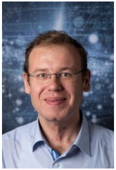
Introduction

A. Maier, V. Christlein, K. Breininger, Z. Yang, L. Rist, M. Nau, S. Jaganathan, C. Liu, N. Maul, L. Folle, K. Packhäuser, M. Zinnen
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
October 1, 2024

Who are we? - Lab Members



Andreas
Maier

Chang
Liu

Alexander
Barnhill

Zijin
Yang

Leonhard
Rist

Merlin
Nau

Noah
Maul

Mathias
Zinnen

Who are we? - Student Tutors

Lukas Hüttner

Majid Sharghi

Mohannad Barakat

Christian Wielenberg

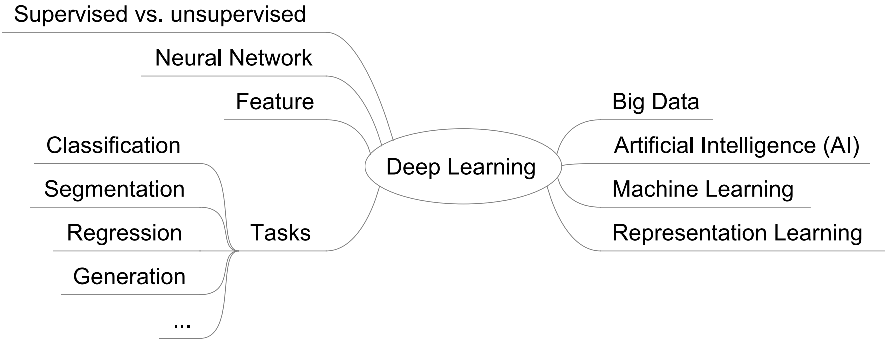
Haiting Huang

Julian Greil

Thakkar Rahul Jayantilal

Mohapatra Maitreya

Deep Learning – Buzzwords



Outline

Motivation

NVIDIA Stock Market

Source: https://www.google.com/finance/quote/

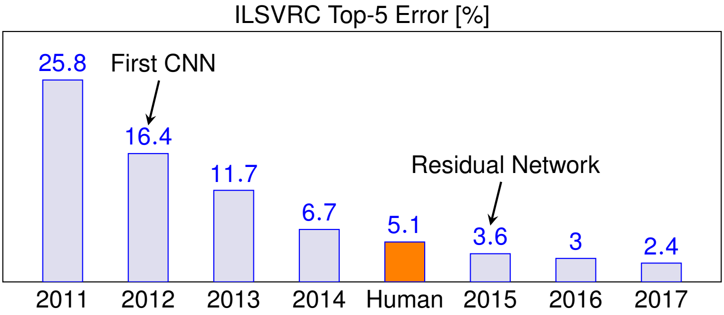
The Big Bang of Deep Learning

ImageNet imagenet Dataset

- ≈ 14 mio. images, labeled into ≈ 20.000 **synonym sets**
- **I**mageNet **L**arge **S**cale **V**isual **R**ecognition **C**hallenge using ≈ 1000 classes
- Images downloaded from the Internet, **single** label per image
- **2012: Breakthrough** by Krizhevsky et al. **Krizhevsky12**

Source: Krizhevsky et al. 2012

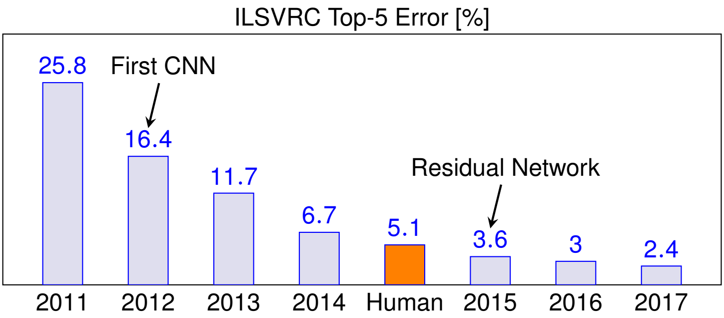
ImageNet Large Scale Visual Recognition Challenge



- First CNN approach now famous as **AlexNet Krizhevsky12**

Source: image-net.org, Russakovsky et al. 2015

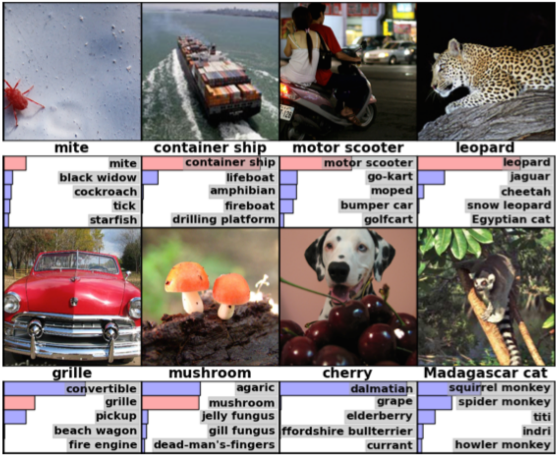
ImageNet Large Scale Visual Recognition Challenge



- First CNN approach now famous as **AlexNet Krizhevsky12**
- “Superhuman” should be Super-Karpathy-an performance

Source: image-net.org, Russakovsky et al. 2015

ImageNet Large Scale Visual Recognition Challenge



Source: Krizhevsky et al. 2012

Deep Learning Users

NETFLIX

Lunit

DAIMLER

SIEMENS

IBM

Google

DeepMind

xerox

SAMSUNG

Microsoft

Playing Go

- **1997**: Deep Blue beats Garry Kasparov
- **Go** as a next challenge
- Large branching factor

Source: https://commons.wikimedia.org/wiki/File:FloorGoban.jpg

Playing Go

- **1997**: Deep Blue beats Garry Kasparov
- **Go** as a next challenge
- Large branching factor
- **2016**: AlphaGo **Silver2016** beats a professional

Source: https://commons.wikimedia.org/wiki/File:FloorGoban.jpg

Playing Go

- **1997**: Deep Blue beats Garry Kasparov
- **Go** as a next challenge
- Large branching factor
- **2016**: AlphaGo **Silver2016** beats a professional
- **2017**: AlphaGoZero **AlphaGoZero** surpasses every human in Go by self-play
- **2017**: AlphaZero **AlphaZero** generalizes to a number of other board games

Source: https://commons.wikimedia.org/wiki/File:FloorGoban.jpg

Playing Go

- **1997**: Deep Blue beats Garry Kasparov
- **Go** as a next challenge
- Large branching factor
- **2016**: AlphaGo **Silver2016** beats a professional
- **2017**: AlphaGoZero **AlphaGoZero** surpasses every human in Go by self-play
- **2017**: AlphaZero **AlphaZero** generalizes to a number of other board games
- **2019**: AlphaStar beats professional StarCraft players

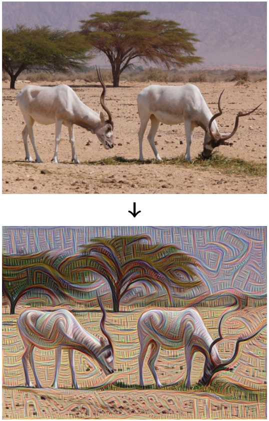
Source: https://commons.wikimedia.org/wiki/File:FloorGoban.jpg

Google DeepDream

Attempt to understand the inner workings of the network: What it "dreams" about when presented with images

Idea:

- Arbitrary image or noise as input
- Instead of adjusting network parameters, tweak image towards high activations
- Different layers enhance different features (low or high level)

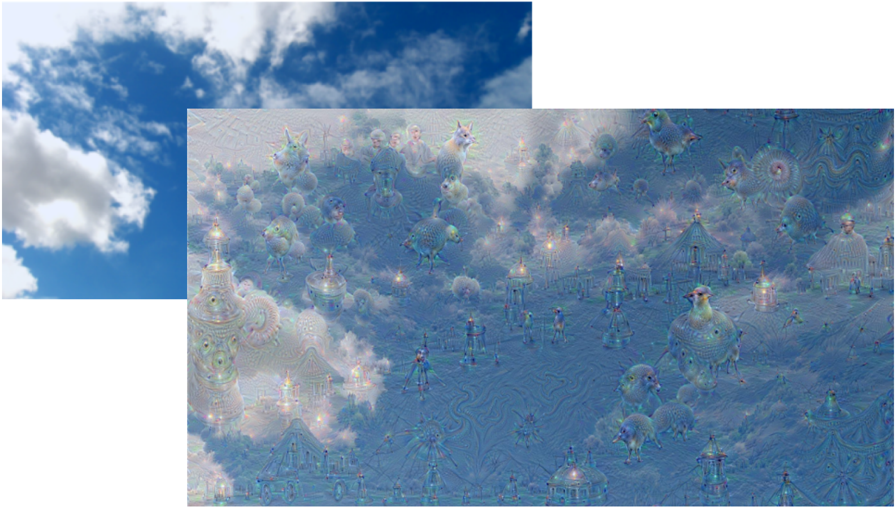


Source: https://research.googleblog.com

Google DeepDream

Source: https://research.googleblog.com

Google DeepDream



Source: https://research.googleblog.com

Google DeepDream

Looking for new animals in the clouds

"Admiral Dog!"

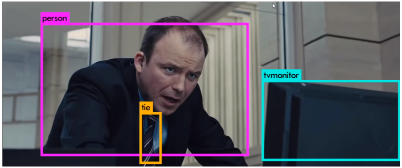
"The Pig-Snail"

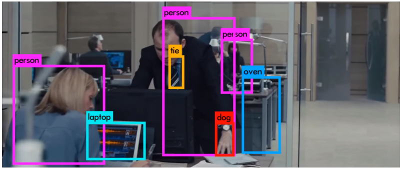
"The Camel-Bird"

"The Dog-Fish"

Source: https://research.googleblog.com

Real-Time Object Detection: YOLO, YOLO9000, YOLOv3 Redmon15-YOL, Redmon16-YOL, Redmon18-YOL





Click for video

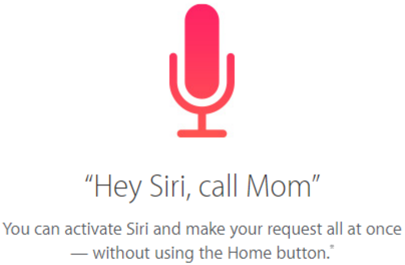
- YOLO: You only ~~live~~ look once
- Prior systems ➔ Use classifiers at multiple locations and scales
- YOLO ➔ Simultaneous regression of bounding box and label
- FAST: 40-90 frames/second on a NVIDIA Titan X

Source: www.youtube.com, Redmon and Farhadi 2016

Every Day Use

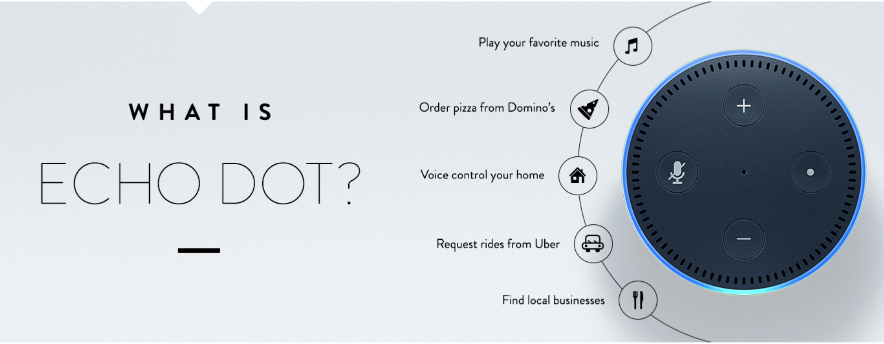
Siri

Siri: Speech Interpretation and Recognition Interface



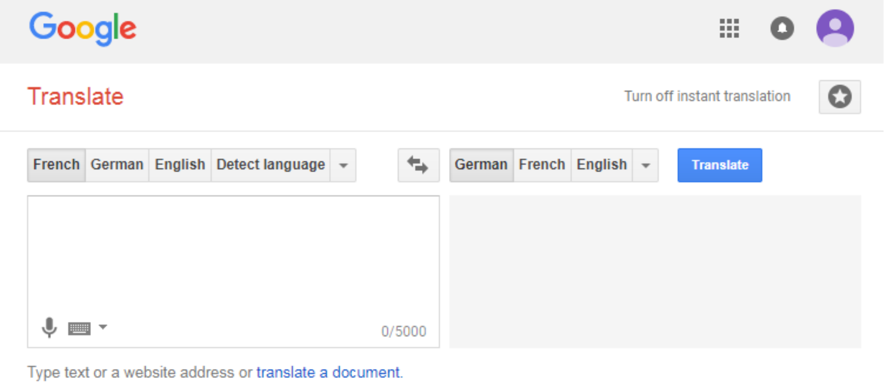
Source: www.apple.com/ios/siri/

Google Echo & Amazon Alexa Voice Service



Source: www.amazon.com

Google Translate



Source: translate.google.de

NEXT TIME

ON DEEP LEARNING

Introduction - Part 2

A. Maier, V. Christlein, K. Breininger, Z. Yang, L. Rist, M. Nau, S. Jaganathan, C. Liu, N. Maul, L. Folle, K. Packhäuser, M. Zinnen
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
October 1, 2024

FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

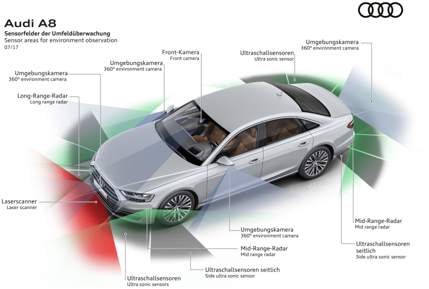
Research at the Pattern Recognition Lab

Assisted and Automated Driving

Goal

Find new ways to train and update deep learning mechanisms in environments with high safety requirements

- Assisted and automatic driving relies on sensor data
- Cameras to detect dynamic objects, driving lanes and free space
- Detection and segmentation tasks
  ➔ deep learning



Source: Audi AG

Assisted and Automated Driving

- Currently: neural networks trained and thoroughly tested before deployment

➔ Requires huge amounts of manually labeled data

- Regular test drives cannot verify system reliability in all traffic scenarios

Click for video

Source: Mobileye N.V.

Assisted and Automated Driving

- Currently: neural networks trained and thoroughly tested before deployment

➔ Requires huge amounts of manually labeled data

- Regular test drives cannot verify system reliability in all traffic scenarios
- **Challenge**: New ways to test algorithms in simulated environments and utilize data collected in production cars equipped with appropriate hardware

Click for video

Source: Mobileye N.V.

Smart Devices

Problem statement

Renewable energy power $\neq$ energy demand

- Underproduction ➔ backup power plants
- Overproduction ➔ energy lost

➔ Real-Time-Pricing to match energy demand and supply

- Needs *smart devices* to shift workload automatically

Smart Devices

Goal

Establish energy equilibrium by predicting energy consumption

- Example: Interrupt fridge cooling cycle when price is high, start washing machine when price is low
- Dependencies between tasks, user information and action necessary (e.g., washer/dryer)
- Task: Identify time-shiftable loads and assess appropriate time frame
- Approach: Train **recurrent neural networks** to identify usage patterns and dependencies between devices

Cloud Detection for Power Forecast Bernecker14-CST

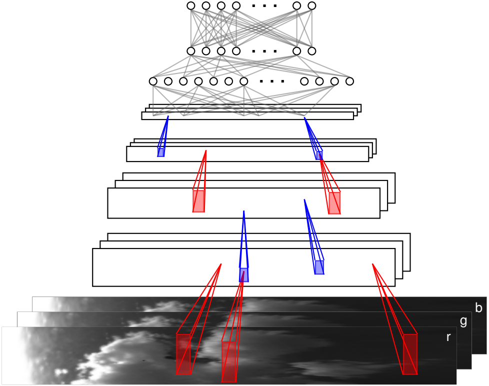
Goal

Power forecast for solar power plants with a high temporal and spatial resolution

Approach

1. Monitor the sky
2. Detect clouds
3. Estimate the cloud motion
4. Establish power forecasts

Cloud Detection for Power Forecast Bernecker14-CST



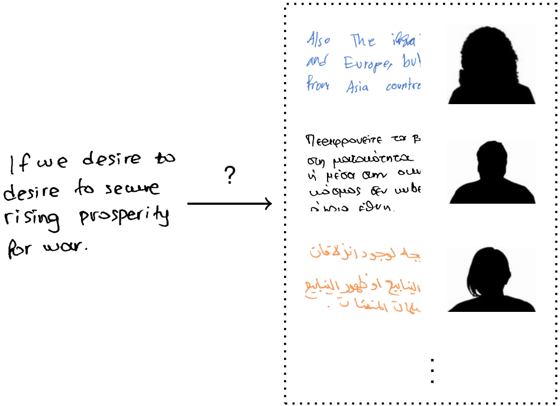
Input: Sky moving towards the sun

Output: Clear Sky Index = values betw. 0 (overcast sky) to 1 (clear sky)

Writer Recognition

Goal

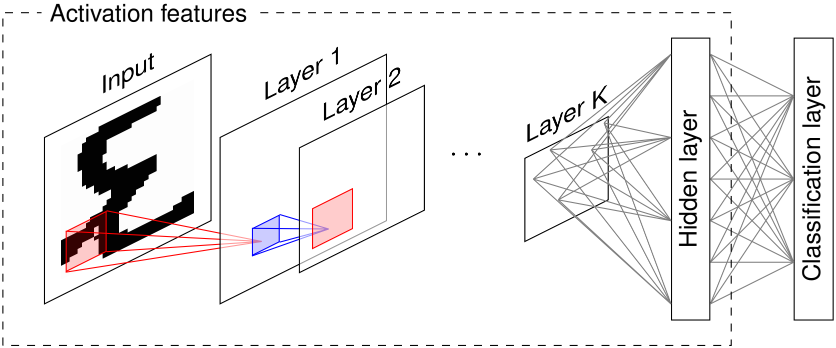
Writer identification with limited training data (few pages per writer)



Source: ICDAR'13 dataset, QUWI'15 dataset, freepik.com

Writer Recognition using CNN Activation Features Christlein17_WIG

Use Neuronal Network for feature extraction



Medical Applications

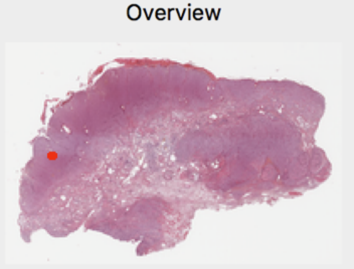
Cell Classification for Tumor Diagnostics Aubreville17-GST

Goal

Identify cells undergoing mitosis to asses tumor proliferation and aggressiveness in histological images

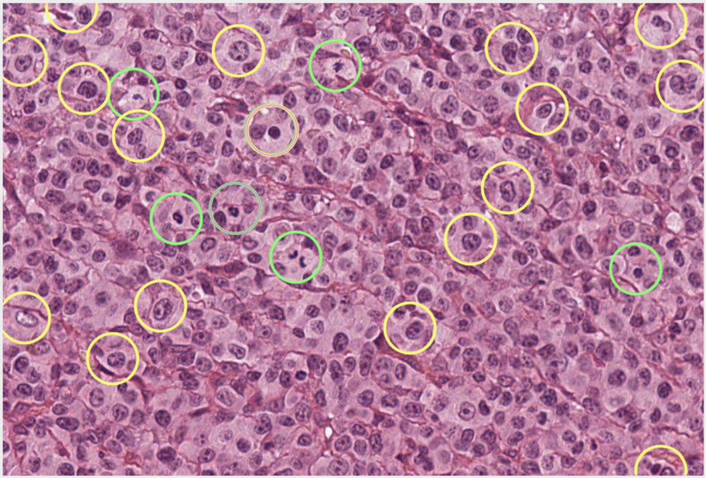
Challenge

- Histological images: large number of cells
- Full annotations not feasible
- Sparse annotations
- Cells vary significantly in size/shape/etc



Source: Aubreville et al. 2017

Cell Classification for Tumor Diagnostics Aubreville17-GST

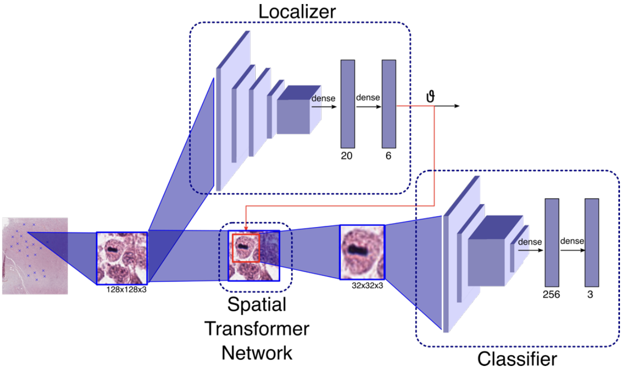


Source: Aubreville et al. 2017

Cell Classification for Tumor Diagnostics Aubreville17-GST

Approach

Use *spatial transformer networks* (STNs) to learn affine transformation **and** classification



Source: Aubreville et al. 2017

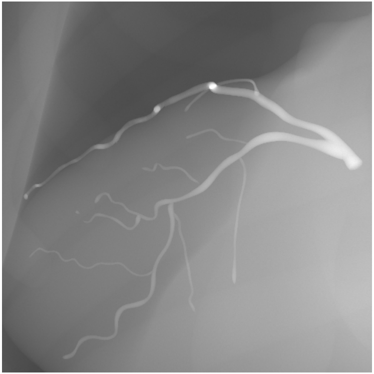
Defect Pixel Interpolation

Goal

- Reconstruction of coronaries based on truncated X-ray images
- Create “virtual” digital subtraction angiography

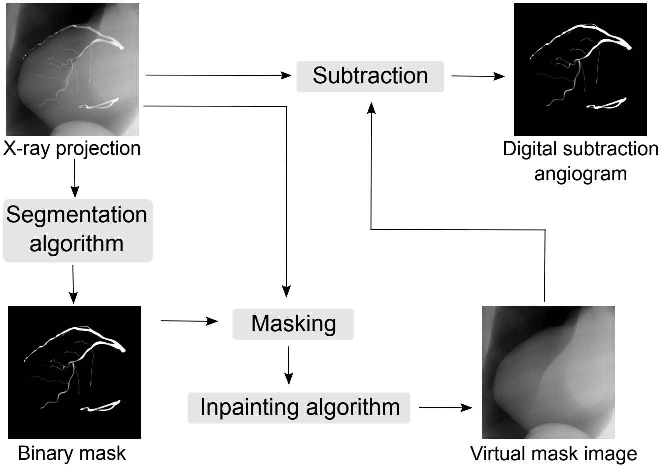
Approach

1. Segment coronary vessels
2. Mask fluoroscopic image
3. Inpaint using U-net
4. Subtract inpainted image to get untruncated data



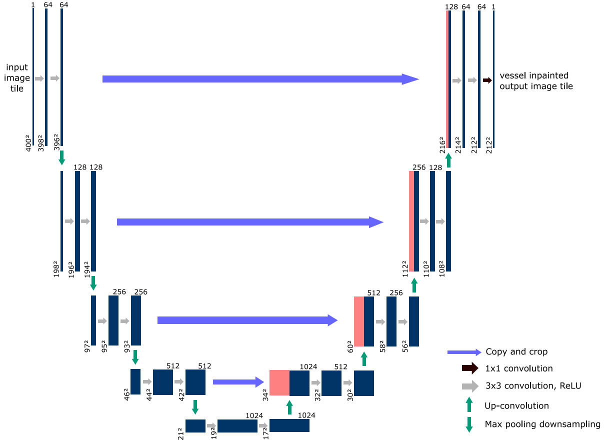
Defect Pixel Interpolation

Processing pipeline



Defect Pixel Interpolation

Deep learning for inpainting



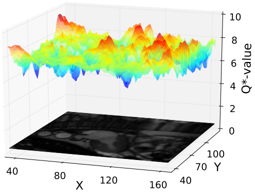
Organ Search Ghesu16-AAA

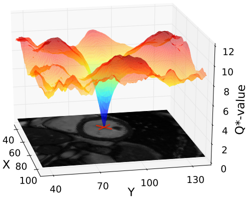
Goal

Locate anatomic structures automatically

Approach

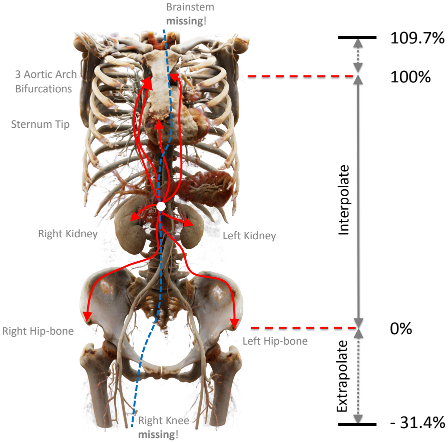
- Deep reinforcement learning
- Learn strategies how to search objects
  ➔ Learn optimal shortest search through image volume to different landmarks
- Hierarchical approach to improve speed and robustness





Source: Ghesu et al. 2016, Ghesu et al. 2017

Organ Search Ghesu16-AAA



3D search trajectories
starting from image center

Body Range

Organ Search Ghesu16-AAA

X-ray-transform Invariant Anatomical Landmark Detection

Goal

- Detect landmarks in X-ray images
- Knowing correspondences enables symbolic reconstruction

➔ Classic computervision reconstruction

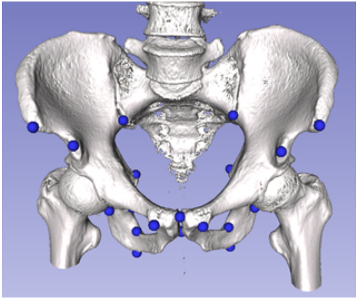
Challenge

- Transmission imaging

➔ Overlap/superposition of structures

➔ High variance due to projection

➔ Artifacts e.g. interventional devices



Source: Bier et al. 2018

X-ray-transform Invariant Anatomical Landmark Detection

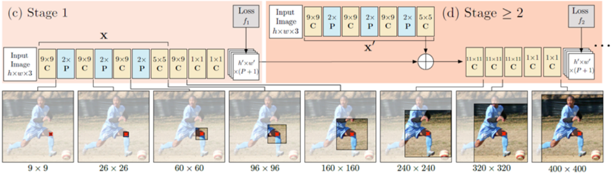
Approach: Convolutional Pose Machine (CPM) Wei2016ConvolutionalPoseMachine

- Sequential prediction framework to detect landmarks

➔ Yields 2D belief maps

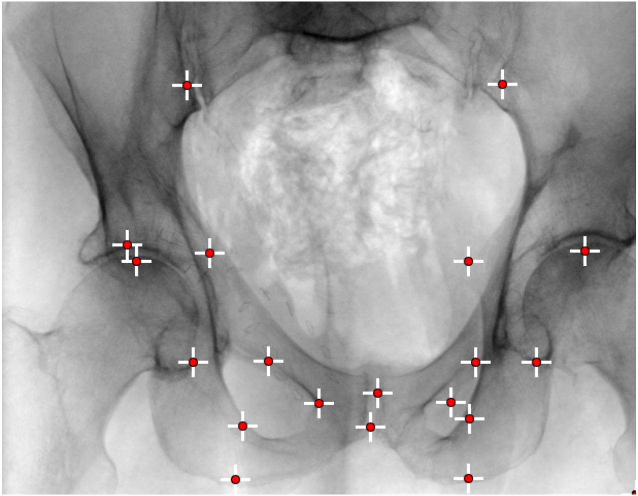
Properties

- Large receptive fields enable learning of configurations
- Estimation is refined over stages



Source: Wei et al. 2016

X-ray-transform Invariant Anatomical Landmark Detection

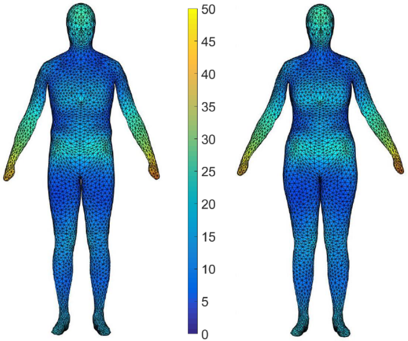


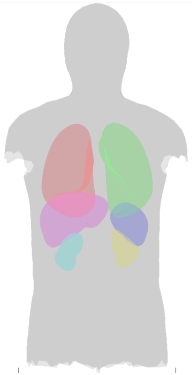
Source: Bier et al. 2018

Organ Prediction

Goal

Estimation of body and organ shapes based on patient's height and weight for X-ray exposure estimation.

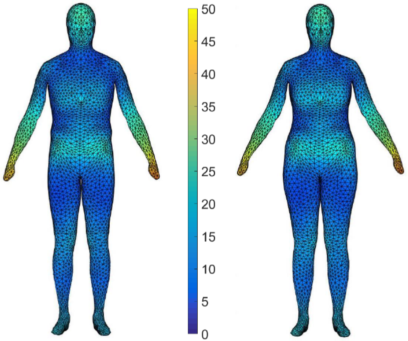


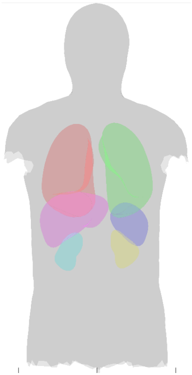


Organ Prediction

Goal

Estimation of body and organ shapes based on patient's height and weight for X-ray exposure estimation.





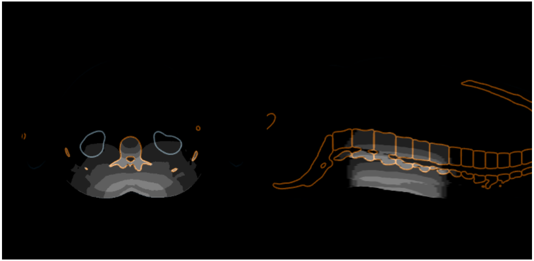
Could we achieve more if we had old CT data of a patient?

Action Learning for 3D Point Cloud Based Organ Segmentation

Goal: Versatile organ segmentation for:

- Use it in computer aided diagnosis
- Treatment planning
- Dose management

Dose estimation in interventions with overlays

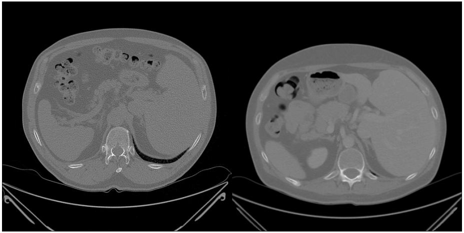


Action Learning for 3D Point Cloud Based Organ Segmentation

Challenges for clinical applications

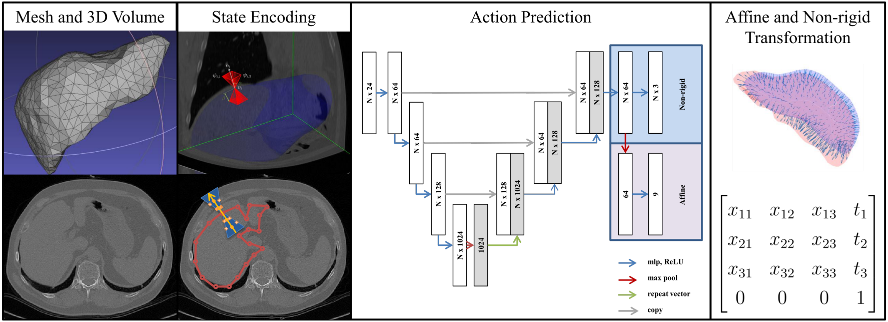
- Robustness w.r.t.
  1. Individual anatomy
  2. Scan protocols
- Time constraints

Pre-operative CT (left) and contrast enhanced CT (right)



Action Learning for 3D Point Cloud Based Organ Segmentation

- Reinforcement learning
- Predict the transformation at given state



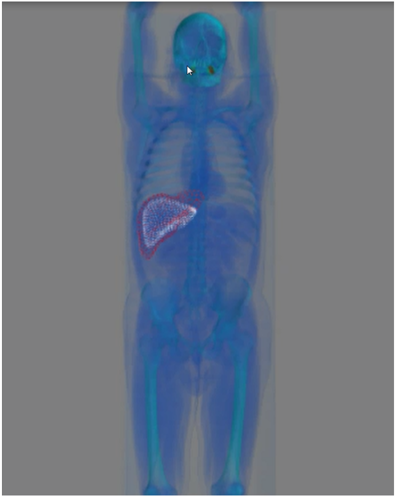
$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & t_1 \\ x_{21} & x_{22} & x_{23} & t_2 \\ x_{31} & x_{32} & x_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Figure: Action prediction pipeline for 3D point cloud based organ segmentation

Source: Zhong et al. 2018

Action Learning for 3D Point Cloud Based Organ Segmentation

- Runtime:
  1. **0.3 - 2.6s per volume**
  2. **50 - 100 speedup** from U-net **christ2016automatic**
- Very accurate
- Robust to:
  1. scan protocol
  2. contrast agent
  3. organ initialization



Source: Zhong et al. 2018

NEXT TIME

ON DEEP LEARNING

Introduction - Part 3

A. Maier, V. Christlein, K. Breininger, Z. Yang, L. Rist, M. Nau, S. Jaganathan, C. Liu, N. Maul, L. Folle, K. Packhäuser, M. Zinnen
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
October 1, 2024

Limitations

Image Captioning

Figure: "baseball player is throwing ball in game."

Figure: "girl in pink dress is jumping in air."

Figure: "man in black shirt is playing guitar."

Source: http://cs.stanford.edu/people/karpathy/deepimagesent

Image Captioning & Generating

GPT4: “An athletically built man with a serious expression stands in a well-lit room with a wooden ceiling. He is wearing a blue t-shirt with the phrase ’SAVE THE CHUBBY UNICORNS’ and an image of a rhinoceros. The man is taking a selfie with a modern smartphone, capturing his reflection in a mirror. In the background, one can see gym equipment including dumbbells and a weight rack. There is also a hanging chandelier with a unique design, illuminating the room”

Source: https://twitter.com/maier_ak/status/1713563211574751494

Image Captioning & Generating

Now DALL-E3, "Create an image for this prompt in different styles"

Figure: "version 1."

Figure: "version 2."

Figure: "version 3."

Source: https://twitter.com/maier_ak/status/1713563211574751494

Image Captioning & Generating

To DALL-E3, "Can you create a few pixel art versions of this?"

Figure: "pixel version 1."

Figure: "pixel version 2."

Figure: "pixel version 3."

Source: https://twitter.com/maier_ak/status/1713563211574751494

Image Captioning & Generating

To DALL-E3, "Show the scene in the style of a massive online roleplaying game"

Figure: "warcraft version 1."

Figure: "warcraft version 2."

Figure: "warcraft version 3."

Source: https://twitter.com/maier_ak/status/1713563211574751494

Image Captioning & Generating

To DALL-E3, "Make a few versions that show the scene in Lego Style"

Figure: "lego version 1."

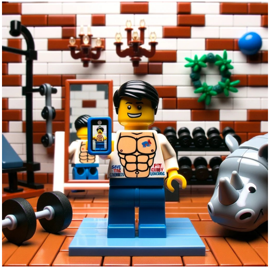


Figure: "lego version 2."

Figure: "lego version 3."

Source: https://twitter.com/maier_ak/status/1713563211574751494

Challenges with Training Data

- Deep learning applications often rely on **huge**, manually-annotated data sets
- Hard to obtain, time-consuming, expensive, ambiguous
- To err is human: Mislabeled ground-truth annotation

➔ May cause a significant drop in performance

Challenges with Training Data

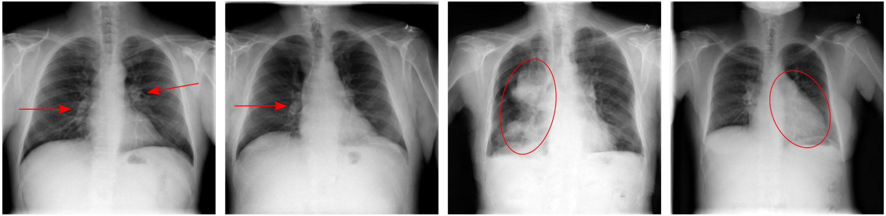
- Deep learning applications often rely on **huge**, manually-annotated data sets
- Hard to obtain, time-consuming, expensive, ambiguous
- To err is human: Mislabeled ground-truth annotation

➔ May cause a significant drop in performance

- Question: How far can we get with simulations?

Generating Synthetic Data

- Sample from trained latent diffusion model

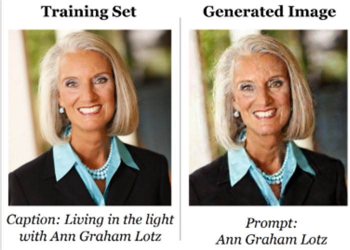


Infiltration Nodule Mass Cardiomegaly

Figure: four chest X-ray images sampled from a trained latent diffusion model. Image generation was done in a conditional way to produce images of specific abnormality classes. The induced abnormality patterns in the synthetic images are highlighted with red arrows and circles.

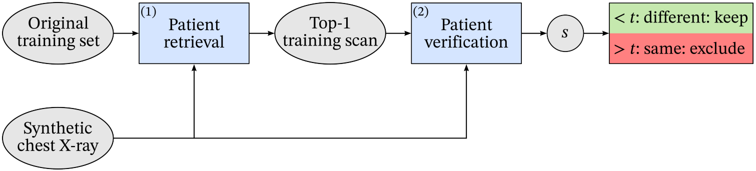
Source: Packhäuser et al. IEEE ISBI 2023

Memorization problem for diffusion models



- Problem: models in practice tend to reproduce patterns/complete images of the used training set.
- Proposed method: uses a pre-trained patient retrieval model to compare a created synthetic scan to all training images (see pipeline below)

Privacy-enhancing Image Sampling Strategy



Challenges with Trust and Reliability

- Verification is mandatory for high risk applications

➔ End-to-end learning prohibits verification of parts

➔ Largely unsolved

Challenges with Trust and Reliability

- Verification is mandatory for high risk applications

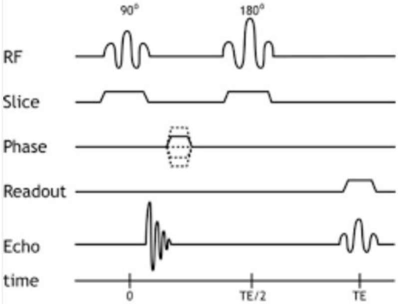
➔ End-to-end learning prohibits verification of parts

➔ Largely unsolved

- Possible solution: Reformulate classical algorithms

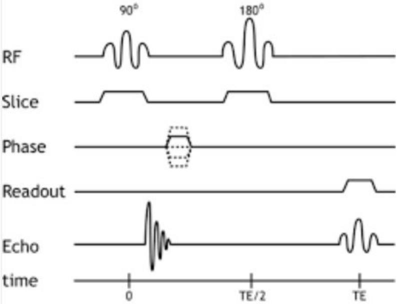
Large Language Models for MRI Scanners

to GPT: Can you implement the sequence show
in the given sequence diagram in pypulseq?



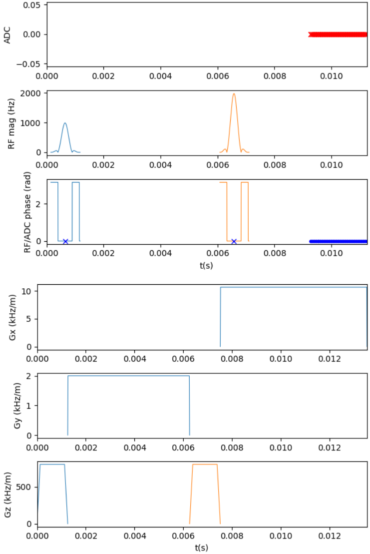
Large Language Models for MRI Scanners

to GPT: Can you implement the sequence show in the given sequence diagram in pypulseq?



GPT: Yes. (On the right)

- Still buggy, works only 50% of the time



Future Directions

Learning of Algorithms

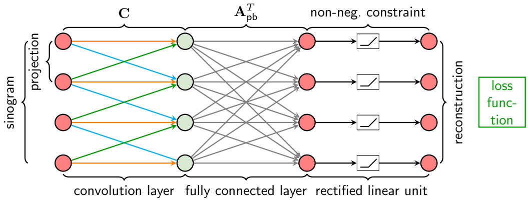
- Computed Tomography
- Efficient solution via filtered back-projection:

$$f(x,y) = \int_0^{\pi} p(s,\theta) * h(s)\big|_{s=x\cos\theta+y\sin\theta} d\theta$$

- Three steps:
  - Convolution along s
  - Back-projection along θ
  - Suppress negative values

Reconstruction Networks

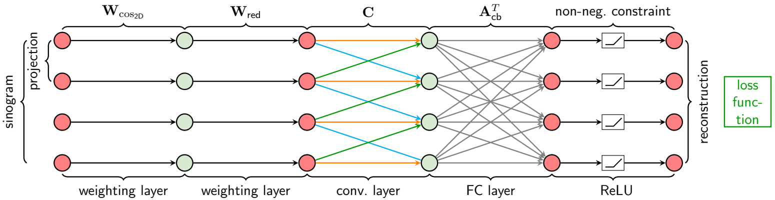
- All three steps can be modeled as a neural network:



- All weights are known from FBP

Reconstruction Networks

- Reconstruction Networks can be expanded



- Embedding of "heuristics" for artifact reduction possible

Application to Incomplete Scans Wuerfl16_DLC

Figure: Reconstruction with 360°

Application to Incomplete Scans Wuerfl16_DLC

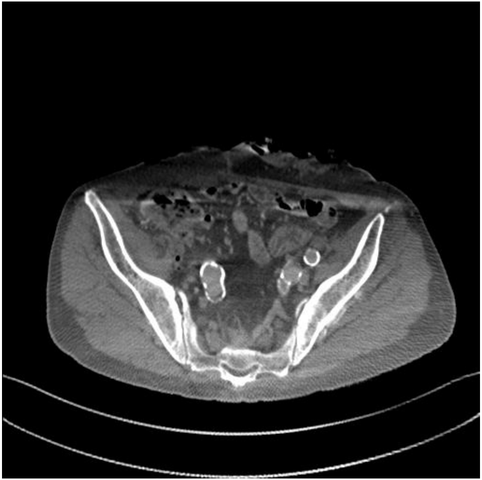


Figure: Reconstruction with 180° (FBP)

Application to Incomplete Scans Wuerfl16_DLC

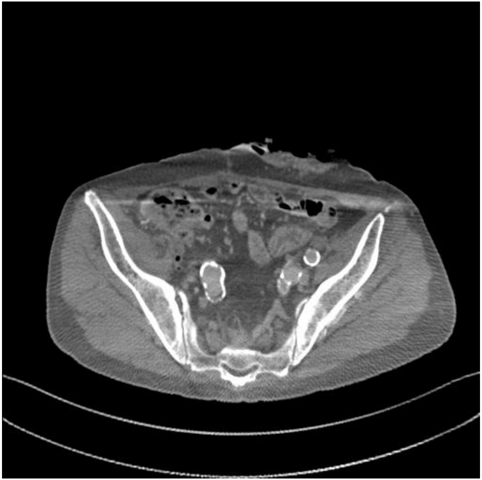


Figure: Reconstruction with 180° (NN)

Application to Incomplete Scans Wuerfl16_DLC

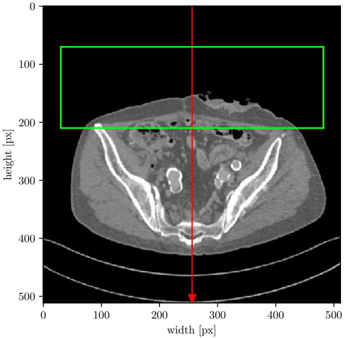


Figure: Location of the lineplot

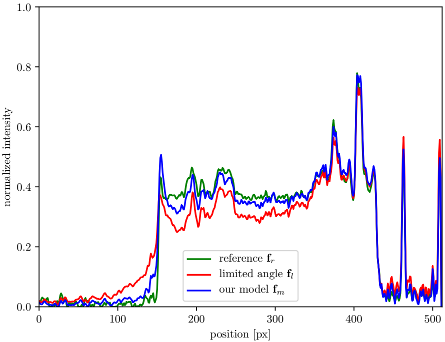


Figure: Lineplot

Parker Weights

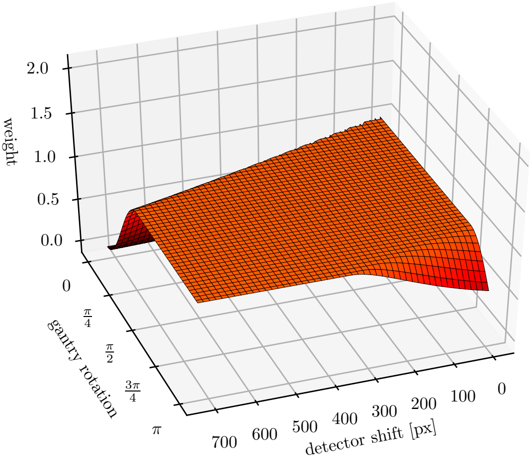


Figure: Parker weights before learning

Parker Weights

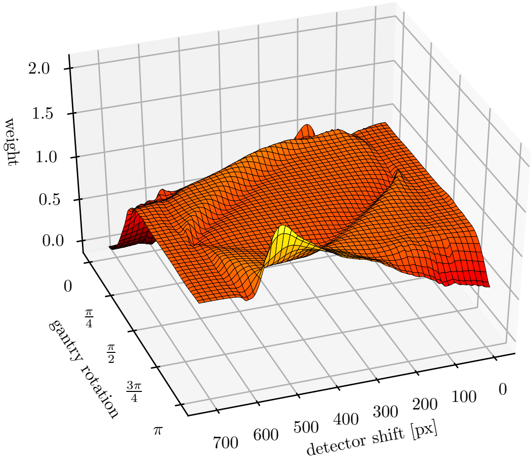
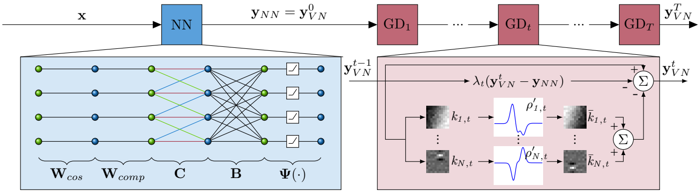


Figure: Parker weights after learning

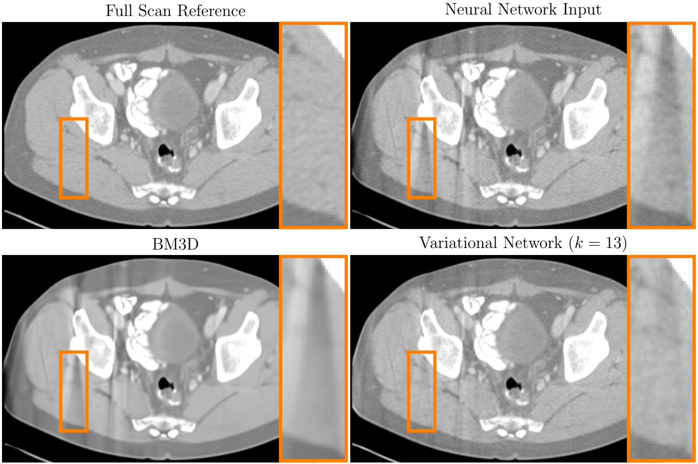
Further Extensions

- Add non-linear de-streaking and de-noising step:



Step 1: Neural network CT reconstruction **Step 2: Variational network non-linear filtering**

Further Extensions



NEXT TIME

ON DEEP LEARNING

Introduction - Part 4

A. Maier, V. Christlein, K. Breininger, Z. Yang, L. Rist, M. Nau, S. Jaganathan, C. Liu, N. Maul, L. Folle, K. Packhäuser, M. Zinnen
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
October 1, 2024

FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
SCHOOL OF ENGINEERING

Machine Learning and Pattern Recognition

Terminology and Notation

Throughout these slides, we will use the following notation:

- Matrices: bold, uppercase, e.g., $\mathbf{M}$, $\mathbf{A}$
- Vectors: bold, lowercase, e.g., $\mathbf{v}$, $\mathbf{x}$
- Scalars: italic, lowercase, e.g., y, w, α
- Gradient of a function: ∇, partial derivative: ∂

Terminology and Notation

Throughout these slides, we will use the following notation:

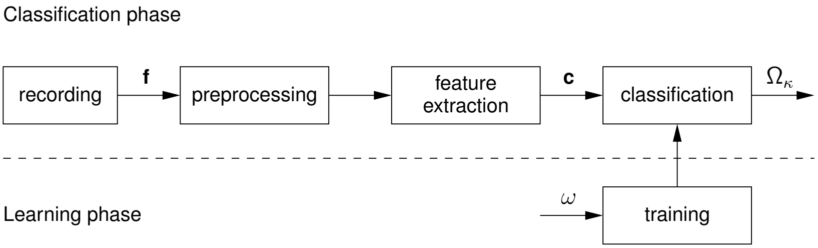
- Matrices: bold, uppercase, e.g., $\mathbf{M}$, $\mathbf{A}$
- Vectors: bold, lowercase, e.g., $\mathbf{v}$, $\mathbf{x}$
- Scalars: italic, lowercase, e.g., y, w, α
- Gradient of a function: ∇, partial derivative: ∂

Notation regarding deep learning:

- Trainable parameters ("weights"): w
- Features/input: $\mathbf{x}$
- Ground truth label/target: y
- Estimated output: $\hat{y}$
- Index denoting iteration will be in superscript, e.g., $\mathbf{x}^{(i)}$

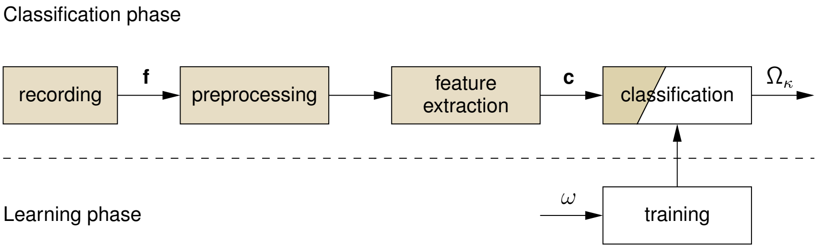
The notation and the terminology will be further developed throughout the lecture.

“Classical” Image Processing Pipeline



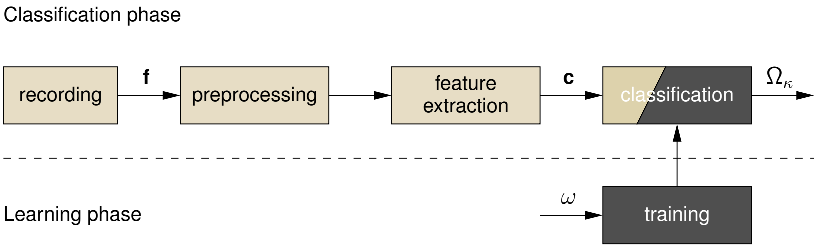
“Classical” Image Processing Pipeline

Lecture Introduction to Pattern Recognition



“Classical” Image Processing Pipeline

Lecture Introduction to Pattern Recognition



Lecture Pattern Recognition

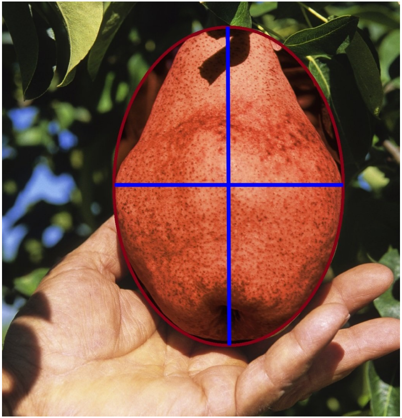
“Classical” Image Processing Pipeline: Apple vs. Pears

Source: https://commons.wikimedia.org

“Classical” Image Processing Pipeline: Apple vs. Pears

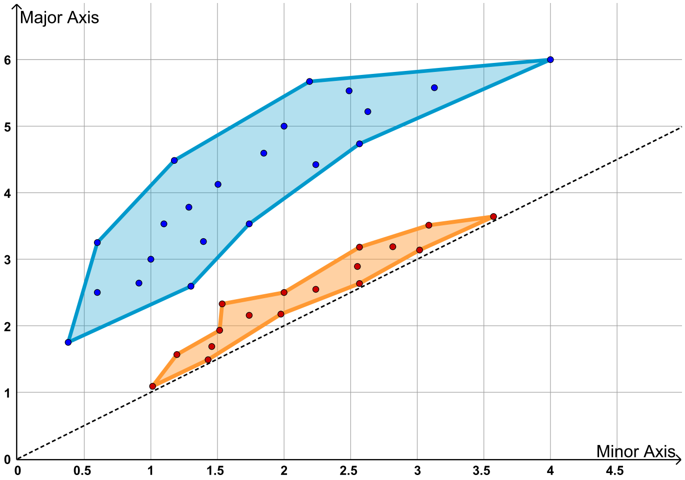
Source: https://commons.wikimedia.org

“Classical” Image Processing Pipeline: Apple vs. Pears

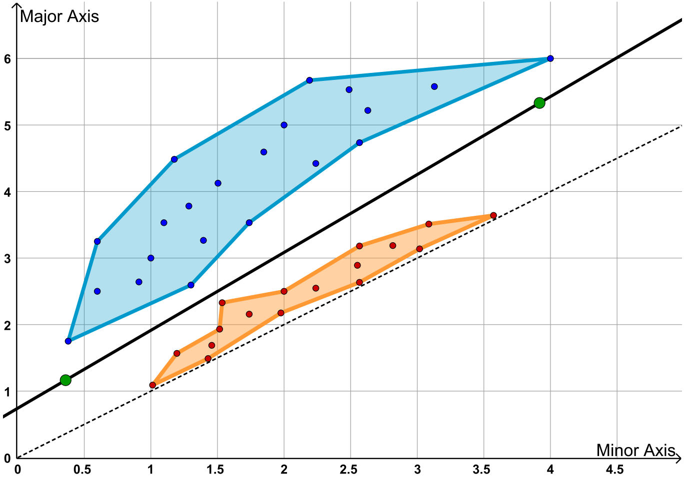


Source: https://commons.wikimedia.org

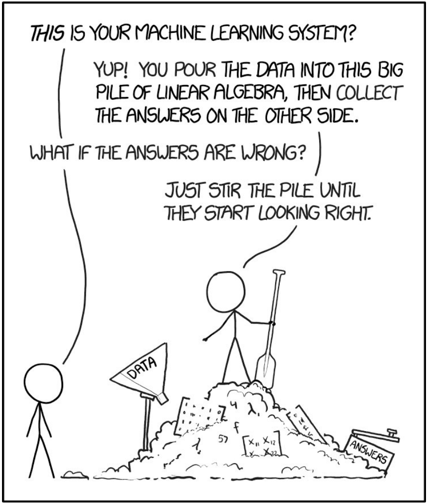
“Classical” Image Processing Pipeline: Apple vs. Pears



“Classical” Image Processing Pipeline: Apple vs. Pears



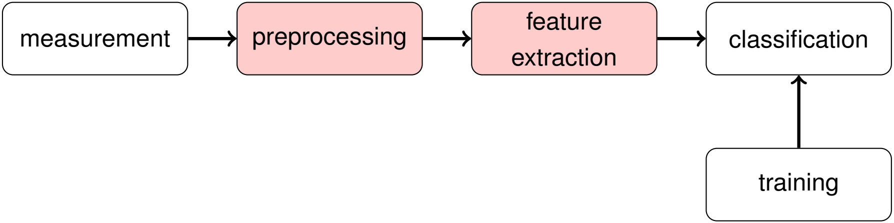
Pipeline in Deep Learning



Source: https://xkcd.com/1838/

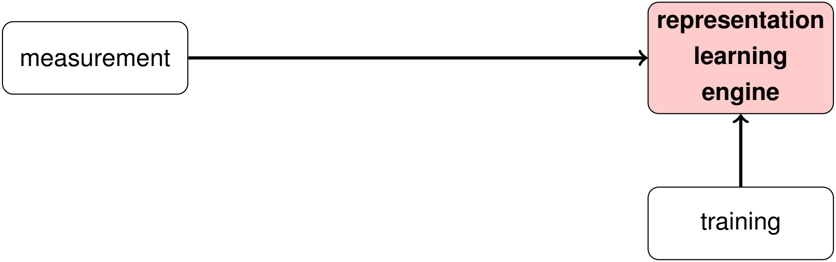
Pipeline in Deep Learning

Reminder



Pipeline in Deep Learning

Now



Postulates for Pattern Recognition

6 Postulates:

1. Availability of a **representative sample** ω of **patterns** ${}^{i}\mathbf{f}(\mathbf{x})$ for the given field of problems Ω

$$\omega = \{{}^{1}\mathbf{f}(\mathbf{x}), \ldots, {}^{N}\mathbf{f}(\mathbf{x})\} \subseteq \Omega.$$

Postulates for Pattern Recognition

6 Postulates:

1. Availability of a **representative sample** ω of **patterns** ${}^{i}\mathbf{f}(\mathbf{x})$ for the given field of problems Ω

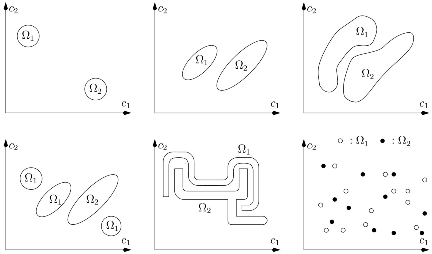
$$\omega = \{{}^{1}\mathbf{f}(\mathbf{x}), \ldots, {}^{N}\mathbf{f}(\mathbf{x})\} \subseteq \Omega.$$

2. A (simple) pattern has **features**, which characterize its membership in a certain class Ω_κ.

Postulates for Pattern Recognition (cont.)

3. Compact domain of features of the same class; domains of different classes are (reasonably) separable.
   - small **intra-class distance**
   - high **inter-class distance**

Example of an increasingly less compact domain in the feature space:



Postulates for Pattern Recognition (cont.)

4. A (complex) pattern consists of **simpler constituents**, which have certain relations to each other. A pattern may be decomposed into these constituents.

Postulates for Pattern Recognition (cont.)

4. A (complex) pattern consists of **simpler constituents**, which have certain relations to each other. A pattern may be decomposed into these constituents.

5. A (complex) pattern $\mathbf{f}(\mathbf{x}) \in \Omega$ has a certain **structure**. Not any arrangement of simple constituents is a valid pattern. Many patterns may be represented with relatively few constituents.

Postulates for Pattern Recognition (cont.)

4. A (complex) pattern consists of **simpler constituents**, which have certain relations to each other. A pattern may be decomposed into these constituents.

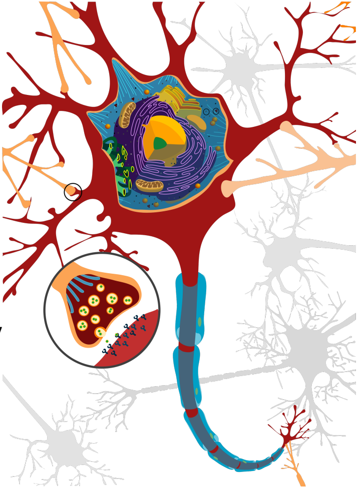
5. A (complex) pattern $\mathbf{f}(\mathbf{x}) \in \Omega$ has a certain **structure**. Not any arrangement of simple constituents is a valid pattern. Many patterns may be represented with relatively few constituents.

6. Two patterns are **similar** if their features or simpler constituents differ only slightly.

Perceptron

Perceptron Biology - Neural Excitation (simplified)

- Neurons are **connected** by synapses / dendrites
- If the **sum** of incoming (excitatory and inhibitory) **activations** is large enough, an action potential is created
- The action potential activates synapses to other neurons, "transmitting" information
- All-or-none response: A **higher** stimulus does **not** cause a **higher** response ➔ "binary classifier"



Source: https://commons.wikimedia.org

Rosenblatt's Perceptron

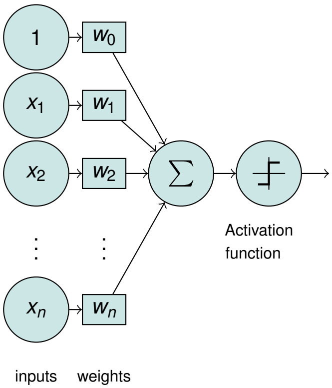
- In 1957, Frank Rosenblatt **Rosenblatt57-P** invented the Perceptron
- Binary classification $y \in \{-1, 1\}$.
- It computes the function

$$\hat{y} = \text{sign}(\mathbf{w}^\mathsf{T}\mathbf{x}),$$

where

$\mathbf{w} = (w_0, \ldots, w_n)$: set of weights ($w_0$=bias)

$\mathbf{x} = (1, x_1, \ldots, x_n)$: input feature vector



Perceptron Objective Function

Task: Find weights that minimize the distance of misclassified samples to the decision boundary

Assumptions

- Let $S = \{(\mathbf{x_1}, y_1), (\mathbf{x_2}, y_2), \dots, (\mathbf{x_m}, y_m)\}$ be a training data set
- Let $\mathcal{M}$ be the set of misclassified feature vectors $y_i \neq \hat{y}_i = \text{sign}(\mathbf{w}^\intercal \mathbf{x_i})$ according to a given set of weights $\mathbf{w}$
- Optimization problem:

$$\underset{\mathbf{w}}{\text{argmin}} \left\{ D(\mathbf{w}) = -\sum_{\mathbf{x_i} \in \mathcal{M}} y_i \cdot (\mathbf{w}^\intercal \mathbf{x}_i) \right\}$$

Perceptron Objective Function – Observations

- Objective function depends on misclassified feature vectors $\mathcal{M}$ ➔ iterative optimization
- In each iteration, the cardinality and composition of $\mathcal{M}$ may change
- The gradient of the objective function is:

$$\nabla D(\mathbf{w}) = -\sum_{x_i \in \mathcal{M}} y_i \cdot \mathbf{x}_i$$

Perceptron Training

- Strategy 1: Process all samples, then perform weight update
- Strategy 2: Take an update step right after each misclassified sample
- Update rule in iteration $(k+1)$ for the misclassified sample $\mathbf{x}_i$ simplifies to:

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + y_i \cdot \mathbf{x}_i$$

- Optimization until convergence or for a predefined number of iterations

NEXT TIME

ON DEEP LEARNING

Introduction - Part 5

A. Maier, V. Christlein, K. Breininger, Z. Yang, L. Rist, M. Nau, S. Jaganathan, C. Liu, N. Maul, L. Folle, K. Packhäuser, M. Zinnen
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
October 1, 2024

Organizational Matters

Grading

- Module consists of lecture **and** exercises (together 5 ECTS)
- 90 min. written exam in the semester break, determines grade
- Exercises are **optional**. 100% exercise completion = 10% grade when you pass the exam

Exercise Content

- Python introduction
- Developing a neural network framework from scratch
  - Feed Forward Neural Networks
  - Convolutional Neural Networks
  - Regularization
  - Recurrent Networks
- Using the PyTorch framework
  - Large scale classification

Exercise Requirements

- Basic knowledge of Python and Numpy
- Linear algebra, -
- Image processing, -
- Pattern recognition fundamentals
- Passion for coding
- Attention to detail
- Time

How it works

- Five exercises throughout the semester
- Unit tests for all but last exercise
- Last exercise: PyTorch + Challenge
- Assistance during exercise sessions
- Personal demonstration of every exercise to get bonus points
- Exercise deadlines are announced in the respective exercise sessions

Summary

- Deep learning more and more present in day to day life
- Huge support and interest from industry
- **Very** active area of research!
- Perceptron as binary classifier motivated by biological neurons

NEXT TIME

ON DEEP LEARNING

Next Lecture Block

- Extending the Perceptron to obtain a universal function approximator
- Gradient based training algorithm for these models
- Efficient automatic computation of gradients

Comprehensive Questions

- What are the six postulates of pattern recognition?
- What is the Perceptron objective function?
- Can you name three applications successfully tackled by deep learning?

Further Reading

- Link - Deep learning book
- Link - Research and publications at the Pattern Recognition Lab
- Link - Google Research Blog with posts on e.g. Deep dream or Alpha Go

Questions?

References

References I