

Deep Learning

Aprendizaje Automático para la Robótica
Máster Universitario en Ingeniería Industrial

Departamento de Automática

Objectives

1. Motivate Deep Learning
2. Introduce main deep architectures
3. Describe state-of-the-art applications

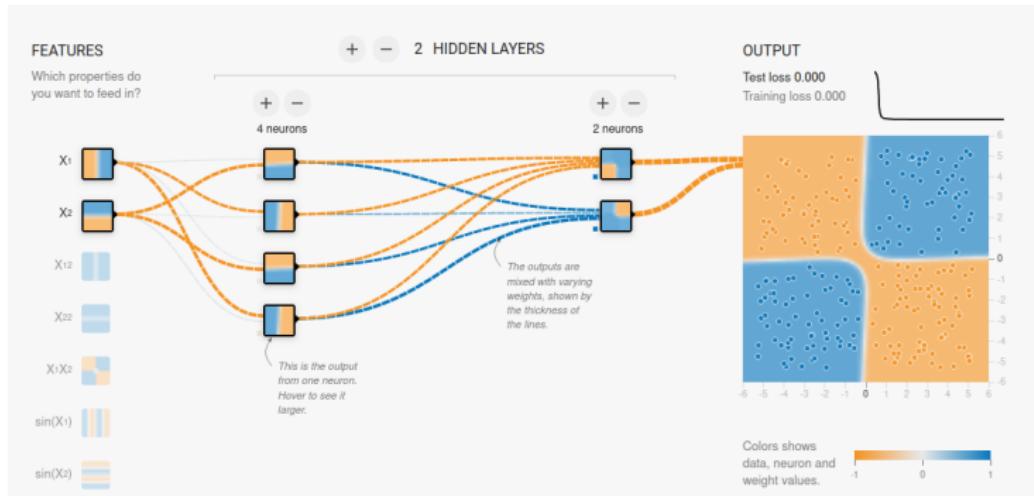
Bibliography

- Géron, Aurélien *Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow*. 2nd edition. O'Reilly. 2019

Table of Contents

- 1. Motivation
- 2. Deep Learning
- 3. Convolutional Neural Networks
 - Biological motivation
 - Convolutional layer
 - Max-pooling layer
 - Dropout layer
 - CNN architectures
 - Other CNN architectures
 - Classification and localization
 - Object detection
 - Object detection: You Only Look Once (YOLO)
 - Semantic segmentation
- 4. Recurrent networks
 - RNNs
 - LTSM networks
 - GRU networks
 - Advanced RNNs
 - RNNs cool applications
- 5. Autoencoders
 - Autoencoders
 - Autoencoders for anomaly detection
 - Autoencoders as generative models
- 6. Advanced topics
 - GAN
 - Advanced topics
 - VAE
 - Generative models state-of-the-art

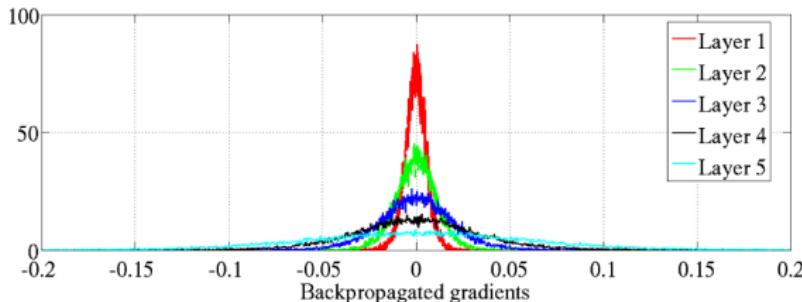
Motivation



Deep Learning (I)

Deep Learning is not just a network with many layers

- High number of parameters to optimize
- More layers \Rightarrow more local optima \Rightarrow more difficult training
- Large datasets
- Gradient vanishing / exploding



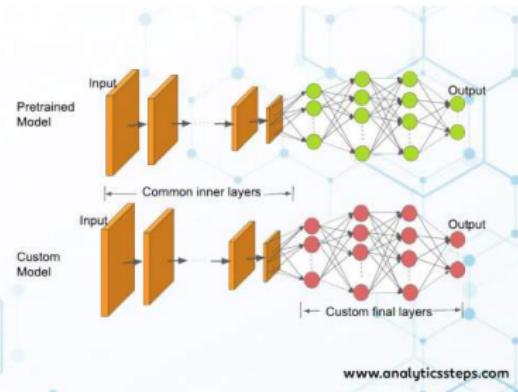
Deep Learning (II)

Need of tricks to train deep networks

- Careful weights initialization
 - Glorot and He initialization, LeCun initialization
- Non-saturating activation functions
 - Sigmoid activation is problematic
 - ReLU and Leaky ReLU activation functions
- Batch normalization
- Transfer learning/ pretrained models

#analytic Steps

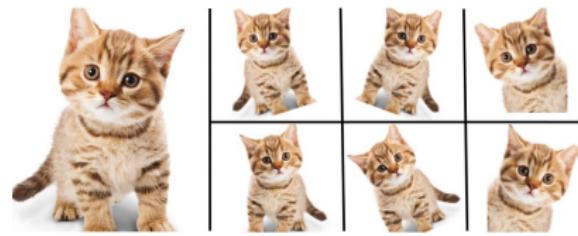
Transfer Learning



(Source)

Deep Learning (III)

- Unsupervised pretraining
- Regularization
 - L₁, L₂ and dropout
- Faster optimizers
 - AdaGrad, Adam, RMSProp
- Data augmentation
 - Creates modified copies of the dataset
 - It increases the number of samples
- Clever design of the network to minimize parameters



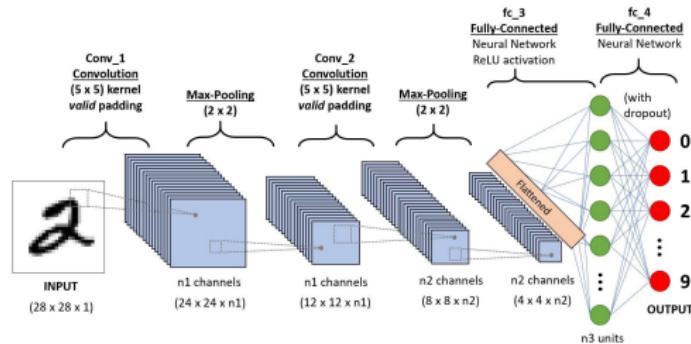
Enlarge your Dataset

(Source)

Deep Learning (IV)

In Deep Learning, we use to think in layers

- Data input layers
- Output layers
- Fully connected (classic)
- Convolutional layers
- Max-pooling
- Recurrent layers
- Dropout layers
- More ...

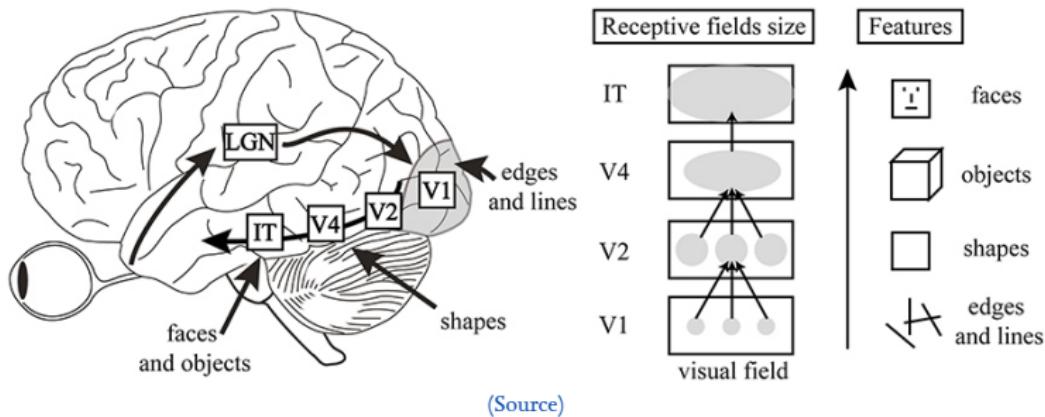


Two popular types of deep architectures

- Convolutional Neural Networks (CNNs) - Image processing
- Long Short-Term Memory (LSTM) - Time-series and NLP

Convolutional Neural Networks

Biological motivation



Convolutional Neural Networks

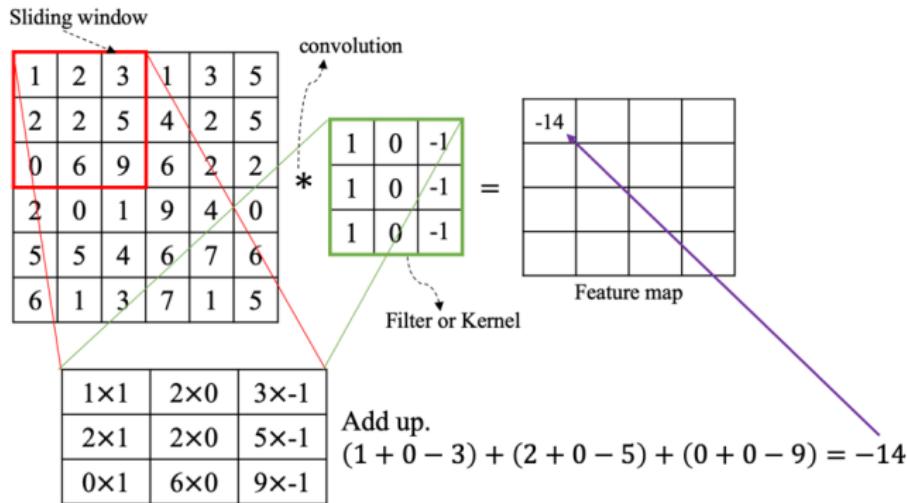
Convolutional layers (I)

CNNs are popular for Computer Vision applications

- Networks with convolutional layers
 - Convolutions are features extractors
 - Its behaviour can be learnt

Convolutional Neural Networks

Convolutional layers (II)



Created by brilliantcode.net

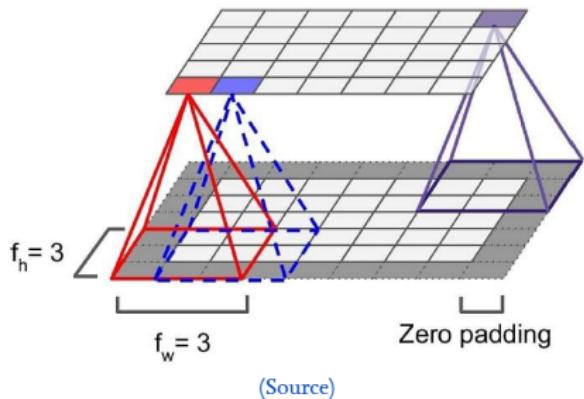
(Source)

(Conv 2D example)

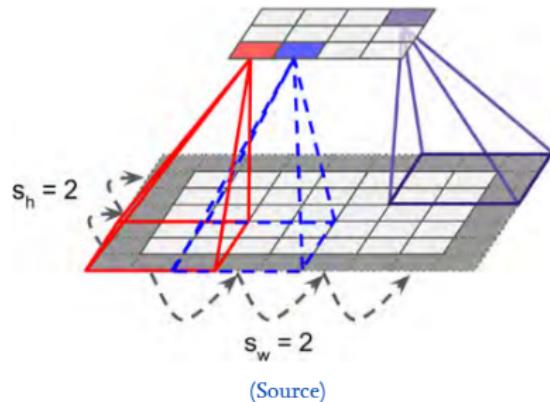
Convolutional Neural Networks

Convolutional layers (III)

Padding



Stride



Convolutional Neural Networks

Convolutional layers (IV)

0	0	0
0	1	0
0	0	0

identity



1	0	-1
2	0	-2
1	0	-1

left sobel



1	2	1
0	0	0
-1	-2	-1

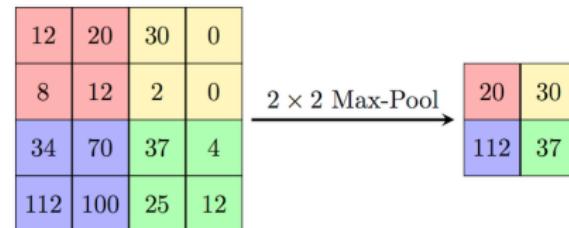
top sobel



(Image kernels)

Convolutional Neural Networks

Max-pooling layer



Max-pooling down-samples data instances

- Given a matrix, it takes its maximum value
- Usually the matrix is $n \times n$ (2D)

Benefits

- Dimensionality reduction
- Filters irrelevant information
- Invariant to scale

Convolutional Neural Networks

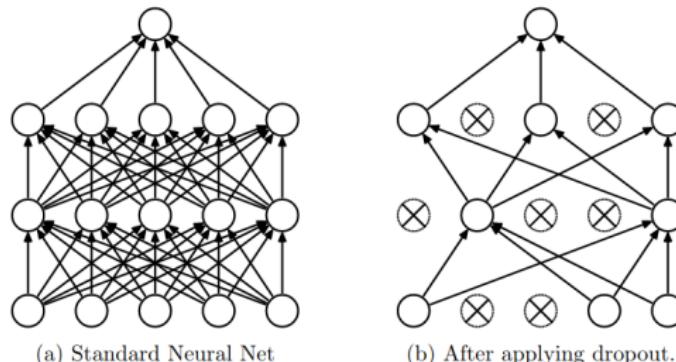
Dropout layer

Dropout is a regularization technique for neural networks

- Dropout deactivates a neuron with probability p for each iteration

Related concept: dense layers

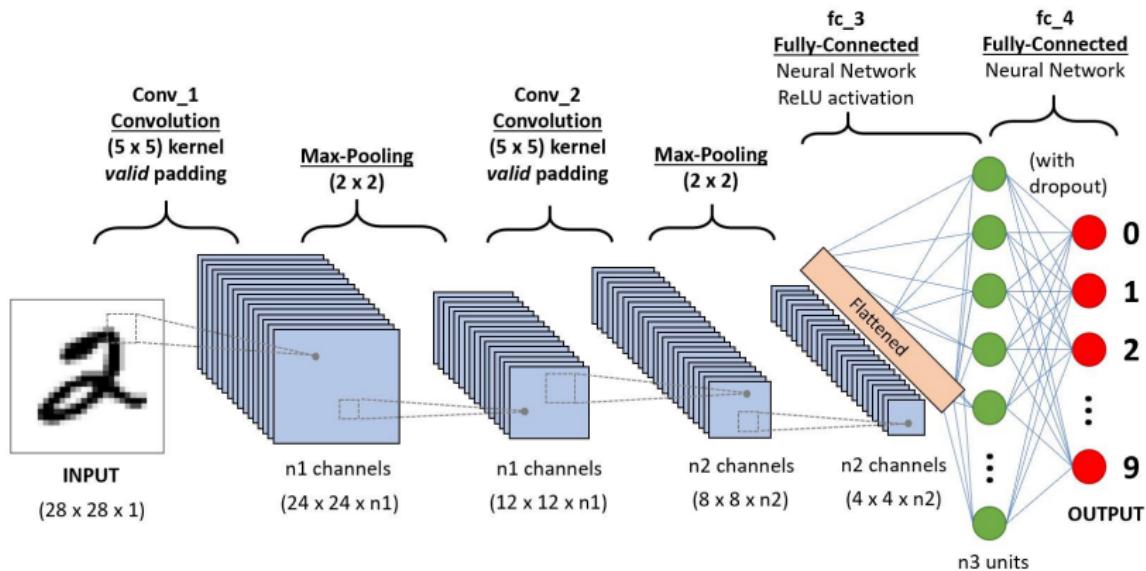
- In Keras, it is just a fully connected layer with regular neurons



(Srivastava et al. (2010))

Convolutional Neural Networks

CNN architectures: standard (I)

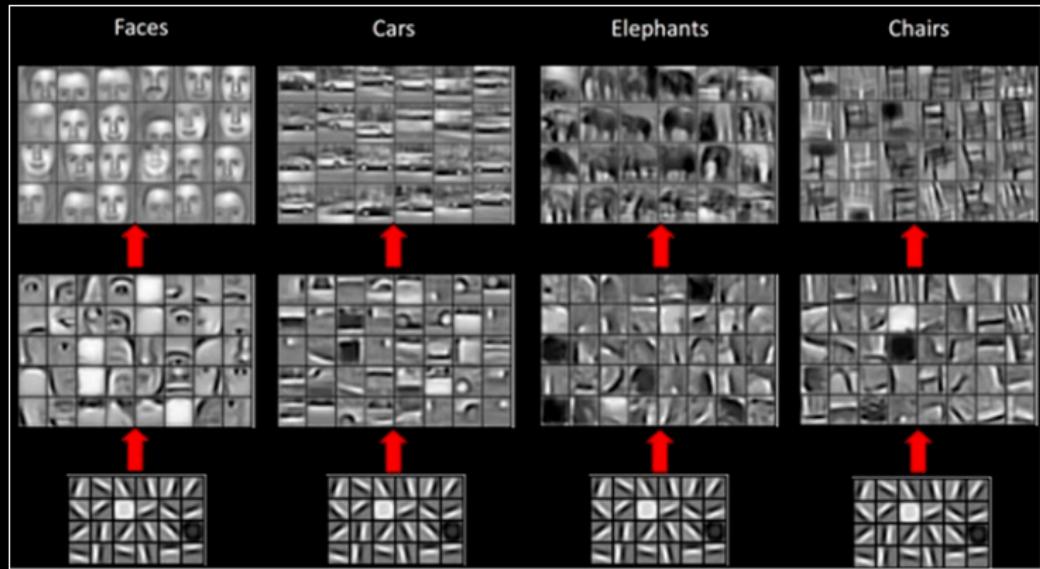


(Source)

(Demo)

Convolutional Neural Networks

CNN architectures: standard (II)



Convolutional Neural Networks

Other CNN architectures

Other CNN architectures

- LeNet-5
- AlexNet
- GoogLeNet
- VGGNet
- ResNet
- Xception
- SENet

Performance on Imagenet

Model	Acc.	Parameters	Depth
VGG16	0,901	138,357,544	23
InceptionV3	0,937	23,851,784	159
ResNet50	0,921	25,636,712	-
Xception	0,945	22,910,480	126

Famous generative deep networks

Model	Parameters
GTP-2	1.5 billion (1,500,000,000)
GTP-3	175 billion (175,000,000,000)
Dall-E 2	12 billion (12.000.000.000)
ChatGPT	Unknown

Convolutional Neural Networks

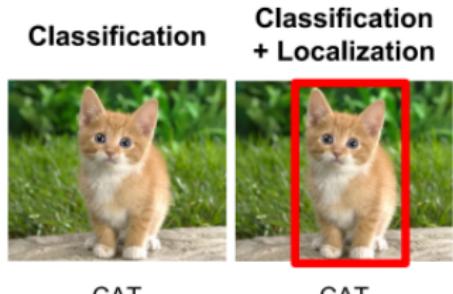
Classification and localization

Localization: Predict a bounding box around the object

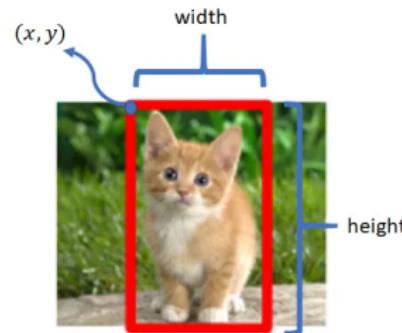
- Classification as a regression problem: X, Y, width and height
- Four output units per class

We need annotation tools

- (VGG Image annotator), (LabelImg), (ImgLab)



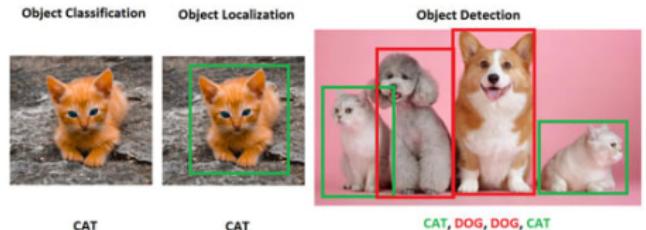
(Source)



Convolutional Neural Networks

Object detection

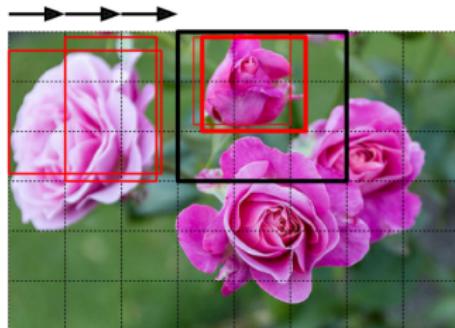
Object detection: Classify and localize multiple objects in an image



(Source)

Old approach: Slide across a single object CNN

- The image must pass through the CNN several times
- Bad for real-time applications

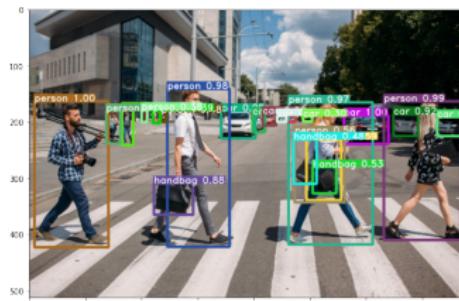


Convolutional Neural Networks

Object detection: You Only Look Once (YOLO)

YOLO is a fast architecture for real-time object detection

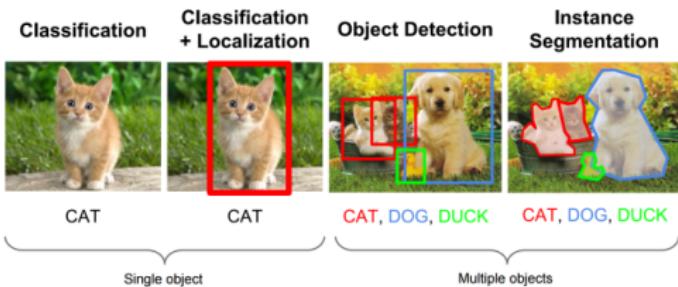
- Several versions proposed by different teams
- Valid for video object detection
- Several implementations
- (Yolo v8)
- YOLO supports image segmentation



Convolutional Neural Networks

Semantic segmentation

Semantic segmentation: Pixel-level classification



(Source)



(Source)

Recurrent networks

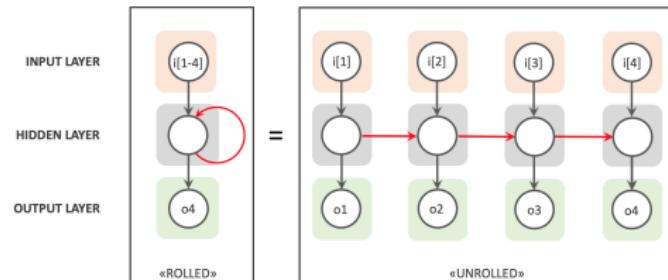
Recurrent neural networks (I)

Recurrent networks have connections pointing backward

- Time-series, NLP, audio, video, ...

Neurons have memory, or **state**

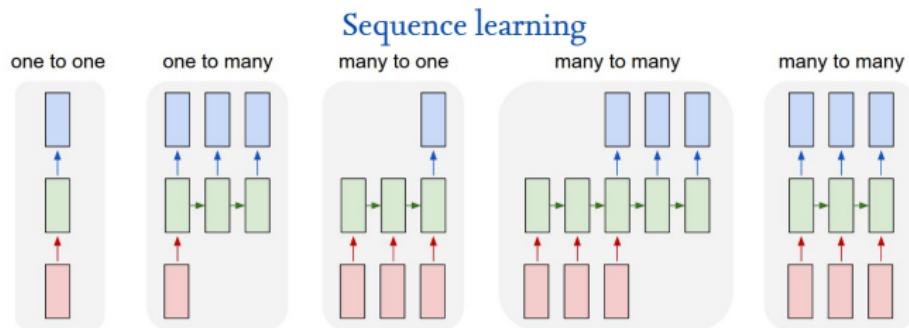
- Named **cells**
- In basic neurons, state and output are the same



(Source)

Recurrent networks

Recurrent neural networks (II)



One to many vec2seq	Many to one seq2vec	Many to many seq2seq
Image description	Spam classification Time series forecasting Sentiment score	Machine translation

Recurrent networks

Recurrent neural networks (III)

RNNs problems

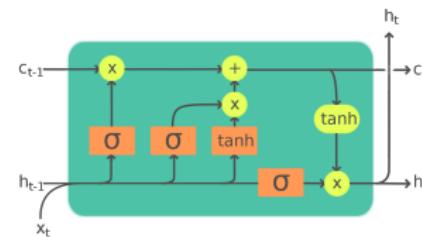
- Gradient instability
 - Smaller learning rate
 - tanh as activation function
 - Usual DL tricks
- Short memory
 - Information vanishes fast
 - Much more difficult solution

Recurrent networks

LSTM networks

LSTM: Long-Short Term Memory

- Complex cell that improves long-term memory
- Two states: short and long terms
- Very much used as a basic cell
- Much better performance
- Lower training time



Legend:

Layer ComponentwiseCopy Concatenate



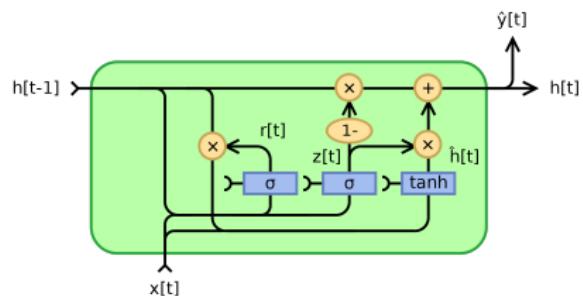
(Source)

Recurrent networks

GRU

GRU: Gated Recurrent Unit

- Simplification of LSTM
- Seems to perform as well as LSTM

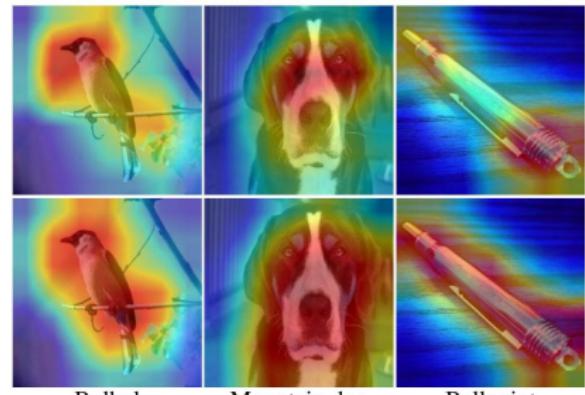


(Source)

Recurrent networks

Advanced architectures

- Attention networks
- Transformers (and visual transformers)
- Encoder-decoder architectures



(Source)

Recurrent networks

RNNs cool applications: language generation with a LSTM

Shakespeare

PANDARUS:

Alas, I think he shall be come
approached and the day
When little strain would be
attain'd into being never fed,
And who is but a chain and
subjects of his death,
I should not sleep.

Second Senator:

They are away this miseries,
produced upon my soul,
Breaking and strongly should
be buried, when I perish
The earth and thoughts of
many states.

Cervantes

-pero me manda mal -respondió
el maererlino, que yo no será
mejor que se ha de ser de don
quijote, que no le
acompañasen, se vengaban de
haber tenido al cabo de los
cuales se venció de su escudero,
y el infierno que todos los
pasaron su hermosura, y a mí
me parece que se cuenta el
yelmo de la caballería, don
gregorio de la similacada al
caballo que de la mano, según la
señora dulcinea del toboso.
-no es bien -respondió sancho-,

Recurrent networks

RNNs cool applications: music synthesis

DeepBach

- (Video)
- (Paper)
- (Code)

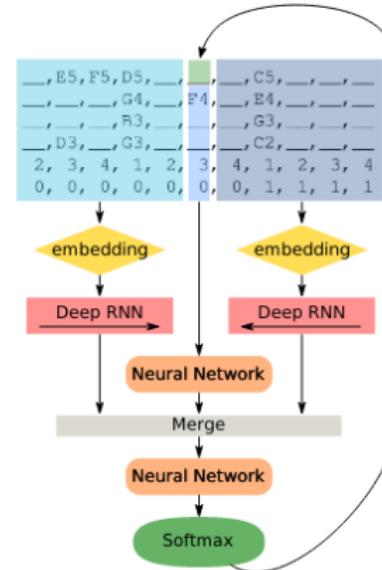
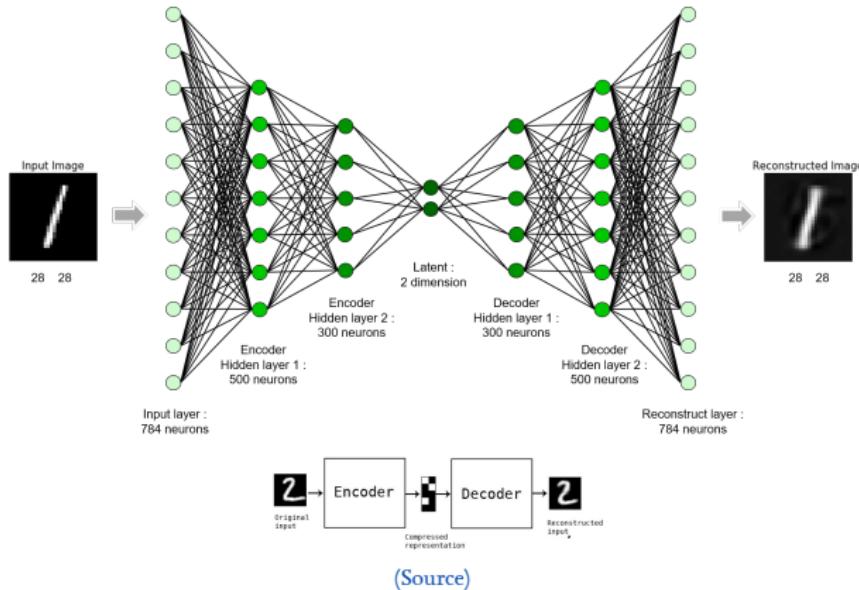


Figure 4. Graphical representations of DeepBach's neural network architecture for the soprano prediction p_1 .

(Source)

Autoencoders

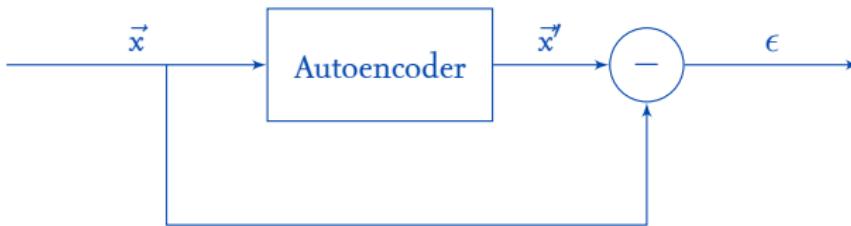
Autoencoders



Important concepts: latent space and latent variables

Autoencoders

Autoencoders for anomaly detection (I)



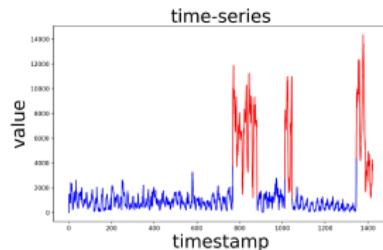
Reconstruction error is an anomaly measure

- A norm can be computed to provide a global measure (MAE/MSE), or ...
- ... keep reconstruction error as vector

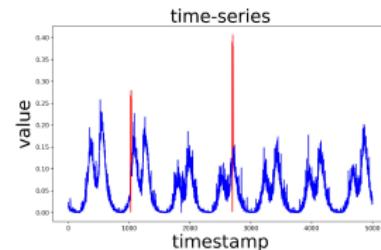
PCA may be used, less powerfull than autoencoders

Autoencoders

Autoencoders for anomaly detection (II)



(a)



(b)

(Source: Niu, Z.; Yu, K.; Wu, X. LSTM-Based VAE-GAN for Time-Series Anomaly Detection. *Sensors* **2020**, *20*, 3738.)

Great flexibility to handle reconstruction error

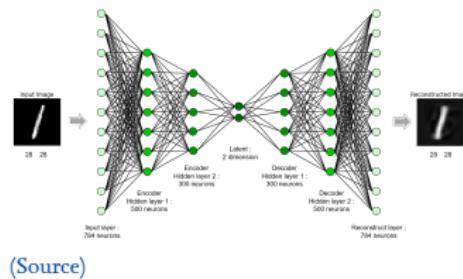
- Trigger an alarm based on a threshold
- Analyze the time-series
- Feed a classifier

Autoencoders

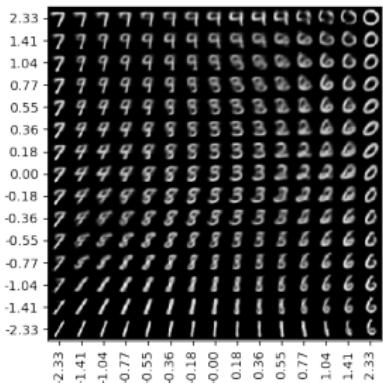
Autoencoders as generative models (I)

Any autoencoder may be used as a generative model

- The decoder can reconstruct an instance from a latent space sample



(Source)

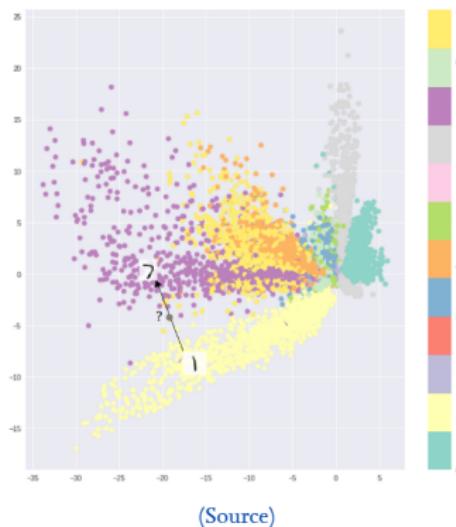


(Source)

Autoencoders

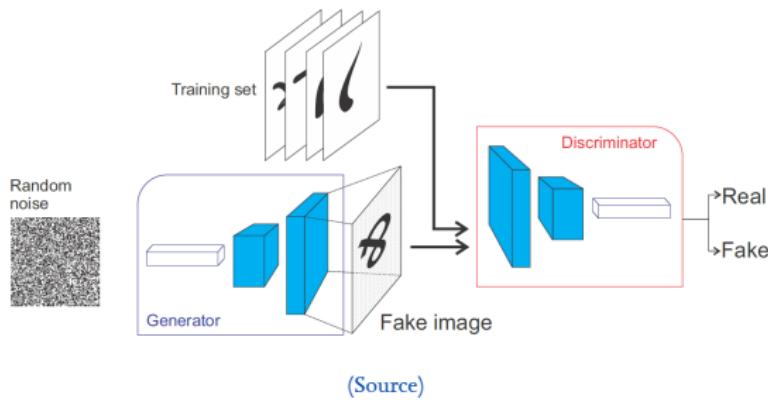
Autoencoders as generative models (II)

Regular autoencoders are not a good choice for generative models



Advanced topics

Generative networks: GAN (I)



Examples:

- (Faces), (art), (Words), (cats)

Advanced topics

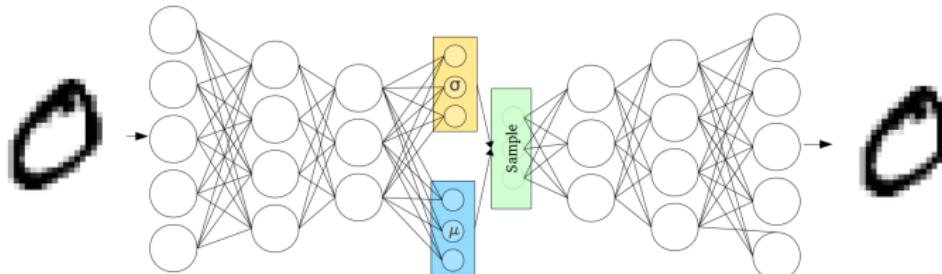
Generative networks: GAN (II)

GauGAN (Demo)



Advanced topics

Variational Autoencoders (I)



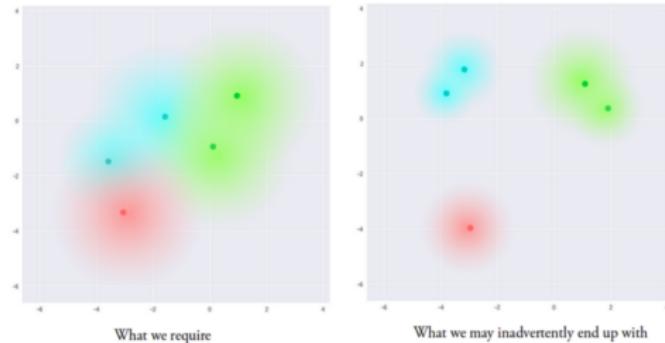
(Source)

VAEs encodes latent variables as probability distributions

- Gaussian distributions with μ and σ
- Decoder sample the distributions

Advanced topics

Variational Autoencoders (II)



(Source)

We want a structured latent space

- Penalty based on Kullback-Leibler (KL) divergence
 - KL measures divergence between two probability distributions

Advanced topics

Variational Autoencoders: semantics (I)

Astonishing VAE feature: latent space has semantics!

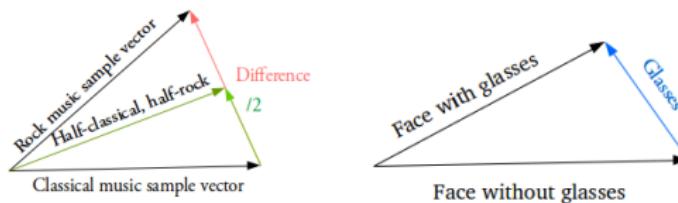


(Source)

Advanced topics

Variational Autoencoders: semantics (II)

Another incredible VAE property: 'semantic' arithmetic operations



(Source)

(Embedding Projector)

Advanced topics

Generative models state-of-the art

Dall-E 2

A raccoon astronaut with the cosmos reflecting on the glass of his helmet dreaming of the stars



(Source)

Cosmic thoughts exploding



(Source)

Advanced topics

Adversarial examples

