



Machine Learning Applications at SDSC

Fernando Perez-Cruz

Chief Data Scientist at Swiss Data Science Center

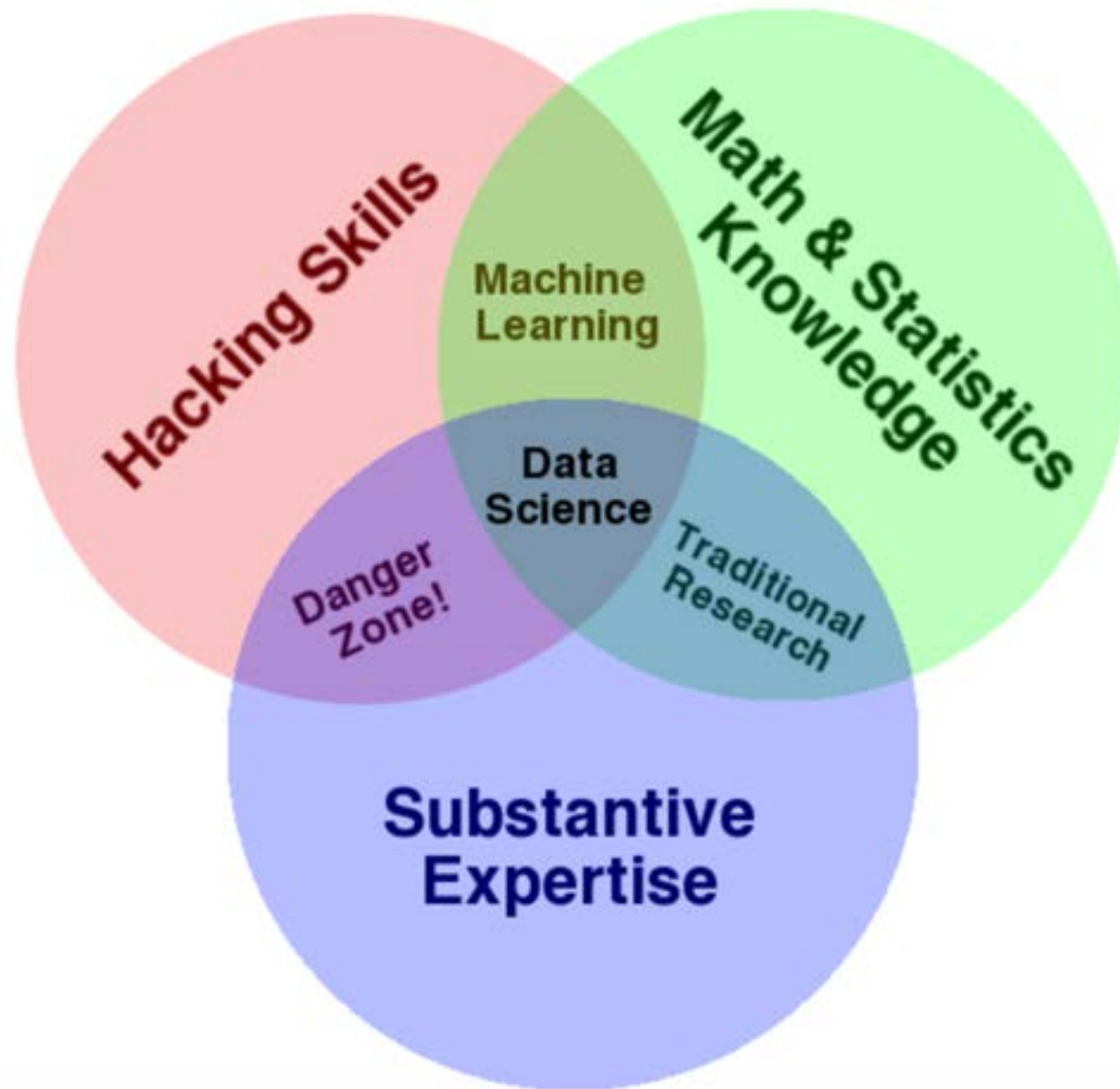
(ETH Zurich and EPFL)

September 24th, 2018

Swiss Data Science Center

- SDSC was established in 2016 by the ETH board to:
“Accelerate the adoption of Data Science in Industry and Academia”
- Why it needs accelerating?
- What are the burdens that impede widespread adoption?
- Why aren't we solving this faster?

Conway's Data Science Venn Diagram



NIPS Sponsor's list 2016



The graphic displays the NIPS 2016 sponsor list, categorized into Platinum, Gold, Silver, and Bronze sponsors. The Platinum Sponsors section features logos for AIG, Apple, Audi, Citadel, and the AWS Collective. The Gold, Silver, and Bronze sections list the names of the respective sponsors. The background of the graphic shows a building, likely the NIPS venue.

PLATINUM SPONSORS

AIG Apple Audi CITADEL AWS COLLECTIVE

DeepMind DIDI facebook Google intel

KLA Tencor Microsoft Tencent 腾讯 WINTON

GOLD SPONSORS

- Adobe Research
- ALIBABA
- Amazon
- AtlaSense
- Baidu
- Criteo
- IBM Research
- NVIDIA
- The Voleon Group

SILVER SPONSORS

- 360 Technology Co.
- Automotive Safety Technologies
- Bloomberg
- Bosch
- Cubist Systematic Strategies
- D. E. Shaw Group
- eBay
- FeatureX
- G-Research
- Hutchin Hill Capital, LP
- Jump Trading
- Man AHL
- Next AI
- Optiver
- Qualcomm
- Renaissance Technologies
- Rosetta Analytics
- SAP SE
- Sentient Technologies
- Two Sigma Investments
- United Technologies Research

BRONZE SPONSORS

- Benevolent AI
- Beijing Institute of Big Data Research
- Cheetah Mobile
- Datatang Technology
- Disney Research
- Invenia Labs
- Maluuba Inc.
- Nokia
- Oracle
- Palantir Technologies
- Panasonic
- PDT Partners
- QuantumBlack
- RBC
- Recursion Pharmaceuticals
- Schibsted Media Group
- SigOpt, Inc.
- Telefonica
- Yandex

SDSC Vision

- SDSC has a team of 20 Software developers and Data Scientists
- We have three lines of actuation:
 - Renku: Open platform for collaborative Data Science.
 - Foster collaboration
 - Open Science
 - Reproducibility
 - Teaching and Mentoring.
 - CAS in Data Science in Lausanne
 - DAS in Data Science in Zurich.
 - Research with labs and industry:
 - Climate Science.
 - Personalized Medicine
 - Physic
 - Architecture
 - ...
 - Predictive maintenance (Buhler and PSA).



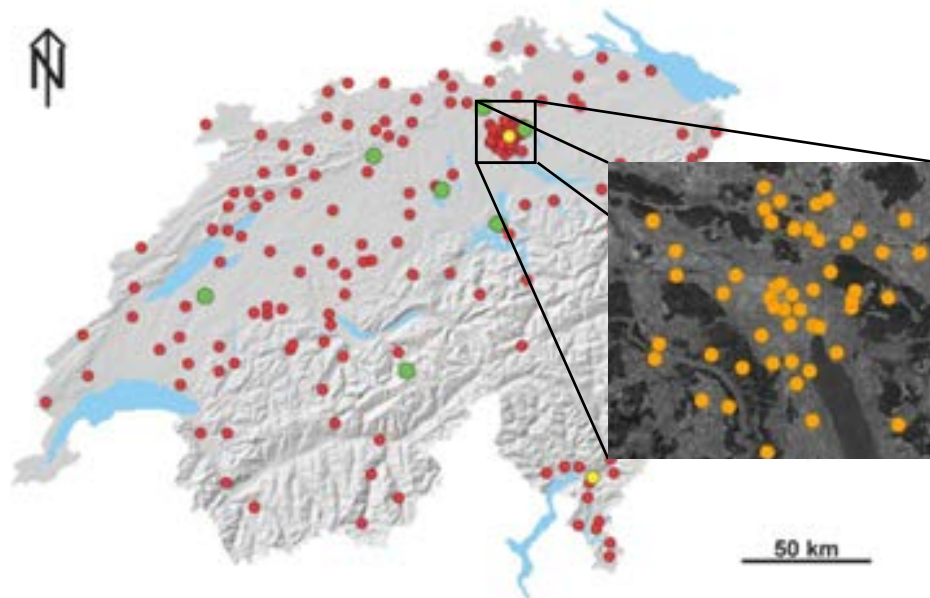
Carbosense4D

w Dominik Brunner

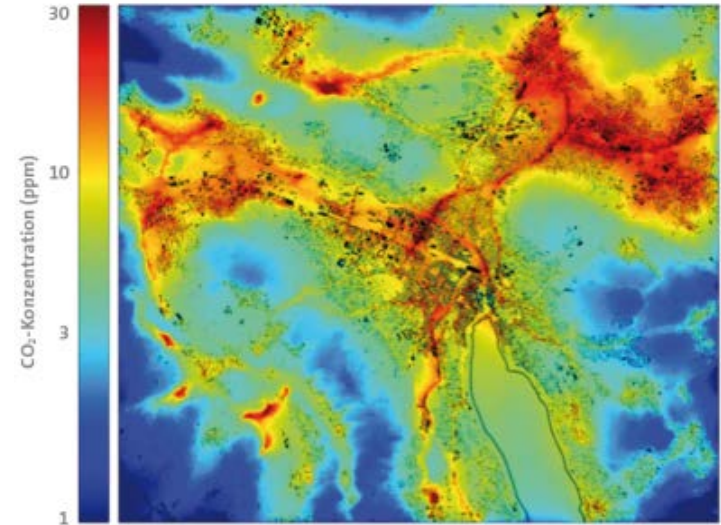
Empa

Goals

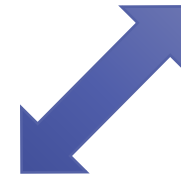
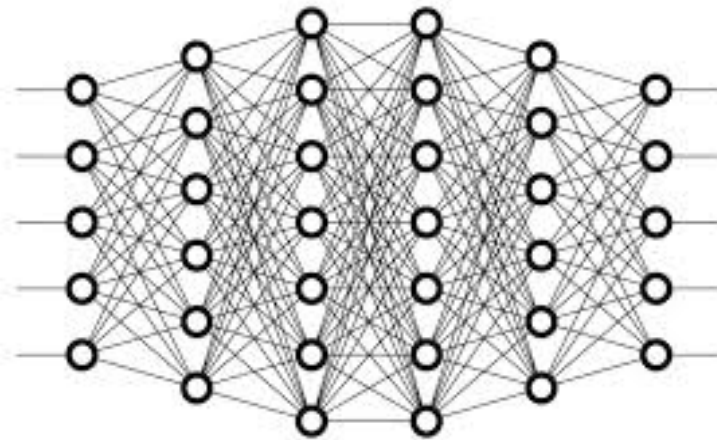
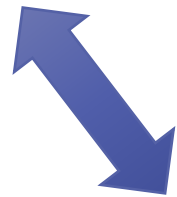
- Determine CO₂ emissions over multiple years.
- Enhance understanding of biospheric CO₂ fluxes over Switzerland.
- Describe accurately the 4-D evolution of CO₂ over Switzerland.



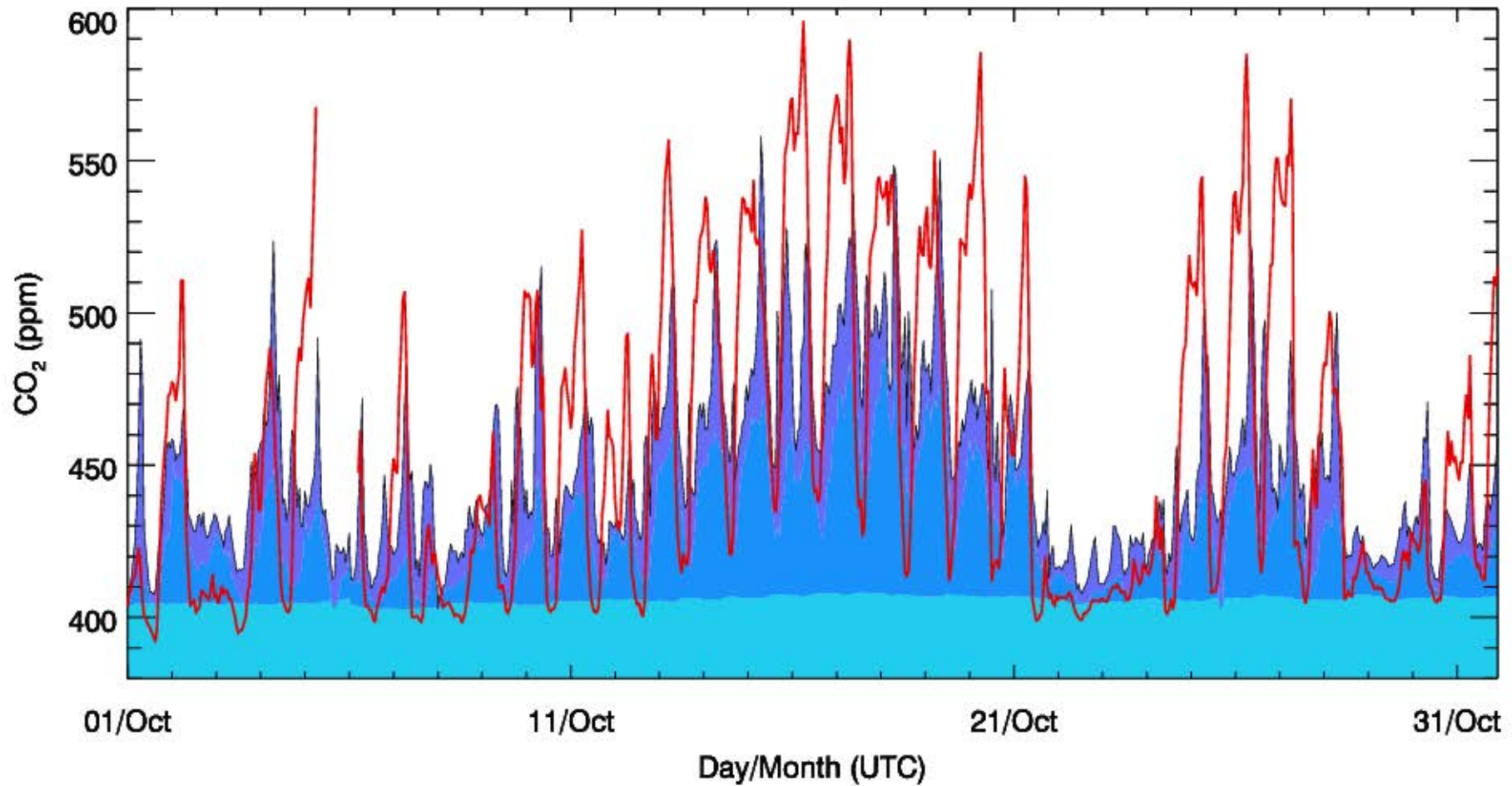
Carbosense4D Sensors



Machine Learning



Observations versus model



- anthrop.
- biospheric
- background
- model
- observations



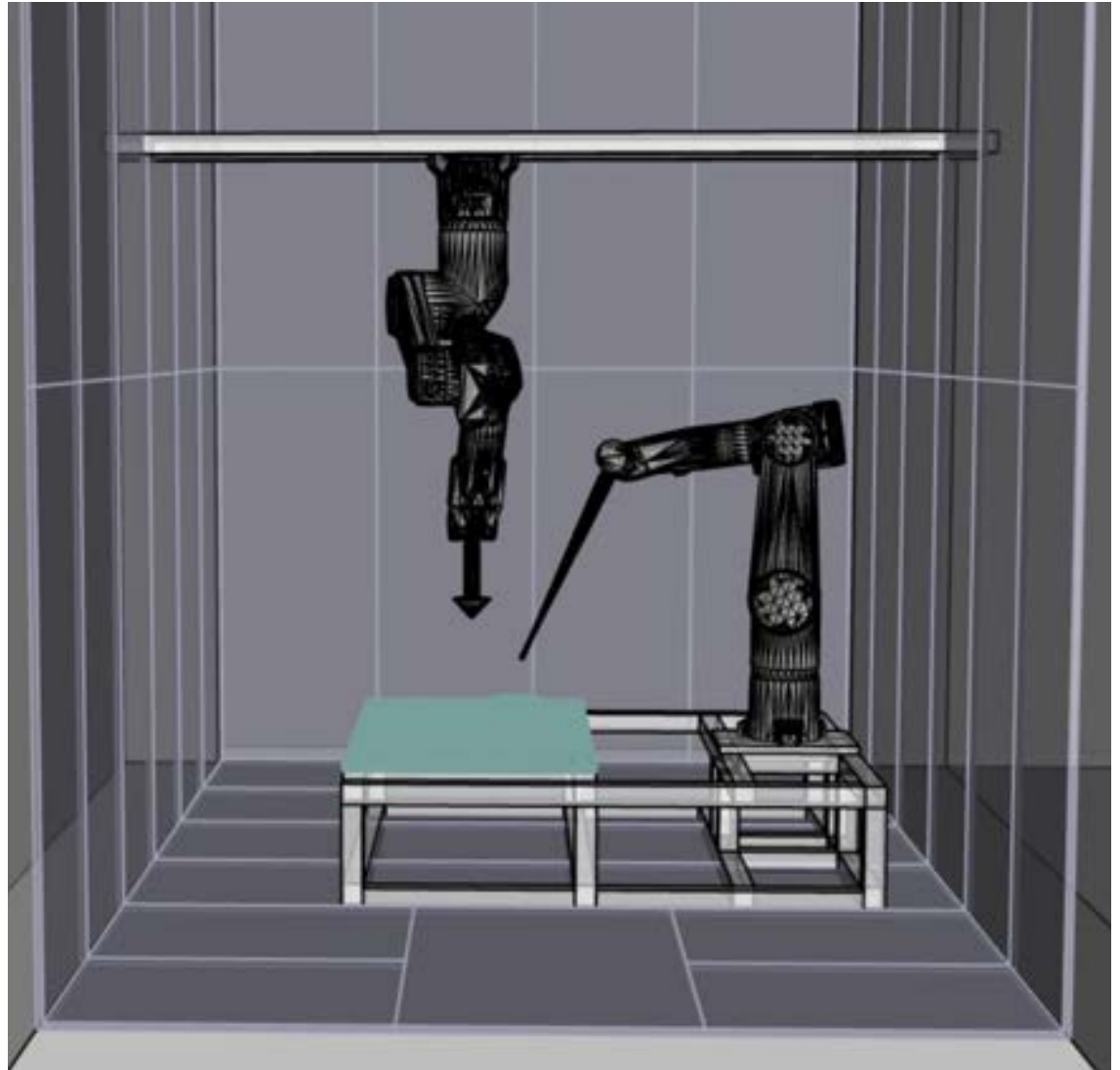
Data Driven Acoustical Design

w Matthias Kohler and Fabio Gramazio

ETHZ - Architecture

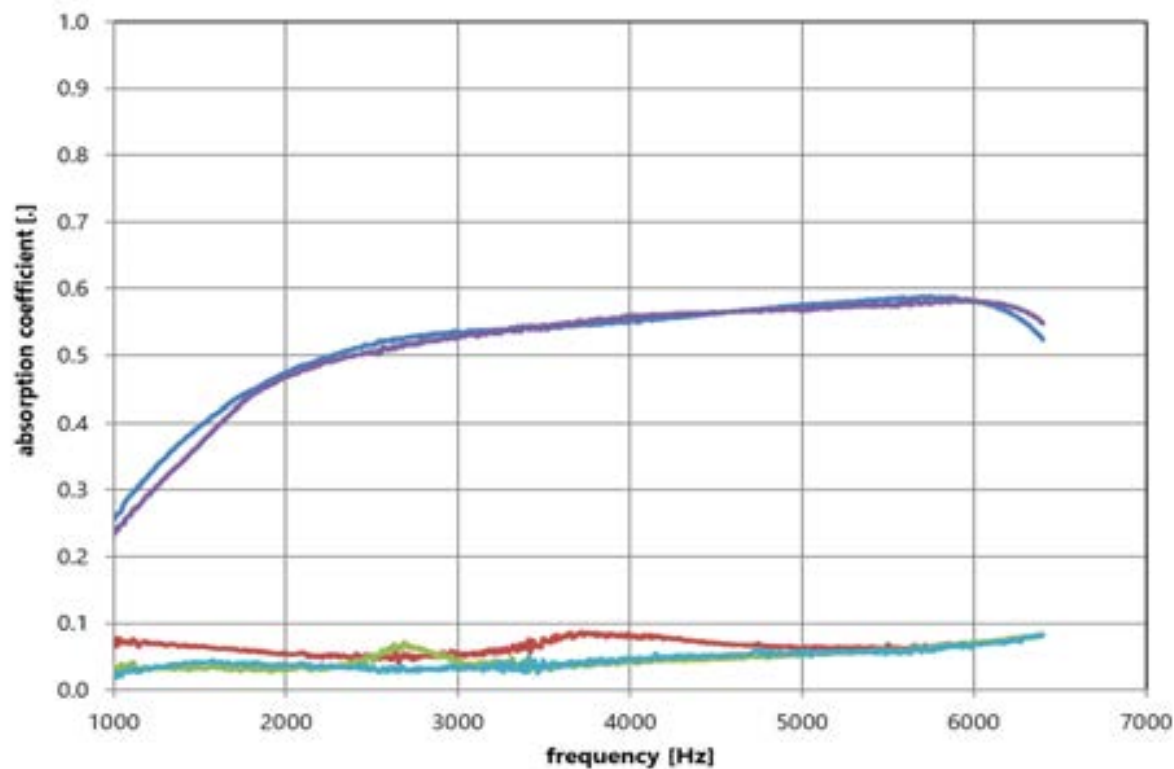
Goals

- Model sound propagation in any wall.
- Use those model to simulate realistic sounds conditions in any room design.
- Architecture for everyday use that is sound resilient.
- Bonus: Improved sound quality of VR environments.



Set up

- 3D print a wall per week 1m x 0.6 m x 0.5m (1:10).
- Measure sound response in over 8 thousand microphone-speaker pairs.
- Every measurement takes over 11s (all just over a day).



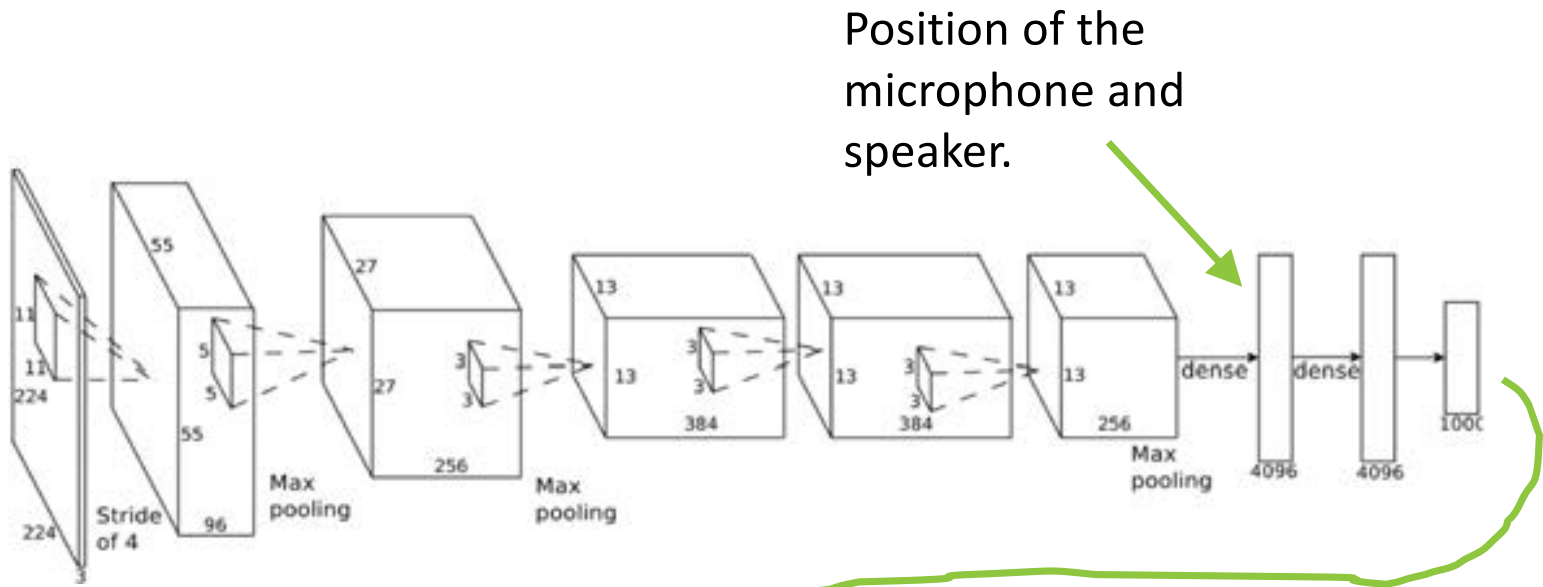
- Original
- Acryl glanz
- Acryl matt
- Baked
- Infiltrate

Convolutional Neural Network

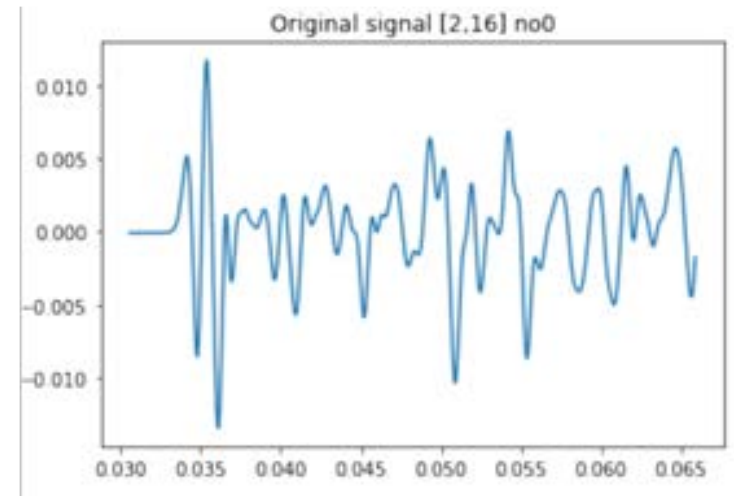
Wall tensor



100×60×10



From the wall, we predict the signal that would be received on each microphone





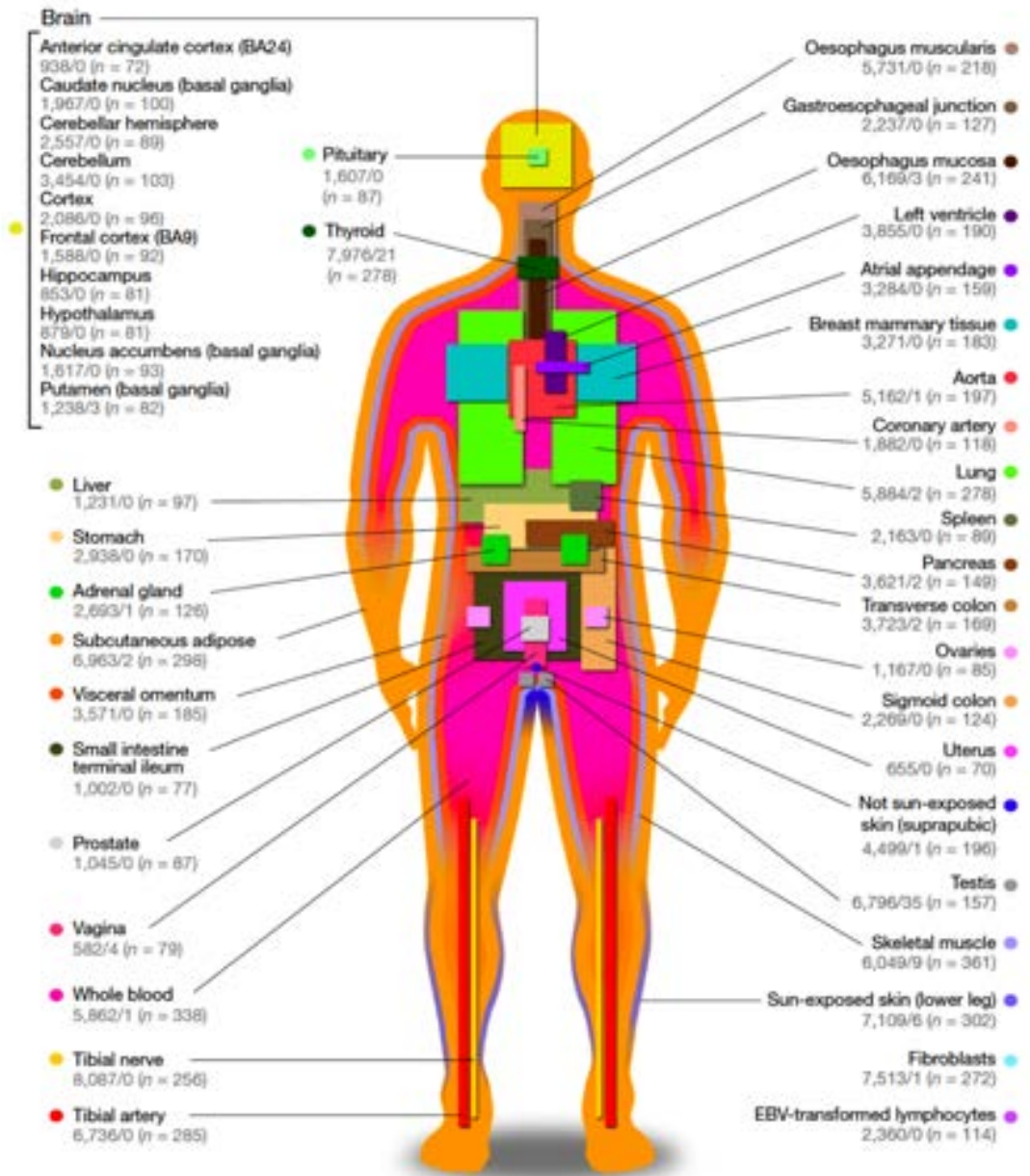
T4Med

w Didier Trono

EPFL – Life Sciences

Goals

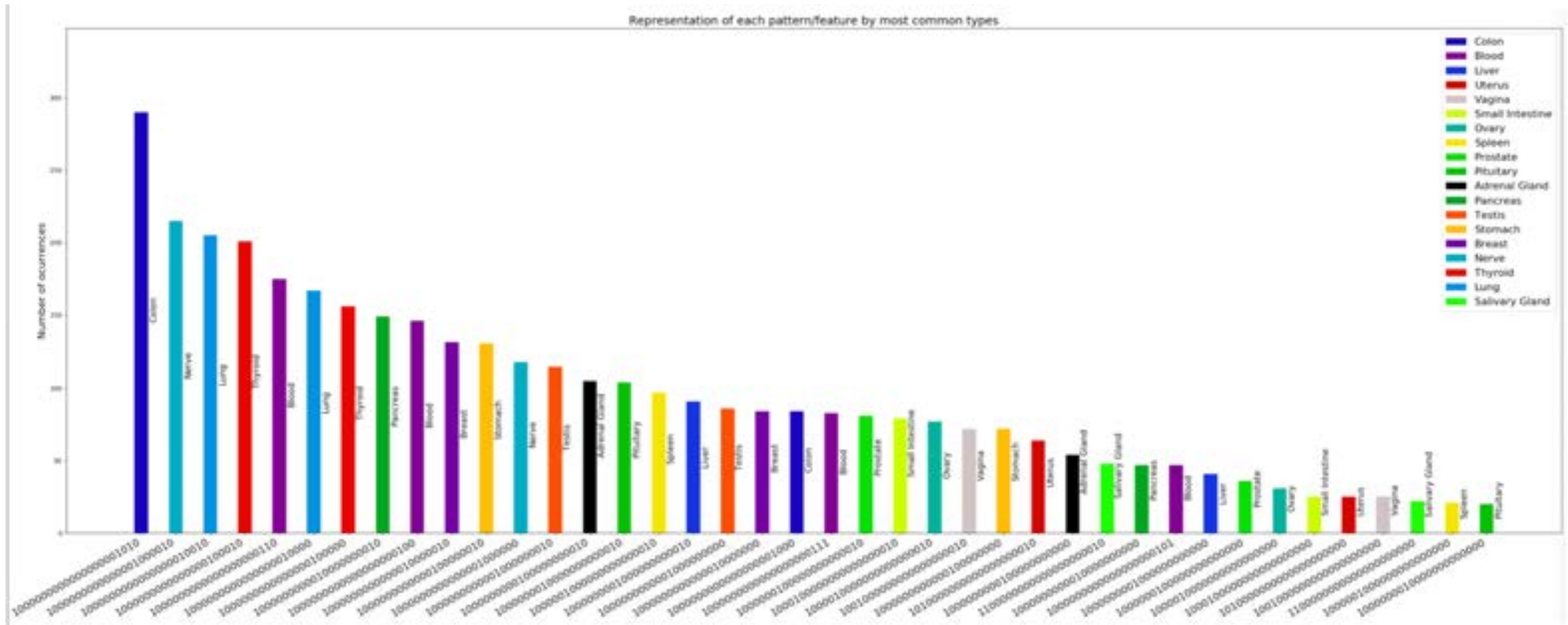
- Can we predict from a tissue gene and TE expression data where it is coming from?
- If so, we might be able to detect metastasis.
- Can we also detect sub-cancer types?
- Data:
 - 9000+ samples from 549 cadavers.
 - 54 body sites.
 - Genes and TE expression data.



Transposable Elements

- TEs are typically referred as junk DNA.
- TEs represent 80% of the human genome.
- They are repetitive and it is thought to be left-over material from the last common ancestor or viruses.
- TEs have regulatory and active in the expression of each cell tissue.
- We are using TEs:
 - Detect tissue type.
 - Detect cancer and sub-cancer type.
 - Detect metastasis.

Promising results



Questions?