MEGAPIXEL IMAGE GENERATION WITH STEP-UNROLLED DENOISNG AUTOENCODERS

Alex F. McKinney[1], **Dr. Chris G. Willcocks**[1,2]

[1]Department of Computer Science, Durham University
[2]Project Supervisor

Fig. 1: 1024×1024 samples from the FFHQ-1024 model. Each sample can be generated in ≈ 2 seconds on a consumer-grade GPU.

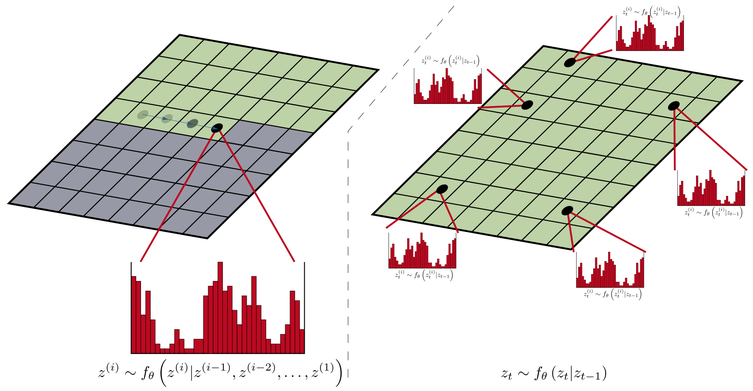

Fig. 2: **Left:** Visualization of AR sampling. AR sampling proceeds one item at a time, resulting in the number of sampling steps being equal to the dimensionality of the input. For each prediction, a probability distribution over possible tokens is predicted and then sampled from. Each prediction can only make use of past context – indicated as a green position – so not to violate the autoregressive property. **Right:** Visualization of NAR sampling. NAR sampling can sample an arbitrary number of items in parallel, including ones previously sampled, allowing for self-correction. It can freely use all context available to it, allowing for flexible inpainting and potentially better predictions.

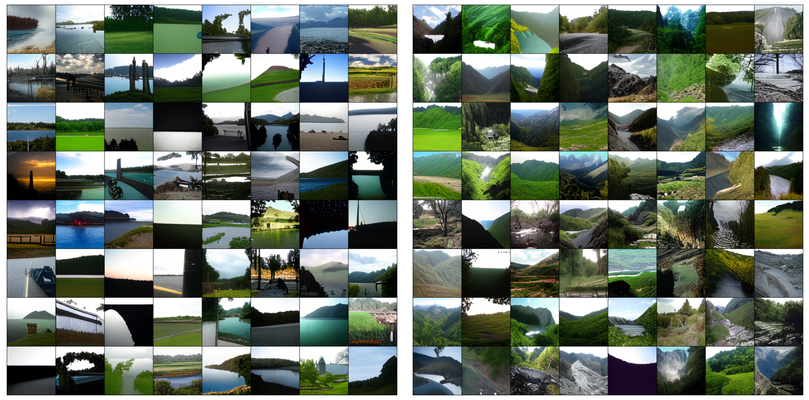

Fig. 3: 1024×1024 samples from the FFHQ-1024 model. Each sample can be generated in ≈ 2 seconds on a consumer-grade GPU.

Fig. 4: 1024×1024 samples from the FFHQ-1024 model. Each sample can be generated in ≈ 2 seconds on a consumer-grade GPU.