TNM112 – Deep Learning for Media Technology

Nithesh Chandher Karthikeyan

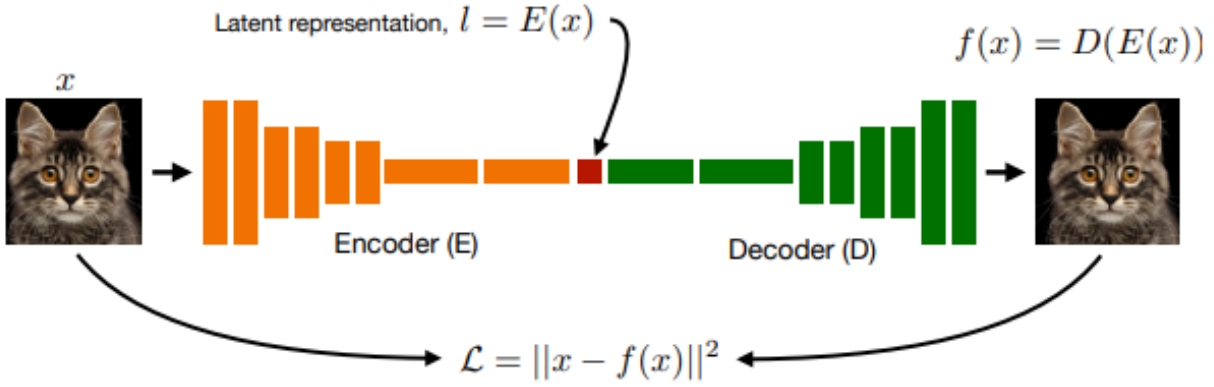
LINKÖPING UNIVERSITY

Dataset

- Select dataset based on the task
- Choose smaller datasets (as it's easy to train)
- Image datasets:
  - MNIST
  - FashionMNIST
  - CIFAR-10
  - Kaggle's CatvsDog
  - STL-10
  - CelebA Faces dataset

Proposal 1 – The Autoencoder

- Autoencoder as two components:
  - Encoder: Image to Latent representation
  - Decoder: Latent representation to Image



Proposal 1 – The Autoencoder

- Encoder:
  - Down-sampling
  - Convolution()
  - Maxpooling()
  - Flatten()
  - Dense()
- Decoder:
  - Up-sampling
  - Reshape()
  - Transpose Conv()
  - Upsampling()

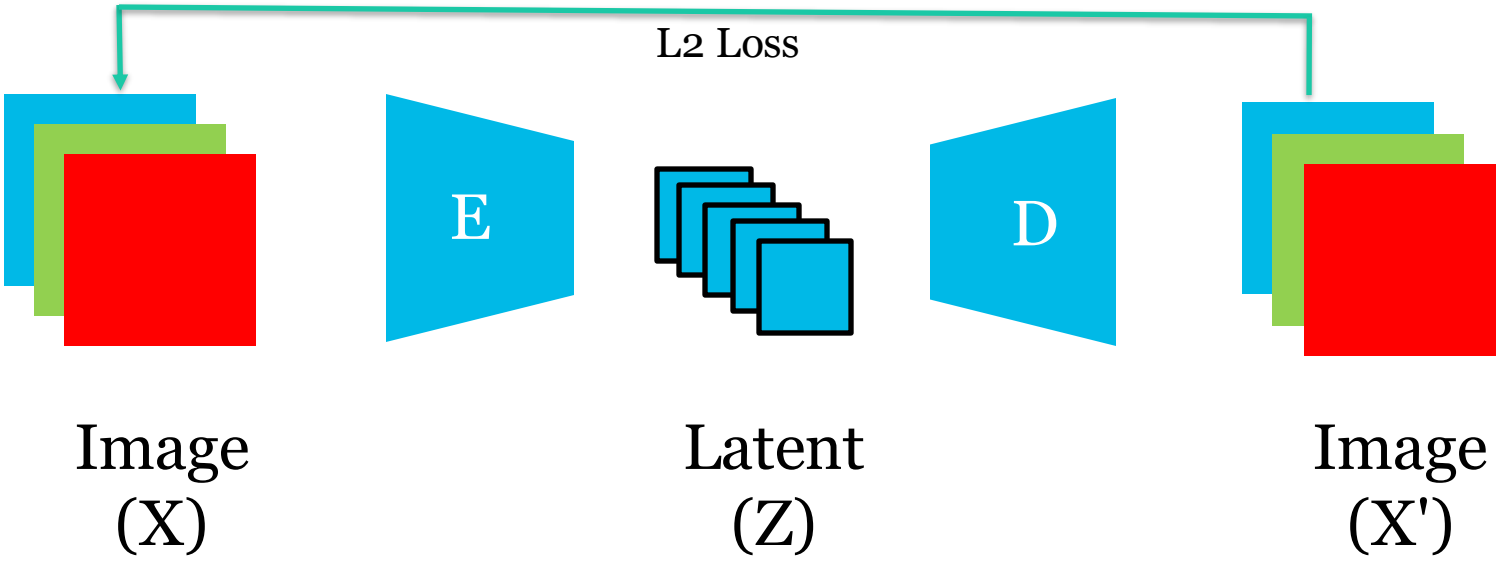
Proposal 1 – The Autoencoder

Tasks:

- Study design choices
- Classification using AE (Latents)
- Pre-training using AE (Encoder+FC+Softmax)
- Visualize Latents

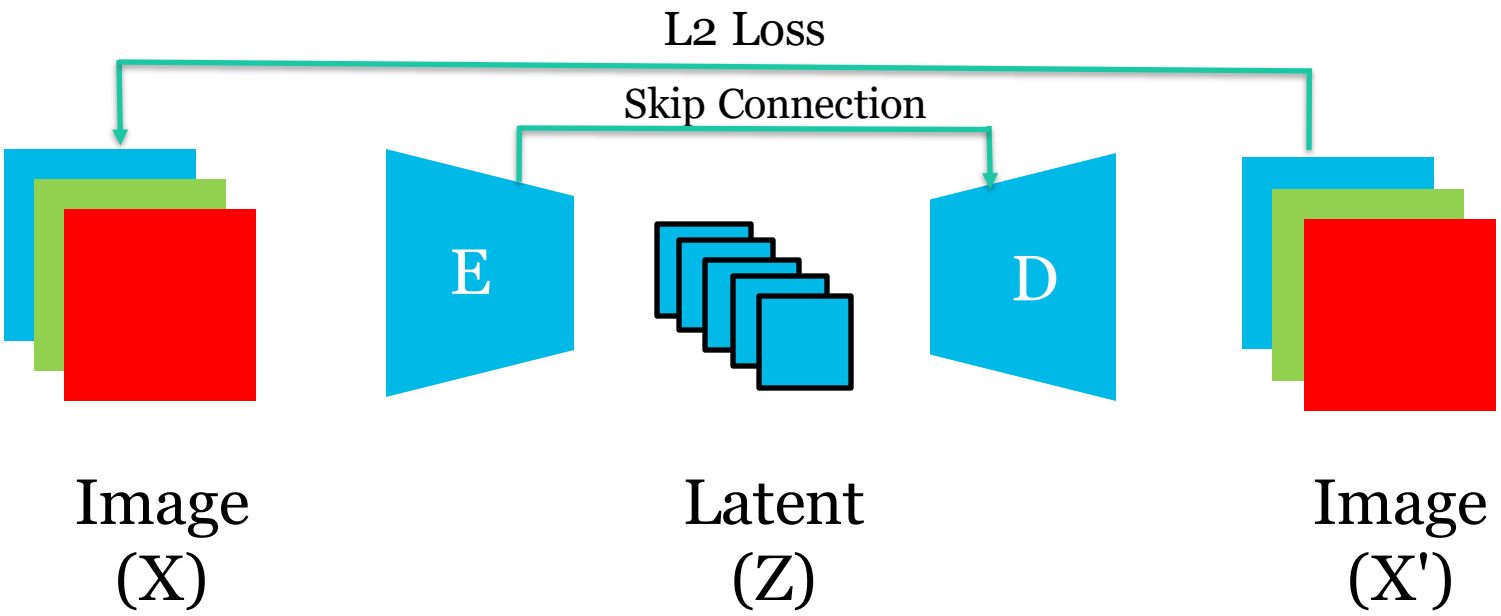
Proposal 2 – Image to Image CNNs

- Fully Convolutional AEs



Proposal 2 – Image to Image CNNs

- Fully Convolutional AEs



Proposal 2 – Image to Image CNNs

Tasks:

- Image Denoising
- Colorization
- Segmentation (Image -> Segmentation Mask)
- Inpainting

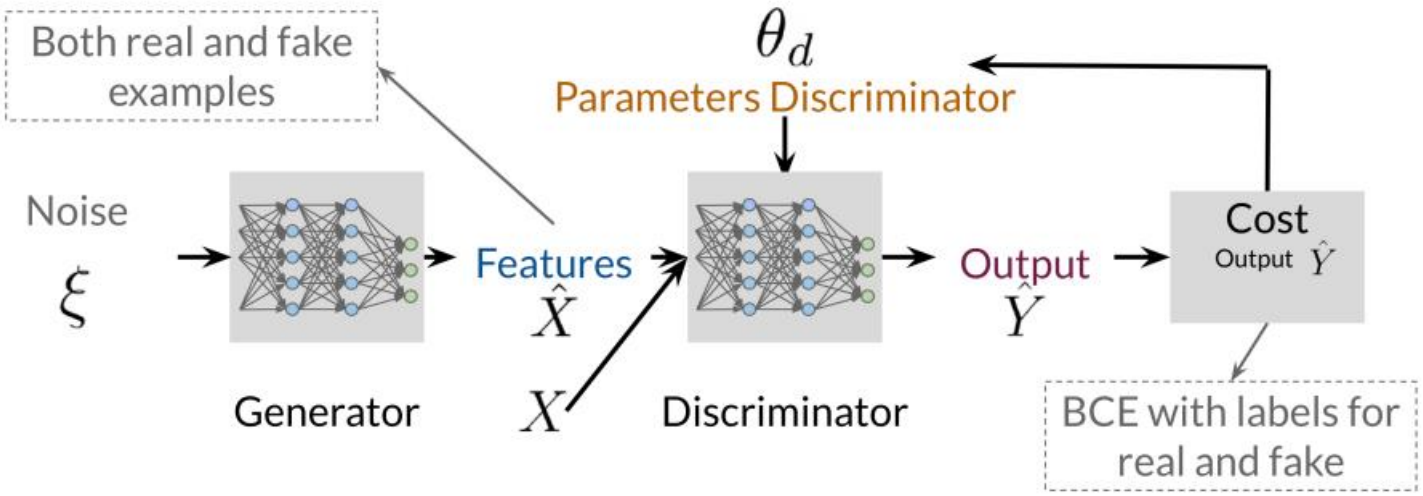
Proposal 3 – Generative Deep Learning

Generative Adversarial Networks:

- GANs learn training data distribution and generates new data
- Two Components:
  - Generator: Takes a random noise vector and generates fake images
  - Discriminator: Classifier that distinguishes real from fake images
- Loss Function: Binary Cross Entropy

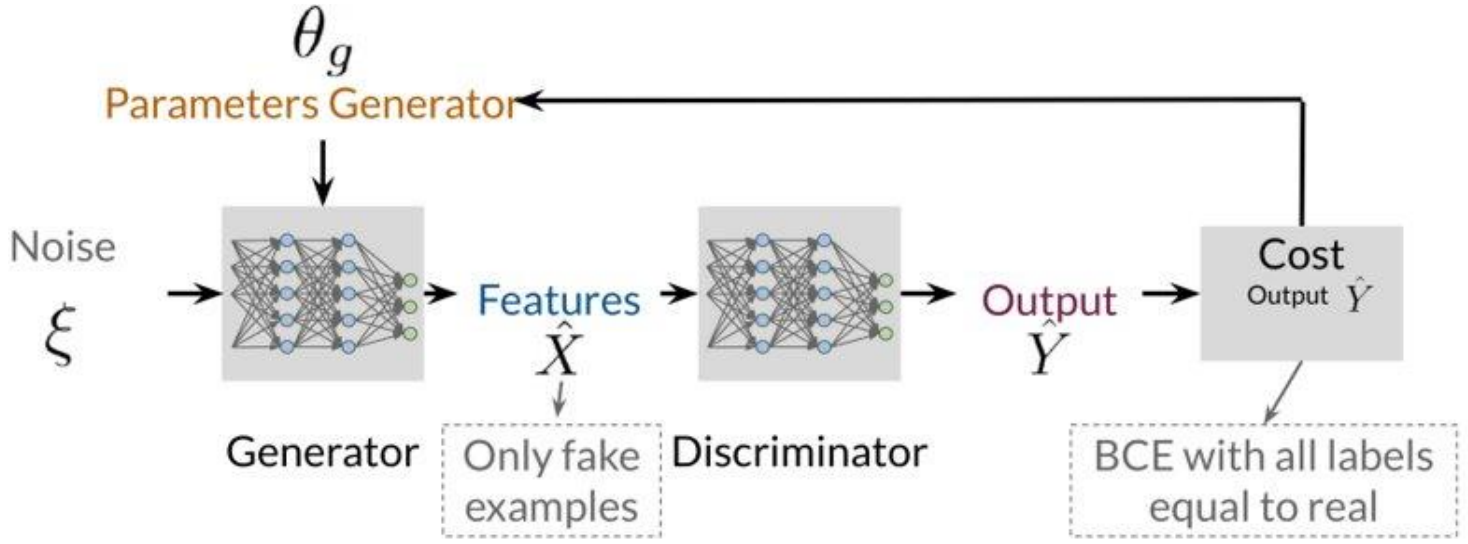
$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right) \right]$$

Proposal 3 – Generative Deep Learning



Training a Discriminator

Proposal 3 – Generative Deep Learning



Training a Generator

Proposal 3 – Generative Deep Learning

Different GAN architectures:

- DC-GAN
- CycleGAN
- Conditional GAN
- WGAN
- StyleGAN

Keras Implementation

Feedback

Kindly send us your feedback