Data Fusion at the Nanoscale: Imaging at Resolutions Better than Wavelength/100

Rieger, Bernd; Stallinga, Sjoerd

Publication date
2017

Document Version
Final published version

Published in
ERCIM News

network are needed to accurately predict those latter high-level processes: while prediction fails when a simple contour-extraction model is used instead.

The artificial network and the brain

To further understand the model, we assess how well each contributing layer of the convolutional net fits each region of the brain image. This process yields a characterisation of the large-scale organisation of the visual system: the lower level layers of the convolutional deep architecture best predict early visual areas such as V1, V2 and V3, while later areas such as V4, lateral occipital and dorsal areas V3A/B, call for deeper layers. The beauty of this result is the continuity in the mapping between brain activity and the convolutional network: the layer preference of brain locations is a smooth mapping, suggesting that the artificial architecture is a relevant abstraction of the biological visual system. This is not unexpected, since convolutional networks have partly been designed as a model of the visual cortex, but it is also known that the wiring of the visual cortex in humans and animals is very different from the purely feedforward model used in artificial networks.

Importantly, this finding is not tied to the experimental data used, but it transposes to a different dataset with different experimental settings: natural video stimulation.

References:

[1] K. N. Kay et al.: "Identifying natural images from human brain activity", Nature 452: 352–355, 2008.

[2] P. Sermanet, et al.: "OverFeat : Integrated Recognition, Localization and Detection using Convolutional Networks" arXiv preprint arXiv:13126229 : 1–15, 2013

[3] M. Eickenberg, et al.: "Seeing it all: Convolutional network layers map the function of the human visual system", NeuroImage, Elsevier, 2016.

Please contact:

Bertrand Thirion

Inria, France

bertrand.thirion@inria.fr

Data Fusion at the Nanoscale: Imaging at Resolutions Better than Wavelength/100

by Bernd Rieger and Sjoerd Stallinga (Delft University of Technology)

Standard fluorescent light microscopy is rapidly approaching resolutions of a few nanometers when computationally combining information from hundreds of identical structures. We have developed algorithms to combine this information taking into account the specifics of fluorescent imaging.

For a hundred years a key rule in microscopic imaging was that the resolution in a microscope is determined by the diffraction limit, which states that the smallest detail separable is given by the wavelength of the light divided by twice the numerical aperture, where the latter is a measure for how much light is captured by the objective lens from the sample. For typical values of visible light, which has a wavelength ~500 nm and high quality immersion objectives, this results in diffraction limits of ~ 200 nm.

With the advent of localisation based super-resolution microscopy around 10 years ago, resolutions in the order of tens of nanometres have been increasingly reported in the literature. As a consequence, in 2014 this particularly successful imaging modality was awarded a Nobel Prize in Chemistry. The basic idea is to localise single isolated fluorescent emitters [1, L1]. This can be realised by photo-chemically altering the properties of the dyes such that most of the time only a very small fraction of all molecules emit light. Many cycles of recording and switching will result in imaging most of the emitters present in the sample. Due to the prior knowledge that only a single emitter is seen, one can fit a model of the system's response function (point spread function) to the data taking into account the different noise sources in a maximum likelihood model and localise the centre with an uncertainty that scales inversely with the square root of the number of detected photons. Typical values result in uncertainties in the tens of nanometres.

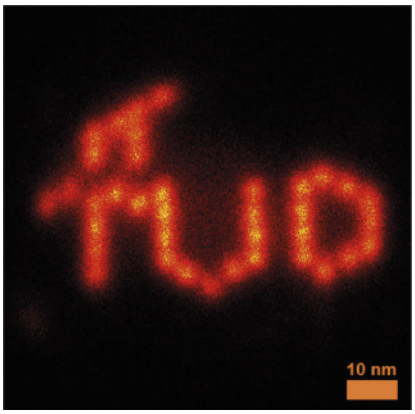

Figure 1: Data fusion image from 456 individual super-resolution reconstruction of a TU Delft logo. The logo is constructed with DNA-Origami and images by PAINT imaging. The final image combines about one million localisations resulting in a computed resolution [1] of 4.0 nm. That is wavelength/140 resolution with the hardware of a conventional light microscope.

However, the overall resolution in an image is a combination of the localisation accuracy and the labelling density of the structure with fluorescent emitters [references in 1]. This problem can be mitigated if instances of the same biological structure are fused (combined) properly. This technique is similar to single particle analysis that has been applied in the field of cryo-electron microscopy for many years [2]. Typically, these particles have arbitrary pose, are degraded by low photon count, false localisations and missing fluorescent labels.

If the structure to be imaged is known a priori, one approach is to register all particles on a template. However, this introduces 'template bias', which can occur if too strong prior knowledge is imposed on the data [2]. We are addressing this problem via a template-free data fusion method. Ideally one

would be able to use an all-to-all registration and then globally optimise the translation and rotation parameters. However, this results in a computational complexity scaling with N2 where N is the number of particles. Many hundreds of particles are needed to improve the effective labelling density and in turn the resolution. Therefore we currently register pairs of particles in a tree structure from leaves to root. This reduces the computational effort to a practically feasible task without the need for supercomputers, i.e., it enables the researcher to perform a full registration on the same time scale as the acquisition (order of 10-240 minutes). For the pair registration, we develop an algorithm, which takes into account the localisation accuracies and can handle missing labels [2]. This is essential as algorithms from electron microscopy work on images rather than on point clouds with associated uncertainties. In Figure 1 we show the result of such a procedure on samples that represent logos of the TU Delft shaped by DNA-origami and imaged with DNA-PAINT [3]. While the initial particles have computed resolution values [1] in the range of 10-30 nm, the final reconstructed particle has 4.0 nm resolution. This is about the best that can be expected, as the distance between the binding sites on the DNA blueprint are 5 nm.

This technique will no doubt have further applications in biology since it is common for many chemically identical copies of a structure to be present within one cell, making imaging an easy task.

This work is partly financially supported by European Research Council grant no. 648580 and National Institute of Health grant no. 1U01EB021238-01.

Links:

[L1] 350 years of light microscopy in Delft: https://www.youtube.com/watch?v=vh3_qOy2uls

References:

[1] B. Rieger, R.P.J. Nieuwenhuizen, S. Stallinga: "Image processing and analysis for single molecule localization microscopy", IEEE Signal Processing Magazine, Special Issue on Quantitative Bioimaging: Signal Processing in Light Microscopy, 32:49-57, 2015.

[2] A. Löschberger, et al.: "Super-resolution imaging reveals eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution", Journal of Cell Science, 125:570-575, 2012.

[3] R. Jungmann, et al.: "Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and exchange-PAINT", Nature Methods, 2014.

Please contact:

Bernd Rieger, Delft University of Technology, The Netherlands
b.rieger@tudelft.nl
http://homepage.tudelft.nl/z63s8/

Computational Processing of Histological Images

by Erwan Zerhouni, Bogdan Prisacari, Maria Gabrani (IBM Zurich) and Qing Zhong and Peter Wild (Institute of Surgical Pathology, University Hospital Zurich)

Cognitive computing (in the sense of computational image processing and machine learning) helps address two of the challenges of histological image analysis: the high dimensionality of histological images, and the imprecise labelling. We propose an unsupervised method of generating representative image signatures that are robust to tissue heterogeneity. By integrating this mechanism in a broader framework for disease grading, we show significant improvement in terms of grading accuracy compared to alternative supervised feature-extraction methods.

Disease susceptibility and progression is a complex, multifactorial molecular process. Diseases, such as cancer, exhibit cellular heterogeneity, impeding the differentiation between diverse stages or types of cell formations, such as inflammatory response and malignant cell transition. Histological images, that visualise tissues and their cell formations, are huge; three to five orders of magnitude larger than radiology images. The high dimensionality of the images can be addressed via summarising techniques or feature engineering. However, such approaches can limit the performance of subsequent machine learning models to capture the heterogeneous tissue microenvironment. Image analysis techniques of tissue specimens should therefore quantitatively capture the phenotypic properties while preserving the morphology and spatial relationship of the tissue microenvironment.

To capture the diverse features of heterogeneous tissues in large tissue images, we enhance the computational framework introduced by Zerhouni et al [1] by also addressing the cellular heterogeneity, without the need for cellular annotation. This reduces the dependency on labels that tend to be imprecise and tedious to acquire. The proposed method is based on an autoencoder-architecture [2] that we have modified and enhanced to simultaneously produce representative image features as well as perform dictionary learning on these features to reduce dimensionality. We call our method DictiOnary Learning Convolutional autoEncoder (DOLCE).

The principle behind this dimensionality reduction is as follows. DOLCE aims to detect and learn the main morphological and colour patterns of cells that appear in tissue, especially in heterogeneous tissue. The objective is then to represent a cell by combining the main patterns that describe it. By using soft assignments that map each patch of a histology image to a set of dictionary elements, we enable a finer-grained representation of the data. To this end, we can quantify both similarities and dif-