Indian Journal of Artificial Intelligence and Neural Networking (IJAINN)
ISSN: 2582-7626 (Online), Volume-2 Issue-6, October 2022

GAN-Generated Terrain for Game Assets

Yogendra Sisodia

Abstract: ***Multimedia applications, such as virtual reality models and video games, are increasingly interested in the ability to generate and author realistic virtual terrain automatically. In this paper, the author proposes a pipeline for a realistic two-dimensional terrain authoring framework that is powered by several different generative models that are applied one after the other. Two-dimensional role-playing games will benefit from this ability to create multiple high-resolution terrain variants from a single input image and to interpolate between terrains while keeping the terrains that are generated close to how the data is distributed in the real world.***

Keywords: ***Deep Learning, Generative Adversarial Networks, Pix2Pix, Procedural Content Generation, Terrain Generation.***

I. INTRODUCTION

The applications of virtual terrain rendering in graphics and computer vision applications can all benefit from accurate representations of real-world topography, which is the goal of terrain modelling. Multimedia applications, such as Virtual Reality models and gaming, are increasingly interested in the automated generation and authoring of realistic virtual terrain. Real-world landscapes are subject to a wide variety of weathering, erosion, and landslide-related natural changes. This results in a wide variety of landforms, from mountains to valleys. Some of these major and distinctive land features are visible at various scales. All these factors combine to make terrain generation and authoring a difficult task. New developments in computer vision for deep learning, especially Generative Adversarial Networks (GAN) [1], have made it possible to learn different terrain features for tasks like changing the terrain, etc. There isn't much written about deep learning-based automated terrain authoring for games. The author's main contribution is to use GANs like Pix2Pix [2] and Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) [3] together to make a nearly infinite number of terrains that can be used in 2-D games.

II. LITERATURE REVIEW

A. Generative Adversarial Networks

The GAN paper suggested a new way to estimate generative models using a process called "adversarial." In this process, we train two competing models at the same time: a generative model G that identifies the distribution of the data and a discriminative model D that estimates the opportunity that a sample came from the training data instead of G.

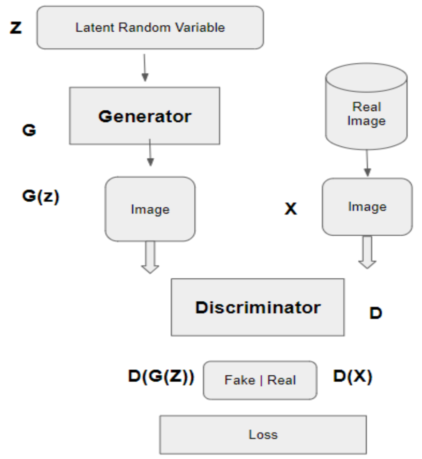

Fig. 1. GAN Architecture

$$V(G, D) = E_{x\sim P\_data}[\log D(x)] + E_{x\sim P\_G}[\log(1-D(x))] \qquad (1)$$

The goal of G's training is to increase the chances that D will make a mistake as given in Value Function (1), so the discriminator is trying to maximize its reward while the generator is trying to make the discriminator's reward as small as possible. This framework is the same as a two-player minimax game. Backpropagation can be used to train the whole system when G and D are both made up of deep multilayer perceptron's.

B. Pix2Pix

Pix2pix is a conditional generative adversarial network (cGAN) that learns a mapping from input images to output images. Pix2pix is not limited to one type of application; it can be used for a wide range of tasks, such as making photos from labelled maps, adding color to black-and-white photos, turning map photos into aerial images, and even turning sketches (Edges) into realistic photos. The idea is to create a pair of sketches and real images for training and to produce realistic images from sketches given as input.

Fig. 2. Pix2Pix for Edges / Sketches

Manuscript received on 25 September 2022 | Revised Manuscript received on 28 September 2022 | Manuscript Accepted on 15 October 2022 | Manuscript published on 30 October 2022.
***** Correspondence Author (s)
Yogendra Sisodia*, Director, Department of Machine Learning, Conga, Thane (Maharashtra), India. E-mail: ysisodia@conga.com, scholarly360@gmail.com

Retrieval Number: 100.1/ijainn.F1060102622
DOI:10.54105/ijainn.F1060.102622
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

C. Esrgan

The Super-Resolution Generative Adversarial Network (SRGAN) [4] is a landmark work that can make realistic textures when super-resolving a single image. Image super-resolution (SR) techniques, in a few words, reconstruct a higher-resolution (HR) image or sequence from a series of lower-resolution (LR) images. For example, a 720p image can be upscaled to 1080p using SR techniques.

One common way to solve this problem is to use deep convolutional neural networks that can take LR images and make them into HD ones. ESRGAN is an SRResNet-based deep neural architecture with residual-in-residual blocks and a mix of context, perceptual, and adversarial losses. Contextual and perceptual losses are used for correct image upscaling, while adversarial loss pushes the deep neural network to the precise image manifold using a discriminator network that has been trained to tell the difference between the super-resolved images and the original photo-realistic images used for training.

Fig. 3. ESRGAN

D. Terrain Generation

There have been attempts, using cGAN [5], to create realistic virtual terrains from hand-drawn user inputs based on a large set of real-world terrain data. The author's work is different because he focuses more on games and making a variety of two-dimensional images with high resolution.

III. METHODOLOGY

The author used a popular Digital Elevation Model (DEM) dataset. This DEM dataset is part of high-resolution DEMs of mountain ranges called the Pyrenees and the Tyrol [6] that are available to the public. As ground truth figures for GAN, DEM patches with 2m/pixel pixels have been used. The DEM tiles were broken up into pieces that were 200x200 pixels. The experiment was set up with 2,101 pictures from the Pyrenees, which were split into 80% train, eval, and holdout for Pix2Pix.

A. Image Pre-processing

The author used Gaussian-Blur and Gray scaling in OpenCV [7] for Sketch Generation. The author further used the eraser synthesiser by drawing random circles within the sketched image. The Eraser Synthesizer is configurable, and one may use as many combinations as possible during inference. The pipeline is shown in Fig.4.

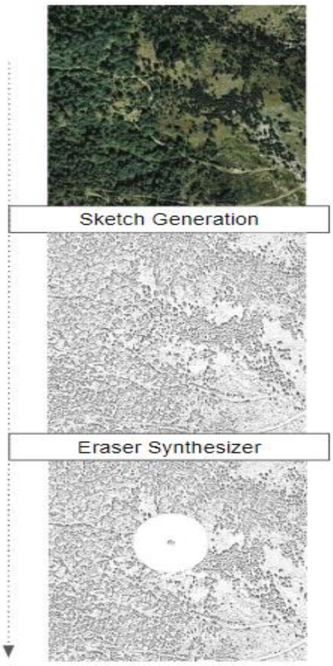

Fig. 4. Image Preprocessing

B. Training Pipeline

Fine-tuning is done on the image pairs (Sketched Erased Image and Original) using the Facades Pix2Pix model trained on [8]. The Author's Fine-Tuned till 200 epochs. The generator has a U-Net-based architecture with 54.414 million Parameters. The discriminator is represented by a convolutional PatchGAN classifier with 2.769 million Parameters.

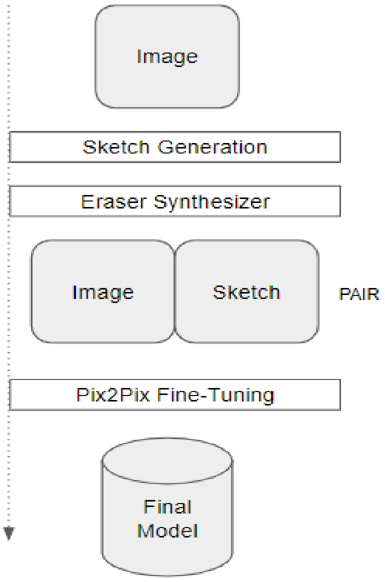

Fig. 5. Training Pipeline

Retrieval Number: 100.1/ijainn.F1060102622
DOI:10.54105/ijainn.F1060.102622
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)

C. Prediction Pipeline

While predicting, the author can generate as many combinations of sketches as possible using the Eraser Synthesizer. These images are then fed into Pix2Pix for prediction and further into ESRGAN to produce a high-quality image.

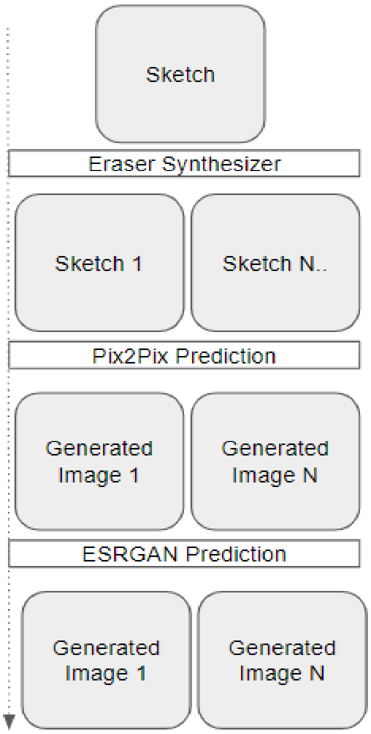

Fig. 6. Inference Pipeline

IV. RESULT AND DISCUSSION

For GAN, the interpretation of results and loss functions is quite difficult. The best way is to see the real generated data.

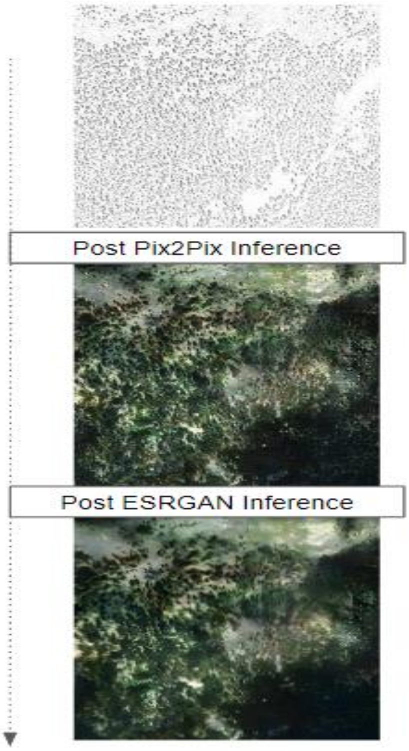

Fig. 7. Inference Pipeline

The author was able to generate high-resolution images or terrains using the input sketches. Since the pipeline includes the Eraser Synthesizer theoretically there could be an infinite number of images that can be generated. All results and enhancements are available here: https://github.com/scholarly360/GanGeneratedTerrain

V. CONCLUSION

The work can be used to generate terrain for 2-D role-playing games, and it can be fine-tuned to generate sprites and characters as well as terrain. The author believes that this will reduce the time of game asset generators. Furthermore, in future work, the author wants to move to 3-D asset generation.

REFERENCES

1. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, DavidWarde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. “Generative adversarial networks”, 2014. Available: https://arxiv.org/abs/1406.2661.
2. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. “Image-to-Image Translation with Conditional Adversarial Networks“, 2017.Available: https://arxiv.org/abs/1611.07004. [CrossRef]
3. Xintao Wang, Liangbin Xie, Chao Dong, Ying Shan. “Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data “, 2021.Available: https://arxiv.org/abs/2107.10833.
4. Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi."Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, 2017.Available: https://arxiv.org/abs/1609.04802.
5. Eric Gu´erin, Julie Digne, Eric Galin, Adrien Peytavie, Christian Wolf, Bedrich Benes, and Benoit Martinez. Interactive example-based terrain authoring with conditional generative adversarial networks. ACM Transactions on Graphics (proceedings of Siggraph Asia 2017), 36(6), 2017 [CrossRef]
6. Oscar Argudo, Antoni Chica, and Carlos Andujar. Terrain super-resolution through aerial imagery and fully convolutional networks. In Computer Graphics Forum, volume 37, pages 101–110. Wiley Online Library, 2018. [CrossRef]
7. Smoothing Images, OpenCV Documentation [Online]. Available: https://docs.opencv.org/4.x/d4/d13/tutorial_py_filtering.html
8. Facades Label2Photo PyTorch Pre Trained Pix2Pix Model [Online]. Available: http://efrosgans.eecs.berkeley.edu/pix2pix/models-pytorch/facades_label2photo.pth

AUTHORS PROFILE

Yogendra Sisodia is Director, Machine Learning at Conga-India. With an MCA from BIT-Mesra, Yogendra has over 15 years of experience in solving complex real-world and big data problems through machine learning. Previously, as an entrepreneur, he worked on various SAAS products in the Fin-Tech, Mar-Tech, and Sales-Tech domains, but he is now focusing on the Legal-Tech domain, where he is empowering attorneys to save time by automating machine-driven annotations for them accurately and promptly. Some areas of research that are interesting right now are semi-supervised learning, adversarial machine learning, large legal language models, and deep learning applications in computer vision and natural language processing.

Retrieval Number: 100.1/ijainn.F1060102622
DOI:10.54105/ijainn.F1060.102622
Journal Website: www.ijainn.latticescipub.com

Published By:
Lattice Science Publication (LSP)