DME, University of Southern Denmark
Jost Adam and Matiyas Korsa

Final Project

Handed out: Wednesday, December 1, 2021
Return by: Sunday, December 19, 2021, by 23:59

Group Number:		(Number)
Pass/Fail:		(Prof.)
Comments:		(Prof.)

Announcements:

Please **fill out your group number** on this sheet and **use it as a cover sheet for your project report**. Hand in the report and your source code (as one zipped file via Black Board), prior to the return date stated above. Please hand in **one report per group**, prepared by **all group members**.

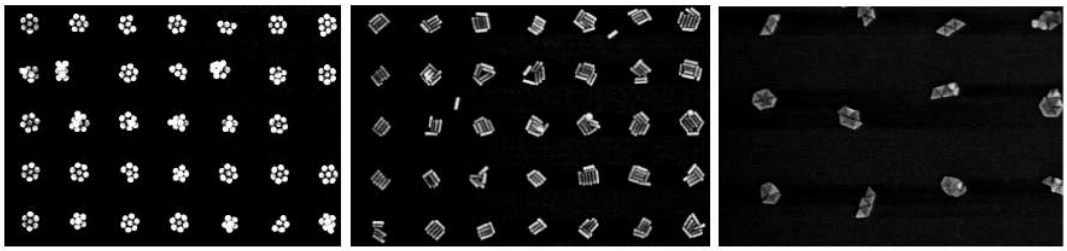

Figure 1: Sample SEM images of Au Nano clusters.

In this project you will work on the statistical image analysis of nano-particle "clusters" in regular, periodic arrangements, as shown in Fig. 1. You are going to derive a combined image processing and optimisation approach to classify the particle types, and number of particles per unit cell (cluster).

Part 1 - Image Processing and Feature extraction:

Use (and extend) you image processing library to automatically extract the following quantities from the images provided to you:

1.1 Extract features of each clusters in constant parameters (E.g., number of foreground pixels, lines and circles).

1.2 Determine number of particles per cluster.

1.3 Discuss the possibility of determining/predicting number of clusters and type of particles from the extracted features, i.e, Machine learning classification techniques.

The questions should be answered in form of a *histogram*, a *box plot* and a *violin plot*, displaying the quantity in question over the (reasonably binned) number of particles/clusters fulfilling this quantity. Furthermore, calculate and state the distribution's ***mean, median, standard deviation, and variance***. Take into account all images provided to create one statistic for all.
For 1.3 The discussion should be supported by various plots of features vs class(number of particles and/or type of particles)[2].
Optional: You can use parts of the clusters for K-nearest neighbors (KNN) algorithm implementations to predict number of particles per clusters and type of particle[1].

Part 2 - Optimization:

Based on the statistical distribution plots created above, fit your statistical data with Gaussian or Lorentzian curves (or sums thereof) to find and characterise all occurring peaks.

2.1 Formulate the related optimisation problem in detail for both cases.

2.2 Pick an appropriate optimisation algorithm for the fitting problem and implement the numerical solution.

2.3 Display the final results, showing the statistics together with your fitting function and print the final parameter values.

2.4 Based on the found function fit, analyse your results: Where are the significant peaks and what does that mean for the number of particles per cluster?

Submission details:

Submit your results in form of a report, containing a detailed description of your approach (either as a Juypyter Notebook or a pdf file), containing *at least*:

- Block diagrams of your approaches (which processing steps did you use to get to the final result?) for the image processing part.
- Figures documenting your processing steps (including the original image).
- The resulting data plots showing the wanted image analysis statistics, together with the final fitting function.

Additionally, you need to provide your source code (which will be tested for functionality).

It is highly recommended to create your report via (a) jupyter notebook(s)! This way, you can hand in your source code via the note book and the report as a (potentially post-processed) pdf-export from your jupyter notebook.

References

[1] *K-nearest-neighbor-classification-scikit-learn.* `http://https://www.datacamp.com/community/tutorials/k-nearest-neighbor-classification-scikit-learn`. Accessed: 2010-09-30.

[2] Pratap Chandra Sen, Mahimarnab Hajra, and Mitadru Ghosh. *Supervised Classification Algorithms in Machine Learning: A Survey and Review.* Vol. 937. Springer Singapore, 2020, pp. 99–111. ISBN: 9789811374029. DOI: 10.1007/978-981-13-7403-6_11. URL: http://dx.doi.org/10.1007/978-981-13-7403-6_11.