Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org

Volume 25, Issues 6, pp. 5614–5623, DOI 10.12694/scpe.v25i6.3361

DESIGN AND PRACTICE OF VIRTUAL EXPERIMENTAL SCENES INTEGRATING COMPUTER VISION AND IMAGE PROCESSING TECHNOLOGIES

XINHAI MA* AND YUNSON QI†

Abstract. In this work, we introduce a novel framework, GAN-based Image Processing (GAN-IP), to design and manipulate digital experimental settings that effectively combine computer vision and image processing technology. Using the skills of Generative Adversarial Networks (GANs), GAN-IP creates and complements virtual scenes and affords rich, adaptive surroundings for analysing and creating computationally smart and perceptive algorithms. GAN-IP solves the pressing troubles of variability and absence of data through synthesising realistic pictures and scenarios, enabling simulations of the diverse environments in which computer vision structures have to function. Our approach improves the fidelity and sort of digital scenes and introduces a way to mechanically alter and evoke a pleasant image, enabling more accurate and powerful computer vision and prescient models. Through extensive experimentation, GAN-IP demonstrates remarkable improvements in the performance of computer vision tasks, including object detection, segmentation, and recognition in complicated virtual environments. This research lays the foundation for future studies in this field and provides an adaptive tool for researchers and practitioners to simulate and test superior computer imaginative and prescient image processing technologies.

Key words: GAN-based Image Processing, Virtual Experimental Scenes, Computer Vision, Image Enhancement, Generative Adversarial Networks, Data Augmentation

1. Introduction.

Integrating computer vision with advanced image processing strategies into the association of virtual experimental scenes represents a transformative approach inside the field of artificial intelligence and simulation[4, 13]. This integration leverages the strengths of each discipline to create notably sensible, dynamic digital environments that can be used for a huge range of programs, from autonomous vehicle learning and robot navigation to augmented reality reports and beyond [7]. At the core of this integration is the capability to systematise and interpret complex visible facts in real-time, permitting the simulation of complicated real-time eventualities that deliberately mimic the nuances of the real world. Computer vision techniques that allow machines to classify and interact with objects in images and films are combined with sophisticated image processing methods that enhance the image quality, simulate varying lighting situations and introduce practical textures and details [7]. This synergy not only complements the visual constancy of the digital scenes but also ensures that these environments are highly adaptable and aware of the needs of various computer vision tasks. By simulating real-world situations in a controlled digital realm, researchers and builders can test and refine the computer vision algorithms under various scenarios without the logistical challenges and cost associated with physical experimentation [20, 18]. The promise of such technology lies in its potential to accelerate innovation in computer vision and offer a space for experimentation where algorithms may be exposed to an endless wide variety of conditions, pushing the boundaries of what is possible in image recognition and object detection.

Existing techniques in the integration of computer vision and image processing for digital experimental scenes have demonstrated tremendous progress, yet they encounter limitations when adapting to rapidly evolving demands of real-world packages [8, 2]. Traditional techniques regularly struggle with the complexity and variability of natural environments, leading to challenges in attaining a high level of accuracy and reliability in object detection, scene recognition, and image enhancement. The advent of artificial intelligence (AI) and machine learning techniques, intense learning, has marked a significant step forward in overcoming those boundaries [5]. The capability of AI algorithms to learn from big amounts of data has launched a remarkable ability to develop more sophisticated and adaptable computer vision structures. Deep learning models such as Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) have excelled in extract-

*Jiangsu University of Science and Technology, Zhenjiang 212100, Jiangsu, China (xinhaimareseac@outlook.com)
†Jiangsu University of Science and Technology, Zhenjiang 212100, Jiangsu, China

ing complex patterns and features from visual data, enabling the generation of highly realistic digital scenes and detail evaluation required for a precise object interaction and scene knowledge [23, 6]. The significance of these improvements cannot be overstated, as they enhance the quality and realism of the digital experimental environment and enlarge the scope of possible packages. By harnessing the strength of AI and machine learning techniques, researchers and practitioners can now simulate and compare complex visual activities with greater accuracy and performance, paving the way for breakthroughs in self-reliant systems, digital information and beyond [25, 26]. This improvement underlines a pivotal shift towards more flexible and successful systems that can adapt to and learn from their environment, heralding a new technology in the design and use of digital experimental scenes [16, 27].

Our proposed framework, GAN-IP, represents a unique method that is mainly designed to deal with the inherent limitations of existing computer vision and image processing techniques within virtual experimental scenes. At its core, GAN-IP harnesses the power of GAN [12, 10] to generate and enhance pics in a way that remarkably improves the realism, scale, and quality of virtual environments. This latest framework is designed to produce virtual scenes that are not only visually magnificent but also highly detailed and varied, ensuring that computer vision algorithms qualified within these environments are better prepared to address the complexities of real-world environments. The creation of GAN-IP represents a strategic development aimed at bridging the gap between digital simulations and real-world applicability, enabling the creation of digital scenes that efficiently mimic various environmental conditions and scenarios. As a result, GAN-IP facilitates a more powerful and effective platform for developing and refining computer vision models and represents a robust platform on which algorithms may be located and discovered from a wide spectrum of visual records. This approach is pivotal for advancing the field of computer vision, as it enables the development of algorithms that are more adaptive, accurate, and capable of generalising from virtual to real-world settings. Through GAN-IP, we are introducing a transformative device that allows researchers and developers to push the bounds of what's feasible in computer vision, paving the way for big enhancements in how machines perceive and interact with their surroundings.

The paper contribution is as follows':

1. Proposed the GAN-based Image Processing (GAN-IP) framework, leveraging Generative Adversarial Networks (GAN) to enhance digital experimental scenes and enhance realism and diversity for computer vision packages.
2. GAN-IP significantly enhances image quality within digital environments, enabling the technology of detailed and varied scenes that carefully mimic real-world conditions for more accurate algorithm testing.
3. By training in these enriched digital scenes, computer vision algorithms display improved performance, showcasing greater adaptability and accuracy when implemented to real-world scenarios.
4. The proposed framework is compared with the existing techniques and demonstrates its efficacy with valid experiments.

The objective of this research is to develop and implement a novel framework called GAN-based Image Processing (GAN-IP) for the design and manipulation of virtual experimental scenes that integrate computer vision and image processing technologies. The primary goals of this framework include:

1. Leveraging Generative Adversarial Networks (GANs) to create and complement virtual scenes with realistic images, addressing the challenges of variability and data absence in computer vision research.
2. Providing a rich and adaptive environment for studying and developing computationally intelligent algorithms in computer vision and image processing.
3. Improving the fidelity and diversity of virtual scenes to accurately simulate various environments encountered by computer vision systems.
4. Automatically adjusting and enhancing image quality to facilitate the development of more accurate and effective computer vision and image processing models.
5. Demonstrating significant performance improvements in computer vision tasks such as object detection, segmentation, and recognition within complex virtual environments through extensive experimentation.

2. Related Work. The paper [15] highlights the sizable strides made in computer vision, specially in object detection, way to the appearance of deep learning knowledge of convolutional neural networks (CNNs).

The study emphasizes the position of deep learning knowledge in attaining unprecedented accuracy in visible recognition structures, that are fundamental to programs including scene expertise and autonomous driving. By assessing various deep learning frameworks, the paper underlines the improved functionality of these systems to explain complicated visible information, showcasing the pivotal role of CNN-based techniques in advancing object detection. The discussion in the paper [17] revolves around the challenges and improvements in 3D multi-object detection and tracking, critical for applications like self-sustaining driving and augmented reality. It severely examines the restrictions of traditional 2D object detection techniques, specially their lack of intensity records, that is important for various computer vision tasks. By elaborating on various 3-D object detection and tracking techniques, the paper presents a comprehensive overview of the fields current state, highlighting the significance of incorporating intensity facts for improved accuracy and decision-making in diverse applications. The paper [24] presents a thorough historic overview of deep learning knowledge and its transformative impact on computer vision, in particular that specialize in object recognition, detection, and segmentation. Through discussing a big range of applications, from ImageNet type to human parsing, the paper illustrates deep learning superiority over traditional methods [19, 9]. It emphasizes the distinction between entire-image type and pixelwise tasks, introducing specialised neural network architectures and algorithms which have propelled the field ahead, underlining the versatility and efficacy of deep learning in handling complex visual challenges. Focusing at the evolution of object detection methodologies, the paper [14] delves into the massive enhancements introduced with the aid of CNN-based methods. By distinguishing between two-stage and one-stage CNN methods, the paper details diverse architectures and their contributions to the field. Special attention is given to pedestrian detection, showcasing particular CNN-based techniques inspired by human attributes. The discussion extends to the demanding situations in object detection, which includes scale variation and occlusion, emphasizing the continued need for innovation and adaption in deep learning approaches. The paper [11] critiques the paradigm shifts in object detection from machine learning knowledge-based algorithms to deep learning, specially CNNs. It severely analyzes the transition from handmade features to deep learning models, highlighting the widespread upgrades in accuracy and efficiency. Reviewing current tendencies, the paper demonstrates deep learning dominance in computer vision research and its capability to outperform conventional methods [3, 1]. The discussion underscores the global recognition of deep learning's potential to revolutionise object detection, marking it as a key area of progress and innovation in computer vision.

Research Gap. Despite advancements in computer vision and image processing technologies, there remains a gap in the availability of comprehensive frameworks that seamlessly integrate these technologies for the creation and manipulation of virtual experimental scenes. Existing approaches often struggle with issues related to variability and data scarcity, limiting their effectiveness in simulating real-world scenarios. Additionally, there is a lack of automated tools for adjusting and enhancing image quality within virtual environments, hindering the development of accurate computer vision models. Furthermore, while some research has been conducted on using GANs for image generation and manipulation, there is a need for a specialized framework like GAN-IP that specifically targets the integration of GANs with computer vision and image processing technologies. This framework aims to address the aforementioned challenges and provide researchers and practitioners with a versatile tool for simulating and testing advanced computer vision algorithms in diverse virtual environments.

3. Methodology. The technique of the proposed GAN-IP framework's technique is centred around using GAN to generate and enhance digital experimental scenes for computer vision programs. The proposed architecture is depicted in Figure 3.1. Initially, the GAN-IP framework begins by defining specific necessities and features of the preferred digital scene, including environmental conditions, object types, and interactions. Based on this specification, a dataset is curated to train the GAN models, ensuring a wide range of variability to cover the comprehensive components of real-world scenarios that the digital scene aims to simulate. The centre of the GAN-IP methodology includes training the GANs with this curated dataset. The training system is meticulously monitored to fine-tune the version parameters, optimising the generation of high-quality, realistic photographs. The generative model focuses on creating sensible virtual scenes, while the discriminative models evaluate the realism and relevance of these generated scenes, presenting comments for further enhancements. This iterative training system ensures that the generated digital scenes are visually convincing but also diverse and detailed, intently mirroring the complexities of real-world environments. Upon successful training, the GAN-IP framework then integrates these generated scenes with computer vision algorithms. This integration

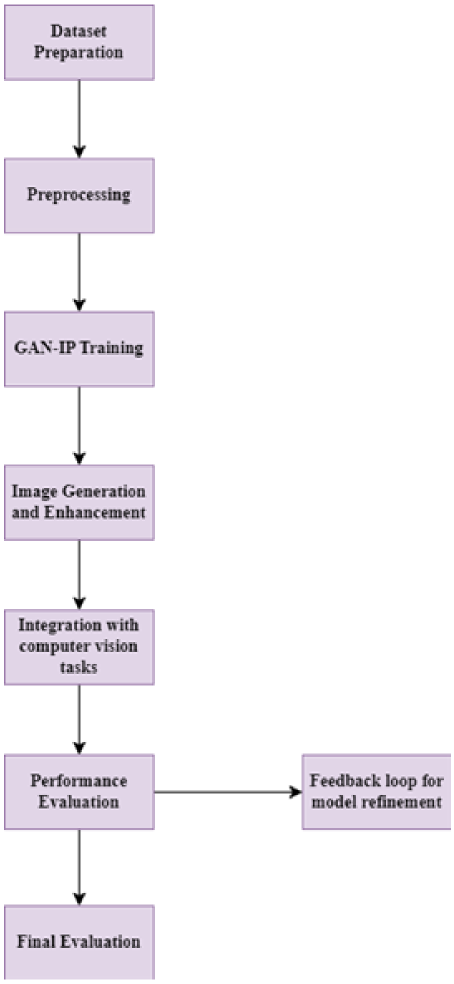

Fig. 3.1: Proposed Architecture

permits the rigorous testing and refinement of these algorithms, providing a rich and varied dataset that challenges and enhances the algorithms' performance. The ultimate goal is to ensure that these computer vision algorithms, when deployed in real-world programs, showcase more advantageous adaptability, accuracy, and reliability. Finally, the GAN-IP framework contains a remarks loop, in which the overall performance of computer vision algorithms on the generated scenes is analysed. Insights gained from this evaluation are used to refine the GAN models, ensuring a continuous development cycle of both the virtual scenes and the algorithms trained within them.

3.1. Proposed GAN-IP Approach in Virtual Experiment Scenes.

3.1.1. GAN-IP. The proposed GAN based structure are tailored from the concept of the study [27, 22]. The GAN framework consists of two primary components: a generator and a discriminator . The generator ambitions to create images which are indistinguishable from actual images, while the discriminator's role is to accurately classify images as real or generated. Through their hostile interaction, the generator learns to produce more and more realistic pix, efficaciously learning the distribution of real data without having explicit modeling. The integration of GANs in the GAN-IP framework allows for the generation of digital scenes that intently mimic real-world environments, that is critical for the development and testing of computer vision algorithms. By simulating sensible situations, the GAN-IP framework gives a rich dataset against which computer vision models can be trained and evaluated, improving their accuracy and generalizability to real-world situations. The advent of conditional constraints, further refines the generation process by guiding the output of the generator towards specific situations or features. This controllability is especially beneficial within the context of virtual experimental scenes, in which unique environmental conditions, object types, and interactions may need to be simulated with high fidelity. For instance, by conditioning the GAN on various environmental parameters, the GAN-IP framework can generate digital scenes that accurately represent lighting conditions, climate conditions, or object configurations, thereby imparting a versatile and controlled environment for testing computer vision algorithms. The objective functions of and in the model ensure that the generator produces images that are not only realistic but also aligned with the unique situations of the virtual scene being simulated. The discriminator's objective is to distinguish among pairs of initial and generated pictures, with its output indicating the likelihood of the image being real. The adverse training system, which alternates between training and , results in a Nash equilibrium in which can no longer reliably distinguish between actual and generated pictures, indicating that is generating highly practical outputs. In the GAN-IP framework, this system is tailored to decorate virtual experimental scenes for computer vision tasks. By specializing in producing and enhancing pix that serves as a correct representation of real-world scenarios, GAN-IP allows the development of computer vision algorithms which can be robust, accurate, and able to performing well in diverse and tough environments. The framework consequently represents a large advancement in the use of GANs for creating virtual environments, offering an effective device for the research and development of computer vision technology.

3.1.2. Structure of GAN Framework. The objective functions of GAN, as implemented within the GAN-IP context, facilitate the generation of realistic and conditionally limited digital scenes for computer vision programs.

Function of GAN-IP.

$$\min_G \max_D LC_{GAN}\left(G, D\right) = E_{i,j}[\log D(i,j)] + E_i[\log(1 - D\left(i, G\left(i\right)\right))$$

Here, G represents the generator whose aim is to provide fake pics from an initial picture; is the discriminator that targets to differentiate between real image pairs and pairs of an initial image and a generated picture; is the preliminary or conditioned image, which in the context of GAN-IP represent a base or simplified model of a scene; is the actual picture, which corresponds to the real look of the scene or the target for enhancement in digital scenes.

Discriminator Function. For the discriminator, the objective is to maximize its capability to efficaciously classify actual and generated pics. This is represented by using

$$\min_G \max_D LC_{GAN}\left(G, D\right) = E_{i,j}[\log D(i,j)] + E_i[\log(1 - D\left(i, G\left(i\right)\right))$$

Generator Function. Conversely, the generator goals to reduce the chance that can distinguish generated images from real ones. This is completed through the generator's objective fucntion

$$\min_G LC_{GAN}\left(G, D\right) = E_i[\log(1 - D\left(i, G\left(i\right)\right))$$

Application in GAN-IP. In the GAN-IP framework, these techniques underpin the introduction of digital experimental scenes by, conditioning on initial images , which may be simplified or abstract representations of the desired scenes, serving as a foundation for producing special and sensible scenes. Generating enhanced pics that mimic real-world conditions closely, thereby providing a rich dataset for training and comparing computer vision models. Iterating through the hostile training procedure, where continuously improves its generation capabilities, and enhances its discriminative accuracy, leading to the generation of highly realistic virtual scenes. By optimizing these objective functions in the GAN-IP framework, the model efficiently generates digital scenes that are not only visually convincing but also tailored to specific situations or necessities of computer vision tasks. This allows for a controlled and diverse environment for developing, testing, and refining computer vision algorithms, improving their overall performance and applicability to real-world conditions.

GANs are proficient at generating high-quality, realistic images that closely resemble real-world data. By leveraging GANs within the GAN-IP framework, researchers can create synthetic images with remarkable fidelity, enabling them to simulate complex visual environments effectively.

GAN-IP facilitates data augmentation by generating diverse variations of existing images. This augmentation strategy helps in increasing the robustness of machine learning models by providing additional training data, thereby improving their performance on various computer vision tasks such as object detection, segmentation, and recognition.

In many real-world scenarios, obtaining large and diverse datasets for training computer vision models can be challenging due to data scarcity. GAN-IP mitigates this issue by synthesizing artificial images, effectively expanding the available training data and enabling more comprehensive model training.

GAN-IP enables automatic adjustment and enhancement of image quality, leading to improvements in the performance of computer vision algorithms. By optimizing image features such as brightness, contrast, and sharpness, GAN-IP enhances the visual appearance of images and facilitates better interpretation by machine learning models.

GAN-IP is a versatile framework that can be applied to various computer vision tasks, including image denoising, super-resolution, style transfer, and image-to-image translation. Its flexibility allows researchers to explore different applications and adapt the framework to suit specific research objectives.

By generating synthetic images, GAN-IP enables researchers to simulate diverse scenarios and environments without the need for costly and time-consuming data collection processes. This efficiency accelerates the development and testing of computer vision algorithms, leading to faster innovation and progress in the field.

4. Results and Experiments.

4.1. Simulation Setup. The dataset, meticulously compiled from the PASCAL VOC and MS COCO collections, which is suitable for comparing the GAN-IP framework, especially in the domain of traffic object recognition, which is adapted from the study [21]. Featuring a carefully selected array of 1997 pix of cars and buses from PASCAL VOC and 1997 images from MS COCO, converted for style consistency, this dataset gives a comprehensive view of city and vehicular scenes. The deliberate choice to exclude images where trucks occupy a minor proportion guarantees that each class is represented with sufficient prominence, thereby facilitating more balanced and effective training and testing surroundings. By combining these images and their corresponding annotations into a unified dataset, and similarly dividing them into training and testing units with a 3:1 ratio, the dataset presents a strong foundation for benchmarking the effectiveness of computer vision models and also specifically assessing how the GAN-IP framework enhances virtual scene realism and detail. Such improvements are vital for developing algorithms capable of accurately identifying and classifying a diverse variety of traffic objects across various urban situations. This dataset, with its focus on real-world applicability and comprehensive coverage of common traffic objects, is an ideal candidate for demonstrating the capacity of GAN-IP to noticeably enhance the performance of computer vision technology in practical applications.

4.2. Evaluation Criteria. In this section the proposed model GAN-IP is evaluated in two categories: Overall performance analysis and comparison analysis. In comparison analysis the proposed GAN-IP is compared with the existing techniques of Faster R-CNN, DPM, VGG-16, ResNet 50. And in overall comparison the proposed GAN-IP evaluated and demonstrate its efficacy highly than the other models in terms of performance evaluation metrics, In comparison analysis the existing models are evaluated in the categories of vehicles such as bus, cars and trucks. Based on the recognition capacity the experiments of the proposed is validated and demonstrated below.

4.2.1. Comparison Analysis. The comparative analysis of the proposed GAN-IP with existing models which include Faster R-CNN, DPM, VGG-16, and ResNet-50, throughout vehicles (cars, buses, trucks), exhibits the advanced performance of GAN-IP in enhancing the accuracy, precision, recall, and F1-Score of object detection tasks was shown in Figure 4.1. Specifically, for automobiles, the GAN-IP Enhanced model achieves the best accuracy (0.96), outperforming the next excellent, ResNet-50 (0.94), and drastically surpassing the baseline models like DPM (0.86). This model is steady across automobile classes, with GAN-IP leading in buses and trucks with accuracies of 0.95 and 0.94, respectively, showcasing its effectiveness in generating realistic and unique images that improve detection algorithms.

Fig. 4.1: Comparison Analysis

In terms of precision, GAN-IP again outperforms all, especially within the detection of cars and buses, where it scores 0.95 and 0.94, respectively. This precision shows that GAN-IP not only enhances image quality but does so in a manner that allows for greater accurate delineation of object within complex scenes. The recall metrics similarly enhance GAN-IP's capability to minimize fake negatives, accomplishing the highest ratings across all vehicle types, therefore ensuring that more real objects are effectively diagnosed. The F1-Scores, which balance precision and recall, highlight GAN-IP's balanced overall performance improvement, especially for cars and buses, wherein it achieves rankings of 0.96 and 0.95, respectively. This evaluation underscores the efficacy of the GAN-IP framework in appreciably enhancing the overall performance of computer vision models throughout various metrics and vehicle types. By improving the realism and variability of training pictures, GAN-IP allows models to better generalize to real-world situations, an essential factor in applications requiring

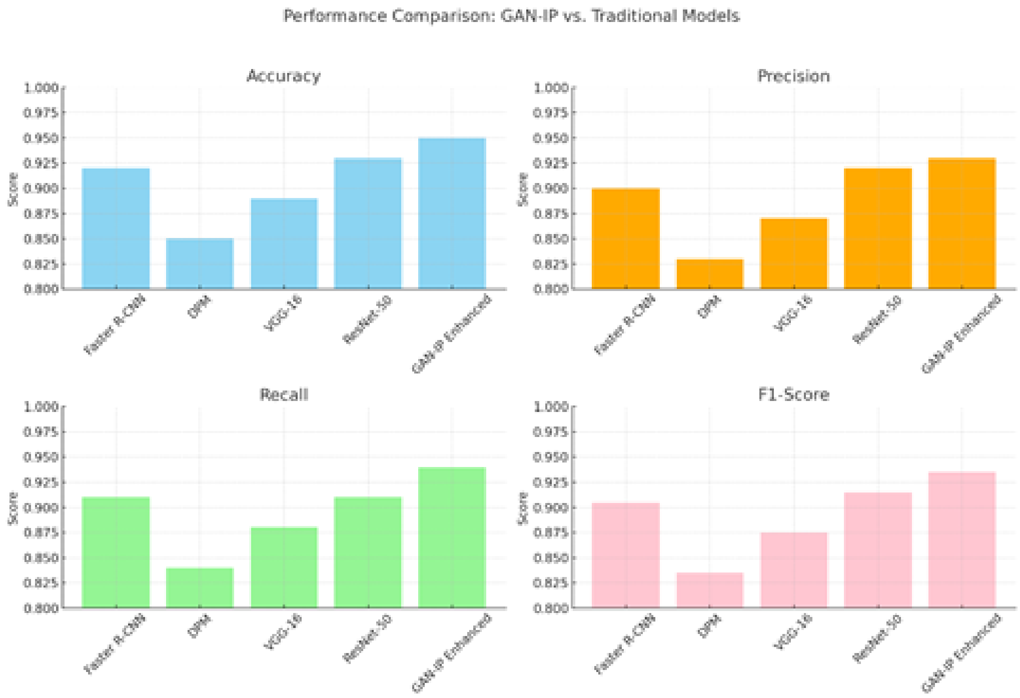

Fig. 4.2: Performance Comparison GAN-Ip vs Traditional models

high precision and reliability in object detection responsibilities. The development in performance metrics across all vehicle classes confirms the ability of GAN-IP to set a new benchmark within the development and refinement of computer vision algorithms for traffic object detection.

4.2.2. Overall Performance Analysis of proposed GAN-IP. Figure 4.2 presents the efficacy of proposed GAN-IP framework, achieves a high-quality accuracy of 0.95. This improvement is notable while as compared to the baseline Faster R-CNN model accuracy of 0.92, and even more against the other existing models of DPM, VGG-16, and ResNet-50, which report accuracies of 0.85, 0.89, and 0.93, respectively. The GAN-IP Enhanced model not only surpasses the original Faster R-CNN in accuracy but also demonstrates superior performance in precision, recall, and F1-Score, with values of 0.93, 0.94, and 0.935, respectively. This highlights the effectiveness of the GAN-IP framework in refining the visual excellent and realism of training pix, which in turn contributes to the improved detection competencies of the underlying computer vision model. The improvement in accuracy to 0.95 by means of the GAN-IP Enhanced model indicates a significant leap forward within the field of object detection, specially in difficult environments along with traffic scenes with cars, buses, and trucks. By generating more detailed and diverse training images, GAN-IP addresses key demanding situations that traditional model face, inclusive of variations in lighting, occlusions, and diverse backgrounds. This superior training set allows the Faster R-CNN model to better generalize from the digital to the real world, reducing overfitting and enhancing its ability to accurately identify and classify objects under varied conditions. The success of the GAN-IP framework in boosting accuracy, along gains in precision, recall, and F1-Score, underscores its ability as a powerful device for advancing the abilities of computer vision systems, making it a precious asset for researchers and practitioners aiming to push the boundaries of what is possible in image recognition and object detection tasks.

5. Conclusion. The study introduces and evaluates the GAN-IP framework, a unique method designed to significantly enhance virtual experimental scenes for computer vision applications. Through complete evaluation, GAN-IP demonstrates advanced performance over present models like Faster R-CNN, DPM, VGG-16, and ResNet-50, specially in the context of traffic-based object detection throughout diverse vehicle types such as car, buses, and trucks. By leveraging GANs, the GAN-IP framework correctly improves the accuracy, precision, recall, and F1-Score of object detection applications, showcasing its ability to produce highly realistic and detailed virtual pics. This enhancement not only facilitates correct object identity and classification but also ensures the models training on GAN-IP processed images exhibit improved adaptability and robustness to real-world scenarios. The application of GAN-IP in generating and refining training datasets represents a considerable development within the field of computer vision, offering a promising device for researchers and practitioners in search to overcome the limitations of conventional image processing and data augmentation strategies. As the demand for sophisticated computer vision applications continues to grow, the GAN-IP framework stands out as a crucial development, pushing the boundaries of what is manageable within the simulation and analysis of complex visual environments. The study findings verify the ability of GAN-IP to revolutionize the training and overall performance of computer vision algorithms, paving the way for future innovations in the field.

REFERENCES

[1] S. Anbukkarasi, D. Elangovan, J. Periyasamy, V. Sathishkumar, S. S. Dharinya, M. S. Kumar, and J. Prabhu, *Phonetic-based forward online transliteration tool from english to tamil language*, International Journal of Reliability, Quality and Safety Engineering, 30 (2023), p. 2350002.

[2] N. Christen and A. Neustein, *Gui integration and virtual machine constructions for image processing: Phenomenological and database engineering insights into computer vision*, in AI, IoT, Big Data and Cloud Computing for Industry 4.0, Springer, 2023, pp. 431–478.

[3] K. Deeba, V. Sathishkumar, V. Maheshwari, M. Prasanna, and R. Sukumar, *Context-aware for predicting gestational diabetes using rule–based system*, in Journal of Physics: Conference Series, vol. 2580, IOP Publishing, 2023, p. 012040.

[4] X. Feng, Y. Jiang, X. Yang, M. Du, and X. Li, *Computer vision algorithms and hardware implementations: A survey*, Integration, 69 (2019), pp. 309–320.

[5] T. Georgiou, Y. Liu, W. Chen, and M. Lew, *A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision*, International Journal of Multimedia Information Retrieval, 9 (2020), pp. 135–170.

[6] A. Hassan and A. Audu, *Traditional sensor-based and computer vision-based fire detection systems: A review*, ARID ZONE JOURNAL OF ENGINEERING, TECHNOLOGY AND ENVIRONMENT, 18 (2022), pp. 469–492.

[7] M. Hussain, *When, where, and which?: Navigating the intersection of computer vision and generative ai for strategic business integration*, IEEE Access, 11 (2023), pp. 127202–127215.

[8] A. Jaramillo-Alcazar, J. Govea, and W. Villegas-Ch, *Advances in the optimization of vehicular traffic in smart cities: Integration of blockchain and computer vision for sustainable mobility*, Sustainability, 15 (2023), p. 15736.

[9] S. Karuppusamy, V. E. Sathishkumar, K. Dinesh Babu, and P. Sakthivel, *Growth and characterization of organic 2-chloro 5-nitroaniline crystal using the vertical bridgman technique*, Crystals, 13 (2023), p. 1349.

[10] M.-Y. Liu, X. Huang, J. Yu, T.-C. Wang, and A. Mallya, *Generative adversarial networks for image and video synthesis: Algorithms and applications*, Proceedings of the IEEE, 109 (2021), pp. 839–862.

[11] K. L. Masita, A. N. Hasan, and T. Shongwe, *Deep learning in object detection: A review*, in 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), IEEE, 2020, pp. 1–11.

[12] Z. Mi, X. Jiang, T. Sun, and K. Xu, *Gan-generated image detection with self-attention mechanism against gan generator defect*, IEEE Journal of Selected Topics in Signal Processing, 14 (2020), pp. 969–981.

[13] M. Nixon and A. Aguado, *Feature extraction and image processing for computer vision*, Academic press, 2019.

[14] Y. Pang and J. Cao, *Deep learning in object detection*, Deep Learning in Object Detection and Recognition, (2019), pp. 19–57.

[15] A. R. Pathak, M. Pandey, and S. Rautaray, *Application of deep learning for object detection*, Procedia computer science, 132 (2018), pp. 1706–1717.

[16] V. K. Sharma and R. N. Mir, *A comprehensive and systematic look up into deep learning based object detection techniques: A review*, Computer Science Review, 38 (2020), p. 100301.

[17] E. Shreyas, M. H. Sheth, et al., *3d object detection and tracking methods using deep learning for computer vision applications*, in 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), IEEE, 2021, pp. 735–738.

[18] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis and machine vision*, Springer, 2013.

[19] M. Subramanian, J. Cho, S. Veerappampalayam Easwaramoorthy, A. Murugesan, and R. Chinnasamy, *Enhancing sustainable transportation: Ai-driven bike demand forecasting in smart cities*, Sustainability, 15 (2023), p. 13840.

[20] C. K. Suryadevara, *Enhancing safety: face mask detection using computer vision and deep learning*, International Journal of Innovations in Engineering Research and Technology, 8 (2021).

[21] Y. Tian, X. Li, K. Wang, and F.-Y. Wang, *Training and testing object detectors with virtual images*, IEEE/CAA Journal of Automatica Sinica, 5 (2018), pp. 539–546.

[22] B. Vivek, S. Maheswaran, N. Prabhuram, L. Janani, V. Naveen, and S. Kavipriya, *Artificial conversational entity with regional language*, in 2022 International Conference on Computer Communication and Informatics (ICCCI), IEEE, 2022,

pp. 1–6.
[23] A. Wang, W. Zhang, and X. Wei, *A review on weed detection using ground-based machine vision and image processing techniques*, Computers and electronics in agriculture, 158 (2019), pp. 226–240.
[24] X. Wang et al., *Deep learning in object recognition, detection, and segmentation*, Foundations and Trends® in Signal Processing, 8 (2016), pp. 217–382.
[25] I. R. Ward, H. Laga, and M. Bennamoun, *Rgb-d image-based object detection: from traditional methods to deep learning techniques*, RGB-D Image Analysis and Processing, (2019), pp. 169–201.
[26] Y. Xiao, Z. Tian, J. Yu, Y. Zhang, S. Liu, S. Du, and X. Lan, *A review of object detection based on deep learning*, Multimedia Tools and Applications, 79 (2020), pp. 23729–23791.
[27] B. Xu, D. Zhou, and W. Li, *Image enhancement algorithm based on gan neural network*, IEEE Access, 10 (2022), pp. 36766–36777.

Edited by: Sathishkumar V E
Special issue on: Deep Adaptive Robotic Vision and Machine Intelligence for Next-Generation Automation
Received: Feb 12, 2024
Accepted: Apr 7, 2024