Cold Diffusion:

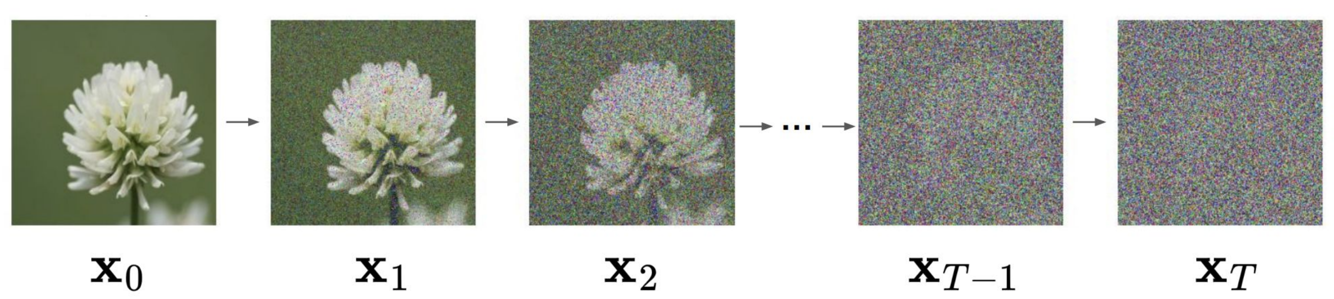
Inverting Arbitrary Image Transforms without Noise

Christopher Cheng*, Marc Davila*, Ryan Lee*, Benjamin Tang*, Oscar Wang*

*Cornell University

MOTIVATION

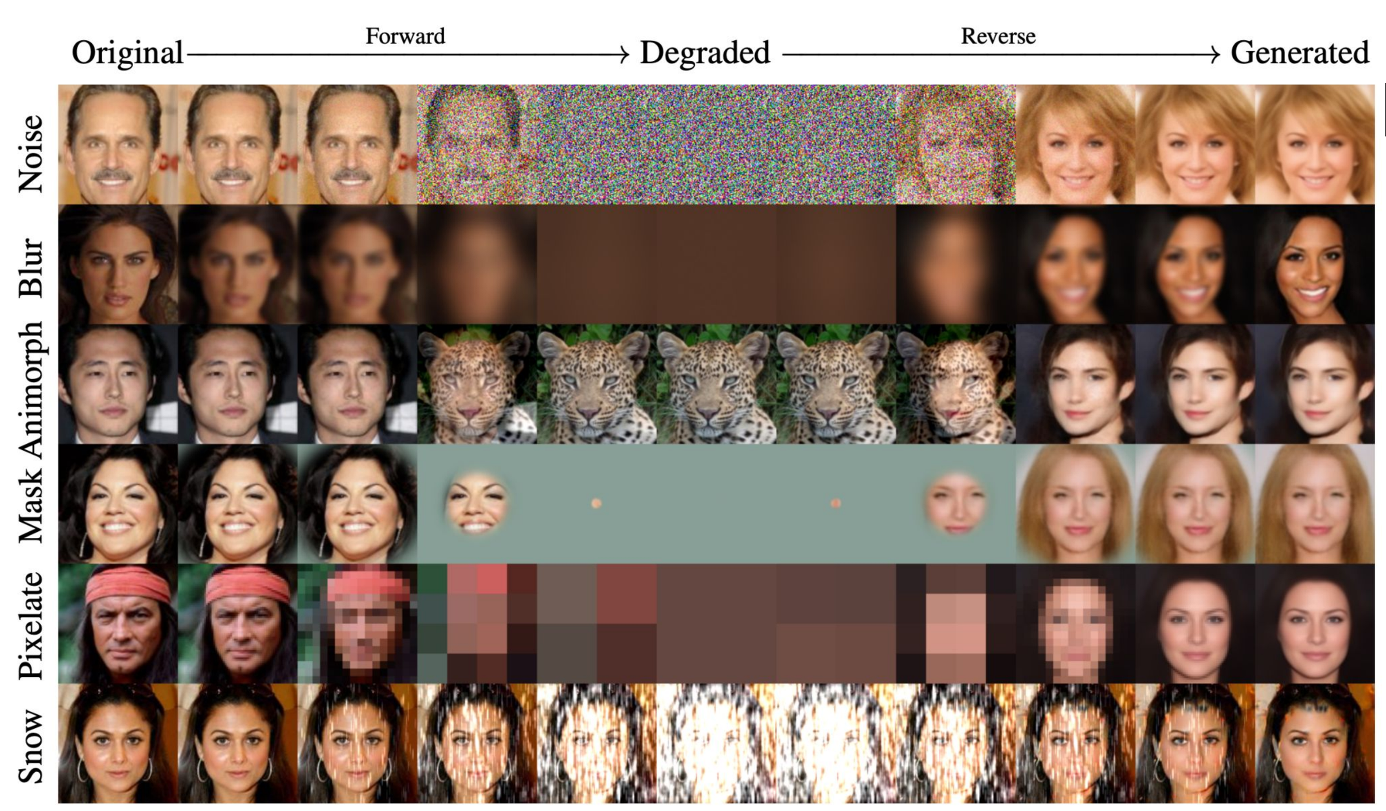
Recall **hot diffusion** = add noise and learn to remove it



[2]

But, why random noise?

Introducing **cold diffusion** = using deterministic degradations can work too!



[1]

Our goal is to **prove that cold diffusion works and can recover high-fidelity images**.

- We don't focus on unconditional generation, nor comparing it with hot diffusion.
- We also verify the efficacy of the new algorithm (Algorithm 2).

RESULTS

Generalized Diffusions

*our focus

Goal: Recover original image given a degraded one.

Deep dive in blurring (100k steps)

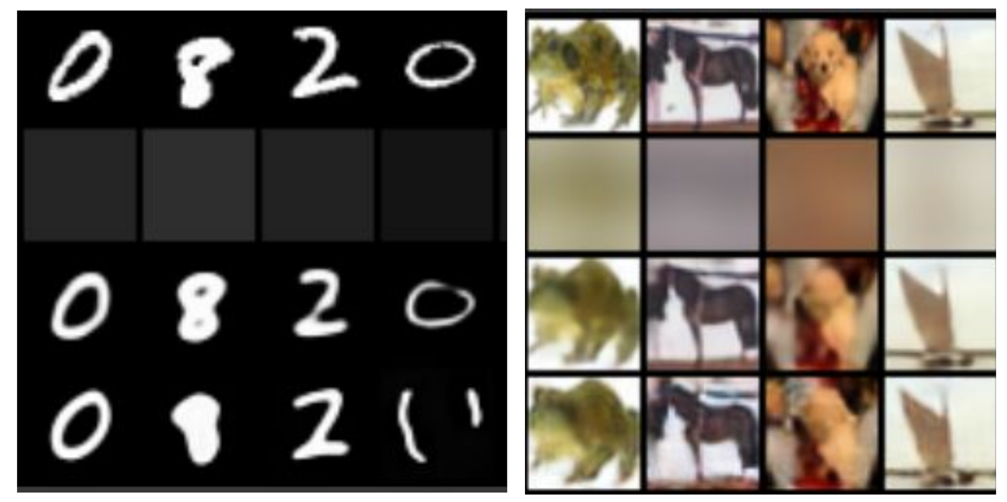


Fig 1. *Blurring* (~100k steps)

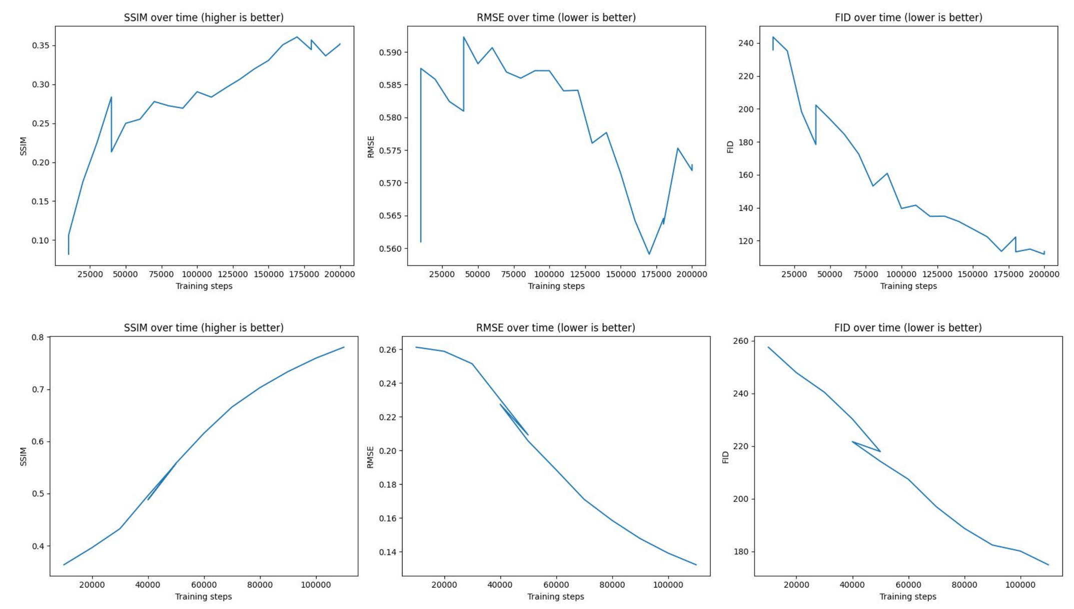


Fig 2. Evolution of FID, SSIM, RMSE

Dataset	Degraded			Direct			Sampled		
	FID	SSIM	RMSE	FID	SSIM	RMSE	FID	SSIM	RMSE
MNIST	234.47	0.059	0.561	**111.72**	0.349	**0.558**	113.47	**0.352**	0.573
CIFAR10	259.31	0.367	0.264	176.89	0.761	0.143	**174.91**	**0.779**	**0.139**

Exploration of other deterministic degradations (10k steps)

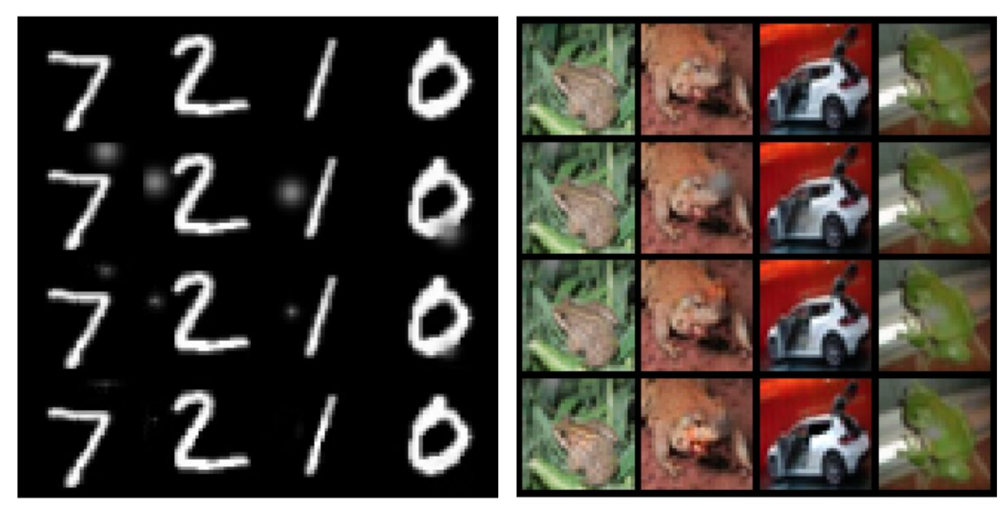


Fig 3. *Inpainting* (~10k steps)

Fig 4. *Super-Res* (~10k steps)

Cold Generation

Goal: Generate a new image from a prior over degraded images.

- Define a prior distribution over fully degraded images, then use Algorithm 2
- Symmetric priors (e.g., uniform blur) → mode collapse
- Adding **small asymmetries** (e.g., slight noise) **restores diversity**.
- Cold generation achieves reasonable FID scores compared to hot in many cases

Dataset	Hot Diffusion		Cold Diffusion	
	Fixed Noise	Estimated Noise	Perfect symmetry	Broken symmetry
CelebA	59.91	23.11	97.00	49.45
AFHQ	25.62	20.59	93.05	54.68

FID scores for CelebA and AFHQ. Breaking symmetry greatly improves results.

METHODOLOGIES

- Training on Google Colab **T4 GPUs**
- Focused on smaller **MNIST** and **CFAR10** datasets
- Metrics: **FID** (Fréchet Inception Distance), **SSIM** (Structural Similarity Index Measure), **RMSE**
- Batch sizes from 32/64 → **256** for faster training
- Decreased degradation and gradient steps until models can be trained in **~3 hours**

Reconstruction types

- **Direct** = $\hat{x}_0 = R(D(x_0, T), T)$
- **Sampled** = Iterative refinement using **Algorithm 2**

Algorithm 2 (novel)

- Necessary because no noise, so errors accumulate

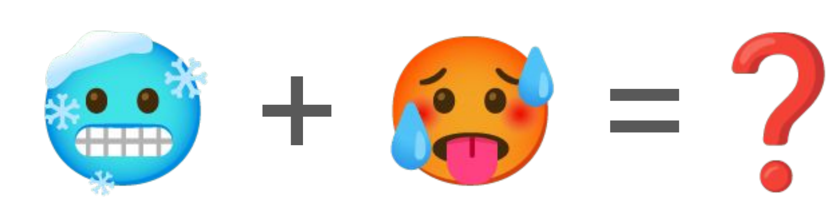
Input: A degraded sample x_t
for $s = t, t-1, \dots, 1$ **do**
  $\hat{x}_0 \leftarrow R(x_s, s)$
  $x_{s-1} = D(\hat{x}_0, s-1) \boxed{- D(\hat{x}_0, s) + x_s}$
end for
Return: x_0

CONCLUSIONS

- **Random noise can be removed entirely from the diffusion model framework**
- Cold diffusion generalizes diffusion models, beyond the Gaussian noise paradigm

Future Research Directions

- Formal framework for when cold diffusion works
- Detailed comparison of cold v.s. hot diffusion
- Cold diffusion in audio/text/video?



REFERENCES

1. Bansal, Arpit, et al. "Cold diffusion: Inverting arbitrary image transforms without noise." *Advances in Neural Information Processing Systems* 36 (2023): 41259-41282.
2. Sun, Jennifer, et al. "Diffusion Models." https://www.cs.cornell.edu/courses/cs4782/2025sp/slides/pdf/week9_1_slides.pdf