8 Additional Results

Varying Multi-Step Look Ahead n_{step}. Section 3 shows that at each denoising step, $\hat{x}_0$ can be predicted in a single step by Equation 1. However, the results displayed in Figure 10 indicate that the 1-step prediction may be blurry and inaccurate. As more DDIM sampling steps are taken, the image quality increases, leading to an improvement in the similarity embedding and the propulsive guidance gradient. Table 5 further illustrates that while ProCreate still outperforms DDIM sampling when the Multi-Step Look Ahead method is not used ($n_{step} = 1$), the method significantly improves FID, KID, and diversity-focused metrics at higher n_{step} settings.

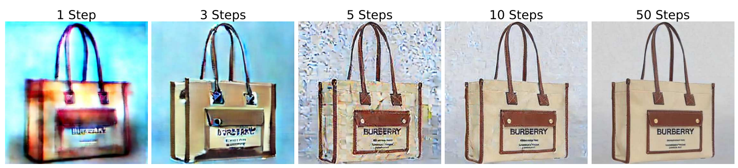

Fig. 10: Predictions for "Burberry Tote Bag" using different numbers of DDIM sampling steps from the evaluated Burberry checkpoint in Table 1.

Table 5: The effect of Multi-Step Look Ahead n_{step} on ProCreate's Performance. $n_{step} = 0$ represents baseline DDIM sampling that does not use ProCreate.

MSLA n_{step}	FID ↓	KID ↓	Precision ↑	Recall ↑	MSS ↓	Vendi ↑	Prompt Fid. ↑
0	35.11 ± 2.60	4.06 ± 0.48	**0.69 ± 0.06**	0.71 ± 0.05	0.18 ± 0.00	26.15 ± 0.56	**0.34 ± 0.00**
1	29.99 ± 1.49	2.64 ± 0.08	0.60 ± 0.06	0.91 ± 0.05	0.13 ± 0.01	30.90 ± 0.69	0.33 ± 0.00
3	17.82 ± 2.99	0.79 ± 0.47	0.64 ± 0.02	0.91 ± 0.04	0.10 ± 0.01	33.51 ± 0.58	0.33 ± 0.00
5	**14.10 ± 1.64**	**0.72 ± 0.11**	0.66 ± 0.06	**0.97 ± 0.02**	**0.09 ± 0.00**	**34.12 ± 0.36**	0.33 ± 0.00

Using Different Base Diffusion Samplers. We experiment with different base diffusion samplers using the standard fine-tuned "Burberry Designs" checkpoint from Table 1. Table 6 shows that although the DDPM and PNDM base samplers does not perform as well as the DDIM sampler on generative modeling metrics, we can apply ProCreate to them to improve their output diversity while maintaining sample quality.

Varying the Number of Training Samples. We conduct a quantitative evaluation for model checkpoints that are trained on 5 and 25 samples of each FSCG-8 subset respectively. In both cases, the models are trained for 2000 iterations. For the 5-shot fine-tuning runs, there are 45 validation samples, while for the 25-shot fine-tuning runs, there are 25 validation samples. We generate the same number of images as the number of validation samples in each evaluation run. The results are presented in Table 7 and 8. As shown in the tables, ProCreate achieves the best performance in terms of matching the validation distribution (FID, KID) and generating diverse samples (Recall, MSS, Vendi Score). ProCreate also remains competitive in quality-focused metrics (Precision and Prompt Fidelity).

Table 6: Quantitative comparison between DDIM, CADS, and ProCreate on different base samplers DDIM, DDPM, and PNDM. The evaluation is performed with standard fine-tuning, "Burberry Designs" images, and a 10 training and 40 validation split.

Sampler	**Method**	**FID ↓**	**KID ↓**	**Precision ↑**	**Recall ↑**	**MSS ↓**	**Vendi ↑**	**Prompt Fid. ↑**
DDIM	Default	35.11 ± 2.60	4.06 ± 0.48	0.69 ± 0.06	0.71 ± 0.05	0.18 ± 0.00	26.15 ± 0.56	**0.34 ± 0.00**
	CADS	34.39 ± 1.97	4.27 ± 0.41	**0.70 ± 0.07**	0.81 ± 0.06	0.18 ± 0.00	25.97 ± 0.20	**0.34 ± 0.00**
	ProCreate	**14.10 ± 1.64**	**0.72 ± 0.11**	0.66 ± 0.06	**0.97 ± 0.02**	**0.10 ± 0.00**	**33.51 ± 0.36**	0.33 ± 0.00
DDPM	Default	37.29 ± 1.97	4.08 ± 0.27	**0.68 ± 0.05**	0.70 ± 0.07	0.19 ± 0.01	26.05 ± 0.68	**0.34 ± 0.00**
	CADS	37.24 ± 1.81	4.36 ± 0.31	**0.68 ± 0.03**	0.73 ± 0.08	0.19 ± 0.01	25.62 ± 0.48	**0.34 ± 0.00**
	ProCreate	**14.80 ± 2.18**	**2.03 ± 0.22**	0.67 ± 0.08	**0.95 ± 0.03**	**0.12 ± 0.01**	**31.74 ± 0.62**	**0.34 ± 0.00**
PNDM	Default	60.86 ± 2.22	8.94 ± 0.53	**0.64 ± 0.09**	0.49 ± 0.10	0.22 ± 0.01	24.51 ± 1.56	**0.30 ± 0.01**
	CADS	63.24 ± 1.47	9.44 ± 0.42	**0.64 ± 0.04**	0.56 ± 0.06	0.23 ± 0.01	23.50 ± 0.54	**0.30 ± 0.00**
	ProCreate	**28.64 ± 2.09**	**2.22 ± 0.34**	0.62 ± 0.05	**0.92 ± 0.04**	**0.11 ± 0.01**	**32.08 ± 0.44**	**0.30 ± 0.00**

DreamBooth Qualitative Samples. In Figure 11, we perform DreamBooth fine-tuning with the experiment setup from Section 5.2 to produce qualitative results with DDIM, CADS, and ProCreate sampling methods. Similarly to the qualitative comparisons in Figure 4, ProCreate generates the most diverse samples that follow the prompts yet do not replicate the Top-1 SSCD matched training images.

Table 7: Quantitative comparison between DDIM, CADS, and ProCreate applied on standard fine-tuning checkpoints for 5-shot learning on various generative modeling metrics. We show the mean and standard deviation values over 5 repeated runs in each cell.

Subset	Method	FID ↓	KID ↓	Precision ↑	Recall ↑	MSS ↓	Vendi ↑	Prompt Fid. ↑
	DDIM	11.27 ± 0.69	2.02 ± 0.18	**0.81 ± 0.03**	0.44 ± 0.12	0.36 ± 0.01	14.32 ± 0.70	**0.34 ± 0.00**
Amedeo	CADS	12.51 ± 0.68	2.43 ± 0.22	0.80 ± 0.04	0.39 ± 0.08	0.36 ± 0.02	14.78 ± 0.96	**0.34 ± 0.00**
	ProCreate	**10.23 ± 0.66**	**0.21 ± 0.03**	0.69 ± 0.06	**0.55 ± 0.17**	**0.19 ± 0.01**	**30.46 ± 0.84**	0.33 ± 0.01
	DDIM	19.68 ± 2.39	0.36 ± 0.13	0.63 ± 0.04	0.70 ± 0.12	**0.11 ± 0.00**	36.11 ± 0.34	**0.30 ± 0.00**
Apple	CADS	27.83 ± 1.34	0.83 ± 0.11	0.48 ± 0.02	0.71 ± 0.12	**0.11 ± 0.01**	36.65 ± 0.75	0.29 ± 0.00
	ProCreate	**14.55 ± 1.49**	**0.12 ± 0.03**	**0.64 ± 0.05**	**0.78 ± 0.10**	**0.11 ± 0.01**	**36.70 ± 0.63**	**0.30 ± 0.00**
	DDIM	10.43 ± 0.58	0.37 ± 0.02	**0.75 ± 0.04**	0.73 ± 0.05	0.30 ± 0.01	16.80 ± 0.45	**0.35 ± 0.00**
Burberry	CADS	12.31 ± 0.69	0.46 ± 0.04	0.69 ± 0.03	0.69 ± 0.04	0.30 ± 0.01	16.70 ± 0.68	**0.35 ± 0.00**
	ProCreate	**9.61 ± 0.51**	**0.27 ± 0.04**	0.69 ± 0.04	**0.89 ± 0.07**	**0.21 ± 0.01**	**23.31 ± 0.56**	**0.35 ± 0.00**
	DDIM	8.76 ± 0.39	1.31 ± 0.11	0.99 ± 0.01	0.52 ± 0.05	0.23 ± 0.00	20.81 ± 0.55	**0.29 ± 0.00**
Frank	CADS	10.06 ± 0.28	1.61 ± 0.04	**1.00 ± 0.01**	0.55 ± 0.07	0.24 ± 0.01	21.18 ± 0.52	**0.29 ± 0.00**
	ProCreate	**5.09 ± 0.33**	**0.57 ± 0.14**	0.93 ± 0.03	**0.62 ± 0.10**	**0.16 ± 0.01**	**29.41 ± 1.66**	0.28 ± 0.00
	DDIM	2.88 ± 0.43	0.03 ± 0.00	0.12 ± 0.05	0.71 ± 0.17	0.48 ± 0.01	11.90 ± 0.44	**0.24 ± 0.00**
Nouns	CADS	3.11 ± 0.46	0.03 ± 0.00	0.11 ± 0.02	0.63 ± 0.12	0.48 ± 0.01	11.78 ± 0.41	**0.24 ± 0.00**
	ProCreate	**2.64 ± 0.31**	**0.02 ± 0.00**	**0.15 ± 0.05**	**0.79 ± 0.11**	**0.44 ± 0.01**	**13.65 ± 0.45**	**0.24 ± 0.00**
	DDIM	9.17 ± 0.21	1.24 ± 0.03	0.72 ± 0.06	0.20 ± 0.03	0.29 ± 0.01	22.56 ± 0.41	**0.30 ± 0.00**
Onepiece	CADS	10.56 ± 0.99	1.52 ± 0.17	0.70 ± 0.05	0.19 ± 0.04	0.30 ± 0.01	22.46 ± 0.51	**0.30 ± 0.00**
	ProCreate	**5.57 ± 0.64**	**0.58 ± 0.10**	**0.73 ± 0.06**	**0.42 ± 0.06**	**0.25 ± 0.01**	**26.67 ± 0.62**	**0.30 ± 0.00**
	DDIM	**13.23 ± 0.43**	0.24 ± 0.02	0.46 ± 0.08	0.84 ± 0.07	0.28 ± 0.01	21.94 ± 0.58	**0.31 ± 0.00**
Pokemon	CADS	17.05 ± 0.40	0.29 ± 0.04	0.44 ± 0.03	0.76 ± 0.12	0.29 ± 0.01	21.98 ± 0.76	**0.31 ± 0.00**
	ProCreate	13.54 ± 0.63	**0.22 ± 0.01**	**0.47 ± 0.09**	**0.88 ± 0.07**	**0.23 ± 0.00**	**23.13 ± 0.22**	**0.31 ± 0.00**
	DDIM	29.56 ± 1.66	7.01 ± 0.51	0.93 ± 0.03	0.45 ± 0.07	0.20 ± 0.00	19.09 ± 0.32	0.30 ± 0.00
Rococo	CADS	33.52 ± 0.63	8.13 ± 0.18	**0.94 ± 0.02**	0.52 ± 0.09	0.21 ± 0.00	18.74 ± 0.57	0.30 ± 0.00
	ProCreate	**23.37 ± 0.36**	**5.07 ± 0.24**	**0.94 ± 0.03**	**0.80 ± 0.04**	**0.17 ± 0.01**	**33.74 ± 0.82**	**0.31 ± 0.00**

Table 8: Quantitative comparison between DDIM, CADS, and ProCreate applied on standard fine-tuning checkpoints for 25-shot learning on various generative modeling metrics. We show the mean and standard deviation values over 5 repeated runs in each cell.

Subset	Method	FID ↓	KID ↓	Precision ↑	Recall ↑	MSS ↓	Vendi ↑	Prompt Fid. ↑
	DDIM	14.87 ± 1.42	3.01 ± 0.37	**0.66 ± 0.07**	0.69 ± 0.19	0.30 ± 0.01	14.07 ± 0.41	**0.34 ± 0.00**
Amedeo	CADS	16.51 ± 0.98	3.48 ± 0.32	0.58 ± 0.08	0.57 ± 0.19	0.32 ± 0.01	13.68 ± 0.26	**0.34 ± 0.00**
	ProCreate	**12.22 ± 1.04**	**2.89 ± 0.24**	0.41 ± 0.07	**0.75 ± 0.05**	**0.29 ± 0.01**	**16.03 ± 0.30**	0.33 ± 0.01
	DDIM	45.35 ± 2.07	2.77 ± 0.22	**0.58 ± 0.03**	0.76 ± 0.08	0.18 ± 0.01	17.85 ± 0.54	**0.31 ± 0.00**
Apple	CADS	30.89 ± 0.35	2.73 ± 0.09	**0.58 ± 0.08**	0.78 ± 0.05	0.19 ± 0.01	17.13 ± 0.52	**0.31 ± 0.00**
	ProCreate	**25.16 ± 1.07**	**2.48 ± 0.14**	**0.58 ± 0.05**	**0.80 ± 0.08**	**0.14 ± 0.01**	**20.01 ± 0.64**	**0.31 ± 0.00**
	DDIM	16.44 ± 0.18	1.39 ± 0.19	**0.95 ± 0.02**	0.82 ± 0.12	0.27 ± 0.01	13.76 ± 0.42	**0.34 ± 0.00**
Burberry	CADS	17.87 ± 1.91	1.51 ± 0.25	0.94 ± 0.03	0.73 ± 0.09	0.26 ± 0.01	14.06 ± 0.42	**0.34 ± 0.00**
	ProCreate	**6.72 ± 1.84**	**0.82 ± 0.26**	0.89 ± 0.07	**0.93 ± 0.07**	**0.16 ± 0.01**	**20.38 ± 0.61**	**0.34 ± 0.00**
	DDIM	6.13 ± 0.28	0.62 ± 0.09	**0.98 ± 0.00**	0.81 ± 0.04	**0.14 ± 0.00**	19.77 ± 0.47	**0.31 ± 0.00**
Frank	CADS	**4.57 ± 0.10**	0.60 ± 0.07	0.96 ± 0.00	0.82 ± 0.04	**0.14 ± 0.00**	19.50 ± 0.21	**0.31 ± 0.00**
	ProCreate	4.90 ± 0.19	**0.34 ± 0.02**	0.93 ± 0.09	**0.83 ± 0.03**	**0.14 ± 0.00**	**21.73 ± 0.26**	**0.31 ± 0.01**
	DDIM	1.81 ± 0.09	0.02 ± 0.00	0.59 ± 0.09	0.95 ± 0.05	0.50 ± 0.01	8.58 ± 0.14	**0.25 ± 0.00**
Nouns	CADS	1.24 ± 0.05	0.02 ± 0.00	**0.66 ± 0.12**	0.97 ± 0.05	0.51 ± 0.01	8.34 ± 0.18	**0.25 ± 0.00**
	ProCreate	**1.22 ± 0.01**	**0.01 ± 0.00**	0.62 ± 0.08	**0.98 ± 0.03**	**0.46 ± 0.01**	**14.55 ± 0.19**	**0.25 ± 0.00**
	DDIM	2.85 ± 0.13	0.10 ± 0.02	**0.83 ± 0.04**	0.69 ± 0.05	**0.25 ± 0.00**	16.86 ± 0.15	0.30 ± 0.00
Onepiece	CADS	3.02 ± 0.16	0.15 ± 0.01	0.78 ± 0.03	0.75 ± 0.02	0.26 ± 0.01	16.81 ± 0.29	**0.31 ± 0.00**
	ProCreate	**2.14 ± 0.11**	**0.09 ± 0.01**	0.82 ± 0.04	**0.79 ± 0.01**	**0.25 ± 0.01**	**18.05 ± 0.45**	0.30 ± 0.00
	DDIM	6.32 ± 0.34	**0.05 ± 0.01**	0.78 ± 0.05	0.78 ± 0.05	0.31 ± 0.01	14.89 ± 0.44	**0.31 ± 0.00**
Pokemon	CADS	8.20 ± 0.30	0.06 ± 0.00	**0.86 ± 0.03**	0.74 ± 0.03	0.31 ± 0.02	14.77 ± 0.67	**0.31 ± 0.00**
	ProCreate	**4.88 ± 0.93**	**0.05 ± 0.00**	0.82 ± 0.05	**0.86 ± 0.03**	**0.28 ± 0.01**	**15.95 ± 0.42**	**0.31 ± 0.00**
	DDIM	10.30 ± 0.58	1.82 ± 0.09	0.83 ± 0.02	0.94 ± 0.02	0.15 ± 0.00	15.08 ± 0.37	**0.34 ± 0.00**
Rococo	CADS	13.36 ± 0.61	1.85 ± 0.11	0.83 ± 0.02	0.92 ± 0.00	0.15 ± 0.00	15.15 ± 0.34	**0.34 ± 0.00**
	ProCreate	**9.41 ± 1.06**	**1.63 ± 0.07**	**0.86 ± 0.03**	**0.97 ± 0.01**	**0.14 ± 0.01**	**15.89 ± 0.84**	**0.34 ± 0.00**

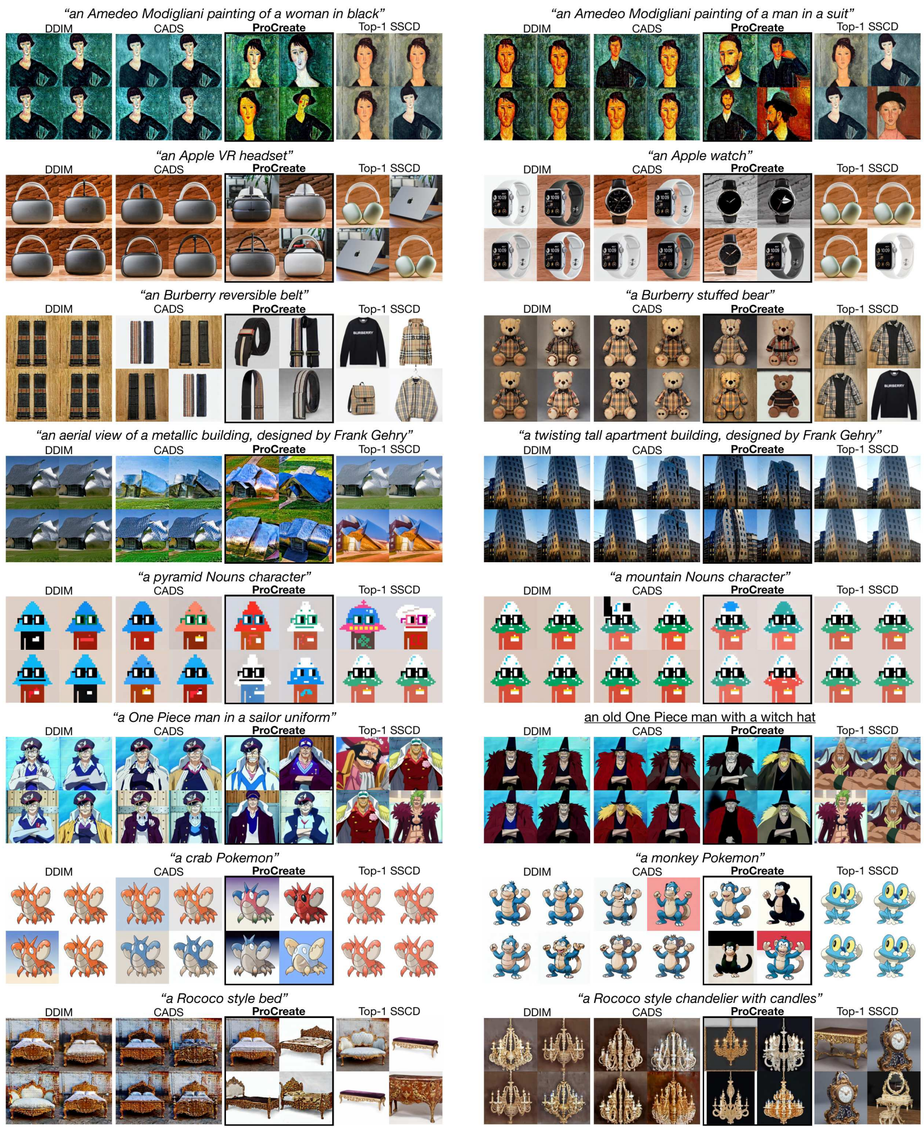

Fig. 11: Qualitative comparison between DDIM, CADS, and ProCreate for few-shot creative generation on FSCG-8 with DreamBooth fine-tuning. For each sampling method, we show two prompts and four generated samples for each prompt. In addition, we match each sample from ProCreate with its closest training image based on the SSCD score [32] between the matched pair.