

# Sensor-Based Systems for Independent Living of Ageing People

Lead Guest Editor: Ivan Miguel Pires

Guest Editors: Nuno M. Garcia, Susanna Spinsante, and Gonçalo Marques





---

# **Sensor-Based Systems for Independent Living of Ageing People**

Journal of Healthcare Engineering

---

## **Sensor-Based Systems for Independent Living of Ageing People**

Lead Guest Editor: Ivan Miguel Pires

Guest Editors: Nuno M. Garcia, Susanna Spinsante,  
and Gonçalo Marques



---

Copyright © 2021 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Healthcare Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Associate Editors

Xiao-Jun Chen , China  
Feng-Huei Lin , Taiwan  
Maria Lindén, Sweden

## Academic Editors

Cherif Adnen, Tunisia  
Saverio Affatato , Italy  
Óscar Belmonte Fernández, Spain  
Sweta Bhattacharya , India  
Prabadevi Boopathy , India  
Weiwei Cai, USA  
Gin-Shin Chen , Taiwan  
Hongwei Chen, USA  
Daniel H.K. Chow, Hong Kong  
Gianluca Ciardelli , Italy  
Olawande Daramola, South Africa  
Elena De Momi, Italy  
Costantino Del Gaudio , Italy  
Ayush Dogra , India  
Luobing Dong, China  
Daniel Espino , United Kingdom  
Sadiq Fareed , China  
Mostafa Fatemi, USA  
Jesus Favela , Mexico  
Jesus Fontecha , Spain  
Agostino Forestiero , Italy  
Jean-Luc Gennisson, France  
Badicu Georgian , Romania  
Mehdi Gheisari , China  
Luca Giancardo , USA  
Antonio Gloria , Italy  
Kheng Lim Goh , Singapore  
Carlos Gómez , Spain  
Philippe Gorce, France  
Vincenzo Guarino , Italy  
Muhammet Gul, Turkey  
Valentina Hartwig , Italy  
David Hewson , United Kingdom  
Yan Chai Hum, Malaysia  
Ernesto Iadanza , Italy  
Cosimo Ieracitano, Italy




Giovanni Improta , Italy  
Norio Iriguchi , Japan  
Mihajlo Jakovljevic , Japan  
Rutvij Jhaveri, India  
Yizhang Jiang , China  
Zhongwei Jiang , Japan  
Rajesh Kaluri , India  
Venkatachalam Kandasamy , Czech Republic  
Pushpendu Kar , India  
Rashed Karim , United Kingdom  
Pasi A. Karjalainen , Finland  
John S. Katsanis, Greece  
Smith Khare , United Kingdom  
Terry K.K. Koo , USA  
Srinivas Koppu, India  
Jui-Yang Lai , Taiwan  
Kuruva Lakshmanna , India  
Xiang Li, USA  
Lun-De Liao, Singapore  
Qiu-Hua Lin , China  
Aiping Liu , China  
Zufu Lu , Australia  
Basem M. ElHalawany , Egypt  
Praveen Kumar Reddy Maddikunta , India  
Ilias Maglogiannis, Greece  
Saverio Maietta , Italy  
M.Sabarimalai Manikandan, India  
Mehran Moazen , United Kingdom  
Senthilkumar Mohan, India  
Sanjay Mohapatra, India  
Rafael Morales , Spain  
Mehrbakhsh Nilashi , Malaysia  
Sharnil Pandya, India  
Jialin Peng , China  
Vincenzo Positano , Italy  
Saeed Mian Qaisar , Saudi Arabia  
Alessandro Ramalli , Italy  
Alessandro Reali , Italy  
Vito Ricotta, Italy  
Jose Joaquin Rieta , Spain  
Emanuele Rizzuto , Italy

Dinesh Rokaya, Thailand  
Sébastien Roth, France  
Simo Saarakkala , Finland  
Mangal Sain , Republic of Korea  
Nadeem Sarwar, Pakistan  
Emiliano Schena , Italy  
Prof. Asadullah Shaikh, Saudi Arabia  
Jiann-Shing Shieh , Taiwan  
Tiago H. Silva , Portugal  
Sharan Srinivas , USA  
Kathiravan Srinivasan , India  
Neelakandan Subramani, India  
Le Sun, China  
Fabrizio Taffoni , Italy  
Jinshan Tang, USA  
Ioannis G. Tollis, Greece  
Ikram Ud Din, Pakistan  
Sathishkumar V E , Republic of Korea  
Cesare F. Valenti , Italy  
Qiang Wang, China  
Uche Wejinya, USA  
Yuxiang Wu , China  
Ying Yang , United Kingdom  
Elisabetta Zanetti , Italy  
Haihong Zhang, Singapore  
Ping Zhou , USA

# Contents

---




**Social Capital and Digital Divide: Implications for Mobile Health Policy in Developing Countries**

Teng Wang , Xitong Guo , and Tianshi Wu 

Research Article (13 pages), Article ID 6651786, Volume 2021 (2021)

**On the Heterogeneity of Existing Repositories of Movements Intended for the Evaluation of Fall**

**Detection Systems**

Eduardo Casilari , José A. Santoyo-Ramón , and José M. Cano-García 

Research Article (36 pages), Article ID 6622285, Volume 2020 (2020)

**Feature Selections Using Minimal Redundancy Maximal Relevance Algorithm for Human Activity Recognition in Smart Home Environments**

Hongqing Fang , Pei Tang, and Hao Si

Research Article (13 pages), Article ID 8876782, Volume 2020 (2020)

## Research Article

# Social Capital and Digital Divide: Implications for Mobile Health Policy in Developing Countries

Teng Wang , Xitong Guo , and Tianshi Wu 

*eHealth Research Institute, School of Management, Harbin Institute of Technology, Harbin, China*

Correspondence should be addressed to Xitong Guo; [xitongguo@gmail.com](mailto:xitongguo@gmail.com)

Received 11 November 2020; Revised 5 January 2021; Accepted 15 January 2021; Published 27 January 2021

Academic Editor: Ivan Miguel Pires

Copyright © 2021 Teng Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Digital divide has been a major obstacle for mobile health services for the elderly in developing countries; to assess the potential solution to narrow digital divide among the elderly, we use data from the China Health and Retirement Longitudinal Study (CHARLS) and test for a causal role of social capital in digital access among elderly individuals in China. To handle endogenous problems associated with social capital, we introduce instrumental variable (IV) estimates in our models. Our data analysis shows that social capital facilitates increased digital access. We distinguish between two digital access patterns, an infrastructure pattern and a personal device pattern, and find that the causal effect of social capital is determined by the personal device pattern. Therefore, since family members and relatives increase digital access among elderly people, we propose a family-centered mobile health policy in developing countries.

## 1. Introduction

Noncommunicable diseases (NCDs) are the leading cause of death globally, but as the treatment and prevention of NCDs are worse in developing countries because resources are insufficient, elderly individuals there face a higher risk [1]. The characteristics of chronic diseases call for long-term and persistent self-management intervention, so patients with chronic disease need to alter their lifestyles. However, due to the fragmented health systems and insufficient funds in developing countries, the prevention and treatment for NCDs are facing more challenges [1, 2].

Mobile health (m-health) offers a potential solution to increase the efficiency of NCD treatment and prevention. Technologies such as apps and wearable devices empower patients self-tracking and self-care, specifically following and receiving instant feedback on their health, movement, and diet, improving health outcomes, and they are becoming increasingly available [3, 4]. Recent empirical research into m-health reveals that feedback, incentive, and social support mechanisms in m-health devices have helped improve health outcomes and self-efficacy among hypertension and diabetes patients [5–7]. Although mobile health is made possible by

widespread mobile technology, younger, more educated, wealthier, and healthier people have an advantage in digital access to m-health; specifically, the affordability, independent usage, and ease of use have been considered as obstacles for elderly users [8, 9]. This digital divide has therefore become a main concern in m-health policy [10, 11], and it is prevalent in the developing world. For instance, in less developed countries, about 89% of urban households but only 63% of rural ones have a mobile phone [12]. One out of five people are online in less developed countries, compared with four out of five in developed countries [13].

We therefore estimate the role of social capital in the digital divide among elderly people in China. We focus on this topic for two reasons: First, although there is a proven positive relationship between social capital and health [14–18], the association between social capital and digital divide is unexamined. More importantly, increasing age means an increasing probability of chronic disease. Due to the incomplete social security system in developing countries, elderly people with chronic diseases generally face greater inequality, but among those countries with a tradition of collectivism, social capital is expected to increase digital access by providing resources for the elderly [19, 20].



Second, quantitative research on m-health policy in developing countries is lacking. ICT researchers have conducted small-sample m-health experiments to examine the feasibility of using ICT for chronic diseases. However, as m-health is a new healthcare service pattern, insufficient attention has been paid to its policy feasibility. Specifically, there is no known path of diffusion for m-health technology among underprivileged populations in developing regions [21]. The population of China is aging rapidly and seeing more NCDs: NCDs caused nearly 80% of deaths among people aged 60 or older in 2012, and in 2013 over 100 million people had at least one chronic NCD [22, 23]. Analysis of data from China will have implications for other developing countries undergoing aging and the transition of disease.

Additionally, previous studies suggested that the effect of social capital on dependent variables poses endogeneity problems, such as the mutually reinforcing relationship between social capital and local public goods [15, 24] and the relationship between living preference and wealth [25]. We measure social capital by reciprocal behavior, which is influenced by personal characteristics like altruism tendency and sympathy, latent variables that cannot be observed or controlled. To avoid the bias of coefficients on social capital and investigate the causal relationship, we use an instrumental variable (IV) approach to test and handle endogeneity problems, introducing two IVs, *Migration* and *Siblings*.

In our main results, we first find that social capital, as measured by mutual reciprocity within strong ties (relatives and friends), does have a causal effect on digital access among elderly people in China. We find two digital access patterns: the infrastructure pattern and the personal device pattern. In the infrastructure pattern, the causal effect of social capital cannot be determined since infrastructure is more sensitive to socioeconomic status (SES) indicators, while in the personal device pattern, the causal role of social capital is valid and significant in increasing digital access. Second, SES is significantly negatively associated with digital access in both patterns, a result consistent with previous studies. Age and chronic diseases are negatively associated with digital access and play a similar role to that of SES in the digital divide. This result implies that age and chronic diseases should be considered structural variables in the digital divide, and m-health in developing countries may face more difficulties than expected. Lastly, based on our findings, we propose a family-centered m-health service system in developing countries. As our results suggest that social capital has a causal effect on personal equipment, family and community can play a critical role in m-health policy to narrow the digital divide for the elderly.

## 2. Social Capital and Digital Divide

Digital divide refers to the gulf in information and communication technology access (e.g., haves or have-nots), capability (e.g., computer skills or ability to find information online), and outcomes (e.g., productivity of IT investment and use) across a variety of demographic, ethnic, geographic, and socioeconomic dimensions [26, 27]. As a measure of social structure, SES is frequently associated with the digital

divide. For some researchers, the digital divide is more than digital haves or have-nots: it is part of the world of structural inequalities [28]. For example, Internet use among the middle-aged and elderly is strongly associated with SES in China, and community resources are also associated with the digital divide [10]. In the US, income, education, age, and family structure are important social determinants of online access, and older respondents were found to have lower Internet access than average [29]. However, since studies on the association between social capital and digital divide are rare and since social capital is a broad notion, we first define social capital from the perspective of private goods. Then we review the literature of social relations and digital access.

*2.1. Social Capital: Private Goods or Public Goods?* The notion of social capital generally includes both social networks and resources embedded in those networks [30]. However, scholars do not agree on how to explain the mechanism and define the function of social capital [31]. Some researchers find that social capital motivates the community self-governance and the pursuit of collective goals by improving cohesiveness at the organization or community level [15, 24, 31, 32]. Studies following this tradition evaluate the effect of social capital on health or health resources. For instance, [17] suggests that social capital affects individual health by influencing access to services and amenities. Similarly, social capital (as measured by kindness and greeting) and social cohesion in a community increase general health [14].

Other studies in developing countries emphasize that social capital as private goods can play a more direct and significant role. In this view, resources from external ties such as relatives and acquaintances play a more critical role than collective action, and individuals rather than communities benefit; in this way, social capital is similar to individual investment [31, 33]. For instance, [20] finds that while social capital measured by organizational membership is unassociated with health outcomes in rural China, support from friends and relatives may contribute more to public goods provision. Using the number of friends as the proxy of social capital, [34] concludes that the more friends, the better one's physical and mental health outcomes. Here, we follow the private goods view in which social capital refers to actual and potential resources embedded in an individual's or social units' social networks [35, 36]. The private goods view suggests that social capital includes social networks and embedded resources. We focus on social networks rather than resources for one important reason: social networks are embedded into culture context, and collectivism is a deeply rooted cultural characteristic in developing countries, shaping individual behavior [19, 37].

*2.2. Social Capital and Digital Access.* Since studies of the association between social capital and digital divide are rare, we review work emphasizing the effect of social networks on digital access. From the private goods perspective, social networks provide resources (e.g., information and influence) that can facilitate digital access. For example, [38] finds a

positive relationship between Internet use and peer effect: individuals in the proximity of others who go online will be influenced to go online themselves. Theoretically, interpersonal interactions could affect technology adoption through social learning, pressure, influence, or support (such as information). A field study in India [39] finds that farmers obtained information from and were influenced by other villagers of advice networks, and advice networks were found to increase the usage of farming information system. In another field experiment lasting for seven years in India, mothers' social networks (both strong and weak ties) were found to affect their use of ICT intervention and further reduce infant mortality [40]. In the literature of health information searching, [41] finds that teen Internet users can be health information seekers for families with low education, suggesting that young members of family may act as a potential source of digital access. Researchers aiming to increase digital access among people with low SES have found that social networks do influence underprivileged populations in developing nations, and they have also been found to increase digital access, even in less developed areas. Therefore, there is a reason to believe that social capital, via social networks, could help narrow the digital divide.

### 3. Data and Measurement

We use two databases to test the causal effect of social capital on digital access: the China Health and Retirement Longitudinal Study (CHARLS 2011) (this data is available at <http://charls.pku.edu.cn/pages/data/2011-charls-wave1/zh-cn.html>) and the Statistical Yearbook of China's Regional Economy (2012) (this data are available at [http://www.stats.gov.cn/tjsj/tjcbw/201303/t20130318\\_451532.html](http://www.stats.gov.cn/tjsj/tjcbw/201303/t20130318_451532.html)). CHARLS is a survey of the middle-aged and elderly in China, based on a sample of households with members aged 45 or older, covering 150 countries/districts and 450 villages/urban communities, and interviewing 17,708 individuals in 10,257 households; our analysis contains 16,316 samples, and 1,392 samples were deleted due to missing variables. This data shows that, of individuals over 60, only 9.74% have Internet access and 63.92% have mobile phone access. Of individuals 45 to 59, only 20.62% have Internet access and 87.86% have mobile phone access. This low adoption rate can be considered a proxy of m-health adoption among the elderly in China today. Digital access could also act as the foundation to predict m-health access in the future because m-health services share similar ICT infrastructure for service providers and similar usage habits for users.

**3.1. Social Capital.** The private good perspective emphasizes the use of accessible resources from external networks to measure social capital in developing countries. We construct a composite score based on whether the respondent has received any help (monetary or nonmonetary) from or provided any help to others (co-resident parents/parents-in-law/children/grandchildren/relatives/nonrelatives). This method is suggested and applied in previous work using CHARLS to measure social

capital [42]. Figure 1 reports the distribution of social capital by SES.

**3.2. Digital Access.** Digital access can be measured by home computer ownership at the individual level, IT investment at the organizational level, and national IT expenditure at the country level [27]. Since smartphones are the current platform for m-health services, we use Internet and mobile phone access to measure digital access at the individual level. For Internet access, we ask, "Does your household have broadband Internet connection?" The majority (84%) of respondents do not. Mobile phone access is measured using the following question: "Do members of your household own the following assets?" We find that 21% of respondents have no mobile phone access at the household level. Distribution of Internet and mobile phone access by SES appears in Figures 2 and 3.

### 3.3. Control Variables

**3.3.1. Socioeconomic Status (SES).** Prior studies measured multiple dimensions of SES: education, occupation, access to goods and services, and household welfare. Because a single proxy could lead to unstable results [43], we use education, wealth, and residence to measure respondents' SES (see Table 1). Education is categorized into four groups (illiterate, primary, medium, and high levels). Since farmers in developing countries have no regular salary and make up a large proportion of respondents (78% have rural hukou and only 22% have urban hukou), it is difficult to accurately measure respondents' income. Expenditure is a better way to gauge income when a person's income is irregular [10, 44]. We therefore used per capita expenditure, which has also been used in other studies with CHARLS (2011). Given the urban/rural structural disparity in developing countries, we use type of hukou (urban or rural) to categorize residence.

**3.3.2. Chronic Diseases.** The CHARLS (2011) contains 14 questions to assess chronic diseases. Comorbidity has been widely observed among chronic disease patients [45]; for instance, one-half of Chinese people aged 70 or older and one-half to two-thirds of Spanish adults older than 65 have two or more chronic conditions [1]. As suggested in previous studies [11], we note three or more cooccurring chronic diseases as an indicator of the severity of chronic disease.

**3.3.3. Demographic Variables.** These include marriage status, sex, and age, which is subdivided into three groups (45–59, 60–74, and 75+).

**3.3.4. Living Arrangements.** Household size and coresidence with children and grandchildren can influence digital access, since younger people are more inclined to purchase digital devices. Household size is measured by the number of people living in the same household, with a member being anyone who lived in the household for over six months in the

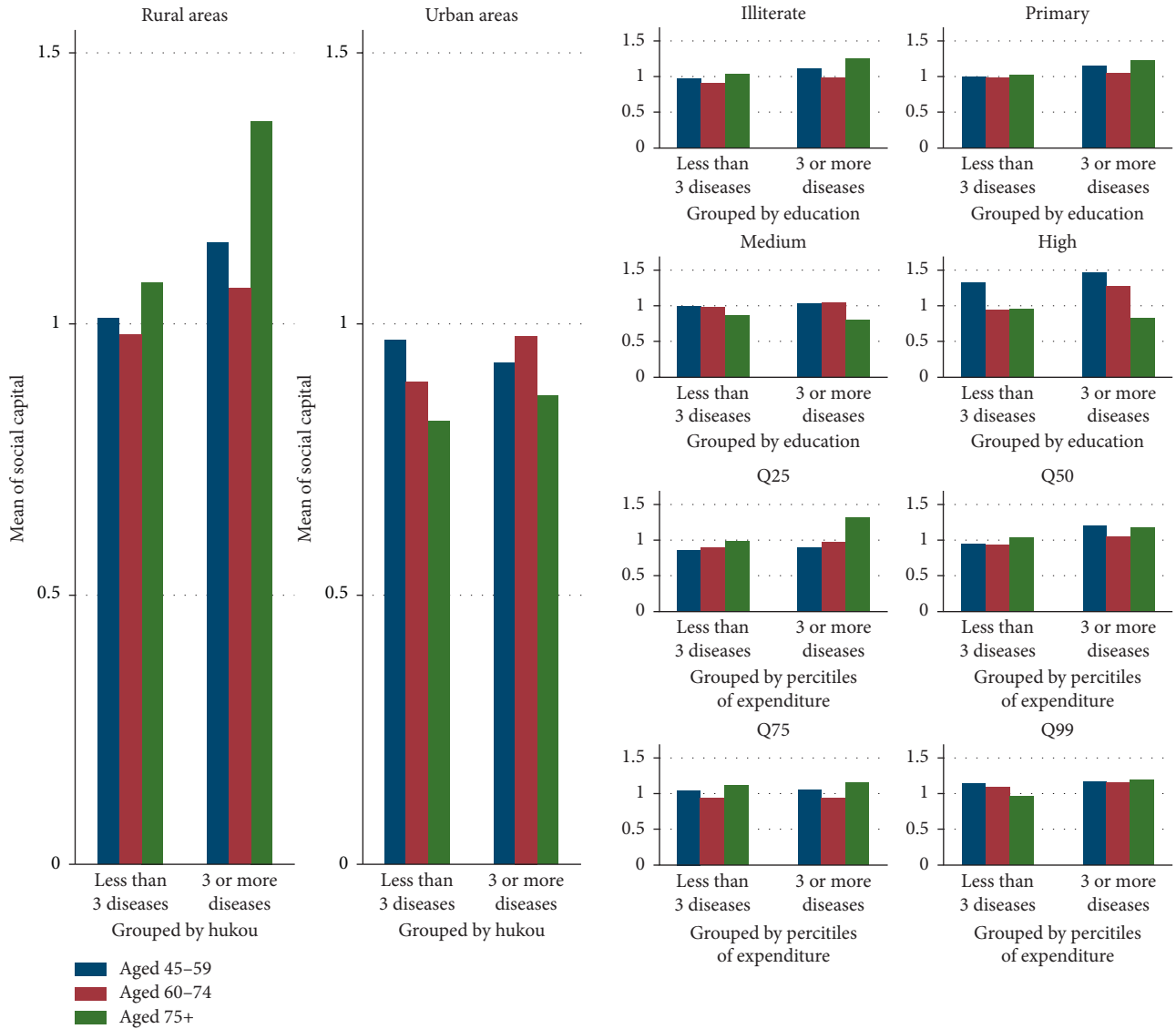


FIGURE 1: Distribution of social capital by residence, education, and wealth. Q25, Q50, Q75, and Q99 represent the respondents' per capita expenditure (Log) in the top 25%, top 26% to top 50%, top 50 to top 75%, and top 51% to top 99%, respectively.

past year. We also track whether respondents live with their adult children or with grandchildren over 16.

**3.3.5. Public Goods.** To control for variables that may affect digital access at community and city level, we introduce public goods as control variables. Public goods related to digital access involve power supply, mobile phone base stations, etc., and we divide them into two levels, community and city. To assess public goods at the community level we follow [42, 46] and make a composite measure by asking respondents, "Does your village/community have the following type of facilities?" Fourteen facility types are listed, including basketball court, swimming pool, and outside exercise facilities. At the city level, we measure public goods by public service budget of local government (100 million yuan/10,000 people), including education, social safety net, and employment efforts. We obtain general budget and

population data from the Statistical Yearbook of China's Regional Economy (2012).

## 4. Empirical Strategy

First, we use probit regression to estimate the effect of social capital. Access to Internet and mobile phone is used to estimate digital access using the following equation:

$$\text{DigitalAccess} = \beta_0 + \beta_1 \text{SES} + \beta_2 \text{ChronicDisease} + \beta_3 \text{SocialCapital} + \beta_4 \text{Control} + \varepsilon \quad (1)$$

The variable SES is socioeconomic status including *wealth* (measured by per capita expenditure), education, and residence (measured by hukou type). *ChronicDisease* equals 1 if the respondent has three or more chronic diseases. We control for demographic characteristics, living arrangements, and public goods, denoted by *Control*. Additionally,

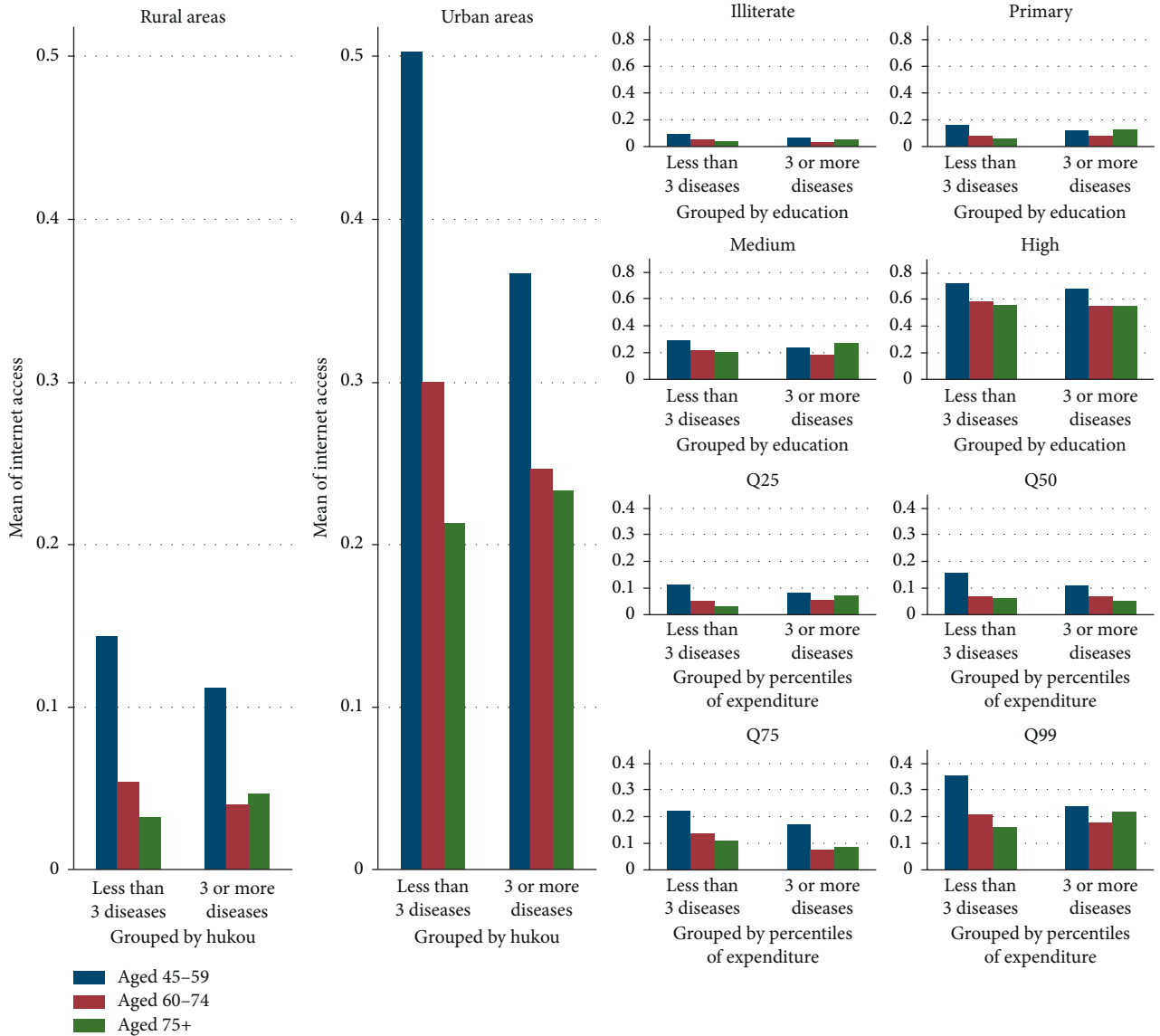


FIGURE 2: Distribution of Internet access by residence, education, and wealth. Q25, Q50, Q75, and Q99 represent the respondents' per capita expenditure (Log) in the top 25%, top 26% to top 50%, top 50 to top 75%, and top 51% to top 99%, respectively.

to assess the endogeneity of social capital, we use the Wald test of exogeneity in IV-probit regression [47].

Second, given the potential endogeneity of social capital, Maximum Likelihood Estimate (MLE) will not be consistent [48], so we introduce two-step sequential estimation in

models with instruments. *Migration* denotes whether the respondent left his/her birthplace and *Siblings* represents number of siblings. Our two-step IV-probit equation is specified as follows:

$$\text{SocialCapital} = \alpha_0 + \alpha_1 \text{Migration} + \alpha_2 \text{Siblings} + \alpha_3 \text{SES} + \alpha_4 \text{Control} + \varepsilon, \tag{2}$$

$$\text{DigitalAccess} = \beta_0 + \beta_1 \text{SES} + \beta_2 \text{ChronicDisease} + \beta_3 \text{Social Capital} + \beta_4 \text{Control} + \varepsilon. \tag{3}$$

**4.1. Instrumental Variables.** Since we require a variable that associates with social capital but without a direct effect on digital access, we use *Migration* (stayers or movers) and *Siblings* as instrumental variables (IVs).

**4.1.1. Migration.** As suggested in recent research [15, 49], *Migration* can be a proxy to measure the causal effect of social capital since it has no direct effect on infrastructure like healthcare resources. If a respondent moves away from his

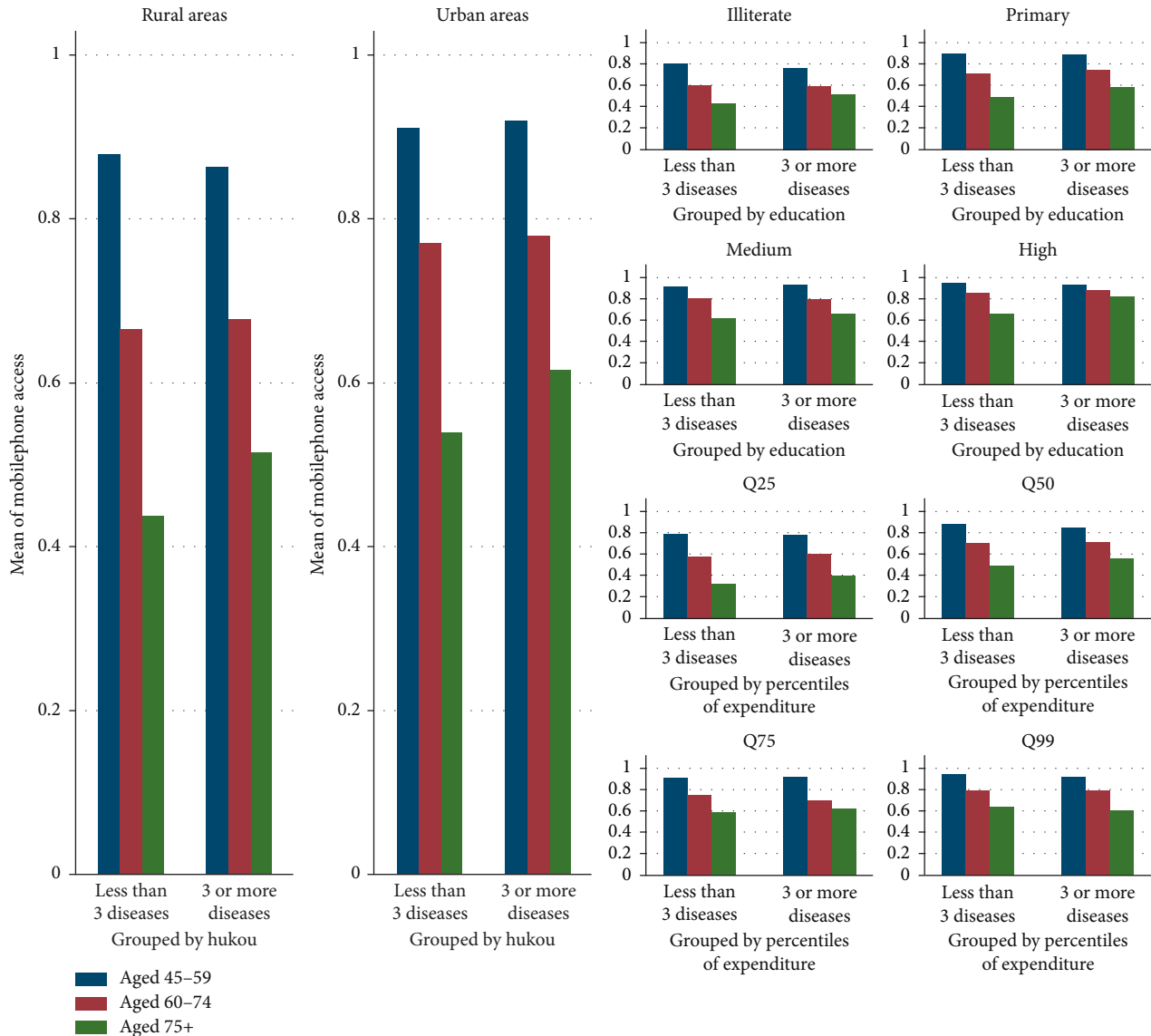


FIGURE 3: Distribution of mobile phone access by residence, education, and wealth. Q25, Q50, Q75, and Q99 represent the respondents' per capita expenditure (Log) in the top 25%, top 26% to top 50%, top 50 to top 75%, and top 51% to top 99%, respectively.

birthplace, the respondent is defined as a mover, otherwise as a stayer. Low population mobility, as is the case overall in China, increases the importance of interpersonal relationships and personal networks with relatives [50]. These social networks in urban communities often involve neighbors who are also work colleagues, as employees of Chinese state-owned businesses could get free housing [51]. In rural communities, people often live with extended family and their neighbors may also be relatives. Since individuals in China access resources from the social network of their place of birth, as is common in countries that are not yet totally industrialized or urbanized, stayers who remain in their birthplace can obtain more resources from existing social networks, while movers must build new social networks.

To identify movers and stayers, we ask whether the respondent's first hukou (usually obtained as an infant or child) is identical to his/her current hukou. If the respondent chooses "same as birthplace," *Migration* equals 0 and the

respondent will be marked as a migrant. According to Chinese policies, migrants need to update their hukou registration information or they may encounter some obstacles in work, education, etc. This is due to differences in social welfare policies between cities, and some welfare policies are only available for people holding local hukou; therefore, hukou can be a reliable variable to measure migration.

**4.1.2. Number of Siblings.** Since social networks and resources are the two components of social capital, a bigger social network should have more resources for an individual. In Chinese culture, the sibling relationship is a strong tie; it involves more obligation and trust than do weaker ties between siblings in western culture [35]. We therefore use number of siblings as a proxy for strong-tie social network size; more siblings indicate a larger social network providing more social capital. For instance, information and influence

TABLE 1: Description of variables.

	Variables	Description	Mean	Std. Dev.	Min	Max
	Dependent variables					
Digital access	Internet access	Dummy variable equals 1 if respondent's household has broadband Internet connection	0.162	0.368	0	1
	Mobile phone access	Dummy variable equals 1 if members of respondent's household own mobile phones	0.782	0.413	0	1
	Independent variables					
	PCE	Household members' total per capita expenditure in the past year is used as the measurement of income	2710.375	6906.97	0	233000
SES	Education	Education is categorized into four groups: illiterate, primary (home school or elementary school), middle (middle school, high school, or vocational school), and high (college and above)				
	IlliterateEDU	Percentage of respondents without formal education	27.29	0.819	0	1
	PrimaryEDU	Percentage of respondents who completed elementary school	39.304	0.819	0	1
	MediumEDU	Percentage of respondents who completed middle school	31.025	0.819	0	1
	HighEDU	Percentage of respondents with higher education	2.378	0.819	0	1
	Residence	Type of hukou: urban hukou is 1 and rural hukou is 0	0.22	0.414	0	1
Chronic disease	Chronic disease	Diagnosed with three or more chronic diseases (yes = 1, no = 0)	0.186	0.389	0	1
Social capital	Social capital	Social capital is measured by a composite score. 12 dichotomous variables are measured by asking whether the respondent or spouse in the past year received/provided any economic supports from/to noncoresident parents/parents-in-law/children/grandchildren/relatives/nonrelatives. These dichotomous values are added to form a composite score	1.008	1.066	0	8
	Sex	Percentage of female respondents (female = 1, male = 0)	0.512	0.500	0	1
	Age	Aged 45 to 59 is middle-age level, 60 to 74 is presenior level, and 75+ is senior level				
Demographic variables	Aged 45–59	Percentage of respondents aged 45 to 59	55.93	0.651		
	Aged 60–74	Percentage of respondents aged 60 to 74	35.32	0.651		
	Aged above 75	Percentage of respondents aged over 75	8.75	0.651		
	Marital	Dummy variable equals 1 if married or separated and 0 if single, divorced, or widowed	0.871	0.335	0	1
Public goods	Community infrastructure	Community infrastructure is measured by a composite score. 14 dichotomous variables are measured by asking whether the community has certain infrastructure (e.g., basketball facilities), organizations (e.g., dance team), and public services (e.g., employment service center)	3.523	3.568	0	14
	Public investment	General budgetary expenditure (education, social safety net and employment effort, medical and healthcare services, agriculture forestry, and water conservation) of respondent's local government (one hundred million RMB per 10,000 people)	3.799	13.13	0.1	94
Living arrangement	HouseholdSize	Number of family members living in the same household	3.725	1.771	2	16
	Coresidence grandchildren	Percentage of respondents living with grandchildren over 16 years old	0.053	0.223	0	1
	Coresidence children	Percentage of respondents living with adult children	0.556	0.497	0	1
Instrumental variables	Migration	Equals 1 if respondent's first hukou (usually obtained as infant or child) is not his/her current hukou	0.49	0.500	0	1
	Siblings	Number of respondents' living siblings	3.140	1.929	0	11
	Observations		16316			

obtained through social contacts help job-seekers secure higher-paying work [52, 53]. The number of siblings is measured by asking, "How many of your siblings are still

alive?" As shown in Table 1, about 48.7% of respondents moved from their birthplace (movers), and respondents have an average of 3.14 *Siblings*.

**4.2. Strength of IVs.** Weak instruments can produce biased IV estimators and fail to solve the endogeneity of social capital. We address this issue by checking the results of the first-stage linear regression. A common rule of thumb considers an instrument strong if the F-statistic in the first stage is over 10 and the coefficients of IVs match theoretical assumption [54]. As shown in Table 2, the F-statistic for *Migration* and *Siblings* is 34.57 ( $p < 0.01$ ) for social capital in both models. Additionally, the coefficients for *Migration* and *Siblings* are -0.051 ( $t = -3.01$ ,  $p < 0.01$ ) and 0.037 ( $t = 8.18$ ,  $p < 0.01$ ), respectively, indicating that *Migration* decreases resources the respondent could obtain through his/her social networks and that *Siblings* is positively associated with social capital. This result is consistent with the assumption that more strong ties result in more social capital. Taken together, we argue that *Migration* and *Siblings* are not weak IVs.

## 5. Results

**5.1. Social Capital and Digital Divide.** As shown in Table 3, the coefficient of social capital in model 1 is not significant. When the IVs are introduced in model 2, the association between social capital and Internet access is significant at 99% confidence level. The difference between these models reveals the possible bias caused by latent variables. In models with IVs, social capital is positively associated with digital access. To be specific, resources obtained from social networks increase digital access. The coefficients in models 2 and 4 are 0.79 ( $t = 3.72$ ,  $p < 0.01$ ) and 1.29 ( $t = 5.88$ ,  $p < 0.01$ ), respectively.

We turn to the Wald test of exogeneity provided in Table 3 to test if social capital is endogenous. In models 3 and 6, the chi-square equals 18.41 ( $p < 0.01$ ) and 53.33 ( $p < 0.01$ ), respectively, and the estimated coefficients of  $\alpha$  are -0.79 ( $t = -4.29$ ,  $p < 0.01$ ) and -1.06 ( $t = -7.30$ ,  $p < 0.01$ ), respectively, indicating a negative relationship between independent variables and unmeasured variables. Hence, we reject the null hypothesis that social capital is exogenous in the 99% confidence intervals and social capital in Internet access and mobile phone access is endogenous. Compared with probit estimates for social capital in model 1 (see Table 3), the two-step IV-probit estimate for social capital is significant in model 2, suggesting that probit estimates ignore the endogenous effect of unobserved latent variables and underestimate the true effect of social capital on digital access. Taken together, we conclude that social capital is an endogenous variable.

To investigate how social capital increases the digital divide, our models in Table 3 include living arrangements. Since our measure of social capital is based on reciprocity within strong ties, household size and living situation (with adult children/grandchildren or not) may directly influence the digital access of elderly people. In the Internet model and mobile phone model, results indicate that household size and living with children/grandchildren can help increase digital access. In model 2, the coefficients of living with children (0.74,  $t = 13.46$ ,  $p < 0.01$ ) are relatively larger than those of living with grandchildren (0.17,  $t = 2.15$ ,  $p < 0.05$ ) and household size (0.06,  $t = 4.91$ ,  $p < 0.01$ ). Similarly, in

model 5, the coefficients are 0.58 ( $t = 10.10$ ,  $p < 0.01$ ) vs 0.28 ( $t = 3.27$ ,  $p < 0.01$ ) and 0.14 ( $t = 6.72$ ,  $p < 0.01$ ); the variance of coefficients implies that living with children can provide more digital access than living with grandchildren and the size of household. We conclude that living with offspring is an important channel by which elderly people benefit from social capital. Altogether, we conclude that social capital has a significant effect on older individuals' digital access, and support from strong ties is a path by which social capital may increase digital access.

**5.2. SES and Digital Divide.** In models 2 and 5 (see Table 3), variables of SES are strongly associated with digital access. Specifically, Internet and mobile phone access are associated with wealth and education. In model 2, a higher education level predicts higher Internet access, with coefficients of primary, medium, and high educational levels being 0.19 ( $t = 4.13$ ,  $p < 0.01$ ), 0.46 ( $t = 9.25$ ,  $p < 0.01$ ), and 0.88 ( $t = 7.67$ ,  $p < 0.01$ ), respectively. In model 5, they are 0.31 ( $t = 7.38$ ,  $p < 0.01$ ), 0.51 ( $t = 10.19$ ,  $p < 0.01$ ), and 0.24 ( $t = 1.73$ ,  $p < 0.01$ ), respectively. Another significant variable to measure SES in developing countries is urban residence, which is positively associated with digital access in both models estimated by IVs. Conversely, lack of infrastructure and of stable income can result in low access to ICT in rural areas.

We note that, in models 2 and 5, chronic diseases have a significant, negative correlation with digital access (-0.22,  $t = -4.97$ ,  $p < 0.01$  vs -0.11,  $t = -2.50$ ,  $p < 0.05$ ). Since healthcare costs for chronic disease patients are higher than for other elderly patients, chronic diseases can decrease the money that the individual can spend on ICT. In addition, both Internet and mobile phone access are negatively associated with age, with older people being more disadvantaged than younger people. Empirical results suggest that chronic diseases and age contribute to the structural inequality surrounding digital access for the elderly. In sum, digital access among the elderly is negatively associated with SES, which is consistent with prior studies [10, 11], and digital access is more than the difference between haves and have-nots; it also represents structural inequality. More importantly, age and chronic diseases broaden the digital divide, potentially preventing the elderly from benefiting from m-health services.

**5.3. Robustness Checks: Instrument Validity.** To satisfy the exclusion-restriction condition, *Migration* and *Siblings* are expected to be associated with social capital but to drive changes in digital access. To prove that IVs are exogenous, we perform a battery of tests. Since we have two instruments (*Migration* and *Siblings*) and one instrumented (Social capital), the estimation of an overidentified model can be performed, we introduce Generalized Method of Moments (GMM) to compare estimators and variance estimates for overidentified models (see Table 4). The results of Hansen's test are shown in Table 4. In model 7, because Hansen's  $J$  chi2 equals 8.04 ( $p < 0.01$ ), we reject the null hypothesis and conclude that the overidentifying restriction (specifically, at least one of the instruments) is not valid. Model 8 supports

TABLE 2: First-stage estimates.

Social capital	Coef.	<i>t</i> -value	<i>p</i> -value	Sig
Migration	-0.05	-3.01	0.000	***
Siblings	0.04	8.18	0.000	***
PCE (Log)	0.05	11.04	0.000	***
PrimaryEDU	0.04	1.98	0.050	*
MediumEDU	0.04	1.66	0.100	
HighEDU	0.29	4.27	0.000	***
Residence	-0.20	-8.34	0.000	***
ChronicDisease	0.08	3.52	0.000	***
Sex	0.04	2.36	0.020	**
Aged 60-74	-0.04	-2.17	0.030	**
Aged above 75	0.11	3.34	0.000	***
HouseholdSize	-0.02	-3.97	0.000	***
CoresidenceChildren	-0.17	-7.17	0.000	***
CoresidenceGrandchildren	0.16	3.87	0.000	***
Marital status	0.08	3.27	0.000	***
Community infrastructure	0.00	1.26	0.210	
PublicInvestment(Log)	-0.07	-9.35	0.000	***
Constant	0.67	13.51	0.000	***
F-test		$F(17, 16298) = 34.57$ prob > $F = 0.0000$		
Observations		16316		
R-squared		0.03		

*t*-values are in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

TABLE 3: Estimates of effect on digital access.

	Internet access			Mobile phone access		
	Probit model (1)	Two-step IV-probit model (2)	IV-probit model (3)	Probit model (4)	Two-step IV-probit model (5)	IV-probit model (6)
PCE(Log)	0.10*** (8.72)	0.06*** (4.08)	0.04*** (2.59)	0.12*** (18.14)	0.06*** (4.62)	0.04*** (3.17)
PrimaryEDU	0.23*** (5.41)	0.19*** (4.13)	0.14*** (3.33)	0.36*** (11.74)	0.31*** (7.38)	0.19*** (5.06)
MediumEDU	0.50*** (11.28)	0.46*** (9.25)	0.34*** (5.51)	0.57*** (14.78)	0.51*** (10.19)	0.32*** (6.03)
HighEDU	1.11*** (12.99)	0.88*** (7.67)	0.66*** (4.28)	0.61*** (5.57)	0.24 * (1.73)	0.15 (1.43)
Residence	0.64*** (18.85)	0.79*** (14.07)	0.60*** (12.29)	0.03 (0.90)	0.26*** (4.20)	0.16*** (5.49)
ChronicDisease	-0.16*** (-4.14)	-0.22*** (-4.97)	-0.17*** (-5.21)	-0.02 (-0.50)	-0.11*** (-2.50)	-0.07*** (-2.65)
SocialCapital	0.02 (1.52)	0.79*** (3.72)	0.64*** (6.50)	0.10*** (7.87)	1.29*** (5.88)	0.81*** (17.19)
Sex	0.14*** (4.99)	0.11*** (3.32)	0.08*** (2.73)	0.07*** (2.83)	0.02 (0.57)	0.01 (0.55)
Aged 60-74	-0.31*** (-9.19)	-0.27*** (-6.68)	-0.20*** (-4.33)	-0.49*** (-17.78)	-0.41*** (-10.54)	-0.26*** (-5.77)
Aged above 75	-0.34*** (-5.06)	-0.38*** (-5.18)	-0.29*** (-4.81)	-0.93*** (-20.72)	-0.99*** (-15.76)	-0.61*** (-8.50)
HouseholdSize	0.04*** (4.31)	0.06*** (4.91)	0.05*** (5.58)	0.11*** (9.13)	0.14*** (9.72)	0.09*** (8.24)
CoresidenceChildren	0.62*** (15.78)	0.74*** (13.46)	0.56*** (10.99)	0.39*** (10.42)	0.58*** (10.10)	0.36*** (10.28)
Coresidence grand children	0.28*** (4.17)	0.17*** (2.15)	0.12* (1.80)	0.45*** (6.96)	0.28*** (3.27)	0.17*** (2.67)
Marital	0.14*** (2.82)	0.07 (1.29)	0.05 (1.15)	0.10*** (2.73)	-0.01 (-0.13)	-0.00 (-0.15)
CommunityInfrastructure	0.09*** (22.42)	0.08*** (18.93)	0.06*** (7.19)	0.02*** (4.86)	0.01*** (2.87)	0.01*** (2.60)
PublicInvestment(Log)	0.08*** (7.54)	0.14*** (7.13)	0.10*** (10.79)	0.05*** (4.68)	0.13*** (6.40)	0.08*** (9.16)



TABLE 3: Continued.

	Internet access			Mobile phone access		
	Probit model (1)	Two-step IV-probit model (2)	IV-probit model (3)	Probit model (4)	Two-step IV-probit model (5)	IV-probit model (6)
Constant	-3.22*** (-29.71)	-3.79*** (-12.50)	-2.89*** (-12.59)	-0.92*** (-9.66)	-1.82*** (-18.25)	-1.13***
athrho			-0.79*** (-4.29)			-1.06*** (-7.30)
Insigma	Wald test of exogeneity: chi <sup>2</sup> (1) = 18.41 prob > chi <sup>2</sup> = 0.0000			Wald test of exogeneity: chi <sup>2</sup> (1) = 53.33 prob > chi <sup>2</sup> = 0.0000		
Obs.	16316	16316	16316	16316	16316	16316
Pseudo R <sup>2</sup>	0.26	.z	.z	0.20	.z	.z

*t*-values are in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

TABLE 4: Regression results GMM.

	Internet access (model 7)		Mobile phone access (model 8)	
	Coef.	t-value	Coef.	t-value
SocialCapital	0.12***	(3.29)	0.32***	(5.97)
PCE (Log)	0.01***	(3.87)	0.02***	(5.51)
PrimaryEDU	0.02 **	(2.32)	0.09***	(8.12)
MediumEDU	0.08***	(9.32)	0.13***	(10.34)
HighEDU	0.29***	(10.28)	0.05 *	(1.58)
Residence	0.18***	(15.47)	0.07***	(4.77)
ChronicDisease	-0.04***	(-5.45)	-0.02**	(-2.14)
Sex	0.02***	(3.55)	0.00	(0.49)
Aged 60-74	-0.04***	(-6.17)	-0.11***	(-11.35)
Aged above 75	-0.06***	(-5.79)	-0.31***	(-17.73)
HouseholdSize	0.01 **	(2.53)	0.04***	(11.43)
CoresidenceChildren	0.14***	(13.81)	0.15***	(10.99)
CoresidenceGrandchildren	0.02	(1.22)	0.10***	(5.25)
Marital	0.01	(0.71)	0.00	(0.33)
Community infrastructure	0.02***	(19.52)	0.00***	(2.55)
Public investment (Log)	0.02***	(7.02)	0.03***	(6.47)
Constant	-0.26***	(-8.11)	0.07	(1.59)
Obs.	16316		16316	
R-squared	0.11		.	
Test of overidentifying restriction:	Hansen's <i>J</i> chi2 (1) = 8.95 ( $p = 0.00$ )		Hansen's <i>J</i> chi2 (1) = 0.62 ( $p = 0.43$ )	

*t*-values are in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

the null hypothesis ( $p > 0.01$ ) and we conclude that IVs are more valid in model 8 than in model 7. Therefore, the causal effect of social capital cannot be guaranteed in the Internet access model as it can in the mobile phone access model.

The difference could be partly explained by the fact that Internet access depends more on massive infrastructure investment at the community and family level than mobile phone access, which depends on an individual's capability to purchase and use personal digital equipment. As shown in Figures 2 and 3, two SES patterns influence digital access. Internet access indicates the infrastructure pattern, which is more sensitive to SES indicators than the personal device pattern represented by mobile phone access. The pattern can be further revealed by comparing the coefficients of social capital and SES in models 2 and 5. For instance, in model 5 predicting mobile phone access, demographic characteristics such as age of 60-74 and age of 75 and

above have greater coefficients than the residence representing SES ( $-0.41$ ,  $t = -10.54$ ,  $p < 0.01$  and  $-0.99$ ,  $t = -15.76$ ,  $p < 0.01$  vs.  $0.26$ ,  $t = 4.20$ ,  $p < 0.01$ ), while it is the opposite in model 2 ( $-0.27$ ,  $t = -0.27$ ,  $p < 0.01$  and  $-0.38$ ,  $t = -5.18$ ,  $p < 0.01$  vs.  $0.79$ ,  $t = 14.07$ ,  $p < 0.01$ ). Considering the disparity in development between urban areas and rural areas in China, residents with high SES enjoy an advantage in access to ICT in the infrastructure pattern. Therefore, social influence (e.g., social capital) is expected to work in personal device pattern in developing countries.

## 6. Conclusion and Discussion

**6.1. Empirical Conclusion.** Our study makes several contributions. First, we determined the causal role of social capital in facilitating increased digital access, but we did not determine a causal effect of social capital on infrastructure

patterns; to put it another way, our findings uncover the boundary of the social influence on digital divide. In line with prior studies, we find association between SES and digital divide among the elderly. However, age and chronic diseases also contribute to the divide, which poses a challenge to m-health policymakers in developing countries. Second, our measure of social capital is built on strong ties (relatives and friends rather than community members), and the causal relationship between social capital and digital access implies that social capital as private goods plays a critical role in China. Third, our results have implications for m-health policy in developing regions. For the elderly in developing world, despite the underprivileged are confronted with the risk of m-health divide, potential resources associated with social capital can help increase digital access. These findings remind policy makers of the importance of unique characteristics of elderly users, and the increase of access to m-health should be a priority due to the potential m-health divide.

To sum up, the main idea suggested by our study is that mobile health is more than adoption of ICT for the elderly; it is also an issue of digital equality. While prior research in the medical and IS fields focuses on urban areas and patients with relatively high SES, the underprivileged in developing countries deserve more attention. Since social capital has a positive effect on digital access, we propose that m-health policy in developing countries should fully exploit local resources including strong ties and community connections.

**6.2. M-Health Policy Discussion.** The digital divide is part of structural disparity, and continuing disparity will decrease m-health access and potentially cause an m-health divide. To avoid potential m-health divide, we suggest a family-centered m-health policy. First, family-centered m-health policy should take digital access for the elderly as a priority. Government or NGOs could provide subsidies to family members living with the elderly to increase m-health technology access. As our results suggested, social network resources should be considered to help bury structural barriers such as the urban-rural disparity and low education among the elderly. Such a family-centered m-health policy would be expected to work in developing countries. Second, family-centered m-health policy should view family as basic unit to receive m-health services to improve health outcomes. As suggested by prior research [21], m-health service implementation depends on both external and internal resources including communities and families; given the disadvantaged digital access and digital capability among the elderly, family members in developing countries are expected to facilitate the implementation of m-health services; for instance, family members could facilitate the management of chronic diseases with the guidance of general physicians.

Our findings will benefit healthcare reforms on supply side in developing countries. Due to the rapid aging and increasing threat of chronic diseases, programs like Health China 2020 in China, Family Medicine Program in Turkey

[55], and Family Health Program in Brazil [56] emphasize the role of general physicians in community hospitals, but it is still unknown how to combine mobile technology and primary service. Particularly, chronic disease management requires a different service pattern compared to infectious disease services [57]; policies should pay more attention to m-health in the future for the ICT empowerment role for elderly patients.

## Data Availability

The data of China Health and Retirement Longitudinal Study (CHARLS 2011) are available at <http://charls.pku.edu.cn/pages/data/2011-charls-wave1/zh-cn.html>. The data of the Statistical Yearbook of China's Regional Economy (2012) are available at [http://www.stats.gov.cn/tjsj/tjcbw/201303/t20130318\\_451532.html](http://www.stats.gov.cn/tjsj/tjcbw/201303/t20130318_451532.html).

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Award nos. 71531007, 71622002, 71471048, 71471049, 71490720, and 71771065).

## References

- [1] WHO, *World Report on Ageing and Health*, World Health Organization, Geneva, Switzerland, 2016.
- [2] WHO, *Global Status Report on Noncommunicable Diseases 2010*, World Health Organization, Geneva, Switzerland, 2011.
- [3] D. Lupton, *Digital Health: Critical and Cross-Disciplinary Perspectives, Critical Approaches to Health*, Routledge, Taylor & Francis Group, London, UK, 2018.
- [4] M. Mars and R. E. Scott, "Global E-health policy: a work in progress," *Health Affairs*, vol. 29, no. 2, pp. 237–243, 2010.
- [5] D. Y. Chao, T. M. Lin, and W.-Y. Ma, "Enhanced self-efficacy and behavioral changes among patients with diabetes: cloud-based mobile health platform and mobile app service," *JMIR Diabetes*, vol. 4, no. 2, p. e11017, 2019.
- [6] M. Stellefson, B. Chaney, A. E. Barry et al., "Web 2.0 chronic disease self-management for older adults: a systematic review," *Journal of Medical Internet Research*, vol. 15, no. 2, p. e35, 2013.
- [7] X. Wu, X. Guo, and Z. Zhang, "The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis," *JMIR mHealth and uHealth*, vol. 7, no. 1, p. e12297, 2019.
- [8] T. Loncar-Turukalo, E. Zdravovski, J. Machado da Silva, I. Chouvarda, and V. Trajkovic, "Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers," *Journal of Medical Internet Research*, vol. 21, no. 9, p. e14017, 2019.
- [9] P. Maresova, O. Krejcar, S. Barakovic et al., "Health-related ICT solutions of smart environments for elderly-systematic review," *IEEE Access*, vol. 8, pp. 54574–54600, 2020.
- [10] Y. A. Hong, Z. Zhou, Y. Fang, and L. Shi, "The digital divide and health disparities in China: evidence from a national

- survey and policy implications,” *Journal of Medical Internet Research*, vol. 19, no. 9, p. e317, 2017.
- [11] Y. Jin, M. Jing, L. Zhang, S. Song, and X. Ma, “Internet access and hypertension management among the elderly population: a nationally representative cross-sectional survey in China,” *Journal of Medical Internet Research*, vol. 21, no. 1, p. e11280, 2019.
- [12] ITU 2018, *ICTs, LDCs and the SDGs: Achieving Universal and Affordable Internet in the LDCs. Thematic Report. In Partnership with the United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States (UN-OHRLS)*, ITU, Geneva, Switzerland, 2018, <https://www.itu.int/en/ITU-D/LDCs/Pages/Publications/LDCs/D-LDC-ICTLDC-2018-PDF-E.pdf>.
- [13] UN 2019, “Digital economy report,” 2019.
- [14] Y. Fujisawa, T. Hamano, and S. Takegawa, “Social capital and perceived health in Japan: an ecological and multilevel analysis,” *Social Science & Medicine*, vol. 69, no. 4, pp. 500–505, 2009.
- [15] G. Hollard and O. Sene, “Social capital and access to primary health care in developing countries: evidence from Sub-Saharan Africa,” *Journal of Health Economics*, vol. 45, pp. 1–11, 2016.
- [16] M. K. Keil, J. Merlo, I. Kawachi, M. Lindström, and U.-G. Gerdtham, “Social capital and health: does egalitarianism matter? A literature review,” *International Journal for Equity in Health*, vol. 5, no. 1, 2006.
- [17] I. Kawachi and L. Berkman, “Social cohesion, social capital, and health,” *Social Epidemiology*, vol. 174, 2000.
- [18] S. Petrou and E. Kupek, “Social capital and its relationship with measures of health status: evidence from the Health Survey for England 2003,” *Health Economics*, vol. 17, no. 1, pp. 127–143, 2008.
- [19] G. Hofstede, G. J. Hofstede, and M. Minkov, *Cultures and Organizations: Software of the Mind. Revised and Expanded*, McGraw-Hill, New York, NY, USA, 3rd edition, 2010.
- [20] W. Yip, S. V. Subramanian, A. D. Mitchell, D. T. S. Lee, J. Wang, and I. Kawachi, “Does social capital enhance health and well-being? Evidence from rural China,” *Social Science & Medicine*, vol. 64, no. 1, pp. 35–49, 2007.
- [21] J. G. Kahn, J. S. Yang, and J. S. Kahn, “‘Mobile’ health needs and opportunities in developing countries,” *Health Affairs*, vol. 29, no. 2, pp. 252–258, 2010.
- [22] X.-Q. Wang and P.-J. Chen, “Population ageing challenges health care in China,” *The Lancet*, vol. 383, no. 9920, p. 870, 2014.
- [23] WHO, *China Country Assessment Report on Ageing and Health*, World Health Organization, Geneva, Switzerland, 2015, <http://www.who.int/ageing/publications/china-country-assessment/en/>.
- [24] L. Rocco, E. Fumagalli, and M. Suhrcke, “From social capital to health - and back,” *Health Economics*, vol. 23, no. 5, pp. 586–605, 2014.
- [25] J. Grytten and I. Skau, “Educational inequalities in access to fixed prosthodontic treatment in Norway. causal effects using the introduction of a school reform as an instrumental variable,” *Social Science & Medicine*, vol. 260, Article ID 113105, 2020.
- [26] R. Hsieh, “Understanding digital inequality: comparing continued use behavioral models of the socio-economically advantaged and disadvantaged,” *MIS Quarterly*, vol. 32, no. 1, p. 97, 2008.
- [27] K.-K. Wei, H.-H. Teo, H. C. Chan, and B. C. Y. Tan, “Conceptualizing and testing a social cognitive model of the digital divide,” *Information Systems Research*, vol. 22, no. 1, pp. 170–187, 2011.
- [28] J.-Y. Jung, J. L. Qiu, and Y.-C. Kim, “Internet connectedness and inequality,” *Communication Research*, vol. 28, no. 4, pp. 507–535, 2001.
- [29] E. P. Bucy, “Social access to the internet,” *Harvard International Journal of Press/Politics*, vol. 5, no. 1, pp. 50–61, 2000.
- [30] P. Bourdieu, “The forms of capital,” in *Readings in Economic Sociology*, N. W. Biggart, Ed., Blackwell Publishers Ltd, Oxford, UK, pp. 280–291, 2002.
- [31] P. S. Adler and S.-W. Kwon, “Social capital: prospects for a new concept,” *The Academy of Management Review*, vol. 27, no. 1, p. 17, 2002.
- [32] R. D. Putnam, *Bowling alone: America’s declining social capital*, in: *Culture and Politics*, Springer, Berlin, Germany, 2000.
- [33] E. L. Glaeser, D. Laibson, and B. Sacerdote, “An economic approach to social capital,” *The Economic Journal*, vol. 112, pp. 437–458, 2002.
- [34] C. Y. Ho, “Better health with more friends: the role of social capital in producing health,” *Health Economics*, vol. 25, no. 1, pp. 91–100, 2016.
- [35] N. Lin, *Social Capital: A Theory of Social Structure and Action*, Cambridge University Press, Cambridge, UK, 2002.
- [36] J. Nahapiet and S. Ghoshal, “Social capital, intellectual capital, and the organizational advantage,” *Academy of Management Review*, vol. 23, no. 2, pp. 242–266, 1998.
- [37] K. Leung, R. S. Bhagat, N. R. Buchan, M. Erez, and C. B. Gibson, “Culture and international business: recent advances and their implications for future research,” *Journal of International Business Studies*, vol. 36, no. 4, pp. 357–378, 2005.
- [38] R. Agarwal, A. Animesh, and K. Prasad, “Social interactions and the ‘digital divide’: explaining regional variations in internet use,” *SSRN Electronic Journal*, vol. 20, no. 2, pp. 277–294, 2005.
- [39] V. Aljafari and T. A. Sykes, “Digital divide initiative success in developing countries: a longitudinal field study in a village in India,” *Information Systems Research*, vol. 24, no. 2, pp. 239–260, 2013.
- [40] V. Venkatesh, A. Rai, A. Rai, and T. A. Sykes, “Combating infant mortality in rural India: evidence from a field study of eHealth kiosk implementations,” *MIS Quarterly*, vol. 40, no. 2, pp. 353–380, 2016.
- [41] S. Zhao, “Parental education and children’s online health information seeking: beyond the digital divide debate,” *Social Science & Medicine*, vol. 69, no. 10, pp. 1501–1505, 2009.
- [42] Y. Shen, D. E. Yeatts, T. Cai, P. Q. Yang, and C. M. Cready, “Social capital and self-rated health among middle-aged and older adults in China,” *Research on Aging*, vol. 36, no. 4, pp. 497–521, 2013.
- [43] S. Kolenikov and G. Angeles, “Socioeconomic status measurement with discrete proxy variables: is principal component analysis a reliable answer?” *Review of Income and Wealth*, vol. 55, no. 1, pp. 128–165, 2009.
- [44] J. Strauss and D. Thomas, “Chapter 54 health over the life course,” *Handbook of Development Economics*, vol. 4, pp. 3375–3474, 2007.
- [45] S. Moussavi, S. Chatterji, E. Verdes, A. Tandon, V. Patel, and B. Ustun, “Depression, chronic diseases, and decrements in health: results from the world health surveys,” *The Lancet*, vol. 370, no. 9590, pp. 851–858, 2007.
- [46] L. W. Li, J. Liu, Z. Zhang, and H. Xu, “Late-life depression in Rural China: do village infrastructure and availability of

- community resources matter?" *International Journal of Geriatric Psychiatry*, vol. 30, no. 7, pp. 729–736, 2015.
- [47] A. C. Cameron and P. K. Trivedi, *Microeconometrics Using Stata*, Stata Press, College Station, TX, USA, 2010.
- [48] R. C. Hill, W. E. Griffiths, and G. C. Lim, *Principles of Econometrics*, John Wiley & Sons, Hoboken, NJ, USA, 2018.
- [49] C. Herberholz and S. Phuntsho, "Social capital, outpatient care utilization and choice between different levels of health facilities in rural and urban areas of Bhutan," *Social Science & Medicine*, vol. 211, pp. 102–113, 2018.
- [50] X. Fei, G. G. Hamilton, and Z. Wang, *From the Soil, the Foundations of Chinese Society: A Translation of Fei Xiaotong's Xiangtu Zhongguo, with an Introduction and Epilogue*, University of California Press, Berkeley, CA, USA, 1992.
- [51] Y. Zhao, "Measuring the social capital of laid-off Chinese workers," *Current Sociology*, vol. 50, no. 4, pp. 555–571, 2002.
- [52] Y. Bian, "Bringing strong ties back in: indirect ties, network bridges, and job searches in China," *American Sociological Review*, vol. 62, no. 3, p. 366, 1997.
- [53] Y. Bian, X. Huang, and L. Zhang, "Information and favoritism: the network effect on wage income in China," *Social Networks*, vol. 40, pp. 129–138, 2015.
- [54] J. H. Stock and M. Yogo, "Testing for weak instruments in linear IV regression," in *Identification and Inference for Econometric Models*, D. W. K. Andrews, Ed., pp. 80–108, New York: Cambridge University Press, New York, NY, USA, 2002.
- [55] R. Cesur, P. M. Güneş, E. Tekin, and A. Ulker, "The value of socialized medicine: the impact of universal primary healthcare provision on mortality rates in Turkey," *Journal of Public Economics*, vol. 150, pp. 75–93, 2017.
- [56] R. Rocha and R. R. Soares, "Evaluating the impact of community-based health interventions: evidence from Brazil's Family Health Program," *Health Economics*, vol. 19, no. S1, pp. 126–158, 2010.
- [57] M. Philip, *Public and Private Roles in Health Theory and Financing Patterns*, World Bank Discussion Paper, Washington, DC, USA, 1996.

## Research Article

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

Eduardo Casilari , José A. Santoyo-Ramón , and José M. Cano-García 

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

Correspondence should be addressed to Eduardo Casilari; [ecasilari@uma.es](mailto:ecasilari@uma.es)

Received 30 October 2020; Accepted 15 November 2020; Published 8 December 2020

Academic Editor: Ivan Miguel Pires

Copyright © 2020 Eduardo Casilari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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 accidental 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, synopsis their basic properties and the testbeds (employed sensors, characteristics of the experimental users, and topology 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 MEBIOMECC [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.	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) <sup>3</sup>	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.) <sup>4</sup>	[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 <sup>1</sup>
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 87 (A)	Tens of meters ±2 g (A)
MobiFall and MobiAct	1	3 (A, G, O)	Thigh (trouser pocket)	1 smartphone	100 (G,O)	±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 <sup>3</sup>	5	$\pm 2$ g (A) $\pm 360^\circ$ (O)
UMAFall	5	3(A, G, M)	Ankle, chest, thigh, and waist	1 smartphone <sup>4</sup>	100 (SP)	$\pm 16$ g (A) $\pm 256^\circ/s$ (G)
FARSEEING	1	2 (A,G)	Wrist	4 external IMUs	20 (IMUs)	$\pm 4800 \mu T$ (M)
SisFall	1	3 (A, A, G)	Waist or thigh	1 external IMU	100	$\pm 6$ g (A) $\pm 100^\circ/s$ (G)
UniMiB SHAR	1	1 (A)	Waist	2 accelerometers and a gyroscope in a single mode	200	$\pm 16$ g (A1) $\pm 8$ g (A2) $\pm 2000^\circ/s$ (G)
SMotion	1	A, G	Thigh (left or right trouser pocket)	1 smartphone	50	$\pm 2$ g (A)
IMUFD	1	A, G	Waist	1 external IMU	51	$\pm 4$ g (A2) $\pm 500^\circ/s$ (G) $\pm 16$ g (A)
CGU-BES	1	3(A, G, M)	Chest, head, left ankle, left thigh, right ankle, right thigh, and waist	7 external IMUs	128	$\pm 2000^\circ/s$ (G) $\pm 800 \mu T$ (M)
CMDFALL <sup>1</sup>	2	2 (A, G)	Chest	1 sensing mote with a gyroscope and accelerometer	200	$\pm 3.6$ g (A) $\pm 400^\circ/s$ (G)
DU-MD	1	1 (A)	Left wrist and left hip	1 external IMU	50	$\pm 16$ g (A)
SmartFall	1	1 (A)	Wrist	1 external IMU	33	$\pm 4$ g (A)
Smartwatch	1	1 (A)	Wrist	1 external IMU	31.25	$\pm 16$ g (A)
UP-Fall	1	1 (A)	Wrist (left hand)	Smartwatch (MS band)	31.25	$\pm 8$ g (A)
DOFDA	5	2 (A, G)	Ankle, neck, and thigh (pocket)	5 external IMUs	14	$\pm 8$ g (A)
			Waist and wrist			$\pm 2000^\circ/s$ (G) $\pm 16$ g (A) $\pm 2000^\circ/s$ (G)
	1	4 (A, G, O, M)	Waist	1 external IMU	33	(G) $\pm 800 \mu T$ (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 ( $\pm 2$ g), 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

$$A_{w_{\text{diff}}}[m] = \sqrt{(A_{X_{\text{max}}}[m] - A_{X_{\text{min}}}[m])^2 + (A_{Y_{\text{max}}}[m] - A_{Y_{\text{min}}}[m])^2 + (A_{Z_{\text{max}}}[m] - A_{Z_{\text{min}}}[m])^2}, \quad (2)$$

where  $A_{X_{\text{max}}}[m]$ ,  $A_{Y_{\text{max}}}[m]$ , and  $A_{Z_{\text{max}}}[m]$  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

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

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

$$A_{w_{\text{diff}}(\text{max})} = A_{w_{\text{diff}}}[k_o] = \max(A_{w_{\text{diff}}}[m]), \quad (4)$$

$$\forall m \in [1, N - N_W + 1],$$

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

$$\mu_{\text{SMV}} = \frac{1}{N_W} \cdot \sum_{i=k_o}^{k_o+N_W-1} \text{SMV}[i], \quad (5)$$

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:

$$N_W = t_W \cdot f_s, \quad (1)$$

where  $f_s$  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 ( $A_{w_{\text{diff}}}[m]$ ) as

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

$$\text{SMV}[i] = \sqrt{(A_X[i])^2 + (A_Y[i])^2 + (A_Z[i])^2}. \quad (6)$$

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

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

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

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

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

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

- (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:

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

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  $\mathbf{A}_{w, \text{diff}(max)}$ , 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:

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

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

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

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

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

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

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

- (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]:

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

where  $\text{FFT}_X[i]$ ,  $\text{FFT}_Y[i]$ , and  $\text{FFT}_Z[i]$ , 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

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

for  $i = 0, 1, \dots, N_W - 1$ .

- (12) Mean of the autocorrelation function ( $\mu_R$ ) of the acceleration magnitude captured during the observation interval:

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

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

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

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

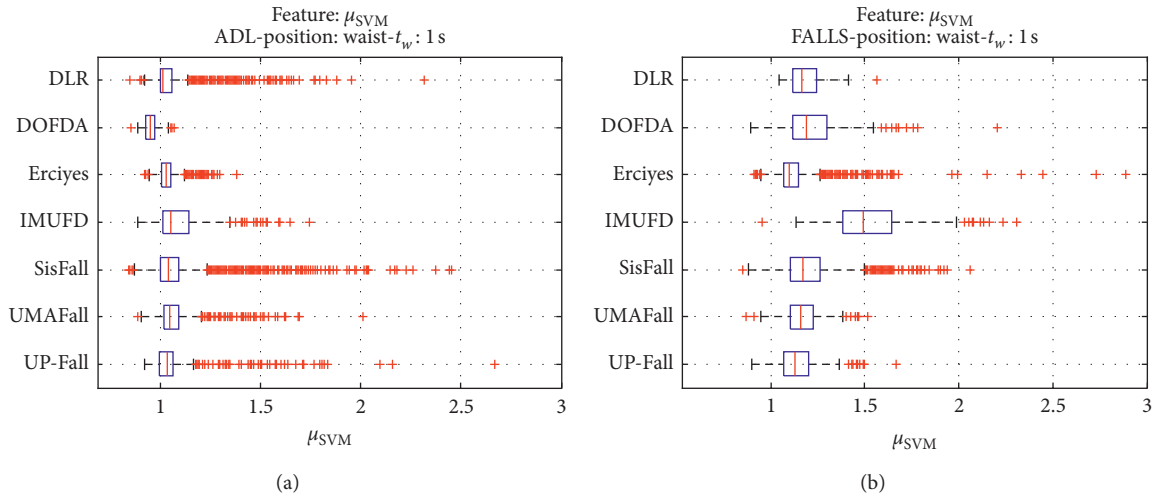


FIGURE 1: Boxplots of the mean Signal Magnitude Vector ( $\mu_{SVM}$ ) for the ADLs (left column) and falls. (a) ADLs. (b) Falls (right column) of all datasets.

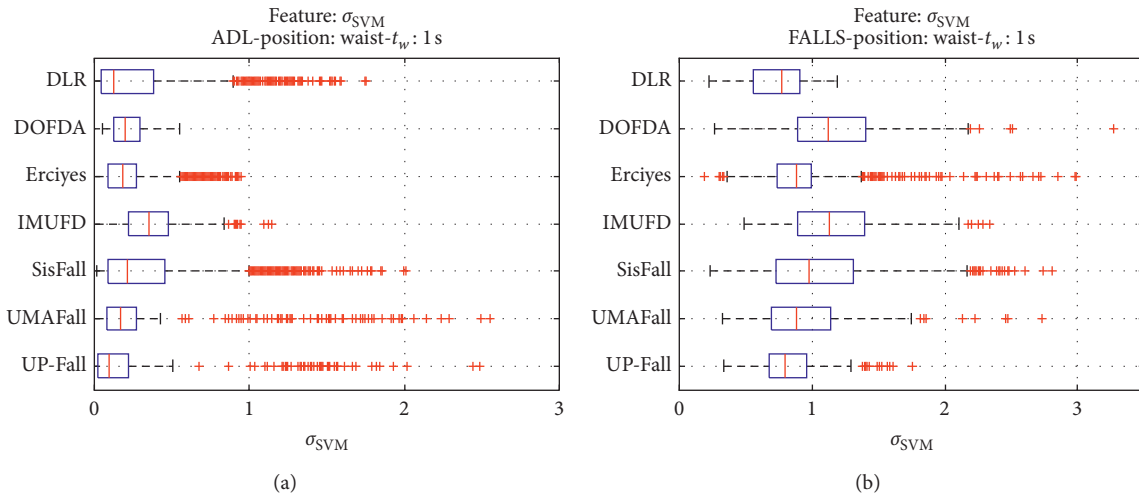


FIGURE 2: Boxplots of the maximum variation of the standard deviation of the Signal Magnitude Vector ( $\sigma_{SVM}$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

(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

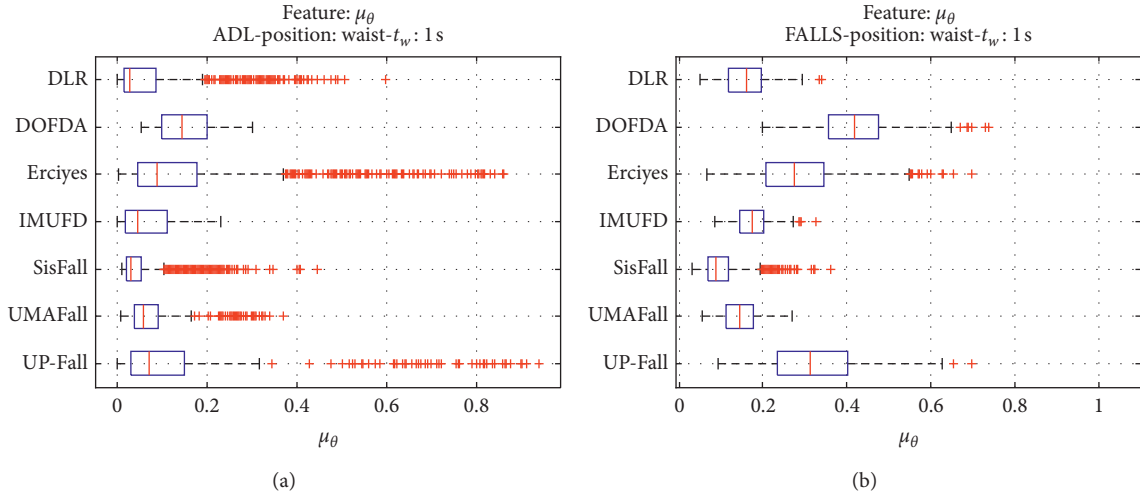


FIGURE 3: Boxplots of the mean rotation angle ( $\mu_\theta$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

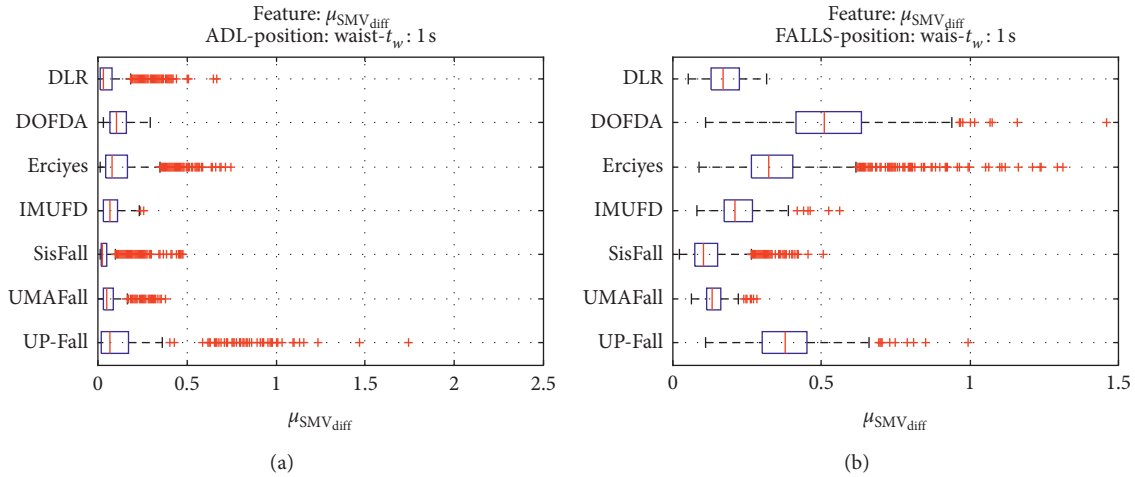


FIGURE 4: Boxplots of the mean absolute difference between consecutive samples ( $\mu_{SMV_{diff}}$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

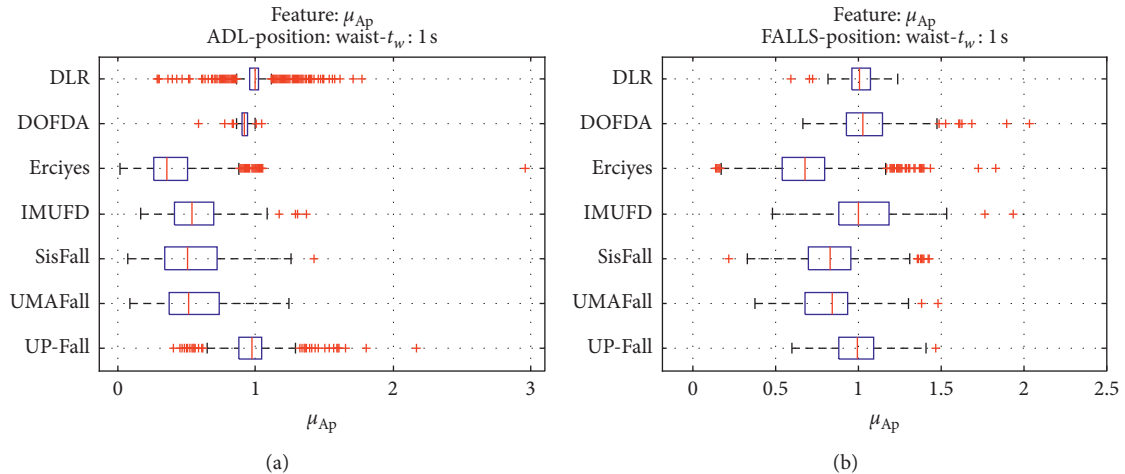


FIGURE 5: Boxplots of the mean module of the not perpendicular acceleration components ( $\mu_{Ap}$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

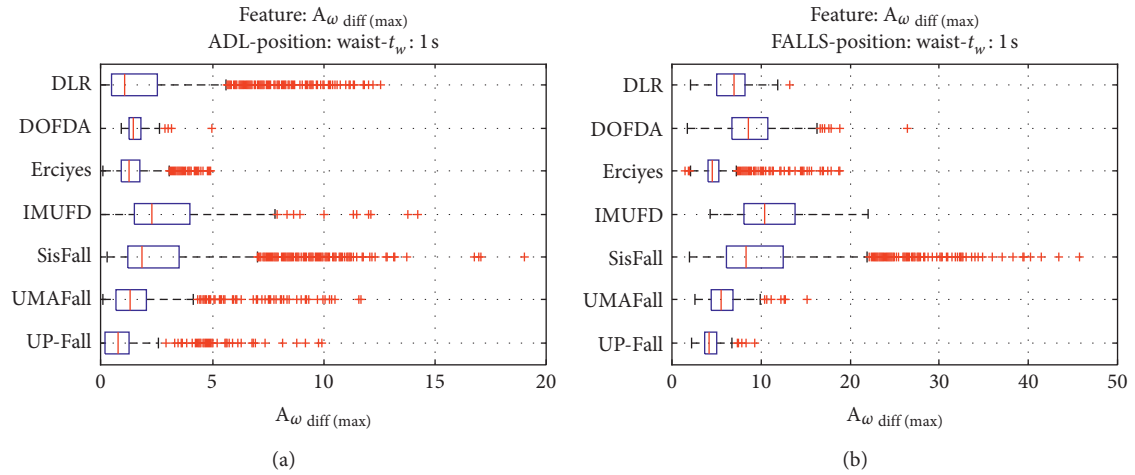


FIGURE 6: Boxplots of the maximum variation of the acceleration components ( $A_{\omega \text{ diff(max)}}$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

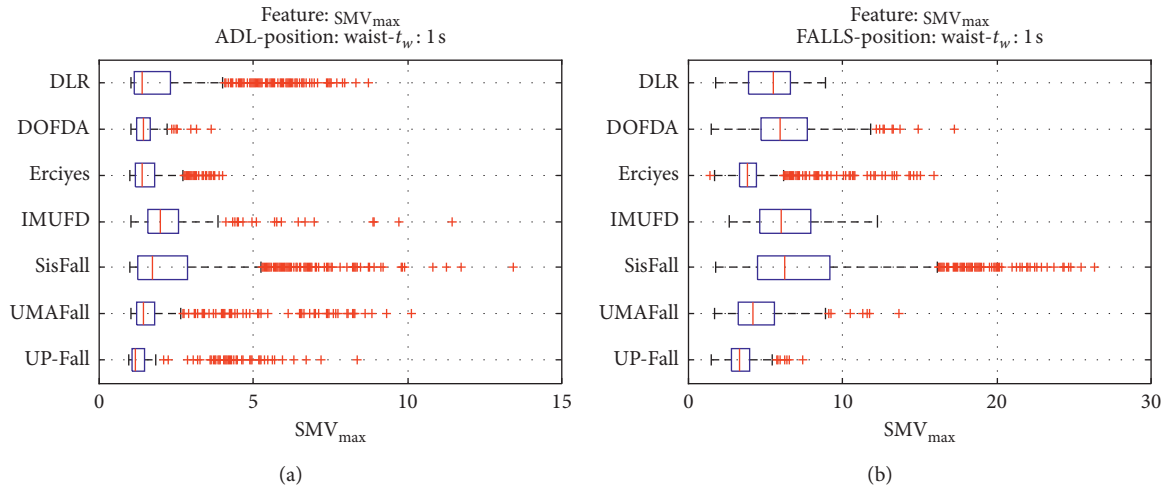


FIGURE 7: Boxplots of the maximum ( $SMV_{\text{max}}$ ) of the SMV for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

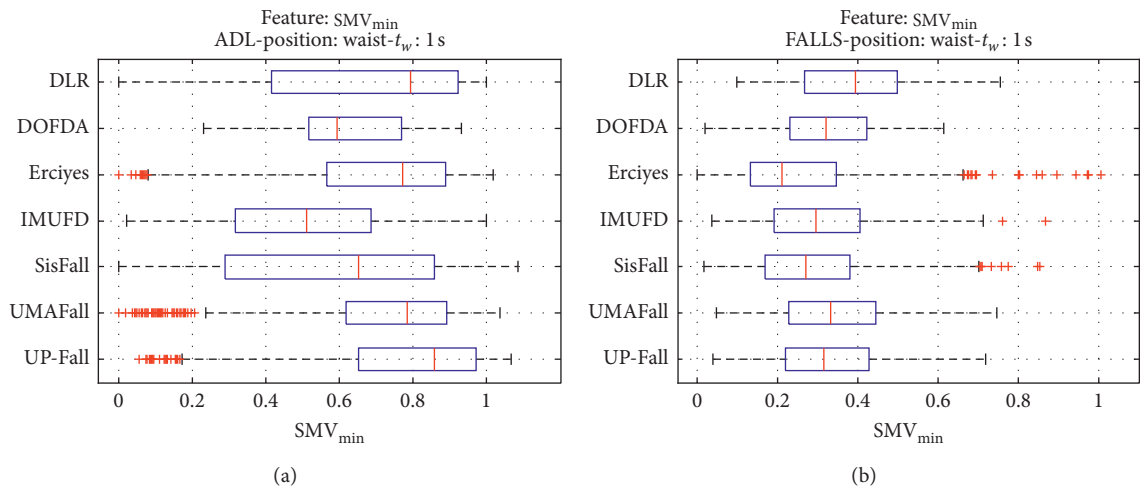


FIGURE 8: Boxplots of the minimum ( $SMV_{\text{min}}$ ) of the SMV for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.



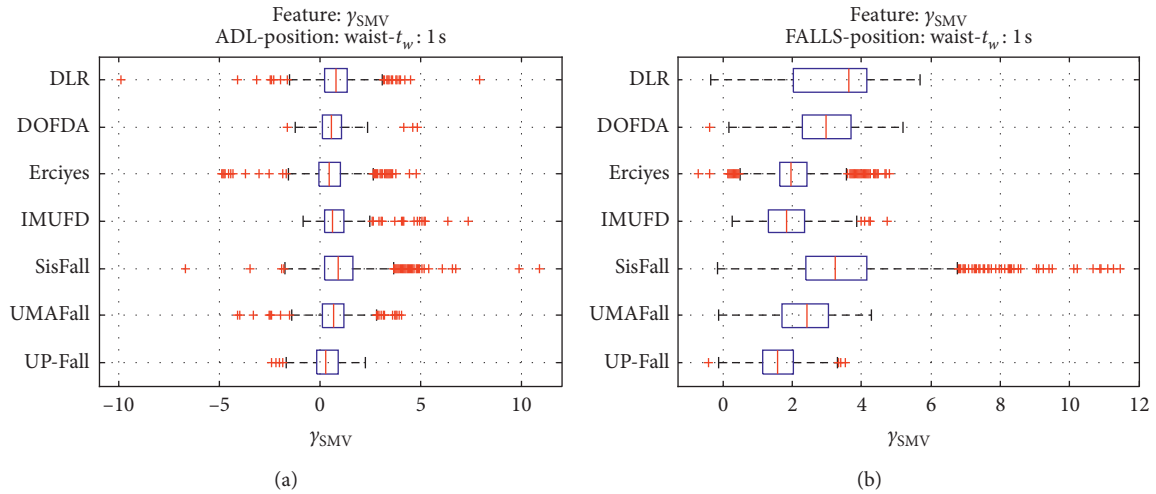


FIGURE 9: Boxplots of the skewness of SMV ( $\gamma_{SMV}$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

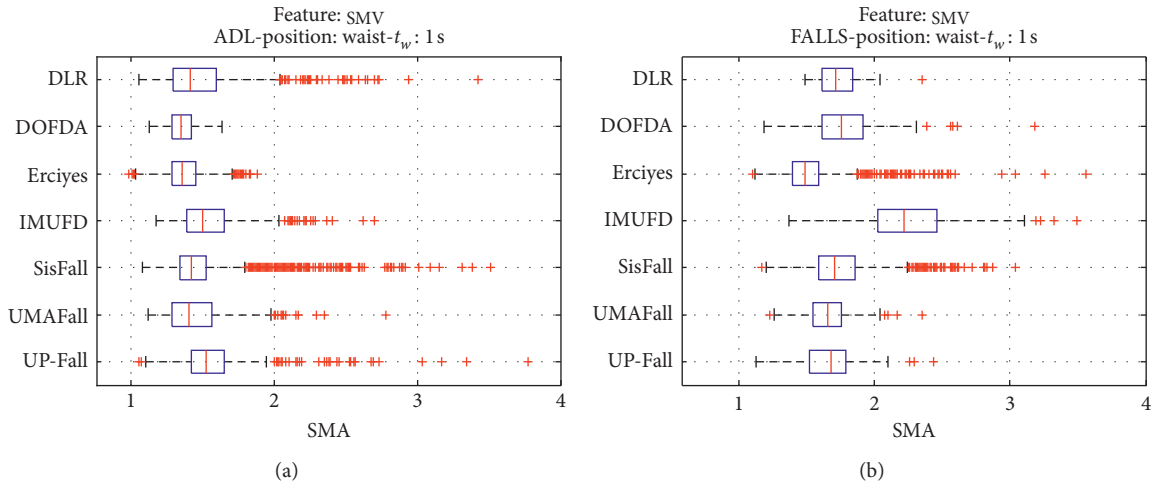


FIGURE 10: Boxplots of the Signal Magnitude Area (SMA) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

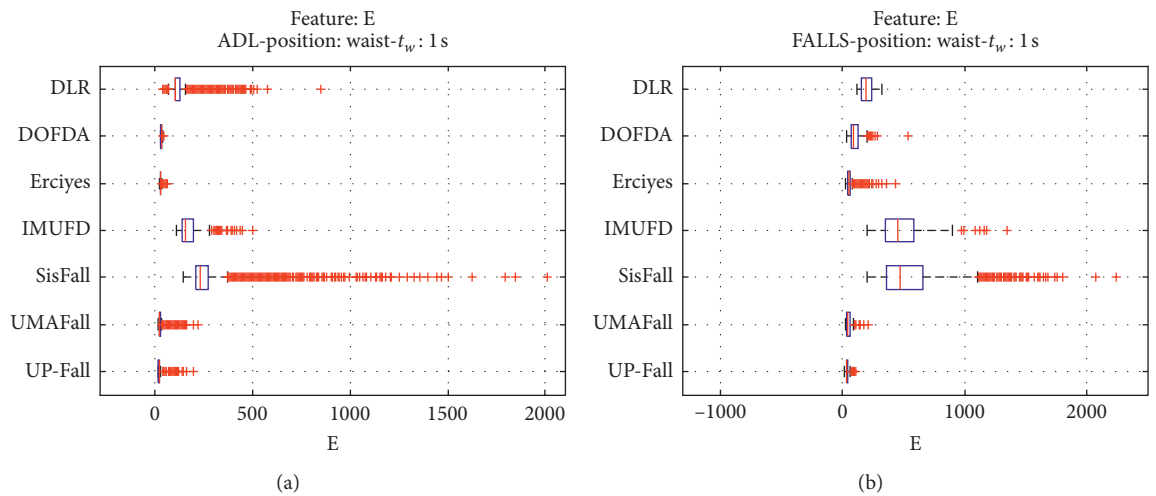


FIGURE 11: Boxplots of the sum of the energy (E) estimated in the three axes for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

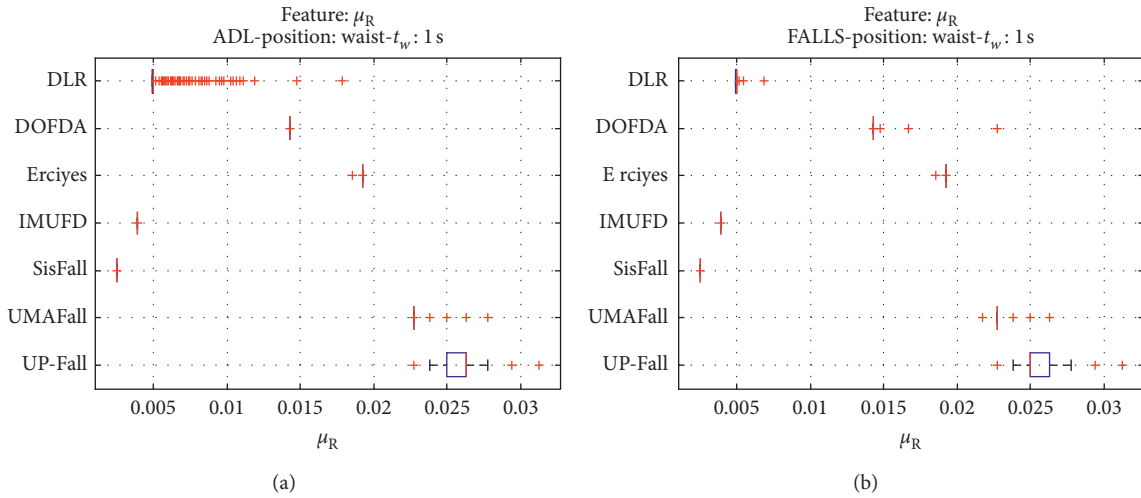


FIGURE 12: Boxplots of the mean of the autocorrelation function ( $\mu_R$ ) for the ADLs (left column) and falls (right column) of all datasets. (a) ADLs. (b) Falls.

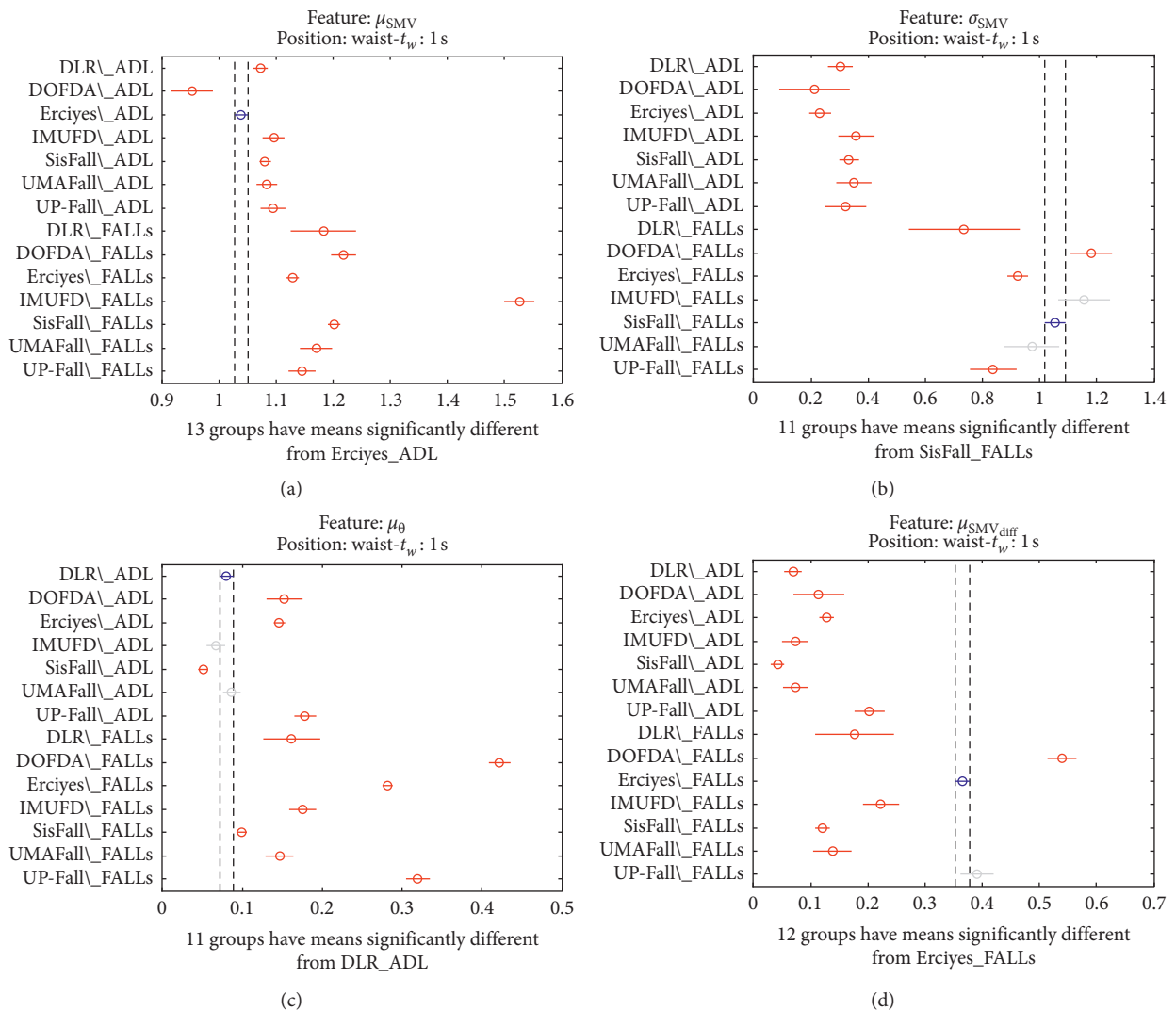


FIGURE 13: Continued.

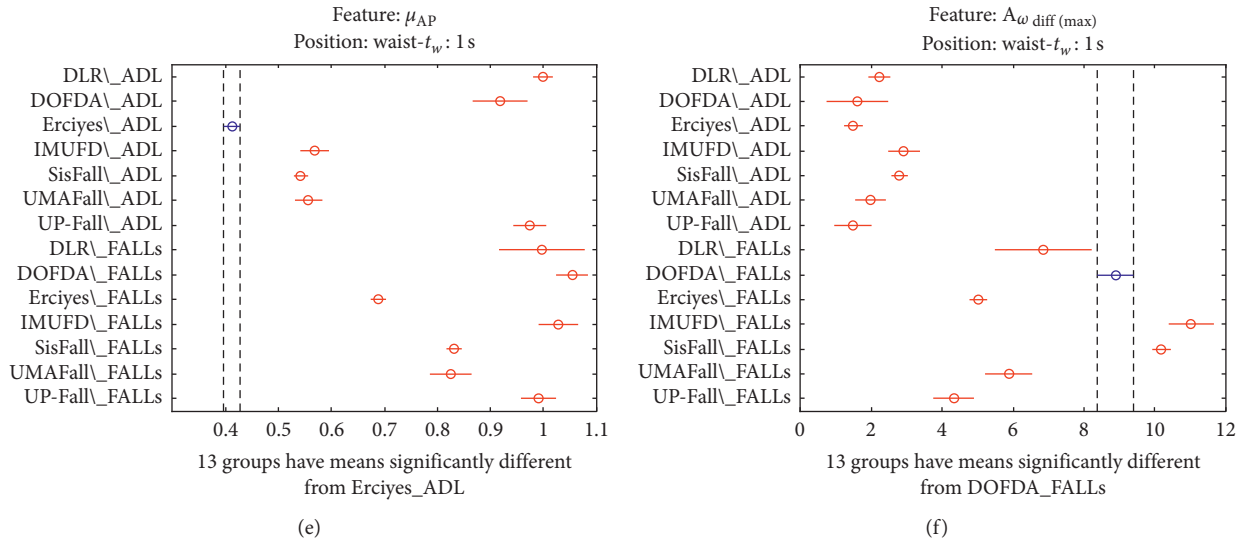


FIGURE 13: Multiple comparison test of the means of the following statistical characteristics of the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV \text{ (diff)}}$ . (e)  $\mu_{AP}$ . (f)  $A_{\omega \text{ diff(max)}}$ .

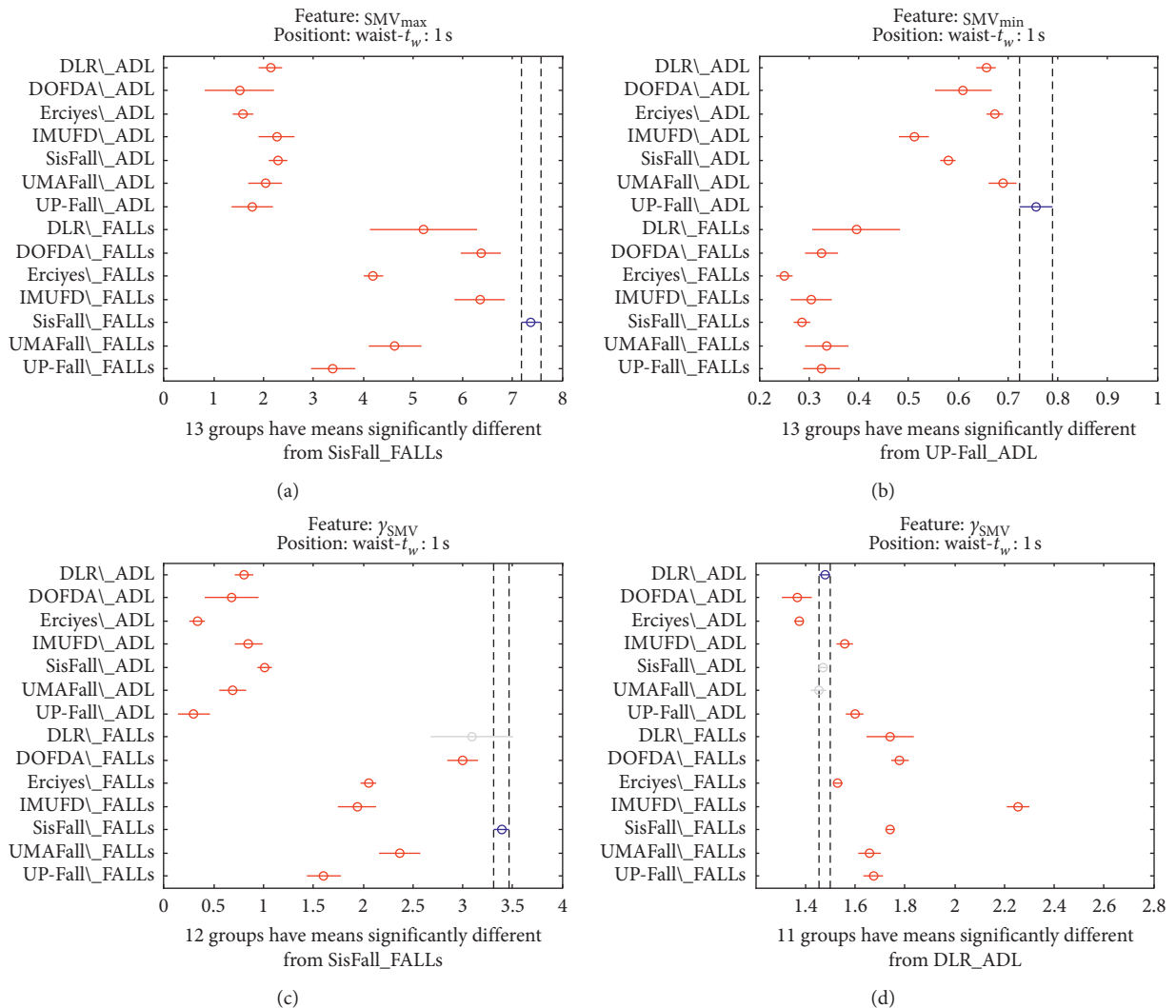


FIGURE 14: Continued.

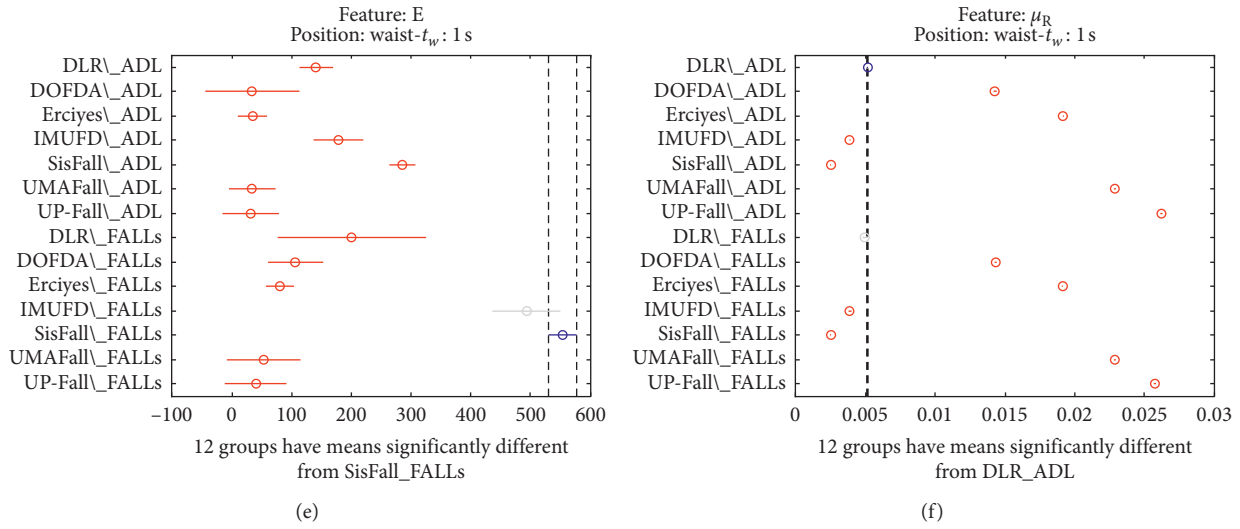


FIGURE 14: Multiple comparison test of the means of the following statistical characteristics of the datasets: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ .

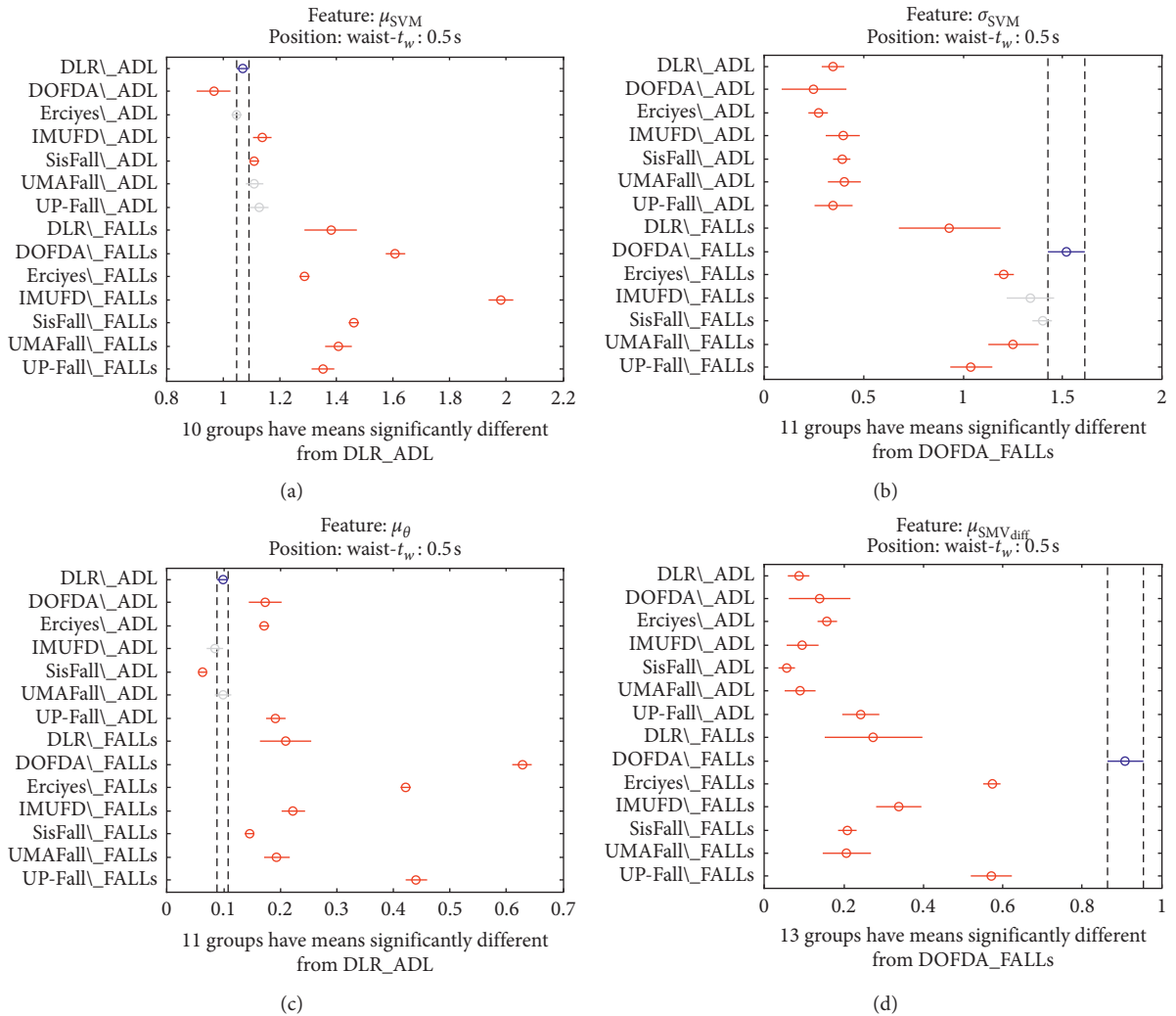


FIGURE 15: Continued.

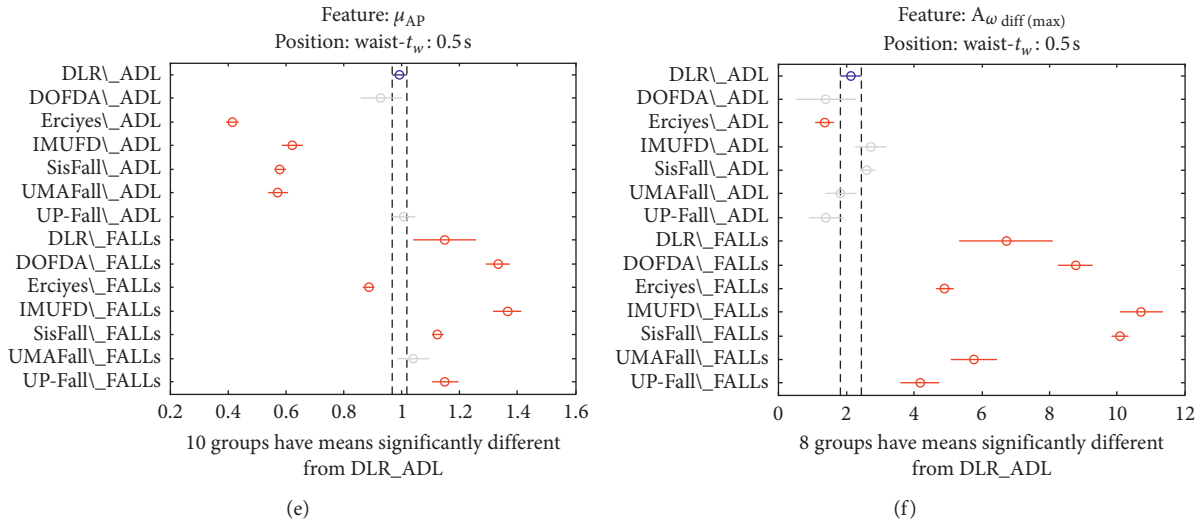


FIGURE 15: Multiple comparison test of the means of the statistical characteristics of the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV \text{ diff}}$ . (e)  $\mu_{AP}$ . (f)  $A_{\omega \text{ diff (max)}}$ . Observation window of 0.5 s.

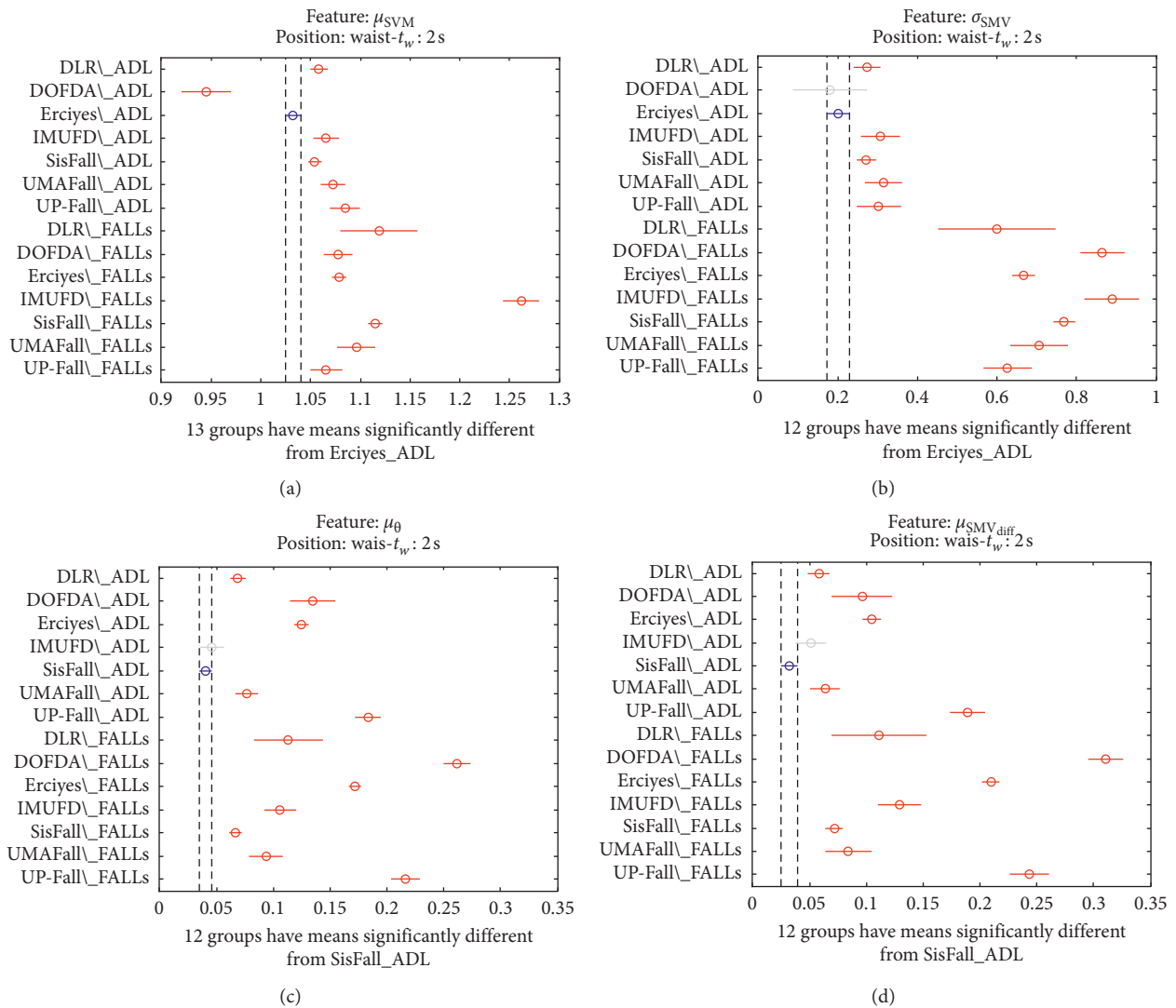


FIGURE 16: Continued.

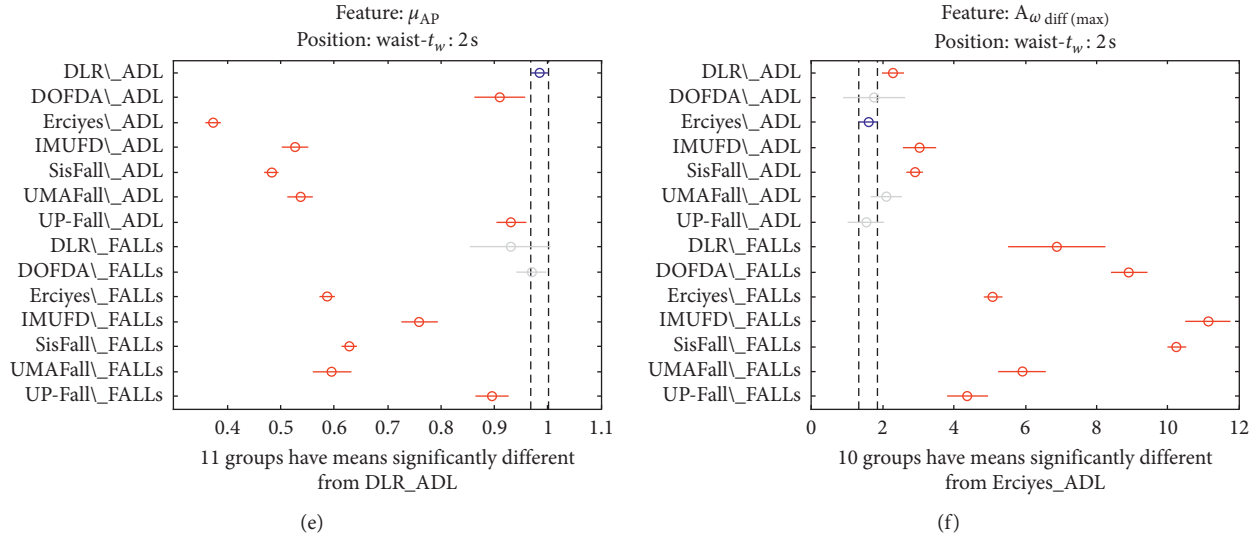


FIGURE 16: Multiple comparison test of the means of the statistical characteristics of the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV_{(diff)}}$ . (e)  $\mu_{AP}$ . (f)  $A_{\omega_{diff(max)}}$ . Observation window of 2 s.

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 ( $\sigma_{SMV}$ ,  $A_{\omega_{diff(max)}}$ , and  $SMV_{max}$ ) 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

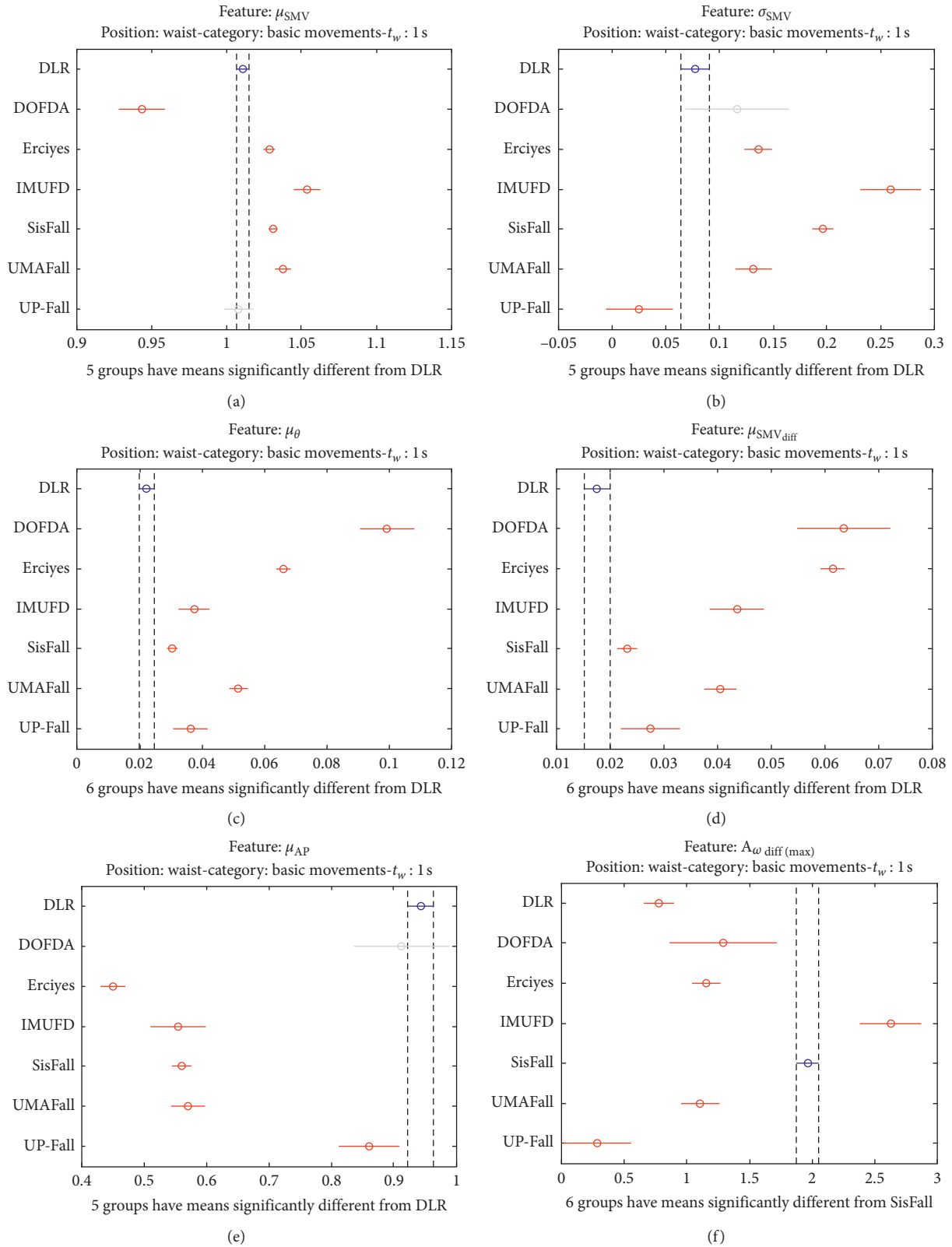


FIGURE 17: Multiple comparison test of the means of the statistical characteristics of the basic movements in the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV_{diff}}$ . (e)  $\mu_{Ap}$ . (f)  $A_{\omega_{diff(max)}}$ . Basic movements.

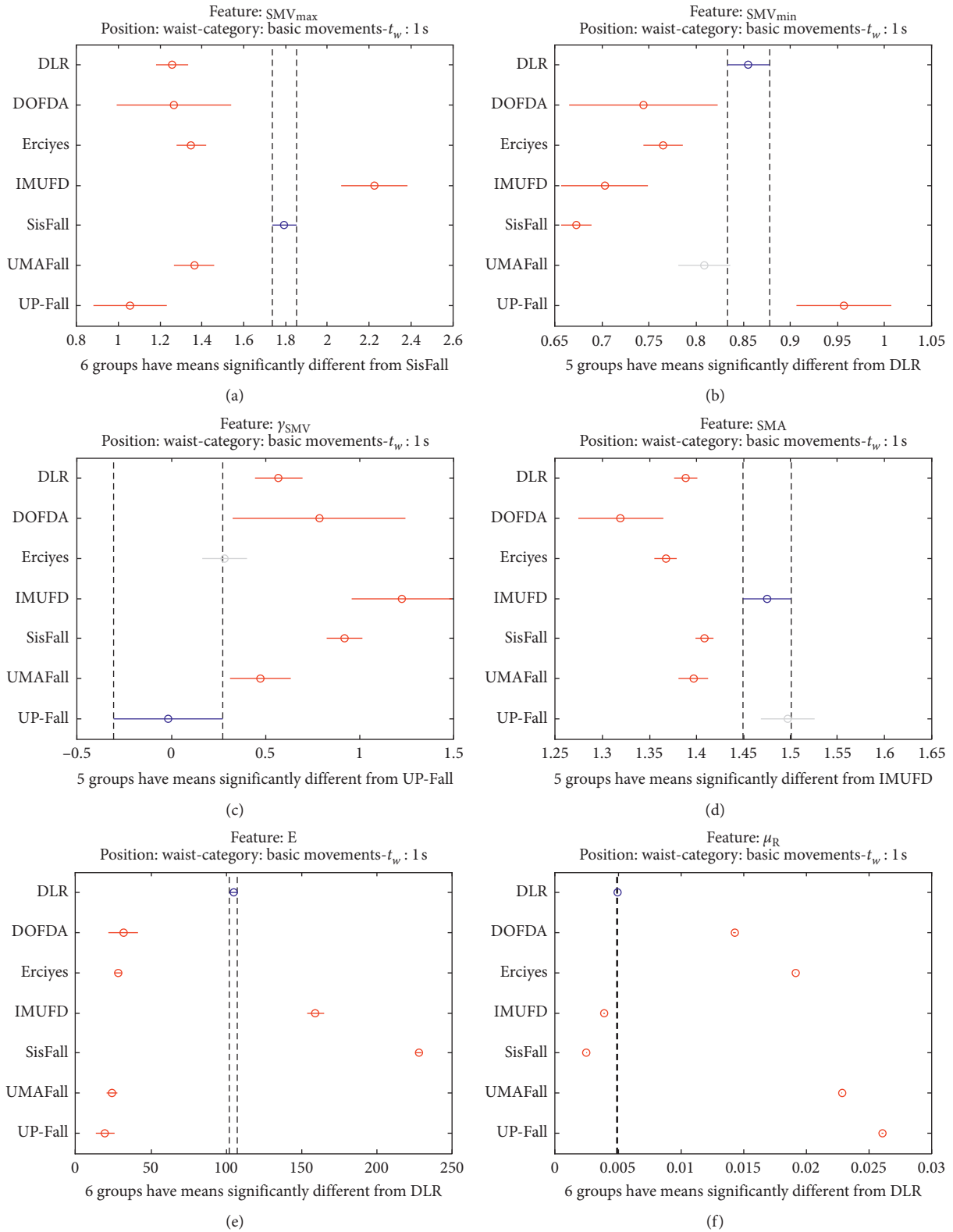


FIGURE 18: Multiple comparison test of the means of the statistical characteristics of the basic movements in the datasets: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ . Basic movements.



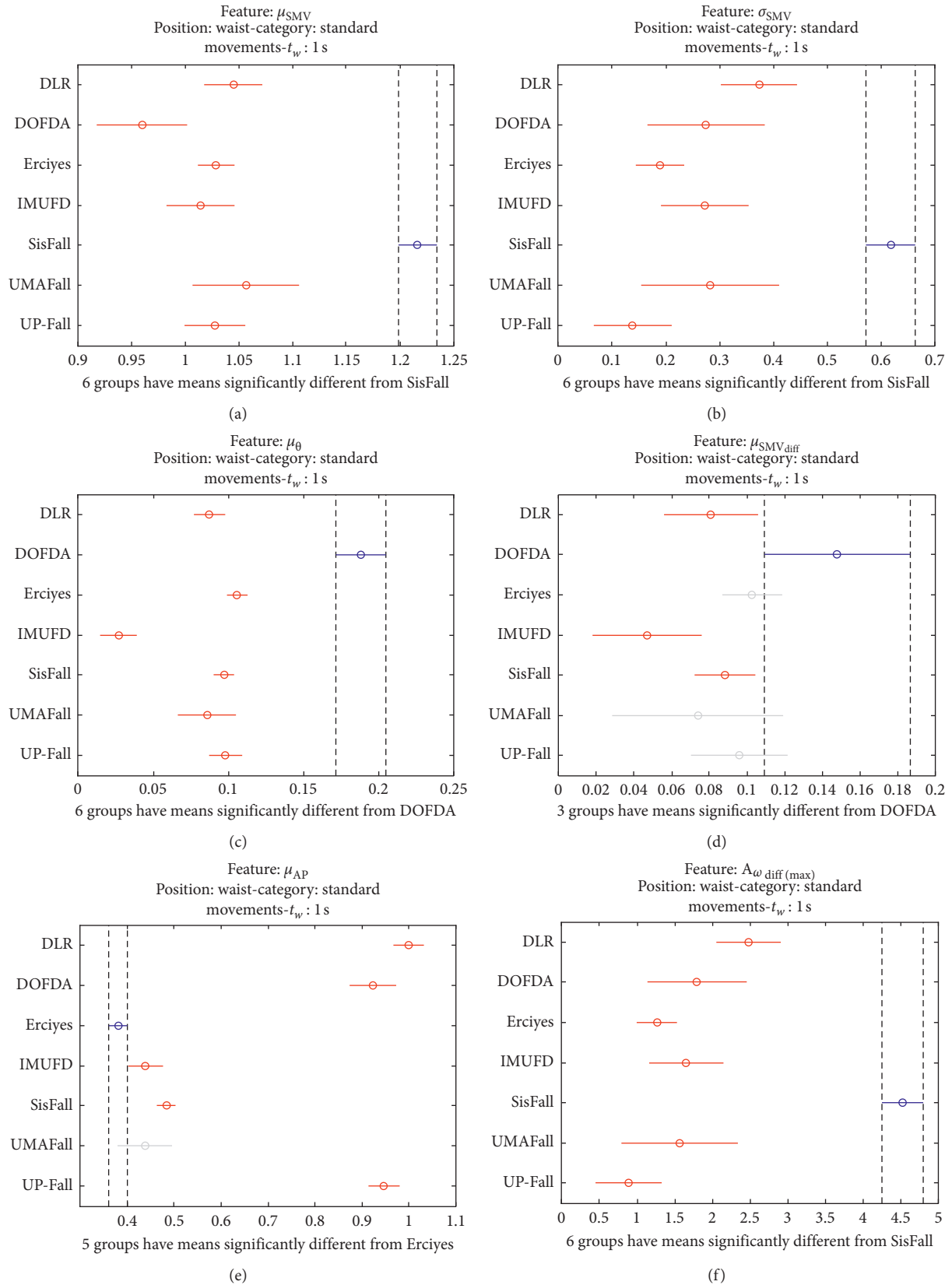


FIGURE 19: Multiple comparison test of the means of the statistical characteristics of the standard movements in the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV_{diff}}$ . (e)  $\mu_{Ap}$ . (f)  $A_{\omega_{diff(max)}}$ . Standard movements.

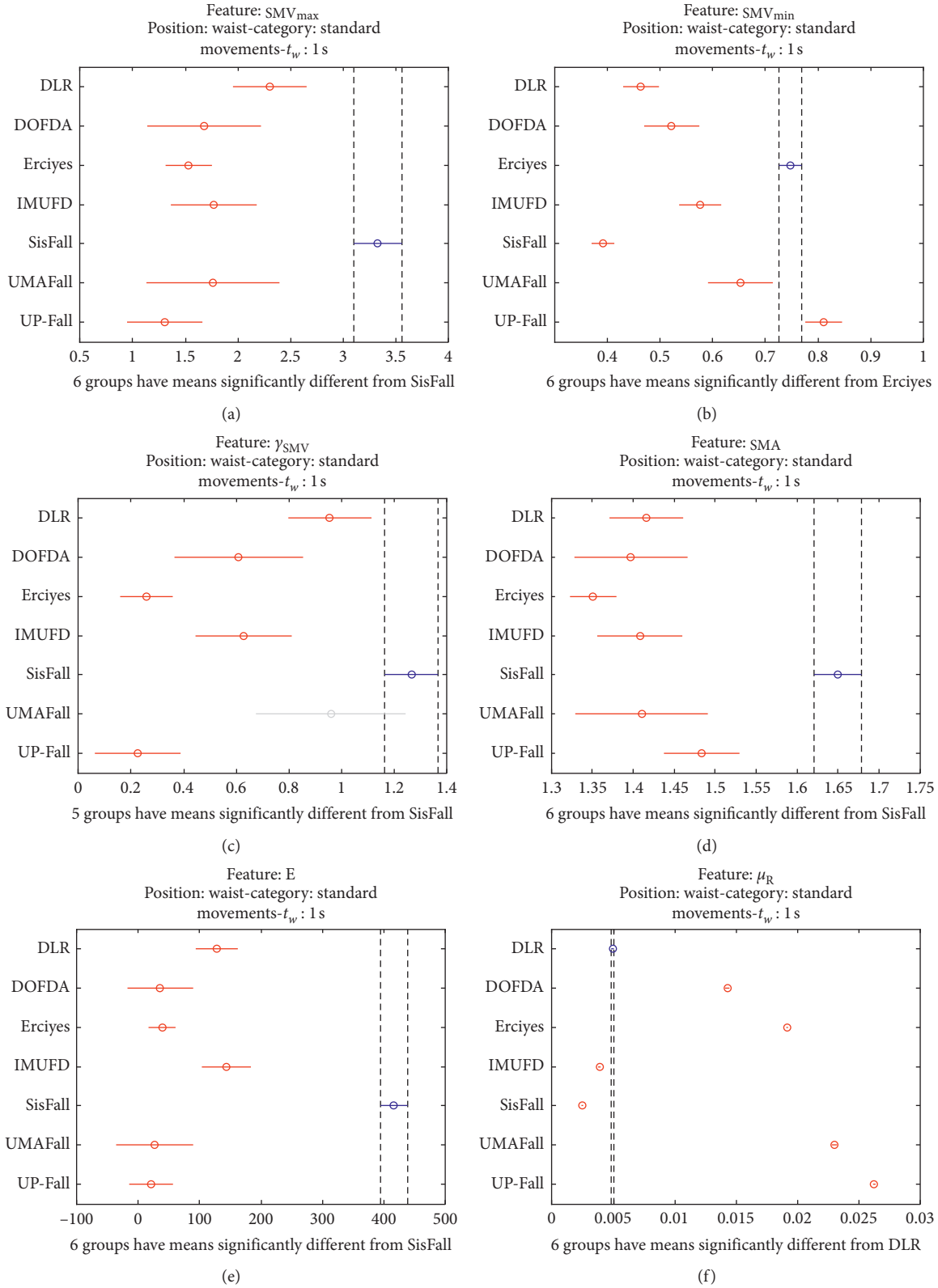


FIGURE 20: Multiple comparison test of the means of the statistical characteristics of the standard movements in the datasets: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ . Standard movements.

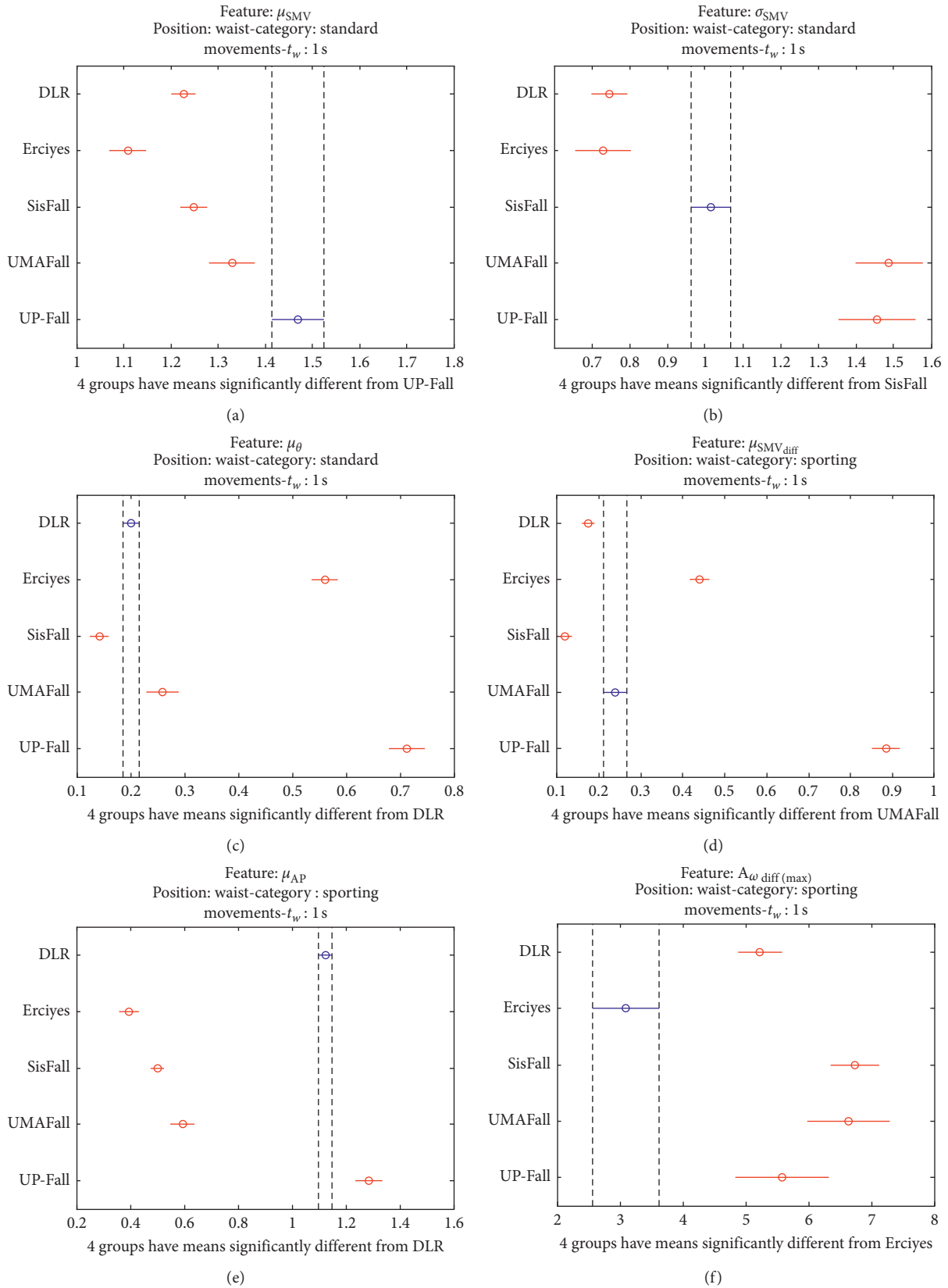


FIGURE 21: Multiple comparison test of the means of the statistical characteristics of the sporting movements in the datasets: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_\theta$ . (d)  $\mu_{SMV_{diff}}$ . (e)  $\mu_{AP}$ . (f)  $A_{\omega_{diff(max)}}$ . Sporting movements.

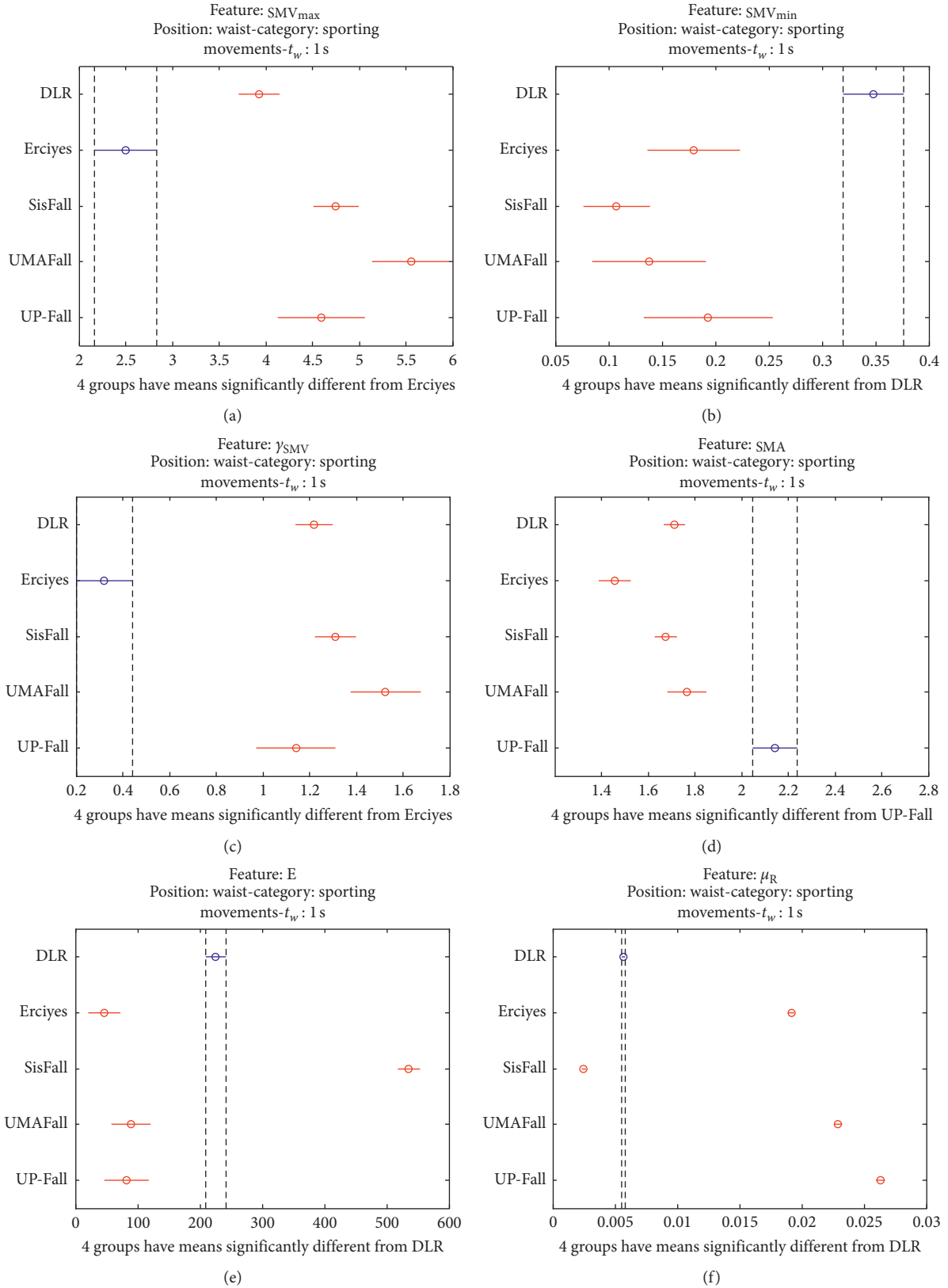


FIGURE 22: Multiple comparison test of the means of the statistical characteristics of the sporting movements in the datasets: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ . Sporting movements.

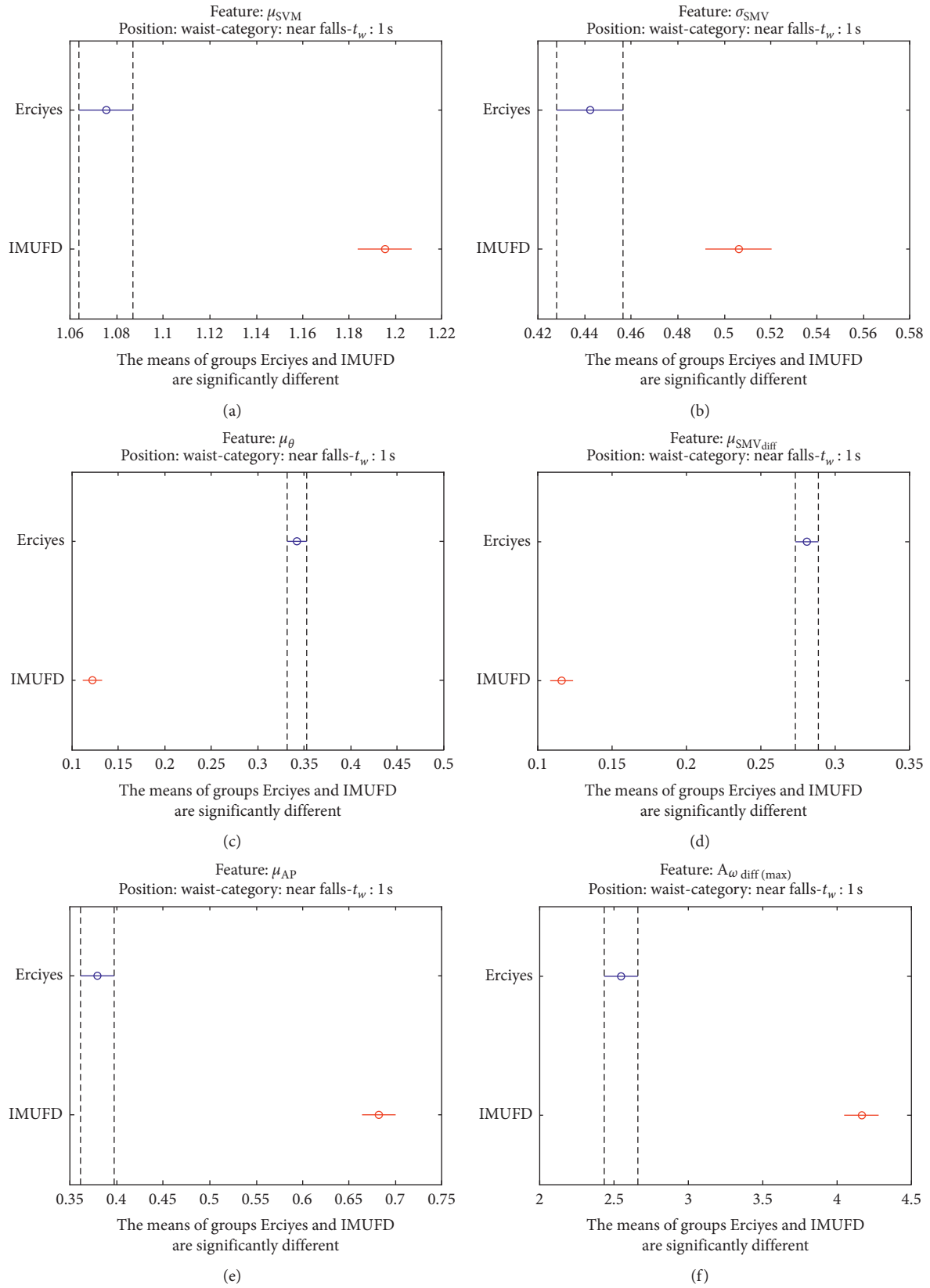


FIGURE 23: Multiple comparison test of the means of the statistical characteristics of the “near falls” in the two datasets that contain this type of movement: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_\theta$ . (d)  $\mu_{SMV_{diff}}$ . (e)  $\mu_{AP}$ . (f)  $A_{\omega_{diff(max)}}$ .

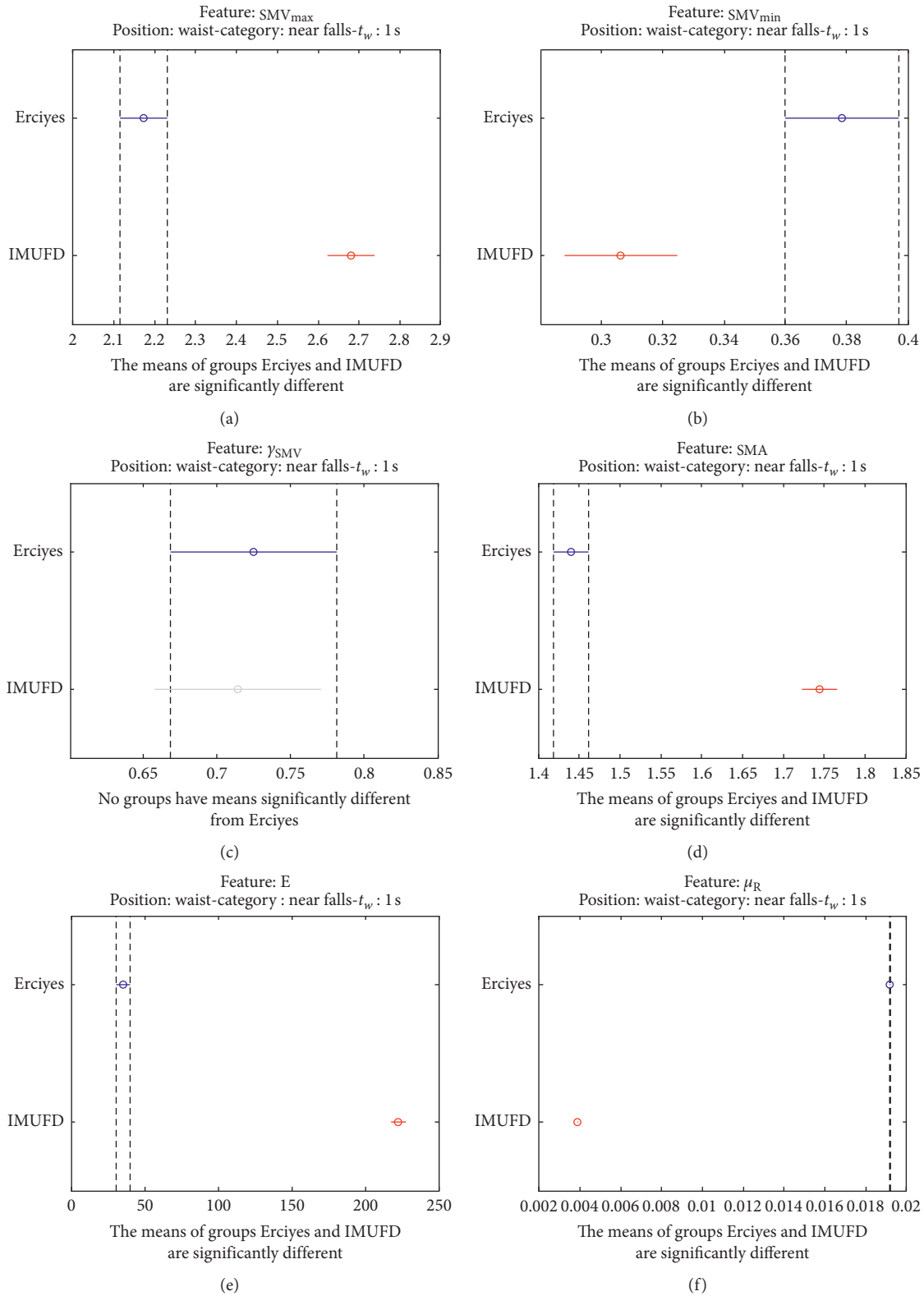


FIGURE 24: Multiple comparison test of the means of the statistical characteristics of the “near falls” in the two datasets that contain this type of movement: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ . Near falls.

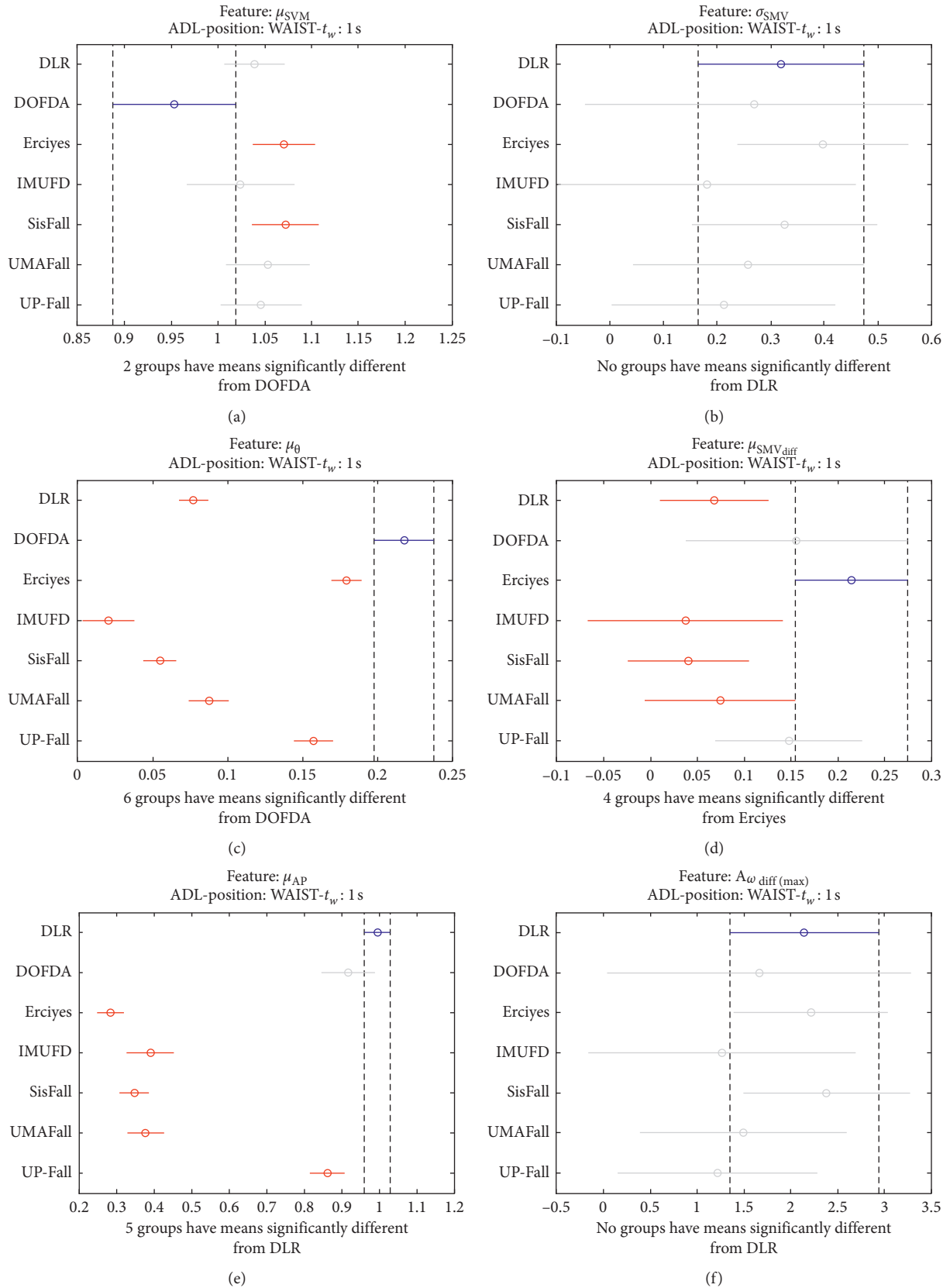


FIGURE 25: Multiple comparison test of the means of the statistical characteristics of the movements labelled as “walking” in the seven datasets that contain this type of ADL (sensor located on the waist): (a)  $\mu_{SVM}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SVM,diff}$ . (e)  $\mu_{AP}$ . (f)  $A\omega_{diff(max)}$ .

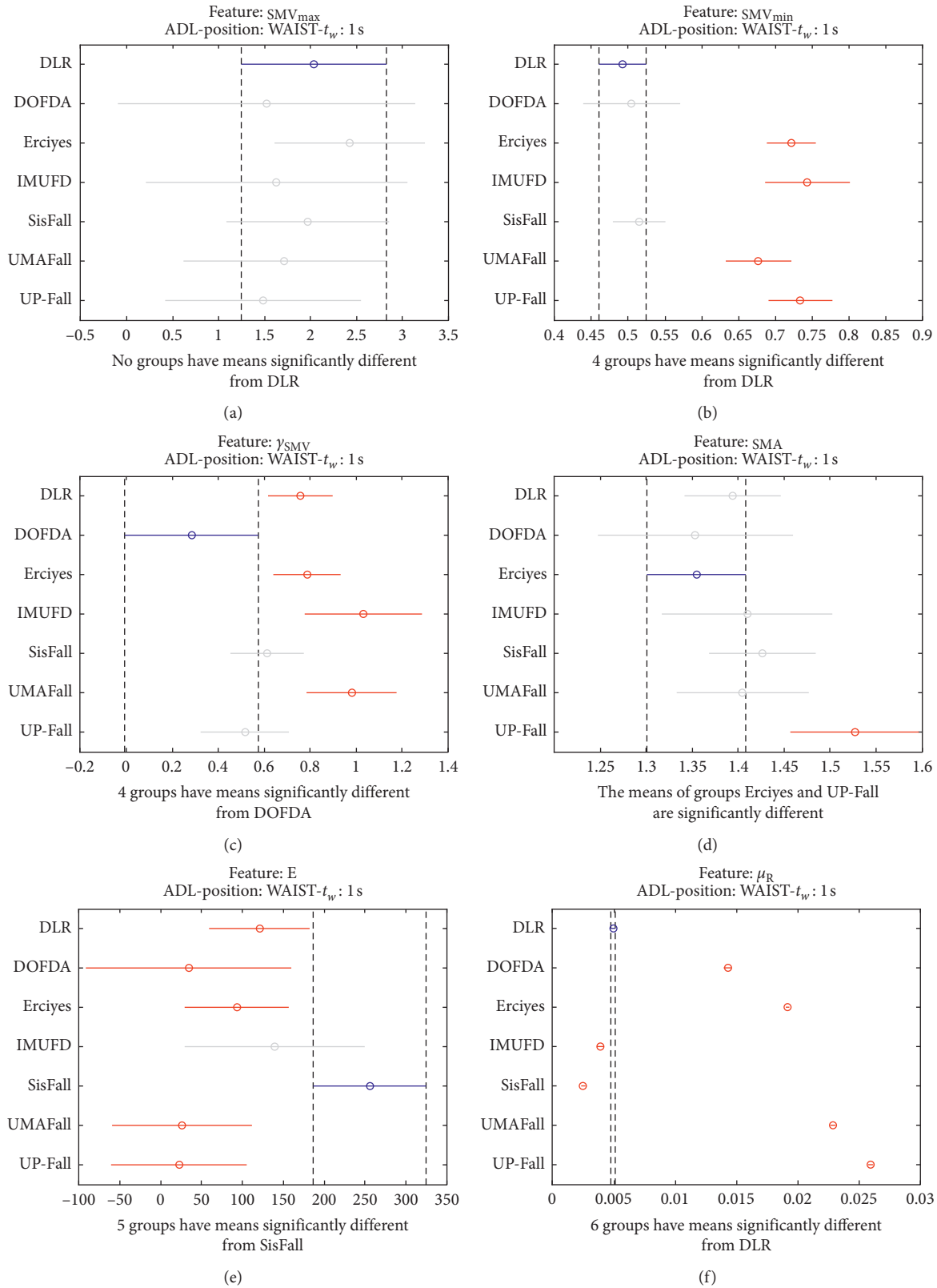


FIGURE 26: Multiple comparison test of the means of the statistical characteristics of the movements labelled as “walking” in the seven datasets that contain this type of ADL (sensor located on the waist): (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ .



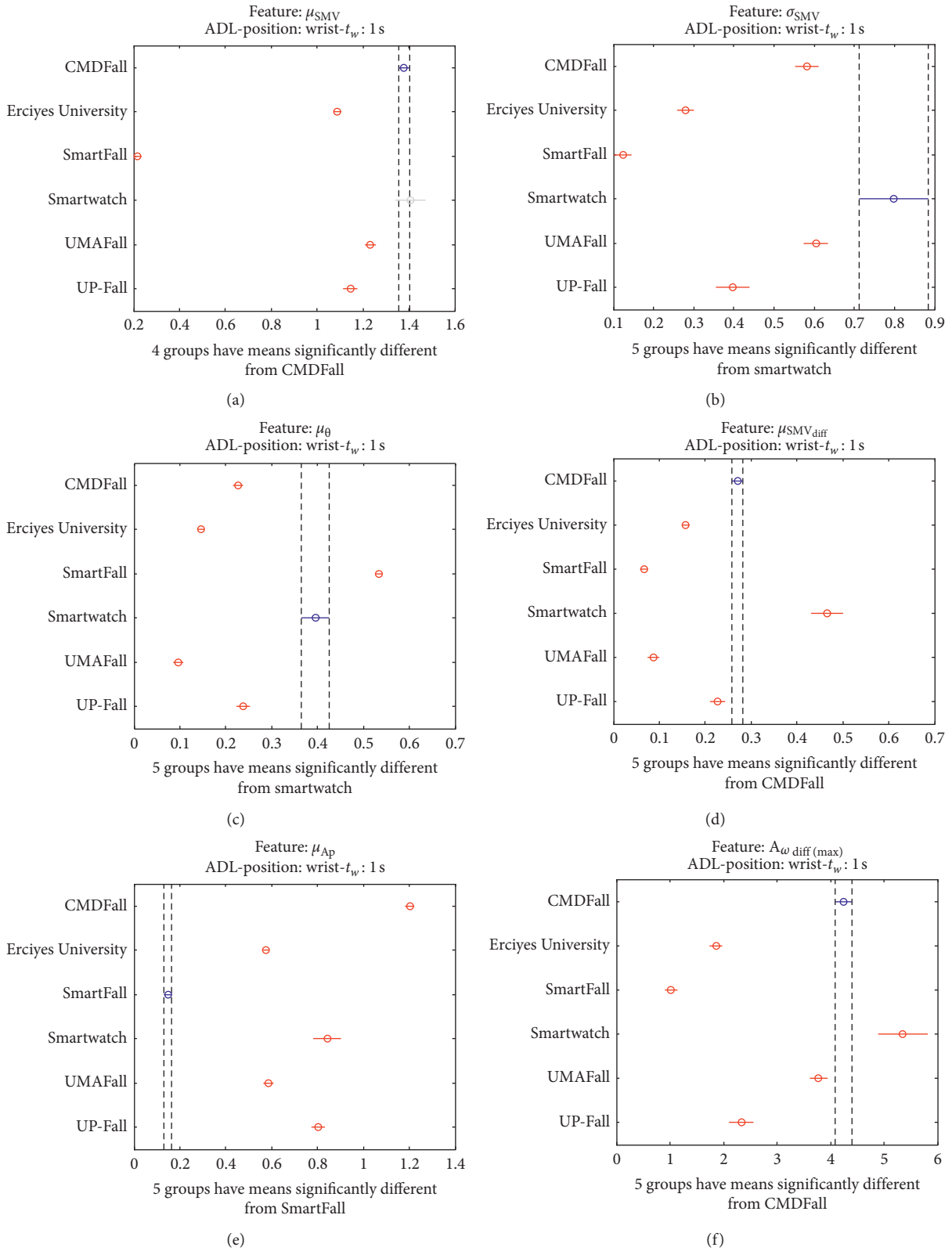


FIGURE 27: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_\theta$ . (d)  $\mu_{SMV,diff}$ . (e)  $\mu_{Ap}$ . (f)  $A_{\omega,diff(max)}$ .

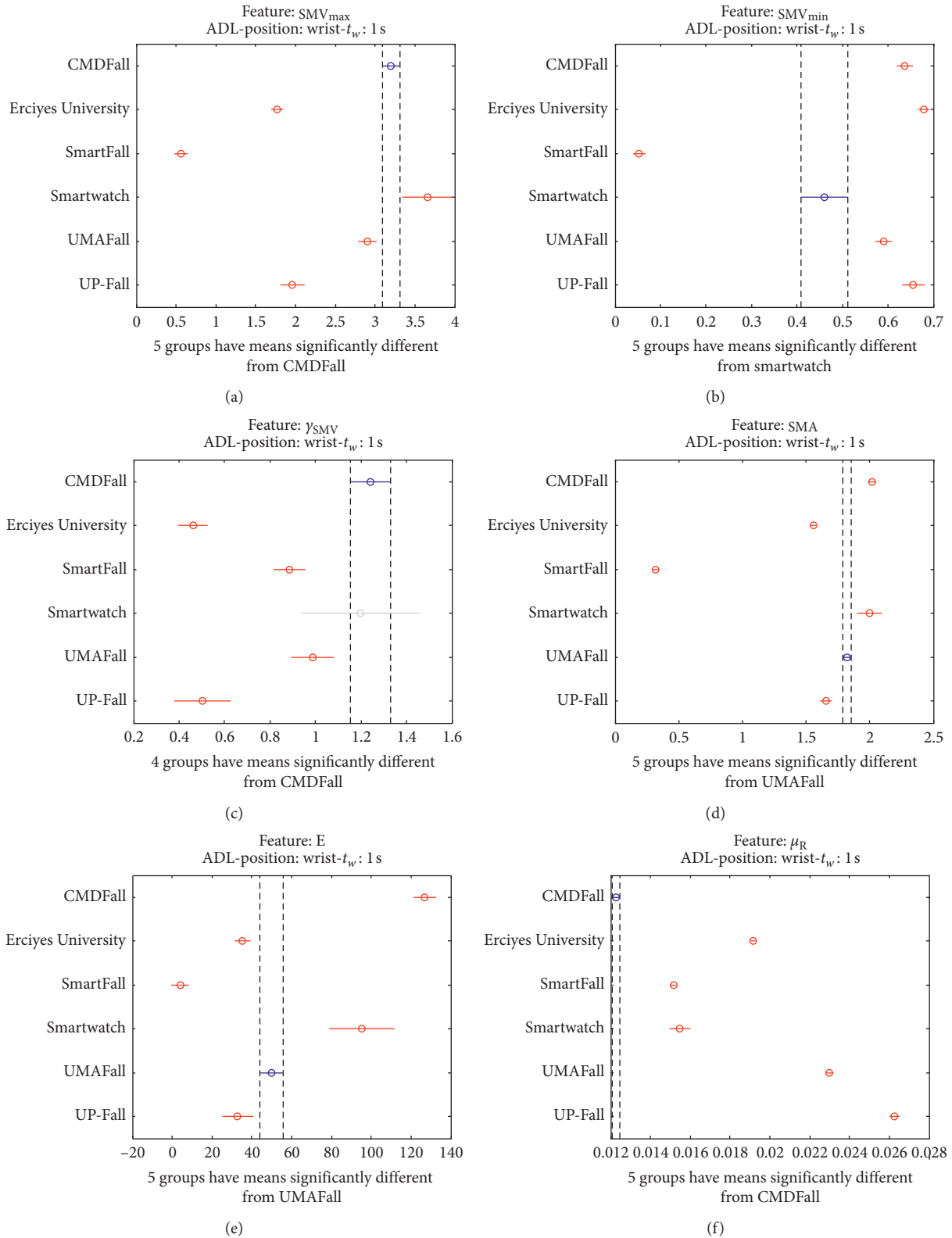


FIGURE 28: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ .

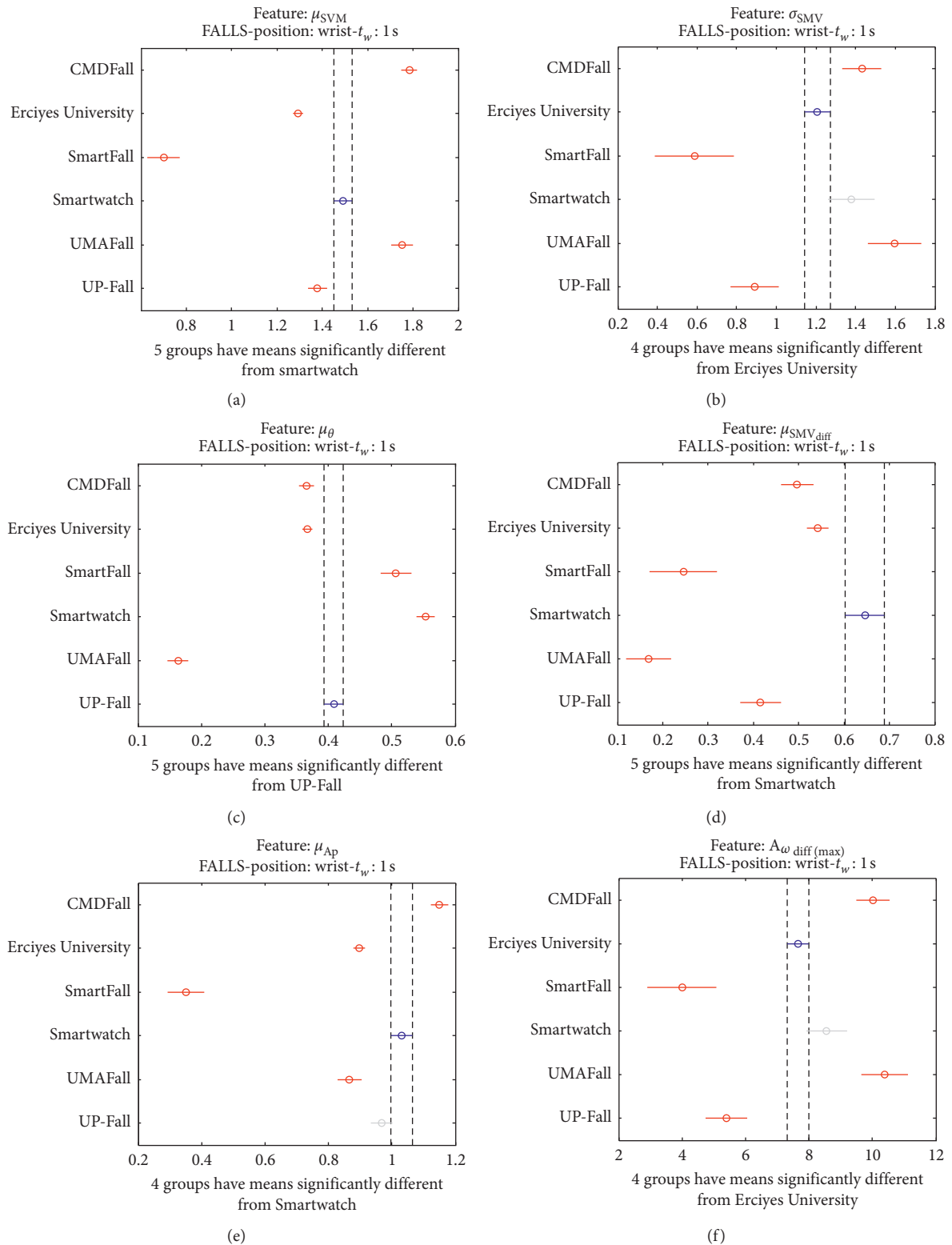


FIGURE 29: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and fall movements: (a)  $\mu_{SMV}$ . (b)  $\sigma_{SMV}$ . (c)  $\mu_{\theta}$ . (d)  $\mu_{SMV_{diff}}$ . (e)  $\mu_{Ap}$ . (f)  $A_{\omega_{diff(max)}}$ .

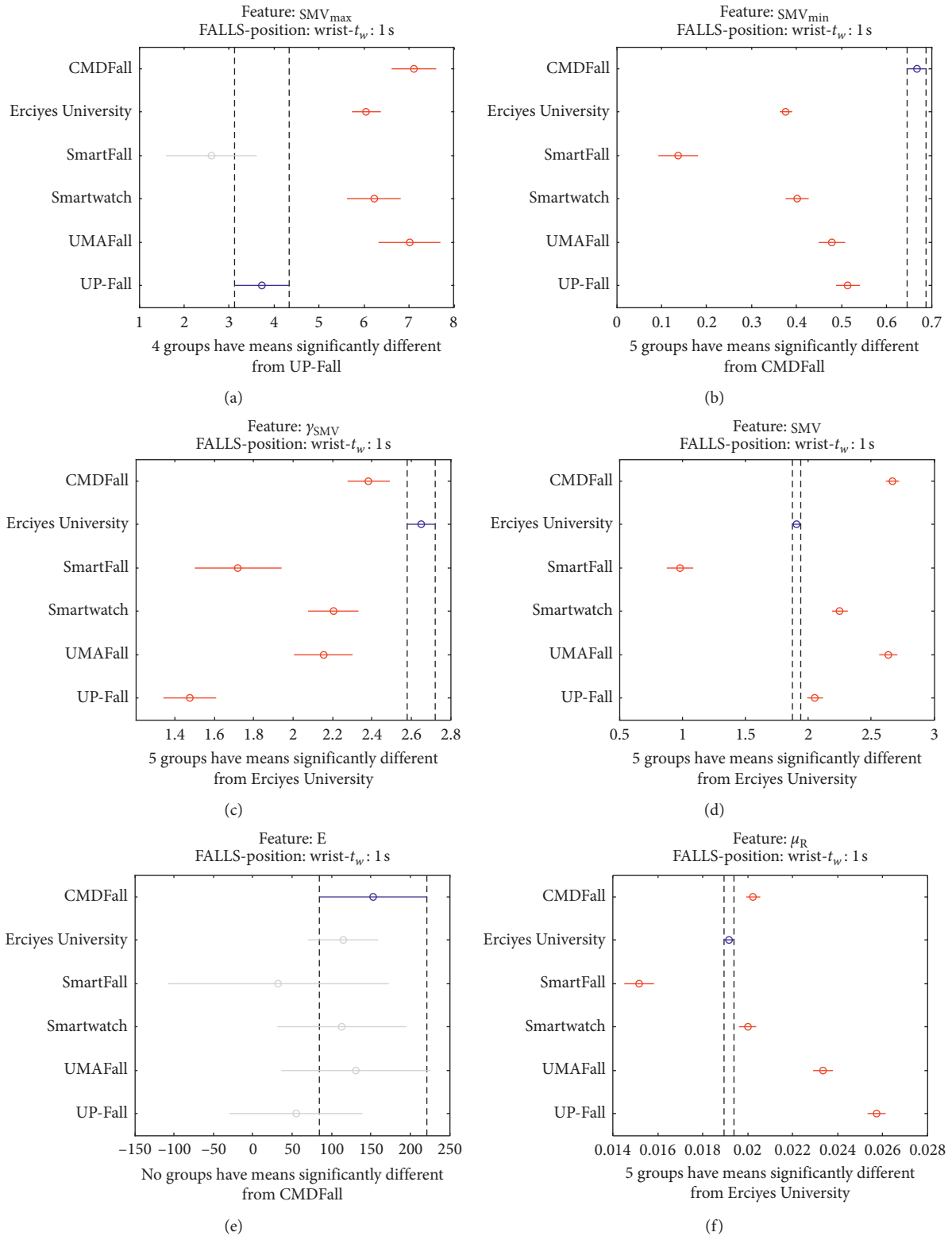


FIGURE 30: Multiple comparison test of the means of the statistical characteristics of the datasets for the measurements on the wrist and ADL movements: (a)  $SMV_{max}$ . (b)  $SMV_{min}$ . (c)  $\gamma_{SMV}$ . (d) SMA. (e) E. (f)  $\mu_R$ .

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 pre-determined 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

This work was supported by FEDER Funds (under grant UMA18-FEDERJA-022) and Universidad de Málaga, Campus de Excelencia Internacional Andalucía Tech.

## 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," *International Journal of Gerontology*, vol. 7, no. 2, pp. 80–84, 2013.
- [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 Ada-boost-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. Debar, 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] MEBIOMECE (Universidad Politécnica de Valencia, "Fall detection testing dataset," 2020.

- [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. Schulhaus, 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. Padiaditis, E. G. Spanakis, and M. Tsiknakis, "The MobiFall dataset: an initial evaluation of fall detection algorithms using smartphones," in *Proceedings 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. Padiaditis, "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. Napolitano, "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.



## Research Article

# Feature Selections Using Minimal Redundancy Maximal Relevance Algorithm for Human Activity Recognition in Smart Home Environments

Hongqing Fang , Pei Tang, and Hao Si

College of Energy and Electrical Engineering, Hohai University, Nanjing, Jiangsu 211100, China

Correspondence should be addressed to Hongqing Fang; fanghongqing@gmail.com

Received 13 September 2020; Revised 17 October 2020; Accepted 9 November 2020; Published 29 November 2020

Academic Editor: Ivan Miguel Pires

Copyright © 2020 Hongqing Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, maximal relevance measure and minimal redundancy maximal relevance (mRMR) algorithm (under D-R and D/R criteria) have been applied to select features and to compose different features subsets based on observed motion sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been evaluated and the activity recognition accuracy rates have been compared with two probabilistic algorithms: naïve Bayes (NB) classifier and hidden Markov model (HMM). The experimental results show that not all features are beneficial to human activity recognition and different features subsets yield different human activity recognition accuracy rates. Furthermore, even the same features subset has different effect on human activity recognition accuracy rate for different activity classifiers. It is significant for researchers performing human activity recognition to consider both relevance between features and activities and redundancy among features. Generally, both maximal relevance measure and mRMR algorithm are feasible for feature selection and positive to activity recognition.

## 1. Introduction

The aging of population and the increasing number of the elderly who chooses to live on their own [1–3] is an indisputable social reality. To implement the goals, smart home technology can play an important role to detect and analyze health events [4] and to provide corresponding medical assistant and caregiver for frail elderly and disabled people who are unable to live independently for a long period of time and in their home far away from hospital or their families, e.g., to remind them of time to take medicine, to see the doctor, to assist them in cutting off the water, turning off the oven, etc.

Actually, accurate assessment of human Activities of Daily Livings (ADLs) is the prerequisite for smart home to yield the correct service, whether it is for the elders or for the severe disabilities of health monitoring, or provide them with other relevant helps. Once the dangerous behavior is

detected, smart home itself can cope with it and eliminate as much of the inhabitant's risk as possible. Therefore, the accurate recognition of human activity in smart home is of great significance and gives a pattern for the realization of healthcare for solitary elderly or disabled as well, which is the most important process in incorporating ambient intelligence into smart environments [5–8].

Recently, human activity discovery and recognition has gained a lot of interest due to its enormous potential in context aware computing systems, including smart home environments. The primary objective of human activity recognition in smart home environment is to find the interesting patterns of behavior from gathered sensor data and to recognize such patterns. Currently, one of the primary challenges of human activity recognition is the choosing of machine learning algorithms which perform better in the same sequence of sensor data collected by smart home environment during the activity. In the last years, several

intelligent algorithms applied for human activity recognition in smart home have been reported. Singla, Crandall, and Cook et al. described the applications of some probability-based algorithms that include naïve Bayes (NB) classifier, Markov model (MM), and hidden Markov model (HMM) [9, 10] to train the partially labeled motion sensor events data to obtain the values of the prior parameters and then to validate the performances of the algorithms by testing the rest of labeled sensor data. Liu et al. presented a Bayesian network-based probabilistic generative framework based on Allen's temporal relations over primitive events to characterize the structural variabilities of complex activities [11]. Gayathri et al. proposed a statistical relational learning approach which augments ontology based activity recognition with probabilistic reasoning through Markov Logic Network (MLN) [12]. Kim et al. proposed a discriminative and generative probabilistic model, conditional random field (CRF), as a more flexible alternative to HMM [13]. Zhu et al. presented a two-layer CRF model to represent the action segments and activities in a hierarchical manner, which allows the integration of both motion and various context features at different levels and automatically learns the statistics that capture the patterns of the features [14]. Chen et al. introduced a knowledge-driven approach to continuous activity recognition based on multisensor data streams in smart homes [15]. Fahim et al. introduced a novel Evolutionary Ensembles Model (EEM) that values both minor and major activities by processing each of them independently, which is based on a Genetic Algorithm (GA) to handle the nondeterministic nature of activities [16]. Fleury et al. proposed support vector machine- (SVM-)based multimodal classification of ADLs in health smart homes [17]. Wen and Wang combined Latent Dirichlet Allocation (LDA) and AdaBoost to jointly train a general activity recognition model with partially labeled data [18]. Hong et al. composed a hybrid model of Bayesian networks and support vector machines to accurately recognize human activity [19].

Besides the suitable choosing of machine learning algorithms, another key point for human activity recognition in smart home is to select valid features from sensor events datasets collected in smart home environment. Usually, sensor events datasets include a large number of observed sensor events generated by various activities and any activity annotated in the dataset has various features, even redundant and irrelevant features [20]. However, these features are selected in one method in all tests, almost, and the influences of these features on the performance of classifiers are seldom addressed. Actually, the features which are irrelevant to activity recognition and redundant between initial features need to be removed prior to activity recognition. Furthermore, feature selection means to select the features subset which is the most favorable for activity recognition and compressing of data successfully.

Recently, minimal redundancy maximal relevance (mRMR) feature selection algorithm has been widely applied in many researching fields, which aims to achieve the best classification performance by reducing redundancy among the selected features and maximizing their relevance to the

target class. Mohamed et al. implemented the mRMR filter and a metaheuristic approach as a feature selection process for drug response microarray classification [21]. Che et al. presented a novel mutual information feature selection method based on the normalization of the maximum relevance and minimum common redundancy for nonlinear classification or regression problems [22]. Xu et al. proposed a new distributed monitoring scheme which integrates mRMR, Bayesian inference, and principal component analysis for plant-wide processes [23]. Li et al. provided a granular feature selection method with an mRMR criterion based on mutual information (MI) for multilabel learning [24]. Escalona-Vargas et al. proposed a method that uses mRMR as criteria to automatically select references for the frequency-dependent subtraction method to attenuate maternal and fetal magnetocardiograms of fetal magnetoencephalography recordings [25]. Mallik et al. developed a new framework of identifying statistically significant epigenetic biomarkers using mRMR criterion-based feature (gene) selection for multiomics dataset [26]. Tiwari calculated the weighted distance to improve the prediction performance of G-protein coupled receptors families and their subfamilies by using sequence derived properties, and the feature selection method based on fusion of mRMR and other supervised filter were provided [27]. Chen and Yan developed an optimized multilayer feedforward network by using mRMR-partial mutual information clustering integrated with least square regression to construct a soft sensor for controlling naphtha dry point [28]. Wang et al. presented a multiobjective evolutionary algorithm which employs the Pareto optimality to evaluate candidate feature subsets and finds compact feature subsets with both the maximal relevance and the minimal redundancy [29]. Morgado and Silveira proposed a multivariate procedure capable of selecting nonredundant subsets of features significantly faster than other similar methods to the diagnosis of Alzheimer's disease and related disorders which is inspired in mRMR algorithm [30]. Kamandar and Ghassemian used a modified mRMR as a criterion for feature extraction for hyperspectral images classification based on information theoretic learning [31]. Kandaswamy et al. extracted the best features using mRMR feature selection algorithm and used the random forest algorithm to predict extracellular matrix proteins [32]. Jin et al. proposed a novel method for health monitoring and anomaly detection for cooling fans in electronic products based on Mahalanobis distance with mRMR features [33]. Unler et al. presented a hybrid filter-wrapper feature subset selection algorithm based on particle swarm optimization for support vector machine classification. The filter model is based on the mutual information and is a composite measure of feature relevance and redundancy with respect to the feature subset selected [34]. Zdravevski et al. proposed a generic feature engineering method for selecting robust features from a variety of sensors, which is from the originally recorded time series and some newly generated time series (i.e., magnitudes, first derivatives, delta series, and fast Fourier transformation-(FFT-) based series); a variety of time and frequency domain features are extracted [35, 36].

Inspired by the ideas mentioned above, in this paper, maximal relevance measure and mRMR feature selection algorithm (under D-R and D/R criteria) have been applied to select features and to compose different features subsets based on the observed motion sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been evaluated, and the activity recognition accuracy rates have been compared with two probabilistic algorithms: NB classifier and HMM.

The rest of the paper is organized as follows: Section 2 describes the smart apartment testbed, the data collection, and 13 features of observed sensor events for human activity recognition. Section 3 presents the concepts of information entropy, mutual information (MI), and minimal redundancy maximal relevance (mRMR) feature selection algorithm. Section 4 gives the training and testing activities, the features subsets selected through analyzing of maximal relevance measure, and mRMR feature selection algorithm. Finally, we present the comparison results of activity recognition accuracy rates with the selected features subsets and the performance measures of NB classifier as well as HMM. Section 5 summarizes the main contributions.

## 2. Smart Apartment Testbed and Data Collection for Human Activity Recognition

**2.1. Smart Apartment Testbed and Data Collection.** The smart apartment testbed for this research is located on Washington State University campus and is maintained as part of the Center for Advanced Studies in Adaptive Systems (CASAS) smart home project, which includes three bedrooms, one bathroom, a kitchen, and a living/dining room. The smart apartment is equipped with motion sensors distributed approximately 1 meter apart throughout the space on the ceilings, as shown in Figure 1. In addition, other sensors installed provide ambient temperature readings and custom-built analog sensors provide readings for hot water, cold water, and stove burner use. Sensor data is captured using a sensor network that was designed in-house and is stored in a Structured-Query-Language (SQL) database. After collecting data from the smart apartment testbed, the sensor events are annotated for ADLs, which are used for training and testing the activity recognition algorithms [3, 4, 7, 9, 10].

The data gathered by CASAS smart home is represented by the following parameters, which specify the number of features that are used to describe the observed sensor events. The generalized syntax of the dataset (Cairo Dataset, 2009) is

*Date Time SensorID SensorValue <label>*

An example of the dataset of Night\_wandering activity is:

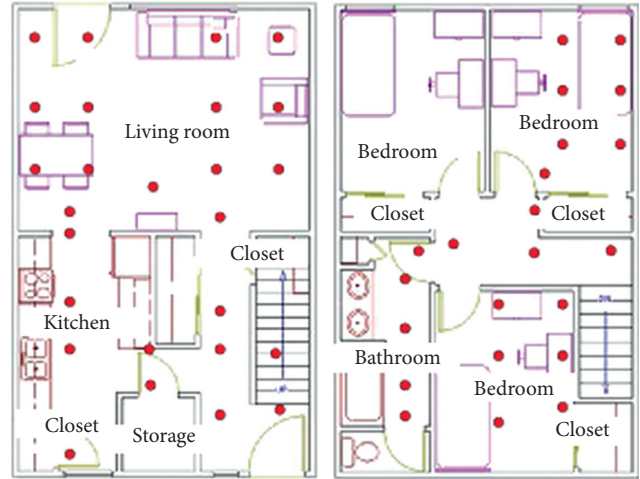
{

2009-06-10 03:20:59.08 M006 ON Night\_wandering begin

2009-06-10 03:25:19.05 M012 ON

2009-06-10 03:25:19.08 M011 ON

2009-06-10 03:25:24.05 M011 OFF



● Motion sensor

FIGURE 1: The smart apartment testbed and sensors in the apartment to monitor motion.

```
2009-06-10 03:25:24.07 M012 OFF Night_wandering
end
}
```

This example shows that the observed sensor events correspond to the Night\_wandering activity with concrete Date, Time, Sensor ID, Sensor Value, and activity label parameters.

**2.2. Features of Observed Sensor Events.** Considering the actual situation, each activity has 13 features of the observed sensor events:

- (1) The means of logical values of Sensor IDs of each activity's sensor events is  $f_1$ .

Considering that the place where each activity happens is relatively stable, therefore, selecting the average of Sensor IDs means the focus area where the activity occurs. The equation is

$$\bar{S}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} S_{ik}, \quad (1)$$

where  $\bar{S}_i$  is the means of Sensor IDs of activity  $i$ ,  $n_i$  is the number of sensors, and  $S_{ik}$  is the  $k^{\text{th}}$  Sensor ID of activity  $i$ .

- (2) The variance of all Sensor IDs triggered by the current activity,  $f_2$ , is

$$S_i^2 = \frac{1}{n_i} \sum_{k=1}^{n_i} (S_{ik} - \bar{S}_i)^2. \quad (2)$$

- (3) Day of the week, which is converted into a value in the range of 0 to 6, is  $f_3$ .

- (4) Previous activity, which represents the activity that occurred before the current activity, is  $f_4$ .
- (5) Activity length, which represents the length of the current activity measured in a number of sensor events, is  $f_5$ .
- (6) The logical value of the first Sensor ID triggered by the current activity is  $f_6$ .
- (7) The logical value of the last Sensor ID triggered by the current activity is  $f_7$ .
- (8) The duration of the current activity, which indicates the time interval between the last sensor and the first sensor triggered by the current activity, is  $f_8$ .
- (9) The beginning time of the current activity is  $f_9$ .
- (10) The ending time of the current activity is  $f_{10}$ .
- (11) Next activity, which represents the activity that occurred after the current activity, is  $f_{11}$ .
- (12) The mode value of the Sensor IDs triggered by the current activity is  $f_{12}$ .
- (13) The median value of the Sensor IDs triggered by the current activity is  $f_{13}$ .

Usually, the optimal features subset contains the least number of dimensions that contribute to higher recognition accuracy rate. Therefore, it is necessary to remove the remaining and unimportant features.

### 3. Minimal Redundancy Maximal Relevance Feature Selection Algorithm

**3.1. Entropy.** Information entropy is a more abstract mathematical concept which can be understood as the probability of the emergence of some specific information. Generally, the higher of the probability of a kind of information indicates that it is spread more widely or more highly cited. Therefore, information entropy is able to represent the value of information. The information source is the source of the message and the message sequence. For example, the simplest discrete information source is  $X = \{x_1, x_2, \dots, x_n\}$ , and  $p(x_i)$  is the probability of a given  $x_i$ ; then, the entropy of  $X$  can be defined as

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i). \quad (3)$$

#### 3.2. Conditional Entropy, Joint Entropy, and Mutual Information

**3.2.1. Conditional Entropy.** In information theory, the conditional entropy quantifies the amount of information needed to describe the outcome of a random variable  $X$  given that the value of another random variable  $Y$  is known. If  $X$  and  $Y$  are dependent on each other, then, in the condition of  $\{Y = y\}$ , the conditional entropy of  $X$  is

$$H(X|Y = y) = \sum_x p(x|y) \log \frac{1}{p(x|y)}. \quad (4)$$

If  $Y$  is given, then the conditional entropy of  $X$  is

$$\begin{aligned} H(X|Y) &= \sum_y p(y) H(X|Y = y), \\ &= \sum_{x,y} p(x,y) \log \frac{1}{p(x|y)}, \end{aligned} \quad (5)$$

where  $p(x, y)$  is the joint probability of  $x$  and  $y$ .

**3.2.2. Joint Entropy.** Joint entropy is a measure of the uncertainty associated with a set of random variables. Supposing two random variables  $X$  and  $Y$  and each of them is given a limited value, then, the joint entropy is

$$H(X, Y) = - \sum_{x,y} p(x, y) \log p(x, y). \quad (6)$$

The joint entropy is a measure of the correlation of  $X$  and  $Y$ . If  $X$  and  $Y$  are independent, then, the joint entropy is

$$H(X, Y) = H(X) + H(Y). \quad (7)$$

**3.2.3. Mutual Information.** Mutual information (MI) is a quantity that measures the level of similarity as well as correlation of random variables [33, 37]. Supposing two random variables  $X$  and  $Y$  and  $Y$  contains some information of  $X$ , then MI between  $X$  and  $Y$  can be defined as

$$I(X; Y) = \sum_{x,y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}. \quad (8)$$

MI is typically defined as the measure of the mutual dependence of two random variables. A larger value of MI means a closer relationship between the two random variables which have larger correlation. If the value of MI is zero, it means that the two random variables are uncorrelated and independent. Therefore, in this paper, MI can be used to measure the similarity among features and the correlation between feature and activity.

**3.3. Minimal Redundancy Maximal Relevance (mRMR) Algorithm.** Although MI is widely applied in the feature selection fields, there still exist some deficiencies. Most of the feature selection algorithms only consider the relationship between features and classification categories but ignore the mutual influence among features. Instead, mRMR feature selection algorithm applied in this paper considers not only the amount of information provided by these features for categorical attributes but also the influence of interaction among features on classification [37].

MI can weigh the quantity of information between feature variables  $X$  and  $Y$ ; furthermore, it can measure how much information quantity that  $X$  can provide to  $Y$  to

classify activity as well. Therefore, MI can not only show the ability of each feature of identifying the activities, but also measure whether there is redundancy among features. According to the traits of MI, two different criteria can be extended to evaluate the features: redundancy and relevance.

**3.3.1. Redundancy Measure.** Redundancy measure utilizes the quantity of MI between features. If the value of MI is large, it means that there is a large amount of information duplication between the two features; i.e., there are redundancies between the two features. A lower value of redundancy measure indicates a better feature selection criterion. Utilizing redundancy measure is to find the feature which has the minimal value of MI among all features.

According to the idea that the smaller the value of redundancy of information between features is, the more beneficial it is to activity classification, which can be expressed by minimizing the MI among features, the minimal redundancy condition is

$$\min R(S), R = \frac{1}{|s|^2} \sum_{x_i, x_j \in S} I(x_i; x_j), \quad (9)$$

where  $|s|$  is the number of features in features subset  $S$  and  $I(x_i; x_j)$  is MI between feature  $i$  and  $j$ .

**3.3.2. Relevance Measure.** Relevance measure utilizes the value of MI between the feature and the target activity. If the value of MI is small, it indicates that there is a weak correlation between the feature and the target activity. On the contrary, the larger value of MI means that the feature has greater amount of information to classify the activity. Therefore, it is necessary to select the maximum value of MI between the features and the target activity, the maximal relevance criterion, which can be expressed as

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c), \quad (10)$$

where  $c$  is the target activity and  $I(x_i; c)$  is the MI between the feature  $i$  and the target activity  $c$ .

Feature selection based on the maximum value of relevance measure is to compose the optimal features subset by selecting MI to target activity.

**3.3.3. Algorithm Designing.** mRMR feature selection algorithm is based on the relevance measure and redundancy measure described above. It combines the relevance between the features and the target activity as well as the redundancy among the features [33, 37]. According to (9) and (10), the influences of relevance measure and redundancy measure have been taken consideration in feature selection, comprehensively. mRMR feature selection algorithm has two evaluation criteria, which optimizes  $D$  and  $R$ , simultaneously, as

$$\max \Phi_1(S, c), \Phi_1 = D - R, \quad (11)$$

$$\text{or } \max \Phi_2(S, c), \Phi_2 = \frac{D}{R}. \quad (12)$$

Supposing that there is a features subset  $S_m$  which is composed of  $m$  features, the next step is to extract the optimal  $(m+1)^{\text{th}}$  feature from the features subset  $\{S - S_m\}$  according to (11) or (12), through

$$\max_{x_i \in S - S_m} \left[ I(x_i; c) - \left( \frac{1}{m} \right) \sum_{x_j \in S_m} I(x_i; x_j) \right], \quad (13)$$

or

$$\max_{x_i \in S - S_m} \left[ \frac{I(x_i; c)}{(1/m) \sum_{x_j \in S_m} I(x_i; x_j)} \right]. \quad (14)$$

The incremental procedures of mRMR feature selection algorithm are as follows:

Step 1. In the original features set  $\Omega$ , the optimal feature  $x_i$  can be selected by  $I(x_i; c)$  and then put into the optimal features subset  $S$ ;

Step 2. In the features subset  $\Omega_S = \Omega - S$ , the next optimal feature  $x_j$  is selected which satisfies (11) or (12);

Step 3. Repeat Step 2 to find the optimal features subset  $S$  which meets the size requirement finally.

## 4. Experimental Results

**4.1. Training and Testing Activities.** A total of 10 activities were performed in the CASAS smart apartment by 2 volunteers to provide physical training data for NB classifier and HMM. These activities include both basic and more complex ADLs that are found in clinical questionnaires. These activities are the following:

- (1) Bed\_to\_toilet (activity 0): transition happens between bed and toilet in the night time (30 instances)
- (2) Breakfast (activity 1): the resident has breakfast (48 instances)
- (3) Bed (activity 2): this is the activity of sleeping in bed (207 instances)
- (4) C\_work (activity 3): the resident works in the office space (46 instances)
- (5) Dinner (activity 4): the resident has dinner (42 instances)
- (6) Laundry (activity 5): the resident cleans clothes using the laundry machine (10 instances)
- (7) Leave\_home (activity 6): the resident leaves smart home (69 instances)
- (8) Lunch (activity 7): the resident has lunch (37 instances)
- (9) Night\_wandering (activity 8): the resident wanders during night time (67 instances)

- (10) R\_medicine (activity 9): the resident takes medicine (44 instances)

The data have been collected in the CASAS smart apartment testbed for 55 days, resulting in total of 600 instances of these activities and 647, 485 collected motion sensor events. The 3-fold cross validation is applied in the data for NB classifier and HMM under the same conditions to ensure that the experimental comparison is fair [38].

**4.2. Feature Selection with Maximal Relevance Measure.** Firstly, maximal relevance (MR) measure of the 13 features has been analyzed. Usually, MI is employed as the criterion of evaluation. Since the feature with larger value of MI has the closer relationship to the target activity, therefore, the values of MI between features and activities must be calculated. By analyzing relevance measure, the ranks of the 13 features according to the values of MI are listed in Table 1, and then the 13 features have been sorted in a descending order successively, composed of 13 features subsets, respectively [39].

It can be found that feature  $f_{10}$  has the largest value of MI, which means that  $f_{10}$  is the most important feature to activity recognition and has the closest relationship to the target activity. On the contrary,  $f_3$  has the smallest value of MI, which means that  $f_3$  is not discriminatory to activity recognition and has the weakest link to the target activity.

Then, the 13 features subsets with MR measure are the following:

- Features subset 1: ( $f_{10}$ )
- Features subset 2: ( $f_{10}, f_9$ )
- Features subset 3: ( $f_{10}, f_9, f_6$ )
- Features subset 4: ( $f_{10}, f_9, f_6, f_8$ )
- Features subset 5: ( $f_{10}, f_9, f_6, f_8, f_7$ )
- Features subset 6: ( $f_{10}, f_9, f_6, f_8, f_7, f_1$ )
- Features subset 7: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}$ )
- Features subset 8: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4$ )
- Features subset 9: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}$ )
- Features subset 10: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5$ )
- Features subset 11: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2$ )
- Features subset 12: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2, f_{13}$ )
- Features subset 13: ( $f_{10}, f_9, f_6, f_8, f_7, f_1, f_{12}, f_4, f_{11}, f_5, f_2, f_{13}, f_3$ )

The activity recognition accuracy rate can be divided into two categories: the individual activity recognition accuracy rate (IARAR) and the total activities recognition accuracy rate (TARAR). The definitions are

$$\text{IARAR} = \frac{\text{correct sample number of this activity}}{\text{the sample number of this activity}}, \quad (15)$$

$$\text{TARAR} = \frac{\text{total correct sample number}}{\text{total sample number}}.$$

The results of activity recognition accuracy rates with the 13 features subsets of NB classifier based on MR measures

are shown in Table 2, and the results of HMM are shown in Table 3, respectively.

From Tables 2 and 3, the 13 features impact on the classifiers' recognition accuracy rate can be found differently. Similarly, for NB classifier and HMM, TARARs vary from different features subsets. Generally, with the increasing of size of features subset, activity recognition accuracy rate increases. TARAR tends to be stable as the size of features subset becomes larger finally. However, the trend is not monotonic. Concretely, for NB classifier, TARAR does not improve with features subset 7 (88.7%) which introduces feature  $f_7$  in features subset 6 (88.8%). It means that not all the features are positive to activity recognition obviously and different combination of features has different effect on activity recognition accuracy rate.

It also can be found that the maximal value of TARAR of NB classifier is slightly higher than that of HMM. With features subset 12, the TARAR of NB classifier reaches the maximal value 90.3%, while, for HMM, the maximal value of TARAR is 88.0% with features subset 13.

Furthermore, it can be seen that the trends of IARARs of NB classifier and HMM are similar as well; i.e., with the increasing of the number of features in features subset, IARAR increases until a specific number. Again, the trends are not monotonic. For example, for activity 1, IARARs of NB classifier and HMM have risen along with the sizes of features subset. For NB classifier, the optimal value of IARAR is 100.0%, with features subset 8, and then drop down with the increasing of size of features subset. For HMM, IARAR reaches the optimal value of 93.8% with features subset 6 and then drop down with the increasing of size of features subset. Moreover, from IARARs of activity 1, it can be seen that the performances of NB classifier and HMM are different by introducing the same feature, such as feature  $f_4$ , compared with features subsets 7 (97.9%) and 8 (100.0%); NB generates a positive result with this feature; however, the recognition performance of HMM does not improve, from 89.6% decreasing to 83.3%. Even with same features subset, the recognition accuracy rates may be quite different from the performances of classifiers, which can be found much more from the recognition accuracy rates for each activity. For example, the performances of NB classifier and HMM differ greatly for activities 0 with features subset 1 (0.0% vs. 90.0%) and activity 2 with features subset 3 (88.4% vs. 39.1%).

Furthermore, from Table 2, it is indicated that the relatively better features subset is different for each activity; e.g., for activity 3, features subset 10 yields the best result, and the accuracy rate is 76.1%, while features subset 2 generates the best result of activity 2, and the accuracy rate is 100.0%, for NB classifier. With features subset 10, IARARs of activities 3, 4, 5, and 6 are better than or equal to those of the other features subsets; the proportion is 40% for all activities. However, with features subsets 12 and 13, the IARARs of activities 0 through 2, and 4, and 6 through 8 are better than or equal to those of features subset 10; the proportions are 70%. Therefore, the optimal results of TARARs (90.3%) are of features subsets 12 and 13. Relatively, features  $f_3$  and  $f_{13}$  are not discriminatory to activity recognition with NB classifier.

TABLE 1: The values of mutual information (MI) and the ranks of the 13 features.

Feature	1	2	3	4	5	6	7	8	9	10	11	12	13
MI	1.1879	0.6885	0.0472	1.1734	0.8628	1.4903	1.2003	1.3213	1.5084	1.5350	1.1675	1.1832	0.5304
Rank	6	11	13	8	10	3	5	4	2	1	9	7	12

TABLE 2: The recognition accuracy rates with all features subsets of NB classifier based on maximum relevance selection.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.433	0.433	0.300	0.300	0.567	0.600	0.567	0.600	0.600	0.633	0.667	0.667
1	0.792	0.729	0.917	0.938	0.917	0.979	0.979	1.000	0.979	0.979	0.979	0.979	0.979
2	0.700	0.676	0.884	0.918	0.937	0.947	0.952	0.957	0.952	0.957	0.957	0.961	0.966
3	0.000	0.022	0.630	0.587	0.696	0.717	0.717	0.739	0.739	0.761	0.739	0.739	0.739
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	0.000	0.300	0.200	0.200	0.300	0.300	0.300	0.200	0.300	0.300	0.300	0.200
6	0.000	0.058	0.072	0.884	0.899	0.899	0.899	0.928	0.971	0.986	0.986	0.986	0.986
7	1.000	0.919	1.000	1.000	1.000	1.000	1.000	0.973	0.973	0.973	0.973	0.973	0.973
8	0.881	0.687	0.687	0.701	0.716	0.851	0.806	0.821	0.806	0.806	0.806	0.821	0.821
9	0.000	0.045	0.136	0.818	0.818	0.886	0.886	0.886	0.909	0.886	0.864	0.864	0.864
Total	0.535	0.528	0.680	0.827	0.843	0.888	0.887	0.893	0.895	0.900	0.898	0.903	0.903

TABLE 3: The recognition accuracy rates with all features subsets of HMM based on maximum relevance selection.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.900	0.700	0.700	0.767	0.767	0.767	0.767	0.700	0.733	0.667	0.667
1	0.000	0.000	0.104	0.521	0.896	0.938	0.896	0.833	0.771	0.792	0.750	0.750	0.750
2	0.000	0.005	0.391	0.667	0.797	0.860	0.870	0.884	0.913	0.918	0.928	0.942	0.961
3	0.000	0.174	0.739	0.804	0.804	0.783	0.783	0.804	0.761	0.761	0.761	0.761	0.761
4	0.000	0.000	0.000	0.524	0.929	0.976	0.976	0.976	0.976	1.000	1.000	0.976	0.976
5	0.000	0.000	0.000	0.200	0.300	0.300	0.300	0.300	0.200	0.400	0.400	0.400	0.200
6	1.000	0.928	0.942	0.986	0.986	0.986	1.000	0.986	0.986	0.986	0.986	0.986	0.971
7	0.000	0.000	0.000	0.514	0.973	0.973	0.973	0.946	0.973	0.919	0.919	0.973	0.946
8	0.090	0.149	0.179	0.313	0.313	0.701	0.701	0.701	0.731	0.776	0.776	0.806	0.821
9	0.000	0.386	0.432	0.750	0.932	0.568	0.568	0.659	0.909	0.932	0.864	0.864	0.864
Total	0.170	0.212	0.405	0.643	0.790	0.837	0.838	0.843	0.865	0.873	0.872	0.878	0.880

For HMM, it can be found from Table 3 that features subset 10 generates the highest proportion of optimal IARARs for all activities, 40%. Concretely, with features subset 10, IARARs of activities 4, 5, 7, and 9 are better than or equal to those of the other subsets. However, features subset 13 yields the optimal TARAR of 88.0%, which is slightly higher than that of features subset 10, 87.3%. Again, it also can be concluded that feature  $f_3$  is not discriminatory to activity recognition with HMM.

**4.3. Feature Selection with mRMR Algorithm.** mRMR feature selection algorithm can be adopted to sort the 13 features and reduce the number of dimensions of features subset. This is done by analyzing the relevance of the features to the target activities on the original features set to remove the irrelevant features and keep the features which have strong correlations to the target activities and then analyzing the redundancy of the selected features to reduce the number of dimensions of features subset.

According to (11) to get the value of information of different features, the results of iterations are listed in Table 4 with D-R criterion. At the beginning, the MI equation has been applied to get the first optimal feature and the result is  $f_{10}$ . The next interaction, still according to (11), feature  $f_6$  is the optimal feature among the other remaining 12 features. And then, the ranks of each feature have been obtained successively.

Finally, the order of the sorted features is  $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}$ , and  $f_3$ . Again, it can be found that feature  $f_{10}$  is the most important feature and feature  $f_3$  is not discriminatory to activity recognition.

Similarly, the information values of different feature groups can be obtained according to (12), and the results of iterations are listed in Table 5 under D/R criterion.

The order of the sorted features is  $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9$ , and  $f_3$ .

Although the orders of the sorted features with D-R and D/R criteria are different, the first feature still is  $f_{10}$ , the most important feature, the second feature is  $f_6$ , and the last

TABLE 4: The information values of feature groups under the D-R criterion.

Feature													
Iteration	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1.188	0.688	0.047	1.173	0.863	1.490	1.200	1.321	1.508	1.535	1.168	1.183	0.530
1	0.476	0.213	-0.160	-0.009	0.441	0.774	0.310	0.146	-1.989	—	-0.032	0.607	0.134
2	0.376	0.230	-0.075	0.281	0.434	—	0.306	0.343	-0.594	—	0.261	0.363	0.151
3	0.520	0.288	-0.042	0.430	—	—	0.440	0.338	-0.041	—	0.453	0.520	0.195
4	0.425	0.268	-0.030	0.488	—	—	0.451	0.428	0.206	—	0.518	—	0.198
5	0.474	0.299	-0.023	0.452	—	—	0.456	0.458	0.238	—	—	—	0.216
6	—	0.173	-0.022	0.486	—	—	0.454	0.479	0.330	—	—	—	0.196
7	—	0.206	-0.019	—	—	—	0.484	0.482	0.326	—	—	—	0.207
8	—	0.207	-0.016	—	—	—	—	0.498	0.365	—	—	—	0.202
9	—	0.196	-0.025	—	—	—	—	—	0.364	—	—	—	0.188
10	—	0.196	-0.040	—	—	—	—	—	—	—	—	—	0.187
11	—	—	-0.040	—	—	—	—	—	—	—	—	—	0.190

TABLE 5: The information values of feature groups under the D/R criterion.

Feature													
Iteration	1	2	3	4	5	6	7	8	9	10	11	12	13
0	1.188	0.688	0.047	1.173	0.863	1.490	1.200	1.321	1.508	1.535	1.168	1.183	0.530
1	1.668	1.449	0.227	0.992	2.047	2.080	1.349	1.124	0.431	—	0.974	2.054	1.339
2	1.463	1.502	0.385	1.314	2.013	—	1.342	1.350	0.717	—	1.288	1.443	1.397
3	1.777	1.719	0.530	1.579	—	—	1.578	1.344	0.973	—	1.634	1.784	1.580
4	1.556	1.635	0.613	1.712	—	—	1.603	1.479	1.158	—	1.797	—	1.596
5	1.663	1.769	0.674	1.626	—	—	1.612	1.531	1.187	—	—	—	1.686
6	1.510	—	0.659	1.811	—	—	1.715	1.620	1.321	—	—	—	1.689
7	1.588	—	0.695	—	—	—	1.773	1.619	1.310	—	—	—	1.733
8	1.586	—	0.732	—	—	—	—	1.644	1.351	—	—	—	1.694
9	1.664	—	0.753	—	—	—	—	1.727	1.462	—	—	—	—
10	1.659	—	0.667	—	—	—	—	—	1.445	—	—	—	—
11	—	—	0.671	—	—	—	—	—	1.487	—	—	—	—

feature is  $f_3$ , which is not discriminatory to activity recognition.

According to the orders of the sorted features with D-R and D/R criteria, 13 features subsets can be obtained through adding one feature to previous features subset each time successively.

The 13 features subsets under D-R criterion are the following:

- Features subset 1: ( $f_{10}$ )
- Features subset 2: ( $f_{10}, f_6$ )
- Features subset 3: ( $f_{10}, f_6, f_5$ )
- Features subset 4: ( $f_{10}, f_6, f_5, f_{12}$ )
- Features subset 5: ( $f_{10}, f_6, f_5, f_{12}, f_{11}$ )
- Features subset 6: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1$ )
- Features subset 7: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4$ )
- Features subset 8: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7$ )
- Features subset 9: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8$ )
- Features subset 10: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9$ )
- Features subset 11: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2$ )
- Features subset 12: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}$ )
- Features subset 13: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_1, f_4, f_7, f_8, f_9, f_2, f_{13}, f_3$ )

And the 13 features subsets under D/R criterion are the following:

- Features subset 1: ( $f_{10}$ );
- Features subset 2: ( $f_{10}, f_6$ )
- Features subset 3: ( $f_{10}, f_6, f_5$ )
- Features subset 4: ( $f_{10}, f_6, f_5, f_{12}$ )
- Features subset 5: ( $f_{10}, f_6, f_5, f_{12}, f_{11}$ )
- Features subset 6: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2$ )
- Features subset 7: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4$ )
- Features subset 8: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7$ )
- Features subset 9: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}$ )
- Features subset 10: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8$ )
- Features subset 11: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1$ )
- Features subset 12: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9$ )
- Features subset 13: ( $f_{10}, f_6, f_5, f_{12}, f_{11}, f_2, f_4, f_7, f_{13}, f_8, f_1, f_9, f_3$ )

The selected features subsets have been validated by NB classifier and HMM to recognize activities. TARARs of NB classifier and HMM, with each features subset, are listed in Table 6, respectively.

It can be found that although D-R and D/R criteria are different, the features subsets which have the same number



TABLE 6: Comparison results of TARARs under two mRMR evaluation criteria (D-R and D/R).

Features subset		1	2	3	4	5	6	7	8	9	10	11	12	13
NB	D-R	0.535	0.695	0.870	0.870	0.870	0.890	0.885	0.902	0.903	0.900	0.898	0.903	0.903
	D/R	0.535	0.695	0.870	0.870	0.870	0.880	0.877	0.892	0.888	0.887	0.903	0.903	0.903
HMM	D-R	0.170	0.255	0.447	0.523	0.638	0.745	0.770	0.832	0.837	0.873	0.872	0.878	0.880
	D/R	0.170	0.255	0.447	0.523	0.638	0.747	0.768	0.817	0.837	0.858	0.853	0.878	0.880

of features are similar, even the same, e.g., if the number of features is little than 6. Therefore, it can be observed that the results of the same classifier with the same number of features are very slightly different, under the two criteria. If the number of features is smaller than 6 or larger than 11, the results of TARARs are the same.

Moreover, the performances of NB classifier and HMM are different even with the same features subset. Generally, with the increasing of size of features subset, TARAR increases until the size reaches a certain value. Still, the trend is not monotonic. For example, the optimal results are with features subsets 9, 12, and 13 under D-R criterion (90.3%), and the optimal result is with features subset 13 under D/R criterion (88.0%), for NB classifier. For HMM, the optimal result is with features subset 13 under D-R (88.0%) and D/R (88.0%) criteria. The results show that NB classifier and HMM yield relatively higher TARARs with features subset 5 through 13 and TARARs of NB classifier are slightly higher than those of HMM with features subsets 8 through 13, but much higher with 1 through 7, under the same criterion.

Basically, it can be concluded that both D-R and D/R criteria can be applied to sort the features and then to compose features subset effectively.

Tables 7 and 8 (under D-R criterion) show that the optimal IARARs of activities 0, 3, 5, 6, and 9 of HMM are higher than those of NB classifier, while the optimal IARARs of the activities 1, 2, 7, and 8 of NB classifier are higher than those of HMM. The optimal IARARs of activity 4 are the same of the two classifiers. Generally, the trends of the two classifiers are nearly the same with the increasing of size of features subset. Again, the trends are not monotonic. Moreover, it can be observed that different features have different effects on activity recognition. For example, features subset 5 means introducing feature  $f_{11}$  to features subset 4; then, IARARs have been improved of some activities, such as activities 0, 2, 3, 4, and 7 through 9; however, IARARs of other activities are degraded or not improved, of NB classifier, as shown in Table 7.

HMM has a similar result. It also can be found that, according to the different performance traits of NB classifier and HMM, the adding of the same feature has different effects on IARARs. For example, by introducing of the feature  $f_5$  to features subset 2 to compose features subset 3, IARAR of activities 0 has been improved in NB classifier apparently but degraded in HMM. By introducing of feature  $f_{11}$  to features subset 4 to compose features subset 5, IARAR of activity 9 of NB classifier has been degraded but improved in HMM.

Furthermore, from Table 7, it is also shown that, with features subset 13, IARARs of activities 0, 1, 2, 4, and 6 are better than or equal to those of the other subsets; the proportion is 50% for all activities. With features subset 9, the proportion is 40% (activities 1, 5, 6, and 9). However, with features subset 9, IARARs of activities 1, 3, 5, 6, 7, and 9 are better than or equal to those of features subset 13. Therefore, features subset 9 yields the optimal TARAR of 90.3%, which is the same as that of features subsets 12 and 13. Still, features  $f_3$  is not discriminatory to activity recognition with NB classifier.

It also can be found from Table 8 that, with features subset 13, IARARs of activities 2 and 8 are better than or equal to those of the other features subsets; the proportion is 20% for all activities, while features subset 13 gives the optimal TARAR of 88.0%, which is slightly higher than that of features subset 12. Moreover, with features subset 12, IARARs of activities 0, 1, and 3 to 7 and 9 are better than or equal to those of the features subset 13; the proportion is 80%. Again, it also can be concluded that feature  $f_3$  is not discriminatory to activity recognition with HMM.

Under the D/R criterion, Tables 9 and 10 show the activity recognition accuracy rates with each features subset of NB classifier and HMM. From the experimental results, it can be found that the performances of NB classifier and HMM with the features subsets obtained under D/R criterion are similar to those of D-R criterion, respectively.

For example, for NB classifier, the optimal TARARs are 90.3% with features subsets 9, 12, and 13 under D-R criterion, which are the same to those of features subsets 11, 12, and 13 under D/R criterion. And for HMM, the optimal TARAR is 88.0% with features set 13 under D-R criterion, which is the same as that under D/R criterion.

From Table 9, for NB classifier, it is indicated that, with features subsets 12 and 13, IARARs of activities 1, 4, 6, and 9 are better than or equal to those of the other features subsets; the proportion is 40%, for all activities. However, with features subset 11, IARARs of activities 1, 3, 5, 6, and 9 are better than or equal to those of features subsets 12 and 13; the proportion is 50%. Features subsets 11, 12, and 13 yield the same optimal TARARs of 90.3%. Again, features  $f_3$  and  $f_{13}$  are not discriminatory to activity recognition with NB classifier.

It also can be found from Table 10, for HMM, that with features subset 13, IARARs of activities 1, 2, 4, and 8 are better than or equal to those of the other features subsets. Features subset 13 yields the optimal TARAR of 88.0%, which is slightly higher than that of features subset 12. Actually, with features subset 12, IARARs of activities 0, 1, 3,

TABLE 7: The recognition accuracy rates with each features subset of NB classifier under the D-R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.000	0.400	0.400	0.400	0.667	0.667	0.667	0.600	0.600	0.633	0.667	0.667
1	0.792	0.979	0.938	0.958	0.958	0.958	0.958	0.958	0.979	0.979	0.979	0.979	0.979
2	0.700	0.899	0.913	0.913	0.913	0.908	0.899	0.937	0.947	0.957	0.957	0.961	0.966
3	0.000	0.848	0.848	0.761	0.761	0.761	0.783	0.783	0.804	0.761	0.739	0.739	0.739
4	1.000	0.976	1.000	1.000	1.000	1.000	1.000	1.000	0.976	1.000	1.000	1.000	1.000
5	0.000	0.100	0.300	0.600	0.600	0.600	0.400	0.500	0.600	0.300	0.300	0.300	0.200
6	0.000	0.072	0.913	0.928	0.957	0.957	0.957	0.957	0.986	0.986	0.986	0.986	0.986
7	1.000	1.000	0.973	0.973	0.973	0.973	0.946	0.973	0.973	0.973	0.973	0.973	0.973
8	0.881	0.881	0.821	0.806	0.821	0.836	0.851	0.851	0.806	0.806	0.806	0.821	0.821
9	0.000	0.045	0.864	0.864	0.795	0.886	0.886	0.886	0.886	0.886	0.864	0.864	0.864
Total	0.535	0.695	0.870	0.870	0.870	0.890	0.885	0.902	0.903	0.900	0.898	0.903	0.903

TABLE 8: The recognition accuracy rates with each features subset of HMM under the D-R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.800	0.800	0.833	0.867	0.833	0.833	0.700	0.700	0.733	0.667	0.667
1	0.000	0.000	0.292	0.271	0.250	0.229	0.083	0.500	0.458	0.792	0.750	0.750	0.750
2	0.000	0.005	0.184	0.256	0.483	0.657	0.768	0.860	0.908	0.918	0.928	0.942	0.961
3	0.000	0.870	0.935	0.957	0.978	0.935	0.891	0.804	0.826	0.761	0.761	0.761	0.761
4	0.000	0.000	0.524	0.524	0.714	0.952	1.000	0.952	0.929	1.000	1.000	0.976	0.976
5	0.000	0.000	0.400	0.700	0.400	0.700	0.500	0.500	0.600	0.400	0.400	0.400	0.200
6	1.000	0.986	1.000	1.000	1.000	0.971	0.971	0.986	0.986	0.986	0.986	0.986	0.971
7	0.000	0.000	0.297	0.703	0.892	0.973	0.892	0.946	0.838	0.973	0.919	0.973	0.946
8	0.090	0.119	0.239	0.373	0.388	0.701	0.701	0.701	0.701	0.731	0.776	0.806	0.821
9	0.000	0.205	0.614	0.705	0.886	0.773	0.886	0.909	0.955	0.932	0.864	0.864	0.864
Total	0.170	0.255	0.447	0.523	0.638	0.745	0.770	0.832	0.837	0.873	0.872	0.878	0.880

TABLE 9: The recognition accuracy rates with each features subset of NB classifier under the D/R criterion.

Features subset													
Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.000	0.000	0.400	0.400	0.400	0.700	0.700	0.700	0.733	0.600	0.600	0.667	0.667
1	0.792	0.979	0.938	0.958	0.958	0.938	0.958	0.958	0.958	0.958	0.979	0.979	0.979
2	0.700	0.899	0.913	0.913	0.913	0.908	0.903	0.947	0.947	0.961	0.961	0.961	0.966
3	0.000	0.848	0.848	0.761	0.761	0.783	0.783	0.783	0.783	0.804	0.804	0.739	0.739
4	1.000	0.976	1.000	1.000	1.000	0.976	1.000	0.976	0.976	0.952	0.952	1.000	1.000
5	0.000	0.100	0.300	0.600	0.600	0.600	0.400	0.400	0.400	0.500	0.600	0.300	0.200
6	0.000	0.072	0.913	0.928	0.957	0.971	0.942	0.942	0.942	0.971	0.986	0.986	0.986
7	1.000	1.000	0.973	0.973	0.973	0.946	0.946	0.946	0.946	0.946	0.946	0.973	0.973
8	0.881	0.881	0.821	0.806	0.821	0.836	0.851	0.851	0.821	0.791	0.806	0.821	0.821
9	0.000	0.045	0.864	0.864	0.795	0.750	0.750	0.773	0.750	0.727	0.864	0.864	0.864
Total	0.535	0.695	0.870	0.870	0.870	0.880	0.877	0.892	0.888	0.887	0.903	0.903	0.903

4, 5, 6, 7, and 9 are better than or equal to those of the features subset 13; the proportion is 80%. Again, features  $f_3$  and  $f_{13}$  are not discriminatory to activity recognition with HMM.

Table 11 shows the comparison results of the optimal IARARs and TARARs of NB classifier and HMM under the three evaluation criteria (MR, D-R, and D/R) and the required minimal size of features subset, respectively.

It indicates that although the features subsets are different under the three evaluation criteria, however, the optimal results of IARARs and TARARs are similar to the

same classifier. For example, for NB classifier, the optimal IARARs of activities 0, 2, 4, 6, 7, and 8 under MR criterion are the same as those under D-R criterion; furthermore, the optimal IARARs of activities 2, 4, 6, 7, and 8 are the same even under the three evaluation criteria, and even the optimal TARARs are the same.

Similarly, for HMM, the optimal IARARs of activities 0, 2, and 6 through 8 are the same even under the three evaluation criteria. Besides, the optimal TARARs are also the same with the three evaluation criteria. Additionally, it can be found that the results of the required minimal size of

TABLE 10: The recognition accuracy rates with each features subset of HMM under the D/R criterion.

Activity	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0.900	0.900	0.800	0.800	0.833	0.833	0.800	0.800	0.800	0.667	0.600	0.667	0.667
1	0.000	0.000	0.292	0.271	0.250	0.688	0.417	0.625	0.625	0.667	0.500	0.750	0.750
2	0.000	0.005	0.184	0.256	0.483	0.652	0.758	0.865	0.879	0.918	0.937	0.942	0.961
3	0.000	0.870	0.935	0.957	0.978	0.935	0.913	0.804	0.804	0.826	0.826	0.761	0.761
4	0.000	0.000	0.524	0.524	0.714	0.738	0.857	0.857	0.833	0.881	0.952	0.976	0.976
5	0.000	0.000	0.400	0.700	0.400	0.500	0.400	0.500	0.500	0.600	0.600	0.400	0.200
6	1.000	0.986	1.000	1.000	1.000	0.986	0.986	0.986	0.986	1.000	0.986	0.986	0.971
7	0.000	0.000	0.297	0.703	0.892	0.919	0.919	0.919	0.865	0.892	0.811	0.973	0.946
8	0.090	0.119	0.239	0.373	0.388	0.522	0.522	0.522	0.716	0.746	0.806	0.806	0.821
9	0.000	0.205	0.614	0.705	0.886	0.886	0.932	0.955	0.932	0.909	0.909	0.864	0.864
Total	0.170	0.255	0.447	0.523	0.638	0.747	0.768	0.817	0.837	0.858	0.853	0.878	0.880

TABLE 11: Optimal recognition accuracy rates under three evaluation criteria (MR, D-R, and D/R) and the required minimal size of features subset (in brackets).

Activity	Classifier	Criterion	0	1	2	3	4	5	6	7	8	9	Total
NB	MR		0.667 (12)	1.000 (8)	0.966 (13)	0.739 (10)	1.000 (1)	0.300 (3)	0.986 (10)	1.000 (1)	0.881 (1)	0.909 (9)	0.903 (12)
	D-R		0.667 (6)	0.979 (2)	0.966 (13)	0.848 (2)	1.000 (1)	0.600 (4)	0.986 (9)	1.000 (1)	0.881 (1)	0.886 (6)	0.903 (9)
	D/R		0.733 (9)	0.979 (2)	0.966 (13)	0.848 (2)	1.000 (1)	0.600 (4)	0.986 (11)	1.000 (1)	0.881 (1)	0.864 (3)	0.903 (11)
HMM	MR		0.900 (1)	0.938 (6)	0.961 (13)	0.804 (4)	1.000 (10)	0.400 (10)	1.000 (1)	0.973 (5)	0.821 (13)	0.932 (5)	0.880 (13)
	D-R		0.900 (1)	0.792(10)	0.961 (13)	0.978 (5)	1.000 (7)	0.700 (4)	1.000 (1)	0.973 (6)	0.821 (13)	0.955 (9)	0.880 (13)
	D/R		0.900 (1)	0.750 (12)	0.961 (13)	0.978 (5)	0.976 (12)	0.700 (4)	1.000 (1)	0.973 (12)	0.821 (13)	0.955 (8)	0.880 (13)

features subset according to the optimal IARAR of each activity and the optimal TARAR are also similar to the same classifier. For example, for NB classifier, the required minimal sizes of features subset of activity 1 to 5 and 7 to 8, under D-R criterion, are the same as those under D/R criterion. Moreover, for activities 2, 4, 7, and 8, the required minimal sizes of features subset are the same under the three evaluation criteria. For HMM, the required minimal sizes of features subset of activities 0, 2, 6, and 8 are the same under the three evaluation criteria; the proportion is 40%. Besides, the required minimal sizes of features subset of TARARs are also the same under the three evaluation criteria.

Obviously, maximal relevance measure and mRMR feature selection algorithm (under D-R and D/R criteria) are effective to feature selection for human activity recognition.

## 5. Conclusions

This paper has applied maximal relevance measure and minimal redundancy maximal relevance (mRMR) algorithm (under D-R and D/R criteria) to select features and to compose different features subsets based on the observed sensor events for human activity recognition in smart home environments. And then, the selected features subsets have been validated by NB classifier and HMM to

recognize human activities. Through the experimental results, it is shown that not all the features are beneficial to activity recognition, as estimated. Different combinations of features lead to different activity recognition results. Furthermore, even the same features subset has different effect on the activity recognition accuracy rate for different activity classifiers. It can be found that feature  $f_{10}$  (the ending time of the current activity) is the most important feature to activity recognition and feature  $f_3$  (day of week) is not discriminatory to activity recognition. Therefore, the suitable features subset must be selected in advance, and the selection of unsuitable feature sets increases the computational complexity and degrades the human activity recognition accuracy rate. Moreover, it is significant for activity recognition to consider both relevance between features and activities and redundancy between features. Generally, maximal relevance measure and mRMR algorithm are beneficial to feature selection and positive to activity recognition of NB classifier and HMM.

## Data Availability

The data were collected from the smart home testbed located on Washington State University campus, which can be downloaded from Dr. Cook's homepage.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

Thanks are due to Dr. Diane J. Cook whose positive and constructive comments helped to enhance the quality and presentation of this paper. Also, the authors gratefully acknowledge the data preparation work that was carried out by Dr. Cook's graduate students.

## References

- [1] L. Liu, E. Stroulia, I. Nikolaidis, A. Miguel-Cruz, and A. Rios Rincon, "Smart homes and home health monitoring technologies for older adults: a systematic review," *International Journal of Medical Informatics*, vol. 91, pp. 44–59, 2016.
- [2] M. R. Alam, M. B. I. Reaz, and M. A. M. Ali, "A review of smart homes-past, present, and future," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 1190–1203, 2012.
- [3] D. J. Cook, "How smart is your home?" *Science*, vol. 335, no. 6076, pp. 1579–1581, 2012.
- [4] G. Sprint, D. J. Cook, R. Fritz, and M. Schmitter-Edgecombe, "Using smart homes to detect and analyze health events," *Computer*, vol. 49, no. 11, pp. 29–37, 2016.
- [5] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.
- [6] T. Luor, H. Lu, H. Yu, and Y. Lu, "Exploring the critical quality attributes and models of smart homes," *Maturitas*, vol. 82, no. 4, pp. 377–386, 2015.
- [7] J. Dahmen, B. Thomas, D. Cook, and X. Wang, "Activity learning as a foundation for security monitoring in smart homes," *Sensors*, vol. 17, no. 4, p. 737, 2017.
- [8] H. M. S. Hossain, M. A. A. H. Khan, and N. Roy, "Active learning enabled activity recognition," *Pervasive and Mobile Computing*, vol. 38, no. 2, pp. 312–330, 2017.
- [9] G. Singla, D. Cook, and M. Schmitter-Edgecombe, "Tracking activities in complex settings using smart environment technologies," *International Journal of BioSciences, Psychiatry and Technology*, vol. 1, no. 1, pp. 25–35, 2009.
- [10] A. S. Crandall and D. J. Cook, "Coping with multiple residents in a smart environment," *Journal of Ambient Intelligence and Smart Environments*, vol. 1, no. 4, pp. 323–334, 2009.
- [11] L. Liu, S. Wang, G. Su, Z.-G. Huang, and M. Liu, "Towards complex activity recognition using a Bayesian network-based probabilistic generative framework," *Pattern Recognition*, vol. 68, pp. 295–309, 2017.
- [12] K. S. Gayathri, K. S. Easwarakumar, and S. Elias, "Probabilistic ontology based activity recognition in smart homes using markov logic network," *Knowledge-Based Systems*, vol. 121, pp. 173–184, 2017.
- [13] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Computing*, vol. 9, no. 1, pp. 48–53, 2010.
- [14] Y. Zhu, N. M. Nayak, and A. K. Roy-Chowdhury, "Context-aware activity modeling using hierarchical conditional random fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 7, pp. 1360–1372, 2015.
- [15] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 6, pp. 961–974, 2012.
- [16] M. Fahim, I. Fatima, S. Lee, and Y.-K. Lee, "EEM: evolutionary ensembles model for activity recognition in smart homes," *Applied Intelligence*, vol. 38, no. 1, pp. 88–98, 2013.
- [17] A. Fleury, M. Vacher, and N. Noury, "SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 274–283, 2010.
- [18] J. Wen and Z. Wang, "Learning general model for activity recognition with limited labelled data," *Expert Systems with Applications*, vol. 74, pp. 19–28, 2017.
- [19] J.-H. Hong, J. Ramos, and A. K. Dey, "Toward personalized activity recognition systems with a semipopulation approach," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 101–112, 2016.
- [20] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: understanding human behavior," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 81–94, 2016.
- [21] N. S. Mohamed, S. Zainudin, and Z. Ali Othman, "Meta-heuristic approach for an enhanced mRMR filter method for classification using drug response microarray data," *Expert Systems with Applications*, vol. 90, pp. 224–231, 2017.
- [22] J. Che, Y. Yang, L. Li, X. Bai, S. Zhang, and C. Deng, "Maximum relevance minimum common redundancy feature selection for nonlinear data," *Information Sciences*, vol. 409–410, pp. 68–86, 2017.
- [23] C. Xu, S. Zhao, and F. Liu, "Distributed plant-wide process monitoring based on PCA with minimal redundancy maximal relevance," *Chemometrics and Intelligent Laboratory Systems*, vol. 169, pp. 53–63, 2017.
- [24] F. Li, D. Miao, and W. Pedrycz, "Granular multi-label feature selection based on mutual information," *Pattern Recognition*, vol. 67, pp. 410–423, 2017.
- [25] D. Escalona-Vargas, E. R. Siegel, P. Murphy, C. L. Lowery, and H. Eswaran, "Selection of reference channels based on mutual information for frequency-dependent subtraction method applied to fetal biomagnetic signals," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1115–1122, 2017.
- [26] S. Mallik, T. Bhadra, and U. Maulik, "Identifying epigenetic biomarkers using maximal relevance and minimal redundancy based feature selection for multi-omics data," *IEEE Transactions on NanoBioscience*, vol. 16, no. 1, pp. 3–10, 2017.
- [27] A. K. Tiwari, "Prediction of G-protein coupled receptors and their subfamilies by incorporating various sequence features into Chou's general PseAAC," *Computer Methods and Programs in Biomedicine*, vol. 134, pp. 197–213, 2016.
- [28] C. Chen and X. Yan, "Optimization of a multilayer neural network by using minimal redundancy maximal relevance-partial mutual information clustering with least square regression," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1177–87, 2015.
- [29] Z. Wang, M. Li, and J. Li, "A multi-objective evolutionary algorithm for feature selection based on mutual information with a new redundancy measure," *Information Sciences*, vol. 307, pp. 73–88, 2015.
- [30] P. M. Morgado and M. Silveira, "Minimal neighborhood redundancy maximal relevance: Application to the diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 155, pp. 295–308, 2015.
- [31] M. Kamandar and H. Ghassemian, "Linear feature extraction for hyperspectral images based on information theoretic

- learning,” *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 702–706, 2013.
- [32] K. K. Kandaswamy, G. Pugalenthi, K.-U. Kalies, E. Hartmann, and T. Martinetz, “EcmPred: prediction of extracellular matrix proteins based on random forest with maximum relevance minimum redundancy feature selection,” *Journal of Theoretical Biology*, vol. 317, pp. 377–383, 2013.
- [33] X. Jin, E. W. M. Ma, L. L. Cheng, and M. Pecht, “Health monitoring of cooling fans based on Mahalanobis distance with mRMR feature selection,” *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, pp. 2222–2229, 2012.
- [34] A. Unler, A. Murat, and R. B. Chinnam, “MR2PSO: a maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification,” *Information Sciences*, vol. 181, no. 20, pp. 4625–4641, 2011.
- [35] E. Zdravevski, P. Lameski, V. Trajkovik et al., “Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering,” *IEEE Access*, vol. 5, pp. 5262–5280, 2017.
- [36] E. Zdravevski, B. Risteska Stojkoska, M. Standl, and H. Schulz, “Automatic machine-learning based identification of jogging periods from accelerometer measurements of adolescents under field conditions,” *PLoS ONE*, vol. 12, no. 9, p. e0184216, 2017.
- [37] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
- [38] H. Fang, L. He, H. Si, P. Liu, and X. Xie, “Human activity recognition based on feature selection in smart home using back-propagation algorithm,” *ISA Transactions*, vol. 53, no. 5, pp. 1629–1638, 2014.
- [39] P. Tang, H. Fang, and H. Si, “Maximal relevance feature selection for human activity recognition in smart home,” in *Proceedings of the 2018 Chinese Control and Decision Conference. (CCDC)*, pp. 4264–4268, Shenyang, China, June 2018.