



## OPEN ACCESS

EDITED AND REVIEWED BY  
Rashid Ibrahim Mehmood,  
Islamic University of Madinah, Saudi Arabia

\*CORRESPONDENCE  
Bo Huang  
✉ huangbo0326@cqu.edu.cn

RECEIVED 05 November 2025  
ACCEPTED 17 November 2025  
PUBLISHED 27 November 2025

CITATION  
Huang B, Zhang D and Liu Q (2025) Editorial:  
Advances and challenges in AI-driven visual  
intelligence: bridging theory and practice.  
*Front. Artif. Intell.* 8:1740331.  
doi: 10.3389/frai.2025.1740331

COPYRIGHT  
© 2025 Huang, Zhang and Liu. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Editorial: Advances and challenges in AI-driven visual intelligence: bridging theory and practice

Bo Huang<sup>1\*</sup>, Dawei Zhang<sup>2</sup> and Qiao Liu<sup>3</sup>

<sup>1</sup>College of Optoelectronic Engineering, Chongqing University (CQU), Chongqing, China,

<sup>2</sup>Department of Computer Science and Technology, Zhejiang Normal University, Jinhua, China,

<sup>3</sup>National Center for Applied Mathematics, Chongqing Normal University, Chongqing, China

## KEYWORDS

AI-driven visual intelligence, adversarial attacks, semi-supervised learning, 3D reconstruction, non-invasive diagnostic

## Editorial on the Research Topic

[Advances and challenges in AI-driven visual intelligence: bridging theory and practice](#)

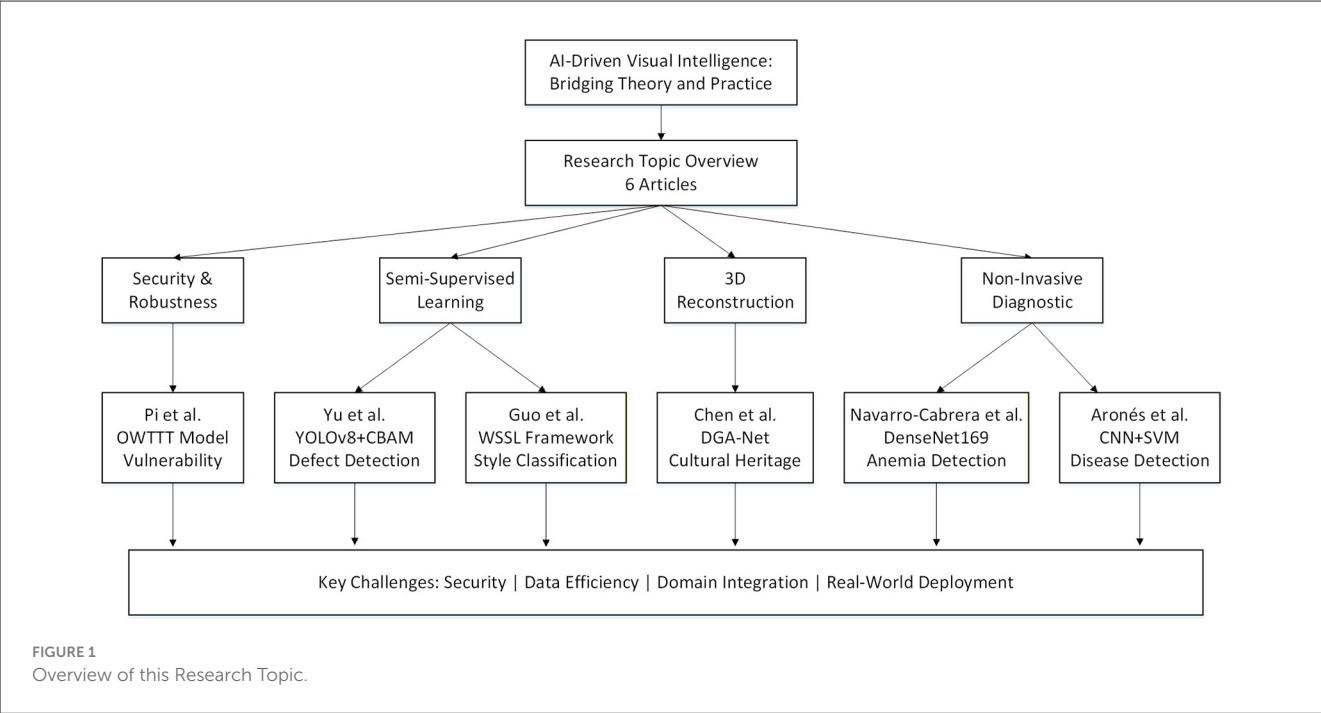
Visual intelligence has become a fundamental research area in artificial intelligence, driven by rapid advances in deep learning and increasing demands from practical applications. This Research Topic brings together cutting-edge research that addresses both the remarkable progress and the persistent challenges in implementing AI-driven visual intelligence systems in real-world applications. The overview of this Research Topic is shown in [Figure 1](#). The six articles published in this collection exemplify the breadth and depth of current research, spanning from fundamental security concerns to practical applications across diverse fields.

## Robustness against adversarial attacks

As AI systems increasingly operate in dynamic, open-world scenarios, their vulnerability to adversarial manipulation presents significant concerns. [Pi et al.](#) investigate test-time poisoning attacks targeting open-world test-time training (OWTTT) models. Their findings reveal that these models can be compromised with merely 100 queries using single-step query-based attack strategies. More critically, the affected models show limited recovery capability even when subsequently exposed to normal samples. This work serves as a crucial reminder that as we pursue adaptive AI systems, we must simultaneously address their security vulnerabilities to ensure safe deployment in critical applications.

## Semi-supervised learning frameworks

The scarcity of labeled training data remains a persistent bottleneck in visual intelligence applications. Two articles tackle this limitation through innovative semi-supervised approaches. [Yu et al.](#) propose an enhanced YOLOv8 framework that embeds the CBAM attention mechanism in high-level networks while incorporating Mean Teacher semi-supervised learning to address data labeling challenges. Their method achieves robust detection of micron-scale defects in industrial polymer films. [Guo et al.](#) introduce a weight-aware semi-supervised self-ensembling



framework (WSSL) for interior decoration style classification. By employing an adaptive weighting module based on truncated Gaussian functions, their approach selectively leverages reliable unlabeled data while mitigating confirmation bias from unreliable pseudo-labels. Both works demonstrate that semi-supervised learning can substantially reduce annotation costs without compromising performance.

### 3D reconstruction and cultural heritage

The preservation of cultural landscapes presents unique challenges that require accurate 3D reconstruction capabilities. [Chen et al.](#) develop DGA-Net, which combines deep feature extraction with graph structure representation for reconstructing historical garden landscapes. The architecture incorporates attention mechanisms to emphasize ecologically significant features, enabling more accurate restoration planning and cultural heritage protection. This work exemplifies how AI-driven visual intelligence can contribute to environmental conservation and the preservation of cultural heritage sites.

### Non-invasive diagnostic applications

Two articles demonstrate the potential of visual intelligence for accessible healthcare and agricultural monitoring. [Navarro-Cabrera et al.](#) apply DenseNet169 for detecting iron deficiency anemia in university students through fingernail image analysis. Their approach achieves 71.08% accuracy with 74.09% AUC, offering a cost-effective alternative to conventional blood tests, particularly relevant for developing regions. [Aronés et al.](#) present App2, a mobile application combining CNN and SVM architectures for apple leaf disease detection. The system

achieves 95% accuracy on clear images and maintains 80% performance under real-field conditions after model adaptation. Notably, their inclusion of validation filters to verify leaf presence reduces false detections, demonstrating attention to practical deployment considerations. Both works highlight how AI-based visual intelligence can address accessibility gaps in healthcare and agriculture.

### Open challenges and future directions

The collected works reveal several persistent challenges in visual intelligence research. First, the security of adaptive systems requires greater attention, particularly as models operate in adversarial environments. The difficulty in recovering from poisoning attacks suggests that defensive mechanisms must be incorporated at the architectural level. Second, efficient learning from limited labeled data continues to demand innovation, with semi-supervised and self-supervised methods showing promise but requiring careful handling of pseudo-label reliability. Third, the integration of domain knowledge with deep learning architectures remains an underexplored avenue that could enhance both performance and interpretability. Finally, the gap between laboratory performance and real-world deployment—exemplified by the accuracy drop in field conditions—indicates that robustness to environmental variations needs greater emphasis during model development.

The progress documented in this Research Topic demonstrates that effective visual intelligence systems require more than algorithmic sophistication. Success depends equally on addressing practical constraints including computational efficiency, data availability, security vulnerabilities, and deployment accessibility. As the field advances, bridging the gap between theoretical capabilities and practical requirements will necessitate continued

collaboration across disciplines and careful consideration of real-world operating conditions.

that could be construed as a potential conflict of interest.

## Author contributions

BH: Writing – original draft, Writing – review & editing. DZ: Writing – review & editing. QL: Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the National Natural Science Foundation of China (Nos. 62306049, 92471207 and W2421089), the General Program of Chongqing Natural Science Foundation (No. CSTB2023NSCQ-MSX0665), and the Fundamental Research Funds for the Central Universities (No. 2024CDJXY008).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

## Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.