

Computer Vision Winter School

Graz, 2023

1 Setup

Google Colab

1. Create account for Google Colab / sign in: <https://colab.research.google.com/>
2. File > Open Notebook > Github: Input our link.

Tropy

Please install the Tropy software (for exploring photos and tags):
<https://trophy.org/>

2 Basic Linux Commands / Navigating the terminal

`pwd` Show current directory (print working directory)
`cd directoryName` Change directory
`cd .` Reference to current directory
`cd ..` Reference to parent directory
`ls` List files

3 Reading CSV & using Pandas dataframes

Pandas

The Pandas library is built on NumPy and provides easy-to-use data structures and data analysis tools for the Python programming language.

Install and Import convention

```
1 pip install pandas
2 import pandas as pd
```

Read and Write to CSV

```
1 df = pd.read_csv('filename.csv', encoding='utf8')
2 df.to_csv('filename.csv', index=False, encoding='utf8')
```

Select the index of a row

```
1 df[].index
```

Drop a row or column in a dataframe

```
1 df.drop()
```

Drop rows based on certain condition

```
1 df.drop(df[(df['XY'] > 400) & (df['YX'] < 600)].index)
```

Create new empty column

```
1 df.loc['ROW', 'COLUMN']
```

Select rows and columns by names

```
1 df['NAME'] = ''
```

Iteration

```
1 df.iteritems()
2 df.iterrows()
```

4 Drive & Colab

Mounting Drive in Colab

```
1 import glob
2 from google.colab import drive
3 #if len(glob.glob("/mnt/drive/*")) == 0:
4 drive.mount("/mnt/drive", force_remount=True)
```

5 CV Terms Glossary

Computer Vision Glossary (Vocabulaire International du Deep Learning)
Prepared by A. Nicolaou

Artificial Intelligence (AI)

A noun phrase that was banned for decades from funding applications and marketing

Machine Learning (ML)

An artificial intelligence paradigm built around back-error propagation algorithms for training statistical models

Groundtruth

An uncontested rigidly defined desired output for a ML system coupled with the respective input that should generate it

Groundtruthing

Manually annotating data in order to create ground truth

Loss (error) function

A differentiable function applied on the output of a model demonstrating how the output for a specific input could have been better. Usually utilizes the groundtruth. **Typical:** Cross-entropy, Mean Square Error, Focal, etc...

Supervised Learning

Training a statistical model by constructing error gradient from triplets input, observed output, and desired output.

Unsupervised Learning

Training a statistical model without knowing the desired output for our data. **Examples:** Clustering, self-supervision, Metric-learning

Reinforcement Learning

Training a statistical model with an abstract notion of achieving a goal in the long-run instead of with groundtruth. It is a kind of semi-supervised learning. Is also used when we can't construct a differentiable error function.

Support Vector Machine

Shallow classifiers that are usually employed on top of hand-crafted features

Pattern Recognition

The field of study of applying ML on to data of specific domains'

Information Theory

Paradigm invented by Claude Shannon in the 40's quantifying information through statistics. A major contributor to modern computer science.

Signal Processing

Field of electrical engineering applying information theory to natural world measurements. **1D:** Sound/Text. **2D:** Images. **3D:** Videos/Volumetric.

Pixel

One or more numerical value representing measured light at one or more wavebands. In deep learning it can hold more abstract information organised in a grid

Chanel

A part of a pixel representing a specific number for that pixel. **Eg:** Green from an RGB image.

Image
A 2D grid of equidistant pixels
Convolution
A local weighed sum operator over a signal with specific weights, can be multidimensional: 1D, 2D, 3D
Frequency Domain
Opposite: Time Domain (also for 2d and 3d). A representation where each number of a signal represents how much a frequency contributes to the total energy
Neural Network
A statistical model processing information structured in layers that process information in parallel through elementary operations
Deep Neural Network
A neural network with more than two layers, usually 10ths, 100s, or 1000s.
Weights
All variables required for inference and computed during training. AKA: model parameters
Vanishing/Exploding Gradients
In deep networks the error gradient can converge to zero (vanishing) or diverge to infinity (exploding)
Convolutional Neural Network
A neural network processing information that is structured in pixel format. Fully convolutional: Only convolutional layers (the information is always in pixel structure)
Residual Network
CNN that adds the output of layers to their input greatly facilitating back-error propagation
Recurrent Neural Network
A neural network organised in Cells where the output of a Cell applied on a part of a signal is feed into Cell along with the next part of the signal: typical RNNs: LSTM, GRU; training algorithm: Back-Propagation Through Time.
Convolutional Neural Network
A neural network with some convolutional layers
Deep Learning Framework
A library allowing the description of a network architecture so that we get the error backpropagation automatically: Pytorch, Tensorflow, Google-JAX, Theano, Caffe
Optimization
Finding the optimal parameters of a model in order minimize a mathematically defined function of it. Linear is easier than convex which is easier than Continuous
Model Training
Optimization of the weights (parameters) of a neural network in order to minimize the error over the train-set

Gradient Descent
Optimisation algorithm that jumps on the direction of the lowest error at every step. Think of walking down a mountain with fog
Stochastic Gradient Descent
Gradient descent on a subset of the train-set resampled randomly at every iteration
Regularization
Making a representation more simple but less exact. See also: bias-variance trade-off.
(Mini)batch
A random subsample of the train dataset stacked as a 4D tensor that propagates through the network in one step
Vector (space)
The location of a point expressed in a multidimensional space
Matrix
A MxN 2D grid of numbers where rows and columns can be perceived as vectors
Tensor
Generalisation of the concept of vectors to more than 2 dimensions
Metric Space
All possible vectors of a given size equipped with a distance function between any two members of the space
Whitening
Statistically preprocessing input samples so that they share the same properties. Not as important with pretrained models. Typical methods: PCA, ZCA
Baseline
A method used as a point of reference over which we expect to demonstrate improvement. The reasonable (non-innovative) way to solve a problem.
State-of-the-art
The consensus about which method is the best to solve a given problem
Generative Adversarial Network
A network that learned to generate images by trying to fool another network
Adversarial Samples
Samples that can fool classifiers
Autoencoder
A network with an hourglass architecture that usually learns to compress samples

Variational Autoencoder
A network with an hourglass architecture that was forced to learn a meaningful representation in an unsupervised way
Regression
Learning to predict a continuous variable
Morphological operators
Tools processing locally binary images. Popular in the 80s, 90s.
Image segmentation
Classifying all pixels of an image
Fundamentals
<ul style="list-style-type: none"> Image Pixel Convolution Optimization Algebra Vector Matrix Tensor Distance
Typical CV tasks
<ul style="list-style-type: none"> Classification Regression Object Detection Segmentation Image retrieval Texture OCR/HTR
Data (pre) Processing
<ul style="list-style-type: none"> Colorspace Manipulation Data Whitening Data Augmentation Data Synthesis
Training
<ul style="list-style-type: none"> Regularization dropout dropconnect Knowledge transfer Knowledge distillation