

WINGED HORSES WITH A GAN

Matthew Chapman

ABSTRACT

This paper proposes using a deep convolutional generative adversarial network (DCGAN) to generate images that look like a Pegasus. This abstract should be short and concise, about 8-10 lines long.

1 METHODOLOGY

First, we trained a deep convolutional generative adversarial network (DCGAN) [] on only the images with the ‘horse’ label in the CIFAR-10 dataset, by minimising the GAN loss function,

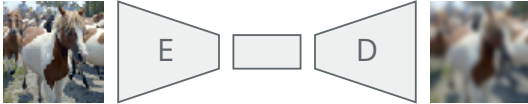
$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (1)$$

where x is the data representing an image, $D(x)$ is the discriminator network output, z is a latent space vector sampled from a standard normal distribution, $G(z)$ is the output of the generator, and p_{data} is the distribution the training data comes from. We did this with the hope of generating images of horses which we can later transform into pegasi. We trained the network for 5, 50, 100, 150, and 200 epochs separately, with a batch size of 64, and generated fake images each time (see Appendix A). By qualitative comparison of the results, it was decided that training for 100 epochs generated the most realistic images (see Section 3 for the limitations of using DCGANs).

Second, we tried fine-tuning the pre-trained network on only the images with the ‘bird’ label in the CIFAR-10 dataset. We did this thinking that the network might generate images of horses with some bird features, thus resembling a pegasi. We fine-tuned the network for 5, 25, and 50 epochs separately, with a batch size of 64, and generated fake images each time (see Appendix B). By inspection of the results, it was decided that this was not a suitable approach for producing pegasi.

Third, we tried fine-tuning the pre-trained network on a hand-selected image (see Appendix C). We did this hoping that the network would recognise the key feature shared by birds and pegasi — wings — and reflect this in its output. We fine-tuned the network for 5, 25, and 50 epochs separately, with a batch size of 1, and generated fake images each time (see Appendix D). By inspection of the results, and given time constraints, it was decided that this method would have to suffice.

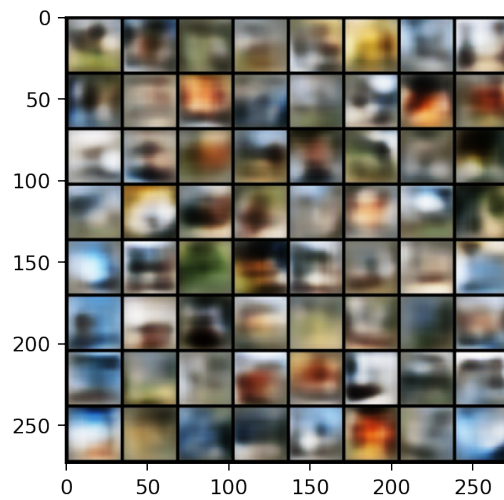
The architectural diagram for our approach is as follows:



The architectural diagram above was created using Inkscape and exported to a PDF. This was then uploaded to the figures directory on the left.

2 RESULTS

The results look very blurry, where the best batch of images looks like this:



From this batch, the most Pegasus-like image (with quite a stretch of the imagination) is:



3 LIMITATIONS

It's very difficult to see anything that looks like a Pegasus. In the future, this could be improved by training for more than 10 epochs, although this was not possible due to the time constraints.

BONUSES

This submission has a total bonus of -4 marks (a penalty), as it is trained only on CIFAR-10, and the Pegasus has a dark body colour.