Journal of Biomedical Informatics 84 (2018) 148–158

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

A sequence-to-sequence model-based deep learning approach for recognizing activity of daily living for senior care

Hongyi Zhu*, Hsinchun Chen, Randall Brown

Department of Management Information Systems, University of Arizona, Tucson, AZ, United States

ARTICLE INFO

Keywords:
Activity of daily living
ADL recognition
Deep learning
Activity state representation
Sequence-to-sequence model

ABSTRACT

Ensuring the health and safety of independent-living senior citizens is a growing societal concern. Researchers have developed sensor based systems to monitor senior citizens' Activity of Daily Living (ADL), a set of daily activities that can indicate their self-caring ability. However, most ADL monitoring systems are designed for one specific sensor modality, resulting in less generalizable models that is not flexible to account variations in real-life monitoring settings. Current classic machine learning and deep learning methods do not provide a generalizable solution to recognize complex ADLs for different sensor settings. This study proposes a novel Sequence-to-Sequence model based deep-learning framework to recognize complex ADLs leveraging an activity state representation. The proposed activity state representation integrated motion and environment sensor data without labor-intense feature engineering. We evaluated our proposed framework against several state-of-the-art machine learning and deep learning benchmarks. Overall, our approach outperformed baselines in most performance metrics, accurately recognized complex ADLs from different types of sensor input. This framework can generalize to different sensor settings and provide a viable approach to understand senior citizen's daily activity patterns with smart home health monitoring systems.

1. Introduction

The rapid advancements of medicine and accessibility of healthcare in the past decade has resulted in a steadily increasing aging population in the US and Europe. In 2016, it was estimated that 49.2 million of US citizens (15.2%) and 98.0 million of EU residents (19.2%) were over 65 years old (i.e., senior citizens) [1,2]. According to the Administration on Aging [3], 85% of senior US citizens live independently. Chronic conditions and their syndromes, such as frailty and dementia, may affect these independent-living senior citizens' health, safety, and quality of life. Ensuring a healthy and safe life for independent-living senior citizens is a growing societal concern. To evaluate the self-care ability of senior citizens, researchers and practitioners have aimed to monitor Activity of Daily Living (ADL).

ADL is a set of self-caring activities needed for independent living senior citizens. Two main ADL types exist: Basic and Instrumental. Basic ADLs (BADLs) include tasks that engage simple physical movements, such as functional mobility, self-feeding, and toilet hygiene. Instrumental ADLs (IADLs) are complex tasks that require more cognitive resources, such as preparing meals, taking prescribed medications, shopping, etc. [4–6]. According to the American Occupational Therapy Association's (AOTA) practice framework, BADLs and IADLs are considered two of eight broad categories of "areas of occupation" in which people engage [7]. The independence level of these activities positively correlates with the person's health and wellness [8]. Past studies have indicated physical and cognitive deficits deteriorate senior citizens' BADL and IADL performance due to aging or disease progression [9–11]. As a result, 44% of US senior citizens have difficulties performing one or more Basic or Instrumental ADLs [12]. Furthermore, IADLs rely more on cognitive capacity; the performance of IADLs can be impaired before the onset of dementia [13]. Thus, monitoring and recognizing the decline of IADL performance can help diagnose mild cognitive impairment (MCI), an intermediate stage between the expected cognitive decline of normal aging and the more-serious decline of dementia [13–15].

To understand aging and disease progression, clinicians and therapists often monitor a senior citizen's ADL performance using various evaluation tools (e.g., Katz ADL scale) in clinical settings [6] or self-reported activities at home [16]. However, the examination during clinic visits cannot provide timely data for early intervention or preventive care. Self-reported data (e.g., via activity diary) may not be reliable, especially if the senior citizen's cognitive function is

* Corresponding author at: Department of Management Information Systems, Eller College of Management, The University of Arizona, McClelland Hall, Room 430, P.O. Box 210108, Tucson, AZ 85721-0108, United States.

E-mail address: zhuhy@email.arizona.edu (H. Zhu).

https://doi.org/10.1016/j.jbi.2018.07.006
Received 25 February 2018; Received in revised form 9 June 2018; Accepted 8 July 2018
Available online 10 July 2018

deteriorating. To address some of these limitations, clinicians have started turning to remote home ADL monitoring systems employing sensor technologies (e.g., wearable motion, environment, cameras) [17,18]. Wearable motion sensors record data such as accelerations (i.e., speed changes, shocks, and impacts) of the monitored humans to assess their motion (e.g., intensity and stability). Environment sensors are usually attached to objects, furniture, and home appliances (e.g., pillbox, cabinet, and fridge door accordingly) to collect environment and object-triggered events' information. Cameras capture image/video data with finer-grained details, including the user's locomotion during ADLs and environment information such as room settings. ADLs are recognized and analyzed from the sensor generated data via classic machine learning approaches (e.g., Naive Bayes (NB), Support Vector Machine (SVM), and Hidden Markov Models (HMMs)) or deep-learning-based approaches such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) to help healthcare providers to offer timely patient support and closely monitor senior citizen's health progressions.

Despite their benefits over traditional clinical visits, current ADL recognition (ADLR) systems have several drawbacks. First, prevailing methods use a particular type of sensor, which cannot generalize to handle arbitrary sensor combinations. Second, classic machine learning techniques rely on heavy manual feature engineering, which is inflexible to setting changes. Third, deep-learning-based systems only recognize simple tasks such as locomotion and gesture/posture, regardless of the temporal patterns that constitute complex ADLs. In light of these limitations, we aim to develop a Sequence-to-Sequence model-based deep-learning framework leveraging a generalizable activity state representation to recognize complex ADLs with arbitrary sensor combinations and configurations.

The remainder of the paper is structured as follows. Section 2 provides a research background of prior studies. Section 3 describes our proposed research design. Section 4 presents evaluation results and discussions. Section 5 concludes this paper.

2. Related work

To form the basis of this research, we examine two areas of literature: (1) general ADLR studies to understand ADLR tasks and common sensors used for ADLR and (2) machine learning approaches for ADLR.

2.1. ADL Recognition (ADLR)

As the basis of ADL monitoring systems, ADLR aims to identify ADL events and their patterns from system observations. Daily activities are performed sequentially or interwoven, which contains temporal patterns. Furthermore, complex activities can be imagined as sequential combinations of simple activities. Thus, a temporal and hierarchical ADL decomposition approach can potentially help to understand the characteristics and semantics of ADLs [19]. Three ADL recognition tasks correspond with the hierarchy: locomotion, gesture, and complex activity. The complexity of these tasks as well as information needed for recognizing the ADLs increase from locomotion to complex activity. Locomotion depicts the human's body motion states, such as walking, standing, and sitting. Only body motion information is required to recognize locomotion. Gestures, also referred to as middle-level ADLs (ML-ADLs), are comprised of physical reciprocal actions between human and objects. One such example is closing a drawer, where a human will push and slide in the drawer. Temporal interaction patterns of human and object motion information aid in recognizing ML-ADLs. Complex activities, also referred to as high-level ADLs (HL-ADLs), consist of temporal-dependent sequences of gestures. One such activity is coffee time, where a human will fetch coffee powder and cups, operate the coffee maker, pick-up and drink the coffee, and wash the cup. HL-ADL recognition tasks are more difficult to identify, as they require motion semantics (i.e., the temporal pattern of ML-ADLs and/or locomotion). Contextual information (i.e., environment information of the HL-ADL) such as object locations, room setting, temperature, and luminance is also key to recognize HL-ADL. Modern sensing technology provides promising solutions to collect these various types of information for ADLR tasks.

Three types of sensors are used in the past ADLR studies: environment/object, motion, and cameras. Environment sensors, such as object triggers, on/off switches, infrared door sensor, and pressure sensors, are attached to the objects or placed in the environment (e.g., on the floor) to record object states or environment changes [20,21]. Environment sensors provide the location or object information during the ADL. Motion sensors, such as accelerometer, gyroscope, and magnetometer, can be used to record object or human movement. Typically, one object motion sensor is sufficient to collect the movement of the object of interest. For humans, multiple wearable sensors can be attached to different body locations to collect comprehensive human motion data [19,22,23]. Cameras capture complete motion semantics required for recognizing all levels of ADLs [24]. However, they are often too intrusive to implement in real-world settings pervasively [25].

These sensors' ability to collect detailed activity information also brings several challenges. Each sensor can collect a large volume of data with a high velocity. For example, a wearable motion sensor can record 20–100 data points per second, result in 1.7–8.6 million records per day, a rate and an amount beyond human's cognitive processing capability. Beyond the high volume and velocity challenges, data collected by environment/object and motion sensors have unique characteristics that must be considered when creating an ADLR system. Most environment sensors generate single-channel data (e.g., on/off state and pressure in Fig. 1) while most wearable and object motion sensors collect multi-channel data (e.g., acceleration on x-, y-, and z-axis). The single-channel data value can be either discrete (e.g., on/off states in Fig. 1) or continuous (e.g., the pressure reading in Fig. 1). Most multi-channel sensors data are continuous. Data on different channels jointly depict the same activity. Due to the differences among sensor and data types, past studies focus on developing ADLR systems for a specific sensor type and configuration. Thus, such systems are less generalizable to other sensor types and configurations. For example, ADLR systems based on counting discrete sensor events cannot be applied to multi-channel sensors recording continuous values (e.g., accelerometer).

Irrespective of sensor type or combination, prevailing ADLR studies often rely on machine learning algorithms to recognize ADLs. Such methods are preferred as they can capture patterns within high volume data that are difficult to detect manually. We next review machine learning approaches for ADLR.

2.2. Machine learning approaches for ADLR

Traditional machine learning approaches for ADLR follow three steps: (1) data preprocessing, (2) feature extraction, and (3) pattern recognition. Data preprocessing deals with missing value and data

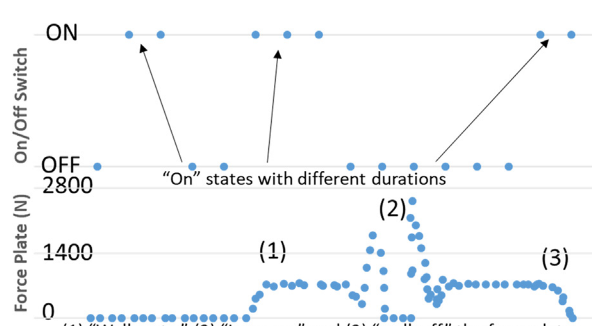

Fig. 1. Discrete and continuous sensor data collected from an on/off switch and a force plate.