ISSN: 2581-3404 (Online) IF: 5.86 (SJIF)
IJIRTM, Volume-10, Issue-2, March-2026

Emerging Trends and Challenges in Image Processing and Computer Vision for Agriculture: A Multidisciplinary Review

Monika Vishwakarma[1], Dr. Rachana Dubey [2]
Ph.D Research Scholar[1], Assistant Professor[2]
Department of Computer Science[1,2]
Pandit Shambhu Nath Shukla Vishwavidyalaya Shahdol, Madhya Pradesh[1,2]
monikasdl22@gmail.com[1], rachanamishra812@gmail.com[2]

Abstract. *Advances in image processing and computer vision have significantly enabled the shift from traditional large-area farming to data-driven, ultra-precision agriculture. Crop monitoring, disease detection, yield forecasting, and sustainable resource management in agricultural fields will require deep learning, edge computing, and multimodal sensing-powered visual intelligence systems by 2026. This multidisciplinary review paper provides a comprehensive synthesis of recent breakthroughs as well as long-standing challenges in image processing and computer vision in agriculture. Key advances include those investigated in edge-to-cloud architectures, multimodal data fusion, convolutional and transformer-based models, and vegetation indices. This complete review paper highlights ongoing challenges with data robustness and real-world applications, particularly for aquatic crops and this review paper concludes by outlining future research directions and provides valuable insights for researchers in the field of smart agriculture.*

Keywords: Image Processing, Computer Vision, Precision Agriculture, Vision Transformers, Multimodal Fusion, Edge Computing, AIoT, Sustainable Farming.

Introduction

The agricultural sector is experiencing an obscure but profound revolution. Farming practices have been developing away from traditional, experience-based decision-making and toward data-driven and technologically supported systems as digital sensing, artificial intelligence, and automation acquire popularity [1,2]. As a result of this change, image processing (IP) and computer vision (CV) have transformed from supporting instruments to essential technologies that are reshaping how agricultural areas are viewed and managed [3]. To address these constraints, computerized and vision-based systems for monitoring have gained popularity. Improvements in autonomous aerial vehicles (UAVs), advanced satellite imaging, and land-based robotic platforms now allow agricultural landscapes to be monitored on a regular basis and at various spatial scales [4]. When integrated with artificial intelligence and deep learning techniques, visualizations can be used to generate actionable insights about the health of plants, soil health, pest activity, and agricultural productivity [2,3]. Deep learning as well as machine learning are developing quickly, considerably improving the accuracy of agricultural image processing systems. Deep learning models, particularly convolutional neural networks (CNNs), have shown excellent accuracy in crop disease diagnosis, weed identification, yield prediction, and plant phenotyping [1], [4]. These approaches minimized the need for handmade techniques for feature extraction and allowed automatic complete learning from agricultural photos. The most recent developments indicate a shift from standard CNN designs to transformer-based vision models. The Vision Transformer (ViT) architecture included a self-attention mechanism capable of detecting long-range relationships in images [6]. Transformer-based techniques have been effectively applied to remote sensing as well as agricultural monitoring, providing better feature representation and scalability [7]. Another new approach is multimodal image fusion, which combines data from many imaging sources, such as RGB cameras, ultraviolet (UV) sensors, and satellite imagery, to improve resilience and decision accuracy. Multimodal deep learning architectures have enhanced the efficacy of sustainable plant care and agricultural monitoring systems [5], [17], and [19]. These systems use spatial, spectral, and temporal picture features to overcome the constraints of single-source vision models. Furthermore, the combination of edge computing and IoT with computer vision allows for real-time agricultural surveillance in resource-constrained rural contexts [9], [16], [18]. By installing lightweight vision models at the edge, latency and bandwidth constraints can be reduced, allowing for faster field-level decision-making. Although these advances, significant hurdles remain, including dataset scarcity, domain adaption issues, computational limits, model explainability, and scalability across a range of agricultural circumstances. As a result, a thorough multidisciplinary study of new trends and problems in image processing and computer vision for agriculture is required.

Related Work

The use of remote sensing and UAV-based imaging technologies opened the way for agricultural computer vision. Zhang and Kovacs [3] conducted one of the first assessments of tiny unmanned aerial systems for precision agriculture, emphasizing applications in crop monitoring and field variability analysis. Mogili and Deepak [13] investigated drone-based systems in greater detail, stressing the significance of aerial imagery for precision farming. Satellite-based image calibration and radiometric correction have also increased the accuracy of remote sensing data for agricultural studies [14]. Convolutional neural networks became the leading method for agricultural image processing as deep learning gained popularity. Kamilaris and Prenafeta-Boldú [1] conducted a survey on deep learning applications in agriculture, revealing their usefulness in plant disease detection and yield estimate. Similarly, Liakos et al. [4] examined machine learning and deep learning approaches in agriculture, demonstrating that CNNs outperform classical classifiers in image-related tasks. More recently, El Sakka et al. [17], [19] examined CNN applications in smart agriculture utilizing multimodal picture data, demonstrating better accuracy through data fusion. Transformer-based vision models mark a fundamental leap in computer vision research. Dosovitskiy [6] proposed the Vision Transformer architecture, which substitutes convolution with self-attention processes. Aleissaee et al. [7] examined transformer uses in remote sensing and found them suitable for large-scale agricultural images and land-use classification. These structures offer better contextual representation, which is especially relevant in diverse agricultural situations. Multimodal AI systems are being more incorporated into precision agriculture frameworks. Yang et al. [5] examined deep learning in multimodal fusion for sustainable plant care, focusing on the integration of spectral, spatial, and environmental data. Ayanlade et al. [11] presented multimodal artificial intelligence for ultra-precision agriculture, which combines imaging data with environmental and sensor-based inputs to improve crop monitoring performance. Advanced edge-based computer vision systems are becoming increasingly important. Shi et al. [9] first proposed the concept of edge computing to facilitate low-latency applications. Ibrahim et al. [16] suggested fog-based architectures for reducing communication latency in distributed networks. Akhtar et al. [19] presented edge-based smart sensing solutions for agricultural monitoring. Miller et al. [18] examined IoT and AI-driven sensor systems that assist vision-based smart farming. In addition to crop surveillance, computer vision now includes livestock and behavioral analytics. Higaki et al. [12] used supervised machine learning on multimodal sensor data to identify estrus in cattle, highlighting the broad utility of vision-driven AI systems. As a whole, the literature shows a clear transition from classical image processing techniques to CNN-based models, transformer structures, and multimodal fusion systems that incorporate edge computing. However, issues such as generalization, data diversity, computing efficiency, and real-world application remain open research topics.

- **Multidisciplinary Foundations of Agricultural Computer Vision**

Agricultural computer vision is intrinsically multidisciplinary, resulting from the intersection of agronomy, computer science, robotics, and environmental science. Agronomy supplies the biological framework for interpreting visible signs such leaf discolouration, canopy framework, and growth variability in terms of plant health, nutritional status, and stress situations [3], [13]. Without this domain expertise, image-based evaluation would lack contextual significance in agricultural settings. Machine learning and deep learning algorithms help to turn raw visual data into useful insights. Convolutional neural networks (CNNs) and transformer-based architectures allow for automated feature extraction and large-scale image analysis in crop monitoring and disease detection [1], [4], [6], and [7]. Multimodal fusion approaches improve model resilience by combining spatial and spectral data from various imaging sources [5, 17]. Automation and edge-enabled systems enable real-time implementation of vision methods in dynamic field environments [9], [16]. Meanwhile, environmental science assures that these vision-driven systems promote sustainable farming practices through improved managing resources and ecological monitoring [5], [11], and [18]. Thus, agricultural computer vision develops through interdisciplinary collaboration, combining biological knowledge, computational intelligence, and sustainability practices to enable thoughtful and accountable farming systems.

Image Processing Techniques in Agriculture

The first part of agricultural computer vision is image processing techniques. These approaches translate raw visual observations into quantitative indicators for crop monitoring and condition assessment. While deep learning is increasingly popular and effective in modern systems, classical image processing methods continue to play a significant role especially in preprocessing and large-scale spectral analysis [3], [13].

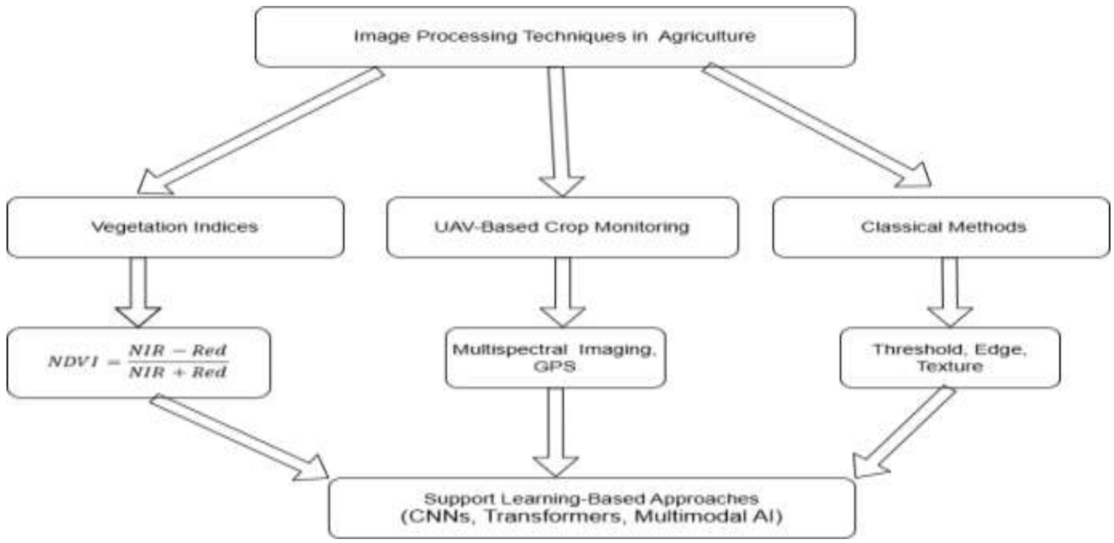

Figure1. Conceptual Framework of Image Processing Techniques Supporting Agricultural Computer Vision.

Vegetation Indices

Vegetation indices were among the first measures used to monitor plant health. They are based on the fact that healthy vegetation reflects NIR light faster than red light [3]. The most widely used index is the Normalized Difference Vegetation Index (NDVI):

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

where $R_{NIR}$is near-infrared reflectance and $R_{Red}$is red-band reflectance. NDVI levels vary from -1 to +1, with higher values indicating healthy vegetation and lower values suggesting stress or bare soil [3], [5]. Although successful, NDVI is sensitive to environmental factors and is frequently integrated with advanced vision models to increase accuracy [17].

UAV-Based Crop Monitoring

Unmanned Aerial Vehicles (UAVs) fitted with multispectral sensors improve crop monitoring by gathering high-resolution imagery across various spectral bands [3] and [13]. These systems can survey huge agricultural fields in a single flight and generate georeferenced data using integrated GPS units. UAV-based NDVI mapping allows farmers to identify stressed zones and implement targeted interventions, integrating traditional image processing with precision agriculture approaches.

Classical Image Processing Methods

Modern image processing techniques including as thresholding, edge detection, texture analysis, and morphological procedures have been widely used in leaf segmentation, weed detection, and fruit counting [4]. These approaches extract basic structural information and perform computationally efficient preprocessing. While less versatile than deep learning algorithms, they are nonetheless useful components of agricultural vision pathways [1], [4].

Deep Learning Approaches in Agricultural Computer Vision

Deep learning has considerably enhanced agricultural computer vision by allowing for automatic the extraction of features and high-quality image analysis. Unlike traditional image processing, deep models learn complicated patterns directly from image input, making them appropriate for real-world variability [1], [4].

Convolutional Neural Networks (CNNs)

CNNs are commonly used to detect plant diseases, identify weeds, count fruits, and classify crops [1], [4]. They extract spatial information like both color and texture variations automatically. According to recent research, combining multicolored and multimodal data boosts CNN performance even more [17], [19].

Transformer-Based Models

Vision Transformers (ViT) depend on self-attention processes to acquire global visual context [6]. These models have demonstrated outstanding outcomes in imagery analysis and agricultural image interpretation, particularly in complicated field contexts [7].

Multimodal and Edge-Based Learning

Multimodal deep learning uses RGB, multispectral, and data on the environment to increase crop monitoring precision [5], [11]. Additionally, edge and fog computing frameworks allow for the actual time deployment of deep learning models in agricultural areas [9], [16], [18].

As a whole, deep learning has helped agricultural vision systems make more scalable and intelligent decisions, however obstacles such as input variety and computing expense remain.

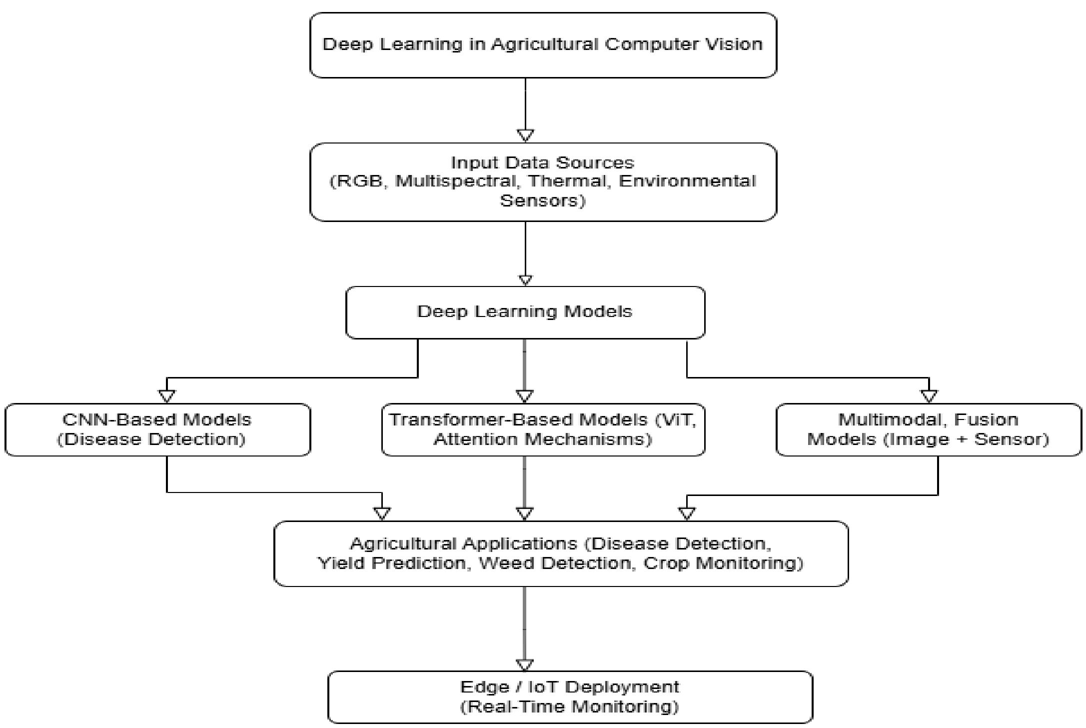

Figure 2. Block diagram of deep learning frameworks in agricultural computer vision.

Key Challenges in Agricultural Image Processing and Computer Vision

Although significant improvements in deep learning, multimodal fusion, and edge computing, various difficulties remain, limiting the wide-ranging implementation of agricultural computer vision systems [1], [5], [9]. Real-world farming conditions are highly dynamic, resulting in challenges with precision, scalability, and ongoing acceptance.

Limited Data Availability and Annotation Complexity

Although significant improvements in deep learning, multimodal fusion, and edge computing, various difficulties remain, limiting the wide-ranging implementation of agricultural computer vision systems [1], [5], [9]. Real-world farming conditions are highly dynamic, resulting in challenges with precision, scalability, and ongoing acceptance.

Generalization and Robustness

Models trained under controlled conditions often fail when applied to different regions, lighting conditions, or crop varieties [3], [4]. Environmental unpredictability and sensor variances impede transferability, limiting practical application across a wide range of agricultural contexts.

Interpretability and Farmer Trust

Deep learning models frequently operate as "black boxes," making decisions difficult to explain [1]. Pesticide, irrigation, and crop management decisions in agriculture have economic and environmental effects, necessitating the deployment of transparent and explainable AI systems to develop user trust (5).

Environmental Variability

Weather variations, shadows, occlusions, and sensor noise all have an impact on agricultural landscapes, reducing image quality and model fidelity [3], [13]. Robust preprocessing and adaptive learning methods are required to handle such real-world heterogeneity.

Research Gaps in Minor and Aquatic Crop Monitoring

Most studies focuses on large terrestrial crops, while tiny or aquatic crops receive little attention [5], [11]. The lack of specific datasets and multimodal frameworks for such crops demonstrates a significant research need in sustainable agriculture.

Future Research Directions

Based on the focused-on trends and obstacles, future research in agricultural image processing and computer vision should concentrate on producing more robust, scalable, and trustworthy systems.

Explainability and Transparency

The next models must include explainable AI (XAI) techniques in order to generate interpretable forecasts. Transparent decision-making will boost farmer trust and enable informed decisions about pesticide use, irrigation, and crop management.

Agriculture-Specific Foundation Models

Large-scale foundation models must be pre-trained on a variety of agricultural datasets that encompass numerous crops, seasons, geographic locations, and sensing conditions. Such models can enhance resilience, transferability, and generalizability across diverse farming contexts.

Real-Time Multimodal AIoT Frameworks

Combining computer vision with environmental sensors, edge computing, and cloud computing systems can provide dependable real-time monitoring and intelligent decision assistance in actual farming applications.

Systematic Evaluation and Benchmarks.

Future research should provide standardized datasets, evaluation metrics, and benchmarking methodologies to ensure that agricultural vision models can be compared fairly and consistently across applications.

Involvement of Underrepresented Farming Systems.

To promote equitable and sustainable agricultural development, research should go beyond major terrestrial crops and encompass minor and aquatic crops, as well as resource-efficient and climate-resilient farming methods.

Conclusion

This review paper discusses how image processing and computer vision have become essential tools in modern agriculture, allowing more precise, data-driven, and sustainable farming operations. Improvements in deep learning, vision transformers, multimodal fusion, and edge computing have improved crop monitoring, plant health assessment, and adaptive field management in a variety of agricultural situations. However, issues like as data scarcity, model generalization, interpretability, and real-world implementation continue to impede large-scale adoption. Addressing these difficulties involves both technical innovation and interdisciplinary teamwork. Future advancement is dependent on combining agronomic skill, powerful artificial intelligence, and ecologically friendly methods. With ongoing research and careful application, agricultural computer vision systems have the potential to significantly increase resilience, resource efficiency, and long-term global food security.

References

1. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and electronics in agriculture, 147, 70-90.
2. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming–a review. Agricultural systems, 153, 69-80.
3. Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. Precision agriculture, 13(6), 693-712.
4. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674.
5. Yang, Z. X., Li, Y., Wang, R. F., Hu, P., & Su, W. H. (2025). Deep Learning in Multimodal Fusion for Sustainable Plant Care: A Comprehensive Review. *Sustainability (2071-1050)*, *17*(12).
6. Dosovitskiy, A. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
7. Amer Aleissaee, A., Kumar, A., Anwer, R. M., Khan, S., Cholakkal, H., & Xia, G. S. (2022). Transformers in Remote Sensing: A Survey. arXiv e-prints, arXiv-2209.

8. Xia, J., Wang, F., Meng, Y., Wu, Z., Wang, L., Lin, W., ... & Li, G. (2018). A computational method for longitudinal mapping of orientation-specific expansion of cortical surface in infants. Medical image analysis, 49, 46-59.
9. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE internet of things journal, 3(5), 637-646.
10. Khan, H. U., Minallah, N., Masood, A., Khalil, A., Frnda, J., & Nedoma, J. (2021). Performance analysis of sphere packed aided differential space-time spreading with iterative source-channel detection. Sensors, 21(16), 5461.
11. Ayanlade, T. T., Jones, S. E., Laan, L. V. D., Chattopadhyay, S., Elango, D., Raigne, J., ... & Sarkar, S. (2024). Multi-modal AI for Ultra-Precision Agriculture. In Harnessing Data Science for Sustainable Agriculture and Natural Resource Management (pp. 299-334). Singapore: Springer Nature Singapore.
12. Higaki, S., Okada, H., Suzuki, C., Sakurai, R., Suda, T., & Yoshioka, K. (2021). Estrus detection in tie-stall housed cows through supervised machine learning using a multimodal tail-attached device. Computers and Electronics in Agriculture, 191, 106513.
13. Mogili, U. R., & Deepak, B. B. V. L. (2018). Review on application of drone systems in precision agriculture. Procedia computer science, 133, 502-509.
14. Lee, L., Wu, C., Qi, C., Hu, X., Yuan, M., Gu, M., ... & Zhang, P. (2021). Solar contamination on HIRAS cold calibration view and the corrected radiance assessment. Remote Sensing, 13(19), 3869.
15. Liu, Y., Chen, Y., Chen, G., & Zhang, J. (2022). Schatten Graph Neural Networks. IEEE Access, 10, 56482-56492.
16. Ibrahim, A. H., Fayed, Z. T., & Faheem, H. M. (2021). Fog-based CDN framework for minimizing latency of web services using fog-based HTTP browser. Future Internet, 13(12), 320.
17. El Sakka, M., Ivanovici, M., Chaari, L., & Mothe, J. (2025). A review of CNN applications in smart agriculture using multimodal data. Sensors, 25(2), 472.Gong, L., et al. (2024). Vision-based precision systems for agriculture. IEEE Transactions on Industrial Informatics.
18. Miller, T., Mikiciuk, G., Durlik, I., Mikiciuk, M., Łobodzińska, A., & Śnieg, M. (2025). The IoT and AI in Agriculture: The Time Is Now—A Systematic Review of Smart Sensing Technologies. Sensors (Basel, Switzerland), 25(12), 3583.
19. El Sakka M, Ivanovici M, Chaari L, Mothe J. A Review of CNN Applications in Smart Agriculture Using Multimodal Data. Sensors (Basel). 2025 Jan 15;25(2):472. doi: 10.3390/s25020472. PMID: 39860841; PMCID: PMC11768470.
20. Srivastava, N., Patil, A., Srivastava, A., Singh, S., & Arya, M. (2023). AZOLLA: AN IMPORTANT BIOFERTILIZER USED TO ENHANCE THE CROP PRODUCTION. IJASS, 859.
21. Luo, X., Xiong, S., Jia, X., Zeng, Y., & Chen, X. (2025). AIoT-enabled data management for smart agriculture: a comprehensive review on emerging technologies. IEEE Access.
22. Jácome-Galarza, L. R. (2022, October). Multimodal deep learning for crop yield prediction. In Doctoral Symposium on Information and Communication Technologies (pp. 106-117). Cham: Springer International Publishing.
23. Adli, H., Remli, M., Wong, K., Ismail, N., González-Briones, A., Corchado, J., & Mohamad, M. (2023). Recent advancements and challenges of AIoT application in smart agriculture: A review. *Sensors*, *23*. https://doi.org/10.3390/s23073752
24. Sharma, K., & Shivandu, S. (2024). Integrating Artificial Intelligence and Internet of Things (IoT) for enhanced crop monitoring and management in precision agriculture. *Sensors International*. https://doi.org/10.1016/j.sintl.2024.100292
25. Liu, L., Cheng, W., & Kuo, H. (2025). A narrative review on smart sensors and IoT solutions for sustainable agriculture and aquaculture practices. *Sustainability*. https://doi.org/10.3390/su17125256

26. Mansoor, S., Iqbal, S., Popescu, S., Kim, S., Chung, Y., & Baek, J. (2025). Integration of smart sensors and IoT in precision agriculture: Trends, challenges and future prospectives. *Frontiers in Plant Science*, *16*. https://doi.org/10.3389/fpls.2025.1587869
27. Dutta, M., Gupta, D., Tharewal, S., Goyal, D., Sandhu, J., Kaur, M., Alzubi, A., & Alanazi, J. (2025). Internet of Things-based smart precision farming in soilless agriculture: Opportunities and challenges for global food security. *IEEE Access*, *13*, 34238–34268. https://doi.org/10.1109/access.2025.3540317
28. Huang, Y., & Khabusi, S. (2025). Artificial Intelligence of Things (AIoT) advances in aquaculture: A review. *Processes*. https://doi.org/10.3390/pr13010073
29. Leong, Y., Lim, E., Subri, N., & Jalil, N. (2023). Transforming agriculture: Navigating the challenges and embracing opportunities of Artificial Intelligence of Things. In *2023 IEEE International Conference on Agrosystem Engineering, Technology & Applications (AGRETA)* (pp. 142–147). https://doi.org/10.1109/agreta57740.2023.10262747.
30. Senoo, E., Anggraini, L., Kumi, J., Karolina, L., Akansah, E., Sulyman, H., Mendonça, I., & Aritsugi, M. (2024). IoT solutions with artificial intelligence technologies for precision agriculture: Definitions, applications, challenges, and opportunities. Electronics. https://doi.org/10.3390/electronics13101894.