Generative Adversarial Networks

Part 1

Jacob Cybulski

What are GANs ?

- **Generative-Discriminative** approach to machine learning relies on a model training strategy which involves probabilistic generation of a sample distribution and discriminative evaluation of a generated sample quality.
- **Generative-Adversarial Networks (GANs)** are machine learning systems that deploy a pair of learning models pitted one against the other in a zero-sum game of creation and critique of a sample distribution.
- Deep Learning models are best suited GAN implementation as they are capable of discovering and relying on the latent representation of high-volume and high-dimensional data, e.g. images or text.

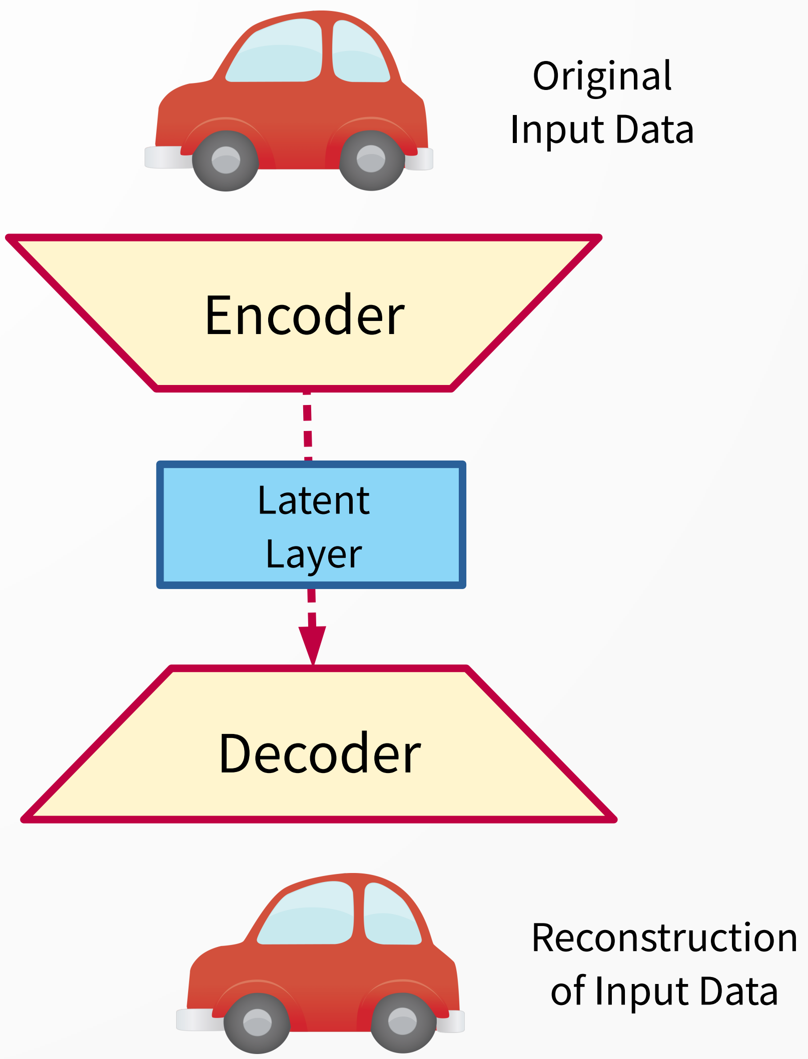
First... Autoencoders (AE)

Autoencoders are neural networks which allow learning of the latent (hidden) representation of data.

An AE is a deep neural network with input matching output and a much smaller middle layer. The network is then trained to reconstruct its input (e.g. a picture, text, music, speech, or business data).

When the network performance improves, the middle layer seems to effectively compress its input and decompress it on output, i.e. it learns the essential features of input data.

The mid-layer becomes a latent representation of data.



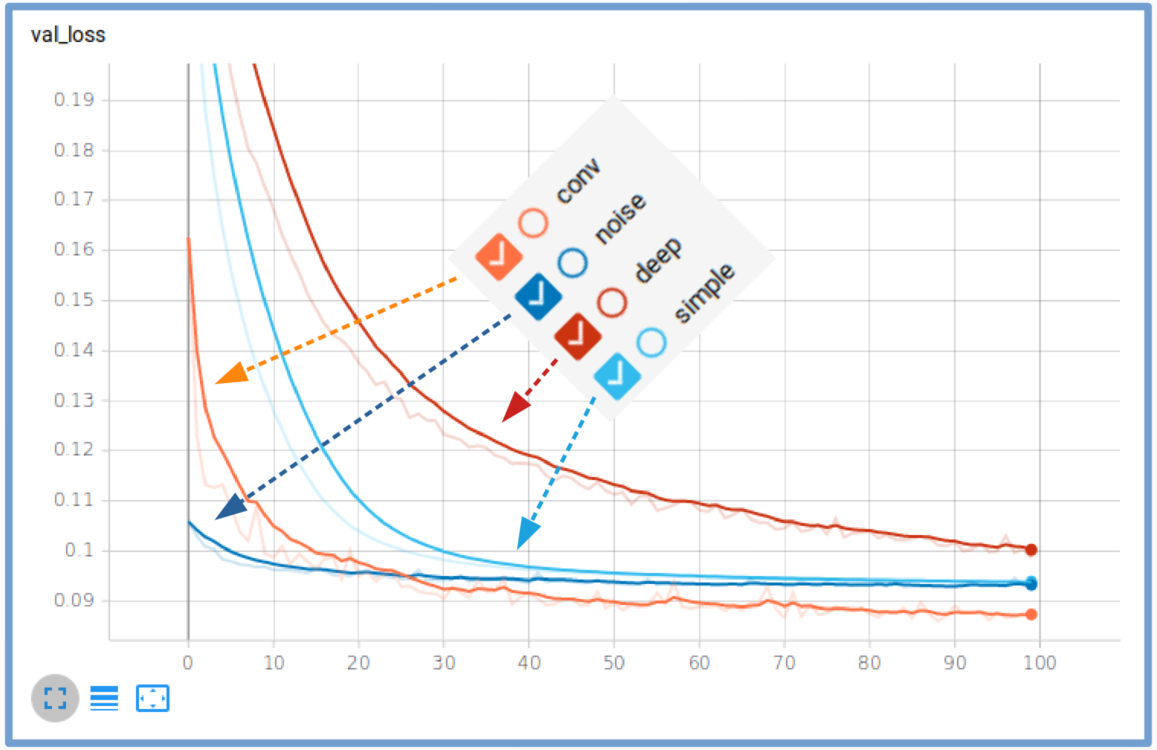
Autoencoders in Action

For AE implementation, it is common to use deep neural nets.

We tested AEs on MNIST data set of handwritten digits.

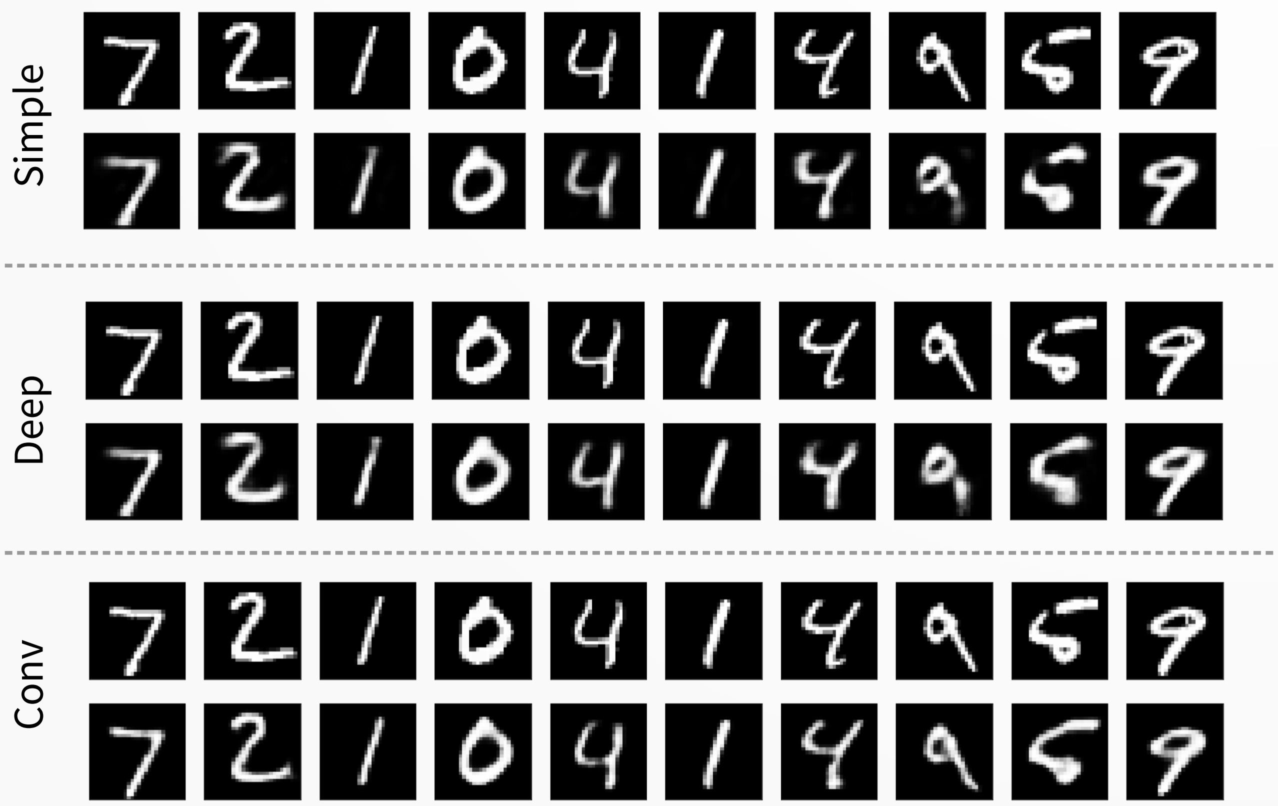
Four AEs were compared for their performance, i.e. a simple multi-layer perceptron, a deep fully-connected network, convolutional neural net, and a convolutional net trained on noisy data.

While training the network, the validation loss was recorded.



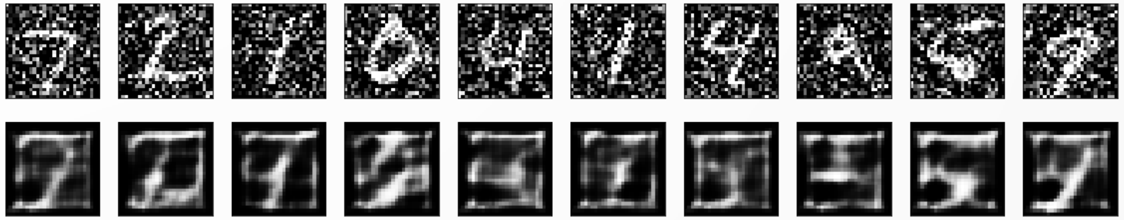
Autoencoders in Action

Autoencoder performance, however, can be best assessed visually. Neural net's depth did not improve its performance. However, knowledge of pixel proximity did.

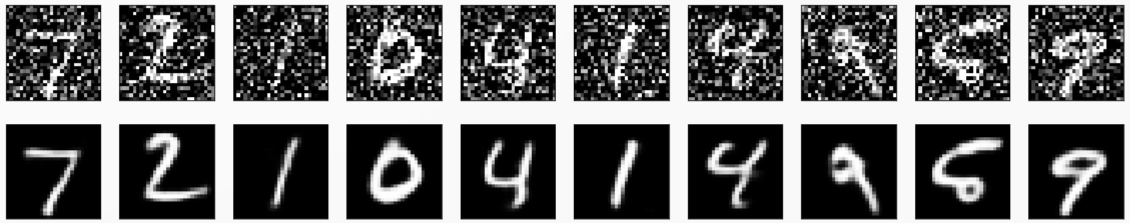


NOISE

Interestingly, even after extensive training CNN was not able to cope with noisy images.



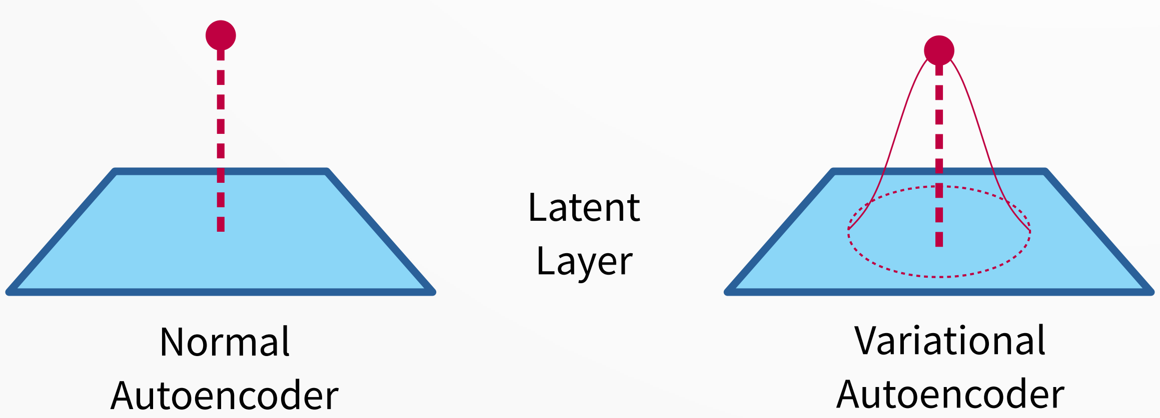
However, extending the number of filters for feature detection and retraining the CNN on noisy data allowed the net to reconstruct the test data from noisy images.

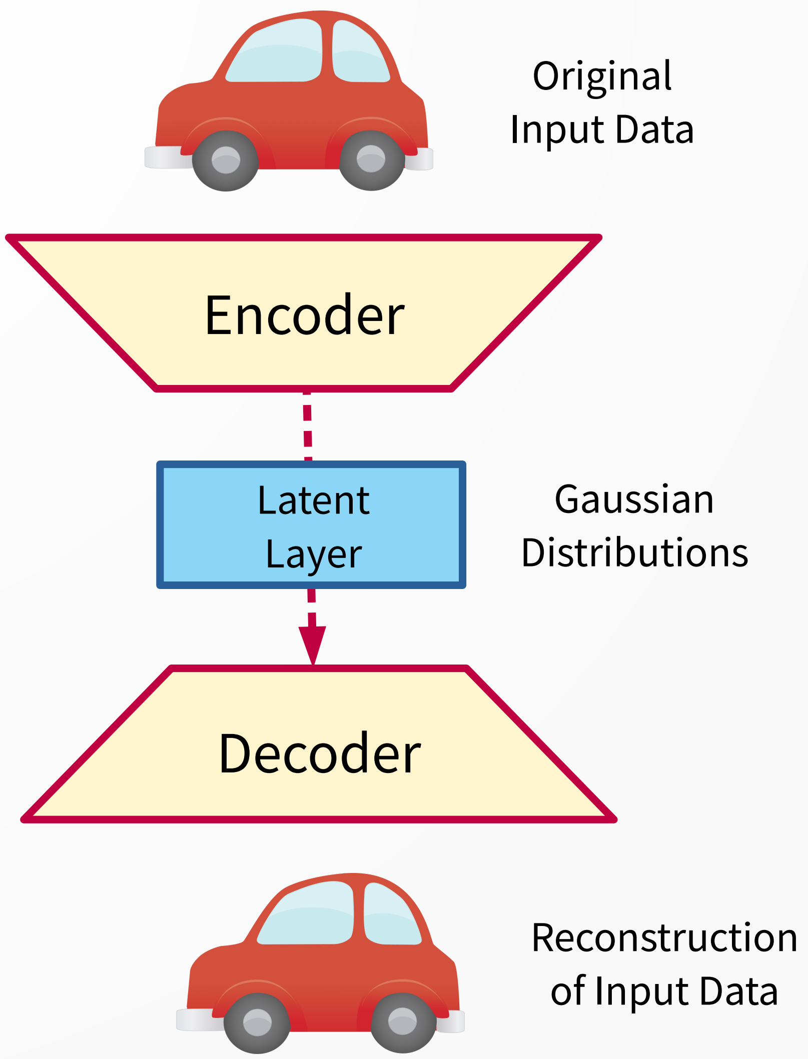


Second… Variational AEs

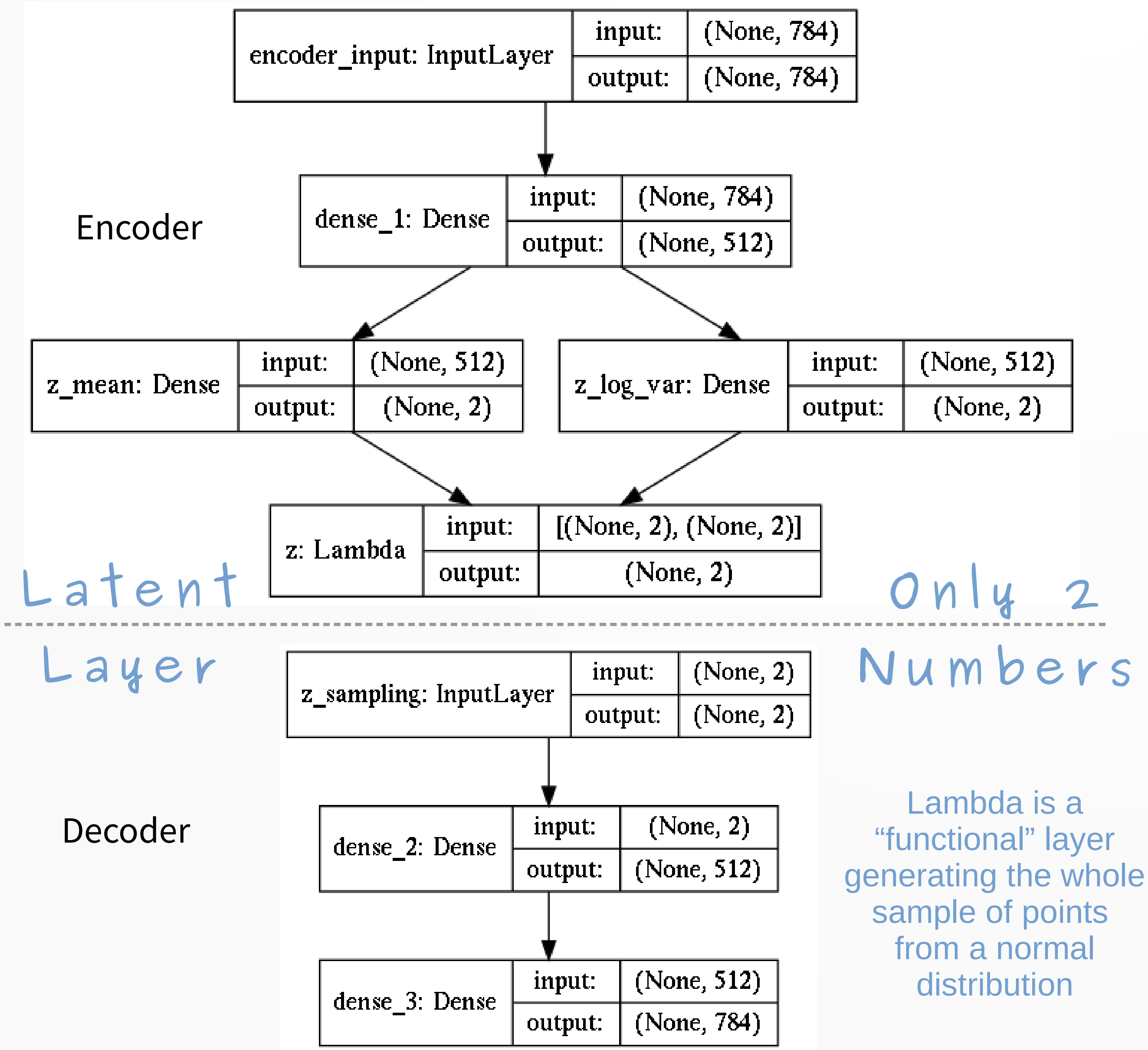
Var AEs borrow an idea from normal AEs that noise in training helps performance.

The main difference between AEs and Var AEs is that instead of mapping AE’s input to a single point in a latent space, it is mapped into a “noisy” Gaussian distribution around a point (mean and log variance).





Variational Autoencoders - Model



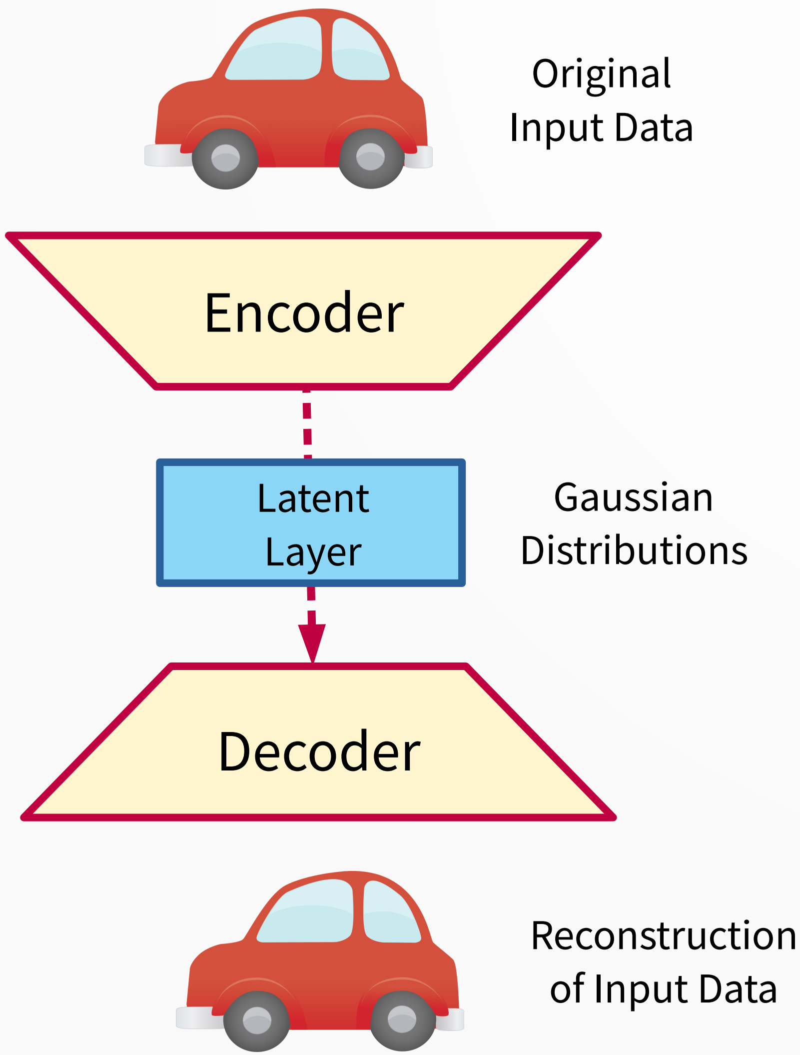
Lambda is a "functional" layer generating the whole sample of points from a normal distribution

Var AEs use a "re-parameterisation trick" by sampling from an "isotropic" unit Gaussian, i.e. it is assumed that all dimensions are independent and variance is the same in all directions.

Encoder's output is a probability distribution, so its error is measured with "Kullback–Leibler divergence", which is an entropy measure of difference between two probability distributions.

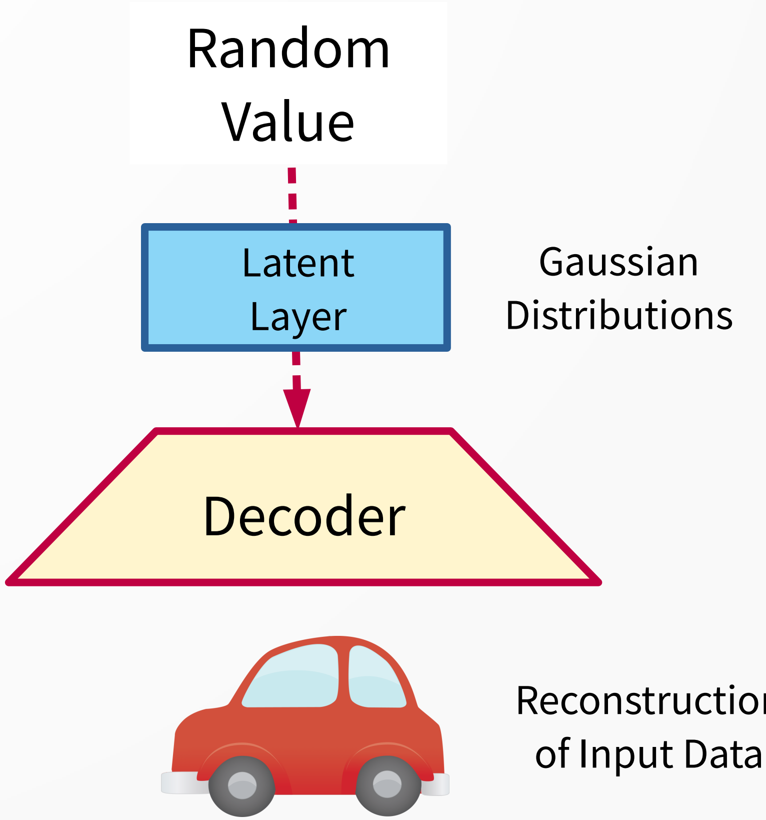
When noise is generated at the latent layer, decoder generates "fake" output.

Variational AEs – What if?

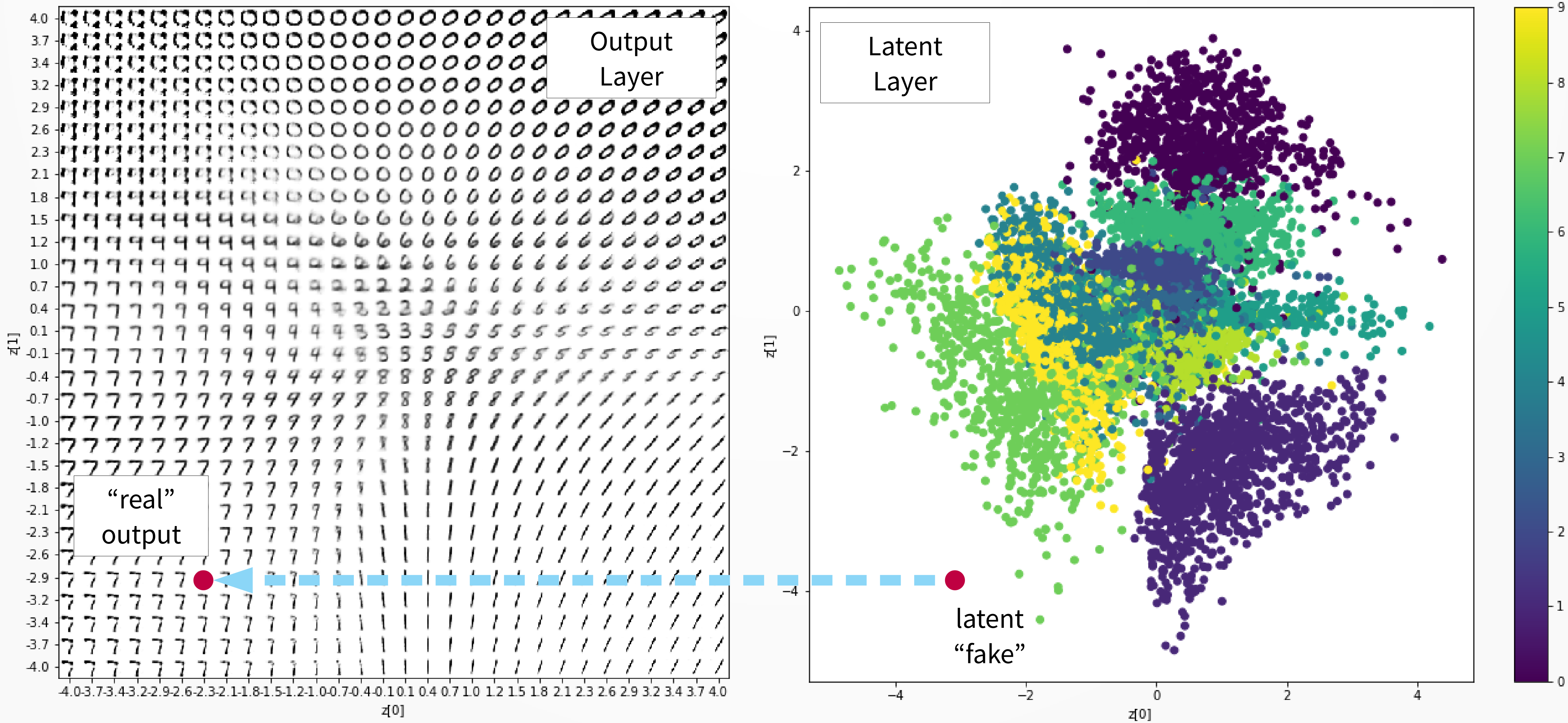


What if after training a Var AE we take away the input layer and its encoder and then enter some random input into a latent layer?

The decoder will generate some output, which can be one of the examples used in training or some “fake” example!



Var AEs – Results – Latent Space



An arbitrary point from the latent space generates an instance of output data – including "fakes"

Finally… GANs

The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection.

Discriminative model is analogous to the police, trying to detect the counterfeit currency.

Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable from the genuine articles.

(Goodfellow et al, 2014)

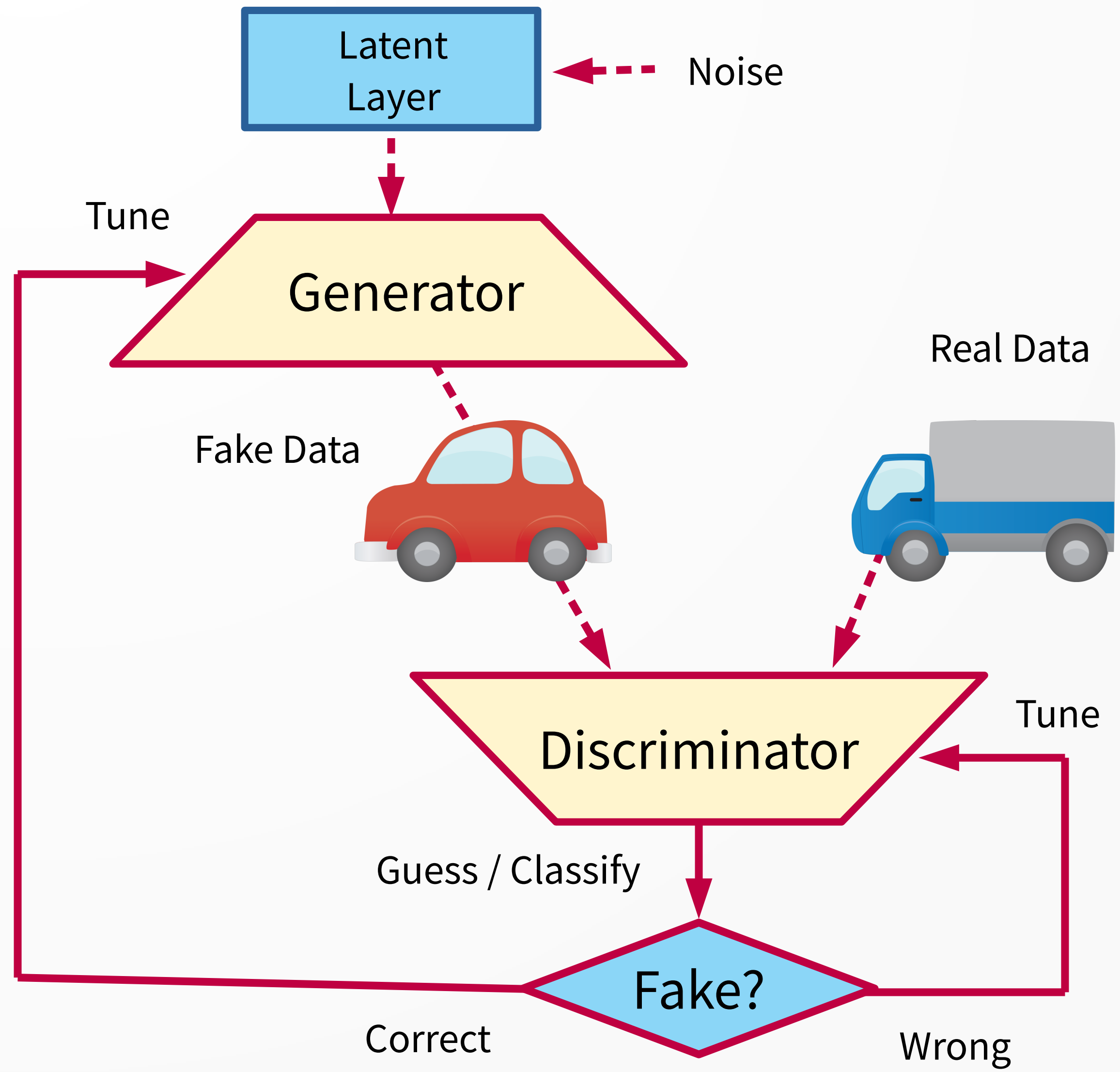
GANs - How They Work ?

We borrow the idea from variational AEs that when noise is generated at the latent layer it creates “fakes”.

The AE architecture is “inverted” so that a generator creates samples from noise produced at latent space and a discriminator is evaluating them.

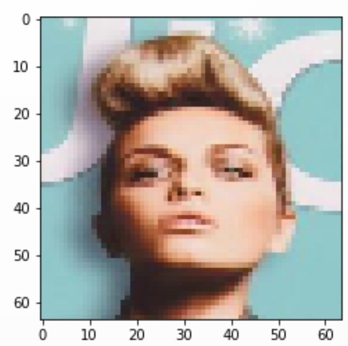
Discriminator learns from real and fake data, however, both networks are trained in this process.

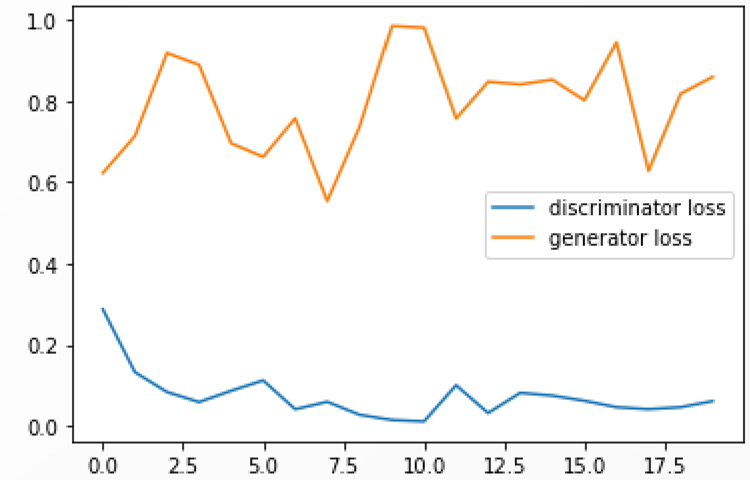
When the discriminator fails to detect a generated fake, only its weights are modified in learning. Otherwise only generator’s weight are modified.



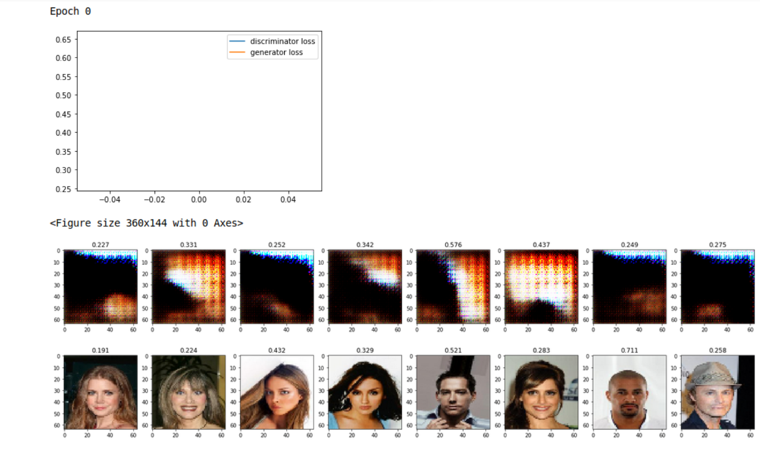
GANs - Example

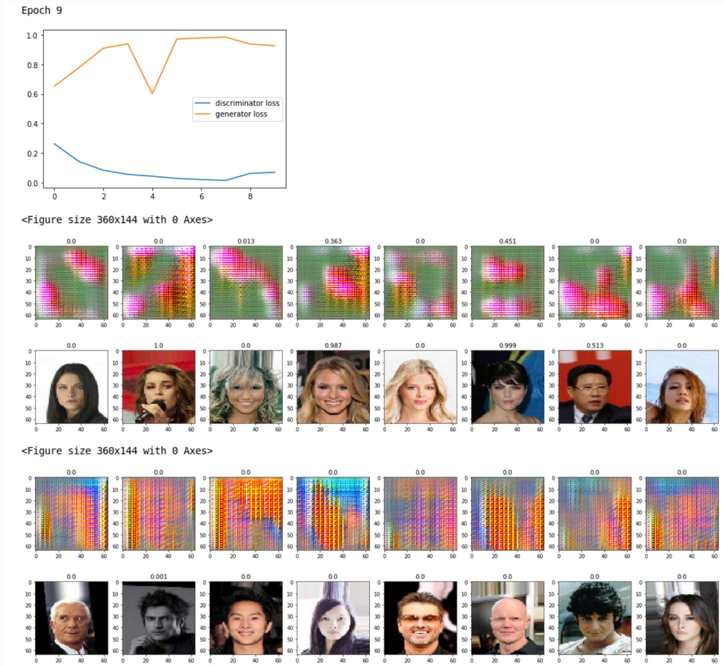
- CelebsA is a data set of over 200K colour photo images
- Only 10K were used here.
- Generator consisted of a Dense layer projecting noise and four fanning out Deconv2D layers responsible for data upsampling
- Discriminator consisted of eight fanning in Conv2D layers and a Dense layer responsible for binomial detection of fakes

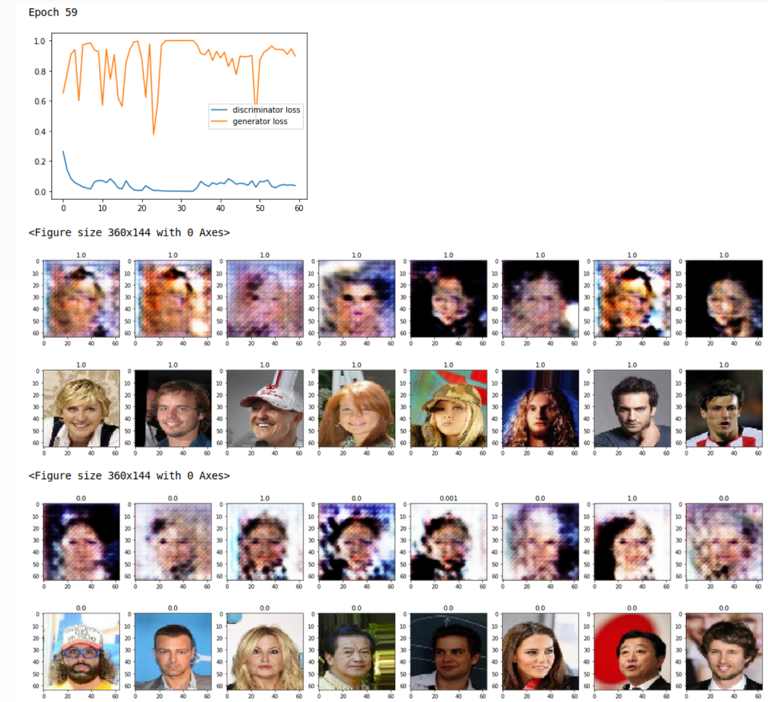




During the cycle of training and classification the model performance was monitored, as well as sample of generated and real images were produced

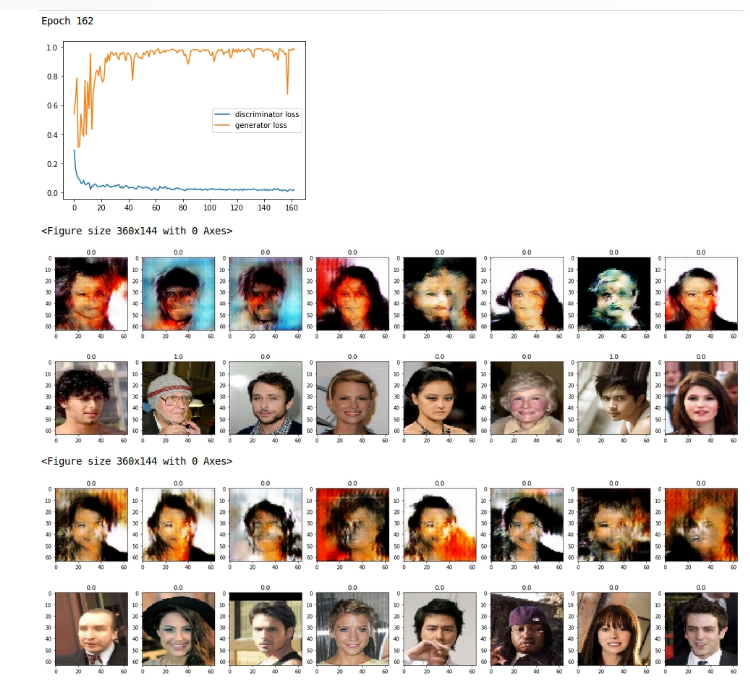


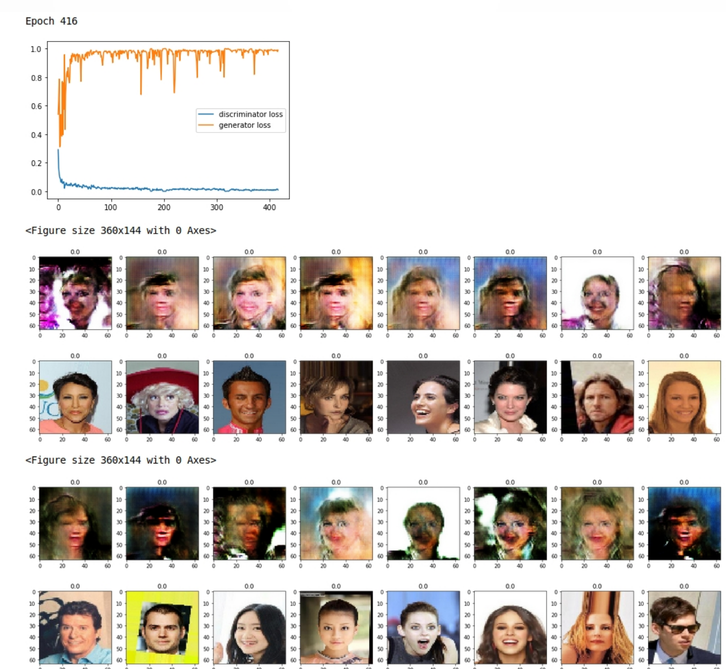


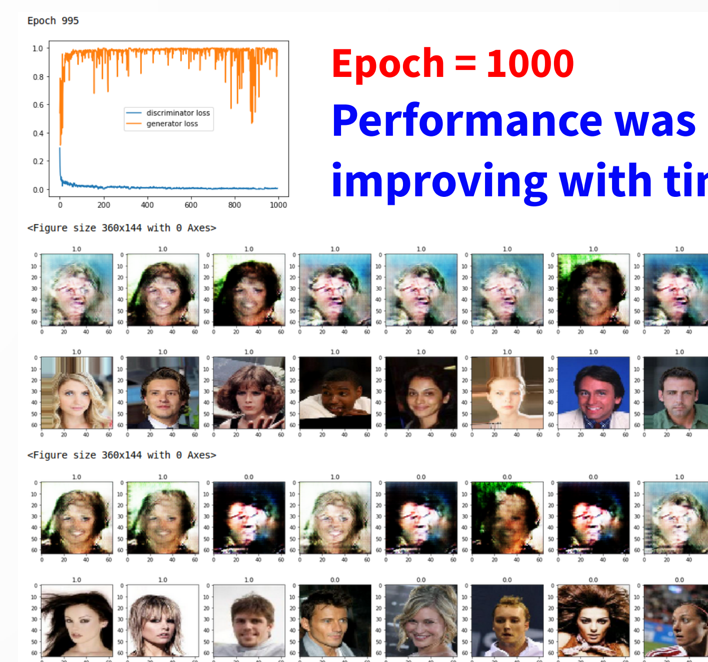


Epoch = 0

The initial performance of both the generator and discriminator was very poor...



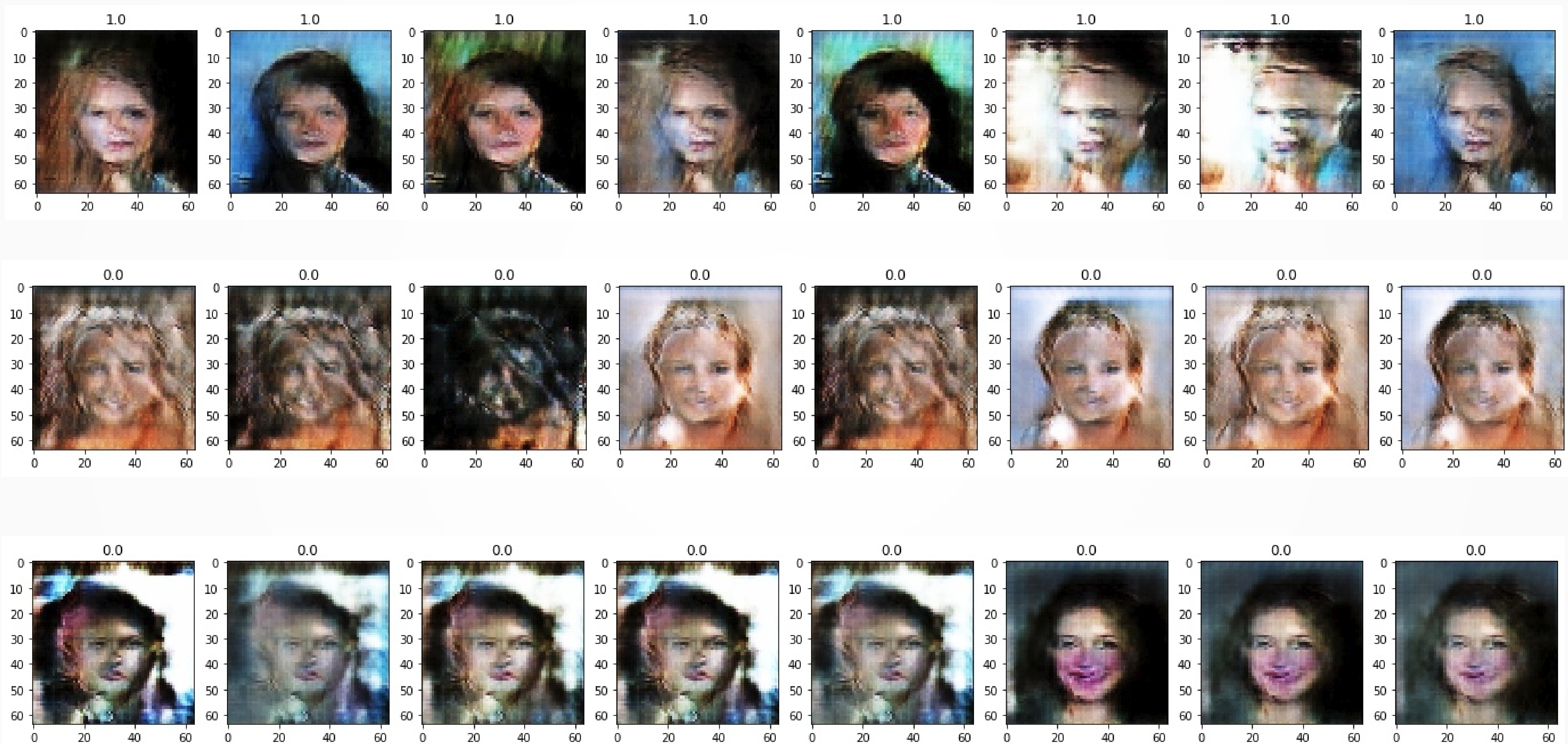




Epoch = 1000

Performance was improving with time!

How good can we get?



Conclusions

- Generative deep learning is one of the most exciting areas of research in Machine Learning
- At the moment most of the applications are limited to vision, sound and text processing
- Business applications are almost non-existent!
- Can we think of those business areas which could benefit from this technology?
- Questions? Comments? Discussion?

References and Discussions

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. “Generative Adversarial Nets.” In *Advances in Neural Information Processing Systems*, 2672–2680, 2014.

Also see Quora discussions:

- What are the typical businesses of generative adversarial networks (GANs)?
- What are some exciting future applications of Generative Adversarial Networks?
- What are the (existing or future) use cases where using Generative Adversarial Network is particularly interesting?
- What industry/sector would you have particular interest in seeing application of generative adversarial networks?

Some GAN ideas:

- Gaming industry, e.g. generation of textures
- Creating model of the world to use for reinforcement learning / motion planning
- Data augmentation for imbalanced classification
- Anomaly detection
- Semi-supervised learning, where only a small subset of data is labelled
- Creating Infographics from text, creating animations for rapid development of marketing content, generating website designs are to name a few
- Scraping data from the web without being detected
- Fashion design and adjustments
- Generation of medical images