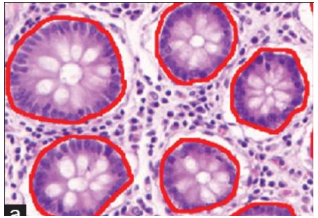
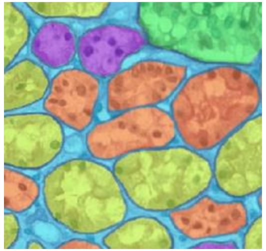
Deep learning for AI-based diagnosis and treatment planning in medicine

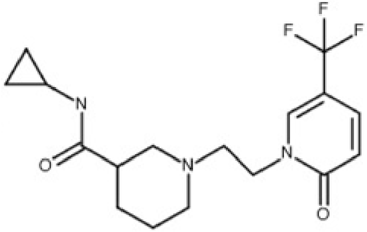
A Tack, K Zeigeler, G Armbrecht, K-G Hermann, S Zachow

AI-based diagnosis and treatment planning

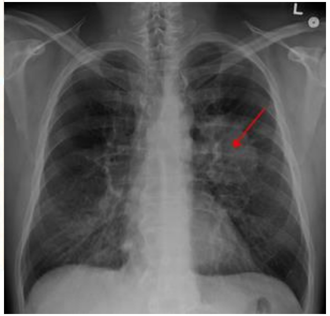


Pathology



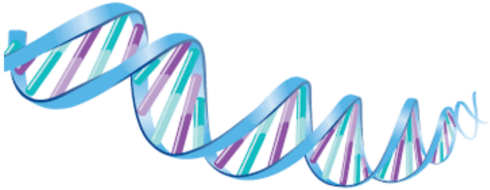


Drug discovery



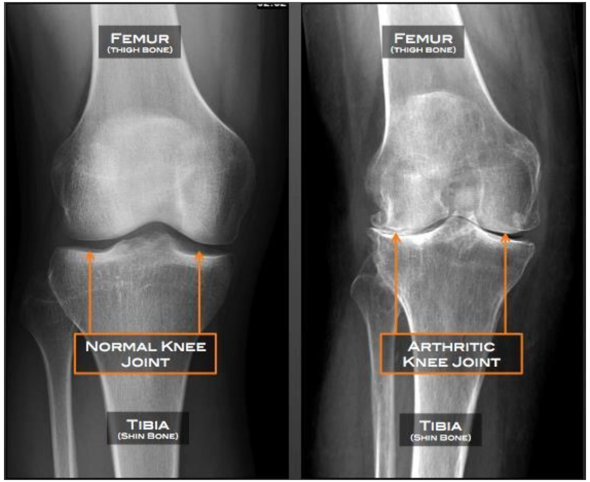
Oncology

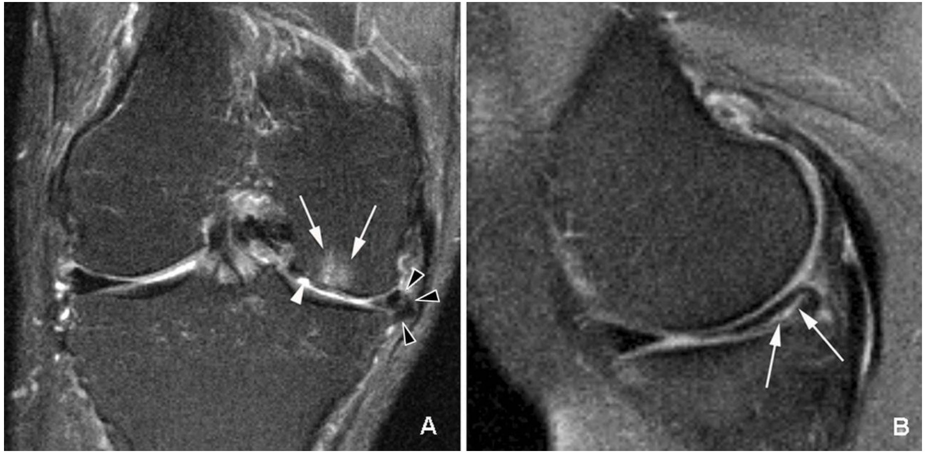
Omics

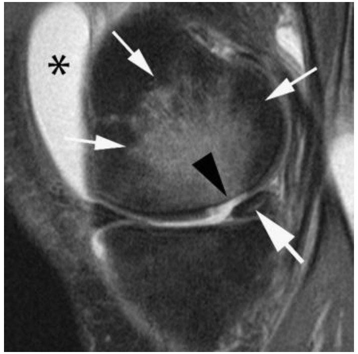


AI-based diagnosis and treatment planning

Our field of research:
Computer-aided diagnosis and treatment planning for Knee Osteoarthritis (OA)







Cohort Studies for Investigation of Osteoarthritis

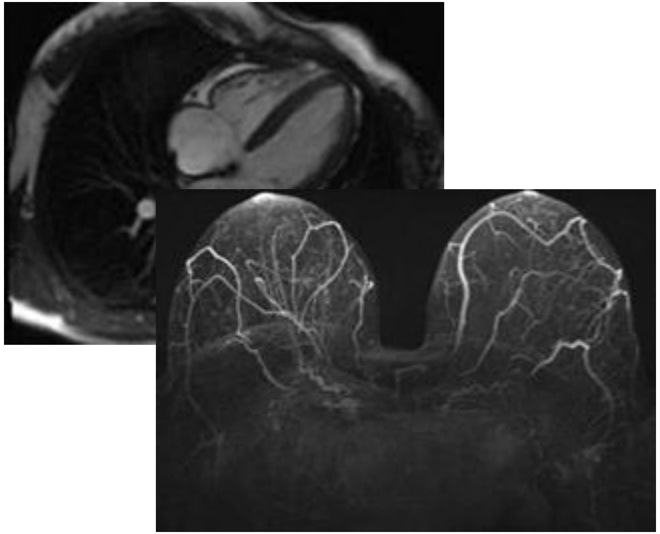
Large databases

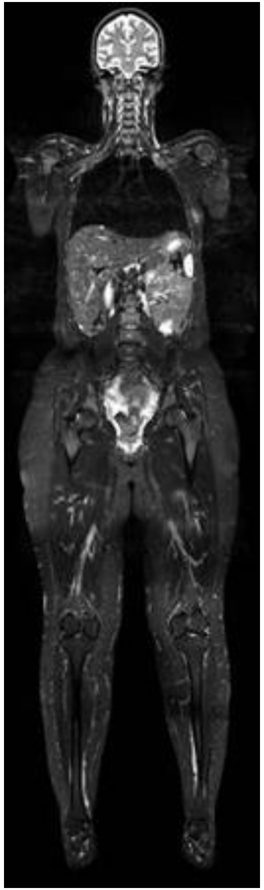
Study of Health in Pomerania (SHIP)

From 1997-today
≈ 2,000 subjects,
2 time points,

Multiple MRI sequences,
serum/urine data, a lot of additional information

Anatomies: whole body, head, neck, thorax, pelvis, spine, …



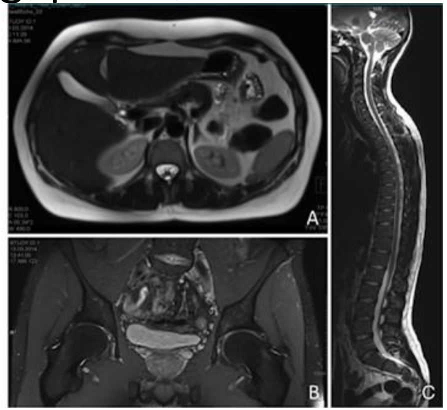


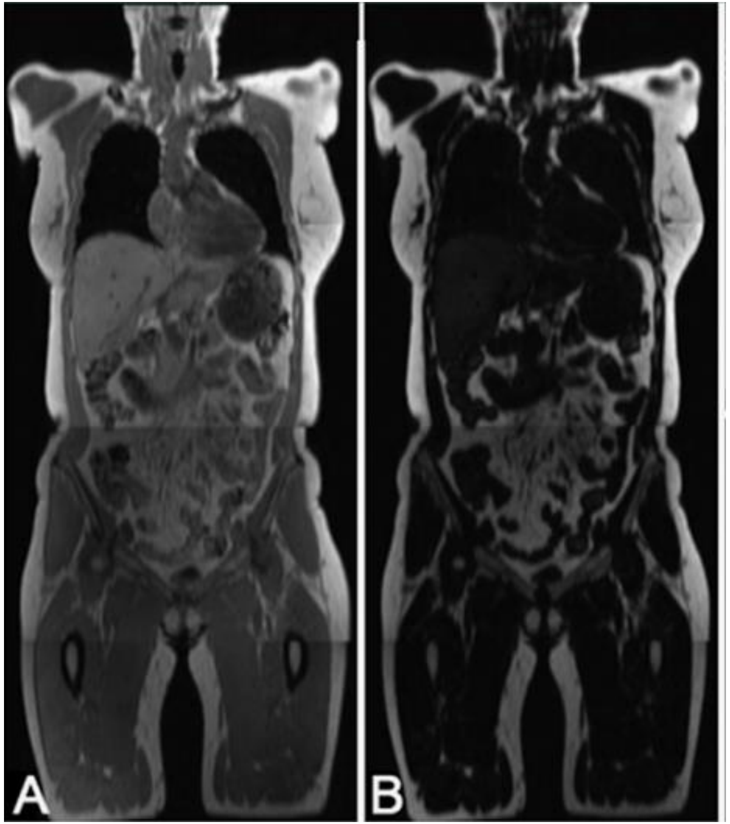
Large databases

The German National Cohort (GNC)

NAKO GESUNDHEITS-STUDIE

From 2014-today
≤ 30,000 subjects,
2 time points,
Several 3T MRI sequences
Anatomies: brain, spine, lung, pelvis, …



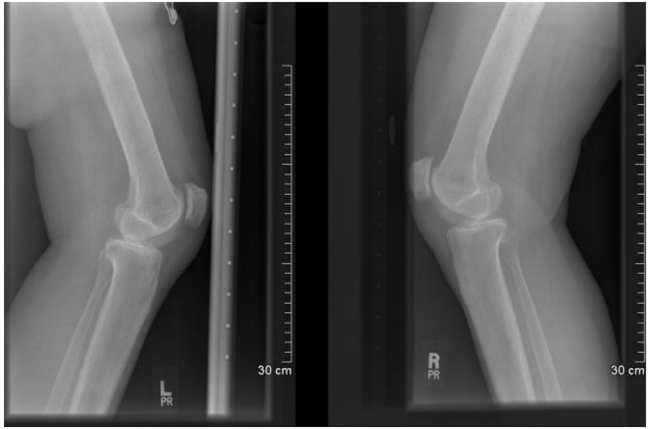


Large databases

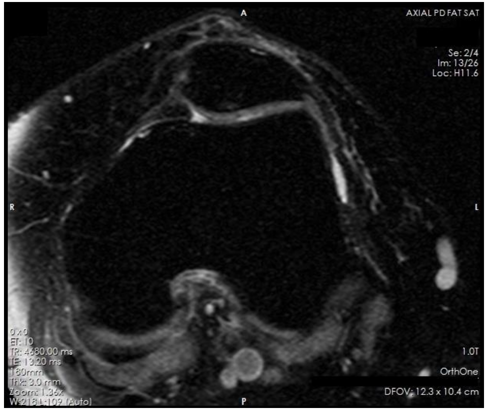
Multicenter Osteoarthritis Study (MOST)

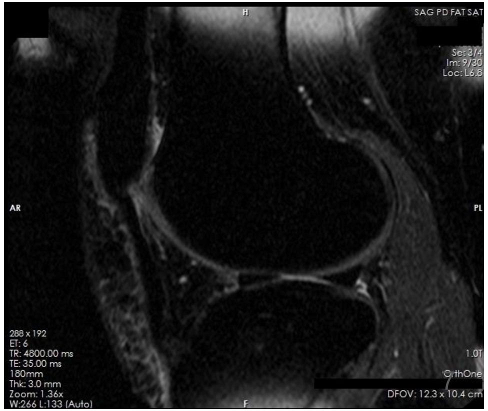
From 2003-2010
≈ 3,000 subjects,
6 time points,
Multiple **1T** MRI sequences, X-ray images, serum/urine data, image assessment studies, a lot of additional information

Anatomies: hip, knee



MULTICENTER OSTEOARTHRITIS STUDY
PUBLIC DATA SHARING
MOST





Large databases

The Osteoarthritis Initiative (OAI)

From 2004-today
≈ 5,000 subjects,
7 time points,
Multiple **3T** MRI sequences, X-ray images, serum/urine data, image assessment studies, a lot of additional information

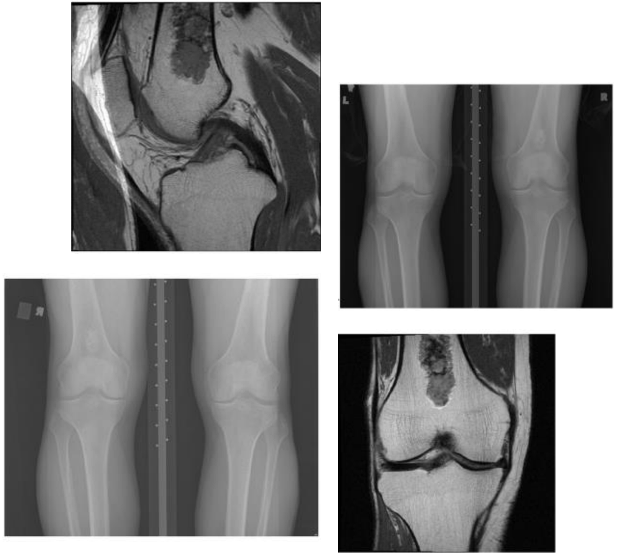
Anatomies: hand, hip, knee

X-rays

MRIs

Deep Learning from Large Databases

Deep learning for 3D medical image data



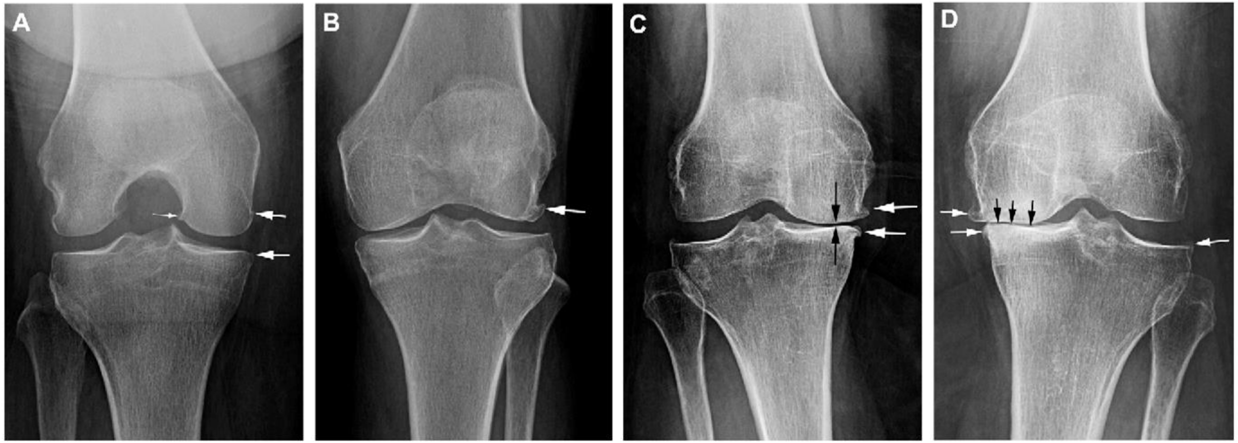
How can we use large cohorts for deep learning-based diagnosis and treatment planning?

What has been done so far?

What are the challenges?

What is the outlook?

Examples: findings in X-ray data



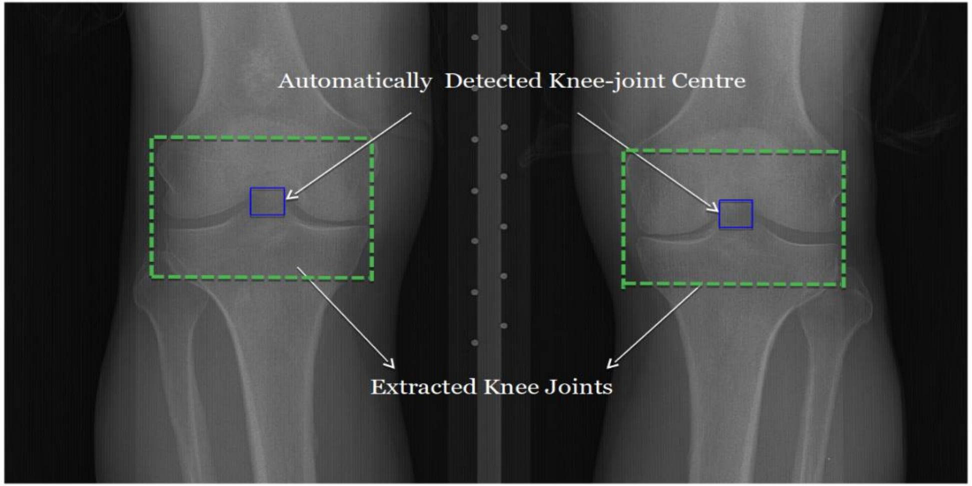
- Osteophytes
- Joint Space Narrowing
- Sclerosis
- Deformation of the bones

→ **Kellgren-Lawrence Grade (KLG)**

Image from: D Hayashi, FW Roemer, A Guermazi: *Imaging for Osteoarthritis* (2016)

Deep learning for 3D medical image data

Diagnosing knee osteoarthritis from X-Ray



Antony et al. *Quantifying Radiographic Knee Osteoarthritis Severity using Deep Convolutional Neural Networks*. (2016)
→ VGG16

Suresha et al. *Automated Staging of Knee Osteoarthritis Severity using Deep Neural Networks*. (2018)
→ ImageNet

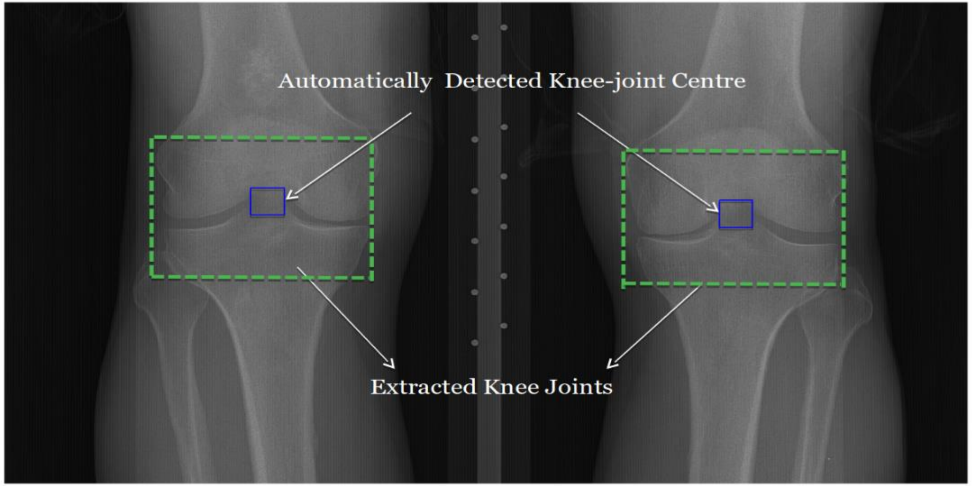
Tiulpin et al. *Automatic Knee Osteoarthritis Diagnosis from Plain Radiographs: A Deep Learning-Based Approach.* (2018)
→ Siamese CNN

… and many more ...

Deep learning for 3D medical image data

Diagnosing knee osteoarthritis from X-Ray

- 2D image data
- Annotations often routinely available
- Small “Region of Interest”
- Clearly defined pathological changes



- X-rays are just a 2D projection
- No soft tissue visible
- Low accuracy in multi-class settings
- Low sensitivity to progression
- Limited explanation of knee pain

Deep learning for 3D medical image data

Our aims for knee OA:

Efficient handling of 3D MRI data to...

… diagnose the OA grade

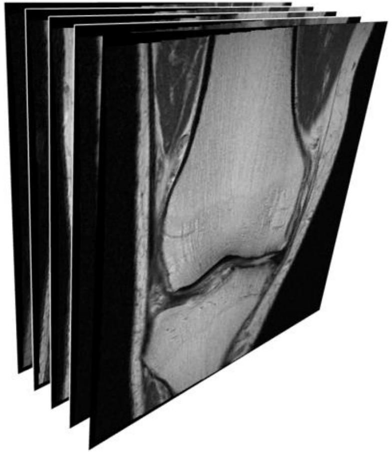
… predict incident OA

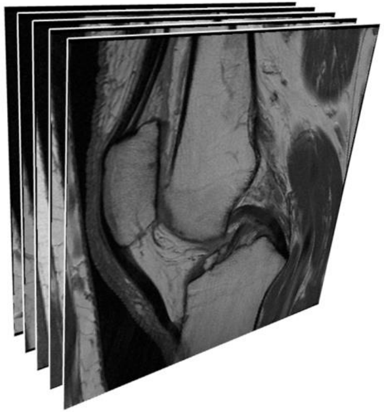
… predict OA progression

… understand pain

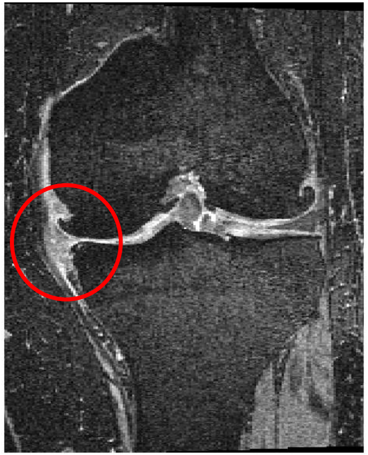
… understand and identify phenotypes

… evaluate treatment programmes

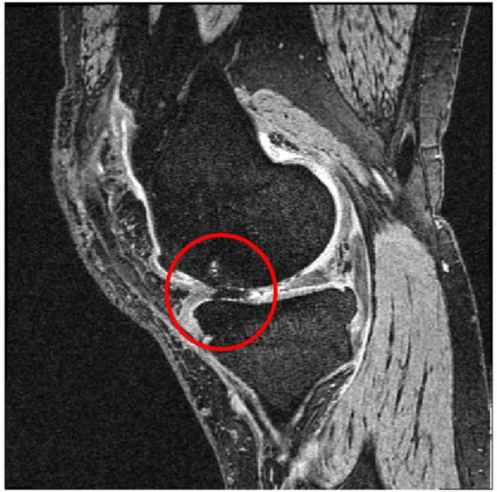




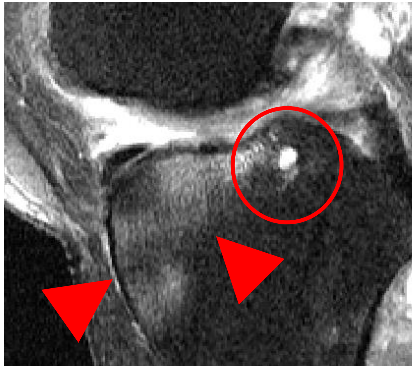
Examples: findings in MRI data



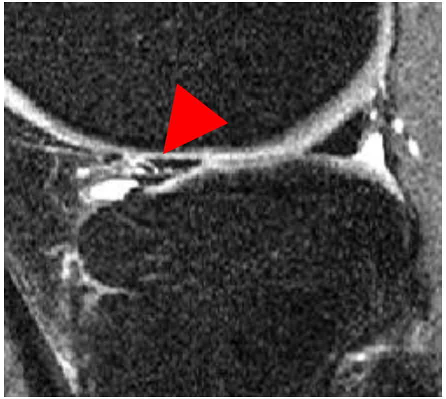
Meniscal extrusion



Cartilage defects

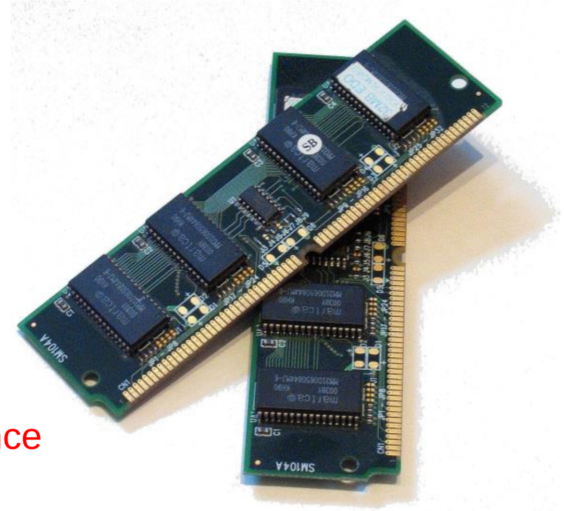


Bone marrow lesions and cysts

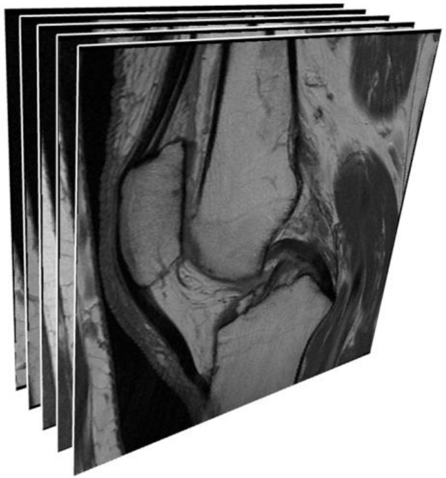


Meniscal tears

Deep learning for 3D medical image data



Challenges of handling 3D image data



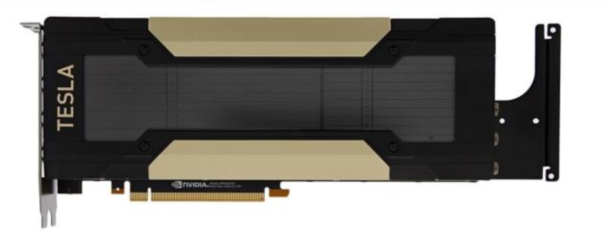
- Small pathological changes
- Weak annotations
- Large volumes

- High run time of training and inference

- GPU memory

- Main memory

- Storage capacity



Deep learning for 3D medical image data

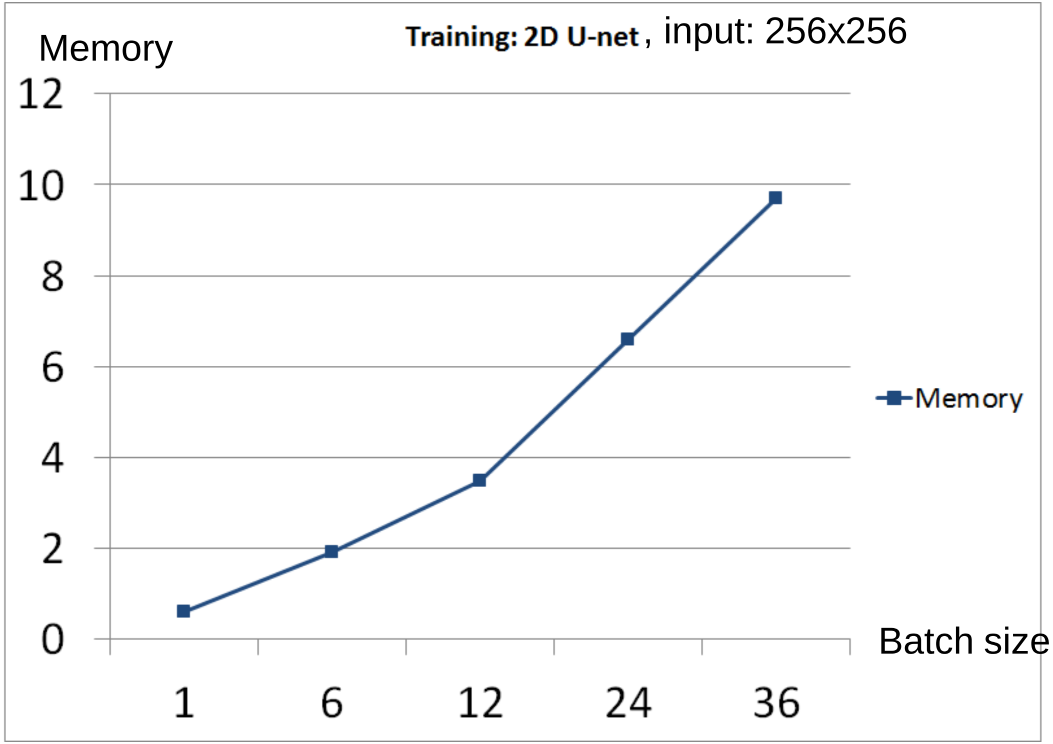
Challenge: **GPU memory**

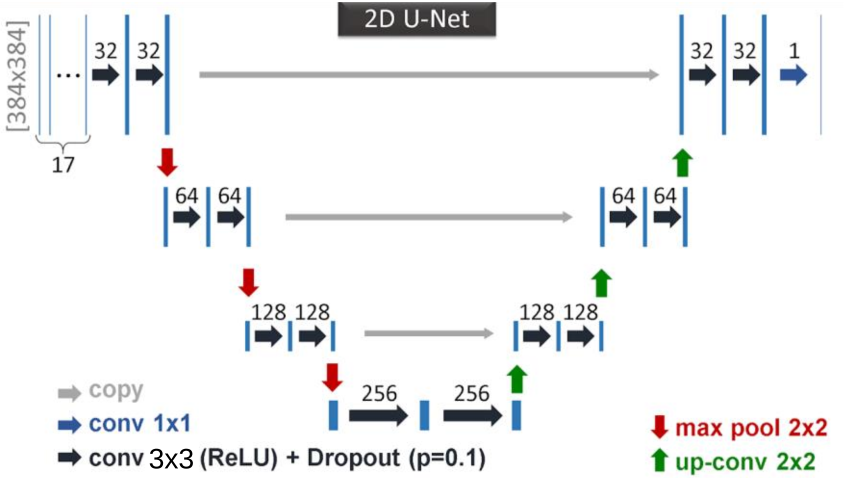
The amount of memory needed for training a CNN is mostly influenced by:

- The number of neurons
- The activations and gradients for each neuron
- The size of the batches used for training
- The size of the input images

Deep learning for 3D medical image data

Example: Training a 2D U-net for segmentation of 2D data

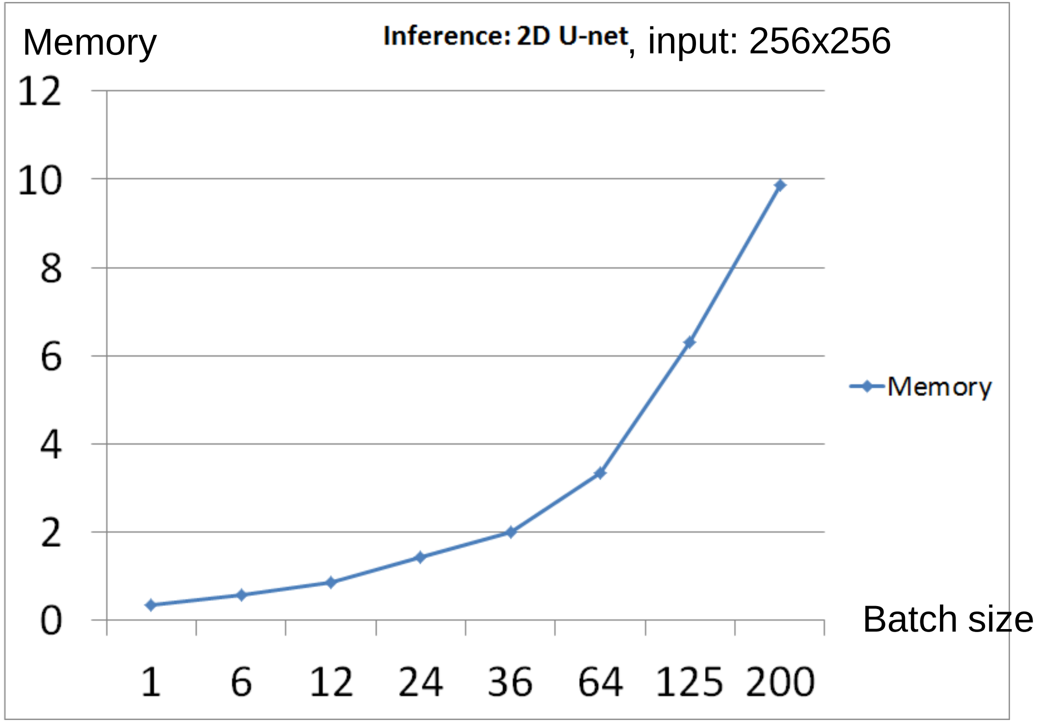


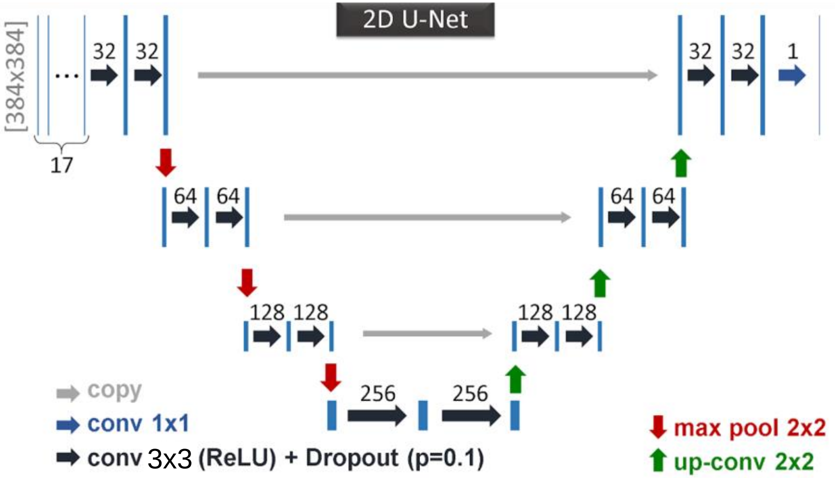


2D U-net architecture

Deep learning for 3D medical image data

Example: Inference using a 2D U-net for segmentation of 2D data

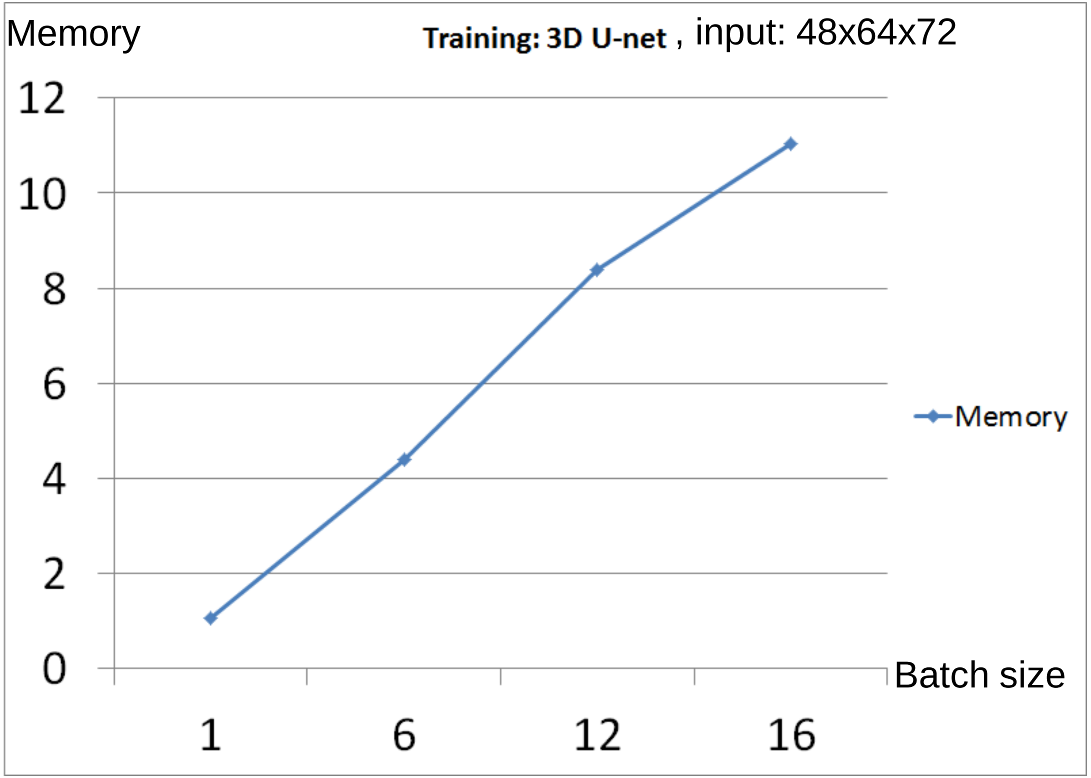


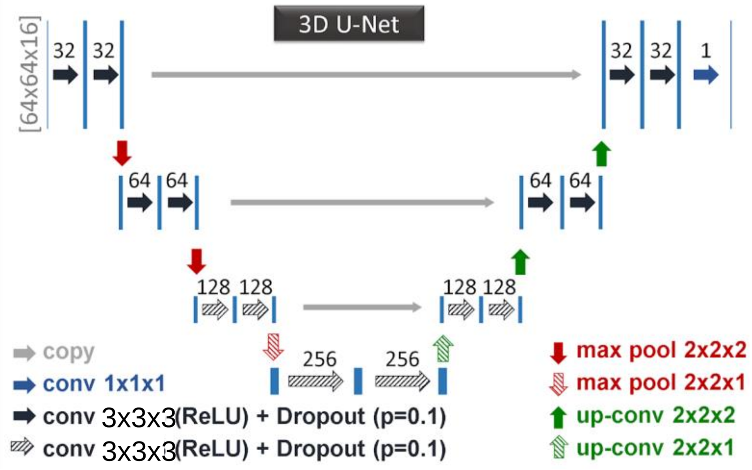


2D U-net architecture

Deep learning for 3D medical image data

Example: Training a 3D U-net for segmentation of 3D data

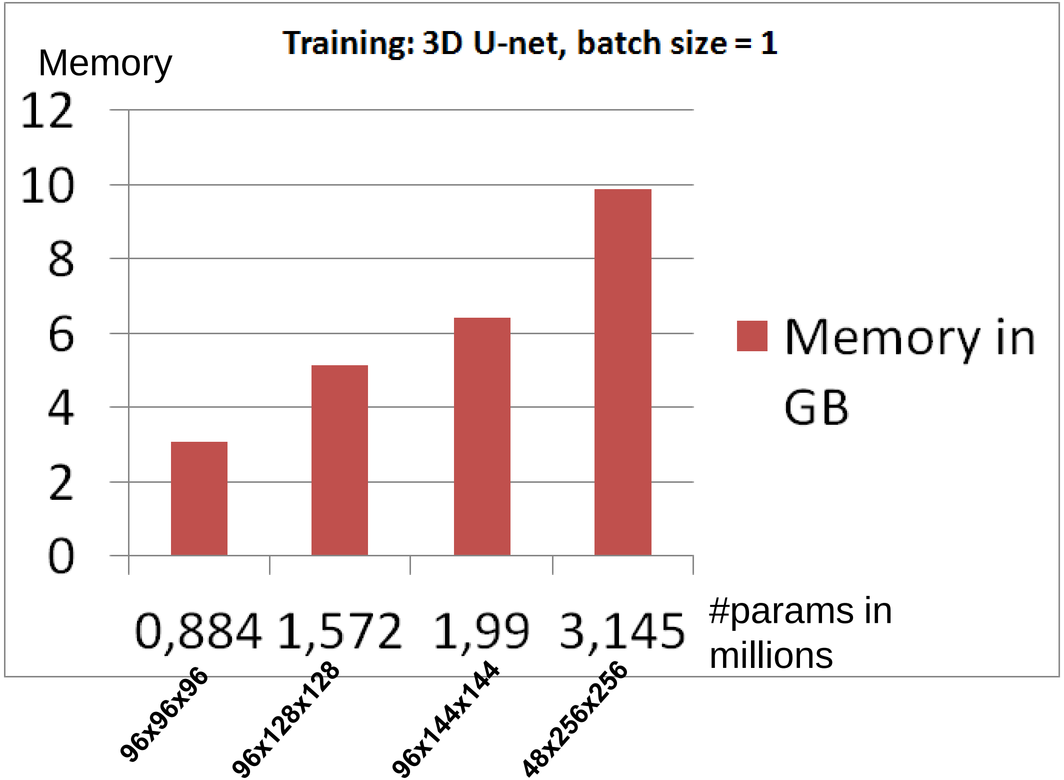


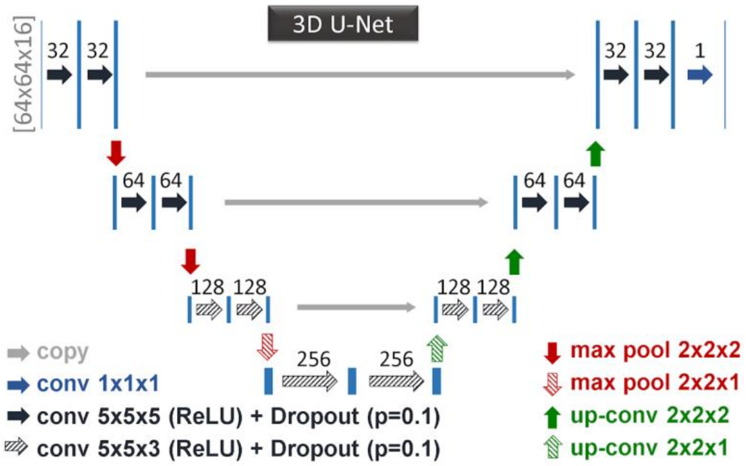


3D U-net architecture

Deep learning for 3D medical image data

Example: Training a 3D U-net for segmentation of 3D data,
Which ROI can we choose with batch size fixed to 1?





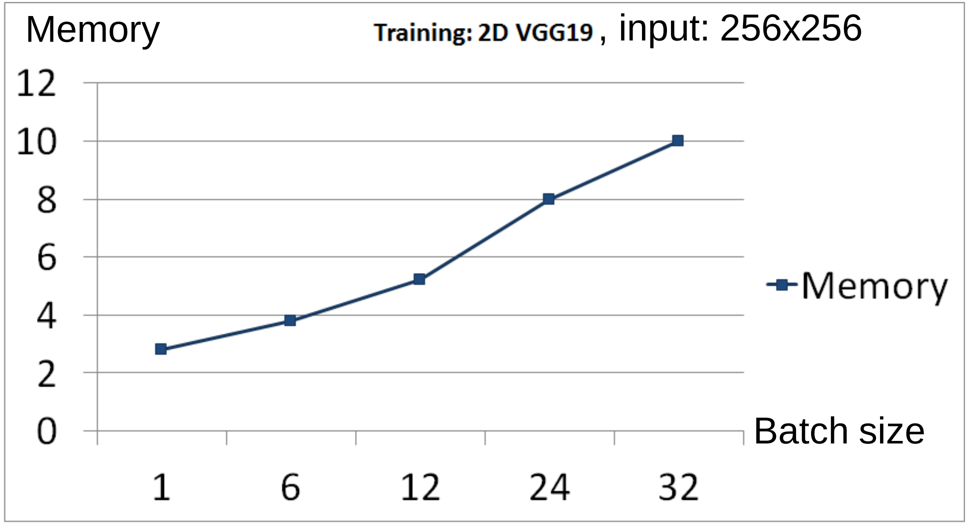
3D U-net architecture

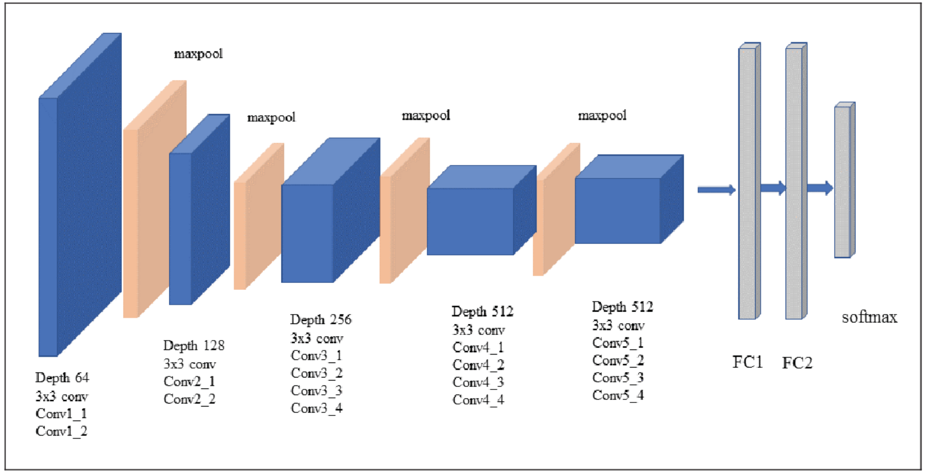
Input size of **48x256x256** and comparable is possible.

We want: **160x384x384**
→ **23,592,960 parameters** → **≈ 40 GB!**

Deep learning for 3D medical image data

Example: training a VGG19-like CNN for classification of 2D data



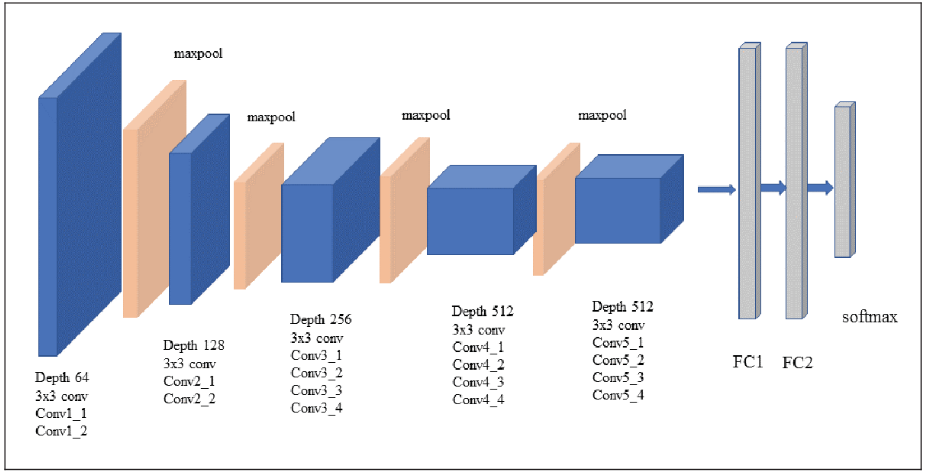


VGG19 architecture

Deep learning for 3D medical image data

Example: training a VGG19-like CNN for classification of 2D data

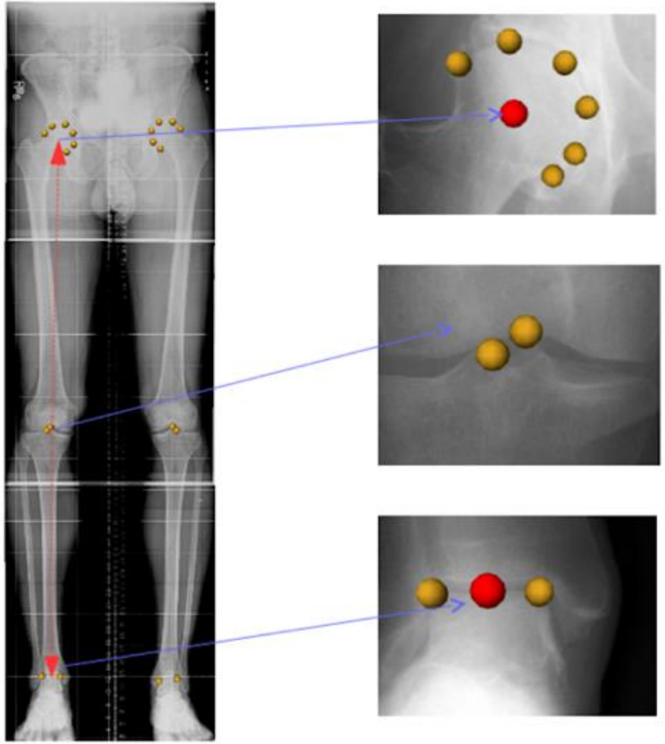
Example input image	Input-size [in thousands]	#parameters [in millions]	Memory [in GB]
224x224	50	139	2.3
256x256	65	171	2.8
320x320	102	246	4.0
384x384	147	338	5.4
512x512	265	573	9.1



VGG19 architecture

Deep learning for 3D medical image data

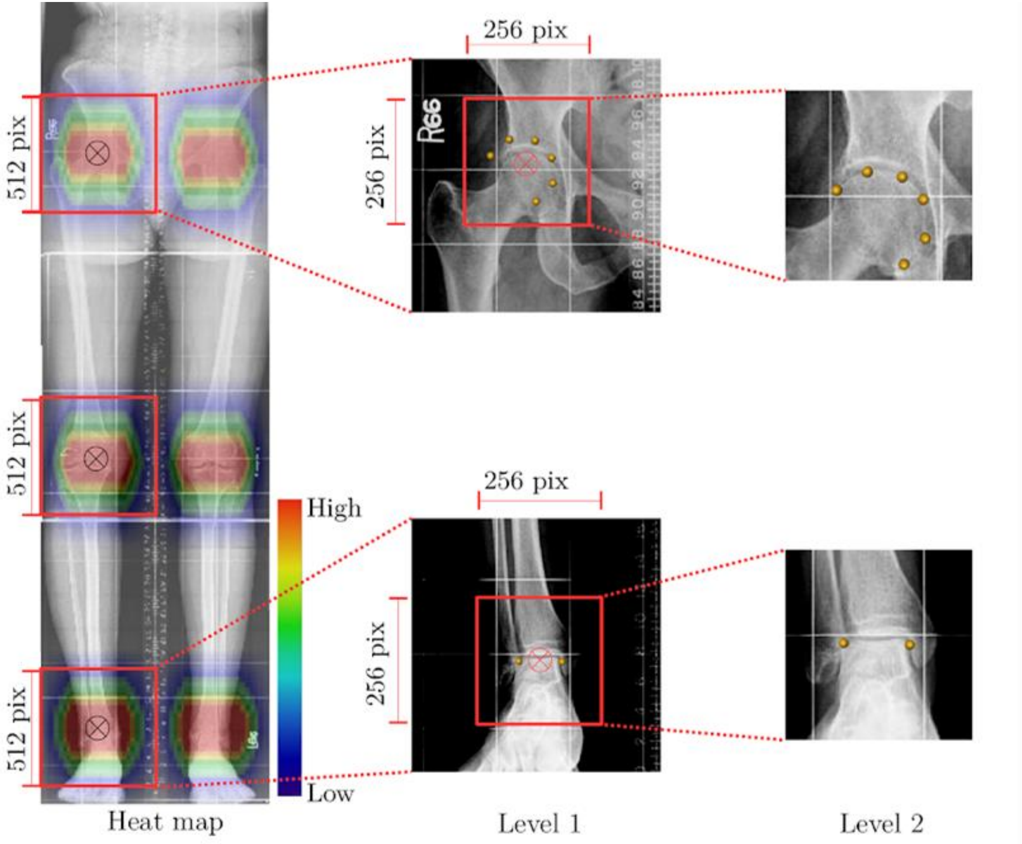
Computer-aided diagnosis of leg alignment using full-leg X-ray images



1000 x 4000 image size!

Deep learning for 3D medical image data

Computer-aided diagnosis of leg alignment using full-leg X-ray images

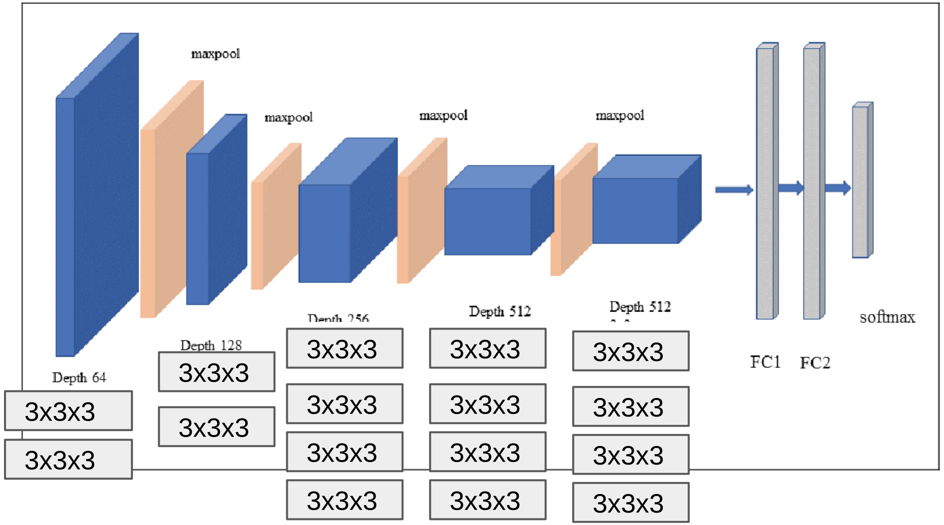


→ **cascaded approach to account for the large image size**

Deep learning for 3D medical image data

Example: Training a VGG19-like 3D CNN for classification of 3D data
Batch size = 1

Example input image	Input-size [in thousands]	#parameters [in millions]	Memory [in GB]
80x80x80	512	93	2.5
96x96x96	884	133	4.1
128x128x128	2,097	211	8
128x144x144	2,654	211	8.6



3D VGG19 architecture

We want: 160x384x384!

Input-size: 23,592,000
#parameters: 1,586,000,000
Memory: ≈ 80 GB

Deep learning for 3D medical image data

Conclusions

- 2D segmentation easy, but not very good results for 3D data
- 3D segmentation can't be done for the whole image, but for smaller sub-regions

- 2D classification has limits wrt. image size
- 3D classification not really possible

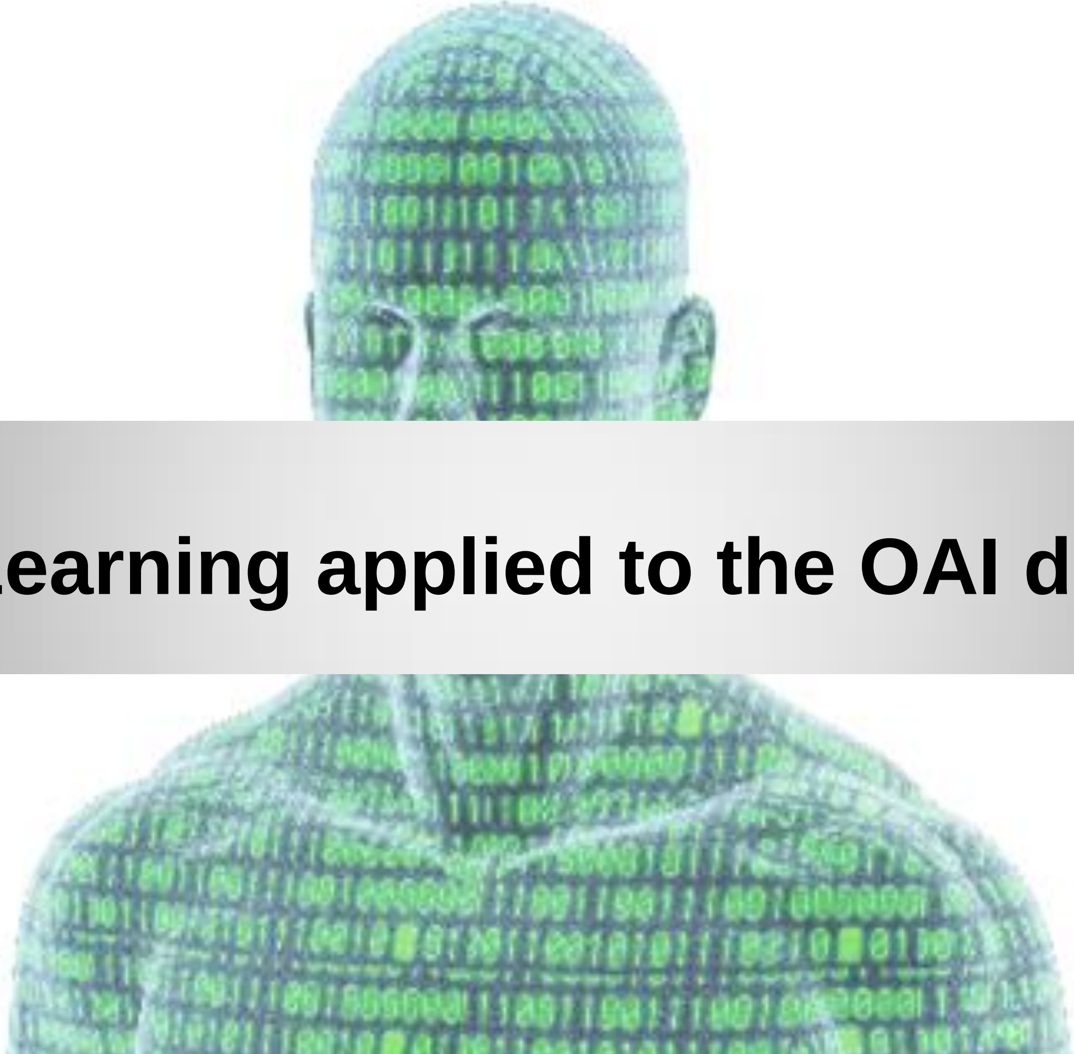
How to handle...

large 2D or 3D images?

4D data (3D+t)?

multiple MRI sequences?

<u>What about the run time?</u>



Deep Learning applied to the OAI database

Recap: OAI Database

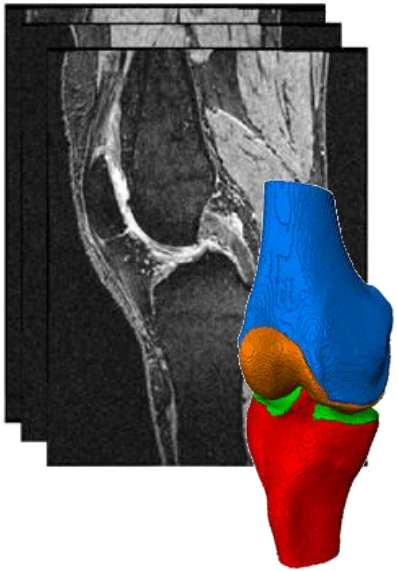
The Osteoarthritis Initiative (OAI)

7 time points
≈ 5.000 subjects
two knees per subject
Multiple MRI sequences

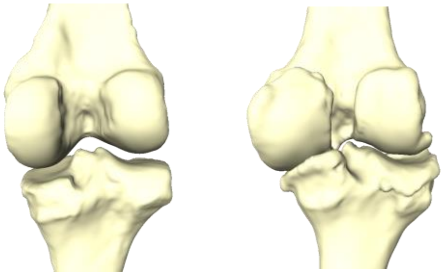
→ using just one MRI sequence: **approx. 50.000 datasets in total**

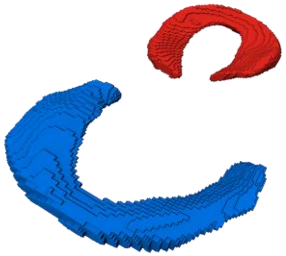
High performance training and inference

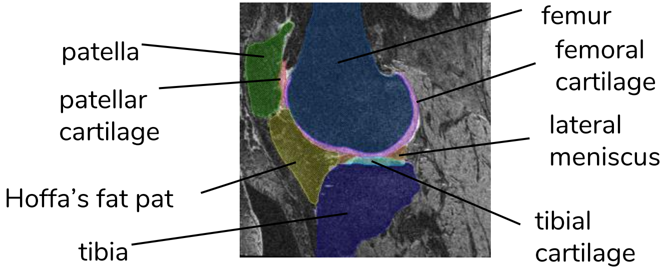
1x NVIDIA GeForce 1080 Ti



Segmentation of 50.000 MRI datasets would take 12 months![2,3]







[2] A Tack, A Mukhopadhyay, S Zachow: *Knee Menisci Segmentation using Convolutional Neural Networks: Data from the Osteoarthritis Initiative* (2018)
[3] F Ambellan, A Tack, M Ehlke, S Zachow: *Automated Segmentation of Knee Bone and Cartilage combining Statistical Shape Knowledge and Convolutional Neural Networks: Data from the Osteoarthritis Initiative* (2019)

High performance training and inference

HLRN-III system “Konrad”

Nodes: 1872
CPUs: 3744
Main memory: 129 TByte
Storage capacity: 4.2 PByte

Inference using 1000 nodes:

5,000 MRIs → ≈30 hours
50,000 MRIs → ≈300 hours

→ segmentation of the whole OAI database would take ≈13 days.

(rough estimation based on experiments using 1000 nodes of the HLRN in parallel)

High performance training and inference

HLRN-III system “Konrad”
Challenges for deep learning

Memory per Node:
64 GB main memory for 24 cores

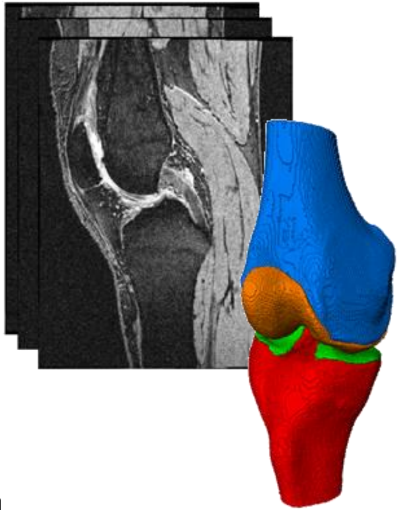
Each CNN-based process for inference needs ≈11GB main memory.

In a “trivial parallel”-setting 5 single threaded individual processes can be started on one node. Each process on one core of the node.

→ **Effectively: Only 5/24 cores are used per node!**

High performance training and inference

NVIDIA
DGX-1



Segmentation of the whole database took 8 days using 8 GPUs of the DGX-1.

Many thanks to: MDC MAX-DELBRÜCK-CENTRUM FÜR MOLEKULARE MEDIZIN IN DER HELMHOLTZ-GEMEINSCHAFT

(Preliminary estimation. Computations will be finished next week.)

High performance training and inference

NVIDIA DGX-1

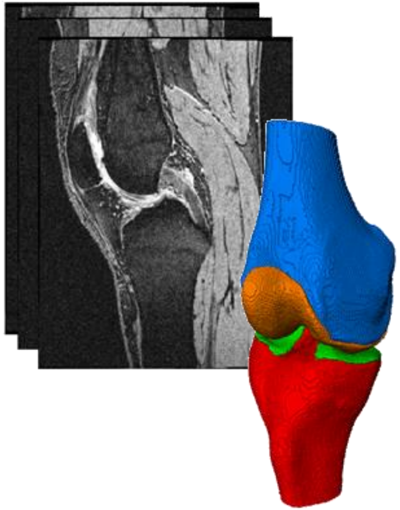
→ **Image data of the OAI database: 11 TB !**

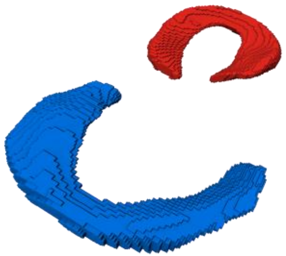
DGX-1 storage capacity: 7 TB.

... supported by a AI200 system by DDN, which increased the hard disk memory of the DGX-1 to 26 TB!

Vision: Deep Learning for Diagnosis and Treatment Planning

Vision: Deep learning to diagnose OA

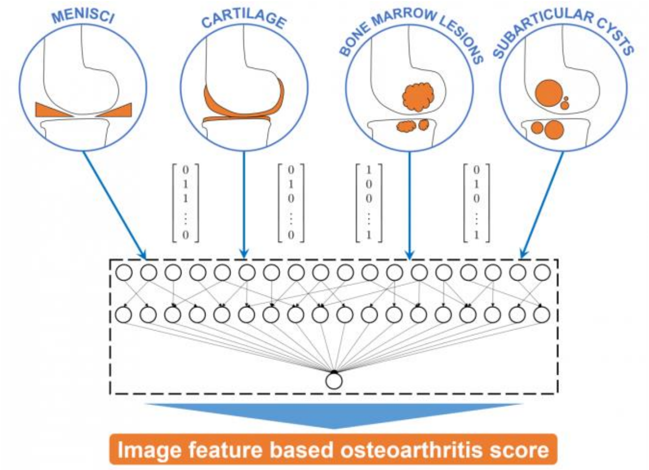


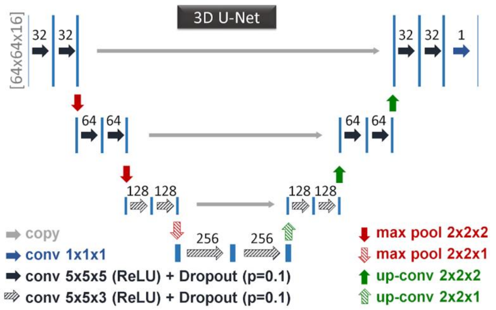


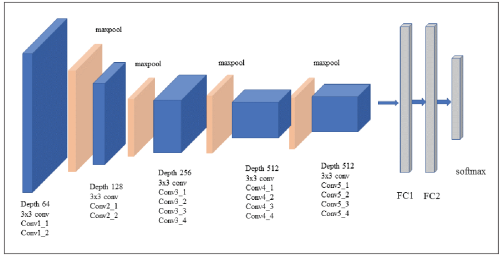
Use OAI database segmentations for

- Training new 3D segmentation methods
- Deep learning-based diagnosis of images (classification)

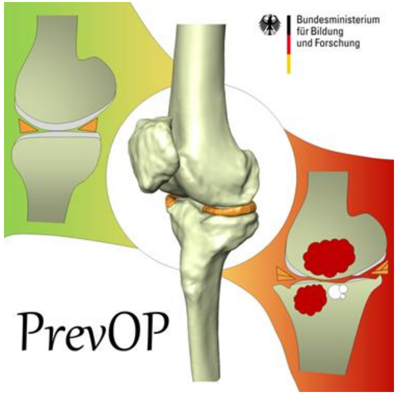
- Computation of image-based biomarkers
→ Combination to a holistic score







Vision: Deep learning for treatment planning



Use deep learning-based OA score to evaluate treatment success after 2 years.

≈ 240 subjects,
4 time points,
Multiple MRI sequence

Thank you!

Questions & Discussion

High performance training and inference

HLRN-III system “Konrad”
Challenges for deep learning

Training of CNNs

- Split the CNN to multiple nodes/cores?
- Duplicate the CNN to multiple nodes/cores?
    - Efficiency?

→ We will soon perform further investigations using the HLRN IV

High performance training and inference

Training and Inference using GPUs: NVIDIA 1080 Ti vs. NVIDIA DGX-1

Training

13 hours

x9 faster!

1.5 hours

Inference

8 hours

x16 faster!

0.5 hours

NVIDIA GeForce 1080 Ti

NVIDIA DGX-1