An Integrated Framework for Human Activity Recognition

Hong Cao[1], Minh Nhut Nguyen[1], Clifton Phua[1], Shonali Krishnaswamy[1,2], Xiao-Li Li[1]
[1]Institute for Infocomm Research, A*STAR
[2]Faculty of Information Technology, Monash University
{hcao, mnnguyen, cwcphua, spkrishna, xlli}@i2r.a-star.edu.sg

ABSTRACT
This poster presents an integrated framework to enable using standard non-sequential machine learning tools for accurate multi-modal activity recognition. Our framework contains simple pre- and post-classification strategies such as class-imbalance correction on the learning data using structure preserving oversampling, leveraging the sequential nature of sensory data using smoothing of the predicted label sequence and classifier fusion, respectively, for improved performance. Through evaluation on recent publicly-available OPPORTUNITY activity datasets comprising of a large amount of multi-dimensional, continuous-valued sensory data, we show that our proposed strategies are effective in improving the performance over common techniques such as One Nearest Neighbor (1NN) and Support Vector Machines (SVM). Our framework also shows better performance over sequential probabilistic models, such as Conditional Random Field (CRF) and Hidden Markov Models (HMM) and when these models are used as meta-learners.

Author Keywords Activity recognition, Data mining, Imbalance, Learning, Ubiquitous computing, Smoothing, Wearable computing.

General Terms Algorithms, Design

INTRODUCTION

The widespread availability and easy accessibility of real-time sensory information brings a new maturity to the field of activity recognition evidenced by the emerging trend of benchmarking initiatives such as the OPPORTUNITY Activity Recognition. While sequential classifiers, e.g. HMM and CRF, are common for learning from sequences, in this poster, we proposed a framework with pre- and post-classification techniques that enable using standard non-sequential learning techniques for accurate activity recognition. This is motivated by our finding that the non-sequential techniques, such as SVM and 1NN, have good competitiveness and scalability on large-dimensional and continuous-valued activity sensory data. We consider the following key factors in the context of activity recognition:

1. Class Imbalance - This refers to the naturally occurring characteristic in activity data wherein some activities recur at a considerably higher frequency and with a longer duration, while some others are relatively infrequent and short. We propose to use structure preserving oversampling technique for addressing the class imbalance at the fundamental data level with apparently improved learning performance.

2. Smoothing and fusion – Leveraging on the fact that sensor sampling rate is at significantly higher frequency than the rate of change of human activities, we propose and develop a simple yet effective smoothing technique that results in improved recognition accuracy. Our fusion further combines our smoothed results from different classifiers with improved performance in activity recognition.

Contrary to many existing works that apply individual learning techniques in isolation for activity recognition, our work shows not only the incremental gains due to each proposed component, but also the good framework performance after integration. Additional experiments in our paper also show that our framework improves the learning performance when compared with sequential learning techniques such as Hidden Mark Models (HMM) and Conditional Random Fields (CRF), and when these techniques are used as meta-learners.

PROPOSED INTEGRATED FRAMEWORK

Figure 1 shows our integrated framework. In the training phase, we perform the necessary preprocessing on the activity data sequence including filling up the missing sensory readings and normalization. On the preprocessed data, we further correct class imbalance through structure preserving oversampling of the minority classes. The balanced data is then used to learn standard classifiers. On a separate branch, we also analyze the sequence of activity labels for determining a best filter length for smoothing. In the testing phase, with similar preprocessing, we classify the multi-modal activity instances using our trained classifiers. Their predicted label sequences are then smoothed and fused to achieve reliable performance.

Class Imbalance Correction and Learning

Previously, several works have proposed to deal with data imbalance through manipulating operating point on the receiver operating characteristic (ROC) curve for a probabilistic classifier [3] and through oversampling the minority class through a non-parametric bootstrapping technique. We propose to adapt structure preserving oversampling (SPO) for imbalance correction on multi-class and high-dimensional activity sensory data. SPO generate synthetic samples by preserving the main covariance structure of the minority class. Besides this, the synthetic distribution also intelligently adds buffering

UbiComp'12, September 5-8, 2012, Pittsburgh, USA.
ACM 978-1-4503-1224-0/12/09.

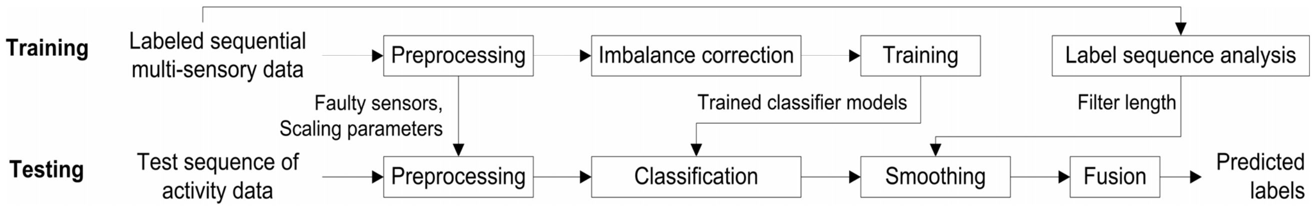

Figure 1. Block Diagram of our Proposed Integrated Framework

variances in the regularized eigen dimensions [4]. In the case of sparse distribution of the minority class, the synthetic instances also effectively disperse into the void space within and in the vicinity of the minority class to form a more comprehensive new class territory. Moreover, the cleaning mechanism ensures that the synthetic instances from SPO do not worsen the current class overlapping by introducing new Tomek links. This reduces the inherent risk that the minority classes do not follow Gaussian distribution. Such desired features are seldom provided in conventional oversampling methods, such as SMOTE [5]. After oversampling, we use sensory values in balanced dataset as features to train the non-sequential classification models, such as SVM and 1NN [1], for classifying the test instances.

Smoothing and Fusion

By applying our trained classifiers to a test sequence of activity instances, we readily obtain several predicted label sequences. As shown in Figure 1, we further introduce two post-processes, i.e. smoothing and fusion, to enhance the classification performance. Given that instance sampling is performed on increasing time dimension, our smoothing operates along the temporal axis and utilizes the fact that neighboring instances in a sensor sequence often share similar activity class labels. This is due to the high sensor sampling rate in comparison with the frequency of human activity changes. We thus propose fusion as a mechanism to integrate the smoothed label sequences from different classifiers to have more robust prediction outcomes.

EXPERIMENT AND DISCUSSION

We evaluated our proposed integrated framework on OPPORTUNITY activity recognition datasets [1]. Three classification tasks including locomotion (4 classes), segmentation (2 classes) and gesture (18 classes) are considered based on activity recognition data from three subjects. Note that activity learning sequence data for each subject contains large amount of data including over 50,000 instances with each instance containing over 100 sensory data dimensions. Based on such data, our framework learns the subject-dependent predictive models for the three different activity classification tasks. The comparison results demonstrated apparent gain corresponding to each of our proposed pre- and post- classification modules. Our imbalance correction module contributes to the most significant gain, subsequently followed by the smoothing and the fusion modules. For example, the improvement (measured in terms of an average of three evaluation criteria) for Gesture classification can be attributed to imbalance correction (+5.8%), smoothing (+5.0%), and fusion (+2.9%) for 1NN, and to imbalance correction (+4.8%), smoothing (+1.9%) and fusion (+1.6%) for SVM.

We have also compared our framework with state-of-the-art sequential classifiers such as HMMs (Hidden Markov Models) and CRFs (Conditional Random Fields), and when these classifiers are used as meta-learners [6]. We find our framework is significantly better than these alternatives. For instance, our integrated framework improves by 24% over the second best (CRF) in terms of the average of three evaluation measures and improves 1.9% when CRF is used as a meta learner. This is likely due to the incompetency of using these sequential classifiers to learn from a large amount of high-dimensional and continuous-valued sensory data. For such type of emerging sensory data, our proposal thus introduces a new highly competent framework to activity recognition community than the conventional approaches.

In view of the growing prospect of large-dimensional body sensor network, our future work would involve sourcing new datasets or data collection for comprehensive study in the modeling of human activities.

REFERENCES

1. Sagha, H., Digumarti, S.T., et al.: Benchmarking Classification Techniques using the Opportunity Human Activity Dataset. In: IEEE Int. Conf. on Systems, Man, and Cybernetics (2011)
2. Albinali, F., Davies, N. and Friday, A. Structural Learning of Activities from Sparse Datasets. In: PerCom'07, 221-228, (2007)
3. Stager, M, Lukowicz, P. and Troster, G. Dealing with Class Skew in Context Recognition, Proc. ICDCSW'06, pp. 58-63, (2006)
4. Cao, H., Li, X.-L., Woon Y.-K., Ng S.-K.: SPO: Structure Preserving Oversampling for Imbalanced Time Series Classification. In: Proc. IEEE Int. Conf. on Data Mining, (2011)
5. Chawla, N.V., Bowyer, K.W., et al.: SMOTE: Synthetic Minority Over-Sampling Technique. J. Artificial Intelligence, 16, 321-357, (2002)
6. Lester, J., Choudhury, T., et al.: A Hybrid Discriminative/Generative Approach for Modeling Human Activities. In: IJCAI'05, 766-772, (2005)