Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors

Ming Zeng, Le T. Nguyen, Bo Yu, Ole J. Mengshoel, Jiang Zhu, Pang Wu, Joy Zhang
Department of Electrical and Computer Engineering
Carnegie Mellon University
Moffett Field, CA, USA
Email: {zeng, le.nguyen, bo.yu, ole.mengshoel, jiang.zhu, pang.wu, joy.zhang}@sv.cmu.edu

***Abstract*—A variety of real-life mobile sensing applications are becoming available, especially in the life-logging, fitness tracking and health monitoring domains. These applications use mobile sensors embedded in smart phones to recognize human activities in order to get a better understanding of human behavior. While progress has been made, human activity recognition remains a challenging task. This is partly due to the broad range of human activities as well as the rich variation in how a given activity can be performed. Using features that clearly separate between activities is crucial. In this paper, we propose an approach to automatically extract discriminative features for activity recognition. Specifically, we develop a method based on Convolutional Neural Networks (CNN), which can capture local dependency and scale invariance of a signal as it has been shown in speech recognition and image recognition domains. In addition, a modified weight sharing technique, called partial weight sharing, is proposed and applied to accelerometer signals to get further improvements. The experimental results on three public datasets, Skoda (assembly line activities), Opportunity (activities in kitchen), Actitracker (jogging, walking, etc.), indicate that our novel CNN-based approach is practical and achieves higher accuracy than existing state-of-the-art methods.**

***Keywords*—Activity Recognition, Deep Learning, Convolutional Neural Network**

I. Introduction

In the recent years, the rapid spread of mobile devices with sensing capabilities has created a huge demand for human activity recognition (AR). Applications that can benefit from AR include daily lifelogging, healthcare, senior care, personal fitness, etc. [7], [32], [31], [9]. As a result, many approaches were proposed for the recognition of a wide range of activities [8], [15], [10], [23].

Feature extraction is one of the key steps in AR, since it can capture relevant information to differentiate among various activities. The accuracy of AR approaches greatly depends on features extracted from raw signals such as accelerometer readings [34]. Many existing AR approaches often rely on statistical features such as mean, variance, entropy or correlation coefficients [3]. Feature extraction is proposed from the frequency domain using FFT [17]. Prior works have shown that some of these heuristically-defined features perform well in recognizing one activity, but badly for others [15]. Therefore, given an application scenario and a set of target activities, one can select a subset of features to optimize the activity recognition performance [34], [15].

Designing hand-crafted features in a specific application requires domain knowledge [23]. This problem is not unique to activity recognition. It has been well-studied in other research areas such as image recognition [22], where different types of features need to be extracted when trying to recognize a handwriting as opposed to recognizing faces. In recent years, due to advances of the processing capabilities, a large amount of Deep Learning (DL) techniques have been developed and successful applied in recognition tasks [2], [28]. These techniques allow an automatic extraction of features without any domain knowledge.

In this work, we propose an approach based on Convolutional Neural Networks (CNN) [2] to recognize activities in various application domains. There are two key advantages when applying CNN to AR:

- **Local Dependency:** CNN captures local dependencies of an activity signals. In image recognition tasks, the nearby pixels typically have strong relationship with each other. Similarly, in AR given an activity the nearby acceleration readings are likely to be correlated.
- **Scale Invariance:** CNN preserves feature scale invariant. In image recognition, the training images might have different scales. In AR, a person may walk with different paces (e.g., with different motion intensity).

We summarize the key contributions of this work as follows:

- We propose an approach based on CNN to extract human activity features without any domain knowledge.
- The proposed approach can capture the local dependencies and scale-invariant features of activity signals. Thus, variations of the same activity can be effectively captured through the extracted features.
- We present the experimental results on three public datasets collected in different domains. The results shown that the proposed approach outperforms the state-of-the-art methods.

The rest of this paper is organized as follows: Section 2 presents related work; Section 3 describes our CNN-based method for activity recognition and improvement; Section 4 presents our experimental results to demonstrate its applications. Finally, we conclude the study in Section 5.

978-1-63190-024-2 © 2014 ICST

DOI 10.4108/icst.mobicase.2014.257786