Illuminating the Space of Beatable Lode Runner Levels Produced By Various Generative Adversarial Networks

Kirby Steckel
Southwestern University
Georgetown, Texas, USA
steckelk@southwestern.edu

Jacob Schrum
Southwestern University
Georgetown, Texas, USA
schrum2@southwestern.edu

ABSTRACT

Generative Adversarial Networks (GANs) are capable of generating convincing imitations of elements from a training set, but the distribution of elements in the training set affects the difficulty of properly training the GAN and the quality of the outputs it produces. This paper looks at six different GANs trained on different subsets of data from the game Lode Runner. The quality diversity algorithm MAP-Elites was used to explore the set of quality levels that could be produced by each GAN, where quality was defined as being beatable and having the longest solution path possible. Interestingly, a GAN trained on only 20 levels generated the largest set of diverse beatable levels while a GAN trained on 150 levels generated the smallest set of diverse beatable levels, thus challenging the notion that more data is always better when training GANs.

CCS CONCEPTS

• **Computing methodologies** → **Neural networks**; **Generative and developmental approaches**; **Learning latent representations**; **Genetic algorithms**.

KEYWORDS

Lode Runner, Generative Adversarial Networks, Quality Diversity

ACM Reference Format:
Kirby Steckel and Jacob Schrum. 2021. Illuminating the Space of Beatable Lode Runner Levels Produced By Various Generative Adversarial Networks. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille, France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459440

1 INTRODUCTION

Generative Adversarial Networks are a type of neural network trained in an unsupervised way to produce imitations of elements in a training set. Breakthrough work by Volz et al. [5] demonstrated the success of latent variable evolution for a GAN trained to produce Mario levels. However, the size and composition of the training set influences the quality and diversity of the results from the GAN.

This paper focuses on the challenge of producing beatable Lode Runner levels using GANs, and shows how the size and composition of the training set influences the results. A full-length version of this paper is available online [2][1].

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459440

2 LODE RUNNER

The game features 150 levels, which are available as part of the Video Game Level Corpus (VGLC [3]). Lode Runner is a game where the player traverses platforms and ropes, climbs ladders, and falls off cliffs. The goal of the game is to collect all of the treasure in each level while avoiding enemies.

A previous approach to generating Lode Runner levels used standard and variational autoencoders [4]. Our research uses GANs.

3 APPROACH

The challenge addressed in this paper is the generation of quality beatable Lode Runner levels. First, data from the original game is collected to train several GANs. Latent vectors are evolved for each GAN using MAP-Elites [1], which collects a diversity of quality levels and measures quality using A* search. Fitness is the maximum of the A* solution path length and connectivity; longer solution paths are better than those with shorter or no solution path.

The GANs associated with each training set are called `On5Levels`, `On20Levels`, `On50Levels`, `On100Levels`, and `On150Levels` respectively. Also, `WordsPresent` is trained on 13 levels that depict words.

4 RESULTS

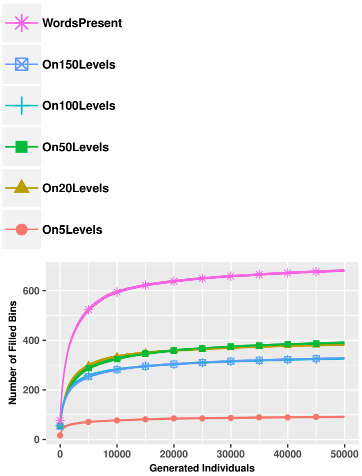
Performance of each GAN is depicted in Fig. 1. The `On20Levels` GAN ended with the highest percentage of beatable levels followed by the `On5Levels` GAN, but `On20Levels` had the highest *number* of beatable levels as well. The `On150Levels` and `WordsPresent` GANs generated the smallest percentage of beatable levels, which could be due to the complexity of the levels in the training sets. However, even though the percentage of beatable `WordsPresent` levels is small, it filled the most bins.

Generated levels mainly fall into three categories: copied levels, merged levels, and levels that are unstructured and chaotic[2].

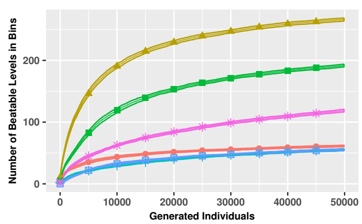
5 CONCLUSION

This paper shows that GANs can generate Lode Runner levels using levels from the original game. GANs trained with small to moderate sized subsets of data produced a greater diversity of levels as well as more beatable levels than GANs trained with larger data sets. Using as much data as possible seems preferable when training a GAN, but these results indicate that others may want to reconsider how much available data they use with these methods.

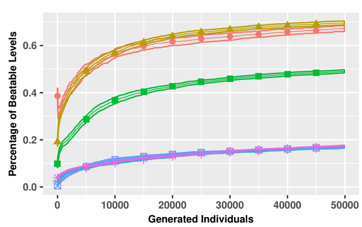
[1] https://arxiv.org/abs/2101.07868
[2] southwestern.edu/~schrum2/SCOPE/loderunner.php



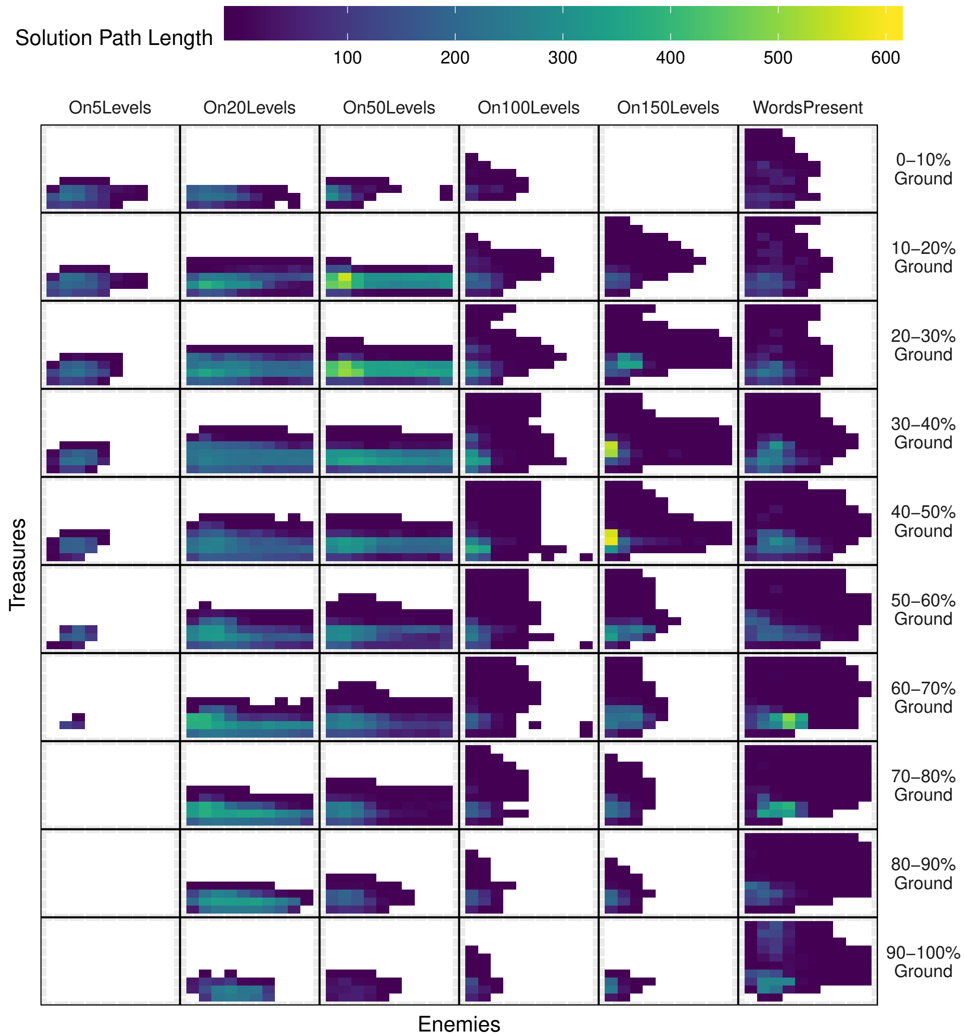
(a) Occupied Bins



(b) Bins With Beatable Levels



(c) Percent Beatable



(d) Average MAP-Elites Fitness Heat Maps For Each GAN

Figure 1: Average Results Across 30 Runs of MAP-Elites. (a) Average number of archive bins occupied by a level. (b) Average number of levels that are beatable according to A*. (c) Percentage of levels that are beatable. The 95% confidence intervals are shown for each line, but are very thin. `WordsPresent` fills the most bins, but `On20Levels` produces the most beatable levels. `On5Levels` produces a comparable percentage of beatable levels to `On20Levels`, but both the total number of beatable levels and total number of occupied bins is very small. (d) Heat map of average archive. Each column averages fitness values across MAP Elites runs using a particular GAN. Rows correspond to the percentage of the level occupied by ground tiles. Each large grid cell is further divided into a grid based on treasure count and enemy count (treasure count increases from bottom to top, enemy count increases left to right). White space lacks any evolved levels, and dark purple corresponds to unbeatable levels. `WordsPresent` best covers the space of possible levels, but primarily with unbeatable levels. `On5Levels` has trouble covering the space. The bright colors for `On20Levels` and `On50Levels` indicate a large number of beatable levels.

ACKNOWLEDGMENTS

Research made possible by the SCOPE program for undergraduate research at Southwestern University.

REFERENCES

[1] Jean-Baptiste Mouret and Jeff Clune. 2015. Illuminating Search Spaces by Mapping Elites. *arXiv:1504.04909* (2015).
[2] Kirby Steckel and Jacob Schrum. 2021. Illuminating the Space of Beatable Lode Runner Levels Produced By Various Generative Adversarial Networks. arXiv:2101.07868
[3] Adam James Summerville, Sam Snodgrass, Michael Mateas, and Santiago Ontañón. 2016. The VGLC: The Video Game Level Corpus. In *Procedural Content Generation in Games*. ACM.
[4] Sarjak Thakkar, Changxing Cao, Lifan Wang, Tae Jong Choi, and Julian Togelius. 2019. Autoencoder and Evolutionary Algorithm for Level Generation in Lode Runner. In *Conference on Games*. IEEE, 1–4.
[5] Vanessa Volz, Jacob Schrum, Jialin Liu, Simon M. Lucas, Adam M. Smith, and Sebastian Risi. 2018. Evolving Mario Levels in the Latent Space of a Deep Convolutional Generative Adversarial Network. In *Genetic and Evolutionary Computation Conference* (Kyoto, Japan). ACM.