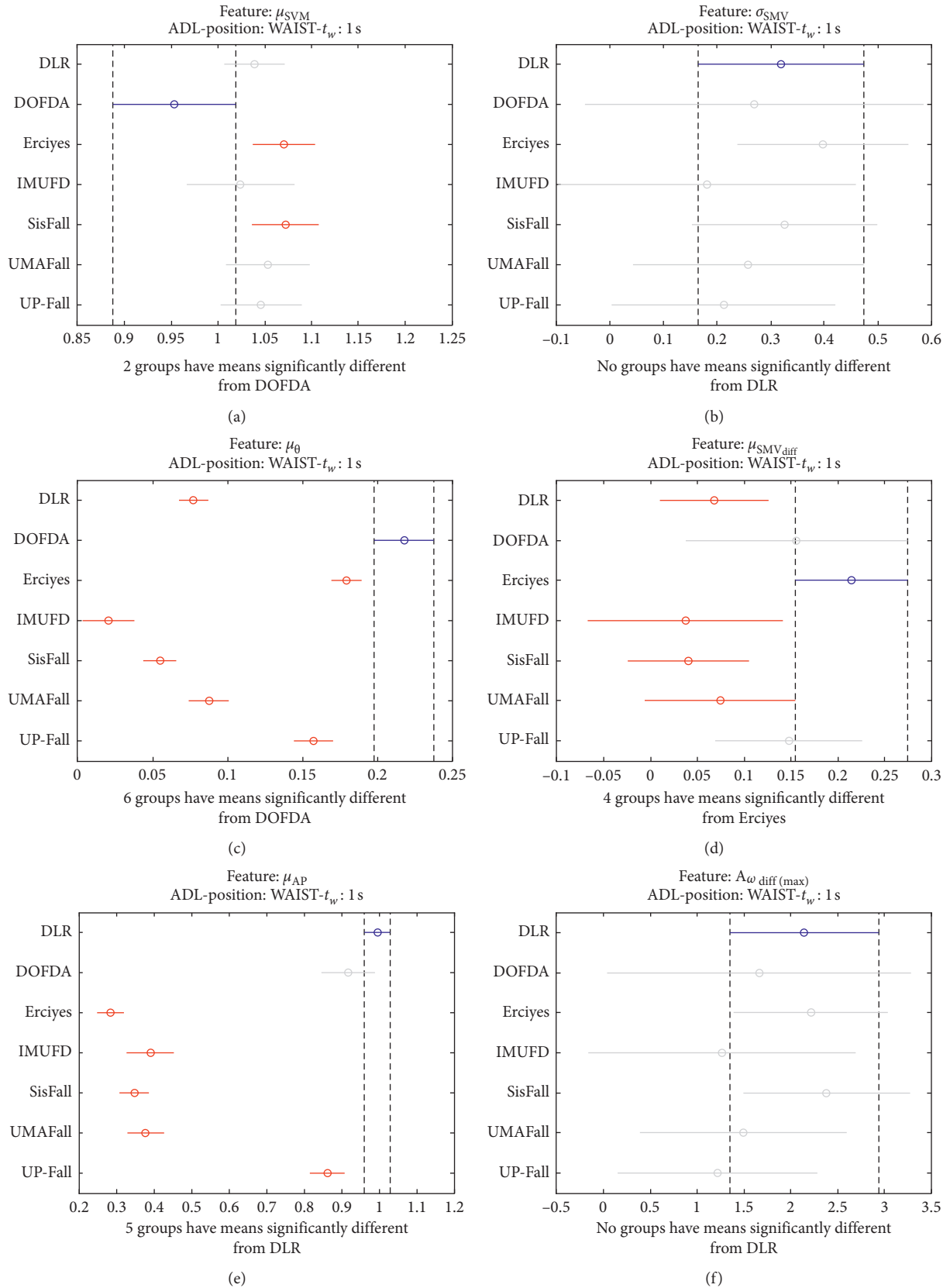
FIGURE 25: Multiple comparison test of the means of the statistical characteristics of the movements labelled as "walking" in the seven datasets that contain this type of ADL (sensor located on the waist): (a) $\mu_{\text{SMV}}$. (b) $\sigma_{\text{SMV}}$. (c) $\mu_\theta$. (d) $\mu_{\text{SMV}_{(\text{diff})}}$. (e) $\mu_{\text{Ap}}$. (f) $A_{\omega_{\text{diff}(\text{max})}}$.
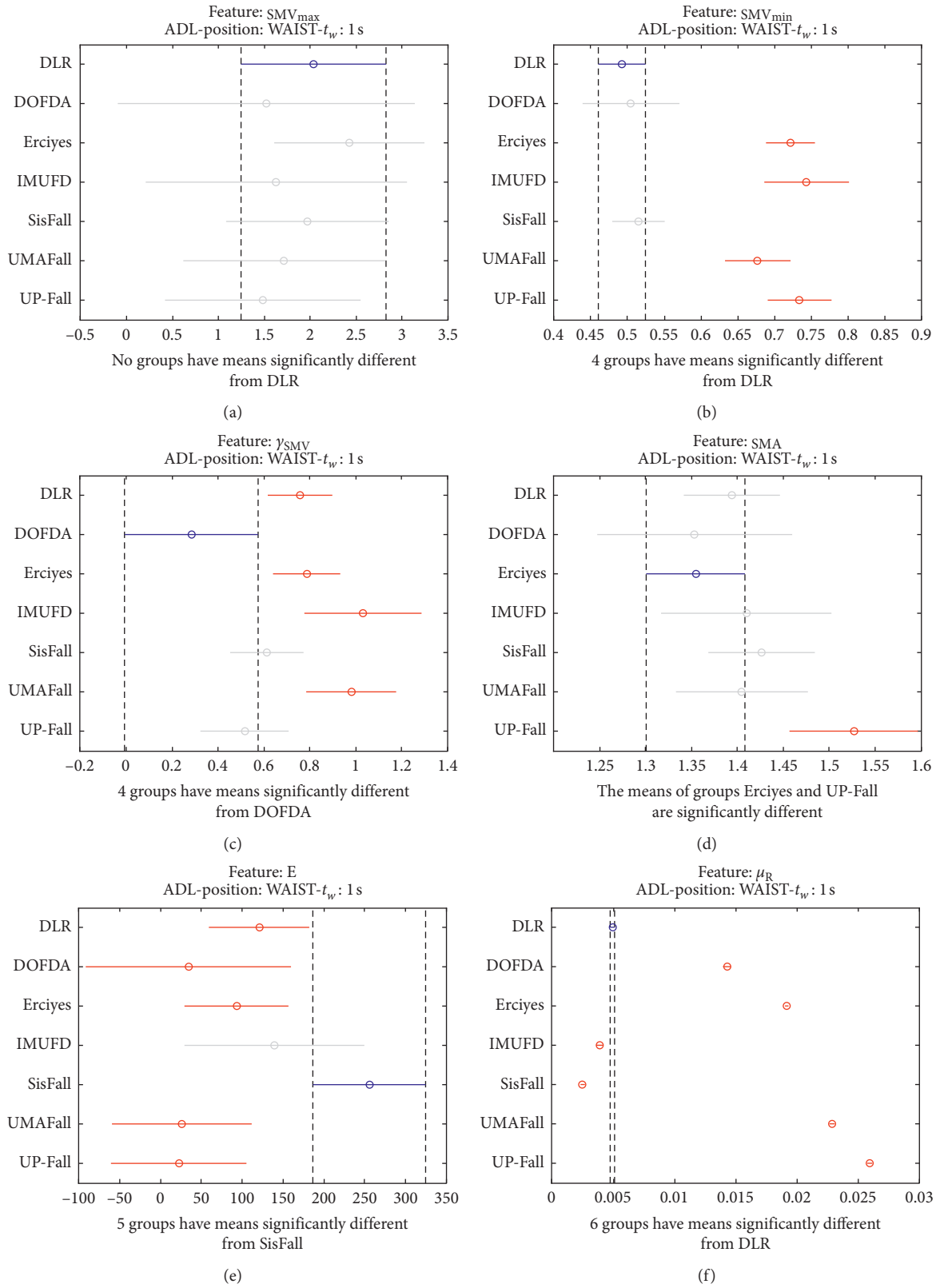
Figure 26: Multiple comparison test of the means of the statistical characteristics of the movements labelled as "walking" in the seven datasets that contain this type of ADL (sensor located on the waist): (a) $\mathbf{SMV_{max}}$. (b) $\mathbf{SMV_{min}}$. (c) $\gamma_{\mathbf{SMV}}$. (d) SMA. (e) $\mathbf{E}$. (f) $\mu_{\mathbf{R}}$.
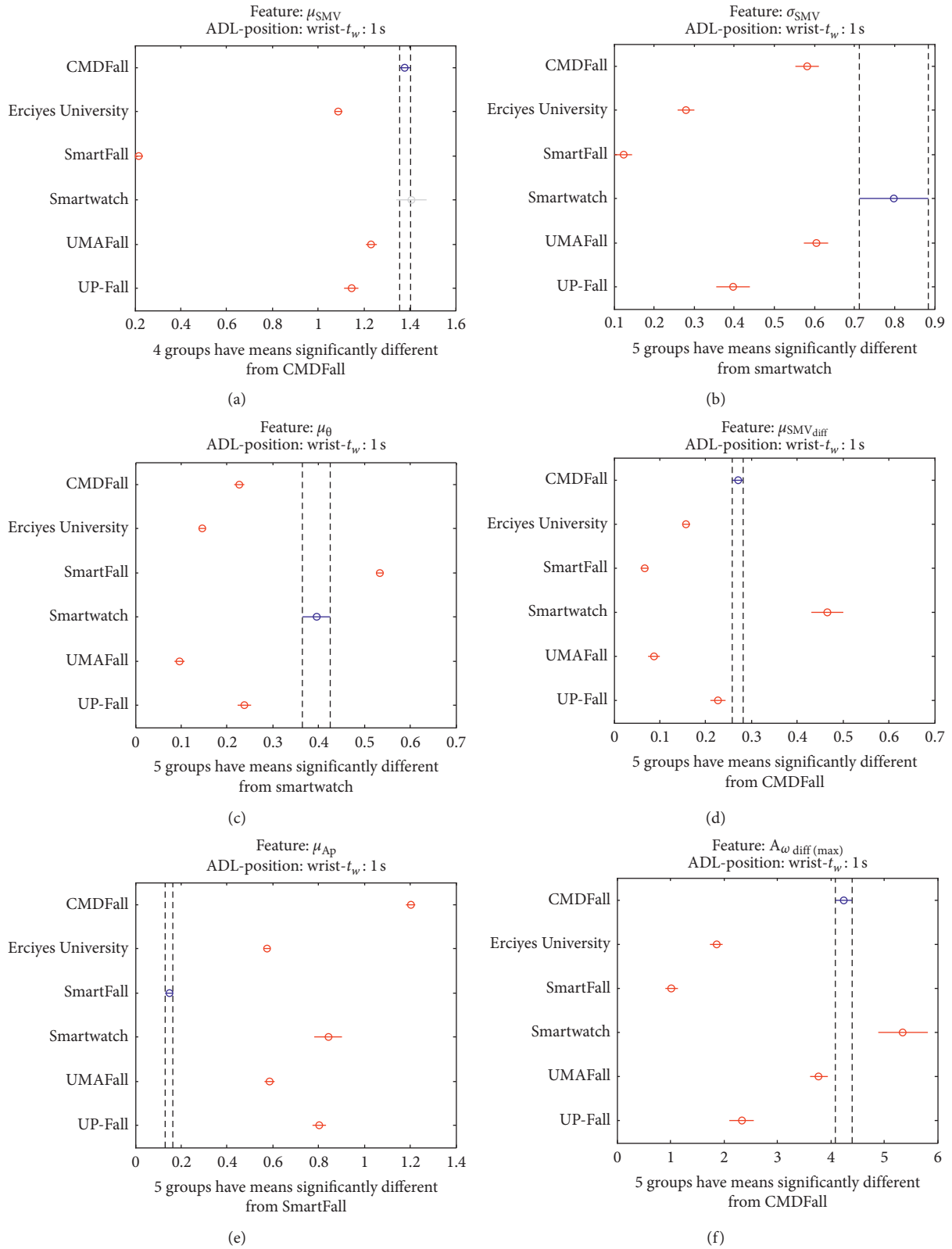
Figure 27: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a) $\mu_{\text{SMV}}$. (b) $\sigma_{\text{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\text{SMV}_{(\text{diff})}}$. (e) $\mu_{\text{Ap}}$. (f) $A_{\omega_{\text{diff(max)}}}$.

Feature: $SMV_{max}$
ADL-position: wrist-$t_w$: 1 s

5 groups have means significantly different
from CMDFall

(a)

Feature: $SMV_{min}$
ADL-position: wrist-$t_w$: 1 s

5 groups have means significantly different
from smartwatch

(b)

Feature: $\gamma_{SMV}$
ADL-position: wrist-$t_w$: 1 s

4 groups have means significantly different
from CMDFall

(c)

Feature: SMA
ADL-position: wrist-$t_w$: 1 s

5 groups have means significantly different
from UMAFall

(d)

Feature: E
ADL-position: wrist-$t_w$: 1 s

5 groups have means significantly different
from UMAFall

(e)

Feature: $\mu_R$
ADL-position: wrist-$t_w$: 1 s

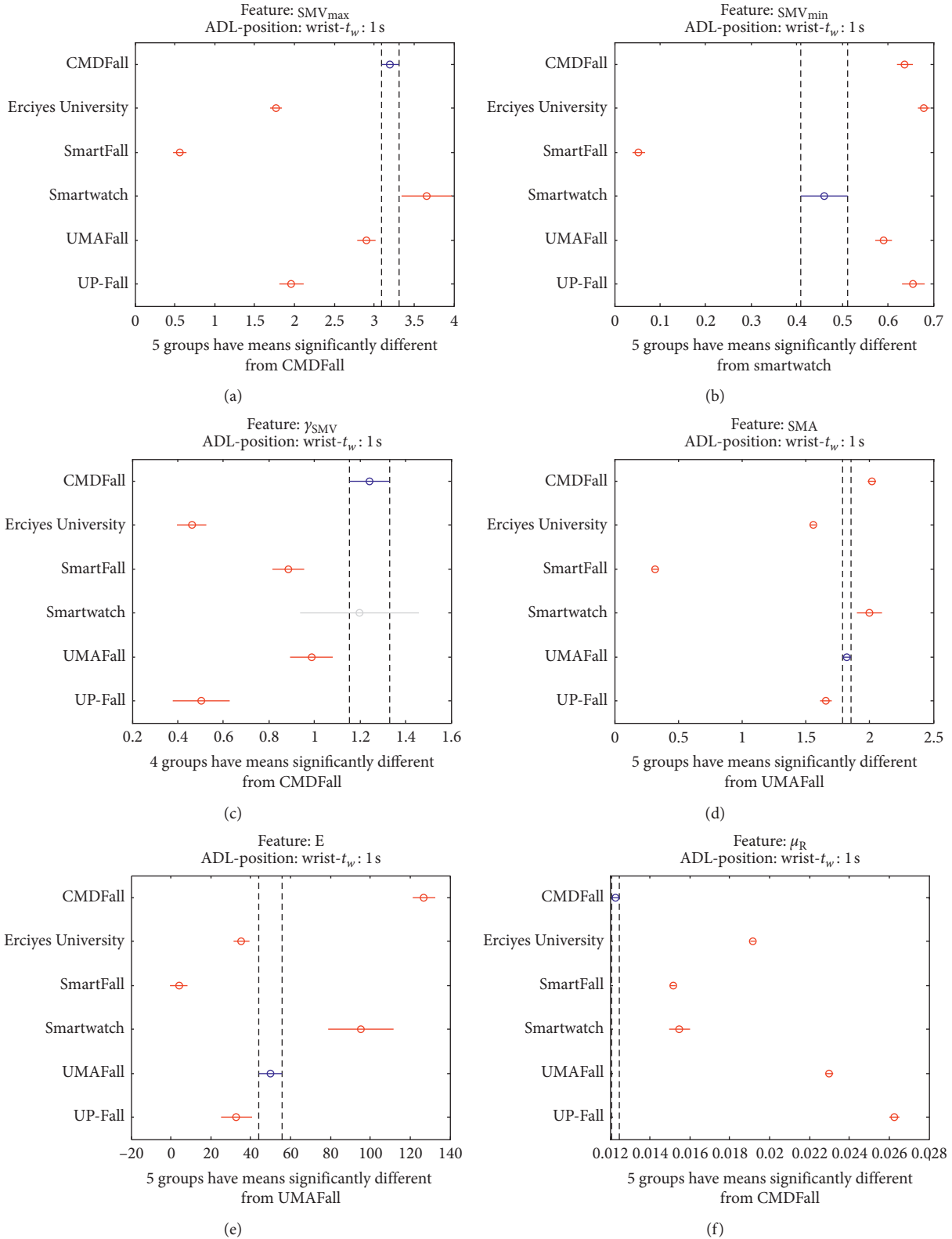5 groups have means significantly different
from CMDFall

(f)

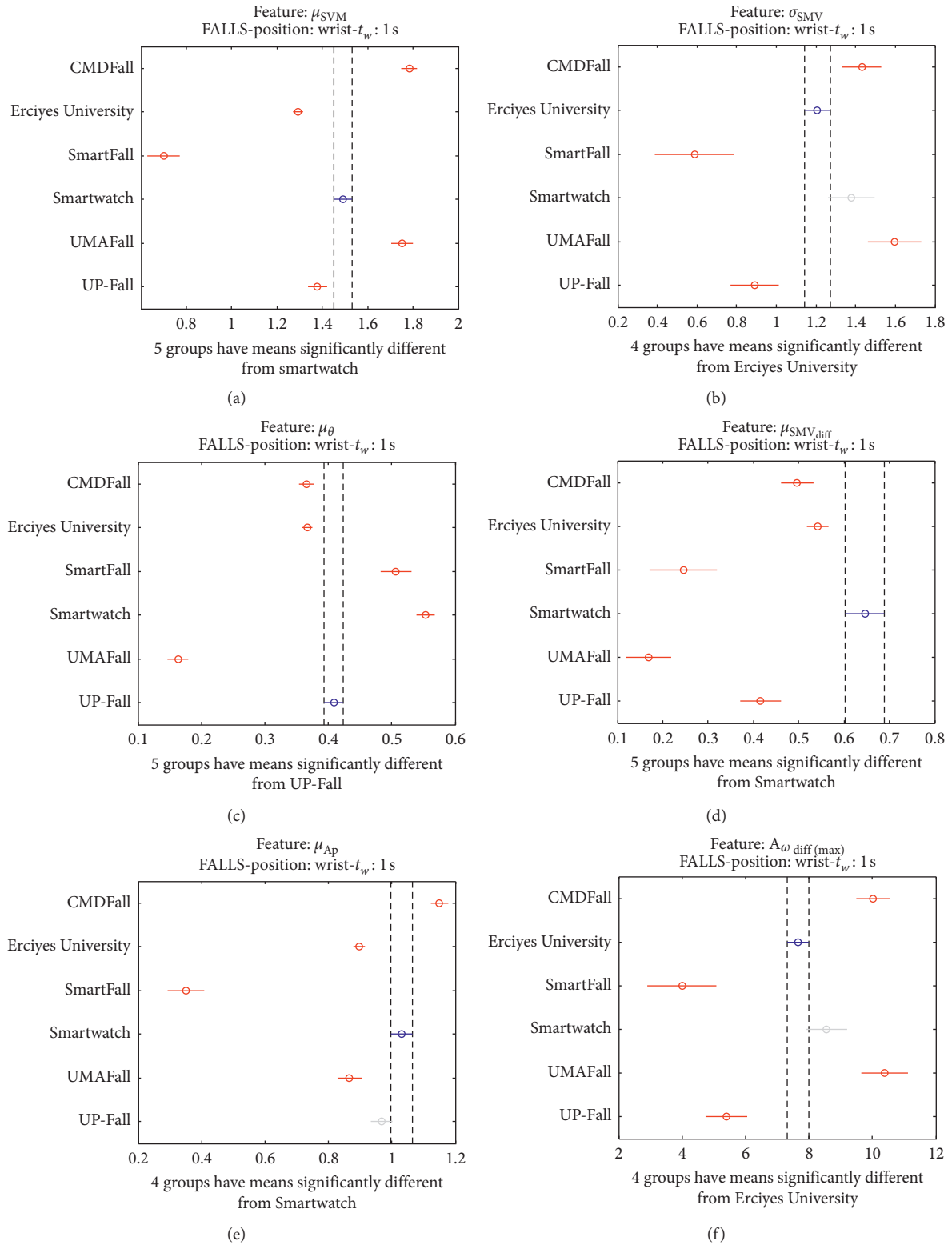FIGURE 28: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a) $\mathbf{SMV_{max}}$. (b) $\mathbf{SMV_{min}}$. (c) $\gamma_{\mathbf{SMV}}$. (d) SMA. (f) $\mathbf{E}$. (f) $\mu_{\mathbf{R}}$.
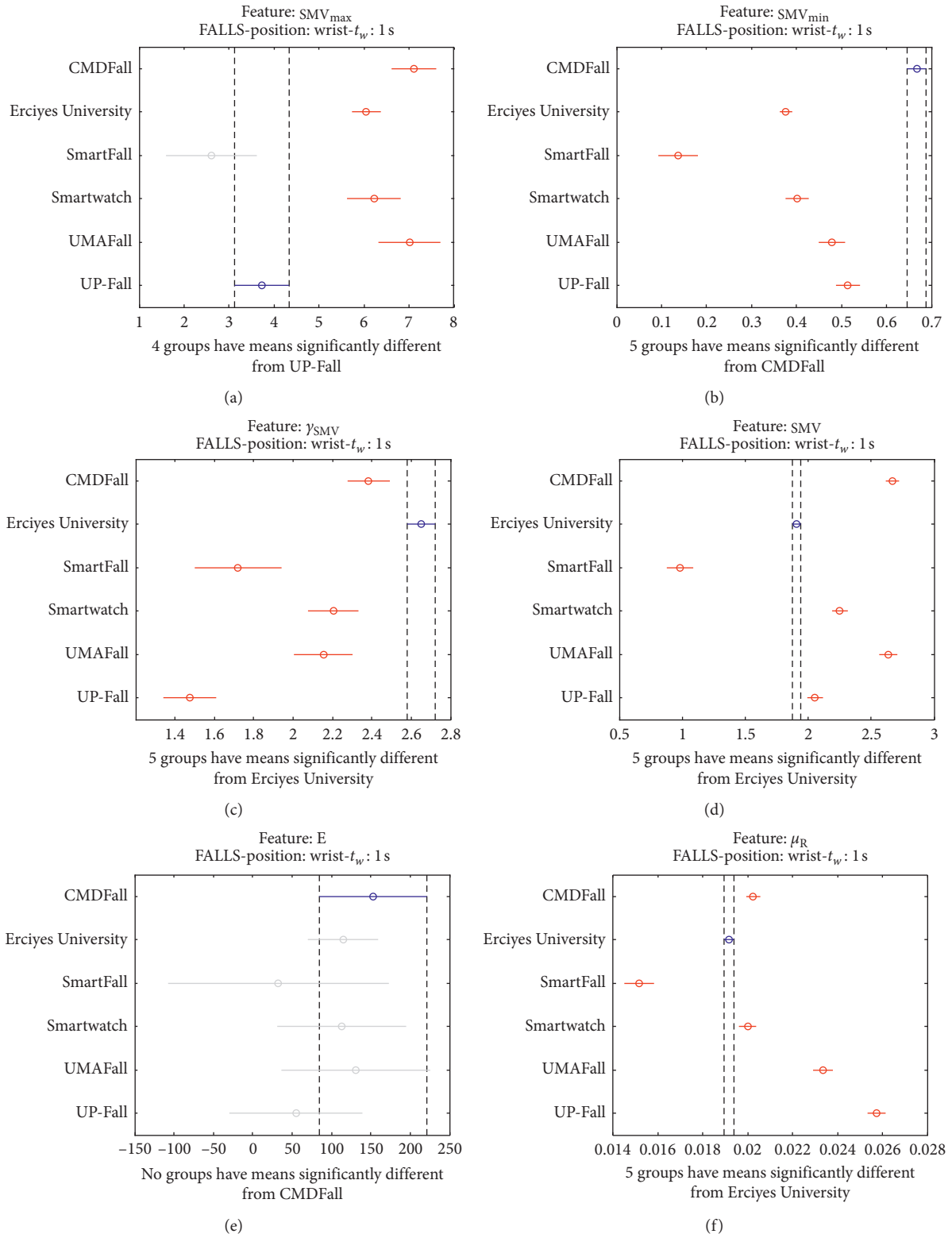
FIGURE 29: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and fall movements: (a) $\mu_{\mathbf{SMV}}$. (b) $\sigma_{\mathbf{SMV}}$. (c) $\mu_{\theta}$. (d) $\mu_{\mathbf{SMV_{(diff)}}}$. (e) $\mu_{\mathbf{Ap}}$. (f) $\mathbf{A}_{\omega_{\mathbf{diff(max)}}}$.

FIGURE 30: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a) **SMV$_{max}$**. (b) **SMV$_{min}$**. (c) $\gamma_{SMV}$. (d) SMA. (f) **E**. (f) $\mu_R$.

characteristics (for example, note the absence of overlapping intervals in the graphs corresponding to $\mu_\theta$ or $\mu_R$) the post hoc tests show that all or almost all datasets are significantly different.

*4.3. Results for the Measurements on the Wrist.* To corroborate the previous results, we apply the previous analysis to the datasets containing measurements captured on a completely different body position: the wrist. In spite of the particular (and independent) mobility of the wrist, this position has been selected in a significant number of studies on FDSs as the position to locate the detection sensor. The wrist offers to the user better ergonomics than other typical placements as humans are already habituated to wear watches. Moreover, commercial smartwatches (which are natively provided with inertial measurement units) can be employed to deploy the FDS without obliging the user to transport any supplementary device. In some articles that consider systems with more than one sensing mote, the wrist-sensor can be used as a backup node to confirm the detection decision taken from the measurements obtained on another body area.

To extend the study to the wrist-based measurements, we repeat the selection process described in Section 3 and select only those datasets that employed a sensor on that position in the datasets (see Table 3). Thus, six datasets were selected: Erciyes, UP-Fall, and UMAFall (already utilized in the previous analysis of the traces obtained from the waist), as well as CMDFall, SmartFall, and Smartwatch datasets.

The results of the ANOVA analysis of the series of the twelve statistical features of these six datasets (when an observation window is contemplated) are represented in Figures 27 and 28 (for ADLs) and Figures 29 and 30 (for the fall movements).

As expected, the graphs show even a higher disparity between the datasets than those obtained on the waist.

The way in which the volunteers are instructed to execute the ADLs and falls may particularly determine the position and movements of the hands during the activities. Thus, the measured dynamics may be extremely dependent on the testbed, which reduces the suitability of the traces for being extrapolated to other scenarios.

*4.4. Discussion.* This heterogeneity of the repositories can be motivated by very different factors, which we could group as follows:

(i) Technological factors: inertial sensor problems and limitations (biases, calibration issues, and range) can affect the measurements

(ii) Ergonomic factors: although we have compared datasets where the measurements were taken in a similar body area (the waist), measurements could be altered by the exact position of the sensor, the discomfort that the sensing device can cause in the user (which could influence the naturalness of the movements), or the firmness with which the device is adjusted to the body

(iii) Factors determined by the design of the testbed: the variability of the datasets could be clearly justified not only by the intrinsic variability (in number and types) of the performed movements but also by the particularities of the physical setting in which the movements take place: the route of the subjects during the execution of each activity, the external elements (stairs, chairs, and beds) used in the routines, or the mechanisms used to cushion the impact of the falls (mattresses, elbow pads, and helmets)

(iv) Human factors: finally, the data could be affected not only by the criteria for choosing the subjects (especially the age) but also by the particular training (or orders) that the volunteers receive to carry out the activities (in particular the falls)

## 5. Conclusions

This paper has presented a thorough study of the existing public repositories employed in the validation of Fall Detection Systems (FDS) based on wearables. The paper compares and summarizes the main basic characteristics of up to 25 available datasets used as benchmarking tools in the evaluation of FDSs.

Due to the difficulties of obtaining inertial measurements of actual falls, all these databases (except one) were created by groups of volunteers that executed a predetermined set of ADLs (Activities of Daily Living) and mimicked falls in a controlled lab-type environment. In this regard, most works in the literature evaluate their proposals by analyzing their behavior when they are applied to just one (or at most two) of these datasets. In order to indirectly assess the validity of testing a certain FDS with a single dataset, we have systematically compared the statistical characteristics of the series contained in seven of these repositories. The selection criterion of the analyzed datasets was founded on the election of a common position (waist) in which the sensor was located and on the cardinality of the measurement sets. In any case, by also analyzing the movements captured on the wrist, we also showed that conclusions could be extrapolated if other body locations with a higher degree of movement autonomy are considered.

The study, which was restricted to the accelerometry signals (as they are massively employed by the related literature on FDSs), defined and computed twelve statistical features to characterize different properties of the human mobility for each activity during the observation window (of fixed duration) in which the maximum variation of the acceleration magnitude is detected. The analysis was repeated with up to three different observation intervals without identifying a strong coherence in the characteristics obtained from the analysis of the different traces.

In particular, by means of an ANOVA analysis, we compared the means of the different statistics taking into account the nature (falls or ADLs) of the activity. This comparison was repeated after clustering the ADLs into three subcategories (basic, standard, and sporting activities) depending on the physical effort that they demand. In all

cases, a significant difference of the means was found for almost all the datasets and features. Same conclusions were drawn even when a unique and simple type of standard movement (walking) was selected to compare the databases.

The divergence of the datasets could be justified by the complex interaction of a wide set of factors: the typology and number of activities (even for those in the same subcategory), the method to execute the programmed movements, the characteristics of the experimental subjects, the range, quality, and ergonomics of the sensors, the way in which the sensing device is fastened, and the elements employed to cushion the falls. In this sense, the study reveals an evident lack of consensus on the procedure followed to define the experimental testbeds in which the datasets are generated. For example, just one of the studied datasets includes (as nonlabelled ADLs) samples captured while monitoring the actual daily routines of the volunteers.

In any case, the heterogeneity of the datasets highlighted by this investigation calls into question the results of all those studies that test the FDS against a single repository. Thanks to the sophisticated methods currently used by the literature, normally based on machine learning or deep learning techniques, some studies have achieved quality metrics (sensitivity and specificity) in the recognition of ADLs and falls very close to 100%. However, these works do not normally evaluate the capability of these methods to extrapolate these positive results when using other datasets than those considered during the training and initial validation of the FDS.

With this in mind, we should not ignore either that the credibility of the research on FDS systems is still undermined by the lack of datasets with a representative number of real falls of older people (the target population of these emergency systems), which could be utilized to benchmark the detection methods in a more realistic scenario.

## Data Availability

Datasets employed in this paper are publicly available in the Internet. URLs to access the data are provided by their authors in the corresponding references (see References).

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

## References

[1] World Health Organization (WHO), "Falls (facts sheet, 16 January 2018)," 2018.

[2] S. R. Lord, C. Sherrington, H. B. Menz, and J. C. T. Close, *Falls in Older People: Risk Factors and Strategies for Prevention*, Cambridge University Press, Cambridge, UK, 2007.

[3] Y.-C. Ku, M.-E. Liu, Y.-F. Tsai, W.-C. Liu, S.-L. Lin, and S.-J. Tsai, "Associated factors for falls, recurrent falls, and injurious falls in aged men living in taiwan veterans homes,"

[4] C. Becker and L. Schwickert, "Proposal for a multiphase fall model based on real-world fall recordings with body-fixed sensors," *Zeitschrift für Gerontologie und Geriatrie*, vol. 45, no. 8, pp. 707–715, 2012.

[5] M. Kangas, I. Vikman, L. Nyberg, R. Korpelainen, J. Lindblom, and T. Jämsä, "Comparison of real-life accidental falls in older people with experimental falls in middle-aged test subjects," *Gait Posture*, vol. 35, no. 3, pp. 500–505, 2012.

[6] J. Klenk, C. Becker, F. Lieken et al., "Comparison of acceleration signals of simulated and real-world backward falls," *Medical Engineering & Physics*, vol. 33, no. 3, pp. 368–373, 2011.

[7] F. Bagalà, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PLoS One*, vol. 7, no. 5, 2012.

[8] O. Aziz, "Validation of accuracy of SVM-based fall detection system using real-world fall and non-fall datasets," *PLoS One*, vol. 12, no. 7, 2017.

[9] FARSEEING, "Fall repository for the design of smart and self-adaptive environments prolonging independent living project," 2015.

[10] UCI Machine Learning Repository, "Human activity recognition using smartphones data set," 2020.

[11] K. A. Davis and E. B. Owusu, "Smartphone dataset for human activity recognition-dataset by uci | data.world," 2020.

[12] D. Anderson, R. Luke, J. M. Keller, and M. Skubic, "CIRL fall recognition resources," 2008.

[13] I. Charfi, J. Miteran, J. Dubois, M. Atri, and R. Tourki, "Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification," *The Journal of Electronic Imaging*, vol. 22, no. 4, 2013.

[14] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, and Y. Li, "Depth-based human fall detection via shape features and improved extreme learning machine," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, p. 1915, 2014.

[15] Z. Zhang, C. Conly, and V. Athitsos, "Evaluating depth-based computer vision methods for fall detection under occlusions," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8888, pp. 196–207, 2014.

[16] F. Riquelme, C. Espinoza, T. Rodenas, J.-G. Minonzio, and C. Taramasco, "eHomeSeniors dataset: an infrared thermal sensor dataset for automatic fall detection research," *Sensors*, vol. 19, no. 20, p. 4565, 2019.

[17] E. Auvinet, C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Multiple cameras fall dataset," *DIRO-Université Montréal (Canada)*, vol. 1350, 2010.

[18] G. Baldewijns, B. Vanrumste, G. Debard, T. Croonenborghs, and G. Mertes, "Bridging the gap between real-life data and simulated data by providing a highly realistic fall dataset for evaluating camera-based fall detection algorithms," *Healthcare Technology Letters*, vol. 3, no. 1, pp. 6–11, 2016.

[19] F. C. Czygan and V. Athitsos, ""Synthetical" aiptasia mutabilis RAPP (coelenterata) (author's transl)," *Zeitschrift Fur Naturforschung*, vol. 31, 1976.

[20] M. Aslan, Y. Akbulut, A. Şengür, and M. C. İnce, "Eklem tabanlı etkili düşme tespiti," *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, vol. 32, no. 4, pp. 1025–1034, 2017.

[21] MEBIOMEC (Universidad Politécnica de Valencia, "Fall detection testing dataset," 2020.

*International Journal of Gerontology*, vol. 7, no. 2, pp. 80–84, 2013.

[22] G. Mastorakis and D. Makris, "Fall detection system using Kinect's infrared sensor," *Journal of Real-Time Image Processing*, vol. 9, no. 4, pp. 635–646, 2014.

[23] K. Adhikari, H. Bouchachia, and H. Nait-Charif, "Activity recognition for indoor fall detection using convolutional neural network," in *Proceedings of the 15th IAPR International Conference on Machine Vision Applications*, pp. 81–84, New York, NY, USA, 2017.

[24] S. B. Khojasteh, J. R. Villar, C. Chira, V. M. González, and E. de la Cal, "Improving fall detection using an on-wrist wearable accelerometer," *Sensors*, vol. 18, no. 5, 2018.

[25] H. Leutheuser, D. Schuldhaus, B. M. Eskofier, Y. Fukui, and T. Togawa, "Hierarchical, multi-sensor based classification of daily life activities: comparison with state-of-the-art algorithms using a benchmark dataset," *PLoS One*, vol. 8, no. 10, 2013.

[26] J. R. Villar, P. Vergara, M. Menéndez, E. de la Cal, V. M. González, and J. Sedano, "Generalized models for the classification of abnormal movements in daily life and its applicability to Epilepsy convulsion recognition," *International Journal of Neural Systems*, vol. 26, no. 6, 2016.

[27] R. Igual, C. Medrano, and I. Plaza, "A comparison of public datasets for acceleration-based fall detection," *Medical Engineering & Physics*, vol. 37, no. 9, pp. 870–878, 2015.

[28] E. Casilari, R. Lora-Rivera, and F. García-Lagos, "A study on the application of convolutional neural networks to fall detection evaluated with multiple public datasets," *Sensors*, vol. 20, no. 5, p. 1466, 2020.

[29] K. Frank, M. J. Vera Nadales, P. Robertson, and T. Pfeifer, "Bayesian recognition of motion related activities with inertial sensors," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pp. 445-446, New York, NY, USA, 2010.

[30] B. Kaluža, V. Mirchevska, E. Dovgan, and M. Luštrek, "An agent-based approach to care in independent living," in *Proceedings of the 1st International Joint Conference on Ambient Intelligence 2010 (AmI-10)*, pp. 177–186, New York, NY, USA, 2010.

[31] G. Vavoulas, M. Pediaditis, E. G. Spanakis, and M. Tsiknakis, "The MobiFall dataset: an initial evaluation of fall detection algorithms using smartphones," in *Proceedigs of the IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE 2013)*, pp. 1–4, Berlin, Germany, 2013.

[32] G. Vavoulas, C. Chatzaki, T. Malliotakis, and M. Pediaditis, "The mobiact dataset: recognition of activities of daily living using smartphones," in *Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and E-Health (ICT4AWE)*, Berlin, Germany, 2016.

[33] S. Kozina, H. Gjoreski, M. Gams, and M. Luštrek, "Three-layer activity recognition combining domain knowledge and meta-classification," *Journal of Medical and Biological Engineering*, vol. 33, no. 4, pp. 406–414, 2013.

[34] S. Gasparrini, E. Cippitelli, S. Spinsante, and E. Gambi, "A depth-based fall detection system using a Kinect sensor," *Sensors*, vol. 14, no. 2, pp. 2756–2775, 2014.

[35] C. Medrano, R. Igual, I. Plaza, and M. Castro, "Detecting falls as novelties in acceleration patterns acquired with smartphones," *PLoS One*, vol. 9, no. 4, 2014.

[36] B. Kwolek and M. Kepski, "Human fall detection on embedded platform using depth maps and wireless accelerometer," *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 489–501, 2014.

[37] A. Özdemir and B. Barshan, "Detecting falls with wearable sensors using machine learning techniques," *Sensors*, vol. 14, no. 6, pp. 10691–10708, 2014.

[38] O. Ojetola, E. Gaura, and J. Brusey, "Data set for fall events and daily activities from inertial sensors," in *Proceedings of the 6th ACM Multimedia Systems Conference (MMSys'15)*, pp. 243–248, London, UK, 2015.

[39] T. Vilarinho, "A combined smartphone and Smartwatch fall detection system," in *Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM)*, pp. 1443–1448, London, UK, 2015.

[40] A. Wertner, P. Czech, and V. Pammer-Schindler, "An open labelled dataset for mobile phone sensing based fall detection," in *Proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MOBIQUITOUS 2015)*, pp. 277-278, Berlin, Germany, 2015.

[41] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection," *PLoS One*, vol. 11, 2016.

[42] J. Klenk, "The FARSEEING real-world fall repository: a large-scale collaborative database to collect and share sensor signals from real-world falls," *European Review of Aging and Physical Activity*, vol. 13, no. 1, p. 8, 2016.

[43] A. Sucerquia, J. D. López, and J. F. Vargas-bonilla, "SisFall : A fall and movement dataset," *Sensors*, vol. 198, no. 52, pp. 1–14, 2017.

[44] D. Micucci, M. Mobilio, and P. Napoletano, "UniMiB SHAR: a new dataset for human activity recognition using acceleration data from smartphones," *Applied Science*, vol. 7, no. 10, 2017.

[45] M. Ahmed, N. Mehmood, A. Nadeem, A. Mehmood, and K. Rizwan, "Fall detection system for the elderly based on the classification of shimmer sensor prototype data," *Healthcare Informatics Research*, vol. 23, no. 3, pp. 147–158, 2017.

[46] O. Aziz, M. Musngi, E. J. Park, G. Mori, and S. N. Robinovitch, "A comparison of accuracy of fall detection algorithms (threshold-based vs. machine learning) using waist-mounted tri-axial accelerometer signals from a comprehensive set of falls and non-fall trials," *Medical & Biological Engineering & Computing*, vol. 55, no. 1, pp. 45–55, 2017.

[47] F. T. Wang, H. L. Chan, M. H. Hsu, C. K. Lin, P. K. Chao, and Y. J. Chang, "Threshold-based fall detection using a hybrid of tri-axial accelerometer and gyroscope," *Physiological Measurement*, vol. 39, no. 10, 2018.

[48] T. H. Tran, "A multi-modal multi-view dataset for human fall analysis and preliminary investigation on modality," in *Proceedings of the International Conference on Pattern Recognition*, Berlin, Germany, 2018.

[49] S. S. Saha, S. Rahman, M. J. Rasna, A. K. M. Mahfuzul Islam, and M. A. Rahman Ahad, "DUMD: An open-source human action dataset for ubiquitous wearable sensors," in *Proceedings of the 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, pp. 567–572, Berlin, Germany, 2018.

[50] T. Mauldin, M. Canby, V. Metsis, A. Ngu, and C. Rivera, "SmartFall: a smartwatch-based fall detection system using deep learning," *Sensors*, vol. 18, no. 10, p. 3363, 2018.

[51] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, "UP-fall

detection dataset: a multimodal approach," *Sensors*, vol. 19, no. 9, 2019.

[52] V. Cotechini, A. Belli, L. Palma, M. Morettini, L. Burattini, and P. Pierleoni, "A dataset for the development and optimization of fall detection algorithms based on wearable sensors," *Data BR*, vol. 19, no. 9, 2019.

[53] G. Zhao, Z. Mei, D. Liang et al., "Exploration and implementation of a pre-impact fall recognition method based on an inertial body sensor network," *Sensors*, vol. 12, no. 11, pp. 15338–15355, 2012.

[54] H. Gjoreski, M. Luštrek, and M. Gams, "Accelerometer placement for posture recognition and fall detection," in *Proceedings of the 2011 7th International Conference on Intelligent Environments*, pp. 47–54, Berlin, Germany, 2018.

[55] J. Dai, X. Bai, Z. Yang, Z. Shen, and D. Xuan, "PerFallD: A pervasive fall detection system using mobile phones," in *Proceedings of the 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 292–297, New York, NY, USA, 2010.

[56] M. Kangas, A. Konttila, P. Lindgren, I. Winblad, and T. Jämsä, "Comparison of low-complexity fall detection algorithms for body attached accelerometers," *Gait & Posture*, vol. 28, no. 2, pp. 285–291, 2008.

[57] S.-H. Fang, Y.-C. Liang, and K.-M. Chiu, "Developing a mobile phone-based fall detection system on android platform," in *Proceedings of the Computing, Communications and Applications Conference (ComComAp)*, pp. 143–146, New York, NY, USA, 2010.

[58] E. Casilari, M. Álvarez-Marco, and F. García-Lagos, "A Study of the use of gyroscope measurements in wearable fall detection systems," *Symmetry*, vol. 12, no. 4, p. 649, 2020.

[59] S.-H. Liu and W.-C. Cheng, "Fall detection with the support vector machine during scripted and continuous unscripted activities," *Sensors*, vol. 12, no. 9, pp. 12301–12316, 2012.

[60] J. Santoyo-Ramón, E. Casilari, and J. Cano-García, "Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection with supervised learning," *Sensors*, vol. 18, no. 4, p. 1155, 2018.

[61] S. Abbate, M. Avvenuti, F. Bonatesta, G. Cola, P. Corsini, and A. Vecchio, "A smartphone-based fall detection system," *Pervasive and Mobile Computing*, vol. 8, no. 6, pp. 883–899, 2012.

[62] A. K. Bourke, K. J. O'Donovan, J. Nelson, and G. M. ÓLaighin, "Fall-detection through vertical velocity thresholding using a tri-axial accelerometer characterized using an optical motion-capture system," in *Proceedings of the 10th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS'08)*, pp. 2832–2835, New York, NY, USA, 2010.

[63] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 1, pp. 156–167, 2006.

[64] O. Ojetola, E. I. Gaura, and J. Brusey, "Fall detection with wearable sensors--safe (smart fall detection)," in *Proceedings of the 7th International Conference on Intelligent Environments*, pp. 318–321, New York, NY, USA, 2010.

[65] I. P. E. S. Putra and R. Vesilo, "Genetic-algorithm-based feature-selection technique for fall detection using multi-placement wearable sensors," in *Proceedings of the 12th International Conference on Body Area Networks (BodyNets 2017)*, pp. 319–332, New York, NY, USA, 2010.

[66] A. T. Özdemir, "An analysis on sensor locations of the human body for wearable fall detection devices: principles and practice," *Sensors (Switzerland)*, vol. 16, no. 8, 2016.

[67] P. Vallabh, R. Malekian, N. Ye, and D. C. Bogatinoska, "Fall detection using machine learning algorithms," in *Proceedings of the 2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–9, New York, NY, USA, 2010.

[68] C. Wang, S. J. Redmond, W. Lu, M. C. Stevens, S. R. Lord, and N. H. Lovell, "Selecting power-efficient signal features for a low-power fall detector," *IEEE Transactions on Bio-Medical Engineering*, vol. 64, no. 11, pp. 2729–2736, 2017.

[69] A. O. Kansiz, M. A. Guvensan, and H. I. Turkmen, "Selection of time-domain features for fall detection based on supervised learning," in *Proceedings of the World Congress on Engineering and Computer Science*, pp. 23–25, New York, NY, USA, 2010.

[70] A. Sucerquia, J. D. López, and J. F. Vargas-Bonilla, "Real-life/real-time elderly fall detection with a triaxial accelerometer," *Sensors*, vol. 18, no. 4, 2018.

[71] S. Chernbumroong, S. Cang, and H. Yu, "Genetic algorithm-based classifiers fusion for multisensor activity recognition of elderly people," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 282–289, 2015.

[72] S. Bersch, D. Azzi, R. Khusainov, I. Achumba, and J. Ries, "Sensor data acquisition and processing parameters for human activity classification," *Sensors*, vol. 14, no. 3, pp. 4239–4270, 2014.

[73] P. Vallabh and R. Malekian, "Fall detection monitoring systems: a comprehensive review," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 6, pp. 1809–1833, 2018.

[74] X. Xi, M. Tang, S. M. Miran, and Z. Luo, "Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors," *Sensors (Switzerland)*, vol. 17, no. 6, 2017.

[75] K.-H. Chen, J.-J. Yang, and F.-S. Jaw, "Accelerometer-based fall detection using feature extraction and support vector machine algorithms," *Instrumentation Science & Technology*, vol. 44, no. 4, pp. 333–342, 2016.

[76] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Analysis of public datasets for wearable fall detection systems," *Sensors*, vol. 17, no. 7, 2017.