

Special issue on concept detection with big data

Shih-Fu Chang¹ · Thomas S. Huang² · Michael S. Lew³ · Bart Thomee⁴

Published online: 17 March 2015
© Springer-Verlag London 2015

One of the grand challenges of machine intelligence and pattern recognition for the past decade has been bridging the semantic gap, that is, determining how to translate the low-level features from images, video and audio to the high-level concepts of humans. Concept detection is an important approach toward bridging the semantic gap by allowing computers to understand imagery using the conceptual vocabulary of humans. By exploiting big data, the current generation of algorithms has contributed and developed both advances in accuracy and computational efficiency as well as new paradigms and techniques in concept detection.

This special issue provides a focus on the state-of-the-art in concept detection with big data. We received 21 submissions of which 16 were selected for the triple peer-review process. In this special issue, we are pleased to present six research papers on concept detection with big data that present the latest advances in this field:

- Instead of learning concept detectors in advance, which limits users to query only for concepts known to the sys-

tem, the paper, “On-the-fly Learning for Visual Search of Large-scale Image and Video Datasets” by K. Chatfield, R. Arandjelović, O. Parkhi and A. Zisserman, presents a framework that enables a suitable concept detector to be generated in real-time given a keyword query. The approach considers as positive training examples the top retrieval results obtained by issuing the query to a standard web search engine, and mixes them with a fixed pool of negative examples before learning the classifier. By varying the kind of positive search images downloaded from the search engine, the real-time classifier can be tailored to support different modalities like objects and faces.

- Concept detectors can improve themselves over time by incrementally learning from additional examples. In the paper, “Learning to detect concepts with Approximate Laplacian Eigenmaps in large-scale and online settings” by E. Mantziou, S. Papadopoulos and Y. Kompatsiaris, the authors propose an inductive manifold learning approach that achieves significant speedups on training of concept detectors without noticeably degrading the detection accuracy. Their approach computes an embedding for incoming images while avoiding having to rebuild the underlying latent representation of the manifold.
- The performance of concept detectors is often dependent on the choice and tuning of their underlying kernels. In the paper, “ImageCLEF Annotation with Explicit Context-Aware Kernel Maps” by H. Sahbi, the author designs a new continuous, symmetric and positive semi-definite kernel in which context is integrated, such that pairs of visually and semantically similar images are mapped together. This is achieved by considering an image collection as a graph and the kernel recursively

✉ Michael S. Lew
mlew@liacs.nl

Shih-Fu Chang
sfchang@ee.columbia.edu

Thomas S. Huang
t-huang1@illinois.edu

Bart Thomee
bthomee@yahoo-inc.com

¹ Columbia University, New York, NY, USA

² University of Illinois, Urbana-Champaign, IL, USA

³ Leiden University, Leiden, The Netherlands

⁴ Yahoo Labs, San Francisco, CA, USA

diffuses similarity between neighborhoods of connected images.

- One of the challenges in concept detection is gathering enough, reliable, annotated training examples for building accurate classifiers. In the paper, “Building effective SVM concept detectors from clickthrough data for large-scale image retrieval” by I. Sarafis, C. Diou and A. Delopoulos, the authors analyze the click logs of the Bing image search engine to determine the degree of relevance of an image for a particular concept using IR models, consequently enabling training sets to be automatically generated and used for the construction of noise-resilient SVM classifiers.
- Manual labeling of training examples may yield accurate annotations, but it is a time-consuming effort. In the paper, “Large image modality labeling initiative using semi-supervised and optimized clustering” by S. Vajda, D. You, S. Antani and G. Thoma, the authors propose a multi-view clustering method to reduce the amount of manual labeling needed while still building a reliable classifier. The method projects biomedical images into multiple feature spaces, each of which is separately clustered and where only the cluster centers require manually labeling. The assigned labels are then propagated back to all other images in each cluster provided that a majority of feature spaces agree on the label to assign to an image.
- The multiple modalities underlying cross-media data can not only bring advantages, in terms of being able to exploit additional and complementary signals, but may also bring disadvantages when these signals are different and conflicting. In the paper, “Distributed Cross-Media Multiple Binary Subspace Learning” by X. Zhao, C. Zhang and Z. Zhang, the various modalities of such media are mapped into a common, binary subspace in which a semi-supervised concept detector can be learned efficiently.