

TOAD-GAN: Coherent Style Level Generation from a Single Example

Maren Awiszus, Frederik Schubert, Bodo Rosenhahn

Institut für Informationsverarbeitung
 Leibniz University
 Hanover, Germany

{awiszus, schubert, rosenhahn}@tnt.uni-hannover.de

Abstract

In this work, we present **TOAD-GAN** (Token-based One-shot Arbitrary Dimension Generative Adversarial Network), a novel Procedural Content Generation (PCG) algorithm that generates token-based video game levels. TOAD-GAN follows the SinGAN architecture and can be trained using only one example. We demonstrate its application for *Super Mario Bros.* levels and are able to generate new levels of similar style in arbitrary sizes. We achieve state-of-the-art results in modeling the patterns of the training level and provide a comparison with different baselines under several metrics. Additionally, we present an extension of the method that allows the user to control the generation process of certain token structures to ensure a coherent global level layout. We provide this tool to the community to spur further research by publishing our source code.

1 Introduction

Level design is a key component of the game creation process. The designer has to consider aspects like physics, playability and difficulty during the creation process. This task can require a lot of time for even a single level.

Procedural Content Generation (PCG) has the potential to assist the designer by automating parts of the process. Early works in PCG used the co-occurrence of tokens (e.g. a single enemy or ground block) in existing game levels to identify patterns (Dahlskog and Togelius 2012) and combined them using simple statistical models (Snodgrass and Ontanón 2013). The quality of these algorithms critically depends on the extracted patterns and co-occurrence relations, which have to be defined manually. Recent approaches used PCG via Machine Learning (PCGML) (Summerville et al. 2018) to learn the patterns and relations from the data automatically (Summerville and Mateas 2016; Volz et al. 2020). However, simply applying them to the level generation task comes with several drawbacks. The Machine Learning algorithms need many examples to extract the patterns. As manual level design is a costly process, there is usually a very limited number available for training. Even if the amount is sufficient, the generated levels are mixtures of all example levels and do not have a coherent style. Finally,

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

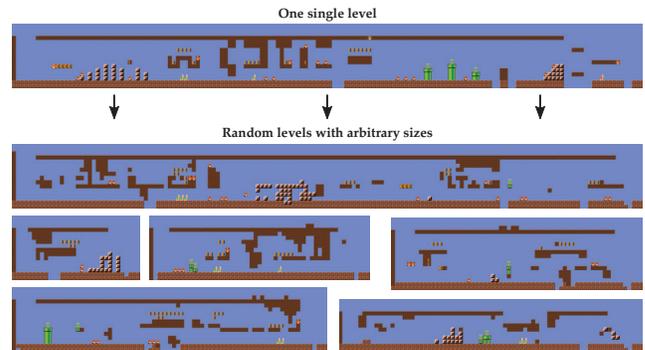


Figure 1: Generated levels based on Super Mario Bros. level 1-2. The new levels can be generated in any size and preserve the token distribution of the original level, while also containing new, previously unseen patterns.

most recent PCGML algorithms are black boxes and do not allow for high-level control of their generated content.

In this paper, we introduce TOAD-GAN as a solution to these problems. Our work is inspired by SinGAN (Shaham, Dekel, and Michaeli 2019), a recent Generative Adversarial Network (GAN) (Goodfellow et al. 2014) architecture that learns a generative model given only one example image. This is achieved by learning patch-based features on different spatial scales. Since SinGAN was developed for natural RGB images, it is unable to generate convincing video game levels that are based on 2D token maps. Our method, on the other hand, was developed exactly with this purpose in mind. As shown in Fig. 1, TOAD-GAN is able to generate new levels in the style of one given example level.

In summary, our **contributions** are:

- With TOAD-GAN, we present a novel generative model that allows for level generation following the one-shot training approach of SinGAN.
- We introduce a downsampling algorithm for token-based levels specifically designed to preserve important tokens.
- An extension of our method enables authoring of the global level structure. We show this for example levels generated for Super Mario Kart.
- We visualise our generated content in a latent space to

compare it with the original levels.

- We enable further research by publishing our source code.¹

2 Related Work

Generating Super Mario Bros. (SMB) levels was one of the first challenges proposed to the PCG community (Shaker et al. 2011). Since its introduction, many approaches were presented that tried to capture the patterns from the original levels and combine them in novel ways. This section only covers a selection, due to the vast amount of PCG approaches. For a review of pattern-based level generators for SMB see Khalifa et al. (2019).

Super Mario Bros. Level Generation

Dahlskog and Togelius (2012) identified and analyzed patterns with different themes, such as enemies, gaps or stairs. They assessed the difficulty of the patterns to human players and outlined how those patterns could be combined and varied to create new levels. In their continued work (Dahlskog and Togelius 2014), they additionally defined micro- (vertical slices) and macro-patterns (sequences of patterns). Using an Evolutionary Algorithm (EA), they generated levels by selecting micro-patterns, with a fitness function based on the occurrence of patterns and macro-patterns.

Search-based PCG (Togelius et al. 2011) was applied to SMB by Summerville, Philip, and Mateas (2015). The authors used Monte Carlo Tree Search (MCTS) (Coulom 2006) to guide the sampling process from a Markov Chain model of tokens. The reward of the MCTS was computed based on the solvability, number of gaps, number of enemies and rewards (coins or power-ups).

Neural Networks for PCG

Recent PCGML approaches that use Neural Networks have also been applied extensively to the SMB level generation problem. Hoover, Togelius, and Yannakis (2015) trained a neural network with an EA to generate new levels. The network predicts the height of a token in a level slice, given the heights of all tokens in the previous slices.

Summerville and Mateas (2016) trained their model on levels by predicting the tokens sequentially. They used a neural network architecture based on Long Short-Term Memory (LSTM) cells (Hochreiter and Schmidhuber 1997) to predict the next token, given a context of previous tokens in the unrolled level.

Recently, GANs were used to create SMB levels. In (Volz et al. 2018) the authors train a GAN on slices of the original levels and use an EA to search the space of generated levels by scoring the fraction of enemy and ground tokens.

A similar approach was taken by Torrado et al. (2019) who used the Self-Attention GAN (SAGAN) (Zhang et al. 2019) architecture and conditioned their generation process on a feature vector that contained the targeted token distributions of the generated levels. This conditioning increased the variability of their generated content.

¹<https://github.com/Mawiszus/TOAD-GAN>



Figure 2: Example of the mode collapse of a WGAN-GP trained solely on 16×16 slices from SMB level 1-1. This slice size is chosen because square images are preferable for GANs and 16 is the default height of all SMB levels. The generator stops creating anything but the ground and sky.

3 Method

Limited training data is one of the key problems of PCGML algorithms (Torrado et al. 2019; Bontrager and Togelius 2020). Therefore, the goal of our work is the generation of new levels for SMB from very little training data. With TOAD-GAN, we take this problem to the extreme regime of learning from only one single training level. Similar to other recent publications (Volz et al. 2018; Torrado et al. 2019; Volz et al. 2020), TOAD-GAN is based on the GAN architecture.

Generative Adversarial Networks

GANs are able to generate samples from a given training distribution (Goodfellow et al. 2014). They consist of two adversaries: The generator G maps random noise vectors z to samples \tilde{x} , which the discriminator D is trying to distinguish from real samples x . In the end, G produces \tilde{x} that are indistinguishable from real x . However, this process can become unstable. In the low-data regime, the discriminator might be able to memorize the distribution of real samples and stops providing useful gradients for the generator. Many different extensions to the basic architecture were proposed to stabilize the training process, for example minimizing the Wasserstein distance (Arjovsky, Chintala, and Bottou 2017) and penalizing the norm of the gradients of the discriminator (Gulrajani et al. 2017). The resulting Wasserstein GAN with Gradient Penalty (WGAN-GP) is able to model a variety of distributions, but it is still prone to failures like mode collapse. This is the case when the generator produces samples that contain only a few features (or modes) of the data which the discriminator cannot classify correctly. Then, the generator will never learn to produce the missed modes. See Fig. 2 for an example.

Even though GANs have shown promising results, their success depends on the availability of a lot of samples from the real distribution. Additionally, long range correlations in an image can only be modeled by convolutional GANs with many layers. For longer levels, this increases the number of parameters that have to be optimized, which further complicates the training process.

SinGAN

SinGAN (Shaham, Dekel, and Michaeli 2019) is a novel GAN architecture that enables learning a generative model from a single image. This is achieved by using a cascade of generators and discriminators that act on patches from differently scaled versions of the image. The weights of the models at each scale are initialized with those from the

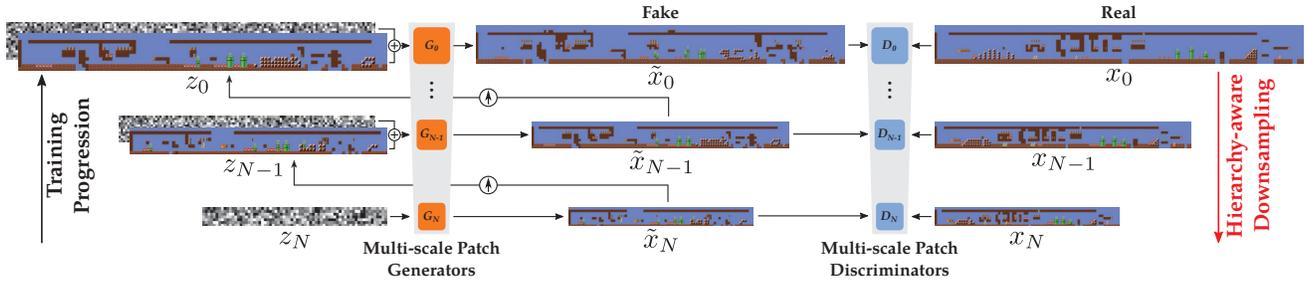


Figure 3: Generation process of TOAD-GAN on Super Mario Bros. level 1-2. The architecture is adapted from SinGAN (cf. Fig. 4 of (Shaham, Dekel, and Michaeli 2019)). We use a downsampling method on a one-hot encoded version of the level that preserves small but important structures which would be lost when performing simple spatial downsampling. The upwards arrow between the scales represents bilinear upsampling.

scale below. This initialization bootstraps the training of the GANs in upper scales and stabilizes the training process.

To generate a new sample, a noise map z_N of arbitrary size is fed into the generator G_N at the lowest scale. At each subsequent scale $N < n \leq 0$, the output \tilde{x}_{n+1} of the previous generator G_{n+1} is up-sampled (\uparrow) and added to a new random noise map $z_n \sim \mathcal{N}(0, \sigma^2)$. The variance σ^2 controls how much information from the lower scales is passed through to the upper layers of SinGAN. Both are used by the generator G_n to produce an output \tilde{x}_n for the current scale

$$\tilde{x}_n = \tilde{x}_{n+1} \uparrow + G_n(z_n + \tilde{x}_{n+1} \uparrow). \quad (1)$$

The discriminators receive either a scaled real image or the output of their respective generator. Generator and discriminator only act on patches and are fully-convolutional. This means that the size of the output is determined by the size of the initial noise map at the lowest scale.

For a more in-depth explanation please refer to the original SinGAN paper by Shaham, Dekel, and Michaeli (2019).

TOAD-GAN

Fig. 3 shows the pipeline of TOAD-GAN for the generation of SMB levels. There are 15 original SMB levels provided by the Video Game Level Corpus (Summerville et al. 2016), each with different characteristics. The levels are placed in three worlds (overworld, underground, floating platforms) with different global structure and token patterns. For training, one level is sampled down to N different scales. We choose N such that the receptive field of the convolutional filters in our generators and discriminators is able to cover at minimum half of the height of the levels at the lowest scale. This ensures that larger structures are modeled correctly, but allows for variation in their global position.

Interpreting each token as one pixel of an image and then downsampling naively results in lost information, as aliasing would make important tokens disappear at lower scales. To keep most of the information from the original level, we propose a downsampling method which preserves important tokens. This method is inspired by TF-IDF weighting (Manning, Raghavan, and Schütze 2008) in Natural Language Processing where the importance of a term is defined by its term frequency multiplied by its inverse document fre-

Table 1: Token Hierarchy

Group	Tokens
0 Sky	
1 Ground	
2 Pyramid	
3 Platforms	
4 Pipes	
5 Enemies	
6 Special Enemies	
7 Special Blocks	
8 Hidden Blocks	

quency. In our case, terms are tokens and documents are levels. Tokens that occur often and in multiple levels, like the sky and ground blocks, are of lower importance than rare tokens, such as the hidden and special blocks. The complete token hierarchy can be found in Tab. 1.

The steps of this process are as follows. First, bilinear downsampling is used on the one-hot encoded training level to create the base levels of the chosen scales. For each pixel in each scale, the tokens with a value greater than zero are selected. From that list, the tokens with the highest rank in our hierarchy are kept and the remaining tokens are set to zero. Finally, a Softmax is applied over all channels per pixel. In Fig. 3 on the right, two downsampled versions of level 1-2 can be seen. Later, we also need to sample the outputs on lower scales up. For this we use bilinear up-sampling.

On natural images, SinGAN uses zero-centered gaussian spatial noise that is constant over all color channels for a given pixel, i.e. it only changes the brightness of that pixel. This places a prior on the hue of the up-sampled pixels and increases the similarity of the generated samples between the different scales. In our case, the channels represent the tile types. Because these are independent from each other, we apply the noise to all channels individually.

TOAD-GAN can be extended to perform level authoring by injecting a predefined input into the generator cascade. The generators fill in the details and produce a sample which

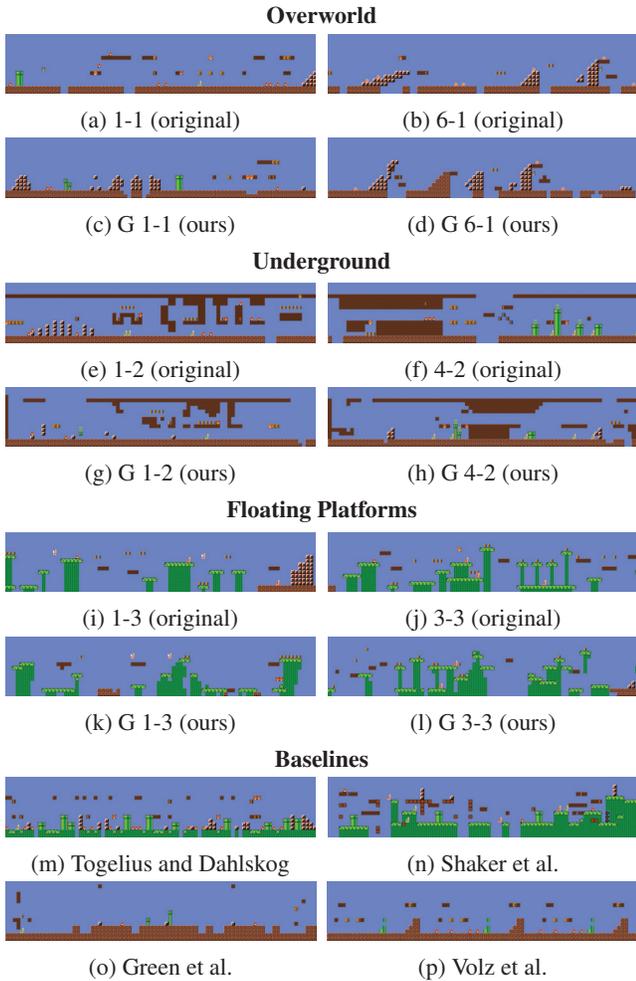


Figure 4: Example levels generated by TOAD-GAN for the three different level types found in Super Mario Bros. in comparison to other generators. TOAD-GAN is able to capture the style of the level it was trained on.

follows the structure of the injected input but has a similar style as the training sample. This application is particularly interesting for PCG as the designer can describe a desired level or layout for a given token and the generators create variants of it. In our experiments, we inject a new, differently structured map for a specific token after the very first generation step. This basic structure is preserved and expanded upon by the following generator steps, which results in a level with the desired structure that consists of the patterns learned by the generators.

4 Experiments

Our experiments are split into two parts. First, we perform a qualitative evaluation of the generated levels by presenting a number of samples to highlight capabilities of our approach. Then, our generated levels are assessed with regards to their Tile Pattern KL-Divergence (TPKL-Div) (Lu-

cas and Volz 2019) and visualised using an embedding of the level slices that is inspired by the Fréchet Inception Distance (FID) (Heusel et al. 2017). In the second part, we show the generality of TOAD-GAN by applying it to *Super Mario Kart* and present an example of level authoring.

We use the same hyperparameters for all experiments. The samples were generated at scales 0.5, 0.75, 0.88 and 1.0 of the original training sample size. The remaining hyperparameters and specific architectures will be published with our source code.

Level Generation Evaluation

Qualitative Examples In this experiment, TOAD-GAN is trained on each of the levels provided by the SMB benchmark. Fig. 4 shows randomly generated samples for levels of different types. To increase comparability, all levels were cut (others), or generated (ours) to the same length. The style of our generated samples matches that of the level they were trained on. For example, the hidden 1-Up block in (b) is placed very similarly in (d). However, the patterns in our samples are combined differently than in their training level (e.g. the three small platforms with coins from (j) are transformed to different heights in (l)). All this while the general structure of the generated levels is similar to a SMB level. We tested the validity of our generated content using the A* agent by Baumgarten (Togelius, Karakovskiy, and Baumgarten 2010), who was able to win 65% of randomly sampled levels compared to the 52% of the original levels (Green et al. 2020).

While we focus on learning one generator for one level, other methods create one for all training levels. The levels (m-p) are some example results of such generators. In them, different kinds of levels are mixed (e.g. pyramids are created on floating platforms) and the level style is not captured. The levels (o) and (p) depict recognizable overworld levels, however the sample cut from (o) was trained on 4-2 and should therefore be more similar to an underground level. The closest to a convincing overworld level is (p) which was also created using a GAN-based approach (Volz et al. 2018). However, this method relies on small samples that are stitched together and can result in repeating patterns.

Tile Pattern KL-Divergence We use the TPKL-Div by Lucas and Volz (2019) to evaluate the similarity of our generated patterns to the originals. Fig. 5 shows the results for all original SMB levels. As expected, the values on the main diagonal, where the generated samples are compared to the level they were trained with, are very small. This indicates that TOAD-GAN is able to model the original pattern distributions for any type of level. Also noticeable are spots in the matrix where a very low value occurs for a different level than the one trained on. This happens because these levels are of a similar style. For example, levels 1-3, 3-3, 5-3, and 6-3 are all levels with floating platforms, as shown in Fig. 4(i)-(l).

Tab. 2 shows our resulting divergences compared to those reported by Lucas and Volz (2019) and Green et al. (2020). For our results, we computed the mean 2×2 , 3×3 , and 4×4 pattern divergences with $w = 1.0$. As we train 15

Table 2: Average Tile Pattern KL-Divergence

Algorithm	TPKL-Div.
ELSGAN ²	1.58
GAN ²	1.70
ETPKLDiv 3x3 ²	0.88
Evolution World ³	1.70
TOAD-GAN (ours)	0.33

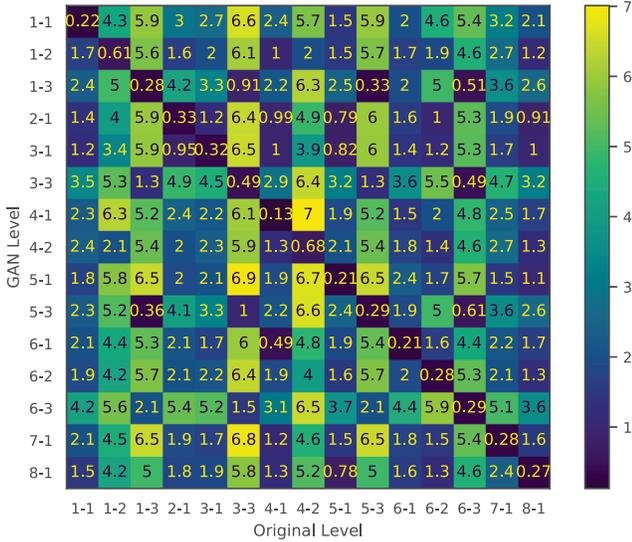


Figure 5: Mean Tile-Pattern-KL-Divergence between 100 generated levels and the original levels. Values are averaged over 2×2 , 3×3 and 4×4 patterns. Each row represents a TOAD-GAN that was trained on the labelled SMB level.

separate generators, we average their values to get the result in Tab. 2. We generated 1000 sample levels with a size of 200×16 tokens for each generator. On average, levels generated by TOAD-GAN produce a lower and therefore better TPKL-Div. However, as the TPKL-Div measures only the differences for patterns already present in the original level, newly generated patterns are not taken into account. Visual inspection of the generated levels (compare Figs. 1 and 4) indicates that existing patterns are not only reproduced, but combined in novel ways and new patterns are generated.

Fig. 2 indicates that GANs tend to produce very similar or even the same output. We tested the variability of our generated content by computing the uniqueness of structures in our generated samples. For that, we randomly picked 100000 square 16×16 slices evenly from our previously generated samples and found that an average of 90.62% of them were unique.

²Results by Lucas and Volz (2019), on level 1-1 averaged over 2×2 , 3×3 , and 4×4 patterns with $w = 1.0$

³Results by Green et al. (2020), averaged over level 1-1, 4-2 and 6-1, only 3×3 patterns

Level Embeddings Even though the TPKL-Div captures some aspects of the similarities between SMB levels, it is limited to patterns of fixed sizes. We propose a new method that additionally results in an easily interpretable visualization. Unlike the TPKL-Div, our distance metric is independent of the size of the patterns. Similar to FID, we train a convolutional classifier c on slices s of the original levels to predict the level they are from.

$$\text{level} \in \arg \max c(s) = \arg \max W\phi(s) + b \quad (2)$$

The slice representation in the penultimate layer $\phi(s)$ of the classifier is a vector of non-linear features that is mapped linearly to the predicted level.

In Fig. 6(a), we visualize the distribution of these representations by projecting them to two dimensions with the Universal Manifold Approximation and Projection (UMAP) (McInnes, Healy, and Melville 2018) method. This reveals the three level types which SMB levels fall into: Overworld to the top right, underground to the left and floating platforms at the bottom.

Fig. 6(b) shows the distribution of level slices generated by our 15 generators by mapping them with the same transformation learned in (a). Our generators are indeed generating slices very close but not limited to the manifold of their original level. In some cases, our generators even learn to create slices that are not in their training distribution, but still within the level manifold. Examples are G 1-3, G 5-3, and G 6-3 that generate slices in the same space as 3-3. This experiment highlights the variability of the produced slices of a generator while still being within the intended constraints of being a SMB level of the same style.

Level Authoring

TOAD-GAN enables us to perform authoring of the global level structure. This is made possible by editing the token maps in any of the scales, which results in the edit being represented in the generated level. Fig. 7 shows examples of this application for seeding a track layout in Super Mario Kart. TOAD-GAN is only trained on the original sample track and will generate track layouts similar to that. Because our method does not yet take playability into account, dead ends and unconnected track pieces can be generated. Seeding a layout can not only ensure a connected and working racing track, it also allows the track to have a significantly different structure than the original sample. Each seed can, depending on the noise in the other token maps and the other scales, generate an infinite amount of levels with the given structure. As the Super Mario Kart levels are much larger than the SMB levels, we used 5 convolutional layers instead of 3 and chose 9 scales (0.2, 0.3, ..., 1.0). The token hierarchy is (from low to high) ground, wall, road and special (coins etc.).

5 Conclusion and Future Work

With this paper, we propose TOAD-GAN, a Procedural Content Generation algorithm that is able to generate token-based levels (as shown for Super Mario Bros. and Super Mario Kart) while being trained on a single example. We

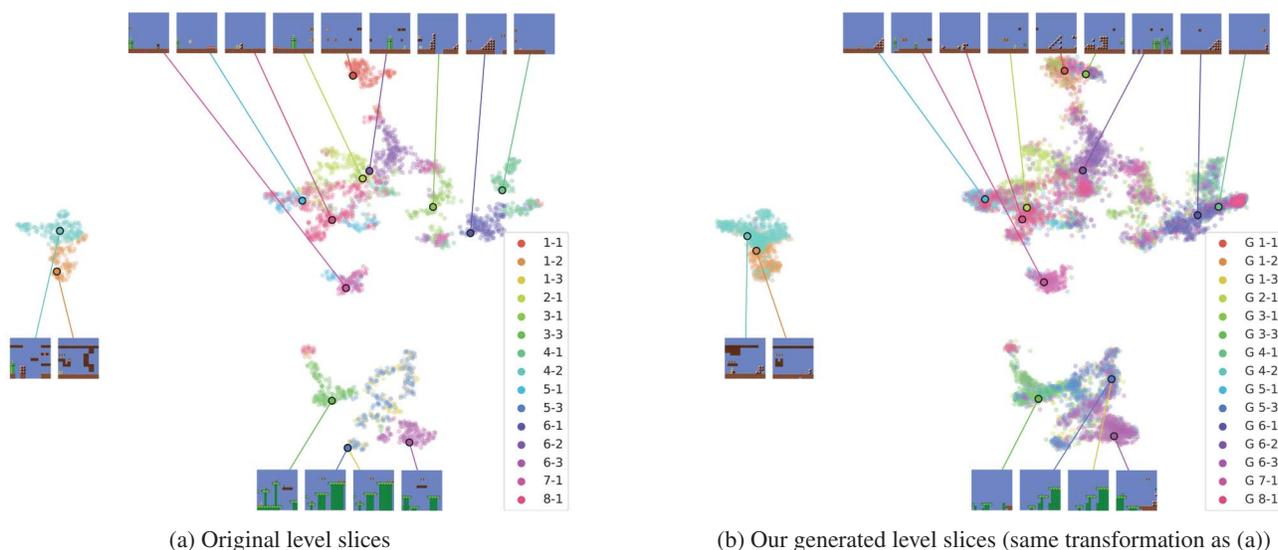


Figure 6: Level slice representations of a level classifier projected to two dimensions. Each point represents a 16×16 slice of a Super Mario Bros. level. The marked points are the ones closest to the mean of their respective level. For visualization purposes, a small amount of noise was added to the points, as some would otherwise overlap. The generated slices are close to the original slices of their respective level, with some slices being similar to other levels of the same style.

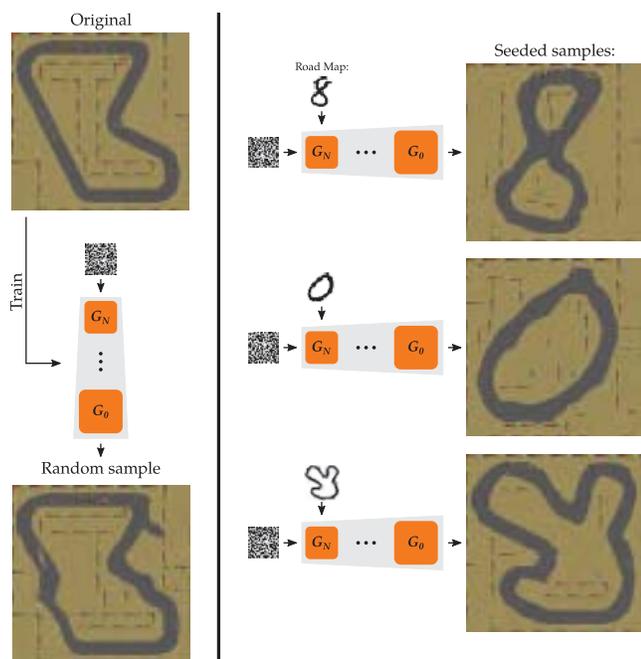


Figure 7: TOAD-GAN can be applied to levels of arbitrary token-based games, e.g. Super Mario Kart. We can enforce a predetermined track layout by conditioning the generation process in the lowest scale. The examples shown are two different hand-drawn digits from the MNIST dataset (LeCun et al. 1998) and an additional layout made for this example.

expand on the novel SinGAN architecture to generate token-based levels instead of natural images. The generated levels are evaluated qualitatively by computing their Tile Pattern KL-Divergence. Their visualization as slice embeddings offers a new way of comparing them with the original levels without specifying the pattern dimensions. By seeding a predefined basic level layout, it is possible to generate new levels while still keeping the style that TOAD-GAN was trained on. An example of this is shown by using hand drawn tracks for generating Super Mario Kart levels. Our experiments demonstrate how TOAD-GAN is able to capture the patterns of its training input and generate consistent variations of it.

We intend to improve our approach in the future by also taking gameplay mechanics into account during the generation process. Samples generated with TOAD-GAN are visually convincing Super Mario Bros. levels, but a proper study with human participants will help to assess the output quality in more depth. Another future direction will be the application of TOAD-GAN to voxel-based games (e.g. Minecraft) or to maze games with a non-linear level structure.

TOAD-GAN is a step towards using PCG via Machine Learning during the game design process due to its low requirements for the amount of data necessary and its extension to Level Authoring.

6 Acknowledgment

This work has been supported by the Federal Ministry for Economic Affairs and Energy under the Wipano programme "NaturalAI" 03THW05K06.

References

Arjovsky, M.; Chintala, S.; and Bottou, L. 2017. Wasserstein generative adversarial networks. In *International Con-*

- ference on Machine Learning, 214–223.
- Bontrager, P., and Togelius, J. 2020. Fully Differentiable Procedural Content Generation through Generative Playing Networks. *arXiv:2002.05259 [cs]*.
- Coulom, R. 2006. Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search. In *International conference on computers and games*, 72–83. Springer.
- Dahlskog, S., and Togelius, J. 2012. Patterns and procedural content generation: Revisiting Mario in world 1 level 1. In *Proceedings of the First Workshop on Design Patterns in Games - DPG '12*. Raleigh, North Carolina: ACM Press.
- Dahlskog, S., and Togelius, J. 2014. A multi-level level generator. In *2014 IEEE Conference on Computational Intelligence and Games*. Dortmund, Germany: IEEE.
- Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *Advances in Neural Information Processing Systems*.
- Green, M. C.; Mugrai, L.; Khalifa, A.; and Togelius, J. 2020. Mario Level Generation From Mechanics Using Scene Stitching. *arXiv:2002.02992 [cs]*.
- Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; and Courville, A. C. 2017. Improved Training of Wasserstein GANs. In *Advances in neural information processing systems*, 5767–5777.
- Heusel, M.; Ramsauer, H.; Unterthiner, T.; Nessler, B.; and Hochreiter, S. 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium. In *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, 6629–6640. Red Hook, NY, USA: Curran Associates Inc.
- Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. *Neural computation* 9(8):1735–1780.
- Hoover, A. K.; Togelius, J.; and Yannakis, G. N. 2015. Composing Video Game Levels with Music Metaphors through Functional Scaffolding. In *First Computational Creativity and Games Workshop. ACC*.
- Khalifa, A.; Green, M. C.; Barros, G.; and Togelius, J. 2019. Intentional computational level design. In *Proceedings of The Genetic and Evolutionary Computation Conference*, 796–803.
- LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–2324.
- Lucas, S. M., and Volz, V. 2019. Tile Pattern KL-Divergence for Analysing and Evolving Game Levels. *Proceedings of the Genetic and Evolutionary Computation Conference*.
- Manning, C. D.; Raghavan, P.; and Schütze, H. 2008. *Introduction to information retrieval*. Cambridge university press.
- McInnes, L.; Healy, J.; and Melville, J. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426 [cs, stat]*.
- Shaham, T. R.; Dekel, T.; and Michaeli, T. 2019. Singan: Learning a generative model from a single natural image. In *Proceedings of the IEEE International Conference on Computer Vision*, 4570–4580.
- Shaker, N.; Togelius, J.; Yannakakis, G. N.; Weber, B.; Shimizu, T.; Hashiyama, T.; Sorenson, N.; Pasquier, P.; Mawhorter, P.; Takahashi, G.; Smith, G.; and Baumgarten, R. 2011. The 2010 Mario AI Championship: Level Generation Track. *IEEE Transactions on Computational Intelligence and AI in Games* 3(4).
- Snodgrass, S., and Ontanón, S. 2013. Generating maps using markov chains. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*.
- Summerville, A., and Mateas, M. 2016. Super Mario as a String: Platformer Level Generation Via LSTMs. *1st International Joint Conference of DiGRA and FDG*.
- Summerville, A. J.; Snodgrass, S.; Mateas, M.; and Ontañón, S. 2016. The VGLC: The Video Game Level Corpus. *Proceedings of the 7th Workshop on Procedural Content Generation*.
- Summerville, A.; Snodgrass, S.; Guzdial, M.; Holmgård, C.; Hoover, A. K.; Isaksen, A.; Nealen, A.; and Togelius, J. 2018. Procedural Content Generation via Machine Learning (PCGML). *IEEE Transactions on Games* 10(3).
- Summerville, A. J.; Philip, S.; and Mateas, M. 2015. MCM-CTS PCG 4 SMB: Monte Carlo Tree Search to Guide Platformer Level Generation. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*.
- Togelius, J., and Dahlskog, S. 2013. Patterns as Objectives for Level Generation. In *Proceedings of the Second Workshop on Design Patterns in Games*. ACM.
- Togelius, J.; Yannakakis, G. N.; Stanley, K. O.; and Browne, C. 2011. Search-Based Procedural Content Generation: A Taxonomy and Survey. *IEEE Transactions on Computational Intelligence and AI in Games* 3(3).
- Togelius, J.; Karakovskiy, S.; and Baumgarten, R. 2010. The 2009 Mario AI Competition. In *IEEE Congress on Evolutionary Computation*. Barcelona, Spain: IEEE.
- Torrado, R. R.; Khalifa, A.; Green, M. C.; Justesen, N.; Risi, S.; and Togelius, J. 2019. Bootstrapping Conditional GANs for Video Game Level Generation. *arXiv:1910.01603 [cs]*.
- Volz, V.; Schrum, J.; Liu, J.; Lucas, S. M.; Smith, A.; and Risi, S. 2018. Evolving mario levels in the latent space of a deep convolutional generative adversarial network. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 221–228.
- Volz, V.; Justesen, N.; Snodgrass, S.; Asadi, S.; Purmonen, S.; Holmgård, C.; Togelius, J.; and Risi, S. 2020. Capturing Local and Global Patterns in Procedural Content Generation via Machine Learning. *arXiv:2005.12579 [cs]*.
- Zhang, H.; Goodfellow, I.; Metaxas, D.; and Odena, A. 2019. Self-Attention Generative Adversarial Networks. *arXiv:1805.08318 [cs, stat]*.