Synthetic data generation to improve object detection in x-ray radiographs

Heidi Komkov, Matthew Marshall, Erik Brubaker

Sandia National Laboratories

Sandia ML/DL Workshop

July 2023

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2023-02778C

Overview

- We always need more data in deep learning.
  - Some domains have data sharing problems.
- Diverse training data helps model generalizability.

- How do the latest techniques to create synthetic data perform to create synthetic radiographs?

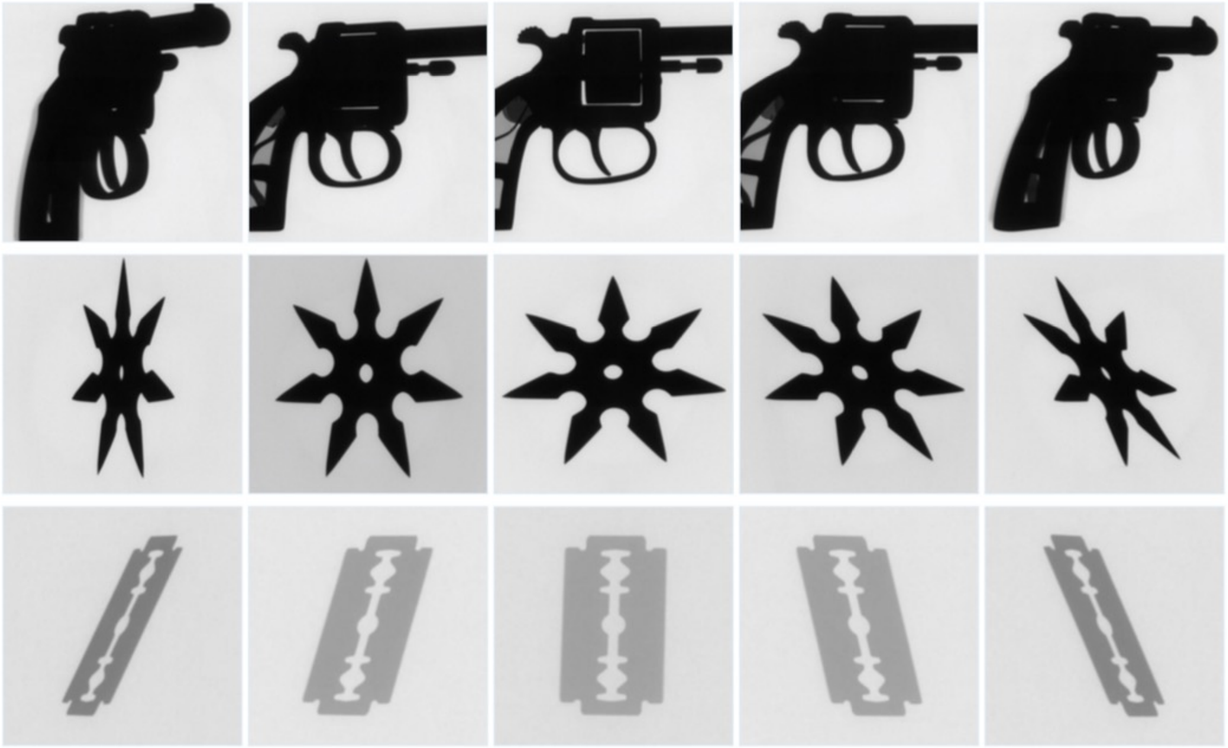
A public dataset: GDXray

https://domingomery.ing.puc.cl/material/gdxray/

8150 images in baggage category

3 main categories of threat objects

Fig. 6 Some X-ray images of a bag containing handguns, *shuriken* and razor blades (group Baggage series B0048).

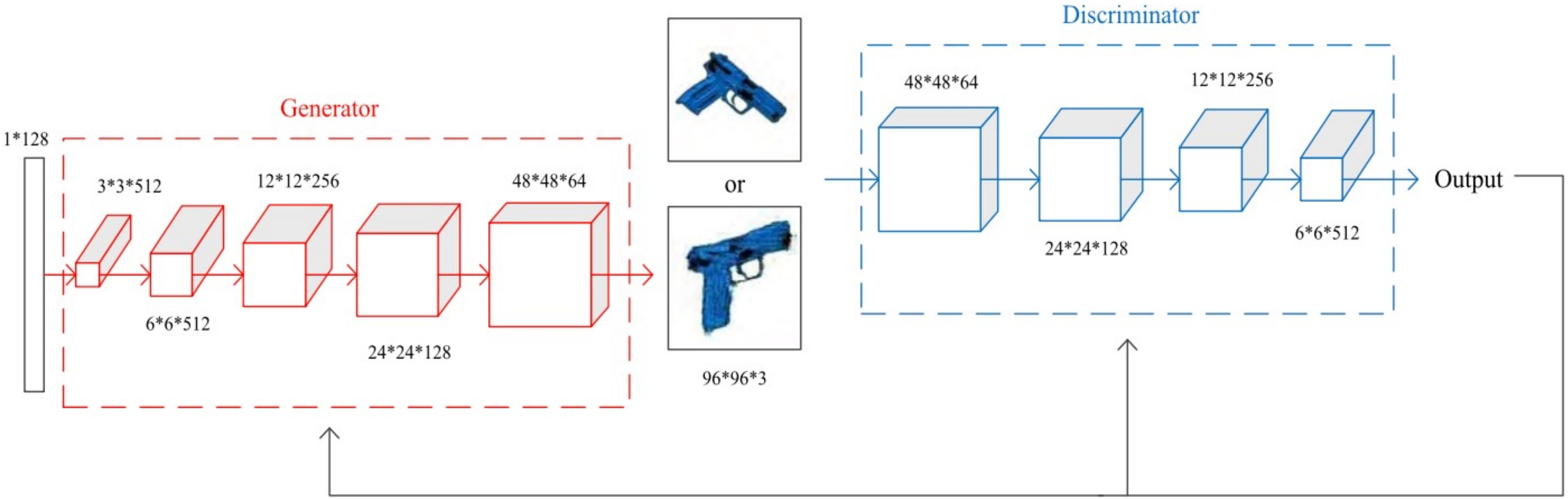


Generative Models in General

Generative models learn the **underlying distribution** of the data, $q(x)$, from a set of examples $\boldsymbol{x}$, and generate an approximate distribution $p(x)$ which is sampled to create new data points $\widehat{\boldsymbol{x}} \sim p(x)$.

Deep generative models use deep neural networks to approximate the distribution of data.

Background: Generative Adversarial Networks

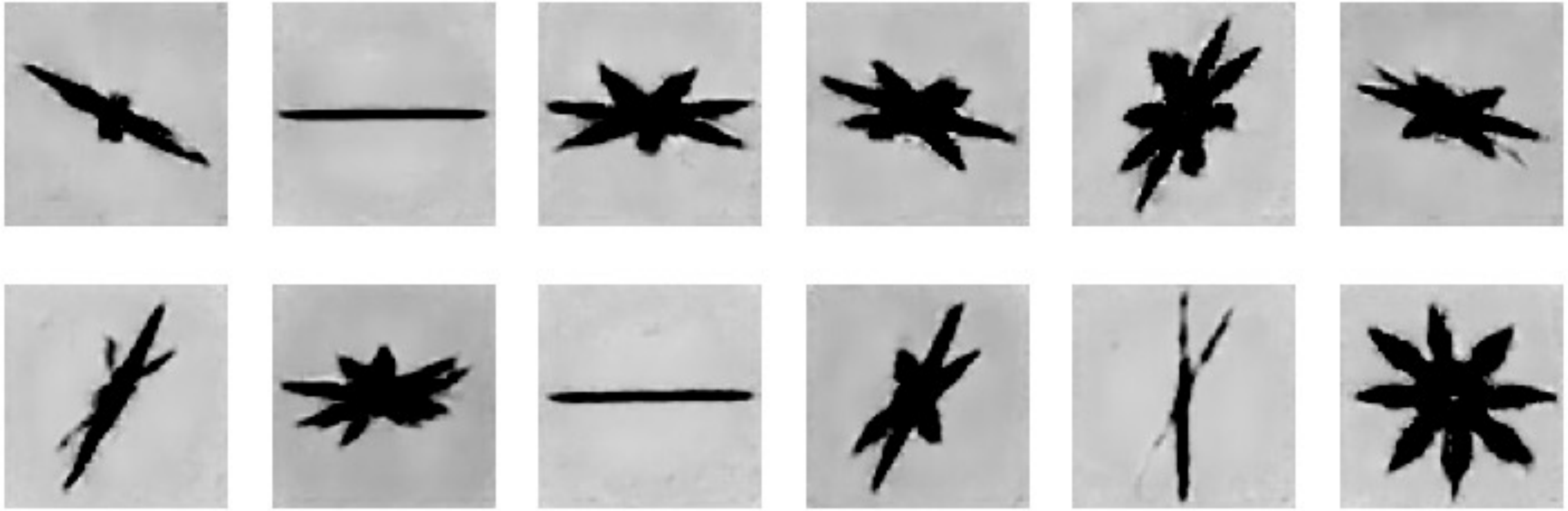


Yang, J., et al. (2019). "Data Augmentation for X-Ray Prohibited Item Images Using Generative Adversarial Networks." IEEE Access **7**: 28894-28902.

Results: Deep Convolutional GAN (DCGAN)

Trained on:

Results:



Inspiration for further investigation

Diffusion Models Beat GANs on Image Synthesis

Prafulla Dhariwal* OpenAI prafulla@openai.com

Alex Nichol* OpenAI alex@openai.com

Abstract

We show that diffusion models can achieve image sample quality superior to the current state-of-the-art generative models. We achieve this on unconditional image synthesis by finding a better architecture through a series of ablations. For conditional image synthesis, we further improve sample quality with classifier guidance: a simple, compute-efficient method for trading off diversity for fidelity using gradients from a classifier. We achieve an FID of 2.97 on ImageNet 128×128, 4.59 on ImageNet 256×256, and 7.72 on ImageNet 512×512, and we match BigGAN-deep even with as few as 25 forward passes per sample, all while maintaining better coverage of the distribution. Finally, we find that classifier guidance combines well with upsampling diffusion models, further improving FID to 3.94 on ImageNet 256×256 and 3.85 on ImageNet 512×512. We release our code at https://github.com/openai/guided-diffusion.

1 Introduction

Figure 1: Selected samples from our best ImageNet 512×512 model (FID 3.85)

Background: Diffusion models

DALL-E (OpenAI)

Text prompt: "Teddy bears working on new AI research underwater with 1990s technology"

https://en.wikipedia.org/wiki/DALL-E

Stable Diffusion (Stability AI)

Text prompt: ”A photograph of an astronaut riding a horse"

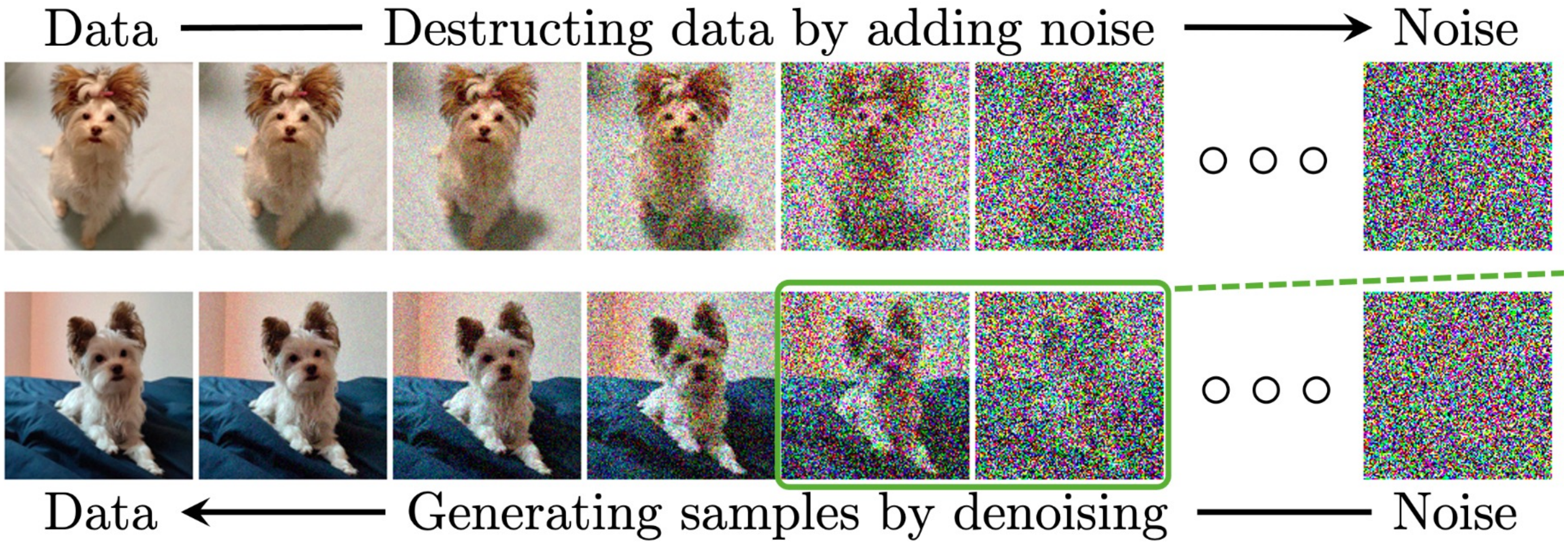
https://en.wikipedia.org/wiki/Stable_Diffusion

Also:

Imagen
https://imagen.research.google/

Midjourney
https://www.midjourney.com/

Background: Diffusion model theory



Yang, L., et al. (2022). "Diffusion models: A comprehensive survey of methods and applications." arXiv preprint arXiv:2209.00796.

Diffusion model theory: noising

Given a data distribution $q(x_0)$, a Markov process generates a sequence of random variables $x_1, x_2, \ldots, x_T$ with a transition kernel $q(x_t|x_{t-1})$. A common choice of transition kernel is a Gaussian:

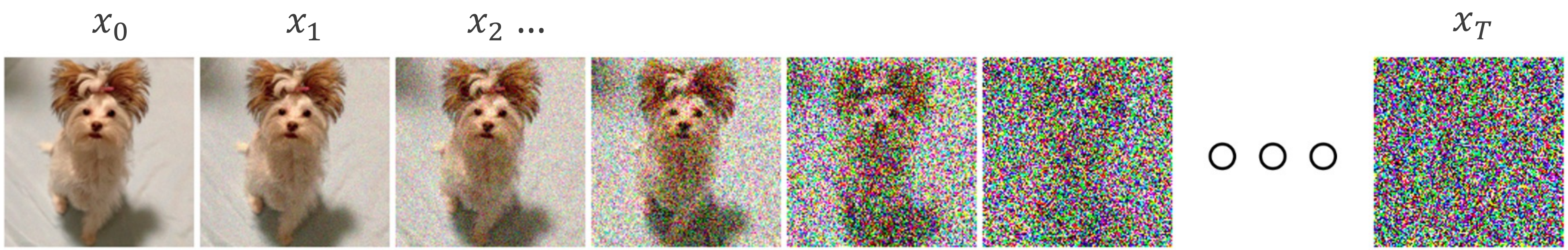
$$q(x_t|x_{t-1}) = N(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I) \ ,$$

where $\beta_t \in (0,1)$ and is varied according to a schedule. We can do a reparametrization trick and pre-compute the noise at time step t.

$$\text{Define: } \alpha_t = 1 - \beta_t \text{ and } \bar{\alpha}_t = \prod_{s=0}^{t} \alpha_s$$

$$\text{Reparametrized: } q_t(x_t|x_0) = N\left(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I\right)$$

Given x_0, we can easily obtain a sample of x_t by sampling a Gaussian vector $\epsilon \sim N(0, I)$ and applying the transformation $x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$.



Diffusion model theory: denoising

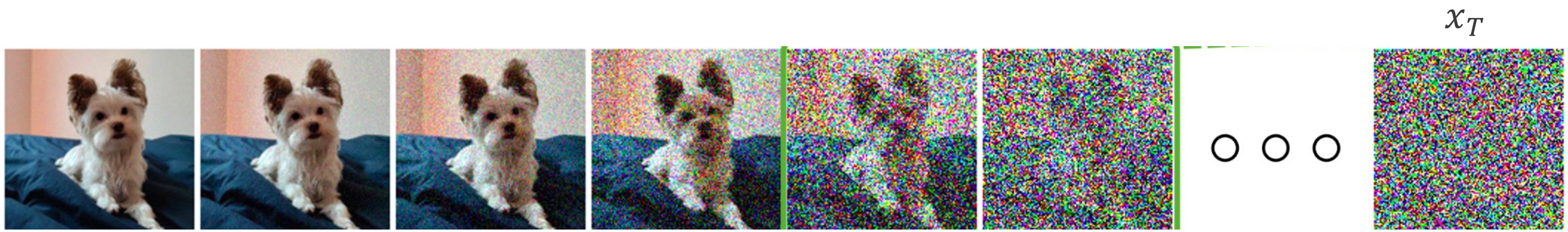
The denoising process uses a Markov chain in the opposite direction, with learnable transition kernels $p_\theta(x_{t-1}|x_t)$. Here, θ denotes model parameters, and the mean $\mu_\theta(x_t, t)$ and variance $\Sigma_\theta(x_t, t)$ are parametrized by deep neural networks:

$$p_\theta(x_{t-1}|x_t) = N\big(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)\big)$$

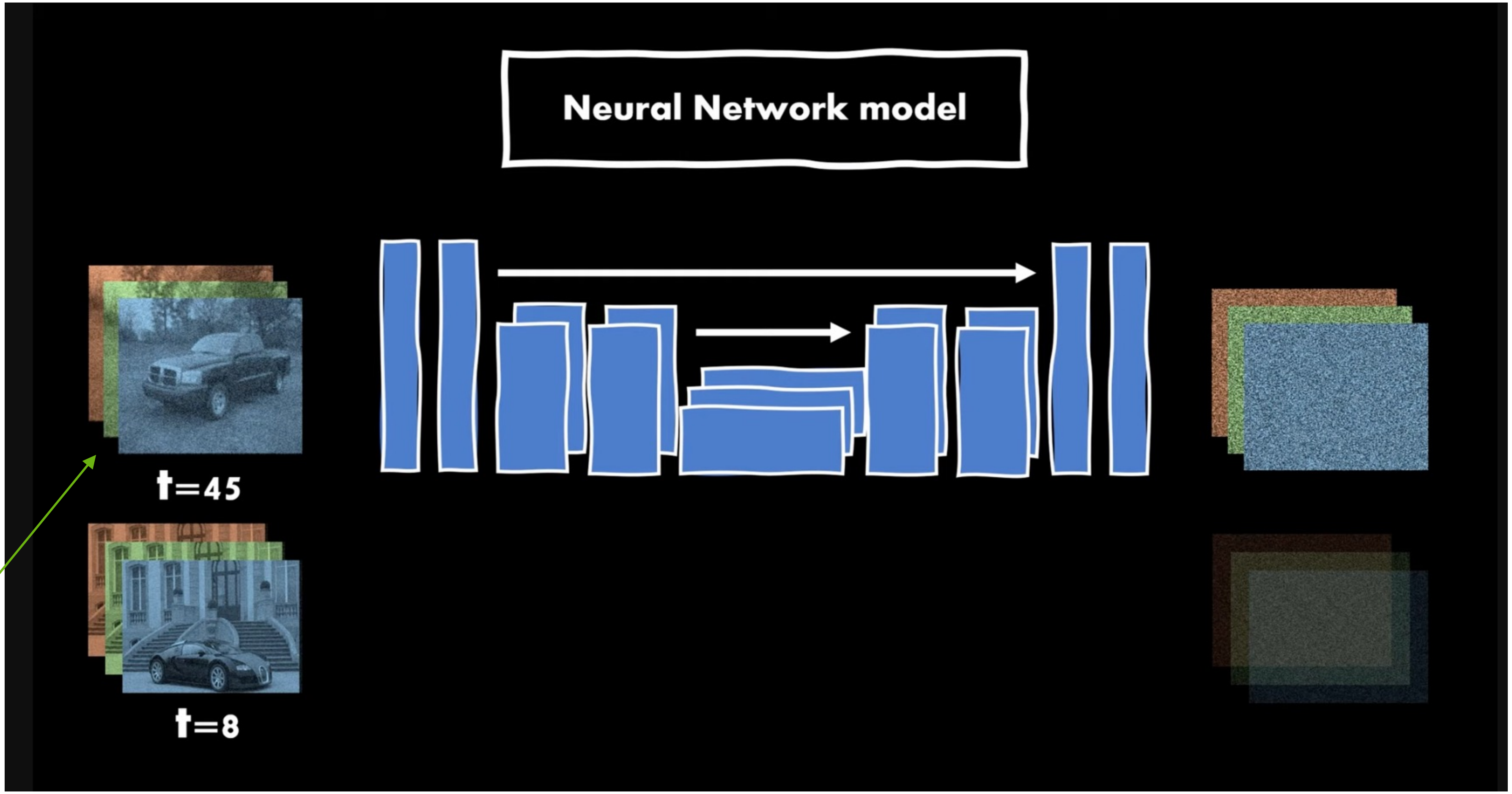
The reverse diffusion process repeatedly applies the kernels until t=0:

$$p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t).$$

Generating a new sample starts with sampling from a noise vector x_T and iteratively sampling from the transition kernel.



U-Net and Training



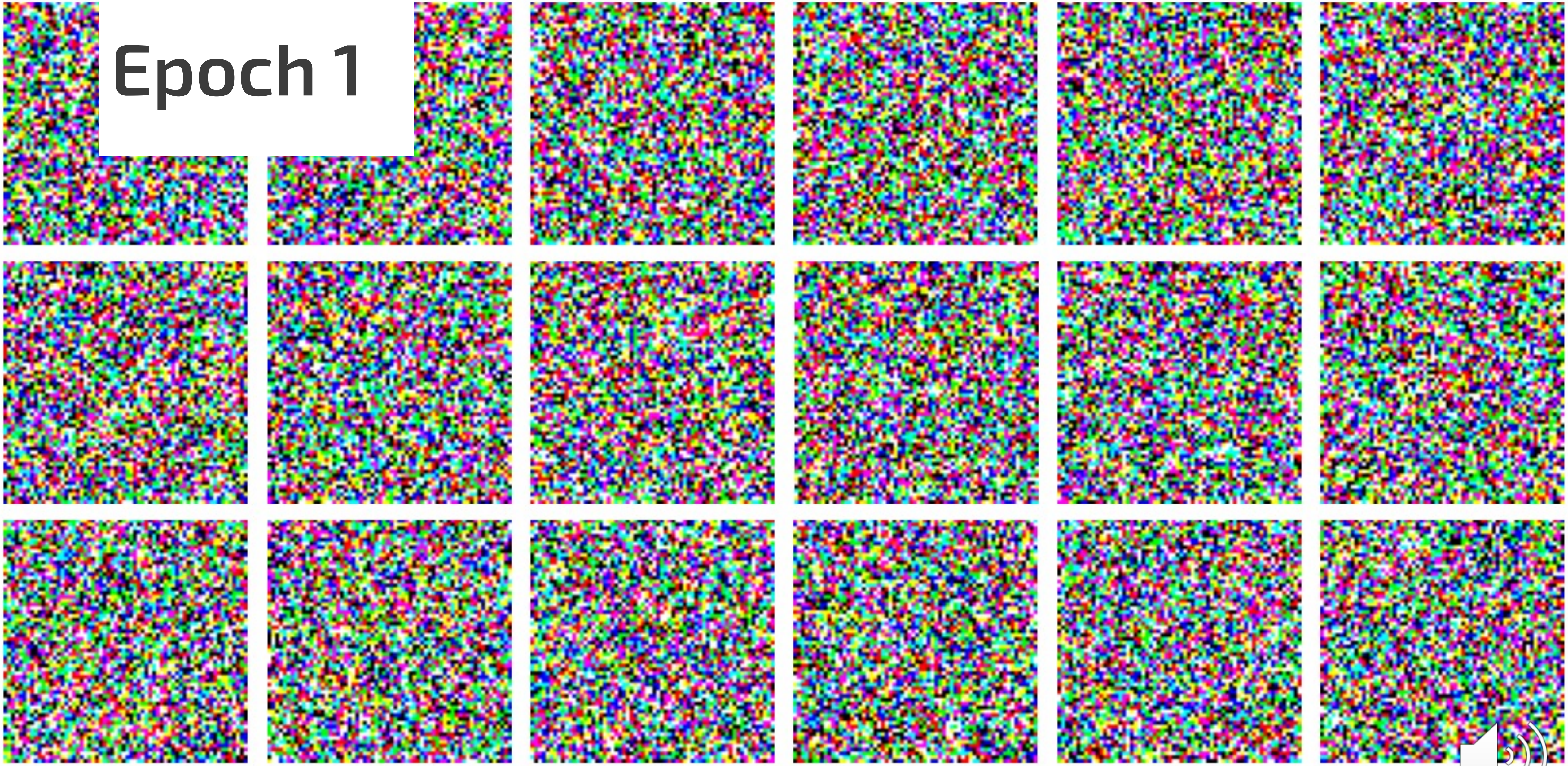
Time embedding added as a channel

https://www.youtube.com/watch?v=a4Yfz2FxXiY

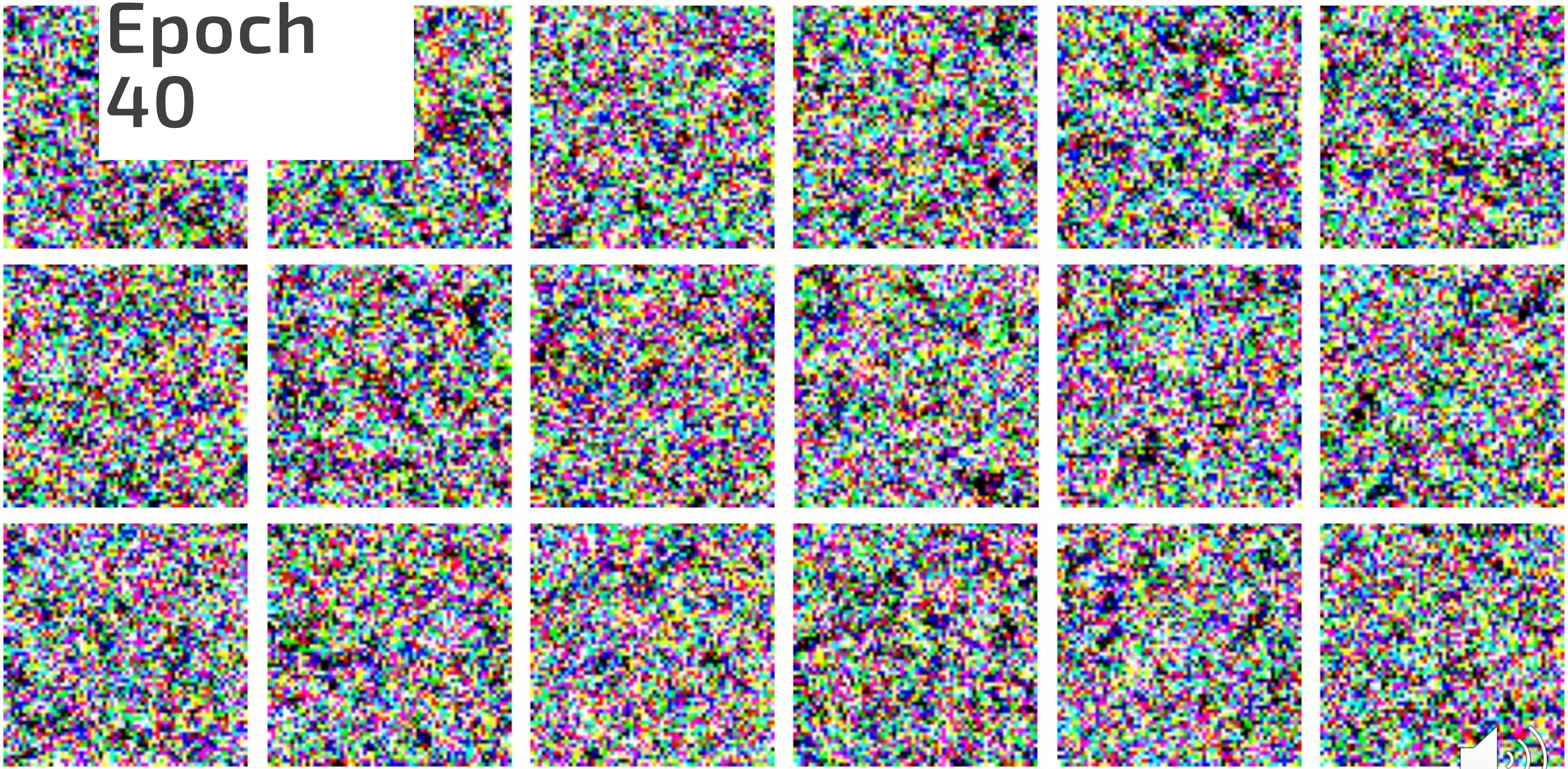
Now we train a diffusion model called Denoising Diffusion Implicit Model (DDIM)…

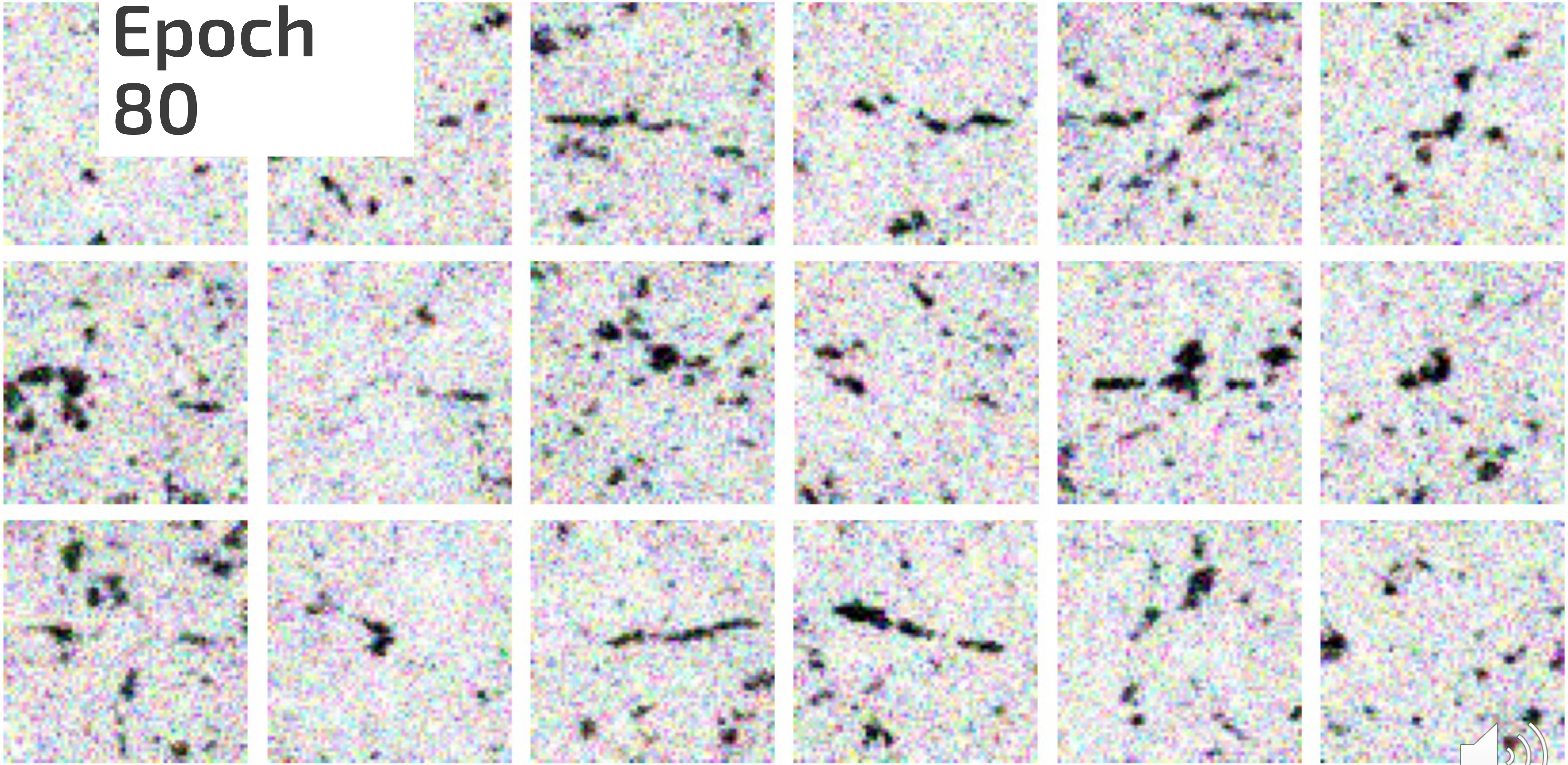
Trained on:

Epoch 1

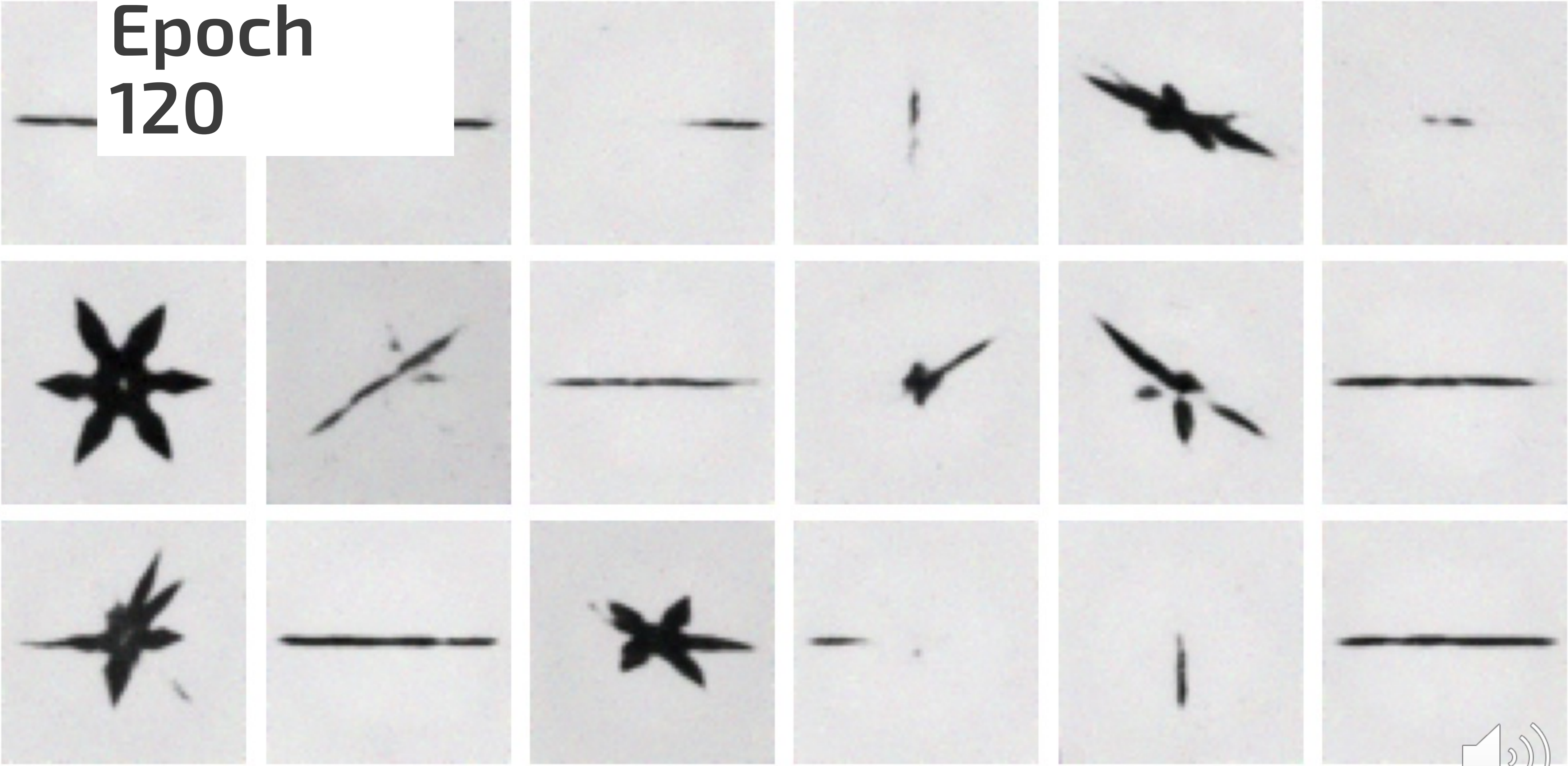


Epoch 40

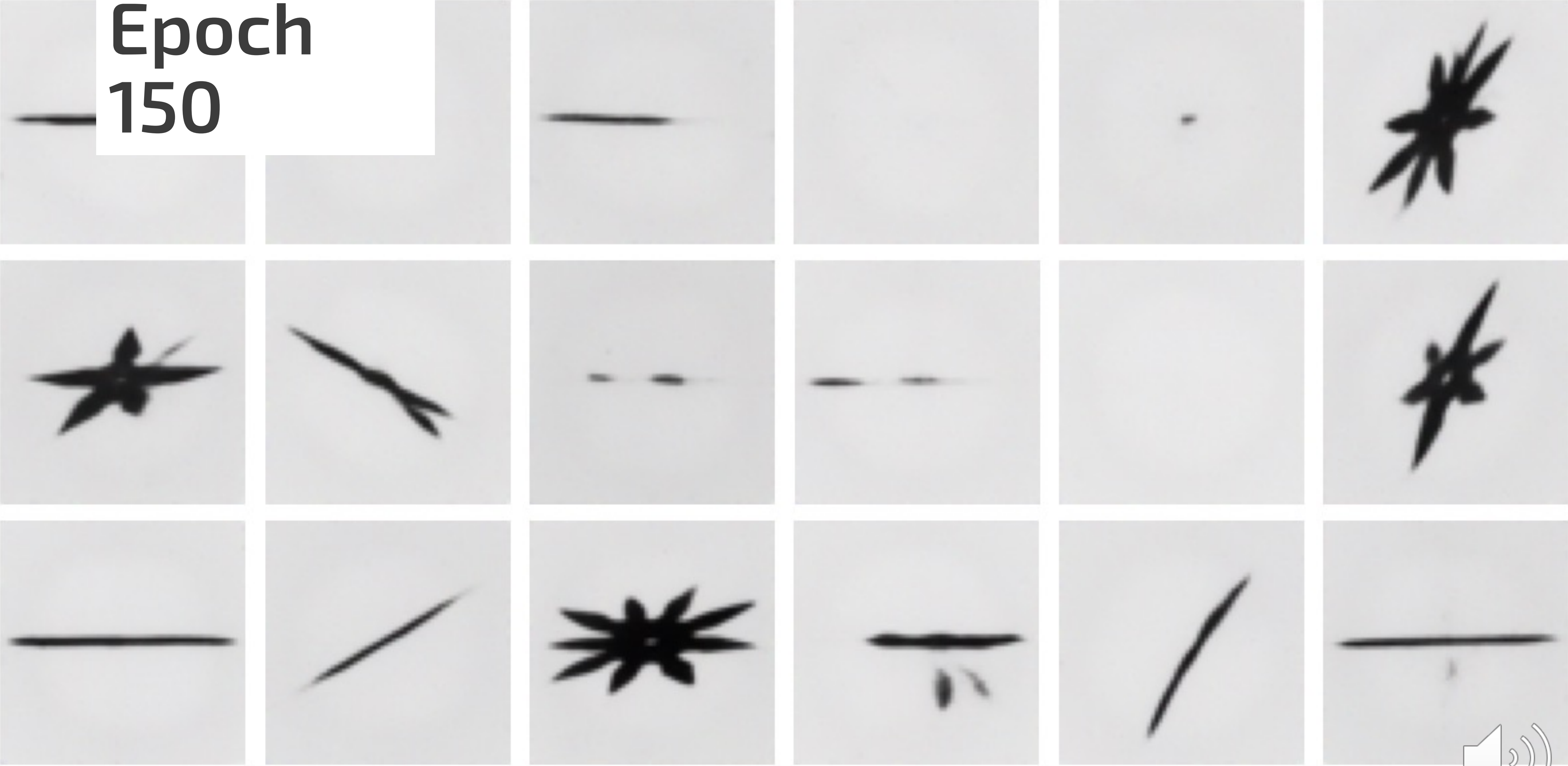



Epoch
80

Epoch 120



Epoch 150



DDIM Generalization to new examples

Remember, this model was trained on 6, 7, and 8-blade shurikens!

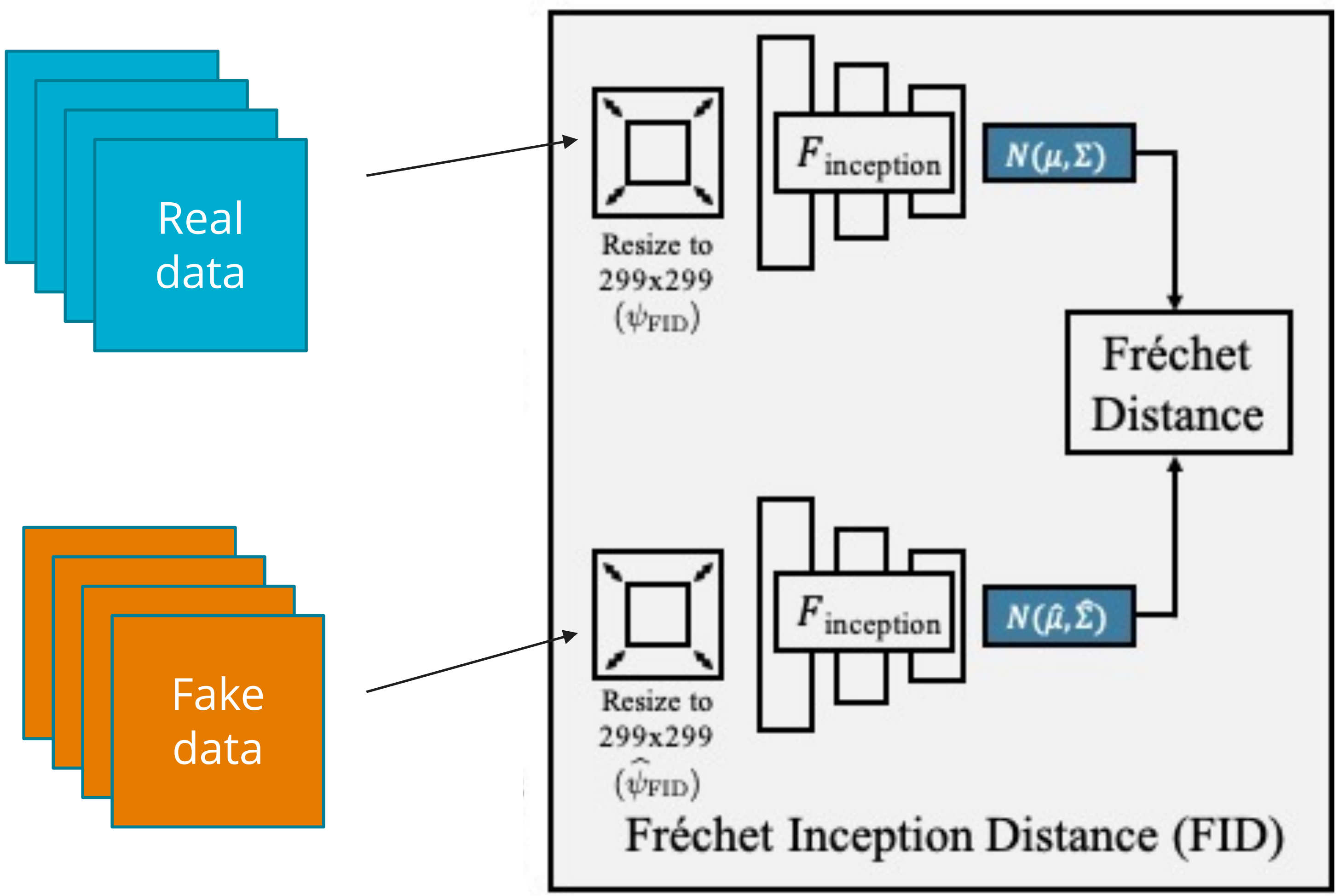
4 blades

5 blades

9 blades

* Only about 30% of data was accepted – the rest was empty or nearly empty.

Background: Quantifying synthetic data quality



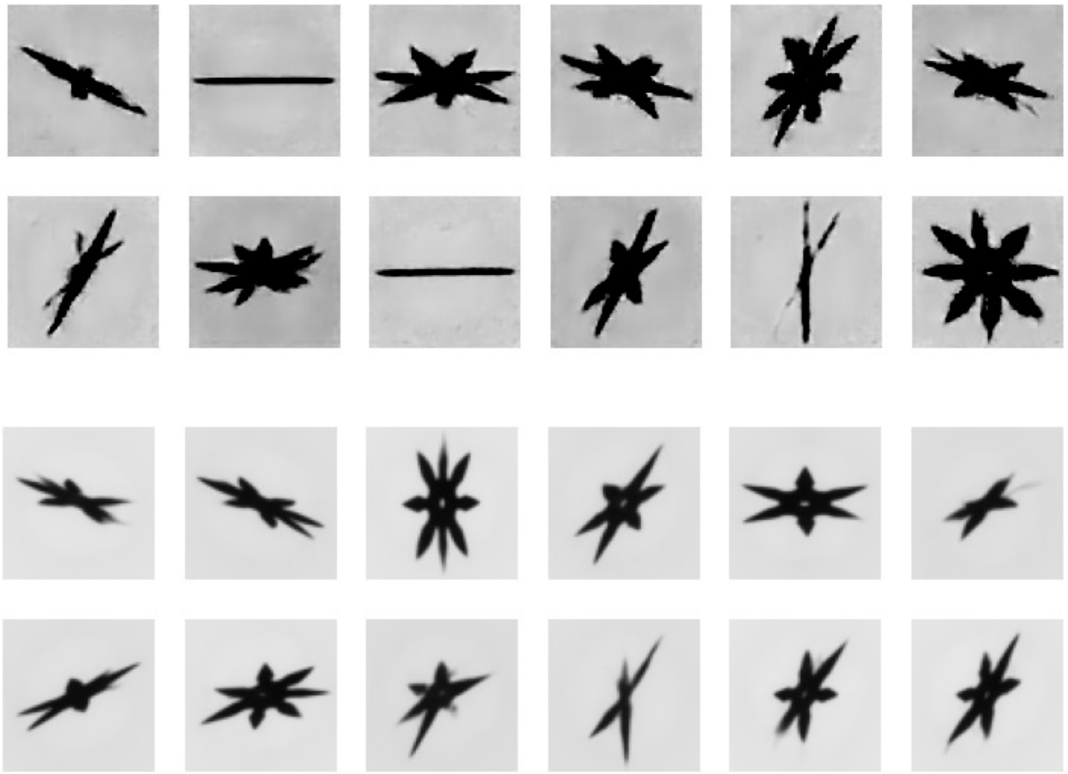
https://pypi.org/project/clean-fid

$$FID = \|\mu_X - \mu_Y\|^2 - Tr\left(\Sigma_X + \Sigma_Y - 2\sqrt{\Sigma_X \Sigma_Y}\right)$$

DCGAN/DDIM comparison

DCGAN

DDIM



DDIM	GAN
• 1,950,627 trainable parameters	• Generator trainable params: 2,039,169 • Discriminator trainable params: 1,717,889.
• Training fast(ish) and stable. • Inference slow, many clearly bad examples.	• Training slow and unstable. • Inference fast.

	DDIM images	DCGAN images
FID score	304.8	1117.4

How much does synthetic data help an object detection model to generalize?

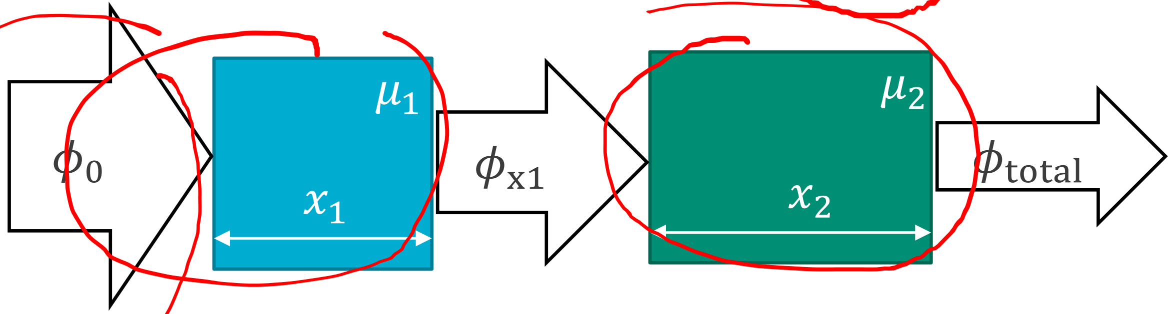
Multiplicative insertion of threats

X-ray intensity after passing through object 1: $\phi_{x_1} = \phi_0 e^{-\mu_1 x_1}$

X-ray intensity after passing through object 2: $\phi_{x_2} = \phi_0 e^{-\mu_2 x_2}$

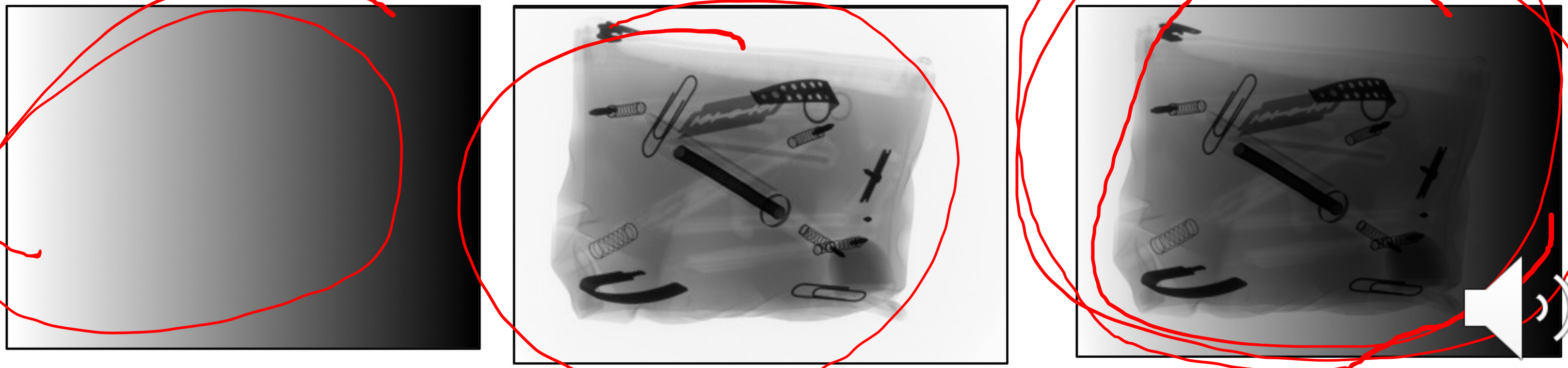
X-ray intensity after passing through object 1 on top of object 2:

$$\phi_{total} = \phi_0 e^{-\mu_1 x_1} e^{-\mu_2 x_2} = \frac{\phi_{x_1}\phi_{x_2}}{\phi_0}$$

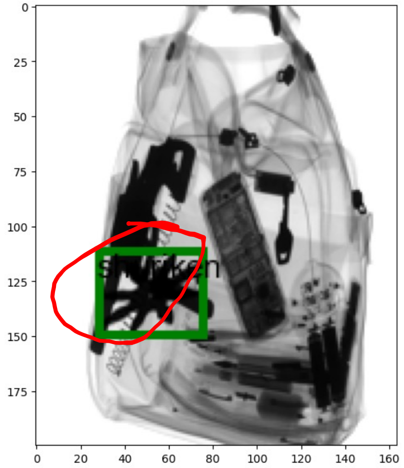


Magnitude of pixel value on detector: $I = A\phi + B$

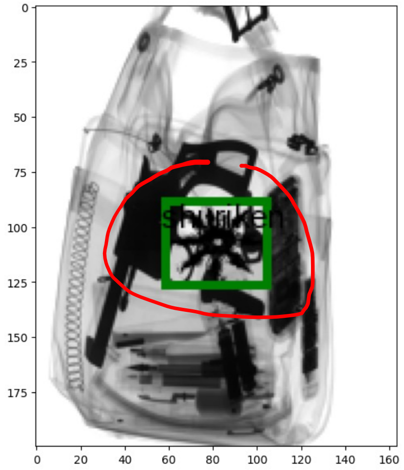
$$I_{total} = \frac{I_{background} \times I_{object}}{255}$$



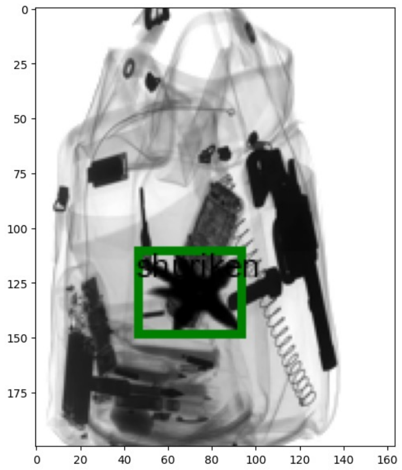
Creating a training set



Real



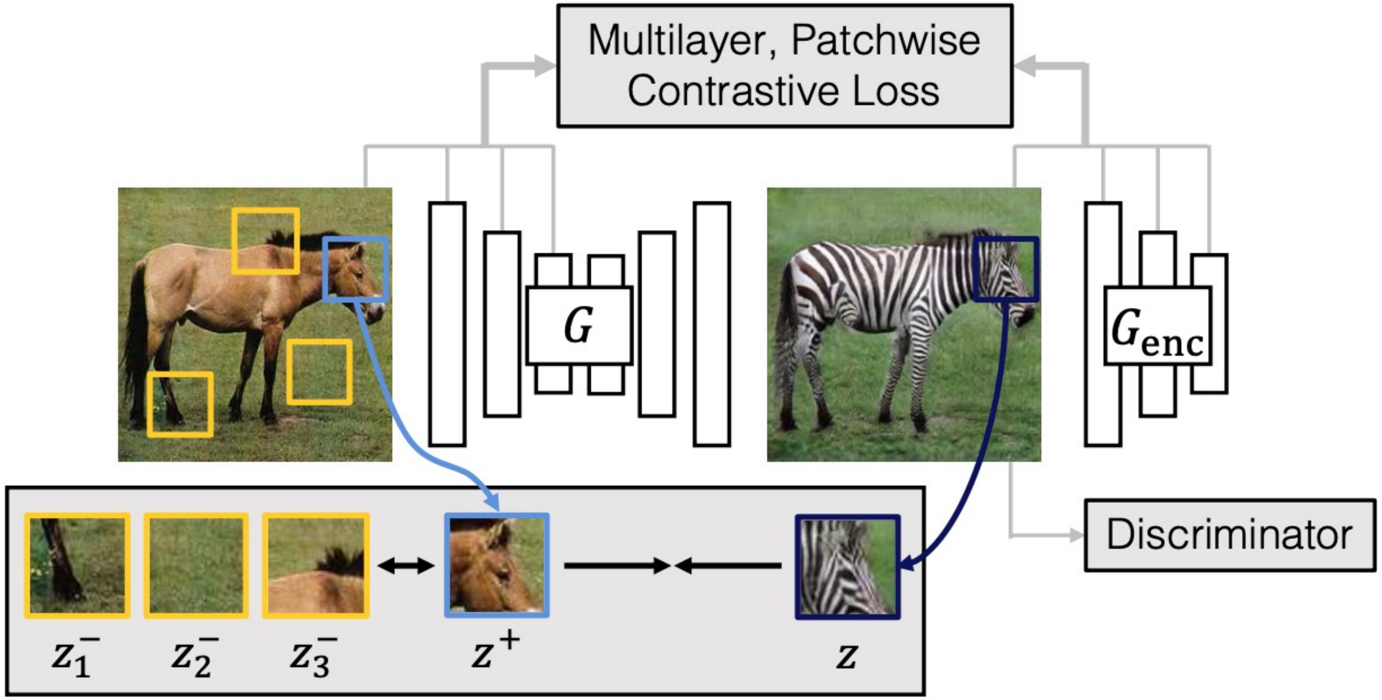
DCGAN



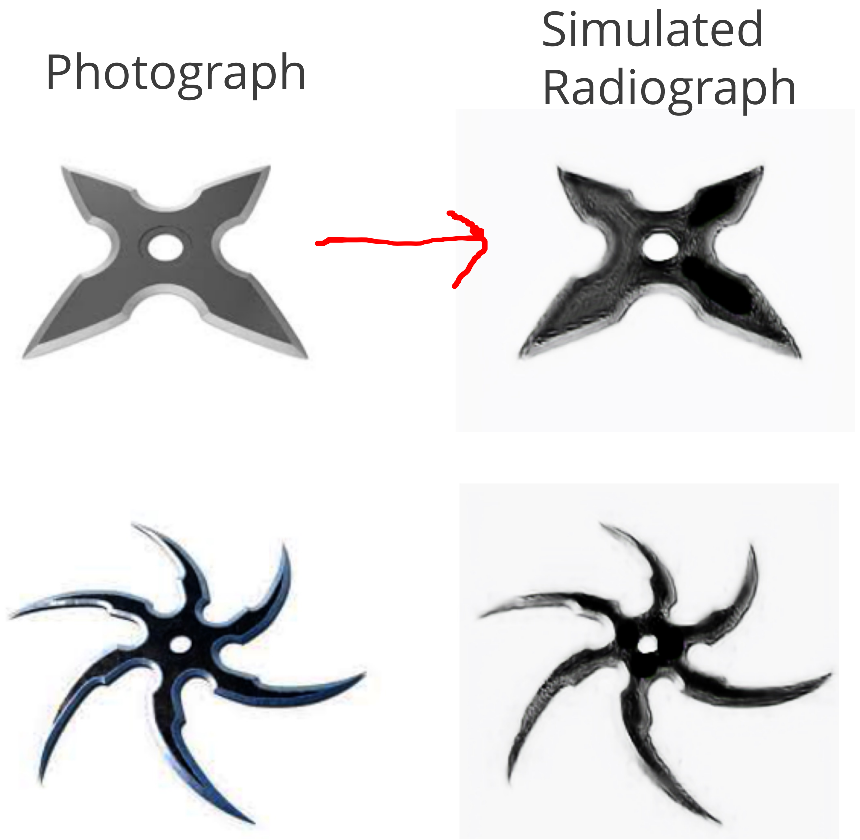
DDIM

- 432 images in each category.
- Inserted into 178 backgrounds to create 10,000 images in each category.
- 10—75% occlusion enforced (occluded pixel < 70).
- Augmented datasets sample real and synthetic data 50/50.

Creating a challenge dataset: Neural style transfer using Contrastive Unpaired Learning



Patchwise Contrastive Learning



Park, T., et al. (2020). Contrastive learning for unpaired image-to-image translation. European conference on computer vision, Springer.

Challenge Dataset and YOLOv5 results

15 challenge items in 1,000 images

	GDXray data	Augmented with DCGAN	Augmented with DDIM
Average precision	0.736	0.843	0.836

Conclusion

- Diffusion models can outperform GANs in realism.
- While DDIM is faster to train, a DCGAN is faster to sample from.
- Synthetic data can improve object detection model generalization.

This work was supported by the Office of Defense Nuclear Nonproliferation Research and Development within the U.S. Department of Energy's National Nuclear Security Administration.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.