Research and Analysis of Generative Adversarial Networks in the Field of Computer Vision

Longxuan Li

School of Electronic Information, Shanghai Dianji University, 300 Shuihua Road Pudong New District Shanghai, China
Longxuanli808@gmail.com

Abstract. The integration of Generative Adversarial Networks (GANs) with various domains in computer vision has become one of the key topics in current research. Researchers have found that GANs have outperformed existing models in vertical applications, significantly enhancing and expanding model training performance and practical requirements across different fields. However, there is still a lack of a unified understanding regarding the systematic summarization and key points of GANs in various vertical domains. Therefore, this paper aims to provide a systematic review of the innovative research progress and breakthroughs made by GANs in the fields of medicine, biomolecules, agriculture, and remote sensing. By detailing the research processes and experimental results of GANs in these four application areas, the paper clarifies their core value and role in each field and establishes that future research in this direction is gradually moving toward goals of lightweight models, high generalization, and wide applicability. This review aims to give researchers a clearer understanding of the applications and impact of GANs across different domains, thereby laying a foundation for their use in even broader application areas in the future.

Keywords: GAN, Computer vision, Vertical application areas.

1 Introduction

Generative Adversarial Networks (GANs) first emerged in 2014, proposed by Goodfellow and others. As soon as it was launched, they quickly became a core of deep generative models in the field of computer vision due to the adversarial training mechanism between the generator and the discriminator [1]. GANs have demonstrated unique advantages in key tasks such as image generation, editing, restoration, and cross-domain translation and greatly enhancing the ability to model high-dimensional visual data. However, their development has always been constrained by core challenges such as training instability, mode collapse, reliance on high-quality data, and limited generalization ability. These bottlenecks have seriously hindered the transition of GANs from theoretical research to practical applications. Research on GANs holds not only significant theoretical value but also practical relevance. At the application level, GANs effectively address issues such as data scarcity, class imbalance, and high-resolution processing in computer vision, providing crucial technical support for

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_95

intelligent upgrading and cross-domain integration in specialized fields such as medical image analysis, agricultural monitoring, biomolecules, and remote sensing mapping. Existing research has systematically explored GANs in terms of technical optimization, core challenge resolution, and application expansion: G. Iglesias and others provided a comprehensive overview of the latest research progress on GANs, categorizing the major variants in overall framework optimization (e.g., DCGAN, StyleGAN) and loss function optimization (e.g., WGAN, LSGAN) [2]. They deeply analyzed common problems during training, such as mode collapse, gradient vanishing, and instability, as well as evaluation metrics such as MS and FID, systematically summarizing GAN applications in image synthesis, cross-domain translation, medical imaging, and agriculture, while also discussing comparisons with emerging architectures like diffusion models and Transformers, offering a comprehensive technical framework for the field; Zhizhong Huang and colleagues pioneered a new approach in medical CT denoising named DU-GAN [3]. Compared to traditional GAN-based denoising methods, they used a U-Net discriminator to perform low-dose CT denoising while maximizing CT image quality, providing pixel-level feedback for the denoising model. The method significantly mitigated stripe artifacts caused by photon starvation and provided confidence maps to assist radiologists in image review and diagnosis; Yan Zhang and colleagues focused on plant disease detection in agricultural image enhancement [4]. Addressing traditional deep learning algorithm shortcomings such as high hardware cost, slow inference speed, and weak generalization ability, they proposed a new Tranvolution monitoring network integrated with a GAN module. This model outperformed all comparable models in performance and experimental results. Furthermore, they packaged the model and built an agricultural intelligent robot based on it for application in real-world agricultural scenarios.Anvita Gupta and colleagues addressed the issue in computational biology where the high computational cost limits extensive exploration of protein conformational landscapes [5]. They proposed an innovative pipeline, MoDyGAN, which combines GANs with molecular dynamics (MD) simulations to accelerate the exploration of protein conformational landscapes, thereby gaining deeper insights into protein dynamics. Yingfei Xiong and others conducted systematic research on key technological breakthroughs of GANs in the field of remote sensing, such as cross-location and cross-sensor remote sensing image super-resolution optimization and noise label suppression techniques in remote sensing classification, as well as the application optimization of super-resolution technology in the extraction and classification of remote sensing features [6]. This further improved the technical application system of GANs in remote sensing image processing. Based on the above research foundation, this paper systematically reviews the research progress of GANs in various vertical domains of computer vision. Following a predefined research framework, it explores three dimensions: fundamental theoretical optimization, core challenge resolution, and practical application. It delves into GANs in different fields, analyzes their deep applications in vertical domains such as medicine, agriculture, biology, and remote sensing, maps out the overall trajectory of technological and methodological developments, clarifies current research bottlenecks, and ultimately provides a comprehensive and systematic reference for further technological innovation and engineering application of GANs in the field of computer vision.

2 Mainstream Technologies and Applications of GANs in Different Fields

2.1 Medical Imaging Direction

The medical imaging direction mainly focuses on two core clinical needs: low-dose CT denoising and high-fidelity generation of fundus images. Zhizhong Huang and Qingshan Hou proposed the DU-GAN and FundusGAN optimization models respectively [3][7]. In this field, GAN technology mainly addresses challenges in traditional medical imaging, such as noise interference, data scarcity, and low accuracy in pathological feature restoration. Through adversarial training mechanisms, it achieves image quality improvement and data augmentation, providing precise support for clinical diagnosis.

DU-GAN. In medical CI imaging, low-dose CT can cause issues such as streak artifacts and photon starvation noise, which severely affect the recognition of key anatomical structures. DU-GAN innovatively introduces a dual-domain U-Net discriminator, achieving improvements in both denoising performance and clinical interpretability.

DU-GAN is based on the core architecture of 'generator denoising and dual-domain discriminator optimization.' The generator module continuously learns the data distribution of the samples to generate new samples that closely resemble the original real samples. RED-CNN is employed as the core, which maps low-dose CT (LDCT) images to their corresponding normal-dose CT images by removing noise from LDCT images, thereby approximating the pixel distribution of normal-dose CT images. The dual-domain U-Net discriminator is introduced in both the image domain and the gradient domain. In the image domain, a low-dose CT denoising framework is designed to process LDCT denoising, improving the visual quality of denoised low-dose CT images， as shown in Fig. 1. In the gradient domain, a new gradient branch is executed using the Sobel operator to mitigate streak artifacts generated during the process.

The introduction of DU-GAN enhances the clarity of low-dose denoised images, while the dual-domain discriminator can capture both global and local information at the same time, expanding the scope of information reception. However, DU-GAN's dual-domain U-Net has a large number of parameters and requires a high-memory GPU for long training periods.

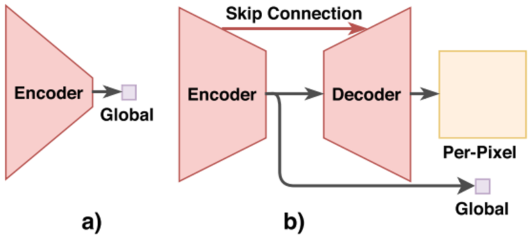

Fig. 1. (a) The traditional discriminator and (b) the dual-domain U-Net discriminator. The dual-domain U-Net discriminator can capture local information while also obtaining global information [3].

FundusGAN. Fundus images are a key basis for diagnosing eye diseases such as diabetic retinopathy and glaucoma. However, in clinical practice, there are issues such as the scarcity of pathological samples and high annotation costs, leading to insufficient training data for ophthalmic AI models. FundusGAN, an improvement based on StyleGAN, synthesizes fundus images with various precise anatomical and pathological features through multi-scale feature extraction and latent vector mapping, which is used to augment training datasets and improve the generalization ability of diagnostic models.

FundusGAN adapts to the structural characteristics of fundus images. The hierarchical encoder uses ResNet as the backbone network and is divided into three levels: low-level feature maps(F_{low}), mid-level feature maps(F_{mid}), and high-level feature maps(F_{high}), which explore information from multiple angles to enhance the ability to express hierarchical features. The improved StyleGAN generator removes redundant upsampling layers and uses mid-level vascular features as the basic input, thereby balancing global structural integrity with local details.

FundusGAN can clearly and comprehensively display subtle lesions in the anatomical process of the eye, greatly improving the precision and success rate of ophthalmic surgeries. At the same time, its pathological features can help doctors focus on high-risk areas and avoid making incorrect judgments. However, FundusGAN's feature pyramid and improved StyleGAN architecture make training far more complex than the basic model, resulting in high computational costs.

2.2 Biological Protein Direction

The biological protein direction mainly focuses on how to efficiently generate new proteins and label protein characteristics. Jingbo Liang, Anvita Gupta, and Musadaq Mansoor proposed three optimized models, MoDyGAN, FBGAN, and GOGAN [5, 8, 9]. These models have played a certain pioneering and leading role in this field.

MoDyGAN. Currently, there is relatively little research on protein conformations in computational biology. MoDyGAN combines GAN with molecular dynamics, which could efficiently generate new protein conformations in the future, fill the gaps in this field, and provide support for drug development and bioengineering.

The 3D-to-2D feature transformation in this model combines the adjacency matrix with node features and adjusts the size through zero-padding to reach a standard dimension. The ProGAN conformation generator can accept input vectors and output paired feature matrices, and after training, this model can distinguish between real and virtual matrices. The dual-discriminator refinement consists of one generator and two discriminators, with the global discriminator ensuring overall structural consistency and the local discriminator focusing on optimizing local details.

FBGAN. Synthetic biology is currently in a stage of continuous development, and GANs can help generate DNA sequences, proteins, and other macromolecules. FBGAN can generate protein-coding sequences and enrich them, replacing manual operations and overcoming the limitations of traditional GANs that lack specific features.

FBGAN applies two core technologies. The first is the WGAN generation module. WGAN is a new model produced by five residual layers after replacement and updating, with an overall architecture that includes a gradient penalty mechanism, making it more stable during training than the conventional GAN formula. The second is the AMP external function analyzer, which serves as a classifier: the input is a gene sequence, and the output is the probability that the gene encodes an antimicrobial peptide. The AMP classifier is trained based on 2,600 proven antimicrobial peptides from the APD3 database [10].

The proposal of FBGAN precisely addresses the current pain points and shortcomings in the field of biology, filling the gap in the area of biological proteins and aligning with the current exploration needs in biology.

GOGAN. Currently, the vast majority of proteins cannot be accurately annotated due to their characteristics being indeterminate, and using deep learning algorithms to characterize proteins through amino acid sequences is too costly and time-consuming. To address these two issues, a new model called GOGAN has been proposed. This model does not require manual extraction of protein features but instead categorizes protein functions into three GO categories: biological processes, molecular functions, and cellular components, automatically extracting features from a large number of unlabelled proteins,as shown in Fig. 2. GOGAN can effectively reduce the time and cost of experimental research while also expanding the scale of executable tasks, providing effective support for image synthesis and text generation problems.

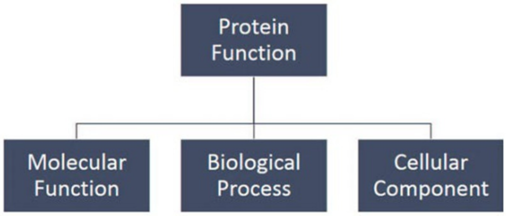

Fig. 2. GO Classification [9].

2.3 Agriculture Direction

The agriculture direction mainly focuses on two core issues: plant disease detection and agricultural pest detection based on computer vision. Yan Zhang and Christophe Karam respectively proposed the Tranvolution architecture and the DCGAN-optimized model [4][11]. In this field, GAN technology primarily addresses the pain points of traditional agricultural image analysis, such as data scarcity, poor environmental adaptability, and weak model generalization [2]. By synthesizing data to expand datasets or enhancing feature extraction, it improves the accuracy of object detection and classification, providing technical support for precision agriculture [12].

Tranvolution Architecture. Currently, plant disease detection is limited by the fact that sample labeling relies on professional agricultural knowledge, and field data collection is constrained by seasonal and regional factors. This results in small and unevenly distributed labeled datasets, directly restricting the generalization capability of deep learning models [13]. Meanwhile, the increasing maturity of computer vision provides more options for plant disease detection. Tranvolution is an architecture proposed by combining a GAN-equipped Tranvolution detection network with CNN. It can improve the accuracy of plant leaf detection while being integrated into a parallel-designed intelligent agricultural robot, merging the model with practical applications.

Tranvolution adds two generative network models in the network: a pre-generated GAN model is placed in front of the backbone network to augment leaf images using WGAN, while a post-positioned GAN model is combined with an attention module to enhance robustness to environmental disturbances; the transformer architecture avoids overly long training times by leveraging its ability to extract global features, as shown in Fig. 3.

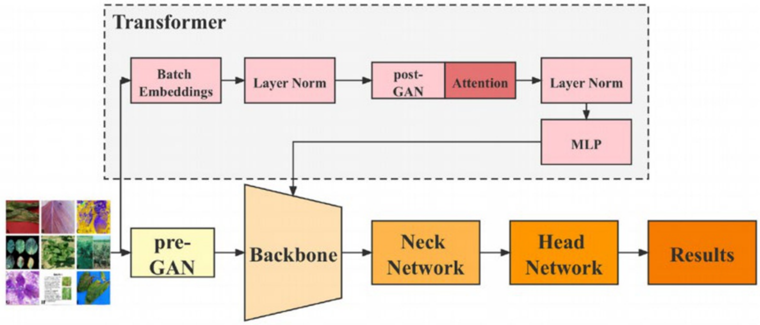

Fig. 3. Tranvolution detection network structure with GAN module [4].

CPB+GAN Pipeline. Agriculture has always been a core pillar in advancing human progress. With the continuous growth of the global population, agriculture is facing increasing pressure, making it more and more important to detect crop pests and diseases and ensure crop yield quality. The proposed CPB GAN pipeline reduces data collection and effectively improves the accuracy of detecting pests on plant leaves, thereby enhancing the overall performance of pest and disease detection, as shown in Fig. 4.

The technical process of the CPB+GAN pipeline is mainly divided into two parts. The DCGAN-based generator consists of three upsampling blocks, each of which contains a sampling layer, a convolutional layer, a batch normalization layer, and a final convolutional layer to output a 32×32 RGB image for enhancing pest masks. The object detection is based on the YOLOv3 model and its lightweight version YOLOv3-tiny[14], which can perform real-time detection at high frame rates. At the same time, this version is still usable on low-resource devices.

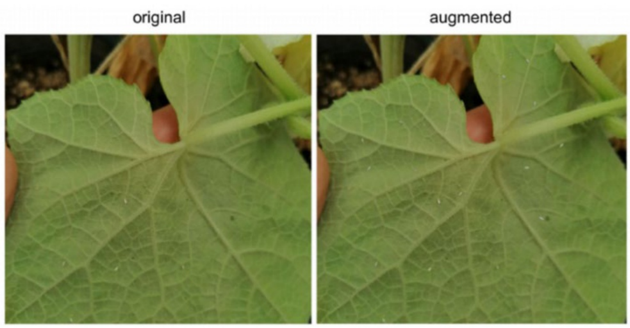

Fig. 4. CPB+GAN Output Comparison [11].

2.4 Remote Sensing Direction

The remote sensing direction mainly focuses on two core issues: super-resolution of remote sensing images and semi-supervised land cover classification of remote sensing images. Yingfei Xiong and others, as well as Taehong Kwak and others, respectively proposed two optimized models, ISRGAN and the CycleGAN EfficientNet framework [6][15]. These models integrate previous technical concepts, significantly improving generalization performance and test accuracy, providing technical support for practical application areas such as urban planning and environmental management.

ISRGAN. The accuracy of land cover and land use spatial changes is a key factor in assessing the ecological environment, thus it is necessary to use computer vision to generate high spatial resolution images. In remote sensing image processing tasks, spectral differences across sensors, texture variations across regional terrains, and the labor costs of annotating high-resolution images limit the generalization ability of models [16]. The introduction of ISRGAN significantly improves the generalization ability across locations and sensors based on many existing models and makes model training more stable.

The original SRGAN is prone to gradient vanishing and mode collapse during training, while the improved super-resolution generative adversarial network modifies the loss function, replacing the traditional KL divergence and JS divergence with the Wasserstein distance, alleviating the issues of gradient vanishing and mode collapse.

CycleGAN EfficientNet Framework. Classifying high-resolution images is a fundamental task in the field of remote sensing and widely covers various application scenarios. Traditional image classification, however, relies on manual work, making the training process labor-intensive and costly, and resulting in relatively poor image performance. The proposed CycleGAN EfficientNet framework aims to establish a semi-supervised learning method to improve the classification performance of remote sensing images, enhancing classification accuracy without increasing annotation costs [17].

Semi-supervised image classification trains unlabeled datasets indirectly by using a cycle consistency loss. During the cycle phase, two generators and two discriminators are used to transmit information from the input data, while in the supervised learning phase, only labeled images and their corresponding ground truth labels are used to reduce the supervised loss. The proposed network architecture introduces a UEnet structure in the generator, replaces the generator of the semi-supervised CycleGAN with EfficientNet-B1 to significantly reduce model complexity, and incorporates a PatchGAN discriminator to judge real/fake patches in the input images.

3 Experiment

3.1 Dataset

Table 1 and Table 2 mainly lists the datasets and processing methods of the four most typical models in the four field directions mentioned above, aiming to provide data support for this paper and also to offer data references for related researchers.

Table 1. Sources and Quantities of Relevant Datasets.

Application Direction	Model Name	Dataset Source	Dataset Quantity
Agricultural Direction	Tranvolution	Indian Institute of Technology PlantDoc Dataset	Total 2567 images; 2328 images in training set, 239 images in test set
Remote Sensing Direction	ISRGAN	1. GF-1 satellite data (China Land Observation Satellite Data Service Platform); 2. Landsat 8 OLI data (US Geological Survey Earth Explorer)	Training set: 3200 256×256 GF-1 images;Test set: Xinjiang GF-1 data, Xinjiang Landsat 8 data
Medical Direction	DU-GAN	1. Simulated dataset: ACR - Mayo Clinic LDCT Grand Challenge Project; 2. Real-world dataset: Piglet CT scans	1. Simulated dataset: 20 training/20 test patients, 300,000/64,000 image patches each; 2. Real-world dataset: 850 images, 60k image patches for training, 12k image patches for testing
Biological Direction	MoDyGAN	Generated by Molecular Dynamics (MD) simulation	1. 30,000+ conformations per system; 2. Generator training set: 10,000; Refiner training set: 10,000 pairs of noisy/clean conformations; 3. Random Forest dataset: 20,000

Table 2. Features of Related Datasets and Data Processing Methods.

Model Name	Dataset Features	Data Processing Methods
Tranvolution	Contains 13 plant species and 27 categories (17 disease types + 10 healthy states); complex image sources (field scenes, solid-color backgrounds, screenshots, etc.); resolution 416×416	1.Basic augmentation: image flipping, translation, scaling, HSV color channel transformation; 2.Advanced augmentation: Mixup, CutOut, CutMix, Mosaic; 3. Erosion + dilation

Model Name	Dataset Features	Data Processing Methods
		operations to remove leaf detail interference
ISRGAN	Cross-location (Guangdong, Xinjiang) and cross-sensor; RGB three bands; GF-1 resolution 8m, Landsat 8 resolution 30m	1. Crop images to 256×256 size; 2. Three downsampling steps to generate 64×64 low-resolution images; 3. Data normalization processing
DU-GAN	1.Simulated dataset: 25% abdominal and 10% chest LDCT data; 2.Real-world dataset: contains 5%-100% dose gradients, resolution 512×512	1. Extract 64×64 image patches, window width [-300, 300]; 2. Linear normalization to [0,1]; 3. Exclude image patches with a large proportion of air
MoDyGAN	Contains 3 rigid proteins (1POA, 2WJ7, 1BMR) and 1 flexible polypeptide (Ala); Ala has 20 conformational states	1. Convert to n×n×3 paired feature matrices, zero-padding normalization; 2. Add [-0.5, 0.5]Å random displacement to generate noisy conformations; 3. Split training/test sets, shuffle to avoid overlap

3.2 Evaluation Metrics

Table 3 lists the evaluation metrics used for training the four models Tranvolution, DU-GAN, MoDyGAN, and ISRGAN, aiming to help peers clearly understand the assessment criteria.

Table 3. Relevant Evaluation Indicators and Specific Values.

Application Direction	Model Name	Evaluation Metrics	Specific Values
Agricultural Direction	Tranvolution	Precision	51.7%
		Recall	48.1%
		mAP	50.3%
		FPS	37
Medical Direction	DU-GAN	PSNR	Mayo-10%:22.3075; Mayo-25%:34.6186;

Application Direction	Model Name	Evaluation Metrics	Specific Values
			Piglet-5%:29.8598
		SSIM	Mayo-10%:0.7489; Mayo-25%:0.9196; Piglet-5%:0.9345
		RMSE	Mayo-10%:0.0802; Mayo-25%:0.0196; Piglet-5%:0.0325
Biological Direction	MoDyGAN	RMSD	1POA:2.00±0.26Å; 2WJ7:1.31±0.24Å; Ala_{10}:≤2Å (partial states)
		RF	1POA:74.05%; 2WJ7:93.81%; 1BMR:33.35%
		EMD	0.02-0.05 (Ala_{10} helical conformations)
		Watson's U^2 Statistic	$U^2<0.2$ and $p>0.05$
Remote Sensing Direction	ISRGAN	PSNR	Cross-location:35.816; Cross-sensor:38.092
		SSIM	0.988
		Impervious Surface Extraction Accuracy Improvement	15%

3.3 Analysis of Experimental Results

Medical Imaging. DU-GAN can provide pixel-level feedback to the denoising network and focus on global structures, while enhancing edge information and reducing stripe artifacts caused by photon starvation. It can also provide confidence maps to help radiologists identify the cause of abnormalities. FundusGAN effectively captures and reconstructs anatomical structures at different scales, consistently leading in SSIM, FID, and KID metrics.

Direction of Biomolecules. MoDyGAN reverses protein conformations from 3D to 2D, providing diversity for using different GANs while significantly improving the quality of generated protein conformations. The type of analyzer used by FBGAN is highly robust, enabling the generator to produce more genes that may possess certain characteristics. The feedback loop mechanism used in this model can integrate smoothly with GANs, making it simpler and easier to implement. GOGAN shows significant improvements in the F1 Score and accuracy of human proteins. This model can obtain important information from protein sequences by training only on a limited number of labeled data.

Agriculture Direction. The Tranvolution architecture adds a preliminary GAN module in front of the backbone network to address the common issue of small sample datasets, and to verify the performance of different combinations of generative models. Experiments found that the model achieved the best performance with the WGAN + SAGAN combination. The CPB + GAN pipeline uses a lightweight model to implement object detection, with development tools that are simple and easy to use while effectively improving generalization capability.

Remote Sensing Direction. Experiments show that the results obtained by ISRGAN are superior to other super-resolution methods. At the same time, experiments and evaluations were conducted on the Guangdong GF1 and Xinjiang GF2 datasets to verify the cross-location generalization ability of the model. The CycleGAN+EfficientNet framework achieved the highest accuracy compared to other baseline methods. By analyzing the impact of labeled and unlabeled data, the applicability of the semi-supervised approach was confirmed.

4 Conclusion

This paper summarizes the recent performance of Generative Adversarial Networks (GANs) in several vertical domains of computer vision, including innovative GAN models and architectures proposed in medicine, biology, agriculture, and remote sensing. It clarifies their core value and role in each field and also systematically compiles datasets from various research teams in each subfield, while summarizing and explaining the evaluation metrics encountered during the process. Furthermore, the analysis of vertical domains in this paper still has certain limitations, particularly in terms of the breadth of the search within vertical application areas.

Considering the current state of research and industry demands, in the future, the application of GANs in real-world scenarios will focus more on aspects such as lightweight models, generalizability, low cost, and adaptability to application scenarios. GANs will further advance iterations and upgrades in computer vision across more emerging vertical domains, helping various industries achieve maximum benefits.

References

1. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y.: Generative adversarial nets. Advances in neural information processing systems, 27. (2014)
2. Iglesias, G., Talavera, E., & Díaz-Álvarez, A.: A survey on GANs for computer vision: Recent research, analysis and taxonomy. Computer Science Review, 48, 100553. (2023)

3. Huang, Z., Zhang, J., Zhang, Y., & Shan, H.: DU-GAN: Generative adversarial networks with dual-domain U-Net-based discriminators for low-dose CT denoising. IEEE Transactions on Instrumentation and Measurement, 71, 1-12. LNCS Homepage, (2023) http://www.springer.com/lncs, last accessed 2023/10/25
4. Zhang, Y., Wa, S., Zhang, L., & Lv, C.: Automatic plant disease detection based on tranvolution detection network with GAN modules using leaf images. Frontiers in Plant Science, 13, 875693. (2022)
5. Gupta, A., & Zou, J.: Feedback GAN (FBGAN) for DNA: a novel feedback-loop architecture for optimizing protein functions. arXiv preprint arXiv:1804.01694. (2018)
6. Xiong, Y., Guo, S., Chen, J., Deng, X., Sun, L., Zheng, X., & Xu, W.: Improved SRGAN for remote sensing image super-resolution across locations and sensors. Remote Sensing, 12(8), 1263. (2020)
7. Hou, Q., Wang, M., Cao, P., Ke, Z., Liu, X., Fu, H., & Zaiane, O. R.: Fundusgan: A hierarchical feature-aware generative framework for high-fidelity fundus image generation. arXiv preprint arXiv:2503.17831. (2025)
8. Liang, J., & Jacobson, B.: MoDyGAN: Combining Molecular Dynamics With GANs to Investigate Protein Conformational Space. arXiv preprint arXiv:2507.13950. (2025)
9. Mansoor, M., Nauman, M., Ur Rehman, H., & Benso, A.: Gene Ontology GAN (GOGAN): a novel architecture for protein function prediction. Soft Computing, 26(16), 7653-7667. (2022)
10. Wang, G., & Li, X.: The antimicrobial peptide database as a tool for research and education. 44, pp. D1087-D1093. (2016) DOI: https://doi. org/10.1093/nar/gkv1278. PMID: https://www. ncbi. nlm. nih. gov/pubmed/26602694.
11. Karam, C., Awad, M., Abou Jawdah, Y., Ezzeddine, N., & Fardoun, A.: GAN-based semi-automated augmentation online tool for agricultural pest detection: A case study on whiteflies. Frontiers in Plant Science, 13, 813050. (2022)
12. Lu, Y., Chen, D., Olaniyi, E., & Huang, Y.: Generative adversarial networks (GANs) for image augmentation in agriculture: A systematic review. Computers and Electronics in Agriculture, 200, 107208. (2022)
13. Sampath, V., Maurtua, I., Aguilar Martin, J. J., & Gutierrez, A.: A survey on generative adversarial networks for imbalance problems in computer vision tasks. Journal of big Data, 8(1), 27. (2021)
14. Farhadi, A., & Redmon, J.: Yolov3: An incremental improvement. In Computer vision and pattern recognition (Vol. 1804, pp. 1-6). Berlin/Heidelberg, Germany: Springer. (2018)
15. Kwak, T., & Kim, Y.: Semi-supervised land cover classification of remote sensing imagery using CycleGAN and EfficientNet. KSCE Journal of Civil Engineering, 27(4), 1760-1773. (2023)
16. Wang, Z., She, Q., & Ward, T. E.: Generative adversarial networks in computer vision: A survey and taxonomy. ACM Computing Surveys (CSUR), 54(2), 1-38. (2021)
17. Saxena, D., & Cao, J.: Generative adversarial networks (GANs) challenges, solutions, and future directions. ACM Computing Surveys (CSUR), 54(3), 1-42. (2021)