DuoDiff: Accelerating Diffusion Modules with a Dual-Backbone Approach

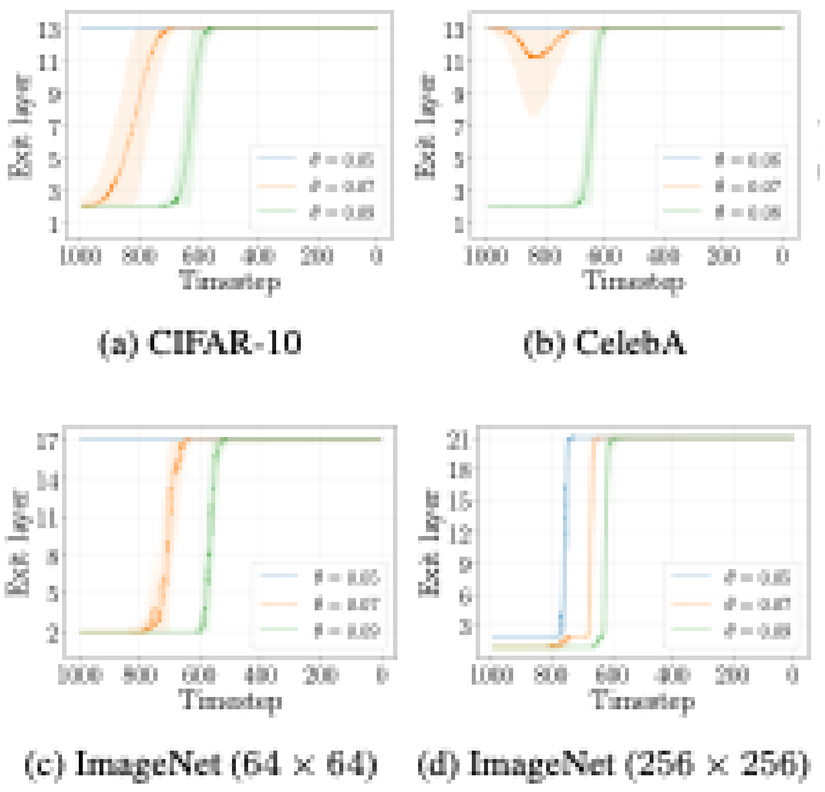
D. Gallo, R. Matişan, A. Monroy, A. Vasilcoiu, J. Partyka, T.H. Veljković, M. Jazbec

University of Amsterdam

NEURAL INFORMATION PROCESSING SYSTEMS

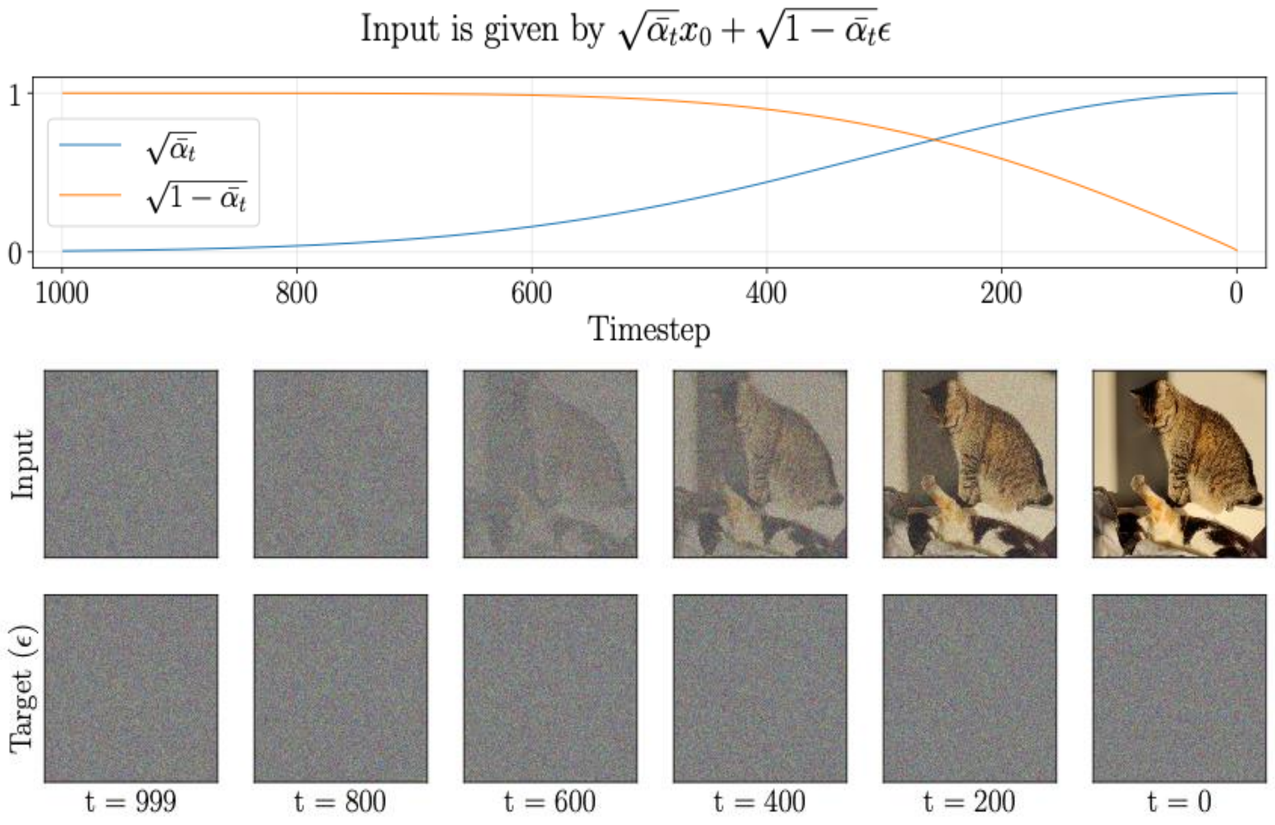
- **Diffusion models** require hundreds of iterations, resulting in **slow inference**
- AdaDiff [1] leverages **early exit** to speed up inference
- We show that **early exit trends are predictable** and introduce a **simpler model (DuoDiff)**

Early exit trends



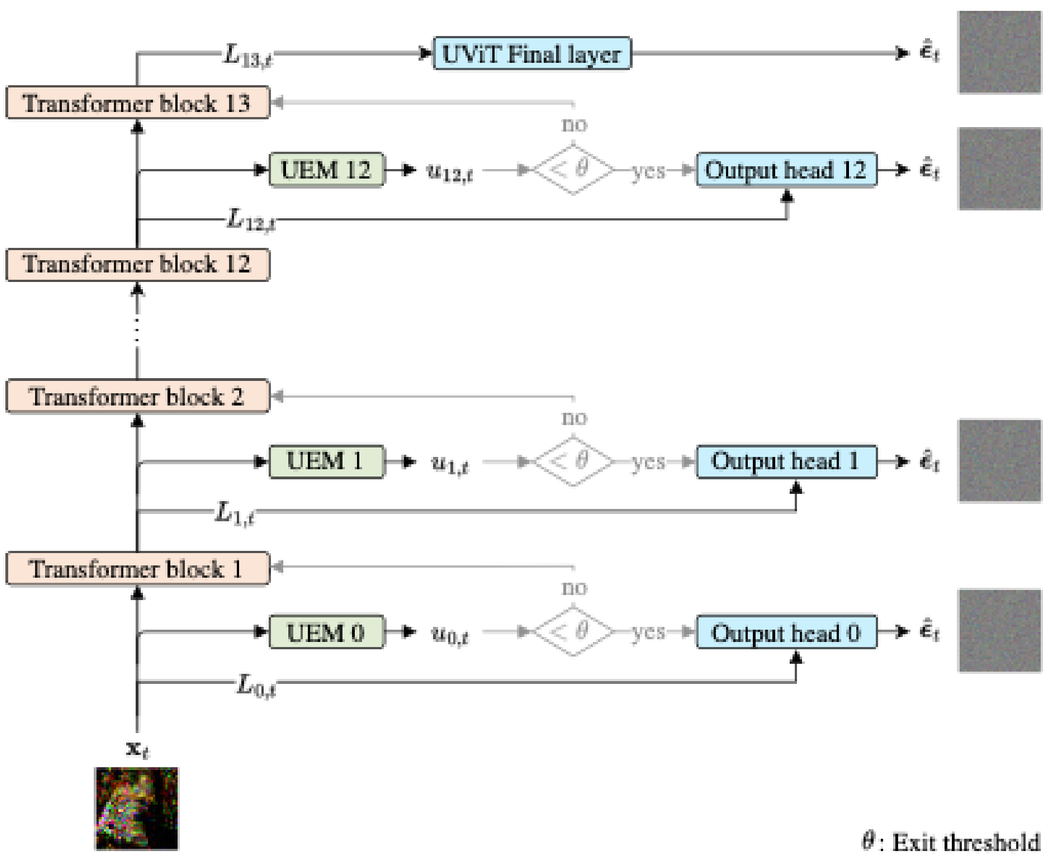
(a) CIFAR-10 (b) CelebA

(c) ImageNet (64 × 64) (d) ImageNet (256 × 256)

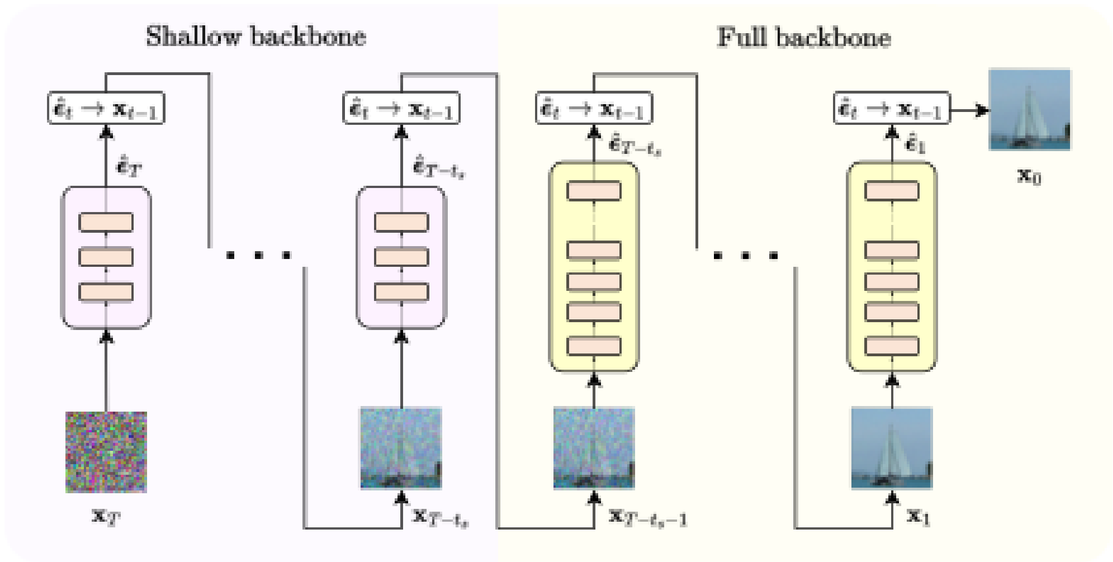


Models

AdaDiff

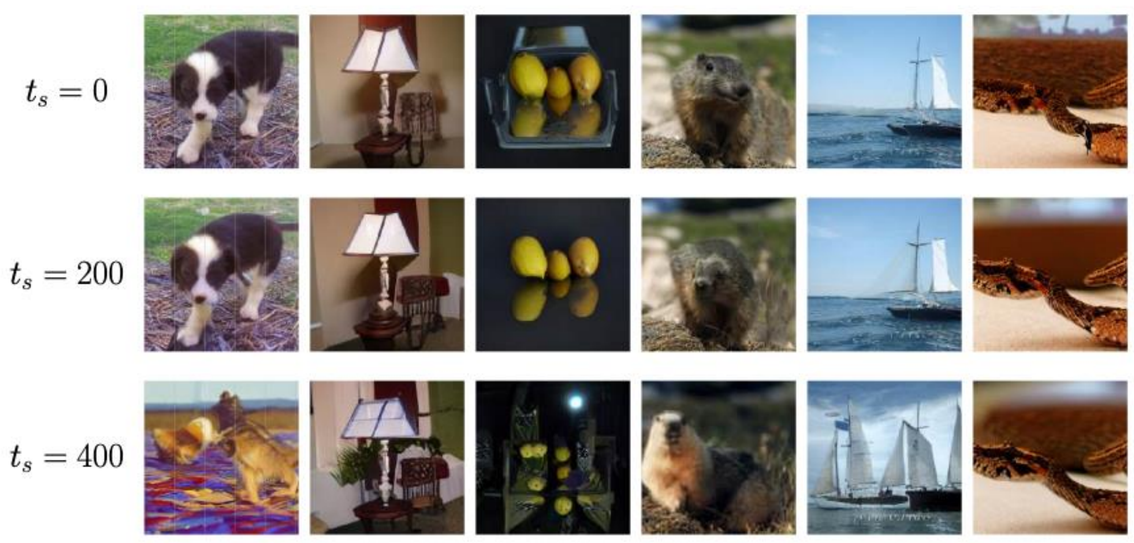


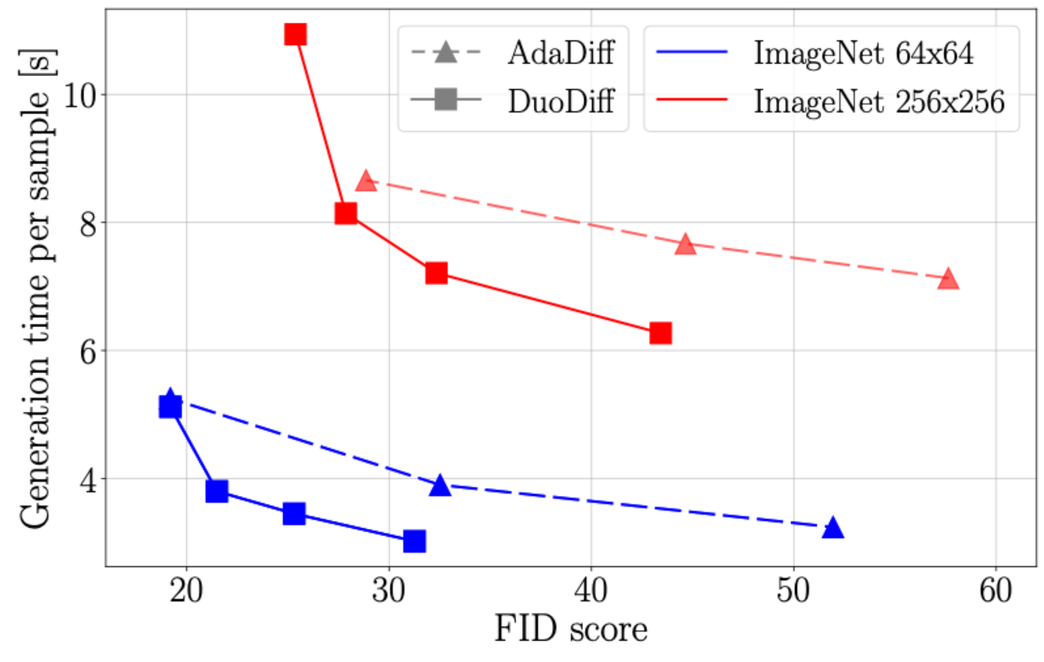
DuoDiff



Results

- ✓ DuoDiff outperforms AdaDiff in both image quality and inference time
- ✓ DuoDiff can be used alongside DDIM and latent space diffusion





💡 **The initial diffusion time-steps are almost-trivial**

💡 We can **replace a 21-layer transformer with a 3-layer transformer** for hundreds of time-steps with almost **no performance hit**!

[1] Zhang H, Wu Z, Xing Z, Shao J, Jiang YG. Adadiff: Adaptive step selection for fast diffusion. arXiv preprint arXiv:2311.14768. 2023 Nov 24.