International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) e-ISSN: 2395-5325
Volume4, Issue 2 (February '2018)

COMPLEX EVENT DETECTION VIA EVENT ORIENTED DICTIONARY LEARNING

B.Sandhiya,
B.Sc(Computer Science),
Department of Computer Science,
Sri Krishna Adithya College of Arts and Science,
Coimbatore-641042,Tamil Nadu,India.

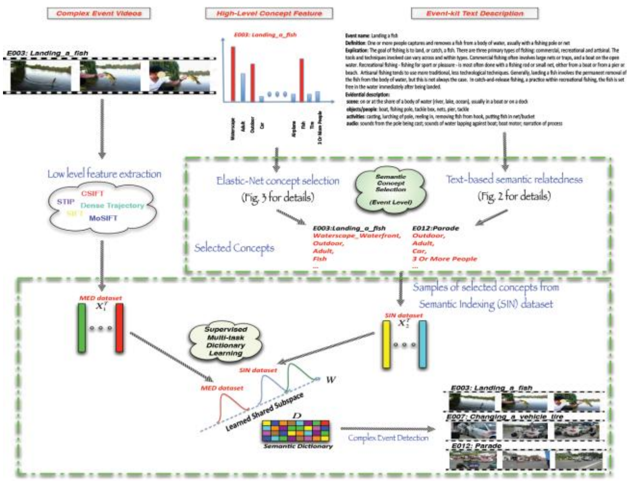
S.ShakthiAkshayaa,
B.Sc(Computer Science),
Department of Computer Science,
Sri Krishna Adithya College of Arts and Science,
Coimbatore-641042,Tamil Nadu,India.

Abstract: Complex event detection is a retrieval task with the goal of finding videos of a particular event in a large scale unconstrained internet video archive, given example videos and text descriptions. Nowadays, different multimodal fusion schemes of low-level and high-level features are extensively investigated and evaluated for the complex event detection task. However, how to effectively select the high-level semantic meaningful concepts from a large pool to assist complex event detection is rarely studied in the literature.

Keywords: Complex Event detection, Multitasking, Compression methods, Supervised

I.INTRODUCTION

Complex event detection in unconstrained videos has received much attention in the research community recently (Tamrakar et al. 2012; Ma et al. 2013; Natarajan et al. 2012). It is a retrieval task with the goal of detecting videos of a particular event in a large-scale internet video archive, given an event-kit. An event-kit consists of example videos and text descriptions of the event. Unlike traditional action recognition of atomic actions, such as 'walking' or 'jumping' from videos, complex event detection aims to detect more complex events such as 'Birthday party', 'Changing a vehicle tire', etc. An event is a higher level semantic abstraction of video sequences than a concept and consists of many concepts.



II.EVENT ORIENTED DICTIONARY LEARNING

After we select semantic meaningful concepts for each event, we can leverage training samples of selected concepts from the SIN dataset into a supervised multi-task dictionary learning framework. In this section, we investigate how to learn an event oriented dictionary representation. To accomplish this goal, we firstly propose our multi-task dictionary learning framework and then introduce its supervised setting.

Multi-task Dictionary Learning

Given K tasks (e.g. K = 2 in our case, one task is the MED dataset and the other task is the subset of SIN dataset where samples are collected from specified selected concepts for each event), each task consists of data samples denoted by Xk = fx1 k; x2 k; :::; xnk k g 2 IRnk_d; (k = 1; :::;K), where xi k 2 IRd is a d-dimensional feature vector and nk is the number of samples in the k-th task. We are going to learn a shared subspace across all tasks, obtained by an orthonormal projectionW2 IRd_s, where s is the dimensionality of the subspace. In this learned subspace, the data distribution from all tasks should be similar to each other. Therefore, we can code all tasks together in the shared subspace and achieve better coding quality. The benefits of this strategy are: (i) we can improve each individual coding quality by transferring knowledge across all tasks. (ii) we can discover the relationship among different datasets via coding analysis.

Supervised Multi-task Dictionary Learning

It is well-known that the traditional dictionary learning framework is not directly available for classification and the learned dictionary has merely been used for signal reconstruction (Mairal et al. 2008). To circumvent this problem, researchers have developed several algorithms to learn a classification-oriented dictionary in a supervised learning fashion by exploring the label information. In this subsection, we extend our proposed multi-task dictionary learning of Eqn.(1) to be suitable for event detection.

Comparison Method

We compare our proposed event oriented dictionary learning method with the following important baselines:

- Support Vector Machine (SVM): SVM has been widely used by several research groups for MED and has shown its robustness (Lan et al. 2013; Oneata et al.

2012), so we use it as one of the comparison algorithms;

- Single Task Supervised Dictionary Learning (ST-SDL): Performing supervised dictionary learning on each task separately;
- Pooling Tasks Supervised Dictionary Learning (PT-SDL): Performing single task supervised dictionary learning by simply aggregating data from all tasks;
- Multiple Kernel Transfer Learning (MKTL) (Jie, Tommasi, and Caputo 2011): A method which incorporates prior features into a multiple kernel learning framework;

Table 1: AP performance for each MED event using Text (T), Visual (V) and Text+Visual (T+V) information for concept selection. The last column shows the number of concepts that are coincided with groundtruth in top 10 concepts.

	Event	T	V	T+V	# in Top 10
P001	Assembling shelter 0	.0331	0.0532	0.613	5
P002	Batting a run	0.4432	0.4653	0.4837	7
P003	Making a cake	0.0502	0.0514	0.0658	9
E001	Attempting board trick	0.1457	0.1675	0.1883	6
E002	Feeding animal	0.0355	0.0361	0.0402	5
E003	Landing fish	0.1721	0.1801	0.1938	5
E004	Wedding ceremony	0.3777	0.3831	0.4104	10

- Dirty Model Multi-Task Learning (DMMTL) (Jalali et al. 2010): A state-of-the-art multi-task learning method imposing `1=`q-norm regularization;
- Multiple Kernel Learning Latent Variable Approach (MKLLVA) (Vahdat et al. 2013): A multiple kernel learning latent variable approach for complex video event detection;
- Random Concept Selection Strategy (RCSS): Performing our proposed supervised multi-task dictionary learning without involving concept selection strategy (leveraging random samples).

III.CONCLUSION

In this paper, we have firstly investigated the possibility of automatically selecting semantic meaningful concepts for complex event detection based on both the MED events-kit text descriptions and the high level concept feature descriptions. Then we attempt to learn an event oriented dictionary representation based on the selected semantic concepts. To this aim, we leverage training samples of selected concepts from the SIN dataset into a novel jointly supervised multitask dictionary learning framework.

IV.REFERENCES

[1]. Aharon, M.; Elad, M.; and Bruckstein, A. 2006. K-SVD:An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. on Image Processing 54(11):4311–4322.

[2]. Argyriou, A.; Evgeniou, T.; and Pontil, M. 2007. Multi-task feature learning. In NIPS.