Image Manipulation with Generative Adversarial Networks - Project Proposal

Pierre Fernandez
pierre.fernandez@polytechnique.edu

Paul Jacob
paul.jacob@polytechnique.edu

Abstract

This report presents our project proposal on Topic K for the course Object Recognition and Computer Vision. First we introduce SinGAN's article[5] in its context, then we detail what will be the main milestones of the project, and finally we detail the proposed extension of SinGANs.

1. Description of the Chosen Topic

Generative Adversarial Networks [2] correspond to an approach for training and using convolutional neural networks as data generators, which have recently attracted a lot of attention in computer vision thanks to their ability to fit complicated image distributions and generate extremely realistic samples. For instance, a GAN trained on photographs can generate new photographs that look at least superficially authentic to human observers, having many realistic characteristics. However, GANs present some limitations such as the need of a large collection of training images for substantial results and hard-to-control characteristics of the generated samples.

Shaham et al. tackle those issues by proposing SinGAN [5], a generative model which is able to learn internal statistics from a single natural image. The proposed model contains a pyramid of fully convolutional GANs, each responsible for learning the patch distribution at a different scale of the image. This allows generating new samples of arbitrary size and aspect ratio, that have significant variability, yet maintain both the global structure and the fine textures of the training image.

SinGAN shows impressive results on various image manipulation tasks like super-resolution, paint-to-image, harmonization and animation from a single image. The original code used in the paper is available here [1].

2. Plan of Work

In this work, we will experiment with SinGANs and propose possible extensions in terms of both model design and its applications. The pipeline of the project is the following:

- First, we will review the article introducing SinGANs and understand the proposed method: we will explore the model's structure, understand how it is trained and how sampling is performed.
- Then, we will reproduce the results presented in the paper.
- Finally, we will propose 2 possible extension of SinGAN. One of the team member will focus on style transfer applications, and the other one will focus on PCG (Procedural Content Generation).

3. Presented SinGAN Extensions

A straightforward application of SinGAN's method would be texture generation. Thanks to its ability to learn internal statistics, SinGAN should create new original frames while respecting the distribution of small patches of the original picture (e.g. leaves, rocks...).

More generally, style transfer from an image to another also comes as a natural extension of SinGANs. They could be used to transfer characteristics of the training image to new images that would be fed to the generative model, not limited to textures but also including images with a larger structure, such as paintings or landscapes chosen from the "Bam!" dataset [7]. Such an approach could be tested qualitatively and quantitatively, using for instance the Single Image Fréchet Inception Distance, adapted from [3], as described in SinGAN's paper.

Another approach would be, as in texture-generation, to leverage how good SinGANs are at learning internal statistics from the training image, to generate maps and video game levels that have the same properties than one chosen instance. Indeed, current techniques need many examples to extract coherent patterns, while there is a very limited number of examples available for training, in addition to being blacks-boxes that do not allow for a high user control. SinGANs could bring a solution to these issues by training on a well-suited single level. To try this extension, we could use as training base the video-game level corpus (VGLC)[6], and obtain quantitative results with methods presented in [4] or [1].

[1] https://github.com/tamarott/SinGAN

References

[1] Steve Dahlskog, Britton Horn, Noor Shaker, Gillian Smith, and Julian Togelius. A comparative evaluation of procedural level generators in the mario ai framework. 04 2014. 1

[2] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014. 1

[3] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. 12 2017. 1

[4] Simon M. Lucas and Vanessa Volz. Tile pattern kl-divergence for analysing and evolving game levels. *Proceedings of the Genetic and Evolutionary Computation Conference*, Jul 2019. 1

[5] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. Singan: Learning a generative model from a single natural image, 2019. 1

[6] Adam James Summerville, Sam Snodgrass, Michael Mateas, and Santiago Ontañón. The vglc: The video game level corpus, 2016. 1

[7] Michael J. Wilber, Chen Fang, Hailin Jin, Aaron Hertzmann, John Collomosse, and Serge Belongie. Bam! the behance artistic media dataset for recognition beyond photography, 2017. 1