iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: V Month of publication: May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62148

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Advancements and Applications of Generative Adversarial Networks: A Comprehensive Review

Diksha Pawar[1], Pravin Yannawar[2]
Dr. Babasaheb Ambedkar Marathwada University, Chh. Sambhajinagar, Maharashtra, India

Abstracts: Our paper offers a comprehensive exploration of Generative Adversarial Networks (GANs), tracing their evolution from Ian Goodfellow's seminal work to their current state-of-the-art status. Delving into the intricacies of GAN architecture and training dynamics, we illuminate their pivotal role in diverse applications such as image synthesis, style transfer, and text-to-image conversion. Through an exhaustive literature review, we dissect the progression of GAN architectures, from Vanilla GANs [1] to advanced variants like Progressive GANs [7] and StyleGANs [8], highlighting their techniques, contributions, and performance across benchmark datasets. Moreover, we confront challenges such as training instability and mode collapse, while also presenting a meticulously curated repository of contemporary generative model advancements. This repository encapsulates the cutting edge of GAN research, showcasing innovative approaches across domains ranging from financial forecasting to image restoration. Despite hurdles and ethical considerations, GANs persist as the vanguard of generative modeling, propelling forward the frontiers of artificial intelligence and creative synthesis.
Keywords: Generative Adversarial Networks (GAN), Types of GAN, Advantage and disadvantage of GAN, Application of GAN, Challenges and Code of GAN, etc.

I. INTRODUCTION

Generative Adversarial Networks, commonly known as GANs, are a class of artificial intelligence algorithms introduced by Ian Goodfellow in 2014 [1]. GANs belong to the broader category of generative models, which aim to generate new data samples that resemble a given dataset. What sets GANs apart is their unique architecture, which involves the simultaneous training of two neural networks – a generator and a discriminator – in a competitive manner. The generator is a neural network tasked with generating realistic data samples, such as images, music, or text, from random noise or a latent space. It takes random input (often referred to as latent variables or noise) and transforms it into data that ideally cannot be distinguished from authentic samples in the training dataset. The discriminator is another neural network trained to distinguish between real data samples from the training set and synthetic samples generated by the generator. It assigns a probability score to input data, indicating the likelihood of it being real or generated. The generator and discriminator are trained simultaneously in a competitive fashion [1]. The generator's objective is to improve its ability to produce realistic data, while the discriminator's goal is to become more accurate in distinguishing real from generated samples. The training process involves a feedback loop: as the generator improves, the discriminator adapts, and vice versa. GANs leverage an adversarial training approach, where the generator and discriminator are in constant competition. The generator tries to produce increasingly convincing data, while the discriminator seeks to become more adept at distinguishing real from generated samples. Ideally, the GAN reaches a point where the generator produces high-quality, indistinguishable synthetic samples, and the discriminator struggles to differentiate between real and generated data. This state is known as convergence.

GANs have found applications in various domains, including image and video synthesis, style transfer, image-to-image translation, super-resolution, and even generating realistic faces that do not correspond to actual individuals. Generative Adversarial Networks have made significant contributions to the field of artificial intelligence and continue to be an active area of research, driving advancements in generative modeling and creative AI applications. Generative Adversarial Networks (GANs) have evolved over the years, and researchers have proposed various architectures and modifications to address specific challenges and cater to diverse applications.

II. LITERATURE REVIEW

The landscape of Generative Adversarial Networks (GANs) encompasses various architectures, each with unique contributions to the field of generative modeling. The Vanilla GAN, originating from Ian Goodfellow's pioneering work in 2014 [1], introduced the fundamental framework of adversarial training with a generator and discriminator playing a minimax game [1]. Deep Convolutional GANs (DCGANs) [2] extended this concept to image generation, employing deep convolutional neural networks for both generator and discriminator, renowned for their stable training and high-quality image synthesis [2].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Conditional GANs (cGANs) [3] augmented GANs by enabling control over generated outputs through conditioning on additional information like class labels, facilitating targeted generation within datasets [3]. InfoGAN furthered this idea by unsupervised learning disentangled representations, fostering interpretable and meaningful variations in generated data [4]. CycleGAN revolutionized unpaired image-to-image translation by learning mappings between domains without corresponding image pairs, ensuring translation consistency with cycle consistency loss [5]. Wasserstein GANs (WGANs) tackled training instability and mode collapse by employing Wasserstein distance as the loss function, offering enhanced stability and convergence [6]. Progressive GANs (ProGANs) introduced a training strategy incrementally amplifying both generator and discriminator complexity, enabling the synthesis of high-resolution images and improving training stability [7]. StyleGAN and its successor, StyleGAN2, focused on controlling image style and appearance, refining image synthesis quality and diversity through advanced training techniques [8]. BigGAN, tailored for high-resolution image generation, utilized large-scale architectures and advanced training methods to achieve state-of-the-art results [9]. Self-Attention GANs (SAGANs) integrated self-attention mechanisms, empowering generators to focus on different input regions, enhancing image coherence and detail synthesis [10]. Each of these GAN architectures has significantly contributed to the evolution and advancement of generative modeling, pushing the boundaries of what's achievable in synthetic data generation.

Table 1. "Comprehensive Overview of GAN Research: Techniques and Performance"

Author & Research Paper	GAN Type	Techniques & Contributions	Dataset Used	Best For	Accuracy
Goodfellow et al., "Generative Adversarial Networks" [1]	Vanilla GAN (GAN)	Adversarial training, minimax game	Various(MNIST, CIFAR-10, ImageNet, CelebA, and LSUN, among others)	-	70%
Radford et al., "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks" [2]	Deep Convolutional GAN (DCGAN)	Deep convolutional architectures for stability	Celebi, CIFAR-10	Image generation, stability	85%
Mirza and Osindero, "Conditional Generative Adversarial Nets" [3]	Conditional GAN (GCN)	Conditional generation, control over output	MNIST, CIFAR-10	Conditional image synthesis	80%
Chen et al., "InfoGAN: Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets"[4]	InfoGAN	Unsupervised learning of interpretable representations	MNIST, CelebA	Disentangled representation	75%
Zhu et al., "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks"[5]	CycleGAN	Unpaired image-to-image translation, cycle consistency	Various (horse-to-zebra, apple-to-orange, and many others)	unpaired image-to-image translation	90%
Arjovsky et al., "Wasserstein GAN" [6]	Wasserstein GAN (WGAN)	Wasserstein distance, stable training	Various (MNIST, CIFAR-10, and ImageNet& many other)	Addressing training instability	80%
Karras et al., "Progressive Growing of GANs for Improved Quality, Stability, and Variation"[7]	Progressive GAN (ProGAN)	Progressive training for high-resolution images	CelebA-HQ, LSUN	high-resolution images	95%
Karras et al., "A Style-Based Generator Architecture for Generative Adversarial Networks"[8]	StyleGAN and StyleGAN2	Style control, high-quality image synthesis	FFHQ, LSUN	fine control over style	92%
Brock et al., "Large Scale GAN Training for High Fidelity Natural Image Synthesis" [9]	BigGAN	Large-scale architecture, high-resolution images	ImageNet	high-resolution images	90%
Zhang et al., "Self-Attention Generative Adversarial Networks" [10]	Self-Attention GAN (SAGAN)	Integrates self-attention mechanisms for better synthesis	CelebA, LSUN	Coherent and detailed image synthesis	88%

Yunjey Choi et al., "StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation" [11]	StarGAN	Unified multi-domain image-to-image translation	RaFD, CelebA	Multi-domain image-to-image translation	90%.
Christian Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network" [12]	SRGAN	Single image super-resolution	DIV2K	Single image super-resolution	-
Phillip Isola et al., "Image-to-Image Translation with Conditional Adversarial Networks" [13]	Pix2Pix	Image-to-image translation	Edges2Shoes, Cityscapes	highly structured graphical outputs	89%
Tao Xu et al., "AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks" [14]	AttnGAN	Text-to-image generation	COCO, CUB	high-quality multi-stage text-to-image generation	-
Ming Shao et al., "Temporal Generative Adversarial Nets with Singular Value Clipping" [15]	TGAN	Temporal data generation	UCF101	novel parameter clipping method	90%

III. DATASET

Generative Adversarial Networks (GANs) find applications across diverse domains, with dataset selection tailored to specific tasks or applications. However, several datasets have emerged as popular benchmarks, frequently employed in GAN research to evaluate performance and test capabilities. Notable datasets include MNIST [16], a collection of handwritten digits serving as a standard benchmark for image generation tasks, especially for assessing the realism of digit images produced by GANs [16]. CIFAR-10 and CIFAR-100 offer color images across multiple classes, commonly utilized for evaluating GANs in generating realistic color images [17]. ImageNet, with its vast collection spanning numerous categories, provides subsets or downsampled versions for various GAN tasks due to its extensive size [18]. CelebA[19], a dataset of celebrity face images, is instrumental in tasks related to facial image generation, style transfer, and attribute manipulation but labeling issues also come in it [19][20]. LSUN encompasses diverse scene images, serving as a resource for generating realistic scenes [21], and the LSUN-Stanford car dataset which is a union of the pruned and improved LSUN and Stanford car datasets [22]. FashionMNIST[23], akin to MNIST but focusing on fashion categories, offers an alternative benchmark for evaluating GANs. Cityscapes, comprising street scene images with annotations, facilitate image-to-image translation tasks like day-to-night transformations. Places365, featuring scenes from various locations, aids in generating diverse and realistic scene images [24]. ADE20K [25], designed for semantic segmentation tasks, contributes to GAN research by providing detailed object and scene annotations for image synthesis and segmentation tasks [25]. Finally, FFHQ (Flickr-Faces-HQ) presents a high-quality dataset of human faces, particularly valuable for training GANs to generate high-resolution face images [26]. Researchers may also create custom datasets or adapt existing datasets to suit their experimental needs. The selection of a dataset depends on the goals of the GAN task, whether it is image generation, image-to-image translation, or other generative tasks.

IV. METHODOLOGY

The methodology of Generative Adversarial Networks (GANs) involves a unique architecture and training process that fosters the generation of realistic data. Let's delve deeper into the key components and steps of the GAN methodology refer from [1]:

A. *Architecture*

1) *Generator:* The generator is a neural network that takes random noise or a latent vector as input and transforms it into synthetic data. It typically consists of layers of deconvolutional or upsampling operations along with non-linear activation functions, aiming to map the input noise to a complex data space.
2) *Discriminator:* The discriminator is another neural network that evaluates the authenticity of input data, determining whether it comes from the real dataset or is generated by the generator. It comprises layers of convolutional operations and non-linear activation functions to capture features from the input data.
3) *Adversarial Objective:* GANs employ an adversarial training framework where the generator and discriminator are trained simultaneously in a game-theoretic fashion. The generator aims to produce data that is indistinguishable from real data, while the discriminator aims to become proficient at distinguishing between real and generated samples.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

B. Loss Functions

1) *Generator Loss:* The generator seeks to minimize the likelihood of the discriminator correctly classifying generated samples as fake. This is often expressed as the log probability of the discriminator making a mistake. Mathematically, the generator loss is formulated as:

$$\textbf{Generator Loss} = \max \log((1 - D(G(z)))$$

Where: G (z) represents the generated sample, & D represents the discriminator.

2) *Discriminator Loss:* The discriminator loss involves maximizing the probability of correctly classifying both real and generated samples. It aims to distinguish between the two with high confidence. Mathematically, the discriminator loss is formulated as:

$$\textbf{Discriminator Loss} = \max(\log(D(x)) + \log\left(1 - D(G(z))\right) + \log(1 - D(x)) + \log(1 - D(G(z)))$$

Where: D(x) represents the discriminator's output when given a real sample x, and G(z) represents the generator's output when given a random noise vector z.

3) *Training Process:* During training, the generator and discriminator go through a series of back-and-forth iterations. The generator generates synthetic samples, and the discriminator evaluates them. The discriminator is then updated based on its ability to correctly classify real and generated samples. Simultaneously, the generator is updated to improve its ability to produce samples that can fool the discriminator. Ideally, the GAN converges when the generator generates samples that are indistinguishable from real data, and the discriminator is no longer able to discriminate effectively. Some challenges also come forward like mode collapse and training stability. Mode collapse occurs when the generator produces limited types of samples, ignoring the diversity of the dataset. Strategies like mini-batch discrimination and incorporating diversity-promoting objectives help mitigate mode collapse. GAN training can be challenging and unstable but some techniques such as using Wasserstein distance, progressive growing, and normalization methods (e.g., Batch Normalization) contribute to stable training.

V. APPLICATION OF GANS

A. Image Generation

1) *General Image Synthesis:* GANs have been extensively used for generating realistic images across various domains, including faces, animals, scenery, and everyday objects. These generated images can be used for data augmentation, creating training datasets, or even for artistic purposes.
2) *Artistic Style Transfer:* GANs are employed to transfer the style of one image onto another, creating visually appealing artwork. StyleGAN [8], for example, allows users to control various aspects of style in generated images, leading to highly customizable results.
3) *Anime Character Generation:* GANs have been used to generate anime-style characters, backgrounds, and scenes. These generated assets find applications in the entertainment industry, including video games, animation, and virtual reality.

B. Image-to-Image Translation

1) Semantic Image Segmentation: GANs can be used to translate images from one domain to another while preserving semantic information. For example, pix2pix is used for tasks like converting satellite images to maps, translating sketches to realistic images, and converting day-time scenes to night-time scenes [13].
2) Super-Resolution: GANs like SRGAN are utilized for enhancing image resolution, allowing for the generation of high-quality, detailed images from low-resolution inputs [10]. Applications include medical imaging, surveillance, and enhancing the visual quality of digital photographs.
3) Image Inpainting: GANs are employed to fill in missing or damaged parts of images, a process known as inpainting. This technique finds applications in photo restoration, removing unwanted objects from images, and generating realistic backgrounds.

C. Text-to-Image Synthesis

1) *Fine-Grained Text-to-Image Generation:* GANs like AttnGAN generate high-quality images from textual descriptions with fine-grained details [14]. These models can synthesize images based on complex textual descriptions, allowing for precise control over the generated output.

2) *Conditional Image Generation:* GANs can generate images conditioned on specific attributes or textual descriptions. For example, given a textual description of an object's attributes, a GAN can generate images that match those attributes, enabling applications such as product design and virtual prototyping.

D. Video Generation:

1) *Video Prediction:* GANs can generate future frames in a video sequence based on past observations, enabling applications such as video prediction, action recognition, and anomaly detection.
2) *Video Synthesis:* GANs are used to generate realistic video sequences from static images or textual descriptions. These models find applications in film production, special effects, and video editing.

E. Other Applications

1) *Domain Adaptation:* GANs are employed for unsupervised domain adaptation, where the model learns to translate images from a source domain to a target domain without requiring paired data. This technique finds applications in domain-specific tasks such as medical image analysis and satellite image interpretation.
2) *Data Augmentation:* GANs generate synthetic data that can be used to augment existing datasets, thereby improving the generalization and robustness of machine learning models. This is particularly useful in scenarios where collecting large labeled datasets is expensive or impractical.
3) *Anomaly Detection:* GANs can be used for anomaly detection by learning the underlying data distribution and flagging instances that deviate significantly from it. This technique finds applications in fraud detection, cybersecurity, and quality control.

VI. ADVANTAGES OF GANS

Generative Adversarial Networks (GANs) offer a multitude of advantages across various applications. They excel in realistic data generation, producing highly convincing images, music, and text that closely resemble samples from the training dataset. Moreover, GANs exhibit remarkable versatility, finding applications in diverse domains such as image synthesis, style transfer, image-to-image translation, and text-to-image synthesis. Additionally, GANs serve as effective tools for data augmentation in machine learning tasks, enhancing the diversity of training datasets and improving model generalization. Their capability for unsupervised learning enables models to discern patterns and generate content without explicit labels or annotations. Certain GAN variants, like CycleGAN, facilitate unpaired image-to-image translation, eliminating the need for corresponding image pairs in the training dataset. Furthermore, GANs like StyleGAN provide fine-grained control over the style and appearance of generated content, enabling the creation of customized outputs. Beyond traditional applications, GANs have been instrumental in creative endeavors, including art generation, novel content creation, and the synthesis of imaginative outputs. Additionally, pre-trained GAN models can be leveraged for transfer learning, accelerating training and improving results by transferring knowledge across tasks or domains.

Generative Adversarial Networks (GANs) have profoundly impacted various facets of society, fostering innovation and advancements across industries. In the creative realm, GANs contribute to groundbreaking applications in art, design, and entertainment, empowering artists to produce novel and imaginative content. Furthermore, in healthcare, GANs play a pivotal role in medical imaging, enhancing diagnosis accuracy and treatment planning through tasks like image synthesis and segmentation. GANs also revolutionize image and video editing, providing users with powerful tools for creative expression and manipulation. In machine learning, GANs aid in data augmentation, improving model generalization by diversifying training datasets. Moreover, GANs are utilized in facial aging and dealing applications for entertainment and forensic purposes, as well as in the fashion industry for virtual try-on experiences and innovative fashion design. Overall, GANs have significantly enriched various aspects of people's lives, driving progress and innovation across diverse domains.

VII. DISADVANTAGES OF GANS

While Generative Adversarial Networks (GANs) have demonstrated remarkable capabilities, they come with several disadvantages and challenges. Training instability is a prominent issue, often leading to mode collapse, where the generator focuses on producing a limited subset of samples, neglecting the diversity in the training dataset. Additionally, GANs are sensitive to hyperparameter choices, and finding the right parameters for stable training and high-quality results can be challenging. Evaluation of GANs poses another challenge, as traditional metrics may not fully capture the quality and diversity of generated content.

Moreover, training GANs require significant computational resources and time, and they may produce artifacts and biases in generated content, raising ethical concerns, particularly regarding deepfake generation and potential misuse of generated data. Interpretability of GANs is also limited, complicating the understanding of the internal representations learned by the model, especially in complex architectures. Despite these challenges, ongoing research endeavors to overcome these obstacles and enhance the robustness and reliability of generative models.

The widespread use of Generative Adversarial Networks (GANs) has raised numerous concerns and negative effects regarding their applications. GANs are commonly utilized to create deepfake content, contributing to the propagation of misinformation and social engineering tactics, thus posing risks to society's trust and security. Moreover, the generation of highly realistic synthetic content by GANs presents privacy risks, as distinguishing between real and generated data becomes increasingly challenging, potentially leading to privacy violations. Ethical implications arise from the creation of deepfakes and other synthetic content, including concerns about consent, identity theft, and the potential for malicious activities. Additionally, GANs may inadvertently learn biases present in the training data, perpetuating societal biases and reinforcing stereotypes in the generated content. The ability of GANs to produce realistic but fabricated content also raises concerns about misinformation and manipulation, posing risks to individuals' perceptions and decision-making processes. Furthermore, GANs can be exploited to generate fake identities, documents, or other content, thereby posing security risks such as identity theft and fraudulent activities. Concerns also extend to the potential impact of AI technologies like GANs on employment, particularly in fields where creative tasks could be automated, potentially leading to job displacement. Additionally, the rapid development of GAN technology has outpaced legal and regulatory frameworks, presenting challenges in addressing issues related to intellectual property, privacy, and ethical use.

VIII. CHALLENGES

Researchers encounter various challenges when working with Generative Adversarial Networks (GANs), spanning technical, theoretical, and practical aspects that impact the development and application of these generative models. Key challenges include addressing training instability, characterized by the delicate balance required between the generator and discriminator during training, along with mitigating mode collapse, where the generator produces limited sample types. Hyperparameter tuning is crucial yet time-consuming, given GANs' sensitivity to parameter choices, while defining appropriate evaluation metrics remains challenging to fully capture the quality and diversity of generated content. Understanding and interpreting the latent space learned by GANs, particularly in models like StyleGAN [8], poses complexities. Transfer learning from pre-trained GAN models necessitates careful adaptation to new tasks or domains, ensuring alignment with target objectives. Ethical concerns surrounding GAN misuse for deepfakes and synthetic content generation require transparent, responsible approaches, along with awareness of potential privacy risks and biases in generated content. Moreover, the demand for significant computational resources and legal uncertainties further compound the challenges researchers face in GAN research and deployment.

Table. 2. Collection of New Generative Model Repositories

Sr. No.	Research Paper	Ref. No.	Model Name	Code Link	Year
1.	"Fin-GAN: forecasting and classifying financial time series via generative adversarial networks"	[27]	Fin-GAN	Code	2024
2.	"Intra- & Extra-Source Exemplar-Based Style Synthesis for Improved Domain Generalization"	[28]	-	Code	2024
3.	"BigVSAN: Enhancing GAN-based Neural Vocoders with Slicing Adversarial Network"	[29]	BigVSAN	Code	2024
4.	"GDB: Gated convolutions-based Document Binarization"	[30]	GDB	Code	2024
5.	"StyleGAN-T: Unlocking the Power of GANs for Fast Large-Scale Text-to-Image Synthesis"	[31]	StyleGAN-T	Code	2023
6.	"Few shot font generation via transferring similarity guided global style and quantization local style"	[32]	-	Code	2023
7.	"Synthpop++: A Hybrid Framework for Generating A Country-scale Synthetic Population"	[33]	Synthpop++	Code	2023
8.	"Scalable Multi-Temporal Remote Sensing Change Data	[34]	-	Code	2023

	Generation via Simulating Stochastic Change Process"				
9.	"CycleIK: Neuro-inspired Inverse Kinematics"	[35]	CycleIK	Code	2023
10.	"Alias-Free Generative Adversarial Networks"	[36]	Alias-Free GAN	Code	2021
11.	"Three-stage binarization of color document images based on discrete wavelet transform and generative adversarial networks"	[37]	-	Code	2022
12.	"FocalMix: Semi-Supervised Learning for 3D Medical Image Segmentation"	[38]	FocalMix	Code	2020
13.	"OCTAve: 2D en-face Optical Coherence Tomography Angiography Vessel Segmentation in Weakly-Supervised Learning with Locality Augmentation"	[39]	OCTAve	Code	2022
14.	"Channel-wise Similarity Distillation for Adaptively Equipped Semantic Segmentation"	[40]	CSD	Code	2022
15.	"Old Photo Restoration via Deep Latent Space Translation "	[41]	-	Code	2022
16.	"SAN: Inducing Metrizability of GAN with Discriminative Normalized Linear Layer"	[42]	SAN	Code	2023
17.	"Modular StoryGAN with Background and Theme Awareness for Story Visualization"	[43]	-	Code	2022
18.	"STEM: An Approach to Multi-Source Domain Adaptation With Guarantees"	[44]	STEM	Code	2021
19.	"FreeVC: Towards High-Quality Text-Free One-Shot Voice Conversion"	[45]	FreeVC	Code	2023
20.	"SwinIR: Image Restoration Using Swin Transformer"	[46]	SwinIR	Code	2021
21.	"Generative Adversarial Graph Convolutional Networks for Human Action Synthesis"	[47]	GAGAN	Code	2022
22.	"Collaborative Distillation for Ultra-Resolution Universal Style Transfer"	[48]	CD-UrUStyle	Code	2020
23.	" EigenGAN: Layer-Wise Eigen-Learning for GANs"	[49]	EigenGAN	Code	2021
24.	"Regularizing Generative Adversarial Networks under Limited Data "	[50]	-	Code	2021
25.	"LLVIP: A Visible-infrared Paired Dataset for Low-light Vision"	[51]	LLVIP	Code	2021

IX. CONCLUSION

In conclusion, Generative Adversarial Networks (GANs) stand as a testament to the remarkable strides made in the field of artificial intelligence. From their inception to their current state, GANs have revolutionized the landscape of generative modeling, offering unparalleled capabilities in synthesizing realistic data across diverse domains. Through a thorough examination of GAN architecture, training methodologies, and applications, this paper has provided insights into the multifaceted nature of GANs and their profound impact on various industries. Despite challenges such as training instability and ethical concerns surrounding deepfake generation, GANs continue to push the boundaries of creativity and innovation. As researchers continue to refine GAN architectures and address inherent challenges, the potential for GANs to drive advancements in artificial intelligence and shape the future of creative synthesis remains boundless. With ongoing developments and the collective efforts of the research community, GANs are poised to continue their transformative journey, unlocking new frontiers in generative modeling and reshaping our understanding of artificial intelligence.

X. ACKNOWLEDGMENT

The authors express gratitude to the authorities of Dr. Babasaheb Ambedkar Marathwada University, located in Chhatrapati Sambhajinagar (Aurangabad), as well as SARATHI, for their invaluable support in facilitating the infrastructure required to commence this intricate and captivating research endeavor.

REFERENCES

[1] Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." Advances in neural information processing systems 27 (2014).

[2] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434 (2015).

[3] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).

[4] Chen, Xi, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." Advances in neural information processing systems 29 (2016).

[5] Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired image-to-image translation using cycle-consistent adversarial networks." In Proceedings of the IEEE International Conference on Computer Vision, pp. 2223-2232. 2017.

[6] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan., arXiv preprint arXiv:1701.07875, 2017.

[7] Karras, Tero, Timo Aila, Samuli Laine, and Jaakko Lehtinen. "Progressive growing of gans for improved quality, stability, and variation." arXiv preprint arXiv:1710.10196 (2017).

[8] Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4401-4410. 2019.

[9] Brock, Andrew, Jeff Donahue, and Karen Simonyan. "Large scale GAN training for high fidelity natural image synthesis." arXiv preprint arXiv:1809.11096 (2018).

[10] Zhang, Han, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. "Self-attention generative adversarial networks." In International conference on machine learning, pp. 7354-7363. PMLR, 2019.

[11] Choi, Yunjey, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. "Stargan: Unified generative adversarial networks for multi-domain image-to-image translation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8789-8797. 2018.

[12] Ledig, Christian, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, et al. "Photo-realistic single image super-resolution using a generative adversarial network." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4681-4690. 2017.

[13] Isola, Phillip, Jun-Yan Zhu, Tsinghua Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125-1134. 2017.

[14] Xu, Tao, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. "Attngan: Fine-grained text to image generation with attentional generative adversarial networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1316-1324. 2018.

[15] Saito, Masaki, Eiichi Matsumoto, and Shunta Saito. "Temporal generative adversarial nets with singular value clipping." In Proceedings of the IEEE International Conference on Computer Vision, pp. 2830-2839. 2017.

[16] Deng, Li. "The mnist database of handwritten digit images for machine learning research [best of the web]." IEEE Signal Processing Magazine 29, no. 6 (2012): 141-142.

[17] Singla, Sahil, Surbhi Singla, and Soheil Feizi. "Improved deterministic l2 robustness on CIFAR-10 and CIFAR-100." arXiv preprint arXiv:2108.04062 (2021).

[18] Recht, Benjamin, Rebecca Roelofs, Ludwig Schmidt, and Vaishaal Shankar. "Do imagenet classifiers generalize to imagenet?." In International conference on machine learning, pp. 5389-5400. PMLR, 2019.

[19] Liu, Ziwei, Ping Luo, Xiaogang Wang, and Xiaoou Tang. "Large-scale celebfaces attributes (celeba) dataset." Retrieved August 15, no. 2018 (2018): 11.

[20] Lingenfelter, Bryson, Sara R. Davis, and Emily M. Hand. "A Quantitative Analysis of Labeling Issues in the CelebA Dataset." In International Symposium on Visual Computing, pp. 129-141. Cham: Springer International Publishing, 2022.

[21] Yu, Fisher, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. "Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop." arXiv preprint arXiv:1506.03365 (2015).

[22] Kramberger, Tin, and Božidar Potočnik. "LSUN-Stanford car dataset: enhancing large-scale car image datasets using deep learning for usage in GAN training." Applied Sciences 10, no. 14 (2020): 4913.

[23] Xiao, Han, Kashif Rasul, and Roland Vollgraf. "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms." arXiv preprint arXiv:1708.07747 (2017).

[24] Shen, Kevin, Bernhard Jobst, Elvira Shishenina, and Frank Pollmann. "Classification of the Fashion-MNIST Dataset on a Quantum Computer." arXiv preprint arXiv:2403.02405 (2024).

[25] Hassani, Ali, and Humphrey Shi. "Dilated neighborhood attention transformer." arXiv preprint arXiv:2209.15001 (2022).

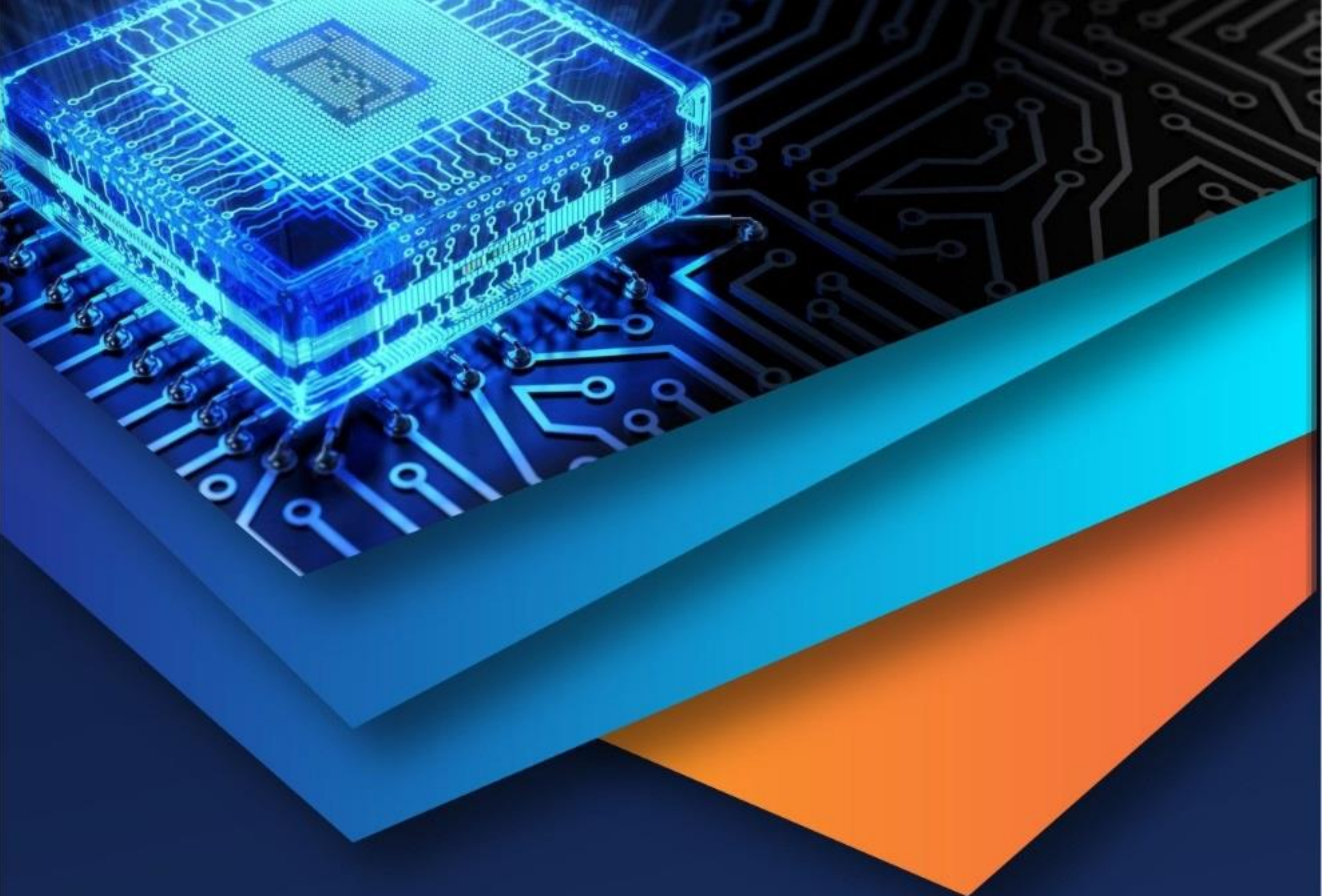
[26] Cabani, Adnane, Karim Hammoudi, Halim Benhabiles, and Mahmoud Melkemi. "MaskedFace-Net–A dataset of correctly/incorrectly masked face images in the context of COVID-19." Smart Health 19 (2021): 100144.

[27] Vuletić, Milena, Felix Prenzel, and Mihai Cucuringu. "Fin-gan: Forecasting and classifying financial time series via generative adversarial networks." Quantitative Finance (2024): 1-25.

[28] Li, Yumeng, Dan Zhang, Margret Keuper, and Anna Khoreva. "Intra-& extra-source exemplar-based style synthesis for improved domain generalization." International Journal of Computer Vision 132, no. 2 (2024): 446-465.

[29] Shibuya, Takashi, Yuhta Takida, and Yuki Mitsufuji. "Bigvsan: Enhancing gan-based neural vocoders with slicing adversarial network." In ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 10121-10125. IEEE, 2024.

[30] Yang, Zongyuan, Baolin Liu, Yongping Xiong, and Guibin Wu. "GDB: gated convolutions-based document binarization." Pattern Recognition 146 (2024): 109989.
[31] Sauer, Axel, Tero Karras, Samuli Laine, Andreas Geiger, and Timo Aila. "Stylegan-t: Unlocking the power of gans for fast large-scale text-to-image synthesis." In International conference on machine learning, pp. 30105-30118. PMLR, 2023.
[32] Pan, Wei, Anna Zhu, Xinyu Zhou, Brian Kenji Iwana, and Shilin Li. "Few shot font generation via transferring similarity guided global style and quantization local style." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 19506-19516. 2023.
[33] Neekhra, Bhavesh, Kshitij Kapoor, and Debayan Gupta. "Synthpop++: A Hybrid Framework for Generating A Country-scale Synthetic Population." arXiv preprint arXiv:2304.12284 (2023).
[34] Zheng, Zhuo, Shiqi Tian, Ailong Ma, Liangpei Zhang, and Yanfei Zhong. "Scalable multi-temporal remote sensing change data generation via simulating stochastic change process." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 21818-21827. 2023.
[35] Habekost, Jan-Gerrit, Erik Strahl, Philipp Allgeuer, Matthias Kerzel, and Stefan Wermter. "CycleIK: Neuro-inspired Inverse Kinematics." In International Conference on Artificial Neural Networks, pp. 457-470. Cham: Springer Nature Switzerland, 2023.
[36] Karras, Tero, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. "Alias-free generative adversarial networks." Advances in neural information processing systems 34 (2021): 852-863.
[37] Ju, Rui-Yang, Yu-Shian Lin, Yanlin Jin, Chih-Chia Chen, Chun-Tse Chien, and Jen-Shiun Chiang. "Three-stage binarization of color document images based on discrete wavelet transform and generative adversarial networks." arXiv preprint arXiv:2211.16098 (2022).
[38] Wang, Dong, Yuan Zhang, Kexin Zhang, and Liwei Wang. "Focalmix: Semi-supervised learning for 3d medical image detection." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3951-3960. 2020.
[39] Chinkamol, Amrest, Vetit Kanjaras, Phattarapong Sawangjai, Yitian Zhao, Thapanun Sudhawiyangkul, Chantana Chantrapornchai, Cuntai Guan, and Theerawit Wilaiprasitporn. "Octave: 2D En face optical coherence tomography angiography vessel segmentation in weakly-supervised learning with locality augmentation." IEEE Transactions on Biomedical Engineering (2022).
[40] Chao, Chen-Hao, Bo-Wun Cheng, and Chun-Yi Lee. "Rethinking ensemble-distillation for semantic segmentation based unsupervised domain adaption." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2610-2620. 2021.
[41] Wan, Ziyu, Bo Zhang, Dong Chen, Pan Zhang, Fang Wen, and Jing Liao. "Old photo restoration via deep latent space translation." IEEE Transactions on Pattern Analysis and Machine Intelligence 45, no. 2 (2022): 2071-2087.
[42] Takida, Yuhta, Masaaki Imaizumi, Takashi Shibuya, Chieh-Hsin Lai, Toshimitsu Uesaka, Naoki Murata, and Yuki Mitsufuji. "SAN: Inducing Metrizability of GAN with Discriminative Normalized Linear Layer." arXiv preprint arXiv:2301.12811 (2023).
[43] Szűcs, Gábor, and Modafar Al-Shouha. "Modular StoryGAN with background and theme awareness for story visualization." In International Conference on Pattern Recognition and Artificial Intelligence, pp. 275-286. Cham: Springer International Publishing, 2022.
[44] Nguyen, Van-Anh, Tuan Nguyen, Trung Le, Quan Hung Tran, and Dinh Phung. "Stem: An approach to multi-source domain adaptation with guarantees." In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9352-9363. 2021.
[45] Li, Jingyi, Weiping Tu, and Li Xiao. "Freevc: Towards high-quality text-free one-shot voice conversion." In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1-5. IEEE, 2023.
[46] Liang, Jingyun, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. "Swinir: Image restoration using swin transformer." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 1833-1844. 2021.
[47] Degardin, Bruno, Joao Neves, Vasco Lopes, Joao Brito, Ehsan Yaghoubi, and Hugo Proença. "Generative adversarial graph convolutional networks for human action synthesis." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1150-1159. 2022.
[48] Wang, Huan, Yijun Li, Yuehai Wang, Haoji Hu, and Ming-Hsuan Yang. "Collaborative distillation for ultra-resolution universal style transfer." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 1860-1869. 2020.
[49] He, Zhenliang, Meina Kan, and Shiguang Shan. "Eigengan: Layer-wise eigen-learning for gans." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 14408-14417. 2021.
[50] Tseng, Hung-Yu, Lu Jiang, Ce Liu, Ming-Hsuan Yang, and Weilong Yang. "Regularizing generative adversarial networks under limited data." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 7921-7931. 2021.
[51] Jia, Xinyu, Chuang Zhu, Minzhen Li, Wenqi Tang, and Wenli Zhou. "LLVIP: A visible-infrared paired dataset for low-light vision." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 3496-3504. 2021.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY