Metal cover generation

Christoper Hansen and Niko Schmidt

12. Dezember 2021

Overwiew of the topics

- Introduction
- Startingpoint
- Data
- Problems
- Results
- Outlook
- Summary

Introduction

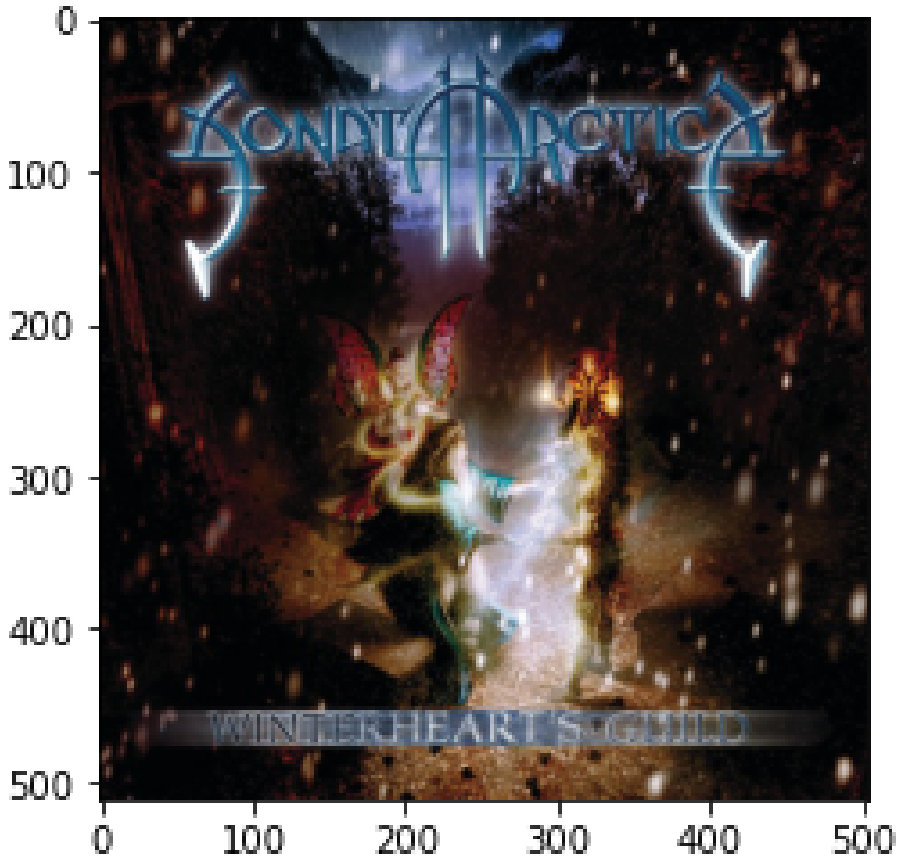
What?
Generating metal covers.
Why?
It would be nice to have an automatic cover generation corresponding to the music style. That could save bands some money.
How?
We use GANs. That means we use two neural networks. One tries to generate the covers and one tries to detect them.

Startingpoint

We startet from Course 1 week 3 of the Coursera notebook, which is the Wasserstein GAN with gradient penalty.
We adapted it for other picture sizes.

Data for the critic

We have the real data for the critic from here:
https://www.kaggle.com/benjamnmachn/metal-album-artwork-dataset-intro
One picture is for example:

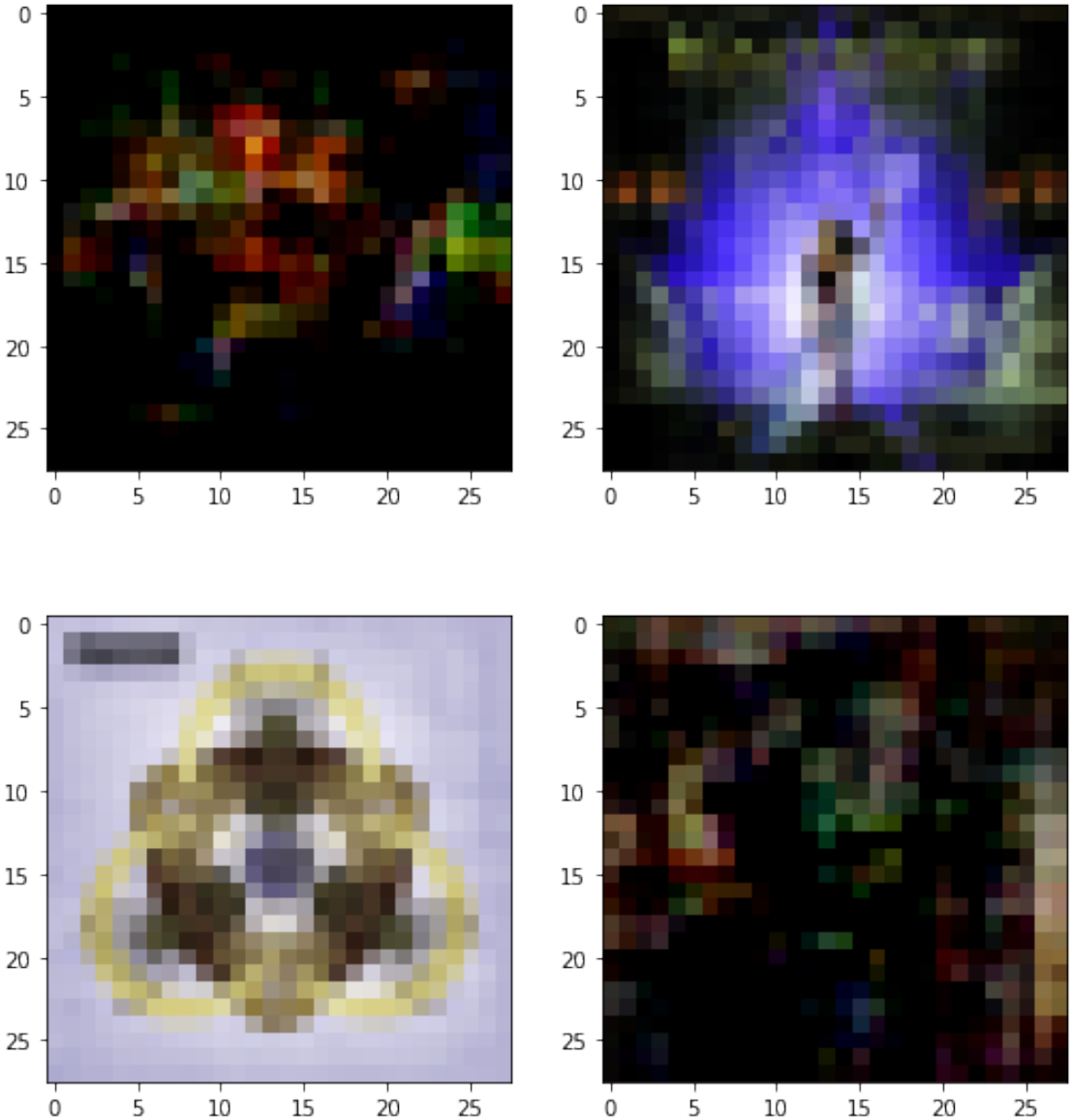


Problems while programming

We had to handle some problems:

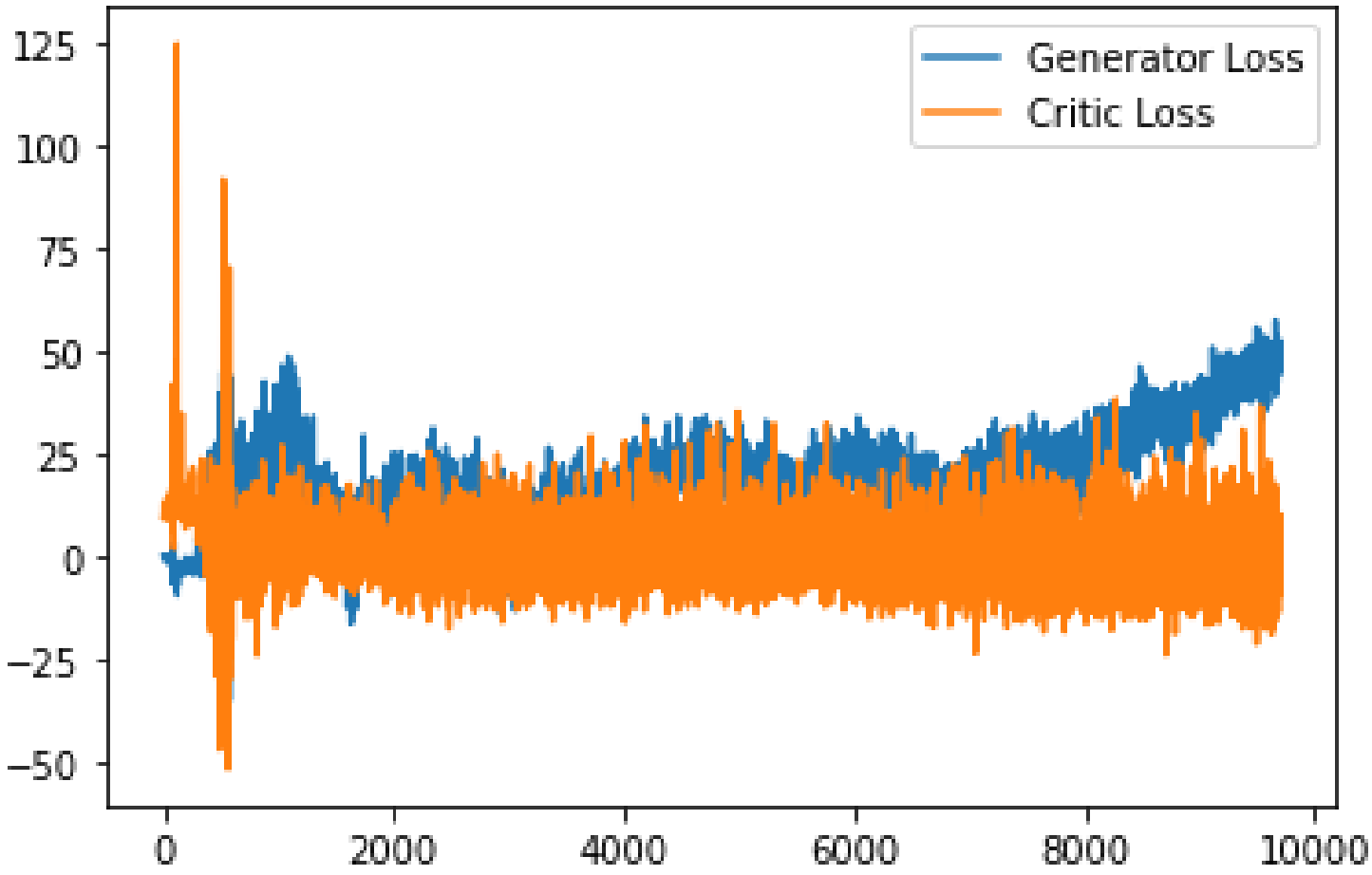
1. We had one picture with another pixel size. We removed it.
2. Corresponding to the WGAN architecture we needed to have the shape of the picture in the right order. But one picture had the same shape, but it was in the format (x, y, z) instead of $(x, y, z, 1)$. Due to that issue we had different cases for changing the order of the tuple.
3. The computation time for GANs is really high. Therefore we computed everything on a graphic card.
4. As the computation still needed to much time, we had to make the size of the pictures smaller.
5. Not only the computation of each iteration was slower with more pictures, it also needed more iterations for higher resolutions.

Can you decide, which is real and which is fake?

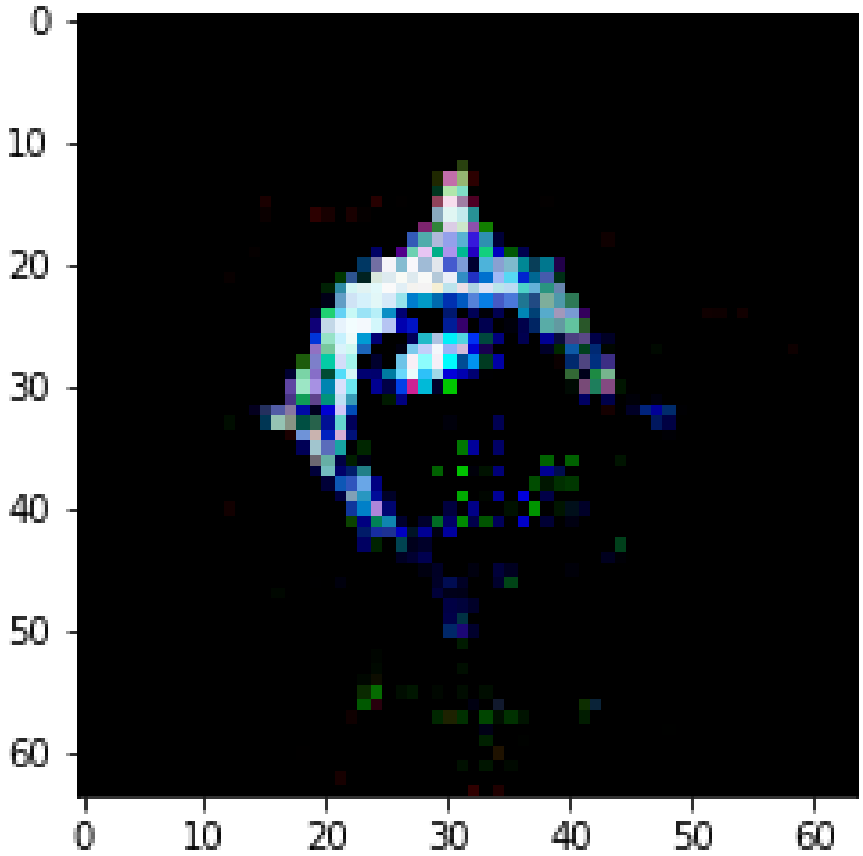


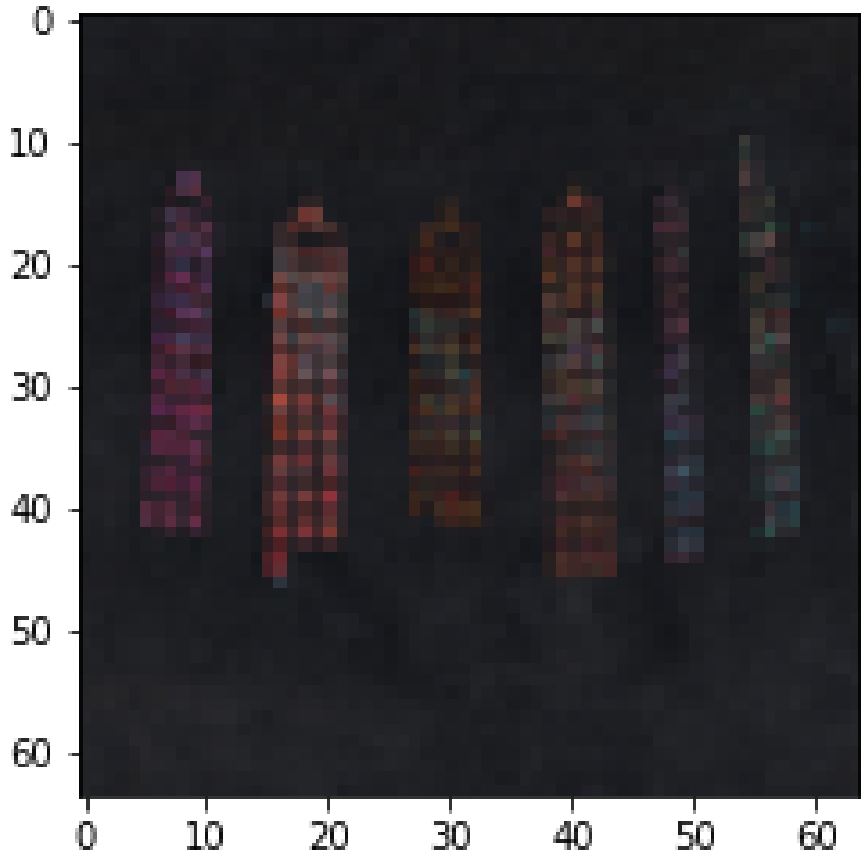
Answer

The upper right and bottom left are real. The other two are fake. The loss for images of size 28×28 is



Pictures with 64 pixel resolution

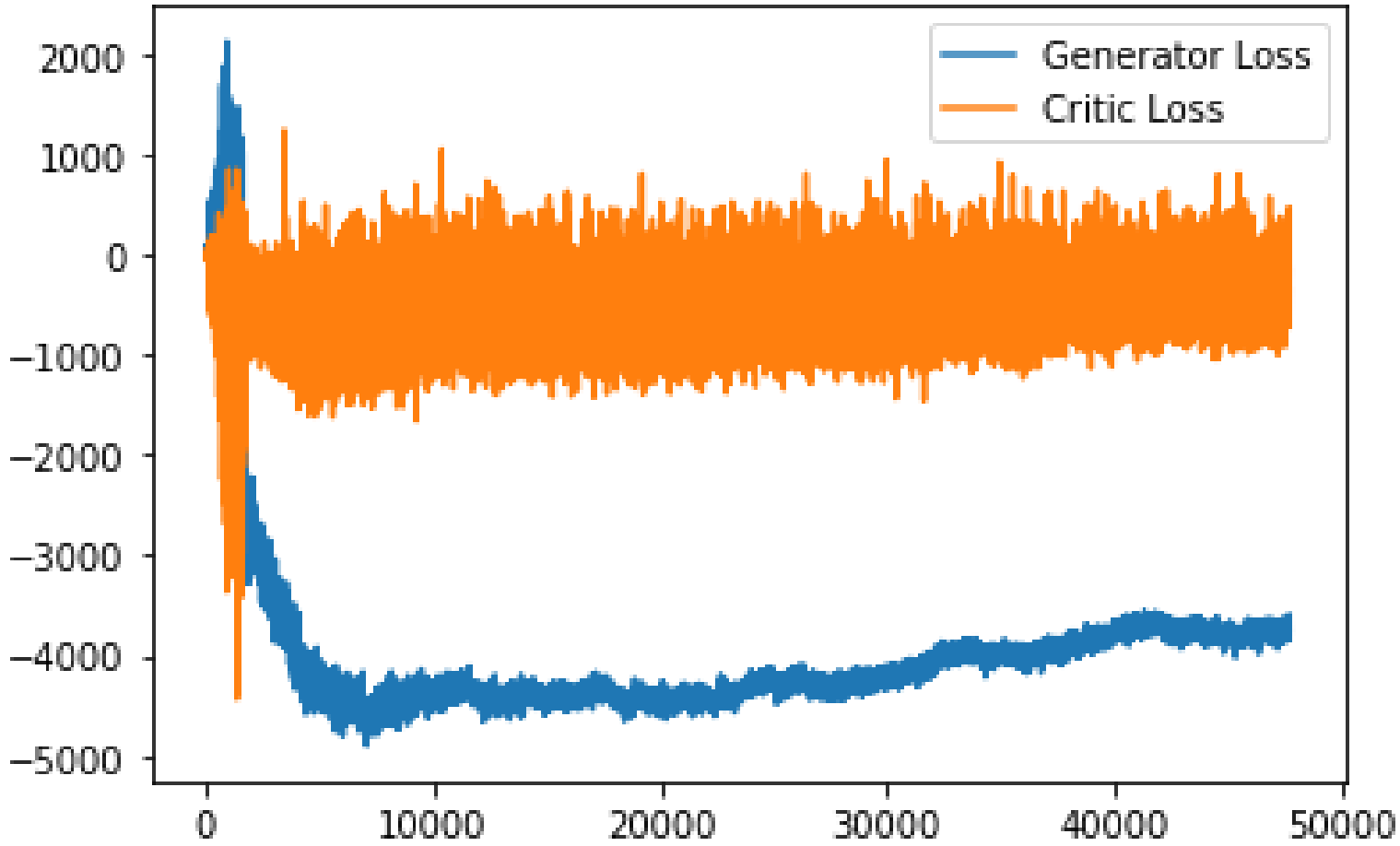




Which one is real and which one is fake?

Answer

The left is fake and the right is real. The loss for images of size 64×64 is



Computation time

The computation time on a graphic card was just really high:

- 28×28 pixel: 6 h
- 64×64 pixel: 50 h
- 128×128 pixel: too much

Summary

- The generator in the WGAN can produce metal cover pictures.
- But the training of the network takes a lot of resources.
- Also higher resolutions are not so easy to generate.

Outlook

- Higher resolutions would be good. For this aim one could maybe implement upscaling.
- Something like GauGAN would make the generation more individual and therefore artists would be able to generate their covers really easy.