
Zero-Shot Learning for Human Action Recognition

Nicole Rafidi
nrafidi@cs.cmu.edu

Yuzuko Nakamura
yuzi@cs.cmu.edu

Danny Zhu
dannyz@cs.cmu.edu

1 Introduction

2 Related Work

Zero-shot learning (ZSL) enables one to classify input into classes not yet seen in training data by making use of a semantic knowledge base as an intermediate step between the features and classes [1]. The main objective of this project is to see whether zero-shot learning yields good performance in human action recognition. Human action recognition is a domain in which there are many possible class labels (actions), making it intractable to provide examples of each action in a training set [2]. This makes it a good candidate for improvement by zero-shot learning.

3 Methods

In order to do this, we must establish two steps of classification on the videos in the data sets. First, we must find a low-dimensional set of core features that can be used to distinguish the labels (running, walking, etc). It is possible that we will find a pre-established feature set based on text corpus statistics or psycholinguistic surveys. Alternatively, we could design our own feature set, based on our knowledge of human actions. Second, we will explore ways of converting the raw data (videos) to points in this feature space. To test the effectiveness of this two-step method over direct video-to-label mappings, we can perform leave-two-out cross-validation, and demonstrate the ZSL classifiers ability to distinguish between two previously unseen class labels based on their video data, a task that direct mappings are unable to perform.

3.1 Materials

Data was taken from the following two data sets:

1. KTH: <http://www.nada.kth.se/cvap/actions/>
2. Weizmann: <http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html>

4 Preliminary Results

5 Future Work

References

- [1] M. Palatucci, D. Pomerleau, G. Hinton, and T. Mitchell, "Zero-shot learning with semantic output codes," in *Neural Information Processing Systems (NIPS)*, December 2009.
- [2] R. Poppe, "A survey on vision-based human action recognition," *Image and Vision Computing*, vol. 28, no. 6, pp. 976 – 990, 2010.