Fraunhofer IAIS

UNIVERSIDAD EAFIT

Recognition of Multi-Level Human Activities by Implementing an "Activity Recognition Chain" on a Benchmark Dataset

Agustín Vargas Toro

Medellin, July 31 2018

Problem definition: Human Activity Recognition (HAR)

Motivation

Figure: Ubiquitous sensing via *wearable* sensors[1].

Human Activity Recognition (HAR) [1]

- Active and assisted living systems for smart homes.
- Indoor and outdoor surveillance systems.
- Healthcare monitoring applications.
- Tele-immersion and interactive applications.

[1] Image available in: *Develop Wearable Devices on the IoT Cutting Edge*, **Electronic Design.**
URL: `https://www.electronicdesign.com/iot/develop-wearable-devices-iot-cutting-edge`

The REACH project

Figure: Example of a REACH use case [2].

Objectives

General Objective

Develop a Human Activity Recognition system that recognizes locomotion, low, and middle level activities from a pool of activities contained in a public benchmark dataset.

Specific Objectives (summarized)

- Select a publicly available dataset that can be used for HAR.
- Implement a method of missing data imputation.
- Implement a method of segmentation of time series.
- Extract a collection of features from time series data and implement a methodology of feature selection.
- Implement a classifier and evaluate the performance of the complete system on the selected dataset.

The Activity Recognition Chain (ARC) [3], [4]

Raw data (time series)

ARC: Training and testing stages

Training set

Data used for modeling a classifier.

Testing set

Data used only for evaluating the performance of the classifier.

The separation into training/testing stages allows to understand and quantify how well does a model perform and how well it might perform in the future with *yet-to-be-seen* data [5].

Preprocessing

What is?

The sensor recordings may undergo several data transformation processes oriented to [6]:

- Eliminate redundant information.
- Improve the quality of the data.

- Artifact removal.
- Data cleaning.
- Missing value imputation.
- Synchronization and resampling.
- Calculation of new predictors [7]:

$$acc_{Mag} = \sqrt{acc_x^2 + acc_y^2 + acc_z^2}$$

Segmentation

Segmentation: review

Window Length	Overlap	References
500 ms	50%	[8]
24 samples	50%	[9]
1 s	0%	[10],[4], [11]
1 s	50%	[12]
2 s	50%	[13], [14], [15]
2.56 s	50%	[16], [17]
3 s	66%	[7]
5 s	50%	[18], [19]
5.12 s	4.12 s	[20]
6 s	0%	[21]
6.7 s	50%	[22]
12 s - 20 s	50%	[19]

Feature extraction

Classification: supervised modeling

Classifier	Reference
Support Vector Machine	[1, 2, 4, 7, 11, 14]
K-nearest neighbors	[1, 5, 7, 8, 10, 11, 13, 14]
Random Forest	[2, 3, 15]
Naïve Bayes	[4, 7, 10, 13, 14]
Discriminative Analysis	[4]
Hidden Markov Models	[5, 9]
AdaBoost	[6, 11]
Decision Tree	[7 - 10, 12 - 15]

Classification: decision trees

Figure: A decision tree of 3 regions. [23]

Figure: A more complex decision tree (6 regions) [23]

$$Gini = \sum_{k=1}^{K} p_{mk}(1 - p_{mk})$$

Classification: decision trees

Figure: A partitioned feature space. [23]

Figure: Tree plot based in the partitions. [23]

Classification: random forests

- Combines knowledge from several decision trees to make a prediction.
- Uses **bootstraped** training datasets to create the decision trees.

Advantages

✓ Reduces variance by decorrelating the trees.

✓ The *out-of-bag* error is an estimate of the test error.

✓ Has a measure to estimate the importance of features.

Limitations

✗ Reduces interpretability as the 'tree' graphs cannot be used.

Implementation in R

The OPPORTUNITY Dataset [8]: generalities

This dataset stood out between 7 reviewed datasets in terms of:

- Data quality.
- Rigorously labeling of data.
- Sensor arrangement.
- Activities of interest for REACH.

The dataset is composed of:

- 24 experimental runs from 4 subjects (6 runs per subject).
- An collection of data from 19 sensors.
- Data from a **mourning routine** experiment.

The OPPORTUNITY Dataset [8]: generalities

The OPPORTUNITY Dataset [8]: sensor configuration

(a) Inertial Measurement Units [8]

(b) Triaxial accelerometers [8]

The OPPORTUNITY Dataset [8]: activities

- **Modes of locomotion:** stand, walk, sit, and lie.
- **Middle level actions:** open door, close door, open fridge, close fridge, open dishwasher, close dishwasher, open drawer, close drawer, clean table, drink from cup, and toggle switch.
- **High level activities:** relaxing, coffee time, early morning, cleanup, and sandwich time).

Performance measures

	Target activity is NOT being performed	Target activity is being performed
Classifier predicts: target activity is NOT happening	True Negatives (TN)	False Negatives (FN)
Classifier predicts: target activity is happening	False Positives (FP)	True Positives (TP)

The prediction for each segment is compared with it's respective ground truth label

Performance measures: F-measure

- **Recall:** measures the proportion of elements belonging to the target class that were correctly classified:

$$Recall = \frac{TP}{FN + TP}$$

- **Precision:** measures how often the classifier is correct when it predicts that a feature vector belongs to the target class:

$$Precision = \frac{TP}{TP + FP}$$

- The F-measure (also named F1-score) takes into account the precision and recall for each class and can provide a better assessment of performance than computing the accuracy

$$F_1 = \sum_i 2 * w_i \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i}$$

The proposed ARCs: summary

ARC step	Comments
Data	OPPORTUNITY challenge
Target activities	Locomotion and hand gestures
NA imputation	Linear interpolation
Magnitude	Yes
Window length	0.5s, 1.5s, 3s
Window overlap	0%, 50%, 66%
Features	Mean, SD, median, IQR, ZCR, MAD, RMS, max, min, energy
Classifier	Random forest
Performance	F-measure

Finding an optimal ARC

Benefits of a random forest: reduction of error

Benefits of a random forest: feature selection

Experiments

OPPORTUNITY challenge	User-dependent	User-independent
Training: 14 runs from 3 subjects Test: 4 runs from 2 of the 3 previous subjects	One model per subject. 4 runs used as training and 2 as testing.	One model per non-considered subject. 3 subjects used as training and the remaining as testing.

Results: challenge

Method	F - measure	
	Locomotion	Gestures
LDA	0.59	0.69
QDA	0.68	0.53
NCC	0.53	0.51
1-NN	0.84	0.87
3-NN	0.85	0.85
UP	0.60	0.64
NStar	0.61	0.84
SStar	0.64	0.86
CStar	0.63	0.88
This work (no FS)	0.87	0.88
This work (FS)	0.87	0.90

Results: user dependent (locomotion)

Method	F-measure			
	Subject 1	Subject 2	Subject 3	Subject 4
LDA	0.62	0.64	0.68	0.43
QDA	0.67	0.66	0.71	0.45
NCC	0.60	0.58	0.56	0.45
1-NN	0.84	0.85	0.83	0.76
3-NN	0.85	0.86	0.83	0.77
NStar	-	0.58	0.66	-
SStar	-	0.61	0.68	-
CStar	-	0.60	0.65	-
MI	-	0.85	0.81	-
Zhu [7]	0.88	0.88	0.80	0.85
This work	0.90	0.91	0.81	0.88

Results: user dependent (gestures)

Method	F-measure			
	Subject 1	Subject 2	Subject 3	Subject 4
LDA	0.65	0.63	0.70	0.62
QDA	0.60	0.57	0.69	0.64
NCC	0.48	0.48	0.51	0.35
1-NN	0.85	0.89	0.86	0.84
3-NN	0.85	0.89	0.86	0.88
NStar	-	0.84	0.83	-
SStar	-	0.87	0.84	-
CStar	-	0.88	0.87	-
NAGS	-	-	-	0.71
Zhu	0.87	0.89	0.87	0.85
This work	0.84	0.91	0.78	0.84

Results: user independent

Test set	Locomotion	Hand gestures
Subject 1	0.81	0.78
Subject 2	0.82	0.68
Subject 3	0.74	0.62
Subject 4	0.78	0.78
Overall	0.79	0.72

Use case: drinking

HAR-specific metrics introduced by Ward et. al [24]

Assigns extended categories to the errors made by the classifier (*i.e.,* false negatives and false positives).

Use case: drinking

HAR-specific metrics introduced by Ward et. al [24]

Assigns extended categories to the errors made by the classifier (*i.e.*, false negatives and false positives).

Use case: drinking ($w_{len} = 0.5s$)

- This ARC used a window length of 0.5s.
- The biggest causes of error are underfill and fragmentation.

Use case: drinking ($w_{len} = 6s$)

- This ARC used a window length of 6s.
- ✓ Fragmentation errors are reduced from 11.6% (windows) to 0% and from 42.9% (events) to 0%.
- ✓ Underfill reduced from 20.4% to 7.7% (windows).
- ✗ Deletion rate is increased from 8% to 35.4% (windows) and from 11.8% to 35.3%.

Discussion

- There is a visible trade-off between underfill/fragmentation errors and deletion errors when choosing a window length for the system.
- Target activities should be grouped by: relevant sensors and average duration.
- Extended metrics give more information about the type of errors and help in the decision of selecting the best-performing classifier but require a more detailed analysis to achieve a fully understanding of the performance.

Conclusions I

1. An activity recognition methodology (*i.e.,* Activity Recognition Chain, ARC) was implemented to recognize human activities using sensor data [4].
2. The linear interpolation method gave the best results for the trade-off of computing speed and performance in missing value imputation.
3. Nine ARCs were implemented to test the influence of the ARC parameters on the system's performance (best results: overlap 66% and window length 0.5s).
4. Ten features were calculated for each sensor measurement, yielding to a total of 1220 features.
5. Recursive feature selection improved the results by selecting the best 20% of the features (244).

Conclusions II

6. The performance of the implemented ARCs surpassed the baselines of the OPPORTUNITY challenge.
7. User-dependent experiments showed that this methodology achieves state-of-the-art results when recognizing locomotion activities.
8. The user-independent scenario exhibited an overall decrease in the system's performance.
9. Drinking events are complex and can be addressed by rule-based classifiers.

Future Work

1. (Ongoing) Create an 'R package' dedicated for HAR.
2. Explore methods that cope with the user-independent HAR challenge.
3. Test the performance of the ARC with other classifiers, as well as enhancements to the implemented classifier (*e.g.,* boosting).
4. Extend the methodology of activity recognition to deep learning approaches [25].
5. Construct *rule-based* classifiers that take as input the predictions from the ARC [26].
6. (Ongoing and proposed by the directives of the REACH project) Construct data collection experiments that address target activities within the REACH scenario and test the performance of the proposed methodology.

Acknowledgments

- To supervisors, Dr. Sebastian Konietzny and Dr. Alejandro Marulanda Tobón who made this work possible.
- To my parents who educated me and ‘gave me wings’.
- To Universidad EAFIT and DAAD who opened to me the doors of the world.
- To Fraunhofer IAIS who became my home during the past year.
- To all of you for assisting today.

References I

Ranasinghe, Suneth, Fadi Al MacHot, and Heinrich C. Mayr (2016). "A review on applications of activity recognition systems with regard to performance and evaluation". In: *International Journal of Distributed Sensor Networks* 12.8. ISSN: 15501477. DOI: 10.1177/1550147716665520.

Konietzny, Sebastian, Agustín Vargas Toro, Pearl Pu, Igor Kulev, Mirana Randriambelonoro, Caroline Perrin, Thomas Linner, Amir Kabouteh, and Henning Andersen (2018). *REACH - Deliverable D3.1: Data collection requirements, ethnographic studies, and a validated, significant set of patient data*. Tech. rep. URL: http://reach2020.eu/?page{_}id=1190.

References II

- Roggen, Daniel, Gerhard Tröster, Mangenat Stéphane, and Markus Waibel (2011). “Designing and sharing activity recognition systems across platforms: methods from wearable computing”. In: *Stephane.Magnenat.Net*. DOI: 10.3929/ethz-a-010001702.
- Bulling, Andreas, U Blanke, and Bernt Schiele (2014). “A tutorial on human activity recognition using body-worn inertial sensors”. In: *ACM Computing Surveys (CSUR)* 46.June, pp. 1–33. ISSN: 03600300. DOI: http://dx.doi.org/10.1145/2499621. URL: http://dl.acm.org/citation.cfm?id=2499621.

References III

Kuhn, Max and Kjell Johnson (2013). *Applied Predictive Modeling*. Springer, p. 620. ISBN: 1461468485. DOI: 10.1007/978-1-4614-6849-3. URL: http://www.amazon.com/Applied-Predictive-Modeling-Max-Kuhn/dp/1461468485/ref=pd{_}bxgy{_}b{_}img{_}z.

Kotsiantis, S B, D Kanellopoulos, and P E Pintelas (2006). "Data preprocessing for supervised learning". In: *International Journal of Computer Science* 1.2, pp. 111–117. ISSN: 1306-4428. DOI: 10.1080/02331931003692557. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.5127{\&}rep=rep1{\&}type=pdf.

References IV

- Zhu, Jiadong, Ruben San-Segundo, and Jose M Pardo (2017). "Feature extraction for robust physical activity recognition". In: *Human-centric Computing and Information Sciences* 7, pp. 1–16. ISSN: 2192-1962. DOI: 10.1186/s13673-017-0097-2.
- Chavarriaga, Ricardo, Hesam Sagha, Alberto Calatroni, Sundara Tejaswi Digumarti, Gerhard Tröster, José Del R Millán, and Daniel Roggen (2013). "The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition". In: *Pattern Recognition Letters* 34.15, pp. 2033–2042. ISSN: 01678655. DOI: 10.1016/j.patrec.2012.12.014. URL: http://dx.doi.org/10.1016/j.patrec.2012.12.014.

References V

- Plötz, Thomas, Nils Y Hammerla, and Patrick Olivier (2011). "Feature Learning for Activity Recognition in Ubiquitous Computing". In: *Proceeding IJCAI'11 Proceedings of the Twenty-Second international joint conference on Artificial Intelligence* Volume 2, pp. 1729–1734. ISSN: 10450823. DOI: 10.5591/978-1-57735-516-8/IJCAI11-290. URL: http://dl.acm.org/citation.cfm?id=2283516.2283683.
- Bannach, David and Paul Lukowicz (2008). "Rapid prototyping of activity recognition applications". In: *IEEE Pervasive Computing*. DOI: 10.1109/MPRV.2008.36.

References VI

- Karantonis, Dean M, Michael R Narayanan, Merryn Mathie, Nigel H Lovell, and Branko G Celler (2006). “Implementation of a Real-Time Human Movement Classifier Using a Triaxial Accelerometer for Ambulatory Monitoring”. In: *IEEE Transactions on information technology on biomedicine* 10.1, pp. 156–167.
- Liu, Kai-chun, Chien-yi Yen, Li-han Chang, Chia-yeh Hsieh, and Chia-tai Chan (2017). “Wearable Sensor-Based Activity Recognition for Housekeeping Task”. In: *IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, pp. 67–70. ISBN: 9781509062447. URL: http://ieeexplore.ieee.org/document/7936009/.

References VII

Stisen, Allan, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen (2015). "Smart Devices are Different: Assessing and MitigatingMobile Sensing Heterogeneities for Activity Recognition". In: *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pp. 127–140. DOI: 10.1145/2809695.2809718. URL: http://dl.acm.org/citation.cfm?doid=2809695.2809718.

Zhao, Zhongtang, Zhenyu Chen, Yiqiang Chen, Shuangquan Wang, and Hongan Wang (2014). "A Class Incremental Extreme Learning Machine for Activity Recognition". In: *Cognitive Computation* 6.3, pp. 423–431. ISSN: 18669964. DOI: 10.1007/s12559-014-9259-y.

References VIII

- Wen, Jiahui and Zhiying Wang (2017). "Learning general model for activity recognition with limited labelled data". In: *Expert Systems with Applications* 74, pp. 19–28. ISSN: 09574174. DOI: 10.1016/j.eswa.2017.01.002. URL: http://dx.doi.org/10.1016/j.eswa.2017.01.002.
- Kavitha, R. and N. J. Sijo Antony (2017). "Human Activity Recognition from Sensor data using Random Forest Algorithm". In: *International Journal of Advanced Research in Computer Science* 8.3, pp. 2015–2018. URL: http://www.ijarcs.info/index.php/Ijarcs/article/view/3009.

References IX

Anguita, Davide, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L Reyes-Ortiz (2013). "A Public Domain Dataset for Human Activity Recognition Using Smartphones". In: *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* April, pp. 24–26. URL: http://www.i6doc.com/en/livre/?GCOI=28001100131010.

Leutheuser, Heike, Dominik Schuldhaus, and Bjoern M. Eskofier (2013). "Hierarchical, Multi-Sensor Based Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms Using a Benchmark Dataset". In: *PLoS ONE* 8.10. ISSN: 19326203. DOI: 10.1371/journal.pone.0075196.

References X

- Lara, Óscar D., Alfredo J. Pérez, Miguel A. Labrador, and José D. Posada (2012). “Centinela: A human activity recognition system based on acceleration and vital sign data”. In: *Pervasive and Mobile Computing* 8.5, pp. 717–729. ISSN: 15741192. DOI: 10.1016/j.pmcj.2011.06.004. URL: http://dx.doi.org/10.1016/j.pmcj.2011.06.004.
- Reiss, Attila and Didier Stricker (2012). “Creating and benchmarking a new dataset for physical activity monitoring”. In: *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments* February, p. 1. ISSN: 15504816. DOI: 10.1145/2413097.2413148. URL: http://dl.acm.org/citation.cfm?doid=2413097.2413148.

References XI

Baños, Oresti, Miguel Damas, Héctor Pomares, Ignacio Rojas, Máté Attila Tóth, and Oliver Amft (2012). "A benchmark dataset to evaluate sensor displacement in activity recognition". In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, p. 1026. ISSN: 1617-4909. DOI: 10.1145/2370216.2370437. URL: http://dl.acm.org/citation.cfm?doid=2370216.2370437.

Bao, Ling and Stephen S. Intille (2004). "Activity Recognition from User-Annotated Acceleration Data". In: pp. 1–17. ISSN: 03029743. DOI: 10.1007/978-3-540-24646-6_1. arXiv: 9780201398298. URL: http://link.springer.com/10.1007/978-3-540-24646-6{_}1.

References XII

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani (2013). *An Introduction to Statistical Learning*. P. 618. ISBN: 9780387781884. DOI: 10.1016/j.peva.2007.06.006. arXiv: arXiv:1011.1669v3. URL: http://books.google.com/books?id=9tv0taI8l6YC.

Ward, Jamie A, Paul Lukowicz, and Hans-Werner Gellersen (2011). "Performance metrics for activity recognition". In: *Acm Tist* 2.1, pp. 111–132.

Wang, Jindong, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu (2018). "Deep learning for sensor-based activity recognition: A Survey". In: *Pattern Recognition Letters* 0, pp. 1–9. ISSN: 01678655. DOI: 10.1016/j.patrec.2018.02.010. arXiv: 1707.03502. URL: https://doi.org/10.1016/j.patrec.2018.02.010.

References XIII

Liu, Ye, Liqiang Nie, Li Liu, and David S. Rosenblum (2016). "From action to activity: Sensor-based activity recognition". In: *Neurocomputing* 181, pp. 108–115. ISSN: 18728286. DOI: 10.1016/j.neucom.2015.08.096. URL: http://dx.doi.org/10.1016/j.neucom.2015.08.096.