



ERCIM "ALAIN BENSOUSSAN"  
FELLOWSHIP PROGRAMME



## Scientific Report

First name / Family name

Ky Trung Nguyen

Nationality

Vietnam

Name of the *Host Organisation*

VTT Finland

First Name / family name  
of the *Scientific Coordinator*

Tuomo Tuikka

Period of the fellowship

01/08/2020 to 31/07/2021

### I – SCIENTIFIC ACTIVITY DURING YOUR FELLOWSHIP

My scientific activity can be split into two main strand:

1. The first stage of my scientific activity focused on human activity recognition (HAR) from wearable sensors using machine learning such as conventional machine learning (Multilayer Perceptron, Support Vector Machines, etc.) and deep learning (Convolutional Neural Network, Recurrent Neural Network, etc.). In HAR system, many challenges have been pointed out in researches in recent years such as feature extraction methods/feature engineering, imbalanced class problem, flexibility, etc. in order to enhance a classifier performance to more accurately recognize human activities. Taking this into account, we developed a novel effective Calibration for the Oversampling Active Learning framework, called COAL, which further enhances the HAR performance by solving the posterior probability of training set

to reflect the true prior probability of testing set. This work was recently submitted in Journal of Personal Ubiquitous Computing and it is under review.

2. The second activity was to automatically generate story/script for better understand human behaviour during their daily life activities from data acquired from wearable sensors in order to automatically recommended systems to support human decision making or assisted lifestyle. The objective of this research activity is to experiment with multiple AI-based systems to develop an automatic AI-based system to generate human understandable story on human behavior. Therefore, we successfully conducted a novel system that translated sensor data (e.g., accelerometers, gyroscope, etc.) to texts which have known as, Data Story-Telling System, allows to automatically generate a story/script in term of text. Due to using human activity events are detected by a deep learning network as input (source) for sequence to sequence with attention model to generate the story (target). The purpose of the story/script is to summarize the human behaviors in their daily life activities, which assists to give instructions for human well-being and healthier life. This work was accepted for publication in IEEE International Conference on Biomedical and Health Informatics (BHI'21).

## II – PUBLICATION(S) DURING YOUR FELLOWSHIP

### Submitted

1. Ky Trung Nguyen, Son T. Mai, Jani Mantyjarvi, Francois Portet, Anh Le, Van-Hau Nguyen. "An Effective Calibration Oversampling Active Learning Framework for Class Imbalance Problem in Human Activity Recognition" (in Journal of Personal and Ubiquitous Computing).

**Abstract:** Over the past few years, human activity recognition (HAR) has been an emerging research topic that provides us with a better understanding of human behavior through their day-to-day activities. Many supervised approaches have been proposed to classify human activities (e.g., neural networks and support vector machines). However, these classifiers are very sensitive to the class imbalance problem, as most discriminative models, they suffered from the mis-classification between majority and minority classes due to in favor of majority classes. To tackle this problem, in this paper we proposed a

novel effective calibration on the a posteriori probability of an oversampling active learning framework based on neural networks so that it can appropriately reflect to the true a priori probability of the target class. In addition, we introduce a novel weighted entropy that make active learning is possible to select effectively minority samples for oversampling techniques. Experiments demonstrate superior classification accuracy of our method than state-of-the-art techniques on two available human activity datasets.

### **Accepted**

2. Ky Trung Nguyen, Jani Mantyjarvi, Tran Thi Ngoc Nguyen. "Automatic Storytelling from Wearable Sensor Data for Health and Wellness Applications", in IEEE International Conference on Biomedical and Health Informatics (BHI'21), 2021.

**Abstract:** Storytelling describes our daily living activities in many ways. It assists us to understand what we have done to advance our health and wellbeing. In this paper, we present our novel approach to generate scripts from events, which are detected from wearable sensors data. First, we use Deep Neural Network (DNN) to recognize semantic concepts such as gesture, activity, and location for generating a chronological sequence of events. Second, we apply a sequence to sequence (SEQ2SEQ) model consisting of two recurrent neural networks (RNNs) to generate human-understandable stories. The results show that our method can improve the performance of script generation (SG) by using SEQ2SEQ with 0.972 BLEU-1 score.

### **To be submitted**

3. Ky Trung Nguyen, Jani Mantyjarvi, Tran Thi Ngoc Nguyen. "Automatic Telling Time Story from Wearable Sensor Data", to be submitted to a Journal.

## **III – ATTENDED SEMINARS, WORKHOPS, CONFERENCES**

I attend IEEE-EMBS International Conference on Biomedical and Health Informatics July 27 - 30, 2021, Greece.

## **IV – RESEARCH EXCHANGE PROGRAMME (REP)**

I did the REP with Professor Ole Jakob Mengshoel from the Department of Computer Science at the NTNU in June 2021. Mengshoel research interest is in automatic storytelling

from wearable sensor data using machine learning. Currently, he has been working human activity recognition, which is first stage to my work. During this time, we met through online Zoom meetings. I presented my current projects that I was doing for the postdoctoral position at VTT Finland.