

Editorial

Guest Editorial: Foreword to the Special Issue on Advanced Research and Applications of Deep Learning and Neural Network in Image Recognition

Ganggang Dong ^{1,*} , Yuanxin Ye ²  and Zhongling Huang ³ ¹ National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China² Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China; yeyuanxin@home.swjtu.edu.cn³ School of Automation, Northwestern Polytechnical University, Xi'an 710072, China; huangzhongling@nwpu.edu.cn

* Correspondence: dongganggang@nudt.edu.cn; Tel.: +86-029-88201246

Over the last two decades, the realm of image recognition has undergone a remarkable transformation, characterized by an astonishing pace of advancement. Object detection performance, once stagnant at around 30-percentage in mean average precision (mAP), has now impressively soared to 90-percentage and beyond, particularly in benchmarks like PASCAL VOC. Equally remarkable is the surpassing of human capabilities by contemporary learning algorithms, achieving unprecedented accuracy in image classification tasks such as those found in the ImageNet dataset. These strides in image classification signal a new era with profound implications for practical applications, spanning areas like video surveillance, autonomous driving, intelligent healthcare, remote sensing image interpretation, and artificial intelligence.

At the core of this revolutionary progress in image recognition lie deep learning algorithms, whose success is rooted in two critical factors: the availability of extensive training datasets and the computational prowess of state-of-the-art platforms. Notably, deep neural networks consistently outperform meticulously crafted manual image features across a spectrum of image tasks. However, amidst the resounding success achieved by deep learning in image recognition, numerous challenges persist, emphasizing the necessity for further exploration and innovation.

This Special Issue goes beyond mere acknowledgment of these challenges; it is devoted to showcasing novel solutions poised to overcome these hurdles. By delving into the intricacies of these issues, our aim is to contribute to the ongoing discourse and advancement in the ever-evolving field of image recognition. This collection of research endeavors not only commemorates achievements thus far but also charts the course for the next frontier in the continuous evolution of image recognition technologies.

For this Special Issue, which was open to all researchers, 12 papers have undergone a meticulous review process and have been ultimately accepted for publication. These papers cover a broad spectrum of vision task topics as follows.

- Point cloud deep learning. The efficacy of 3D vision surpasses that of 2D vision due to its enhanced spatial and depth perception capabilities. Point cloud representation, commonly utilized in relevant applications, preserves the original geometric information in three-dimensional space without discretization. However, the inherent disorder of point clouds poses challenges for their integration into deep learning frameworks. In 2017, deep convolutional neural networks were used to address the sparsity and disorder issues associated with point clouds, forming the famous models, PointNet and PointNet++. Various kinds of improved models were presented subsequently, leading to the expanded application of point clouds in areas such as 3D object detection (Contributor 1), pose estimation (Contributor 2), and more (Contributor 3).



Citation: Dong, G.; Ye, Y.; Huang, Z. Guest Editorial: Foreword to the Special Issue on Advanced Research and Applications of Deep Learning and Neural Network in Image Recognition. *Electronics* **2024**, *13*, 557. <https://doi.org/10.3390/electronics13030557>

Received: 21 January 2024

Accepted: 26 January 2024

Published: 30 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

- Pixelwise semantic segmentation. Semantic segmentation is a fundamental task in computer vision field. It refers to classify each pixel in an image, enabling segmentation into semantically distinct regions. In 2015, the full convolutional network (FCN) was presented, in which the task of semantic segmentation were revolutionized by employing transposed convolutions. The fully connected layers were replaced by the convolutional layers to achieve end-to-end segmentation. The presented fashion have been expanded across various fields. Weng et al. enhanced the DeepLabV3+ model, addressing railway track extraction errors with morphological algorithm optimization (Contributor 4). Zheng et al. applied multi-scale semantic segmentation to fire smoke, incorporating global information (Contributor 5). These advancements showcase the widespread impact of semantic segmentation in diverse fields.
- Zeroshot learning. The common supervised learning tasks were often struggled with the limited or the unlabeled examples. This challenge was tackled by zero-shot learning, in which the transferable representations can be exploited. The discriminative and semantic-relevant features can be therefore encompassed into the learned representations. In some researchers, the semantic-relevant representations were emphasized through visual-semantic alignment. Likewise, the discrimination techniques for broader generalization were focused on in the other studies. Recently, the shared representations between these sub-tasks were targeted. In this special issue, Wang et al. introduces a novel partially-shared multi-task representation method preserving complementary knowledge, namely PS-GZSL (Contributor 6). The emerging algorithms like federated learning and contrast learning are used too to offer new solutions to zero-sample learning (Contributor 7).
- Model optimization. A significant challenge in the current advancement of deep learning lies in the extensive computation and parameters involved. It is infeasible to deploy the resource-intensive convolutional neural networks (CNNs) on computing devices with limitations, such as embedded systems and mobile devices (Contributor 8). To tackle these issues, considerable research efforts have been dedicated to compression techniques, including channel pruning, low-rank decomposition, and weight quantization. In this special issue, a new trick via dynamic pruning and layer fusion is presented to optimize the deep model (Contributor 9). Through the incorporation of knowledge distillation and short-long fine-tuning, the redundant layers with minimal accuracy loss can be eliminated. The primary objective is to reduce memory access more significantly than reducing computational complexity.
- Multimodal applications. In the real-world visual tasks, a multi-dimensional framework is involved, encompassing spatial, temporal, and modal dimensions. Spatially, tasks may be spanned from image-level and region-level to pixel-level assignments. Temporally, the challenges are then extended beyond the static images to include the processing of time-series videos. In terms of modalities, the inputs and the outputs can be displayed in a variety of fashions, such as images, text, videos, or other types like body poses (Contributor 10) and depth maps (Contributor 11). So, it refers to another important research fields, the data engineering (Contributor 12). Given the diverse range of application scenarios (Contributor 13), it is challenging to achieve the universality for model design. Consequently, the future development trajectory of deep visual systems will focus on constructing more versatile models capable of accommodating a wide array of input and output types, effectively addressing the varied demands arising from different scenarios.

In the end, we would like to express our gratitude to these authors who have dedicated their efforts to in-depth research in the field of computer vision. Their contributions are of significant importance in addressing current challenges. Additionally, we extend our thanks to all the reviewers for their time, dedication, and valuable insights during the evaluation process. This helps ensure the selection of high-quality papers in accordance with standards.

Funding: This work was supported by the National Natural Science Foundation of China under Grant 61971324, 61525105, and the foundation of National Key Laboratory of Radar Signal Processing under Grant JKW202310.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Contributions

1. Zhang, L.; Meng, H.; Yan, Y.; Xu, X. Transformer-Based Global PointPillars 3D Object Detection Method. *Electronics* **2023**, *12*, 3092.
2. Wang, Q.; Lei, H.; Qian, W. Siamese PointNet: 3D Head Pose Estimation with Local Feature Descriptor. *Electronics* **2023**, *12*, 1194.
3. Wang, Q.; Qian, W. Z.; Lei, H.; Chen, L. Siamese Neural PointNet: 3D Face Verification under Pose Interference and Partial Occlusion. *Electronics* **2023**, *12*, 620.
4. Weng, Y.; Li, Z.; Chen, X.; He, J.; Liu, F.; Huang, X.; Yang, H. A Railway Track Extraction Method Based on Improved DeepLabV3+. *Electronics* **2023**, *12*, 3500.
5. Zheng, Y.; Wang, Z.; Xu, B.; Niu, Y. Multi-Scale Semantic Segmentation for Fire Smoke Image Based on Global Information and U-Net. *Electronics* **2022**, *11*, 2718.
6. Wang, G.; Tang, S. Generalized Zero-Shot Image Classification via Partially-Shared Multi-Task Representation Learning. *Electronics* **2023**, *12*, 2085.
7. Yu, L.; Huang, J. Cyclic Federated Learning Method Based on Distribution Information Sharing and Knowledge Distillation for Medical Data. *Electronics* **2022**, *11*, 4039.
8. Zhao, M.; Li, M.; Peng, S.L.; Li, J. A Novel Deep Learning Model Compression Algorithm. *Electronics* **2022**, *11*, 1066.
9. Li, Q.; Li, H.; Meng, L. Deep Learning Architecture Improvement Based on Dynamic Pruning and Layer Fusion. *Electronics* **2023**, *12*, 1208.
10. Gong, Z.; Zhang, Y.; Lu, D.; Wu, T. Vision-Based Quadraped Pose Estimation and Gait Parameter Extraction Method. *Electronics* **2022**, *11*, 3702.
11. Han, H.; Liu, J.; Wang, W.; Gao, C.; Shi, J. An Improved CNN for Polarization Direction Measurement. *Electronics* **2023**, *12*, 3723.
12. Liu, S.; Zhang, L.; Liu, W.; Hu, J.; Gong, H.; Zhou, X.; Gong, D. RERB: A Dataset for Residential Area Extraction with Regularized Boundary in Remote Sensing Imagery for Mapping Application. *Electronics* **2022**, *11*, 2790.
13. Chen, X.; Feng, X.; Li, Y. An Image Unmixing and Stitching Deep Learning Algorithm for In-Screen Fingerprint Recognition Application. *Electronics* **2023**, *12*, 3768.

Short Biography of Authors



Ganggang Dong received the M.S. and Ph.D. degrees in information and communication engineering from National University of Defense Technology, Changsha, China, in 2012 and 2016, respectively. His research interests include, but not limited to deep learning, SAR imaging, radar target detection and recognition, cognitive radio, radar image interpretation. He has authored more than 40 scientific papers in peer-reviewed journals and conferences, including IEEE TIP, IEEE TCYB, IEEE TGRS, IEEE TIM, and Pattern Recognition. Dr. Dong is currently an associate professor with Xidian University. He received more than 1410 citations in Google Scholar. He was awarded the 2017 Excellent Doctoral Thesis of the Chinese Institute of Electronics. He served as a reviewer for some top-tier journals on remote sensing and image processing.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.