

This Dinosaur Does Not Exist

Year Project

Matej Kukurugya & Václav Krňák

January 15, 2026

What We Wanted to Achieve

Inspiration

<https://thispersondoesnotexist.com/>



What Was Needed

- Create a model that would generate images of dinosaurs

What Was Needed

- Create a model that would generate images of dinosaurs
 - “State-of-the-art”

What Was Needed

- Create a model that would generate images of dinosaurs
 - “State-of-the-art”
 - Explore and compare different models

What Was Needed

- Create a model that would generate images of dinosaurs
 - “State-of-the-art”
 - Explore and compare different models
- Use of the school’s HPC cluster for training on graphics cards

What Was Needed

- Create a model that would generate images of dinosaurs
 - “State-of-the-art”
 - Explore and compare different models
- Use of the school’s HPC cluster for training on graphics cards
- Creation of a web application

Data Collection

To create training data, we used <https://www.kaggle.com/>. Specifically:

- <https://www.kaggle.com/datasets/larserikrishholm/dinosaur-image-dataset-15-species> – 2448 images - 510MB
- <https://www.kaggle.com/datasets/cmglonly/simple-dinosurus-dataset> – 200 images - 58MB
- <https://www.kaggle.com/datasets/antaresl/jurassic-park-dinosaurs-dataset> - 4364 images - 4GB
- <https://www.kaggle.com/datasets/caokhoihuynh/jurassic-world-dinosaur> - 5835 images - 1.6GB

Pre-processing

- Images are rescaled to 256x256 px
- Given the size of the dataset (12848 images total), we decided not to include any augmentation techniques

Model Selection

Stable Diffusion

- Iterative noise removal
- Too computationally demanding for our use

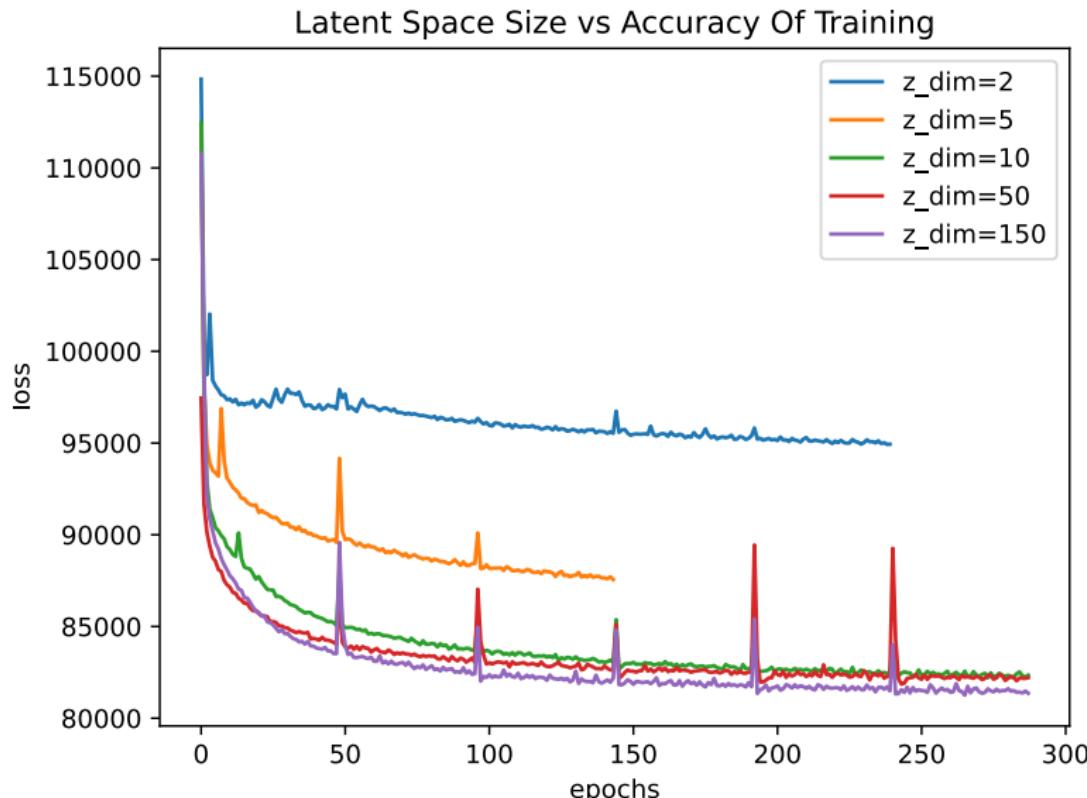
DC-GAN

- The model has two parts: Generator and discriminator
- Too difficult a task, the discriminator always achieves 100% accuracy

VAE

- Encoder -> Latent space -> Decoder
- Computationally undemanding

Training



Results

- 48 epochs

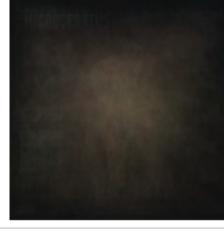
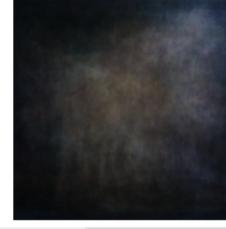
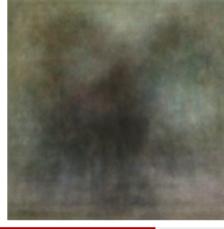
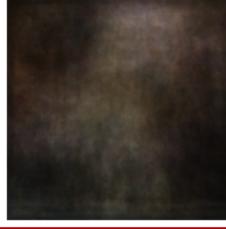
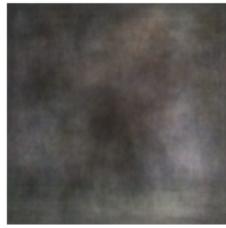
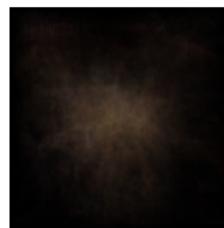
150

50

10

5

2



Results

- 96 epochs

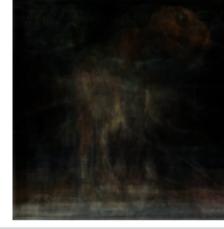
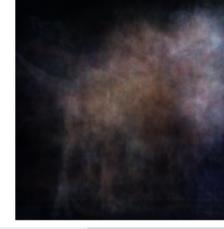
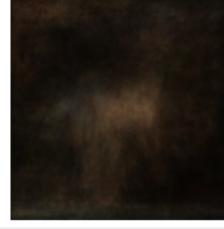
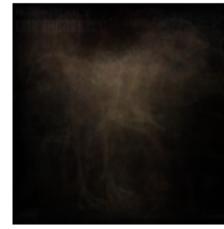
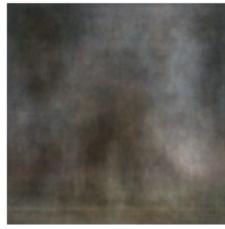
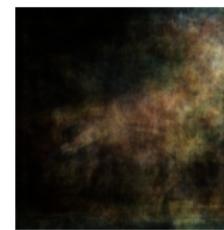
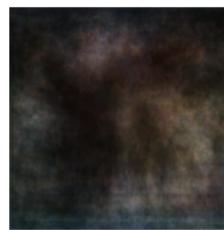
150

50

10

5

2



Results

- 144 epochs

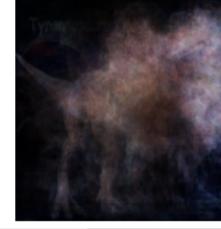
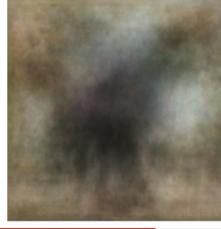
150

50

10

5

2



Results

- 192 epochs

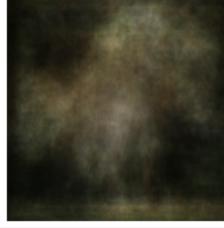
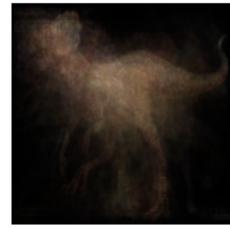
150

50

10

5 (144 ep.)

2



Results

- 240 epochs

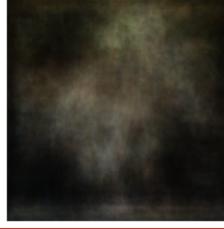
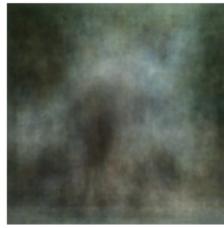
150

50

10

5 (144 ep.)

2



Results

- 288 epochs

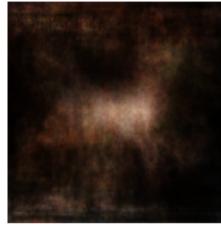
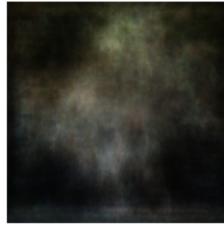
150

50

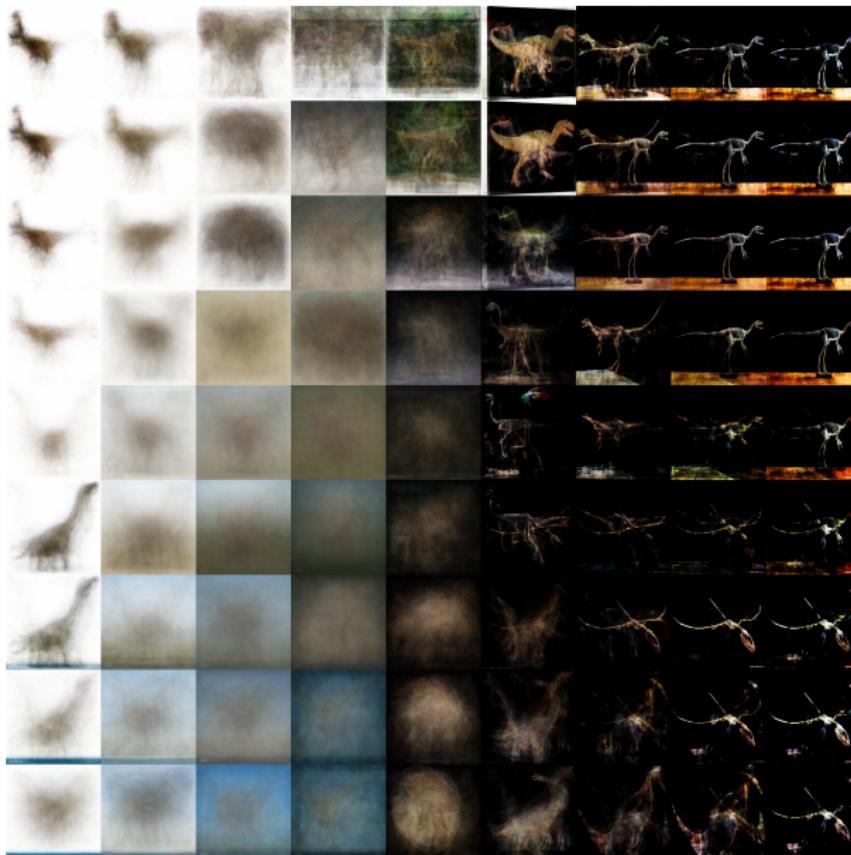
10

5 (144 ep.)

2 (240 ep.)



Results



Discussion

- The models trained properly, but did not have sufficient capacity

Discussion

- The models trained properly, but did not have sufficient capacity
- Models with large latent space were a dead end