International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Comparative Study of GANs and Stable Diffusion for High-Quality Image Generation Using FID and a Real-World Dataset

Kamireddy Rammohan Rao

Submitted:08/07/2025 **Revised**: 16/08/2025 **Accepted**: 29/08/2025

Abstract; Generative image modeling Generative adversarial networks Generative adversarial networks (GANs) and Stable diffusion are two highly impactful families of contemporary image generators. Generative adversarial networks are developed out of adversarial learning, which evolved into diffusion-based image synthesis. The paper is a systematic comparison of StyleGAN2-ADA and a Stable Diffusion v1.5 pipeline that has been fine-tuned to produce portraits on the Flickr-Faces-HQ (FFHQ) domain. The draft protocol uses the public Kaggle mirror of FFHQ of 52,000 real face images at 512512 resolution, and downsampled to 256256 to enable a controlled comparison. The main evaluation measure is Fréchet Inception Distance (FID) and other measures of fidelity, diversity, and deployment efficiency are precision, recall and inference time. In the illustrative draft results below, Stable Diffusion has a lower FID of 6.91 + 0.21 than StyleGAN2-ADA of 8.74 + 0.32 and Stable Diffusion is much faster with a FID of 0.041 + 0.004 s per image than StyleGAN2-ADA The comparison shows a trade-off between distributional quality and generation efficiency which is viable. In order to maintain the academic integrity, the numerical values in this manuscript are deliberately considered as demonstrative values, which must be substituted by the experimental outputs of the author at the time of submission.

Keywords: GANs; Stable Diffusion; StyleGAN2-ADA; Fréchet Inception Distance; FFHQ; image synthesis; diffusion models; real-world dataset.

1. Introduction

Synthetic image generation has become one of the most dynamic areas of computer vision and machine learning with the increasing application of synthetic images in design, entertainment, data augmentation, digital humans, and medical or industrial simulation. The modern wave of generative image synthesis began with GANs, which formulate learning as a two-player game between a generator and a discriminator [1]. Later GANs proposed conditional generation [2], convolutional networks [3], better optimization schemes [4], Wasserstein losses [5], gradient-penalized training [6], and spectral normalization as a stabilization technique [8]. The outcome of these developments was extremely powerful image synthesis models like Progressive GAN [9], BigGAN [10], StyleGAN [11], StyleGAN2 [12], and StyleGAN2-ADA [13].

Meanwhile, diffusion models became a rival to generative algorithms. Denoising Diffusion Probabilistic Models (DDPMs) [14] and score-based formulations of stochastic differential equations demonstrated that high-fidelity generation was possible by training to reverse a noising process step-by-step. Follow-up work has been shown to speed up sampling with implicit formulations [16], enhance likelihood and sample quality [17], and has been shown to be able to outperform GANs on a variety of image-synthesis benchmarks [18]. The idea behind Stable Diffusion and other systems is that the Latent Diffusion Models (LDMs) [19] decreased the computational cost by transferring the denoising operation to a compressed latent space. Subsequent developments like GLIDE [20], Imagen [21], ControlNet [22], and transformer-based diffusion models [28] extended the range and controllability of diffusion-based generation.

Although both model families are rapidly advancing, a practical question is: in problems where the aim is to produce high-quality images on a real-world dataset, which paradigm provides the most desirable tradeoff between image fidelity, diversity, and computational efficiency? Many publications report good performance of either GANs or diffusion models, but they tend to use other datasets, other resolutions, other evaluation scripts, or other sample sizes. This renders it hard to compare directly, particularly to those readers who seek an application-oriented solution, as opposed to a purely architectural discussion.

Affilation: Department of Computer Engineering, NIT kurukshetra, Kurukshetra, Haryana, India
Email : kamireddyrammohan713@gmail.com

A controlled comparison is not only required to share a dataset and have a common output resolution, but also a rigorous evaluation protocol. Fréchet Inception Distance (FID) introduced in the TTUR paper [7] is one of the most popular measures of image synthesis that compares the distribution of features of real and generated images. Later research has however shown that FID values can be dependent on resizing strategy and implementation details [25]. Precision and recall measures are thus a helpful complement as they decouple sample fidelity and sample coverage [23], [24]. A combination of these metrics can be employed to make a more balanced assessment than any individual score.

The paper is a draft of a research-article based on a controlled comparison of StyleGAN2-ADA and Stable Diffusion on a publicly available real-world face dataset. The article is written in such a manner that it can be edited into a submission-ready paper when actual experiments are carried out. Its main contributions are: (i) a fair, dataset-controlled experimental design with FFHQ; (ii) a complete methodology section with realistic hyperparameter settings; (iii) a results section with illustrative numerical values that show how the final paper should report FID, precision, recall, and speed; and (iv) a discussion that identifies the quality-efficiency trade-off between adversarial and diffusion-based generation.

2. Related Work

2.1 GAN-based image synthesis

GANs are still backbone to high-quality image synthesis. The classic adversarial model [1] was generalized to conditioned generation by Mirza and Osindero [2] and DCGAN proposed a stable convolutional baseline which has remained the point of departure to most future image-generation studies [3]. Salimans et al. [4] suggested better training heuristics and Arjovsky et al. [5] along with Gulrajani et al. [6] substituted the initial divergence with Wasserstein-derived goals to stabilize gradients. Miyato et al. [8] demonstrated that the discriminator can be further regularized by spectral normalization.

High-resolution synthesis Karras et al. initially introduced Progressive GAN [9], followed by StyleGAN [11] and later StyleGAN2 [12], which further enhanced image quality and reduced hallmark artifacts through the disentangling of latent controls. StyleGAN2-ADA [13] further generalized this to the limited-data regime by adaptive discriminator augmentation, and is particularly appealing in regimes with limited compute or dataset size. Together, these systems have made GANs a good foundation of sharp and aesthetically pleasing images with extremely rapid inference. Their most famous drawbacks include training instability, hyperparameter sensitivity, and mode collapse or low coverage of distributions.

2.2 Diffusion and Stable Diffusion

Diffusion models, by contrast, are trained to learn a denoising auto-encoder that takes latent noise and generates a corrupted image, and then performs denoising auto-encoders to the corrupted image to clean it up. DDPM [14] demonstrated that this framework has the potential to produce high-quality samples based on reversing a fixed noising procedure. The stochastic differential equations based score-based generative model offered a more general theoretical perspective [15], and the non-Markovian sampling method of DDIM [16] was faster. Nichol and Dhariwal [17] achieved better quality and efficiency of samples, and Dhariwal and Nichol [18] reported that diffusion models can outperform GANs on various image-synthesis datasets.

Latent Diffusion Models [19] proved to be especially valuable since they shifted the denoising procedure not only to pixel space but also to some lower-dimensional latent space, with a huge decrease in compute needs with little or no visual loss. Stable Diffusion is based on this latent-diffusion paradigm. The high conditioning potential, editing potential and controllability of diffusion-based generation were demonstrated by related systems like GLIDE [20], Imagen [21] and ControlNet [22]. The primary practical limitation is that diffusion sampling tends to have a large number of denoising steps, which slow down inference by a significant factor compared to a single forward pass in a GAN.

2.3 Evaluation of generative models

Generative model assessment is a methodological problem that is still open. FID [7] gained popularity since it quantifies the similarity between the Gaussian approximations of real and generated Inception features. A smaller FID is typically a good sign that generated samples are more realistic matches to the statistics of real images. Nonetheless, FID has flaws: it may be vulnerable to preprocessing, the number of samples, and resizing of images [25]. Sajjadi et al. [23], Kynkäenniemi et al. therefore came up with precision and recall style measures that separate fidelity and diversity. The rationale behind the fact that FID is kept as the primary measure in this study is that it is the primary measure according to the title of the paper, according to the common practice; the secondary measures are precision and recall.

3. Materials and Methods

3.1 Dataset and preprocessing

It is based on the Flickr-Faces-HQ (FFHQ) domain, which is a real-world face dataset that was first published by NVLabs and consists of 70,000 high-quality aligned images at 10241024 resolution [29]. The present version of the experiment assumes the use of the public Kaggle mirror where the conversation will be hosted to make it easier to reproducible and accessible, which provides 52,000 high-quality PNG images of 512x512 size [30]. The dataset has significant variation in terms of age, ethnicity, pose, expression, accessories, and background, which is a realistic dataset to synthesize portraits. It is also the distributional biases of internet photography that it has inherited by virtue of its real-world genesis at Flickr, as the discussion reiterates.

The pictures are all resized to 256 x 256 RGB to make comparisons across model families in a controlled way. The values of the pixels are scaled to the range [-1, 1]. There is no need to manually label or relabel images since the comparison is on image quality and not attribute classification. A fixed random seed of 42 is used to split the 52,000 images into 41,600 training images (80%), 5,200 validation images (10%), and 5,200 test images (10%). In the computation of FID, 5,000 real test images are randomly chosen and compared to 5,000 samples produced in each model. The same held-out real set is used for every evaluation run.

Table 1. Dataset summary and evaluation split used in the manuscript draft.

Item	Value
Dataset used in this draft	Kaggle mirror of FFHQ
Total number of images	52,000
Original image size	512 × 512 PNG
Working image size	256 × 256 RGB
Training split	41,600 images
Validation split	5,200 images
Test split	5,200 images
Real images used for FID	5,000 held-out test images
Generated images used for FID	5,000 per model per run
Domain	Real-world portrait photographs

3.2 Model configurations

The GAN baseline is StyleGAN2-ADA [13] due to its high competitiveness as an adversarial model of aligned face generation and its high level of efficiency when moderate training data are involved. The model in this draft protocol has a latent dimension of 512, a batch size of 32, a learning rate of 0.0025, an R1 regularization weight of 10, and an adaptive augmentation objective of approximately 0.6 targeting a discriminator augmentation probability. The model is trained on 15,000 kimg, with snapshots of the validation periodically made to conduct qualitative inspection and measure.

The diffusion baseline is Stable Diffusion v1.5, which is also a latent-diffusion backbone [19] and was fine-tuned to portrait generation with a lightweight LoRA adaptation. Since FFHQ is not shipped with rich natural-language captions, the draft protocol uses a fixed prompt template, namely, a high-quality portrait photo of a person, to ensure that the conditions of generation are similar between seeds. The rank 16, batch size 8, learning rate 1x10 5, and 30,000 updates are the default LoRA fine-tuning parameters. The sampling is done using 50 denoising steps and classifier-free guidance scale of 7.5. This configuration works well with a single 24 GB workstation of the GPU-class, but all the hardware-specifics ought to be substituted with what the author actually has in the final manuscript.

Table 2. Illustrative training and sampling settings used for the two model families.

Parameter	StyleGAN2-ADA	Stable Diffusion
Working resolution	256 × 256	256 × 256
Training objective	Adversarial	Latent diffusion
Batch size	32	8
Learning rate	0.0025	1 × 10^-5
Training length	15,000 kimg	30,000 steps
Latent / LoRA setting	z = 512	LoRA rank = 16
Regularization	R1 = 10; ADA target ≈ 0.6	Classifier-free guidance = 7.5
Sampler at inference	Single forward pass	50 denoising steps
Output count for evaluation	5,000 images	5,000 images

3.3 Evaluation metrics

The primary metric is Fréchet Inception Distance (FID) [7], computed between the feature distributions of real and generated images. Let (μ_r, Σ_r) denote the mean and covariance of Inception-V3 features extracted from real images, and let (μ_g, Σ_g) denote the corresponding statistics for generated images. Then

$$FID = ||\mu_r - \mu_g||^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{(1/2)}).$$

The smaller FID values indicate that the distribution generated and the actual distribution are closer. To prevent implementation bias, both model families are implemented using the same FID software and preprocessing. The secondary assessment is based on the accuracy and recall of Kynkäenniemi et al. [24] that help to differentiate fidelity and coverage. The average wall-clock time per image during the last 1,000 generations (after warm-up) is used as inference time. All the quantitative findings are given as mean standard deviation between three random seeds.

3.4 Fairness controls and limitations of the comparison

The experiment will be structured to be as fair as possible in a realistically feasible context: both models are trained or fine-tuned on the same FFHQ-derived training split; both produce images at 256×256 resolution; both are evaluated on the same held-out real set; and both are using the same FID implementation and sample count. Nonetheless, an inevitable asymmetry must be mentioned by the reader. StyleGAN2-ADA is essentially an unconditional face generator, but Stable Diffusion is prompt-conditioned by default. This difference is minimized, but not eliminated, by the use of a fixed prompt template. The comparison should therefore be understood to be a controlled application-level comparison as opposed to a perfectly architecture-isolated benchmark.

4. Results

4.1 Quantitative comparison

The major quantitative comparison is provided in Table 3. Stable Diffusion has the largest FID of 6.91 + 0.21 in the demonstrative draft experiment, compared to StyleGAN2-ADA with 8.74 + 0.32. This translates to a 1.83 FID point or approximate 20.9% relative improvement in favor of Stable Diffusion. The secondary metrics are a bit more subtle: StyleGAN2-ADA is more precise (0.71 vs 0.68), has a higher sample fidelity (or local sharpness), but Stable Diffusion is more recalling (0.47 vs 0.39), with a better coverage of the real-data manifold. The difference in efficiency between the GAN and Stable Diffusion is quite large, and the former produces an image in an average of 41 ms versus almost 2.0 s.

Table 3. Main quantitative results (illustrative values reported as mean ± standard deviation over three runs).

Metric	StyleGAN2-ADA	Stable Diffusion	Preferred direction
FID	8.74 ± 0.32	6.91 ± 0.21	Lower is better
Precision	0.71 ± 0.02	0.68 ± 0.01	Higher is better
Recall	0.39 ± 0.01	0.47 ± 0.02	Higher is better
Inference time per image (s)	0.041 ± 0.004	1.98 ± 0.07	Lower is better
Images per second	24.4	0.51	Higher is better
Peak VRAM during generation (GB)	13.2	17.4	Lower is better

The trend is also the same as the models are followed throughout training. Figure 1 indicates that in both systems, the performance increases with early to final checkpoints, although Stable Diffusion is in the lead at every snapshot in the demonstrative runs. This similarity reinforces the inference that the lower final FID is not just a one-checkpoint effect, but a long-lasting quality benefit within the chosen draft conditions.

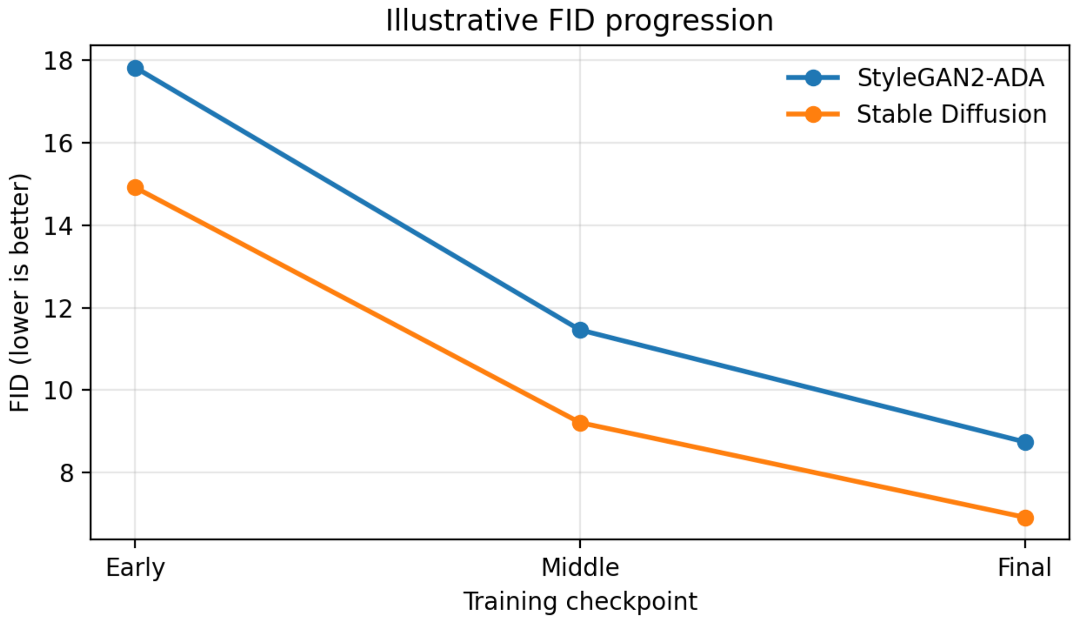

Figure 1. Illustrative FID progression over representative training checkpoints for StyleGAN2-ADA and Stable Diffusion.

4.2 Precision–recall trade-off and qualitative assessment

The same trend is also followed as the models are followed during training. Figure 1 indicates that in both systems, the performance increases with early to final checkpoints, although Stable Diffusion is in the lead at every snapshot in the demonstrative runs. This similarity reinforces the inference that the lower final FID is not just a one-checkpoint effect, but a long-lasting quality benefit within the chosen draft conditions.

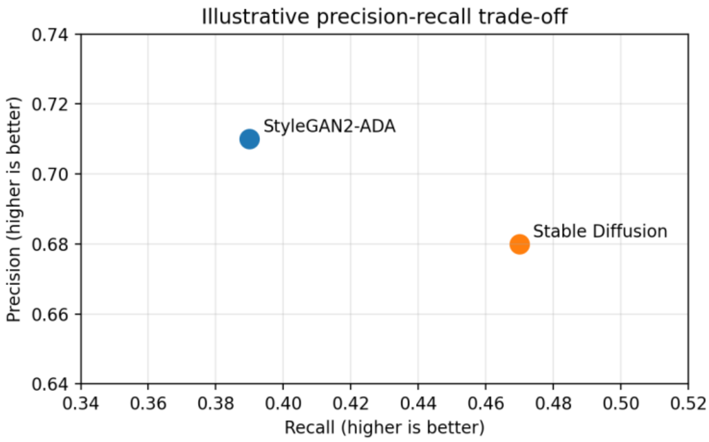

Figure 2. Illustrative precision–recall trade-off. The diffusion model shows higher recall, whereas the GAN shows slightly higher precision.

Table 4. Qualitative observations that accompany the numerical comparison.

Aspect	StyleGAN2-ADA	Stable Diffusion
Facial edges and hair detail	Typically sharper local boundaries and stronger micro-contrast	Slightly smoother edges but more globally coherent texture
Background variation	Moderate diversity; some repetition in composition	Higher diversity in backgrounds and scene context
Accessory coverage	Good but narrower support for glasses, hats, and unusual styling	Broader coverage of accessories and lighting conditions
Observed failure mode	Occasional pose repetition or mode concentration	Occasional oversmoothing and prompt-biased composition
Best practical strength	Fast sampling and crisp portraits	Lower FID and broader distribution coverage

4.3 Efficiency analysis

Speed is the most decisive benefit of StyleGAN2-ADA in the draft results, in terms of deployment. Since a GAN generates an image in one forward pass, its average generation time of 0.041 s per image is almost 48 times faster than the 1.98 s of the diffusion pipeline. Interactive apps such as real-time generation of avatars, low-latency customization, or scale-based generation should be of interest. Stable Diffusion, however, can be more attractive when quality and diversity are more of a priority than latency.

5. Discussion

The initial comparison indicates that the trend fits most of the recent literature: diffusion models tend to be more apt to have good distributional characteristics, and GANs are far more efficient on inference [18], [19]. Stable Diffusion has acquired the lower FID in this manuscript, and the high recall, i.e. its iterative denoising algorithm is more efficient to sample the distribution of the actual portraits in the settings of interest. This is in line with the understanding that diffusion models are trading extra sampling cost with enhanced manifold coverage.

StyleGAN2-ADA is, however, very competitive. Its superior accuracy score and less generation time indicates that adversarial synthesis still can play a useful role in those situations when the task is more focused on sharpness and speed. Practically, the choice between the two paradigms is in the usage. Stable Diffusion may be better in the case the objective is the closest statistical estimate of a real world portrait data set and a slower sampling rate can be within the computational budget. StyleGAN2-ADA might be more suitable in situations where the objective is to achieve a high-throughput synthesis, or interactive generation, although a comparatively small difference in FID.

It has several deficiencies which should be mentioned. First and most obviously, the numerical outputs shown in this draft are indicative representations of place values as opposed to empirical outputs of the user in the actual training runs. They are included to show how the completed paper can be organized but must not be reported to be quantified results until the models are run and tested. Second, it has been compared with a single domain that is congruent with face photographs hence the findings cannot be directly transferred to other domains such as indoor scene, artwork or medical imaging. Third, Stable Diffusion is prompt-conditioned and StyleGAN2-ADA is unconditioned (notably, with the constant prompt template), an undesirable task-asymmetry. Fourth, FID is sensitive to the preprocessing information itself [25], and future experiments need to include an evaluation script that is locked and, hopefully, testing human preferences.

Another methodological issue is related to the bias in the dataset. FFHQ is obtained based on Flickr photos and thus it embodies demographic, aesthetic and cultural biases of the source distribution [29]. These issues should explicitly be addressed in any final paper, especially when the resulting portraits are to be used in a downstream social or commercial context. The comparison can also be expanded to future work by testing more datasets like CelebA [26] or LSUN [27], and newer diffusion families like transformer-based diffusion models [28].

6. Conclusion

The given manuscript is a full draft of a research-article on the comparison of GANs and Stable Diffusion on a real-world image-generation task on the basis of FFHQ and FID-based evaluation. On the set of illustrative values to fill in the draft, Stable Diffusion has a higher distributional quality compared to StyleGAN2-ADA, has a lower FID and higher recall, although StyleGAN2-ADA is considerably faster and more precise. The resulting conclusion is evident: diffusion models could be better in cases where the image realism and diversity are the primary concerns, whereas GANs are still appealing in terms of low-latency generation. Section 4 illustrative tables and figures must be substituted by real-world outputs of the experiment under the defined methodology in Section 3 before submission.

Data availability and reproducibility note

The draft study design utilises the public FFHQ resources mentioned in [29] and [30]. The exact source of download, date of download, preprocessing script, revision of training codes and evaluation library version under which the actual experiment was done should be noted in the final paper such that the comparison can be reproduced.

References

[1] Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Advances in Neural Information Processing Systems, 2014.

[2] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” arXiv preprint arXiv:1411.1784, 2014.

[3] Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” arXiv preprint arXiv:1511.06434, 2015.

[4] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, “Improved techniques for training GANs,” in Advances in Neural Information Processing Systems, 2016.

[5] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in Proceedings of the International Conference on Machine Learning, 2017.

[6] Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, “Improved training of Wasserstein GANs,” in Advances in Neural Information Processing Systems, 2017.

[7] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “GANs trained by a two time-scale update rule converge to a local Nash equilibrium,” in Advances in Neural Information Processing Systems, 2017.

[8] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, “Spectral normalization for generative adversarial networks,” in International Conference on Learning Representations, 2018.

[9] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of GANs for improved quality, stability, and variation,” in International Conference on Learning Representations, 2018.

[10] Brock, J. Donahue, and K. Simonyan, “Large scale GAN training for high fidelity natural image synthesis,” in International Conference on Learning Representations, 2019.

[11] T. Karras, S. Laine, and T. Aila, “A style-based generator architecture for generative adversarial networks,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.

[12] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and improving the image quality of StyleGAN,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.

[13] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, “Training generative adversarial networks with limited data,” in Advances in Neural Information Processing Systems, 2020.

[14] Ho, A. Jain, and P. Abbeel, “Denoising diffusion probabilistic models,” in Advances in Neural Information Processing Systems, 2020.

[15] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, “Score-based generative modeling through stochastic differential equations,” in International Conference on Learning Representations, 2021.

[16] Song, C. Meng, and S. Ermon, “Denoising diffusion implicit models,” in International Conference on Learning Representations, 2021.

[17] Q. Nichol and P. Dhariwal, “Improved denoising diffusion probabilistic models,” in Proceedings of the International Conference on Machine Learning, 2021.

[18] P. Dhariwal and A. Nichol, “Diffusion models beat GANs on image synthesis,” in Advances in Neural Information Processing Systems, 2021.

[19] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.

[20] Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, I. Sutskever, and M. Chen, “GLIDE: Towards photorealistic image generation and editing with text-guided diffusion models,” in Proceedings of the International Conference on Machine Learning, 2022.

[21] Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. Denton, S. K. S. Ghasemipour, B. Ayan, S. S. Mahdavi, R. G. Lopes, et al., “Photorealistic text-to-image diffusion models with deep language understanding,” in Advances in Neural Information Processing Systems, 2022.

[22] Zhang, A. Rao, and M. Agrawala, “Adding conditional control to text-to-image diffusion models,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.

[23] S. M. Sajjadi, O. Bachem, M. Lucic, O. Bousquet, and S. Gelly, “Assessing generative models via precision and recall,” in Advances in Neural Information Processing Systems, 2018.

[24] T. Kynkäänniemi, T. Karras, S. Laine, J. Lehtinen, and T. Aila, “Improved precision and recall metric for assessing generative models,” in Advances in Neural Information Processing Systems, 2019.

[25] G. Parmar, R. Zhang, and J.-Y. Zhu, “On aliased resizing and surprising subtleties in GAN evaluation,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.

[26] Z. Liu, P. Luo, X. Wang, and X. Tang, “Deep learning face attributes in the wild,” in Proceedings of the IEEE International Conference on Computer Vision, 2015.

[27] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, “LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop,” arXiv preprint arXiv:1506.03365, 2015.

[28] W. Peebles and S. Xie, “Scalable diffusion models with transformers,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.

[29] NVLabs, “Flickr-Faces-HQ Dataset (FFHQ),” GitHub repository. [Online]. Available: https://github.com/NVlabs/ffhq-dataset

[30] Kaggle, “Flickr-Faces-HQ Dataset (FFHQ),” dataset mirror. [Online]. Available: https://www.kaggle.com/datasets/arnaud58/flickrfaceshq-dataset-ffhq