Editorial: Learning With Fewer Labels in Computer Vision

I. Introduction

UNDOUBTEDLY, Deep Neural Networks (DNNs), from AlexNet to ResNet to Transformer, have sparked revolutionary advancements in diverse computer vision tasks. The scale of DNNs has grown exponentially due to the rapid development of computational resources. Despite the tremendous success, DNNs typically depend on massive amounts of training data (especially the recent various foundation models) to achieve high performance and are brittle in that their performance can degrade severely with small changes in their operating environment. Generally, collecting massive-scale training datasets is costly or even infeasible, as for certain fields, only very limited or no examples at all can be gathered. Nevertheless, collecting, labeling, and vetting massive amounts of practical training data is certainly difficult and expensive, as it requires the painstaking efforts of experienced human annotators or experts, and in many cases, prohibitively costly or impossible due to some reason, such as privacy, safety or ethic issues.

Despite the recent progress of foundation models, DNNs still lack the ability to learn from limited exemplars and quickly generalize to new tasks. However, real-world computer vision applications often require models that are able to (a) learn with few annotated samples, and (b) continually adapt to new data without forgetting prior knowledge. By contrast, humans can learn from just one or a handful of examples (i.e., few-shot learning), can do very long-term learning, and can form abstract models of a situation and manipulate these models to achieve extreme generalization. As a result, one of the next big challenges in computer vision is to develop learning approaches that are capable of addressing the important shortcomings of existing methods in this regard. Therefore, in order to address the current inefficiency of machine learning, there is a pressing need to research methods, (1) to drastically reduce requirements for labeled training data, (2) to significantly reduce the amount of data necessary to adapt models to new environments, and (3) to even use as little labeled training data as people need.

This special issue gathers recent advances in the area of visual learning with fewer labels via techniques including selfsupervised learning, semisupervised learning, weakly supervised learning, few shot learning, zero shot learning, meta-learning, continual learning, and domain adaptation. As guest editors of this special section, we were happy to receive 162 submissions to our special issue. After a careful review process, we accepted 37 papers for publication. We thank the reviewers who provided detailed, insightful, and timely reviews, leading to the high quality of accepted papers. We also thank *TPAMI* EIC and Associate EICs for recognizing the widespread interest in this field, which warrants this special issue. The accepted 37 papers are grouped into six different main categories and described in the following six sections.

II. Unsupervised or Selfsupervised Learning

The paper [A1] by Chen et al. provides a comprehensive survey of recent advances in exploring unlabeled data to enhance deep model generalization, covering recent SemiSupervised Learning (SemiSL) and Unsupervised Learning (UL) methods, their applications in visual recognition, open challenges, and emerging trends. The survey distinguishes itself from previous surveys by providing a unified taxonomy of SemiSL and UL and drawing underlying connections among different methods, which offers researchers a systematic overview.

The paper [A2] by Ding et al. proposes a novel selfsupervised pretraining method that is designed for improving spatially sensitive downstream tasks such as object detection. Specifically, a contrastive patch matching task is proposed that exploits multiple levels of the feature hierarchy, as required for downstream detection tasks.

The paper [A3] by Huang et al. presents a training algorithm, i.e., Selfadaptive training, that dynamically leverages model predictions to guide and enhance the training process. The authors claim that their method can advance both supervised learning and SelfSupervised Learning (SelfSL) and can improve the generalization of deep networks under noisy data, such as data corrupted by label noise, random noise, and adversarial examples.

The paper [A4] authored by Yang et al. addresses the challenge of removing rain streaks and rain accumulation from videos. The proposed solution is an augmented Selflearned Deraining Network (SLDNet+), which eliminates the need for ground-truth images during training. SLDNet+ exploits the temporal correlation and consistency of adjacent frames to predict a clean version of the current frame. In addition, priors of rain videos are incorporated to guide the network to focus on rain regions and remove rain accumulation.

The paper [A5] by Wang et al. presents a novel variational inference model, i.e., a selfsupervised latent space optimization method called nebula variational coding, to improve the deep learning architectures. It forms clusters in the latent space through the additional variables trained in an unsupervised way and employs metric learning to separate the clusters in a selfsupervised way further. It can be applied to different architectures

Date of current version 5 February 2024.
Digital Object Identifier 10.1109/TPAMI.2023.3341723

such as VAEs, CNNs, RNNs, and GANs and used to solve different problems, including text sequence, images, 3D point clouds and volumetric data.

The paper [A6] by Xinwang Liu investigates the potential of Multiple Kernel Alignment (MKA) maximization, an effective method for multiple kernel clustering (MKC), for the problem of Incomplete Multiple Kernel Clustering (IMKC). The authors propose to integrate the imputation of incomplete kernel matrices and MKA maximization into a unified learning framework and alternately optimize them until convergence. Theoretical analysis and empirical results indicate the effectiveness of the proposed algorithm.

III. Few Shot Learning and Continual Learning

Four papers study the problem of FSL based on meta-learning. The paper [A7] by Ye et al. addresses two potential weaknesses of meta-learning based few-shot classification and proposes a novel meta-learning paradigm to introduce rich supervision to encourage the few-shot learner to generate classifiers that perform like strong classifiers. The paper [A8] by Baik et al. addresses the limitations of Model-Agnostic Meta-Learning (MAML) and proposes a new task adaptive weight update rule to enhance the fast adaptation process, instead of searching for a better initialization. The paper [A9] by Tsutsui et al. addresses the problem of one-shot fine-grained recognition, and presents a meta-learning framework to combine synthesized images with generative models and original images. The "off the shelf" image generator is updated by a few training instances of novel classes, and a Meta Image Reinforcing Network is proposed to conduct one-shot fine-grained recognition as well as image reinforcement. The paper [A10] by Du et al. proposes a novel method called MetaKernel for FSL by introducing kernel approximation based on random Fourier features into the meta-learning framework.

There are three papers studying the problem of FSL without using meta learning. The paper [A11] by Parket al. addresses the problem of few-shot font generation, and presents a novel font generation method that learns componentwise style representations, instead of universal style representations. To reduce the number of required reference glyphs, the authors represent componentwise styles by a product of component and style factors inspired by low-rank matrix factorization. The article [A12] explores the idea of recognizing complex objects by their components and presents a method called Recognition as Part Composition (RPC) that decomposes images into salient parts and then learns to represent each part instance as a mixture of a few concepts. RPC is beneficial for low-shot generalization tasks, including Few-Shot Learning (FSL), Zero-Shot Learning (ZSL), and UDA. The paper [A13] by Willes et al. studies the problem of few shot open world recognition which extends the scope of the existing open-world recognition setting to include learning with limited labeled data, and presents a highly flexible framework which combines Bayesian nonparametric class priors with an embedding-based pretraining scheme.

Two papers study the problem of Zero Shot Learning (ZSL). The paper [A14] by Yan et al. studies the problem of Zero shot object detection (ZSD) and proposes a novel semantic-guided contrastive detection framework (ContrastZSD) using contrastive learning. ContrastZSD incorporates two subnetworks that contrast between (region, category) and (region, region) pairs, respectively. The paper [A15] by Mancini et al. addresses the problem of open world Compositional ZSL (CZSL), i.e., recognizing unseen compositions of state and object visual primitives seen during training, and propose a novel model, i.e., Compositional Cosine Graph Embeddings, which models the relationship between primitives and compositions through a graph CNN.

Two papers study two scenarios of class incremental learning. The paper titled [A16] by Mazumder et al. addresses the problem of task incremental learning under the vanilla and zero-shot settings, and presents a novel approach, i.e.; Rectification-based Knowledge Retention (RKR), that applies weight rectifications and affine transformations for adapting the model to any task. The paper [A17] by Zhao et al. addresses the problem of Few Shot Class Incremental Learning by introducing a fast and slow learning process to manage the balance between learning new concepts and remembering old ones. Distinct feature spaces are introduced with different learning rates, and frequency-aware distillation is applied after decomposing the feature space by DCT.

IV. Semisupervised Learning

The paper [A18] by Zhu et al. addresses the annotation bottleneck in semantic segmentation by introducing new self-training based techniques for semi-supervised learning. The authors adopt a hard-pseudo label teacher-student approach and introduce centroid sampling as a technique to deal with class imbalance and associated bias amplification. They then introduce a coarse-to-fine training schedule to enable large scale semi-supervised learning for segmentation to be conducted quickly and efficiently.

The paper [A19] by Ali Akbari et al. aims to address the accuracy drop of facial age estimation in data domain shifts. propose a novel method, namely Robust Age Estimation (RAgE), to improve the robustness and reduce the uncertainty of age estimates based on a new similarity-preserving pseudo-labeling algorithm and a noise-tolerant consistency regularisation technique. The authors claim their method is robust to confounding external factors, including variations in head pose, illumination, and expression.

V. Weakly Supervised Learning

Two papers address the problem of weakly supervised semantic segmentation with image-level supervision only. The paper [A20] by Lee et al. presents AdvCAM (Adversarial Class Activation Map), an attribution map of an image that is manipulated to increase the classification score (produced by a classifier before the final softmax layer), allowing it to identify more regions occupied by an object rather than limited to a small discriminative region of the object. The paper [A21] by Wang et al. explores the value of cross-image semantic relations for comprehensive object pattern mining rather than using intra-image information as in existing works. The authors propose a co-attention classification network where a co-attention is exploited to mine the common semantics within an image pair, and a contrastive co-attention is utilized to emphasize exclusive and unshared semantics for capturing complementary supervision cues.

The paper [A22] by Yu et al. brings the variational information bottleneck (VIB) principle to graph convolutional network learning. Specifically, they propose a bilevel optimization algorithm to identify a compact sub-graph that has minimum mutual information with the input graph while being maximally informative about the output labels. This informative sub-graph recognition strategy requires no subgraph annotation and can be used for a variety of tasks, including improving graph classification, graph denoising, and 3D segmentation.

VI. Cross Domain Learning

Three papers address the problem of Unsupervised Domain Adaptation (UDA) based semantic segmentation. The paper [A23] by Dong et al. addresses the problem by considering "what", "where" and "how" to transfer, and presents a novel Knowledge Aggregation induced Transferability Perception Adaptation Network (KATPAN) based on three key modules to determine "what", "where", and "how" respectively. The paper [A24] by Jing et al. presents a feature representation method that can encapsulate both the knowledge from the source domain and the target domain and minimize the domain gap. In addition, a reliable pseudo-label retraining strategy is also proposed. The paper [A25] by Ma et al. proposes a unified framework named "Image to Feature" (I2F) by performing both image level and feature level adaptation. The authors propose a global photometric alignment module and a global texture alignment module to address image-level domain shifts and perform global manifold alignment to address feature level domain shifts.

The paper [A26] by Shiyu Xuan and Shiliang Zhang addresses the problem of unsupervised person Reidentification by explicitly reducing the domain discrepancy across different cameras. The proposed method progressively learns features that are robust to intracamera and intercamera variations by decomposing the training into two stages: the intracamera training stage and the intercamera training stage.

The paper [A27] by Xu et al. addresses the problem of Multisource Domain Adaptation (MSDA) for image classification by exploring graphical models to model the graphical structure among different domains effectively. The authors explore two types of graphical models for MSDA, i.e., Conditional Random Field and Markov Random Field for MSDA, and claim that MRF is more expressive and computationally efficient than CRF.

The paper [A28] by Zhang et al. studies the problem of domain adaptive object detection under across different weather conditions and proposes to introduce causal intervention to explore weather condition invariant feature representations. The authors first construct a confounder dictionary which stores prototypes of representative object features under various scenarios, and then use the causal intervention to discover the causal association between objects.

The paper [A29] by Gu et al. brings together recent ideas in spherical learning (i.e., mapping the data onto a sphere by magnitude normalizing the embeddings) with domain adaptation. Specifically, the seminal adversarial domain adaptation strategy is extended with spherical discriminators and classifiers, and pseudo-label based domain adaptation is extended with a robust loss that weights pseudo-labels by estimated probability of correctness.

The paper [A30] by Peng et al. addresses the problem of single domain generalization. A method named meta learning based Adversarial Domain Augmentation is proposed. Adversarial domain augmentation with Wasserstein auto-encoder aims at creating multiple augmented domains from the source domain to mimic unseen domains. Uncertainty quantification is used for broad and safe domain generalization to avoid risk from out-of-domain. Finally, Meta-Learning is used for training the source domain and augmented domains.

The paper [A31] by Wang et al. addresses the annotation bottleneck in monocular 6D object pose estimation through self-supervision. The authors pretrain a fully supervised pose estimator using simulated data, and then use a variety of self-supervised techniques to adapt the model to the real-data domain without using any labels. They consider a teacher-student unsupervised adaptation framework with losses based on the implied mask and perceptual appearance as well as geometric alignment based on point matching and chamfer distance.

VII. Applications

The paper [A32] by Yin et al. brings together ideas from few-shot meta-learning to improve the efficiency of black-box adversarial attacks. They establish a correspondence between images and tasks and attacks with outcomes as training examples. Then they consider learning an image-conditional adversarial perturbation generator to generate the attacks. In this way, the perturbation generator can be learned from historical images/tasks and then rapidly and efficiently fine-tuned for each new image in order to generate strong attacks with few queries.

The paper [A33] by Su et al. studies the problem of using brain images to accurately screen mental illness. Brain images from patients are limited, whereas brain images from healthy people are abundant. To deal with the class imbalance and domain shift problems in diagnostic classification, the authors propose few-shot domain-adaptive anomaly detection to achieve cross-site anomaly detection of brain images. The proposed method achieves effective cross-site diagnostic classification based on only a few labeled samples.

The paper [A34] by Liu et al. presents a unified learning paradigm that can simultaneously deal with long-tailed realistic data and actively explore novel data with humans in the loop. The authors formally define the Open LongTailed Recognition++ (OLTR++) that can handle imbalanced classification, few-shot learning, open-set recognition, and active learning in one integrated algorithm.

The paper [A35] by Feng Liu and Xiaoming Liu presents a novel paradigm for unsupervised dense 3D object shape correspondence by leveraging the implicit function representation. The implicit function produces a probabilistic embedding to represent each 3D point in a part embedding space and implements dense correspondence through an inverse function mapping from the part embedding vector to a corresponding 3D point.

The paper [A36] by Weijian Deng and Liang Zheng investigates the problem of estimating classifier accuracy on unlabeled test datasets, and casts it as the problem of Automatic model Evaluation (AutoEval). To address this, the authors formulate AutoEval as a dataset-level regression problem.

The paper [A37] by Gao et al. casts the visual object tracking from a single initial exemplar in the testing phase as a few shot online adaptation problem, and proposes a simple yet effective recursive *least squares estimator* aided online learning approach for few-shot online adaptation without requiring offline training.

Li Liu, *Guest Editor*
College of Electronic Science and Technology
National University of Defense Technology
Changsha, Hunan 410073, China
dreamliu2010@gmail.com

Timothy Hospedales
University of Edinburgh
EH8 9YL Edinburgh, U.K.
t.hospedales@ed.ac.uk

Yann LeCun
New York University
New York, NY 10012 USA
yann@fb.com

Mingsheng Long
Tsinghua University
Beijing 100190, China
mingsheng@tsinghua.edu.cn

Jiebo Luo
University of Rochester
Rochester, NY 14627 USA
jluo@cs.rochester.edu

Wanli Ouyang
Shanghai AI Laboratory
Shanghai 200240, China
wanli.ouyang@sydney.edu.au

Matti Pietikäinen
Center for Machine Vision and Signal Analysis
University of Oulu
90570 Oulu, Finland
matti.pietikainen@oulu.fi

Tinne Tuytelaars
KU Leuven
3000 Leuven, Belgium
tinne.tuytelaars@esat.kuleuven.be

Acknowledgments

This work was supported by the National Key Research and Development Program of China No. 2021YFB3100800, the National Natural Science Foundation of China under Grant 62376283, and the Key Stone grant (JS2023-03) of the National University of Defense Technology (NUDT).

Appendix
Related Articles

[A1] Yanbei Chen, Massimiliano Mancini, Xiatian Zhu, and Zeynep Akata, “Semi-Supervised and Unsupervised Deep Visual Learning: A Survey,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1327–1347, Mar. 2024.

[A2] Jian Ding, Enze Xie, Hang Xu, Chenhan Jiang, Zhenguo Li, Ping Luo, and Gui-Song Xia, “Deeply Unsupervised Patch Re-Identification for Pre-Training Object Detectors,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1348–1361, Mar. 2024.

[A3] Lang Huang, Chao Zhang, and Hongyang Zhang, “Self-Adaptive Training: Bridging Supervised and Self-Supervised Learning,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1362–1377, Mar. 2024.

[A4] Wenhan Yang, Robby T. Tan, Shiqi Wang, Alex C. Kot, and Jiaying Liu, “Learning to Remove Rain in Video With Self-Supervision,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1378–1396, Mar. 2024.

[A5] Yida Wang, David Joseph Tan, Nassir Navab, and Federico Tombari, “Self-Supervised Latent Space Optimization With Nebula Variational Coding,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1397–1411, Mar. 2024.

[A6] Xinwang Liu, “Incomplete Multiple Kernel Alignment Maximization for Clustering,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1412–1424, Mar. 2024.

[A7] Han-Jia Ye, Lu Ming, De-Chuan Zhan, and Wei-Lun Chao, “Few-Shot Learning With a Strong Teacher,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1425–1440, Mar. 2024.

[A8] Sungyong Baik, Myungsub Choi, Janghoon Choi, Heewon Kim, and Kyoung Mu Lee, “Learning to Learn Task-Adaptive Hyperparameters for Few-Shot Learning,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1441–1454, Mar. 2024.

[A9] Satoshi Tsutsui, Yanwei Fu, and David Crandall, “Reinforcing Generated Images via Meta-Learning for

One-Shot Fine-Grained Visual Recognition," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1455–1463, Mar. 2024.

[A10] Yingjun Du, Haoliang Sun, Xiantong Zhen, Jun Xu, Yilong Yin, Ling Shao, and Cees G. M. Snoek, "MetaKernel: Learning Variational Random Features With Limited Labels" *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1464–1478, Mar. 2024.

[A11] Song Park, Sanghyuk Chun, Junbum Cha, Bado Lee, and Hyunjung Shim, "Few-Shot Font Generation With Weakly Supervised Localized Representations," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1479–1495, Mar. 2024.

[A12] Samarth Mishra, Pengkai Zhu, and Venkatesh Saligrama, "Interpretable Compositional Representations for Robust Few-Shot Generalization," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1496–1512, Mar. 2024.

[A13] John Willes, James Harrison, Ali Harakeh, Chelsea Finn, Marco Pavone, and Steven L. Waslander, "Bayesian Embeddings for Few-Shot Open World Recognition," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1513–1529, Mar. 2024.

[A14] Caixia Yan, Xiaojun Chang, Minnan Luo, Huan Liu, Xiaoqin Zhang, and Qinghua Zheng, "Semantics-Guided Contrastive Network for Zero-Shot Object Detection," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1530–1544, Mar. 2024.

[A15] Massimiliano Mancini, Muhammad Ferjad Naeem, Yongqin Xian, and Zeynep Akata, "Learning Graph Embeddings for Open World Compositional Zero-Shot Learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1545–1560, Mar. 2024.

[A16] Pratik Mazumder, Pravendra Singh, Piyush Rai, and Vinay P. Namboodiri, "Rectification-Based Knowledge Retention for Task Incremental Learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1561–1575, Mar. 2024.

[A17] Hanbin Zhao, Yongjian Fu, Mintong Kang, Qi Tian, Fei Wu, and Xi Li, "MgSvF: Multi-Grained Slow versus Fast Framework for Few-Shot Class-Incremental Learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1576–1588, Mar. 2024.

[A18] Yi Zhu, Zhongyue Zhang, Chongruo Wu, Zhi Zhang, Tong He, Hang Zhang, R. Manmatha, Mu Li, and Alexander Smola, "Improving Semantic Segmentation via Efficient Self-Training," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1589–1602, Mar. 2024.

[A19] Ali Akbari, Muhammad Awais, Soroush Fatemifar, Syed Safwan Khalid, and Josef Kittler, "RAgE: Robust Age Estimation Through Subject Anchoring With Consistency Regularisation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1603–1617, Mar. 2024.

[A20] Jungbeom Lee, Eunji Kim, Jisoo Mok, and Sungroh Yoon, "Anti-Adversarially Manipulated Attributions for Weakly Supervised Semantic Segmentation and Object Localization," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1618–1634, Mar. 2024.

[A21] Wenguan Wang, Guolei Sun, and Luc Van Gool, "Looking Beyond Single Images for Weakly Supervised Semantic Segmentation Learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1635–1649, Mar. 2024.

[A22] Junchi Yu, Tingyang Xu, Yu Rong, Yatao Bian, Junzhou Huang, and Ran He, "Recognizing Predictive Substructures With Subgraph Information Bottleneck," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1650–1663, Mar. 2024.

[A23] Jiahua Dong, Yang Cong, Gan Sun, Zhen Fang, and Zhengming Ding, "Where and How to Transfer: Knowledge Aggregation-Induced Transferability Perception for Unsupervised Domain Adaptation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1664–1681, Mar. 2024.

[A24] Jing Li, Kang Zhou, Shenhan Qian, Wen Li, Lixin Duan, and Shenghua Gao, "Feature Re-Representation and Reliable Pseudo Label Retraining for Cross-Domain Semantic Segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1682–1694, Mar. 2024.

[A25] Haoyu Ma, Xiangru Lin, and Yizhou Yu, "I2F: A Unified Image-to-Feature Approach for Domain Adaptive Semantic Segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1695–1710, Mar. 2024.

[A26] Shiyu Xuan and Shiliang Zhang, "Intra-Inter Domain Similarity for Unsupervised Person Re-Identification," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1711–1726, Mar. 2024.

[A27] Minghao Xu, Hang Wang, and Bingbing Ni, "Graphical Modeling for Multi-Source Domain Adaptation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1727–1741, Mar. 2024.

[A28] Hua Zhang, Liqiang Xiao, Xiaochun Cao, and Hassan Foroosh, "Multiple Adverse Weather Conditions Adaptation for Object Detection via Causal Intervention," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1742–1756, Mar. 2024.

[A29] Xiang Gu, Jian Sun, and Zongben Xu, "Unsupervised and Semi-Supervised Robust Spherical Space Domain Adaptation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1757–1774, Mar. 2024.

[A30] Xi Peng, Fengchun Qiao, and Long Zhao, "Out-of-Domain Generalization From a Single Source: An Uncertainty Quantification Approach," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1775–1787, Mar. 2024.

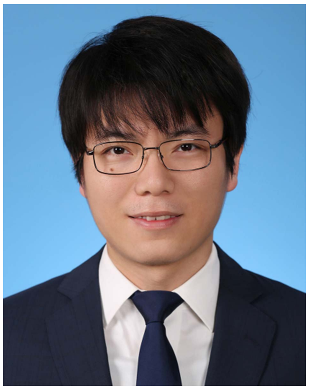
[A31] Gu Wang, Fabian Manhardt, Xingyu Liu, Xiangyang Ji, and Federico Tombari, "Occlusion-Aware Self-Supervised Monocular 6D Object Pose Estimation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1788–1803, Mar. 2024.

[A32] Fei Yin, Yong Zhang, Baoyuan Wu, Yan Feng, Jingyi Zhang, Yanbo Fan, and Yujiu Yang, "Generalizable Black-Box Adversarial Attack With Meta Learning," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1804–1818, Mar. 2024.
[A33] Jianpo Su, Hui Shen, Limin Peng, and Dewen Hu, "Few-Shot Domain-Adaptive Anomaly Detection for Cross-Site Brain Images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1819–1835, Mar. 2024.
[A34] Ziwei Liu, Zhongqi Miao, Xiaohang Zhan, Jiayun Wang, Boqing Gong, and Stella X. Yu, "Open Long-Tailed Recognition in a Dynamic World," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1836–1851, Mar. 2024.
[A35] Feng Liu and Xiaoming Liu, "Learning Implicit Functions for Dense 3D Shape Correspondence of Generic Objects," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1852–1867, Mar. 2024.
[A36] Weijian Deng and Liang Zheng, "AutoEval: Are Labels Always Necessary for Classifier Accuracy Evaluation?" *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1868–1880, Mar. 2024.
[A37] Jin Gao, Yan Lu, Xiaojuan Qi, Yutong Kou, Bing Li, Liang Li, Shan Yu, and Weiming Hu, "Recursive Least-Squares Estimator-Aided Online Learning for Visual Tracking," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 46, no. 3, pp. 1881–1897, Mar. 2024.

Li Liu received the PhD degree from the National University of Defense Technology (NUDT), China, in 2012. She has held research visits at the University of Waterloo (Canada), the Chinese University of Hong Kong and the University of Oulu (Finland). She served as the leading guest editor of the special issues for *IEEE Transactions on Pattern Analysis and Machine Intelligence* and *International Journal of Computer Vision*, and currently serves as an associate editor for *IEEE Transactions on Circuits and Systems for Video Technology*, *IEEE Transactions on Geoscience and Remote Sensing*, and *Pattern Recognition*. Her research interests include computer vision, remote sensing image analysis, pattern recognition, and machine learning. Her papers currently have more than 12,000 citations, according to Google Scholar.

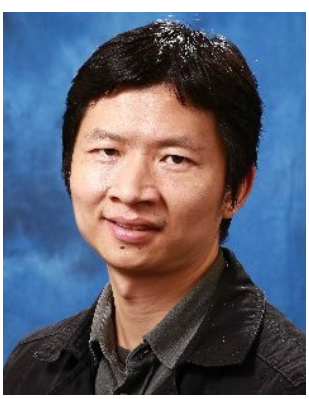
Timothy Hospedales (Senior Member, IEEE) is a professor with the University of Edinburgh and head of Samsung AI Research Centre, Cambridge. He is an associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and has served as area chair of major venues including CVPR, ICCV, ECCV, NeurIPS, and ICML, program chair of BMVC'18, and Associate Program chair of AAAI'22. He he has published extensively data-efficient learning, including few-shot learning, domain adaptation, meta-learning, and so on. His work has won prizes or nominations at ICML, BMVC, ICPR, and ICML AutoML.

Yann LeCun is VP and chief AI Scientist with Facebook and Silver professor with NYU affiliated with the Courant Institute and the Center for Data Science. He was the founding director of Facebook AI Research and of the NYU Center for Data Science. His research interests include machine learning and artificial intelligence, with applications to computer vision, natural language understanding, robotics, and computational neuroscience. He is best known for his work in deep learning and the invention of the convolutional neural network which has been widely used for image, video, and speech recognition. He has received numerous important awards. He is a member of the US National Academy of Engineering. He is the recipient of the 2018 ACM Turing Award (with Geoffrey Hinton and Yoshua Bengio) for "conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing".



Mingsheng Long received the BE and PhD degrees from Tsinghua University, in 2008 and 2014 respectively. He is an Associate Professor (with tenure) in the School of Software, Tsinghua University. He was a visiting researcher with UC Berkeley from 2014 to 2015. He serves as an associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence* and the *Artificial Intelligence Journal*, and as (Senior) Area Chairs of major machine learning conferences, including ICML, NeurIPS, and ICLR. His research is dedicated to machine learning theory, algorithms, and models, with special interests in transfer learning and domain adaptation, deep learning and foundation models, scientific learning and world models.

Jiebo Luo is professor of computer science with the University of Rochester which he joined in 2011 after a prolific career of fifteen years at Kodak Research Laboratories. He has authored nearly 600 technical papers and holds more than 90 U.S. patents. His research interests include computer vision, NLP, machine learning, data mining, computational social science, and digital health. He has served on the editorial boards of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Multimedia*, *IEEE Transactions on Circuits and Systems for Video Technology*, *IEEE Transactions on Big Data*, *ACM Transactions on Intelligent Systems and Technology*, and *Pattern Recognition*. He was the EIC of the IEEE TMM from 2020-2022. He is a fellow of ACM, AAAI, IEEE, SPIE, and IAPR.



Wanli Ouyang received the PhD degree in the Department of Electronic Engineering with the Chinese University of Hong Kong. He is now a professor with Shanghai AI Lab. His research interests include pattern recognition, machine learning, and AI for Science. Before that, he was an Associate Professor with The University of Sydney. He served as the associate editor of IJCV and PR, the senior area chair of CVPR, and the guest editor of TPAMI.

Matti Pietikäinen received the DrSc (Eng.) degree from the University of Oulu, Finland. He is now an emeritus professor with the Center for Machine Vision and Signal Analysis, with the University of Oulu. He has made fundamental contributions, e.g., to Local Binary Pattern (LBP) methodology, texture based image and video analysis, and facial image analysis. He has been associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Information Forensics and Security*, *etc*. He has also been a guest editor for a special session of TPAMI and a special issue for IJCV. In 2014, his research on LBP based face description was awarded the Koenderink Prize for fundamental contributions to computer vision. He was the recipient of the IAPR King-Sun Fu Prize 2018 for fundamental contributions to texture analysis and facial image analysis. He is a life fellow of the IEEE. His papers have more than 87,000 citations in Google Scholar.

Tinne Tuytelaars (Member, IEEE) received the master of electrical engineering degree from the KU Leuven, Belgium, in 1996. She is a professor with KU Leuven Belgium. She has been one of the program chairs for ECCV 2014 and one of the general chairs of CVPR 2016. She served as associate editor in chief of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and is a member of the editorial board of computer vision and image understanding. She served as General cochair of CVPR2016, Program cochair of ECCV2014 and CVPR2021, and area chair for ECCV(2008, 2010, 2012, 2018), CVPR (2009, 2011, 2013), ICCV (2011, 2013), and BMVC (2017). She received the Koenderink Award with ECCV 2016 for fundamental contributions in computer vision that stood the test of time, and the CVIU Most Cited Paper Award in 2011. Her research interests are object recognition, action recognition, multimodal analysis, and image representations.