Terrain Design Prototyping With Diffusion Models

Progress of My Graduation Research (2024/12/11)

s1290102 Hajime Fukai

Background

- Modern 3D games require a vast expanse of world map
  - Designing such a world map by hand is a demanding task
 - It is difficult to design natural terrain for nonexperts

Objective

- Our goal is to facilitate the prototyping of a world map
  - By creating a terrain generation model
  - Can be guided by hand-drawn sketch
- Use diffusion models primarily; instead of traditional GAN
- Previous study
  - T. Wang and S. Kurabayashi, "Sketch2Map: A Game Map Design Support System Allowing Quick Hand Sketch Prototyping"
    - 2020 IEEE Conference on Games (CoG)

Models to Use

- Diffusion Model

  - "Palette: Image-to-Image Diffusion Models" introduced by Chitwan Saharia et al.

- GAN (Generative Adversarial Network)

  - pix2pix: "Image-to-Image Translation with Conditional Adversarial Networks" introduced by Phillip Isola et al.

Dataset Creation

- We created dataset from NASADEM_HGT v001

- NASADEM_HGT consists of elevation data and water body data
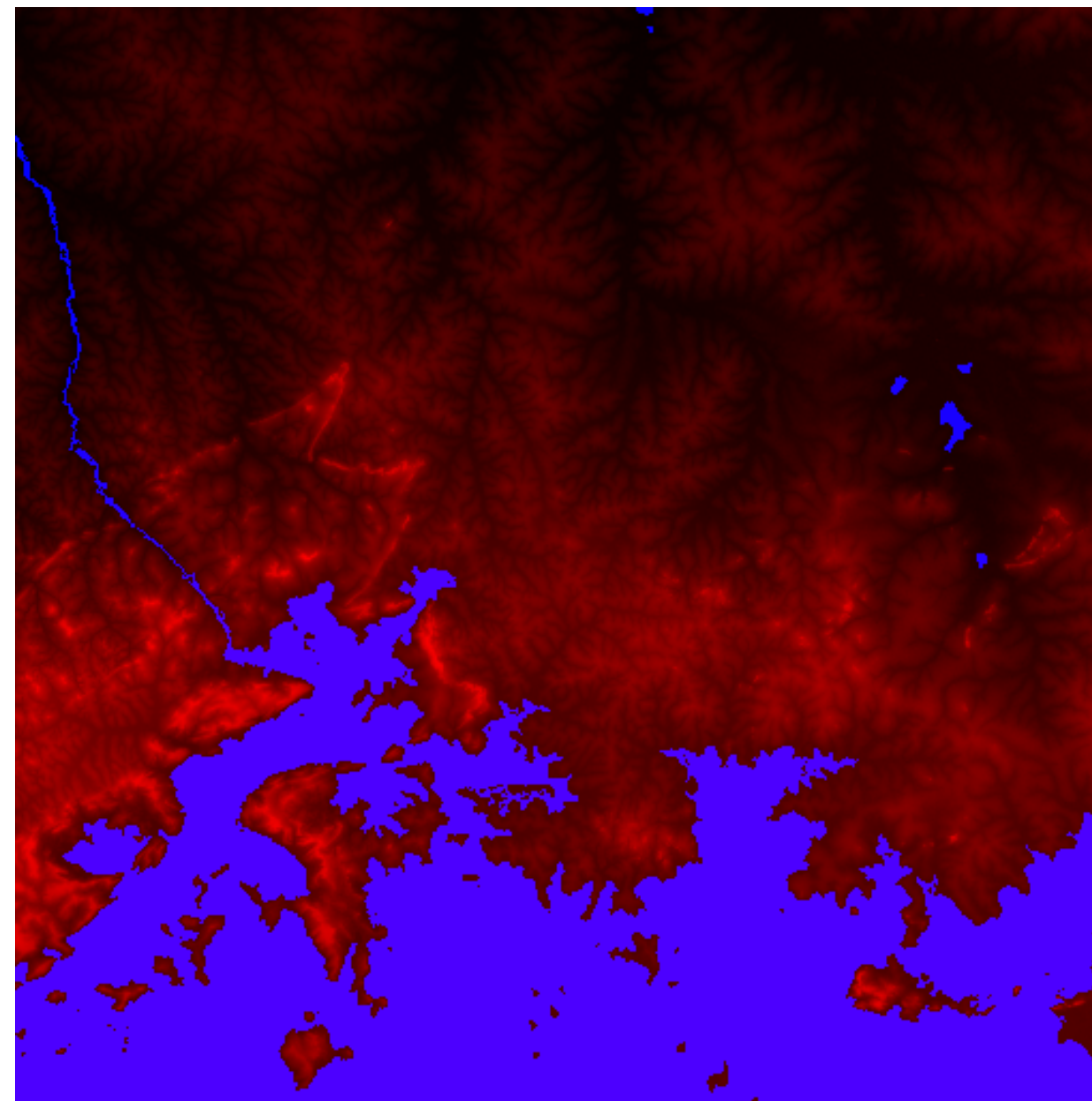
Height map from
NASADEM_HGT

Updated Dataset Creation

- Previously the height-maps represented its height by a single 8-bit channel (red)

  - The height information was compressed from 16 bits to 8 bits, resulting in lower precision

- The height is now represented by two channels (red and green) and each channel is 16-bit

  - The green area is higher than the base-height

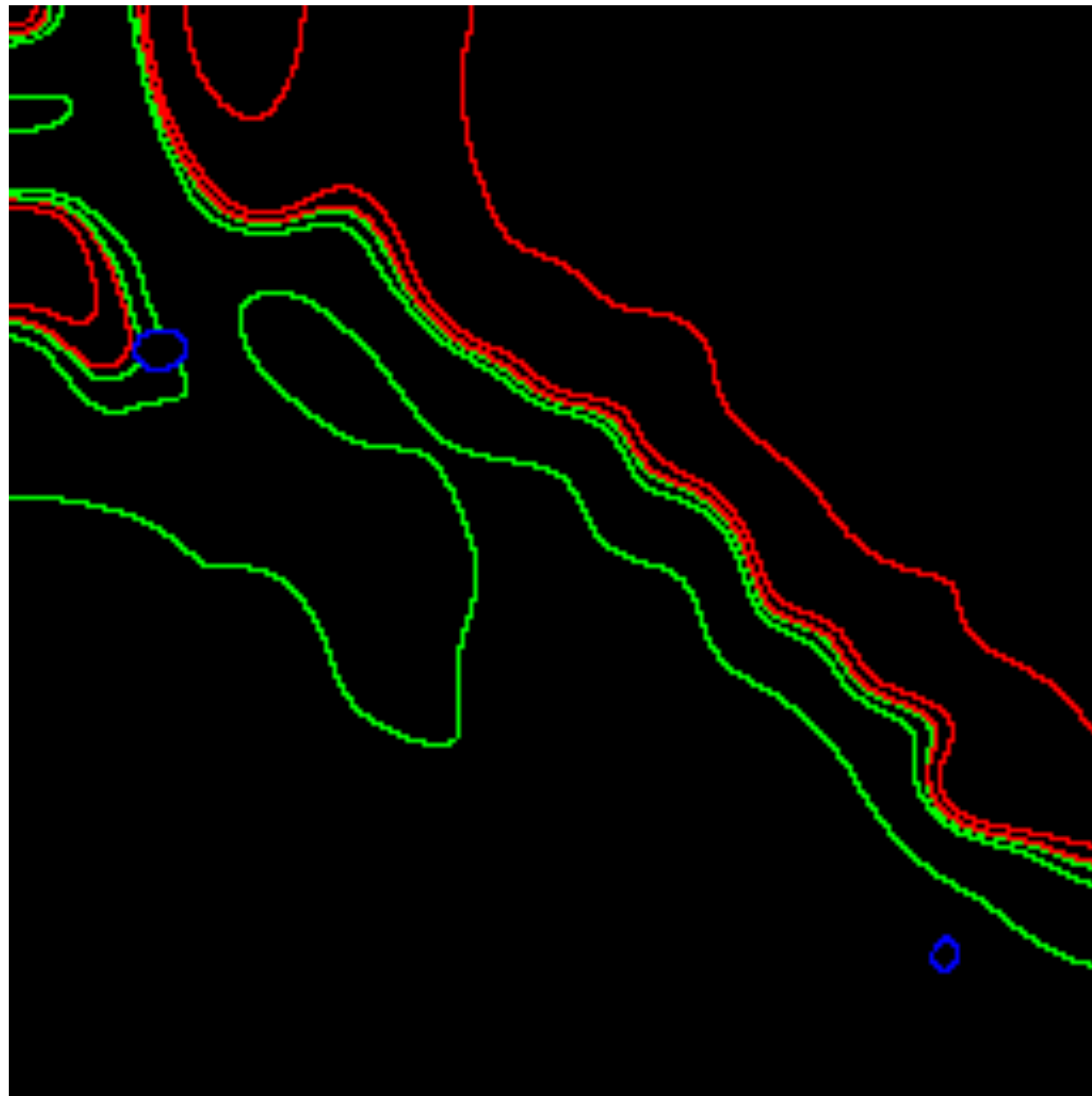 - The red area is lower than the base-height
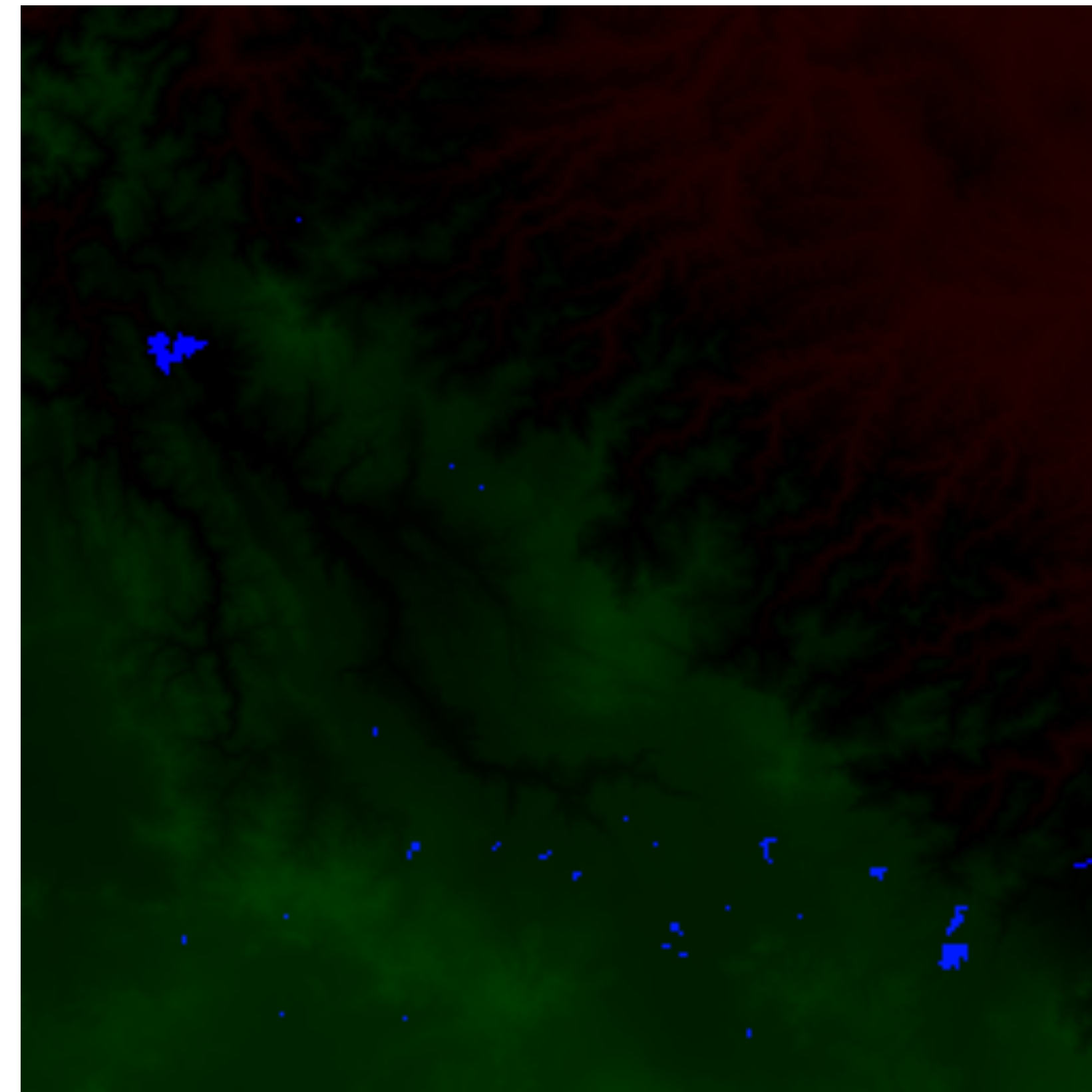
Updated Dataset Creation
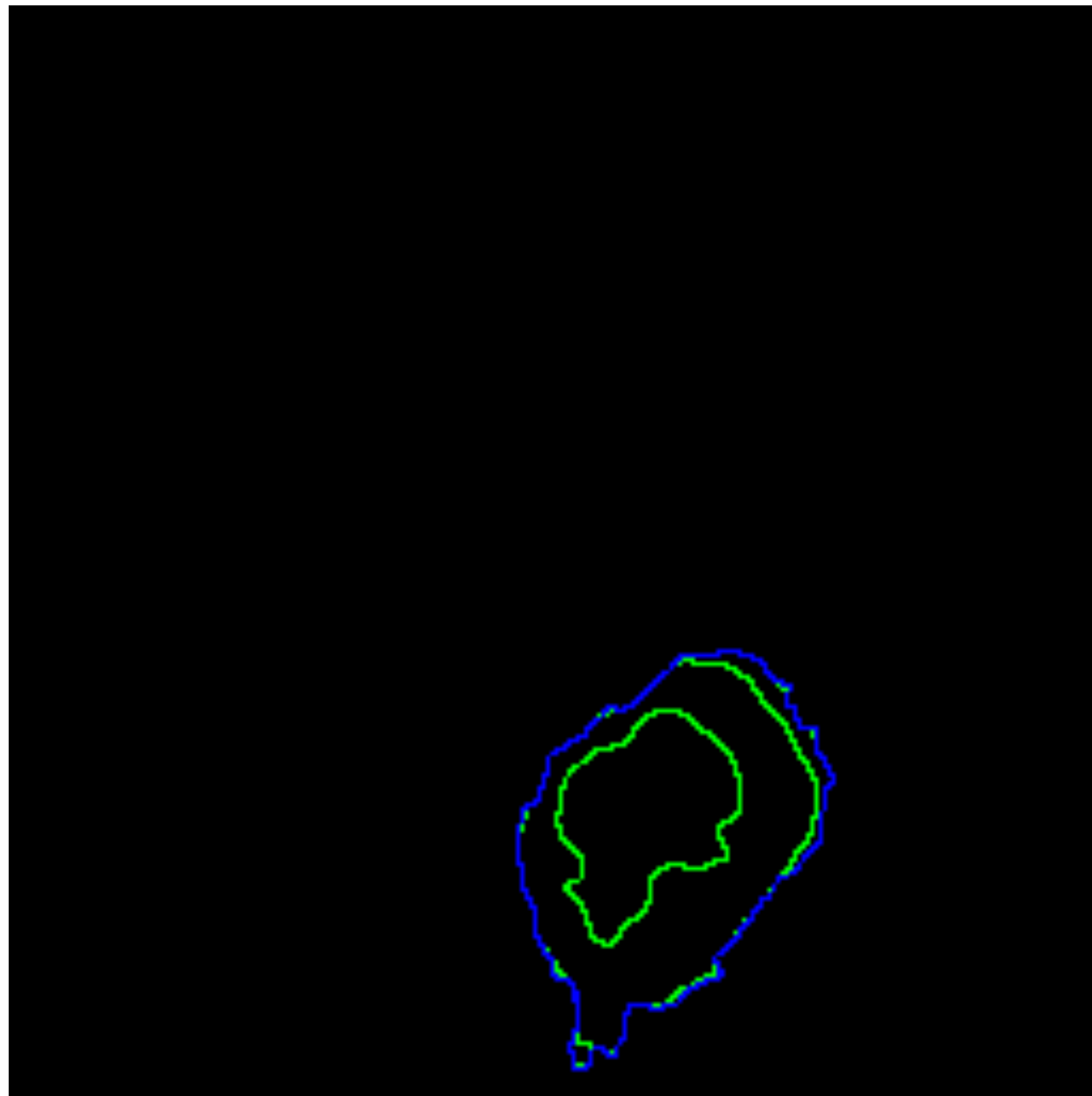
Contour-Map

Height-Map

Updated Dataset Creation

Contour-Map

Height-Map

Training Environment

- Intel Core i7 6700K

- 32GB DDR4

- NVIDIA GeForce RTX 4060 Ti 16GB

Model Implementations

- pix2pix
  - "Image-to-Image Translation in PyTorch"
  - https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix
- Palette
  - "Unofficial implementation of Palette: Image-to-Image Diffusion Models by Pytorch"
  - https://github.com/Janspiry/Palette-Image-to-Image-Diffusion-Models

Training of pix2pix Model

- Completed training the pix2pix model

  - Trained for 1200 epochs
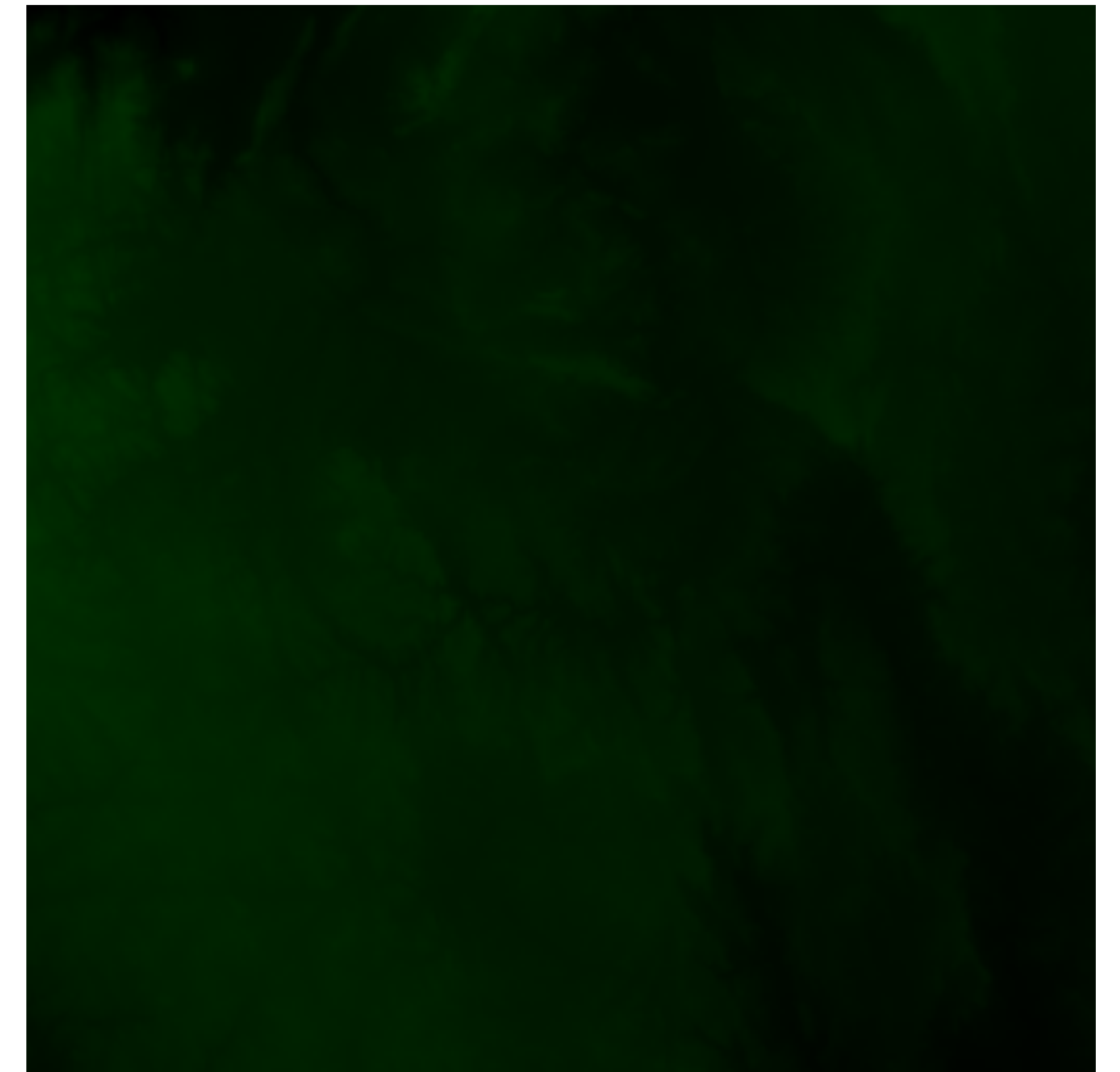
  - ≈280 seconds per epoch

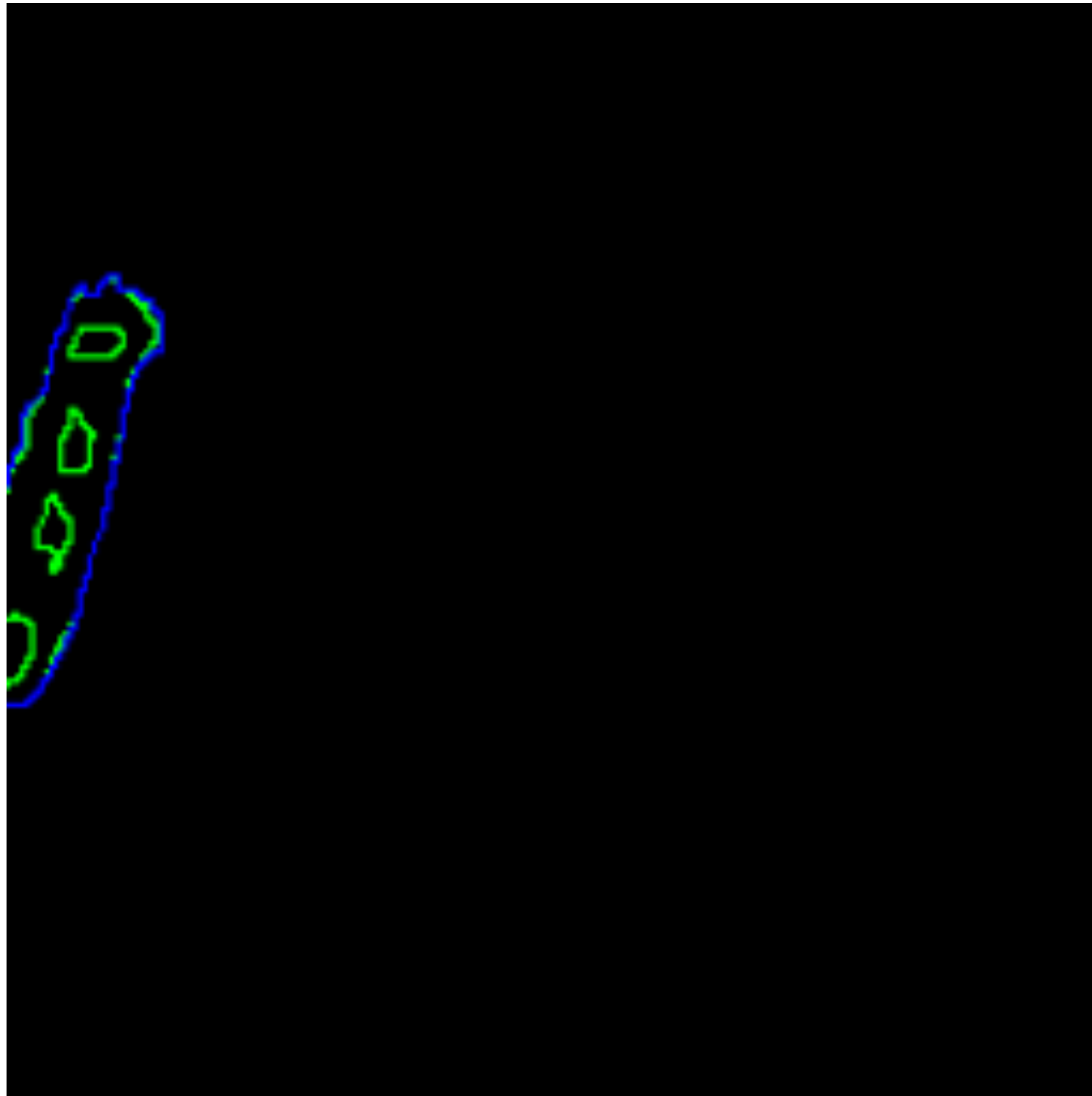 - ≈93 hours in total

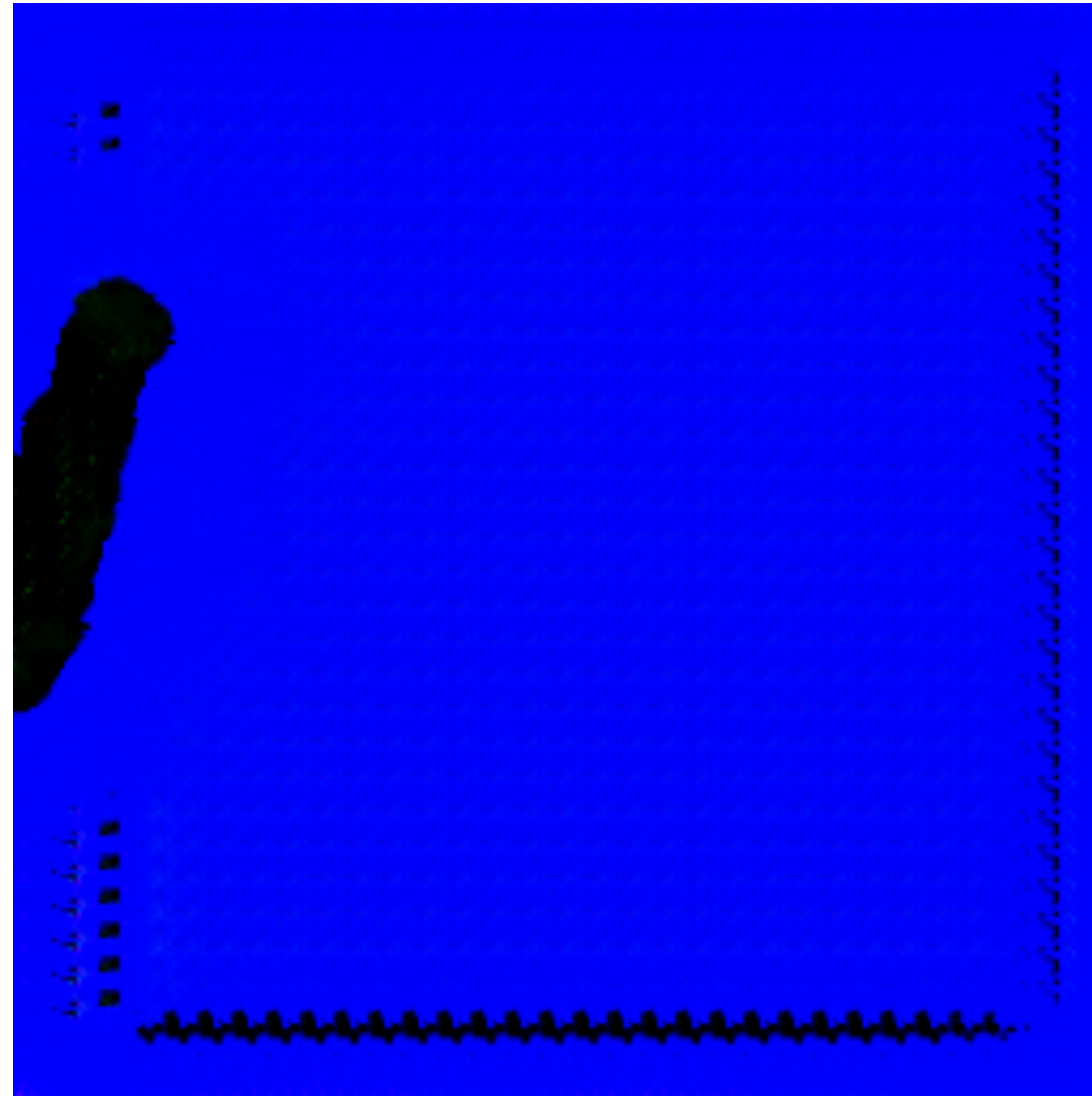# Results of pix2pix Model
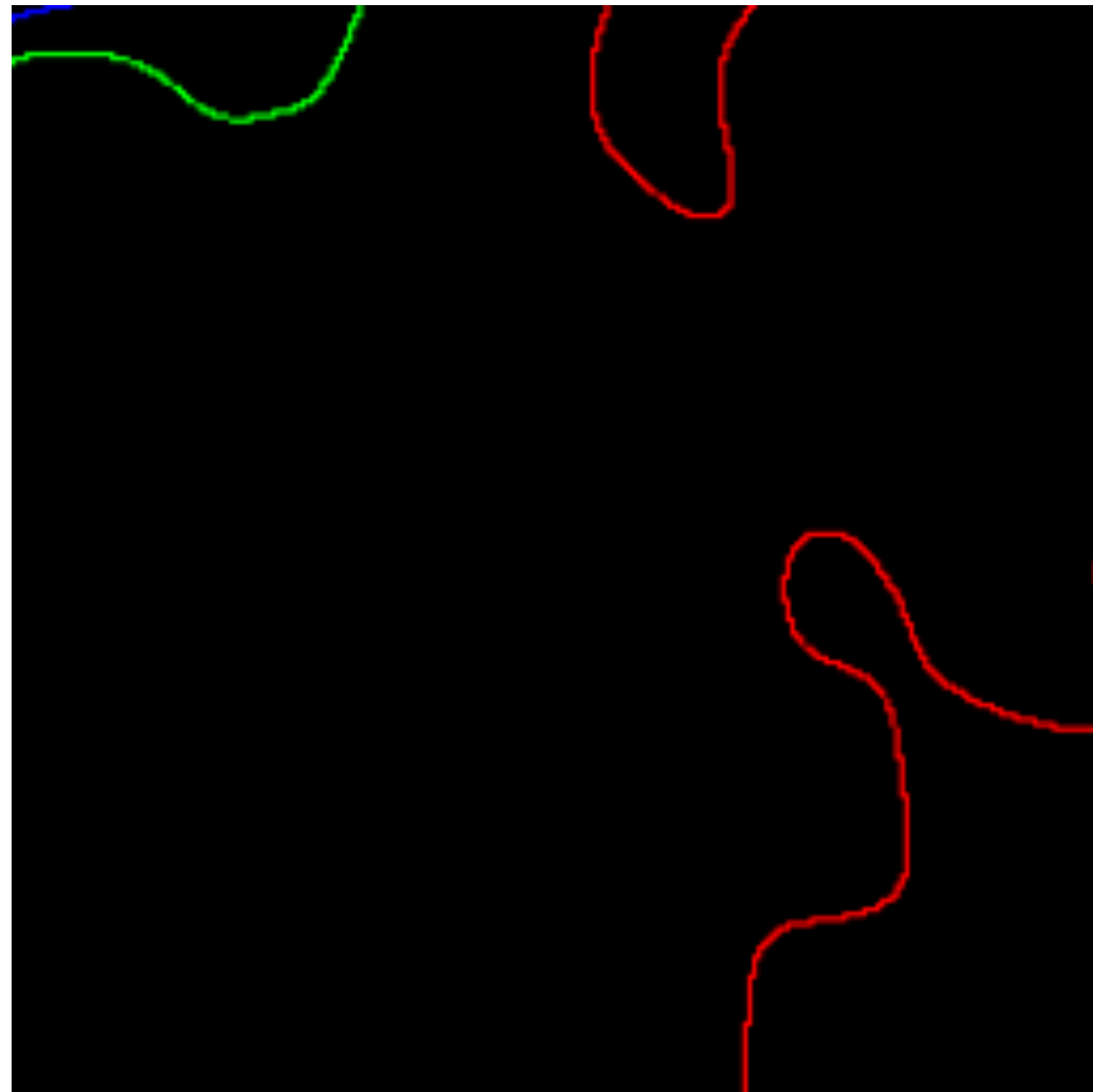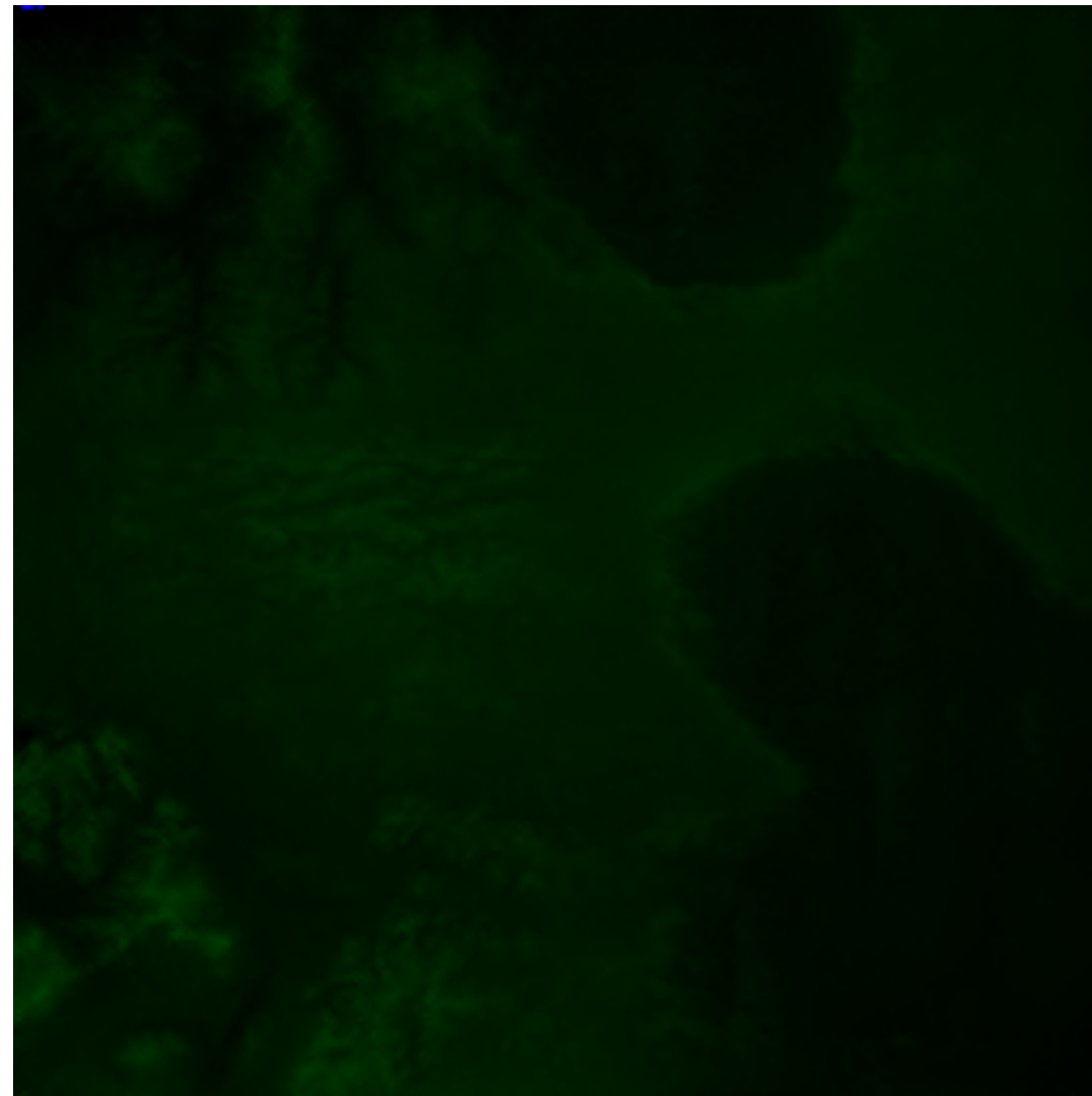
Input

Output (1200 epoch)

Ground Truth

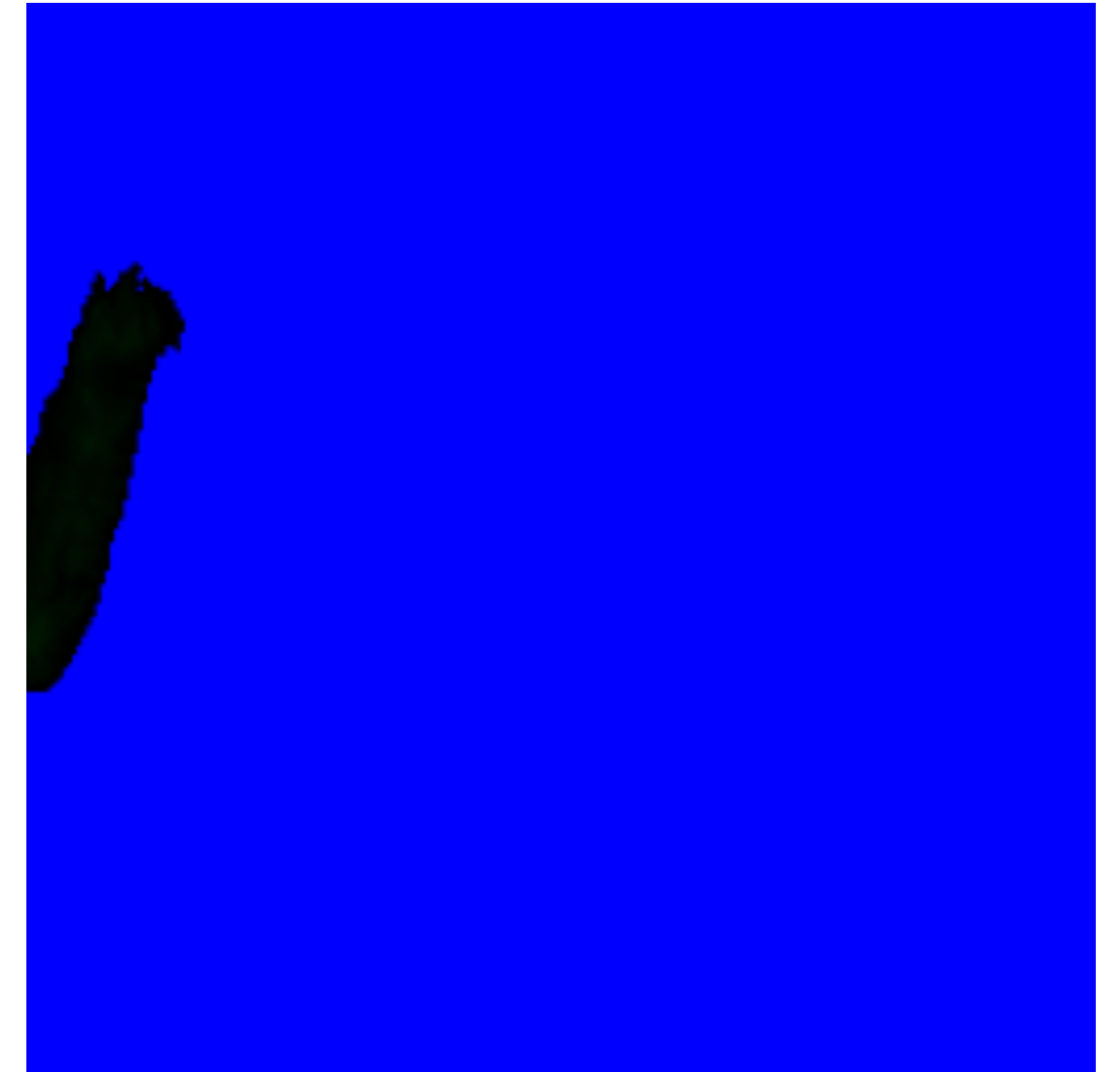*Terrain Design Prototyping With Diffusion Models*

Results of pix2pix Model

Input

Output (1189 epoch)
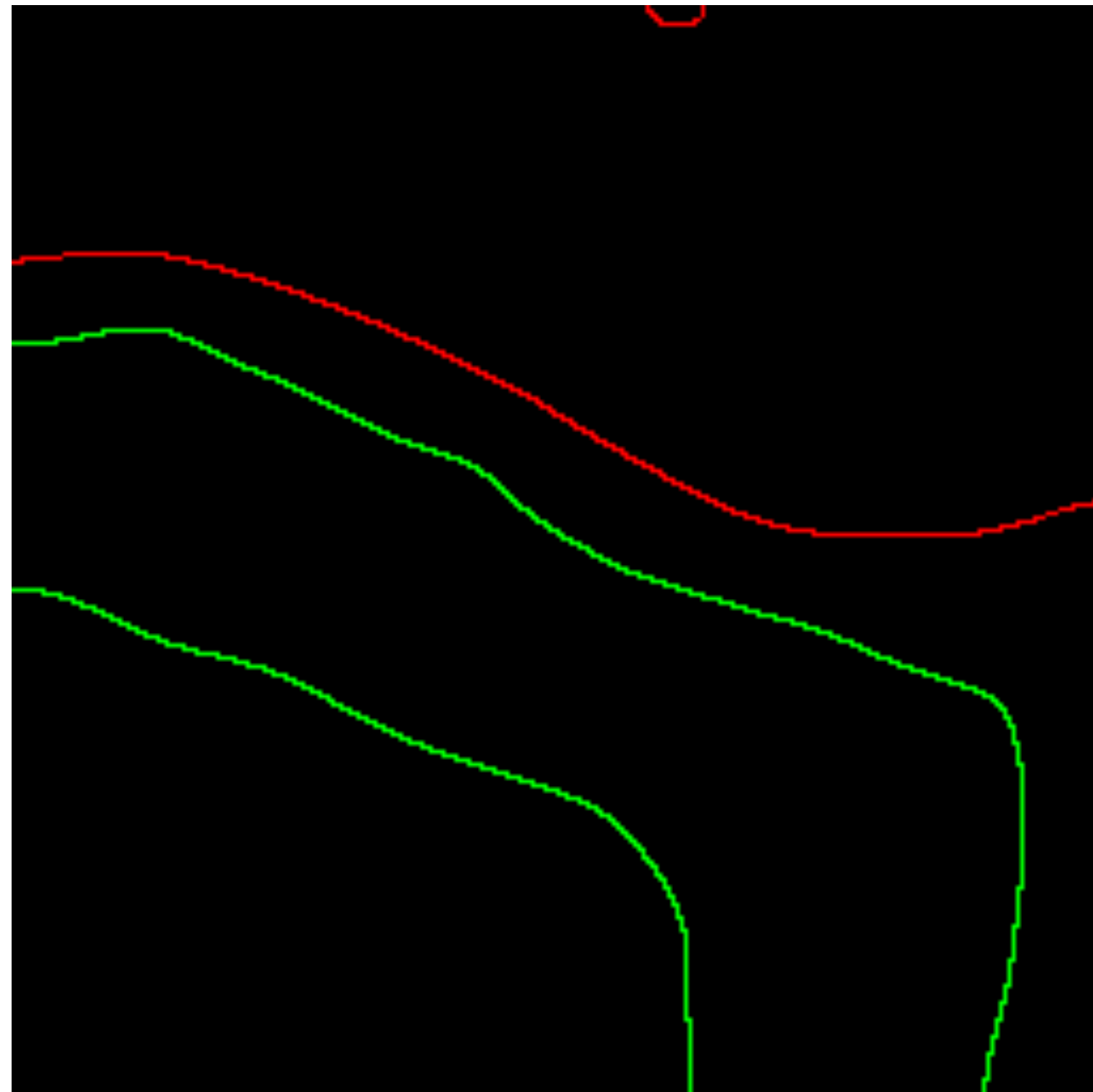
Ground Truth

Results of pix2pix Model

- Problem: Sometimes the output has an unnatural pattern

  - Recognizable especially when the area is full of water

- Potential cause

  - The lack of training epoch
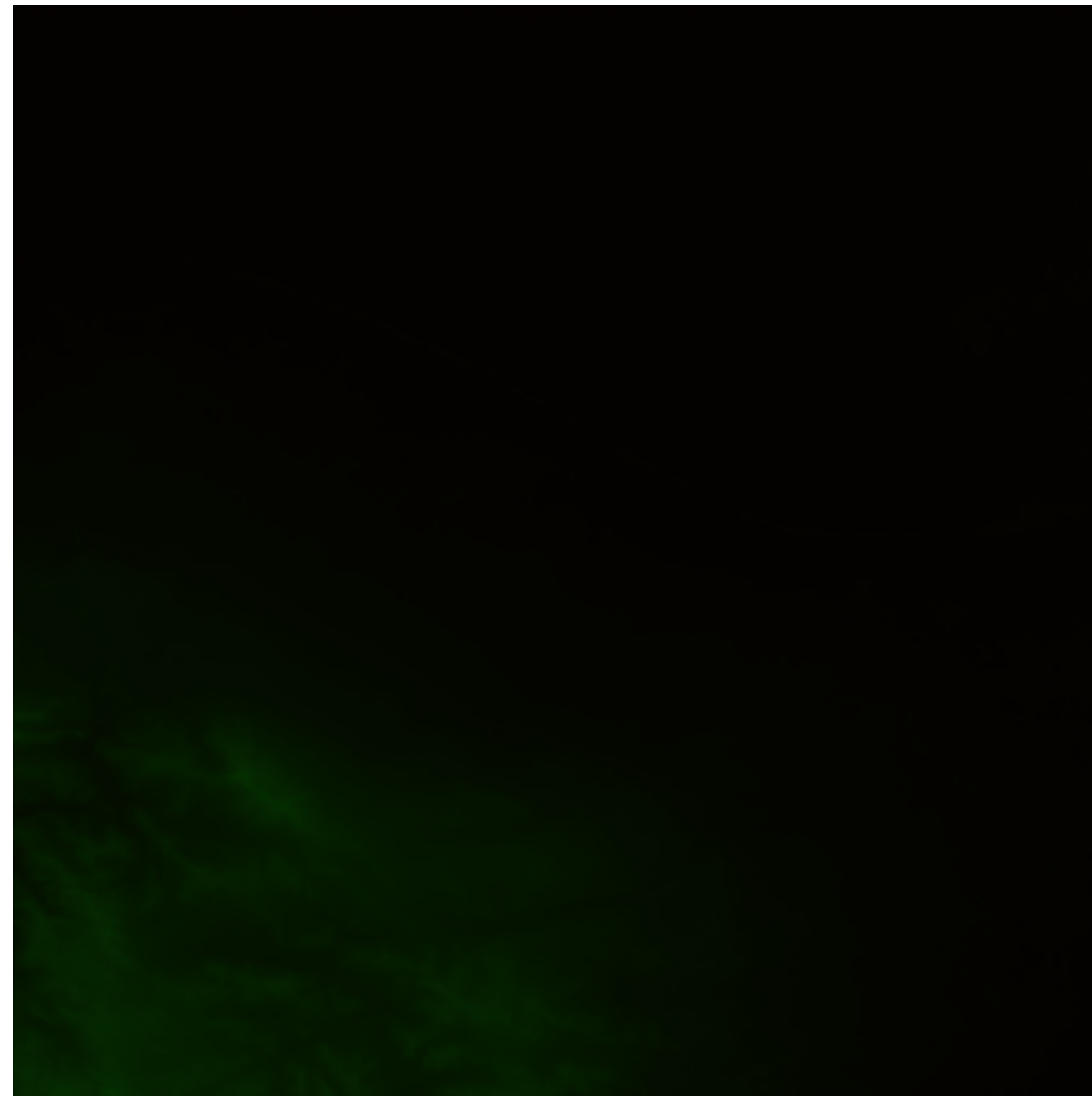
  - The limitation of pix2pix

Training of Palette Model

- Now we are training the Palette model

  - 520/1200 epochs completed

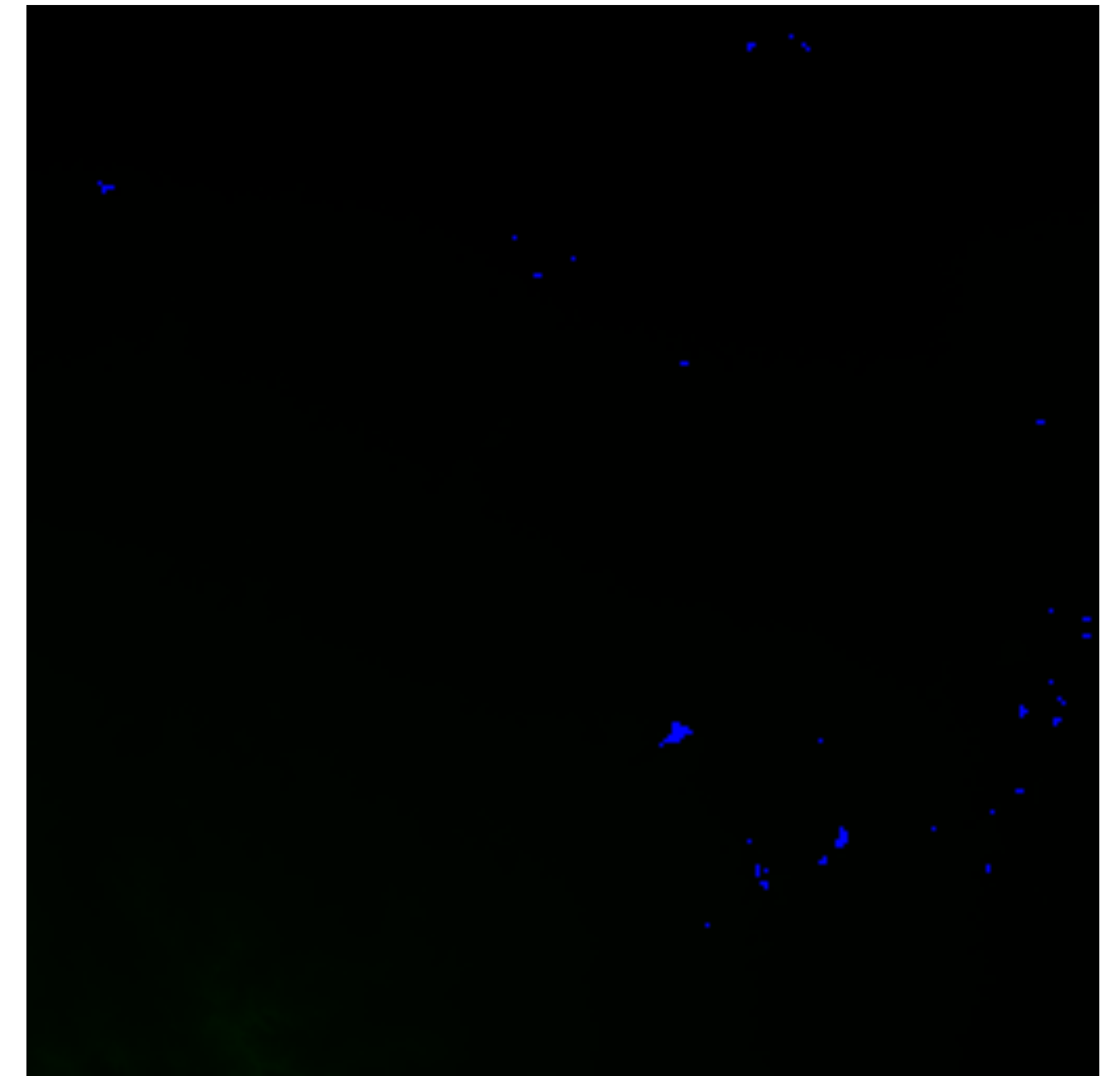 - ≈24 minutes per epoch

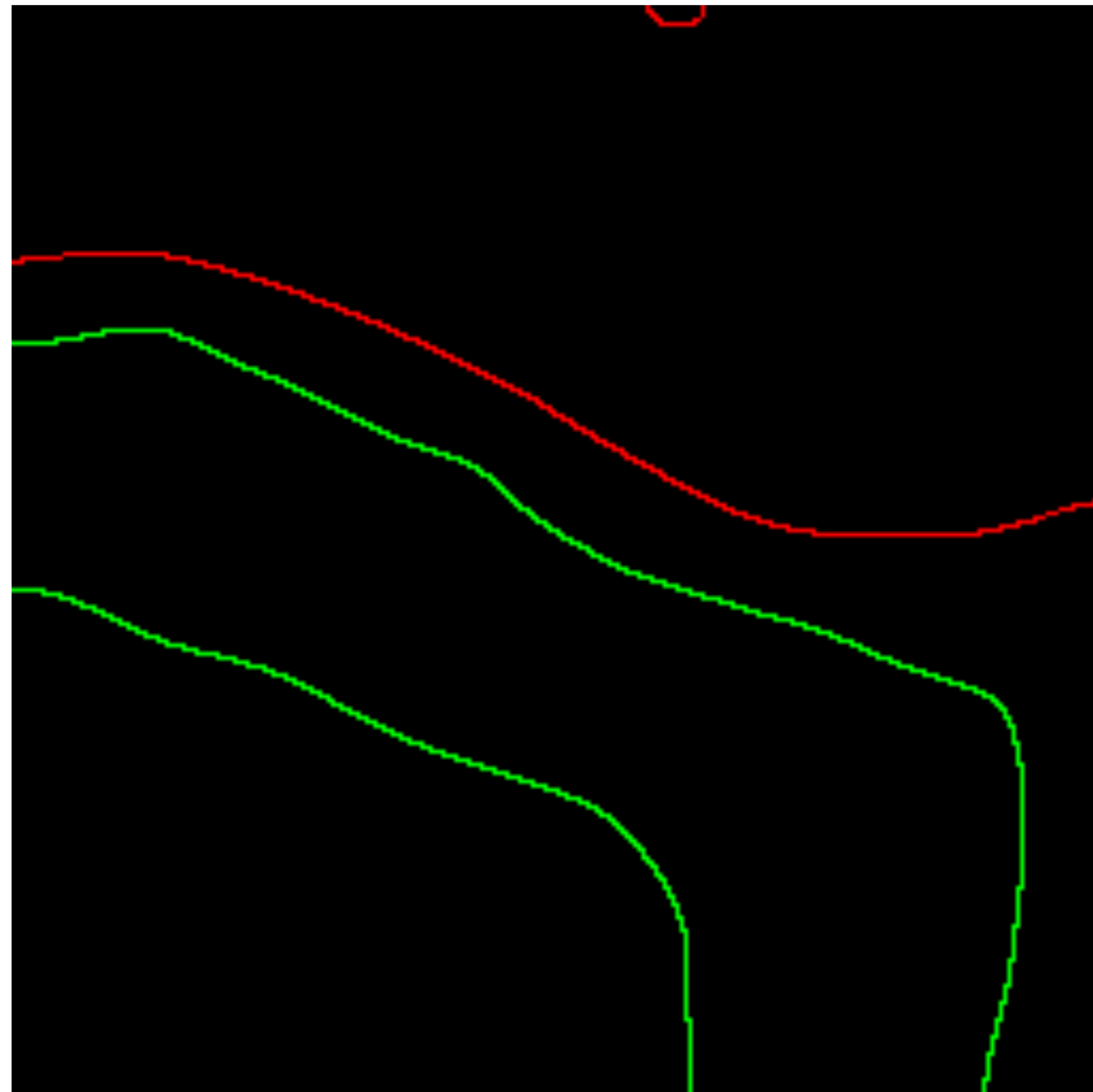 - ≈20 days in total, ≈11 days to complete
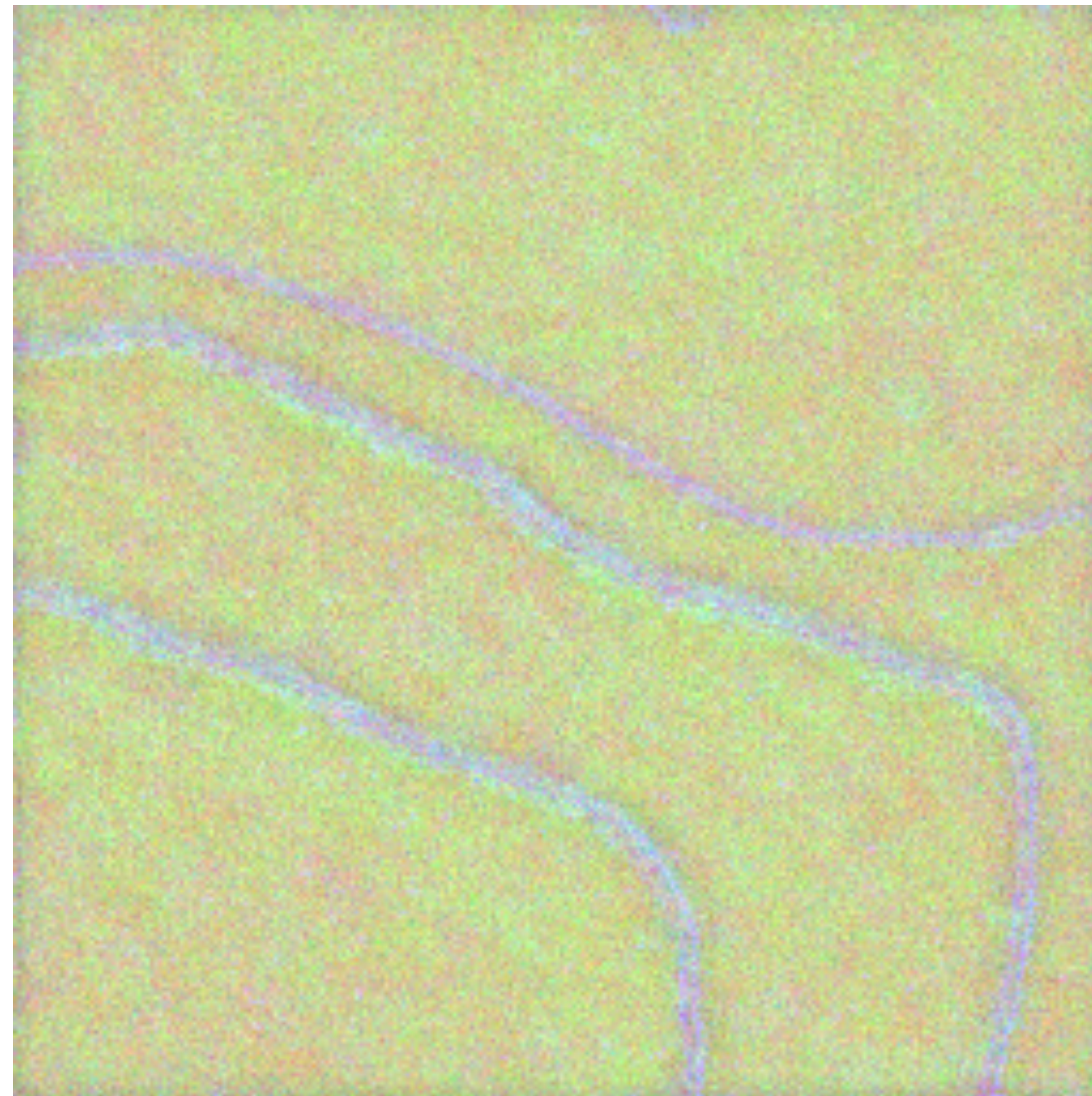
Results of Palette Model

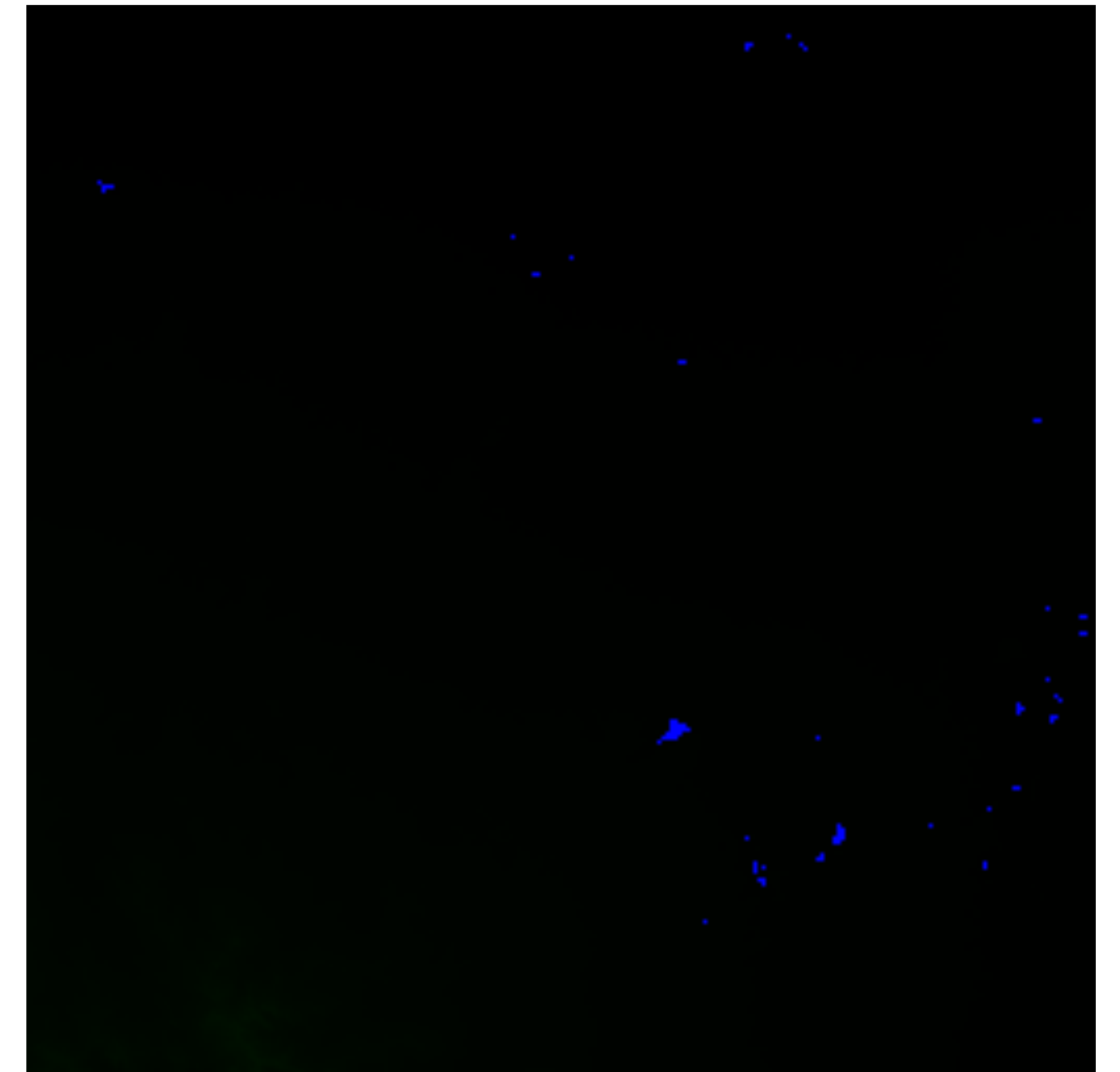Input

Output (470 epoch)

Ground Truth

Results of Palette Model

Input   Output (485 epoch)   Ground Truth

Results of Palette Model

- Problem: Training is unstable. Most of the time the output is completely full of noise

- Potential cause

  - The lack of training epoch

Thank you for your attention