
One Step Diffusion via Flow Fitting

Anonymous Author(s)

Affiliation

Address

email

Table 1: Performance comparison of various training paradigms applied to a unified model architecture (**DiT-B**) on the class-conditioned ImageNet-256 dataset. FID-50k scores are reported for different denoising schedules: 128-step, 4-step, and single-step. Lower scores denote better fidelity. Our method, FlowFit, produces high-quality outputs in a single forward pass, reducing the performance gap with methods utilizing multi-stage distillation. Parentheses indicate out-of-distribution evaluations.

	ImageNet-256 (Class-Conditional)		
	128-Step	4-Step	1-Step
Two phase training			
Progressive Distillation	(201.9)	(142.5)	35.6
Consistency Distillation	132.8	98.01	136.5
Reflow	16.9	32.8	44.8
End-to-end (single training run)			
Diffusion	39.7	(464.5)	(467.2)
Flow Matching	17.3	(108.2)	(324.8)
Consistency Training	42.8	43.0	69.7
Live Reflow (ours)	46.3	95.8	58.1
Shortcut Models	15.5	28.3	40.3
FlowFit (ours)	-	-	34.4

1 A key question in FlowFit is whether the proposed basis expansion reliably approximates the target
2 flow. The following proposition confirms that this is indeed the case.

3 **Proposition 1** (Universal Approximation of Flow Trajectories). *Let $\psi : (\mathbb{R}^d \times [0, 1]) \rightarrow \mathbb{R}^d$ be a*
4 *continuous trajectory from an initial point x_0 to a target point x_1 . Then, for any $\varepsilon > 0$, there exists*
5 *a sufficiently large integer $N > 0$, a set of basis functions $\{\gamma(t)\}_{i=1}^N$, and coefficients $\{W_i(x)\}_{i=1}^N$*
6 *such that*

$$\left\| \psi(x, t) - \left(\sum_{i=1}^N \gamma_i(t) \cdot W_i(x) \right) \right\| < \varepsilon, \quad \forall t \in [0, 1].$$

7 *Proof.* The results is immediately obtained by applying the Stone–Weierstrass theorem [2, 1], and
8 because \mathbb{R}^d and $[0, 1]$ are both locally compact Hausdorff spaces, and the basis functions and the
9 coefficients are all continuous. \square

10 1 Evaluation of Class-Conditional Generation on ImageNet-256

11 Table 1 showcases the performance of various generative learning objectives under class-conditioning
12 on ImageNet-256. Our proposed approach, FlowFit, attains strong performance in the single-step
13 regime, surpassing the end-to-end counterparts. For ImageNet experiments, we use the classifier-free
14 guidance (CFG) [3] is employed to enhance conditional generation.



Figure 1: Unfiltered samples generated on the unconditional CelebA-HQ dataset at a resolution of 256×256 . These images were produced in a single forward pass using a DiT-B model trained for 400,000 iterations.

2 Qualitative Samples

Figures 1 and 2 present sample outputs from models trained on CelebA-HQ (unconditional) and ImageNet (class-conditioned), respectively, using our proposed training procedure.

3 Implementation Details

Detailed training configurations and hyperparameters corresponding to the results presented in the main text (Table 1) and supplementary (Table 1) are provided in Table 2.

References

- [1] Errett Bishop. A generalization of the stone-weierstrass theorem. 1961.

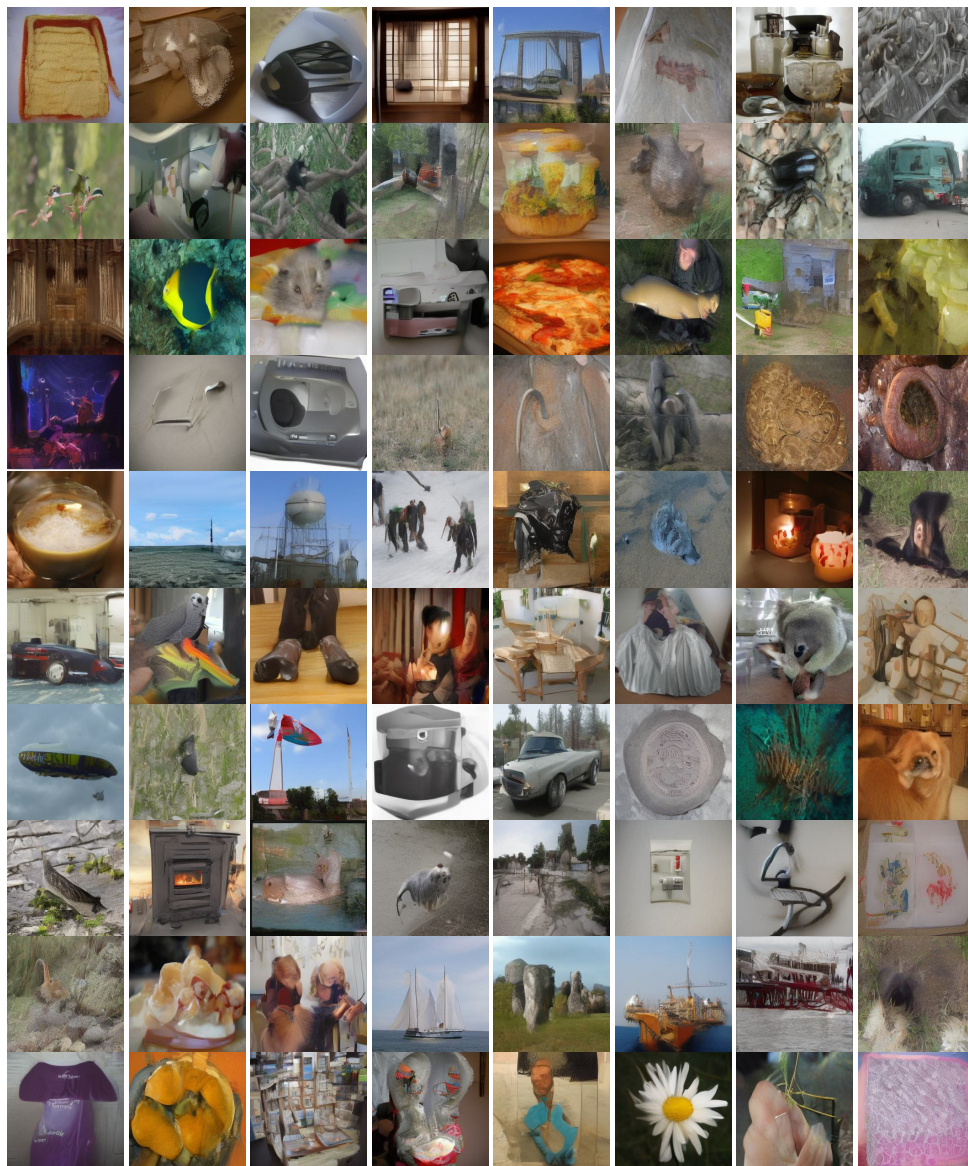


Figure 2: Unfiltered samples generated on the unconditional ImageNet dataset at a resolution of 256×256 . These images were produced in a single forward pass using a DiT-B model trained for 800,000 iterations.

Batch Size	64 (CelebA-HQ), 256 (Imagenet)
Training Steps	400,000 (CelebA-HQ), 800,000 (Imagenet)
Latent Encoder	sd-vae-mse-ft
Latent Downsampling	8 (256x256x3 to 32x32x4)
Classifier Free Guidance	0 (CelebA-HQ), 1.5 (Imagenet)
Class Dropout Probability	0 (CelebA-HQ), 0.1 (Imagenet)
EMA Parameters Used For Evaluation?	Yes
EMA Ratio	0.999
Optimizer	AdamW
Learning Rate	0.00004
Weight Decay	0.0
Hidden Size	768
Patch Size	2
Number of Layers	12
Attention Heads	12
MLP Hidden Size Ratio	4
Basis	Polynomial
Basis order	8

Table 2: Default hyperparameter settings used during training.

- 23 [2] Louis De Branges. The stone-weierstrass theorem. *Proceedings of the American Mathematical*
24 *Society*, 10(5):822–824, 1959.
- 25 [3] Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint*
26 *arXiv:2207.12598*, 2022.