Image and Vision Computing 114 (2021) 104268

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Advances in domain adaptation for computer vision

In daily routines, humans not only learn and apply knowledge for visual tasks but also have intrinsic abilities to transfer knowledge between related vision tasks. For example, if a new vision task is relevant to any previous learning, it is possible to transfer the learned knowledge for handling the new vision task. In developing new computer vision algorithms, it is desired to utilize these capabilities to make the algorithms adaptable. Generally, traditional computer vision methods do not automatically adapt to a new task and have to learn the new task from the beginning. These methods do not consider that the two visual tasks may be related and the knowledge gained in one way may be applied to learn the other one efficiently in lesser time. Domain adaptation for computer vision is the area of research, which attempts to mimic this human behavior by transferring the knowledge learned in one or more source domains and use it for learning the related visual processing task in the target domain. Recent advances in domain adaptation, particularly in co-training, transfer learning, and online learning have benefited computer vision research significantly. For example, learning from high-resolution source domain images and transferring the knowledge to learning low-resolution target domain information. This special issue is focused on the recent advances in domain adaptation for different computer vision tasks and collects 12 papers.

In "Infrared and Visible Image Fusion via Global Variable Consensus", the authors proposed a general fusion framework to solve the global variable consensus optimization problem by altering the direction method of multipliers. This method expands the fusion algorithm's compatibility by tackling various settings including dimensionality, data types, and styles. The paper "An Unsupervised Domain Adaptation Scheme for Single-Stage Artwork Recognition in Cultural Sites" reported a new dataset containing both synthetic and real images of 16 different artworks. Moreover, a comprehensive investigation on different domain adaptation techniques such as one-stage and two-stage object detectors, image-to-image translation, and feature alignment is conducted. The authors in "Multimodal Image Fusion based on Point-Wise Mutual Information" proposed a novel multimodal image fusion algorithm that focuses on both transferring the salient structures and maintaining spatial consistency. The proposed algorithm selects features to be transferred into fusion results by a graph cut algorithm, in which the spatial varying smoothness cost is formulated between local features measured by point-wise mutual information. The paper "Cuepervision: Self-supervised learning for continuous domain adaptation without catastrophic forgetting" aims at developing a self-supervised continuous domain adaptation method. In the proposed model, a pre-trained network is used to generate pseudo labels on the samples of an adjacent target domain. The pseudo labels and samples enable one to fine-tune the model, which, as a result, is adapted into the intermediate domain. In "Camera pose estimation in multi-view environments: from virtual scenarios to the real world", the authors proposed a domain adaptation strategy to efficiently train network architectures for estimating the relative camera pose in multi-view scenarios.

The paper "Motion saliency based multi-stream multiplier ResNets for action recognition" aims at developing a Motion Saliency based multi-stream Multiplier ResNets (MSM-ResNets) for action recognition. The proposed MSM-ResNets model consists of three interactive streams: the appearance stream, motion stream, and motion saliency stream. Similar to conventional two-stream CNNs models, the appearance stream and motion stream are responsible for capturing the appearance and motion information, respectively, while the motion saliency stream is responsible for capturing the salient motion information. In "DA-SACOT: Domain Adaptive-Segmentation guided Attention for Correlation-based Object Tracking", the authors proposed an instance segmentation method as an attention mechanism in an object-tracking framework. The core contribution of this paper is threefold: (i) a region proposal module (RPM) based on instance segmentation to focus on search proposals having a high probability of being the target, (ii) a target localization module (TLM) to localize the final target using a correlation filter and (iii) a domain adaptation technique in both RPM and TLM modules to incorporate target specific knowledge and strong discrimination ability. In "DeepSegment: Segmentation of Motion Capture Data Using Deep Convolutional Neural Network", a novel framework was proposed to segment 3D human motion capture data into distinct behaviors.

The authors in "Knowledge Distillation Methods for Efficient Unsupervised Adaptation across Multiple Domains", proposed a progressive approach for unsupervised single and multi-target domain adaptation in CNNs. In "Improved Multi-Source Domain Adaptation by Preservation of Factors", the authors investigated how domain adaptation affects different computer vision tasks. The paper "PDA: Proxy-based Domain Adaptation for Few-shot Image Recognition", aims at developing a simple yet effective method, called proxy-based domain adaptation, to optimize the pre-trained representation and a target classifier simultaneously. In "Deep Domain Adaptation with Ordinal Regression for Pain Assessment Using Weakly-Labeled Videos", the authors introduced a new deep learning model for weakly supervised domain adaptation with ordinal regression that can be adapted using target domain videos with coarse labels provided on a periodic basis. The model enforces ordinal relationships among the intensity levels assigned to target sequences, and associates multiple relevant frames to sequence-level labels (instead of a single frame). This model learns discriminant and domain-invariant feature representations by integrating multiple instance learning with deep adversarial domain adaptation, where soft Gaussian labels are used to efficiently represent the weak ordinal sequence-level labels from the target domain.

https://doi.org/10.1016/j.imavis.2021.104268
0262-8856/

Finally, the guest editors wish to thank the Editors-in-Chief Prof. Sinisa Todorovic, for providing the opportunity to edit this special issue. We would like to thank the authors for submitting contributions, and all the reviewers for their most helpful and constructive comments. We hope that the readers benefit from this special issue.

Pourya Shamsolmoali
Institute of Image Processing & Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China
*Corresponding author.
E-mail address: pshams@sjtu.edu.cn

Salvador García
Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain
E-mail address: salvagl@decsai.ugr.es

Huiyu Zhou
School of Informatics, University of Leicester, Leicester, UK
E-mail address: hz143@leicester.ac.uk

M. Emre Celebi
Department of Computer Science and Engineering, University of Central Arkansas, Conway, AR, USA
E-mail address: ecelebi@uca.edu

Available online 6 August 2021