SeamlessGAN: Self-Supervised Synthesis of Tileable Texture Maps

2022 IEEE Transactions on Visualization and Computer Graphics

Carlos Rodríguez-Pardo[1,2], Elena Garcés[1,3]

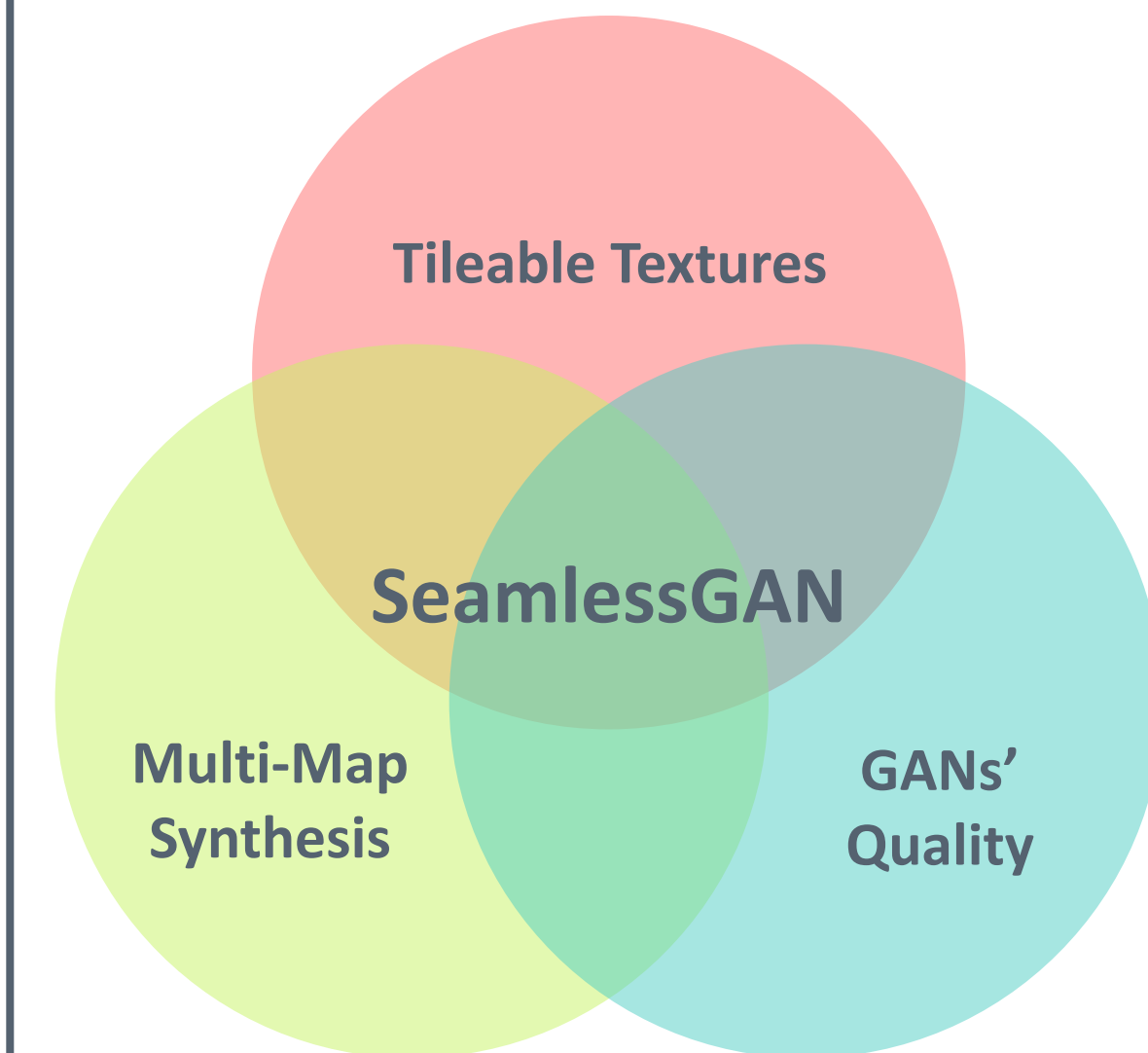SEDDI[1], Universidad Carlos III[2], Universidad Rey Juan Carlos[3]

Introduction

We present SeamlessGAN, a texture synthesis method which:
- Generates tileable textures
- Works with one or multiple texture maps
- Works with many types of textures
- Obtains SOTA results
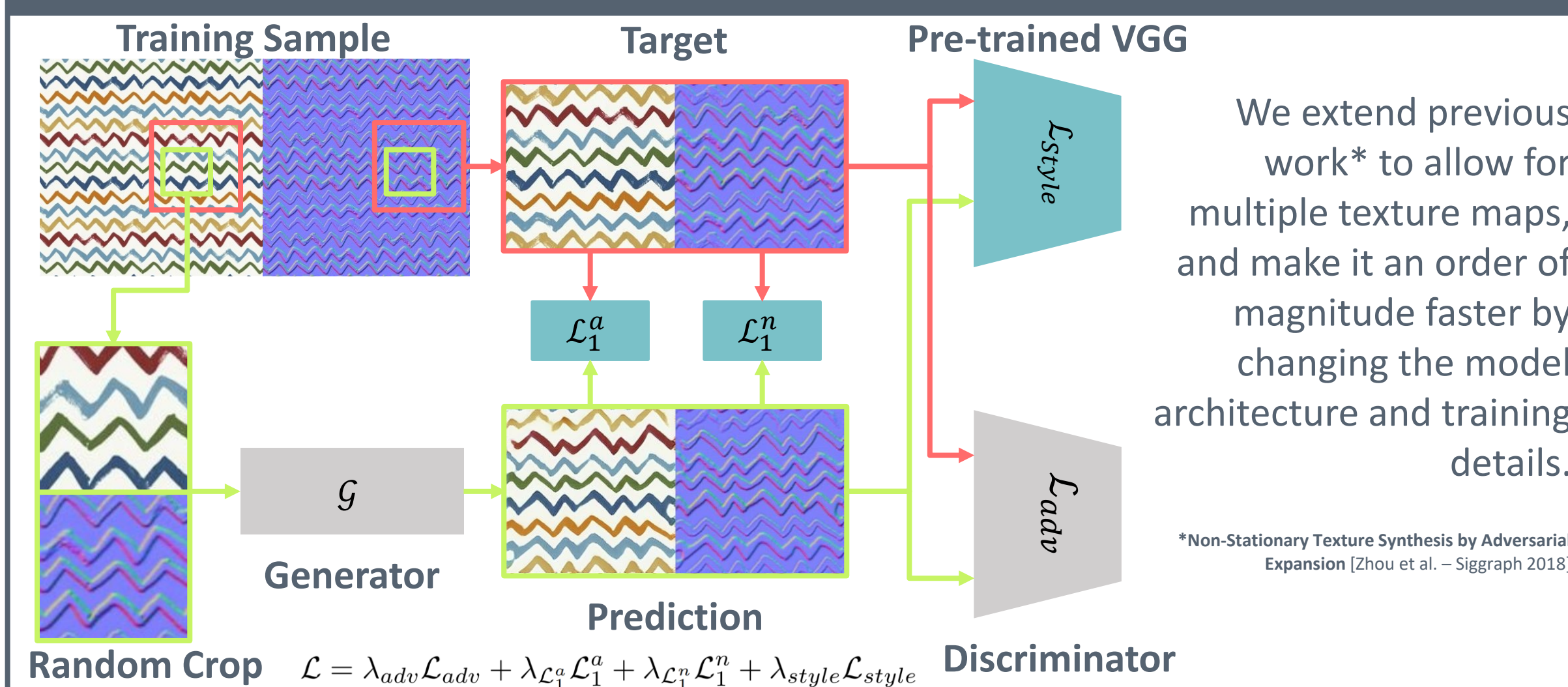- Works faster than previous methods.
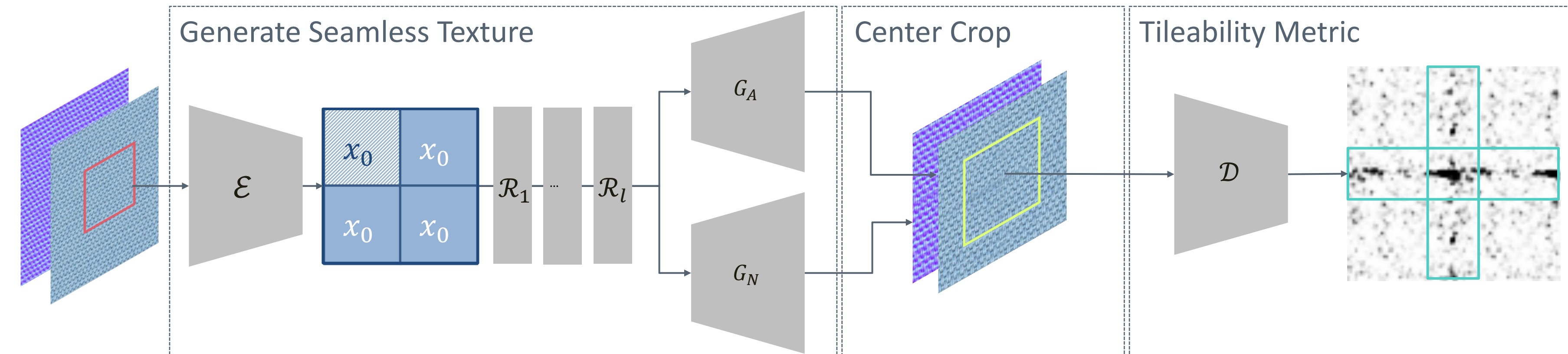

Why Tileable Texture Stacks?

Single Texture Map | Multiple Texture Maps
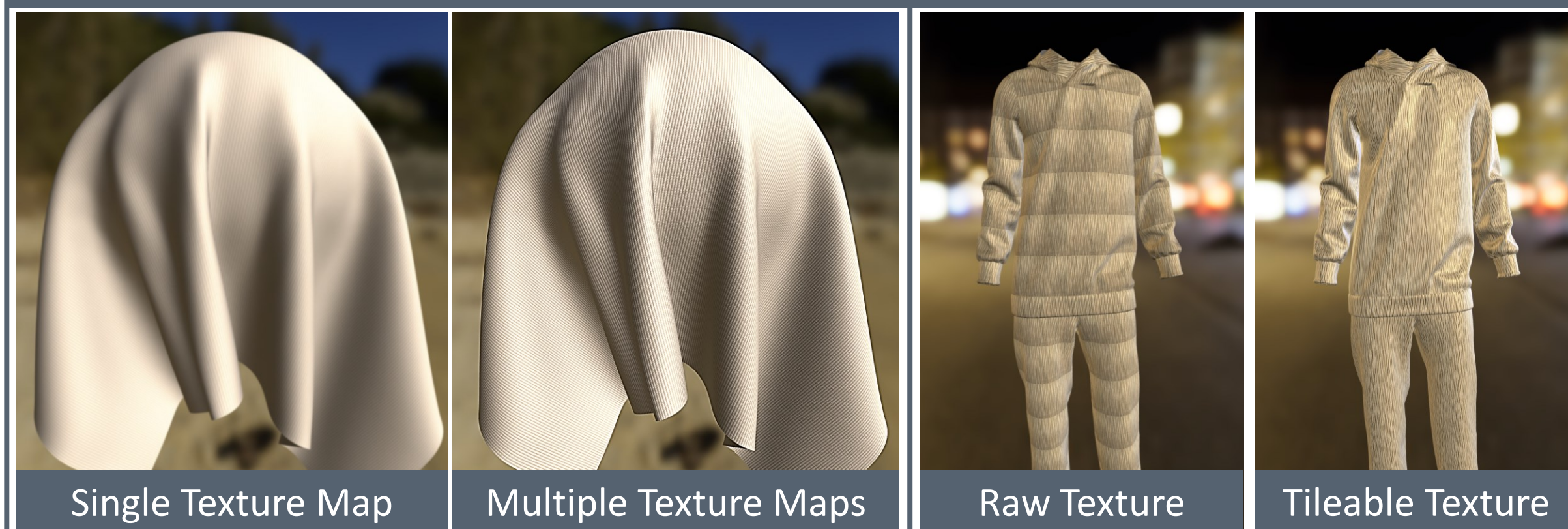
Multiple texture maps allow for enhanced realism.

Raw Texture | Tileable Texture

Tileable textures represent materials by repeating a portion of them.

Model Training using Adversarial Expansion

Training Sample | Target | Pre-trained VGG

We extend previous work* to allow for multiple texture maps, and make it an order of magnitude faster by changing the model architecture and training details.

*Non-Stationary Texture Synthesis by Adversarial Expansion [Zhou et al. – Siggraph 2018]

Random Crop | Generator | Prediction | Discriminator

$\mathcal{L} = \lambda_{adv}\mathcal{L}_{adv} + \lambda_{\mathcal{L}_1^a}\mathcal{L}_1^a + \lambda_{\mathcal{L}_1^n}\mathcal{L}_1^n + \lambda_{style}\mathcal{L}_{style}$

SeamlessGAN: Sampling

Generate Seamless Texture | Center Crop | Tileability Metric
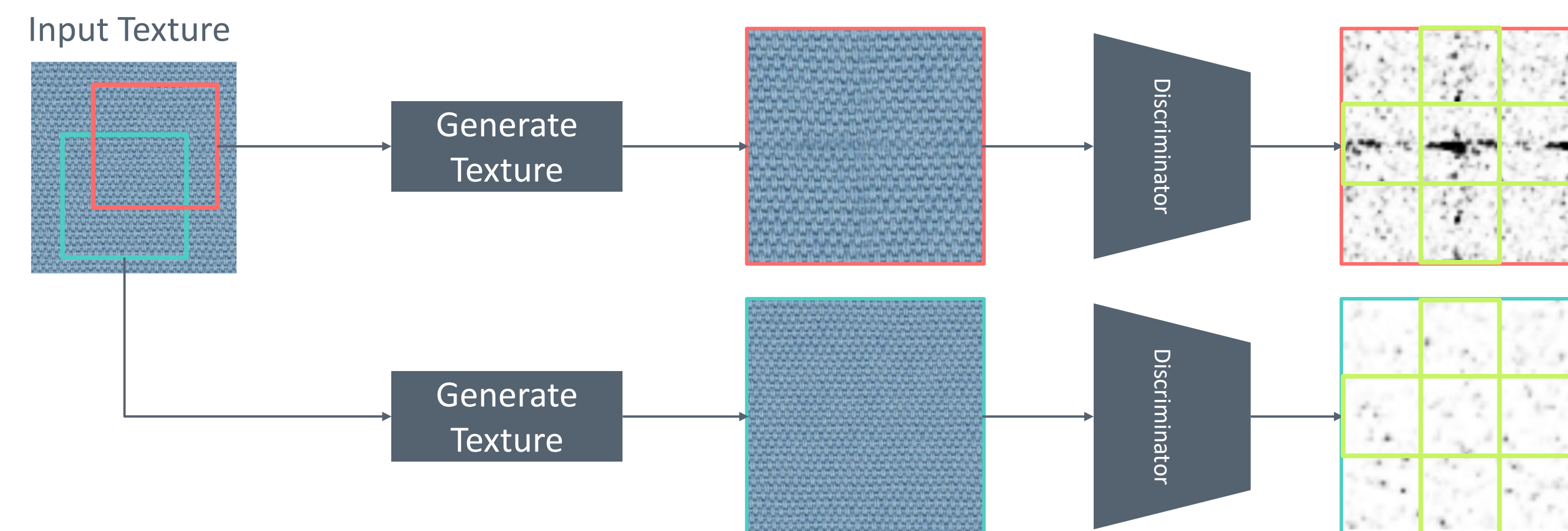
To generate a tileable texture stack, we concatenate (tile) the latent space within the generator. We crop the output, which contains 4 copies of the tileable texture. Using the discriminator at test time, we can evaluate if the generated image is tileable.
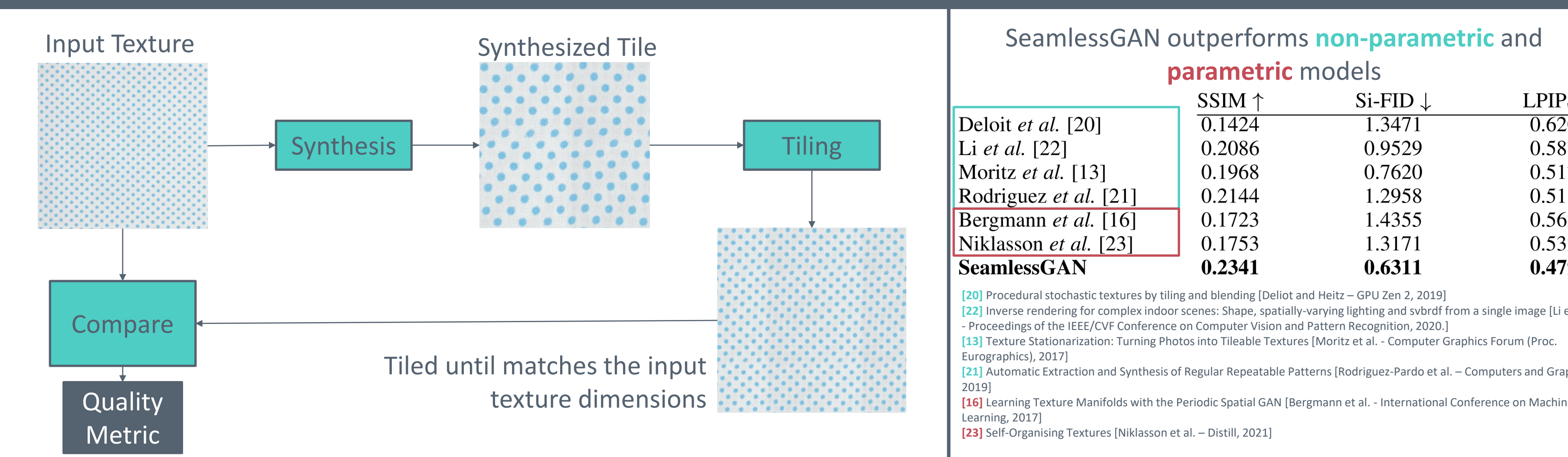
Reusing Discriminator at Test-Time as a Quality Function

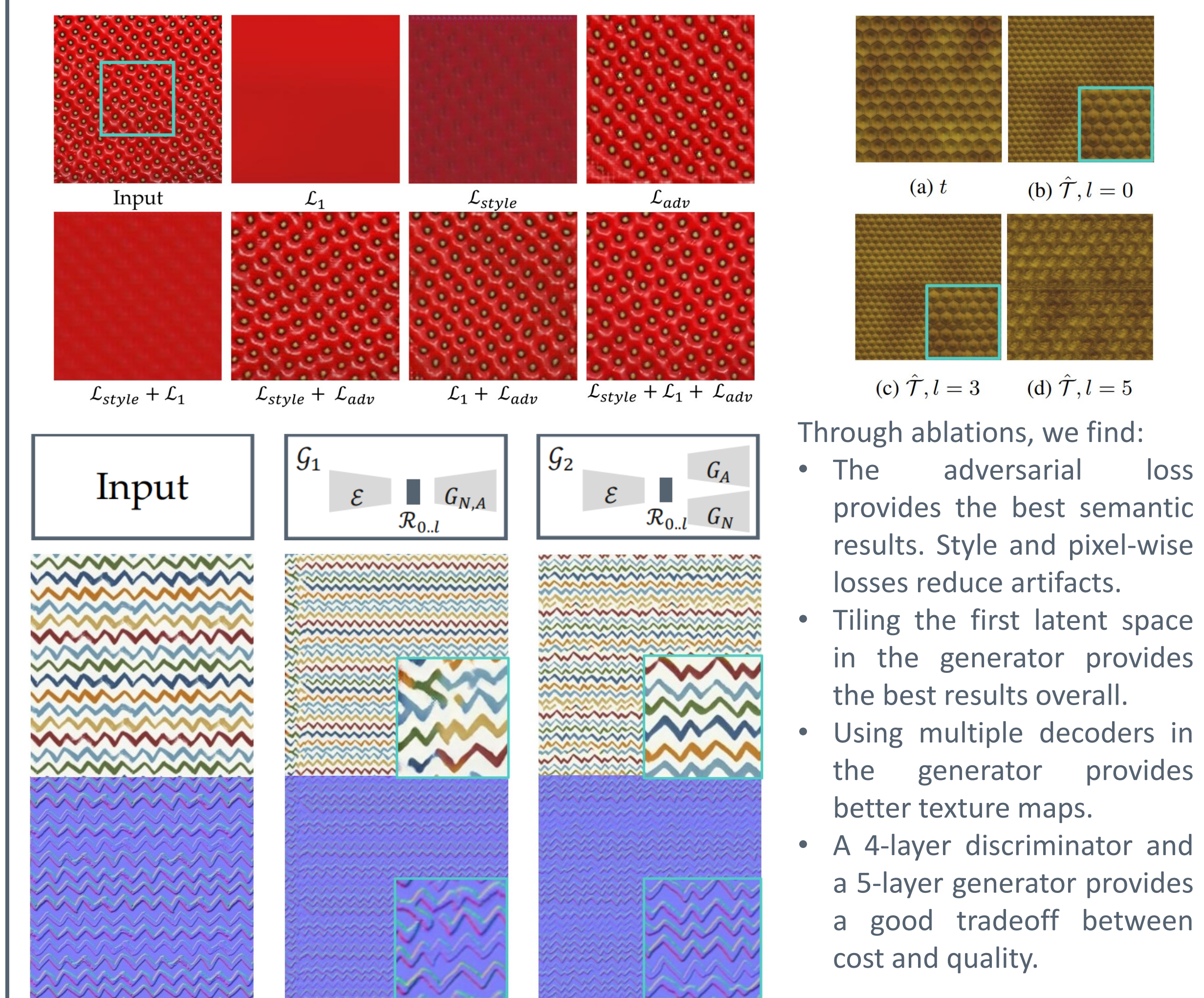Input Texture | Generate Texture | Discriminator

The generator does not always create high-quality textures. The latent space tiling operation may introduce artifacts in the center area of the textures.

At test time, we use the discriminator to find artifacts on the generated textures. If artifacts are found on the central area, the texture is discarded and another sample is generated.

Quantitative Comparison

Input Texture | Synthesis | Synthesized Tile | Tiling

Tiled until matches the input texture dimensions

Compare | Quality Metric

SeamlessGAN outperforms non-parametric and parametric models

	SSIM ↑	Si-FID ↓	LPIPS ↓
Deloit et al. [20]	0.1424	1.3471	0.6207
Li et al. [22]	0.2086	0.9529	0.5818
Moritz et al. [13]	0.1968	0.7620	0.5171
Rodriguez et al. [21]	0.2144	1.2958	0.5137
Bergmann et al. [16]	0.1723	1.4355	0.5624
Niklasson et al. [23]	0.1753	1.3171	0.5328
SeamlessGAN	**0.2341**	**0.6311**	**0.4792**

[20] Procedural stochastic textures by tiling and blending [Deliot and Heitz – GPU Zen 2, 2019]
[22] Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting and svbrdf from a single image [Li et al. – Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.]
[13] Texture Stationarization: Turning Photos into Tileable Textures [Moritz et al. – Computer Graphics Forum (Proc. Eurographics), 2017]
[21] Automatic Extraction and Synthesis of Regular Repeatable Patterns [Rodriguez-Pardo et al. – Computers and Graphics, 2019]
[16] Learning Texture Manifolds with the Periodic Spatial GAN [Bergmann et al. - International Conference on Machine Learning, 2017]
[23] Self-Organising Textures [Niklasson et al. – Distill, 2021]

Ablation Studies

Input | $\mathcal{L}_1$ | $\mathcal{L}_{style}$ | $\mathcal{L}_{adv}$

$\mathcal{L}_{style} + \mathcal{L}_1$ | $\mathcal{L}_{style} + \mathcal{L}_{adv}$ | $\mathcal{L}_1 + \mathcal{L}_{adv}$ | $\mathcal{L}_{style} + \mathcal{L}_1 + \mathcal{L}_{adv}$

(a) t | (b) $\hat{\mathcal{T}}, l = 0$

(c) $\hat{\mathcal{T}}, l = 3$ | (d) $\hat{\mathcal{T}}, l = 5$

Input | $\mathcal{G}_1$ | $\mathcal{G}_2$

Through ablations, we find:
- The adversarial loss provides the best semantic results. Style and pixel-wise losses reduce artifacts.
- Tiling the first latent space in the generator provides the best results overall.
- Using multiple decoders in the generator provides better texture maps.
- A 4-layer discriminator and a 5-layer generator provides a good tradeoff between cost and quality.

Qualitative Results

Input | Tiled Output | Input | Tiled Output | Input | Tiled Output