

On Distillation of Guided Diffusion Models

Chenlin Meng¹

chenlin@cs.stanford.edu

Robin Rombach²

robin@stability.ai

Ruiqi Gao³

ruiqi@google.com

Diederik Kingma³

durk@google.com

Stefano Ermon¹

ermon@cs.stanford.edu

Jonathan Ho³

jonathanho@google.com

Tim Salimans³

salimans@google.com

¹Stanford University

²Stability AI & LMU Munich

³Google Research, Brain Team

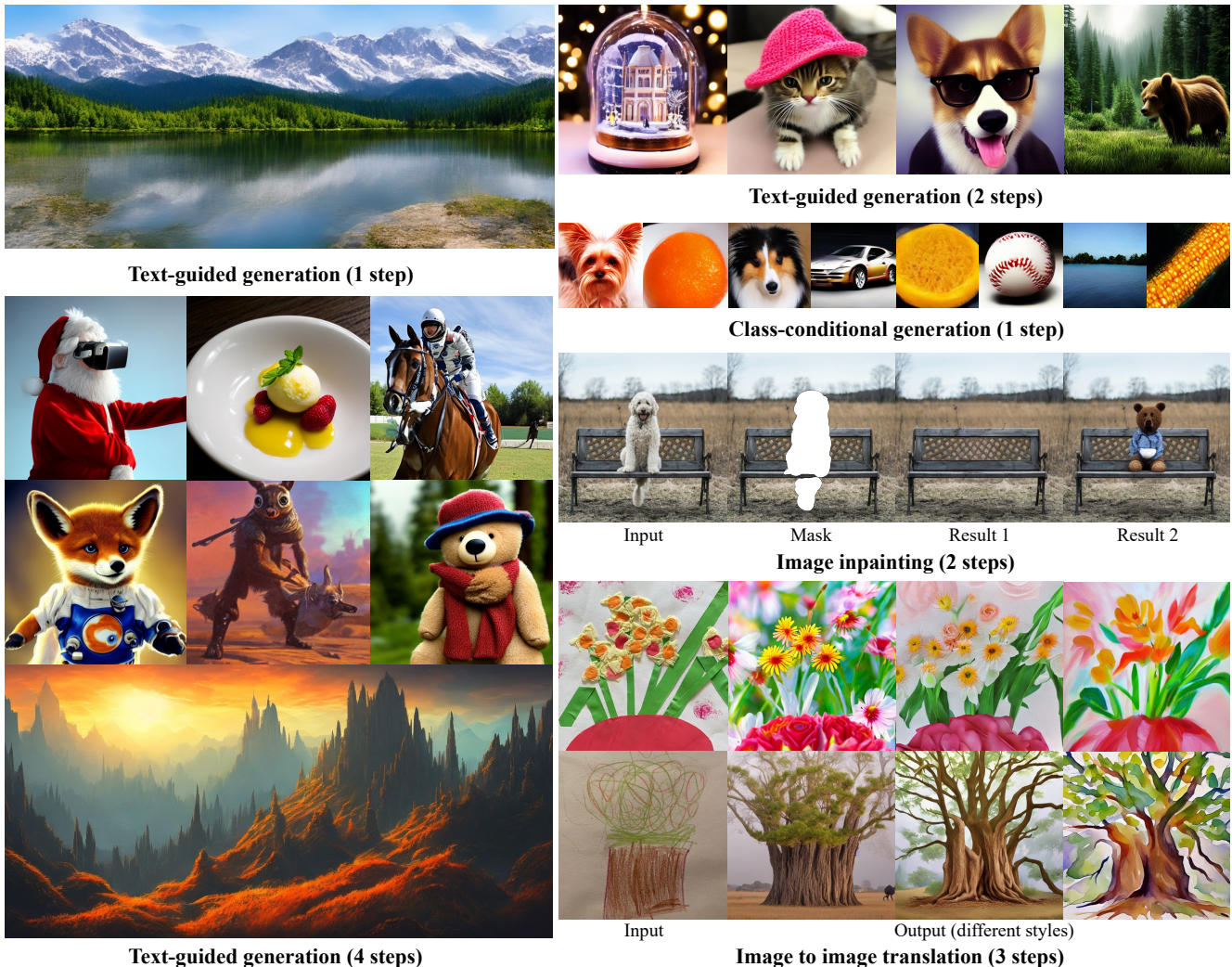


Figure 1. Distilled Stable Diffusion samples generated by our method. Our two-stage distillation approach is able to generate realistic images using only 1 to 4 denoising steps on various tasks. Compared to standard classifier-free guided diffusion models, we reduce the total number of sampling steps by at least 20 \times .

Abstract

Classifier-free guided diffusion models have recently been shown to be highly effective at high-resolution image generation, and they have been widely used in large-scale diffusion

frameworks including DALL-E 2, Stable Diffusion and Imagen. However, a downside of classifier-free guided diffusion models is that they are computationally expensive at inference time since they require evaluating two diffusion models, a class-conditional model and an unconditional model, tens to hundreds of times. To deal with this limitation, we pro-

*Work partially done during an internship at Google