JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 10 Number 1 | June 2025: 205-215
DOI: 10.15575/join.v10i1.1576

Generative Adversarial Networks in Object Detection: A Systematic Literature Review

Anis Farihan Mat Raffei[1]*, Sinung Suakanto[2], Faqih Hamami[3], Mohd Arfian Ismail[4], Ferda Ernawan[5]
[1,4,5]Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pahang, Malaysia
[2,3]Department of Information Systems, Faculty of Industrial Engineering, Telkom University, Bandung, Indonesia

Article Info

Article history:

Received February 19, 2025
Revised April 17, 2025
Accepted April 24, 2025
Published June 05, 2025

Keywords:

Computer Vision
GANs
Image Processing
Object Detection
Occlusion

ABSTRACT

The intersection of Generative Adversarial Networks (GANs) and object detection represents one of the most promising developments in modern computer vision, offering innovative solutions to longstanding challenges in visual recognition systems. This review presents a systematic analysis of how GANs are transforming these challenges, examining their applications from 2020 to 2025. The paper investigates three primary domains where GANs have demonstrated remarkable potential: data augmentation for addressing data scarcity, occlusion handling techniques designed to manage visually obstructed objects, and enhancement methods specifically focused on improving small object detection performance. Analysis reveals significant performance improvements resulting from these GAN applications: data augmentation methods consistently boost detection metrics such as mAP and F1-score on scarce datasets, occlusion handling techniques successfully reconstruct hidden features with high PSNR and SSIM values, and small object detection techniques increase detection accuracy by up to 10% Average Precision in some studies. Collectively, these findings demonstrate how GANs, integrated with modern detectors, are greatly advancing object detection capabilities. Despite this progress, persistent challenges including computational cost and training stability remain. By critically analyzing these advancements and limitations, this paper provides crucial insights into the current state and potential future developments of GAN-based object detection systems.

Corresponding Author:

Anis Farihan Mat Raffei,
Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah,
Pekan, Pahang, Malaysia. 26600
Email: anisfarihan@umpsa.edu.my

1. INTRODUCTION

Computer vision has grown rapidly in recent years, making object detection essential in fields such as self-driving cars, medical imaging, and factory automation [1], [2]. The evolution of object detection systems has been driven by the demand for accurate and reliable visual recognition in complex environments. These systems process vast amounts of visual data to identify and localize objects, making them integral to applications ranging from surveillance to quality control in manufacturing [3]. However, the increasing complexity of real-world scenarios, along with the need for robust performance across diverse conditions, presents persistent challenges, including data scarcity, real-time processing demands, and difficulties in detecting small objects.

Deep learning has revolutionized object detection, driving a shift from traditional systems based on handcrafted features and classical machine learning algorithms to more advanced architectures [4]. Convolutional Neural Networks (CNNs) have played a pivotal role in this transformation by enabling automated feature learning, resulting in significant improvements in detection accuracy and generalization across diverse applications [5]. Modern object detection frameworks, such as two-stage detectors (e.g., Faster R-CNN) and single-stage detectors (e.g., YOLO and SSD), have further optimized performance by balancing accuracy and computational efficiency [2], [5]. Despite these advancements, challenges such as handling occlusions, detecting small objects, and maintaining robust performance under varying environmental conditions persist [1], [6]. Additionally, data scarcity remains a critical issue, particularly in specialized domains where annotated datasets are limited [4]. Addressing these limitations has driven interest in innovative approaches, including the integration of generative models like GANs, which offer new possibilities for enhancing object detection capabilities.

The emergence of Generative Adversarial Networks (GANs) presents a promising framework for addressing fundamental object detection challenges through innovative data generation and feature learning approaches [7]. With the global generative AI market projected to reach USD 967.65 billion by 2032 [8], and object detection being a key application area, GANs have become increasingly significant in practical applications. By using the adversarial training paradigm, GANs offer sophisticated solutions that enhance detection capabilities across multiple dimensions. GANs have significantly advanced object detection through diverse methodological applications. Researchers have leveraged GANs to generate synthetic training data to address scarcity, improve generalization, and enhance robustness [4], [9]. Additionally, GANs serve as powerful enhancement methods, including specific occlusion handling strategies to better manage partially obscured objects and techniques focused on improving small object detection performance, often by enhancing feature representations [10]. These applications demonstrate GANs' fundamental contributions to object detection advancement.

While extensive research has examined object detection and GAN architectures independently, a comprehensive analysis comparing their intersection in recent practical applications remains limited. This review addresses this research gap by systematically examining and comparing GAN applications in object detection reported between 2020 and 2025. The analysis focuses specifically on work within three key application domains where GANs address fundamental challenges: Data Augmentation, Occlusion Handling, and Small Object Detection. Within each of these domains, this review compares the methodologies and performance evaluation frameworks reported by different researchers.

The contributions of this review are as follows:

1. This review examines applications of GANs in object detection between 2020-2025.
2. This paper analyzes GAN applications in object detection, categorizing them based on their role in addressing the three key challenge domains (Data Augmentation, Occlusion Handling, Small Object Detection) and providing insights into the specific GAN architectures employed. key challenges and providing insights into the specific GAN architectures employed.
3. This paper presents a comparative analysis of performance metrics for object detection tasks when employing GAN-based approaches, highlighting the effectiveness of different methods across various applications.
4. This review identifies key challenges and proposes future directions to improve object detection using GANs.

2. METHOD

The application of deep learning in object detection has gained significant attention from the research community, resulting in numerous survey papers examining various aspects of this field. While these surveys have contributed valuable insights into object detection techniques, the specific role of

GANs in addressing object detection challenges deserves focused attention. This section reviews relevant surveys to contextualize the current work and highlight its unique contribution, as summarized in Table 1.

Table 1. Summary of relevant surveys

Reference	Year	Primary focus	GAN Coverage	Main Contribution
[11]	2024	Small Object Detection Challenges	Moderate	Analysis of key challenges in small object detection
[12]	2023	Evolution of Object Detection	Limited	Comprehensive Chronological Review
[13]	2023	Deep Learning Object Detection	Limited	Systematic analysis of detection architectures
[14]	2023	GANs in Computer Vision	Extensive	Overview of GAN developments in computer vision
[15]	2023	Underwater Object Detection	Limited	Analysis of underwater-specific detection challenges
[16]	2023	GANs for Occlusion Handling	Extensive	Analysis of GANs in handling occlusion problems
[17]	2021	Deep Learning in Self-Driving Cars	Limited	Analysis of deep learning for autonomous vehicle perception
[18]	2020	Small Object Detection	Extensive	Comprehensive review of small object detection methods

Several surveys have focused on general object detection approaches and their evolution. Kaur and Singh [13] provided a comprehensive analysis of deep learning-based object detection, examining both two-stage and one-stage detectors, along with various backbone networks. Their work emphasized the performance trade-offs between detection accuracy and speed but only briefly touched upon GAN applications. Similarly, Gupta and colleagues [17] explored deep learning applications in autonomous driving, highlighting the importance of robust object detection systems while addressing challenges in real-world scenarios. Small object detection, a particularly challenging aspect of object detection, has been the focus of multiple surveys. Tong et al. [18] conducted a systematic analysis of deep learning approaches for small object detection, notably including GAN-based detection methods among their five key aspects. Their work highlighted the effectiveness of approaches like Perceptual-GAN and MTGAN in generating enhanced representations of small objects. Wei and colleagues [11] further expanded on this by examining four fundamental challenges in small object detection, incorporating GAN-based solutions as part of their comprehensive analysis. Domain-specific survey have provided valuable insights into specialized application. Xu et al. [15] reviewed deep learning-based underwater object detection, addressing unique challenges like color distortion, low contrast, blur effects, and domain shift. They examined image preprocessing, detection architectures, and domain adaptation strategies. While image enhancement preprocessing didn't always improve detection accuracy, they explored transfer learning and domain adaptation solutions. The survey highlighted critical research directions, emphasizing the need for reliable public datasets, improved preprocessing, and robust domain generalization strategies.

Regarding GAN-specific surveys, two notable works stand out. Iglesias et al. [14] presented a comprehensive review of GANs in computer vision, analyzing architectural developments, optimization strategies, and various applications. More specifically, Saleh et al. [16] focused on GANs for addressing occlusion in images, examining applications such as a modal segmentation and content completion, while highlighting challenges in GAN training and evaluation. While extensive literature exists on object detection methodologies, GAN architectures, and their domain specific applications, there remains a significant research gap in the comprehensive analysis of GAN-based approaches within the broader context of object detection tasks. Previous survey efforts have predominantly focused on either general object detection frameworks, providing only cursory examination of GAN architectures, or have extensively analyzed GAN methodologies without sufficiently investigating their transformative impact across the spectrum of object detection challenges. These challenges encompass critical areas including data augmentation strategies, robust detection under adverse conditions, domain adaptation methodologies, and weakly supervised learning paradigms.

This review addresses this critical research gap by presenting a methodologically rigorous analysis of GAN-based approaches in object detection. Through comprehensive examination of recent technological advancements in GAN-based object detection methodologies, this review provides deeper analytical insights and broader coverage compared to existing surveys, offering a nuanced understanding of their contributions to advancing the field of computer vision. The review's scope extends beyond traditional boundaries, incorporating cutting-edge developments in GAN architectures and their applications to object detection tasks. By analyzing these advanced methodologies, this review provides researchers and practitioners with a thorough understanding of current capabilities, limitations, and future research directions in this rapidly evolving domain.

3. RESULT AND DISCUSSION

The evolution of GAN-based data augmentation has seen several notable advancements. Bosquet et al. [10] addressed the challenge of small object detection by developing a comprehensive pipeline centered around their Downsampling GAN (DS-GAN), designed to generate small objects from larger ones. Their methodology, which integrates object segmentation, image inpainting, and image blending techniques, alongside optical flow analysis for optimal object placement, demonstrated substantial performance improvements on the Unmanned Aerial Vehicle Detection and Tracking (UAVDT) and instance Segmentations in Aerial Images Dataset (iSAID) datasets. Building upon the need to handle scarce data, Xu et al. [19] introduced Scarcity-GAN, a novel approach for augmenting datasets in industrial defect detection scenarios. Scarcity-GAN leverages external datasets containing analogous defect characteristics and employs a clustering module for relevant data selection, which helps the GAN learn the specific features of defects even when training data is limited. The architecture combines an Encoder-Decoder generator with a Fusion Patch-Embedding module, ensuring precise defect localization. The model demonstrated significant performance improvements on datasets for O-ring inspection and Metal Iron Sheet analysis, showcasing its ability to generalize across different types of defects.

Further contributing to the field of defect detection, Wang and colleagues [3] developed conditional TransGAN (cTransGAN) for PCB electronic component inspection augmentation. This architecture synthesizes conditional GAN capabilities with Transformer-based architectures, incorporating class embedding conditioning in both generator and discriminator networks. The methodology was validated through extensive experimentation with PCB component and defect generation, including capacitors, diodes, optocouplers, and various defect types. Comparative analysis demonstrated consistent performance improvements across multiple object detection architectures, including Faster R-CNN, YOLOv3, and SCNet. Similarly, the study by Haruna et al. [20] utilized StyleGAN2-ADA to augment datasets for rice leaf disease detection. This work highlighted the challenges of training GANs with limited data and the importance of generating high-quality synthetic images that closely resemble real data. By combining StyleGAN2-ADA, an architecture known for its stability and image quality, with a variance of the Laplacian filter to discard blurry or poorly generated images, the authors improved the performance of Faster-RCNN and SSD models in detecting rice leaf diseases. This approach not only addressed the issue of data scarcity but also improved the models' ability to generalize across different types of rice leaf diseases.

Further extending the application of GANs to new domains, Ayub et al. [21] introduced ColorGAN, a model designed for vehicle color augmentation. ColorGAN utilizes a modified StarGAN-v2 architecture to generate diverse images of vehicles with different colors, demonstrating the feasibility of using GANs for targeted attribute modification. This method significantly improved the detection performance of a CNN model when trained with augmented images, showcasing the potential of GANs in scenarios where specific object attributes need to be varied. Taguchi and colleagues [22] investigated GAN-based synthetic-to-real domain adaptation for improving wood ear mushroom detection, comparing CycleGAN and RetinaGAN approaches against pure computer-generated imagery (CGI) and real images. Their study demonstrated that GAN-generated training data improved detection accuracy by approximately 6% in F2-score compared to using CGI alone, with RetinaGAN showing better feature preservation than CycleGAN (13% vs 21% feature degradation rate). However, when combining generated data with real images, no significant improvement was observed overusing real images alone, suggesting limitations in the approach for this specific use case. The study highlighted important considerations for GAN-based data augmentation, including the challenges of preserving object features during domain adaptation, the impact of image resolution (limited to 256×256 pixels), and the importance of computational resources for generating high-quality synthetic data.

These studies collectively demonstrate the significant advancements in GAN-based data augmentation techniques for object detection. From addressing data scarcity and improving model generalization to enabling precise control over synthesized image attributes, GANs have proven to be a powerful tool for enhancing the performance of object detection models across various applications. To provide a comparative overview of the performance of these GAN-based data augmentation methods, Table 2 summarizes the key metrics reported in each study.

Table 2. Performance summary of GAN-based data augmentation methods for object detection

Ref	Method	Dataset	Results		Drawbacks
			mAP (%)	F1-score (%)	
[3]	SCNet ResNet101 + cTransGAN	Self-developed PCB dataset	96.2		Manual selection of generated samples required
	YOLOV3 DarkNet + cTransGAN	DeepPCB	96.1		
	Faster R-CNN ResNet101 + cTransGAN	DeepPCB	98.8		
[10]	DS-GAN	UAVDT	11.9 (@.5 improvement)		limited to video datasets for optimal performance
		iSAID	4.7 (@.5 improvement)		
[19]	Scarcity-GAN	O-ring	-	85.2	Threshold selection for clustering
		Metal Iron Sheet	-	92.1	
[20]	Faster-RCNN + GAN-based augmentation	Rice leaf disease	93		Manual parameter tuning needed
[21]	ColorGAN	HACKATHON	76.1%		Resource-intensive training process
[22]	CycleGAN	1235 GAN-generated images from CGI source		72.28 (+5.51% over CGI)	GAN-based methods improved over CGI but couldn't match real image performance

Occlusion, where objects are partially or fully hidden behind other objects or elements in the scene, poses a significant challenge to object detection systems. The ability to detect and reason about occluded objects is crucial for a variety of applications, including autonomous driving, robotics, and surveillance. GANs have emerged as a promising approach to address the challenge of occlusion handling in object detection. The key idea is to leverage the generative power of GANs to imagine or hallucinate the missing parts of occluded objects, thereby enabling the detector to recognize them despite the missing visual information. One notable example of a GAN-based approach for occlusion handling is the work of Cai and colleagues [23], who proposed an occlusion-aware GAN for natural face de-occlusion. Their method, called Occlusion-Aware Generative Adversarial Network (OA-GAN), is designed to address the challenging task of restoring faces that are partially occluded by objects such as glasses, masks, or hands. Unlike traditional inpainting methods that rely on pixel-level manipulations, OA-GAN employs a two-stage generator network. The first stage predicts an occlusion mask, which identifies the regions of the face that are occluded. The second stage then uses this mask, along with the visible parts of the face, to generate a complete, de-occluded face image. By incorporating both adversarial and attribute-preserving losses, OA-GAN generate de-occluded face images that are perceptually similar to the original, non-occluded faces.

Cao et al. [24] proposed a novel approach for recognizing occluded retail products using a combination of GANs with prior inference and spherical clustering. The method addresses two key challenges: the lack of product features due to hand occlusions and high similarity between different products. The architecture uses DarkNet53 as the backbone network and incorporates semantic segmentation to locate occluded regions, followed by a prior-guided GAN for feature restoration and expansion. The system employs a multi-scale spatial attention (MSSA) mechanism combined with effective channel attention (ECA) to select discriminative features and uses von Mises-Fisher (vMF) distribution-based metric learning to improve feature distinction through spherical clustering. The approach achieves 96.34% average accuracy on occluded retail products and improves mAP by 1.2% over baseline methods. Abuhussein et al. [25] proposed a novel two-stage approach combining GANs with segmentation networks to address occlusion challenges in thermal imaging. The method first employs a modified U-Net architecture to segment occluded regions, followed by a Pix2Pix GAN-based network for inpainting the obscured areas. The system uses a reduced U-Net variant with only 485,000 parameters (compared to the standard 40.4 million) for efficient segmentation, while the Pix2Pix component handles the restoration of occluded regions through patch-wise processing. The approach achieved impressive results with PSNR of 32.5dB and SSIM of 0.89, outperforming traditional dehazing and denoising methods like BM3D and BM4D in both accuracy and processing speed. While the method shows strong performance in handling moderate occlusions, it faces challenges with dense, large-scale occlusions. The system's practical utility is demonstrated through its efficient runtime of 0.3 seconds per frame, making it suitable for real-world surveillance and security applications where thermal imaging is crucial.

Another approach to handling occlusions with GANs is presented by Meng and colleagues [26] who proposed a novel GAN-based approach called Soil Surface Occlusion Image Inpainting Network (SOI NET) for handling occlusions in soil surface images. The network employs a two-stage architecture consisting of a Rough Network and a Refinement Network, with the latter incorporating a coherent semantic attention (CSA) module to improve context information utilization. The system includes a vehicle-mounted detection component that identifies occlusions (like weeds, crop residue, and soil aggregates) and generates corresponding masks for removal. The method achieved superior performance compared to traditional and other deep learning approaches, with the highest average PSNR (21.139) and SSIM (0.754) scores across different occlusion levels. When applied to real-world farmland scenarios, the system effectively removed occlusions and improved subsequent soil moisture content estimation accuracy, increasing R^2 from 0.637 to 0.667 and reducing RMSE from 1.916% to 1.822%. The approach is particularly notable for its practical implementation in agricultural applications and its ability to handle various types of occlusions while maintaining image quality and detail preservation. Sun et al. [27] propose Ganster R-CNN, a novel architecture combining GAN with Faster R-CNN to improve occluded object detection. The model introduces two key components: Region Feature Pyramid Network (RFPN) which integrates feature maps from different scales to capture both semantic and spatial information, and IGAN (Improved GAN) which generates occluded samples to enhance the detector's capability in handling occlusions. The IGAN module employs an adversarial learning strategy where the generator creates occluded samples while the detector learns to improve detection accuracy through iterative learning. The architecture uses ResNet-101 as the backbone and incorporates a mask-based approach in the generator to create realistic occlusions while preserving object features.

While these methods have shown promising results in handling occlusions in object detection, they also have some limitations. For example, many of these methods rely on paired data for training, which can be difficult to obtain in real-world scenarios. Additionally, some methods may struggle to generate realistic and consistent completions for objects with complex shapes or textures. Table 3 provides a summary of GAN-based methods for handling occlusions.

Table 3. Performance summary of GAN-based methods for handling occlusions in object detection

Ref	Method	Dataset	Result	Drawbacks
[23]	Pix2Pix	KAIST2	PSNR: 32.5dB SSIM: 0.89	Struggles when large portions of the frame are occluded
[24]	Prior GAN with vMF clustering	Self-made	PSNR: 0.7743 SSIM: 0.0183	Real-time performance needs improvement for practical retail applications
[25]	OA-GAN	CelebA	PSNR: 22.61 dB SSIM: 0.787	Limited ability to handle extreme occlusions or multiple occlusions simultaneously
[26]	SOI NET	6000 training soil surfaces	PSNR: 21.139dB SSIM: 0.754	Cannot effectively identify objects with similar color to soil
[27]	Ganster R-CNN	MS COCO	+10.3 AP over Faster R-CNN	Poor performance on small objects like tableware

Small object detection remains a significant challenge in computer vision, particularly when objects appear small in images due to distance or environmental conditions. This challenge is further amplified in low-light conditions, where limited training data and poor visibility can severely impact detection performance. Recent approaches have leveraged GANs to address these limitations by augmenting training datasets and enhancing feature representation for small object detection. GANs have demonstrated promising capabilities in enhancing feature representation, augmenting training datasets, and improving the overall detection performance for small objects across various domains and applications. Table 4 summarizes the performance of GAN-based methods for small object detection.

Wu et al. [28] addressed small object detection challenges in nighttime conditions through an enhanced CycleGAN architecture. Their approach tackled the limitations of insufficient training data and poor visibility by incorporating ResNeXtBlocks for improved feature extraction and an optimized upsampling module. The enhanced model demonstrated substantial performance gains, achieving increases of 13.34% in precision, 45.11% in recall, and 56.52% in average precision, while improving the Fréchet Inception Distance (FID) score by 29.7% compared to the original CycleGAN architecture. Bashir and Wang [29] proposed SRCGAN-RFA-YOLO, a framework combining super-resolution with object detection to address small object detection challenges in low-resolution remote sensing images. Their approach integrates three key innovations: a residual feature aggregation (RFA) network in the generator for better feature extraction, a cyclic GAN architecture for improved training stability, and a detector-agnostic framework that can work with various object detectors. The method demonstrated significant improvements in detection metrics when dealing with highly downsampled images (scale factor of 16), achieving a 54.3% increase in Average Precision compared to bicubic interpolation and a 29.4% improvement over EDSR-RFA baseline.

Salaudeen and Çelebi [30] proposed a two-stage approach combining Enhanced Super-Resolution GAN (ESRGAN) with object detection networks (YOLOv5 and EfficientDet) to improve small pothole detection in challenging conditions. Their method addresses the limitations of detecting small objects in low-resolution images by first enhancing image quality through super-resolution before detection. The approach demonstrated significant improvements, with the ESRGAN+YOLOv5 model achieving 32% mAP (at IoU 0.5:0.95) on super-resolution images compared to 12% on low-resolution images, while ESRGAN+EfficientDet achieved 26% mAP versus 10.6% on low-resolution images. Ren et al. [31] proposed Region Super Resolution Generative Adversarial Network (RSRGAN), a novel two-stage approach combining region proposal with super-resolution for infrared small target detection. Their method first employs a lightweight Region Context Network (RCN) to efficiently extract potential target regions, followed by an enhanced SRGAN with a distribution transformation generator to improve small target detection. The approach addresses two key challenges: computational efficiency through region-based processing, and the distribution gap between real small targets and downsampled training data. The method demonstrated superior performance, achieving 0.980 mAP@.5 and outperforming state-of-the-art detectors.

Rabbi et al. [32] proposed EESRGAN-FRCNN, an end-to-end architecture combining edge-enhanced super-resolution with object detection to improve small object detection in low-resolution satellite imagery. Their approach integrates three key components: an enhanced super-resolution GAN (EESRGAN) with residual-in-residual dense blocks, an edge enhancement network (EEN) to preserve crucial edge information, and a detector network (Faster R-CNN/SSD). The method demonstrated significant improvements in detection metrics, achieving 95.5% AP with end-to-end training compared to 64% AP using standard detectors on low-resolution images. A key innovation was the introduction of edge consistency loss alongside image consistency loss to generate clearer edges and reduce noise in enhanced images.

Table 4. Performance summary of GAN-based methods for small object detection

Ref	Method	Performance	Drawbacks
[28]	Enhanced CycleGAN with ResNeXtBlocks and optimized upsampling module	Precision: +13.34% Recall: +45.11% Average Precision: +56.52% F1Score: +30.52% FID Score reduction: 29.7%	Cannot completely eliminate mode collapse
[29]	SRCGAN-RFA-YOLO	AP increase: 54.3% over bicubic interpolation AP increase: 29.4% over EDSR-RFA	Higher inference time compared to other object detectors.
[30]	ESRGAN	ESRGAN + YOLOv5:mAP: 32% ESRGAN + EfficientDet: mAP: 26%	Higher computational cost
[31]	RSRGAN	mAP@.5: 0.980	Computationally expensive
[32]	EESRGAN-FRCNN	COWC Dataset: AP: 95.5% OGST Dataset: AP: 83.2%	Dataset-specific performance variations

3.1 *Challenges and Future Directions*

Despite the significant advances in GAN-based approaches for object detection, several critical limitations persist that warrant attention from the research community. The computational complexity and resource requirements remain a significant barrier to widespread adoption, as training GANs demands substantial computational resources and time, particularly for high-resolution image generation and complex architectures. This computational burden becomes especially pronounced when combining GAN-based enhancement with detection pipelines, making real-time performance challenging and limiting deployment on resource-constrained devices. Training stability and convergence continue to pose significant challenges in GAN implementation. Mode collapse remains a persistent issue, particularly in domain-specific applications, while achieving stable training across different weather conditions and lighting scenarios requires careful consideration. The delicate balance required between generator and discriminator performance necessitates careful hyperparameter tuning, adding complexity to the development process.

Data dependencies also represent another crucial limitation, as many approaches rely heavily on paired training data, which is often scarce or expensive to obtain. This challenge is particularly severe in domain-specific applications where annotated training data may be limited. While synthetic data generation offers a potential solution, generated data may not fully capture the complexity and variations present in real-world scenarios, potentially limiting model generalization. Performance gaps continue to exist across different aspects of object detection. Detection accuracy for small objects consistently lags that of larger objects, especially in adverse conditions. Integration challenges further complicate the practical implementation of GAN-based approaches in object detection systems. There are persistent difficulties in seamlessly combining GAN-based enhancement with existing detection pipelines, often requiring complex trade-offs between image enhancement quality and detection speed. The limited standardization in evaluation metrics and benchmarking procedures also makes it challenging to compare different approaches effectively and establish best practices.

Looking ahead, several promising research directions emerge that could address current limitations and advance the field of GAN-based object detection. In terms of architectural innovations, future research should focus on developing lightweight GAN architectures optimized for resource-constrained environments. This includes investigating hybrid approaches combining GANs with transformers for improved feature learning and exploring multi-task architectures that can jointly handle enhancement and detection tasks efficiently. Training methodologies represent another crucial area for future research. There is a pressing need to develop self-supervised and weakly-supervised training approaches that reduce reliance on paired data. Additionally, research into adaptive training strategies that can dynamically adjust to different environmental conditions could significantly improve model robustness. The investigation of transfer learning techniques to improve generalization across domains also holds promise for enhancing model performance.

Performance enhancement remains a critical focus area, particularly in addressing specific challenges like small object detection through specialized GAN architectures. Future research should prioritize developing more robust approaches for extreme weather conditions and severe occlusions, while also investigating domain-specific optimization techniques for targeted applications. This includes exploring novel loss functions and architectural modifications that can better preserve fine details while maintaining computational efficiency. Efficiency improvements represent another vital research direction, focusing on model compression and quantization techniques for GAN-based detection systems. The development of adaptive inference pipelines that can balance performance and computational cost could make these systems more practical for real-world applications. Research into efficient architectures specifically designed for real-time applications could help bridge the gap between laboratory demonstrations and practical deployments.

Finally, standardization and evaluation frameworks require significant attention from the research community. The development of comprehensive evaluation frameworks for GAN-based detection systems, along with standardized benchmarks across different environmental conditions, could help establish more reliable comparisons between different approaches. Creating unified metrics that consider both enhancement quality and detection performance would provide better insights into the practical utility of different methods. These research directions collectively aim to address current limitations while advancing the field toward more robust, efficient, and practical implementations of GAN-based object detection systems.

4. CONCLUSION

This comprehensive review has examined the intersection of GANs and object detection, highlighting significant advancements and innovations between 2020-2025. Through systematic analysis of various applications and methodologies, the review has demonstrated how GANs have transformed object detection capabilities across multiple dimensions, from data augmentation and occlusion handling to small object detection. While substantial progress has been made, with some approaches showing remarkable improvements in detection metrics, several challenges persist, including computational complexity, training stability, and data dependencies. The integration of GANs with modern detection architectures continues to evolve, suggesting promising directions for future research in areas such as lightweight architectures, self-supervised learning, and adaptive training strategies. As the field matures, future developments are expected to yield more practical implementations that further push the boundaries of what's possible in computer vision applications, particularly in addressing real-world challenges such as resource constraints, extreme environmental conditions, and domain-specific requirements.

ACKNOWLEDGEMENTS

This research was supported by Universiti Malaysia Pahang Al-Sultan Abdullah (Grant No: RDU242712 & UIC241514). The authors also acknowledge the support from the International Matching Grant Scheme provided by Telkom University, Indonesia.

REFERENCES

[1] M. Ahmed, K. A. Hashmi, A. Pagani, M. Liwicki, D. Stricker, and M. Z. Afzal, “Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments,” *Sensors*, vol. 21, no. 15, p. 5116, Jul. 2021, doi: 10.3390/s21155116.

[2] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, “A survey of modern deep learning based object detection models,” *Digit Signal Process*, vol. 126, p. 103514, Jun. 2022, doi: 10.1016/j.dsp.2022.103514.

[3] C. Wang, G. Huang, Z. Huang, and W. He, “Conditional TransGAN-Based Data Augmentation for PCB Electronic Component Inspection,” *Comput Intell Neurosci*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/2024237.

[4] Y. Lu, D. Chen, E. Olaniyi, and Y. Huang, “Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review,” *Comput Electron Agric*, vol. 200, p. 107208, Sep. 2022, doi: 10.1016/j.compag.2022.107208.

[5] L. Jiao *et al.*, “New Generation Deep Learning for Video Object Detection: A Survey,” *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 8, pp. 3195–3215, Aug. 2022, doi: 10.1109/TNNLS.2021.3053249.

[6] Y. Liu, P. Sun, N. Wergeles, and Y. Shang, “A survey and performance evaluation of deep learning methods for small object detection,” *Expert Syst Appl*, vol. 172, p. 114602, Jun. 2021, doi: 10.1016/j.eswa.2021.114602.

[7] K. Zhang, X. Yang, L. Xu, J. Thé, Z. Tan, and H. Yu, “Enhancing coal-gangue object detection using GAN-based data augmentation strategy with dual attention mechanism,” *Energy*, vol. 287, p. 129654, Jan. 2024, doi: 10.1016/j.energy.2023.129654.

[8] Fortune Business Insights, “Generative AI Market Size, Share, Growth | Forecast [2030].” Accessed: Mar. 01, 2025. [Online]. Available: https://www.fortunebusinessinsights.com/generative-ai-market-107837

[9] G. Yang, C. Song, Z. Yang, and S. Cui, “Bubble detection in photoresist with small samples based on GAN augmentations and modified YOLO,” *Eng Appl Artif Intell*, vol. 123, p. 106224, Aug. 2023, doi: 10.1016/j.engappai.2023.106224.

[10] B. Bosquet, D. Cores, L. Seidenari, V. M. Brea, M. Mucientes, and A. Del Bimbo, “A full data augmentation pipeline for small object detection based on generative adversarial networks,” *Pattern Recognit*, vol. 133, p. 108998, Jan. 2023, doi: 10.1016/j.patcog.2022.108998.

[11] W. Wei, Y. Cheng, J. He, and X. Zhu, “A review of small object detection based on deep learning,” *Neural Comput Appl*, vol. 36, no. 12, pp. 6283–6303, Apr. 2024, doi: 10.1007/s00521-024-09422-6.

[12] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, “Object Detection in 20 Years: A Survey,” *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257–276, Mar. 2023, doi: 10.1109/JPROC.2023.3238524.

[13] R. Kaur and S. Singh, “A comprehensive review of object detection with deep learning,” *Digit Signal Process*, vol. 132, p. 103812, Jan. 2023, doi: 10.1016/j.dsp.2022.103812.

[14] G. Iglesias, E. Talavera, and A. Díaz-Álvarez, “A survey on GANs for computer vision: Recent research, analysis and taxonomy,” *Comput Sci Rev*, vol. 48, p. 100553, May 2023, doi: 10.1016/j.cosrev.2023.100553.

[15] S. Xu, M. Zhang, W. Song, H. Mei, Q. He, and A. Liotta, “A systematic review and analysis of deep learning-based underwater object detection,” *Neurocomputing*, vol. 527, pp. 204–232, Mar. 2023, doi: 10.1016/j.neucom.2023.01.056.

[16] K. Saleh, S. Szénási, and Z. Vámossy, “Generative Adversarial Network for Overcoming Occlusion in Images: A Survey,” *Algorithms*, vol. 16, no. 3, p. 175, Mar. 2023, doi: 10.3390/a16030175.

[17] A. Gupta, A. Anpalagan, L. Guan, and A. S. Khwaja, “Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues,” *Array*, vol. 10, p. 100057, Jul. 2021, doi: 10.1016/j.array.2021.100057.

[18] K. Tong, Y. Wu, and F. Zhou, “Recent advances in small object detection based on deep learning: A review,” *Image Vis Comput*, vol. 97, p. 103910, May 2020, doi: 10.1016/j.imavis.2020.103910.

[19] C. Xu *et al.*, “Scarcity-GAN: Scarce data augmentation for defect detection via generative adversarial nets,” *Neurocomputing*, vol. 566, p. 127061, Jan. 2024, doi: 10.1016/j.neucom.2023.127061.

[20] Y. Haruna, S. Qin, and M. J. M. Kiki, “An Improved Approach to Detection of Rice Leaf Disease with GAN-Based Data Augmentation Pipeline,” *Applied Sciences*, vol. 13, no. 3, p. 1346, Jan. 2023, doi: 10.3390/app13031346.

[21] A. Ayub and H. Kim, “GAN-Based Data Augmentation with Vehicle Color Changes to Train a Vehicle Detection CNN,” *Electronics (Basel)*, vol. 13, no. 7, p. 1231, Mar. 2024, doi: 10.3390/electronics13071231.

[22] H. Taguchi, R. Matsumura, and H. Kitakaze, “The effectiveness of GAN-based synthetic-to-real domain adaptation methods in training a wood ear mushroom detection model,” *Artif Life Robot*, vol. 30, no. 2, pp. 310–316, May 2025, doi: 10.1007/s10015-024-00998-9.

[23] J. Cai, H. Han, J. Cui, J. Chen, L. Liu, and S. K. Zhou, “Semi-Supervised Natural Face De-Occlusion,” *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1044–1057, 2021, doi: 10.1109/TIFS.2020.3023793.

[24] M. Cao *et al.*, “Recognition of Occluded Goods under Prior Inference Based on Generative Adversarial Network,” *Sensors*, vol. 23, no. 6, p. 3355, Mar. 2023, doi: 10.3390/s23063355.

[25] M. Abuhussein, I. Almadani, A. L. Robinson, and M. Younis, “Enhancing Obscured Regions in Thermal Imaging: A Novel GAN-Based Approach for Efficient Occlusion Inpainting,” *J (Basel)*, vol. 7, no. 3, pp. 218–235, Jun. 2024, doi: 10.3390/j7030013.

[26] C. Meng, W. Yang, Y. Bai, H. Li, H. Zhang, and M. Li, “Research of soil surface image occlusion removal and inpainting based on GAN used for estimation of farmland soil moisture content,” *Comput Electron Agric*, vol. 212, p. 108155, Sep. 2023, doi: 10.1016/j.compag.2023.108155.

[27] K. Sun, Q. Wen, and H. Zhou, “Ganster R-CNN: Occluded Object Detection Network Based on Generative Adversarial Nets and Faster R-CNN,” *IEEE Access*, vol. 10, pp. 105022–105030, 2022, doi: 10.1109/ACCESS.2022.3211394.

[28] F. Wu *et al.*, “An Enhanced Cycle Generative Adversarial Network Approach for Nighttime Pineapple Detection of Automated Harvesting Robots,” *Agronomy*, vol. 14, no. 12, p. 3002, Dec. 2024, doi: 10.3390/agronomy14123002.

[29] S. M. A. Bashir and Y. Wang, "Small Object Detection in Remote Sensing Images with Residual Feature Aggregation-Based Super-Resolution and Object Detector Network," *Remote Sens (Basel)*, vol. 13, no. 9, p. 1854, May 2021, doi: 10.3390/rs13091854.

[30] H. Salaudeen and E. Çelebi, "Pothole Detection Using Image Enhancement GAN and Object Detection Network," *Electronics (Basel)*, vol. 11, no. 12, p. 1882, Jun. 2022, doi: 10.3390/electronics11121882.

[31] K. Ren, Y. Gao, M. Wan, G. Gu, and Q. Chen, "Infrared small target detection via region super resolution generative adversarial network," *Applied Intelligence*, vol. 52, no. 10, pp. 11725–11737, Aug. 2022, doi: 10.1007/s10489-021-02955-6.

[32] J. Rabbi, N. Ray, M. Schubert, S. Chowdhury, and D. Chao, "Small-Object Detection in Remote Sensing Images with End-to-End Edge-Enhanced GAN and Object Detector Network," *Remote Sens (Basel)*, vol. 12, no. 9, p. 1432, May 2020, doi: 10.3390/rs12091432.