Uniosun Journal of Engineering and Environmental Sciences (UJEES), Vol. 8 No. 2, Sept., 2026
DOI: 10.64980/ujees.v8i2.002 **Adebisi et al.**
Published: September 30, 2026

ADVANCES IN MACHINE LEARNING FOR IMAGE ANALYSIS IN MECHATRONIC SYSTEMS: A CROSS-DOMAIN REVIEW AND RESEARCH ROADMAP

Oluwadare Adepeju ADEBISI[1*], Ibiyemi Iyabo ADEAGA[2], Samson Oladayo AYANLADE[3], Khadijah LAWAL[4], Isaac Adewale OJEDOKUN[5], Olusola Joel OYEDELE[6]

[1,5,6]Department of Electrical and Electronics Engineering, Bowen University, College of Agriculture, Engineering, and Science, Iwo, Nigeria.
[2]Department of Computer Engineering, The Polytechnic Ibadan, Nigeria.
[3]Department of Electrical and Electronics Engineering, College of Engineering and Environmental Studies, Olabisi Onabanjo University, Ago-Iwoye, Nigeria.
[4]Department of Mechatronics Engineering, Abiola Ajimobi Technical University, Ibadan, Nigeria

[*]Corresponding Author: oluwadare.adebisi@bowen.edu.ng

Abstract
Image analysis has become a fundamental component of modern mechatronic systems. It enables intelligent perception, decision-making, and autonomous operation across a wide range of industrial applications. However, existing literature remains largely fragmented, with most reviews concentrating on specific application domains or individual algorithmic developments, thereby limiting a comprehensive understanding of how machine learning-based image analysis is advancing across the broader mechatronics landscape. This review provides a cross-domain synthesis of machine learning techniques employed in image analysis for mechatronic systems. A comparative assessment is presented covering conventional machine learning approaches and deep learning methods. The study highlights the strengths, limitations, data requirements, and computational characteristics of these techniques. Major challenges affecting practical deployment are examined. Emerging research directions are also discussed, with particular emphasis on synthetic data generation, foundation models, self-supervised learning, multimodal intelligence, edge artificial intelligence, digital twin technologies, and trustworthy artificial intelligence frameworks. By consolidating current developments and identifying future research priorities, this review provides a structured reference for researchers and practitioners seeking to advance the development of intelligent, adaptive, and reliable mechatronic systems.

Keywords
Machine Learning; Image Analysis; Mechatronic Systems; Computer Vision; Deep Learning; Vision Transformers

1. INTRODUCTION

Mechatronics engineering integrates mechanical, electrical, computer science, and control engineering disciplines. Sensors, actuators, and control units are frequently integrated by mechatronics systems to accomplish complex tasks in industries [1]. In modern mechatronics systems, image analysis is essential because it allows machines to recognize and interpret visual information from their surroundings [2]. Mechatronic systems can perform a variety of functions using image analysis [3]. The ability of traditional image processing techniques to handle complex and variable visual data may be limited because they frequently rely on manually created features and pre-established criteria [4]. Machine learning techniques, on the other hand, are more reliable and adaptable for image analysis tasks because they can automatically learn features [5].

Figure 1 presents a workflow of machine learning-based image analysis. The stages involve transforming raw image data into real-world intelligent applications. Images are acquired from multiple sources such as cameras, industrial vision systems, medical imaging devices, UAV platforms, and sensor-based imaging systems. Acquired images undergo preprocessing operations to improve data quality and consistency. Following preprocessing, image segmentation and region of interest (ROI) extraction are performed to isolate relevant objects and structures within the images. At the preprocessing stage, both traditional handcrafted features and deep learned representations are generated. Images are classified based on the features by machine learning models.

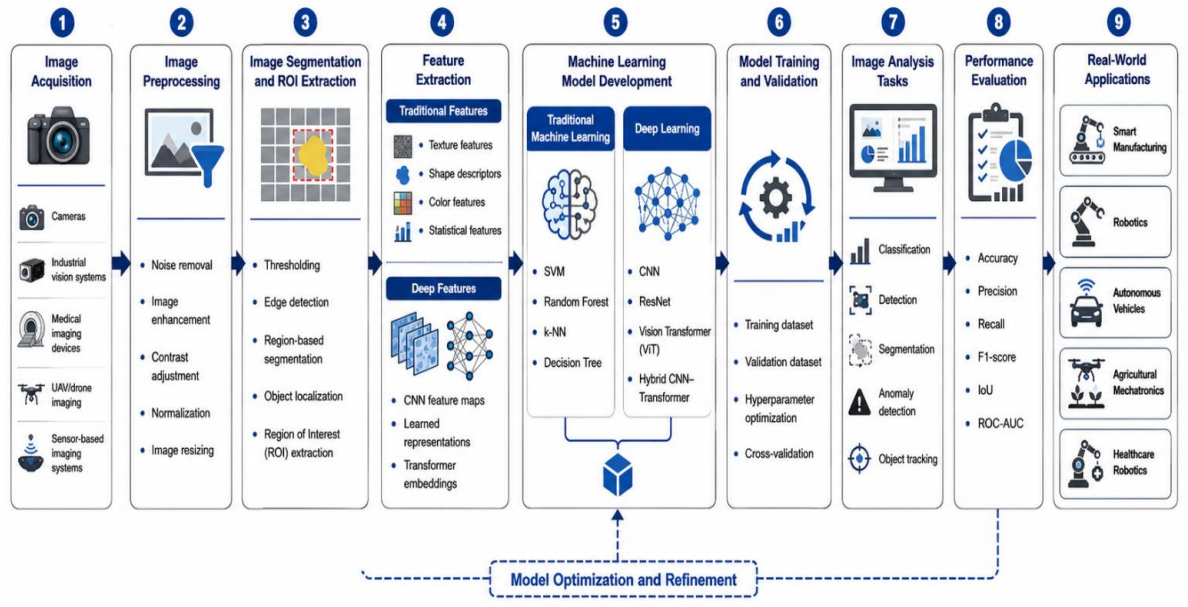


Figure 1: General Workflow of Machine Learning-Based Image Analysis

Advanced machine learning and deep learning algorithms play a major role in integrating image analysis into mechatronics by enabling systems to process and interpret images with better accuracy and efficiency. Mechatronics systems benefit greatly from the integration of machine learning and image analysis since they can improve the systems' perception, processing, and response capability [6]. Mechatronics systems are now able to obtain useful information from images and make deft decisions by using machine learning algorithms, especially deep learning models [7]. Mechatronics systems can become more accurate and efficient at identifying patterns, objects, and characteristics in images by utilizing machine learning approaches. Moreover, mechatronics systems can gradually enhance and optimize their performance through machine learning [7]. These systems can learn from their experiences and improve their algorithms to better handle novel and difficult scenarios by utilizing machine learning techniques [8]. This flexibility is especially important in dynamic settings like robots and autonomous systems, where situations might change suddenly. These systems achieve better levels of performance, adaptability, and intelligence by integrating the power of machine learning algorithms with the rich information offered by visual data [9]. This combination will spur innovation and breakthroughs across a wide range of industries.

Despite the progress achieved in machine learning-based image analysis, many literatures remain fragmented in both scope and depth. Most surveys focus on a single application domain, such as industrial defect inspection, robotic vision, autonomous driving, or medical image analysis [10][11]. While these studies provide valuable domain-specific insights, they often fail to establish connections between developments occurring across different mechatronic fields. Therefore, there is limited understanding of how machine learning techniques perform under diverse operational conditions and application requirements. As intelligent systems converge across different fields, a broader cross-domain assessment is necessary.

Another limitation of the most literature is the disproportionate emphasis placed on deep learning techniques while traditional machine learning methods receive comparatively less attention. Although deep learning has demonstrated remarkable success in many images analysis tasks, conventional machine learning approaches continue to offer practical advantages in scenarios characterized by limited data availability, constrained computational resources, and requirements for model interpretability. Existing reviews rarely provide a systematic comparison of these approaches [12][13]. Furthermore, recent developments involving Vision Transformers, hybrid CNN–Transformer architectures, foundation models, multimodal learning frameworks, edge intelligence, and explainable artificial intelligence have introduced new opportunities and challenges. There is a need for a more holistic and forward-looking assessment that not only examines established methodologies but also evaluates emerging technologies that are expected to shape the future of intelligent mechatronic systems.

Motivated by these observations, the review presents: a Cross-Domain Applications of Machine Learning-Based Image Analysis in Mechatronics, Comparative Analysis of Machine Learning Techniques for Image Analysis in Mechatronics, and future research roadmap encompassing large-scale datasets, synthetic data generation.

2. METHOD

2.1. Fundamentals Of Mechatronics And Machine Learning

Mechatronics systems are designed to combine many different fields of engineering to accomplish particular goals, such as improving robotics capabilities, streamlining manufacturing procedures, and developing sophisticated driver-assistance systems. Mechatronics systems primarily include sensors, actuators, controllers, and communication interfaces. Sensors collect and translate electrical signals from environmental data, such as pressure, temperature, and visual input [1]. Actuators are devices, such as motors or hydraulic systems, that move physically in response to commands from the controller to ensure that the system runs smoothly and effectively, controllers process sensor data and use it to generate control signals for the actuators. Communication interfaces provide smooth integration and interoperability by facilitating data transmission between components and external systems [14].

Machine learning is a subdivision of artificial intelligence (AI) that involves the development of models and techniques that allow computers to learn from data and enhance their functionality without the need for explicit programming. This technology is crucial to modern applications including robotics, image processing, and mechatronics because it enables systems to recognize patterns and make data-driven decisions. The key concepts in machine learning are labels, features, and data. Training models are built using data, which can be either unstructured (text or images) or structured (tabular data). Features are the qualities or traits of the data that the model makes use of to classify or forecast [5]. Labels serve as a representation of the desired results or classifications for training models in supervised learning [14].

Machine learning involves various methods such as supervised learning, unsupervised learning, and reinforcement learning. Supervised learning employs neural networks and linear regression among other algorithms, and trains models on labeled data [15]. Unsupervised learning techniques, like clustering, use unlabeled data to identify trends or groupings. In reinforcement learning, an agent can learn through interaction with its surroundings and the imposition of rewards for its actions [7]. Deep learning is a subset of machine learning that is commonly used in image and video analysis. It models relationships within data by using neural networks with multiple layers [16].

2.2 Machine Learning Techniques For Image Analysis In Mechatronics

Supervised learning methods are essential in image analysis as they provide reliable strategies for tasks like object recognition, categorization, and detection. A widely used approach in supervised learning is the Convolutional Neural Network (CNN), which has shown exceptional performance in many image-processing tasks [17]. CNNs are highly effective in analyzing visual data because they possess the capability to autonomously acquire hierarchical features from images. This allows them to accurately recognize intricate patterns and structures [17]. CNNs are frequently used in mechatronics applications, particularly in activities such as defect identification in manufacturing [18]. They are capable of reliably identifying and categorizing defects in products by utilizing labeled training data [19]. Table 1 reveals a progression from conventional machine learning techniques toward deep learning and transformer-based models. While advanced architectures achieve superior feature learning and predictive performance, they demand significantly greater computational resources and data availability. The selection of an appropriate technique depends on the specific application requirements, dataset size, and available computational infrastructure.

The support vector machine (SVM) is a commonly used supervised learning method in mechatronics for image processing. SVMs are very effective algorithms used for classification tasks. They have the ability to accurately divide data into separate classes by utilizing labeled samples [20]. SVMs are commonly used in mechatronics systems for tasks like object recognition and image classification. They are effective in distinguishing between distinct items or categories by analyzing their attributes [21]. SVMs possess benefits such as their ability to handle noise and outliers effectively, rendering them well-suited for a wide range of practical applications across many industries [20].

Furthermore, decision trees are another supervised learning technique applied in image analysis for mechatronics applications. Decision trees are models that use features to divide data into smaller subsets, allowing them to make decisions through a sequence of binary splits [21]. Decision trees are commonly employed in mechatronics systems for tasks like object detection and recognition. They are particularly useful in generating clear and understandable rules for classifying things in images [22]. Decision trees have features such as transparency and ease of interpretation, making them attractive tools in applications where understanding the decision-making process is vital [21].

Unsupervised learning methods are valuable tools in image analysis for mechatronics, offering techniques for tasks such as clustering, dimensionality reduction, and anomaly detection. One commonly used algorithm in

unsupervised learning is k-means clustering, which partitions data into a predetermined number of clusters based on their similarity [23]. In mechatronics applications, k-means clustering is utilized in tasks such as image segmentation, where it can partition an image into distinct regions based on pixel similarity [24]. By grouping similar pixels together, k-means clustering enables the identification of different objects or regions within an image without the need for labeled training data [23].

Another unsupervised learning method widely employed in image analysis for mechatronics is Principal Component Analysis (PCA) [25]. PCA is a method used to reduce the number of dimensions in a dataset. It achieves this by transforming the data from a high-dimensional space to a lower-dimensional space, while still retaining the most significant information. In mechatronics systems, PCA is often used for feature extraction and compression in tasks such as image compression and pattern recognition [26]. By reducing the dimensionality of the data, PCA enables more efficient storage and processing of image data while retaining critical information for analysis and decision-making [25]. Additionally, autoencoders are another unsupervised learning technique utilized in image analysis for mechatronics applications. Autoencoders are neural network architectures designed to learn efficient representations of input data by reconstructing it at the output layer [27]. In mechatronics systems, autoencoders are used for tasks such as image denoising, where they can remove noise and artifacts from images while preserving essential features [28]. By learning a compact representation of the input data, autoencoders enable more robust and efficient image analysis in mechatronics applications [27].

Deep learning models have emerged as effective tools for image processing and analysis in mechatronics, offering exceptional performance in numerous tasks such as object detection, segmentation, and recognition. Convolutional neural networks (CNNs) are the cornerstone of deep learning-based image analysis, notable for their capacity to automatically learn hierarchical features from raw data. In mechatronics applications, CNNs are extensively deployed for tasks like as defect identification in manufacturing, where they can effectively identify and classify faults in products based on visual inspection [29]. By employing deep architectures with numerous layers, CNNs may capture complex patterns and structures inside images, enabling more accurate and robust image analysis in mechatronics systems.

Another class of deep learning models often utilized in image processing for mechatronics is Recurrent Neural Networks (RNNs). RNNs are specialized structures built to handle sequential data by capturing dependencies between input items [30]. In mechatronics applications, RNNs are applied for tasks like as video analysis and gesture recognition, where they can represent temporal correlations within sequences of images or frames [31]. By processing sequences of images over time, RNNs enable mechatronics systems to understand dynamic visual inputs and execute complicated tasks such as action identification and prediction.

Furthermore, generative adversarial networks (GANs) have gained importance in image processing for mechatronics, enabling capabilities for image production, augmentation, and manipulation. GANs consist of two neural networks, a generator and a discriminator, trained adversarial to produce realistic images [32].

Table 1**:** Comparative Analysis of Machine Learning Techniques for Image Analysis in Mechatronics

Technique	**Learning Type**	**Feature Extraction**	**Data Requirement**	**Computational Cost**	**Strengths**	**Limitations**	**Recent Verified References (2022–2026)**
Support Vector Machine (SVM)	Supervised	Manual	Low to Medium	Low	Effective with small datasets, robust in high-dimensional spaces	Performance depends on handcrafted features, limited scalability	[33][34]
k-Nearest Neighbor (k-NN)	Supervised	Manual	Low	Low	Simple implementation, no training phase required	Sensitive to noise and data dimensionality	[35]
Random Forest (RF)	Supervised	Manual	Medium	Medium	Robust against overfitting,	Requires feature engineering, limited deep	[36][37]

Technique	Learning Type	Feature Extraction	Data Requirement	Computational Cost	Strengths	Limitations	Recent Verified References (2022–2026)
					interpretable	feature representation	
Artificial Neural Network (ANN)	Supervised	Semi-Automatic	Medium	Medium	Learns nonlinear relationships effectively	Limited feature extraction capability compared to deep learning	[38][39]
Convolutional Neural Network (CNN)	Supervised	Automatic	High	High	Excellent feature extraction and classification performance	Requires large datasets and computational resources	[11][40]
Transfer Learning	Supervised	Automatic	Medium	Medium	Reduces training time and data requirements	Performance depends on source model suitability	[41][11]
Autoencoder	Unsupervised	Automatic	Medium	Medium	Effective for anomaly detection and dimensionality reduction	Reconstruction quality affects performance	[42]
Generative Adversarial Network (GAN)	Unsupervised	Automatic	High	Very High	Generates synthetic data, addresses data imbalance	Training instability and mode collapse	[43][11]
Vision Transformer (ViT)	Supervised	Automatic	Very High	Very High	Captures long-range dependencies and global features	Requires extensive training data and computational power	[44][45]
Hybrid CNN-ViT Models	Supervised	Automatic	Very High	Very High	Combines local and global feature learning	Increased architectural complexity	[11][44]

3. RESULTS AND DISCUSSION

3.1 Applications of Image Analysis in Mechatronics

Image analysis is used in many different sectors in mechatronics to improve system capabilities and change processes. Table 2 shows the widespread adoption of intelligent image analysis technologies across diverse mechatronic domains. It underscores the transition from traditional machine learning methods toward deep learning and transformer-based frameworks. It also shows research focus on developing robust, explainable, and computationally efficient vision. In manufacturing, image analysis is crucial for quality control and inspection, employing cameras and algorithms to detect faults and assure product consistency [46]. Additionally, in robotics, image analysis helps robots to see and interact with their surroundings, facilitating tasks such as object detection, navigation, and manipulation [47]. In the automobile sector, image analysis plays a vital part in advanced driver-assistance systems (ADAS) and autonomous vehicles, where cameras and sensors detect impediments, recognize traffic signs, and monitor the vehicle's surroundings in real-time [48].

Furthermore, image analysis in mechatronics has applications in healthcare, particularly in medical imaging and robotic surgery. Medical imaging technologies, such as MRI and CT scanners, utilize image analysis to provide detailed images for diagnosis and treatment planning [49]. Robotic surgery systems employ real-time imaging and analysis to guide surgical tools with precision, enabling minimally invasive treatments and

increasing patient outcomes [50]. Additionally, image analysis adds to developments in agriculture, where it is utilized for crop monitoring, disease detection, and yield estimation, helping farmers to optimize resource usage and enhance agricultural productivity [51].

The integration of image analysis in mechatronics systems has also led to innovations in surveillance and security, consumer electronics, aircraft, and marine exploration. Surveillance systems apply image analysis for monitoring and identifying security risks, while consumer electronics incorporate it for features such as facial recognition and gesture control [52]. In aerospace, image analysis aids in navigation, reconnaissance, and target tracking, contributing to the efficiency and safety of operations [53]. Moreover, in marine exploration, image analysis aids mapping of the seafloor, monitoring of marine life, and assessment of marine ecosystems, giving vital insights for research and conservation activities.

Table 2: Cross-Domain Applications of Machine Learning-Based Image Analysis in Mechatronics

Application Domain	Image Analysis Task	ML Techniques	Datasets	Performance Metrics	Major Challenges	References
Smart Manufacturing	Surface defect detection and quality inspection	CNN, ViT, YOLO, Transformer-based detectors	NEU-CLS, DAGM, MVTec AD	Accuracy, F1-score, mAP	Class imbalance, tiny defects, noisy environments	[33][37]
Robotics	Object detection, object recognition, scene understanding	YOLOv8, Faster R-CNN, ViT, SAM	COCO, ImageNet, Open Images	mAP, IoU, Accuracy	Occlusion, dynamic scenes, lighting changes	[54][55]
Autonomous Vehicles	Object detection, lane detection, semantic segmentation	CNN, Transformer, BEV models	ZOD, Argoverse 2, aiMotive Dataset	mAP, IoU, Recall	Adverse weather, real-time processing	[56] [57]
Predictive Maintenance	Fault diagnosis and condition monitoring	CNN, Autoencoder, Transfer Learning	CWRU, Paderborn	Accuracy, F1-score, ROC-AUC	Domain shift, noisy signals	[58][59]
Agricultural Mechatronics	Plant disease detection and crop monitoring	CNN, EfficientNet, ViT	PlantVillage, PlantDoc	Accuracy, Precision, Recall	Background clutter, varying illumination	[60][61]
Healthcare Robotics	Medical image classification and robotic-assisted diagnosis	CNN, U-Net, Vision Transformer	BraTS, ISIC, ChestX-ray datasets	Accuracy, Sensitivity, Dice Score	Data scarcity, explainability	[62]

3.2 Challenges and Suggested Solutions

Table 3 provides a structured synthesis of the main challenges, existing solutions, research gaps, and future directions in machine learning-based image analysis for mechatronic applications. It spans critical issues that repeatedly appear across domains such as smart manufacturing, robotics, autonomous systems, agriculture, and healthcare imaging.

One significant difficulty is the availability and quality of data for training machine learning models. To overcome this, efforts should focus on data augmentation approaches, which entail generating synthetic data to supplement current datasets [63]. Additionally, collaboration between industry and academia can facilitate the collection of labeled datasets through efforts such as open data sharing platforms and joint research projects [64]. Researchers may train more robust machine learning models capable of handling numerous real-world scenarios by providing access to diverse and high-quality datasets.

Another key challenge is the interpretability and explainability of machine learning models, particularly deep learning models. To boost model interpretability, researchers may explore strategies such as attention mechanisms, which highlight significant sections of input data that contribute to the model's predictions [65]. Additionally, adding model explanation tools into existing development frameworks can help users evaluate model decisions and grasp the reasons behind them [66]. Stakeholders may establish trust in their capabilities and ease their adoption in safety-critical applications by increasing the openness and interpretability of machine learning models.

Real-time processing and computing efficiency are critical for implementing image analysis solutions in mechatronic systems, particularly in applications where rapid decision-making is required. One option to overcome this difficulty is through hardware optimization, employing specialized hardware accelerators such as graphics processing units (GPUs) or field-programmable gate arrays (FPGAs) to accelerate computationally intensive tasks. Additionally, researchers might examine strategies such as model compression and quantization, which lower the computational complexity of deep learning models without significantly sacrificing performance [67]. By optimizing algorithms and exploiting hardware acceleration, developers can achieve real-time processing and computational efficiency in mechatronic systems.
Integration and implementation issues in mechatronic systems require assuring seamless integration of image analysis algorithms with existing hardware and software platforms. To solve this, researchers can design standardized interfaces and protocols for communication between different components of mechatronic systems [68]. Additionally, adopting modular and scalable designs can ease the integration of image analysis technologies into varied mechatronic systems [68]. By overcoming integration difficulties, developers can streamline the implementation of image analysis technologies in real-world applications, enabling their wider adoption

Table 3: Current Challenges, Research Gaps, and Future Research Directions in Machine Learning-Based Image Analysis for Mechatronics

Challenge Area	Impact on Mechatronic Image Analysis Systems	Current Approaches	Research Gap	Future Research Direction	References
Data scarcity and limited labeled datasets	Reduces generalization performance and increases overfitting in industrial and medical vision systems	Data augmentation, transfer learning, synthetic data generation	Lack of large-scale domain-specific annotated datasets for mechatronics applications	Self-supervised learning and foundation models for label-efficient learning	[33][11]
High computational cost of deep models	Limits deployment on embedded systems and edge devices used in robotics and manufacturing	Model compression, pruning, lightweight CNNs	Inefficient balance between accuracy and computational efficiency	Edge AI, TinyML, and efficient transformer architectures	[37][69]
Lack of explainability in AI decisions	Reduces trust and adoption in safety-critical systems such as healthcare robotics and autonomous driving	Post-hoc XAI methods (Grad-CAM, SHAP)	Limited integration of explainability into model design	Explainable-by-design AI frameworks for mechatronic systems	[70][71]
Domain shift and poor generalization	Performance drops when models are deployed in new environments (factory, road, field, hospital)	Transfer learning, domain adaptation	Lack of robust cross-domain adaptation methods for real-world mechatronics	Domain generalization and continual learning approaches	[11][69]
Real-time processing constraints	Delays in decision-making affect robotics control, autonomous navigation, and industrial inspection	Lightweight CNNs, edge inference optimization	Trade-off between speed and accuracy remains unresolved	Real-time vision transformers and hardware-aware neural design	[33][69]
Limited multimodal integration	Loss of contextual information in complex environments (vision + sensor fusion)	Early fusion and late fusion techniques	Lack of unified frameworks for multimodal learning in mechatronics	Vision-language-action models and multimodal transformers	[72]
Safety and reliability in autonomous systems	Risk of failure in robotics, autonomous vehicles, and medical systems	Rule-based safety layers and validation testing	Lack of certified AI systems for safety-critical deployment	Trustworthy AI and formal verification for deep learning systems	[73]

.3.3 Future Trends and Directions

Current research is increasingly focused on developing more scalable, efficient, and intelligent frameworks. The future of machine learning-based image analysis is therefore expected to be driven by advances in data-centric methodologies, foundation models, multimodal intelligence, edge computing, and trustworthy Artificial Intelligence. The future research roadmap for intelligent mechatronic systems is shown in Figure 2. As image analysis systems become more sophisticated, the emphasis is g shifting from model-centric optimization toward data-centric artificial intelligence. Future research is expected to focus on the development of large-scale domain-specific repositories that support reproducible experimentation and fair benchmarking. Open-source initiatives will play a critical role in accelerating innovation by enabling broader access to representative datasets. The adoption of comprehensive benchmark datasets will also improve model generalization across different operating environments.

The growing demand for large training datasets has stimulated considerable interest in synthetic data generation techniques. Synthetic data has emerged as a promising alternative for addressing data scarcity while supporting the development of more robust machine learning models. Recent advances in generative artificial intelligence, particularly Generative Adversarial Networks (GANs) and diffusion models, have demonstrated remarkable capabilities in producing realistic images that closely resemble real-world data. Digital twin technology further extends these capabilities by creating virtual representations of physical systems. The combination of virtual environments and advanced generative models offers a scalable pathway toward more reliable, cost-effective, and adaptable image analysis solution. Self-supervised learning has emerged as a promising solution to this challenge by enabling models to learn meaningful representations directly from unlabeled data [74]. By leveraging unlabeled data, self-supervised frameworks can improve learning efficiency while reducing dependence on costly labeling processes.

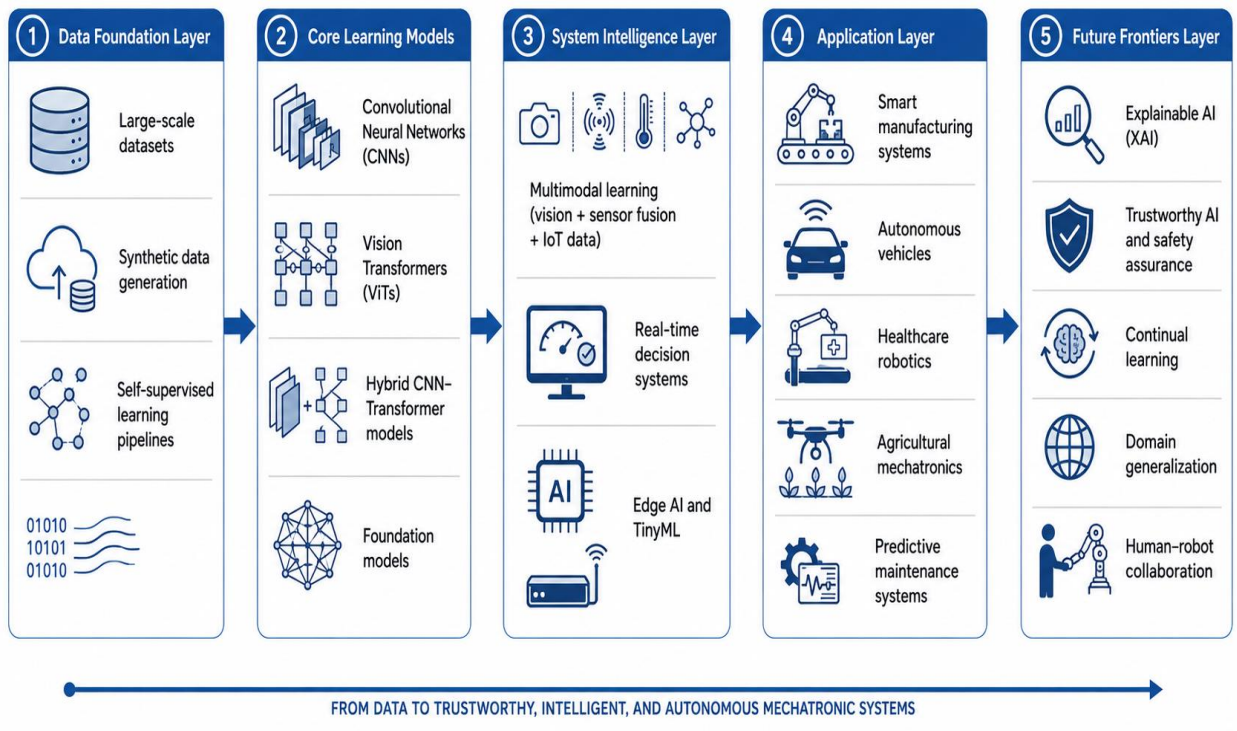


Figure 2: Future research roadmap for intelligent mechatronic systems,

4. CONCLUSION

Machine learning-based image analysis has become a central technology in modern mechatronic systems. As mechatronic applications continue to evolve toward greater levels of autonomy and environmental awareness, the ability to extract meaningful information from visual data has become increasingly important across a broad range of industrial and societal domains. The convergence of machine learning, computer vision, sensing technologies, and mechatronic design has created new opportunities for developing systems This review demonstrate that machine learning-driven image analysis is now widely adopted across diverse mechatronic applications. Although the operational objectives vary among these domains, a common trend is the growing reliance on intelligent vision systems to improve efficiency and adaptability.

A clear technological evolution can be observed in the methods used for image analysis. Traditional machine learning approaches established important foundations for automated visual interpretation and remain valuable in applications However, the growing complexity of visual data has accelerated the adoption of deep

learning architectures capable of automatically learning hierarchical feature representations. Convolutional Neural Networks have become the dominant framework for many image analysis tasks because of their strong feature extraction capabilities and classification performance. More recently, transfer learning techniques have reduced the dependence on large labeled datasets, while Generative Adversarial Networks have expanded opportunities for data augmentation and synthetic data generation. The emergence of Vision Transformers and hybrid CNN-Transformer architectures represents a significant step forward, combining local feature extraction with global contextual understanding to address increasingly challenging image analysis problems.

Despite these advances, several obstacles continue to limit the widespread deployment of intelligent image analysis systems in real-world mechatronic environments. Emerging research directions offer promising pathways for addressing many of these limitations. Foundation models and self-supervised learning frameworks have the potential to reduce dependence on extensive labeled datasets while improving model generalization across different operational environments. Synthetic data generation through advanced generative models and digital twin technologies may help overcome data scarcity and support more robust training procedures. At the same time, multimodal learning frameworks capable of integrating visual, sensory, and contextual information can provide richer environmental understanding for intelligent mechatronic systems. Advances in edge artificial intelligence are expected to facilitate efficient deployment in resource-constrained environments, while the development of trustworthy and explainable AI frameworks will be critical for ensuring transparency, safety, and user confidence.

Machine learning-based image analysis has progressed from a supporting computational tool to a fundamental component of intelligent mechatronic systems. Its continued development will play a decisive role in shaping the next generation of autonomous, adaptive, and interconnected technologies. By addressing current limitations and embracing emerging paradigms such as foundation models, multimodal intelligence, edge AI, and trustworthy artificial intelligence, the field is well positioned to deliver more capable and resilient mechatronic systems.

REFERENCES

[1] Archana, R., and Jeevaraj, P. S. E. (2024), "Deep learning models for digital image processing: a review," Artif. Intell. Rev., vol. 57, no. 1, p. 11, Jan., doi: 10.1007/s10462-023-10631-z.

[2] Iqbal, S., Qureshi, A. N., Li, J., and Mahmood, T. (2023), "On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks," Arch. Comput. Methods Eng., vol. 30, no. 5, pp. 3173–3233, Jun., doi: 10.1007/s11831-023-09899-9.

[3] Guerrero-Osuna, H. A., Nava-Pintor, J. A., Olvera-Olvera, C. A., Ibarra-Pérez, T., Carrasco-Navarro, R., and Luque-Vega, L. F. (2023), "Educational Mechatronics Training System Based on Computer Vision for Mobile Robots," Sustainability, vol. 15, no. 2, p. 1386, Jan., doi: 10.3390/su15021386.

[4] Yadav, N., and Schledjewski, R. (2023), "Review of in-process defect monitoring for automated tape laying," Compos. Part A Appl. Sci. Manuf., vol. 173, p. 107654, Oct., doi: 10.1016/j.compositesa.2023.107654.

[5] Galić, I., Habijan, M., Leventić, H., and Romić, K. (2023), "Machine Learning Empowering Personalized Medicine: A Comprehensive Review of Medical Image Analysis Methods," Electronics, vol. 12, no. 21, p. 4411, Oct., doi: 10.3390/electronics12214411.

[6] Salawu, G., and Glen, B. (2026), "Integrating Artificial Intelligence into Mechatronics: A Comprehensive Study of Its Influence on System Performance, Autonomy, and Manufacturing Efficiency," Technologies, vol. 14, no. 3, p. 143, Feb., doi: 10.3390/technologies14030143.

[7] Adebisi, O., Afolayan, A., Ayoade, I., Adejumobi, P., and Adejumobi, I. (2024), "Integration of Deep Learning Techniques in Mechatronic Devices and Systems: Advancement, Challenges, and Opportunities," in 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), IEEE, Apr., pp. 1–6. doi: 10.1109/SEB4SDG60871.2024.10630414.

[8] Jhaveri, R. H., Revathi, A., Ramana, K., Raut, R., and Dhanaraj, R. K. (2022), "A Review on Machine Learning Strategies for Real-World Engineering Applications," Mob. Inf. Syst., vol. 2022, pp. 1–26, Aug., doi: 10.1155/2022/1833507.

[9] Alsubaei, F. S., Almazroi, A. A., and Ayub, N. (2024), "Enhancing Phishing Detection: A Novel Hybrid Deep Learning Framework for Cybercrime Forensics," IEEE Access, vol. 12, pp. 8373–8389, doi: 10.1109/ACCESS.2024.3351946.

[10] Sistaninejhad, B., Rasi, H., and Nayeri, P. (2023), "A Review Paper about Deep Learning for Medical Image Analysis," Comput. Math. Methods Med., vol. 2023, no. 1, Jan., doi: 10.1155/2023/7091301.

[11] Ma, Y., Yin, J., Huang, F., and Li, Q. (2024), "Surface defect inspection of industrial products with object detection deep networks: a systematic review," Artif. Intell. Rev., vol. 57, no. 12, p. 333, Oct., doi: 10.1007/s10462-024-10956-3.

[12] Talaei Khoei, T., Ould Slimane, H., and Kaabouch, N. (2023), "Deep learning: systematic review, models, challenges, and research directions," Neural Comput. Appl., vol. 35, no. 31, pp. 23103–23124, Nov., doi: 10.1007/s00521-023-08957-4.

[13] Münke, F. R., Schützke, J., Berens, F., and Reischl, M. (2024), "A review of adaptable conventional image processing pipelines and deep learning on limited datasets," Mach. Vis. Appl., vol. 35, no. 2, p. 25, Mar., doi: 10.1007/s00138-023-01501-3.

[14] Kittisares, S., Hirota, Y., Nabae, H., and Suzumori, K. (2023), "Open-looped pressure and volume control of multi-degree-of-freedom hydraulic systems," Mechatronics, vol. 95, p. 103068, Nov., doi: 10.1016/j.mechatronics.2023.103068.

[15] Adebisi, O. A., Ojo, J. A., Adegbola, O. A., Oseni, O. F., and Oyeniran, O. A. (2023), "Development of Computer Aided Diagnosis System for Detection of COVID-19 Using Transfer Learning," pp. 213–230. doi: 10.1007/978-3-031-28631-5_11.

[16] Valente, J., António, J., Mora, C., and Jardim, S. (2023), "Developments in Image Processing Using Deep Learning and Reinforcement Learning," J. Imaging, vol. 9, no. 10, p. 207, Sep., doi: 10.3390/jimaging9100207.

[17] Adebisi, O. A., Busari, O., Oyewola, Y., and Adeaga, I. (2020), "Automatic classification of lung nodules on computed tomography images using a pre-trained convolutional neural network," Int. J. Eng. Sci. Invent., vol. 9, no. 1, pp. 63–66.

[18] Kollu, G. S., and Kittur, J. (2025), "Integrating Deep Learning Into Mechanical Engineering: A Systematic Review of Applications and Educational Implications," Comput. Appl. Eng. Educ., vol. 33, no. 3, May, doi: 10.1002/cae.70048.

[19] Adebisi, O. A., Ajagbe, S. A., Ojo, J. A., and Oladipupo, M. A. (2022), "Computer Techniques for Medical Image Classification: A Review," in Intelligent Healthcare, Singapore: Springer Nature Singapore, pp. 19–36. doi: 10.1007/978-981-16-8150-9_2.

[20] Cortes, C., and Vapnik, V. (1995), "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Sep., doi: 10.1007/BF00994018.

[21] Shah Hosseini, R., Entezari, I., Homayouni, S., Motagh, M., and Mansouri, B. (2011), "Classification of polarimetric SAR images using Support Vector Machines," Can. J. Remote Sens., vol. 37, no. 2, pp. 220–233, Apr., doi: 10.5589/m11-029.

[22] Quinlan, J. R. (1986), "Induction of decision trees," Mach. Learn., vol. 1, no. 1, pp. 81–106, Mar., doi: 10.1007/BF00116251.

[23] Hartigan, J. A., and Wong, M. A. (1979), "Algorithm AS 136: A K-Means Clustering Algorithm," Appl. Stat., vol. 28, no. 1, p. 100, doi: 10.2307/2346830.

[24] Shi, Jianbo, and Malik, J. (2000), "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, doi: 10.1109/34.868688.

[25] Jolliffe, I. T., and Cadima, J. (2016), "Principal component analysis: a review and recent developments," Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., vol. 374, no. 2065, p. 20150202, Apr., doi: 10.1098/rsta.2015.0202.

[26] Duong, T. D. X., and Duong, V. N. (2008), "Non-negative Sparse Principal Component Analysis for Multidimensional Constrained Optimization," pp. 103–114. doi: 10.1007/978-3-540-89197-0_13.

[27] Hinton, G. E., and Salakhutdinov, R. R. (2006), "Reducing the Dimensionality of Data with Neural Networks," Science (80-.)., vol. 313, no. 5786, pp. 504–507, Jul., doi: 10.1126/science.1127647.

[28] Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008), "Extracting and composing robust features with denoising autoencoders," in Proceedings of the 25th international conference on Machine learning - ICML '08, New York, New York, USA: ACM Press, pp. 1096–1103. doi: 10.1145/1390156.1390294.

[29] Ramteke, S. M., Chelladurai, H., and Amarnath, M. (2022), "Diagnosis and Classification of Diesel Engine Components Faults Using Time–Frequency and Machine Learning Approach," J. Vib. Eng. Technol., vol. 10, no. 1, pp. 175–192, Jan., doi: 10.1007/s42417-021-00370-2.

[30] Graves, A. (2014), "Generating Sequences With Recurrent Neural Networks," Jun., [Online]. Available: http://arxiv.org/abs/1308.0850

[31] Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T. (2015). Long-term recurrent convolutional networks for visual recognition and description. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2625-2634).

[32] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012), "Improving neural networks by preventing co-adaptation of feature detectors," Jul., [Online]. Available: http://arxiv.org/abs/1207.0580

[33] Liu, Y., Zhang, C., and Dong, X. (2023), "A survey of real-time surface defect inspection methods based on deep learning," Artif. Intell. Rev., vol. 56, no. 10, pp. 12131–12170, Oct., doi: 10.1007/s10462-023-10475-7.

[34] Majeed, H. L., Hassen, O. A., Farhan, D. A., Gromov, Y. Y., Sheoran, K., and Dhand, G. (2025), "Computer Vision of Smile Detection Based on Machine and Deep Learning Approach," J. Cybersecurity Inf. Manag., vol. 16, no. 1, pp. 208–230, doi: 10.54216/JCIM.160115.

[35] Cunningham, P., and Delany, S. J. (2022), "k-Nearest Neighbour Classifiers - A Tutorial," ACM Comput. Surv., vol. 54, no. 6, pp. 1–25, Jul., doi: 10.1145/3459665.

[36] Wang, P. (2024), "Prediction of the Groundwater Levels Based on Random Forest Regression Algorithm," in 2024 IEEE 4th International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), IEEE, Dec., pp. 214–217. doi: 10.1109/ICIBA62489.2024.10869277.

[37] Tercan, H., and Meisen, T. (2022), "Machine learning and deep learning based predictive quality in manufacturing: a systematic review," J. Intell. Manuf., vol. 33, no. 7, pp. 1879–1905, Oct., doi: 10.1007/s10845-022-01963-8.

[38] Aldrini, J., Chihi, I., and Sidhom, L. (2024), "Fault diagnosis and self-healing for smart manufacturing: a review," J. Intell. Manuf., vol. 35, no. 6, pp. 2441–2473, Aug., doi: 10.1007/s10845-023-02165-6.

[39] Chauhan, B. V. S., Vedrtnam, A., Wyche, K. P., and Verma, S. (2025), "Review on Artificial Intelligence in the Environmental Monitoring," in Prospects of Artificial Intelligence in the Environment, Singapore: Springer Nature Singapore, pp. 29–60. doi: 10.1007/978-981-96-6863-2_2.

[40] Cumbajin, E., Rodrigues, N., Costa, P., Miragaia, R., Frazão, L., Costa, N., ... & Pereira, A. (2023). A systematic review on deep learning with CNNs applied to surface defect detection. Journal of Imaging, 9(10), 193.

[41] Bareja, R., Ismail, M., Martin, D., Nayate, A., Yadav, I., Labbad, M., ... & Tiwari, P. (2024). nnU-Net–based segmentation of tumor subcompartments in pediatric medulloblastoma using multiparametric MRI: a multi-institutional study. *Radiology: Artificial Intelligence*, *6*(5), e230115.

[42] Gautam, P. (2026), "AI-Driven Data Center Infrastructure, Mechanical Piping Systems, and Reliability Engineering: Energy Optimization, Fault Prediction," Int. J. Emerg. Res. Eng. Technol., vol. 7, pp. 135–145, doi: 10.63282/3050-922X.IJERET-V7I1P119.

[43] Ju, Z., Chen, Y., Qiang, Y., Chen, X., Ju, C., and Yang, J. (2024), "A systematic review of data augmentation methods for intelligent fault diagnosis of rotating machinery under limited data conditions," Meas. Sci. Technol., vol. 35, no. 12, p. 122004, Dec., doi: 10.1088/1361-6501/ad7a97.

[44] Kim, J. W., Khan, A. U., and Banerjee, I. (2025), "Systematic Review of Hybrid Vision Transformer Architectures for Radiological Image Analysis," J. Imaging Informatics Med., vol. 38, no. 5, pp. 3248–3262, Jan., doi: 10.1007/s10278-024-01322-4.

[45] Alammar, Z., Alzubaidi, L., Zhang, J., Li, Y., Zhao, Z., and Gu, Y. (2026), "Improving robustness in X-ray image classification through attention mechanisms in convolutional neural networks," PeerJ Comput. Sci., vol. 12, p. e3447, Jan., doi: 10.7717/peerj-cs.3447.

[46] Mohammed, M. Q., Kwek, L. C., Chua, S. C., Al-Dhaqm, A., Nahavandi, S., Eisa, T. A. E., ... & Alandoli, E. A. (2022). Review of learning-based robotic manipulation in cluttered environments. Sensors, 22(20), 7938.

[47] Pak, M., and Kim, S. (2017), "A review of deep learning in image recognition," in 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT), IEEE, Aug., pp. 1–3. doi: 10.1109/CAIPT.2017.8320684.

[48] Hassan, M., Kabir, M. E., Islam, M. K., Alam, E., Rambe, A. H., Jusoh, M., & Sameer, M. (2025). Mapping the machine learning landscape in autonomous vehicles: a scientometric review of research trends, applications, challenges, and future directions. IEEE Access.

[49] Kumar, R., Anil, M., Panda, S., and Raj, A. (2023), "Medical imaging: Challenges and future directions in AI-Based systems," p. 020147. doi: 10.1063/5.0154355.

[50] Kossowsky, H., and Nisky, I. (2022), "Predicting the Timing of Camera Movements From the Kinematics of Instruments in Robotic-Assisted Surgery Using Artificial Neural Networks," IEEE Trans. Med. Robot. Bionics, vol. 4, no. 2, pp. 391–402, May, doi: 10.1109/TMRB.2022.3156635.

[51] Bargoti, S., and Underwood, J. P. (2017), "Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards," J. F. Robot., vol. 34, no. 6, pp. 1039–1060, Sep., doi: 10.1002/rob.21699.

[52] Raghu, M., and Schmidt, E. (2020), "A Survey of Deep Learning for Scientific Discovery," pp. 1–48, [Online]. Available: http://arxiv.org/abs/2003.11755

[53] Emimi, M., Khaleel, M., and Alkrash, A. (2023), "The Current Opportunities and Challenges in Drone Technology," Int. J. Electr. Eng. Sustain., vol. 1, no. 3, pp. 74–89, Jul., doi: 10.65998/ijees.v1i3.47.

[54] Hussain, A., Farwa, U. E., Ali, S., and Kim, H.-C. (2026), "The Rise of Foundation Models: Opportunities, Technology, Applications, Challenges, Recent Trends, and Future Directions," Appl. Syst. Innov., vol. 9, no. 2, p. 35, Jan., doi: 10.3390/asi9020035.

[55] Wu, X., Cheng, M., Ren, X., Hu, Q., Chen, J., Huang, Y., ... & Le Traon, Y. (2026). Foundation models for autonomous driving systems: An initial roadmap. ACM Transactions on Software Engineering and Methodology.

[56] Alibeigi, M., Ljungbergh, W., Tonderski, A., Hess, G., Lilja, A., Lindström, C., ... & Petersson, C. (2023, October). Zenseact open dataset: A large-scale and diverse multimodal dataset for autonomous driving. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 20121-20131). IEEE.

[57] Wilson, B., Qi, W., Agarwal, T., Lambert, J., Singh, J., Khandelwal, S., ... & Hays, J. (2023). Argoverse 2: Next generation datasets for self-driving perception and forecasting. arXiv preprint arXiv:2301.00493.

[58] Nguyen, V.-T., and Diep, Q. B. (2025), "Vibration-based gearbox fault diagnosis using a multi-scale convolutional neural network with depth-wise feature concatenation," PLoS One, vol. 20, no. 7, p. e0324905, Jul., doi: 10.1371/journal.pone.0324905.

[59] Srinivaas, A., Sakthivel, N. R., and Nair, B. B. (2025), "Machine Learning Approaches for Fault Detection in Internal Combustion Engines: A Review and Experimental Investigation," Informatics, vol. 12, no. 1, p. 25, Feb., doi: 10.3390/informatics12010025.

[60] Wang, D., Cao, W., Zhang, F., Li, Z., Xu, S., and Wu, X. (2022), "A Review of Deep Learning in Multiscale Agricultural Sensing," Remote Sens., vol. 14, no. 3, p. 559, Jan., doi: 10.3390/rs14030559.

[61] Cao, Z., Sun, S., and Bao, X. (2025), "A Review of Computer Vision and Deep Learning Applications in Crop Growth Management," Appl. Sci., vol. 15, no. 15, p. 8438, Jul., doi: 10.3390/app15158438.

[62] Beyaz, S., Özgözen, A. L., Turgut, N., and Ölke, H. C. (2025), "Artificial intelligence and robotic surgery in clinical medicine: progress, challenges, and future directions," Futur. Sci. OA, vol. 11, no. 1, Dec., doi: 10.1080/20565623.2025.2540742.

[63] Wu, L., Qiu, S., and Chen, Z. (2024), "Real-time semantic segmentation network based on parallel atrous convolution for short-term dense concatenate and attention feature fusion," J. Real-Time Image Process., vol. 21, no. 3, p. 74, Jun., doi: 10.1007/s11554-024-01453-5.

[64] Lee, M., Choi, M., Yang, T., Kim, J., Kim, J., Kwon, O., & Cho, N. (2024). A study on the advancement of intelligent military drones: Focusing on reconnaissance operations. *IEEE Access*, *12*, 55964-55975.

[65] Simonyan, K., Vedaldi, A., and Zisserman, A. (2014), "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps," Apr., [Online]. Available: http://arxiv.org/abs/1312.6034

[66] Ribeiro, M. T., Singh, S., and Guestrin, C. (2016), "'Why Should I Trust You?,'" in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA: ACM, Aug., pp. 1135–1144. doi: 10.1145/2939672.2939778.

[67] Han, S., Mao, H., and Dally, W. J. (2016), "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," Feb., doi: 1510.00149.

[68] Cheah, C. C., Bock, T., Cao, J., Linner, T., Liu, Y. H., Yamashita, A., & Yang, L. (2021). Guest editorial introduction to the focused section on mechatronics and automation for constructions. *IEEE/ASME Transactions on Mechatronics*, *26*(6), 2819-2825.

[69] Zhang, Y., and Chen, F. (2024), "A Fault Diagnosis and Predictive Maintenance Algorithm for Mechanical Systems Based on Deep Learning," in 2024 International Conference on Electrical Drives, Power Electronics & Engineering (EDPEE), IEEE, Feb., pp. 760–765. doi: 10.1109/EDPEE61724.2024.00146.

[70] Kong, X., Xing, Y., Tsourdos, A., Wang, Z., Guo, W., Perrusquia, A., & Wikander, A. (2024). Explainable interface for human-autonomy teaming: A survey. arXiv preprint arXiv:2405.02583.

[71] Islam, M. R., Ahmed, M. U., Barua, S., and Begum, S. (2022), "A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks," Appl. Sci., vol. 12, no. 3, p. 1353, Jan., doi: 10.3390/app12031353.

[72] Sapkota, R., Cao, Y., Roumeliotis, K. I., and Karkee, M. (2026), "Vision-Language-Action (VLA) Models: Concepts, Progress, Applications and Challenges," Jan., doi: 2505.04769.

[73] Perez-Cerrolaza, J., Abella, J., Borg, M., Donzella, C., Cerquides, J., Cazorla, F. J., ... & Flores, J. L. (2024). Artificial intelligence for safety-critical systems in industrial and transportation domains: A survey. ACM Computing Surveys, 56(7), 1-40.

[74] Eldele, E., Ragab, M., Chen, Z., Wu, M., Kwoh, C. K., Li, X., & Guan, C. (2023). Self-supervised contrastive representation learning for semi-supervised time-series classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(12), 15604-15618.