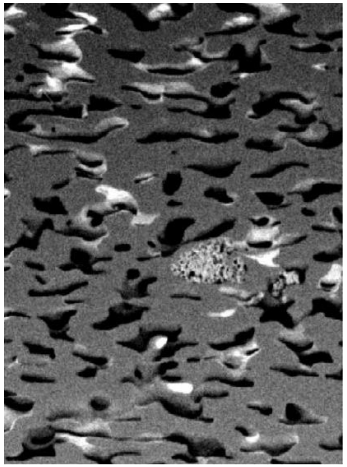
Analysis of microscopy data

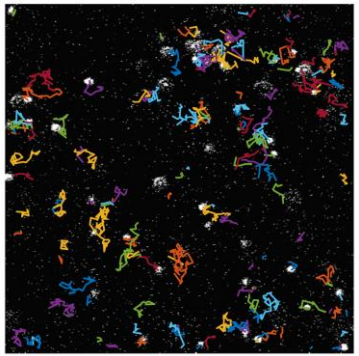
FIB-SEM, particle tracking, FRAP

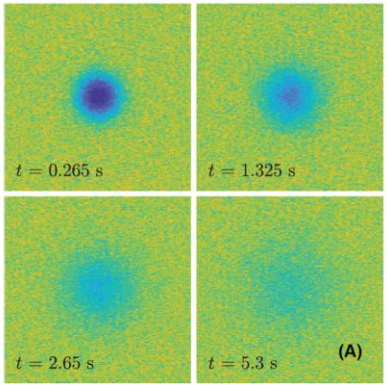
Magnus Röding

Guest lecture in Spatial statistics and image analysis

2021-05-10







Very briefly about me…

- Scientist at RISE Research Institutes of Sweden, Gothenburg (Agriculture and Food department)
- Adjunct Associate Professor at Chalmers University of Technology, Gothenburg (Department of Mathematical Sciences)
- Applied math/stats - image analysis, spatial/stochastic modeling, machine learning, with applications in materials science and microscopy

Today

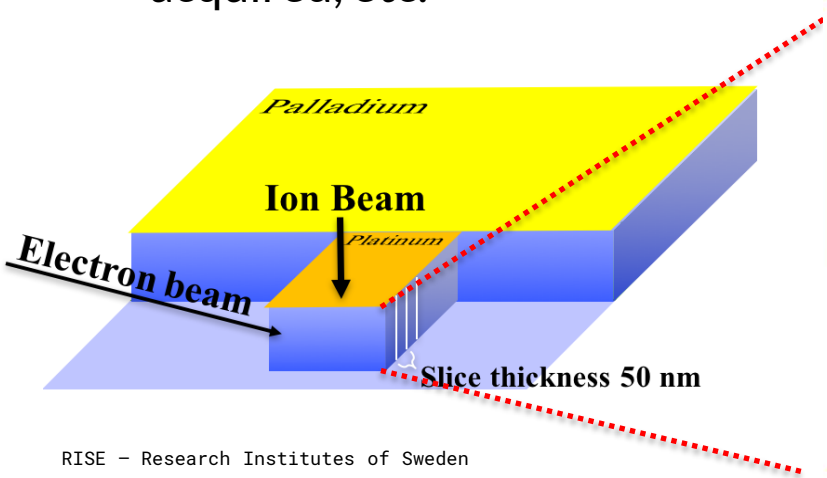
- We will cover three different types of microscopy/imaging techniques and the related image analysis

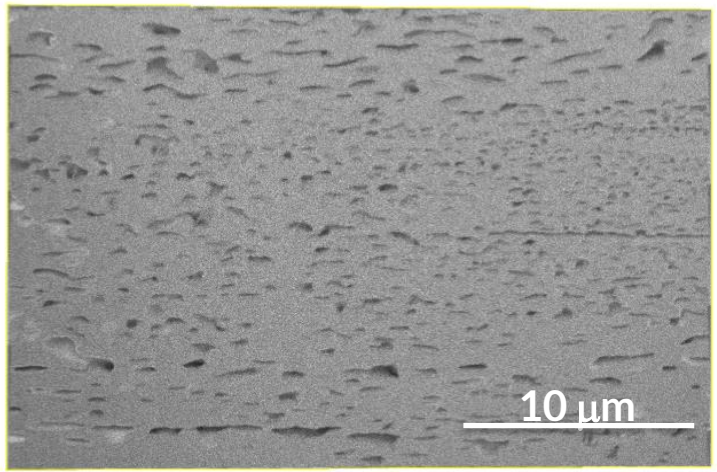
    1) Focused ion beam scanning electron microscopy (FIB-SEM)
    2) Diffusion theory
    3) Particle tracking of diffusing particles
    4) Fluorescence recovery after photobleaching (FRAP)

Focused ion beam scanning electron microscopy (FIB-SEM)

Acquiring FIB-SEM data

- Focused Ion Beam Scanning Electron Microscopy (FIB-SEM)
- One slice (~50 nm) is peeled off using the FIB, then one image is acquired with the SEM, then another slice is peeled off, another is acquired, etc.



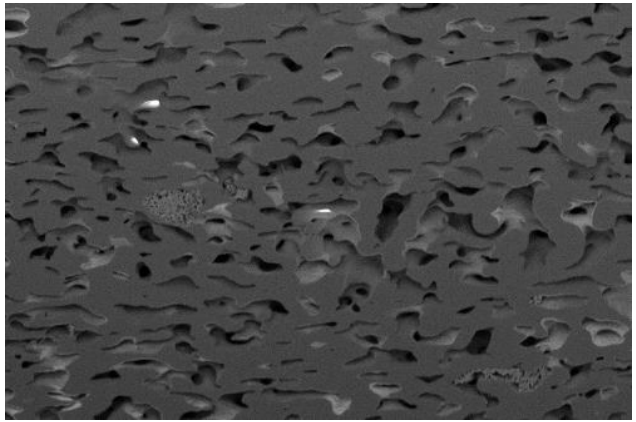


Acquiring FIB-SEM data

- Each of these slices is 2D, but you can actually see ”into the pores”, making it 2.5D in a sense
- It is a destructive technique, because you destroy the material with the FIB
- Alternatives:
  - X-ray tomography (lower resolution)
  - Transmission electron microscopy (TEM)-based tomography (only for very thin samples)

Example: EC/HPC films

- Ethyl cellulose hydroxypropyl cellulose (EC/HPC) two-phase polymeric film used for controlled drug release in pharmaceutical industry
- One phase is leached out, yielding solid-pore structure for imaging
- Image analysis is a challenge because you can ”see into” the material (2.5-dimensional slices)

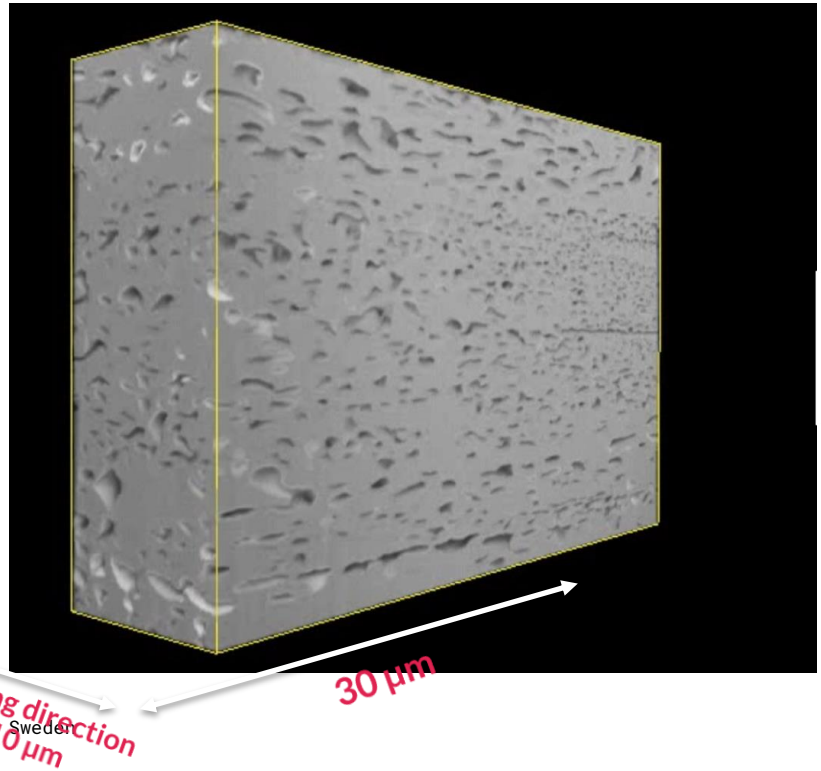


The data

- Three datasets, each is 3000 x 2000 pixels and 200 slices
- We want to perform *semantic segmentation* i.e. differentiate between solid and pore
- Why? Because we want to understand how the drug is transported through the coating, therefore we need to characterize the geometry of the porous network

3D visualization

Transport direction: 22 µm



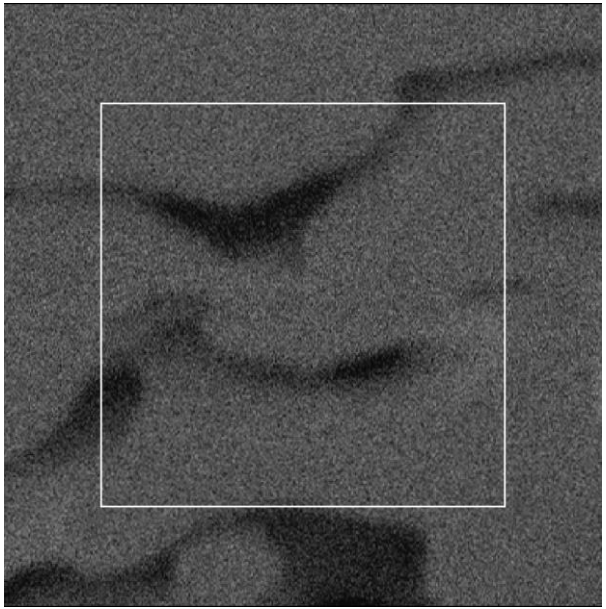
Slice thickness:50 nm
200 slices

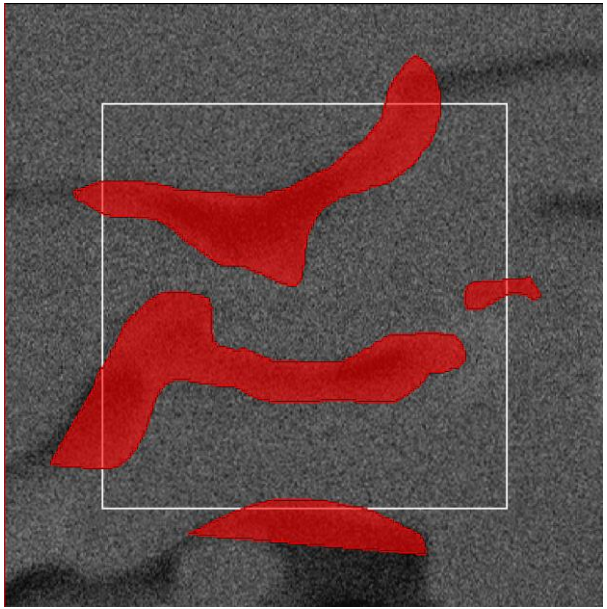
Manual segmentation

- Need annotated/labeled data for validating any automatic method and assessing its performance
- Data are too large for manual segmentation of it all (would take months)
- Solution
  - Manual segmentation in 100 randomly placed square regions of size 256x256 by an expert (~0.5 % of the data)
  - Takes more like a day instead of several months

Manual segmentation

- 3 x 100 of these

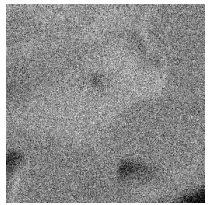




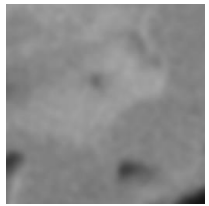
Segmentation, approach I

- What information do we use to classify one pixel?
- Use information not only in that pixel but in the surrounding neighborhood
  - Intensity information from 5 slices in each direction
  - Information at different scales (original data + Gaussian smoothed data at different scales (sigma = 1, 2, 4, 8, 16, 32, 64, 128 pixels), so-called scale space) -> averages values in different size regions
  - In total we use 99 variables to predict solid or pore

Original



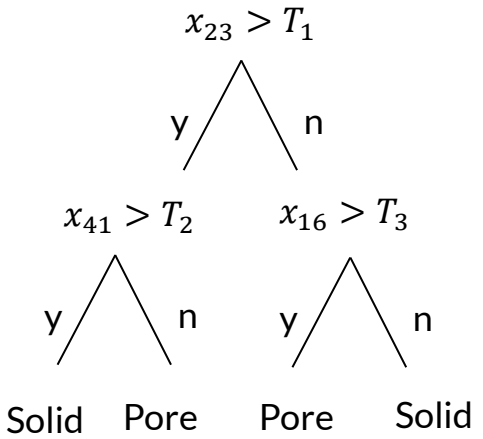
sigma = 4



sigma = 16

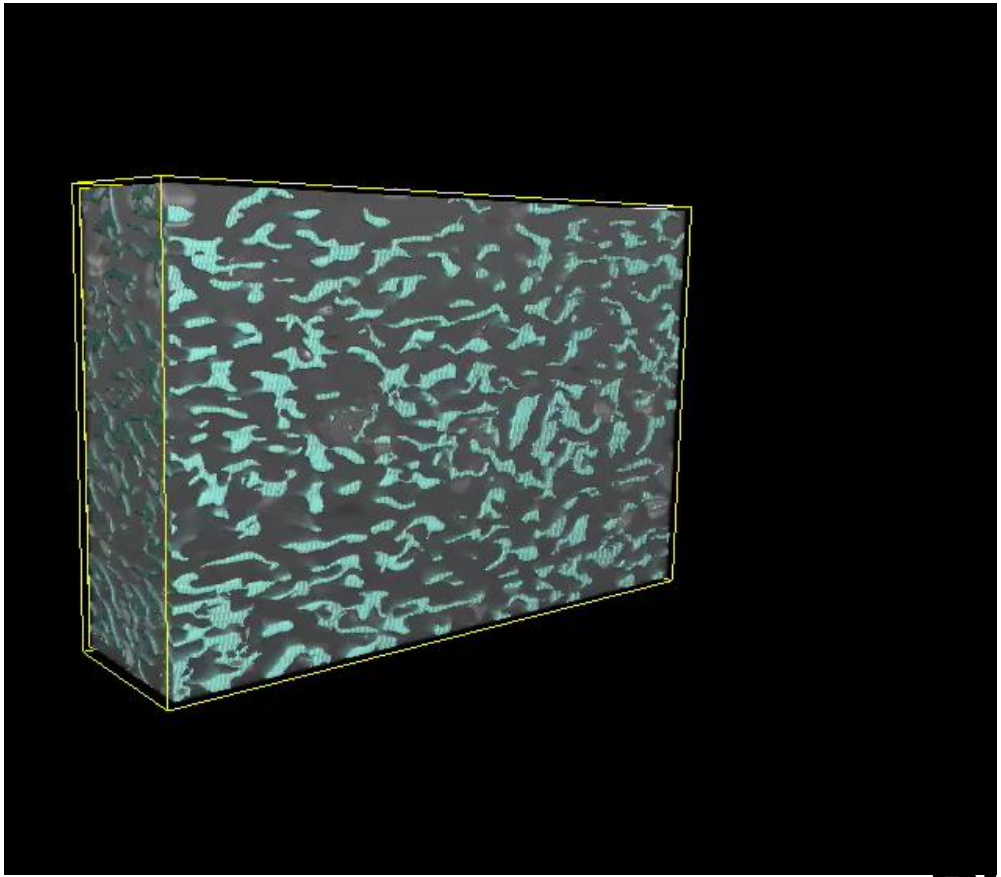
Segmentation, approach I

- Assigning each voxel to on of the classes (solid + pore), a *classification* problem
- Can use all sorts of machine learning algorithms for this
- We pick decision trees (see fig to right), a very simple method based on thresholding:
  - If we have many different variables describing an image, say x1, x2, x3, …. Try to make a decision of class membership based on choosing threshold values for these
  - Combine a large number of decision trees -> **random forest algorithm**
  - Combining simple decision trees by averaging their predictions
  - Averaging of n = 151 decision trees (odd number, no ties)



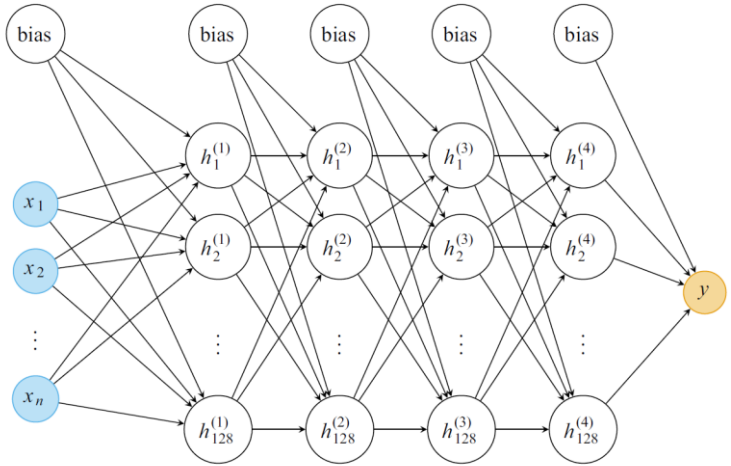
Example result

- Classification accuracy
  - Training data: 98.0 %
  - Test data: 92.2 %
- Porosity
  - Expected: 30.0 %
  - Obtained: 29.8 %



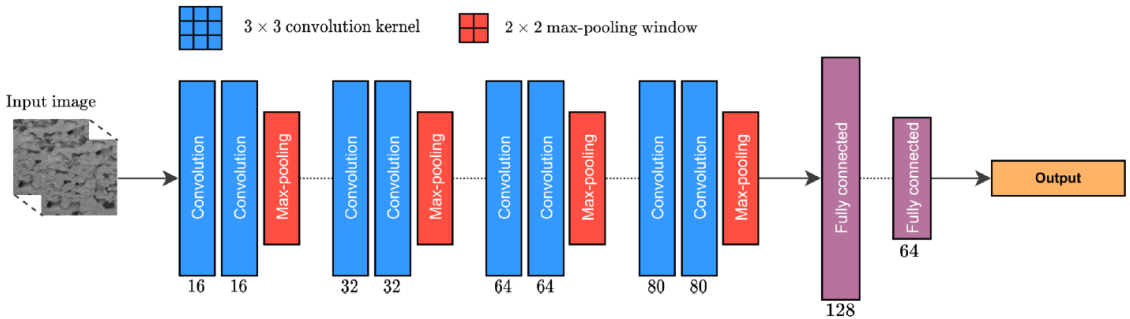
Segmentation, approach II

- Convolutational neural networks (CNNs)
- **Crash course in CNNs**
- Conventional (artificial) (fully-connected) neural networks (ANNs) are based on a number of layers, each with a number of nodes corresponding to nonlinear mappings that together form a very complex nonlinear mapping. This can be used to solve classification problems



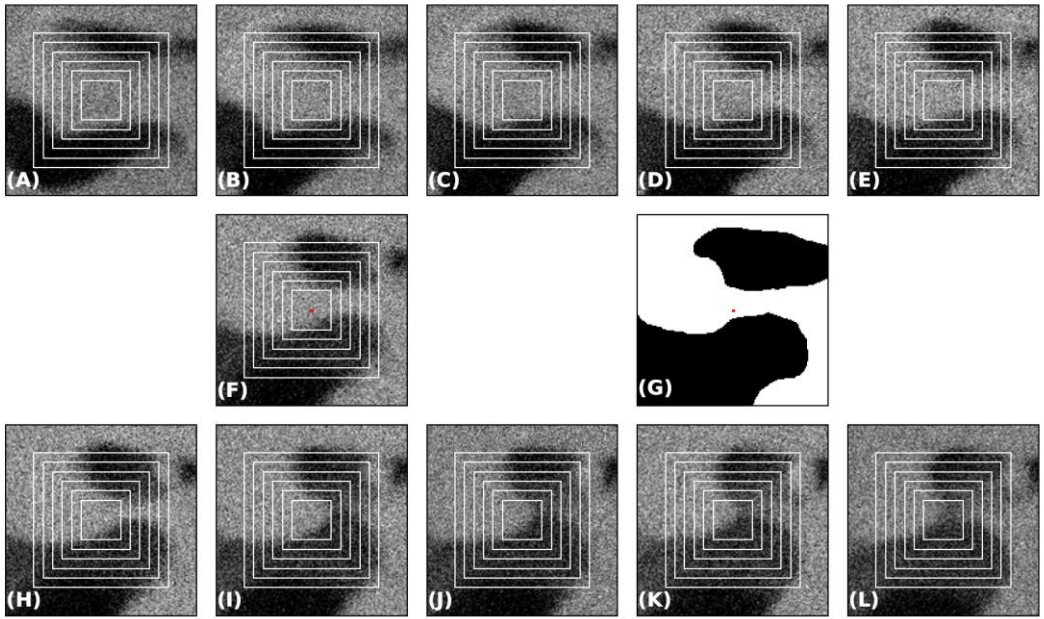
Segmentation, approach II

- Convolutational neural networks (CNNs), on the other hand, are built up using convolutional filters.
- During training, the CNN learns the weights of the filters.
- Pooling reduced dimension of the image.
- Finish with a fully connected part.



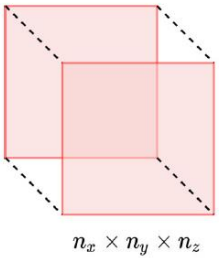
Segmentation, approach II

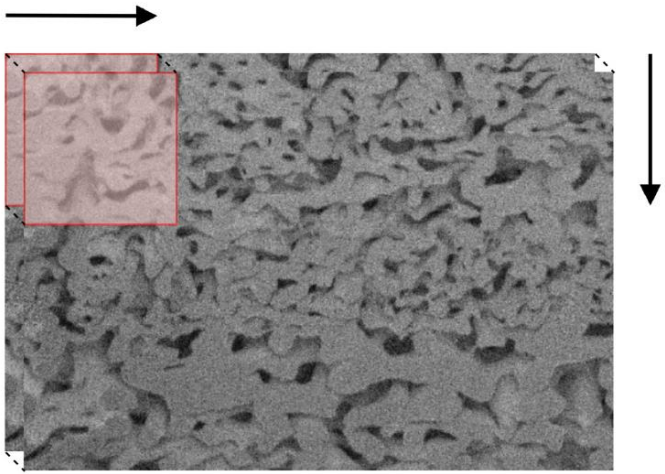
- We extract rectangular neighborhoods of different sizes around the voxel of interest to find the optimal size



Segmentation, approach II

- We got best results for a 113x113x3 voxel neighborhood
- Classification of all voxels is done with a sliding-window approach, looping over all parts of the dataset





Comparison on test data to random forest:

RF: mIoU 0.7616, accuracy 0.8825

CNN: mIoU 0.7980, accuracy 0.9024

RI. SE

In summary

- CNNs perform slightly better but at the cost of much more training time and hyperparameter optimization etc.
- Actually, when we added derivatives of the images (not published) as features in the random forest approach, we got an improvement by a couple of percent in accuracy, so close to CNNs…

M. Röding, C. Fager, A. Olsson, C. von Corswant, E. Olsson, N. Lorén. Three-dimensional reconstruction of microporous polymer films from FIB-SEM nanotomography data using random forests. Journal of Microscopy, 281, 76-86, 2021.

F. Skärberg , C. Fager , F. Mendoza-Lara, M. Josefson, E. Olsson, N. Lorén, M. Röding. Convolutional neural networks for segmentation of FIB-SEM nanotomography data from porous polymer films for controlled drug release. Accepted in Journal of Microscopy.

Theory of diffusion

What is diffusion?

- Diffusion is essentially random transport of particles (and concentrations of particles) driven by thermal energy
- Can be understood from a
  - microscopic point of view (a stochastic model)
  - macroscopic point of view (a partial differential equation)
- Both these perspectives will be important for the microscopy techniques we will cover today

Brownian motion

- Brownian motion is the random motion of particles suspended in a medium (like a fluid)
- Even at thermodynamic equilibrium when the system is macroscopically stable, random fluctuations in the fluid will cause random motion of the suspended particles
- First documented by Robert Brown (1827)

Brownian motion

- Polystyrene particles (diameter ~ 500 nm) in water imaged by confocal laser scanning microscope
- Pushed in random directions by individual water molecules
- Note: They disappear when they move out of focus

Brownian motion

- Brownian motion can be modelled as a Gaussian random walk, with normal distributed increments

$$\Delta x = x(t + \Delta t) - x(t) \sim N(\mu = 0, \sigma^2 = 2D\Delta t)$$

- Where Δx is a 1D displacement, Δt a time step, and D is the **diffusion coefficient**

- Alternatively, in terms of mean squared displacement (MSD)
  $E(\Delta x^2) = 2D\Delta t$ $(E(\Delta x^2 + \Delta y^2) = 4D\Delta t, E(\Delta x^2 + \Delta y^2 + \Delta z^2) = 6D\Delta t)$

- Einstein (1905) and von Schmolukowski (1906)

Diffusion equation

- If an infinite number of Brownian particles would start diffusing at $x = 0, t = 0$, then their distribution will be

$$f(x;t) = \frac{1}{\sqrt{4\pi D t}} \exp\left(-\frac{x^2}{4Dt}\right)$$

- This is also the so-called fundamental solution to the diffusion equation, or Fick's second law (Fick 1855), that describes the evolution of a concentration of diffusing particles in time and space:

$$\frac{\partial c(x,t)}{\partial t} = D \frac{\partial^2 c(x,t)}{\partial x^2}$$

Diffusion & Brownian motion

- The macroscopic description (Fick 1855) came 50 years before the microscopic-scale explanation of the phenomenon (Einstein et al 1905-1906)
- Einsteins derivations were taken as proof that molecules exist

Diffusion & Brownian motion

- Is Brownian motion inherently Gaussian?
- No, it’s a result of the central limit theorem i.e. the sum of a large number of small displacements is approximately Gaussian
- If you zoom in enough, Brownian motion will be locally rather straight (ballistic motion) when it’s pushed by a fluid particle, then changed direction when hit by the next

Particle tracking

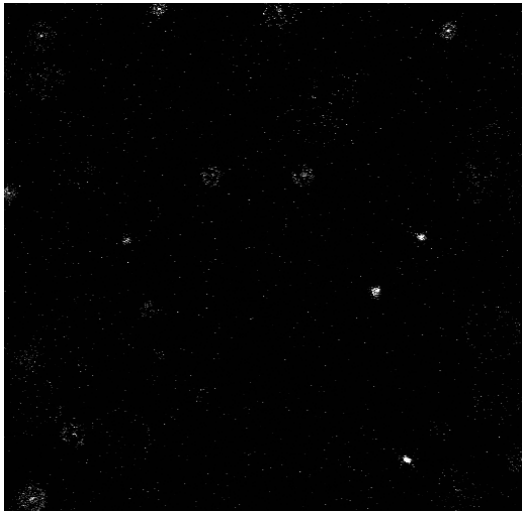
Particle tracking

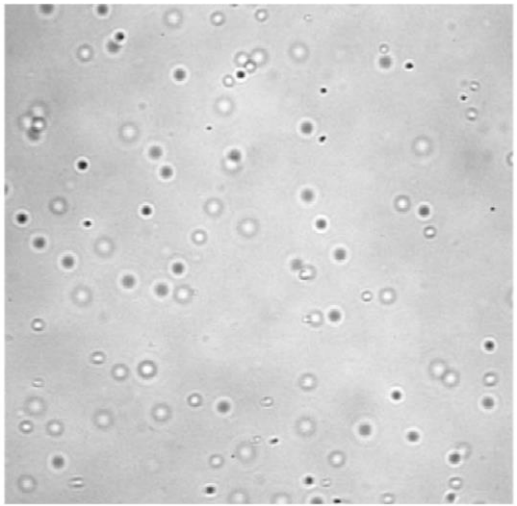
- Particle tracking experiments are performed for various reasons, like
  - Estimating a diffusion coefficient (or a distribution)
  - Estimating mechanical properties of the medium (particle tracking microrheology)
  - Estimating both diffusion and flow velocity at the same time

- …today, only pure diffusion

Particle detection

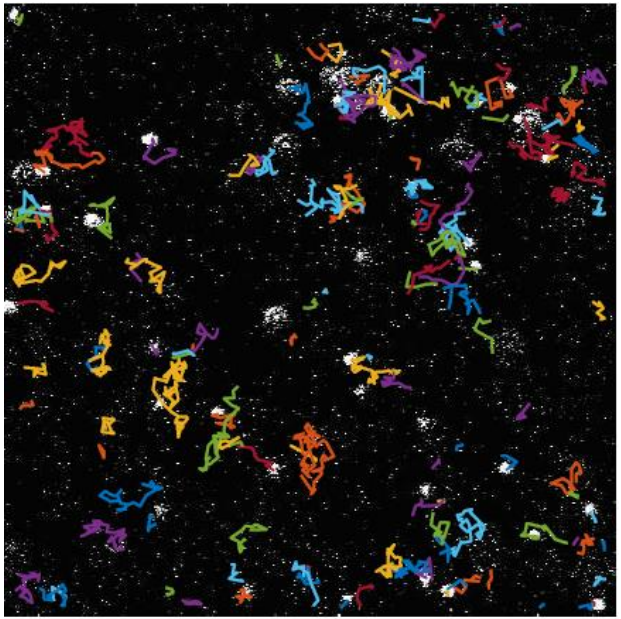
- Classical method: Smooth the image, find intensity maxima and if the maxima are above a threshold value, it is a particle
- Template matching: Fit example particle intensity profiles to the image (in the lower images, particles can be both bright or dark)
- *Depends on microscopic technique*





Particle linking

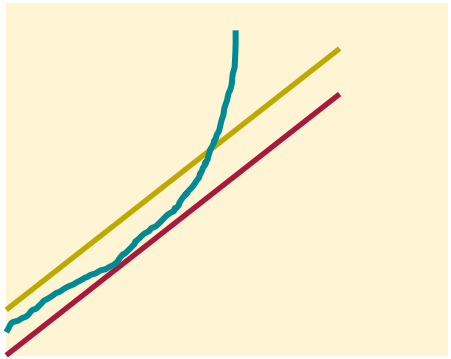
- Classical method: In frame n+1, find the particle positions closest to the ones in frame n. If a particle is sufficiently far away from all in frame n, it's likely the start of a new trajectory
- Distance threshold based on prior information about mobility ("how far is it reasonable to move between two frames?")
- Very short trajectories (1 or 2 frames) are likely just noise



MSD analysis

- Analysis of the mean square displacement (MSD) can be useful to detect the type of motion of the particles

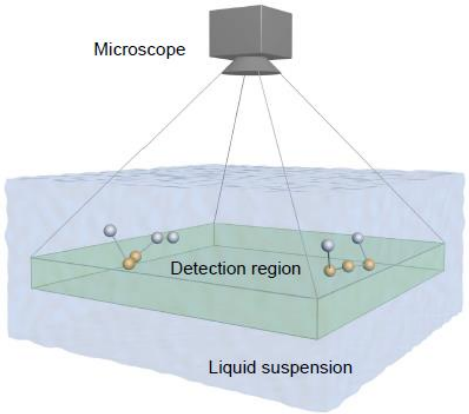
$$\mathrm{MSD} \equiv \langle |\mathbf{x}(t) - \mathbf{x_0}|^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{x^{(i)}}(t) - \mathbf{x^{(i)}}(0)|^2$$

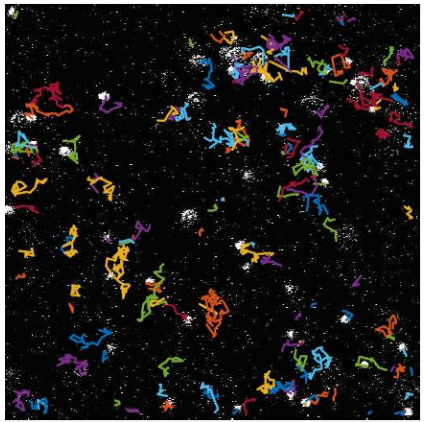


- Pure diffusion $MSD(t) = 4Dt$ (assuming # dimensions = 2)
- Diffusion + localization error ($\sigma = \varepsilon$) $MSD(t) = 4\varepsilon^2 + 4Dt$
- Diff + loc err + flow (vel. v) $MSD(t) = 4\varepsilon^2 + 4Dt + |v|^2 t^2$

Diffusion and concentration

- Image individual diffusing fluorescent particles, performing Brownian motion, at ~ 15 frames per second using a confocal laser scanning microscope (CLSM)
- We observe the particles in a box-shaped detection region, the size of which is determined by field of view (~60 x 60 µm) and other parameters



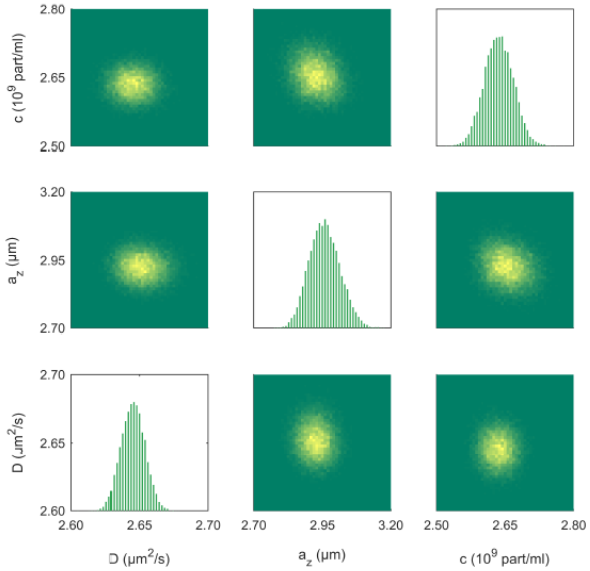


Diffusion and concentration

- The data consists of a list of particle trajectories. Particles move in and out and the same particle can appear more than once
- From them, their durations (number of positions) and estimated diffusion coefficients for each trajectory can be obtained
- We want to estimate diffusion coefficient, D, and (number) concentration, c (and we will obtain also the thickness of the detection region a_z as a bi-product)

Diffusion and concentration

- To write down a model in terms of simple equations is not so simple, but easy to simulate (basically, just Gaussian random walks moving in and out of a box)
- Therefore, simulation-based inference, more precisely Approximate Bayesian computation (ABC), is a very useful tool, for estimating $\theta = (D, a_z, c)$



M. Röding, M. Billeter. "Massively parallel approximate Bayesian computation for estimating nanoparticle diffusion coefficients, sizes and concentrations using confocal laser scanning microscopy." Journal of Microscopy 271, 174-182, 2018.

M. Röding, E. Zagato, K. Remaut, K. Braeckmans. "Approximate Bayesian computation for estimating number concentrations of monodisperse nanoparticles in suspension by optical microscopy". Physical Review E, 93, 063311, 2016.

M. Röding, H. Deschout, K. Braeckmans, M. Rudemo. Measuring absolute number concentrations of nanoparticles using single-particle tracking. Physical Review E, 84, 031920, 2011.

M. Röding, H. Deschout, K. Braeckmans, A. Särkkä, M. Rudemo. Self-calibrated concentration measurements of polydisperse nanoparticles. Journal of Microscopy, 252, 79-88, 2013.

Fluorescence recovery after photobleaching (FRAP)

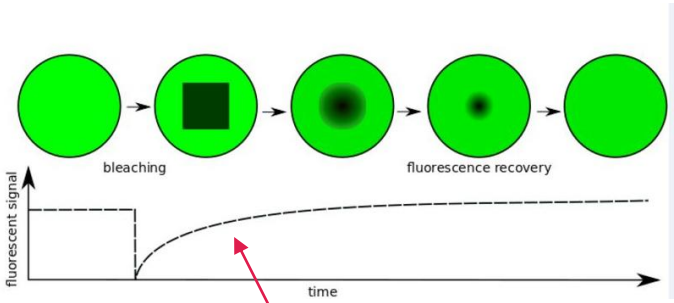
Fluorescence Recovery After Photobleaching (FRAP)

- In FRAP, we move from studying individual particles to a concentration of particles in space and time; we “zoom out”
- That means we move from the Brownian motion/mean square displacement model to the diffusion equation for modeling
- In thermodynamic equilibrium, nothing is happening at the macroscopic scale because the concentration is already uniform in the sample
- To get something to measure, we have to perturb the system out of equilibrium….this is the idea of FRAP

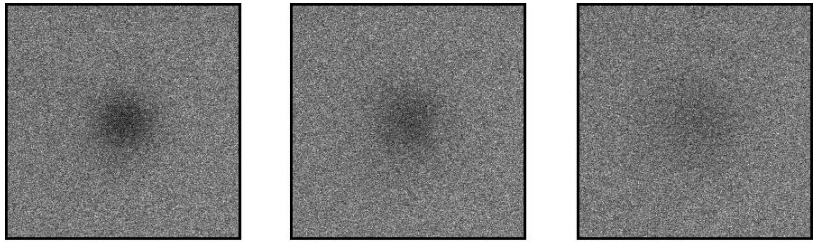
RI.
SE

Fluorescence Recovery After Photobleaching (FRAP)

- Fluorescent particles, uniformly distributed in a sample, are bleached (kill the fluorescence) in a circular or rectangular bleach region using a powerful laser
- The time evolution of the concentration/fluorescence intensity back to equilibrium contains information about the motion of particles
- If we image this process using e. g. confocal laser scanning microscopy (CLSM), we can estimate e. g. diffusion coefficients



Recovery curve i.e. average intensity in bleach region



256x256 pixel images, from Schuster et al. (2014)

RI. SE

Numerical model

- Solve diffusion equation $\frac{\partial c}{\partial t} = D\nabla^2 c$ numerically on a grid corresponding to the image we are simulating (256x256 plus padding around it -> 512x512 grid)
- Direct time-stepping too slow -> use spectral/Fourier domain methods -> analytical time-stepping in Fourier domain
- Initial condition is the bleached concentration, so what we solve is the time evolution back to equilibrium (uniform distribution of concentration)

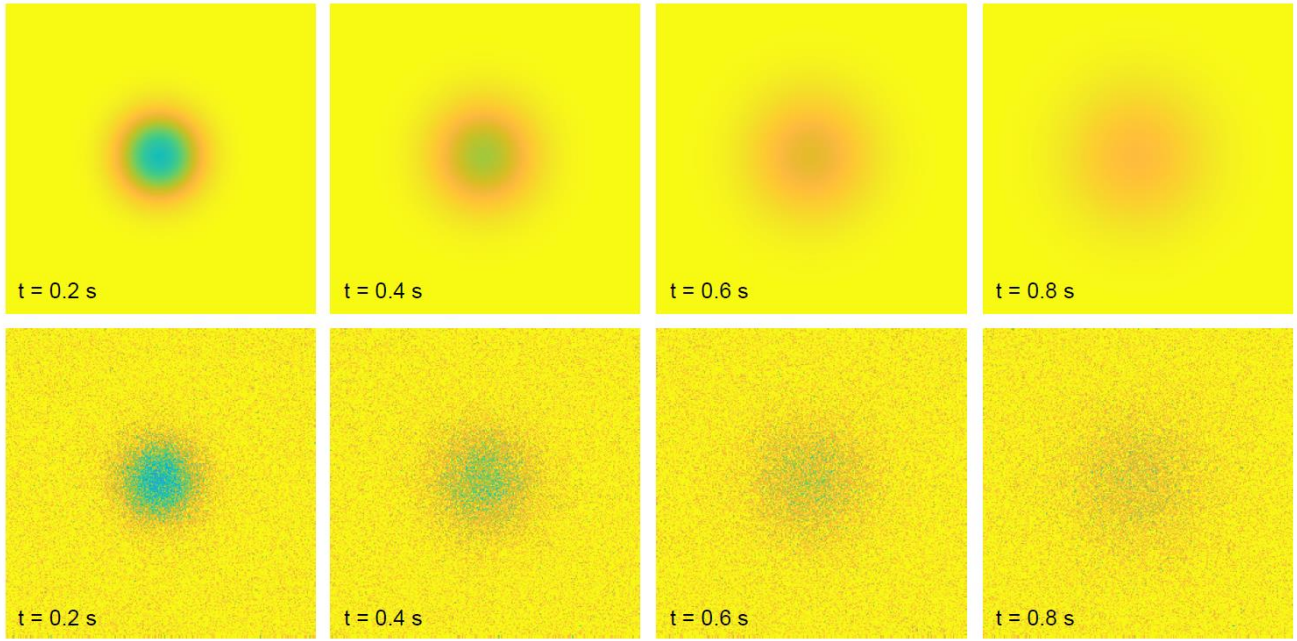
M. Röding, L. Lacroix, A. Krona, T. Gebäck, N. Lorén. A highly accurate pixel-based FRAP model based on spectral-domain numerical methods. *Biophysical Journal*, 116(7), 2019, 1348-1361.

RI. SE

Comparison between ”micro” and ”macro”

Deterministic (numerical solution to diff eq)

Stochastic (Gaussian random walks)



Parameter estimation

- Assume normal distributed noise with constant variance, and do non-linear least squares fitting of the model to either
  - the recovery curve data (just a function of time, most common)
  - the full image/video data (2D + t)
- Sometimes also weighted least squares or maximum likelihood (depending on noise assumptions in the data)

RI.
SE

Computational speed is sometimes a problem

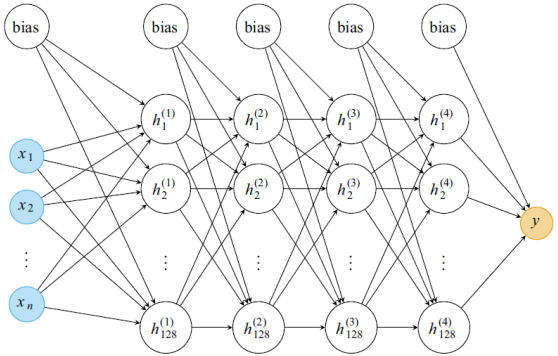
- Detailed numerical models can be heavy to run
- Nonlinear least squares should entail multiple fits with different initial parameter guesses to ensure global optimum
- Hence, analysing large batches of data can take time

- **Can we use neural networks to speed up parameter estimation?**
- **Yes we can…**

RI.
SE

Artificial neural networks (ANNs) for predicting parameters

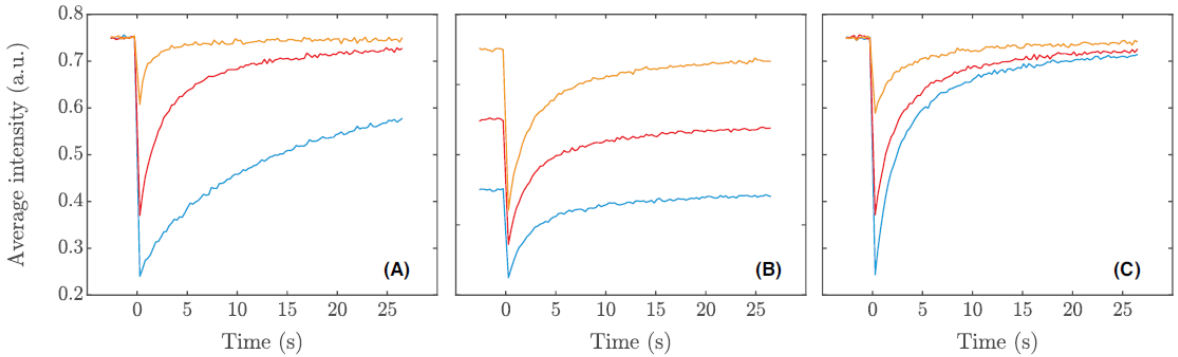
- Classical (fully-connected) ANNs can be used for classification (predicting class memberships) and **regression (predicting numerical values)**
- An ANN is essentially a composition of operations that together form an arbitrarily complex nonlinear mapping from input to output
- In each *layer* of the ANN, a number of *nodes* receives a weighted sum of input, applies a nonlinear *activation function* and sends the result to the nodes of the next layer



Data for ANN

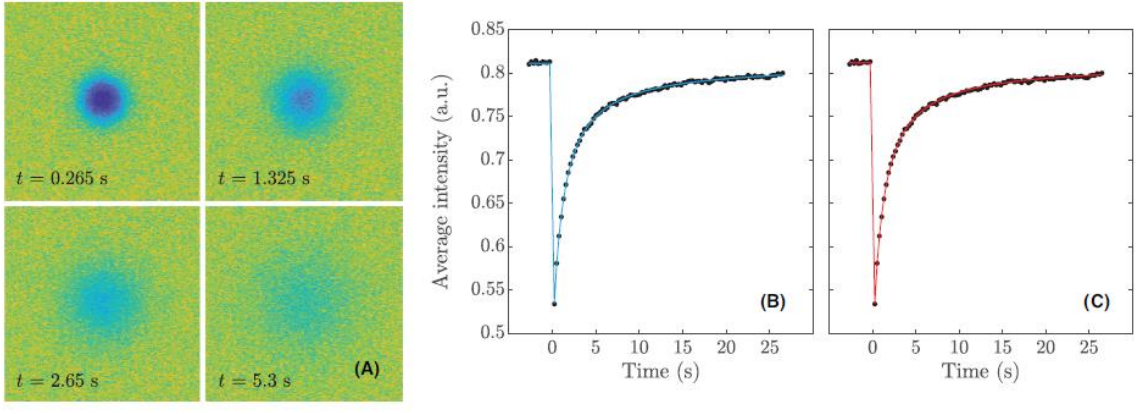
- Three parameters to predict: Diffusion coefficient D, initial intensity/conc c_0, bleaching depth α
- Also, variance of normal distributed image noise, a
- Generate 2^{20} (~1,000,000) for training and 2^{18} (~250,000) for validation and test

Parameter	Distribution
D	Log-uniform in $[10^{-12}, 10^{-9}]$ m^2/s
c_0	Uniform in [0.5, 1]
α	Uniform in [0.45, 0.95]
a	Log-uniform in $[10^{-4}, 10^{-2}]$



Experimental validation

- One example fit to the data (A), showing almost identically the same recovery curve fit for least squares (B) as for neural networks (C)



Skärström, V. W., Krona, A., Lorén, N., & Röding, M. (2020). DeepFRAP: Fast fluorescence recovery after photobleaching data analysis using deep neural networks. Journal of Microscopy. 2021, 282, 146-161.

Finishing off…

- If you want to discuss, get more info or papers sent on these topics today, just contact me at

magnus.roding@ri.se