# In vivo deep learning estimation of diffusion coefficients of nanoparticles

Julius B. Kirkegaard[1,2,5⋆], Nikolay P. Kutuzov[3⋆], Rasmus Netterstrøm[1], Sune Darkner[1], Martin Lauritzen[3,4], and François Lauze[1]

[1] Department of Computer Science, University of Copenhagen
[2] Niels Bohr Institute, University of Copenhagen
[3] Department of Neuroscience, University of Copenhagen
[4] Department of Clinical Neurophysiology, Rigshospitalet Denmark
[5] Corresponding author: juki@di.ku.dk

**Abstract.** Understanding the transport of molecules in the brain *in vivo* is the key to learning how the brain regulates its metabolism, how brain pathologies develop, and how most of the developed brain-targeted drugs fail. Two-photon microscopy – the main tool for *in vivo* brain imaging – achieves sub-micrometer resolution and high image contrast when imaging cells, blood vessels, and other microscopic structures. However, images of small and fast-moving objects, e.g. nanoparticles, are ill-suited for analysis of transport with standard methods, e.g. super-localization, because of (i) low photon budgets resulting in noisy images; (ii) severe motion blur due to slow pixel-by-pixel image acquisition by two-photon microscopy; and (iii) high density of tracked objects, preventing their individual localization. Here, we developed a deep learning-based estimator of diffusion coefficients of nanoparticles directly from movies recorded with two-photon microscopy *in vivo*. We've benchmarked the method with synthetic data, model experimental data (nanoparticles in water), and *in vivo* data (nanoparticles in the brain). Our method robustly estimates the diffusion coefficient of nanoparticles from movies with severe motion blur and movies with high nanoparticle densities, where, in contrast to the classic algorithms, the deep learning estimator's accuracy improves with increasing density. As a result, the deep learning-based estimator facilitates the estimation of diffusion coefficients of nanoparticles in the brain *in vivo*, where the existing estimators fail.

**Keywords:** diffusion · video regression · two-photon microscopy · brain

## 1 Introduction

The low treatment efficiency of progressive and fatal neurodegenerative disorders, e.g., Alzheimer's disease and Parkinson's disease, is partly due to many transport barriers along the delivery routes of the drugs [10]. To identify and

⋆ These authors contributed equally to this work.

study these barriers, one needs to estimate diffusion coefficients (DCs) of potential drugs, e.g., brain-targeted nanoparticles (NPs), locally, near individual cells or blood vessels [7].

A commonly used approach estimates the diffusion coefficient of an NP by fitting a straight line to the graph of its mean-squared displacement as a function of time [11, 7]. Maximum-likelihood estimators give unbiased estimates of DCs with the lowest variance compared to other estimators. This, however, requires accurate localization and identification of the NPs, and using a suitable physical theory for the studied motion, which is only possible in relatively simple systems, e.g., freely diffusing molecules, diffusion of a molecule on a fluctuating substrate [16] or diffusion of a molecule confined in a volume of a simple shape [9].

Recent work has shown that deep learning (DL) can aid prediction of diffusive behavior. For example, DL methods have been shown to able to (i) classify based on particle trajectories if diffusive behavior is anomalous or confined [6, 3], (ii) predict diffusion behavior in porous media directly from the structure of the media [17, 2], (iii) infer time-varying diffusion properties from trajectories [12], and (iv) predict DCs when particles undergo flow and thermal gradients at high signal-to-noise ratios [13].

Here we introduce a deep learning-based estimator (DLE) tailored for estimating DCs of NPs directly from images recorded with two-photon microscopy (2PM) in the brain *in vivo*. Such images are often distorted by high motion blur and high NPs density within the imaging volume. High motion blur occurs because 2PM is scanning microscopy (Fig. 1AB), which records images pixel-by-pixel (much slower than wide-field imaging). High NP density can occur near vessel walls, where their diffusion is limited [8], near a damaged blood-brain barrier (BBB) leading to a release of NPs from blood into the brain, or near a dead-end of the brain extracellular spaces (ECS), where NPs get stuck and cannot escape. High motion blur and high NP densities restrict the use of classical diffusion estimation *in vivo* typically due to the inability to accurately localize NPs on recorded images. Furthermore, linking positions of NPs into trajectories becomes ambiguous and eventually fails when the distance between neighboring NPs on the images becomes comparable or less than an average displacement of NPs between consecutive movie frames.

DLE estimates DCs of NPs directly from recorded movies, skipping the localization of NPs on individual images (Fig. 2A). As a result, DLE, trained on simulated data, estimated DCs more precisely than the standard particle-tracking methods, especially so when images are corrupted by high motion blur and when the density of NPs is high.

## 2 Materials and methods

### 2.1 Two-photon microscopy

We used a commercial (Olympus, Japan) two-photon microscope (FVMPE-RS), equipped with a Ti:Sapphire laser and GaAsP detectors (photomultiplier tubes;

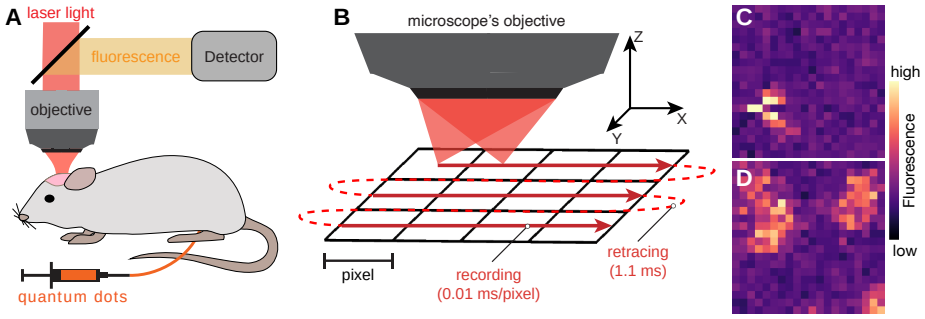


**Fig. 1.** ***Image acquisition with two-photon microscopy (2PM).*** **A**. A schematic of an *in vivo* experimental setup: 2PM records images of quantum dots (QDs) inside the brain of a living, anaesthetised mouse by sweeping the laser beam across the brain and collecting fluorescence emitted by the QDs. **B**. In scanning microscopy (confocal, 2PM) images are recorded one pixel at a time by moving the laser beam along straight lines (solid red). Between the lines, the beam is retraced (dashed red) for unidirectional scanning. **C**. An image of a QD in the brain of a mouse, showing high motion blur: The QD moves between the recording of two consecutive image lines. **D**. An image of several QDs in the brain of a mouse, showing overlapping fluorescence intensity distributions when multiple QDs move close to each other. Pixels on images are 200 nm wide.

PMTs). We used QTracker655 quantum dots (QDs; ThermoFisher Scientific, USA) and FluoSpheres nanobeads (NBs; ThermoFisher Scientific, USA) for *in vivo* brain and *in vitro* (beads in water) experiments, respectively.

### 2.2 Animals and surgical procedures

All protocols involving mice were approved by the Danish National Committee on Health Research Ethics, in accord with the guidelines of the European Council's Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes, and complied with the ARRIVE guidelines. We used five male wild-type mice (C57bl/6j; Janvier-labs) and followed standard acute surgery [7]. We injected QDs intravenously and released them in the brain parenchyma by ablating a small capillary with high laser power.

### 2.3 Synthetic data generation

The simulation of realistic 2PM images (training data) of random walkers consisted of four steps. First, we simulated isotropic free diffusion in 3D space by adding normally distributed displacements with zero expected value and variance (same for each of the three coordinates), $\sigma^2 = 2D\Delta t$, where $D$ is the DC and $\Delta t$ is the time step. Second, we simulated pixel-by-pixel image acquisition, where a pixel with row and column indexes $r$ and $c$, respectively, is recorded at time $t_{rc} = t_0 + (rC + c)/v + \tau r$, where $v$ is the laser beam scanning speed, $\tau$ is the retracing time (Fig. 1B), and $R, C$ are the total numbers of rows and columns

in an image. We trained our network on a small window size, which required a trivial adjustment of $t_{rc}$ depending on the cropping size of the experimental videos. Third, we calculated the expected value of a pixel $I_{rc}$ assuming the 3D Gaussian point-spread function [14] with STDs $\sigma_x = \sigma_y = \sigma_{xy}$ for the focal plane and $\sigma_z$ in the direction of the principle axis of the objective ($z$-axis).

$$I_{rc} = I_{bg} + \sum_p \frac{I_{qd}}{\sqrt{8\pi^3\sigma_{xy}^4\sigma_z^2}} \exp\left(-\frac{(c\Delta x - x_p)^2 + (r\Delta x - y_p)^2}{2\sigma_{xy}^2} - \frac{z_p^2}{2\sigma_z^2}\right), \quad (1)$$

where $I_{bg}$, $I_{qd}$ are the background and QDs intensities, respectively, and the index $p$ runs over all QDs. We used experimentally measured PSF parameters, $\sigma_{xy} = 0.2$ µm and $\sigma_z = 1.4$ µm from [8], but add variation by sampling from a normal distribution with standard deviation 0.05 µm for $\sigma_{xy}$ and 0.2 µm for $\sigma_z$. Forth, we added noise to each pixel by sampling a Poisson-distributed random number with expected value $I_{rc}$ (see above) to simulate shot noise and, finally, added Gaussian noise to simulate other remaining sources of noise (e.g. read-out noise of the detector). See SI for code.

For validation, we have generated two datasets of synthetic movies: A dataset consisting of 25,000 videos with low DCs sampled uniformly between 0 and 1.5 µm$^2$/sec, and a dataset consisting of 25,000 videos with high DCs sampled uniformly between 0 and 50 µm$^2$/sec. In the paper we refer to these two datasets as *low diffusion coefficient* and *high diffusion coefficient* datasets.

### 2.4 Neural architecture

The input to our network is a sequence (a movie) of gray-scale images of shape $(T, W, H)$ (time, width, height). For training, we used $T = 400$ images of size $W = H = 20$ pixels. Our network consists of (Fig. 2A): (i) An auto-encoder (five standard convolutional layers, with GeLU activations) that maps the images to a latent dimension $h = 128$; (ii) A deep 1D resnet [5] which performs convolutions along the temporal dimension (we used seven resnet layers, each with two convolutions, and initial kernel size of seven time steps); (iii) An adaptive average pooling along the temporal dimension, which enables generalization beyond a fixed movie length; and (iv) a simple multi-layered perceptron (MLP), which outputs both an estimate of the DC $\hat{D}$ and of the error of the estimate $\hat{\sigma}$. Both outputs, the diffusion coefficient and the estimated erro,r are constrained to be positive by using the nonlinearity $f(x) = 1 + \mathrm{ELU}(x)$. The full networks are provided as PyTorch code in the SI.

The encoder network is trailed separately to minimize an autoencoder L2 loss. We use an Adam optimizer with learning rate $5 \cdot 10^{-4}$ and take 5,000 steps of batch size 512. We train the predictor network to maximize a Gaussian likelihood, corresponding to minimizing the loss

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{(\hat{D}_i - D_i)^2}{2\hat{\sigma_i}^2} + \log(\hat{\sigma}_i)\right], \quad (2)$$

which combines an L2 loss on the diffusion coefficient estimation while simultaneously punishing mistakes in the error prediction. The network is trained for 25,000 steps with batch size 50 using an Adam optimizer with learning rate 0.01.

## 3 Results

Figure. 2A shows the architecture of DLE (Methods) and compares it with a standard data analysis pipeline for estimating DCs of particles from their images. DLE requires annotated data for training, which is a challenge because a typical *in vivo* experiment generates only small volumes of data, not enough for the training and, most importantly, one cannot annotate DCs, $D$, for *in vivo* data as there is no alternative method to measure $D$. Therefore, we generated realistic 2PM images of a population of random walkers with known DCs (see "Methods: Synthetic data generation"). Figure 2BC illustrates the visual consistency of experimental and synthetic images at both low and high DCs.

First, we tested the performance of DLE on synthetic data against a very common maximum-likelihood, sub-pixel single-particle tracking estimation of the DCs [1], implemented by using the Python *TrackPy* package. We refer to this method as to the "classic" method (estimator) in the paper. Table 1 shows the average performances of DLE for the synthetic movies with low DC, which should be compared to the zero-knowledge baselines (i.e. always guessing the mean of the prior), e.g. $L_1 = \langle |x - 0.75| \rangle_x = 0.375$. DLE outperforms the classic estimator by 50% on the entire dataset, based on $L_1$ loss. In experimental movies with low NP densities many movie frames do not show any NPs and so do the synthetically generated images. Because empty images may comprise up to half of the generated images, we also evaluated the models on the 50% of the dataset where they performed best, chosen according to their own output error $\hat{\sigma}$. Here, DLE improved estimation of the DCs by 134% in $L_1$ and by 500% in $L_2$ loss, compared to the classic estimator (Table 1). For the synthetic data, generated with high DCs, DLE improved the estimation by 322% in $L_1$ and by 1020% in $L_2$ loss in DLE, compared to the classic estimator (Table 1).

Second, we compared the performance of DLE to the classic estimator for an increasing number of NPs seen in the synthetic images. To do that, we calculated the performance as a function of the number of possibly visible NPs in the synthetic videos, for low and high DCs (Fig. 3AB). We evaluate this by, for each frame of the synthetic input video, finding the number of NPs that are within the imaging volume of the microscope's objective, meaning that $x, y$–coordinates of the NPs are within the image $x, y$ ranges and that NPs are in focus, i.e., their z-coordinates are no further away from the focal plane than 1 STD of the microscope's PSF ($\sigma_z$; see Methods). Summing this over the entire video clip estimates the number of possibly visible transitions (although some might be heavily obscured by noise). The number of transitions is proportional to NP density on average, but better captures average information availability in the videos. We will use the NP density and number of transitions interchangeably

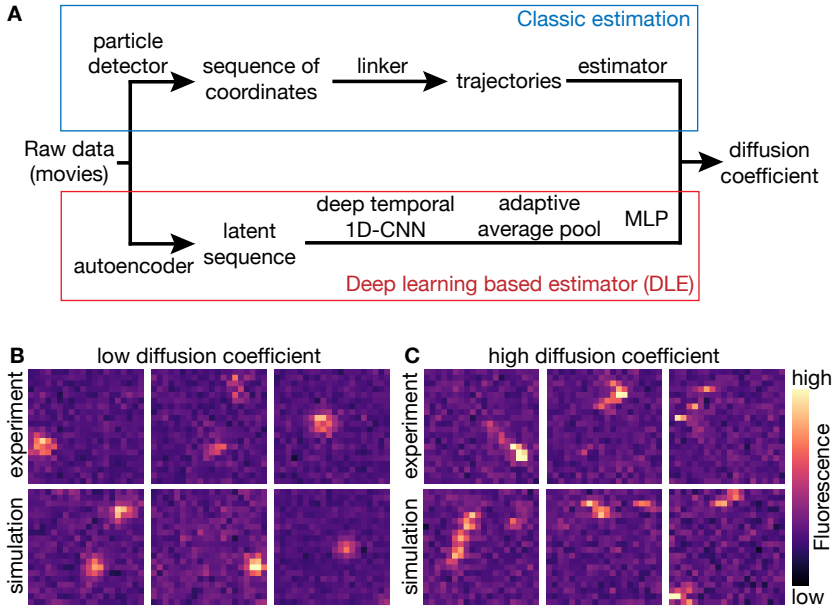


**Fig. 2.** ***Architecture of the deep-learning-based estimator (DLE) and generation of the training data.*** **A.** In classic estimation, the positions of particles are estimated on every image of a movie, the positions are linked into trajectories, and DCs are estimated from these trajectories. DLE estimated DCs directly from raw experimental movies without the need for localization of particles on individual movie frames (see Methods for details). **B,C.** Examples of real experimental images (top rows) and simulated images (bottom row) for low (**B**; QDs in the brain) and high (**C**; nanobeads in water) DCs. Magenta and white show low and high fluorescence intensity, respectively. Pixels on images are 200 nm wide.

below. DLE outperforms the classic estimator over the entire studied range of NP densities (more than two orders of magnitude range). Naturally, both methods perform the worst when there are very few NPs visible in some of the images, i.e., there is very little information about their transport (Fig. 3AB). The low diffusion dataset contains high-density videos, and here the classic approach also performs poorly when there are many NPs visible on single images, which decreases tracking accuracy. In contrast, DLE's performance keeps increasing even at the highest densities (Fig. 3A).

Third, we've shown that DLE returned correct DCs of 50-nm diameter nanobeads in water, where we calculated an estimated true theoretical DC of the nanobeads using the Stokes–Einstein equation (Fig. 3C). The images of nanobeads in water are so severely corrupted by the motion blur (Fig. 2C), classic tracking algorithms, based on blob-detection, fail to localize them because they don't image like blobs. Instead, the images of single nanobeads sometimes show multimodal intensity distributions without a clearly defined center.

| | All | | | Top 50% | | |
|---|---|---|---|---|---|---|
| | $L_1$ | $L_2$ | $\mathcal{L}$ | $L_1$ | $L_2$ | $\mathcal{L}$ |
| | | | Low diffusion coefficients | | | |
| Baseline | 0.375 | 0.187 | -0.33 | / | / | / |
| Classic | 0.163 | 0.0588 | -1.33 | 0.0860 | 0.0138 | -1.71 |
| DLE | 0.109 | 0.0351 | -2.12 | 0.0368 | 0.0023 | -2.82 |
| Difference (%) | 50 | 67 | / | 134 | 500 | / |
| | | | High diffusion coefficients | | | |
| Baseline | 12.5 | 208 | 3.17 | / | / | / |
| Classic | 13.4 | 301 | 56.3 | 12.3 | 294 | 106 |
| DLE | 5.72 | 64.9 | 1.82 | 2.91 | 18.6 | 1.12 |
| Difference (%) | 118 | 220 | / | 322 | 1020 | / |

**Table 1.** Comparison of DLE with classical DC estimation [1], evaluated on synthetically generated movies with high and low DCs. The full dataset includes many empty frames that skew the results of both methods negatively. The top-50% column is evaluated only on half the dataset, self-chosen by $\hat{\sigma}$ of the models. Columns are average $L_1$, $L_2$ losses and Gaussian likelihood loss, $\mathcal{L}$, as defined in Eq. (2).

Finally, we used DLE to estimate DCs of QDs in the brain ECS of living, anaesthetised, mice *in vivo* before and after heart stop. Before stopping the heart, we estimated DCs from three movies recorded in the same brain location: $D = 0.87 \pm 0.08$, $0.78 \pm 0.08$, and $0.80 \pm 0.09$ $\mu$m$^2$/s (first three points in Fig. 3D)). The values are consistent with a single value and fall into the range of DCs measured previously for similar QDs in the brain [4][Fig.2B]. After the heart stop, DCs of the QDs fell to zero, $D = 0.004 \pm 0.012$ $\mu$m$^2$/s, (last point in Fig. 3D), consistent with shrinking brain ECS [15], where QDs diffuse. Last but not least, DLE could estimate DCs of QDs in another region of brain ECS, $D = 0.338 \pm 0.015$ $\mu$m$^2$/s (also consistent with [4][Fig.2B]), despite many overlapping QDs on individual images. Classic estimators rely on a search range parameter (how far away from a current position should we search for a particle on the next image) for linking positions of NPs into trajectories. DCs estimated by the classic estimator were unreliable: The estimates varied by more than one order of magnitude with increasing search range (Fig. 3E).

## 4 Discussion

We have developed a robust and efficient tool for estimating DCs of NPs from microscopic images, where standard particle tracking methods fail, i.e., from images characterized by high motion blur and high particle densities. Such images occur often in *in vivo* experiments: When imaging mouse brains with 2PM, the image sampling rate is initially low due to pixel-by-pixel image acquisition and

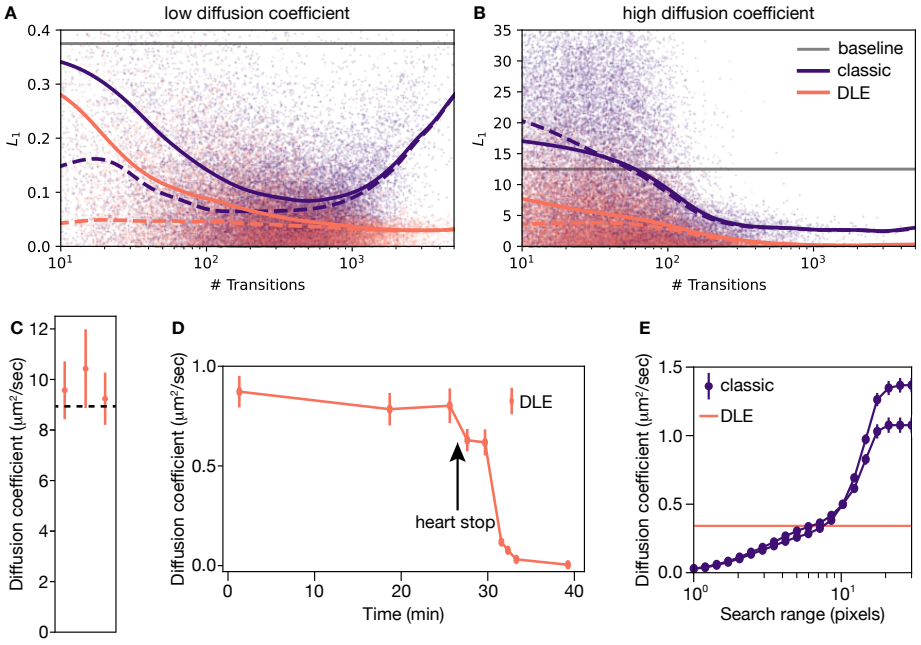


**Fig. 3. Benchmarks and performance of DLE**. **A, B.** $L_1$ loss $|D - \hat{D}|$ (rolling average) as a function of possible number of observable particle transitions (see text). Dashed lines were calculated using half of the data with the lowest error, $\hat{\sigma}$, on DC (as in Table 1). Scatter plots show $L_1$ estimates for individual movies. **C**. Three DLE estimates of the DCs of 50-nm nanobeads in water (from three regions within the same movie) are consistent with the expected DC, $D = 8.8$ $\mu$m$^2$/s (dashed line), estimated with the Einstein–Stokes relation for 50-nm diameter spheres in water at 21°C. **D**. DCs of QDs in brain ECS fall to zero following a heart stop, explained by the shrinking of brain ECS. **E**. At high density of QDs on images (Fig. 1D), DCs (two curves estimated from two parts of the same movie) returned by the classic estimators largely depend on the "search range" parameter (see text), making the estimates unreliable. DLE doesn't have tunable hyper-parameters, and it returns two consistent with each other values of DCs from the two movie parts: $D = 0.332 \pm 0.024$ $\mu$m$^2$/s and $0.343 \pm 0.020$ $\mu$m$^2$/s. The red line shows the weighted average of two values, $D = 0.338 \pm 0.015$ $\mu$m$^2$/s.

is typically decreased even more to collect more photons (many are lost due to scattering), causing severe motion blur of fast-moving NPs.

As expected, DLE outperforms the classic estimators the most when the motion blur or NPs density is high (Table 1). In contrast to classic estimators, DLE has the potential to use the complex, severely blurred intensity distributions of NPs on individual movie frames as additional information for estimating their DCs.

Increasing performance of DLE with increasing NPs density on recorded images makes DLE the method of choice for studying transport *in vivo* under pathological conditions. For example, high densities of NPs will occur near blood

vessels with local disruption of the BBB or within shrinking brain ECS, following swelling of brain cells (e.g. in terminal ischaemia).

By tracking QDs in the brain ECS, we showed that DLE is perfectly suited for high-throughput preclinical studies of brain-targeted drugs. DLE also applies to study transport in any scattering biological tissue imaged with 2PM, e.g., in the heart, liver, kidney, and skin. DLE has the potential to become a standard tool for estimating a very wide range of DCs from any images recorded with scanning confocal or multi-photon microscopes, regardless of the degree of motion blur and particle density.

**Acknowledgements** This work was supported by research grants from Novo Nordisk Foundation, grant number NNF20OC0062047; Danish Cardiovascular Academy, which is funded by the Novo Nordisk Foundation, grant number NNF20SA0067242; The Lundbeck Foundation; the Independent Research Fund Denmark; the Novo Nordisk Foundation.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References


1. John C Crocker and David G Grier. Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.*, 179(1):298–310, 1996.
2. Krzysztof M Graczyk, Dawid Strzelczyk, and Maciej Matyka. Deep learning for diffusion in porous media. *Sci. Rep.*, 13(1):9769, 2023.
3. Naor Granik, Lucien E Weiss, Elias Nehme, Maayan Levin, Michael Chein, Eran Perlson, Yael Roichman, and Yoav Shechtman. Single-Particle diffusion characterization by deep learning. *Biophys. J.*, 117(2):185–192, 2019.
4. Diego Grassi, Agata Idziak, Antony Lee, Ivo Calaresu, Jean-Baptiste Sibarita, Laurent Cognet, U Valentin Nägerl, and Laurent Groc. Nanoscale and functional heterogeneity of the hippocampal extracellular space. *Cell Rep.*, 42(5):112478, 2023.
5. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
6. Patrycja Kowalek, Hanna Loch-Olszewska, and Janusz Szwabiński. Classification of diffusion modes in single-particle tracking data: Feature-based versus deep-learning approach. *Phys Rev E*, 100(3-1):032410, 2019.
7. Krzysztof Kucharz, Kasper Kristensen, Kasper Bendix Johnsen, Mette Aagaard Lund, Micael Lønstrup, Torben Moos, Thomas Lars Andresen, and Martin Johannes Lauritzen. Post-capillary venules are the key locus for transcytosis-mediated brain delivery of therapeutic nanoparticles. *Nat. Commun.*, 12(1):4121, 2021.
8. Nikolay Kutuzov, Henrik Flyvbjerg, and Martin Lauritzen. Contributions of the glycocalyx, endothelium, and extravascular compartment to the blood–brain barrier. *PNAS*, 115(40):E9429–E9438, 2018.
9. Kim I Mortensen, Henrik Flyvbjerg, and Jonas N Pedersen. Confined brownian motion tracked with motion blur: Estimating diffusion coefficient and size of confining space. *Frontiers in Physics*, 8:601, 2021.
10. William M Pardridge. Blood-Brain barrier and delivery of protein and gene therapeutics to brain. *Front. Aging Neurosci.*, 11:373, 2019.
11. H Qian, M P Sheetz, and E L Elson. Single particle tracking. analysis of diffusion and flow in two-dimensional systems. *Biophys. J.*, 60(4):910–921, 1991.
12. Borja Requena, Sergi Masó-Orriols, Joan Bertran, Maciej Lewenstein, Carlo Manzo, and Gorka Muñoz-Gil. Inferring pointwise diffusion properties of single trajectories with deep learning. *Biophys. J.*, 122(22):4360–4369, 2023.
13. Pranshul Sardana and Steven T Werely. Deep particle diffusometry: convolutional neural networks for particle diffusometry in the presence of flows and thermal gradients. *Measurement Science and Technology*, 35, 2024.
14. C J R Sheppard and Min Gu. Image formation in two-photon fluorescence microscopy. *Optik*, 86(3):104–106, 1990.
15. Robert G Thorne and Charles Nicholson. In vivo diffusion analysis with quantum dots and dextrans predicts the width of brain extracellular space. *PNAS*, 103(14):5567–5572, 2006.
16. Christian L Vestergaard, Paul C Blainey, and Henrik Flyvbjerg. Optimal estimation of diffusion coefficients from single-particle trajectories. *Phys. Rev. E*, 89(2):022726, 2014.
17. Haiyi Wu, Wen-Zhen Fang, Qinjun Kang, Wen-Quan Tao, and Rui Qiao. Predicting effective diffusivity of porous media from images by deep learning. *Sci. Rep.*, 9(1):20387, December 2019.