β VARIATIONAL AUTOENCODERS

Weronika Michoń, Krzysztof Bazner

19 November 2019

What is Autocoder?

Autocoders are a family of neural network models aiming to learn compressed latent variables of high-dimensional data.

We will talk about ...

- What is an autoencoder?
- What is the latent space and why regularising it?
- How to generate new data from VAEs?
- What is the link between VAEs and variational inference?
- What is an β VAE?

What is dimensionality reduction?

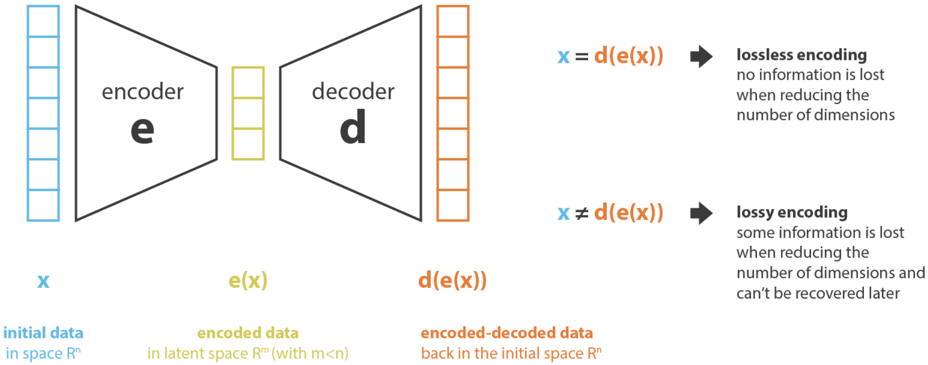


Figure 1: Illustration of the dimensionality reduction principle with encoder and decoder.

Autoencoder

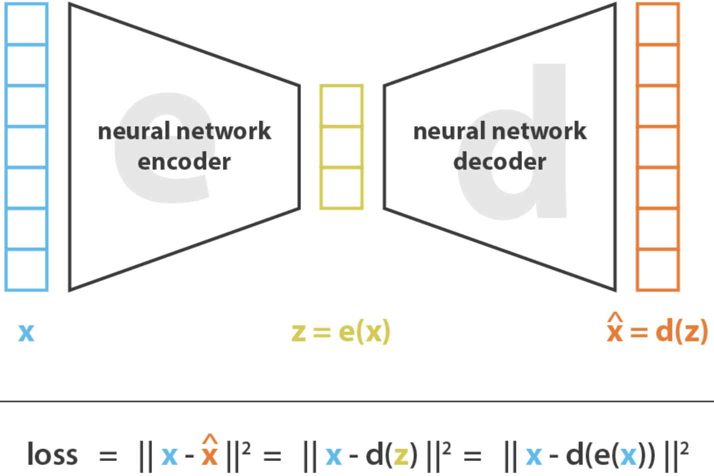
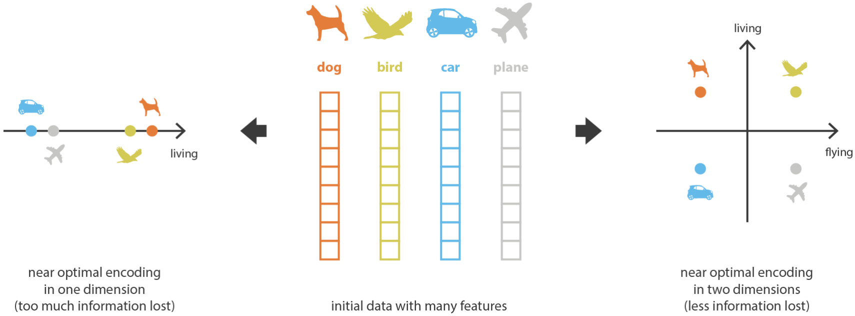


Figure 2: Illustration of an autoencoder with its loss function.

What is worth remembering?(1/3)

1. Dimensionality reduction with no reconstruction loss often comes with a price: the lack of interpretable and exploitable structures in the latent space (lack of regularity).
2. The dimension of the latent space and the “depth” of autoencoders have to be carefully controlled and adjusted depending on the final purpose of the dimensionality reduction.

What is worth remembering?(2/3)



What is worth remembering?(3/3)

Due to overfitting, the latent space of an autoencoder can be extremely irregular (close points in latent space can give very different decoded data, some point of the latent space can give meaningless content once decoded, ...) and, so, we can't really define a generative process that simply consists to sample a point from the latent space and make it go through the decoder to get a new data.

Variational Autoencoder(1/3)

What is it?

Variational Autoencoders (VAEs) are autoencoders that tackle the problem of the latent space irregularity by making the encoder return a distribution over the latent space instead of a single point and by adding in the loss function a regularisation term over that returned distribution in order to ensure a better organisation of the latent space

Variational Autoencoder(2/3)

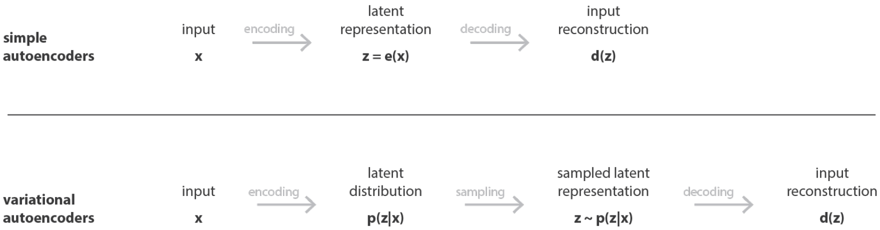
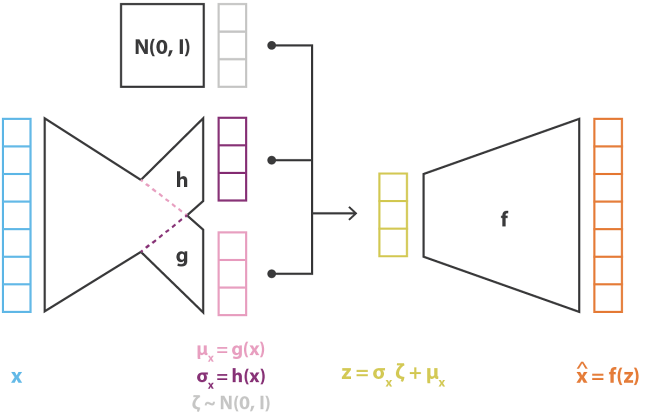


Figure 3: Difference between autoencoder and variational autoencoder

Variational Autoencoder(3/3)



$$\text{loss} = C\,\|x - \hat{x}\|^2 - KL[N(\mu_x, \sigma_x), N(0, I)] = C\,\|x - f(z)\|^2 - KL[N(g(x), h(x)), N(0, I)]$$

Figure 4: Variational Autoencoders representation.

β Variational Autoencoder(1/5)

β VAE is a modification of Variational Autoencoder with a special emphasis to discover disentangled latent factors. Following the same incentive in VAE, we want to maximize the probability of generating real data, while keeping the distance between the real and estimated posterior distributions small.

$$\mathbf{loss}_{\mathbf{VAE}} = C \, || \, x - \hat{x} \, ||^2 - \mathbf{KL}[\, N(\mu_x, \sigma_x), N(0, I) \,] = C \, || \, x - f(z) \, ||^2 - \mathbf{KL}[\, N(g(x), h(x)), N(0, I) \,]$$

$$\mathbf{loss}_{\beta\,\mathbf{VAE}} = C \, || \, x - \hat{x} \, ||^2 - \boldsymbol{\beta}\, \mathbf{KL}[\, N(\mu_x, \sigma_x), N(0, I) \,] = C \, || \, x - f(z) \, ||^2 - \boldsymbol{\beta}\, \mathbf{KL}[\, N(g(x), h(x)), N(0, I) \,]$$

Figure 5: Loss function in VAE and β VAE.

Hyperparameter β balances latent channel capacity and independence constraints with reconstruction accuracy. β VAE with appropriately tuned $\beta > 1$ qualitatively outperforms VAE ($\beta = 1$).

β Variational Autoencoder(2/5)

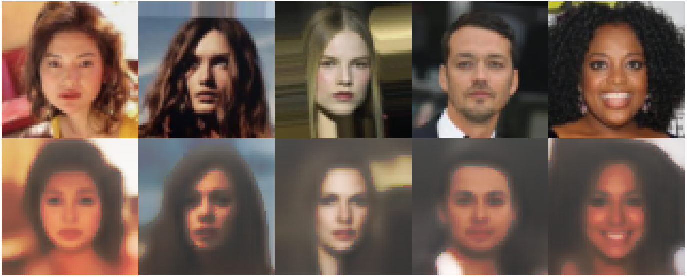


Figure 6: Original Faces (Top) vs. Reconstructed (Bottom).

β Variational Autoencoder(3/5)

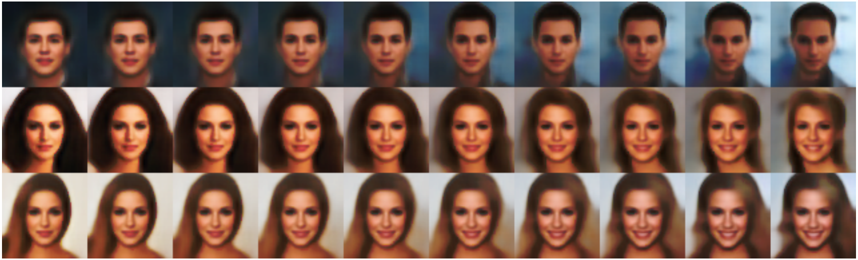


Figure 7: Linear Interpolation from z1 (leftmost) to z2 (rightmost).

β Variational Autoencoder(4/5)

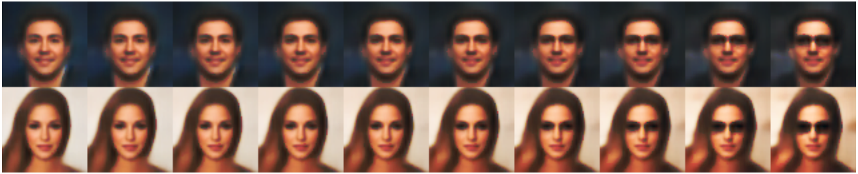


Figure 8: Vector Arithmetic from original (leftmost) to wearing sunglasses (rightmost).

β Variational Autoencoder(5/5)

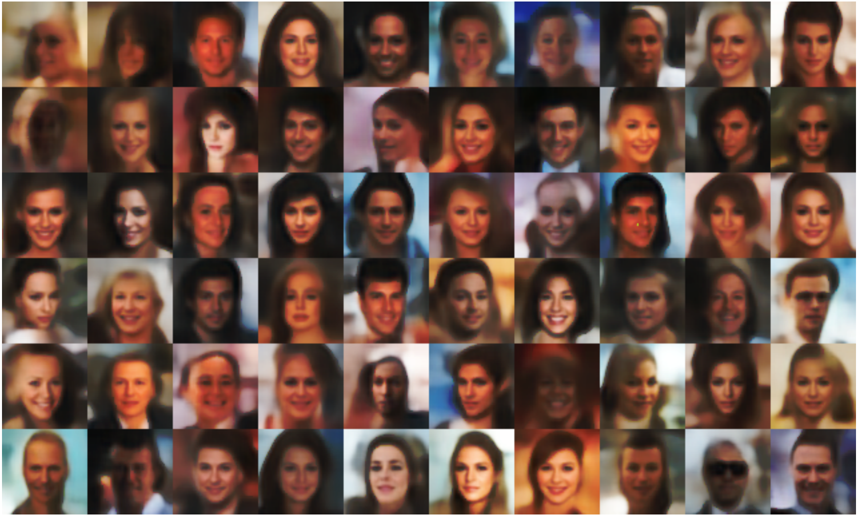
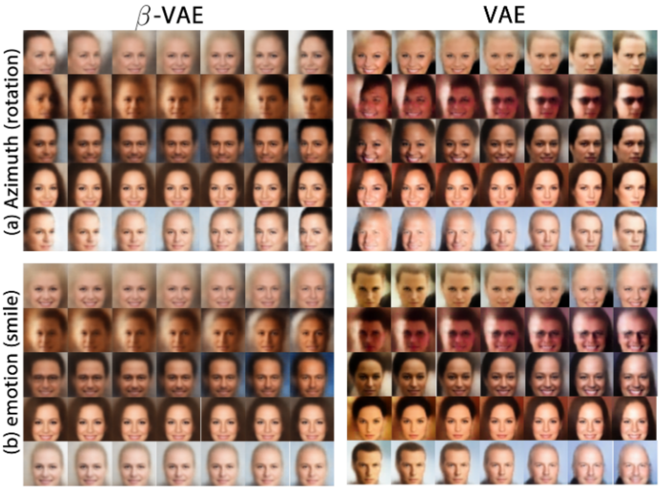


Figure 9: Generated Images with randomly sampled latent z ∼ N(0,1).

Variational Autoencoder vs β Variational Autoencoder



At VAE, we get sharper images, but β VAE does better with disentangled latent factors.