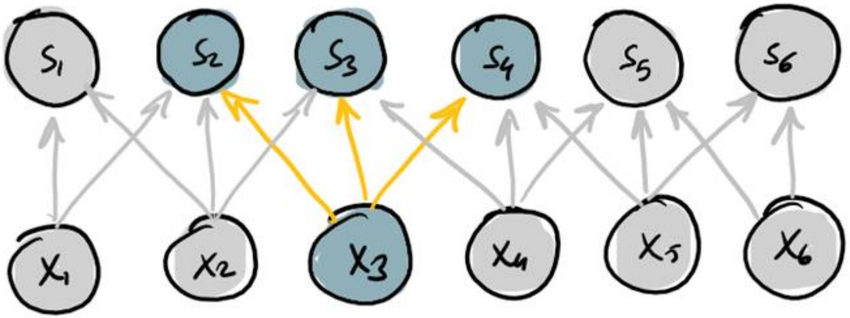
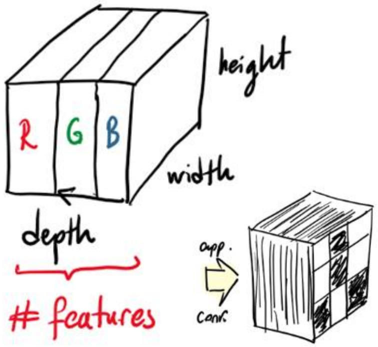
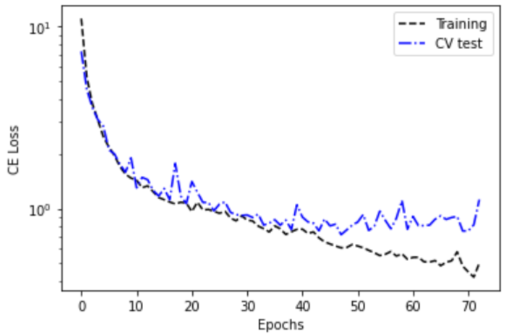
Data Driven Engineering II: Advaced Topics

Image processing and analysis

Institute of Thermal Turbomachinery
Prof. Dr.-Ing. Hans-Jörg Bauer







Term Projects

Welcome to DDE II projects!

If you are interested in the group projects for fun or planning to take the final exam for credits, you need to register to a topic before 14.05.021.Note that each topic has a number of maximum participants. You may find the details in Lecture 1.

Particle Image Density Analysis in PIV Recordings

An object detection study for PIV analysis

Free places: 1

Physical interpretation of LCSs

Data driven model discovery in air blast atomizers

Free places: 5

Time resolved flow field analysis in film cooling

PIV data will be used for flow analysis.

Free places: 5

Others

for HPC access

Period of Event: Today - 14. May 2021

Outline of the week :

Karlsruhe Institute of Technology

Conv. Neural Networks

* What is CNN ?
* Why convolution is useful ?
* Where is it useful ?
* CNN – How does it work ?
* "Hall of Fame" : Popular Arch. } Next week
* Transfer Learning with CNN } Next week

"The soul never thinks without a picture."

Conv. Neural Networks: Basics

Convolutional Neural Network

$\underbrace{\text{Convolutional}}_{\text{Specialized}}$ $\underbrace{\text{Neural Network}}_{\text{Training}}$

Specialized

⇓

{ Organized, Grid-like data }

1D 2D 3D N-D

Training

⇓

Labelled data

⇓

"Backprop."

☑ Flag ship of Deep Learning

☑ Image / Video processing

⇓

Benchmark datasets

⇓

Best solutions ~ weekly / monthly

Conv. Neural Networks: Basics

Karlsruhe Institute of Technology

Datasets

3,749 Machine Learning Datasets

"paperswithcode"

CIFAR-10

The CIFAR-10 dataset (Canadian Institute for Advanced Research, 10 classes) is a subset of the Tiny Images dataset and consists of 60000 32x32 color images. The images are labelled with...

5,634 PAPERS • 44 BENCHMARKS

ImageNet

The ImageNet dataset contains 14,197,122 annotated images according to the WordNet hierarchy. Since 2010 the dataset is used in the ImageNet Large Scale Visual Recognition Challenge...

5,565 PAPERS • 56 BENCHMARKS

MNIST

The MNIST database (Modified National Institute of Standards and Technology database) is a large collection of handwritten digits. It has a training set of 60,000 examples, and a test set of...

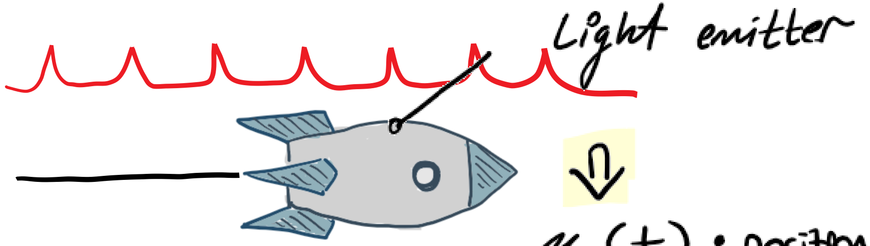
3,881 PAPERS • 36 BENCHMARKS

COCO (Microsoft Common Objects in Context)

The MS COCO (Microsoft Common Objects in Context) dataset is a large-scale object detection, segmentation, key-point detection, and captioning dataset. The dataset consists of 328K im-...

3,804 PAPERS • 58 BENCHMARKS

What is convolution operation?



$x(t)$: position

t	x

$$s(t) = \int x(a)\, w(t-a)\, da$$

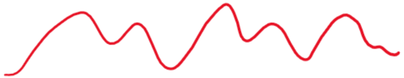
weighted

history

"Convolution"

$W :=$ wisdom

noise



Averaging ~ integration

Recent data is more relevant!

weighted averaging

Convolution in CNN:

Term.

- $x :=$ input
- $w :=$ kernel
- $s :=$ feature map

Data

- Discrete; $\int \rightarrow \sum \Rightarrow s = \sum x(a)\, w(t-a)$
- Multi-dim; 2D $\Rightarrow$ $s(i,j) = \sum_m \sum_n I(m,n)\, K(i-m, j-n)$

$$I(m+i, n+j)\, K(m,n) \; *$$

$$(a+b = b+a)$$

Wait a minute... Is it you MLP?

* What is special here?

Operation ⇒ " Multiplication by a matrix "

* Kernel is smaller than input

$m \times n \Longrightarrow k \times n \; ; \; k \ll m$

$$\underset{I}{\begin{bmatrix} 1 & 3 & 4 \\ 9 & 6 & 7 \\ 5 & 0 & 2 \end{bmatrix}} * \underset{K}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \begin{bmatrix} a+3b+9c+6d & 3a+4b+6c+7d \\ 9a+6b+5c+\emptyset & 6a+7b+\emptyset+2d \end{bmatrix}$$

Wait a minute... Is it you MLP?

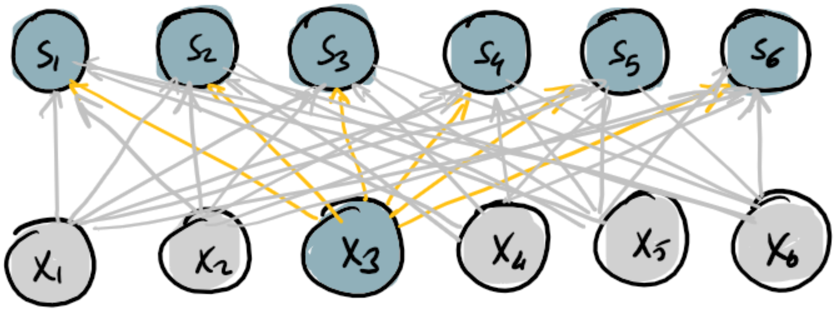
$$I \begin{bmatrix} 1 & 3 & 4 \\ 9 & 6 & 7 \\ 5 & 0 & 2 \end{bmatrix} \dashrightarrow \underbrace{\begin{bmatrix} 1 & 3 & 4 & 9 & 6 & 7 & 5 & 0 & 2 \end{bmatrix}^T}_{I^+}$$

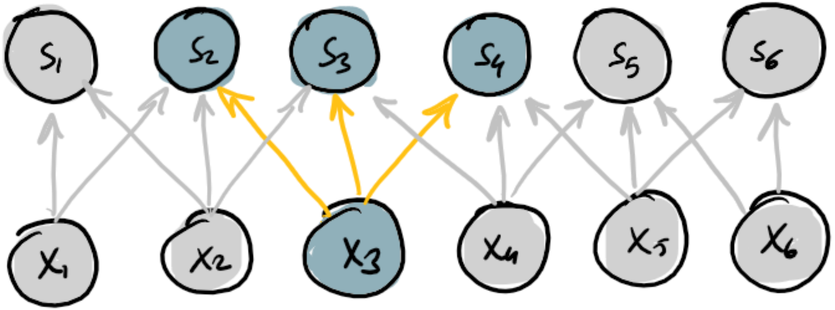
$$K \begin{bmatrix} a & b \\ c & d \end{bmatrix} \dashrightarrow \underbrace{\begin{bmatrix} a & b & \varnothing & c & d & \varnothing & \varnothing & \varnothing & \varnothing \\ \varnothing & a & b & \varnothing & c & d & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing & a & b & \varnothing & c & d & \varnothing \\ \varnothing & \varnothing & \varnothing & \varnothing & a & b & \varnothing & c & d \end{bmatrix}}_{K^+}$$

$$K^+ * I^+ \Rightarrow \underbrace{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \end{bmatrix}^T}_{\text{Reshaping}}$$

$$\begin{bmatrix} a+3b+9c+6d & 3a+4b+6c+7d \\ 9a+6b+5c & 6a+7b+2d \end{bmatrix}$$

Advantage #1 Sparse interactions





MLP: Densly conected

$x_3 \rightarrow (s_1, \ldots s_n)$

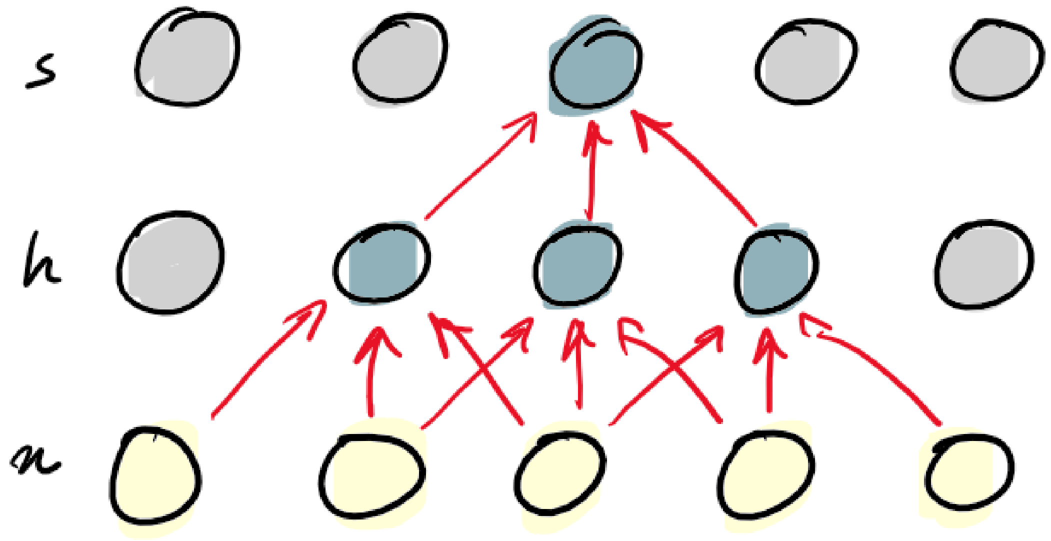
$O(m \times n)$ runtime/example

CNN $\Rightarrow k = 3$ $\quad x_3$ affects only (s_2, s_3, s_4)

$O(k \times n)$ runtime/example

Advantage #1 Sparse interactions

Receptive field in deeper layers

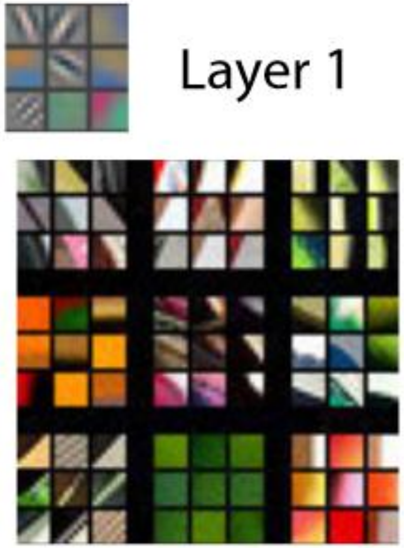


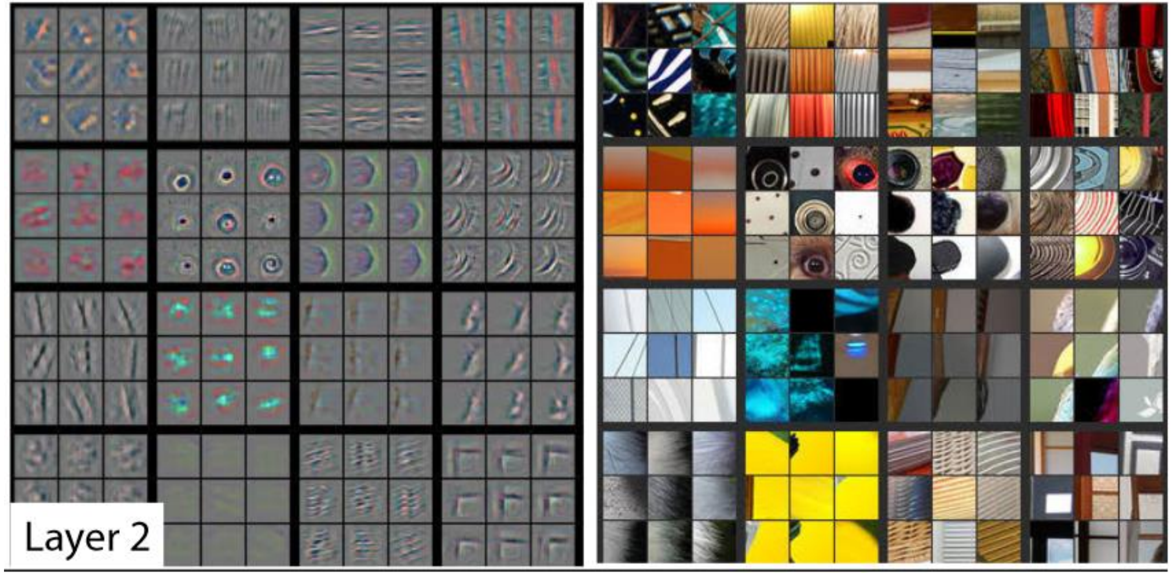
" indirectly connected to most of input layer "

Advantage #1 Sparse interactions

Receptive field in deeper layers

Visualizing and understanding convolutional networks

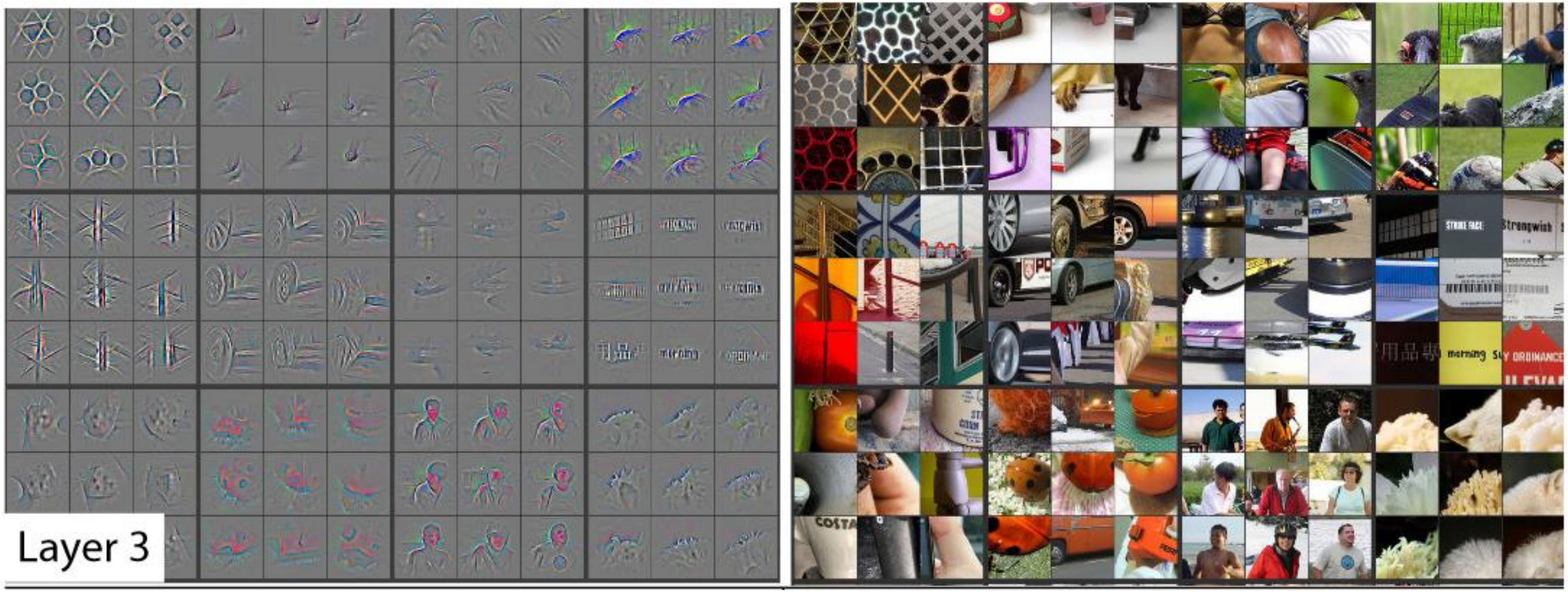




Advantage #1 Sparse interactions

Receptive field in deeper layers

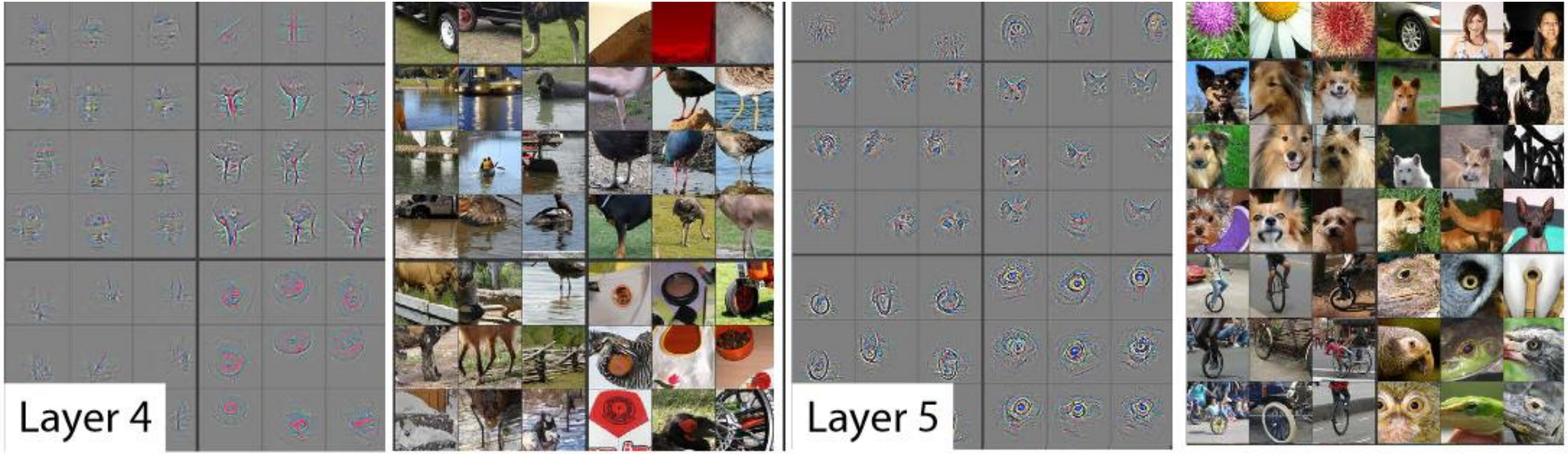
Visualizing and understanding convolutional networks



Advantage #1 Sparse interactions

Receptive field in deeper layers

Visualizing and understanding convolutional networks



Advantage # 2 Shared parameters

* In MLP; each element in matrix W is learnt & used only once

* In CNN; we learn filters (kernels)
  ↳ used at every position of the input

* ↑↑ memory eff. & statistical eff.

Eg
- Image of 280 × 320 px.
- $k = 2$

⇒ 280 × 319 px.

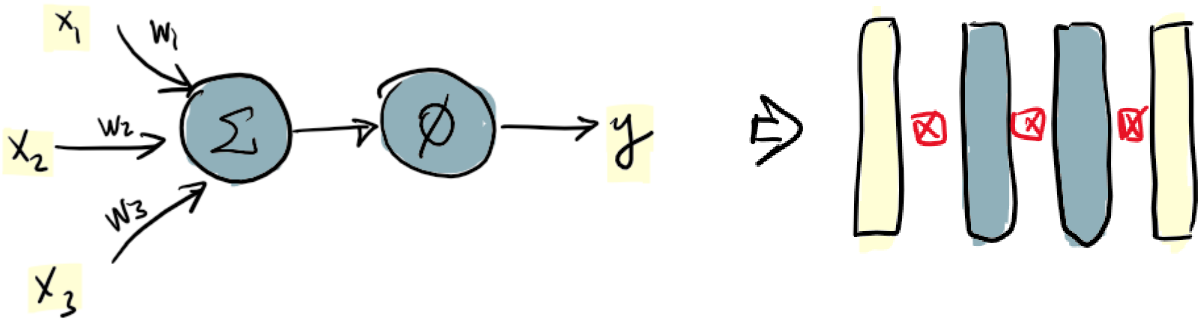
CNN; 280 × 319 × 3 operation (× × +)

MLP; [280 × 320 × 280 × 319]

- CNN ~ 60k compt. more efficient

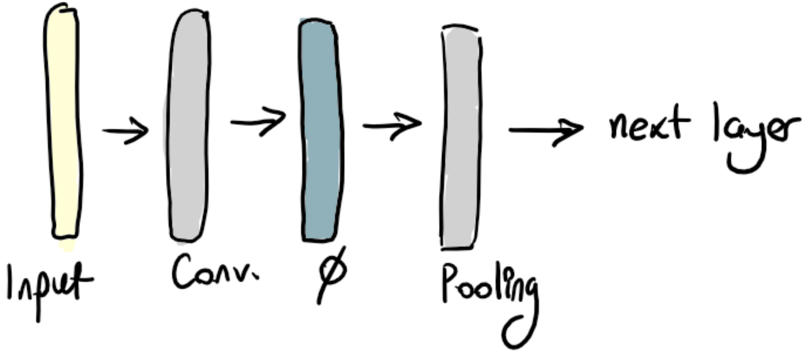
Under the hood: CNN layer

MLP:



CNN:

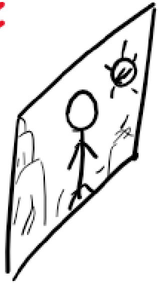
①



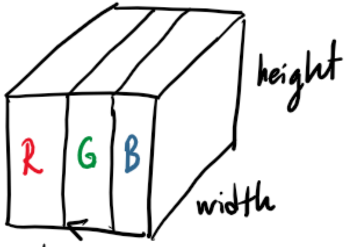
Under the hood: CNN layer

Conv. Layer * "Grid-like" / spatial structure

Eg.:



2D image



$\underbrace{\text{depth}}_{}$

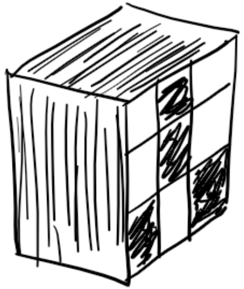
features

(224 x 224 x 3)

batch

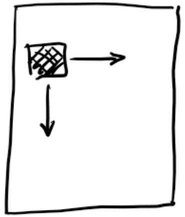
[None, 224, 224, 3]

app. conv.



- 64 kernels (filters)
- 64 features

} (3x3)



How to move the kernel?

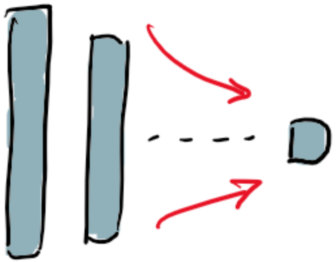
$W \Rightarrow 224 - 3 + 1 = 222$

$H \Rightarrow 224 - 3 + 1 = 222$

Under the hood : CNN layer

Conv. Layer * "Grid-like" // spatial structure

Issue :



spatial space gets smaller ⟹ limit the graph-depth

Solution : padding

6	3	4	4	5	0	3
4	7	4	0	4	0	4
7	0	2	3	4	5	2
3	7	5	0	3	0	7
5	8	1	2	5	4	2
8	0	1	0	6	0	0
6	4	1	3	0	4	5

PAD ⟹

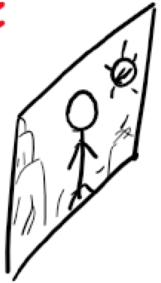
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	6	3	4	4	5	0	3	0	0
0	0	4	7	4	0	4	0	4	0	0
0	0	7	0	2	3	4	5	2	0	0
0	0	3	7	5	0	3	0	7	0	0
0	0	5	8	1	2	5	4	2	0	0
0	0	8	0	1	0	6	0	0	0	0
0	0	6	4	1	3	0	4	5	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

* valid ⟹ no padding
* same ⟹ keep dim. same

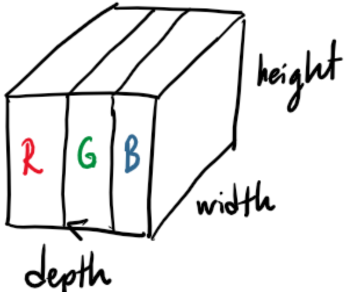
Under the hood: CNN layer

Conv. Layer ✱ "Grid-like" / spatial structure

Eg.:



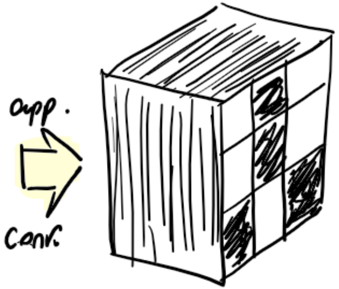
2D image



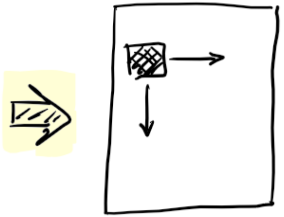
features

(224 x 224 x 3)

batch [None, 224, 224, 3]



- 64 kernels (filters) } (3x3)
- 64 features



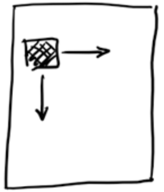
How to move the kernel?

Can I skip some px.?

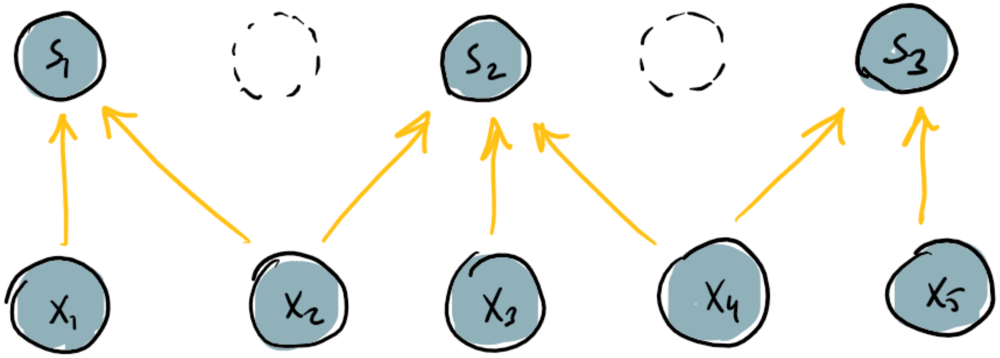
Under the hood: CNN layer

Conv. Layer

* Stride: granularity of conv. operation



How to move the kernel?



- $s > 2$ is rare
- $s = 1$; $s = 2$

$s = 2$

$k = 3$

Output: $(224 - 3)/2 + 1$

Under the hood: CNN layer

Pooling

* Obj: subsample the input data → comp. load, memory, # parameters

* Replace the output with summary statistics

- [x] max.
- [x] Mean
- [x] l_2 norm.
- [x] weighted operations (eg. distance based)

The pooling operation used in convolutional neural networks is a big mistake, and the fact that it works so well is a disaster.

Geoffrey Hinton

Under the hood: CNN layer

Pooling

Max Pooling

$$I = \begin{pmatrix} 6 & 3 & 4 & 4 & 5 & 0 & 3 \\ 4 & 7 & 4 & 0 & 4 & 0 & 4 \\ 7 & 0 & 2 & 3 & 4 & 5 & 2 \\ 3 & 7 & 5 & 0 & 3 & 0 & 7 \\ 5 & 8 & 1 & 2 & 5 & 4 & 2 \\ 8 & 0 & 1 & 0 & 6 & 0 & 0 \\ 6 & 4 & 1 & 3 & 0 & 4 & 5 \end{pmatrix}$$

$\boxed{(3\times 3),\ s=1}$

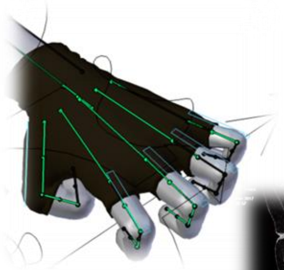
$$\Rightarrow \begin{bmatrix} 7 & 7 & 5 & 5 & 5 \\ 7 & 7 & 5 & 5 & 7 \\ 8 & 8 & 5 & 5 & 7 \\ 8 & 8 & 6 & 6 & 7 \\ 8 & 8 & 6 & 6 & 6 \end{bmatrix}$$

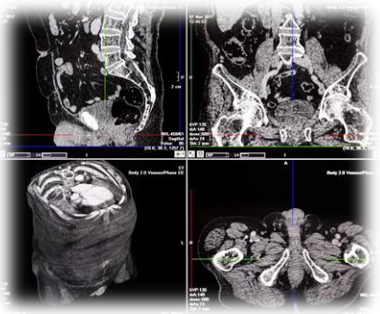
$\boxed{3\times 3,\ s=2}$

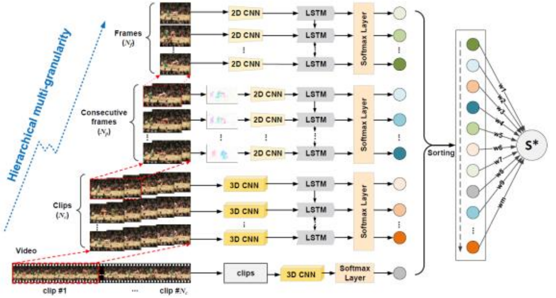
$$\begin{bmatrix} 7 & 5 & 5 \\ 8 & 5 & 7 \\ 8 & 6 & 6 \end{bmatrix}$$

Applications of CNN :

#	Single feature	Multifeature
1D	Audio signal Univariate time-series	Joint - movement Multiv. time-series
2D	$\underbrace{\text{Audio + FFT}}_{(\nu,\, t)}$ Images (intensity)	Color Image Analysis — RGB — St. Transport Phen.
3D	Medical imaging Volumetric data	Color - video analy. Tr. Transport Phen.







Thread "Interesting CNN projects"

Actions

All Topics | Sort by Posts | Order by Date

Add Posting

Cihan Ates | em0787 | 04. May 2021, 14:41
Edited on: 04. May 2021, 14:43 - by Cihan Ates | em0787

Interesting CNN projects

Dear all,

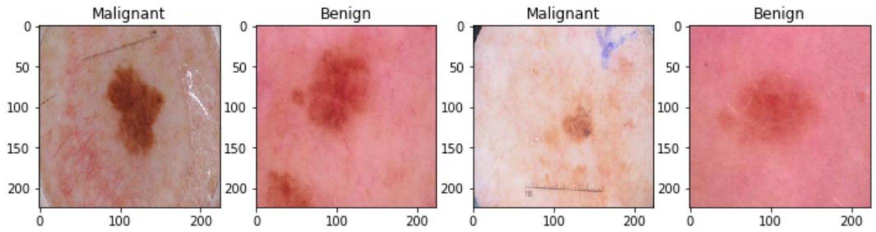
Below you will find links to some interesting / useful CNN projects. I may update the list later on:

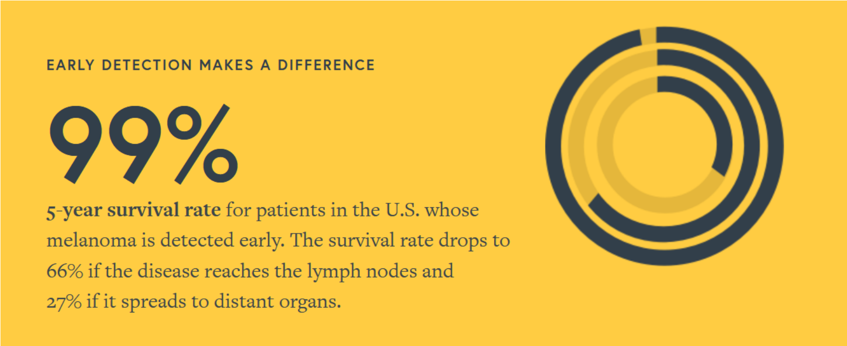
Build a CNN network to predict 3D bounding box of car from 2D image:

food image to recipe with deep convolutional neural networks

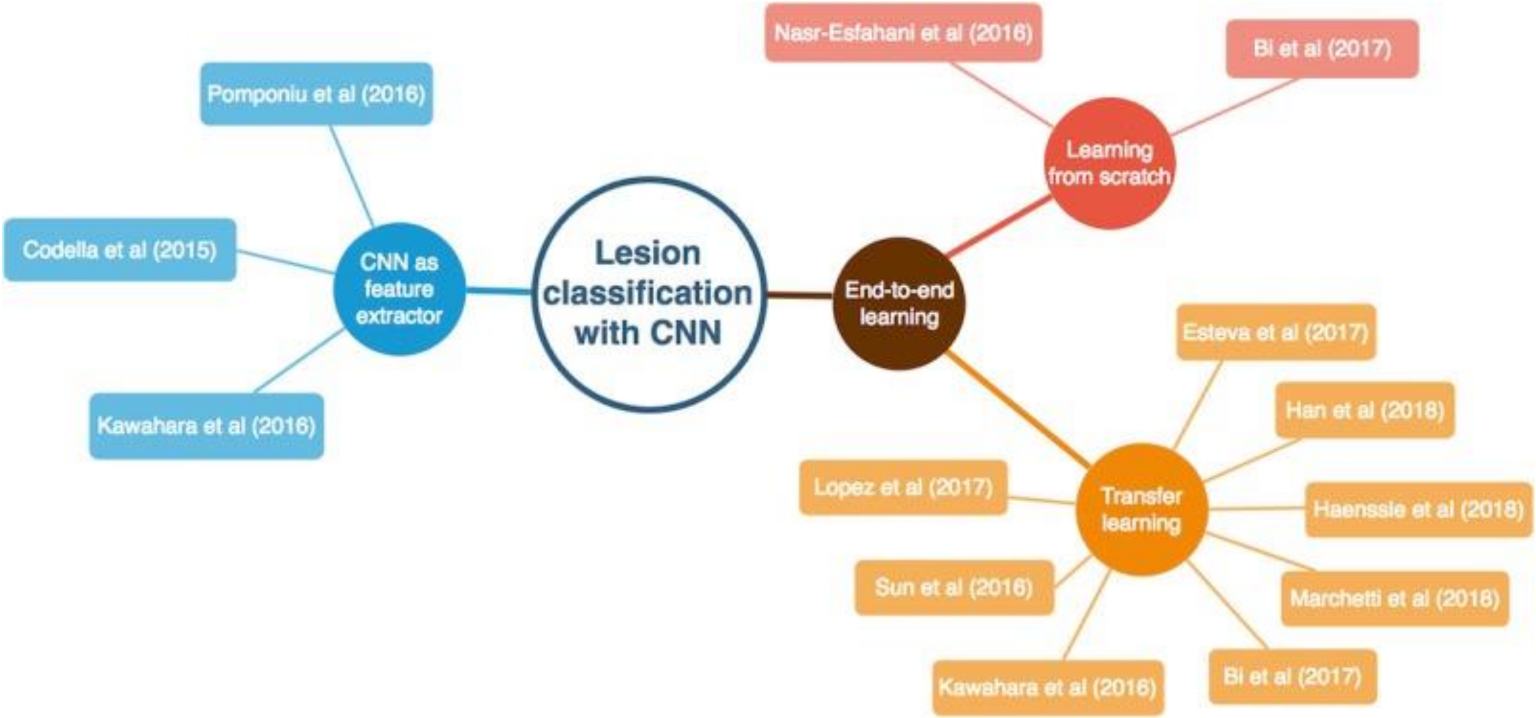
Case study: Melanoma classification

- Annual case ~ ↑ 53% btw. 2008 - 2018
  ↳ UV exposure
- Diagnosis → visual examination by a dermatologist
  ↳ 65 ~ 80% accuracy
- CNN for binary classification

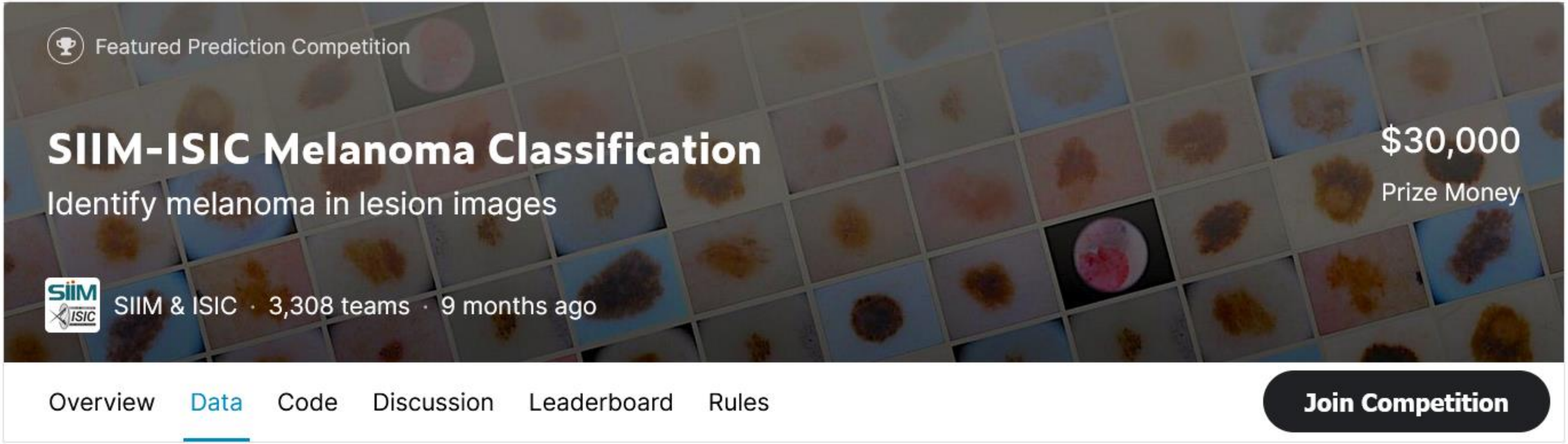




Case study: Melanoma classification



an updated database



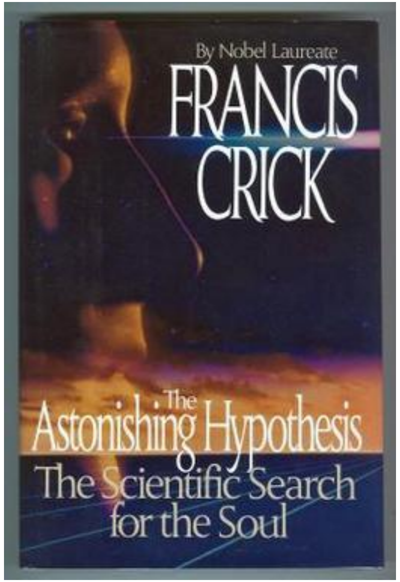
Outline of a ML Project

Basic Steps to Follow:

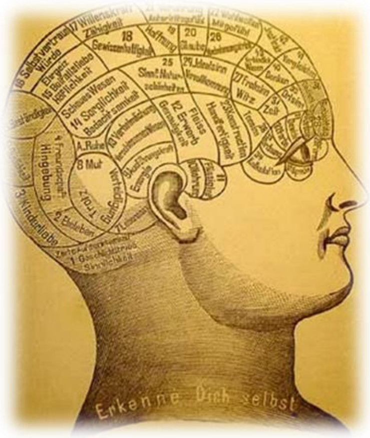
0.) Understand the business/task.
1.) Understand the data.
2.) Explore & prepare the data.
3.) Shortlist candidate models.
4.) Training the model
5.) Evaluate the model predictions.
6.) "Serve" the model!

"Classification"

colab



A Walk-through of the Mammalian Visual System



Imagery Debate: The Role of the Brain