

Flow Matching: Making small LDMs GREAT

github.com/CompVis/fm-boosting

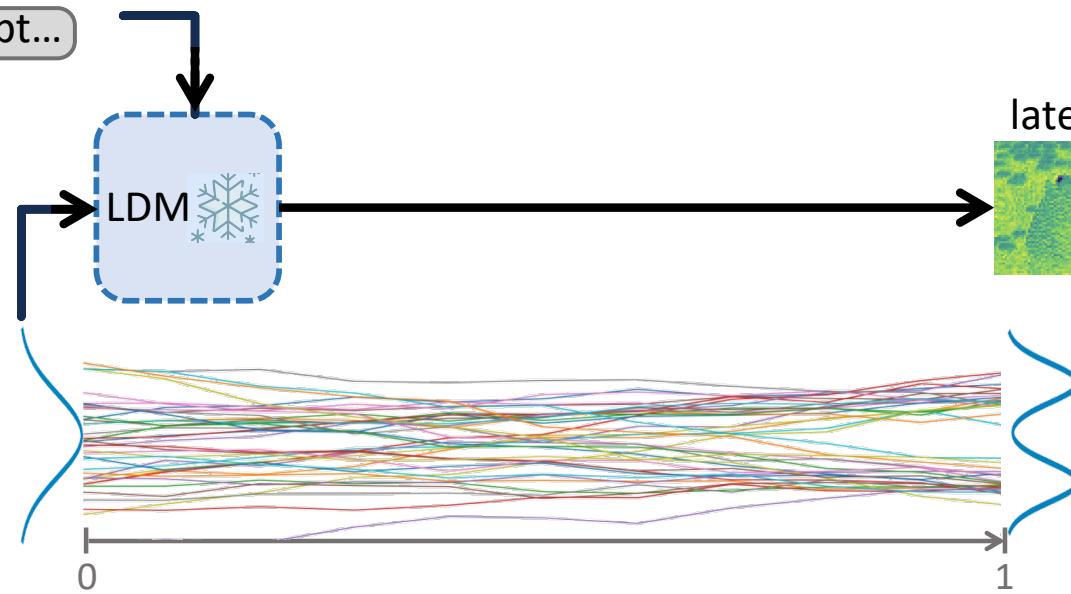
arXiv:2312.07360



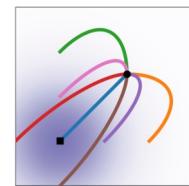
user control:

prompt...

Inference



high-res image



diffusion:
curved trajectories
✓ diversity
✗ slow sampling

WANT:

- high-resolution
- fast inference
- With existing small LDMs
- No costly distillation,
or re-training

conv. autoencoder:
perceptual compression
X only local details
✓ efficient

arXiv:2210.02747 / 2209.03003

Flow Matching: Making small LDMs GREAT

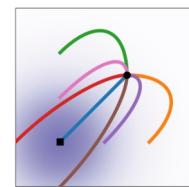
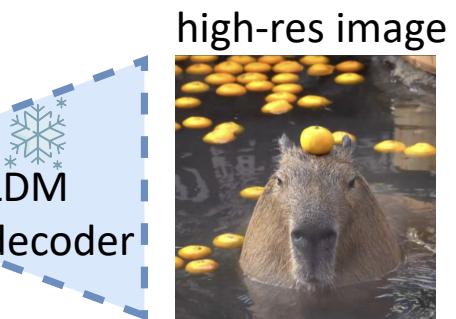
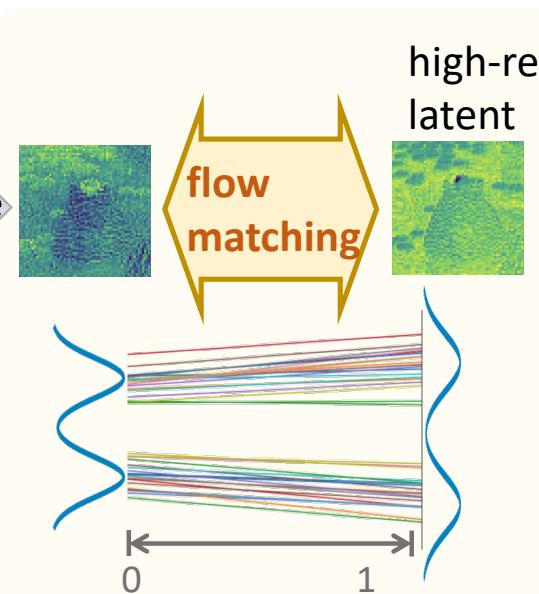
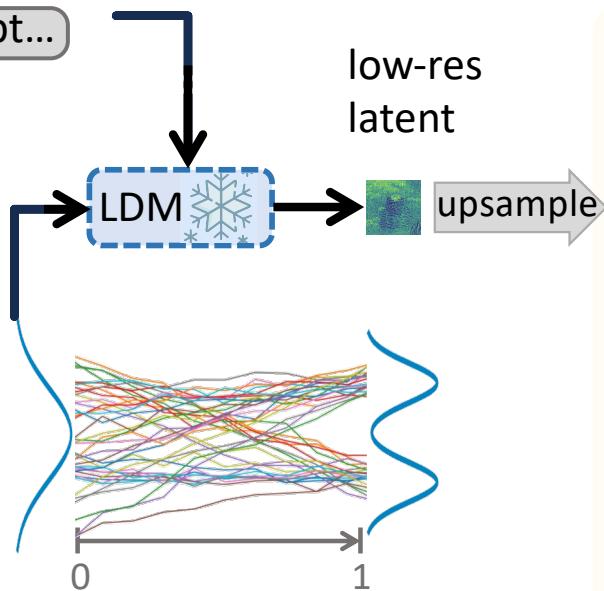
github.com/CompVis/fm-boosting

arXiv:2312.07360

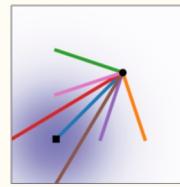
user control:

prompt...

Inference



diffusion:
curved trajectories
✓ diversity
✗ slow sampling



flow matching:
straightened paths
✗ less diversity
✓ fast sampling

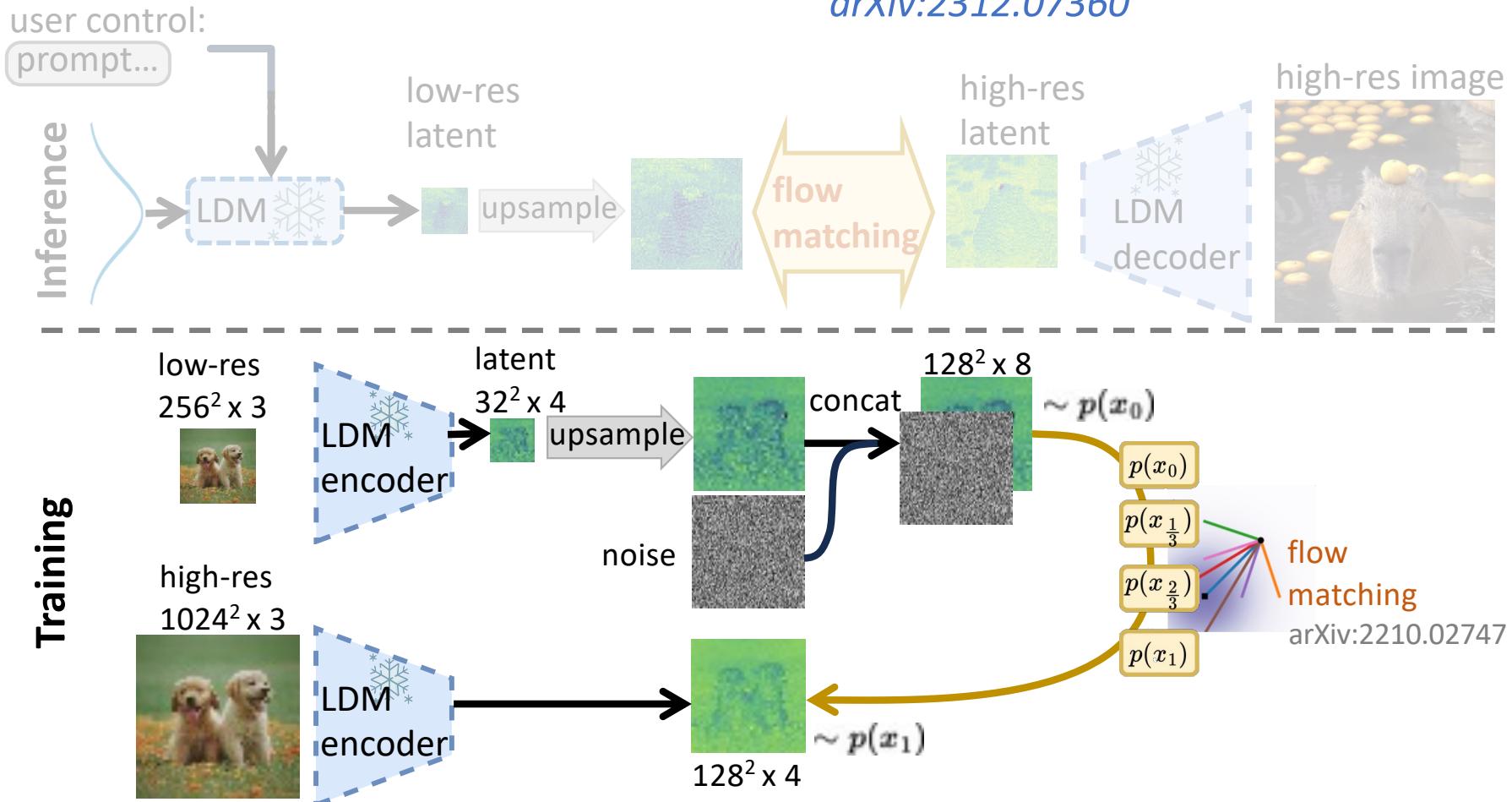
conv. autoencoder:
perceptual compression
✗ only local details
✓ efficient

arXiv:2210.02747 / 2209.03003

Flow Matching: Making small LDMs GREAT

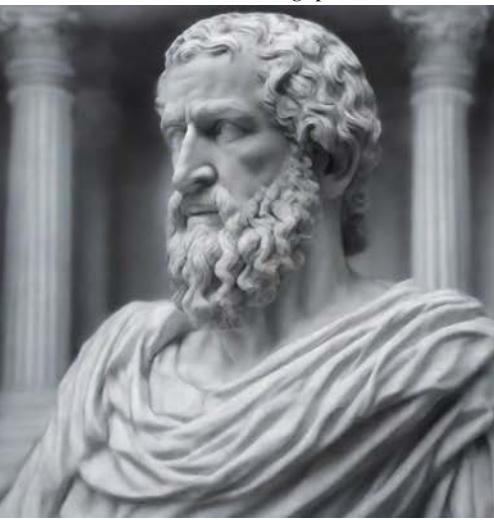
github.com/CompVis/fm-boosting

arXiv:2312.07360



Efficient Inference

LCM (SDXL)
time per image: 0.408s



Closeup Roman statue of Aristotle black and white made of marble, artstation award-winning, photorealistic

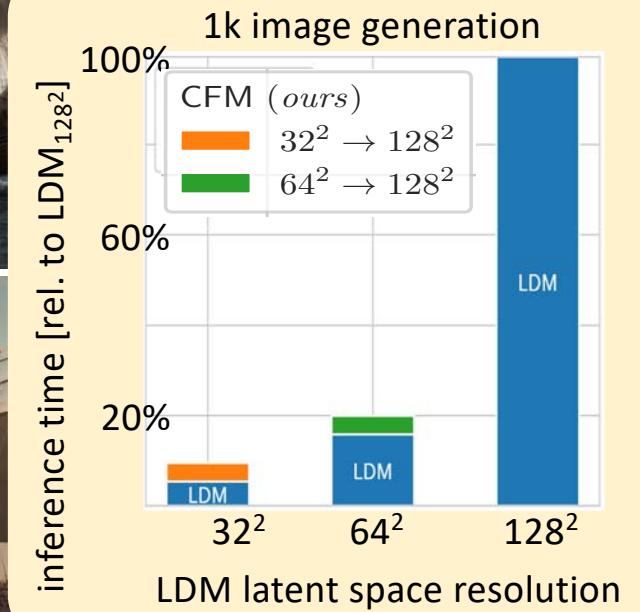
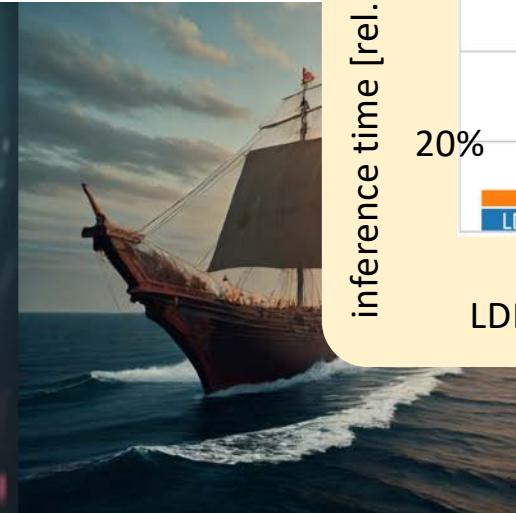


Anime art of ghost in the shell, detailed scene, red, perfect face



Viking ship sailing northeast, detailed, 8k

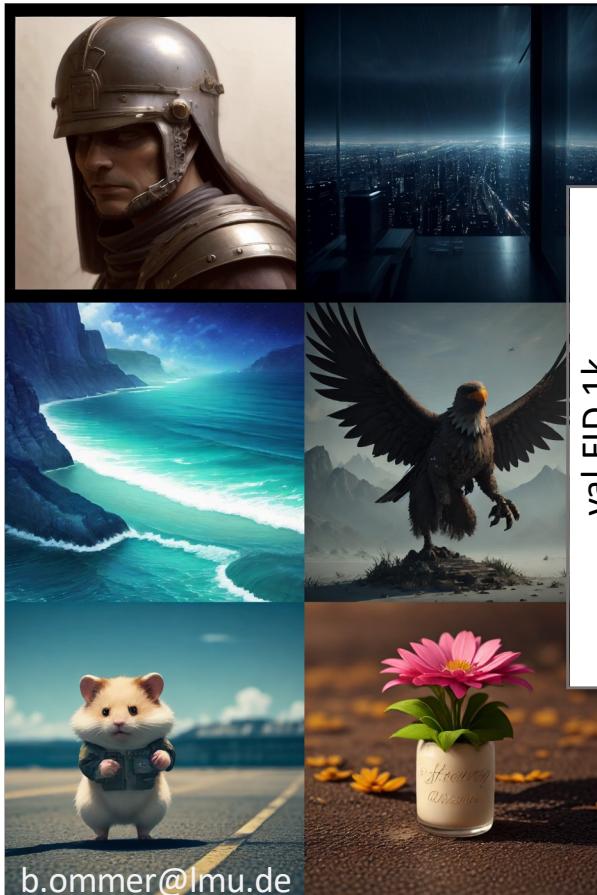
LCM (SD1.5) + CFM (Ours)
time per image: 0.347s



github.com/CompVis/fm-boosting

Sampling at 1k Resolution

LCM-LoRA + our CFM



SD1.5 + our CFM

