Personalized Models in Human Activity Recognition using Deep Learning

Hamza Amrani, Daniela Micucci,and Paolo Napoletano

Introduction

HAR

- Human Activity Recognition (**HAR**): methods and techniques to automatically identify Activities of Daily Living (**ADLs**) from inertial sensors of smartphones and smartwatches

- Recent statistics show that more than 5.19 billion people (67% of the world's population) use **smartwatches**, with user numbers up by 124 million (2.4%) over the past year.

- HAR can be extensively used in several domains. For example, in **healthcare** to keep track of elderly people or the rehabilitation process after accidents and injuries, or in the guide to safe travel

Personalization in HAR

BACKGROUND

- Generally, the recognition of human activities is carried out using the so called ***user-independent recognition models***.
- User-independent models struggle to generalize 1) to new users and 2) to changes in the way a pre-existing user performs an activity due to the ***inter-subject variability*** and ***intra-subject variability***
- These factors make a constantly updated ***personal recognition model*** the ideal solution that can be obtained by training the **subject-independent** model with the data of the new user (*inter-variability*) and keeping the personal model constantly updated with new data for the existing user (*intra-variability*).

Personalization in HAR

AIM

- **Deep learning** techniques have proven to be more efficient than traditional ones in classifying ADLs
- A very interesting approach is that of Siirtola et al. that exploits ***incremental learning*** to generate a personalized model without requiring data from the target user.
- The aim of this work is to experiment the effectiveness of an approach that combines **incremental learning** and **deep learning** techniques.

Materials

DATASETS

- **Anguita**: includes 3-axial linear acceleration, 3-axial angular velocity, and gyroscope sensor data of 11 ADLs recorded with Samsung Galaxy S II (30 subjects)
- **Shoaib**: includes 3-axial acceleration, gyroscope, magnetometer, and linear acceleration sensor data of 7 ADLs recorded with Samsung Galaxy S II (10 subjects)
- **Siirtola**: includes 3-axial acceleration sensor data of 5 ADLs recorded with Nokia N8 smartphones. The activities have been performed by 8 volunteers.

NUMBER OF SEGMENTS AND CLASSES FOR EACH DATASET

Dataset	segment size	# segments	# per user	# classes
Anguita	150x4	9,712	~ 324	11
Shoaib	150x4	41,930	4,193	7
Siirtola	120x4	6,921	~ 865	5

Methods

Incremental Learning

- The **incremental learning** procedure consists in three phases: *(1) data preparation*, *(2) model generation*, and *(3) personalization:*
  - ***Data preparation*:** the subjects, except subject x, are used to train the base model in the *model generation* phase. The subject x is used in the *personalization* phase to adapt the model trained in the previous phase.
  - The ***model generation*** phase consists of obtaining an initial model from all the samples of the dataset except those from the user x. This model is called *user-independent model*.
  - For the ***personalization*** phase we experimented three different approaches: *non-supervised approach*, *semi-supervised*, and *supervised*
- **Proposed CNNS:** *Residual Network (ResNet) , Simplified CNN*
- We compared deep learning with a previous method based on Learn++ with hand-crafted features

Results

..

Performance is measured in terms of **macro average accuracy**

AVERAGE OF THE MACRO AVERAGE ACCURACY OF ALL USERS WHATEVER IS THE METHOD ADOPTED WITH AND WITHOUT THE DATA AUGMENTATION PROCEDURE PROPOSED. BEST RESULTS ARE REPORTED IN BOLD.

Augmentation	User-Independent	Non-Supervised	Semi-Supervised	Supervised
no	89.10 (± 7.56)	89.41 (± 7.76)	94.78 (± 5.95)	98.01 (± 1.98)
yes	**90.32 (± 6.24)**	**90.73 (± 6.24)**	**96.98 (± 3.46)**	**99.53 (± 0.88)**

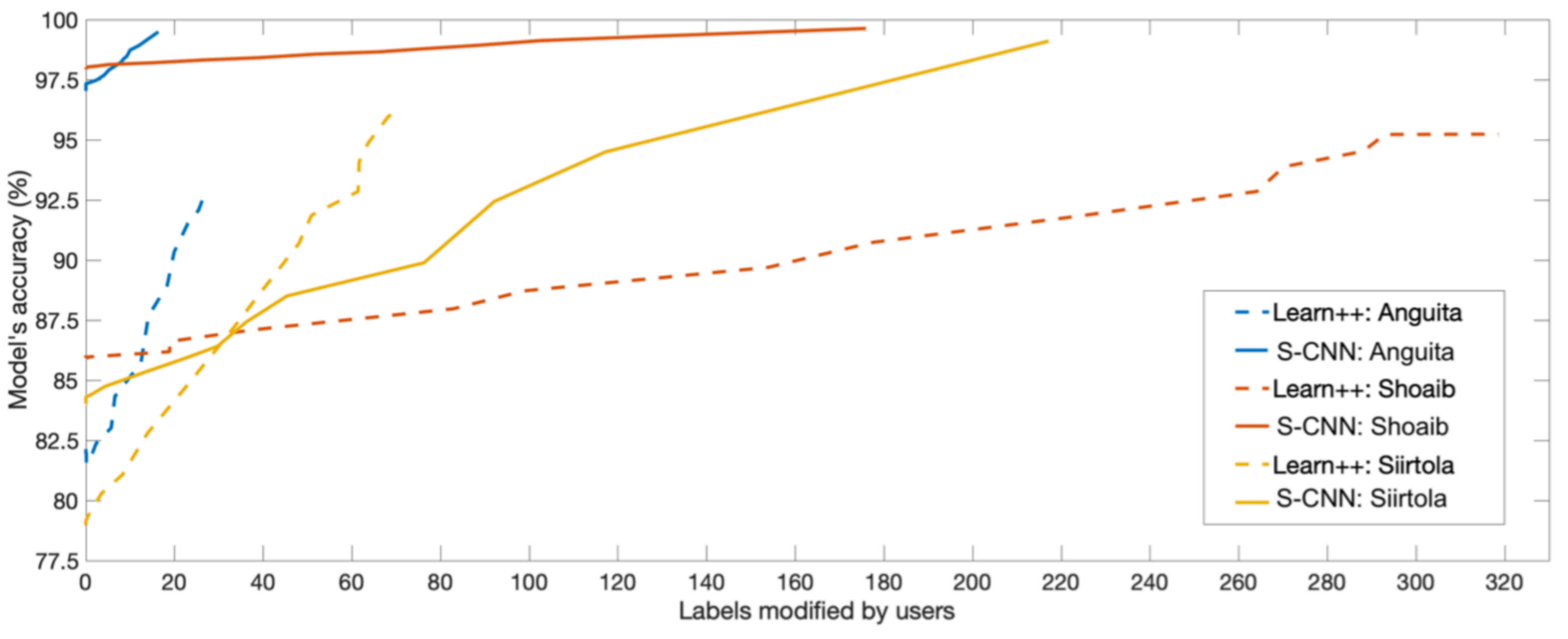
MACRO AVERAGE ACCURACY OF ALL USERS: LEARN++ VS RESNET AND S-CNN. FOR EACH MODEL (USER INDEPENDENT, NON SUPERVISED, SEMI-SUPERVISED AND SUPERVISED), THE BOLD FONT REPRESENTS THE BEST METHOD.

Dataset / Method	User-Independent	Non-Supervised	Semi-Supervised	Supervised
Anguita / Learn++	85.60 (± 7.74)	86.72 (± 7.81)	96.89 (± 4.52)	99.53 (± 0.61)
Shoaib / Learn++	86.79 (± 4.10)	87.95 (± 4.06)	97.71 (± 1.91)	99.02 (± 0.48)
Siirtola / Learn++	75.84 (± 17.51)	76.24 (± 17.53)	95.59 (± 6.08)	97.66 (± 5.73)
Mean	82.74 (± 9.78)	83.64 (± 9.80)	96.73 (± 4.17)	98.74 (± 2.27)
Anguita / ResNet	97.30 (± 2.53)	97.33 (± 2.67)	98.78 (± 2.13)	99.81 (± 0.53)
Shoaib / ResNet	98.45 (± 0.76)	98.52 (± 0.76)	99.07 (± 0.56)	99.96 (± 0.05)
Siirtola / ResNet	86.67 (± 9.41)	87.31 (± 9.25)	92.72 (± 7.31)	99.94 (± 0.04)
Mean	**94.14 (± 4.23)**	**94.39 (± 4.23)**	96.86 (± 3.33)	99.90 (± 0.21)
Anguita / S-CNN	97.49 (± 2.75)	97.61 (± 2.79)	99.06 (± 1.87)	99.99 (± 0.27)
Shoaib / S-CNN	98.32 (± 0.78)	98.37 (± 0.83)	99.22 (± 0.64)	99.93 (± 0.05)
Siirtola / S-CNN	86.54 (± 10.61)	86.50 (± 10.48)	93.74 (± 6.11)	99.90 (± 0.12)
Mean	94.09 (± 4.71)	94.16 (± 4.70)	**97.34 (± 2.87)**	**99.94 (± 0.15)**

Results

..

Learn++ algorithm and S-CNN comparison in semi-supervised cases. The plot shows the trend of classifiers' accuracy in relation to the number of labels adjusted by a user.



Conclusions

DATASETS

- In this paper we evaluated how deep learning can be employed in an **incremental learning** procedure and we compared **deep learning** with a previous method based on Learn++.
- Experiments carried out on 3 different datasets showed that, overall, deep learning outperforms Learn++. In particular, we evaluated two different networks: a ResNet and a simplified CNN. Both CNNs demonstrated to be faster than Learn++ to adapt to a new user thus demonstrating to require less user interaction than Learn++.
- We are now working on the design of a software component that implements the S-CNN and the transfer learning techniques. The component will be deployed on an Android mobile device. Once implemented, we will analyze the performance from the point of view of both power consumption and performance on real scenarios.