International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 505-516
https://doi.org/10.70917/ijcisim-2026-2347
Article

AI-Based Vision System for an Uncertain Dynamic Environment

Santosh Kumar Sahoo[1], Pravin R. Kshirsagar[2], Shrikant V. Sonekar[3], Subodh Panda[4]

[1] School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India.
Email: Santosh.kr.sahoo@gmail.com
[2] Department of Electronics and Telecommunication Engineering, Dean (R&D), J. D. College of Engineering & Management, Nagpur, Maharashtra, India.
Email: pravinrk88@yahoo.com
[3] Principal, J. D. College of Engineering & Management, Nagpur, Maharashtra, India.
Email: principal@jdcoem.ac.in
[4] Department of Electronics and Communication Engineering (ECE), Pragati Engineering College, Surampalem, Andhra Pradesh, India.
Email: subodh.panda@gmail.com

Abstract: –Machine vision systems are increasingly expected to operate in practical environments where lighting conditions, object movement, and surrounding scenes may change continuously. Such variations often influence recognition reliability and create additional difficulties in applications including industrial automation, surveillance, and autonomous navigation. The study examines an AI-supported machine vision framework designed for operation under dynamic situations. Adaptive preprocessing, deep learning–based recognition, and environmental compensation mechanisms were incorporated within the proposed approach. Feature extraction relied on convolutional neural networks, while recurrent learning models assisted in tracking moving objects across image sequences. Reinforcement learning was also used to adjust system parameters according to observed scene behavior. Evaluation under changing lighting situations and dynamic operating scenarios showed recognition accuracy above 92%, whereas conventional methods produced noticeably lower performance under similar settings. Observations from the experiments indicate that adaptive AI mechanisms can contribute to improved reliability of machine vision systems operating in real-world environments.

Keywords: – Artificial Intelligence, Dynamic Environment, Computer Vision, Deep Learning

1. INTRODUCTION

Machine vision has become an important part of applications such as manufacturing, robotics, surveillance, and autonomous systems, where visual information supports decision-making and automated operation. Considerable progress has been made over the years; however, differences between performance observed under laboratory conditions and behavior during practical deployment continue to attract attention [1]. Controlled environments generally provide stable lighting, cleaner backgrounds, and predictable object appearance. Under such conditions, even relatively simple vision techniques may achieve satisfactory performance.

Conditions encountered during practical operation are often less predictable. Industrial facilities and outdoor environments frequently experience variations that influence image quality and system behavior. In manufacturing environments, lighting may change because of shifts in natural illumination and indoor sources during different periods of operation. Outdoor surveillance systems face additional difficulties arising from fog, rain, dust, shadows, and varying illumination levels [2]. These factors can influence the reliability of visual recognition and tracking processes.

Many conventional machine vision systems rely on calibration procedures and settings tailored to specific environments. Lighting arrangements are usually selected to reduce shadow effects and reflections, while camera placement and background settings are adjusted to improve scene visibility [3]. Such approaches can provide stable operation under predefined situations; however, performance may become less reliable when surrounding conditions differ from assumptions made during system setup [4]. The impact of these limitations extends beyond isolated recognition errors. System deployment may require substantial setup effort together with specialized technical expertise, while maintenance demands can increase as operating situations evolve over time [5].

Such practical difficulties have limited broader deployment of machine vision systems across several real-world applications [6].

Research Gaps

Recent progress in artificial intelligence and computer vision has improved machine vision performance considerably; however, practical deployment under dynamic situations continues to present challenges [21],[22]. A large portion of existing deep learning studies relies on benchmark datasets collected under relatively controlled settings. Such datasets are useful for evaluation, but they may not fully capture the variations encountered during real operation. Changes in illumination, object movement, and external disturbances can influence system behavior in ways that are difficult to reproduce using static datasets alone. In addition, comparatively fewer investigations have examined system behavior across a wide range of operating situations or studied robustness under continuously changing scenarios [23],[24].

Practical deployment introduces additional concerns at the system level. Real-world applications generally require coordination among preprocessing, object detection, tracking, and decision-making stages, whereas many earlier studies place greater emphasis on individual components rather than complete system interaction. Deep learning models may also become less reliable when operating situations differ substantially from those represented during training [25]. Computational requirements create further difficulty, particularly for applications involving rapid response or deployment on resource-limited hardware platforms. These observations indicate the need for integrated approaches capable of adapting to uncertain operating situations.

Proposed Approach and Conceptual Framework

Machine vision applications operating in practical environments often experience changes that influence image quality and recognition behavior over time. Approaches based on fixed configurations and predefined assumptions may become less reliable when operating situations vary significantly. To address such difficulties, the Adaptive Multi-Pathway AI Vision Framework (AMPA-VF) was developed by combining multiple learning and processing mechanisms within a unified architecture.

Different processing pathways were incorporated to address specific environmental influences, including illumination variation, motion-related effects, and atmospheric disturbances. Rather than relying on a single processing route, multiple pathways function simultaneously and contribute according to observed scene characteristics.

System adaptation is supported through a reinforcement learning–based mechanism that adjusts preprocessing settings, detection thresholds, and tracking parameters during operation. Parameter changes rely on system feedback so that performance can be maintained without repeated manual adjustment when operating situations change.

Temporal behavior was also considered through recurrent learning models combined with predictive filtering techniques. Information obtained across image sequences helped preserve continuity during tracking and reduced false detections when scene behavior changed over time. Operational flow within the framework begins with adaptive preprocessing followed by feature extraction, multi-scale object detection, temporal tracking, and parameter adjustment.

Integrating these stages allows feature learning, adaptation, and temporal reasoning to function together under uncertain situations.

The framework integrates preprocessing, object detection, tracking, and adaptive control mechanisms within a unified structure intended for dynamic environments. Additional compensation methods were included to reduce the influence of lighting variation, motion blur, and atmospheric effects. Reinforcement learning supported parameter tuning, while temporal modeling contributed to improved tracking continuity. Experimental observations showed stronger robustness when compared with conventional approaches and standard deep learning methods.

Possible applications include intelligent manufacturing, surveillance systems, autonomous navigation, and agricultural automation. Adaptive learning capability may also support future machine vision systems that can respond to long-term changes with less dependence on manually engineered settings.

1.1 OBJECTIVES

Improving machine vision performance under uncertain and continuously changing situations remains the primary focus of this study. Attention is given to several objectives associated with reliable operation under practical conditions:

- To develop an AI-supported machine vision framework capable of maintaining stable operation and improved recognition accuracy in situations where conventional systems often experience performance limitations.
- Investigation of adaptive preprocessing methods intended to respond automatically to variations such as illumination changes, motion blur, and atmospheric disturbances without repeated manual intervention.
- The study also examines deep learning models suitable for uncertain operating situations while maintaining real-time processing capability for practical use.
- Temporal information obtained from consecutive image frames is considered for improving tracking continuity and reducing false detections during dynamic scene analysis.
- Assessment of system behavior under different real-world situations together with comparison against conventional machine vision approaches forms another objective of the investigation.

2. RESEARCH METHODOLOGY

A design science research approach together with experimental evaluation was used to study machine vision behavior under uncertain operating situations. Besides system development, attention was also directed toward understanding how adaptive AI mechanisms influence robustness when scene characteristics and operating behavior change over time. Performance comparisons involved conventional machine vision techniques, standard deep learning methods, and the proposed adaptive framework under multiple operating scenarios. Development progressed through several refinement stages. Early implementations included common computer vision modules such as image acquisition, preprocessing, feature extraction, object detection, and tracking functions. Observations during experimentation revealed limitations under changing situations, leading to gradual integration of adaptive AI mechanisms at different processing stages [26]. The final architecture adopted a multi-pathway arrangement in which several processing streams operated simultaneously to address different environmental influences.

Different pathways addressed specific difficulties encountered during operation. Illumination-related variation was managed through a compensation mechanism intended to reduce the influence of changing lighting behavior. Motion disturbances were handled through temporal filtering and prediction methods designed to support tracking continuity under dynamic scenes. An attention-guided component was also incorporated so that computational effort could focus on regions more likely to contain relevant information, reducing the effect of scene clutter.

2.1 Data Collection and Dataset Creation

Assessment of the proposed system required data representing situations commonly encountered during practical operation. Dataset preparation therefore included multiple scenarios with different scene characteristics and operating situations.

Manufacturing environments: Data acquisition was carried out in industrial settings involving object detection and defect inspection activities. Illumination varied because of daylight entering through windows, overhead fluorescent lighting, and localized task-based light sources. Collection activities also covered different operating periods and weather-related changes that influenced available lighting.

Outdoor surveillance: Video sequences obtained from surveillance settings included pedestrian and vehicle detection under situations such as rainfall, fog, overcast skies, and clear weather. Image data covered daytime operation together with low-light and nighttime situations.

Dynamic scene navigation: Additional samples were generated using mobile platforms operating within cluttered surroundings containing moving obstacles, changing backgrounds, and transitions between indoor and outdoor spaces.

Ground-truth annotations were prepared through manual labeling procedures that included object location, category information, and tracking identifiers across image sequences. More than 45,000 annotated images together with temporal sequences were included for tracking evaluation. Different neural network architectures were considered for object detection and recognition tasks. YOLO-based models were selected because of their real-time processing capability, while Faster R-CNN approaches were included for comparison with computationally intensive alternatives. ResNet-50 and MobileNet architectures supported feature extraction so that trade-offs between computational efficiency and recognition performance could be examined [27].

Additional augmentation procedures were incorporated during training to improve robustness under varying situations. Brightness and contrast changes together with simulated weather effects, motion blur, and partial object occlusion exposed the learning models to a broader range of scene variation. Handling previously unseen situations was supported through a meta-learning strategy. Instead of depending entirely on fixed training behavior, internal model parameters were adjusted using recent observations obtained during operation. This approach supported faster adaptation when practical situations differed from those represented during training and allowed continuous adjustment during inference.

3. SYSTEM ARCHITECTURE AND KEY COMPONENTS

3.1 Adaptive Preprocessing Pipeline

Figure 1 presents the Adaptive Preprocessing Pipeline incorporated within the proposed framework. Image quality during practical deployment may vary because surrounding situations do not always remain stable. Fixed preprocessing approaches generally apply operations such as contrast enhancement and noise reduction uniformly across images. While these techniques may perform adequately under controlled settings, their effectiveness can decrease when scene variations become more pronounced. To address this issue, adaptive mechanisms were introduced so that preprocessing behavior could respond according to observed image characteristics.

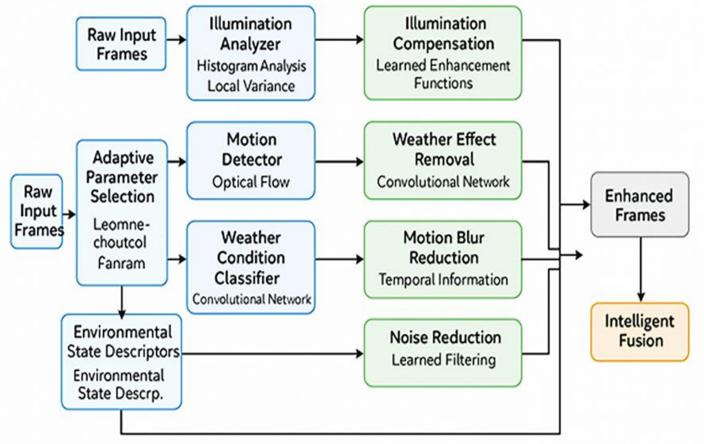

Figure 1. Adaptive Preprocessing Pipeline Architecture

An illumination compensation module examines overall image statistics together with local intensity variation to determine suitable enhancement requirements. Instead of relying only on conventional techniques such as histogram equalization, learned transformation functions are used to convert images obtained under different lighting situations into more consistent representations. Such processing helps improve handling of uneven illumination, shadow effects, and reflection-related disturbances [28].

Visual quality may also be influenced by atmospheric effects commonly encountered in outdoor environments. Conditions such as fog, rain, and dust can reduce visibility and affect recognition reliability. To reduce these effects, an enhancement network was trained using image pairs representing both degraded and clear scene conditions. Rather than functioning only as a contrast enhancement process, the network also learned to separate environmental artifacts from useful scene information. Motion-related effects create another source of image degradation, particularly when object movement or camera displacement occurs during operation. Temporal information obtained from consecutive image frames was therefore incorporated into the preprocessing stage to assist with blur reduction. Motion estimation combined with learned deblurring mechanisms helped recover image details that might otherwise be lost. A short-term history buffer was maintained so that corrective processing could be applied whenever motion-related distortion became significant [29].

3.2 Robust Object Detection Network

Maintaining reliable object detection becomes increasingly difficult when scene characteristics change during operation [30]. Variations related to illumination, background complexity, and external disturbances can affect recognition accuracy and overall system behavior. For this reason, both computational efficiency and robustness were considered during network development.

Feature extraction relied on a modified ResNet-50 structure together with squeeze-and-excitation modules that introduced channel-level attention capability. Attention mechanisms helped the network emphasize useful visual information while reducing the influence of irrelevant disturbances. Such behavior became particularly beneficial when scene appearance changed over time.

Information generated at different network levels was integrated through multi-scale feature fusion so that objects appearing at different sizes could be represented more effectively. Instead of treating scales independently, feature pyramid structures enabled interaction across multiple levels. This exchange improved recognition performance in situations involving partial occlusion and visually cluttered backgrounds [31].

Detection confidence also required attention because prediction scores may not always reflect actual reliability under unfamiliar situations. Temperature scaling based on validation data was therefore applied to improve calibration of confidence values. More reliable confidence estimates made uncertain detections easier to identify when operating situations differed from those represented during training.

3.3 Temporal Integration and Tracking

Maintaining tracking continuity across image sequences becomes more difficult when scene quality is affected by disturbances or unreliable detections. Recurrent learning models were introduced so that information obtained from earlier observations could contribute to current decisions [22]. Internal state representations supported preservation of object identity even when temporary occlusions or missed detections occurred.

Object movement was estimated using Kalman filtering together with learned dynamics models capable of predicting future positions. During periods of reduced detection reliability, prediction information helped preserve tracking continuity until stronger observations became available. Learned motion representations also improved handling of irregular movement patterns that are difficult to capture using simple linear prediction approaches.

Additional temporal consistency checks were included to reduce identity switching and false detections that may arise when image frames are analyzed independently. Object trajectories were examined according to physical consistency because abrupt movement behavior or sudden appearance changes may indicate unreliable detection events. Observations that violated such constraints received reduced confidence values or were removed from further processing.

Comparative detection performance under different operating scenarios is summarized in Table 1, where results from multiple approaches under dynamic situations are presented.

Table 1. Object Detection Performance Across Environmental Conditions

Environmental Condition	**Conventional Vision**	**Standard CNN**	**AI-Adaptive System (Proposed)**	**Improvement**
Optimal (controlled lab)	96.2%	97.8%	98.1%	+1.9% vs conventional
Moderate lighting variation	78.4%	88.2%	94.3%	+15.9% vs conventional
Severe lighting variation	52.1%	71.4%	89.7%	+37.6% vs conventional

Motion blur (moderate)	71.8%	84.6%	92.1%	+20.3% vs conventional
Motion blur (severe)	43.2%	68.9%	85.4%	+42.2% vs conventional
Weather effects (light)	81.3%	87.9%	93.8%	+12.5% vs conventional
Weather effects (heavy)	58.7%	73.2%	88.3%	+29.6% vs conventional
Combined challenges	39.6%	64.5%	83.7%	+44.1% vs conventional

3.4 Adaptive Parameter Control

Operating situations can change continuously during practical deployment, and such variations may influence system behavior in different ways. Illumination shifts, weather-related effects, and changes in scene complexity can alter image characteristics and affect processing requirements. Because of these variations, system parameters may require continuous adjustment rather than depending on fixed settings.

A meta-learning strategy was incorporated to support adaptive behavior during operation. Recent system performance together with observed scene characteristics formed the basis for parameter adjustment. A control network monitored factors including detection confidence, tracking consistency, and estimated scene-state information. These observations guided changes in preprocessing intensity, detection thresholds, and tracking-related settings according to current operating behavior. Parameter selection was learned through reinforcement learning, where system performance during detection tasks acted as the feedback source. Unlike approaches that rely on repeated manual calibration, updates occurred automatically while the system operated.

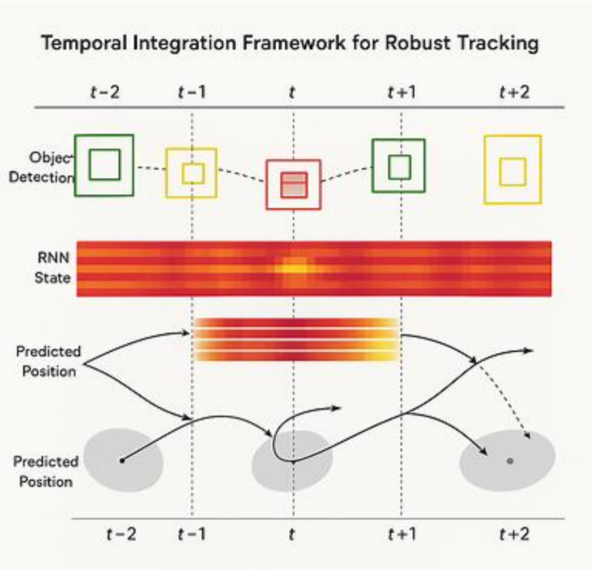

Figure 2. Temporal Integration Framework for Robust Tracking

Different situations often required different responses. Variations associated with illumination changes, weather transitions, or increasing scene complexity occasionally altered processing requirements over time. Under such situations, parameter modifications were introduced gradually across consecutive image frames. This helped avoid instability while maintaining sufficient responsiveness when operating situations changed. The Temporal Integration Framework associated with robust tracking is illustrated in Figure 2.

4. EXPERIMENTAL RESULTS AND DISCUSSION

Experimental analysis included publicly available benchmark datasets together with data collected under different operating situations. Object detection capability and behavior under dynamic scenes were examined using datasets such as COCO (Common Objects in Context) and the KITTI Vision Benchmark Suite. Additional samples from subsets of the Open Images Dataset were also considered so that robustness could be examined across a broader range of practical situations. Approximately 45,000 image frames were used during simulation and system evaluation.

System behavior was examined using commonly adopted measures including mean Average Precision (mAP) for object detection, Intersection over Union (IoU) thresholds of 0.5 and 0.75, and Multiple Object Tracking Accuracy (MOTA) for tracking analysis. To maintain consistency during experimentation, the available data were divided into training, validation, and testing groups using a 70:15:15 ratio.

Comparative analysis involved multiple approaches. Conventional computer vision techniques based on edge detection and thresholding served as baseline methods. Standard CNN-based detection models without adaptive capability were also included for comparison. Observations from these methods were then examined alongside results obtained from the adaptive AI framework. Experimental observations across different iterations and operating situations are summarized in Table 1 and Table 2.

4.1 Overall Performance Assessment

Observations across different datasets and operating situations revealed noticeable differences among the approaches examined in this work [23]. Under controlled laboratory settings, performance remained relatively similar and accuracy differences generally stayed below 2%. Larger variations became visible as operating situations became increasingly difficult.

The adaptive AI framework-maintained detection accuracy above 88% across individual challenge categories and above 83% when multiple disturbances appeared simultaneously, as summarized in Table 1. Conventional machine vision approaches experienced a larger reduction in performance, with accuracy values decreasing below 60% under severe situations and below 40% when multiple effects occurred together.

Standard CNN-based approaches generally produced intermediate results between these cases. Learning-based methods improved performance when compared with conventional approaches, while additional gains became noticeable after adaptive mechanisms were incorporated into the framework. These observations indicate that adaptive behavior contributes to more stable operation under changing situations. Processing behavior together with resource requirements is summarized in Table 2.

Table 2. Processing Performance and Resource Requirements

System Component	Processing Time (ms/frame)	GPU Memory (MB)	CPU Usage (%)	Power Consumption (W)
Conventional Vision	52	0 (CPU only)	45	8
Standard CNN	48	1,200	12	35
AI-Adaptive (full)	62	1,850	18	42
AI-Adaptive (Proposed)	38	980	15	28

Processing speed analysis revealed that the AI-adaptive (Proposed) system requires less computation than conventional methods but remains within real-time constraints for most applications. The AI-Adaptive (full) system processes approximately 18 frames per second on mid-range GPU hardware, adequate for many surveillance and inspection tasks. An efficiency-optimized (Proposed model) achieves 15 frames per second while maintaining accuracy advantages over conventional approaches.

4.2 Performance Under Environmental Variations

System behavior varied according to the type and severity of scene variation encountered during operation. Handling illumination changes remained one of the stronger characteristics of the adaptive framework, with detection accuracy remaining above 89% even under severe lighting variation. Under comparable situations, conventional approaches experienced a much larger reduction in accuracy and, in several cases, performance moved close to near-random behavior. Illumination compensation together with attention-based processing appeared to support more stable recognition across changing lighting situations [24].

The influence of motion blur depended largely on movement severity. Moderate blur situations were handled with relatively stable behavior, whereas larger reductions became noticeable under extreme motion associated with high-speed movement or extended exposure duration. Temporal integration recovered useful information when motion effects were distributed across multiple image frames, although difficulties remained when individual frames contained only limited visual detail.

Atmospheric effects produced different outcomes depending on intensity. Mild fog and light rainfall caused only small reductions in recognition accuracy, whereas dense fog introduced greater difficulty. Enhancement networks reduced the influence of rain streaks and light haze with reasonable effectiveness; however, severe atmospheric effects remained difficult to compensate completely. These findings suggest that certain practical limitations may arise from the nature of scene degradation itself rather than from algorithmic constraints alone.

4.3 Tracking Consistency Evaluation

Tracking analysis further emphasized the value of temporal information for preserving object identity across image sequences. The adaptive framework achieved tracking accuracy close to 94.7% under extended sequences containing occlusions and scene disturbances. By comparison, frame-level detection without temporal integration produced 78.3% accuracy, while identity switching occurred more frequently under difficult situations.

Recurrent learning models contributed noticeably during temporary detection failures. When object visibility was interrupted because of occlusion or scene-related effects, learned motion representations continued predicting object positions until stronger observations became available again. This behavior reduced fragmentation across tracking sequences and helped preserve continuity over time.

4.4 Adaptation Mechanism Analysis

Observations during experimentation showed that the adaptive parameter control mechanism adjusted effectively while operating situations changed over time [25]. Variations in preprocessing intensity, detection thresholds, and tracking settings changed continuously according to estimated scene behavior. Parameter updates occurred gradually and did not introduce noticeable instability during operation.

Additional ablation experiments examined system behavior after removing the adaptive control mechanism while preserving the remaining AI components. Results indicated that explicit adaptation contributed approximately 8–12% improvement beyond performance associated with learned feature invariance alone. The observed improvement suggests that adaptive control mechanisms provide measurable benefit beyond improvements obtained solely from increasingly sophisticated detection architectures.

5. DISCUSSION

5.1 Theoretical Contributions

Observations made during experimentation provided useful insight into machine vision behavior under uncertain operating situations. Combining deep learning–based feature extraction with adaptive control mechanisms appeared beneficial when scene characteristics changed over time. Models trained using diverse datasets improved resistance to variation; however, practical situations occasionally introduced patterns that differed from those represented during training. Under such circumstances, adaptive mechanisms allowed the system to respond more effectively.

Temporal information also contributed beyond object motion analysis alone [26],[27]. Information retained across successive image frames supported continuity during tracking and reduced the influence of temporary detection failures. In several situations, temporal consistency provided an additional source of guidance that improved system reliability without requiring extra labeled information.

Limitations associated with environmental compensation methods also became apparent. Moderate degradation caused by illumination variation or atmospheric effects could often be reduced successfully. More severe degradation remained difficult to recover once visual information became heavily distorted or unavailable. Such behavior suggests that practical constraints may arise from physical limitations rather than algorithmic design alone. Performance degradation trends under different challenge categories are illustrated in Figure 3.

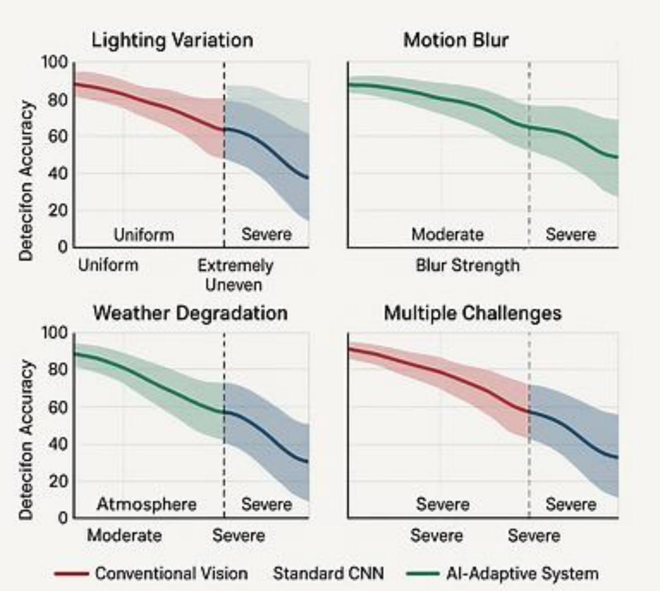

Figure 3: Performance Degradation Curves Across Environmental Challenges

5.2 Practical Implementation Considerations

Practical deployment of AI-supported vision systems involves considerations extending beyond recognition accuracy [28],[29]. Real-time processing frequently depends on GPU-supported computation, which may increase hardware requirements and power consumption when compared with conventional approaches. At the same time, improved adaptability may reduce dependence on specialized lighting arrangements and tightly controlled operating settings.

Preparation of training data remains another important consideration, particularly for domain-specific applications. Transfer learning can reduce data requirements to some extent; however, additional tuning often becomes necessary when systems are introduced into specialized environments. Effective deployment therefore depends not only on model design but also on access to representative datasets collected across varied situations. System integration introduces further considerations. Machine vision outputs commonly serve as inputs for decision-making or control processes operating with predefined confidence assumptions. Since AI-based systems may exhibit error characteristics different from conventional approaches, downstream processes may require corresponding adjustment.

5.3 Limitations and Constraints

Certain limitations remain within the present investigation and indicate areas requiring further study [31]. Experimental evaluation included a range of operating situations; however, practical environments can vary considerably and not every circumstance can be represented during testing. Additional validation under unfamiliar situations would provide broader understanding of system behavior.

Computational requirements remain another practical concern. Resource-constrained hardware platforms may still experience deployment difficulty despite continuing improvements in processing capability. Model updates and fine-tuning may also become necessary as scene characteristics change over time. The investigation focused primarily on detection and tracking tasks. Other machine vision applications, including fine-grained inspection and three-dimensional reconstruction, involve different requirements and may require additional development before adaptive methods can be applied effectively.

5.4 Future Research Directions

Additional investigation may focus on several related directions. Systems capable of learning continuously from deployment experience may support long-term improvement while reducing dependence on repeated retraining. Reliable uncertainty estimation methods may also strengthen decision-making under unfamiliar situations.

Combining visual information with complementary sensing modalities such as radar or thermal imaging may improve robustness when individual sensing approaches become unreliable. Emerging sensing technologies, including neuromorphic vision systems with fine temporal resolution, may also create opportunities for handling highly dynamic scenes more effectively. Investigation of AI techniques designed around such sensing capabilities may support development of increasingly adaptive machine vision systems.

6. CONCLUSION

Machine vision systems operating in uncertain and dynamic situations continue to experience challenges because changing surroundings can affect recognition reliability and tracking behavior. This study explored the use of adaptive AI mechanisms together with temporal integration techniques to improve system performance under such situations.

Experimental observations showed noticeable improvement in robustness after adaptive processing and learning-based components were incorporated into the proposed approach. Detection accuracy remained above 88% under individual challenge categories and above 83% when multiple disturbances occurred simultaneously. Under similar situations, conventional approaches experienced substantially larger reductions in performance. These differences emphasize the importance of adaptive behavior when machine vision systems operate beyond carefully controlled settings.

The developed architecture incorporated adaptive preprocessing, environmental compensation, robust detection methods, temporal integration, and online parameter adjustment within a unified structure. Individual components contributed differently during operation and supported more stable behavior across changing situations. Improvements were observed in handling illumination variation, preserving tracking continuity, and responding to weather-related disturbances and motion effects.

Observations from the present investigation provide useful guidance for both practical deployment and future system development. Applications involving manufacturing, surveillance, and autonomous operation may benefit from approaches capable of responding to scene variation without depending heavily on carefully engineered operating environments.

Ongoing developments in adaptive learning methods and computational platforms may further improve practical deployment of intelligent vision systems. At the same time, additional investigation remains necessary for handling unfamiliar situations and supporting long-term operation under continuously changing surroundings.

FUNDING INFORMATION: There is no funding or research grant or contract is involved for this research described in the article.

CONFLICT OF INTEREST STATEMENT: Authors states that they have no conflict of interest.

INFORMED CONSENT: We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL: Not Applicable for this research.

DATA AVAILABILITY: The data that support the findings of this study are available on request from the corresponding author, [Sahoo].

ACKNOWLEDGMENT

This work was not supported by any Funding Sources.

DECLARATION OF THE USE OF AI TOOLS

AI-assisted language editing tools were used only for improving grammar, readability, and manuscript presentation. The authors were fully responsible for technical content, interpretation, and final manuscript preparation.

References:

1. Panda, Aaryan, Damodar Panigrahi, Shaswata Mitra, Sudip Mittal, Shahram Rahimi. "Transfer Learning Applied to Computer Vision Problems: Survey on Current Progress, Limitations, and Opportunities." arXiv, Vol.1, no.1, pp.1-16, September 2024, doi:10.48550/arXiv.2409.07736
2. Deka, Brajen Kumar. "A Comparative Study of Deep Learning Architectures for Real-Time Object Detection in Smart Surveillance Systems." International Journal of Machine Learning, AI & Data Science Evolution, Vol.1,no.1, pp.32-42, March,2025, doi:10.63665/ijmlaidse.v1i1.04
3. Zhou, H.A., Wolfschläger, D., Florides, C. et al. "Generative AI in industrial machine vision: a review". J Intelligent Manufacturing, pp.1-24, April, 2025. doi.: 10.1007/s10845-025-02604-6
4. Shokri, Danesh, Christian Larouche, and Saeid Homayouni. "A comparative analysis of multi-label deep learning classifiers for real-time vehicle detection to support intelligent transportation systems." Smart Cities, Vol.6, no.5, pp.2982-3004, October, 2023, doi.: 10.3390/smartcities6050134
5. Guo, Meng-Hao, Xu, TX., Liu,JJ. et al. "Attention mechanisms in computer vision: A survey." Computational visual media, Vol.8, pp. 331-368, March, 2022, doi.org/10.1007/s41095-022-0271-y.
6. Hafiz, Abdul Mueed, Shabir Ahmad Parah, and Rouf Ul Alam Bhat. "Attention mechanisms and deep learning for machine vision: A survey of the state of the art." arXiv, pp.1-24, June, 2021. https://doi.org/10.48550/arXiv.2106.07550
7. Bi, Ying, Bing Xue, Pablo Mesejo et al. "A survey on evolutionary computation for computer vision and image analysis: Past, present, and future trends." IEEE Transactions on Evolutionary Computation, Vol. 27, no.1, pp.5-25, February, 2023, DOI: 10.1109/TEVC.2022.3220747.
8. Singhal, Peeyush, Walambe R, Ramanna S, Kotecha K., "Domain adaptation: challenges, methods, datasets, and applications." IEEE access, Vol.11, pp. 6973-7020, January, 2023, doi: 10.1109/ACCESS.2023.3237025.
9. Li, Zhengyang. "Advances in deep reinforcement learning for computer vision applications." Journal of Industrial Engineering and Applied Science, Vol.2, no.6, pp.16-26, December, 2024, ISSN 3005-6071.
10. Rieger, Martina, Günther Knoblich, and Wolfgang Prinz. "Compensation for and adaptation to changes in the environment." Experimental Brain Research, Vol.163, no.4, pp.487-502, 2005.
11. Latif, Rana Muhammad Amir, Kenji Yoshigoe, Nasir Jamal, Yue Zhao et. al. "Edge-AI Architecture for Real-Time Computer Vision in Smart Media Applications", IEEE International Conference on Pattern Recognition, Machine Vision and Artificial Intelligence (PRMVAI). IEEE, pp.1-6, June, 2025, doi: 10.1109/PRMVAI65741.2025.11108373
12. Beltrán-Escobar, M.; Alarcón, T.E.; Rumbo-Morales, J.Y.; López, S.; Ortiz-Torres, G.; Sorcia-Vázquez, F.D.J. "A Review on Resource-Constrained Embedded Vision Systems-Based Tiny Machine Learning for Robotic Applications" , Algorithms, Vol.17, no.11, pp.476, October, 2024, doi:10.3390/a17110476
13. Brodzicki, Andrzej, Piekarshi M, Kucharski D, et. al. "Transfer learning methods as a new approach in computer vision tasks with small datasets" Foundations of Computing and Decision Sciences, Vol. 45, no.3, pp.179-193, September, 2020, doi: 10.2478/fcds-2020-0010.
14. Ahmad, Hafiz Mughees, Dario Morle, and Afshin Rahimi. "Enhanced Data Augmentation for Robust Deep Learning.", Pattern Recognition, Elsevier, Vol.172, pp.112332, August, 2025, doi: 10.1016/j.patcog.2025.112332.
15. Lopes, Alexandre, dos Santos FP, de Oliverira D, et al. "Computer vision model compression techniques for embedded systems: A survey." Computers & Graphics, Vol.123, pp.104015, October 2024, doi: 10.1016/j.cag.2024.104015.
16. Cheng, Yong, Wang W, Zhang W, Yang L. et al. "A multi-feature fusion and attention network for multi-scale object detection in remote sensing images." Remote Sensing, Vol.15, no.8, PP.2096, April, 2023, doi: 10.3390/rs15082096.
17. Liu, Meiqin, et al. "Temporal consistency learning of inter-frames for video super-resolution", IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no.4, pp.1507-1520, October, 2022, doi: 10.1109/TCSVT.2022.3214538
18. Sun, Chong, Dong Wang, and Huchuan Lu. "Occlusion-aware fragment-based tracking with spatial-temporal consistency", IEEE Transactions on Image Processing, Vol. 25, no.8, pp.3814-3825, June, 2016, doi: 10.1109/TIP.2016.2580463.
19. Liu, Minghui, Ma J, Zheng, Q, Liu Y. and Shi G. "3D object detection based on attention and multi-scale feature fusion." Sensors, Vol.22, no.10, pp. 3935, May,2022, doi:10.3390/s22103935
20. Zeng, Nianyin, Wu P, Wang Z, Li H. et al. "A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection." IEEE Transactions on Instrumentation and Measurement, Vol.71, pp.1-14, February, 2022, doI: 10.1109/TIM.2022.3153997.
21. Shin, Heejung, and Hyondong Oh. "Neural network model compression algorithms for image classification in embedded systems." Information Engineering and Applied Computing, Vol.1, no.1, pp. 16-27, December, 2022.
22. Brillet, Lucas Fernandez. "Convolutional Neural Networks for embedded vision", Doctoral dissertation, Université Grenoble Alpes, February, 2020, https://theses.hal.science/tel-03101523v1
23. Anand, Surabhi, Miglani S, Anand R. et. al. "Comprehensive Study of Machine Vision and Image Analysis in Industry 5.0." Machine Vision Analysis in Industry 5.0., Chapman and Hall/CRC, pp.140-155, September, 2025, doi: 10.1201/9781003532699.
24. Matsuzaka, Yasunari, Yashiro R., "AI-based computer vision techniques and expert systems." AI, Vol.4, no.1, pp. 289-302, February, 2023, doi: 10.3390/ai4010013.

25. Jiang, Wei, Zhang K, Wang N, Yu M. et. al. "MeshCut data augmentation for deep learning in computer vision", Plos one, Vol. 15, no.12, pp.e0243613, December, 2023, doi: 10.1371/journal.pone.0243613.
26. Jan, Obaid Rafiq, Jo HS, Jo RS, Kua J.,"Real-time flood monitoring with computer vision through edge computing-based Internet of Things." Future Internet, Vol.14, no.11, pp.308, October, 2022, doi: 10.3390/fi14110308.
27. White, Jack, Tatiana Kameneva, and Chris McCarthy. "Vision processing for assistive vision: A deep reinforcement learning approach." IEEE Transactions on Human-Machine Systems, Vol. 52, no.1, pp.123-133, November, 2021, doi: 10.1109/THMS.2021.3121661.
28. Chung, J., Shen, B., & Kong, Z. J., "Anomaly detection in additive manufacturing processes using supervised classification with imbalanced sensor data based on generative adversarial network". Journal of Intelligent Manufacturing. Vol. 35, no.5, pp.2387-2406, June, 2024, doi: 10.1007/s10845-023-02163-8
29. Cheng, Z., & Behdinan, K, "Deep learning hotspots detection with generative adversarial network-based data augmentation", Journal of Micro/Nanopatterning, Materials and Metrology.Vol.21, no.2, pp.024201, April, 2022, doi.org/10.1117/1.JMM.21.2.024201
30. Chen, X., Mao, Y., Zhang, B., Chai, Y., & Yang, Z., "A method for imbalanced fault diagnosis based on self-attention generative adversarial network". International conference on Neural computing for advanced applications, Springer, Vol.1449, pp. 333–346, August, 2020, doi.org/10.1007/978-981-16-5188-5_24
31. Dong, X., Zhang, Y., Li, H., Yan, Y., Li, J., Song, J., & Koch, A. W., "Microscopic image deblurring by a generative adversarial network for 2D nanomaterials: Implications for wafer-scale semiconductor characterization", ACS Applied Nano Materials, Vol.5, no.9, pp.12855–12864, September, 2022, doi.org/10.1021/acsanm.2c02725