


浙江大学爱丁堡大学联合学院

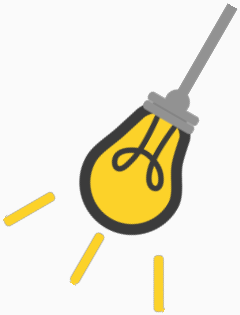
ZJU-UoE Institute

Lecture 18 - Autoencoders

Nicola Romanò - nicola.romano@ed.ac.uk

Learning objectives

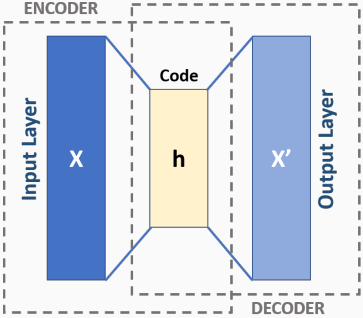
- Discuss autoencoders and their applications.
- Code a denoising autoencoder using Keras.



Introduction

What is an autoencoder?

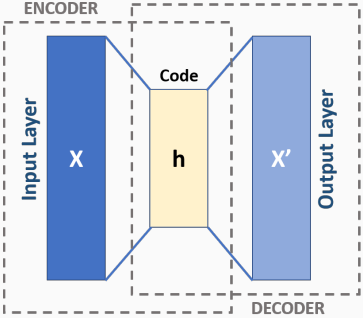
- An **autoencoder** is a neural network that is trained to reconstruct its input.
- It consists of an **encoding** part and a **decoding** part.



Source: Wikipedia

What is an autoencoder?

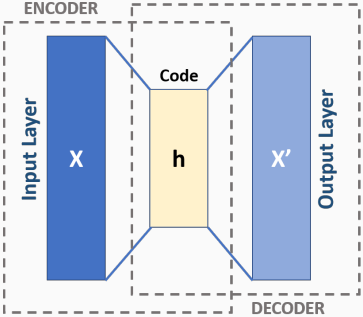
- An **autoencoder** is a neural network that is trained to reconstruct its input.
- It consists of an **encoding** part and a **decoding** part.
- The easiest mapping of an input to itself would be the identity function.
- Autoencoders force the input to be compressed, and then decompressed to reconstruct it.



Source: Wikipedia

What is an autoencoder?

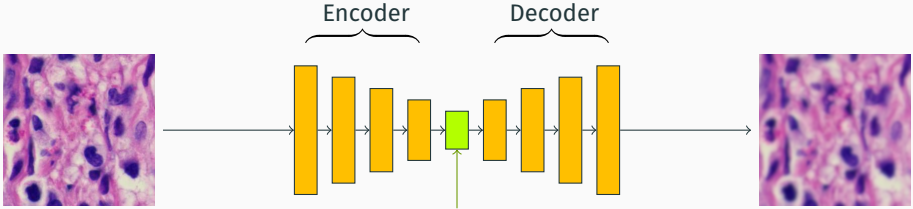
- An **autoencoder** is a neural network that is trained to reconstruct its input.
- It consists of an **encoding** part and a **decoding** part.
- The easiest mapping of an input to itself would be the identity function.
- Autoencoders force the input to be compressed, and then decompressed to reconstruct it.
- It can only reconstruct an approximated version of the input.



Source: Wikipedia

Convolutional autoencoders

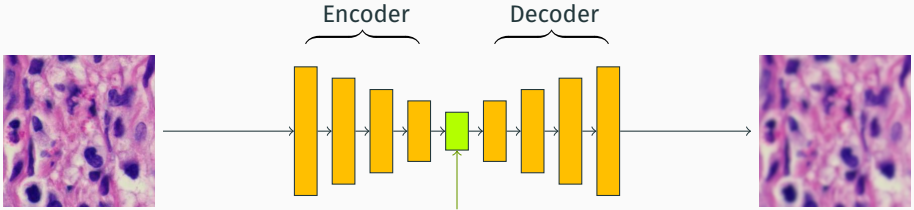
A **convolutional autoencoder** is a convolutional network that is trained to reconstruct its input.



$$X \simeq X'$$

Convolutional autoencoders

A **convolutional autoencoder** is a convolutional network that is trained to reconstruct its input.



$$X \simeq X'$$

While ideally the input is identical to the output, in practice we get a slightly degraded version.

Why would we do this?

Uses of autoencoders

Autoencoders have been used in a variety of applications, including image analysis.

Examples of uses in computer vision include

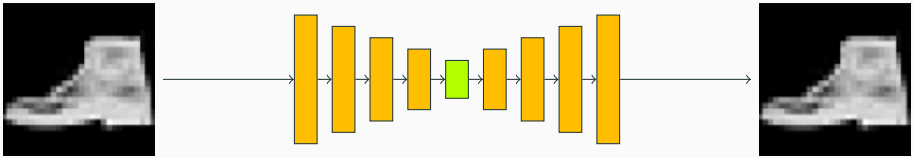
- Noise reduction
- Image colorization
- Image compression
- Image restoration
- Superresolution

Outside of computer vision, autoencoders are used in a variety of applications, including fraud detection, recommendation systems, and text translation.

Building an autoencoder in Keras

We will now build a simple autoencoder that just learns to reconstruct its input (trained on Fashion MNIST).

See `basic_autoencoder.ipynb`



Uses of autoencoders

Uses of autoencoders

Compression

Autoencoders as compressors

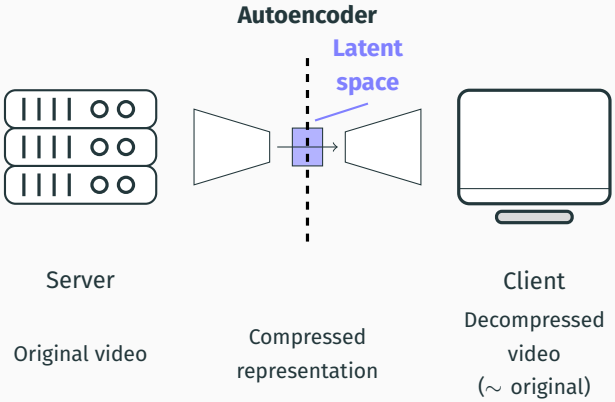
- The **latent space** is a low-dimensional representation of the input.

Autoencoders as compressors

- The **latent space** is a low-dimensional representation of the input.
- We can use it to compress the input.

Autoencoders as compressors

- The **latent space** is a low-dimensional representation of the input.
- We can use it to compress the input.
- Useful e.g. to reduce the size of images/videos for display purposes (e.g. streaming a video).



Autoencoders for feature extraction

We can think of the latent space as an *optimal* low-dimensional representation of the input, therefore we can use it to generate features from the input.

Autoencoders for feature extraction

We can think of the latent space as an *optimal* low-dimensional representation of the input, therefore we can use it to generate features from the input.

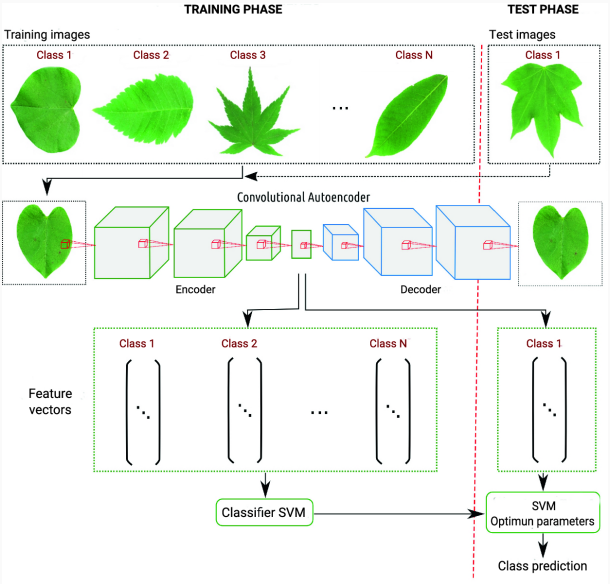
Example:

A Feature Extraction Method Based on Convolutional Autoencoder for Plant Leaves Classification

Authors | Authors and affiliations

Mery M. Paco Ramos, Vanessa M. Paco Ramos, Arnold Loaiza Fabian, Erbert F. Osco Mamani

Classification of leaves using autoencoders

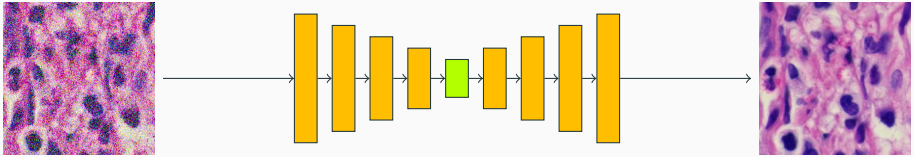


Uses of autoencoders

Denoising

Autoencoders for denoising

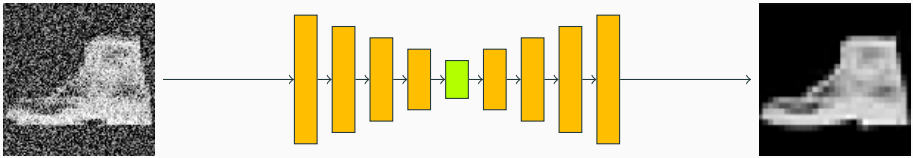
- Noisy image = original image + noise
- Since the autoencoder learns only the *important* features of the image, it will not learn random noise.
- We can *trick* the autoencoder by feeding it pairs of noisy/clean images as source and target, and it will learn to denoise the images!



Denoising using an autoencoder

Let's update our code to make a denoising autoencoder!

See `denoising_autoencoder.ipynb`

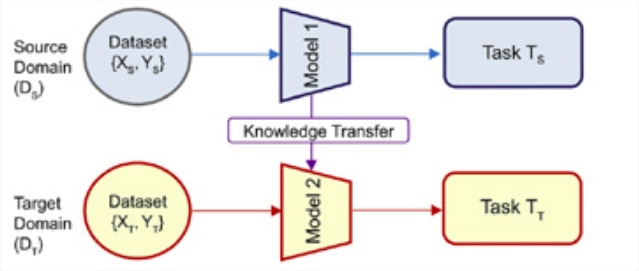


Uses of autoencoders

Domain adaptation

Domain adaptation

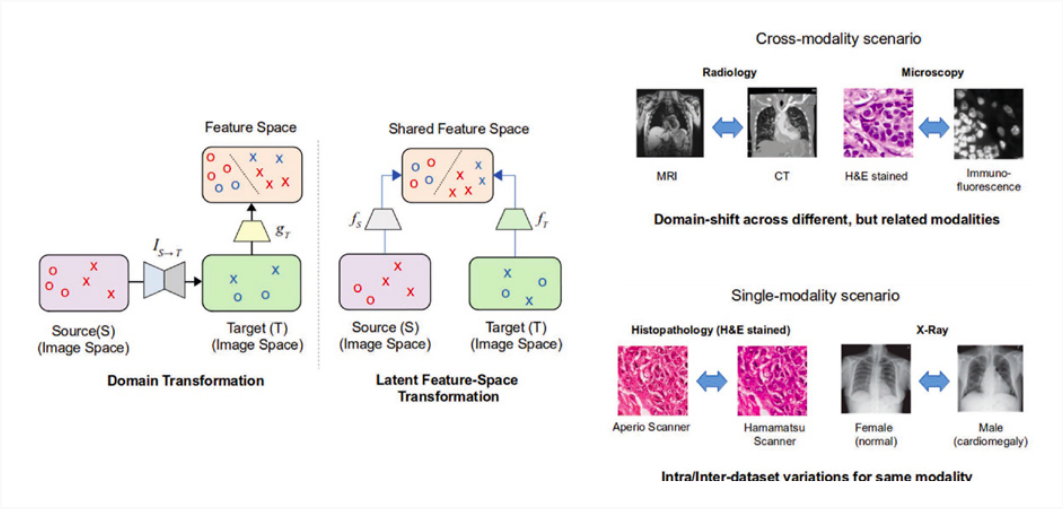
Domain adaptation is an approach to transfer knowledge from one domain to another.



Choudary et al. 2020.

Domain adaptation

Domain adaptation is an approach to transfer knowledge from one domain to another.



Choudary et al. 2020.