

Applying Machine Learning Techniques in Older People Activity Recognition using Wearable and Mobile Devices

Flavio Vinicius Vieira Santana, Bruno Henrique Rasteiro, Larissa Cardoso Zimmermann, Luciana De Nardin, Maria da Graça Campos Pimentel
Universidade de São Paulo, São Carlos-SP, Brasil

flavio.santana@usp.br,bruno.rasteiro@usp.br,larissa.zimmermann@usp.br,lnardin@usp.br,mgp@icmc.usp.br

ABSTRACT

We investigate the potential of the combined use of smartwatch accelerometer data and smartphone apps for online older adults activity recognition. We selected machine learning algorithms which resulted in *a posteriori* recognition accuracy of 98.92%. Our smartphone app, with the selected machine learning algorithms, carried out online recognition from data captured on the smartwatch. These results allow us, as future work, assess the accuracy of online recognition when the system is used by older adults.

Keywords: HAR, smartphone, smartwatch, ML, older people.

1 INTRODUÇÃO

Reconhecer atividades humanas (HAR: *Human Activity Recognition*) e identificar (quebra de) padrões comportamentais de indivíduos é um campo de pesquisa promissor e desafiador, com aplicações em segurança, saúde, entretenimento, entre outras [8]. Apesar do crescimento da população idosa em todo o mundo, poucas pesquisas em HAR empregam *datasets* compostos por atividades realizadas por idosos [3, 9] — o que é problemático dado que diferenças significativas foram identificadas, inclusive, entre grupos de idosos [6]

Alguns pesquisadores investigam HAR a partir de dados de sensores de *smartphones* [6, 7]. Entretanto *smartwatches*, dotados de acelerômetro, giroscópio, microfone e sensor de batimentos cardíacos, permitem coleta informações de forma contínua e não intrusiva [2, 4]. Tendo como público-alvo adultos com idade a partir de 60 anos, o objetivo deste trabalho é investigar o potencial do uso combinado do acelerômetro de *smartwatches* com uma *smartphone app* para o reconhecimento de um conjunto de atividades humanas (andar, correr, subir e descer escada, sentar e levantar) *enquanto* a atividade é realizada (*online activity recognition*).

Utilizando o *dataset* EARDB [12], com dados capturados de 32 idosos enquanto realizam as atividades selecionadas, implementamos algoritmos de aprendizado de máquina que obtiveram acurácia de reconhecimento das atividades de 98.92% em processamento *a posteriori*. Como requisito para realizar o reconhecimento *online* com usuários idosos, investigamos a possibilidade do reconhecimento ser feito *online* por uma *smartphone app* que processa dados capturados do acelerômetro de um *smartwatch*. Implementamos uma *Android app* que, utilizando os algoritmos que apresentaram melhor resultado no reconhecimento *a posteriori*, recebe os dados de um *smartwatch* e realiza com sucesso o reconhecimento *online* no



Figura 1: (esq) eixos do acelerômetro; (dir) smartwatch app

smartphone. Em teste realizado com um usuário jovem, a acurácia caiu para 66.9% — o que era esperado pelo perfil de usuário diferente daquele do EARDB com os quais os algoritmos foram treinados.

Neste texto, a Seção 2 apresenta a EARS Database utilizada, a Seção 3 os trabalhos relacionados, a Seção 4 a metodologia utilizada, a Seção 5 os resultados, e a Seção 6 as considerações finais.

2 EARDB DATASET

Para investigar o uso de sensores de *smartphones* e *smartwatches* no reconhecimento de atividades de pessoas idosas, Zimmermann [12] desenvolveu um conjunto de aplicativos que permitem a coleta de dados para a criação de *datasets* associados. Após aprovação do projeto de pesquisa (Plataforma Brasil CAAE: 57875016.3.0000.5390), em colaboração com uma gerontologista foram coletados dados de 15 mulheres e 17 homens com idade entre 60 e 75 anos enquanto realizavam as atividades de andar, correr, subir e descer escada, sentar e levantar — classificadas como atividades *simples* por Eun et al. [5]. A coleta fez uso de protocolos padronizados na área da saúde (eg *Timed Up and Go*) enquanto os usuários utilizavam um *smartwatch* instrumentado para a coleta (Figura 1). A execução de cada atividade foi capturada por 2 minutos; foi utilizado o *smartwatch LG Urbane* com *Android Wear 2.0*; foram coletados os dados do acelerômetro de três eixos (Fig 1:esq) com frequência de amostragem de 16Hz. O correspondente *dataset* EARDB (*Elderly Activity Recognition DataBase*) foi utilizado no *Elderly Physical Activity Reminder System* (EPARS) [13].

3 TRABALHOS RELACIONADOS

Capela et al. [1] reportam o estudo de características utilizadas em aplicações HAR baseadas em dispositivos vestíveis e reportadas na literatura. A análise foi realizada de forma a não levar em consideração o classificador utilizado na aplicação. O estudo selecionou 76 características e realizou a análise das mesmas a partir de coleta de dados com um *smartphone* preso à cintura dos participantes.

A análise foi feita com três métodos de seleção de características (*ReliefF*, *Correlation based Feature Selection* e *Fast Correlation based Filter*), em três populações diferentes e em sete níveis de atividades. No nível de atividade e na população de interesse deste projeto (pessoas acima de 60 anos) foram selecionadas em média 10 características e os recursos do sinal com *smartphone* nos três grupos de indivíduos (saudáveis, idosos e AVC) foram identificados com sucesso. A aplicação dos três classificadores mostrou ainda que a seleção de um subconjunto de características produziu resultados com precisão bastante semelhante, ou um pouco melhor, do que quando foi utilizado o conjunto completo de características, evidenciando que, o conjunto de características utilizado é independente do classificador. No contexto do trabalho aqui reportado, a principal diferença é com relação a forma com que os dados foram coletados: Capela et al. [1] utilizam *smartphone* preso à cintura enquanto utilizamos um *smartwatch* no pulso do participante.

Zhang and Sawchuk [11] selecionaram características de dois tipos: 87 características estatísticas e 23 características físicas. Participaram da coleta de dados 6 indivíduos de diferentes gêneros, idade, altura e peso. O dispositivo utilizado - *NotionMode*, foi fixado na parte frontal do quadril dos indivíduos. A coleta dos dados foi feita a 100Hz utilizando uma janela de tempo de 2 segundos com 50% de sobreposição. A seleção das características foi feita através de 3 métodos: *ReliefF*, *SFC (Wrapper Method based on Single Feature Classification)* e *SFS (Wrapper Method based on Sequential Forward Selection)*. Utilizando apenas 5 características estatísticas selecionadas pelo SFS, o classificador SVM com *kernel* linear apresentou um erro de 12%, considerado um bom resultado. Em relação ao nosso trabalho, Zhang and Sawchuk [11] se diferencia pelo tipo de aparelho, posicionamento do corpo, características dos participantes, frequência de captura e janela de tempo.

Yang et al. [10] reportam estudo específico para avaliação de risco de queda que comparou (a) informações capturadas por *smartbands* utilizadas durante 7 dias por indivíduos moradores de um residencial para idosos, e (b) dados de uma avaliação padrão utilizada no residencial (RAI-HC: *Resident Assessment Instrument-Home Care*). Os dados dos dois instrumentos foram analisados *a posteriori* utilizando os algoritmos *logistic regression*, *decision tree* e *random forest*. Os autores concluíram que os dados do RAI-HC tiveram performance melhor que os dados da *smartband*, e que o melhor resultado é obtido com a combinação dos dois. Diferentemente do nosso estudo, os autores avaliaram explicitamente o risco de queda, utilizaram o histórico dos pacientes e analisaram dados coletados em ambiente real. Os dados coletados utilizam atividades reconhecidas como repouso e *walking*, *jogging/running* a valores do batimento cardíaco obtido durante essas atividades.

4 CARACTERÍSTICAS: GERAÇÃO E SELEÇÃO

4.1 Geração das características

Esta seção detalha o processo de geração das características e as análises feitas sobre os conjuntos de características obtidos. As características consistem de medidas estatísticas como média e desvio padrão que são utilizadas como entrada para treinar o classificador.

Todas as características foram geradas a partir de quatro dados obtidos pelos sensores: aceleração no eixo X (A_x), aceleração no eixo Y (A_y), aceleração no eixo Z (A_z) e a aceleração linear (Acc). Tendo

Tabela 1: Características Estatísticas

Nome	Descrição
mean_Ax, mean_Ay, mean_Az	Média do sinal nos 3 eixos de aceleração
std_Ax, std_Ay, std_Az	Desvio padrão do sinal nos 3 eixos de aceleração
rms_Ax, rms_Ay, rms_Az	RMS (Raiz média quadrática) do sinal nos 3 eixos de aceleração
acc	Média da aceleração linear
diff	Diferença da aceleração ($A_y - A_z - A_x$)

em vista a frequência de 16 Hz de amostragem do sensor utilizado na coleta dos dados, foi possível então definir uma janela de tempo para sumarizar as instâncias e assim obter as características. Neste trabalho foi utilizado uma janela de 2 segundos o que significa que, a cada 32 instâncias de dados coletados foi calculada e definida uma instância com N características. No caso da características média (*mean*), por exemplo, ela consiste no valor médio das 32 instâncias coletadas no intervalo de 16 Hz para cada um dos eixos.

O processo de geração das características é feito para cada indivíduo e cada atividade, dessa forma se garante que instâncias de indivíduos ou classes distintas não se misturem. A última etapa do processo é a aleatorização (embaralhamento) dos dados obtidos com o objetivo de evitar um possível viés nos dados. As características estatísticas geradas estão especificadas na Tabela 1.

4.2 Seleção das características

Considerando as 11 características descritas na Tabela 1, foi realizada uma análise com o objetivo de reduzir a dimensionalidade. Para isso, foram testadas todas as combinações possíveis de características com a condição de que as 3 características de média ficassem sempre juntas, bem como desvio padrão e RMS. Dessa forma, obteve-se os 31 subconjuntos de características identificados na Tabela 2.

A acurácia foi a métrica de avaliação do desempenho para cada subconjunto em um sistema de avaliação cruzada com 10 *folds*. Os classificadores escolhidos foram Árvore de Decisão e KNN. Uma vez que o algoritmo KNN necessita da definição do parâmetro livre K, que representa o número de vizinhos mais próximos a serem considerados no processo de classificação, uma análise foi realizada onde para cada subconjunto definido: o algoritmo foi executado com o valor de K variando de 1 a 10, e o que obteve melhor acurácia foi definido como K padrão. Os resultados são reportados a seguir.

5 RESULTADOS

5.1 Árvore de decisão

Os resultados obtidos utilizando o algoritmo de Árvore de Decisão podem ser observados na Figura 2, que apresenta no eixo x cada um dos subconjuntos de características identificados na Tabela 2. Cada experimento foi repetido 10 vezes utilizando validação cruzada: os traços verticais no topo das barras indicam o desvio padrão. Observa-se que os diferentes subconjuntos tiveram pouca influência no desempenho do classificador. A melhor acurácia obtida foi de 98.33, utilizando o subconjunto #20 (características mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay e std_Az).

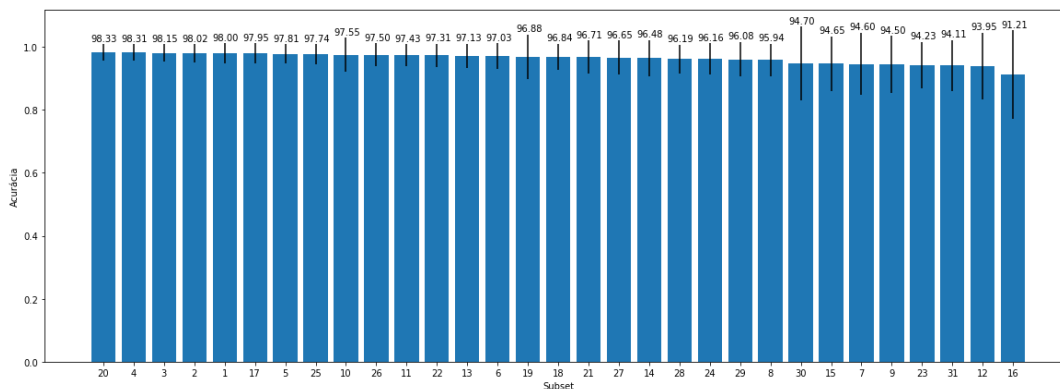


Figura 2: Acurácia nos subconjuntos de características com Árvores de Decisão.

Tabela 2: Subconjuntos (SC) de características estatísticas

SC	características
1	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az, acc
2	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az
3	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, acc
4	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az
5	diff, mean_Ax, mean_Ay, mean_Az, rms_Ax, rms_Ay, rms_Az, acc
6	diff, mean_Ax, mean_Ay, mean_Az, rms_Ax, rms_Ay, rms_Az
7	diff, mean_Ax, mean_Ay, mean_Az, acc
8	diff, mean_Ax, mean_Ay, mean_Az
9	diff, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az, acc
10	diff, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az
11	diff, std_Ax, std_Ay, std_Az, acc
12	diff, std_Ax, std_Ay, std_Az
13	diff, rms_Ax, rms_Ay, rms_Az, acc
14	diff, rms_Ax, rms_Ay, rms_Az
15	diff, acc
16	diff
17	mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az, acc
18	mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az
19	mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, acc
20	mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az
21	mean_Ax, mean_Ay, mean_Az, rms_Ax, rms_Ay, rms_Az, acc
22	mean_Ax, mean_Ay, mean_Az, rms_Ax, rms_Ay, rms_Az
23	mean_Ax, mean_Ay, mean_Az, acc
24	mean_Ax, mean_Ay, mean_Az
25	std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az, acc
26	std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az
27	std_Ax, std_Ay, std_Az, acc
28	std_Ax, std_Ay, std_Az
29	rms_Ax, rms_Ay, rms_Az, acc
30	rms_Ax, rms_Ay, rms_Az
31	acc

5.2 KNN

O algoritmo 1NN obteve maior acurácia para todos os subconjuntos de características (Figura 3). A Figura 4 mostra a acurácia para todos os subconjuntos: a melhor acurácia (98.92) foi do subconjunto #4 (diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az).

Tabela 3: Acurácia online dos classificadores (usuário jovem)

Atividade	Acurácia		
	1NN (%)	3NN (%)	5NN (%)
Sentar	100.0	100.0	100.0
Ficar em pé	96.7	96.7	100.0
Andar	70.0	66.7	40.0
Descer escada	11.1	40.7	46.9
Subir escada	20.5	30.3	14.3
Média	59.7	66.9	60.2

5.3 Online activity recognition

Para medir a acurácia do classificador no reconhecimento online de atividades, foi desenvolvida uma extensão (Figura 5) para um aplicativo de coleta de dados de Zimmermann [12]. Considerando os resultados reportados na seção anterior, na extensão foi implementado o algoritmo KNN com as características diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay e std_Az (subconjunto #4) para classificar as atividades de sentar, ficar em pé, andar, descer escada e subir escada. Foi utilizado um *smartphone Motorola Moto G6 Play*.

Um usuário (um dos autores, aluno de graduação de 22 anos) utilizou um *smartwatch LG Urbane com Android Wear 2.0* para realizar cada uma das atividades por 1 minuto. Como a janela de tempo é de 2 segundos, tivemos 30 classificações feitas por atividade e o desempenho obtido está expresso na Tabela 3 para n=1, 3 e 5.

A Tabela 3 indica que os três classificadores tiveram acurácia alta nas atividades de sentar e ficar em pé enquanto que o 3NN teve o melhor desempenho médio. As atividades de descer e subir escadas tiveram acurácia baixa: o classificador recorrentemente reconheceu essas atividades como “andar”. Essa confusão pode estar associada à combinação do posicionamento do *smartwatch*, no pulso, com o movimento dos braços (*up and down*) durante essas atividades.

Comparados à acurácia reportada na seção anterior, a média de acerto caiu de 98.92% para 66.9%. Uma das razões para essa variação é que o teste foi realizado com um usuário de perfil diferente (jovem universitário) dos que compõem a EARDB (32 idosos de 60 a 75 anos). Assim, não é possível interpretar esse resultado como um indicativo da precisão do classificador. Por outro lado, os resultados indicam ser possível realizar no *smartphone* o reconhecimento online das atividades a partir dos dados coletados no *smartwatch*.

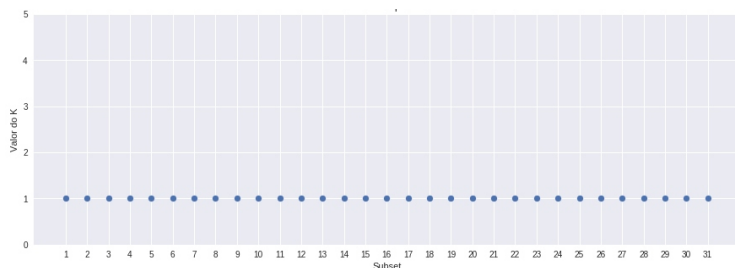


Figura 3: Melhor valor de K para subconjuntos de características.

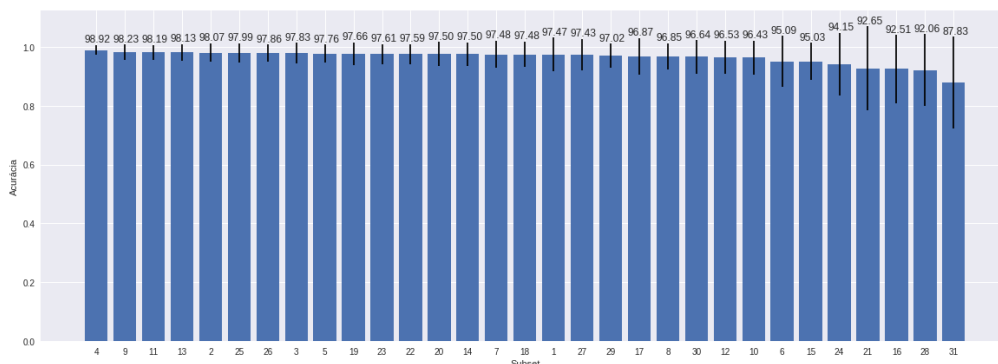


Figura 4: Acurácia nos subconjuntos de características com KNN.

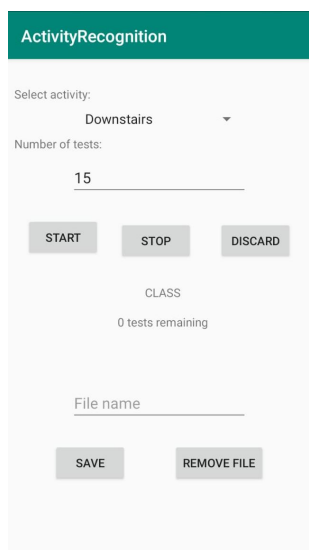


Figura 5: Tela do protótipo para testes dos classificadores

6 CONSIDERAÇÕES FINAIS

O trabalho investigou o potencial do uso de dados capturados do acelerômetro de um smartwatch e processados em um smartphone para o reconhecimento *online* de atividades realizadas por idosos. Algoritmos de *machine learning* executados *a posteriori* resultaram em acurácia de reconhecimento de 98.92%. Os algoritmos foram implementados em uma *smartphone app* e um estudo permitiu realizar, no *smartphone*, o reconhecimento *online* das atividades a partir dos dados capturados no *smartwatch*. Esses resultados abrem portas para trabalhos futuros: o próximo passo é avaliar a acurácia do reconhecimento *online* com usuários idosos.

Agradecimentos: Agradecemos pelo financiamento: Santana (CNPq-PIBIC); Rasteiro (PUB-USP); Zimmermann (MSc CAPES); De Nardin (Programa Permanente de Capacitação Docente PUC Minas); Pimentel (CNPq-PQ; FAPESP 16/00351-6 e 16/50489-4).

REFERÊNCIAS

- [1] Capela, N. A., Lemaire, E. D., and Baddour, N. (2015). Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients. *PLoS one*, 10(4):e0124414.
- [2] Davoudi, A. et al. (2019). Accuracy of Samsung Gear S Smartwatch for Activity Recognition: Validation Study. *JMIR Mhealth Uhealth*, 7(2):e11270.
- [3] De-La-Hoz-Franco, E. et al. (2018). Sensor-based datasets for human activity recognition: A systematic review of literature. *IEEE Access*, 6:59192–59210.
- [4] Dorn, D., Gorzelitz, J., Gangnon, R., Bell, D., Koltyn, K., and Cadmus-Bertram, L. (2019). Automatic identification of physical activity type and duration by wearable activity trackers: A validation study. *JMIR Mhealth Uhealth*, 7(5):e13547.
- [5] Eun, S.-J., Whangbo, T.-K., Park, D. K., and Kim, K.-H. (2017). Development of personalized urination recognition technology using smart bands. *International neurology journal*, 21(Suppl 1):S76.
- [6] Galán-Mercant, A. and Cuesta-Vargas, A. I. (2013). Differences in trunk accelerometry between frail and nonfrail elderly persons in sit-to-stand and stand-to-sit transitions based on a mobile inertial sensor. *JMIR Mhealth Uhealth*, 1(2):e21.
- [7] Jansi, R. and Amutha, R. (2019). Sparse representation based classification scheme for human activity recognition using smartphones. *Multimedia Tools Appl.*, 78(8):11027–11045.
- [8] Lara, O. D. and Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209.
- [9] Lentzas, A. and Vrakas, D. (2019). Non-intrusive human activity recognition and abnormal behavior detection on elderly people: a review. *Artificial Intelligence Review*.
- [10] Yang, Y., Hirdes, J., et al. (2019). Fall risk classification in community-dwelling older adults using a smart wrist-worn device and the resident assessment instrument-home care: Prospective observational study. *JMIR Aging*, 2(1):e12153.
- [11] Zhang, M. and Sawchuk, A. A. (2011). A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 92–98. ICST.
- [12] Zimmermann, L. C. (2019). *Reconhecimento de atividades de pessoas idosas com integração de sensores de smartphones e dispositivos vestíveis*. Master's thesis, PPg-CCMC, USP/São Carlos, Brasil <https://doi.org/10.11606/D.55.2019.tde-10062019-080049>.
- [13] Zimmermann, L. C., Rodrigues, K. R., and Pimentel, M. G. C. (2019). EPARS: Elderly Physical Activity Reminder System Using Smartphone and Wearable Sensors. <http://doi.acm.org/10.1145/3341162.3350845>. In *Adjunct Proceedings UbiComp/ISWC'19*, pages 1139–1145.