

Application of Machine Learning Techniques in Human Activity Recognition using Wearable and Mobile Devices

Flavio Vinicius Vieira Santana
Universidade de São Paulo
São Carlos - SP, Brasil
flavio.santana@usp.br

Bruno Henrique Rasteiro
Universidade de São Paulo
São Carlos - SP, Brasil
bruno.rasteiro@usp.br

ABSTRACT

In this paper, the use of wearable devices and machine learning algorithms is proposed for human activity recognition. In particular, this research is focused on the elderly public. Initially, a database was built from activities executed by the elderly. The data was collected by the sensors of smartwatches and smartphones. From this database, a classifier was built using machine learning algorithms to recognize the activities. Finally, some tests were made in a real scenario.

KEYWORDS

Human Activity Recognition, wearable, mobile, machine learning

RESUMO

Neste artigo é proposto o uso de dispositivos vestíveis e algoritmos de aprendizado de máquina para o reconhecimento de atividades humanas. Particularmente, essa pesquisa tem o foco no público idoso. Inicialmente, foi obtida uma base de dados de atividades executadas por idosos. Esses dados foram coletados pelos sensores de smartwatches e smartphones. A partir dessa base, utilizando aprendizado de máquina foi construído um classificador para reconhecer essas atividades. Então, foram feitos alguns testes em cenário real.

1 INTRODUÇÃO

Reconhecer atividades humanas (HAR - *Human Activity Recognition*) e identificar padrões (ou quebra de padrões) comportamentais de indivíduos é um campo de pesquisa bastante promissor e desafiador para pesquisadores e pode ser aplicada a diversas áreas de conhecimento – setor militar, segurança, saúde, entretenimento, entre outras [4]. Neste contexto, o método ESPIM - *Experience Sampling and Programmed Intervention Method* [5] propõe a coleta de informações, de forma pervasiva através da utilização de intervenções programadas e/ou de sensores móveis vestíveis, de pessoas em seu contexto natural combinando técnicas de ESM - *Experience Sampling Method* [2].

A maioria dos *smartwatches* modernos já são dotados de sensores como acelerômetro, giroscópio, microfone e sensor de batimentos cardíacos, e permitem coletar diversas informações de forma totalmente transparente ao usuário e então utilizá-las para diferentes objetivos. No contexto do reconhecimento de atividades

humanas, este trabalho tem por objetivo investigar a utilização dos *smartwatches* como dispositivos vestíveis para o reconhecimento de atividades humanas simples [3], utilizando recursos já existentes no *smartwatch* e coletando os dados do acelerômetro para que, em seguida, tais dados sejam enviados a um *smartphone*. No *smartphone* foi implementado um sistema utilizando algoritmos de aprendizado de máquina que tem a capacidade de identificar qual atividade está sendo realizada por aquele indivíduo em tempo real.

Este trabalho está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados, a Seção 3 descreve a metodologia utilizada para definição das características a serem utilizadas no reconhecimento, a Seção 4 mostra as características que foram selecionadas, a Seção 5 introduz os resultados dos testes com cada classificador e, por fim, a Seção 6 apresenta as considerações finais.

2 TRABALHOS RELACIONADOS

O componente de reconhecimento de atividades foi desenvolvido como um aplicativo à parte e está em fase de integração com um aplicativo chamado SENSEM.

Como um aplicativo de agendamento e execução de tarefas, o SENSEM é capaz de suportar a implementação do método ESPIM - *Experience Sampling and Programmed Intervention Method* [5], que propõe a coleta de informações de forma pervasiva e ubíqua com o uso de intervenções programadas e/ou de sensores móveis vestíveis, de pessoas em seu contexto natural combinando técnicas de ESM - *Experience Sampling Method* [2].

Outro trabalho relacionado é o trabalho de qualificação de mestrado da Larissa Cardoso Zimmermann¹ que fez a captação de dados das atividades executadas por idosos, conforme a aprovação do Comitê de Ética em Pesquisa da USP, conforme registro na Plataforma Brasil (CAAE: 57875016.3.0000.5390). No mesmo trabalho foi desenvolvidos um conjunto de aplicativos que permitem a coleta de dados para a criação de uma base e o reconhecimento de atividades humanas com o uso de *smartwatch* e *smartphone*.

Em [1] foi feito um estudo para selecionar as melhores características da literatura em aplicações HAR utilizando dispositivos vestíveis, a análise foi feita de forma a não levar em consideração qual classificador será utilizado na aplicação. O estudo considerou 76 características e utilizou um *smartphone* preso a cintura dos envolvidos. A análise foi feita com 3 métodos de seleção de características (*Relief F*, *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

In: XVI Workshop de Trabalhos de Iniciação Científica (WTIC 2019), Rio de Janeiro, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2019.
ISSN 2596-1683

¹<http://lattes.cnpq.br/4945723091500656>

em média 10 características, entretanto, vale ressaltar que o artigo se diferencia deste trabalho principalmente pelo fato de utilizar um *smartphone* preso a cintura e não um *smartwatch* no pulso.

No trabalho [6] foi feito um estudo para selecionar características de dois tipos, 87 características estatísticas e 23 características físicas. A população utilizada é formada por 6 indivíduos de diferentes gêneros, idade, altura e peso. O dispositivo utilizado chama-se *NotionMode* e foi fixado na parte frontal do quadril dos indivíduos, a coleta dos dados foi feita a 100Hz e foi utilizado uma janela de tempo de 2 segundos com 50% de sobreposição. A seleção das características foi feita através de 3 métodos, *Relief F*, *SFC (Wrapper Method based on Single Feature Classification)* e *SFS (Wrapper Method based on Sequential Forward Selection)*. Utilizando apenas as 5 primeiras características estatísticas do método SFS, o classificador SVM com *kernel* linear ficou com um erro de 12%. O artigo mostrou bons resultados ao conseguir uma taxa de erro baixa com apenas 5 características, entretanto ele se diferencia deste trabalho pelo tipo de aparelho, posição, população, frequência de captura e janela de tempo.

3 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, sendo eles: aceleração no eixo X (Ax), aceleração no eixo Y (Ay), aceleração no eixo Z (Az) e a aceleração linear (Acc).

Tendo 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. Nesse trabalho, foi utilizado uma janela de 2 segundos, isso significa que a cada 32 instâncias de dados coletados foi calculada e definida uma instância com N características. No caso da *feature* 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, é garantido que instâncias de indivíduos ou classes distintas não se misturam. 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 a partir dos dados coletados são descritas detalhadamente na Tabela 1.

4 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 features de média ficassem sempre juntas, bem como desvio padrão e RMS. Dessa forma, obteve-se 31 subconjuntos de características. Os subconjuntos foram enumerados a fim de facilitar sua citação e a Tabela 2 exibe os subconjuntos criados e seus respectivos números.

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

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 (Ay - Az - Ax)

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 *subset* definido, o algoritmo foi executado com o valor de K variando de 1 a 10 e o que obteve melhor acurácia foi escolhido como K padrão. Os resultados obtidos são descritos detalhadamente na seção a seguir.

5 RESULTADOS

Esta seção exibe os resultados obtidos na seleção dos subconjuntos de características nos dois classificadores avaliados. Além disso, apresenta o protótipo desenvolvido para realização dos testes práticos, bem como os resultados obtidos.

5.1 Árvore de decisão

Os resultados obtidos utilizando o algoritmo de Árvore de Decisão podem ser observados na Figura 1 que apresenta no eixo x cada um dos subconjuntos de características apresentados na Tabela 2. Cada experimento foi repetido 10 vezes utilizando validação cruzada e, os traços verticais no topo das barras indicam o desvio padrão. Através do gráfico, pode-se observar que os diferentes subconjuntos tiveram pouca influência no desempenho do classificador. A melhor acurácia obtida foi de 98.33, utilizando as características mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay e std_Az.

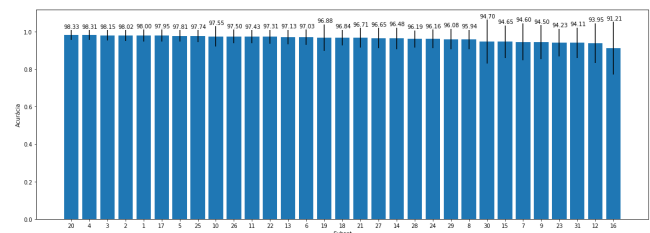


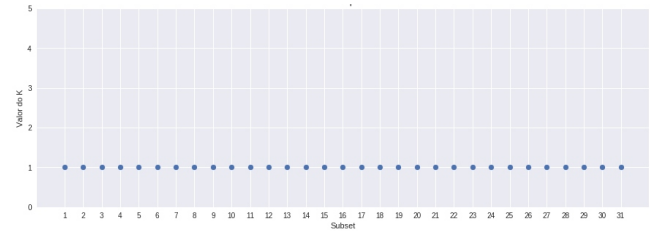
Figura 1: Acurácia nos subconjuntos de características utilizando Árvore de Decisão

Tabela 2: Subconjuntos de características estatísticas

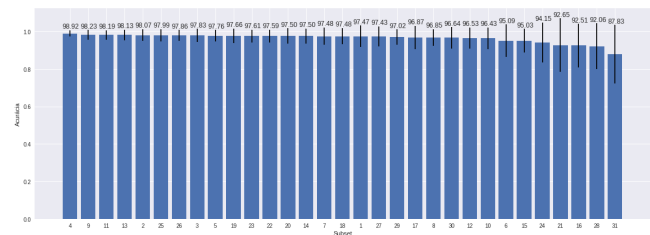
Subconjunto #	características
1	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az, acc
2	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, rms_Ax, rms_Ay, rms_Az
3	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay, std_Az, acc
4	diff, mean_Ax, mean_Ay, mean_Az, std_Ax, 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

A Figura 2 mostra o melhor valor de K para todos os subconjuntos de características, é interessante notar que para todos os subconjuntos o K que obteve maior acurácia foi sempre 1.

**Figura 2: Melhor valor de K para os subconjuntos de características**

O subconjunto que obteve o melhor resultado no KNN foi o das variáveis diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay e std_Az. A acurácia obtida foi 98.92, esse resultado pode ser observado na Imagem 3.

**Figura 3: Acurácia nos subconjuntos de características utilizando o KNN**

5.3 Testes práticos

Após os resultados obtidos com experimentos teóricos foram então realizados testes práticos com o objetivo de medir a acurácia do classificador no ambiente real. Para isso, uma extensão do aplicativo foi desenvolvida e tem a aparência mostrada na Imagem 4.

Utilizando-se como parâmetro os testes preliminares, foi definido que o algoritmo KNN seria utilizado com as características diff, mean_Ax, mean_Ay, mean_Az, std_Ax, std_Ay e std_Az para classificar as atividades de sentar, ficar em pé, andar, descer escada e subir escada.

Cada uma dessas atividades foi executada por 1 minuto por um de nossos pesquisadores. Como a janela de tempo é de 2 segundos, tivemos 30 classificações feitas por atividade e o desempenho obtido está expresso na Tabela 3.

Com base nos dados mostrados na Tabela 3 é possível perceber que os três classificadores tiveram acurácia alta nas atividades de sentar e ficar em pé enquanto que o 3NN teve o melhor desempenho médio. Já as atividades de descer e subir escadas tiveram acurácia baixa e, foi percebido que, na maioria das vezes, o classificador reconhecia essas atividades como "andar".

Figura 4: Tela do protótipo desenvolvido para os testes dos classificadores

Tabela 3: Acurácia dos classificadores na prática

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

Ao comparar à acurácia obtida antes dos testes práticos, vemos que 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 do perfil dos usuários que geraram a base, uma vez que, a captação de dados para a base foi feita em idosos e o teste foi realizado por um jovem. Além disso, os testes, até o momento, foram feitos com apenas um indivíduo. Portanto só é possível interpretar esses resultados como um indicativo da precisão do classificador.

6 CONSIDERAÇÕES FINAIS

Este trabalho abordou o reconhecimento de atividades humanas com o uso de dispositivos vestíveis e móveis. Para isso, inicialmente, uma base de dados foi gerada através da coleta de dados provenientes de diversos sensores enquanto idosos executavam atividades pré definidas. A partir desses dados foram analisadas as melhores características e o melhor classificador para o conjunto de atividades. Então, o classificador foi testado em um cenário real, onde foi

novamente analisada sua acurácia para cada uma das atividades existentes na base.

Os resultados teóricos obtidos foram, consideravelmente, altos obtendo-se uma acurácia média máxima de 98.92% atingida pelo classificador 1NN. Já em um cenário real, o classificador 3NN apresentou o melhor resultado para as mesmas características escolhidas, com uma acurácia média de 66.9%, sendo que, os melhores resultados foram obtidos para as atividades sentar e ficar em pé e um resultado não satisfatório para as atividades de descer e subir escadas.

A partir dos resultados obtidos é possível concluir que novos testes precisam ser realizados para que haja maior confiabilidade nos resultados. Além disso, uma nova vertente a ser investigada é com relação ao desempenho do classificador quando testado com idosos ao invés de jovens, permitindo estabelecer comparações do quanto a base de dados exerce influência no poder do classificador.

AGRADECIMENTOS

Agradecemos à Larissa Cardoso Zimmermann e à Luciana de Nardin² por toda a colaboração durante o projeto de pesquisa e na escrita deste artigo. Também agradecemos à nossa orientadora Maria da Graça Campos Pimentel pela orientação e a oportunidade de participar deste projeto.

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] Csikszentmihalyi, Mihaly, R. L. and Prescott, S. (2017). "The ecology of adolescent activity and experience". *Journal of youth and adolescence* 6:3: 281-294.
- [3] 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.
- [4] 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.
- [5] Zaine, Isabela, e. a. (2016). "ESPIM: An Ubiquitous Data Collection and Programmed Intervention System using ESM and Mobile Devices". *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*.
- [6] 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 (Institute for Computer Sciences, Social-Informatics and ...)

²<http://lattes.cnpq.br/2149904407727510>