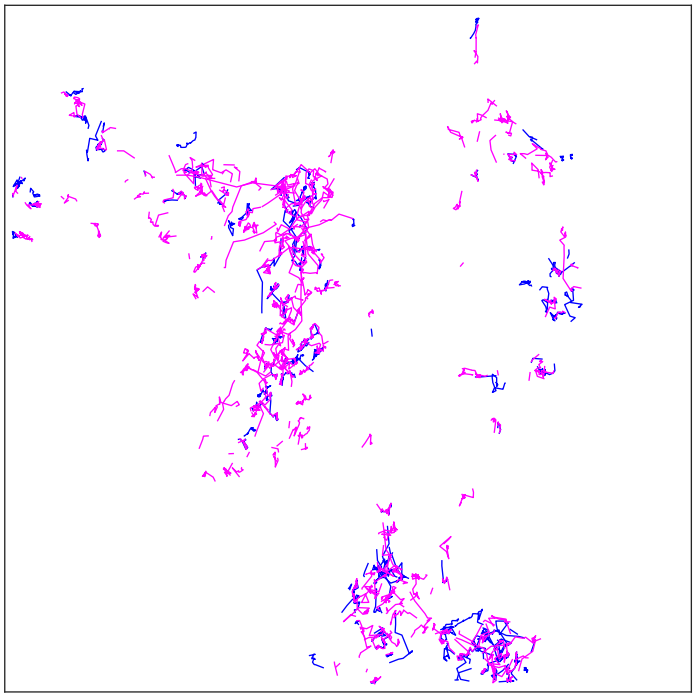
CHALMERS
UNIVERSITY OF TECHNOLOGY



# Exploring correlation between the A$\beta$-peptide accumulation and endosome transport in Alzheimer's disease using particle tracking

Master's thesis in Biomedical Engineering

Alice Deimante Neimantaite

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Master's thesis 2019

# Exploring correlation between the A$\beta$-peptide accumulation and endosome transport in Alzheimer's disease using particle tracking

Alice Deimante Neimantaite


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
*Division of Signal processing and Biomedical engineering*
Chalmers University of Technology
Gothenburg, Sweden 2019

Exploring correlation between the A$\beta$-peptide accumulation and endosome transport in Alzheimer's disease using particle tracking
Alice Deimante Neimantaite



Supervisors:
Elin Esbjörner Winters, Department of Biology and Biological Engineering
Emelie Vilhelmsson Wesén, Department of Biology and Biological Engineering
Daniel Midtvedt, Department of Physics
Examiner: Fredrik Kahl, Department of Electrical Engineering

Master's Thesis 2019
Department of Electrical Engineering
Division of Signal processing and Biomedical engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Lysosome (magenta) and A$\beta$ (blue) movement over 30 frames estimated using the best performing particle tracking method for the application.

Typeset in LaTeX
Gothenburg, Sweden 2019

Exploring correlation between the A$\beta$-peptide accumulation and endosome transport in Alzheimer's disease using particle tracking

Alice Deimante Neimantaite
Department of Electrical Engineering
Chalmers University of Technology

# Abstract

Alzheimer's disease results from pathological changes and degeneration of neurons in the brain. Pathological hallmarks of the disease include accumulation of amyloid beta, A$\beta$ (a small peptide), as well as defects in nerve cell endocytosis and the trafficking of endosomes. It is known that uptake and/or overproduction of A$\beta$ peptides in endosomes result in intraneuronal accumulations that could drive endosome enlargement. The aim of this Thesis is to develop particle tracking software applicable for analysis of the relationship between endosome transport and A$\beta$ trafficking in Alzheimer's disease cell models. The work involved systematic investigation and testing of suitable particle tracking methods using simulated as well as real fluorescence microscopy data; particular focus was on identifying methods for particle localization and trajectory linking. The best performing method was applied on live cell imaging data of A$\beta$ and lysosomes (which are the destination for maturing endosomes in the living cell) to obtain their trajectories; moreover time-dependent effects on A$\beta$ mean displacement and velocity were also deciphered from the trajectories. The imaging data was acquired over 24 hours post addition of A$\beta$ to lysosome-labelled cells. Further, a particle trajectory correlation analysis method was established to explore co-movement of A$\beta$ and lysosomes. The results show that the fraction of A$\beta$ trajectories correlating with lysosome trajectories increased with time, consistent with their biological colocalization. This result was also in good agreement with object-based colocalization analysis on the same data set. Altogether, the results suggest that particle tracking analysis and trajectory correlation could be a promising tool to understand the biological basis of A$\beta$ accumulations in Alzheimer's disease.

Keywords: Alzheimer's disease, amyloid beta, endosomal pathway, particle tracking, dynamical colocalization

# Acknowledgements

I would like to deeply thank my supervisor Associate Professor Elin Esbjörner Winters for creating this project for me, including me in her research group and for all of the support from her during the project. The project and experiences with the research group and the whole department have been better and more rewarding than I could have ever hoped for. I would also like thank my other supervisors Emelie Vilhelmsson Wesén and Daniel Midtvedt, and my examiner Fredrik Kahl for all the help, shared knowledge and support during the project.

I am also thankful to the researchers at the university who provided me with their developed software and guidance to adapt the software to this application, namely Daniel Midtvedt and his research group, Karl Granström, Quentin Lubart, Saga Helgadottir and Aykut Argun.

Lastly I would like to express my gratitude to my co-workers, friends and grandmother for endless support, encouragement and exchange of ideas during the work.

Alice Deimante Neimantaite, Gothenburg, June 2019

# Declaration of contributions

This project was created at the Division of Chemical Biology, Department of Biology and Biological Engineering at Chalmers University of technology. It was done by Alice Deimante Neimantaite in collaboration with Emelie Vilhelmsson Wesén who is a PhD student at the department. The experimental data collection including sample preparation and fluorescence imaging was done by Emelie. The collected data processing, software application, testing and development was done by Alice Deimante.

# Contents

# List of Figures

# List of Tables

# 1
# Introduction

Alzheimer's disease is a growing threat to society as the probability to get the disease increases with age [1], and people are predicted to live longer. At the moment, there is no cure to the disease as its cause and machinery are not completely understood. [2] People are not only at risk of getting Alzheimer's at old age but also as early-onset dementia (before 60-67 years old), which is typically inherited (due to genetic mutations). [3] The disease is progressive, from relatively mild cognitive impairment, to full dementia, personality changes and loss of ability to perform simple tasks like having a conversation. [4]

Alzheimer's disease results from pathological changes and degeneration of the brain. [5] Pathological hallmarks of the disease include accumulation of amyloid beta, A$\beta$ (a small peptide), reduced synaptic function, tau protein aggregation and dysfunctional endocytosis in nerve cells. [6, 7] A comparison between a healthy and Alzheimer's diseased nerve cell can be seen in Figure 1.1. Endosomes are membrane-enclosed organelles that are part of a machinery that all cells use to take up, sort, transport, deliver and degrade proteins and other macromolecular cargoes. This process can be described as a pathway which starts with uptake (endocytosis) and subsequent inward transport of cargo to early endosomes, late endosomes and finally lysosomes. [8] The early endosome is also known to be a production site of the A$\beta$ peptide. It is known that uptake and/or overproduction of A$\beta$ peptides in endosomes lead to intraneuronal accumulation and endosome enlargement. This is one of the earliest pathological alterations to the Alzheimer's disease brain. [7] However, very little is known about how the A$\beta$ peptide is trafficked through different endosomes. It is therefore of interest to examine the relation between the endosomes and A$\beta$ peptides in models that can mimic Alzheimer's disease conditions. This could contribute to the knowledge of Alzheimer's machinery. The relation between A$\beta$ and the endosomal pathway is schematically highlighted in Figure 1.2.

Particles like A$\beta$ peptides, structures of cells and their dynamics can be imaged using fluorescence microscopy. [9] However, their relationships and dynamical changes are difficult to detect by mere observance of microscopy images and movies. Image analysis methods are advantageous for this purpose. For the specific application at hand, the image analysis method particle tracking can be exploited to monitor A$\beta$ and endosomes in time and space on single particle basis. The information obtained from particle tracking can be further analysed by correlation analysis methods to add quantitative information of colocalization and co-movement of the particles.

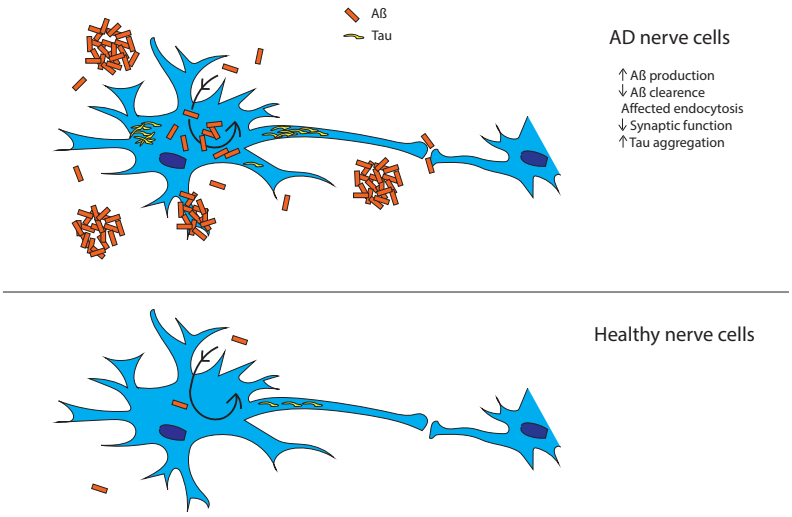



**Figure 1.1:** An overall comparison of a healthy nerve cell and a nerve cell with Alzheimer's disease.

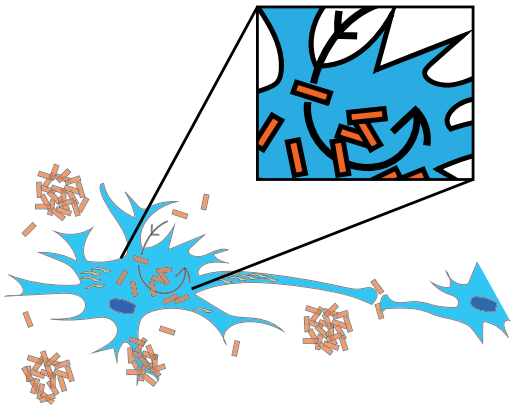


**Figure 1.2:** Investigation focus of Alzheimer's disease in the project: the correlation between endosome transport and A$\beta$ trafficking.

## 1.1 The project

This project aims to develop particle tracking software applicable for analysis of the relationship between endosome transport and A$\beta$ trafficking in Alzheimer's disease. As many different methods for particle tracking in biology and for other applications have been developed, the focus of this project is to analyse and test which one will be best suited this application. Performance tests will be done on both experimental and simulated data; realistic data sets for the latter will also be developed. In order to analyse the relation between the particles of interest from obtained particle trajectories, a method for correlation will be developed building on the tracking method that is judged to perform best for this project. The aim of the correla-

tion method is to determine how much and for how long endosomes and A$\beta$ move together, meaning a measure of correlation in space and time.

### 1.1.1 Limitations

- As there is a broad set of particle tracking methods, all cannot be evaluated
- Online available and locally developed methods will therefore be applied
- The methods/code used will be adapted to suit our specific data
- The adaptation will be done as good as possible within the time frame
- Simulated data will only be created to mimic our specific data and the methods' performance test will be done on that specific simulation, meaning that the performance results are specific for the application and are not general for all types of data

# 2
# Background

## 2.1 Particle tracking

Particle tracking is applicable in a variety of applications, originally created for military and aerospace tracking purposes, and now including live-cell microscopy. [10] There is a huge amount of particle tracking methods and several analyses have been made to find the optimal one(s).[11] The main factor showing from both objective and less objective comparisons is that the optimal method depends on the specific application, even though there are methods performing overall better than others.

Particle tracking can be divided into two steps, namely the identification of particles and association of the particles over time. [10] The two steps will be referred to as localization and linking in this report.

### 2.1.1 Localization

The end result of localization is in the form of x-and y-coordinates of the center positions of the found particles. To end up there, several approaches can be used and in this report they are divided into two classes, namely the algorithmic approach and the data driven approach, as their methodology is quite deviating. One could also make a division between single particle and multiple particle localization, however as in the project considered image data includes multiple particles, methods for multiple particles localization or general methods that can be used for both single and multiple particles will be the mainly considered ones.

The difference between algorithmic and data driven approaches is that algorithmic methods are more specific as one has to e.g. choose the optimal parameters for the algorithm to perform well. Data driven approaches (e.g. deep learning) are more autonomous and don't require manual parameter input, but an as big as possible set of training data with known ground truth is required for good performance. [12]

### 2.1.2 Linking

The method of linking is usually applied on the localization results i.e. the localised particles centers over an image time series (movie). The end result of the method is estimated trajectories of the particles over the time interval. This can be done in several ways. The approaches can be divided into deterministic and probabilistic methods. The deterministic approaches are variants of a nearest neighbor

association i.e. the particles are assumed to move within their neighborhood, while probabilistic approaches are based on hypothesis association, meaning that the trajectories are estimated making one or several hypothesis of where a particle could move and choose the one with the highest probability in the latter case of several hypothesis.

## 2.2 Correlation analysis

Correlation analysis in the context of this application (fluorescence image data) is a form of a colocalization analysis. Colocalization analysis is typically performed for investigation of how well the localised particles in e.g. two images correlate, it is thereby a spatial analysis. In order to analyse co-movement, the temporal aspect needs to be included, and this type of analysis is often called dynamical colocalization analysis.

In dynamical colocalization, one investigates the positioning overlap of particles of interest (endosomes and A$\beta$ in our case) over time. This has been done in several ways and a division of three classes of methods can be made: (1) pixel-based dynamical colocalization, (2) object-based dynamical colocalization, (3) particle tracking and trajectory based correlation.[13]

In the first class one correlates the locations of the identified pixels as particle pixels between each pair of images (in the present case this corresponds to pairs of endosome and A$\beta$ image pixels). The second class is similar to the first one, however not all pixels are compared between each other but all of the found objects, which in this context are the identified particle centers. Class 3 first utilises particle tracking on the time serie images of the particles and then analyse the correlation between particles on the resulting trajectories.

The resulting correlation in all three classes can be quantitatively calculated in different ways; a common measure is Pearson's correlation coefficient. It gives the amount of correlation between two time evolved variables [14] (trajectories in this context).

## 2.3 Previous work

The relationship between endosomes and A$\beta$ has been previously investigated biologically and have shown that A$\beta$ are internalised in the cell using endocytosis.[15] The correlation between the rest of the endosomal pathway and A$\beta$ has however not been investigated thoroughly, as mentioned previously. Above mentioned methods will be utilized to make the investigation in this project.

Particle tracking and trajectory based colocalization analysis have previously been used in similar applications where one is interested in the relationship between biological particles. This method has shown best performance among the mentioned

dynamical colocalization methods. [13] The colocalization analysis applied on the trajectories has been done in different ways. In [16] the obtained trajectory of a particle of interest has been locally colocalised with the localization of the other particle of interest. This method however does not prove that the particles are actually moving together and not just crossing each other by a coincidence at that time point. [17] A different method instead seeks to correlate trajectories spatially [13], where a relation between particles is established if the correlation coefficient of their trajectories is above a certain value. This is however a global measure, which does not take into account the possibility of shorter interaction of the trajectories. The method described in [17] includes both spatial and local correlation of trajectories, by calculating the correlation on smaller windows of the spatially paired trajectories. A similar approach to this will be applied on the particles of interest (A$\beta$ and endosomes) in this thesis work.

# 3
# Methods

## 3.1 Overview

An overview over the whole project can be seen below in Figure 3.1. It shows in which order the different parts of the project were done and what was needed to be done for each part. The upcoming subsections consist of more detailed descriptions of how the different parts were accomplished.

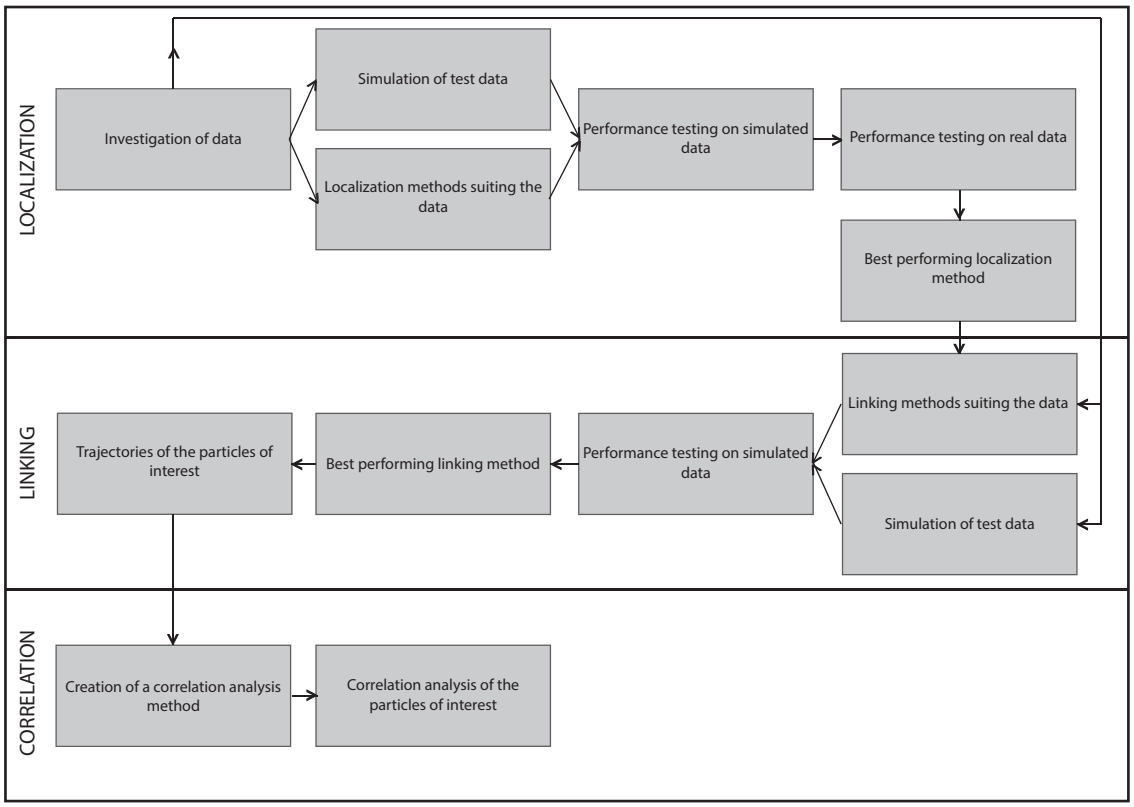



**Figure 3.1:** Overview of the project.

## 3.2 Data acquisition

Two types of data have been obtained: simulated data was created as part of the project and experimental data from live cell imaging was collected at the department

where the project was done. Data in the context of this project means images and movies of A$\beta$ and endosomal dynamics.

### 3.2.1 Experimental data

The experimental data was collected through a collaboration with a PhD student. The data was acquired in parallel with the particle tracking method to obtain the best suited data to draw conclusions about the particles of interest (A$\beta$ and endosomes). The upcoming description (Section 3.2.1.1) of how the data was collected is written by the PhD student who performed the method.

#### 3.2.1.1 Sample preparation and fluorescence imaging

The lyophilized A$\beta$ peptide powders were dissolved in hexafluoro-2-propanol to disrupt any aggregates and monomerize the peptide. The solutions were vortexed briefly and aliquoted at 4°C. The solvent in each aliquot was evaporated at 37°C for 60min using a RVC 2–18 CD Rotational Vacuum concentrator (Martin Christ, Germany). The remaining peptide films were snap frozen in liquid nitrogen and kept at 80°C until further use. For concentration determinations the peptide film was dissolved in 1% ammonium hydroxide (v/v) and the absorption of the dye label was measured on a Cary 4000 UV-Vis Spectrophotometer (Agilent Technologies, Santa Clara, CA, US). An extinction coefficient of 70,000 $M^{-1}cm^{-1}$ at 504nm was used for the HF488 dye label, according to the information provided by the manufacturer. Prior to each experiment one peptide film was dissolved in a small volume 1% ammonium hydroxide (v/v) and diluted with cell culture medium supplemented with 2% B-27 and 30 mM HEPES.

SH-SY5Y cells were grown in a 1:1 mixture of minimal essential medium (MEM) and nutrient mixture F-12 Ham supplemented with 10% heat-inactivated fetal bovine serum, 1% MEM non-essential amino acids and 2 mM L-glutamine. The cells were detached (trypsin-EDTA 0.05%, 5 minutes) and passaged twice a week.

Endolysosomal vesicles were labelled either by pre-incubation with fluorescently labelled dextran or by transfection with plasmids coding for fluorescently labelled vesicle specific proteins. For pre-labelling with dextran, cells were seeded in glass-bottomed culture dishes (MatTek; 25,000 cells/14 mm dish). 24h post seeding, the cells were washed 1x with serum-free medium followed by 4h incubation with 0.5mg/ml AlexaFluor647-labelled dextran 10kDa in complex medium, whereupon the cells were washed 2x in serum-free medium and incubated in complex medium for 20h to allow for the dextran to be transported to (and thus also label) lysosomes. Prior to A$\beta$(1-42) exposure the cells were washed 1x with serum-free medium. The cells were then incubated with 5µM HiLyteFluor488-labelled A$\beta$(1-42) in serum-free medium supplemented with 2% B-27 and 30mM HEPES for 30min followed by 2x wash in serum-free medium. Serum-free medium supplemented with 2% B-27 was added and the cells were imaged immediately. For plasmid transfection, SH-SY5Y cells were passaged two days prior to transfection, grown to 70% confluency and transfected with plasmids encoding for mRFP-Rab5 (Addgene #14437),

dsRed-Rab11 (Addgene #12679), mRFP-Rab7 (Addgene #1443600 or Lamp1-RFP) (Addgene #1817) by electroporation using a Neon Transfection System (Invitrogen, Carlsbad, CA, US), following the protocol provided by the manufacturer and applying a single pulse of 1,100 V with a pulse width of 50 ms. The cells were transfected using 1µg plasmid DNA/100,000 cells in a 10µl Neon Tip and plated immediately after in glass-bottomed culture dishes (MatTek; 50,000 cells/14 mm dish). After 48h the cells were exposed to A$\beta$(1-42) as outlined above, with an exposure time of 30min unless otherwise stated. Confocal images were acquired on an inverted Nikon C2+ confocal microscope equipped with two detector units (a C2-DUVB GaAsP Detector Unit with variable emission bandpass and a second GaAsP PMT), using an oil-immersion 60×1.4 Nikon APO objective and the 488nm, 561nm and 640nm laser lines (Nikon Instruments, Amsterdam, Netherlands). The sample was excited and detected with appropriate excitation laser lines and emission filters, and both fluorophores were excited and detected simultaneously.

#### 3.2.1.2 The collected images

The fluorescence process gives colored (multi-colored if several dyes used on several particles/molecules) image (series) of the wanted particles/molecules. An example of an obtained image can be seen in Figure 3.2, showing the cell and the particles of interest. The images were processed for the upcoming particle tracking by separation of A$\beta$ and endosome channels from the cell background channel and conversion to gray-scale. The processed versions of the image in Figure 3.2 can be seen in Figure 3.3. Image processing was done using ImageJ.

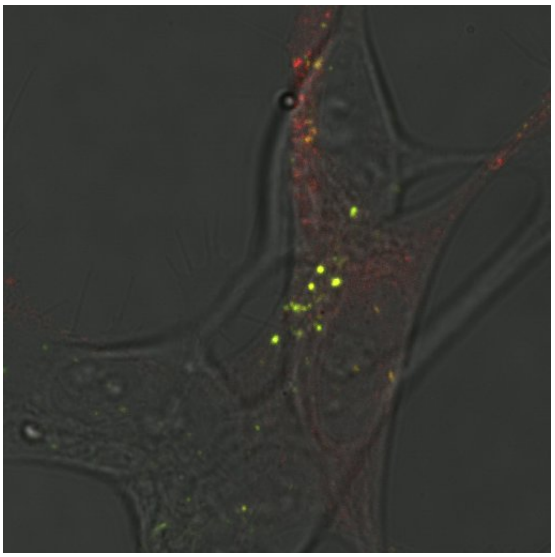
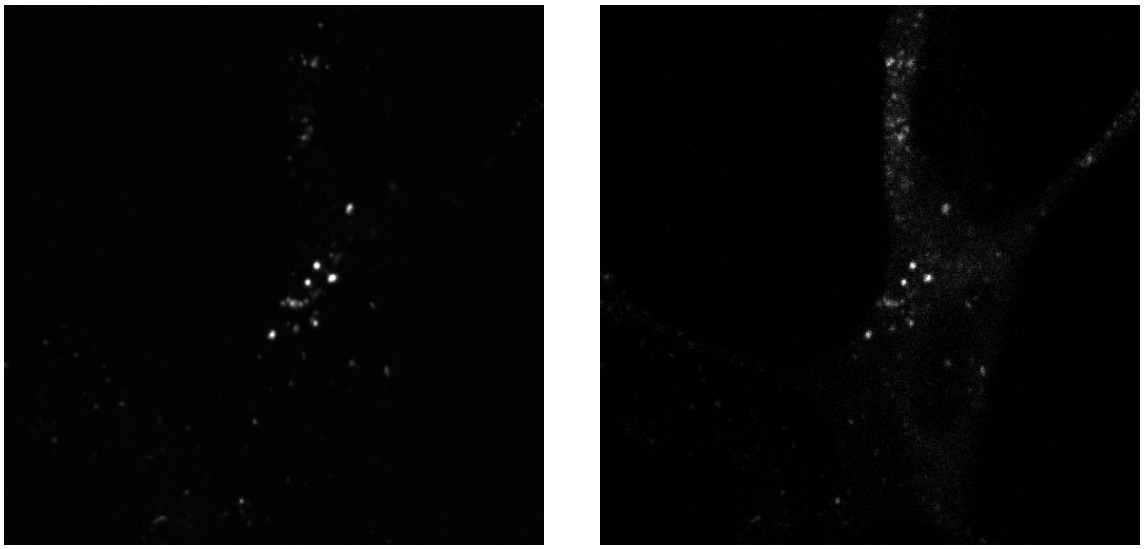


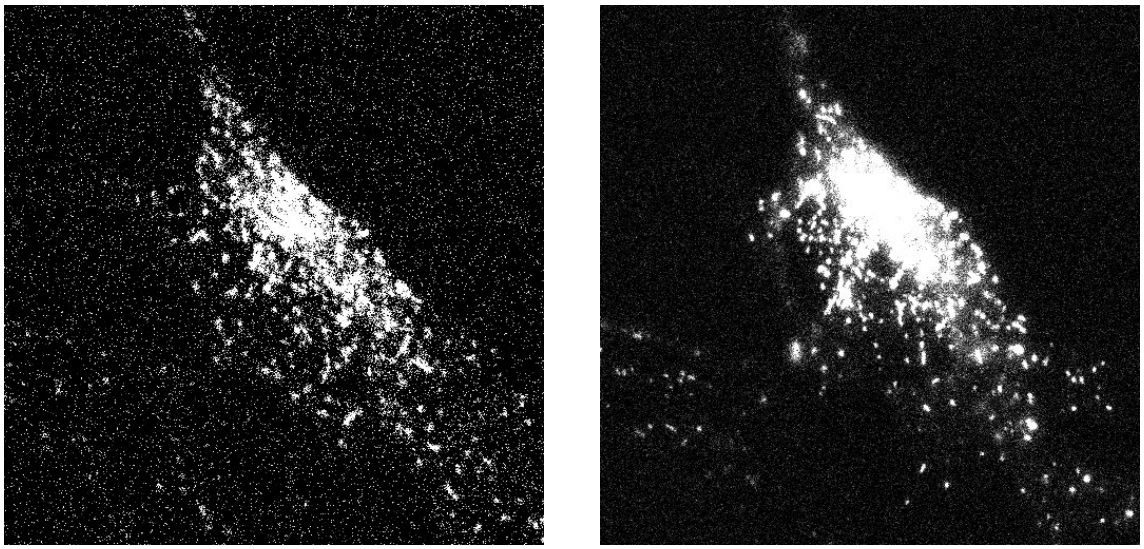
**Figure 3.2:** Example of experimentally obtained image with the particles of interest. The green color visualises the A$\beta$ and the red color the early endosomes (Rab5 labeled).



**(a)** **(b)**

**Figure 3.3:** Processed version of the Figure 3.2 image. a) shows the A$\beta$ channel (green) and b) the endosome channel (red).

In order to determine the optimal experimental setup for this study, fluorescence imaging with the confocal microscope was e.g. done with pinhole 1.2AU (confocally) and fully open. Two examples of processed images obtained with the two settings can be seen in Figure 3.4. Imaging with the open pinhole was tested to increase the amount of signal, however this method also resulted in out of focus-light, i.e. the imaging is no longer confocal.



**(a)** **(b)**

**Figure 3.4:** Example of obtained processed images (of lysosomes labeled with Lamp1) with two different experimental set-ups: a) is obtained with pinhole 1.2AU (confocally) and b) is obtained with a fully open pinhole.

### 3.2.2 Simulated data

The simulated data was created to imitate the experimental data. This was done for the purpose of measuring the performance of the different tracking methods and choose the best option for this project. This is recommended to do in such comparisons because of the unknown ground truth in biological images. [11]

To obtain the simulated data set, a couple of experimental data sets were first acquired and investigated in order to obtain their specific parameters (particle size, distribution, noise and contrast). This was then accounted for in the simulated data. It should however be noted that the experimental data was of differing appearance depending on the specific fluorescence method used and varied between samples. Because of this fact, only the main features of the experimental data were included in the simulated data, meaning that the simulated data does not mimic all samples completely, but instead their average appearance including varying particle size and distribution, amount, different levels of noise and presumed motion. Method testing on this type of simulated data will show which methods that perform best on the overall appearance of the data, but may not give optimal parameter settings for all data because of its varied appearance.

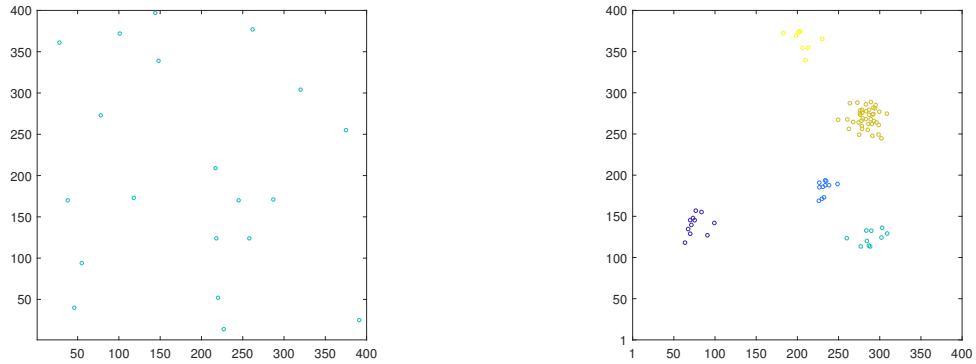
The first type of simulated data was created as single images of particles, aimed to test the performance of the localization methods. The second type was movies, meaning a series of images where the motion of the particles was included.

Simulation of the single particle images first included creation of the fluorescent particles (spots). For this purpose the method and function created by Raghuveer Parthasarathy and used in performance testing of radial symmetry centers method [18] was utilized. The particles were chosen to be of three different sizes to have a variety of sizes (3, 5 and 7 pixels in radius) as the particle sizes (spots) in experimental data vary, as one can see in the different example images above in Figure 3.3 and Figure 3.4. The size of the imaged endosome particles differ because endosomes naturally have a variety of dimensions, and each seen A$\beta$ particle contains many A$\beta$ peptides that have been taken up by an endosome, meaning that those particles will also be shaped as endosomes and have a variety in size.

The positioning of the particles was done both in a random manner and random clusters, i.e. particles were both placed randomly in the image and in random size clusters placed at random positions. See Figure 3.5 for examples of the random positions and the random clusters. The cluster generating function written by Nuno Fachada [19] was utilized. After the particles were placed in the images, noise was added to the image for testing of the localization methods' performance on different signal to noise ratios (SNR). The chosen SNR values were 4, 10, 15 and 20. The value of 4 because it is known to be the critical lowest level of what localization methods can handle [11], and the higher values because the collected data showed a varying SNR between approximately 10 and 15.

In order to simulate data for linking of the particles, the knowledge of statistics

of particle motion is required. From observation of the experimental data, directed random (Brownian) motion (i.e. a motion unfolding in one direction in the long run, but randomly in the short run, an example is shown in Figure 3.6) is hypothesized, because of the particles seeming to move randomly in short time intervals but directed in long intervals in the obtained experimental data.




**(a)** **(b)**

**Figure 3.5:** Examples of a) randomly generated positions and b) randomly generated cluster positions for particles.

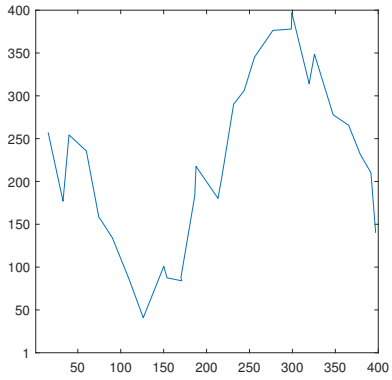



**Figure 3.6:** Example of how simulated directed random motion of particles look like.

For this simulation, 3 particle sizes were used as previously. The particles begin at random positions, both randomly and in random clusters. The particle positions were then changed over frames/images following a directed Brownian motion. Noise was not added to this data as it only affects the localization of particles and not how well the localised particles are linked to trajectories in the linking performance testing.

An additional added feature to this data type is birth and death of the particles during the image series, meaning that the particles have a certain probability to

be born and to die during the time interval. This was indeed observed in the experiments, since we imaged in 2 dimensions while the cells, even if very thin, are 3 dimensional, meaning that the particles could move in/out of the imaged frame. The fluorescent signal also decreases over time due to photobleaching increasing the possibility that some particles could disappear if they get too bleached. The added feature does not however allow for the particles to reappear being gone one or more time steps and continue on their trajectory as they can do in reality, only that new particles can appear and existing ones can die. The chosen values for the birth and death parameters were such that during the image series of 30 images, the average life length of a particle is 20 images. The average life length was chosen in this way to see how well the methods perform on longer trajectories of several particles which induces crossing of trajectories, which should be the most difficult case to do the linking in.

## 3.3 Localization and linking methods

The choice of different localization and linking methods to be tested in this project was made by looking at the factors affecting the choice. These factors were: the appearance of the experimental data and the availability of the methods. The availability of localization methods was better than for linking. There are a lot of available methods and combinations of methods to locate particles and some of the main ones were chosen to be tested on our application. The methods were either available online or were obtained from groups at the university at which this project was done. Availability of linking methods was not as good as for localization methods, especially in our application where the experimental data contains many particles with unknown ground truth number and also disappearance of particles. This led to the tested methods being only the ones obtained from groups at the university.

## 3.4 Performance measures

The performance of the chosen localization and linking methods were tested on simulated data. The best performing localization method was also tested on experimental data for reassurance of the methods performance. This test was done including different types of experimental data and was done by observation by eye by 4 people. It should be noted that we do not know how many particles we should see and can only make an approximation when examining experimental data. The investigation by eye was not possible to do in the same way when evaluating linking as one can not know which particles that move where exactly. In some cases one can see where the more distinct particles move, but not the particles in particles dense areas with similar appearance of the particles. One can also place the obtained particle trajectories on the experimental movies and roughly compare the trajectories, but it is not possible to get a measure as it is in localization.

The upcoming methods used to investigate the performance of localization and linking methods on simulated image data were inspired by methods used for similar

type of investigation of tracking methods [11] and common statistical performance measures.

### 3.4.1 Localization performance measures

In order to measure the performance of different localization methods, the following factors were considered. First, the precision of the estimated particle centers can be evaluated, as it has been done in [12, 18]. In our application however, images contain many particles, and even though the precision of particle center is important especially for the separation of overlapping particles, identification of the location and number of all real particles is of bigger importance, to be able to (as closely as possible) track the overall endosome movement within each cell.

To measure the amount of correctly classified particles, a limit was set for how far away the estimated particle center could be from the real particle center for all particles. This means that instead of measuring particle center precision one measures if the particle center is located within an allowable radius around the real particle center, as it also has been done in [11]. The limit was set to less than 2 pixels as the width of the smallest simulated particle is 3 pixels, meaning that if the center would be classified 2 pixels from the real center point it would actually be outside the particle, where another particle could be placed, making the classification incorrect.

However, only counting the number of correctly classified particles is not enough as one should also consider how many particles were classified incorrectly. Sensitivity and specificity analysis was applied with the purpose of taking both factors into consideration, as described below.

Sensitivity measures the fraction of correctly classified particles (true positives, TP) out of all particles that should be classified as particles, meaning particles classified as particles and also particles that were not classified as particles (false negatives, FN). In our case a correctly classified particle is a particle center estimated to be located within the chosen radius of the simulated particle center. The formula is shown below.

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.1}$$

Specificity gives the fraction of correct classification of non-particles (background and non-center particle pixels) not as particles (true negatives, TN) out of all that should be classified as non-particles, which is all that was classified as non-particles correctly plus the particles that were classified as particles but are not (false positives, FP), see formula below.

$$Specificity = \frac{TN}{TN + FP} \tag{3.2}$$

False discovery rate was also calculated as an aid for the choice of best performing method for the application. It is calculated by the formula below.

$$False\ discovery\ rate = \frac{FP}{FP + TP} \tag{3.3}$$

It gives the relation of falsely classified non-particles as particles to all that were classified as particles.

### 3.4.2 Linking performance measures

Measuring performance of the linking methods is more demanding comparing to the methods of localization as linking involves both space and time aspects. Output from a linking method results in a set of particle trajectories over time. And these are to be compared with the simulated trajectories. In order to do this, one first needs to find matching pairs of trajectories. This is done by calculating a cost for each of the possible trajectory pairs, and the obtained cost matrix is then minimized resulting in best matching trajectory pairs.

Cost calculation and minimization are determined similarly as in [11]. For each pair of simulated and estimated trajectories and for each time point for the pair, the first check is whether both trajectories contain a trajectory point at the current time point. If both do not, the cost is set to zero. If only one of them contains a trajectory point a penalty value is added to the cost. And if both contain a trajectory point, the Euclidean distance for the pair is calculated and then the cost is set to the smallest value of either the penalty or the calculated distance. This is done for the reason that for points that are too far away from each other it does not matter how far away they actually are and a penalty is set instead. The chosen penalty value is set to the same value as used in localization, i.e. 2 pixels, as points farther away cannot be considered as matching. The cost for each pair over the whole time interval is then summed up to obtain the final cost for each pair of trajectories. The obtained cost matrix is then minimized using Munkres algorithm.

The obtained matching pairs are then evaluated using a similar approach as for localization. For each optimal pair of trajectories, each pair of positions over the time interval are set to be true positive if the distance between them is less than 2 pixels, else it is set to a false positive. And if only the real or estimated location exists for some time point, the pair is set to a false positive.

If the number of estimated trajectories is smaller/bigger than the simulated trajectory number, Munkres algorithm will give no match for some of the trajectories. Each existing trajectory location during the time interval of the not match simulated trajectories is set to a false negative, as these particle locations should be matched in the ideal case if the evaluated method would have found all of the wanted trajectories.

Lastly, sensitivity and precision are calculated from the above obtained measures.

Sensitivity in this context is a measure of the fraction of correctly estimated trajectory positions out of the sum of correctly estimated trajectory positions and all trajectory positions of the simulated trajectories that should have been matched with an estimated trajectory, but were not. Precision is calculated as

$$Precision = \frac{TP}{TP + FP}. \tag{3.4}$$

It shows the fraction of correctly estimated trajectory positions out of all trajectory positions of the matched pairs of trajectories.

## 3.5 Correlation methods

Two correlation methods were developed and applied for this thesis work. A trajectory-based method and an object-based method (to draw conclusion of the trajectory-based method's performance).

### 3.5.1 Trajectory-based method

As mentioned in Section 2.3, a modified version of the method including spatial and local correlation of trajectories was applied to our application. The modification was that instead of calculating the correlation coefficient over an iterating window over the paired trajectories and then comparing the obtained coefficient values to a threshold in order to decide on interaction between positions of particles over time, we chose a spatial threshold for two positions of trajectories to be interacting and calculated the number of interacting positions of particles over time for all paired trajectories. This means if two positions of a pair of matching trajectories of A$\beta$ and endosome are within the chosen radius (spatial threshold) then we count them as matching (see Figure 3.7). This analysis results in a fraction of matching particles (positions) of A$\beta$ and endosomes over time. The fraction was calculated by

$$\frac{TP}{TP + FP + FN}, \tag{3.5}$$

where TP are all the matching positions, FP are all non-matching positions of A$\beta$ trajectories and FN are all positions of A$\beta$ of the A$\beta$ trajectories that were not matched with an endosome trajectory. $TP + FP + FN$ corresponds to the total number of A$\beta$ positions on the A$\beta$ trajectories (meaning that only the moving A$\beta$ particles are included). In this way the fraction is calculated with respect to A$\beta$ particle trajectories.

### 3.5.2 Object-based method

The object-based correlation in our case measures the correlation between particle centers of endosome and A$\beta$ images. The correlation was measured in the form of fraction of matching endosome and A$\beta$ particles, as for the trajectory-based method to make a fair comparison. The fraction was calculated by taking all the matching A$\beta$ and endosome particles (centers) divided by the total number of A$\beta$ particles

(centers). This was done for all endosome and A$\beta$ image pairs in the image time series.

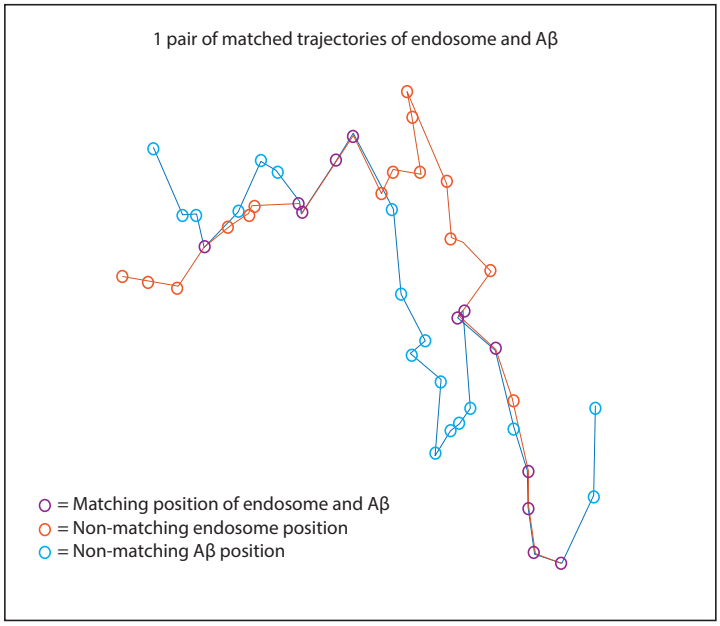



**Figure 3.7:** An illustration of two possible paired trajectories of endosome and A$\beta$ and their matching positions.

### 3.5.3 Experimental data

The experimental data was, as mentioned, collected in different ways resulting in e.g. 30 time frames every 10 minutes or every hour. To get a development of correlating particles over time in sense of every 10 minutes or every hour, the obtained development curves of correlating particles over the 30 frames from both correlation methods were averaged to obtain one value for each 10 min/1 hour slot. The data was also collected by placing the imaging frames on different cells (positions), meaning that each 10 min/1 hour data set (the 30 time frames) was collected on several positions. To again obtain a development of correlating particles over time, the averaged correlating particle values were averaged over the different positions as well.

# 4
# Results

## 4.1 Image simulation

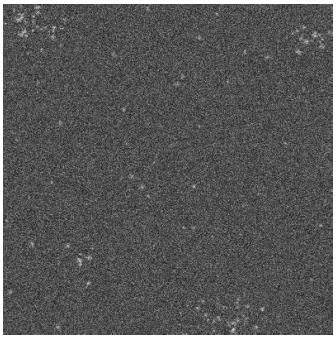
Two types of image sets were simulated for performance testing of tracking methods. One type for testing localization methods and one for testing linking methods. The results from the image simulations are presented in the following subsections.

### 4.1.1 Localization

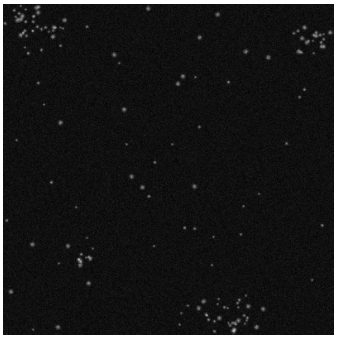
Simulated images with different SNR were created to test the performance of different localization methods. An example image from each SNR image set can be seen in Figure 4.1. In the images with a lower SNR value it becomes more difficult to separate the particles from the background and determine their size, especially in the image where SNR=4. Comparing the experimental data with the simulated localization data one can see that the overall signal to noise ratio in experimental data is not as bad as in the simulated data with the lowest signal to noise ratio (SNR=4). However, in some areas of the experimental images, as the particle dense areas, the SNR could be compared with the level of 4, while other more clear areas could be compared with a SNR level of 15-20.

### 4.1.2 Linking

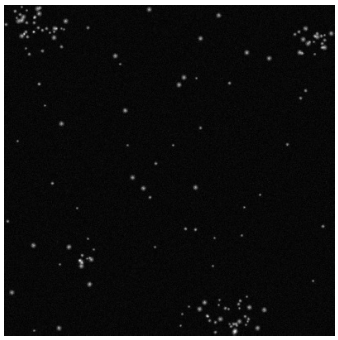
Sets of particle images were created for linking, mimicking the hypothesized directed Brownian motion of the particles over time. An example of the obtained particle trajectories can be seen in Figure 4.2. One can see that the trajectories are overlapping, in order to create challenging linking situations as will be the case for endosome trajectories in live cells. The simulated trajectories are longer than in reality, but it is good for the method to be able to handle cases where some particles could have longer overlapping trajectories as well.



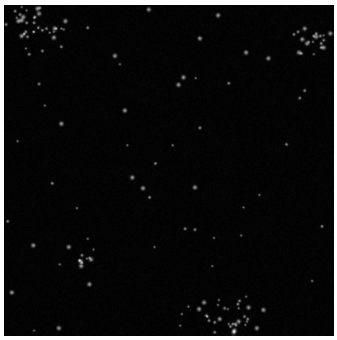
**(a)** SNR=4



**(b)** SNR=10



**(c)** SNR=15



**(d)** SNR=20

**Figure 4.1:** Examples of the simulated localization images with the different signal to noise ratios.

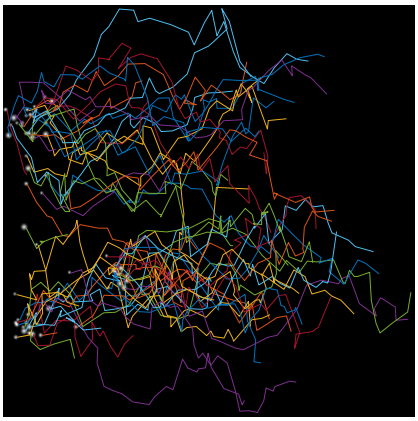


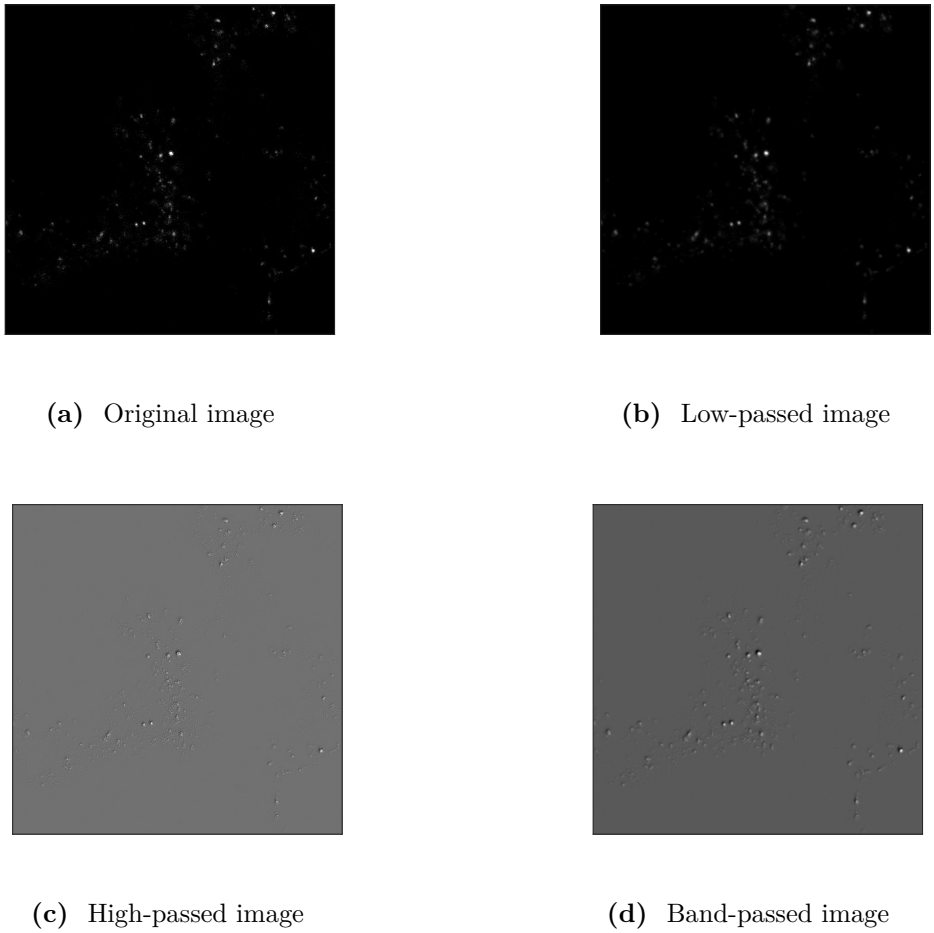
**Figure 4.2:** An example of the simulated particle trajectories.

## 4.2 Particle localization

Particle localization can be considered to consist of 3 steps: (1) filtering, (2) thresholding and (3) finding centers of particles. Different methods for all steps need to be tested to evaluate and identify the most suited combination of methods for this project. The first tests were done on simulated data (see Section 3.2.2 and Section 4.1.1) and combinations of methods for the three steps of localization were tested as it is described below.

Three filtering methods were chosen to be tested: (1) no filter, (2) a band-pass filter including Boxcar and Gaussian filters and (3) a band-pass filter including Gaussian filters. A band-pass filter is both a low-pass and high-pass filter. By applying a low-pass filter you suppress the highest frequencies in the image (noise), and with application of a high-pass filter you suppress the lowest frequencies leading to sharpening of edges in the image. [20, 21] The low-pass filtered image can be obtained by convolution with an appropriate function e.g. Gaussian or Boxcar function, cutting away the higher frequencies in the image. By subtraction of the low-pass filtered image from the original image, you obtain a high-pass filtered image (as you remove the lower frequencies from the image). Lastly, a band-pass filtered image is gained by convolution of the high-pass filtered image with some appropriate function (as mentioned above) to add a low-pass filter also. An example of an image with applied low-, high, and band-pass filters is shown in Figure 4.3, the images are obtained using Gaussian filters as in method (3).

Thereafter, two types of threshold methods were compared: a local threshold which sets a specific threshold for each pixel of the image and an iterative global threshold, meaning that one threshold is set for the entire image. The iterative threshold is however also dynamical by adaptation to each image, even though it does not vary throughout the image. The local threshold is calculated by setting a threshold over windows of selectable size around each pixel. A window of size 9x9 pixels was chosen as if the window would include a particle of the biggest simulated size (7 pixels in radius) it would also include some background around the particle. Deciding a particle threshold over a smaller window could include only the particle and would not give a correct threshold for finding particles within the background, and bigger windows would reduce the local-feature of the method. The iterative global threshold method was chosen among other global methods as it showed best performance in a smaller test on experimental and simulated data by observation by eye. The method was found online in a question thread [22]. The local threshold method is called Sauvola and was written by Jan Motl [23]; this method was the only tested local threshold method in the project. Both threshold methods were performance tested by application on either the filtered image from the above step or on the original image, in order to obtain an optimal approach.



**(a)** Original image

**(b)** Low-passed image

**(c)** High-passed image

**(d)** Band-passed image

**Figure 4.3:** A visualisation of low-, high- and band-pass filters applied on an image (obtained using Gaussian filters).

Lastly, two methods were tested to find the center points of particles out of all pixels identified as (parts of) particles. One needs to find particle centers to enable count of particles, as each identified particle-pixel is not a particle on its own (a particle consist of several pixels), and to separate between particles in particle dense areas. The chosen methods were: a centroid method and a radial center method. The centroid algorithm works by firstly finding local maximas over the entire image as a first estimation, then calculating the centroids of the maximum points. Radial center method calculates the center point of particles by calculating the maximal radial symmetry point over windows of desired size iterating over the whole image.

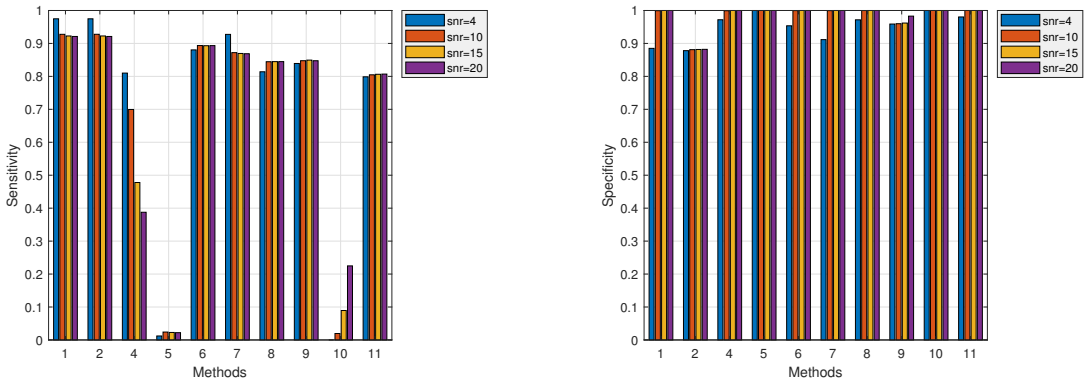
The aim in the beginning of the project was also to apply and investigate the performance of a promising machine learning method (deep learning) for localization. This method has shown to outperform the other tested methods in single particle localization with respect to center point of the particle. [12] This was however not managed because of lack of time.

Performance of the combinations of methods for the three steps of localization is

summarized in Table 4.1; the results are also shown as bar diagrams in Figure 4.4. The performance measures given in the table were calculated as described in Section 3.4.1. The combinations of methods that are not shown in the table resulted in no found particles for all SNR levels. The combinations were build by testing each filter method (no filter, Boxcar and Gaussian) with each threshold method (local and iterative methods applied on either the original image or the filtered image) with each particle center localization method (centroid and radial center).

**Table 4.1:** Particle localization on simulated images test results. Boxes with no sensitivity and specificity indicate that the following combination of methods yielded no found particles. The threshold methods were applied to either the original image (noted as "threshold" in the table) or on the filtered image (noted as "threshold*" in the table).

| Nr. | Method | SNR=4 | SNR=10 | SNR=15 | SNR=20 |
|---|---|---|---|---|---|
| 1. | No filter, iterative threshold, centroid | Sensitivity: 0.9748<br>Specificity: 0.8851 | 0.9277<br>0.9994 | 0.9227<br>1.000 | 0.9212<br>1.000 |
| 2. | No filter, local threshold, centroid | Sensitivity: 0.9748<br>Specificity: 0.8781 | 0.9278<br>0.8811 | 0.9227<br>0.8817 | 0.9212<br>0.8821 |
| 3. | Boxcar, local threshold*, centroid | Sensitivity: 0.0184<br>Specificity: 1.000 | -<br>- | -<br>- | -<br>- |
| 4. | Boxcar, iterative threshold, centroid | Sensitivity: 0.8100<br>Specificity: 0.9719 | 0.6996<br>0.9993 | 0.4779<br>1.000 | 0.3876<br>1.000 |
| 5. | Boxcar, iterative threshold*, centroid | Sensitivity: 0.0121<br>Specificity: 1.000 | 0.0240<br>1 | 0.0228<br>1 | 0.0221<br>1 |
| 6. | Gaussian, iterative threshold*, centroid | Sensitivity: 0.8804<br>Specificity: 0.9536 | 0.8936<br>0.9997 | 0.8931<br>1 | 0.8934<br>1 |
| 7. | No filter, iterative threshold, radial center | Sensitivity: 0.9275<br>Specificity: 0.9116 | 0.8721<br>0.9996 | 0.8696<br>1 | 0.8688<br>1 |
| 8. | Gaussian, iterative threshold, radial center | Sensitivity: 0.8140<br>Specificity: 0.9717 | 0.8445<br>0.9999 | 0.8448<br>1 | 0.8448<br>1 |
| 9. | Gaussian, iterative threshold*, radial center | Sensitivity: 0.8392<br>Specificity: 0.9589 | 0.8476<br>0.9599 | 0.8494<br>0.9619 | 0.8475<br>0.9829 |
| 10. | Boxcar, iterative threshold, radial center | Sensitivity: 0<br>Specificity: 1 | 0.0196<br>1 | 0.0894<br>1 | 0.2248<br>1 |
| 11. | Boxcar, iterative threshold*, radial center | Sensitivity: 0.7987<br>Specificity: 0.9804 | 0.8048<br>1 | 0.8064<br>1 | 0.8070<br>1 |




**(a)** **(b)**

**Figure 4.4:** Another visualisation of the sensitivity and specificity values shown in Table 4.1. The bar diagrams exclude however method 3 performance as it only gave results for one SNR level.

Examining the results in the table/bar-diagrams, four combinations (1, 6, 8 and 11) of methods can be identified as performing better than the others with respect to both sensitivity and specificity. In our application we consider specificity value as more important than sensitivity. This is because in our images there are much more non-particle center pixels than particle center pixels, and specificity measures the fraction of correctly classified non-particle center pixels and sensitivity the fraction of correctly classified particle center pixels. This means that a lower value of specificity implies a much bigger error (more incorrectly classified pixels) than a lower value in sensitivity. This lead to the choice of the best methods to be more affected by a higher specificity value than sensitivity value for all SNR levels. The methods with values close to 1 in specificity for all SNR levels and sensitivity around 0.8 or higher are 1, 6, 7, 8 and 11. Methods 6, 8 and 11 have the highest specificity values. When deciding between methods 1 and 7, specificity does not differ by much but method 1 has higher sensitivity for all SNR levels which led to the choice of method 1.

As one can observe, the sensitivity values decrease in some cases on images with higher SNR, which is the opposite of what one might expect as the localization should get better on images with higher SNR as they are clearer as seen in Figure 4.1. The combinations of methods of decreasing SNR have in common that the threshold of the method is calculated on the original image. The explanation for these methods' decrease in sensitivity is found investigating what actually happens when the methods are applied to the images with different SNR. With lower SNR, i.e. more noise in the image, the distribution of the pixel-intensities of the image do not have a clear separation between particles and background in the image. This leads to that the threshold value is set higher than in the case of less noise, as the threshold methods take into account the mean and the spread of the intensities in different ways. With a higher set threshold, less particles are classified as particle pixels than in the higher-SNR case. Less particle pixels could lead to less errors when calculating the particle centers, as the method has less pixels to choose between; with more particle pixels (in the higher-SNR case) there are more possibilities for particle centers and more room for errors for the particle center methods. This could explain the decrease in sensitivity. One should note that this behaviour happens for all cases when the centroid method is used (both on images with and without filtering) with the threshold calculated on the original image. The behaviour is not seen for the radial center method (applied on the filtered image) with the threshold calculated on the original image. This could imply that the radial center is more robust than the centroid method to a lower set threshold (leading to more particle pixels to choose between when calculating center points), however only if radial center is applied on the filtered version of the image.

When deciding which of the methods is the optimal one for the application at hand, we narrowed the investigation to SNR levels 10 and 15 as the experimental data is in that range. The SNR value in experimental data was calculated as done in [11], namely by

$$\mathrm{SNR} = \frac{I_o - I_b}{\sqrt{I_o}}, \tag{4.1}$$

where $I_o$ is the highest particle pixel value and $I_b$ is the mean of background pixel values. The case of lower signal to noise ratio (SNR=10) was considered further, to be sure that our chosen method can handle noise. To understand what the obtained sensitivity and specificity values mean better we translated them into particle values with the knowledge of how many particles were simulated (150 particles meaning 150 particle centers) and how much non-particle center pixels we have (from the image size, minus the number of particle centers). Method 1 found on average 139 particles (centers) and 95 false particles (centers) (non-particle centers classified as particle centers). Method 6 found 134 particles and 48 false particles, method 8 gave 127 particles and 16 false particles and the last method 120 particles and zero false ones.

As the resulting particle (centers) of the chosen localization method are used in the linking step, it is desirable to have found as many of the correct simulated/real particles as possible, but it is also not a huge concern to have some false positive particles, since if they are false they will probably disappear in the next frame and will not be linked. However, if we count too many false particles it will affect the linking negatively since some particles will be linked to false particles. This means that the percentage of the found false particles with respect to all found particles should not be high. Calculation of this can be found in Table 4.2. Looking for a low percentage of false positives together with a higher count of true positive particles, method 8 was chosen for further work.

**Table 4.2:** Results of false discovery rate calculation of the four best performing methods.

| Nr. | True positives | False positives | False discovery rate |
|---|---|---|---|
| 1. | 139 | 95 | 0.4060 (40%) |
| 6. | 134 | 48 | 0.2637 (26%) |
| 8. | 127 | 16 | 0.1119 (11%) |
| 11. | 120 | 0 | 0 (0%) |

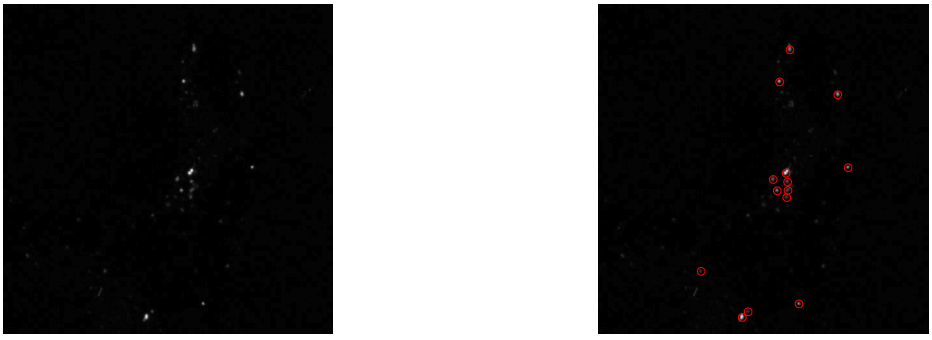
Next, method 8 was applied to experimental data to investigate its accuracy. Several additional factors affecting the procedure and its outcome needed to be considered here. First, as the exact ground truth is not known, one must to some extent make assumptions of which spots in the images that are particles and which are noise or auto-fluorescence. Some particles can have a more elliptical shape than circular, meaning that one does not know if a detected local intensity maximum corresponds to several spherical particles in close proximity or to one elliptical; the endosomes could have an elliptical shape also. Due to the uneven spatial distribution of endosomes within cells, some of the images can have larger dense particle regions in them making it difficult to discern the particles present there. It was also difficult to set a strict criteria on the brightness of the spots to be counted as particles (threshold), as what one would say to be a particle in a darker image could be set as a part of the background in an image containing much brighter spots. The manual by-eye threshold should be adapted to each image as done in the used threshold methods, however it is not certain whether the threshold is adapted as consistently

and equally by the examining participants as it is in the threshold methods. These factors introduce an uncertainty and bias in the outcome.

Four images were chosen from the collected data representing its variance Figure 4.5-Figure 4.8. These images were then evaluated by 4 people by manually counting the number of particles in each. The resulting counts of the particles in each image can be seen in Table 4.3. The table also contains the count results using method 8. However, one should note that when testing, the method's performance varied depending on what type of image it was and some parameter values needed to be adjusted to obtain more satisfying results both regarding the count and localization of the particles. The changed parameter values were the minimum distance between found particles and threshold modifications.

**(a)** **(b)**

**Figure 4.5:** Collected data sample 1 (showing A$\beta$) and localization results.

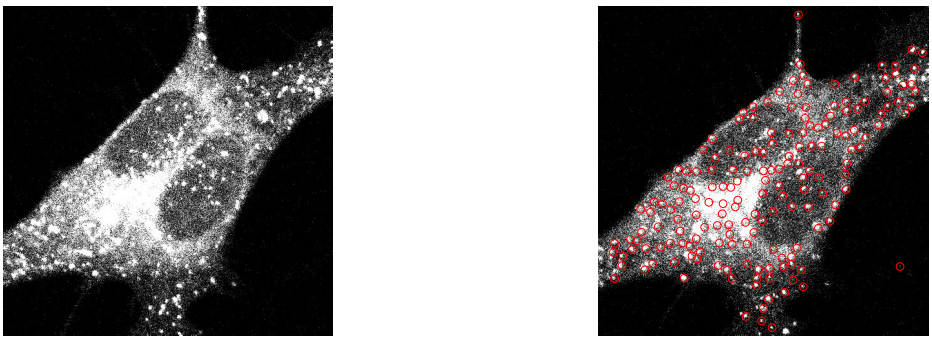


**(a)** **(b)**

**Figure 4.6:** Collected data sample 2 (showing A$\beta$) and localization results.

**(a)** **(b)**

**Figure 4.7:** Collected data sample 3 (showing endosomes labeled with Rab11) and localization results.



**(a)** **(b)**

**Figure 4.8:** Collected data sample 4 (showing endosomes labeled with Rab7) and localization results.

**Table 4.3:** Results of the evaluation of the best localization method's performance on experimental data.

| Image | Manual count | Count by method |
|---|---|---|
| 1. | 12, 19, 21, 21 | 18 |
| 2. | 10, 13, 15, 17 | 14 |
| 3. | 61, 64, 66, 78 | 73 |
| 4. | 155, 158, 180, 198 | 175 |

Particle localization results by method 8 are, in the tested cases, within the range of the manual count, as seen in Table 4.3 (the manual count values are ordered by size in the table). Observing the simpler images as Figure 4.5 and Figure 4.6, one can see that the localized particles indeed look like possible particles. For the more difficult cases in Figure 4.7 and Figure 4.8, one can see that the method has localized some false positives in the wider bright regions where one cannot know if there are any actual particles (it could also be that the method is actually better than the

human eye). A couple of spots that could be particles in the images are also missed. However, small fractions of missed particles and false positives were expected from the method, and as mentioned previously it should not effect the upcoming linking performance.

## 4.3 Particle linking

Two linking methods were applied and tested: one deterministic method and one probabilistic method. The deterministic method is based on minimal distance and temproal hiatus. The probabilistic method is a Poisson multi-Bernoulli mixture filter that is based on setting multiple hypothesis on where each particle could travel in each time and calculating probabilities of how likely each of the hypotheses is. [24] Both methods were acquired from research groups at the university at which the project was done.

Firstly, the deterministic method with varying parameters was tested using the performance measures described in Section 3.4.2. An example of the estimated trajectory evolution over time by the method applied on simulated data is shown in Figure 4.9. Further below in Figure 4.10 one can see a comparison between the simulated and estimated trajectories. When comparing, one can see that the estimated trajectories seem to match at least a part of the simulated ones. The estimated trajectories show also a similar shape even if not matching everything correctly.

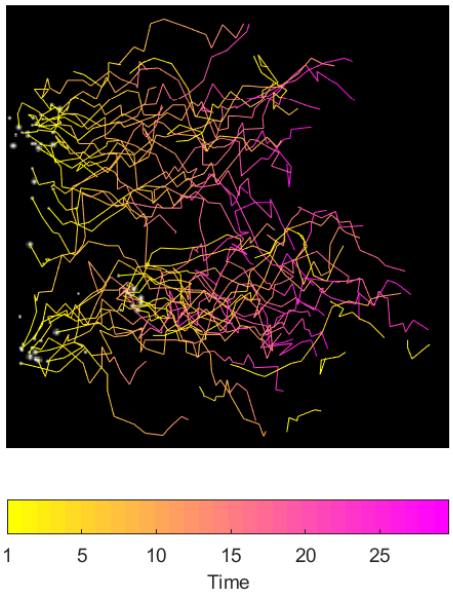
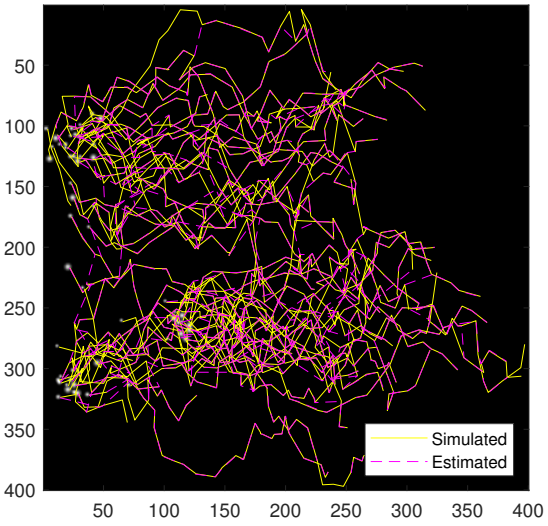
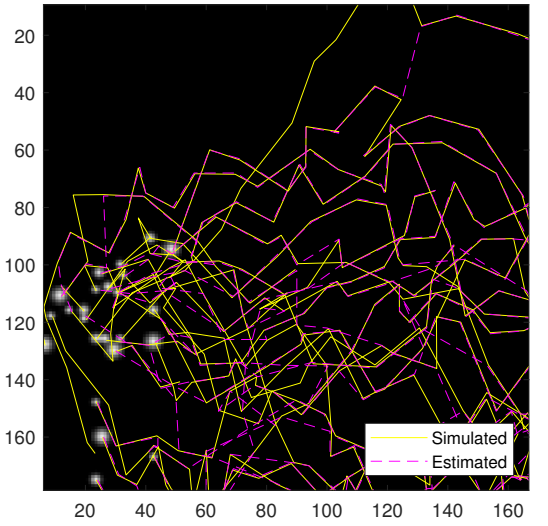



**Figure 4.9:** Example of the estimated trajectory development over time.




**(a)**




**(b)**

**Figure 4.10:** Simulated and estimated trajectory comparison.

The first parameter to be analysed was the flow direction parameter which adds an assumption of the particle flow direction when estimating the trajectories. It was tested to have no assumed flow direction and assumed flow direction of directed Brownian motion (which is consistent with the inward motion of endosomes into cells). Figure 4.11 shows the resulting sensitivity and precision over the y-axis. The x-axis includes a varying displacement threshold parameter in the method. A displacement threshold is a restriction of how far away a particle can move between each frame (time step). One can see that the sensitivity and precision values are almost the same in both cases regarding the flow, precision is slightly better with assumed flow direction. An assumed flow direction was chosen to be used further.

The varied displacement threshold did not seem affect the outcome by much, except at one point for the sensitivity.

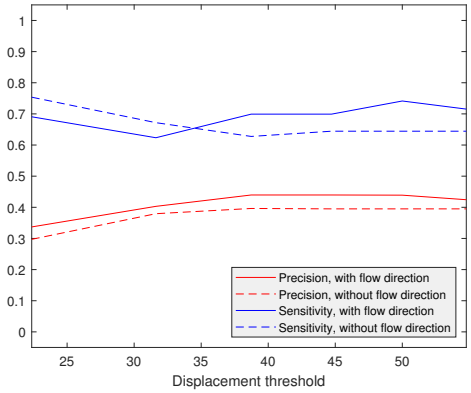
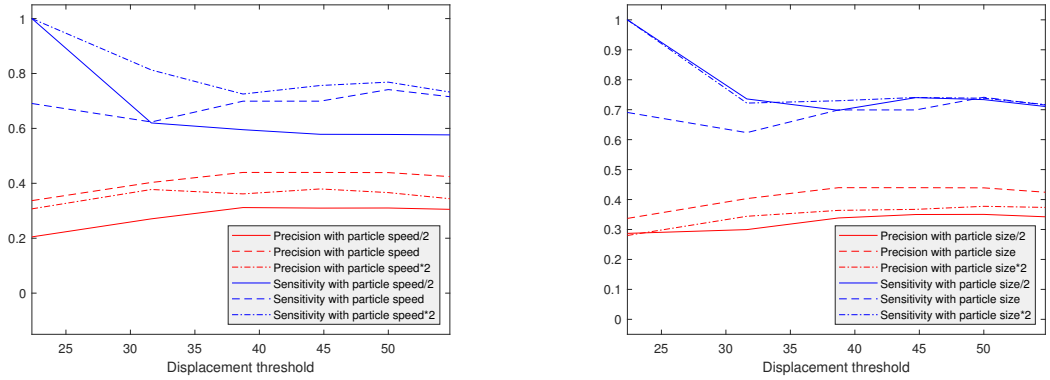



**Figure 4.11:** Linking performance varying the displacement threshold and assumption of flow direction.

While continuing varying the displacement threshold, it was also tested how the method performed when varying the simulated motion parameters, particle size and speed. 3 levels of particle size and speed were chosen for the simulated data and the results for them can be seen in Figure 4.12. Looking at the precision, the method does not show any specific trends for neither size or speed parameters, e.g. precision is lowest for the lowest speed but not highest for the highest speed. Sensitivity varied more over the displacement thresholds for the different size and speed parameter values. This could be explained by the fact that sensitivity is calculated by looking at true positive and false negative values. True positives are positions of the matched trajectories matching the simulated positions. False negatives are all positions from the simulated trajectories that did not have a match within the estimated trajectories. It means that these values depends also on how well the matching algorithm perform (Munkres algorithm), and not just how well the trajectories are estimated, which could induce further errors and affect the sensitivity values more than precision. Precision is easier to interpret as it compares how many true positives there are with respect to the false negatives on the matched trajectories only. The varied displacement threshold did not seem affect the outcome by much in the precision here either.

The last parameter to be tested was the one called *epsilon*. The value of *epsilon* decides how far away an estimated particle position is allowed to be from the simulated position for them to count as matching. As expected both the sensitivity and the precision increased with a larger *epsilon*-value, since there we allow a bigger error, as seen in Figure 4.13. The displacement threshold did not have a distinguishable effect here either. From this it was decided to choose a specific displacement threshold on experimental data directly.




**(a)** **(b)**

**Figure 4.12:** Linking performance varying the displacement threshold and the simulated data type.

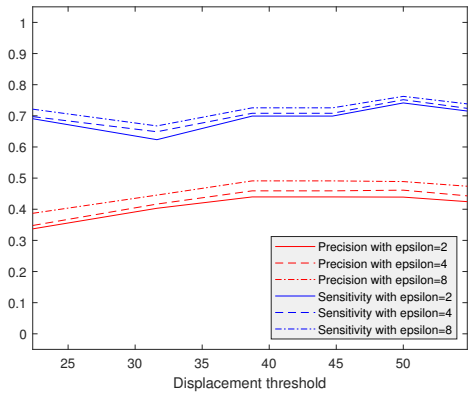
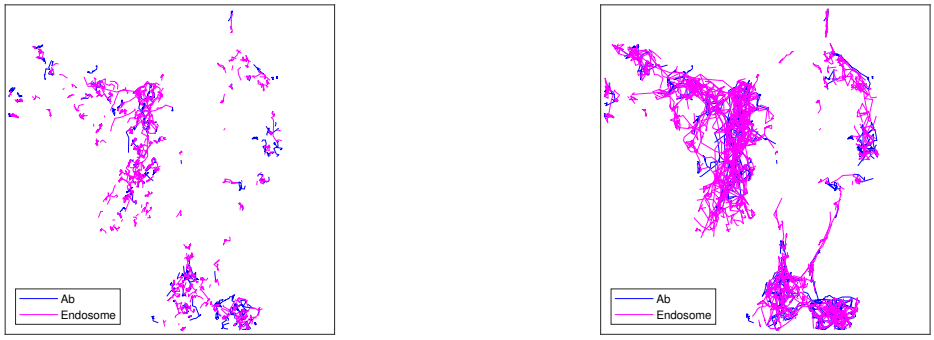


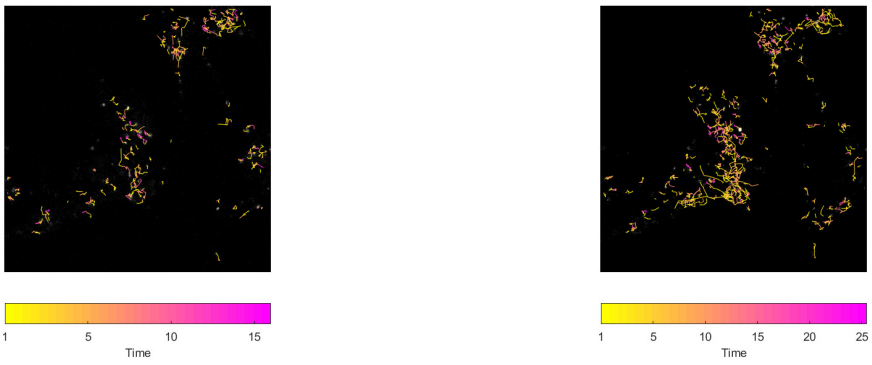

**Figure 4.13:** Linking performance varying the displacement threshold and how far away two positions can be to be counted as a match.

The method was lastly tested on experimental data. One cannot measure the performance here by more than observing and comparing between the estimated trajectories and how the particles move in the collected films. Figure 4.14 shows both the A$\beta$ and lysosome trajectories in the same image, where one can observe their estimated trajectories (to be analysed in the upcoming correlation section). By comparing these images with the experimental data image series it was decided to continue with displacement threshold of 10 pixels as it seemed to mimic the experimental data better. Figure 4.15 shows the same results as in Figure 4.14 a), however with A$\beta$ and lysosome trajectories separately and also showing their evolution over time.




**(a)** **(b)**

**Figure 4.14:** Example of the estimated trajectories of both A$\beta$ and lysosomes (labeled with dextran), using a displacement threshold of a) 10 pixels and b) 22 pixels.



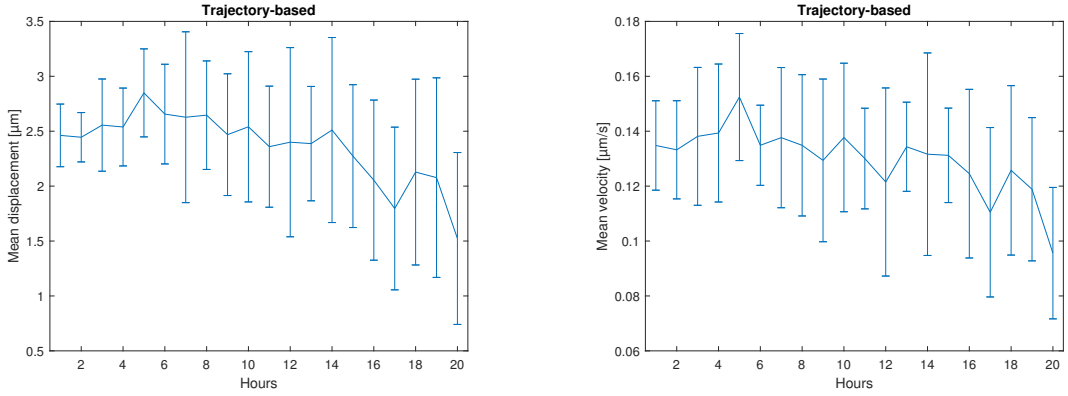

**(a)** **(b)**

**Figure 4.15:** Example of the obtained time evolution of the estimated trajectories of A$\beta$ and lysosomes (labeled with dextran) separately, using a displacement threshold of 10 pixels.

Secondly, the probabilistic method was tested. This was, however, a much more difficult task than with the deterministic method, as the probabilistic method has a more complicated nature with more parameters and more steps involved. After initial parameter testing and runnings on the simulated data it was also discovered that the algorithm does not have the restriction of the estimated trajectories staying inside the image frame. In the experimental data we are observing endosomes within cells, thus the boundaries are absolute. It was not possible to adapt the algorithm within the time frame of the project. It was therefore decided to continue only with the deterministic method.

The linking method was also used to obtain the mean displacement and mean velocity of A$\beta$ (from data sets containing the dynamics of A$\beta$ and lysosomes) from the estimated trajectories. They show how long the particles approximately move and their approximate velocity over 30 frames and the distribution of this over films taken hourly for 20 hours. It was done by calculating the length of the obtained

trajectories and over how many images the trajectories were distributed. The results shown in Figure 4.16 visualise the mean and standard deviation over data sets on different imaged positions each containing several cells. The data sets include 30 frames for each hour and the trajectory length/velocity is calculated by taking the mean of all trajectory lengths/velocities over the 30 frames taken hourly. The results show a slight decrease of A$\beta$ mean displacement and velocity over time which is consistent with the A$\beta$ accumulation in more stationary lysosomes over time.



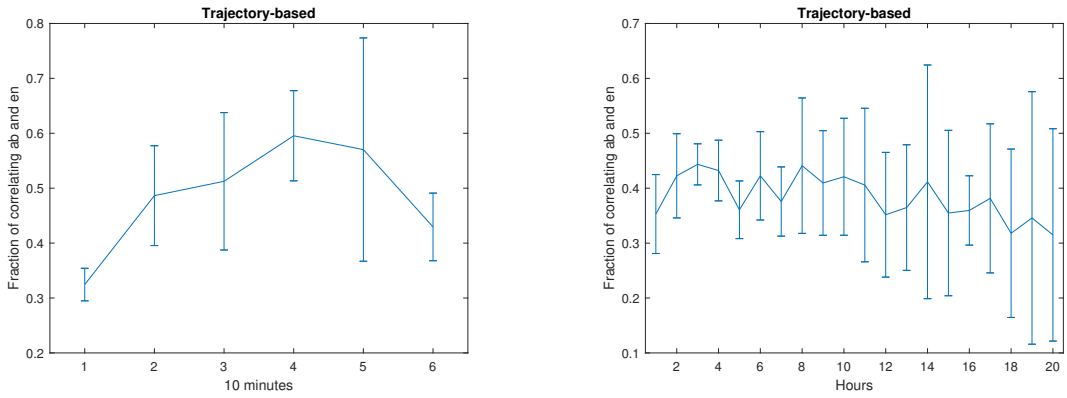

**(a)** **(b)**

**Figure 4.16:** Mean displacement and velocity of A$\beta$.

## 4.4 Correlation analysis

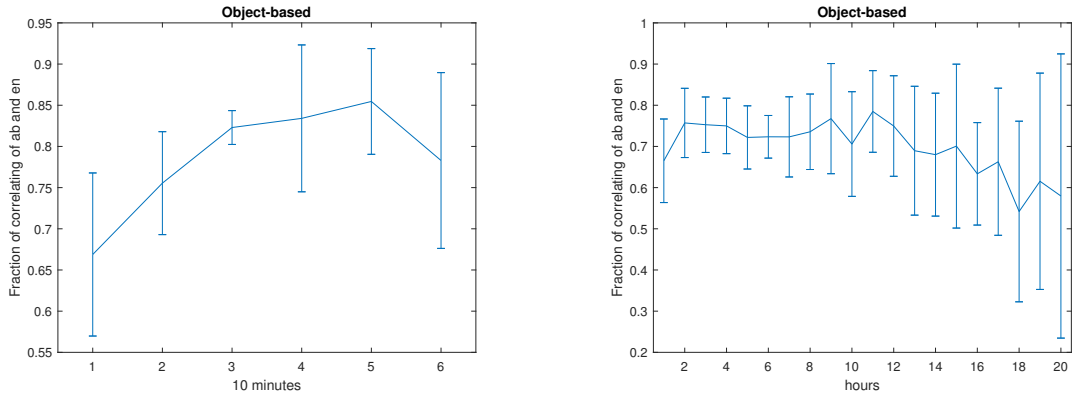
The correlation methods described in Section 3.5 were applied on experimental data of A$\beta$ and lysosomes recorded both with 10 minutes intervals (30 frames of movement recorded after each 10 minutes) and with 1 hour intervals (30 frames of movement recorded after each hour). Both types of data sets were recorded on several positions observing one or several cells. The expected correlation is that of increasing correlation over time, as A$\beta$ is assumed to accumulate in the lysosomes at the end of the endosomal pathway. The time dependence of this process was however not known. The chosen parameters for the methods were the following: displacement threshold of 10 pixels (only relevant and applied in the trajectory-based method), and to count for a match between the endosome and A$\beta$ particles on their paired trajectories, the particles centers were at most 5 pixels from each other (*epsilon*-value).

The trajectory-based method gave the resulting fraction of correlation seen in Figure 4.17. The figures show mean correlation fraction and their standard deviations over films of cells taken at different positions. For comparison, an object-based correlation analysis was also performed; results from which can be seen in Figure 4.18.




**(a)** **(b)**

**Figure 4.17:** Correlation of A$\beta$ and dextran-labeled lysosome results, using a trajectory-based correlation method.




**(a)** **(b)**

**Figure 4.18:** Correlation of A$\beta$ and dextran-labeled lysosome results, using an object-based correlation method.

# 5

# Discussion and conclusion

In the following section the results from the localization, linking and correlation analysis will be discussed and summarised. This will be done both from a software development perspective and from the biological application perspective.

The localization method found to be performing best on this application was the combination of a Gaussian filter method, an iterative global threshold method applied on the original image and the radial symmetry center method. The fact that the radial symmetry method outperformed the centroid method is expected as it has been shown in a more general particle localization analysis to give better results than the centroid method. [18] The Gaussian filter is also a more advanced low-pass filter while the Boxcar is a simpler version and as our image data is of low contrast it makes sense that the more advanced filter gives better results.

The best suiting threshold method was more surprising, considering that it is a global method and that it was applied on the original image (meaning that it sets the threshold using the original image but the set threshold is then applied on the filtered version of the image in this specific combination of methods for localization). It makes more sense that the correctly chosen threshold for the filtered image should come from the application of the threshold method on the filtered image and not the original image. Comparing between the two combinations including Gaussian filter, radial symmetry center and the iterative threshold applied on either the original image or the filtered image, one can see that the biggest difference is in the specificity values, meaning the fraction of correctly classified non-particle center pixels (background and particle pixels that are not particle centers) as non-particle center pixels. For this we can go back to our previous discussion of how the threshold value is set on images with different SNR. On more noisy images the threshold is set higher, meaning that more pixels will be classified as background pixels. When the threshold method is applied on the unfiltered image (containing more noise), the threshold will be set higher, more pixels will be classified as background, and as in our images we have much more background than particle pixels (center and not center), it increases the probability of a higher fraction of correctly classified non-particle center pixels as not particle centers. This could be the reason why the threshold method performs better when applied on the original image.

When comparing the ideas of a global and a local threshold method, a local method should perform better as it adapts more over the entire image. The fact that the global method performs better on the simulated data could have to do with how the

local threshold is set. It could have been set on too small windows not including the whole "image" of the contrast for example and not including the whole distribution of intensities which affects how the threshold is set as mentioned previously. The local threshold is neither calculated in the exactly the same way on the windows as the global method. Also, only one local threshold method was tested in the project, other local methods could show different performance. More global methods could be tested also.

It should also be noted here that the simulated data does not mimic the experimental data completely, but only its overall appearance. It does not include the variance of contrast and noise within the images (in e.g. very particle dense areas) as is present in the experimental images. A local threshold could have a different performance in this situation and could outperform the global threshold. In this project the threshold methods are only tested on simulated data that does not have this variance within the images. Further development of simulated data should be done to include more attributes of the experimental data to draw better conclusions about the performance of threshold methods. The simulated localization data gave us however a measure of how the methods perform on images with many particles, including smaller clusters of particles, varying sizes and with different levels of noise over the entire images.

Looking at the results of the found best performing localization method, one can conclude that the results were good (on both simulated and experimental data), even though the possibly more promising machine learning method was not managed to be applied, meaning that an even better performance could be obtained. The found optimal localization method resulted in as mentioned some false positives and not all true positive particles were localised but it should not be a problem when applying the results in linking. We are interested in the mean movement and mean correlation meaning that a small missed fraction of particles should not have a big effect and false positives should disappear when linking as they should not be linked over frames and therefore should not exist in the resulting particle trajectories.

The performance of particle linking was not as good as the localization performance on simulated data, which is expected because of a much more complicated task of linking. The simulated data that the linking method was tested on was also quite difficult and *epsilon* value was set quite strictly, not allowing space for smaller errors. More methods should be applied and tested (like the probabilistic method) to see if an improved performance could be obtained, with the goal of as high sensitivity and precision values as possible.

There was a difficulty in choosing the displacement threshold for experimental data, as no specific value of the displacement threshold showed clear better performance when testing on simulated data. The simulated data gave however a measure of performance of the method with varying parameters (even though the performance did not differ much between different parameters) and gave a visualisation how well the estimated trajectories follow the simulated ones in the difficult case of many

crossing trajectories. It would also be interesting to do the same tests and obtain a measure of performance on a simpler simulated motion with shorter trajectories and less crossing, which seemed to be the most common case in experimental data. Also, when investigating the experimental data more thoroughly, it was discovered that a big part of particles seemed to move smaller random steps while a smaller fraction took much larger directed steps. This could mean a more switching motion (between random and directed motion) of particles than the assumed directed Brownian motion for all particles. This assumption was made both in the simulated data and in the linking algorithm. Further inspection of the data is necessary to be able to draw a conclusion whether the assumption was wrong or not. This fact also made it difficult to choose a displacement threshold suiting both behaviour types of the particles. The displacement threshold affect how the resulting trajectories look like, and have an effect on the resulting mean displacement, velocity and correlation calculations.

Comparing the correlation curves of the trajectory-based and object-based methods one can see that they are similar in the trend of correlation over time. The object-based method gave quite higher correlation which is expected since here we are only looking at how many of the particles are correlating in space at different time-points, while using the trajectory based method we are looking at correlation in space and time, i.e. how many of the particles move together over time and not just are at the same positions. A smaller fraction is expected in a trajectory-based correlation as you do not include the particles that are at the same positions accidentally and do not actually move together. False positive particles should also be filtered away by looking at the trajectories and not only the positions, which could contribute to the smaller fraction.

The fact that the curves have a similar appearance (besides their overall values which should differ as explained above) is very positive. It shows that the trajectory-based method shows correct results and is working well for the application. This opens up for analysis opportunities with the calculated trajectories as particle mean displacement and velocity as was shown to be calculated in the linking results, one could also determine the direction of the trajectories and do similar measurements that cannot be done by only an object-based analysis.

Overall results of both correlation methods do not look exactly as expected. In their development over 20 hours it is expected that the correlation increase (until maximum correlation) in the beginning and then reach a steady state, as when A$\beta$s reaches the lysosomes they should stay there. In the resulting curves we however see some decrease in the last hours. This was investigated further by looking at the data and results from different parts of the algorithm and it was concluded that this behaviour could have come from the low contrast of A$\beta$ films in the end hours (bleaching). If the particles get too bleached it gets too difficult for the algorithm to set a good threshold value as the intensity of the background and particles approach each other (also when there is noise present in the image, as it was, too bleached particles could be removed as noise if they appear similar to the noise). This could

lead to a threshold value set such that a lot of the background is counted as particles (false positives); a too big fraction of false positives could disturb the tracking results and lead to false trajectories which are then not matched with the lysosomes (the contrast of lysosome particles was much better, with very little amount of false positives) and the correlation fraction decreases. This amount of bleaching was however not present in all position data sets of A$\beta$, which explains the big standard deviation.

One can also observe that a quite large amount of standard deviation of the correlation is present over the whole time interval over different positions, this was also investigated and it was seen that this could come from the fact that the cells in the frames move. In some films at a specific hour the cells had moved out from the frame fully or by half, this lead to a smaller fraction of particles present (almost none in some cases) and if the contrast of the present particles was also bad, it leads to the same case as with bleaching, a lot of false positive particles and the correlation fraction decreases a lot. In some cases the cells came back in the film for the next hour, in some cases they did not, also affecting the variance. These affecting factors could be minimized with further work on the collected data by imaging on a bigger frame including more cells so that even if one cell would leave the frame others would still be seen, however the maximum frame depends on the used microscope. One could also re-calculate the last more bleached time points on data collected in a different way that is not bleached at those time steps and compare the correlation values to see how much effect the bleaching has.

The threshold method was adapted slightly when applied on the experimental data for the above reasons, but it only improved the appearance a little. The changes that were done included minimum and maximum (15 and 200 particles respectively) values of A$\beta$ to be found and a maximum (200 particles) value of lysosomes, if the found number did not meet the criteria the threshold was either decreased or increased. These restrictions were added to reduce the risk of no found particles if their intensity was too low as it often was for the A$\beta$ particles (leading to a too high threshold value) and too many found particles (background classified as particles) if the threshold was set too low. A further improvement that could be made here is to have dynamical (if possible) maximum and minimum values, as the ones set now were only tested on a small amount of images and might not suit all. If for example more than one cell is present in the frame, leading to more particles that should be found, and this amount might be bigger than the maximum value now set. This also affects the outcome of the correlation and could have led to a slight decrease in the correlation fraction. The "correct" extremum values could also be different for A$\beta$ and endosomes/lysosomes depending on the data.

The correlation figures containing 10 minute intervals show the expected behaviour of increase of correlation in the beginning of time, except for the last dip and the big variance, which probably come from the above explained reasons.

To obtain curves for understanding the reality better, further work on both collection of the data as mentioned above and also investigation of other threshold

methods or development a new threshold method that can handle the bleaching better should be done. Other filtering and image enhancement methods could also be investigated to exclude the possibility of bleached particles being removed as noise.

The correlation values could also be improved by a correction including of what is expected as random correlation and what is expected from maximum correlation. Further testing and comparison of the trajectory-based method could also be done by inclusion of other developed methods for correlation. One could also look at the amount of correlating particles instead of the fraction, but that would require additional work as the overall amount of particles present in the images varies over time and cell positions. One could also calculate the mean displacement and velocity of correlating A$\beta$ and lysosomes as was done for the A$\beta$ trajectories in the linking results. This is however more difficult as not all positions on paired A$\beta$ and lysosome trajectories are actually matching. In some cases the beginning of the trajectories are matching where one could look at the matching interval, but in other cases there is only one matching position or two that are not in a row i.e. not forming a trajectory. Without a trajectory you cannot investigate displacement or velocity. One could in that case only look at the matching positions on paired trajectories that are in a row (form a trajectory), but this would not include the single matching positions on trajectories and the fraction of matching positions in a row could be small compared to all matching positions. For further work on the application one could also investigate the correlation of the other endosomes and A$\beta$.

To conclude, the applied and developed methods can and should be improved in several ways to be able to draw completely correct conclusions about the biological application they are applied to. Nevertheless, the methods have an overall good performance and particle tracking and trajectory based colocalization analysis could be a promising tool for investigation of A$\beta$ trafficking in Alzheimer's disease conditions.

# References


[1] 2018 ALZHEIMER'S DISEASE FACTS AND FIGURES Includes a Special Report on the Financial and Personal Benefits of Early Diagnosis;. Available from: https://www.alz.org/media/Documents/facts-and-figures-2018-r.pdf.

[2] Wang XP, Ding HL. Alzheimer's disease: epidemiology, genetics, and beyond. Neurosci Bull. 2008;24(2):105–109. Available from: http://www.neurosci.cn.

[3] Bird TD. Early-Onset Familial Alzheimer Disease. University of Washington, Seattle; 1993. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20301414.

[4] Isik AT. Late onset Alzheimer's disease in older people. Clinical interventions in aging. 2010 10;5:307–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21103401http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2981103.

[5] Funk KE, Kuret J. Lysosomal Fusion Dysfunction as a Unifying Hypothesis for Alzheimer's Disease Pathology. International Journal of Alzheimer's Disease. 2012 8;2012:1–10. Available from: http://www.hindawi.com/journals/ijad/2012/752894/.

[6] Crimins JL, Pooler A, Polydoro M, Luebke JI, Spires-Jones TL. The intersection of amyloid $\beta$ and tau in glutamatergic synaptic dysfunction and collapse in Alzheimer's disease. Ageing research reviews. 2013 6;12(3):757–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23528367http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3735866.

[7] Cataldo AM, Peterhoff CM, Troncoso JC, Gomez-Isla T, Hyman BT, Nixon RA. Endocytic Pathway Abnormalities Precede Amyloid $\beta$ Deposition in Sporadic Alzheimer's Disease and Down Syndrome: Differential Effects of APOE Genotype and Presenilin Mutations. The American Journal of Pathology. 2000 7;157(1):277–286. Available from: https://www.sciencedirect.com/science/article/pii/S0002944010645385.

[8] Elkin SR, Lakoduk AM, Schmid SL. Endocytic pathways and endosomal trafficking: a primer. Wiener medizinische Wochenschrift (1946). 2016 5;166(7-8):196–204. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26861668http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4873410.

[9] Ettinger A, Wittmann T. Fluorescence live cell imaging. Methods in cell biology. 2014;123:77–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24974023http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4198327.

[10] Kalaidzidis Y. Intracellular objects tracking. European Journal of Cell Biology. 2007 9;86(9):569–578. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0171933507000775.

[11] Chenouard N, Smal I, de Chaumont F, Maška M, Sbalzarini IF, Gong Y, et al. Objective comparison of particle tracking methods. Nature Methods. 2014 3;11(3):281–289. Available from: http://www.nature.com/articles/nmeth.2808.

[12] Helgadottir S, Argun A, Volpe G. Digital video microscopy enhanced by deep learning. 2018 12;Available from: http://arxiv.org/abs/1812.02653.

[13] Vercauteren D, Deschout H, Remaut K, Engbersen JFJ, Jones AT, Demeester J, et al. Dynamic Colocalization Microscopy To Characterize Intracellular Trafficking of Nanomedicines. 2011;Available from: www.acsnano.org.

[14] Pearson's Correlation Coefficient - Statistics Solutions;. Available from: https://www.statisticssolutions.com/pearsons-correlation-coefficient/.

[15] Wesén E, Jeffries GDM, Matson Dzebo M, Esbjörner EK. Endocytic uptake of monomeric amyloid-$\beta$ peptides is clathrin- and dynamin-independent and results in selective accumulation of A$\beta$(1–42) compared to A$\beta$(1–40). Scientific Reports. 2017 12;7(1):2021. Available from: http://www.nature.com/articles/s41598-017-02227-9.

[16] Dupont A, Stirnnagel K, Lindemann , Lamb DC. Tracking Image Correlation: Combining Single-Particle Tracking and Image Correlation. Biophysj. 2013;104:2373–2382. Available from: http://dx.doi.org/10.1016/j.bpj.2013.04.005.

[17] Deschout H, Martens T, Vercauteren D, Remaut K, Demeester J, De Smedt S, et al. Correlation of Dual Colour Single Particle Trajectories for Improved Detection and Analysis of Interactions in Living Cells. International Journal of Molecular Sciences. 2013 8;14(8):16485–16514. Available from: http://www.mdpi.com/1422-0067/14/8/16485.

[18] Parthasarathy R. Rapid, accurate particle tracking by calculation of radial symmetry centers. Nature Methods. 2012 7;9(7):724–726. Available from: http://www.nature.com/articles/nmeth.2071.

[19] Generate Data for Clustering - File Exchange - MATLAB Central;. Available from: https://se.mathworks.com/matlabcentral/fileexchange/37435-generate-data-for-clustering.

[20] High-Pass Filtering (Sharpening);. Available from: https://diffractionlimited.com/help/maximdl/High-Pass_Filtering.htm.

[21] Low-Pass Filtering (Blurring);. Available from: https://diffractionlimited.com/help/maximdl/Low-Pass_Filtering.htm.

[22] Iterative threshold selection on an input gray-level image - MATLAB Answers - MATLAB Central;. Available from: https://se.mathworks.com/matlabcentral/answers/345166-iterative-threshold-selection-on-an-input-gray-level-image.

[23] Sauvola local image thresholding - File Exchange - MATLAB Central;. Available from: https://se.mathworks.com/matlabcentral/fileexchange/40266-sauvola-local-image-thresholding.

[24] Granström K, Svensson L, Xia Y, Williams J, Garcia-Fernandez AF. Poisson multi-Bernoulli mixture trackers: continuity through random finite sets of trajectories. 2018 12;Available from: http://arxiv.org/abs/1812.05131.