

# Supplementary materials:

## GeoGuide: Geometric guidance of diffusion models

Mateusz Poleski<sup>1</sup> Jacek Tabor<sup>1</sup> Przemysław Spurek<sup>1,2</sup>

<sup>1</sup>Faculty of Mathematics and Computer Science, Jagiellonian University

<sup>2</sup>IDEAS NCBR

mateusz.poleski@student.uj.edu.pl {jacek.tabor, przemyslaw.spurek}@uj.edu.pl

### 1. Comparison of samples from conditional model

In Figure 1 we compare one-to-one samples generated with ADM-G and GeoGuide. Conditional model trained on ImageNet 256x256 was used. Each pair was sampled using the same random seed and class labels. We can observe that visually improvement is much smaller compared to unconditional case, as models produce mostly similar results.

### 2. Examination of the intermediate denoising steps

Figure 2 compares corresponding intermediate steps during the sampling process with GeoGuide and ADM-G. The same class label and random seed was used for both models. We can observe that samples from GeoGuide are visually more appealing even at earlier stages of the process, which can suggest that GeoGuide brings improvement over vanilla guidance during the entire process, not only at the later stages.

### 3. Examining number of forward passes in GeoGuide

Diffusion models during the training usually uses 1000 diffusion steps, which can be later reduced during the sampling stage. This can significantly accelerate the sampling speed, without making a large sacrifice in quality. Table 1 shows this relation in case of GeoGuide. We can observe similar results for all tested settings. Using 1000 iterations brings very little improvement over 250, and even using just 50 steps can produce satisfying results if we want to significantly reduce sampling time.

Model	Conditional	Iters	Scale	FID	sFID	IS	Precision	Recall
GeoGuide ( $\sqrt{1 - \bar{\alpha}_t}$ )	✗	50	0.15	9.61	13.45	185.63	0.78	0.41
GeoGuide ( $\sqrt{1 - \bar{\alpha}_t}$ )	✗	250	0.15	7.47	8.85	206.79	0.82	0.43
GeoGuide ( $\sqrt{1 - \bar{\alpha}_t}$ )	✗	1000	0.15	7.15	7.73	208.33	0.82	0.43
GeoGuide	✗	50	0.15	9.80	11.41	226.90	0.72	0.40
GeoGuide	✗	250	0.15	7.32	7.98	243.34	0.77	0.42
GeoGuide	✗	1000	0.15	7.14	7.43	245.79	0.78	0.42
GeoGuide	✓	50	0.025	6.10	8.53	191.41	0.79	0.52
GeoGuide	✓	250	0.025	4.06	5.19	206.86	0.82	0.55
GeoGuide	✓	1000	0.025	4.16	4.98	202.65	0.83	0.55

Table 1. Effect of using different number of forward passes during the sampling process with GeoGuide and its variation on ImageNet 256x256. Using 1000 iterations is not providing significant improvement over 250, in contrast to using 50, where difference is much larger. Taking into account computation time, using 250 iterations seems optimal for most use-cases.

### 4. GeoGuide scaled variant

Table 2 is a full version of the table with comparison of GeoGuide with its variant given by  $A_t = \frac{\sqrt{D}}{T} \sqrt{1 - \bar{\alpha}_t} \frac{v(x)}{\|v(x)\|}$ .

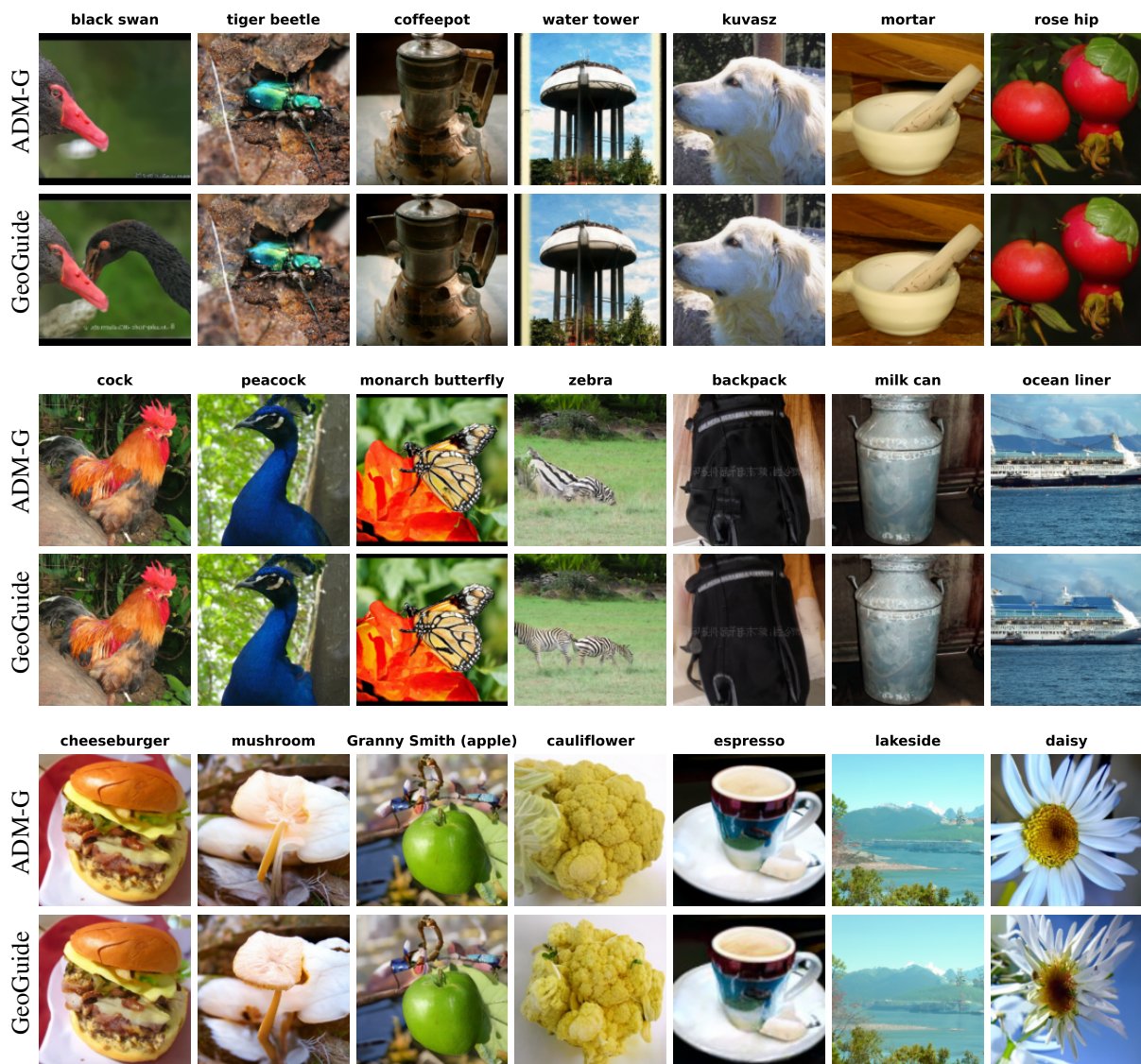


Figure 1. Images generated by conditional guided diffusion using the same noise (random seed) and class label, with a vanilla (FID 4.78, top) and a geometric (FID 4.06, bottom) guidance. Visually both models produce mostly similar results, possibly with small improvement in GeoGuide. Difference in quality is much smaller compared with unconditional model.

Model	Conditional	Scale	FID	sFID	IS	Precision	Recall
GeoGuide ( $\sqrt{1 - \bar{\alpha}_t}$ )	$\times$	0.15	7.47	8.85	206.79	0.82	0.43
GeoGuide	$\times$	0.15	7.32	7.98	243.34	0.77	0.42
GeoGuide ( $\sqrt{1 - \bar{\alpha}_t}$ )	$\checkmark$	0.025	4.78	5.13	174.13	0.80	0.56
GeoGuide	$\checkmark$	0.025	4.06	5.19	206.86	0.82	0.55

Table 2. Comparison of GeoGuide with its variant, where we rescale the basic adjustment additionally by  $\sqrt{1 - \bar{\alpha}_t}$ . Base approach achieves altogether better results. Evaluated on ImageNet 256x256 using 250 iterations during the sampling.



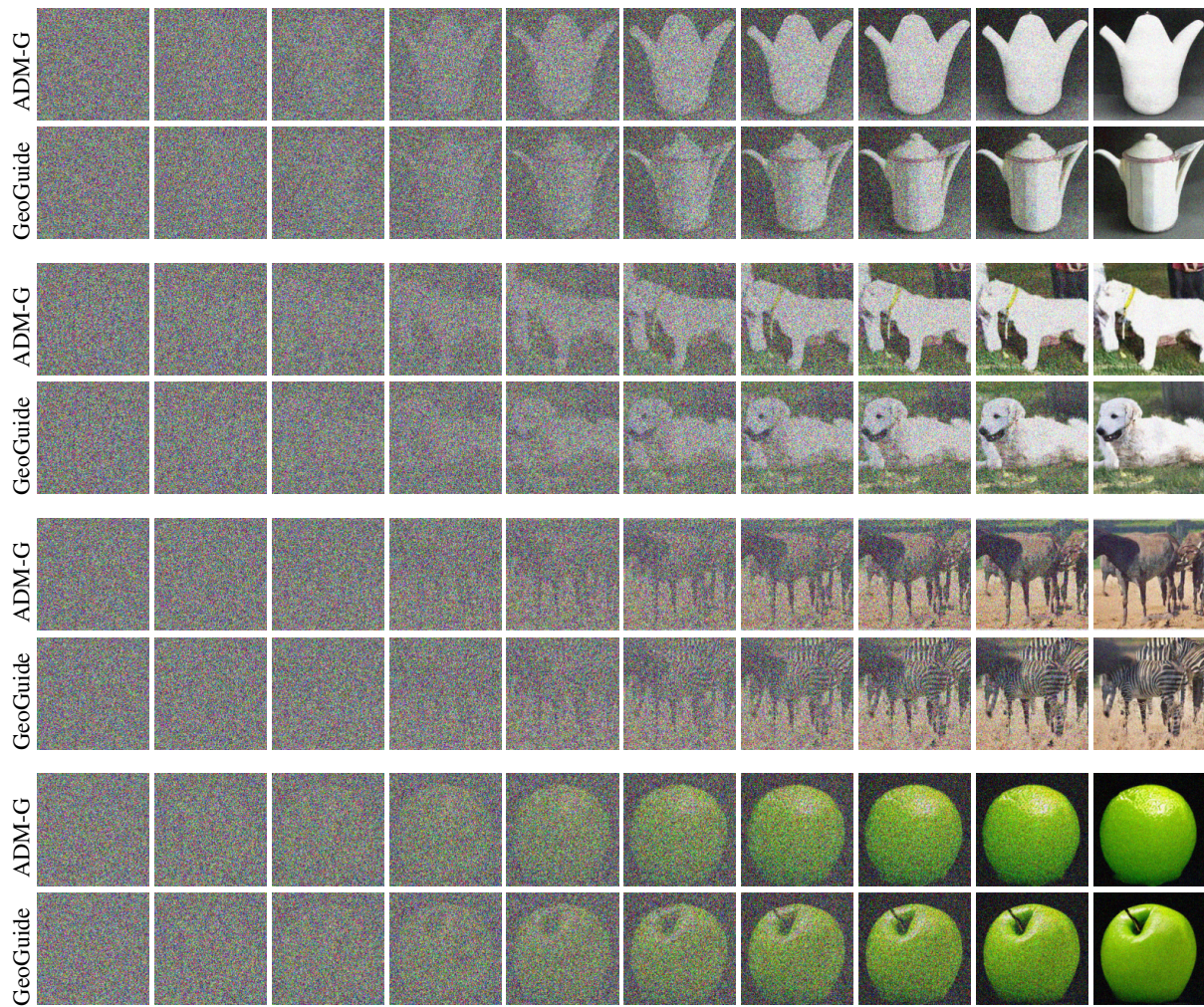


Figure 2. Comparison of the selected snapshots from the denoising processes. GeoGuide improves quality over the entire timeline.





Figure 3. Uncurated samples from ImageNet 256x256 generated by a conditional diffusion model, guided with GeoGuide (FID 4.06).



