UNIVERSITY OF COPENHAGEN

Behavioural Analysis with Big Data & Deep Learning

PhD Course on *animal models of disease and behavioral analysis*

Raghavendra Selvan

Assistant Professor
Dept. of Computer Science (ML Section)
Dept. of Neuroscience (Kiehn Lab)
Data Science Lab
University of Copenhagen
Pioneer Centre for Artificial Intelligence, Denmark

raghav@di.ku.dk
/raghavian
`https://raghavian.github.io`

September 16, 2022

Materials

`https://raghavian.github.io/outreach/`

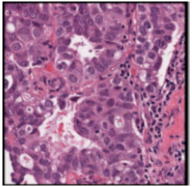
Overview

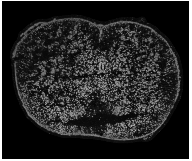
- Videos as Numerical Arrays
- Biomedical Image Analysis
- Model-based methods
- Deep learning based methods

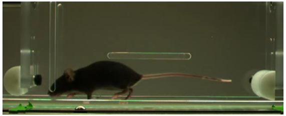
Overview

- Videos as Numerical Arrays
- Biomedical Image Analysis
- Model-based methods
- Deep learning based methods

Biomedical Images & Videos are discretized numerical arrays with physical properties for intensities



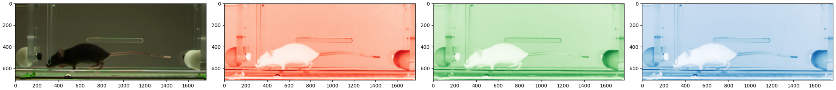




- Microscopy images, Magnetic resonance imaging (MRI), optical videos
- Specialized data containers for storing and processing: DICOM, NIFTI, TIFF
- 2D/3D/4D images with more than one channel
- Videos are stacks of images along time-axis

Storing Intensity information in regular arrays

- Grayscale images are stored in single channel 2D arrays
- Intensity values are stored as integers or floats
- Ex: An 8-bit integer can have 2**8=256 levels
  0 $\rightarrow$ black and 255 $\rightarrow$ white, and other shades of gray
- Color images in multi-channel (3 or 4)
- RGB is a common way of storing images
- Multi-stack images



Overview

- Videos as Numerical Arrays
- Biomedical Image Analysis
- Model-based methods
- Deep learning based methods

Notations

Observed data	$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} : \mathbf{x}_i \in \mathbb{R}^{C \times H \times W \times D}$
Labels / Targets	$\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\} : \mathbf{y}_i \in \mathbb{R}^{M \times H \times W \times D}$
Decision functions/ Models	$f_\theta(\cdot) : \mathbf{X} \mapsto \mathbf{Y}$

Videos as sequence of images

Easiest approach for Video Analysis is to perform frame-by-frame image analysis

Computational methods can reduce the labour intensive processes of Image Analysis

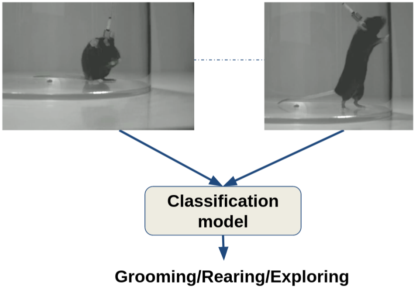
- Expert annotators are expensive to perform mundane tasks
- Tasks like delineation are tedious and cumbersome [1]
- Variability across/within annotators
- Large volumes of high dimensional data
- Discover novel metrics from large volumes

RootPainter3D: Interactive-machine-learning enables rapid and accurate contouring for radiotherapy. AG Smith et al. 2021

Tasks in Biomedical Image Analysis

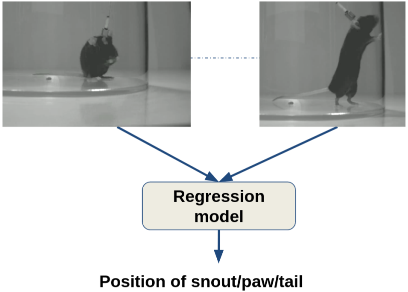
Depending on the nature of the input-output relations several downstream tasks can be formulated

Image Classification



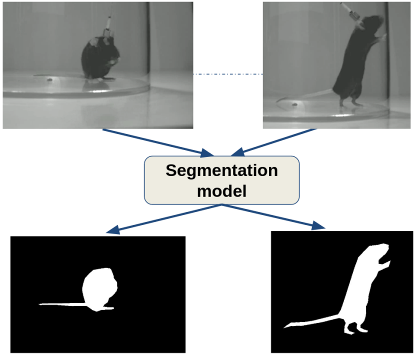
- $f(\cdot) : \mathbf{X} \in \mathbb{R}^{C \times H \times W \times D} \mapsto \mathbf{Y} \in \{0, 1\}^M$
- Image level targets
- Detection of behaviour episodes
- Quality Control for denoising
- Probabilistic predictions that can be thresholded
- Takes entire input image into consideration

Image Regression



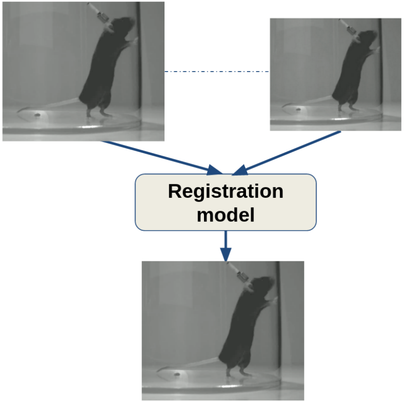
- $f(\cdot) : \mathbf{X} \in \mathbb{R}^{C \times H \times W \times D} \mapsto Y \in \mathbb{R}$
- Image level scores
- Grades of measurements
- Counting number of cells
- Anatomical measurements
- Harder than classification in most cases

Image Segmentation



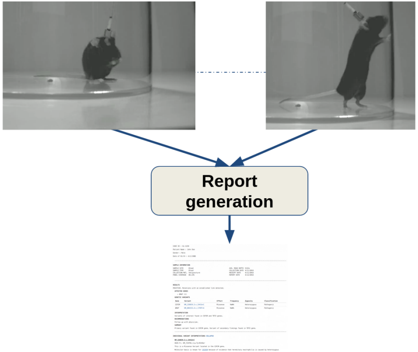
- $f(\cdot) : \mathbf{X} \in \mathbb{R}^{C\times H\times W\times D} \mapsto \mathbf{Y} \in \{0,1\}^{M\times H\times W\times D}$
- Pixel level predictions
- Foreground & backgroud delineations
- Useful for localizing decisions

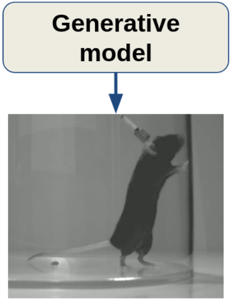
Image Registration



- $f(\cdot) : \mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^{C\times H\times W\times D} \mapsto \mathbf{Y}$
- Aligning images over different time points or to a reference atlas
- Predictions are Transformation matrices or Deformation fields
- Useful for quantifying temporal progression or colocalisation

Image understanding & Generative modeling





DeepLabCut Exercise: Part-1

We need to let the machine learn!

1. Install Anaconda
2. Create new virtual environment
   conda create -n DLC python=3.7
3. Install wxpython
   conda install -c conda-forge wxpython=4.0.7
4. Install DeepLabCut
   pip install "deeplabcut[gui]"

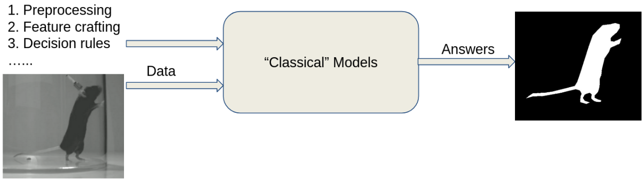
Machine Learning Fundamentals

Overview

- Videos as Numerical Arrays
- Biomedical Image Analysis
- Model-based methods
- Deep learning based methods

Model-based methods are based on hand-crafted rules

Consider the example of segmenting lungs from chest X-ray images.



Can we come up with the simplest decision based segmentation model?

- **Hint**: Preprocessed. Intensity range is 0 (black) - 255 (white)
- Intensity thresholding. $\mathbf{Y} = \mathbb{I}[I_{min} \leq \mathbf{X} \leq I_{max}]$
- Other standard (powerful) segmentation methods ([2], [3])

[2] Seeded region growing. R. Adams, L. Bischof. 1994

[3] Watershed of a continuous function. L. Najman and M. Schmitt. (1994)

Bias-variance trade-off

Encoding strong prior → Low Variance, High Bias

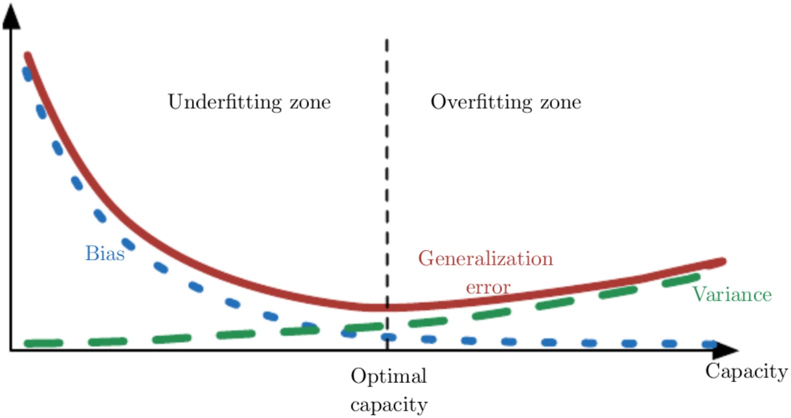
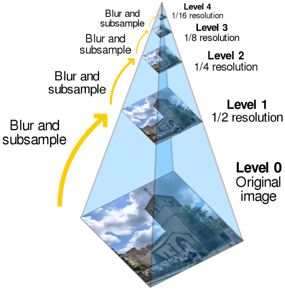


Figure from I. Goodfellow et al., Deep Learning. 2017

Rich classes of model-based methods are based on custom filters



Source [4]

- Hand-crafting features is same as designing filters
- Are high *bias* models
- Example: The Laplacian image pyramid
- Scale-space theory [5]
- + Incorporates prior knowledge
- + Robust and efficient methods
- + Few parameters to tune
- - Cumbersome to tune parameters
- - Transfering domain knowledge can be challenging

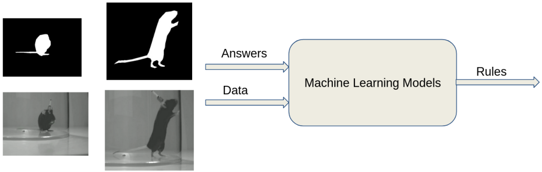
[4] https://en.wikipedia.org/wiki/Pyramid_(image_processing)

[5] T. Lindeberg, Scale-space theory: A basic tool for analyzing structures at different scales. (1994)

Overview

- Videos as Numerical Arrays
- Biomedical Image Analysis
- Model-based methods
- Deep learning based methods

Machine Learning* is the process of *Learning from Data*



Based on Fig. 1.9, from Mostafa et al.

- Hand-crafting gives way to *learning from data*
- Approximating the underlying data distribution from observed data
- Over-parameterised function approximators
- Parameters learned using automatic differentiation

+/- No domain knowledge required

\+ Can learn efficiently from labelled datasets

\- Features and flaws are learned

\- Curating data is highly important

*We basically only focus on **Supervised Deep Learning** which is a small subset of all of ML!

Bias-variance trade-off

Learning from data → Low Bias, High Variance

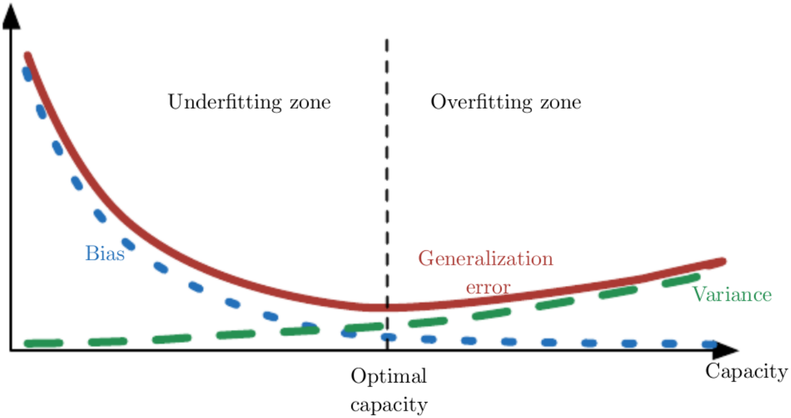
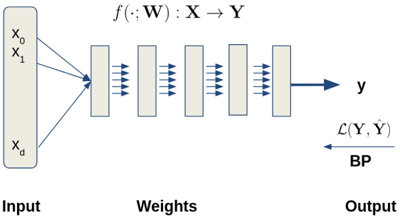


Figure from I. Goodfellow et al., Deep Learning. 2017

One Slide Introduction to Deep Learning

- Over-parameterised function approximator
- Can have thousands of millions of parameters
- Optimising these parameters is difficult
- Magic sauce of DL is Automatic Differentiation
- Implemented in several powerful frameworks (Pytorch, Tensorflow)
- Scalable training on GPUs



A Quick Peek into Convolutional neural networks (CNNs)

- Learnable convolution kernels
- Technically, not convolution but cross-correlation
- Implemented as matrix multiplications
- Kernel flipping can be overcome during learning
- Known to learn general image descriptors
- And also, specialized task-specific kernels

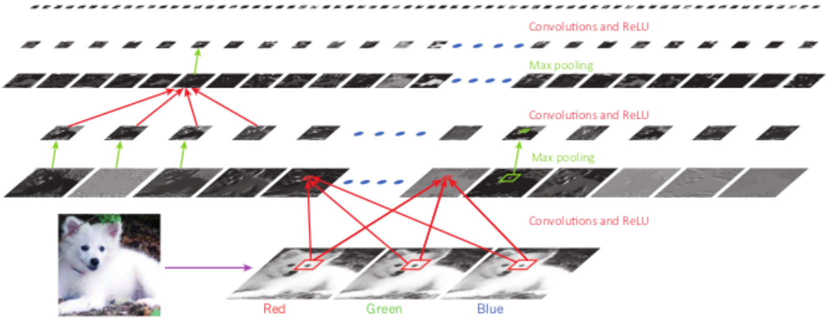
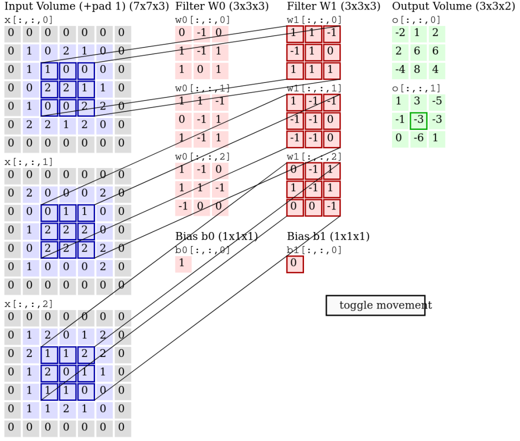


Figure reproduced from Le Cun et al. Deep Learning. 2015

The Convolution Operation in CNNs

Adapted from https://cs231n.github.io/convolutional-networks

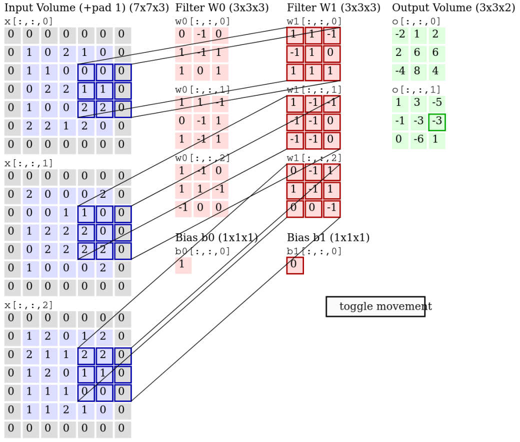
The Convolution Operation in CNNs



Back to Top

Adapted from `https://cs231n.github.io/convolutional-networks`

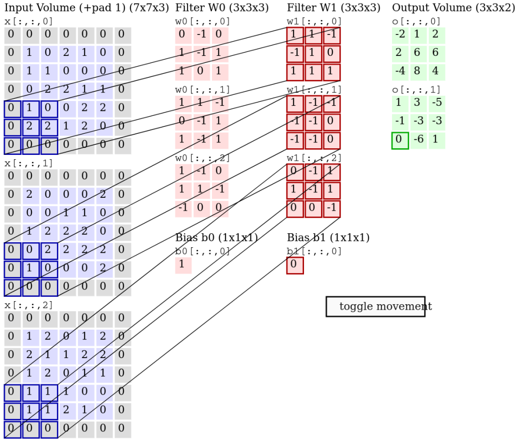
The Convolution Operation in CNNs



Back to Top

Adapted from `https://cs231n.github.io/convolutional-networks`

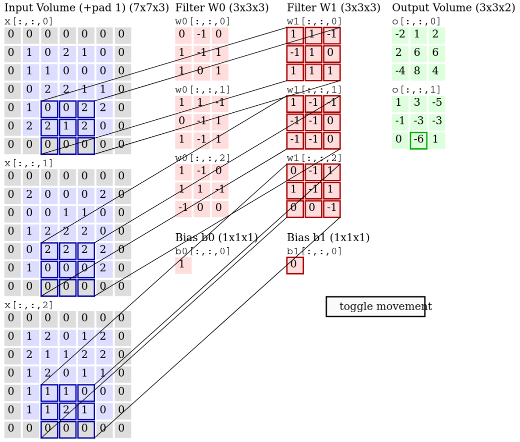
The Convolution Operation in CNNs



Back to Top

Adapted from `https://cs231n.github.io/convolutional-networks`

The Convolution Operation in CNNs



Back to Top

Adapted from `https://cs231n.github.io/convolutional-networks`

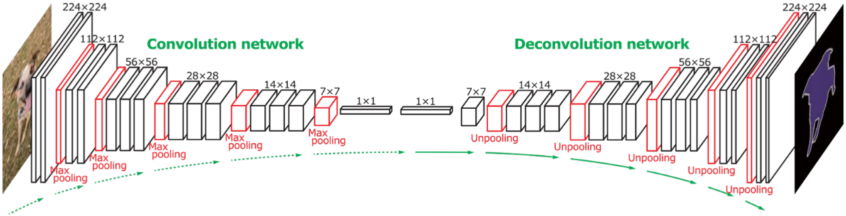
More formally, CNNs exploit some key properties when operating on images

Success of CNNs in computer vision is due to several properties.

Consider a translation operation $\mathcal{T}(x) = (x - v)$

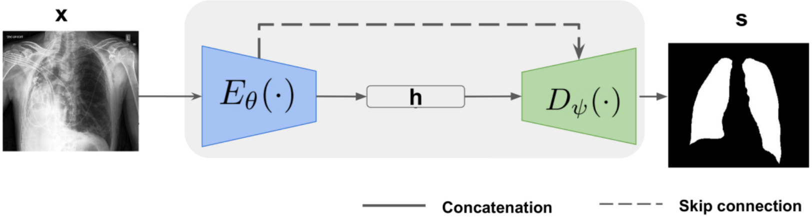
- **Translation invariance**: $y = f(\mathcal{T}(x)) = f(x)$, or
- **Translation equivariance**: $y = f(\mathcal{T}(x)) = \mathcal{T}(f(x))$
- **Scale separation**: Long range dependencies from multi-scale interaction terms
- **Compositionality**: $\mathbf{y} = f_1 \circ f_2 \circ \cdots f_L(\mathbf{x})$
  where $f_i(\cdot)$ are comprised of convolution kernels, non-linearities and sometimes pooling operations.

Fully Convolutional Neural Network (FCNN) for Image outputs



Source: `https://arxiv.org/abs/1411.4038`

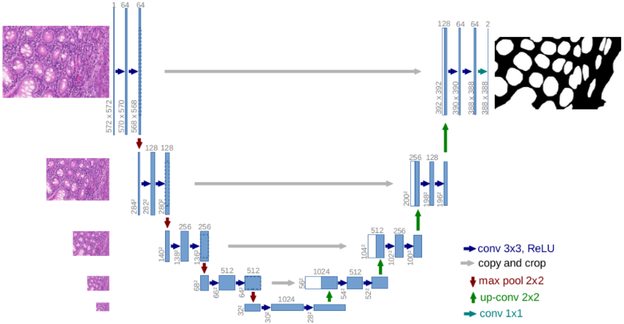
Closer look at the U-net (aka work-horse CNN model for image segmentation)



- Supervised segmentation model
- Encoder-decoder type neural network
- Encoder and Decoder are comprised of several layers of CNNs
- Encoder and decoder paths communicate via *skip connections*

Ronneberger O. et al. U-net for biomedical image segmentaion. 2015

U-net architecture



Source Ronneberger et al. 2015

Laplacian pyramids?

Summary

- Image analysis is a challenging sub-domain of Computer Vision
- Downstream tasks can be of huge consequence
- Automating IA can be beneficial to practitioners
- Opens up novel paradigms of handling large volumes of dat
- Model based methods can be powerful; they encode strong prior
- Trade-off between optimization and learning
- ML models are high variance and learned from data
- Data preparation is the most important step!
- CNNs are indispensable when working with images.
- U-net like models use CNNs in a hierarchical manner for segmentation
- Image analysis is a starting point for video analysis

Resources

https://github.com/DeepLabCut/DeepLabCut
https://github.com/DeepLabCut/DLCutils
https://www.fast.ai/