Brain Imaging Data Analysis with MATLAB: from Pictures to Knowledge

MATLAB EXPO, Bern, June 22, 2017

Dr. Henry Lütcke, Scientific IT Services (SIS), ETH Zürich

Scientific IT Services at ETH Zürich

- Founded in 2013 as part of central IT
- HPC experts, software developers, scientific computing specialists
- 34 team members at 2 sites (Zürich, Basel)
- >50% with PhDs and research experience

“We work closely with ETH researchers to enable research and improve efficiency by providing first-class scientific computing services.”

Outline

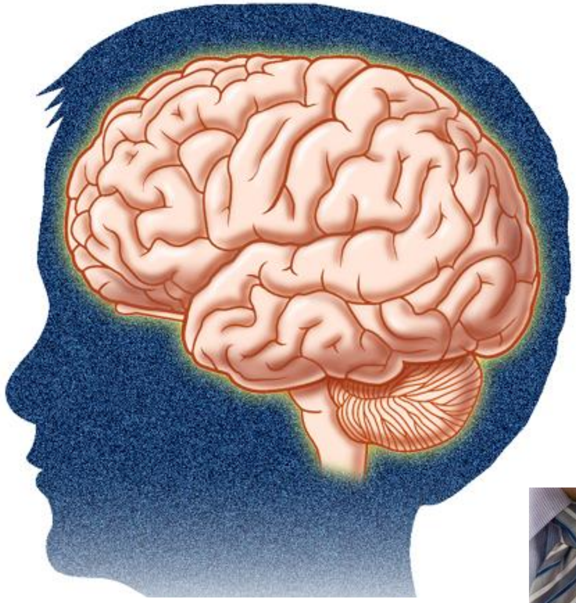
- Importance of quantitative imaging analysis in neuroscience
- Image analysis examples
  - Signal extraction from noisy neuronal activity measurements
  - Machine learning based quantification of neuronal network activity
- From small to Big Data
  - Scalable analysis with cluster computing

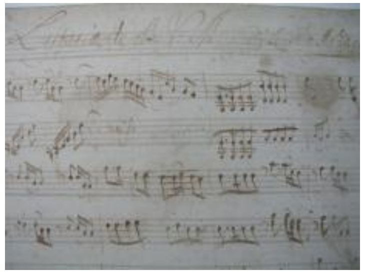
Neuroscience: Understanding the Brain

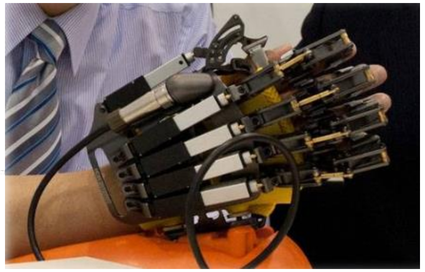
Movement

Art

Language



Thinking

Music



Technical Discoveries

What is the brain made of? How does it work?

Neuroscience: Understanding the Brain

"As long as our brain is a mystery, the universe – as reflection of the structure of the brain – will also remain a mystery."

Santiago Ramón y Cajal (1852-1934)

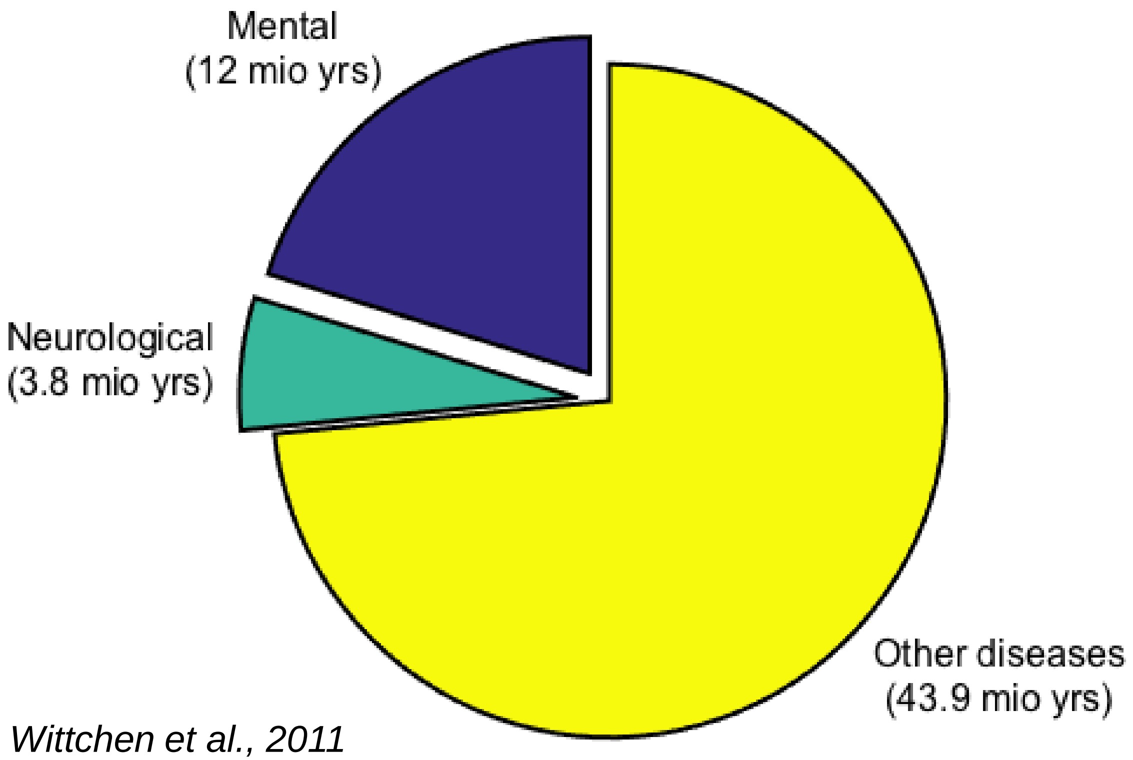
Burden & Cost of Brain Disease

Deep-brain stimulation in Parkinson’s disease

youtube.com/watch?v=mO3C6iTpSGo

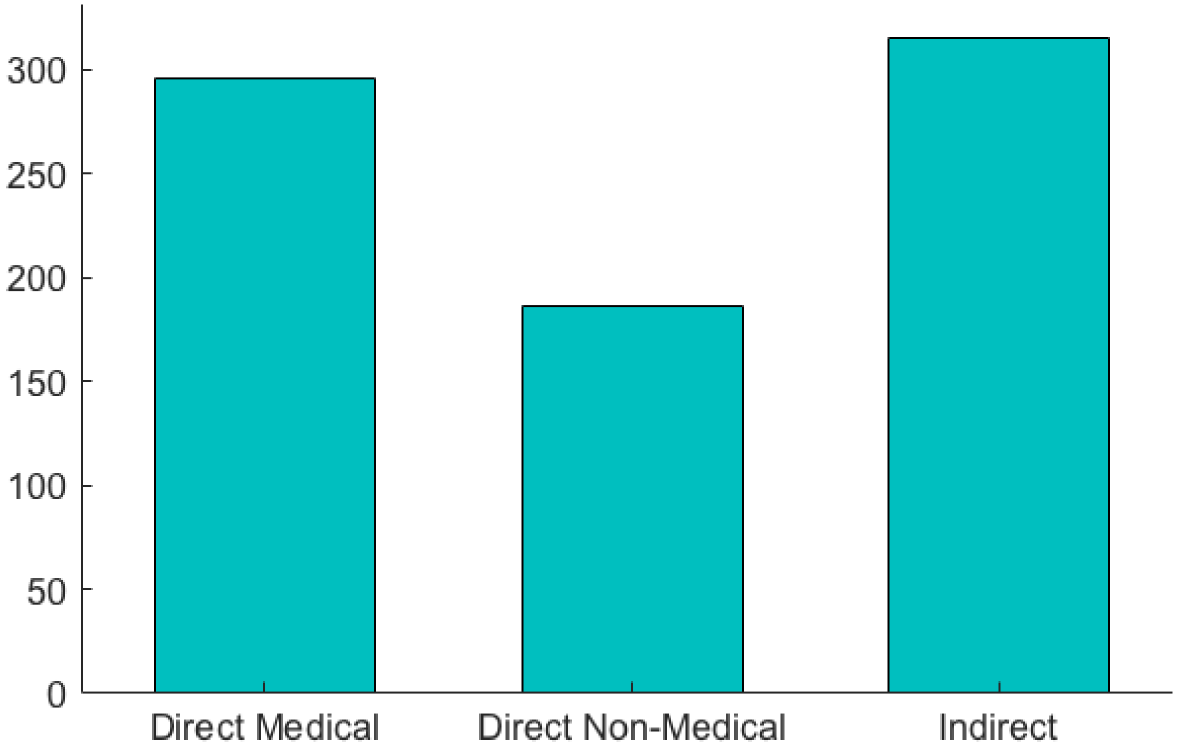
Burden & Cost of Brain Disease

The burden of brain disease in Europe
(Quantified as Disability Adjusted Life Years Lost)



Wittchen et al., 2011

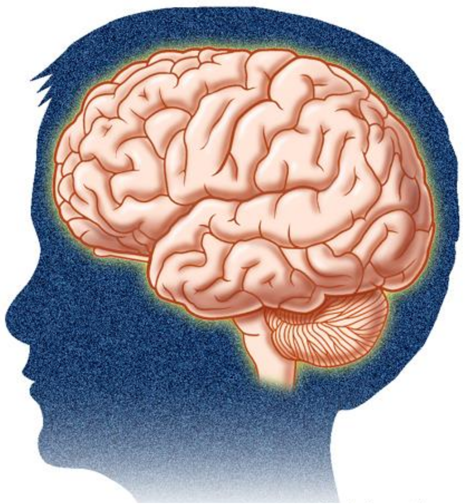
The cost of brain disease in Europe
(In billion €, 2010)



Olesen et al., 2012

Disorders of the brain are extremely disabling and incur enormous costs for patients, relatives and society!

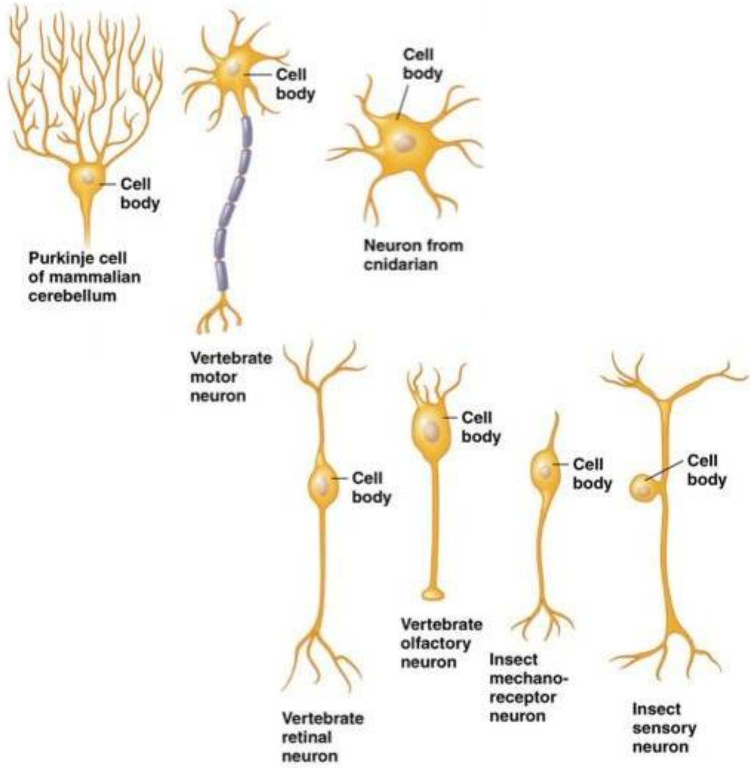
The Brain consists of a Large Network of Neurons



10^{12}

(Trillion)

Neurons



10^{15}

(Quadrillion)

Synapses

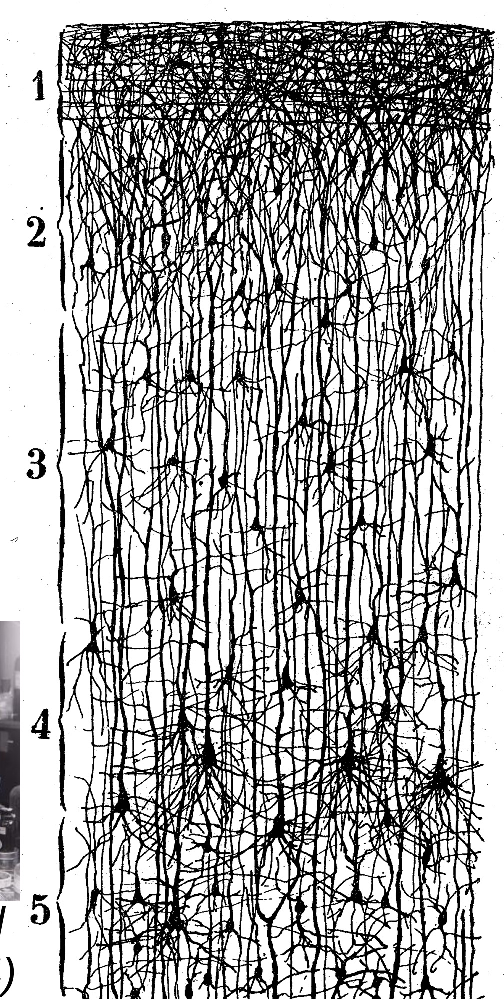
The brain consists of a large number of diverse nerve cells (neurons), which communicate via specialized contacts (synapses).

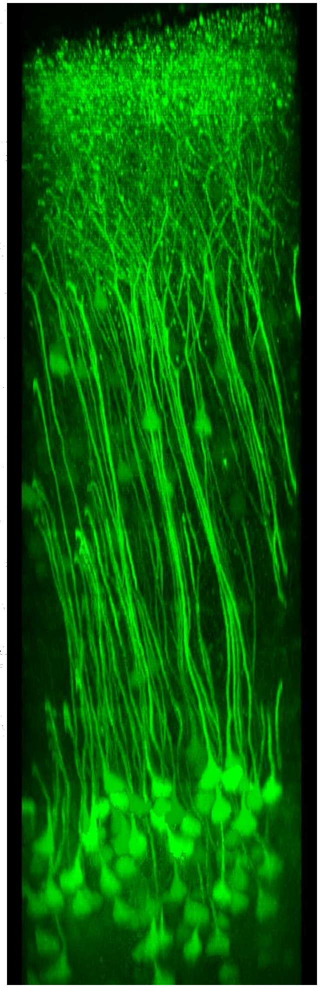
Imaging plays a critical role in revealing brain structure and function.

Importance of Imaging in Neuroscience

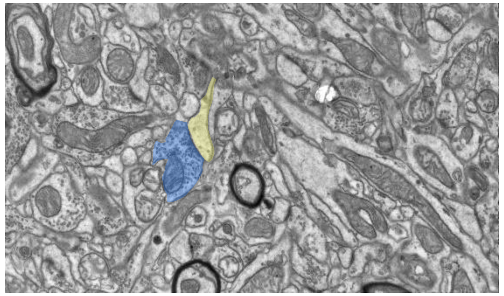
around 1900 around 2000

Ramón y Cajal
(1852-1934)

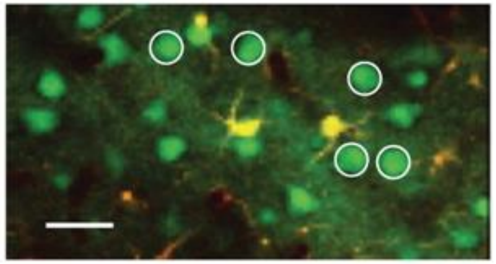


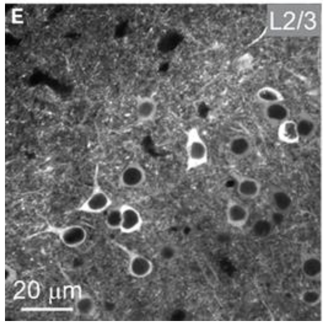


Imaging at different scales

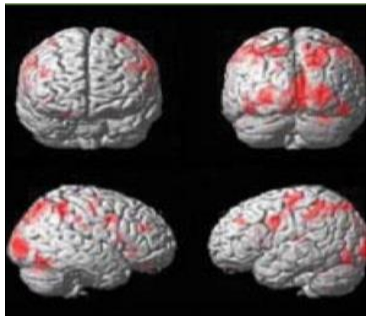


Single cells / sub-cellular (microscopic)





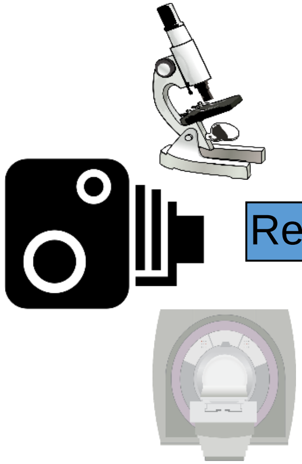
Networks (mesoscopic)



Brain (macroscopic)

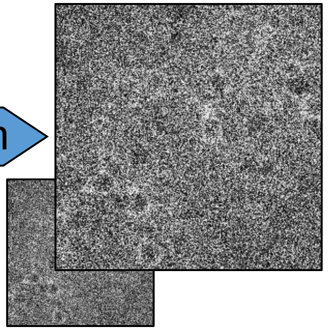
Generic Workflow for Image Analysis

Acquisition



Reconstruction

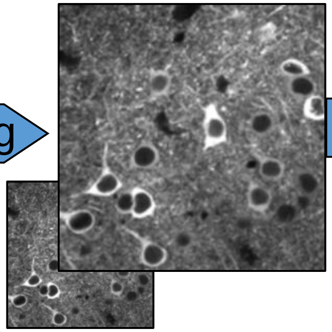
Raw Data



Noise,
Artefacts, …

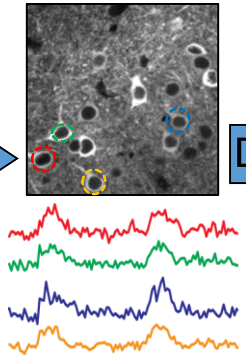
Preprocessing

Filtered Data



Reduction

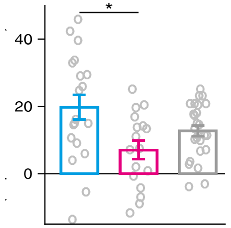
Reduced Data



Regions of interest,
Interpolation, …

Data Analytics

Knowledge



Confirm / reject
hypotheses

Customized
(Manufacturer, in-house)

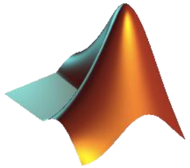


Image Proc.
Signal Proc.
…

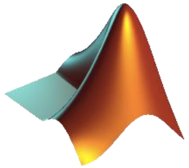
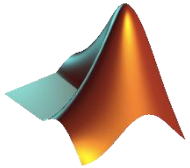
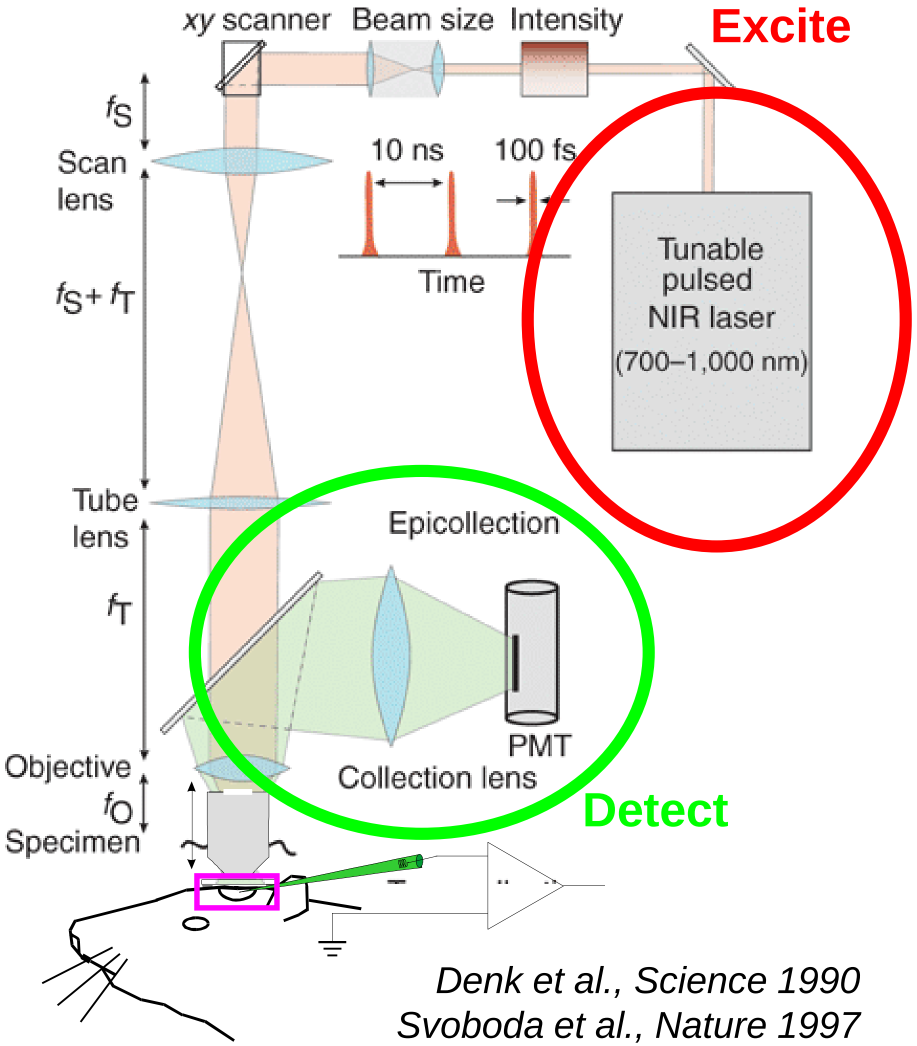


Image Proc.
GUIs
…



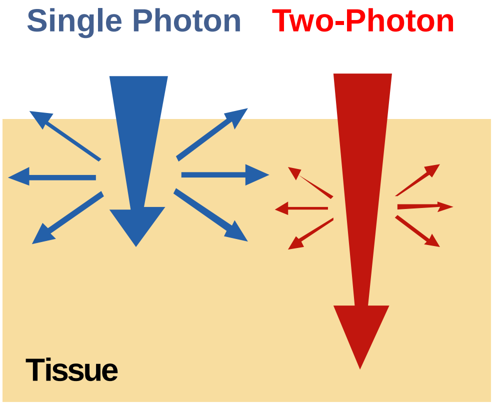
Statistics & ML
Curve Fitting
…

In vivo Two-Photon Microscopy

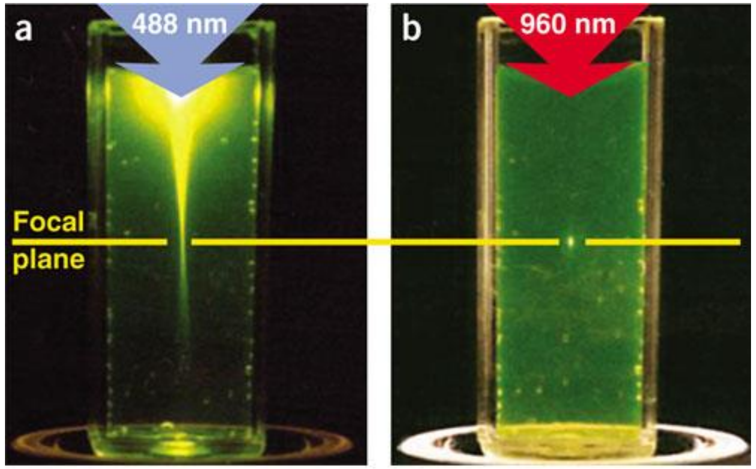


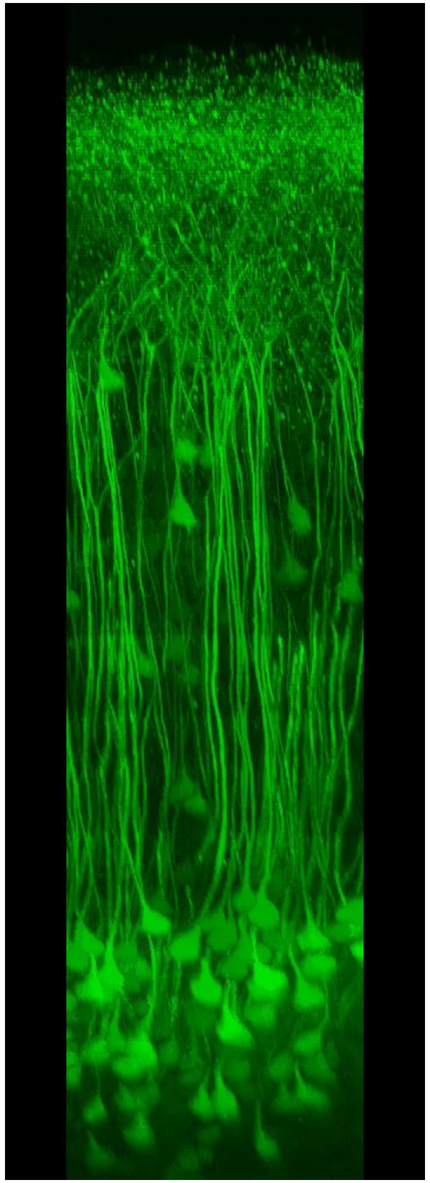
Denk et al., Science 1990
Svoboda et al., Nature 1997

Reduced Scattering

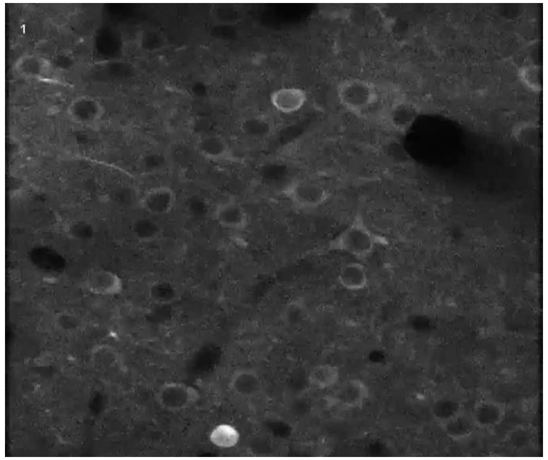


Point Excitation

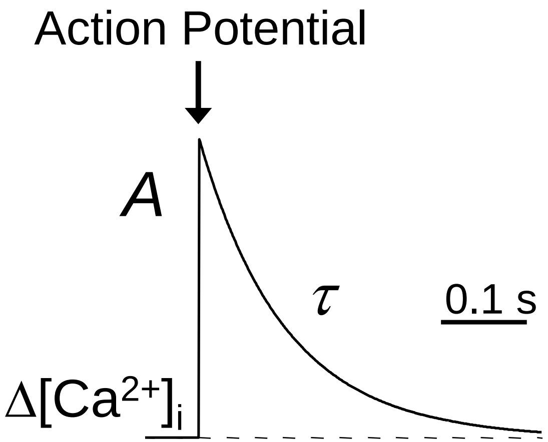




Example 1: Denoising and signal extraction



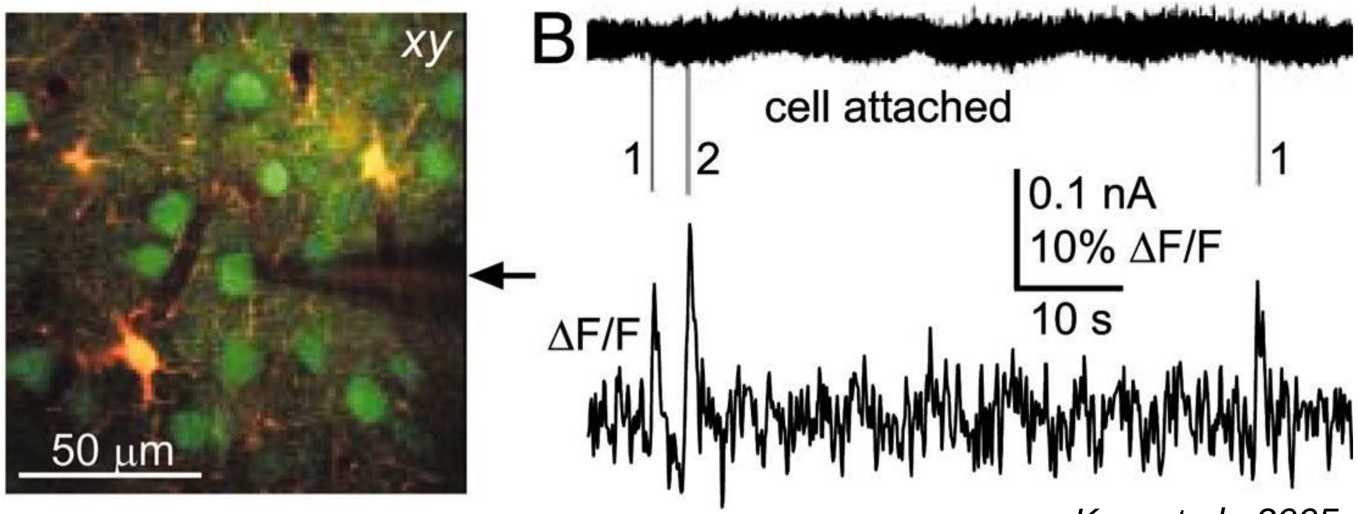
T. Rose, MPI Neurobiology



Automated peeling algorithm for spike train reconstruction

B. Grewe and F. Helmchen, Brain Research Institute, University of Zürich

Effect of noise

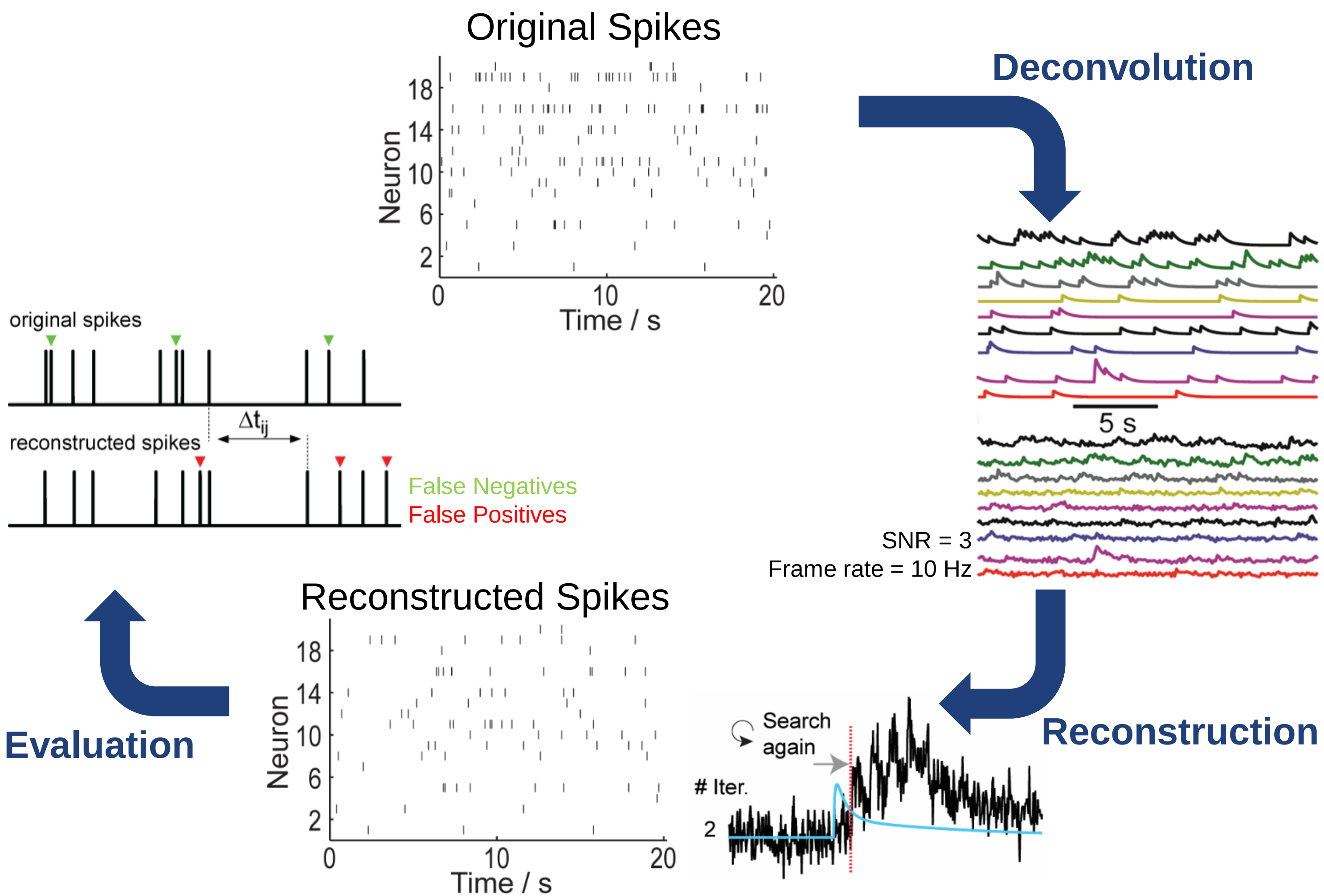


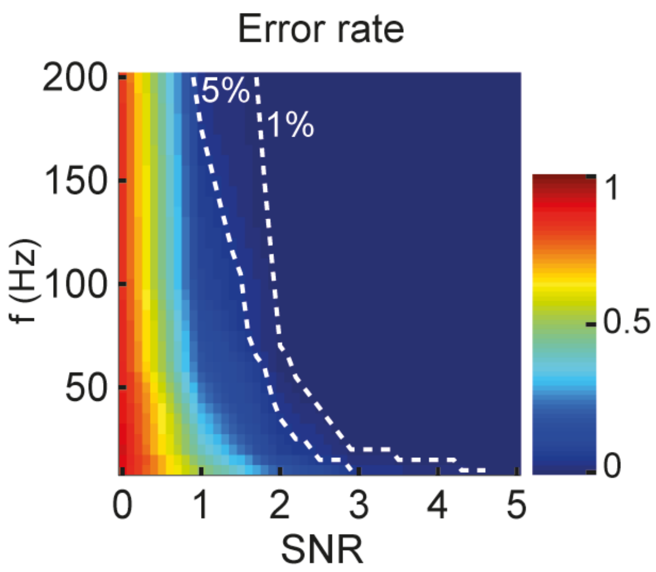
Kerr et al., 2005

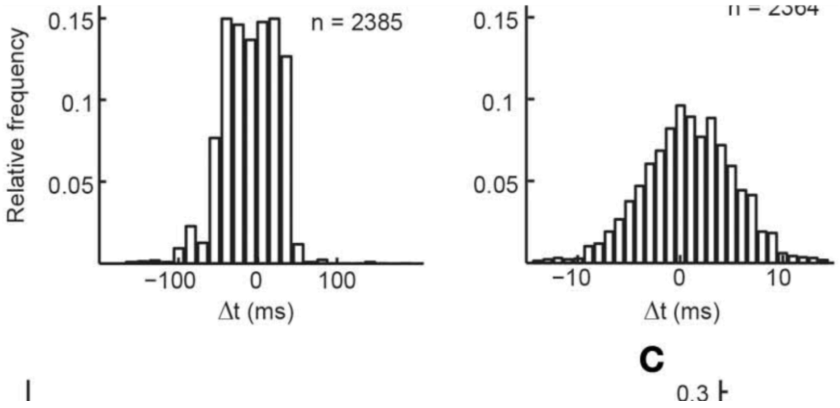
Other algorithms (all MATLAB-based):

- OOPSI (Vogelstein et al., 2010)
- MLspike (Deneux et al., 2016)
- CNMF (Pnevmatikakis et al., 2016)

Denoising and signal extraction



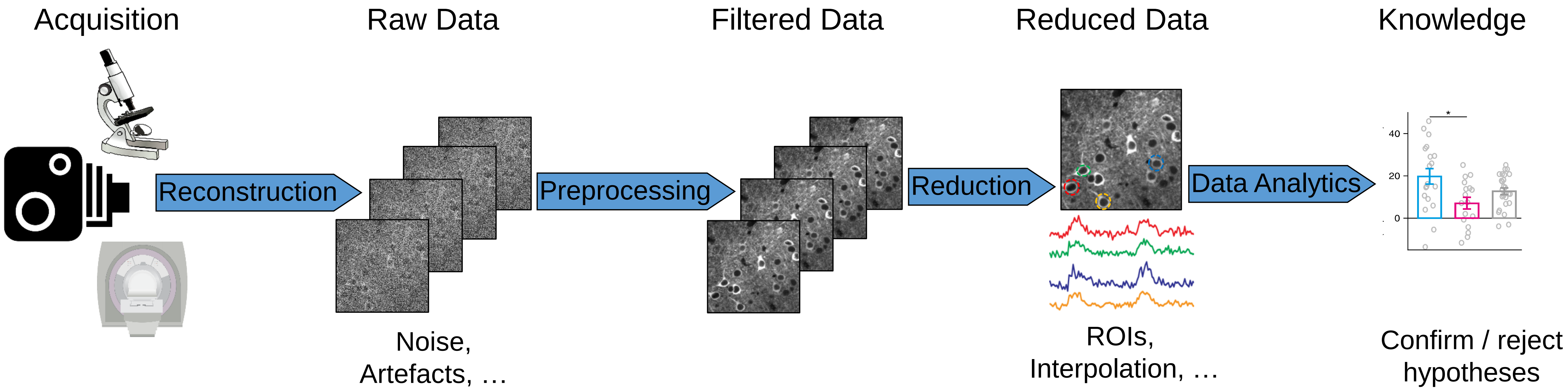


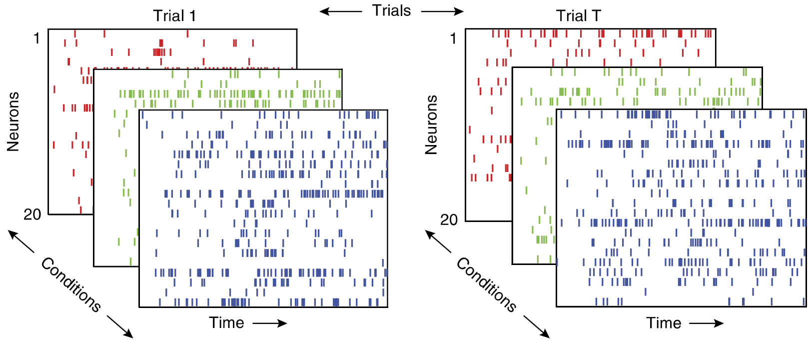


See Lütcke et al., 2013

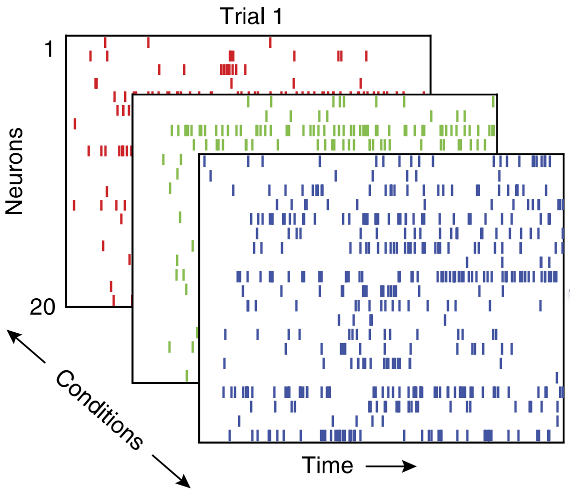
A MATLAB-based simulation framework for systematic evaluation of reconstruction algorithms.

Generic Workflow for Image Analysis



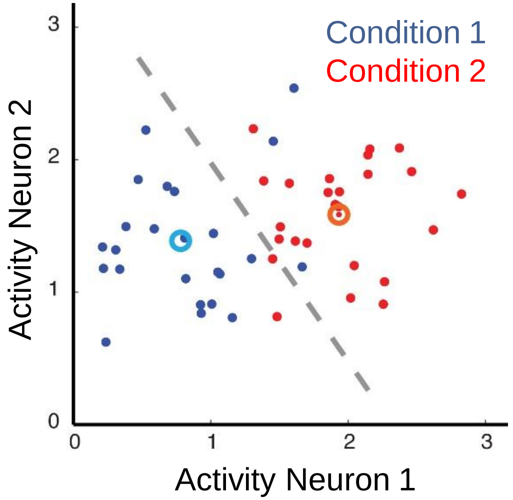


Quantifying Network Activity with Machine Learning



Population vector in N-dimensional space (N ... no. of neurons)

For N = 2:



Classification Algorithms

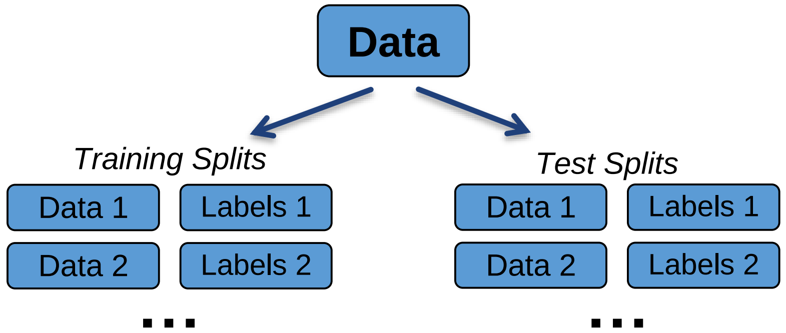
Support Vector Machine

Naive Bayes

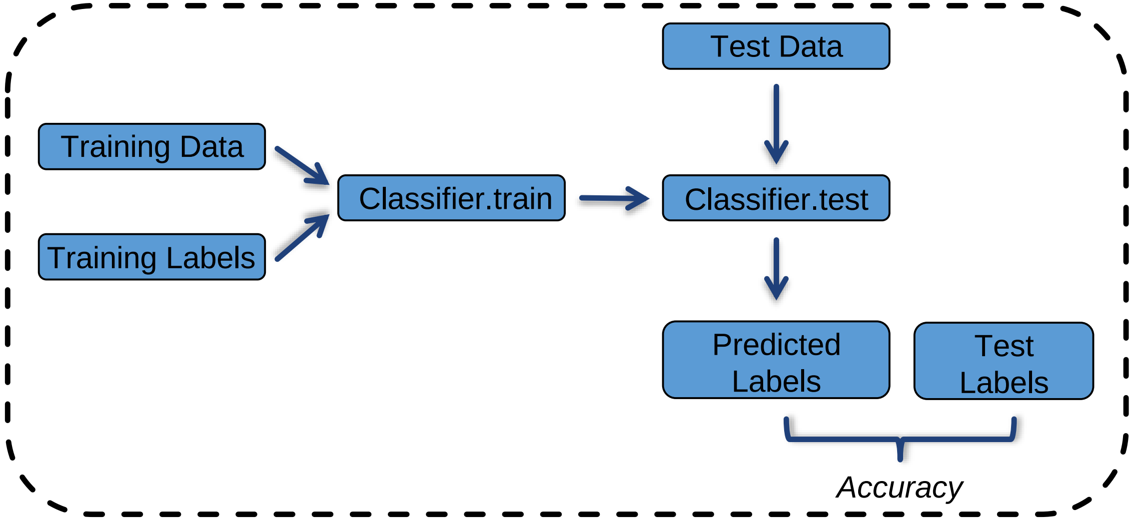
Random Forest

Statistics & Machine Learning Toolbox

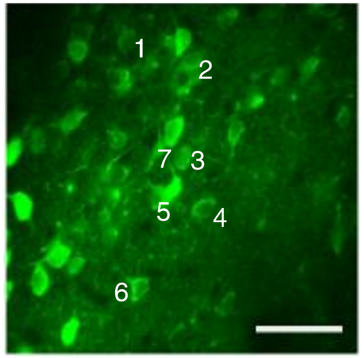
Supervised Learning Approach

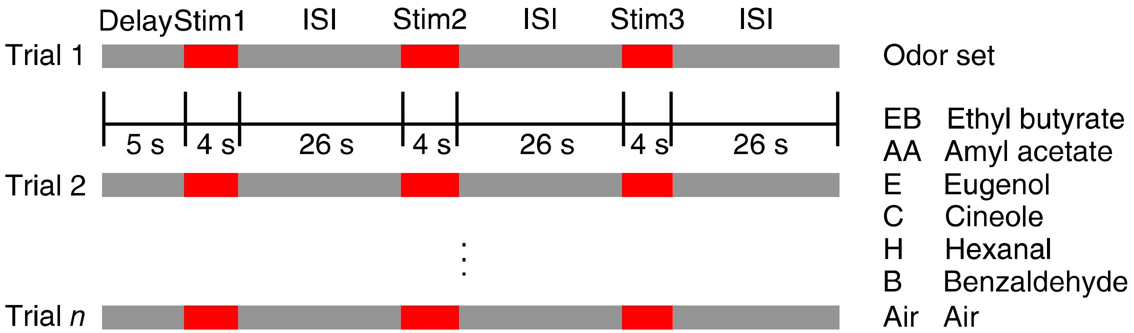


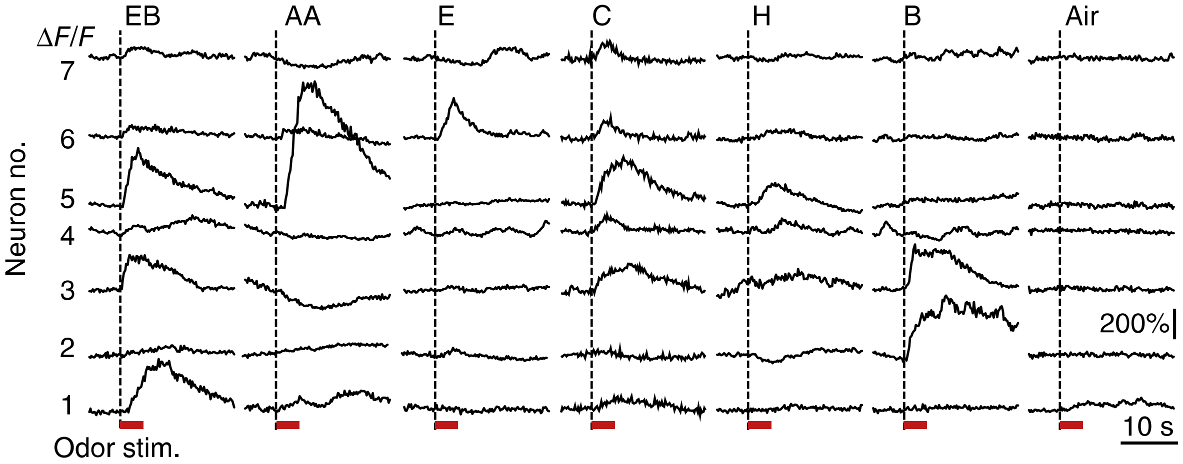
For each cross-validation split

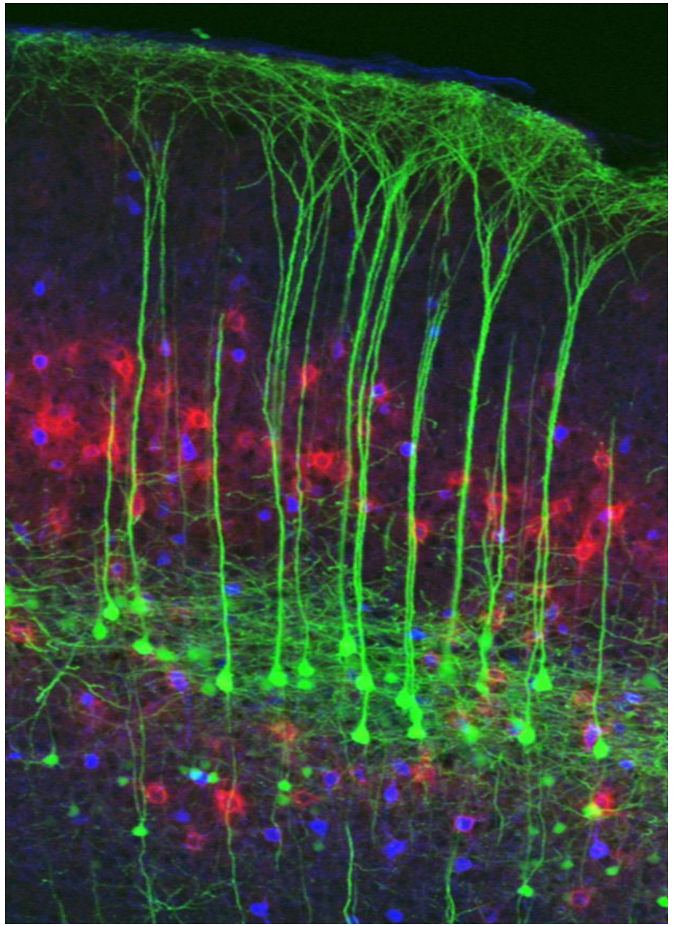


Example 2: Quantifying Network Activity with Machine Learning







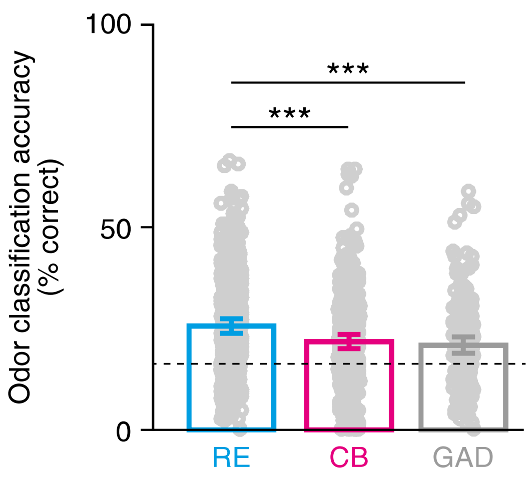


Leitner et al., 2016

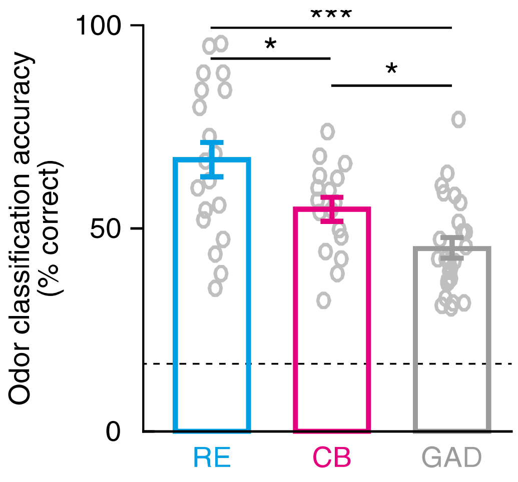
How is odor information encoded by different neuronal sub-networks?

Quantifying Network Activity with Machine Learning

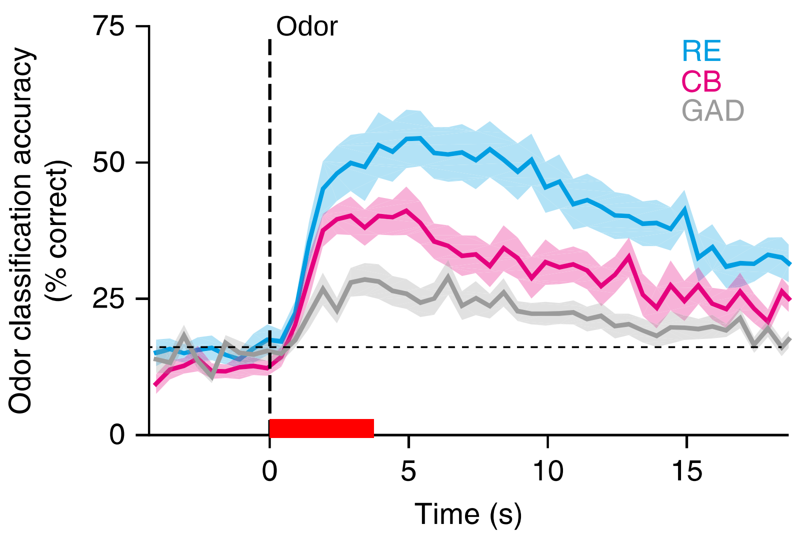
Classification Accuracy – Single Neuron



Classification Accuracy – Network



Temporal Profile of Network Classification Accuracy

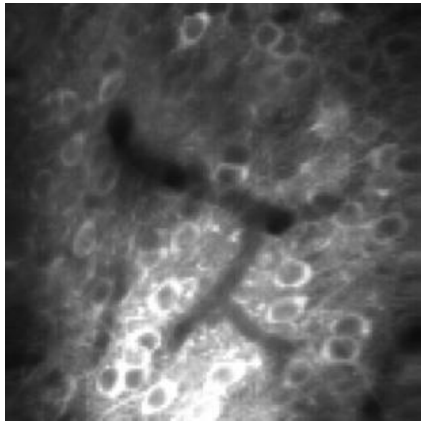


Leitner et al., 2016

Machine learning analysis reveals that odor information is differentially encoded in defined neuronal sub-networks!

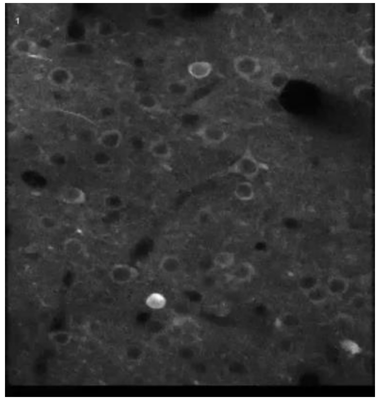
Towards Quantitative Big Imaging Analysis

2009



10 – 50 neurons
100’s of MB / h

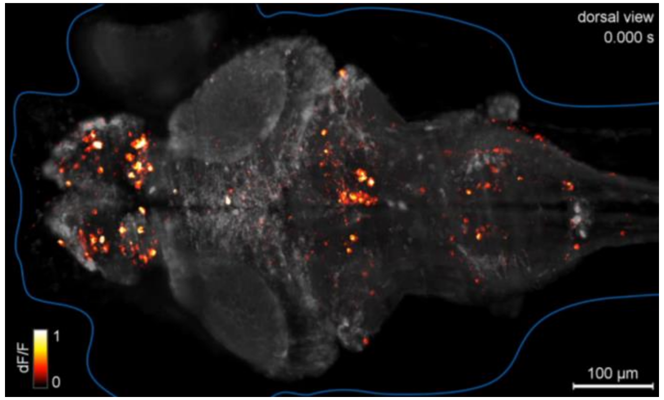
2011



T. Rose, MPI Neurobiology

100s of neurons
10’s of GB / h

Present



Ahrens et al., Nat Meth, 2013

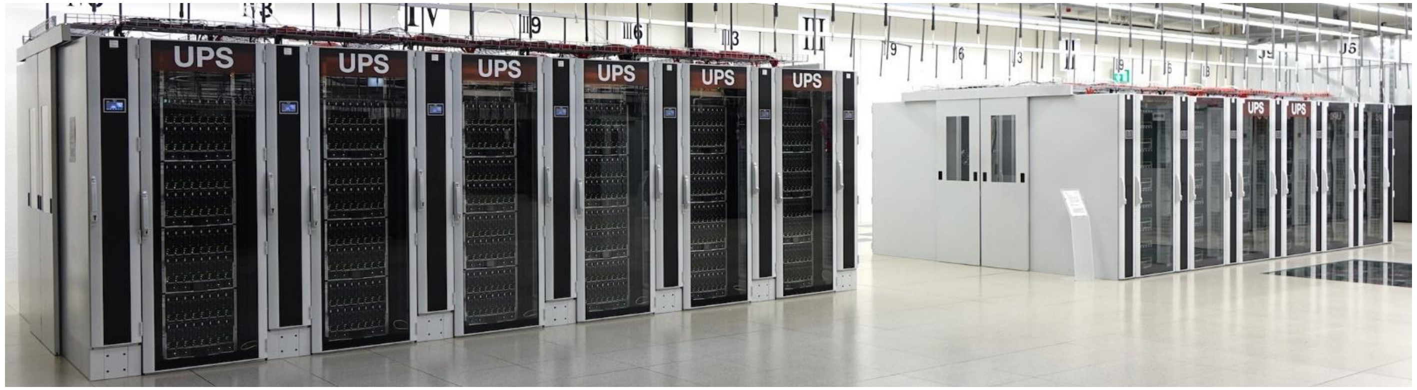
> 10’000 neurons
100’s of GB - TBs / h

- More neurons, better resolution, longer recordings → Increased data size & complexity
- Existing analysis workflows based on desktop PCs scale poorly
- Need for scalable, cluster-based analysis pipelines

MATLAB
Distributed Computing Server

+ Spark

High-Performance Computing @ ETH Zürich



Euler I & II clusters (Euler III added in 2017)

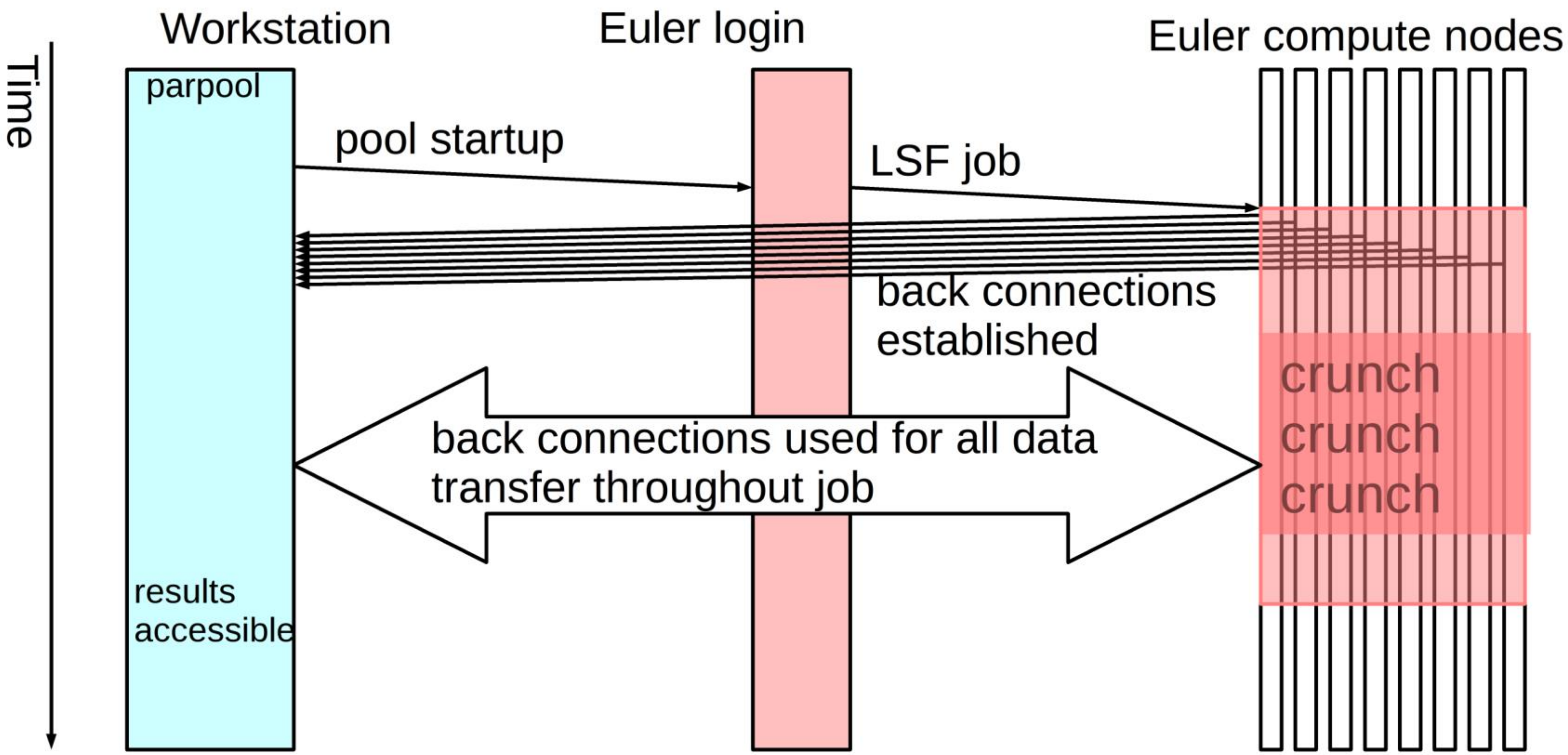
Euler I
448 compute nodes with two **12-core** Intel Xeon E5-2697v2 CPUs
64 - 256 GB RAM
Euler II
768 compute nodes two 12-core Intel Xeon E5-2680v3 CPUs
64 - 512 GB RAM
Euler III
1215 compute nodes with one quad-core Intel Xeon E3-1285Lv5 CPUs
32 GB RAM / 256 GB NVMe flash drive

Big Data Analysis with MATLAB @ ETH Zürich

MATLAB
Distributed Computing Server

Interactive Mode
Parallel for loop

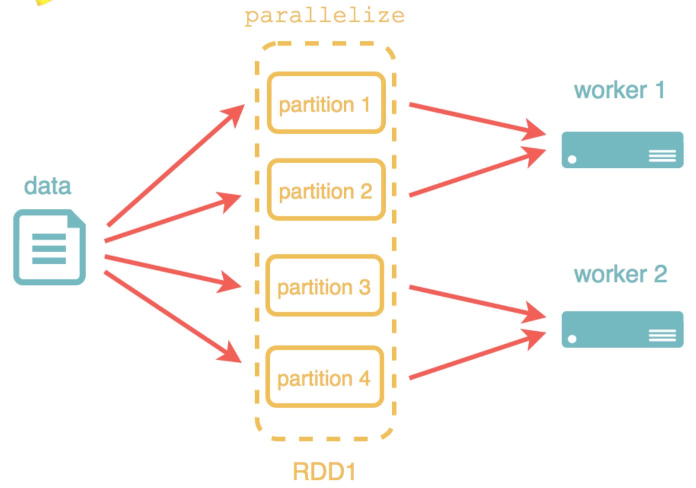
```
cluster = parcluster('Euler');
poolobj = parpool(cluster, 10);
acc = 0;
parfor i = 1:1000
    acc = acc + i^2;
end
```

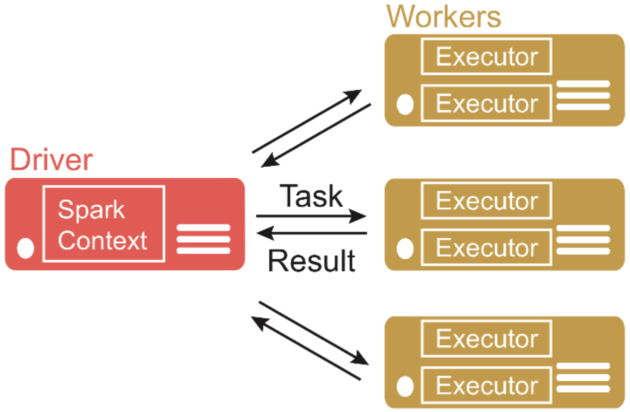


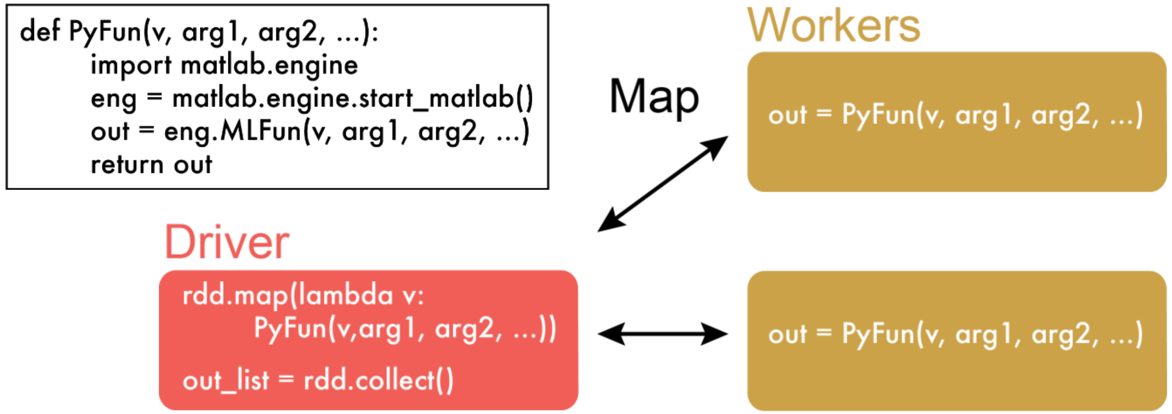
Cluster-scale computing power combined with the convenience of the MATLAB desktop!

Big Data Analysis with MATLAB @ ETH Zürich

+

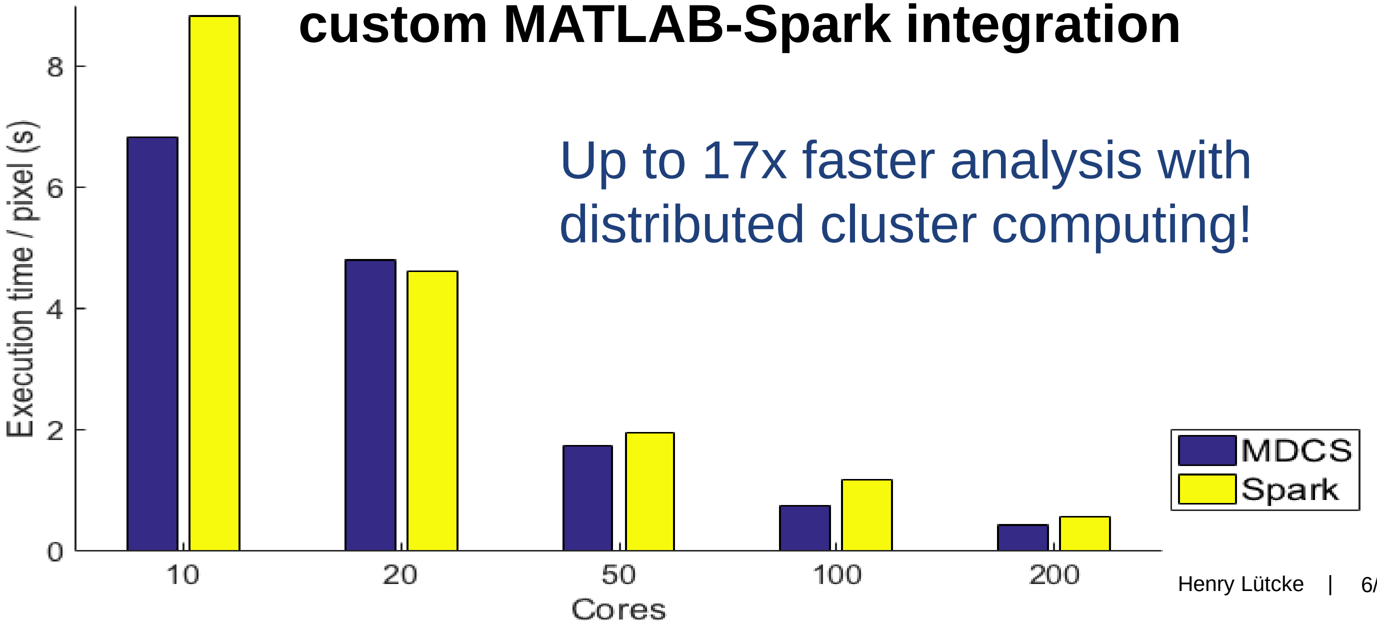






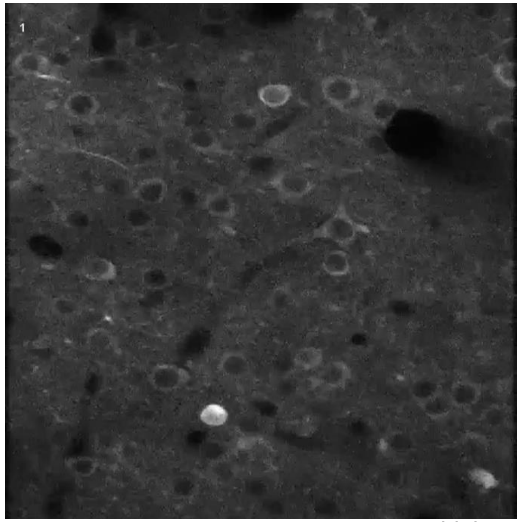
ML-based Image Analysis with MDCS or custom MATLAB-Spark integration

Up to 17x faster analysis with distributed cluster computing!

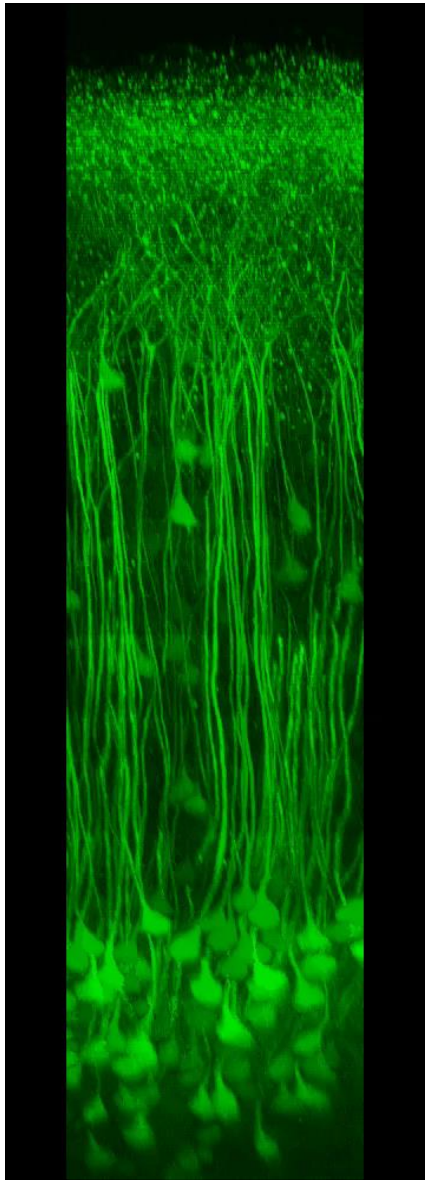


Summary & Conclusions

- Imaging techniques are crucial for understanding the brain and ultimately develop better cures
- Recent shift from qualitative to quantitative imaging
- Image analysis skills & techniques are becoming critical
- MATLAB is applied at all stages and has many advantages
  - Intuitive for novices, powerful for experts
  - Excellent documentation
  - Allows rapid code development / profiling
  - Established in the community
  - Parallelization / scalability

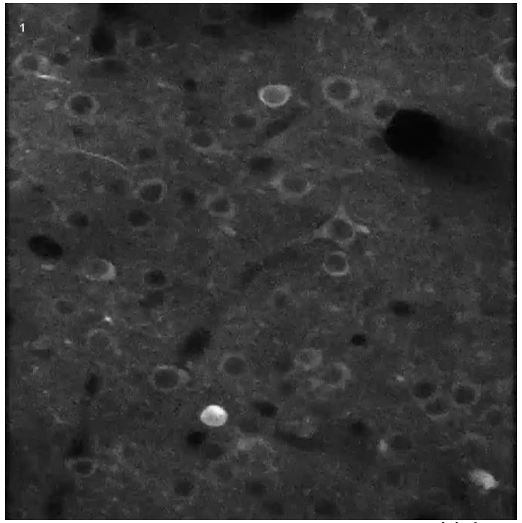


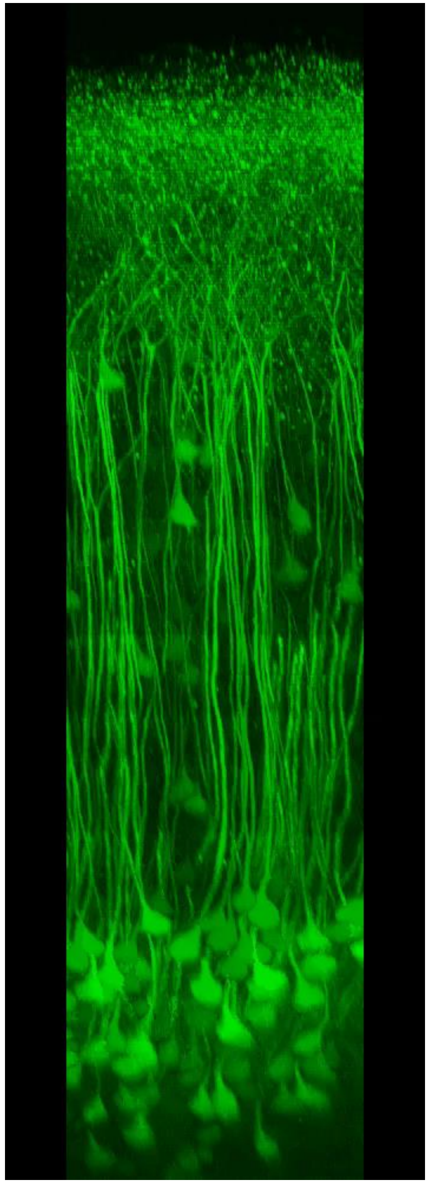
T. Rose, MPI Neurobiology



Future Challenges

- Analysis of millions of neurons
- Real-time analysis and targeted manipulations
- Leverage power of deep-learning approaches
- Further standardization of analysis toolbox





T. Rose, MPI Neurobiology

ETH zürich

Acknowledgments

Scientific IT Service (ETH Zürich)

Rok Roškar
Balazs Laurenczy
Urban Borstnik
Thomas Wüst
Bernd Rinn

Brain Research Institute (University of Zürich)

Fritjof Helmchen

German Cancer Research Center (DKFZ, Heidelberg)

Hannah Monyer
Frauke Leitner

Thank you for your attention!