

SinGAN: Learning a Generative Model from a Single Natural Image

Björn Bebensee (2019–21343)
Topics in Artificial Intelligence

November 19, 2019

Shaham et al. [1] propose *SinGAN*, a generative model which can generate high-quality samples learned from a single training image. In the single training image setting the goal is to learn the distribution of image *patches* in a single image rather than whole images. By sampling this learned distribution later, we can generate images similar to the training example.

SinGAN is a model architecture consisting of a hierarchy of patch-GANs that learn the patch distribution within the image at a fixed scale (coarse-to-fine). At each scale n the patch-GAN G_n learns to generate image samples in which the discriminator D_n can not distinguish the patches between the generated image and the training image downsampled to scale n . The input to the generator G_n at each scale is an upsampled version of the generated image from the previous generator G_{n+1} and Gaussian noise which is added to the image. Each generator G_n adds new details to the image at a higher scale, allowing for generation of high-quality samples.

The authors show that their model can be used for a variety of image manipulation tasks without any modifications. Super-resolution can be achieved by training the model on the image and feeding an upsampled version of the original image into the last generator repeatedly to progressively obtain a higher resolution output with more details added by the generator. In paint-to-image the authors train the model on a single training image and then feed a down-scaled version of the painted image into the coarsest generator. The generators will then add detail found in the image patches at each scale. Likewise for harmonization and editing a naively modified image is pasted into the coarsest generator and the model will be able to add realistic detail at each layer and output a high-resolution image. Lastly, by sampling the patch distribution using a random walk instead of random noise an animation effect on the image can be achieved.

Shaham et al. evaluate SinGAN in a “real/fake” user study and a single-image version of the Fréchet Inception Distance and find it achieves good results (high confusion rate in unpaired setting). SinGAN performs similarly to the best external method in the super-resolution task (which has been exposed to many training images). Overall it can be said that their results are very good in terms of visual quality and better than competing methods in tasks like harmonization and editing.

References

- [1] Shaham, Tamar Rott, Tali Dekel, and Tomer Michaeli. “SinGAN: Learning a Generative Model from a Single Natural Image”. *Proceedings of the IEEE International Conference on Computer Vision*. 2019.