

1 In the supplemental material, we show additional results under other training settings in [App. A](#) and
2 more generated examples in [App. B](#).

3 A Results with Other Experimental Settings

4 We try different training settings for VAR-d24 and LlamaGen-L model and get more results. For
5 **VAR-d24 model**, we continue training from the checkpoint used in our main table. To ensure
6 sufficient training of the generator, we decrease the guidance update frequency to 1 and fix the
7 generator learning rate at $2.8e-5$. We train for an additional 100K guidance iterations and 100K
8 generator iterations. For **LlamaGen-L model**, we perform DD2 from start under a new setting.
9 Specifically, compared to the training configuration used in the main paper, we increase the generator
10 learning rate to $2e-4$ and trained for 35K guidance iterations and 7K generator iterations. We then
11 conduct the performance alignment phase and train the guidance network for an additional 20K
12 iterations. To ensure training stability of the generator, we reduce the generator learning rate to $1e-5$
13 and set the guidance update frequency to 20. Under this configuration, we continue training for
14 another 10K guidance iterations and 500 generator iterations. Results under the new training settings
15 are listed in [Tab. 1](#).

Table 1: Performance of DD2 under new settings.

Type	Model	FID↓	IS↑	Pre↑	Rec↑	#Para	#Step
DD2 (ours)	VAR-d24	4.91	282.2	0.87	0.39	1.04B	1
DD2 (ours)	LlamaGen-L	8.17	225.0	0.77	0.34	335M	1

16 B More Visualization Examples

17 We demonstrate more generated samples by our DD2 model in [Figs. 1 to 4](#). We further present a
18 comparison between the images generated by DD2 and by the teacher model in [Figs. 5 to 8](#).

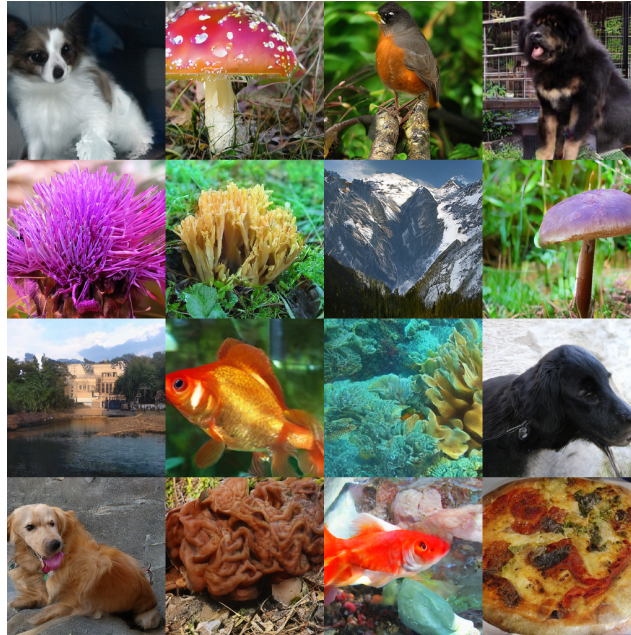


Figure 1: Generated by DD2-VAR-d16 model.



Figure 2: Generated by DD2-VAR-d20 model.



Figure 3: Generated by DD2-VAR-d24 model.



Figure 4: Generated by DD2-LlamaGen-L model.



(a) Images generated by DD2 model.



(b) Images generated by teacher model.

Figure 5: Images generated by VAR-d16 models.



(a) Images generated by DD2 model.



(b) Images generated by teacher model.

Figure 6: Images generated by VAR-d20 models.

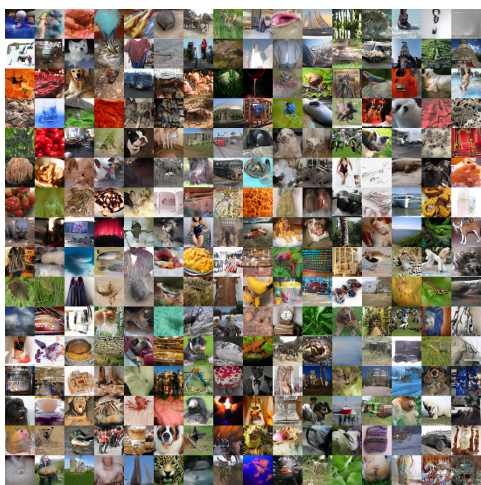


(a) Images generated by DD2 model.

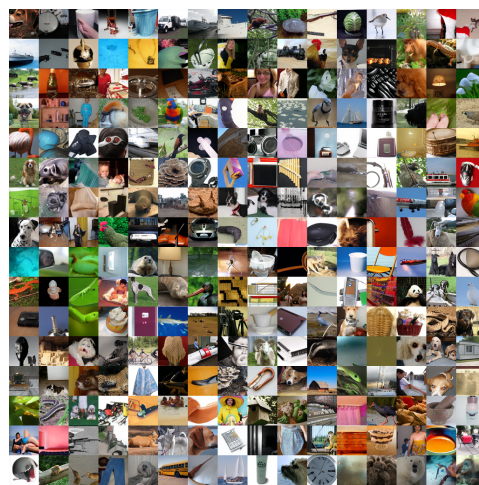


(b) Images generated by teacher model.

Figure 7: Images generated by VAR-d24 models.



(a) Images generated by DD2 model.



(b) Images generated by teacher model.

Figure 8: Images generated by LlamaGen-L models.