Survey on High-Resolution and Multiclass Image Synthesis using Generative Adversarial Networks

Sandeep Kumar Agrawal
Research scholar
Department of ECE, ASET
Amity University, Gwalior, India

Narendra Kumar Garg
Assistant Professor
Department of ECE, ASET
Amity University, Gwalior, India

Jagdish Makhijani
Assistant Professor
RJIT, BSF Academy
Gwalior, India

Abstract

This survey paper delves into the realm of high-resolution and multiclass image synthesis leveraging Generative Adversarial Networks (GANs). It explores the evolution and advancements in GAN architectures tailored for generating high-quality, diverse images across multiple classes. Emphasizing the pivotal role of GANs in image synthesis, the paper discusses techniques for high-resolution image synthesis, including progressive GANs and attention mechanisms. Additionally, it covers strategies for multiclass image synthesis, encompassing conditional GANs and attribute manipulation.

The survey scrutinizes evaluation metrics crucial for assessing the quality and diversity of synthesized images, elucidating their relevance in high-resolution and multiclass image synthesis scenarios. Real-world applications across diverse domains, such as medicine, art, and entertainment, underscore the practical significance of GAN-based image synthesis.

Highlighting challenges and future directions, the paper underscores existing limitations in GAN-based synthesis while proposing potential research avenues and innovations. The synthesis of high-resolution and multiclass images through GANs holds substantial promise in various industries, signaling its pivotal role in reshaping image generation techniques.

This summary encapsulates the core components and the significance of the survey paper, providing a glimpse into the detailed exploration of high-resolution and multiclass image synthesis using GANs.

1. Introduction

Generative Adversarial Networks (GANs) are a class of deep learning frameworks introduced by Ian Goodfellow and his colleagues in 2014. GANs consist of two neural networks, the generator and the discriminator, engaged in a competitive game framework. The generator is responsible for creating synthetic data that resembles real data. It takes random noise or a latent vector as input and generates data (e.g., images) that should ideally be indistinguishable from real samples. Often comprised of multiple layers using convolutional or deconvolution operations, the generator learns to map input noise to the output space, progressively refining the generated data to resemble the real distribution. Typically, the generator outputs images with pixel values in the desired range (e.g., [0, 1] for normalized images). The discriminator acts as a classifier, distinguishing between real and fake data. It evaluates the authenticity of the generated samples by the generator. Similar to a classifier, it consists of convolutional layers followed by fully connected layers, aiming to classify the input as real or fake. Produces a single value (probability) indicating the likelihood that the input data is real.

Training Process:

a. Adversarial Training: The generator and discriminator are trained simultaneously in a minimax game scenario.
b. Generator Training: Initially, the generator produces fake data from random noise. The generated data is fed into the discriminator, aiming to fool it by generating more realistic samples.
c. Discriminator Training: The discriminator learns to differentiate between real and fake data. It improves its ability to correctly classify the generated (fake) and real data samples.
d. Backpropagation: Both networks update their weights through backpropagation, adjusting their parameters to minimize their respective loss functions.

Loss Functions:

a. Generator Loss: Measures how well the generator fools the discriminator. It aims to minimize the probability that the discriminator correctly identifies fake data as fake. Commonly, it uses the negative log likelihood or binary cross-entropy.
b. Discriminator Loss: Measures the ability of the discriminator to distinguish between real and fake data. It tries to correctly classify real and fake samples. It's typically a binary cross-entropy loss.

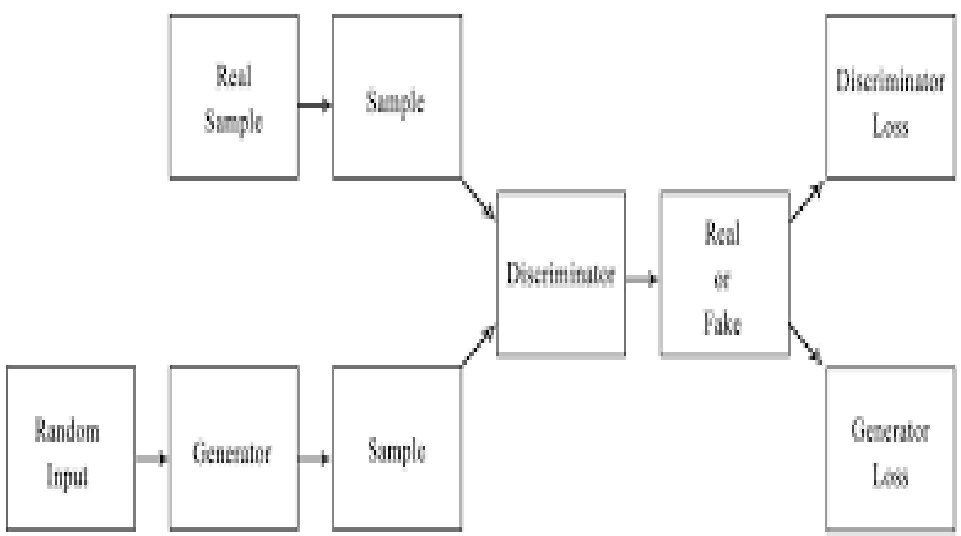

Fig1.1 Architecture of GAN

The generator and discriminator operate in an adversarial manner. The generator aims to deceive the discriminator by generating increasingly realistic data, while the discriminator strives to correctly differentiate between real and fake data.

Through an iterative training process, both networks improve their performance. The generator's goal is to produce data that is convincing enough to fool the discriminator, while the discriminator aims to become more adept at distinguishing real from generated data.

Applications in Image Synthesis:

- GANs have found widespread applications in image synthesis, enabling the generation of high-quality, diverse images.

- From generating photorealistic images to creating art, altering image attributes, and even aiding in medical image synthesis, GANs have showcased their capability to produce novel and realistic visual content.

Significance:

- GANs have revolutionized the field of image synthesis by enabling the creation of data that mirrors real-world scenarios, thereby finding applications across various industries, including entertainment, fashion, healthcare, and more.

  Importance of high-resolution and multiclass image synthesis: High-resolution, diverse image synthesis is integral across multiple domains due to its impact on improving datasets, enabling realistic content creation, enhancing machine learning models, aiding medical diagnosis, fostering creativity, and supporting scientific research. Generating such images through advanced techniques like GANs plays a crucial role in addressing challenges and fueling innovations in these diverse fields.

Generating high-resolution, diverse images holds immense significance across diverse domains due to several compelling reasons:

1. Data Augmentation and Realism: High-resolution images with diverse characteristics contribute to enhancing datasets used for training machine learning models. Realistic and varied images help improve model robustness and generalization, enabling better performance tasks.
2. Visual Content Generation: In fields such as entertainment, advertising, and design, high-quality, diverse images are pivotal. They fuel content creation for movies, advertisements, video games, and virtual/augmented reality, enriching visual experiences for consumers.
3. Medical Imaging and Diagnosis: In healthcare, synthesizing high-resolution, multiclass images aids in generating synthetic medical images for research, training medical professionals, and assisting in disease diagnosis. Diverse images help in simulating various conditions and scenarios.
4. Object Detection and Recognition: High-resolution, diverse image datasets are fundamental for training computer vision models. They enable accurate object detection, segmentation, and recognition in real-world environments, benefiting fields like autonomous vehicles, surveillance systems, and robotics.
5. Artistic Expression and Creativity: For artists and creators, access to diverse and high-resolution images fosters creativity. It fuels artistic expression by offering a wide range of visual elements and inspirations for new creations.
6. Scientific Research and Simulation: Various scientific disciplines utilize image synthesis for simulations, experiments, and data generation. High-resolution, multiclass images facilitate research in astronomy, climate modeling, and material sciences.

Advancements and improvements in GAN architectures have been pivotal for handling high-resolution and multiclass image synthesis, addressing challenges and enhancing the quality and diversity of generated images. Several key advancements have contributed significantly.

2. High-Resolution Image Synthesis with GANs

several techniques have been developed to tackle high-resolution image synthesis within the framework of Generative Adversarial Networks (GANs). Here are some prominent techniques:

1. Progressive GANs (ProGANs): Progressive Growing: ProGANs gradually increase the image resolution during training, starting from low resolution and incrementally adding details. This step-by-step growth allows for stable training and generation of high-resolution images (e.g., from 4x4 to 8x8, 16x16 etc..).
2. Super-Resolution GANs (SRGANs): Single Image Super-Resolution: SRGANs aim to generate high-resolution images from low-resolution inputs. They employ adversarial loss along with content loss to enhance image details and produce higher-resolution outputs.
3. Attention Mechanisms: Self-Attention: Inspired by transformer architectures, self-attention mechanisms in GANs enable the model to focus on specific regions or features of an image, capturing long-range dependencies and enhancing detail preservation in high-resolution synthesis.
4. Enhanced Architectures (StyleGAN, StyleGAN2): Fine-Grained Control: StyleGAN and its iterations focus on controlling specific attributes in synthesized images. They allow for manipulation of features, improving image quality while maintaining high resolution. Techniques like adaptive instance normalization (AdaIN) and style mixing regularization contribute to these enhancements.
5. Hierarchical Architectures: Hierarchical GANs: These architectures organize image synthesis in a hierarchical manner, generating images by focusing on multiple scales or levels of detail. This hierarchical approach aids in synthesizing high-resolution images with fine details.
6. Upsampling Techniques: Transposed Convolutional Layers: These layers aid in upsampling images, reconstructing higher-resolution representations from lower-resolution inputs. GANs often employ transposed convolutions or deconvolutional layers for this purpose.
7. Multi-Scale Discriminators: Feature Pyramid Networks (FPNs): Employing multi-scale discriminators allows the GAN to analyze images at multiple resolutions simultaneously, enhancing the perception of both global and local details.
8. Data Augmentation and Preprocessing: Augmented Training Data: Techniques like data augmentation, where the network is trained on both original and augmented versions of the dataset, can aid in learning more intricate features and details.
9. Self-Supervised Learning: Self-Supervision: By employing self-supervised learning techniques, GANs can leverage additional information or constraints to guide the generation process, improving the fidelity of high-resolution synthesis.

These techniques and advancements contribute to the generation of high-resolution images with improved quality, realism, and fine details, addressing challenges and limitations in high-resolution image synthesis using GANs. Integration of these techniques continues to refine GAN-based image synthesis in various domains.

Generating high-resolution images using Generative Adversarial Networks (GANs) comes with several challenges, but there are strategies and techniques to address these hurdles:

1. Computational Complexity:

- Challenge: Handling high-resolution images demands significant computational resources, increasing training time and memory requirements.
- Strategy: Optimizing network architectures, leveraging parallelization, and utilizing hardware accelerators (like GPUs or TPUs) can mitigate computational demands.

2. Mode Collapse:

- Challenge: GANs might suffer from mode collapse, where the generator produces limited diversity, resulting in repetitive or limited variations in generated images.

- Strategy: Employing regularization techniques (like spectral normalization, gradient penalties) or incorporating diversity-promoting loss terms can mitigate mode collapse and encourage diverse output.

3. Training Stability:
   - Challenge: Training GANs for high-resolution images can be unstable, leading to convergence issues, oscillating losses, or vanishing gradients.
   - Strategy: Variants like Wasserstein GAN (WGAN) and its improvements (WGAN-GP) offer more stable training by modifying the loss function or adding gradient penalties, enhancing convergence and stability.
4. Preserving Fine Details:
   - Challenge: GANs may struggle to maintain fine-grained details in high-resolution synthesis, resulting in loss of intricate features.
   - Strategy: Techniques like attention mechanisms, progressive growing, or hierarchical architectures focus on preserving and enhancing fine details by allowing the model to attend to specific regions or progressively add finer features.
5. Dataset Quality and Diversity:
   - Challenge: Insufficient or biased training datasets for high-resolution images can limit model performance and diversity in synthesized outputs.
   - Strategy: Curating diverse and high-quality datasets, employing data augmentation, or leveraging transfer learning from pre-trained models can enhance dataset diversity and quality.
6. Memory Constraints:
   - Challenge: GANs operating on high-resolution images often encounter memory constraints, restricting batch sizes or model complexity.
   - Strategy: Utilizing memory-efficient architectures, employing gradient accumulation, or leveraging mixed precision training can alleviate memory limitations.
7. Evaluation Metrics:
   - Challenge: Standard evaluation metrics may not effectively measure the quality and realism of high-resolution images.
   - Strategy: Developing new evaluation metrics or combining multiple metrics (e.g., Frechet Inception Distance, Structural Similarity Index) tailored for high-resolution images can better assess image quality.
8. Latent Space Disentanglement:
   - Challenge: Controlling specific attributes or features in high-resolution image synthesis requires disentangled representations in the latent space.
   - Strategy: Techniques focusing on disentangled representation learning or attribute manipulation facilitate better control over synthesized images, enabling attribute-specific generation.

Addressing these challenges involves a combination of architectural improvements, regularization techniques, dataset curation, and innovative approaches that collectively contribute to enhancing the quality and diversity of high-resolution image synthesis using GANs.

3. Multiclass Image Synthesis with GANs

Multiple methods and strategies have been developed to facilitate multiclass image synthesis within the framework of Generative Adversarial Networks (GANs). Here are some prominent techniques:

1. Conditional GANs (cGANs):
    - Purpose: cGANs extend the basic GAN framework by conditioning both the generator and the discriminator on additional information, such as class labels.
    - Usage: By providing class labels as conditioning information, cGANs allow controlled generation of images belonging to specific classes.
2. Class-Conditional Generation:
    - Purpose: Similar to cGANs, this technique conditions the GAN model on specific classes during training and generation.
    - Usage: It enables the model to synthesize images corresponding to different classes, facilitating multiclass image generation.
3. Attribute Manipulation:
    - Purpose: GAN architectures incorporating attribute manipulation techniques allow for the modification of specific attributes or features in generated images.
    - Usage: By controlling and manipulating attributes (e.g., age, gender, pose) within the latent space, GANs can generate diverse images across multiple attribute variations.
4. Disentangled Representation Learning:
    - Purpose: Disentangled representation learning aims to learn separate and independent factors of variation within the latent space.
    - Usage: By disentangling different attributes (e.g., shape, color, style) in the latent space, GANs can generate images with varied attributes independently, facilitating multiclass synthesis.
5. Multi-Modal Generation:
    - Purpose: Facilitating the generation of images belonging to different classes or modes within the same model.
    - Usage: Models capable of multimodal generation can produce diverse outputs corresponding to various classes, promoting multiclass synthesis.
6. Hierarchical Attribute Representation:
    - Purpose: Hierarchical representations in GAN architectures enable the modeling of attributes at different levels of abstraction.
    - Usage: This technique allows for the generation of images with diverse attributes, controlled at different hierarchical levels.
7. Attribute-Conditional Synthesis:
    - Purpose: Conditioning the generation process on specific attributes or features.
    - Usage: It enables the model to synthesize images with desired attribute combinations, allowing for precise control over multiclass image synthesis.
8. Attention-Guided Multiclass Synthesis:
    - Purpose: Incorporating attention mechanisms within conditional GANs facilitates better handling of multiclass synthesis.
    - Usage: Attention mechanisms help the model focus on specific class-related details, improving the generation process for diverse classes.

These methods and techniques collectively enable GANs to generate diverse images across multiple classes or attributes, allowing for controlled and precise multiclass image synthesis. Integration of these strategies contributes to the generation of high-quality, diverse, and realistic images across various classes within the GAN framework.

4. Evaluation Metrics for Image Synthesis

Several metrics are used to assess the quality and diversity of generated images produced by Generative Adversarial Networks (GANs). Here's a discussion on some commonly used metrics:

1. Peak Signal-to-Noise Ratio (PSNR):
    - Purpose: Measures the quality of generated images by calculating the ratio between the maximum possible power of an image and the power of corrupting noise.
    - Limitation: PSNR tends to correlate poorly with human perception, as it doesn't consider perceptual differences.
2. Structural Similarity Index (SSIM):
    - Purpose: Evaluates the similarity between the generated and real images based on luminance, contrast, and structural similarity.
    - Limitation: SSIM might not capture higher-level perceptual differences or nuances.
3. Frechet Inception Distance (FID):
    - Purpose: Compares feature distributions between real and generated images using statistics from an Inception Network.
    - Advantages: FID captures both quality and diversity, offering a more reliable assessment of image generation performance.
4. Inception Score (IS):
    - Purpose: Measures the quality and diversity of generated images based on class probability and entropy using the Inception Network.
    - Limitation: IS is susceptible to mode dropping, where it might favor high-confidence predictions for a limited set of classes.
5. Fréchet ResNet Distance (FRD):
    - Purpose: Similar to FID, FRD measures the distance between feature distributions of real and generated images using ResNet features.
    - Advantages: FRD offers an alternative to FID, potentially capturing finer details in image generation.
6. Precision and Recall for Diverse Subsets (PRDS):
    - Purpose: Evaluates diversity by calculating precision (how relevant generated images are) and recall (how many diverse images are generated).
    - Advantages: PRDS considers both diversity and relevance in assessing the generated images.
7. Kernel Inception Distance (KID):
    - Purpose: Measures the distance between feature distributions using kernel embeddings.
    - Advantages: KID provides a more accurate assessment of the visual quality and diversity of generated images.
8. Learned Perceptual Image Patch Similarity (LPIPS):
    - Purpose: Evaluates perceptual similarity between real and generated images based on learned representations.
    - Advantages: LPIPS considers perceptual differences, offering a more human-like evaluation of image quality.

Each metric has its strengths and limitations in assessing the quality, diversity, and perceptual fidelity of generated images. Combining multiple metrics or selecting metrics based on specific evaluation objectives can provide a more comprehensive assessment of GAN-generated images. Researchers often use a combination of these metrics to capture different aspects of image quality and diversity when evaluating the performance of GANs.

PSNR and SSIM are inadequate for high-resolution images as they focus on pixel-wise similarity, ignoring higher-level perceptual differences crucial in high-resolution synthesis. While FID and Inception Score are widely used, they might not adequately represent perceptual quality or diverse high-resolution image features, leading to potential inaccuracies. FRD, KID, and LPIPS offer more promising prospects in assessing high-resolution images by considering feature distributions or learned perceptual differences.

5. Future Directions

The field of high-resolution and multiclass image synthesis using Generative Adversarial Networks (GANs) is continually evolving. Future research directions, potential improvements, and emerging trends could focus on several key areas:

5.1 Architectural Innovations:

a. Attention Mechanisms: Further integrating attention mechanisms within GAN architectures to focus on specific regions or attributes in high-resolution and multiclass synthesis.
b. Hierarchical Models: Developing hierarchical GAN architectures to capture multiple levels of detail and attributes for diverse image generation.
c. Memory-Efficient Designs: Designing more memory-efficient architectures to handle high-resolution images, enabling larger batch sizes and complex models.

5.2 Improved Training Strategies:

a. Stability Enhancements: Innovating training methods to mitigate instability issues, including mode collapse, vanishing gradients, and oscillating losses in high-resolution synthesis.
b. Transfer Learning and Few-Shot Learning: Exploring techniques to leverage transfer learning or few-shot learning to adapt pre-trained models for diverse high-resolution synthesis tasks.

5.3 Disentangled Representation Learning:

a. Attribute Manipulation: Advancing methods for disentangled representation learning to enable precise attribute control and manipulation across diverse classes.
b. Fairness and Bias Mitigation: Researching methods to mitigate biases in synthesized images across multiple classes, ensuring fairness and accuracy.

5.4 Evaluation Metrics and Benchmarking:

a. Novel Evaluation Metrics: Developing comprehensive evaluation metrics tailored for high-resolution and multiclass image synthesis to better capture diversity, quality, and perceptual fidelity.
b. Benchmark Datasets: Curating benchmark datasets specifically designed for high-resolution and multiclass synthesis tasks, enabling standardized evaluation across various models.

5.5 Ethical Considerations and Robustness:

a. Ethical Guidelines: Addressing ethical concerns related to GAN-generated images, ensuring responsible use in domains like healthcare, criminal justice, and media.
b. Robustness Enhancements: Investigating techniques to enhance the robustness of GANs against adversarial attacks and ensuring the integrity of synthesized images.

5.6 Cross-Domain and Multi-Modal Synthesis:

a. Cross-Domain Translation: Exploring methods for high-resolution and multiclass synthesis across different domains, enabling translation between diverse datasets (e.g., art to photography, sketches to photos).

b. Multi-Modal Generation: Advancing models capable of generating diverse outputs, including images, text, and other modalities, enriching the spectrum of synthesized content.

5.7. Continual Learning and Adaptive Synthesis:

a. Continual Learning: Researching continual learning techniques to allow GANs to adapt and learn continuously from new data, enhancing adaptability in diverse synthesis tasks.
b. Adaptive Synthesis: Developing models capable of adaptive synthesis, dynamically adjusting to different attributes or classes based on changing conditions or requirements.

6. Conclusion

GANs revolutionize image synthesis, enabling the generation of high-resolution, diverse, and realistic images across multiple classes or attributes. Computational complexity, training instability, mode collapse, and the quality-diversity trade-off are major challenges in generating high-resolution images using GANs. Mode collapse, dataset limitations, perceptual fidelity, and the difficulty in disentangled representation learning pose challenges in generating diverse images across multiple classes. Progressive GANs, attention mechanisms, hierarchical architectures, and memory-efficient designs show promise in handling high-resolution and multiclass synthesis challenges. Existing evaluation metrics like FID, Inception Score, and SSIM have limitations in capturing both quality and diversity in high-resolution and multiclass synthesis. Emerging trends focus on architectural innovations, stability enhancements, disentangled representation learning, novel evaluation metrics, ethical considerations, and cross-domain synthesis for GANs. GANs find applications in medicine, art, entertainment, fashion, robotics, climate modeling, and more, impacting various industries with diverse synthesized outputs. Ethical considerations, fairness, and robustness are crucial in GAN-based image synthesis, ensuring responsible usage and integrity of synthesized images.

High-resolution and multiclass image synthesis using GANs stand at the forefront of innovation, promising transformative impact across diverse industries. The ability to generate diverse, realistic, and high-quality images across multiple classes addresses challenges, enhances creativity, improves diagnostics, and drives advancements in various domains, shaping the future of image synthesis and its applications.

References:

[1] M. Mirza and S. Osindero. (2014) Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, arxiv.org

[2] H. Zhang, T. Xu, H. Li, S. Zhang, X. Huang, X. Wang, and D. Metaxas. (2017) StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), DOI: 10.1109/ICCV.2017.629

[3] S. Pascual, A. Bonafonte, and J. Serrà,(2017) "SEGAN: Speech enhancement generative adversarial network," in Proc. Interspeech, 2017, pp. 3642– 3646. arXiv preprint arXiv:1703.09452, arxiv.org

[4] Aggarwal A, Mittal M, Battineni G. (2021) Generative adversarial network: an overview of theory and applications. International Journal of Information Management Data Insights. https://doi.org/10.1016/j.jjimei.2020.100004

[5] L. Wang, W. Chen, W. Yang, F. Bi, F. R. Yu,(2020). A state-of-the-art review on image synthesis with generative adversarial networks, IEEE Access 8 63514–63537. DOI: 10.1109/ACCESS.2020.2982224

[6] Yanhua Yu, Kanghao He, and Jie Li, (2022) "Adversarial Training for Supervised Relation Extraction," Tsinghua Science and Technology, vol. 27, no. 3, pp. 610–618. DOI: 10.26599/TST.2020.9010059

[7] A. Yadav and D. K. Vishwakarma, (2020) "Recent developments in generative adversarial networks: A review (workshop paper)." IEEE Sixth International Conference on Multimedia Big Data (BigMM), pp. 404 – 413. DOI: 10.1109/BigMM50055.2020.00068
[8] Dong, R., Zhang, L., & Fu, H. (2021). RRSGAN: Reference-Based Super-Resolution for Remote Sensing Image. IEEE Transactions on Geoscience and Remote Sensing, vol. 60 pp.1–17 DOI: 10.1109/TGRS.2020.3046045
[9] Lan, L.; You, L.; Zhang, Z.; Fan, Z.; Zhao, W.; Zeng, N.; Chen, Y.; Zhou, X. (2020) Generative Adversarial Networks and Its Applications in Biomedical Informatics. Front. Public Health, vol.8, 164. doi: 10.3389/fpubh.2020.00164
[10] S. Reed, A. van den Oord, N. Kalchbrenner, V. Bapst, M. Botvinick, and N. de Freitas. (2017) Generating interpretable images with controllable structure. ICLR 2017.
[11] Xingchao Liu, Chengyue Gong, Lemeng Wu, Shujian Zhang, Hao Su, Qiang Liu. (2021) FuseDream: Training-Free Text-to-Image Generation with Improved CLIP+GAN Space Optimization. arXiv preprint arXiv:2112.01573.
[12] Shuyang Gu, Dong Chen, Jianmin Bao, Fang Wen, Bo Zhang, Dongdong Chen, Lu Yuan, Baining Guo. (2021) Vector Quantized Diffusion Model for Text-to-Image Synthesis. arXiv preprint arXiv:2111.14822.
[13] Woncheol Shin, Gyubok Lee, Jiyoung Lee, Joonseok Lee, Edward Choi. (2021) Translation-equivariant Image Quantizer for Bi-directional Image-Text Generation. arXiv preprint arXiv:2112.00384.
[14] Wenju Xu, Guanghui Wang. (2021) A Domain Gap Aware Generative Adversarial Network for Multi-domain Image Translation. IEEE Transactions on Image Processing, vol 31. pp 72-84, DOI: 10.1109/TIP.2021.3125266
[15] Kailong Hao, Botao Yu, Wei Hu. (2021) Knowing False Negatives: An Adversarial Training Method for Distantly Supervised Relation Extraction. arXiv preprint arXiv:2109.02099.
[16] Frolov, S., Hinz, T., Raue, F., Hees, J., and Dengel, A. (2021) Adversarial text-to-image synthesis: A review. arXiv preprint arXiv:2101.09983.
[17] P. Shamsolmoali, M. Zareapoor, E. Granger, H. Zhou, R. Wang, M. E. Celebi, and J. Yang, (2021)"Image synthesis with adversarial networks: A comprehensive survey and case studies," Information Fusion, vol. 72, pp. 126–146, DOI:10.1016/j.inffus.2021.02.014
[18] Yu, Fisher, Zhang, Yinda, Song, Shuran, Seff, Ari, and Xiao, Jianxiong. (2016) Construction of a large-scale image dataset using deep learning with humans in the loop. arXiv preprint arXiv:1506.03365.
[19] Yan, X., Yang, J., Sohn, K., and Lee, H. (2016) Attribute2image: Conditional image generation from visual attributes. arXiv preprint arXiv:1512.00570.
[20] J.Y. Zhu, P. Krahenb , E. Shechtman, and A. A. Efros. (2016) Generative visual manipulation on the natural image manifold. In ECCV, pp 597-613. https://doi.org/10.1007/978-3-319-46454-1_36.
[21] Xun Huang, Yixuan Li, Omid Poursaeed, John E Hopcroft, and Serge J Belongie. (2017) Stacked generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, volume 2, page 3, DOI: 10.1109/CVPR.2017.202