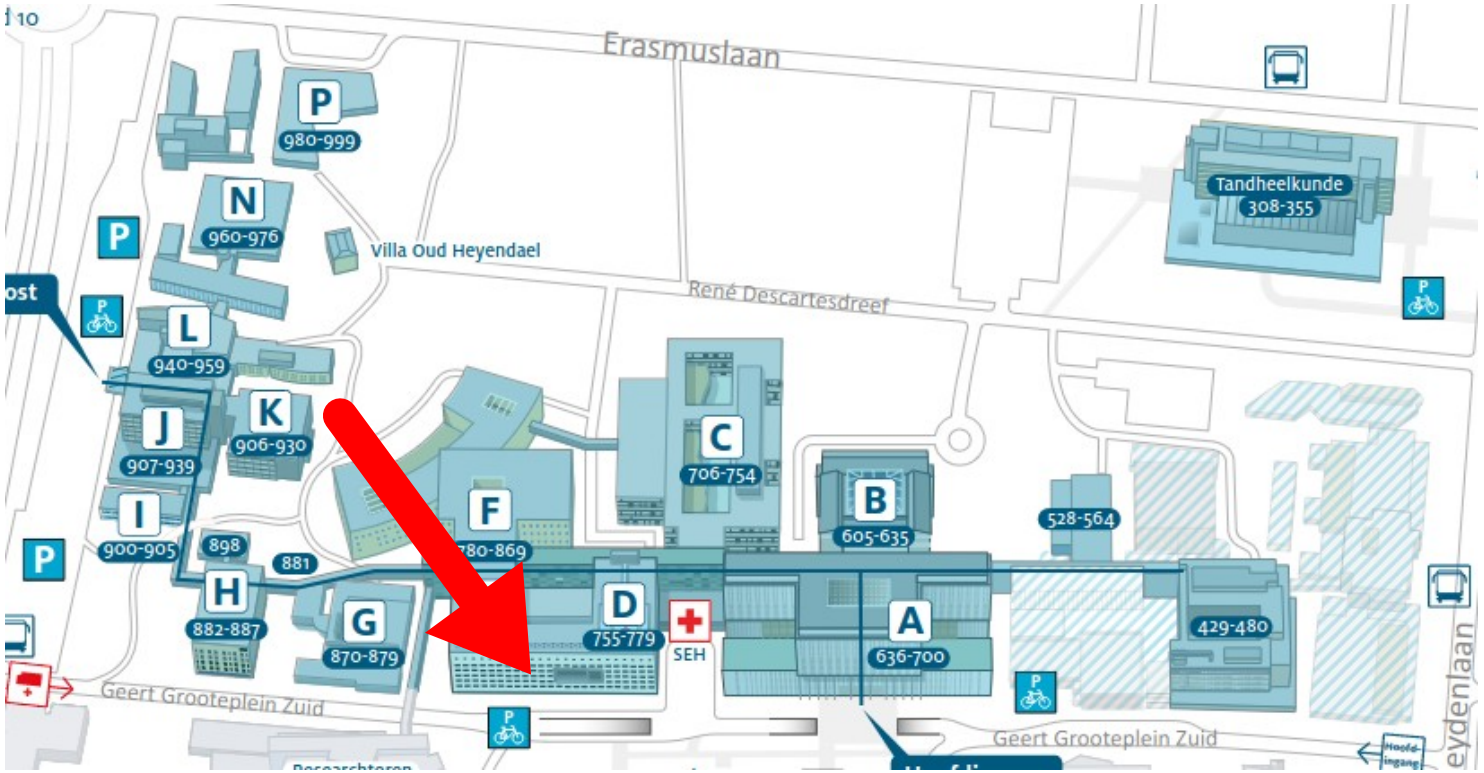
Building and Replacing On-prem Deep Learning Infrastructure for Medical Image Analysis

Paul Konstantin Gerke, MSc
Research Software Engineer
Diagnostic Image Analysis Group (DIAG)

Radboud umc
university medical center

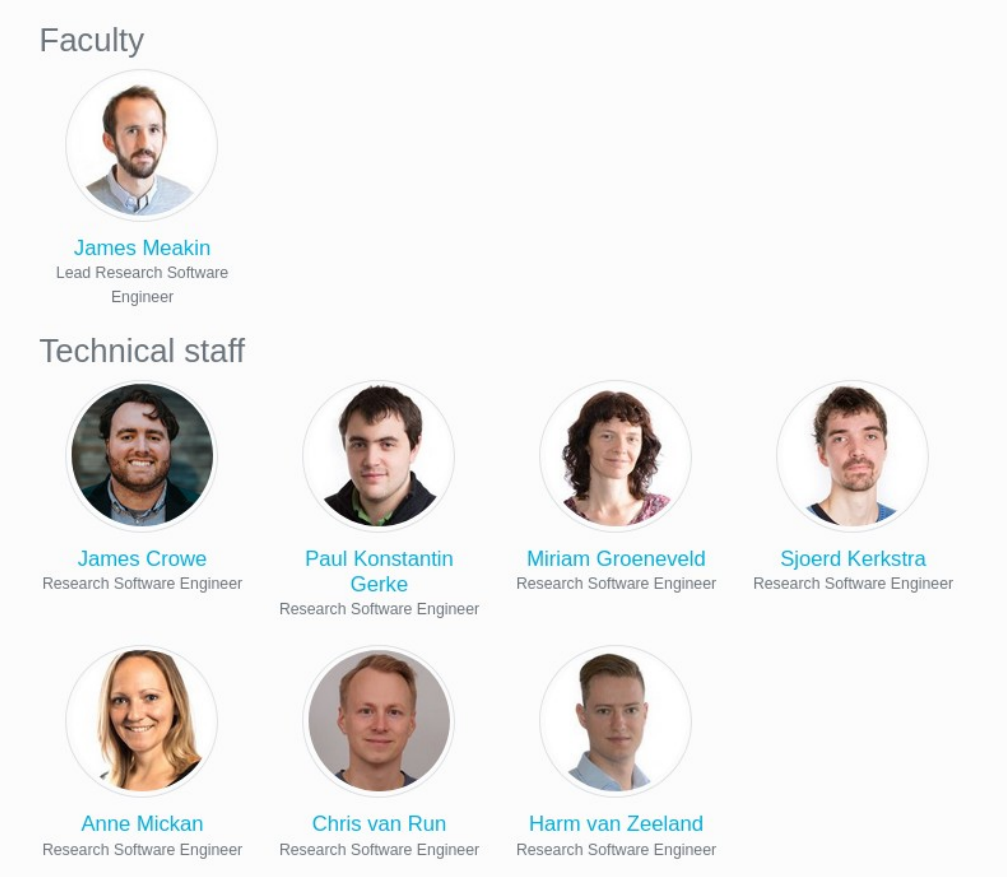
DIAG, the RSE team, and myself

- AI-based medical image analysis from inside the Radiology Department



Radboudumc
university medical center

DIAG, the RSE team, and myself



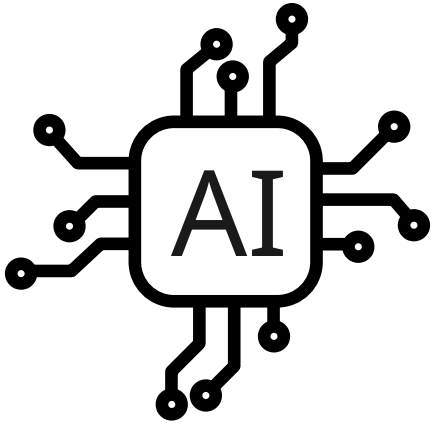
- RSE = Research Software Engineer
- Fully remote software engineering team
- Development of grand-challenge.org

Radboudumc
university medical center

DIAG, the RSE team, and myself

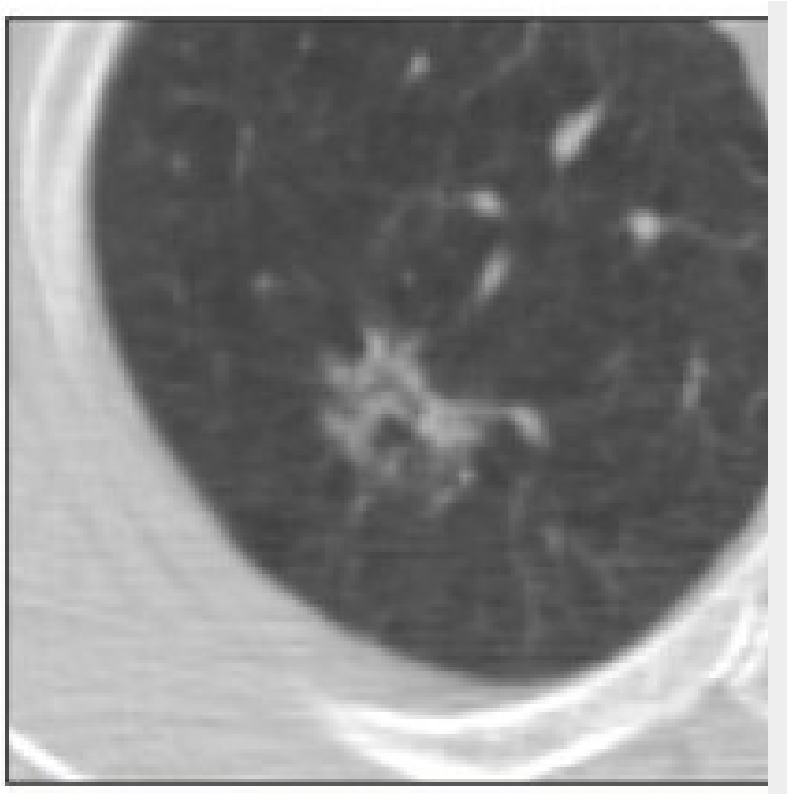


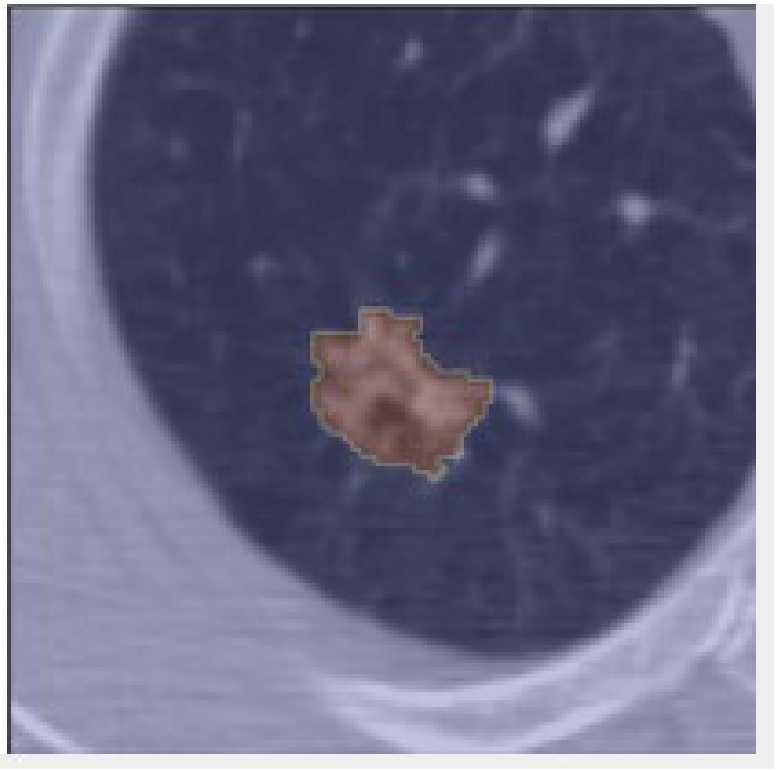




Diagnostic Image Analysis Group



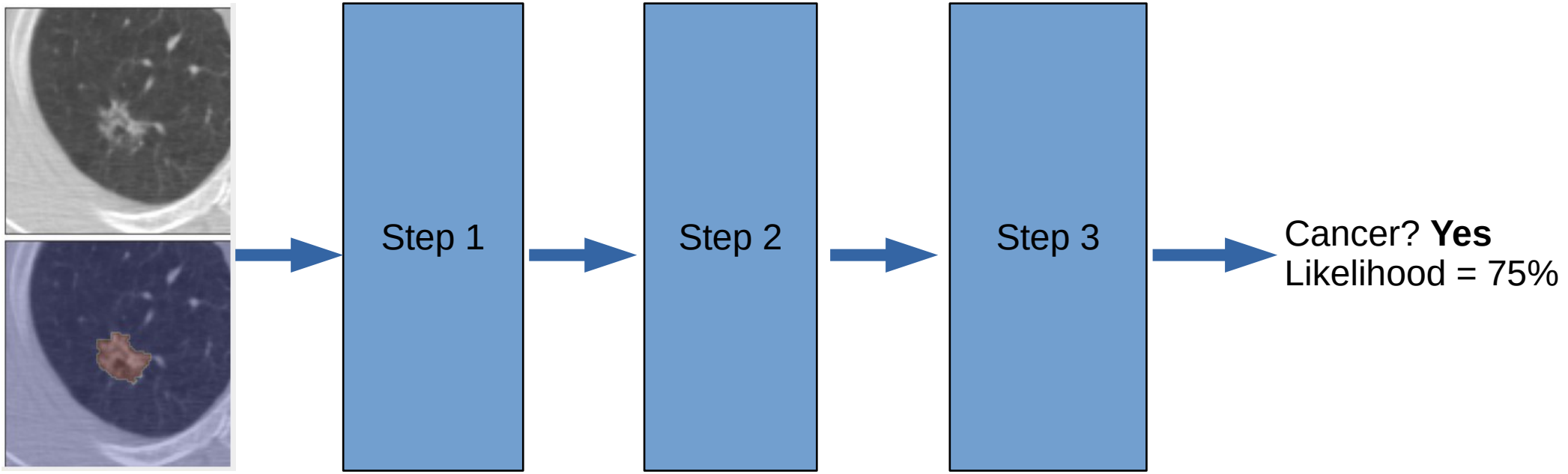




Radboudumc
university medical center

Artificial Intelligence

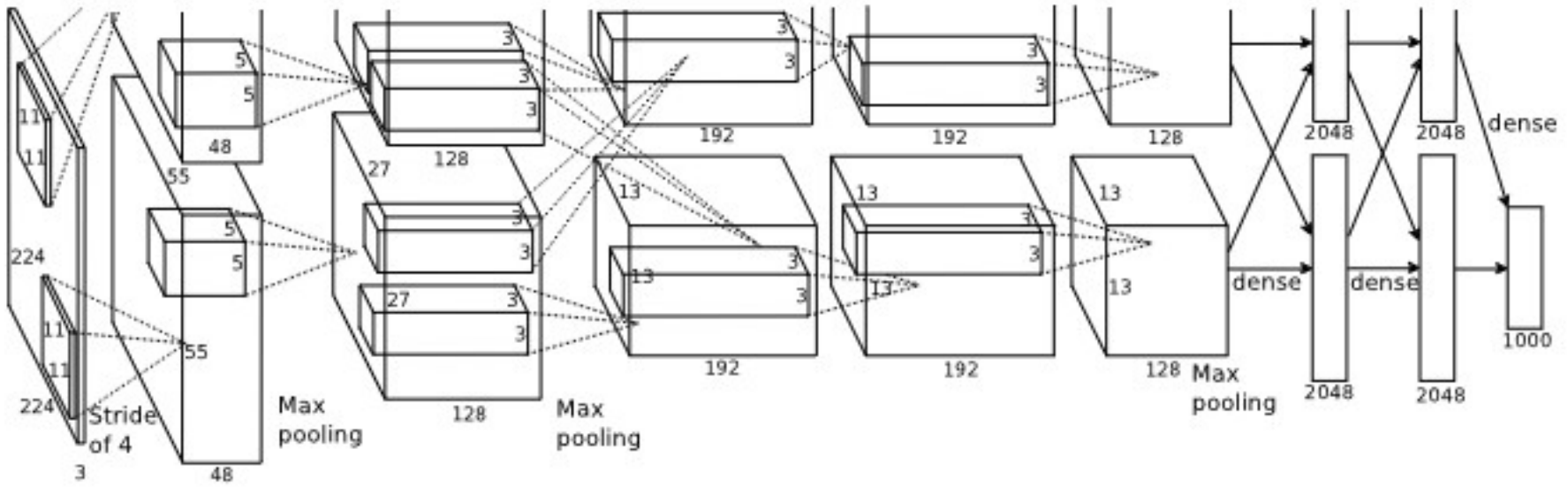
- Before 2015 – classical image processing, classic feature extraction, etc.



Deep learning
A new path for AI

Deep Learning

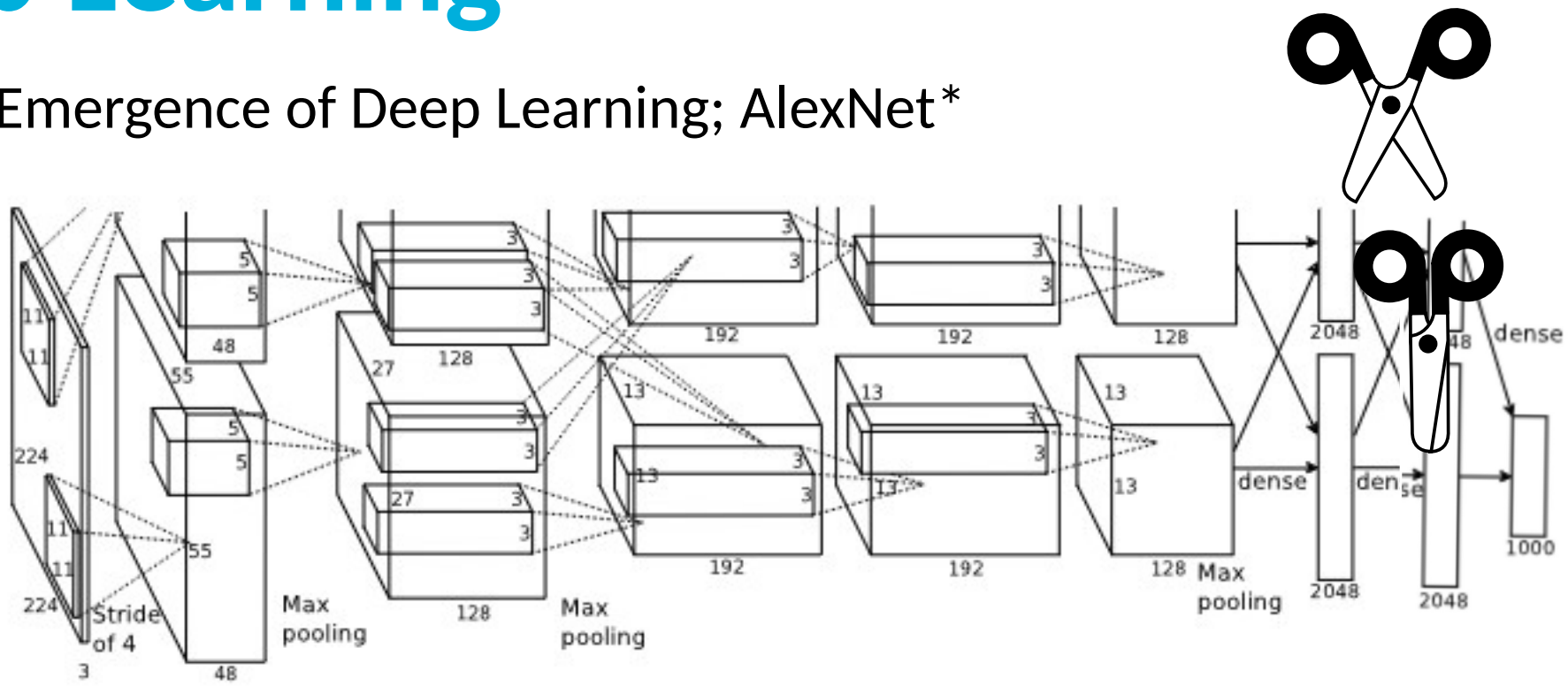
- 2014 – Emergence of Deep Learning; AlexNet*



*: “ImageNet Classification with Deep Convolutional Neural Network”; 2012; Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton; Advances in neural information processing systems

Deep Learning

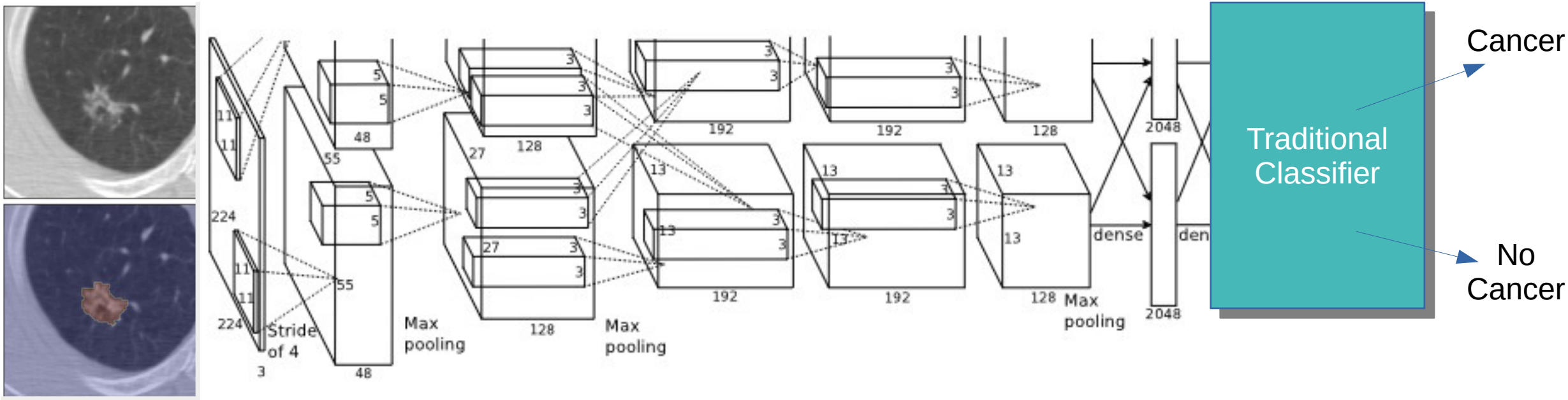
- 2014 – Emergence of Deep Learning; AlexNet*



*: “ImageNet Classification with Deep Convolutional Neural Network”; 2012; Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton; Advances in neural information processing systems

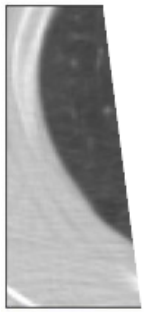
Deep Learning

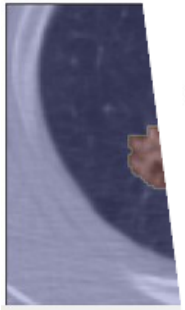
- 2014 – Emergence of Deep Learning; AlexNet*



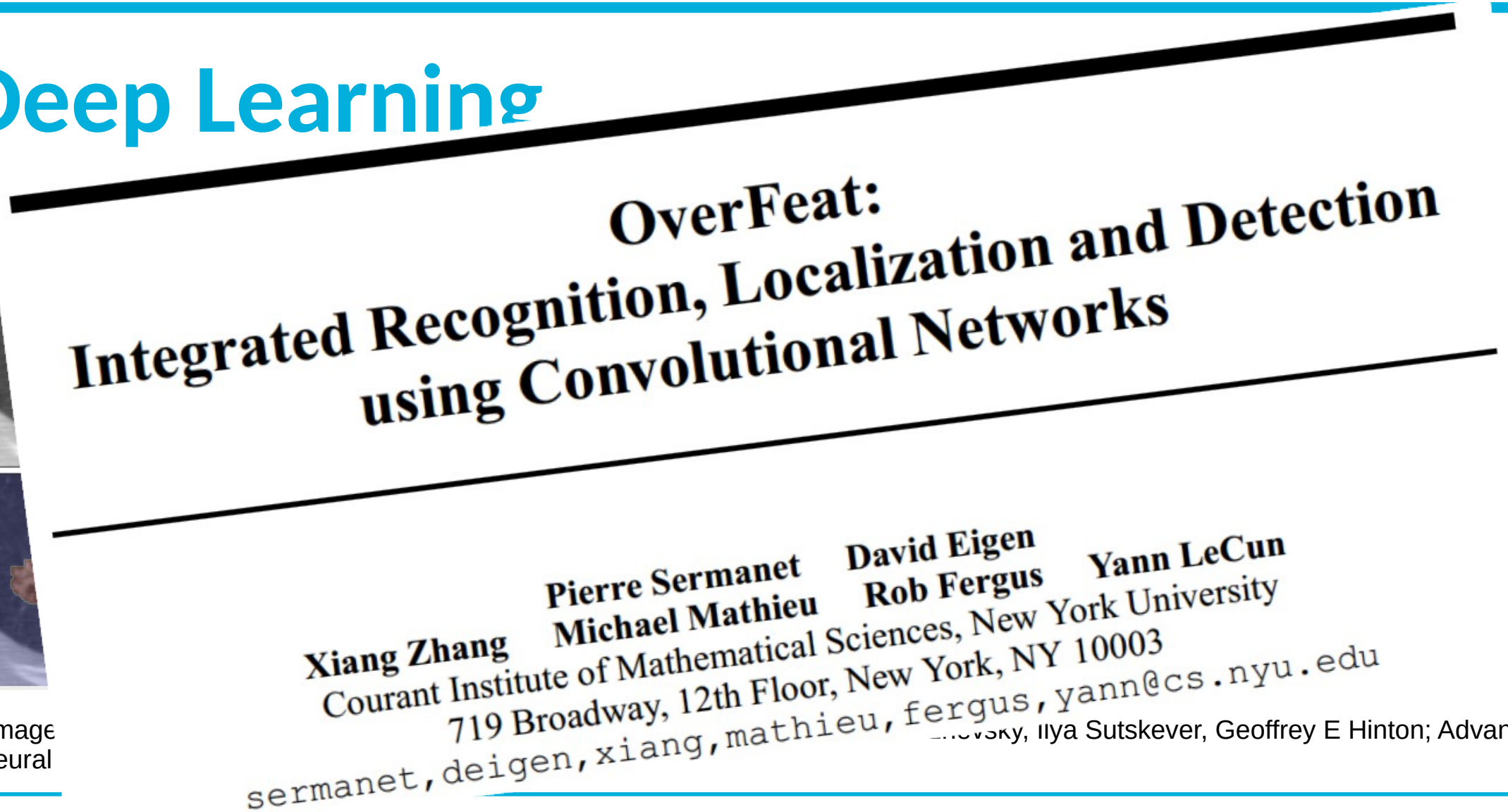
*: “ImageNet Classification with Deep Convolutional Neural Network”; 2012; Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton; Advances in neural information processing systems

Deep Learning





*: “Image
in neural



OverFeat:
Integrated Recognition, Localization and Detection using Convolutional Networks

Pierre Sermanet David Eigen
Xiang Zhang Michael Mathieu Rob Fergus Yann LeCun
Courant Institute of Mathematical Sciences, New York University
719 Broadway, 12th Floor, New York, NY 10003
sermanet,deigen,xiang,mathieu,fergus,yann@cs.nyu.edu

Cancer

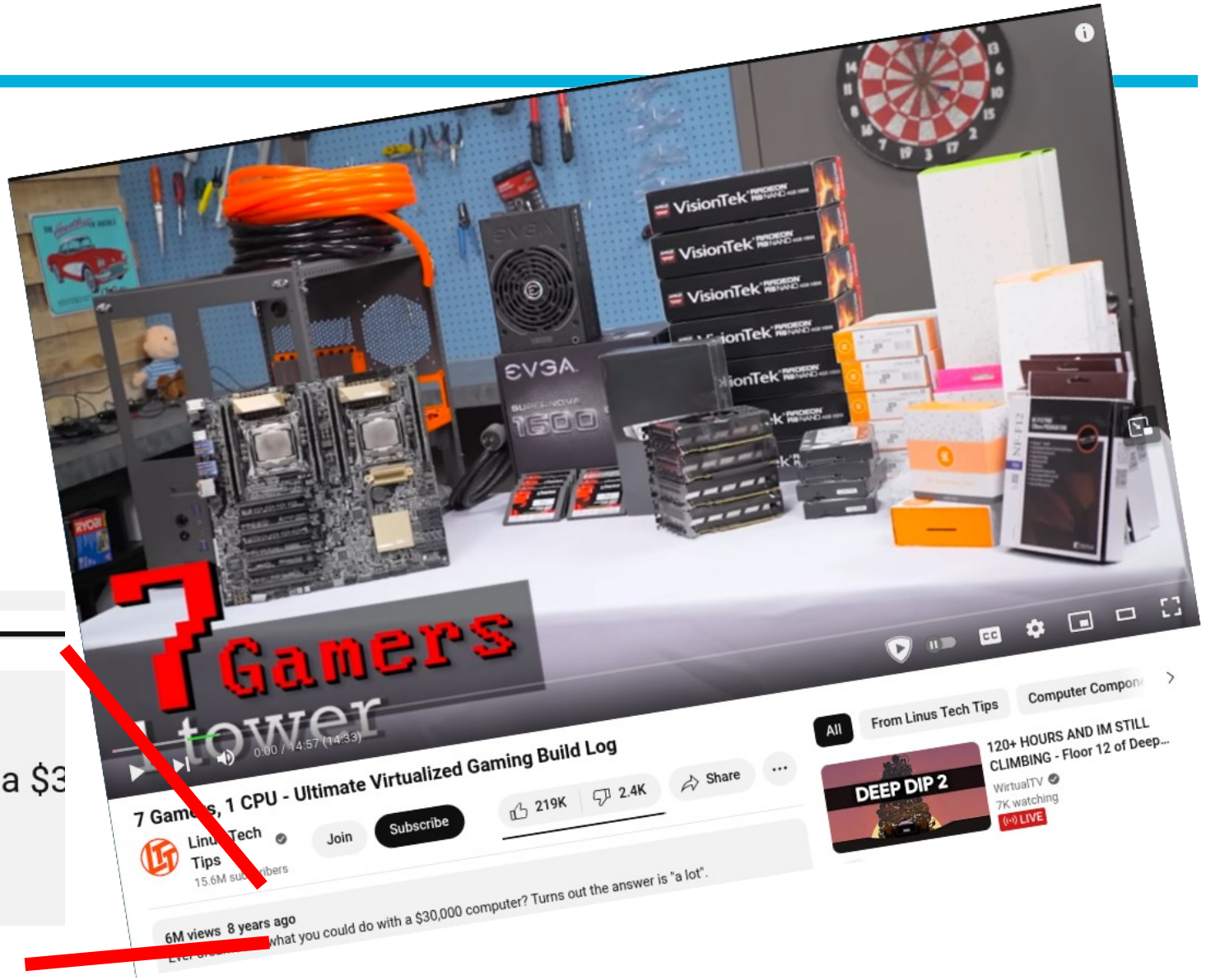
No
ancer

…ovsky, Ilya Sutskever, Geoffrey E Hinton; Advances

Radboudumc
university medical center

Deep Learning

- We needed NVIDIA GPUs
  - How?



2016

6M views 8 years ago

Ever dreamed of what you could do with a $3

answer is "a lot".

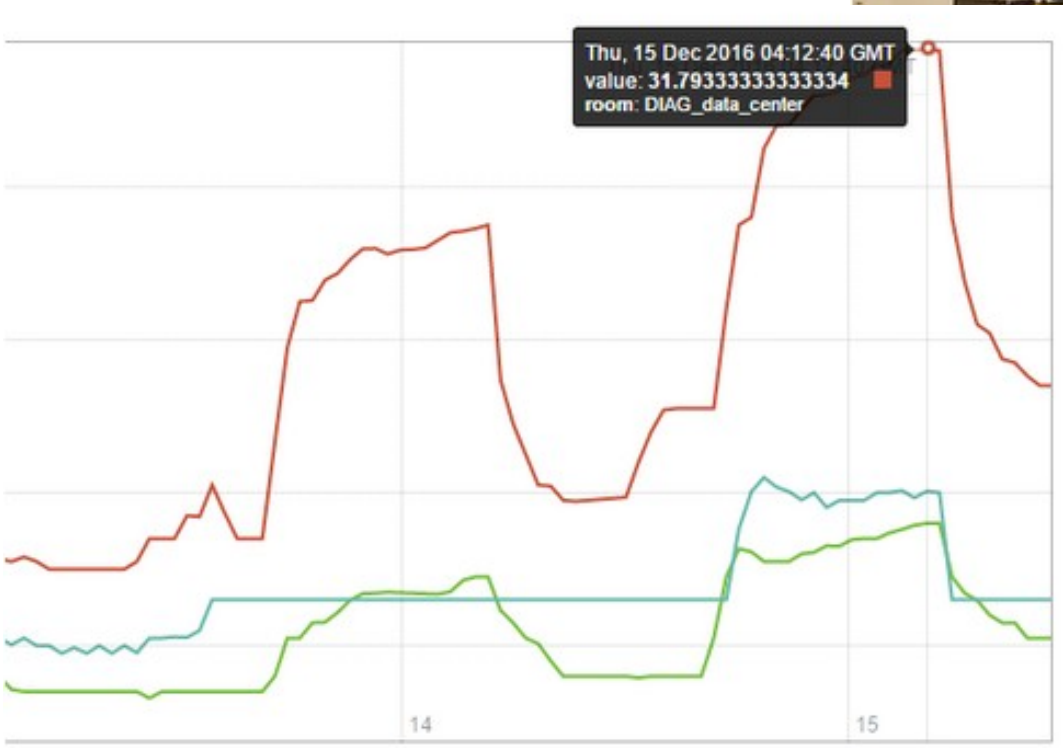
...more

Radboudumc
university medical center

Our first cluster

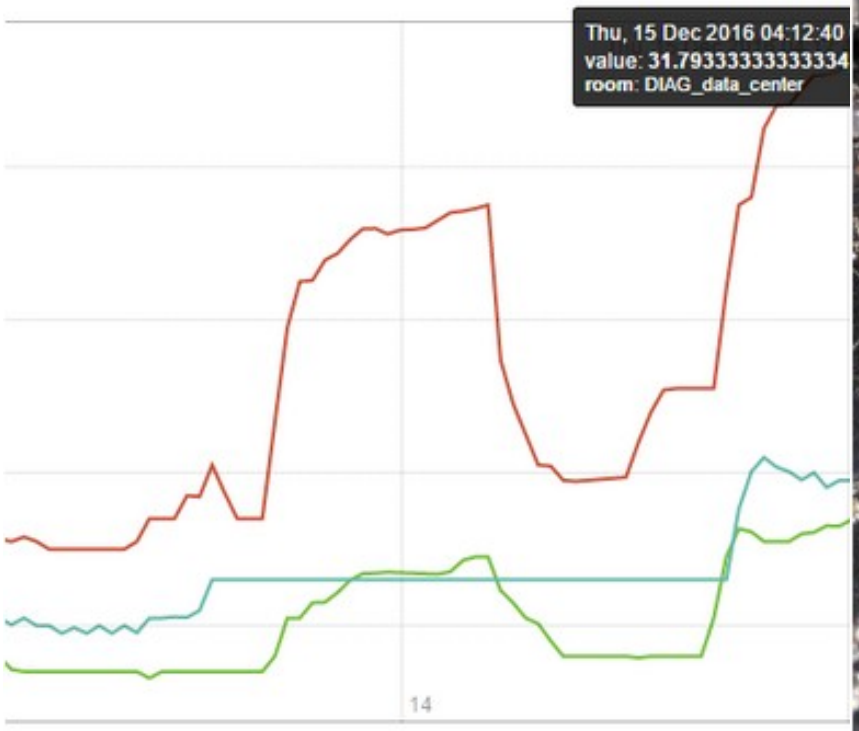


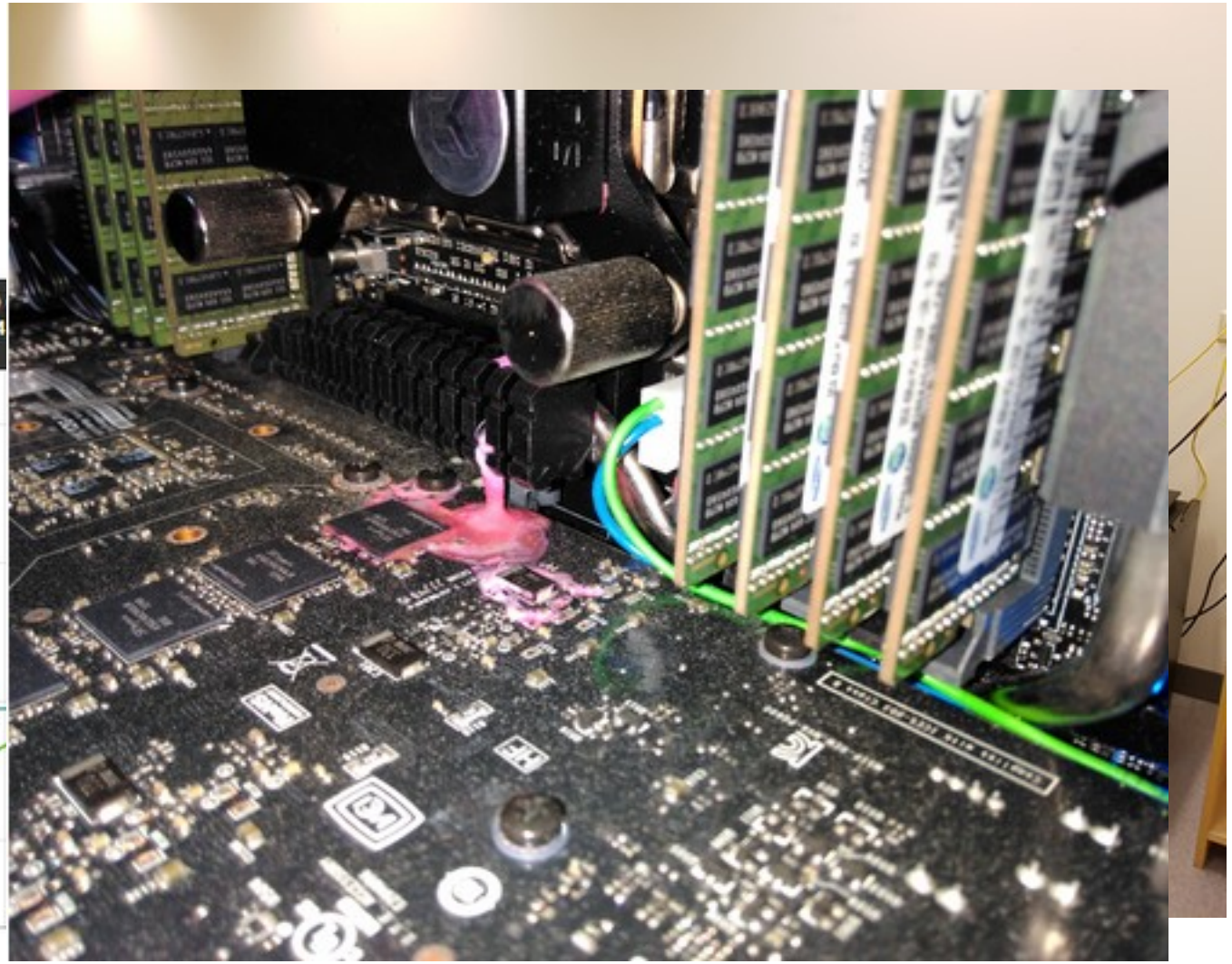
Our first cluster





Our first cluster





Radboudumc
university medical center



Radboudumc
university medical center

Cluster specs

- 100-ish GPUs of various kinds
  - GTX TITAN, GTX 1080, RTX2080, Tesla
  - Up to 8 GPUs per system
- ~1 PB of NAS storage
- User group ~50 active PhD students
- Running:
  - Linux with **docker** container service
  - **Slurm** as a job server, experiments are packaged in containers

Deep learning algorithms
Quirks, properties, and more

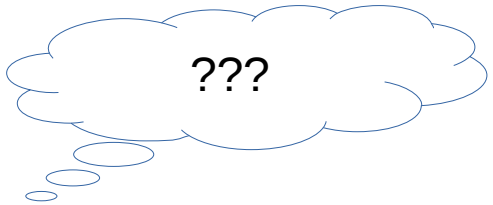
Deep learning quirks

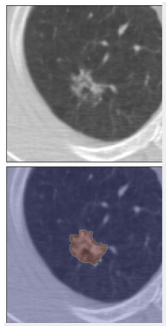
- Code is married to their hardware generations. Code for a GTX980s does not necessarily work on an RTX2080 anymore.
- Algorithms generally depend on external hardware, namely NVIDIA GPUs.
  - Many cutting-edge libraries do not provide CPU implementations anymore.
- Everything is CUDA, and CUDA only.
- Vendor lock-in!

=>



Deep learning quirks





Deep Cancer Analysis Tool

Cancer? **Yes**
Likelihood = 75%

1. Make sure you have a computer with 4 RTX 2080s
2. Make sure that CUDA is installed and working
3. Install docker on the machine
4. Download the experiment docker image from DIAG's repository
5. …..

Grand-challenge
An integrated approach

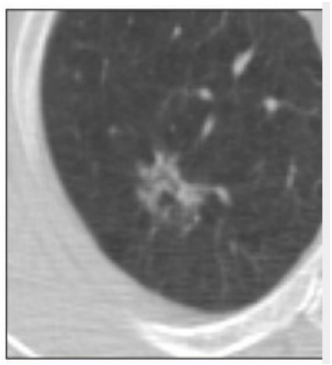
Faculty

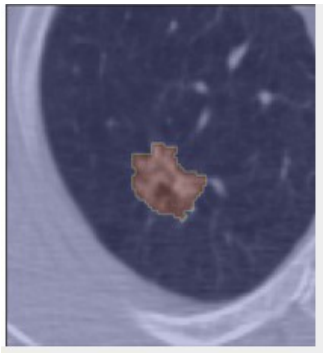


Radboudumc
university medical center

Medical professional's view







Cancer?

- Has data
- Has a task
- Wants something accessible to solve the task for the given data

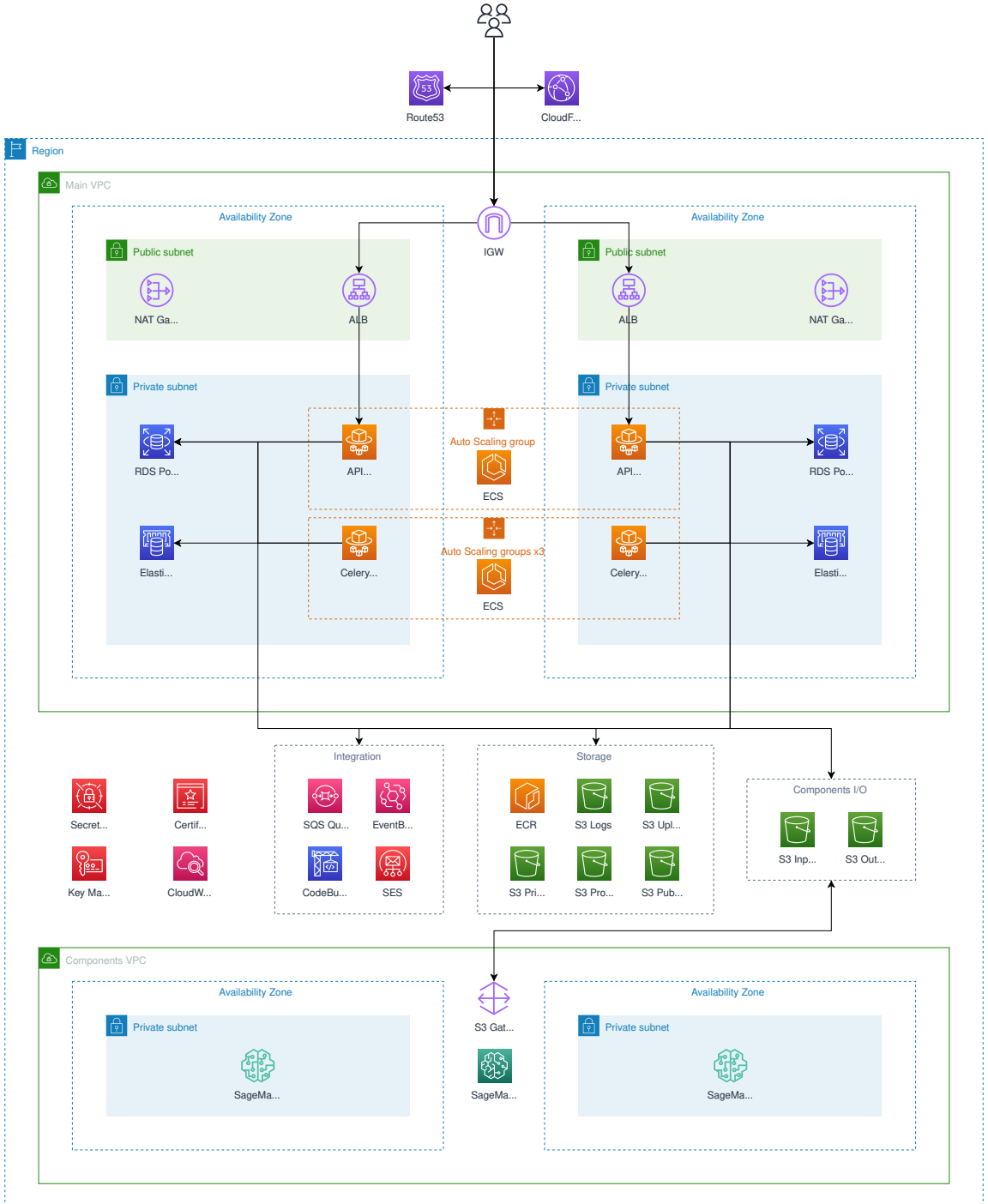
Grand-challenge

- https://grand-challenge.org/
  - Source code: https://github.com/comic/grand-challenge.org/
  - Apache 2.0 license
  - Running on Amazon Web Services
- Viewer component is *not* open source
- ***NOT FOR MEDICAL USE***



Grand-challenge

- Platform for running research experiments
- Arbitrary code execution
  - Sandboxed
- GPU resources

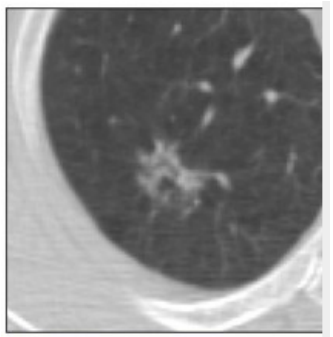


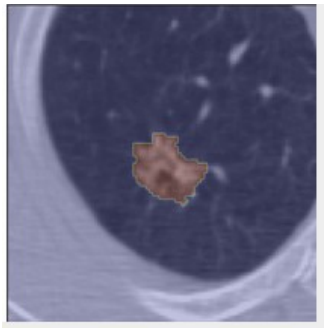
Medical professional's view



Has data

- Has a task
- Wants something accessible to solve the task for the given data





Cancer?

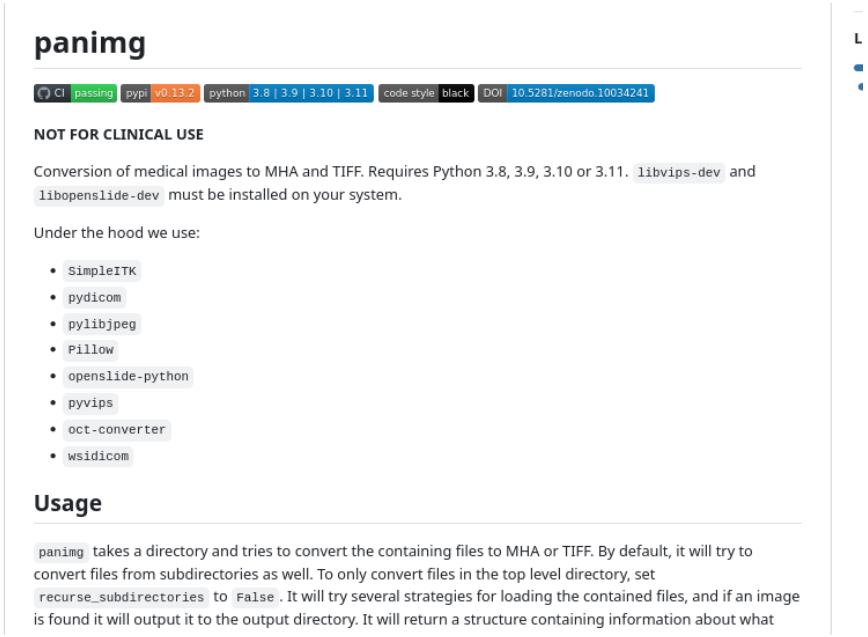
Data



- Upload data to grand-challenge!
  - CT images, X-Ray images, Timeseries data, Videos, Pathology images
- ***Legal issues to be solved by the healthcare institutes:***
  - ***Privacy of patient data***

panimg

- https://github.com/DIAGNijmegen/rse-panimg
- Conversion tool for medical images into simpler formats for researchers
  - MetaIO for most image types
  - TIFF + DZI for multi resolution images



panimg

CI passing | pypi v0.13.2 | python 3.8 | 3.9 | 3.10 | 3.11 | code style black | DOI 10.5281/zenodo.10034241

NOT FOR CLINICAL USE

Conversion of medical images to MHA and TIFF. Requires Python 3.8, 3.9, 3.10 or 3.11. `libvips-dev` and `libopenslide-dev` must be installed on your system.

Under the hood we use:

- `SimpleITK`
- `pydicom`
- `pylibjpeg`
- `Pillow`
- `openslide-python`
- `pyvips`
- `oct-converter`
- `wsidicom`

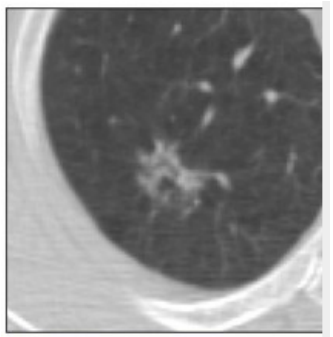
Usage

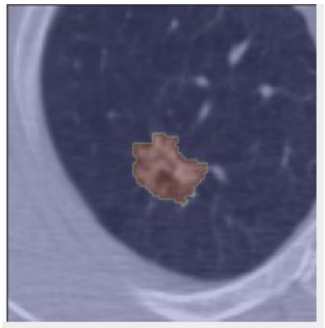
`panimg` takes a directory and tries to convert the containing files to MHA or TIFF. By default, it will try to convert files from subdirectories as well. To only convert files in the top level directory, set `recurse_subdirectories` to `False`. It will try several strategies for loading the contained files, and if an image is found it will output it to the output directory. It will return a structure containing information about what

Medical professional's view



- Has data
- **Has a task**
- Wants something accessible to solve the task for the given data

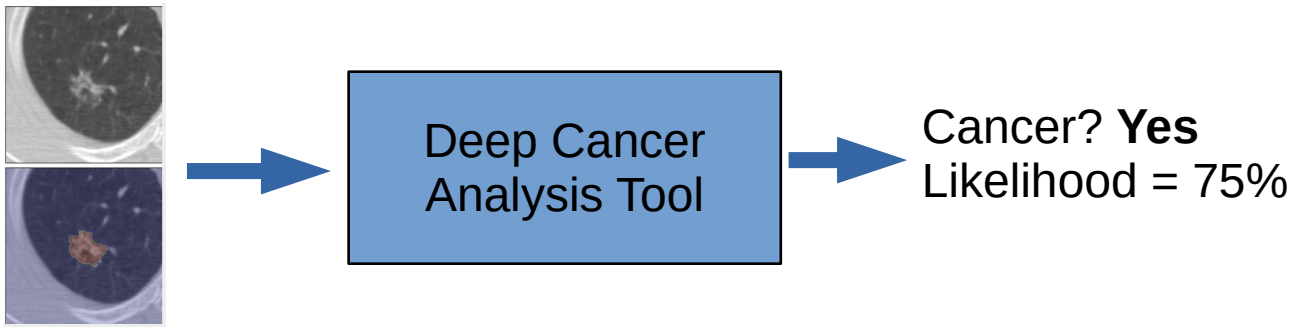




Cancer?

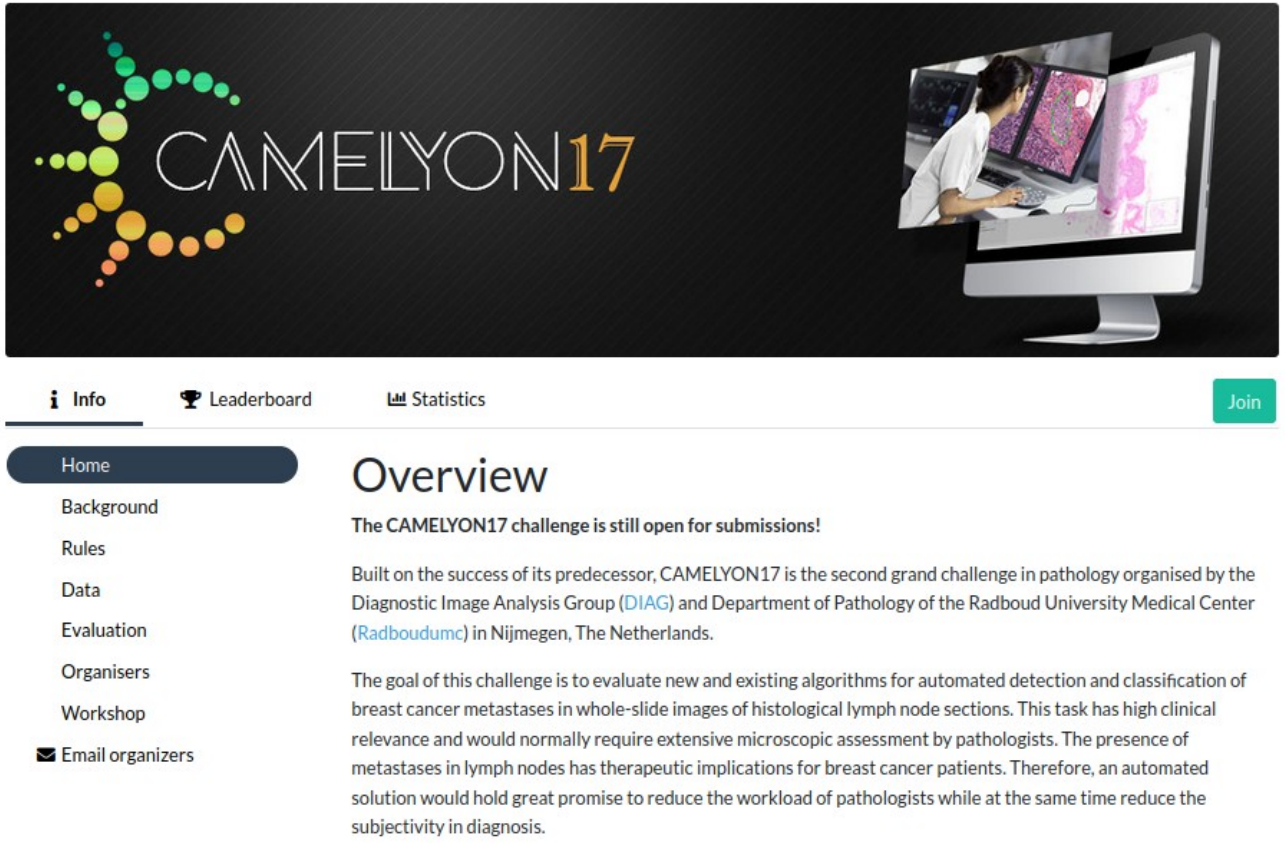
Challenges and algorithms

- Algorithms take medical data and compute predictions



- Challenges are public comptetitions that allow for benchmarking of algorithms

Challenge example: CAMELYON17

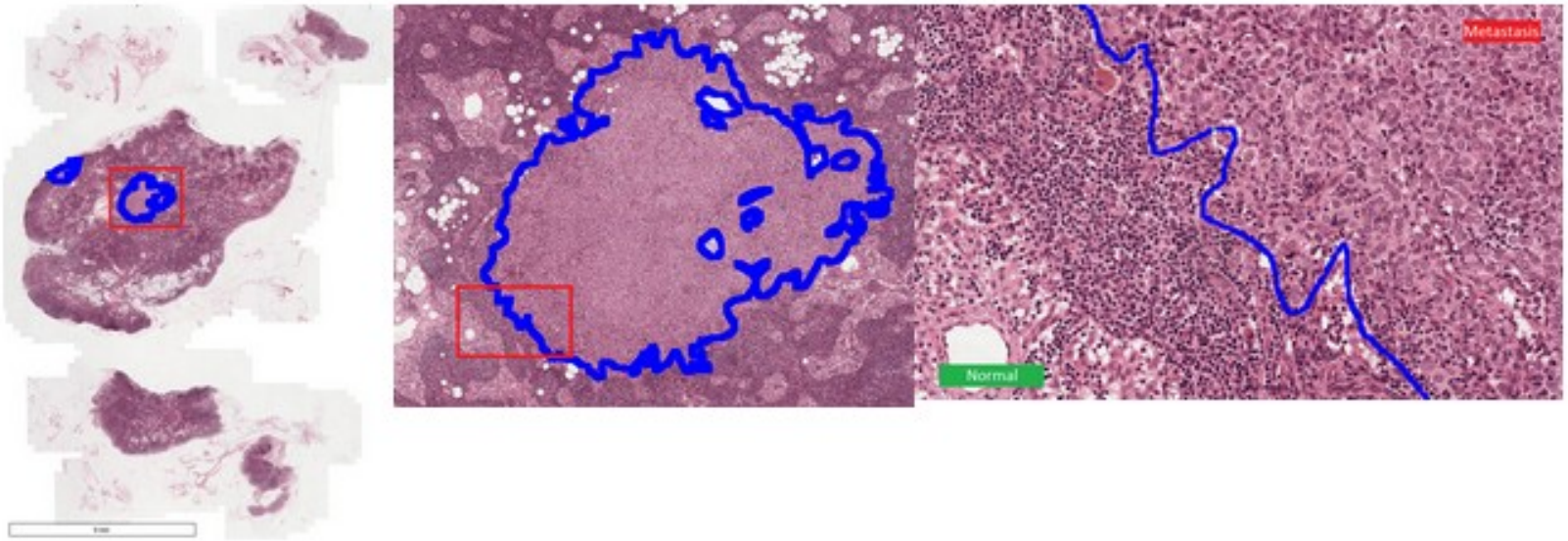


https://camelyon17.grand-challenge.org/

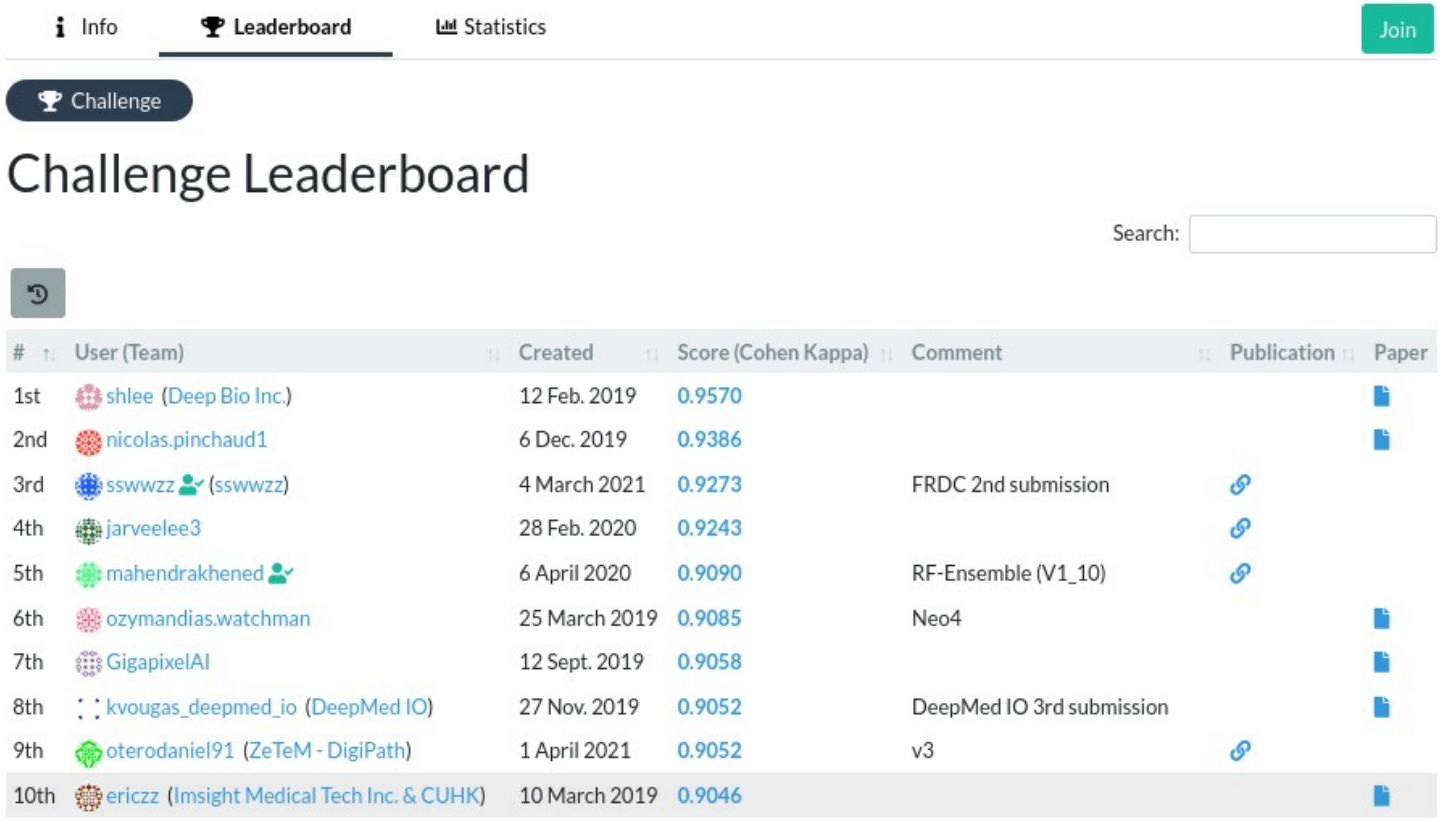
Radboudumc
university medical center

Challenge example: CAMELYON17

Example of a metastatic region



Challenge example: CAMELYON17



Info | Leaderboard | Statistics | Join

Challenge

Challenge Leaderboard

Search:

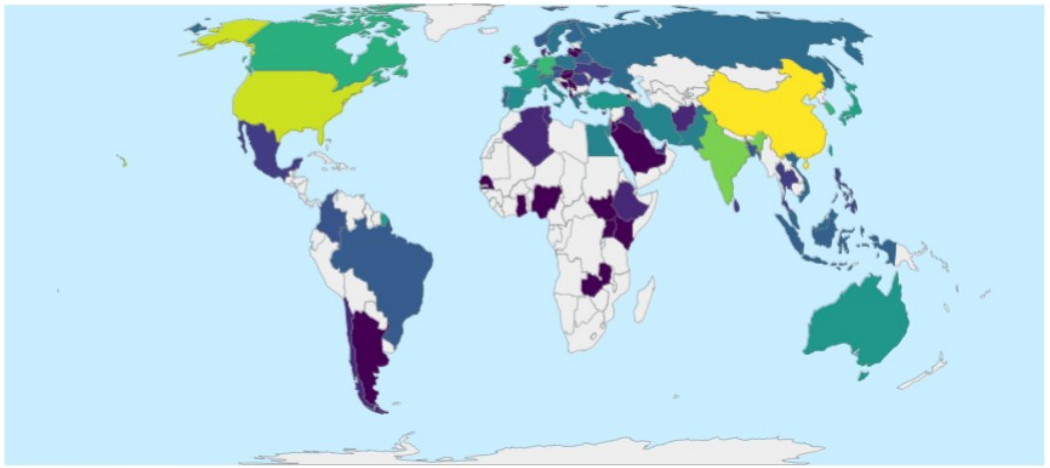
#	User (Team)	Created	Score (Cohen Kappa)	Comment	Publication	Paper
1st	shlee (Deep Bio Inc.)	12 Feb. 2019	0.9570			
2nd	nicolas.pinchaud1	6 Dec. 2019	0.9386			
3rd	sswwzz (sswwzz)	4 March 2021	0.9273	FRDC 2nd submission		
4th	jarveelee3	28 Feb. 2020	0.9243			
5th	mahendrakhened	6 April 2020	0.9090	RF-Ensemble (V1_10)		
6th	ozymandias.watchman	25 March 2019	0.9085	Neo4		
7th	GigapixelAI	12 Sept. 2019	0.9058			
8th	kvougas_deepmed_io (DeepMed IO)	27 Nov. 2019	0.9052	DeepMed IO 3rd submission		
9th	oterodaniel91 (ZeTeM - DigiPath)	1 April 2021	0.9052	v3		
10th	ericzz (Imsight Medical Tech Inc. & CUHK)	10 March 2019	0.9046			

Radboudumc
university medical center

Challenge example: CAMELYON17

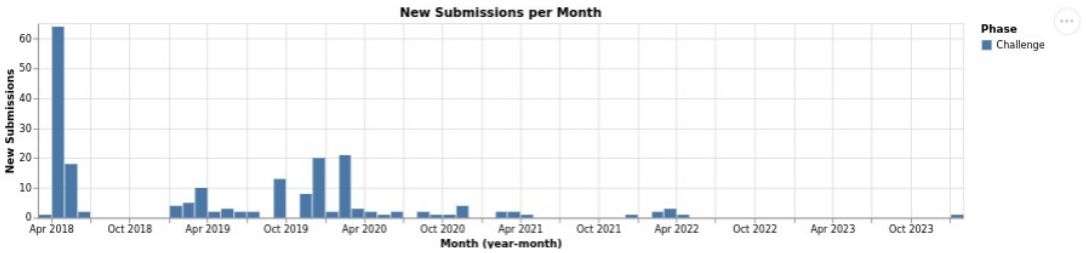
Statistics

Participants



Total participants: 2,025

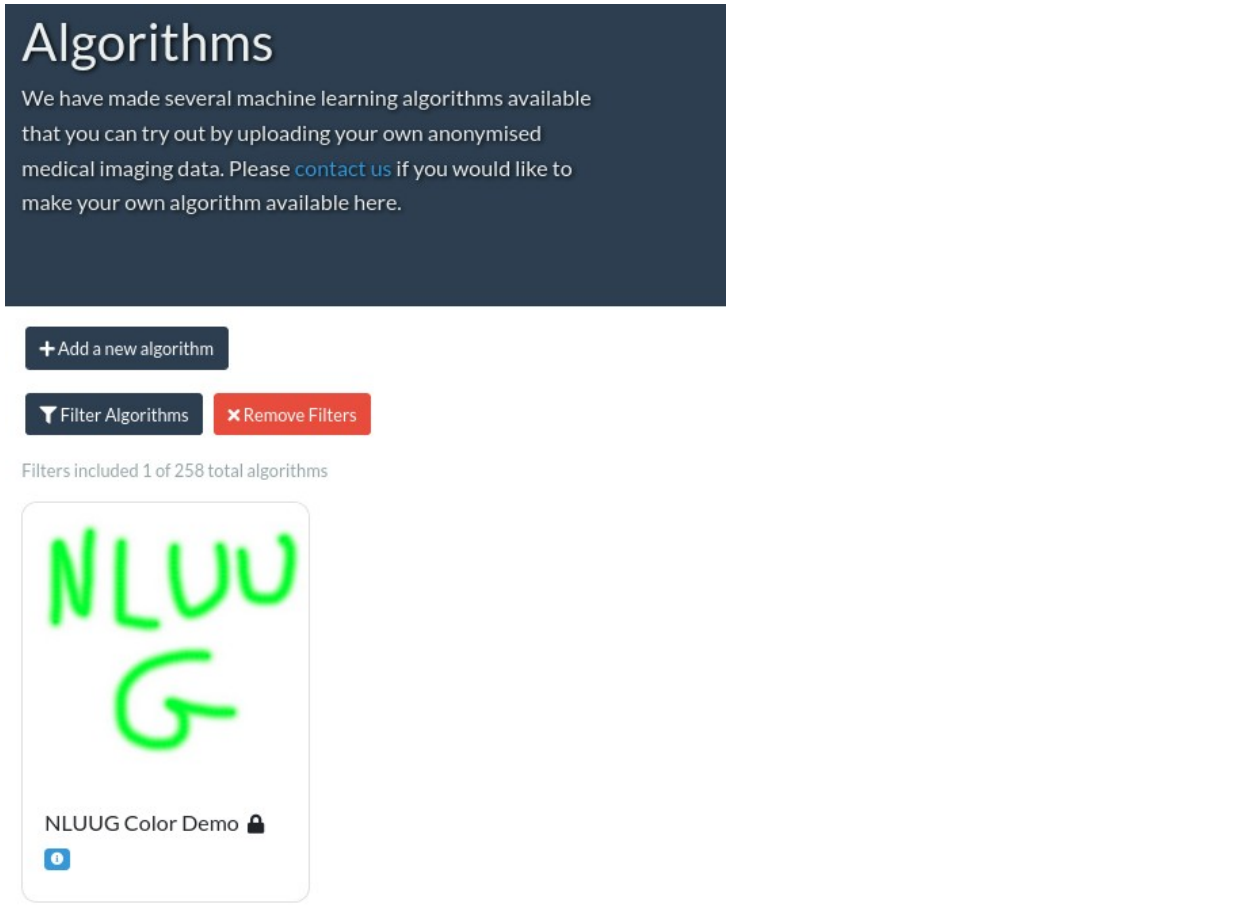
Submissions



Radboudumc
university medical center

Algorithms

- Can be part of a challenge
- Can be run separately



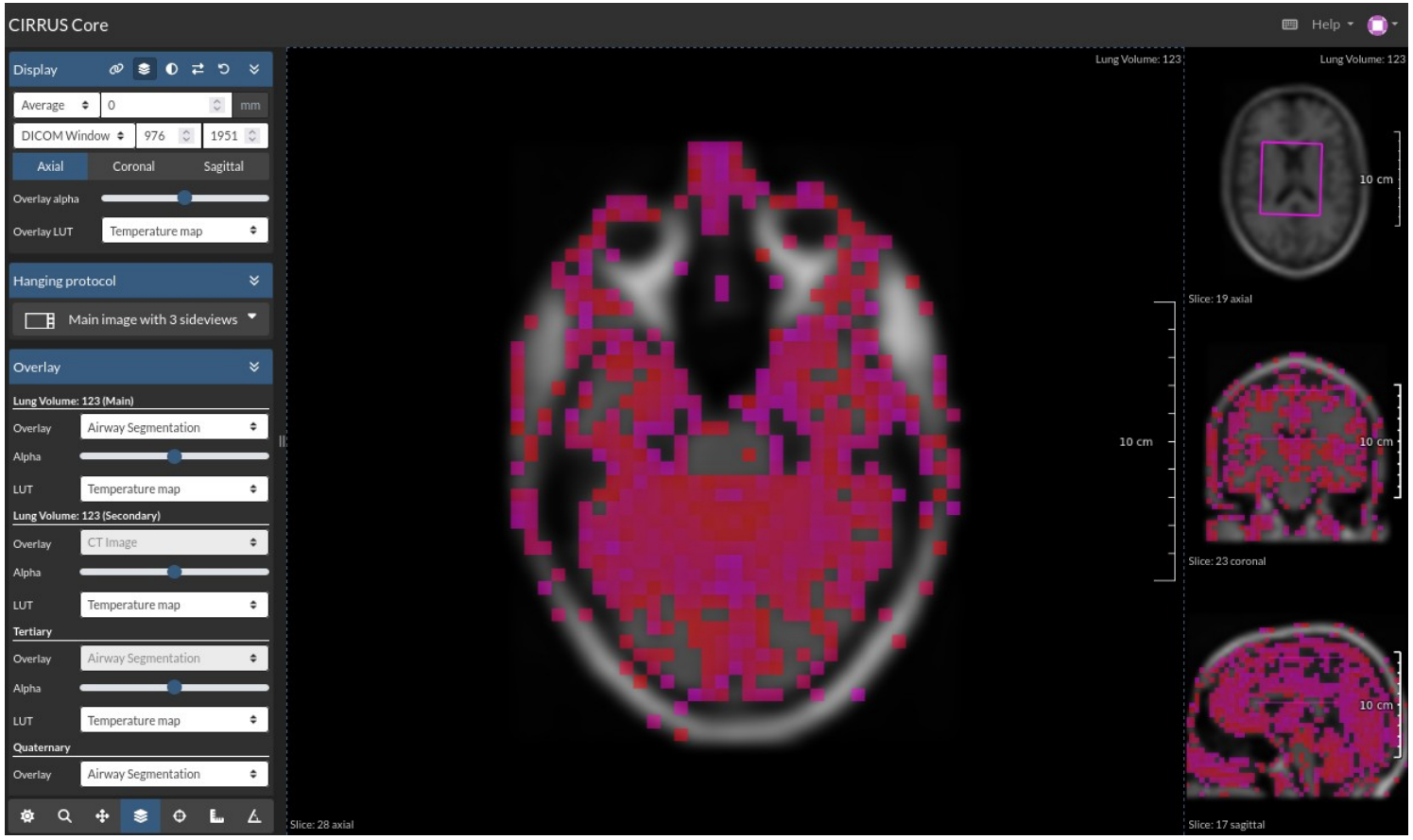
https://grand-challenge.org/algorithms/nluug-color-demo/

Radboudumc
university medical center

(Demo)

Viewer

- Can be used to collect data
- Or can be used to inspect results of deep learning algorithms

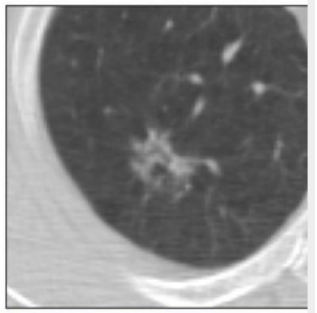


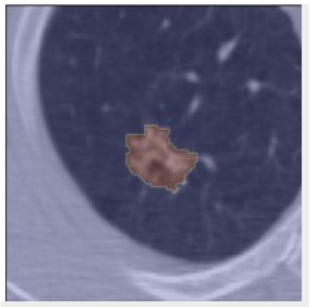
Radboudumc
university medical center

(Demo)

Medical professional's view







Cancer?

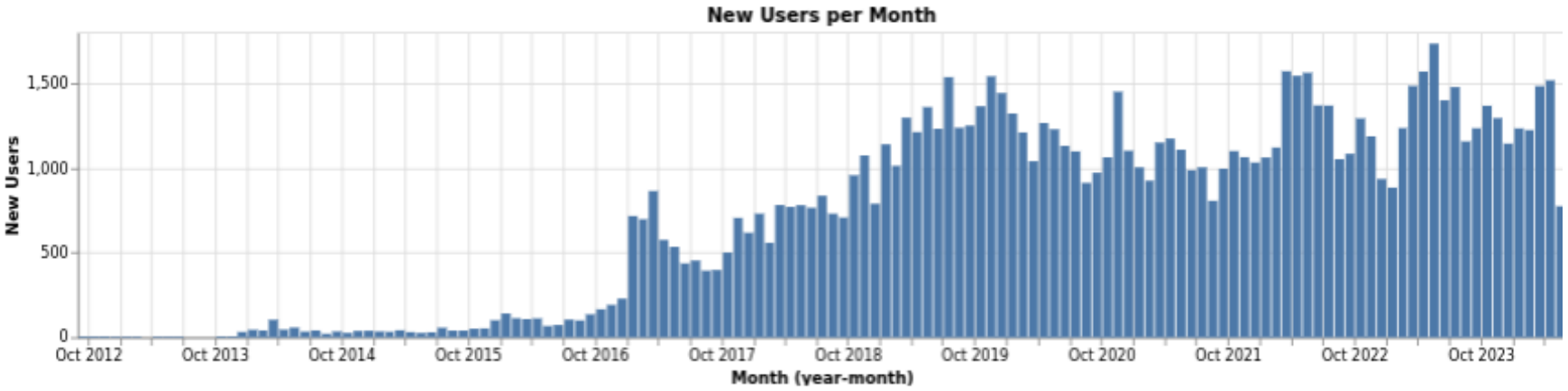
- Has data
- Has a task
- Wants something accessible to solve the task for t
  given data



Radboudumc
university medical center

Deep learning and applied science

- Including yout target audience is very important for deep learning.
- grand-challenge.org is our solution for making deep learning research accessible to a wider audience.



Total users: 98,266
Users logged in (last 30 days): 3,656