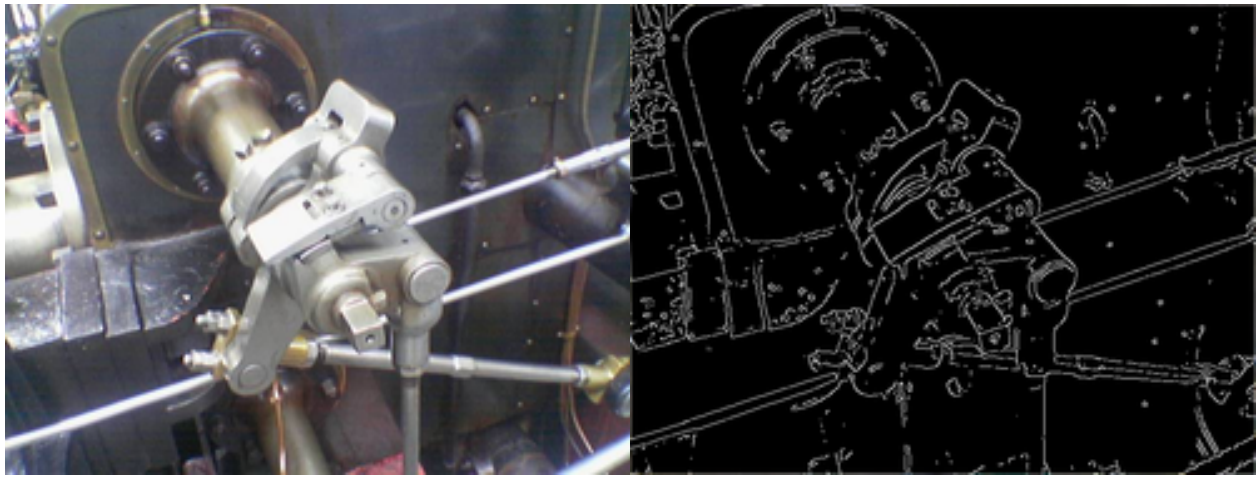
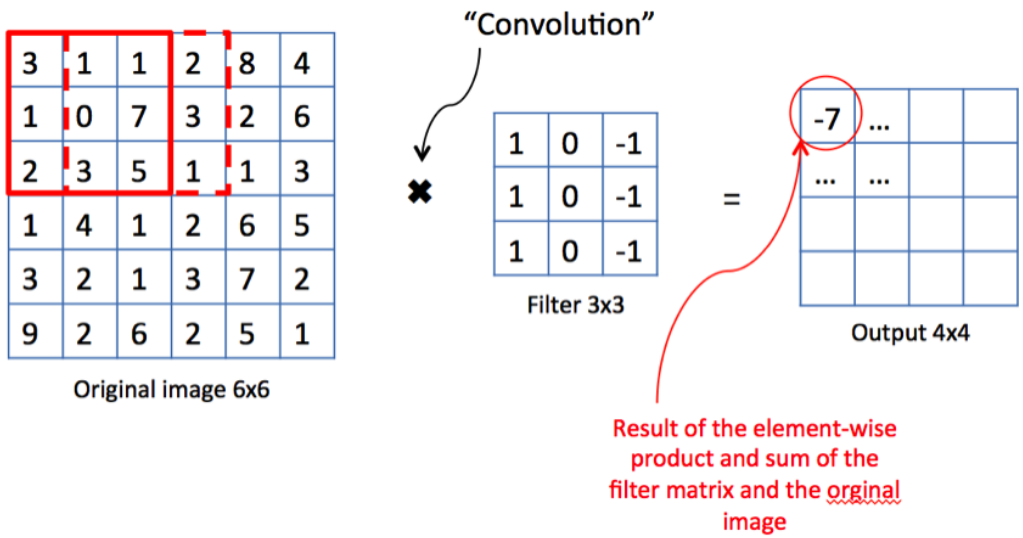
Images

Convolution

How to do image processing with a neural network?

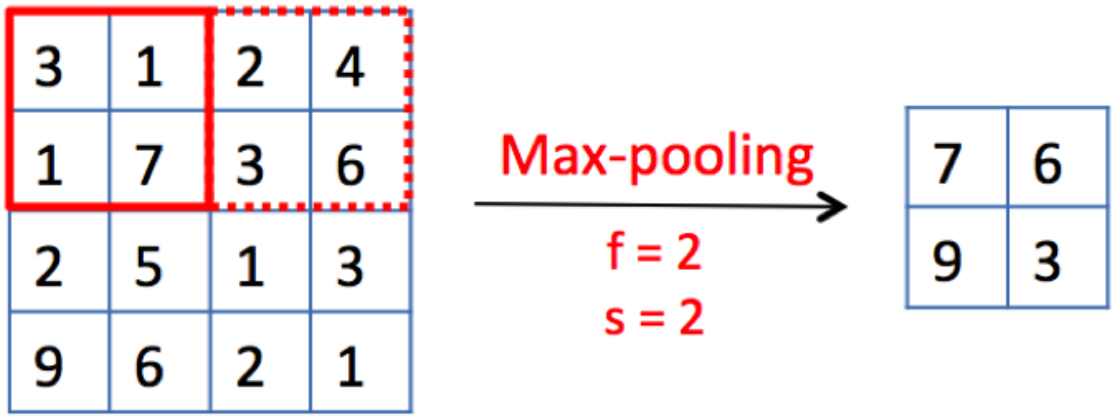




https://en.wikipedia.org/wiki/Canny_edge_detector

https://medium.com/machine-learning-bites/deeplearning-series-convolutional-neural-networks-a9c2f2ee1524

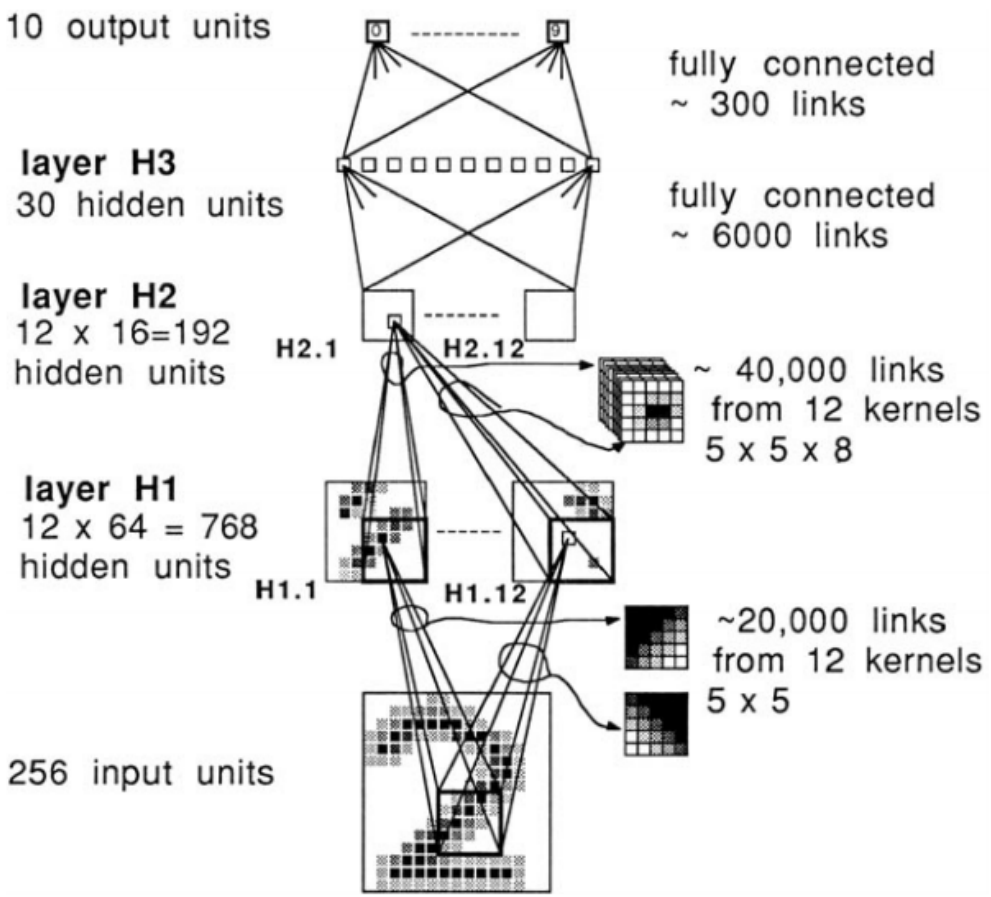
Pooling

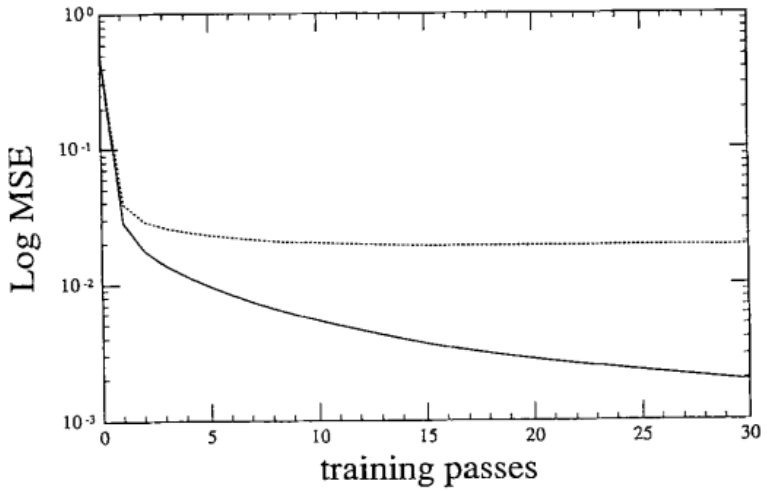


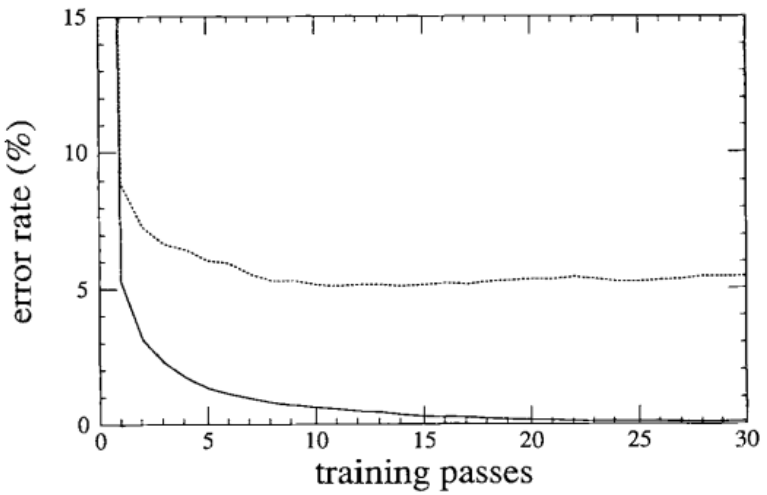
https://medium.com/machine-learning-bites/deeplearning-series-convolutional-neural-networks-a9c2f2ee1524

LeCun et al (1989)

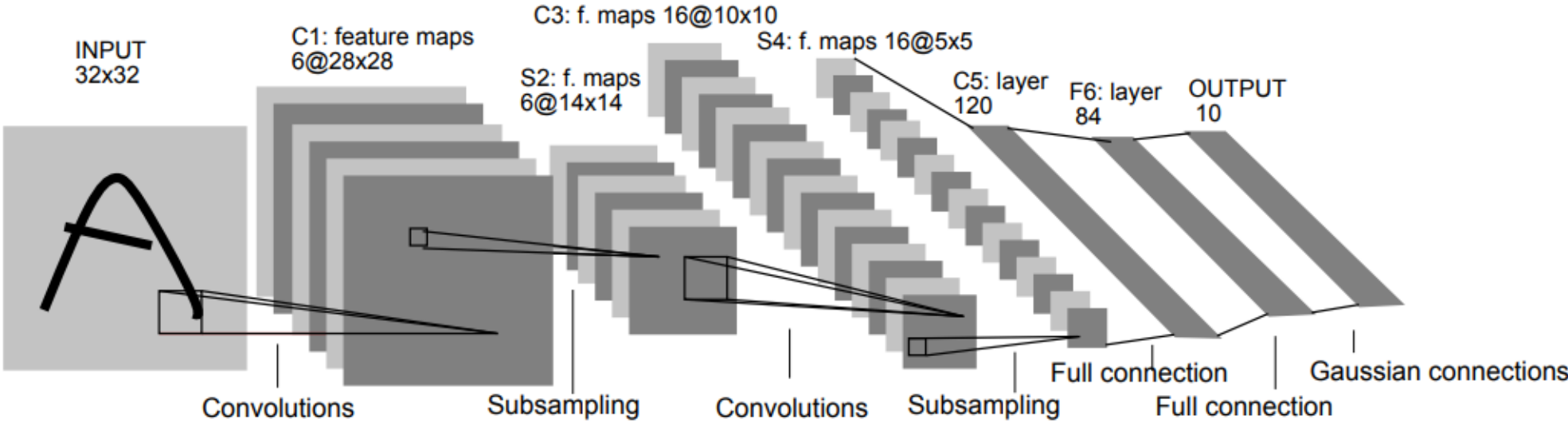
LeCun Y et al (1989) Backpropagation Applied to Handwritten Zip Code Recognition. Neural Com put 1:541–551.







LeCun et al (1998)



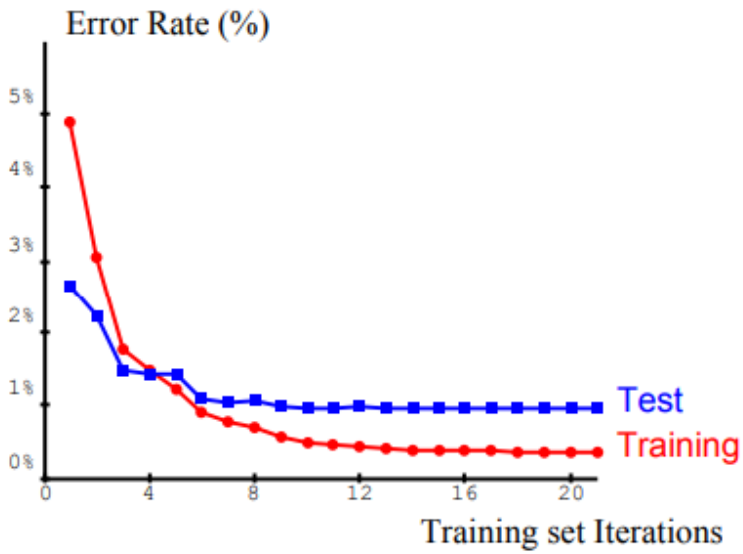
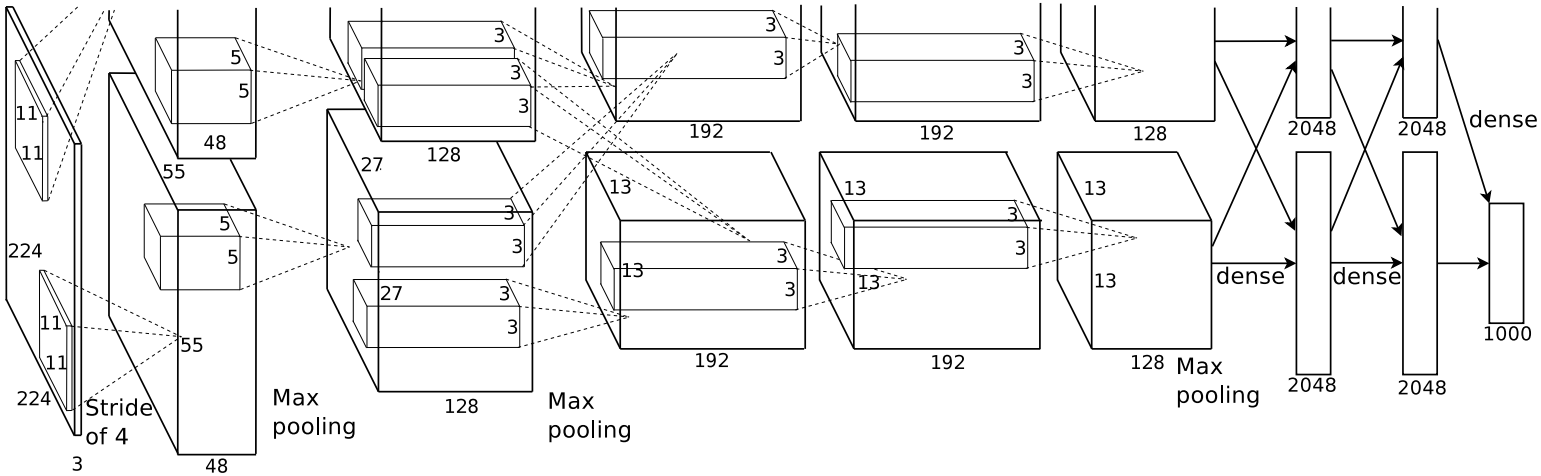


Image classification (Krizhevsky et al 2012)

- Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems 25, Burges CJC, Bottou L, Weinberger KQ, eds), pp 1097–1105.
- Annual competition. 1,000 examples of 1,000 classes.
- Rotating/translating of training images to make 'new' samples.
- Top-1 (37.5%) and top-5 (17.0%) error rate exceed previous state of the art. Won 2012 competition with top-5 test error rate of 15%; second had error rate of 26%.

ImageNet architecture (Krizhevsky et al. 2012)



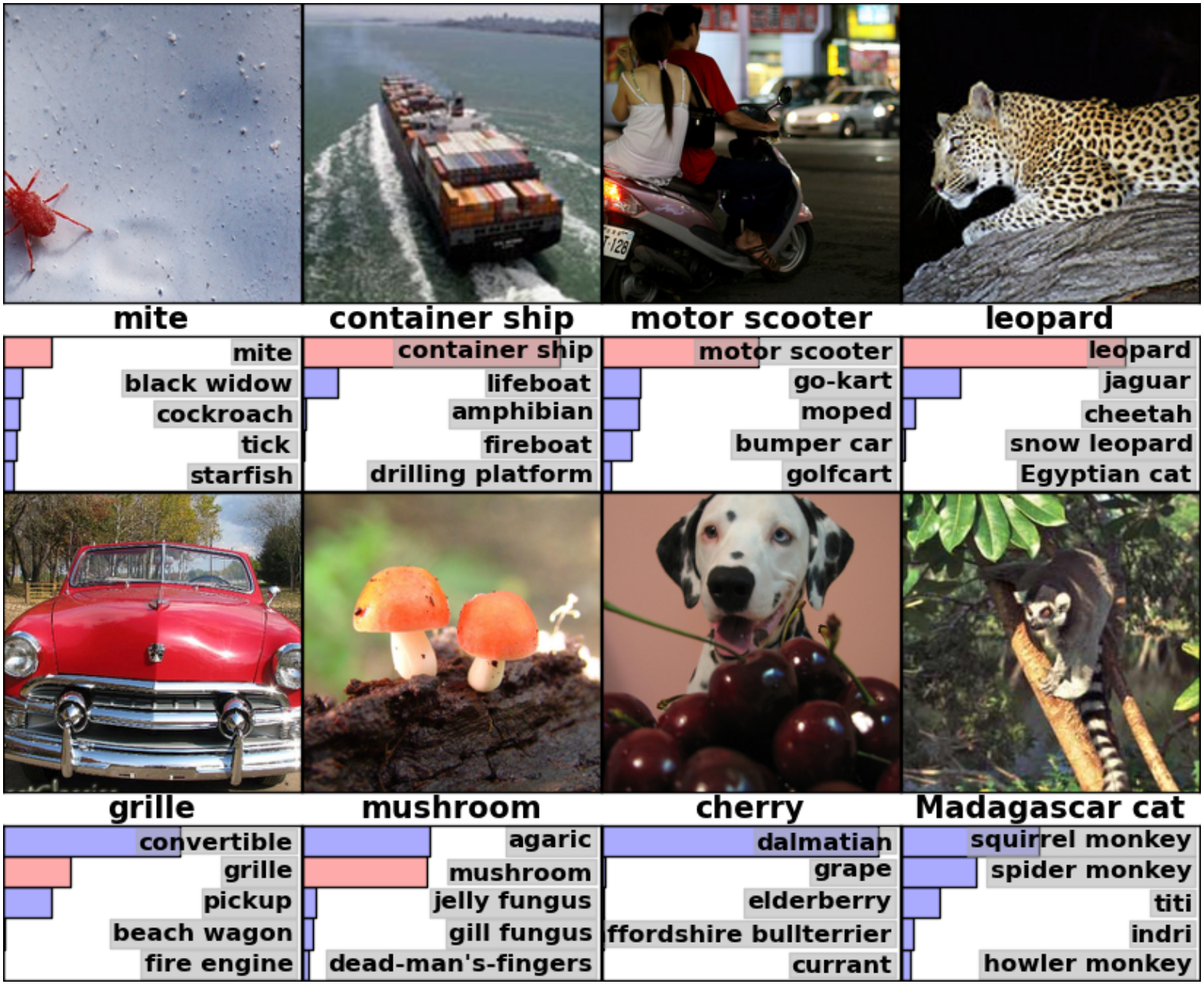
Split on 2 GPUs.

150K pixels (224 x 224 x 3) into 8 layer network:

253K-186K-65K-65K-43K-4K-4K-1K.

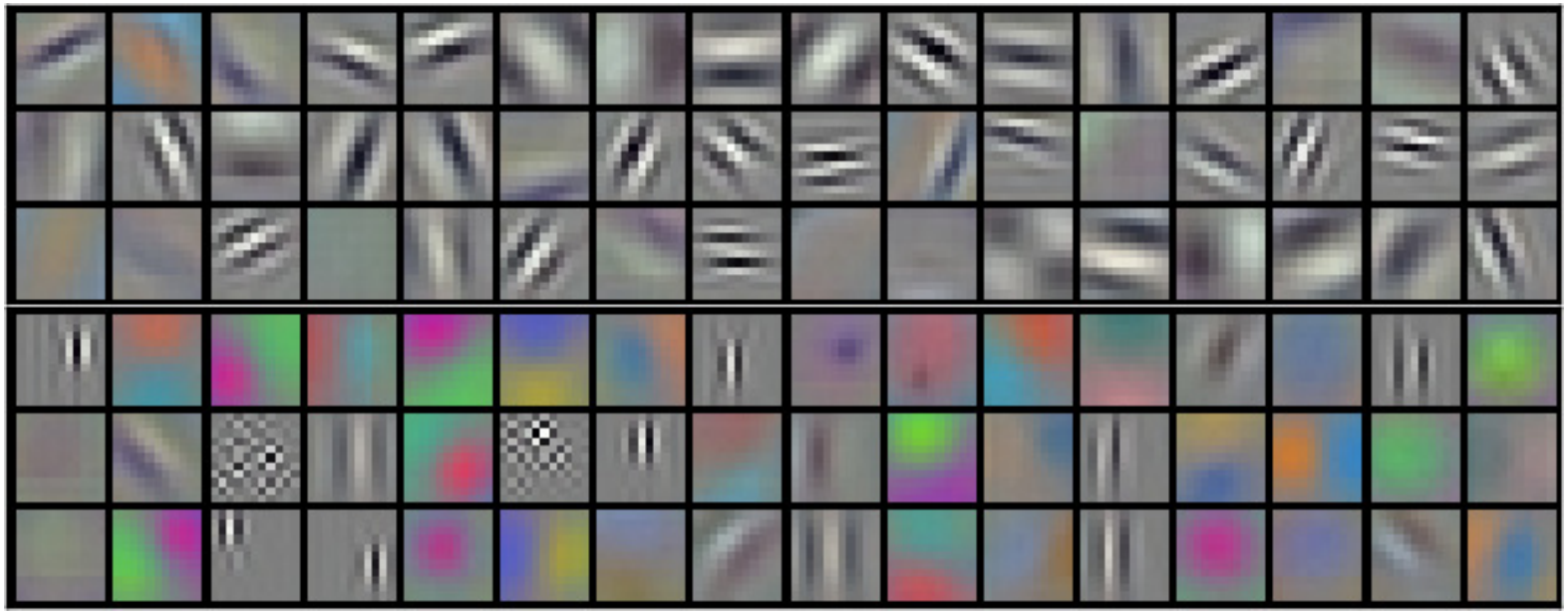
i.e. 650,000 neurons (excluding input), 60 million parameters.

ImageNet performance



ImageNet close-matches

ImageNet RFs



“It is notable that our network’s performance degrades if a single convolutional layer is removed. For example, removing any of the middle layers results in a loss of about 2% for the top-1 performance of the network. So the depth really is important for achieving our results.”

Life since AlexNet

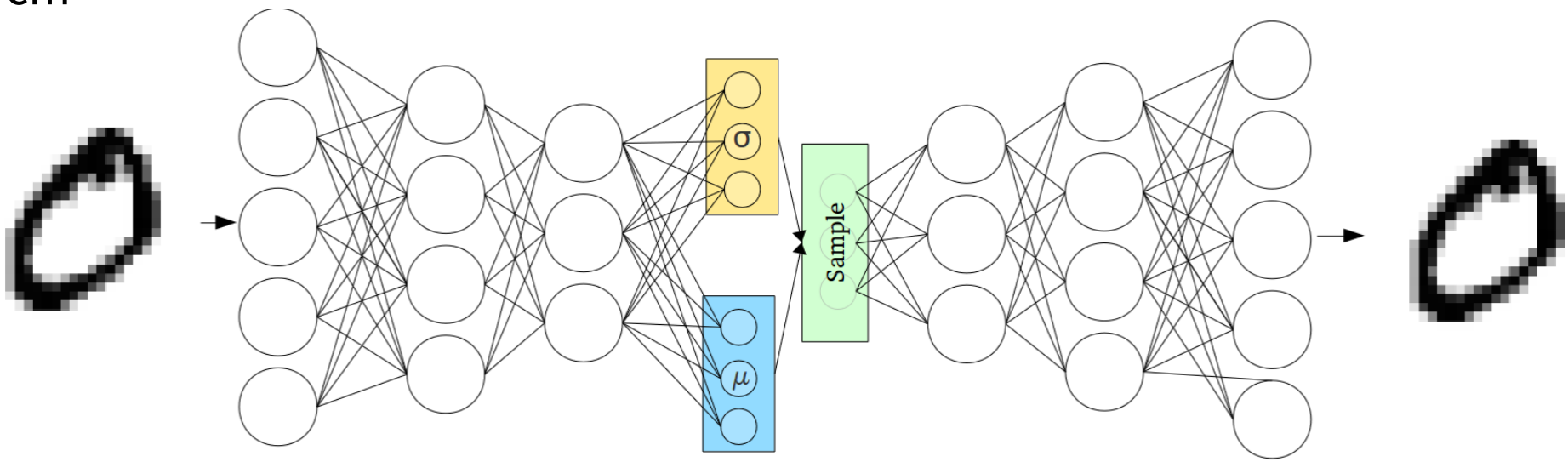
1. 2012: AlexNet (Krizhevsky et al 2012). Top-5 error of 17% (top-1 of 37%).
2. 2014: GoogLeNet / “Inception” (Szegedy et al 2014). Top-5 error of 6.67%.
3. 2015: ResNet (He et al 2015). First to beat human, top-5 error of 3.75%. 152 layers.
4. ResNet with 1001 layers (He et al 2016).

Transfer learning using these networks to provide features: freeze main network and learn layers on top.

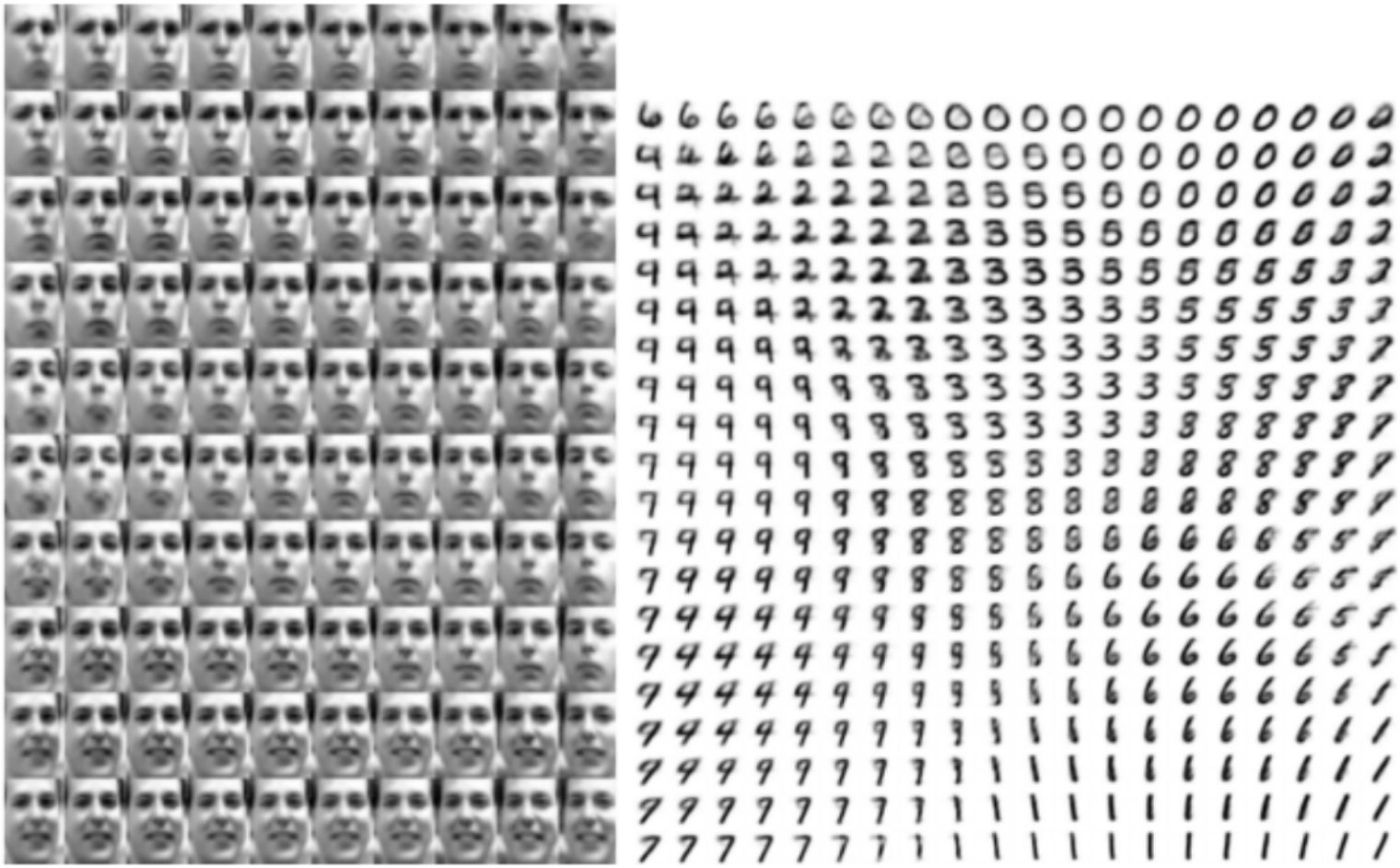
Variational autoencoders

https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf

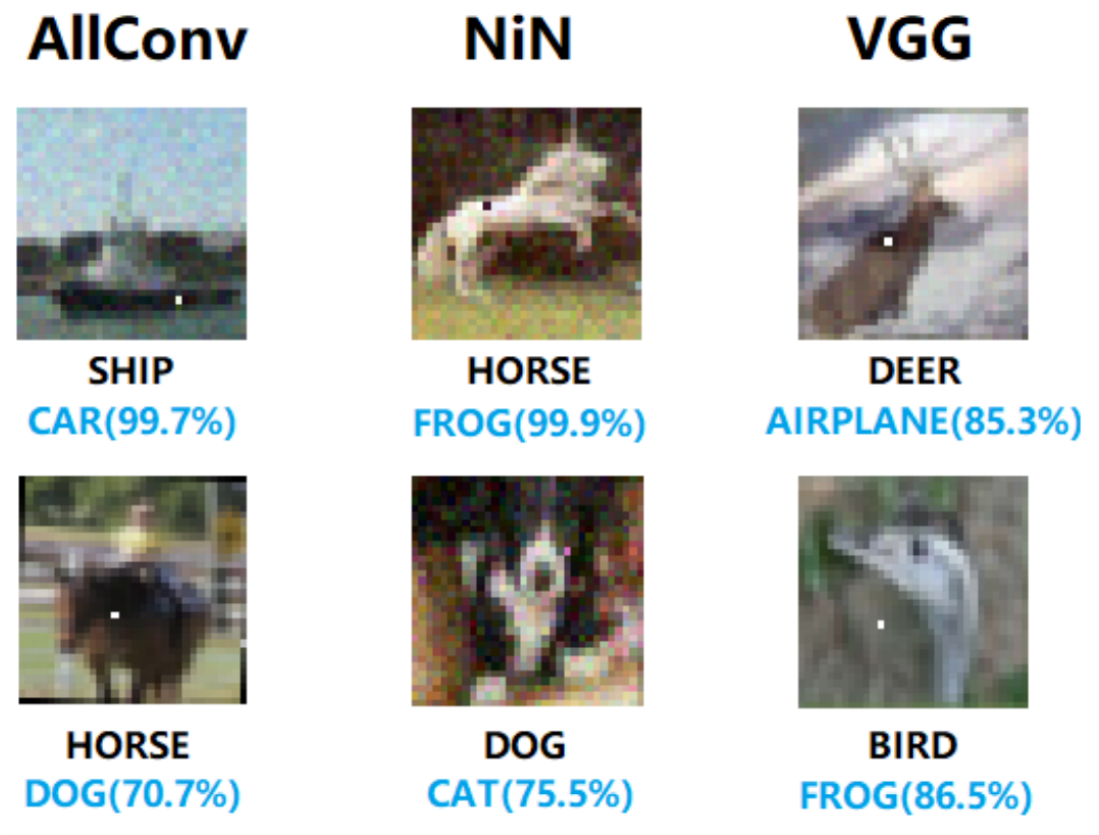
Ucm



Sampling the latent space (Goodfellow , Figure 20.6)



Fooling Deep Networks with adversarial samples



Su et al (2017)

See also `https://arxiv.org/pdf/1707.08945.pdf` for robust attacks on stop signs.

Generative adversarial neworks (Goodfellow et al 2014)

Architecture of a GAN

Discriminator network: spot real vs fake training samples. Adjust weights to increase discrimination.

Generator network: Adjust weights to generate images that are more likely to be classified as training images.

`http://bamos.github.io/2016/08/09/deep-completion/#step-1-interpreting-images-as-samples-from-a-probability-distr`

Radford et al. (2015), figure 4

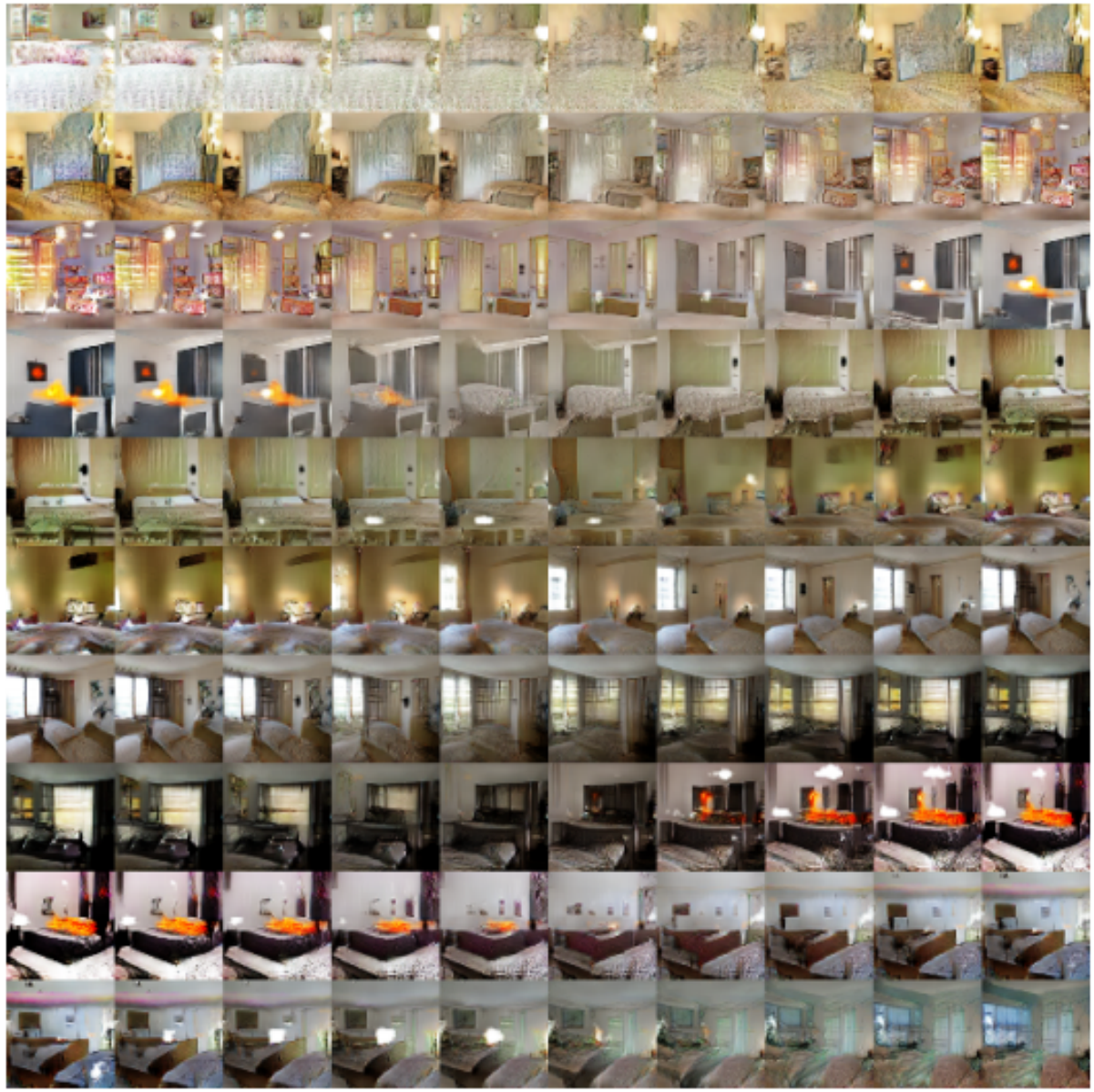


Figure 4: Top rows: Interpolation between a series of 9 random points in Z show that the space learned has smooth transitions, with every image in the space plausibly looking like a bedroom. In the 6th row, you see a room without a window slowly transforming into a room with a giant window. In the 10th row, you see what appears to be a TV slowly being transformed into a window.