electronics

Editorial

Future AI and Robotics: Visual and Spatial Perception Enhancement and Reasoning

Wenfeng Zheng [1,*], Mingzhe Liu [2], Chao Liu [3] and Dan Wang [4]

[1] School of Automation, University of Electronic Science and Technology of China, Chengdu 610054, China
[2] College of Computer Science and Cyber Security, Chengdu University of Technology, Chengdu 610059, China; liumz@cdut.edu.cn
[3] LIRMM, UMR5506, University of Montpellier-CNRS, 34095 Montpellier, France; liu@lirmm.fr
[4] Department of Internal Medicine, Division of Nephrology, The Ohio State University, Columbus, OH 43210, USA; wang.15506@osu.edu
* Correspondence: winfirms@uestc.edu.cn

Over the past several decades, artificial intelligence (AI) has been tremendously boosted by new algorithm designs, exponentially increased computing power, and an immense volume of calculation materials (i.e., data). Nevertheless, appropriate feature fusion and high-level, abstract forms of knowledge representation are required to help AI to achieve better results, as the primary goal of AI research is to enable machines to perform complex tasks that would typically require human intelligence.

Restoration and perception enhancement techniques are active research areas in robotics which play essential roles in helping us to perceive and understand the world. Their applications include human activity recognition, surgical medicine, geoinformatics, and remote sensing analysis.

Artificial intelligence based on computer vision has been greatly strengthened and developed, becoming one of the most important developing areas in robotics. Object recognition, classification, segmentation, topology, network, efficiency, navigation, and search based on spatial attributes are also anticipated to become important and valuable fields of development in artificial intelligence and robotics in the future.

Recently, intelligent reasoning has been widely used to address the significant technical issues involved in implementing AI in real-world applications, such as intelligent medical care, environmental analysis and prediction, autonomous driving, intelligent transportation, text classification, recommended systems, machine translation, and analog dialogues.

In this Special Issue, we present groundbreaking research and case studies that demonstrate the future applications of and advances in artificial intelligence and robotics, especially regarding visual and spatial perception enhancement and reasoning.

Yungyao Chen et al. (Contribution 1) introduce HDRFormer, an innovative framework designed to enhance high dynamic range (HDR) image quality in edge cloud-based video surveillance systems. Leveraging advanced deep learning algorithms and Internet of Things (IoT) technology, HDRFormer employs a unique architecture comprising a feature extraction module (FEM) and a weighted attention module (WAM). The FEM leverages a transformer-based hierarchical structure to adeptly capture multiscale image information. In addition, guided filters are utilized to steer the network, thereby enhancing the structural integrity of the images. On the other hand, the WAM focuses on reconstructing saturated areas, improving the perceptual quality of the images and rendering natural, saturated reconstructed HDR images. In addition, the framework exhibits outstanding performance in multiscale structural similarity (MS-SSIM) and HDR visual difference predictor (HDR-VDP2.2). The proposed method not only outperforms the existing HDR reconstruction techniques [1,2], but also offers better generalization capabilities, laying a robust foundation for future applications in smart cities.

check for updates

Citation: Zheng, W.; Liu, M.; Liu, C.; Wang, D. Future AI and Robotics: Visual and Spatial Perception Enhancement and Reasoning. *Electronics* **2023**, *12*, 4787. https://doi.org/10.3390/electronics12234787

Received: 21 November 2023
Accepted: 24 November 2023
Published: 26 November 2023

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

Jiawei Tian et al. (Contribution 2) developed a novel calibration algorithm that capitalizes on the unique attributes of binocular endoscopes [3]. By integrating principles of monocular camera calibration, their proposed algorithm effectively eliminates vertical disparity while retaining horizontal disparity in stereo images. This not only simplifies the subsequent stereo matching operation, but also meets stringent accuracy standards, as evidenced by robust experimental validation. Moreover, their investigation into the 3D cardiac soft tissue surface reconstruction method has yielded promising results. The utilization of a stereo endoscope vision system [4] in conjunction with dense parallax images has facilitated accurate 3D coordinate acquisition within the left endoscope coordinate system. The subsequent surface reconstruction process, employing a dual-pass filter and the Delaunay triangulation method, has proven highly effective in generating a detailed and accurate representation of the surface of cardiac soft tissue. Their experimental validation demonstrated that the reconstructed 3D spatial points closely align with manually obtained coordinates, with deviations falling well within acceptable error margins.

Yanhua Liu et al. (Contribution 3) proposed a dose image reconstruction method based on tensor sparse dictionary learning. Specifically, they combine tensor coding with compressed sensing data, extend 2D dictionary learning to 3D by using a tensor product, and then utilize the spatial information of X-ray acoustic signals more efficiently [5,6]. To reduce the artifacts of reconstruction images caused by spare sampling, they design an alternate iterative solution of the tensor sparse coefficient and tensor dictionary. In addition, they build an X-ray-induced acoustic dose image reconstruction system, simulate X-ray acoustic signals based on information from patients from the Sichuan Cancer Hospital, and then create simulated datasets. Compared to typical, state-of-the art imaging methods, their experimental results demonstrate that this method can significantly improve the quality of reconstructed images and the accuracy of dose distribution.

Runxin Liu et al. (Contribution 4) propose a heterogeneous quasi-continuous spiking cortical model (HQC-SCM) method for neutron and gamma-ray pulse shape discrimination. Their method utilizes specific neural responses for different features inside radiation pulse signals [7], fully extracting features present in the falling edge and delayed fluorescence parts. Subsequently, the contributions of the HQC-SCM's parameters to its discrimination performance are studied to identify an automatic parameter selection strategy for the HQC-SCM. Since the HQC-SCM is a chaotic system that cannot be optimized by traditional algorithms like gradient descent [8], a GA-based parameter optimization method is proposed. Experiments are then conducted to evaluate the performance of this optimization method in finding the local optima of the HQC-SCM's parameter solutions. It is found that the GA can optimize solutions in a way closer to stochastic searching, which is suitable for local optima searching in a chaotic system like the HQC-SCM. This GA-based optimization method is efficient and robust, locating local optima in just a few iterations of evolution.

Shengliang Cai et al. (Contribution 5) propose an energy-based semantic augmented segmentation (ESAS) model, a new approach to cross-modality image segmentation [9,10] which leverages the energy of the support modality's latent semantic features to generate semantic comparative modality information. This is a novel, generalizable method that can be applied to most unpaired multimodal image learning tasks. To achieve it, they developed a framework that involves the use of a pre-trained model with shared parameters, which is then used to train an energy-based model that leverages modality-shared knowledge. They conducted experiments on the MM-WHS 2017 dataset to evaluate the performance of this method. The results of their experiments demonstrate that their proposed approach is effective in improving the segmentation performance of query modality images by incorporating prior knowledge from supporting modality images. Overall, their novel framework could provide a valuable contribution to the field of cross-modality image segmentation and can potentially be applied to a range of medical imaging applications.

Qiuxiang Gu et al. (Contribution 6) construct a parallel platform and its mechanical structure design. Based on a microprogrammed control unit (MCU) + pre-driver chip + three-phase full bridge solution, they complete the circuit design of the motor driver.

The programs of the MCU and the the parallel platform control center are developed to drive six parallel robotic arms, and system joint debugging is also completed, achieving a closed-loop control effect in the parallel platform workspace. As a result, they create a physical platform with low costs and a flexible structure. It can easily replace the component of the parallel platform control center [11,12] to create other algorithms and test their effectiveness.

Feng Xiong et al. (Contribution 7) presents a novel generalized knowledge distillation framework to overcome the limitations of missing modalities in glioma segmentation, particularly in unimodal scenarios. Their framework successfully extracts rich knowledge from a multimodal segmentation model [13,14] and transfers it to a unimodal segmentation model, enhancing its performance. They introduce two knowledge distillation strategies—segmentation map distillation and cascade region attention distillation—to effectively transfer multimodal knowledge from the teacher model. The segmentation map distillation strategy enables the student model to mimic the teacher's output and acquire segmentation capabilities. In contrast, the cascade region attention distillation strategy employs label masks to concentrate on local features and allows the student model to focus on essential knowledge without being distracted by superfluous feature information. Notably, their proposed framework requires less training effort than alternative methods and demonstrates superior segmentation performance in unimodal scenarios.

We would like to thank all the authors and reviewers for their contributions to this Special Issue. We hope that this Special Issue and the seven articles that comprise it will help readers to better understand artificial intelligence and robotics and inspire further research and new results within this exciting domain.

Author Contributions: Conceptualization, W.Z., M.L., C.L. and D.W.; writing—original draft preparation, W.Z. and C.L.; writing—review and editing, W.Z., M.L., C.L. and D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

List of Contributions:

1. Chen, Y.-Y.; Hsia, C.-H.; Jhong, S.-Y.; Lai, C.-F. Attention-Guided HDR Reconstruction for Enhancing Smart City Applications. *Electronics* **2023**, *12*, 4625. https://doi.org/10.3390/electronics12224625.
2. Tian, J.; Ma, B.; Lu, S.; Yang, B.; Liu, S.; Yin, Z. Three-Dimensional Point Cloud Reconstruction Method of Cardiac Soft Tissue Based on Binocular Endoscopic Images. *Electronics* **2023**, *12*, 3799. https://doi.org/10.3390/electronics12183799.
3. Liu, Y.; Liu, M.; Jiang, X.; Liu, X.; Liu, M. Dose Images Reconstruction Based on X-ray-Induced Acoustic Computed Tomography. *Electronics* **2023**, *12*, 2241. https://doi.org/10.3390/electronics12102241.
4. Liu, R.; Liu, H.; Yang, B.; Gu, B.; Yin, Z.; Liu, S. Heterogeneous Quasi-Continuous Spiking Cortical Model for Pulse Shape Discrimination. *Electronics* **2023**, *12*, 2234. https://doi.org/10.3390/electronics12102234.
5. Cai, S.; Shen, C.; Wang, X. Energy-Based MRI Semantic Augmented Segmentation for Unpaired CT Images. *Electronics* **2023**, *12*, 2174. https://doi.org/10.3390/electronics12102174.
6. Gu, Q.; Tian, J.; Yang, B.; Liu, M.; Gu, B.; Yin, Z.; Yin, L.; Zheng, W. A Novel Architecture of a Six Degrees of Freedom Parallel Platform. *Electronics* **2023**, *12*, 1774. https://doi.org/10.3390/electronics12081774.
7. Xiong, F.; Shen, C.; Wang, X. Generalized Knowledge Distillation for Unimodal Glioma Segmentation from Multimodal Models. *Electronics* **2023**, *12*, 1516. https://doi.org/10.3390/electronics12071516.

References

1. Banterle, F.; Ledda, P.; Debattista, K.; Chalmers, A. Inverse Tone Mapping. In Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia, Perth, Australia, 1–4 December 2006; pp. 349–356.
2. Landis, H. Production-Ready Global Illumination. In Proceedings of the International Conference on Computer Graphics and Interactive Techniques, San Antonio, TX, USA, 21–26 July 2002; pp. 93–95.
3. Marr, D.; Poggio, T.; Hildreth, E.C.; Grimson, W.E.L. A computational theory of human stereo vision. In *From the Retina to the Neocortex: Selected Papers of David Marr*; Vaina, L., Ed.; Birkhäuser Boston: Boston, MA, USA, 1991; pp. 263–295.
4. Zhang, Y.-J. Camera calibration. In *3D Computer Vision: Principles, Algorithms and Applications*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 37–65.
5. Lee, D.; Park, E.Y.; Choi, S.; Kim, H.; Min, J.j.; Lee, C.; Kim, C. GPU-accelerated 3D volumetric X-ray-induced acoustic computed tomography. *Biomed. Opt. Express* **2020**, *11*, 752–761. [CrossRef] [PubMed]
6. Xiang, L.; Han, B.; Carpenter, C.; Pratx, G.; Kuang, Y.; Xing, L. X-ray acoustic computed tomography with pulsed X-ray beam from a medical linear accelerator. *Med. Phys.* **2013**, *40*, 10701. [CrossRef] [PubMed]
7. Roush, M.L.; Wilson, M.A.; Hornyak, W.F. Pulse shape discrimination. *Nucl. Instrum. Methods* **1964**, *31*, 112–124. [CrossRef]
8. Liu, H.; Liu, M.; Xiao, Y.; Li, P.; Zuo, Z.; Zhan, Y. Discrimination of neutron and gamma ray using the ladder gradient method and analysis of filter adaptability. *Nucl. Sci. Tech.* **2022**, *33*, 159. [CrossRef]
9. Han, Z.; Chen, Q.; Zhang, L.; Mo, X.; You, J.; Chen, L.; Fang, J.; Wang, F.; Jin, Z.; Zhang, S.; et al. Radiogenomic association between the t2-flair mismatch sign and idh mutation status in adult patients with lower-grade gliomas: An updated systematic review and meta-analysis. *European Radiol.* **2022**, *32*, 5339–5352. [CrossRef] [PubMed]
10. Zhou, C.; Ding, C.; Lu, Z.; Wang, X.; Tao, D. One-pass multi-task convolutional neural networks for efficient brain tumor segmentation. In *Part III 11, Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, 16–20 September 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 637–645.
11. Galván-Pozos, D.; Ocampo-Torres, F. Dynamic analysis of a six-degree of freedom wave energy converter based on the concept of the Stewart-Gough platform. Renew. *Energy* **2020**, *146*, 1051–1061. [CrossRef]
12. Dai, X.; Song, S.; Xu, W.; Huang, Z.; Gong, D. Modal space neural network compensation control for Gough-Stewart robot with uncertain load. *Neurocomputing* **2021**, *449*, 245–257. [CrossRef]
13. Wang, Y.; Zhang, Y.; Hou, F.; Liu, Y.; Tian, J.; Zhong, C.; Zhang, Y.; He, Z. Modality-pairing learning for brain tumor segmentation. In Proceedings of the International MICCAI Brainlesion Workshop, Lima, Peru, 4 October 2020; pp. 230–240.
14. Dolz, J.; Gopinath, K.; Yuan, J.; Lombaert, H.; Desrosiers, C.; Ayed, I.B. HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation. *IEEE Trans. Med. Imaging* **2018**, *38*, 1116–1126. [CrossRef]