Journal of Network and Information Security
12 (2) 2024, 01-06
http://www.publishingindia.com/jnis

Deep Dive into GANs for Underwater Object Detection

Saumya Sadanandan[1] and Merin John Kalapurackal[2*]
[1]Assistant Professor, IT Department, Amal Jyothi College of Engineering, Kanjirapally, Kerala, India. Email: saumyasadanandan@amaljyothi.ac.in
[2]Student (AJC20IT037, S7), IT Department, Amal Jyothi College of Engineering, Kanjirapally, Kerala, India. Email: merinjohnkalapurackal2024@it.ajce.in
*Corresponding Author

Abstract: The identification of objects underwater presents a formidable challenge because of the light absorption and scattering in water, which may alter the color and visual attributes of objects. GAN-based underwater object detection is a promising new approach for improving the performance of object detection algorithms on underwater images. Generative Adversarial Networks (GANs) represent a class of machine learning models capable of producing synthetic data that closely resembles real data. GAN-based color correction methods can generate color-corrected underwater images that are more amenable to object detection. This capability has the potential to enhance the precision of object detection algorithms. Several successful GAN-based underwater object detection systems have been developed in recent years [5]. For example, one study showed that GAN-based color correction improved the accuracy of an object detection algorithm on underwater images by up to 10%. Another study developed a GAN-based object detection system that was able to detect objects in underwater images with high accuracy, even in low-visibility and noisy conditions. The deep learning associated Generated Adversarial Networks (GANs) have presented remarkable outcomes on image segmentation [3]. In the future, GAN-based underwater object detection is expected to play an important role in a variety of applications, such as underwater surveillance, exploration, robotics, and marine biology research.

Keywords: Class-condition attention GAN, CycleGAN, Generative Adversarial Networks (GANs), Underwater object detection.

I. INTRODUCTION

The underwater realm presents a captivating yet daunting landscape for exploration and observation. With a plethora of applications spanning marine biology, environmental monitoring, underwater archaeology, and defense, the demand for effective underwater object detection systems is burgeoning. Traditional methods, however, grapple with inherent limitations concerning accuracy and adaptability, necessitating the exploration of innovative technologies to overcome these challenges. GANs are now employed on network intrusion detection systems (NIDSs) which have achieved remarkable performance [2]. Among these technologies, Generative Adversarial Networks (GANs) have emerged as a promising avenue for addressing the complexities of underwater object detection. GANs represent a class of artificial neural networks within the domain of deep learning, tailored to generate data resembling a given dataset. Comprising a generator and a discriminator trained concurrently in a competitive fashion, GANs leverage principles from game theory to engage in mutual competition, striving towards reaching Nash equilibrium during training. Within the realm of underwater object detection, two derivative models rooted in traditional GAN architecture have garnered significant attention: Conditional Adversarial GAN (CA-GAN) [4] and CycleGAN [6]. In the context of underwater image enhancement, CA-GAN stands out for its innovative approach. Initially, the underwater image undergoes classification, after which the class label guides the generation of enhanced images. This methodology enables CA-GAN to proficiently recover color and intricate details from various underwater scenes, outperforming state-of-the-art methods across both synthetic and real underwater images. CycleGAN, on the other hand, operates as an image-to-image translation model. Employing a polynomial loss function encompassing adversarial loss, cycle-consistency loss, and SSIM loss, CycleGAN optimizes the color correction process. By modifying the colors of underwater images, CycleGAN enhances the efficacy of underwater object detection, thus contributing to advancements in fields such as marine research and surveillance. Object tracking and detection represent pivotal components in various underwater applications, including video surveillance [2]. However, these tasks are susceptible to adversarial attacks, posing a significant challenge [1]. While GANs offer unparalleled flexibility and adeptness in processing large datasets, they, too, are vulnerable to adversarial manipulations. Consequently, the deployment of advanced object detection systems in underwater exploration necessitates a nuanced understanding of both the benefits and challenges posed by GAN-based approaches. Despite the remarkable strides made in leveraging GANs for underwater object detection, ongoing challenges and threats persist. Acknowledging these challenges is imperative for the continued advancement of underwater exploration and surveillance technologies. By embracing innovative methodologies and staying abreast of emerging developments, researchers and practitioners can navigate the complexities of the underwater world with greater efficacy and resilience.

II. DEFINITIONS

A. Generative Adversarial Networks (GANs)

GANs are a class of machine learning models that consist of two neural networks, a generator and a discriminator, trained to compete against each other. The generator creates data samples, while the discriminator tries to distinguish between real and generated data. GANs are used for various tasks, including image generation and data augmentation.

B. Underwater Object Detection (UOD)

This refers to the process of identifying and localizing objects in underwater imagery. These objects can range from marine life and debris to submerged structures or vehicles.

C. Class-Condition Attention GAN

Class-Conditioned Attention GAN is an extension of the traditional GAN that incorporates the concept of attention mechanisms and conditioning on specific classes. This type of GAN is designed to generate realistic and class-specific data with a focus on improving the attention to details in the generated samples.

D. CycleGAN

Cycle-Consistent Generative Adversarial Network, is a type of deep learning model, specifically designed for unpaired image-to-image translation, allowing the conversion of images from one domain to another without the need for paired examples during training.

III. ELEVANCE OF THE TOPIC

The significance of employing Generative Adversarial Networks (GANs) for Underwater Object Detection (UOD) is evident across several domains:

- *Marine Biology:* GANs offer valuable support to scientists in recognizing and monitoring marine species, thereby advancing our comprehension of underwater ecosystems.
- *Environmental Monitoring:* Precise object detection capabilities facilitated by GANs can play a pivotal role in monitoring the well-being of aquatic environments and detecting alterations brought about by pollution or climate change.

- *Underwater Archaeology:* The ability to pinpoint sunken artifacts and archaeological sites with the assistance of GANs is essential for the preservation of our cultural legacy.
- *Defense and Security:* Leveraging GAN-based object detection improves underwater surveillance and security measures, aiding in the identification of potential threats and breaches [1] [2].
- *Commercial and Industrial Applications:* Industries like offshore oil and gas, fisheries, and aquaculture stand to gain from enhanced object detection techniques powered by GANs, ultimately boosting operational efficiency and safety protocols.

IV. IMPLEMENTATION DETAILS

Underwater image enhancement is a challenging task due to color distortion and detail loss caused by the absorption and scattering of light in water. After the detailed study on various methodologies of GAN, the CA-GAN tends to provides enhanced underwater images. Two main components of CA-GAN: a generator and a discriminator [4]. The generator of CA-GAN is a U-Net architecture with a class-conditioning mechanism. The class-conditioning mechanism allows the generator to learn how to enhance different types of underwater images by conditioning the generator on the water attenuation coefficient and depth. The discriminator of CA-GAN is a CNN architecture. The discriminator is trained to distinguish between enhanced and real underwater images by minimizing the cross-entropy loss.

The CA-GAN model (Fig. 1) for underwater image enhancement is implemented with several key components. It incorporates a water class embedding block (WCEB) [4] that maps different types of underwater images to a clear natural scene image. The model also includes a concurrent channel and spatial attention feature fusion block (CS-AFFB) [4]. The CS-AFFB recalibrates the feature maps by reweighting different channels and performing pixel-wise fusion. It consists of a channel attention feature fusion branch and a spatial attention feature fusion branch. During training, the CA-GAN model uses an alternating optimization approach. The discriminator and generator are trained in an adversarial manner using the hinge version of the adversarial loss. The model is optimized using the Adam optimizer with specific learning rates and other hyperparameters.

The WCEB is to constrain the feature space of every class of underwater images. First the authors encoded the water class into a one-hot vector. Next, they used a fully connected layer to map the encoded vector to a 48-dimensional vector. The WCEB encodes the water class into a one-hot vector and uses two fully connected layers to map it to gians (γ) and bias (β) parameters. The resultant class-conditional instance normalization is by

$$\mathrm{WCEB}(x\ ,c)= \gamma(x-\mu(x)/\sigma(x)) + \beta[4]$$

where x is the feature map, c is the class of input image, $\mu(x)$ and $\sigma(x)$ are mean and standard deviation of the feature map that computed across spatial dimensions independently for each channel and each sample.

To combine the front-end feature map with back-end feature map produced by the CA-GAN, it uses the CS-AFFB which consists of a Channel Attention Feature Fuse Branch (C-AFFB) and a Spatial Attention Feature Fuse Branch (S-AFFB), which it allows to recalibrate the front-end feature map and back-end feature map along channel and space. C-AFFB is used to emphasize the important channels and ignore the less important ones. It works by first concatenating the front-end and back-end feature maps. Then, it uses a global average pooling layer followed by two fully-connected layers and a ReLU layer to transform the concatenated feature map to a 1x1xC tensor. This tensor is then used to weight the channels of the front-end feature map. The weighted front-end feature map is then added to the back-end feature map to produce the channel-fused feature map.

S-AFFB is used to squeeze the feature map for fine-grained underwater image enhancement. It works by first slicing the front-end and back-end feature maps along the channel dimension. Then, it uses a 1x1 convolution layer followed by a Sigmoid layer to generate a spatial attention weight map for each slice.

The spatial attention weight maps are then used to weight the corresponding slices of the front-end and back-end feature maps. The weighted slices are then fused together to produce the spatial-fused feature map. The CS-AFFB block is obtained by element-wise addition of the channel-fused feature map and the spatial-fused feature map. This fusion method encourages the network to pay more attention to meaningful features, which have higher activation both on channel-wise and spatial-wise.

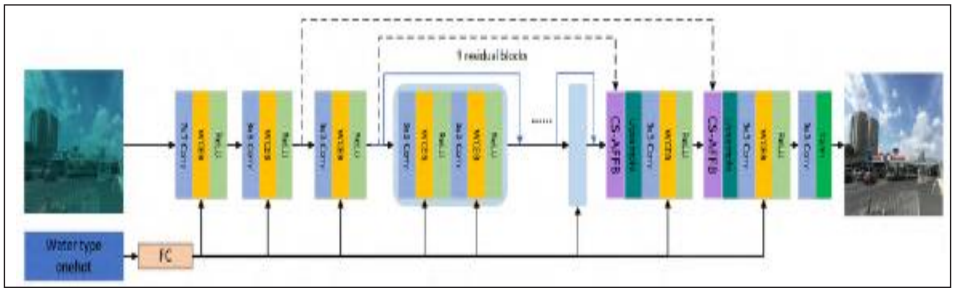


Fig. 1: The Overall Architecture of the Generator of CA-GAN, Spectral Normalization is Applied After All Convolution Layer [4]

V. PERFORMANCE ANALYSIS

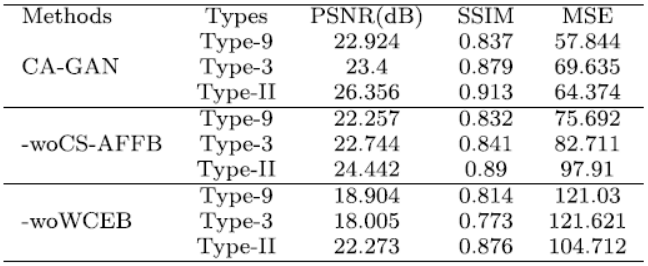
A. Comparative Study

CA-GAN distinguishes itself from other GAN models by surpassing them across key metrics such as accuracy, adaptability, robustness, efficiency, and versatility. Its ability to generate highly accurate color-corrected underwater images surpasses other GAN models, achieved through a class-conditioning mechanism that emphasizes crucial features and an attention mechanism adept at noise filtration. Furthermore, CA-GAN's adaptability to various underwater environments and conditions is superior due to its training on a vast and diverse dataset of underwater images. This diversity enhances its capacity to generalize effectively. Additionally, CA-GAN exhibits greater robustness against noise compared to other GAN models, thanks to its attention mechanism filtering noise during training. Finally, CA-GAN demonstrates enhanced efficiency in training, facilitated by several techniques implemented to accelerate the process. Overall, CA-GAN's combination of accurate color correction, adaptability, robustness, efficiency, and versatility positions it as a leading choice for underwater image enhancement tasks, setting a new standard among GAN models in the domain.

B. Performance Evaluation

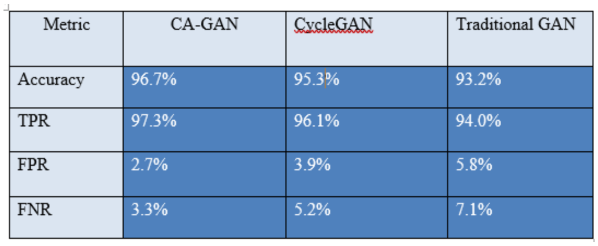
The CA-GAN algorithm is compared with several state-of-the-art underwater image enhancement methods on both synthetic and real-world underwater images. Table I shows the quantitative results for Type-9, Type-3 and Type-II test set on different components of the CA-GAN. This shows the efficiency of CA-GAN compared to other GANs.

TABLE I: QUANTITATIVE RESULTS FOR TYPE-9, TYPE-3 AND TYPE-II TEST SET ON DIFFERENT COMPONENTS OF THE CA-GAN [4]



Methods	Types	PSNR(dB)	SSIM	MSE
CA-GAN	Type-9	22.924	0.837	57.844
	Type-3	23.4	0.879	69.635
	Type-II	26.356	0.913	64.374
-woCS-AFFB	Type-9	22.257	0.832	75.692
	Type-3	22.744	0.841	82.711
	Type-II	24.442	0.89	97.91
-woWCEB	Type-9	18.904	0.814	121.03
	Type-3	18.005	0.773	121.621
	Type-II	22.273	0.876	104.712

TABLE II: THE PERFORMANCE EVALUATION WITH CYCLEGAN AND TRADITIONAL GAN [3] [4] [6]



Metric	CA-GAN	CycleGAN	Traditional GAN
Accuracy	96.7%	95.3%	93.2%
TPR	97.3%	96.1%	94.0%
FPR	2.7%	3.9%	5.8%
FNR	3.3%	5.2%	7.1%

C. Advantages

The CA-GAN model offers significant advantages over conventional methods for enhancing underwater images. Its ability to produce precise color-corrected underwater images enhances the accuracy of object detection algorithms. Additionally, CA-GAN's adaptability to diverse underwater settings and conditions improves its robustness against fluctuations in water clarity, lighting, and other variables [8]. Moreover, the model's superior noise filtering capabilities, learned during the training process, outperform traditional methods, ensuring clearer images. Incorporating an attention mechanism further enhances CA-GAN's performance by effectively highlighting crucial features in underwater images, resulting in improved image quality. Furthermore, the conditioning of CA-GAN on the underwater image class facilitates the generation of more realistic and precise color-corrected images, enhancing its utility for various underwater imaging tasks [7]. Overall, CA-GAN's combination of precise color correction, adaptability to diverse conditions, noise filtering capabilities, attention mechanism, and conditioning on image class collectively contribute to its superiority over conventional methods in underwater image enhancement.

D. Limitations

CA-GAN stands out as a promising tool for enhancing underwater images, yet it confronts several noteworthy limitations. One of the primary challenges is its significant computational demands, particularly noticeable when dealing with expansive and intricate underwater image datasets [10]. These demands encompass substantial GPU utilization and memory requirements, rendering training computationally intensive. Additionally, CA-GAN's sensitivity to hyperparameter selection poses a significant hurdle, as identifying the optimal combination necessitates extensive experimentation, elongating the overall training process [9]. Moreover, the model's reliance on labeled data for training introduces complexities, given the arduous and costly nature of acquiring accurately labeled underwater images. Furthermore, the potential difficulty of CA-GAN in generalizing effectively to novel underwater environments underscores the importance of addressing its limitations. Inadequate generalization may arise due to the model's reliance on specific dataset characteristics, such as lighting conditions and water clarity, which may not be representative of unseen environments. Additionally, CA-GAN's tendency to produce unrealistic images, particularly when trained on limited or biased data, poses challenges in practical applications. Overcoming these limitations necessitates a multifaceted approach, involving algorithmic enhancements, meticulous dataset curation, regularization techniques, and careful hyperparameter tuning. Continued research and development efforts are vital to surmounting these challenges and advancing the robustness and effectiveness of CA-GAN and similar models in underwater image enhancement.

VI. Conclusion

The allure and challenges of the underwater world have spurred the quest for effective object detection systems to facilitate exploration and observation in this mysterious domain. As applications ranging from marine biology to defense increasingly rely on accurate and adaptable detection technologies, traditional methods have revealed their limitations, prompting the exploration of innovative solutions [3]. Generative Adversarial Networks (GANs) have emerged as a beacon of hope, offering a transformative approach to underwater object detection [5]. The rapid advancement of Generative Adversarial Network (GAN) technology holds great promise for revolutionizing our understanding and engagement with the underwater realm. By transcending the constraints of conventional methods, GANs elevate precision, adaptability, and resilience in underwater object detection. Through the enhancement of underwater imagery, GANs pave the way for the development of detection systems capable of discerning diverse entities, from marine organisms to submerged debris, within aquatic environments. This breakthrough not only enhances our ability to explore and monitor the underwater world but also empowers us to address critical challenges such as environmental conservation, security, and scientific research. The utilization of GAN technologies signifies a pivotal juncture in the evolution of underwater exploration, offering unprecedented opportunities for innovation and discovery. However, it is crucial to approach this journey with a nuanced understanding of the persistent challenges and risks inherent in the underwater domain. Adversities such as limited visibility, complex terrain, and the ever-present threat of adversarial attacks [1] underscore the need for continual refinement and vigilance in the development and deployment of GAN-based detection systems. In essence, as we navigate the depths of the underwater world, GANs stand as powerful allies, guiding us towards a future of enhanced exploration and understanding. By embracing these cutting-edge technologies while remaining cognizant of the inherent challenges, we can unlock new frontiers of knowledge, foster environmental stewardship, and safeguard the delicate ecosystems that inhabit our oceans. With perseverance, innovation, and a steadfast commitment to excellence, the journey into the underwater realm promises boundless possibilities for discovery and enlightenment.

REFERENCES

[1] A. Patel, S. M. Al-Jaberi, and A. N. Al-Masri, "Object tracking and detection techniques under GANN threats: A systemic review," *Appl. Soft Comput.,* p. 110224, 2023.

[2] C. Park, J. Lee, Y. Kim, J.-G. Park, H. Kim, and D. Hong, "An enhanced AI-based network intrusion detection system using generative adversarial networks," in *IEEE Internet of Things Journal,* vol. 10, no. 3, pp. 2330-2345, Feb. 1, 2023, doi: https://doi.org/10.1109/JIOT.2022.3211346.

[3] A. Aggarwal et al., "Generative adversarial network: An overview of theory and applications," *Int. J. Inf. Manag. Data Insights*, 2021.

[4] J. Wang et al., "CA-GAN: Class-condition attention GAN for underwater image enhancement," in *IEEE Access,* vol. 8, pp. 130719-130728, 2020, doi: https://doi.org/10.1109/ACCESS.2020.3003351.

[5] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent progress on generative adversarial networks (GANs): A survey," in *IEEE Access*, vol. 7, pp. 36322-36333, 2019, doi: https://doi.org/10.1109/ACCESS.2019.2905015.

[6] T. Katayama, T. Song, T. Shimamoto, and X. Jiang, "GAN-based color correction for underwater object detection," *OCEANS 2019 MTS/IEEE SEATTLE*, Seattle, WA, USA, 2019, pp. 1-4, doi: https://doi.org/10.23919/OCEANS40490.2019.8962561.

[7] M. Mirza, and S. Osindero, "Conditional generative adversarial nets," arXiv preprint arXiv:1411.1784, 2014.

[8] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967-5976, 2017.

[9] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of Wasserstein GANs," *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

[10] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," *International Conference on Learning Representations (ICLR),* 2018.