Proposal: Exploring Methods of Procedural Terrain Generation

Laura Hulley ID: 201277571

12/11/2020

0.1 Project Description

0.1.1 Procedural Terrain Generation

Procedural Generation has been an important part of video games throughout their existence. An early example of such being the entirely ASCII-based 1980 game, Rogue [1]. The level maps, treasure and monsters are procedurally generated each time the game is played. Rogue inspired future games and spawned the category of Rogue-like games which now contains hundreds, if not thousands,[1] of procedurally generated games.

In more recent history, similar techniques are applied in games such as No Man's Sky (2016) [4] to create 18 quintillion unique planets that the user can explore. These are generated from a single 64bit source number and are therefore the same for each user. This technique of generation saves designers the impossible task of designing 18 quintillion separate planets and adds an element of limitlessness and indefiniteness for the user. While procedural generation is used here to cut down design time, it also massively reduces the amount of memory used by the game at any one time. The program can generate the planets when necessary and does not need to store any data other than the area of a planet currently visible to a player[2].

0.1.2 Algorithmic techniques

There are a variety of terrain generation techniques used in video games. One traditional technique still in use today is midpoint displacement, specifically the diamond square algorithm. This technique relies on fractal algorithms to replicate the fractal properties of mountain ranges. There are a variety of other techniques that use noise such as perlin noise [5] to add an element of natural randomness to terrain. One main downfall of these techniques is that landscapes often look too repetitive or too random to the natural eye. In this project, I will explore a modern approach to these issues using Generative Adversarial Networks (GAN).

0.1.3 Generative Adversarial Networks

GANs consist of two competitive networks: A generative network and an adversarial network. The adversarial network is trained (in this case) on real world terrain data an aims to classify whether an input is an example of real satellite imagery and elevation or a fake. Alongside the training data, the adversarial network also receives data from the generator network. This data is initially random but over time, the generator "learns" and which data can "fool" the adversarial network into thinking it is data from the original set.

Each time the adversarial network falsely classifies the generator's output as real, it is marked down. If the generator produces an output that doesn't "fool" the adversarial network, it is marked down. This technique means the generative and adversarial networks improve their ability to produce realistic

[1] Figure estimated from Steam search results [2] and the following list [3]. There is also the assumption that there are unpublished video games developed using rogue-like techniques.

[2] This does not include the many other elements saved such as the user's inventory and changes made by the user to the planet.

terrain and to identify which terrains are fake (produced by the generative network). Therefore, improving the overall realism of the terrain outputted by the Generator.

0.1.4 Terrain Generation using GANs

The adversarial network of a GAN is trained on real world data. Therefore, with this technique, this project aims to discuss whether GANs can solve the problem of unrealistic terrain generated by more traditional fractal and noise techniques. GANs also come with their own negatives, long training times and large processing power to name a few. These will all form a comparison between techniques to better understand the future of terrain generation in video games.

0.2 Aims and Objectives

0.2.1 Aims

There are two main aims of this project which are both dependent on each other. They must therefore both be achieved for completion of the intended project.

1. Demonstrate the ability of a GAN to produce realistic and efficient[3] terrain elevation and rendering.
2. Form a discussion that compares the effectiveness of a GAN technique to more traditional algorithmic techniques.

0.2.2 Objectives

The objectives below are measurable statements defining the necessary steps to be taken if the above aims are to be achieved. Since there are two aims, the objectives are split accordingly. The tests for each objective are described in more detail in section 0.6.

Aim 1

a. Source satellite image and elevation data of the planet that is legally available for educational use.

b. Source a sufficient amount [4] of training data for the GAN to enable a fair comparison. The definition of a sufficient amount of training data for the discriminator would need to be defined in later stages of the project after initially testing the success of the model.

c. Define a measure of success for the GAN, using the methods laid out in section 0.6.1-Aim 1-c.

[3]The definition of what constitutes realistic and efficient in this project is described in section 0.6

[4]In this case, sufficient could be described as any value (although likely above a certain threshold). The amount of data needed is dependent on a variety of factors including - but not limited to - the model architecture, data quality, desired accuracy and complexity of the subject.

d. Develop a GAN with the aim of producing results that are indistinguishable by humans from those in the original data set. This objective is a desirable requirement of the model developed, however the closer the model is to reaching this standard, the better we are able to argue the merits of this approach. The process of developing a GAN will also increase my personal understanding of the subject and allow a more informed discussion.

Aim 2

The second aim is harder to write objectively measurable objectives for, as instead of developing a system, the aim is to form a comparative analysis. The quality of this style of writing is a much more subjective interpretation and is largely down to the view of the reader.

The objectives for this aim are informed by an article by Dr Walk of Harvard University [6]. She states that "To write a good compare-and-contrast paper, you must take your raw data — the similarities and differences you've observed — and make them cohere into a meaningful argument." This forms the basis of my first two objectives. The remaining objectives relate to the points outlined in the article which will further reinforce the first two objectives.

a. Outline the main similarities and differences between the use of GANs and traditional approaches to terrain generation.

b. Form a meaningful argument. As this project is an investigation into how one technique compares to another, there is no bias to have a definite 'winner'. To be classified as a meaningful argument, in line with the aforementioned article, this objective would be satisfied by completion of objectives c-f.

c. Frame of reference: Define the context in which these techniques are being measured. For example, in the intended context of game design, it is worth noting why we would place more weight on the efficiency of the program at the expense of realism.

d. Grounds for comparison: Explain the rationale behind the choice of terrain generation techniques used as examples in the project.

e. Clearly define the relationship between the techniques: "Do they extend, corroborate, complicate, contradict, correct, or debate one another?" [6]

f. Form the comparison around a clear organisational scheme.

0.3 Key Literature

This section elaborates on the descriptions set out in section 0.1.

0.3.1 Generative Adversarial Networks

The GAN model is a fairly recent development in AI that was first proposed in a 2014 paper by Ian Goodfellow et. al. [7]. GANs have revolutionised generation and creation in a variety of areas; from creating art in the style of Van Gogh [8] to generating faces of people who do not exist[9] [5]. While the aforementioned approaches seem lighthearted, they showcase the power of GANs to fool even the human eye.

Figure 1: Shows the results of satellite river generation in [10]

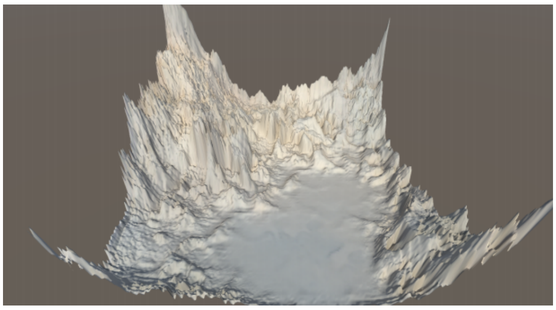

Figure 2: Shows the resulting terrain map from the work in [11]

Evidence shows the success of GANs in various aspects of terrain generation. Figure 1 shows the results of a PGGAN (progressive GAN) trained on satellite images of US rivers [10]. The success of this study opens up the argument for the use of GANs for satellite image generation. As for elevation, a study by Beckham et.al.[11] shows similar success generating terrain height maps using NASA elevation data, as seen in Figure 2.

[5] I highly recommend looking at this site if you haven't seen it before: https://thispersondoesnotexist.com/ Every page refresh creates a new person.

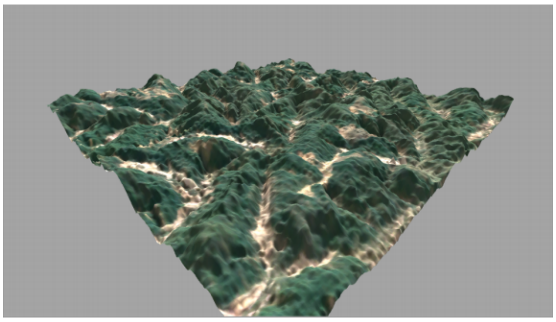

Figure 3: Shows results of GAN terrain generation [12]

Combining the success of these approaches, in this project I will aim to produce a satellite image with an associated height map using a GAN. This approach has been used before [12] to varying degrees of success (Figure 3). I intend to build upon this approach by using StyleGAN [13] where they used a spatial GAN [14]. The motivation behind the choice of StyleGAN is detailed in section 0.3.2.

0.3.2 StyleGAN

The original GAN model is limited in that it cannot provide control over individual features in the data, e.g freckles or glasses. It also falls short in that, like many ML models, it is largely governed by the data provided. This means that often, common themes in the data will become a standard of the output. For example, in a data set of satellite images of the planet - the model (depending on it's specific architecture) may only produce images of the sea.

StyleGAN introduces the concept of "styles". This allows the generator to form it's result using both high and low level features (e.g. face shape and eye colour respectively) of an image. Noise is also added to the image after each convolutional layer. This helps ensure there is some deviation between the generated images and you would not have results such as the satellite images of the sea mentioned earlier. There is evidence of success of the use of StyleGAN on a variety of data sets [15] Figure 4.

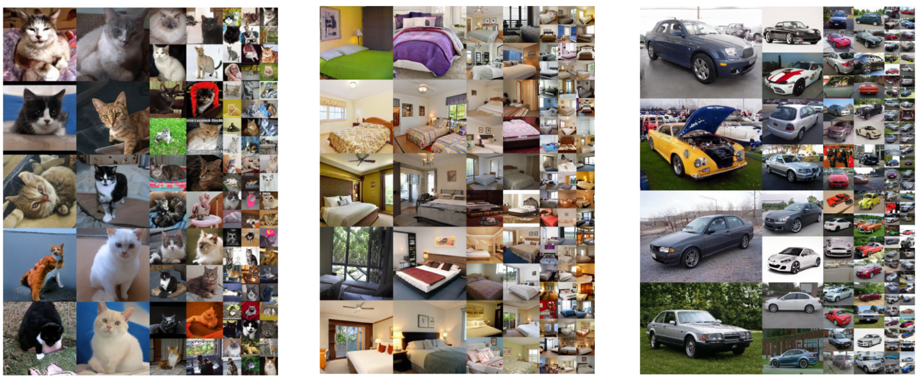

Figure 4: Shows the results of [14] on various datasets

0.4 Development Process

The development process for this project will largely be a rapid prototyping approach. GANs are an area I have only recently been able to study in depth and I am planning to learn more as I build and test models. The models and data can then be fine tuned. More information on my process can be seen in my Gantt chart (section 0.10).

GitHub will be used for version control. The StyleGAN architecture is written with Tensorflow and I have already installed this and tested it on my current python environment. All work will be done in a virtual environment to allow it to be run easily on another PC.

0.5 Data Sources

To satisfy the aims, the training data will need to be a combination of satellite images and elevation data. The most appropriate data sets currently seem to be NASA Global Imagery Browse Services [16] and Shuttle Radar Topography Mission (SRTM) [17] respectively. Both of these data sets are owned by NASA who support an open data policy. This means the data is available for educational use providing it is correctly acknowledged.

These may not be the final data sets used in the project as better ones may become clear with more research. However, any new data set should also comply to being available for free, educational use.

When manipulating the data for training, the current method intended is to add the elevation data to the RGB data of the image. Each image then has the associated elevation data with each pixel (or similar) and can be used to create a 3d visualisation of the terrain.

0.6 Testing and Evaluation

The following tests are split by the aims and objectives they relate to.

0.6.1 Aim 1 - GAN

a. The ethical and legal considerations surrounding the data are described in section 0.7. To satisfy this objective, the data should satisfy the data usage guidelines described there.

b. As mentioned in this objective, the appropriateness of the data can only truly be evaluated after the model architecture has been designed. However, there are steps that can be taken to ensure the choice of data set gives the largest chance of success. These are outlined in section 0.5.

c. As the aim of this project is a comparison, the GAN does not need to be state-of-the-art to achieve this. Indeed, conclusions, comparisons and a proof of concept could even be drawn from the results of a GAN that produces what are considered unrealistic results. Therefore, regardless of the quality of the generated data, valid conclusions will still be drawn.

With this said, metrics to measure the success of the GAN in various aspects are still necessary to form a meaningful comparison. Some of the tests would need to be the same for both the GAN and the opposing algorithm to ensure a fair comparison. These tests could be carried out by myself and/or results gathered from researching the results of similar studies. The evaluation of the results against the following aspects would provide information for a conclusion. However, more informative metrics may become apparent throughout the project alongside those described below.

- Realism. Video games are intended for humans. To test the realism of the results, human verification would be most appropriate. The proposed test would involve presenting people with two images, one generated from the data set and the other by the GAN. They would then be asked which one is "real". Since this method of testing involves human input, ethical guidelines will need to be put in place (section 0.7.3).
- Efficiency. The definite measures for this can only be defined after further research into the context of video game design. Aspects to measure will include the speed, processing power required and memory required.
- Scale of Randomness/ Repetitiveness. This concerns questions such as: Are all the generated results very similar, is there a degree of repetition in the terrain, are there a variety of different terrain styles (e.g mountains, flat lands) within one terrain? The desired outcome to these questions can chance depending on the specific context of the intended game. Therefore, further research is needed in the project to define good parameters for these points. More research will also need to be done into methods of measuring the repetitiveness of terrain.
- Suitability in Context. Mentioned briefly above, this concerns measuring how suitable a GAN would be for terrain generation in video games specifically. If the GAN produces incredibly realistic results but needs an unrealistic amount of processing power - it would not be suitable. Finding the limits of what is suitable will involve research into existing games to define metrics that the GAN should adhere to to be considered a success.

There is currently a lack of quantitative methods of GAN assessment [18]. Further research would be required to complete the assessment methods presented in the aforementioned article. However, assessments of this kind could prove invaluable in selecting the most appropriate model architecture for the data and are therefore worth the time to investigate.

d. This objective simply states that in the course of this project, a GAN should be developed with the aim of generating virtual terrains. Regardless of the resulting data, the completion of this objective would consist of:

   - A GAN architecture consisting of generative and discriminative networks.

- The discriminative network should be trained on relevant data.
- The generative network should should produce some form of output.

0.6.2 Aim 2 - Discussion

a. To best evaluate the similarities and differences between the techniques, similar tests should be run on both. The tests outlined in 0.6.1.c should also be applied to existing techniques.

b. To mark this objective as completed, tests c - f should also be completed. Despite this, even with all the below tests met, whether the conclusion is considered meaningful is down to the reader.

c. Clearly define a context, likely game design in this case. Reference when the techniques proposed are used.

d. This is linked heavily to the test above. When understanding the context, it should become clear why the chosen methods are used.

e. and f. The relationship between the two techniques could be found by going through the important points methodically and comparing them at each point. Another approach would be to explain each approach in detail and define their relationship in a later section. These would be examples of point-by-point or text-by-text from the referenced article [6].

0.7 Ethical Considerations

0.7.1 Job Replacement

It's important to consider the consequences to human jobs when proposing any aspect of job automation. While terrain generation is already a largely automated task, there are still aspects which often involve creative human input, e.g designing land cover styles. Game design is a highly creative role and is unlikely to be automated in the next decades [19]. Evidence shows that game designers can instead use automation to minimise mundane tasks and allow their time to be spent more efficiently. An example of this is Assassin's Creed Unity where designers used city blocks and styles to speed up the process of designing Paris [20].

0.7.2 Data Ownership and privacy

The data sets mentioned in section 0.5 are all available for educational use. The handling of personal privacy is already done by them as any private areas are blurred/ not available and no data would need to close enough that you could identify a person. If at any point in the project other data sets are used, they must also comply with data ownership and privacy regulations.

0.7.3 Human Feedback

As mentioned in the test section, human feedback could be used to evaluate the realism of the results of the GAN. Due to the current lockdown, to do this safely this would be carried out virtually. There is no intent to capture any data relating to the tester or any information that could identify them. Anybody who tests the results will be made completely aware of the purpose of the data and the project.

0.8 BCS Project Criteria

An ability to apply practical and analytical skills gained during the degree programme.

This project shows the application of data science, AI, programming and algorithmic skills learnt throughout my degree.

Innovation and/or creativity.

In my research I have only found one other example of GANs used to generate elevation and satellite data [12]. I am proposing this same technique but using StyleGAN - a purpose for which I can find no evidence of it being used.

Synthesis of information, ideas and practices to provide a quality solution together with an evaluation of that solution.

This is will be evidenced heavily by my project as it is focused around the evaluation of solutions to the given problem of terrain generation.

That your project meets a real need in a wider context

The target context of this project is the game design community. A specific need for this project is defined in section 0.1.4.

An ability to self-manage a significant piece of work.

This will be shown on completion and throughout the project. The majority of deadlines are self set.

Critical self-evaluation of the process

This occurs throughout the project. It will be especially seen when working on the GAN as it will be modified based on previous results. Critical evaluation is also necessary when writing the comparison as it ensures the work is of a good quality and is coherent.

0.9 Risk and Contingency Plans

Risks	Contingencies	Likelihood	Impact
Hardware failure	Since we have access to the library, there are a lot of back up computers available for me to use. It's very unlikely that I would lose access to any computer and in that case, my discussion would have to revolve around existing work instead.	very low	medium
Software Failure	Since the main focus of this project is a discussion, the GAN and any related software should be completed early on allowing lots of time for errors, should they occur.	low	medium
Running out of time	The plan for my project has multiple set deadlines will which ensure I keep on top of timings in the project. Should I run over, my notes throughout the project should form a conclusion, albeit weaker than the original one could have been.	low	high
Programming problems	I have experience in Tensorflow and python but no experience in GANs specifically. Despite this, there is a lot of documentation available so I should be able to develop a GAN to the standards outlined in the test section.	medium	medium
Limited processing power	After reviewing the processing power requirements of other GANs and reading the recommended 8 GPUs for StyleGAN, there is some worry that my laptop will not have enough processing power to deliver quality results. As mentioned in section 0.6.1.c this is a low impact as it does not have to produce state-of-the-art results to form a meaningful discussion.	high	low

References

[1] M. Toy and G. Wichman, "Rogue," freely-distributed executable, 1980.

[2] V. Corporation, "Browsing roguelike," 2020, 125 results. [Online]. Available: https://store.steampowered.com/tags/en/Roguelike/

[3] community written, "List of roguelikes by theme," 2020. [Online]. Available: http://www.roguebasin.com/index.php?title=List_of_Roguelikes_by_Theme

[4] H. Games, "No man's sky," 2016.

[5] K. Perlin, "An image synthesizer," in *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '85. New York, NY, USA: Association for Computing Machinery, 1985, p. 287–296. [Online]. Available: https://doi.org/10.1145/325334.325247

[6] K. Walk, "How to write a comparative analysis," 1998, last accessed 9 November 2020. [Online]. Available: https://writingcenter.fas.harvard.edu/pages/how-write-comparative-analysis

[7] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[8] A. D. et.al., "Gangogh," https://github.com/rkjones4/GANGogh, 2017.

[9] P. Wang, "This person does not exist," 2019. [Online]. Available: https://thispersondoesnotexist.com/

[10] A. Gautam, M. Sit, and I. Demir, "Realistic river image synthesis using deep generative adversarial networks," 2020.

[11] C. Beckham and C. Pal, "A step towards procedural terrain generation with gans," 2017.

[12] r. r. spick and j. walker, "Realistic and textured terrain generation using gans," ser. CVMP '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3359998.3369407

[13] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," 2019.

[14] R. J. Spick, P. Cowling, and J. A. Walker, "Procedural generation using spatial gans for region-specific learning of elevation data," in *2019 IEEE Conference on Games (CoG)*, 2019, pp. 1–8.

[15] T. Karras, "stylegan," https://github.com/NVlabs/stylegan/blob/master/README.md, 2019.

[16] NASA, "Global imagery browse services," ongoing. [Online]. Available: https://earthdata.nasa.gov/eosdis/science-system-description/eosdis-components/gibs

[17] ——, "Usgs eros archive - digital elevation - shuttle radar topography mission (srtm) 1 arc-second global," ongoing. [Online]. Available: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects

[18] D. J. Im, H. Ma, G. Taylor, and K. Branson, "Quantitatively evaluating gans with divergences proposed for training," 2018.

[19] C. B. Frey and M. A. Osborne, "The future of employment: How susceptible are jobs to computerisation?" *Technological forecasting and social change*, vol. 114, pp. 254–280, 2017.

[20] IGN, "Making assassin's creed unity: Part 2 - next generation technology," 2014, youtube video - 5:05 onwards. [Online]. Available: https://www.youtube.com/watch?v=QGA0WZLp_08&ab_channel=IGN