Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.087043

EDITORIAL

Introduction to the Special Issue on Advances in AI-Driven Computational Modeling for Image Processing

Sathishkumar Veerappampalayam Easwaramoorthy*

Faculty of Engineering and Technology, School of Computing and Artificial Intelligence, Sunway University, No. 5, Jalan Universiti, Bandar Sunway, Selangor, Malaysia

*Corresponding Author: Sathishkumar Veerappampalayam Easwaramoorthy. Email: sathishv@sunway.edu.my

Received: 09 June 2026; Accepted: 16 June 2026; Published: 27 July 2026

Image processing has become one of the most consequential frontiers of artificial intelligence (AI). The convergence of deep learning [1], large-scale data, and accelerated computing has transformed how machines perceive, interpret, and reconstruct visual information, enabling capabilities that were unattainable with classical, hand-engineered pipelines. Convolutional neural networks first demonstrated that hierarchical feature learning could surpass manually designed descriptors on large-scale recognition tasks [2], deep residual learning subsequently enabled the stable optimization of substantially deeper networks [3], and the emergence of attention-based and transformer architectures has further reshaped the field by modeling long-range spatial dependencies that earlier models struggled to capture [4,5]. In parallel, generative modeling—propelled by generative adversarial networks (GANs) [6] and, more recently, diffusion-based frameworks [7]—has redefined the boundaries of image synthesis, restoration, and cross-modal translation.

Despite this rapid progress, translating AI-driven models into dependable, deployable systems remains challenging. Real-world imagery is rarely clean: it is degraded by noise, rain, blur, low illumination, and domain shift; annotated data are often scarce, particularly in medical and remote-sensing settings, despite the success of label-efficient architectures such as U-Net [8] and the recent emergence of promptable segmentation foundation models [9]; and high-stakes applications increasingly demand not only accuracy but also robustness, calibrated uncertainty, computational efficiency, and interpretability. Addressing these requirements calls for tighter integration between AI and the computational modeling principles—optimization, probabilistic reasoning, and physically grounded representations—that underpin reliable engineering systems.

This Special Issue, *Advances in AI-Driven Computational Modeling for Image Processing*, was conceived to provide a platform for research that advances this integration. It brings together nine contributions (eight original research articles and one comprehensive review) that collectively span medical image analysis, image restoration and enhancement, object detection in remote sensing, content-based image retrieval, multimodal image generation, and uncertainty quantification. The contributions are unified by a common theme: leveraging AI not as a black box, but as a component of carefully designed computational frameworks that target accuracy, robustness, and practical relevance. In the following, we briefly summarize each contribution and situate it within the broader landscape of the field.

In [contribution 1], the authors introduce TransCP-Net, a transformer-based framework for the early screening of infant cerebral palsy through spatiotemporal pose representation learning. The model employs hierarchical spatial and temporal attention to analyze motion patterns in video sequences and integrates multi-modal information to improve detection. By targeting subtle motor abnormalities that conventional clinical assessment often misses before twelve months of age, the work illustrates how attention-based modeling of pose dynamics can support timely, objective screening for a critical neurodevelopmental condition.

In [contribution 2], the authors present a comprehensive survey of GANs for medical images. Following a structured, PRISMA-guided methodology, the review synthesizes peer-reviewed studies published between 2020 and 2025 and examines the role of GAN architectures in image synthesis, denoising, super-resolution, data augmentation, and cross-modality translation. The article consolidates the advances, applications, and limitations of generative models in medical imaging, offering a valuable reference for researchers seeking to address the perennial challenge of limited annotated data in clinical domains.

In [contribution 3], the authors address single-image deraining with a stage-wise dual-residual network augmented by cross-dimensional spatial attention. The framework progressively restores rain-degraded content while preserving fine structural detail, combining multi-scale feature extraction, channel attention, and ground-truth-guided spatial attention to stabilize training and mitigate feature degradation across layers. The work demonstrates how carefully structured residual architectures can deliver efficient and high-fidelity restoration for a degradation that significantly impairs downstream vision tasks.

In [contribution 4], the authors propose EffNet-CNN, a semantic model that integrates an EfficientNet backbone with a convolutional neural network for image mining and content-based image retrieval. By jointly improving classification and retrieval performance, the approach targets the efficiency and accuracy demands of retrieval systems across healthcare, security, and related domains. The study underscores the continued importance of designing semantically rich representations for retrieval, a task that remains central to organizing and exploiting ever-growing image collections.

In [contribution 5], the authors develop a framework that integrates speech input into image generation by coupling speech-to-text translation with a CLIP + VQGAN model. Recognizing that humans naturally communicate visualization prompts through speech rather than text, the system translates spoken prompts into text that conditions the generative model. The work represents a step toward more natural, accessible human-machine interaction in generative image synthesis and highlights the growing role of multimodal pipelines in creative and assistive applications.

In [contribution 6], the authors propose a nature-inspired framework for glaucoma diagnosis that combines a feature excitation-based dense segmentation network with an improved grey-wolf-optimization-supported support vector machine. A fuzzy-integral technique is used to segment the optic cup and optic disc from retinal images even under varying backgrounds and intensity inconsistencies, after which the optimized classifier performs diagnosis. The two-stage design illustrates how metaheuristic optimization can be combined with deep segmentation to improve robustness in ophthalmic screening.

In [contribution 7], the authors couple YOLOv9 with a transformer head (YOLOv9-TH) for small-object detection in remote-sensing imagery. An additional prediction head is introduced to detect objects of varying sizes, and the original heads are replaced with transformer heads to exploit self-attention. By addressing scale variation, densely packed objects, and motion blur, the work targets persistent obstacles in aerial and satellite image analysis, where small and crowded targets routinely challenge conventional detectors.

In [contribution 8], the authors integrate a Bayesian convolutional neural network for cataract detection from fundus images, with an emphasis on uncertainty quantification. By assigning probability values to predictions and comparing CNN and BCNN models, the study moves beyond point estimates toward distributions over potential outcomes, enabling decisions that explicitly account for predictive uncertainty. The work reflects a broader and increasingly important shift toward trustworthy AI in clinical decision support, where confidence calibration is as critical as raw accuracy.

In [contribution 9], the authors propose an enhanced lung-cancer detection approach based on a dual-model deep learning technique. By combining complementary models, the method aims to capture a broader range of features in medical imaging data than single-model pipelines, improving both the detection and localization of tumors at early stages. Given that early and accurate detection is decisive for patient outcomes, the study highlights the value of ensemble and multi-model strategies in high-stakes diagnostic imaging.

Taken together, these nine contributions reflect both the breadth and the maturity of contemporary AI-driven image processing. Several recurring themes emerge across the issue: the central role of attention and transformer mechanisms in capturing complex spatial and temporal structure; the importance of generative modeling for synthesis, restoration, and data augmentation; the growing emphasis on uncertainty quantification and trustworthiness, particularly in medical applications; and the persistent need to combine learning with classical computational principles such as optimization and probabilistic reasoning. Medical imaging features prominently, underscoring the field's potential to deliver tangible societal benefit, while contributions in remote sensing, image retrieval, and multimodal generation demonstrate the versatility of these methods across application domains.

As Guest Editors, we believe these works will help clarify the current state of AI-driven computational modeling for image processing and stimulate further research toward models that are not only accurate but also robust, efficient, interpretable, and deployable in the real world. We thank all the authors for their valuable contributions, the reviewers for their careful and constructive evaluations, and the editorial staff of *Computer Modeling in Engineering & Sciences* for their support throughout the preparation of this Special Issue.

Funding Statement: The authors received no specific funding for this study.

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

List of Contributions:

1. Ksibi A, Ayadi M, Elmannai H, Hamdi M, Alluhaidan AS, Ksibi I. TransCP-net: transformer-based spatiotemporal pose representation for early screening of infant cerebral palsy. Comput Model Eng Sci. 2026;147(2):1–10. doi:10.32604/cmes.2026.078347.
2. Sagheer SVM, Nimitha U, Ameer PM, Parayangat M, Abbas M, Arunachalam KP. A survey of generative adversarial networks for medical images. Comput Model Eng Sci. 2026;146(2):1–10. doi:10.32604/cmes.2025.067108.
3. Wang T, Hu Z. Efficient image deraining through a stage-wise dual-residual network with cross-dimensional spatial attention. Comput Model Eng Sci. 2025;145(2):2357–81. doi:10.32604/cmes.2025.073640.
4. Thanikachalam R, Muniasamy A, Alasmari A, Thavasimuthu R. EffNet-CNN: a semantic model for image mining & content-based image retrieval. Comput Model Eng Sci. 2025;143(2):1971–2000. doi:10.32604/cmes.2025.063063.
5. Mahajan S, Gite S, Pradhan B, Alamri A, Inamdar S, Shriyansh D, et al. Integrating speech-to-text for image generation using generative adversarial networks. Comput Model Eng Sci. 2025;143(2):2001–26. doi:10.32604/cmes.2025.058456.
6. Latif J, Wajahat A, Tahir A, Bilal A, Zakariah M, Alnuaim A. A nature-inspired AI framework for accurate glaucoma diagnosis. Comput Model Eng Sci. 2025;143(1):539–67. doi:10.32604/cmes.2025.062301.
7. Barr M. Coupling the power of YOLOv9 with transformer for small object detection in remote-sensing images. Comput Model Eng Sci. 2025;143(1):593–616. doi:10.32604/cmes.2025.062264.
8. Muniasamy A, Alasmari A. Integrating Bayesian and convolution neural network for uncertainty estimation of cataract from fundus images. Comput Model Eng Sci. 2025;143(1):569–92. doi:10.32604/cmes.2025.060484.
9. Elhassan SM, Darwish SM, Elkaffas SM. An enhanced lung cancer detection approach using dual-model deep learning technique. Comput Model Eng Sci. 2025;142(1):835–67. doi:10.32604/cmes.2024.058770.

References

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. doi:10.1038/nature14539.
2. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90. doi:10.1145/3065386.
3. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27–30; Las Vegas, NV, USA. p. 770–8. doi:10.1109/CVPR.2016.90.
4. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16 × 16 words: transformers for image recognition at scale. In: Proceedings of the International Conference on Learning Representations (ICLR); 2021 May 3–7; Vienna, Austria.
5. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: a survey. ACM Comput Surv. 2022;54(10s):1–41. doi:10.1145/3505244.
6. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Commun ACM. 2020;63(11):139–44. doi:10.1145/3422622.
7. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. In: Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2022 Jun 18–24; New Orleans, LA, USA. p. 10674–85. doi:10.1109/CVPR52688.2022.01042.
8. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention—MICCAI 2015. Cham, Switzerland: Springer International Publishing; 2015. p. 234–41. doi:10.1007/978-3-319-24574-4_28.
9. Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. Segment anything. In: Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision (ICCV); 2023 Oct 1–6; Paris, France. p. 3992–4003. doi:10.1109/ICCV51070.2023.00371.