IEEE Transactions on Circuits and Systems for Video Technology

Special Issue on Label-Efficient Learning on Video Data

Summary

While huge success for various image tasks is fueled by large-scale, well-annotated datasets, collecting and labeling such video equivalents can be incrementally and notoriously labor-intensive (annotating 30+ frames/sec), posing significant obstacles for video algorithm training and rendering them incapable of application to practical problems. It is therefore a natural desire to develop label-efficient schemes, in order to dramatically alleviate manpower in annotation-expensive video regimes. Although related label-efficient techniques applied to image data have gained much recent attention, principled solutions for video data are still in their infancy.

On the other hand, due to the intrinsic, structured nature, unlabeled video data are valuable assets for label-efficient learning algorithms. Specifically, one of the most complimentary and dominant strategies for improving image label efficiency is through data augmentation (e.g., rotation, flip, rotation, etc.), which alleviates annotation cost by encouraging consistent representations among the original and transformed samples. However, utilizing such naive transformations, though straightforward, generates artifacts during learning, may leading to sub-optimal performance. In contrast, video data naturally capture the appearance and geometric transformations in the physical world, hence providing rich and realistic samples for label-efficient learning.

From a more practical point of view, video tasks have a greater potential to revolutionize society-relevant applications (e.g., autonomous driving, video surveillance, and search-and-rescue), as they are better suited to describe the dynamic reality in which we live. Meanwhile, video tasks pose more challenges due to the inherent requirement of simultaneously comprehending the spatial granularity and temporal coherence that exist in the real-world complex. Thus, it is highly needed to exploit the full potential of label-efficient learning in order to reach a holistic understanding of visual content in both spatial and temporal domains, hence facilitating the wide range of real-world video applications.

The aforementioned discussion merely scratches the surface of the innovative opportunities available in label efficiency in the video domain. This special issue on "**Label-Efficient Learning on Video Data**" aims at promoting cutting-edge research along this direction and offers a timely collection of works to benefit researchers and practitioners.

Scope

We welcome high-quality original submissions addressing important novel theories, methods, applications, and insights centered on label efficiency for video data. The list of possible topics includes, but is not limited to:

- Annotation-efficient learning for video classification, object detection, segmentation, tracking, prediction
- Annotation-efficient learning for video captioning, video question-answering, spatio-temporal visual grounding

- Annotation-efficient learning for emerging video tasks
- Video weakly supervised learning
- Video semi-supervised learning
- Video unsupervised and self-supervised learning
- Video zero-/few-shot learning
- New benchmark datasets and survey papers related to the aforementioned topics

Paper Submission
Paper submissions can be made to the TCSVT manuscript submission website, here. Please select the correct special issue when submitting.

Important Dates:
- Submission deadline: April 1, 2023
- First review notification: May 20, 2023
- Revision submission due: July 1, 2023
- Second round review: July 1 to August 20, 2023
- Notification of acceptance/rejection: August 20, 2023

Guest Editors:
Dongfang Liu, Assistant Professor, Rochester Institute of Technology, USA (Email: dongfang.liu@rit.edu)
Tianfei Zhou, Research Fellow, ETH Zurich, Switzerland (Email: ztfei.debug@gmail.com)
Wenguan Wang, Lecturer and DECRA Fellow, University of Technology Sydney, Australia (Email: wenguanwang.ai@gmail.com)
Zheng (Thomas) Tang, NVIDIA Scientist, USA (Email: tangzhengthomas@gmail.com)
Alexander C. Loui, Professor, Rochester Institute of Technology, USA (Email: acleec@rit.edu)