Quantifying Cell Morphology With Label-free Imaging and Deep Learning

Dr. Johannes Bauer, TRI and Dr. Bruno Chilian, TRI

Introduction

Label-free modes of light-microscopy like bright-field or phase-contrast microscopy are a central pillar of routine workflows in cell culture laboratories. Scientists rely on this method, combined with their deep insight into cell biology and their experience in judging such images, to assess the state of their cell cultures and drive critical decisions. For quantitative measurements however, especially those that go beyond simple measurements of cell covered area, staining techniques combined with specialized image analysis routines are often employed. This raises the question: **If profound information about the cells' state is obviously contained in the label-free images, why can they not be used to extract quantitative results?**

Here, we review an analysis paradigm that employs deep learning methodologies as the key ingredient to turn label-free microscopy into a universal, quantitative assay technique. Artificial Neural Networks (ANN) are ideally suited to capture the specialized knowledge of one or more human experts to produce a robustified, objective and scalable version of the human brain's remarkable capability in structure detection and classification.

Workflow

After acquisition of a representative set of images, the machine-learning process is initiated by generating a set of labeled data for training. For this purpose, the human expert uses a labeling software that works like a simple graphics program: **The structures of interest are highlighted in partly transparent colors on top of the original images.** Multiple colors can be used to label different classes to be detected, e.g. cell nuclei and cytoplasm.

When a representative set of all relevant morphologies has been labeled in this way, a suitable ANN is trained to reproduce the labels. The performance of the ANN is validated on new images, which have not been used in training. This ensures that the ANN has learned to generalize from the specific labeled images and is now capable of detecting the intended morphologies in general. Additionally, quantitative metrics like the prediction accuracy of labeled data, which were not used in the training, may be employed to validate the ANN's performance. If the performance does not meet the requirements, the training process may be enhanced by adding more labeled data or by using a more potent architecture for the ANN.

Requirements for implementation

The approach outlined here is sufficiently generic as to be implemented by any team of researchers with the required skill set in cell culture, microscopy and data science. However, for application in productive routine use, a **high degree in standardization, automation and usability** in all three areas is required to make the method viable and efficient. Ideally, a specialized microscopy solution is employed that is tailored for the specific needs of machine learning rather than generic image inspection. Automated image acquisition and direct incorporation of metadata are essential to **avoid excessive variability and human error**. Furthermore, directly integrating the rather complex but nonetheless tedious steps for data preprocessing, machine learning and result generation with the previous steps is advisable to **eliminate the waste of precious data scientists' time** on repetitive tasks. All analyses shown here were performed with a VAIDR-System by TRI Thinking Research Instruments GmbH.

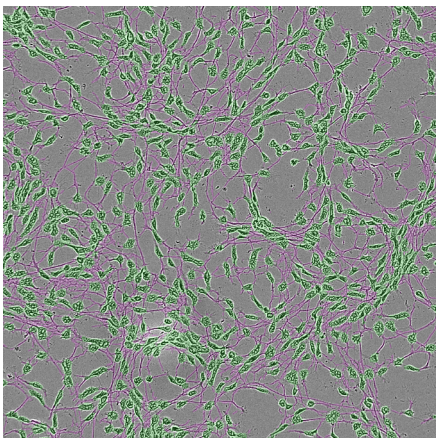

Fig. 1: Neurons imaged analyzed with the VAIDR-System. Colored overlays show segmentation results: Green: Cell bodies, magenta: neurites.

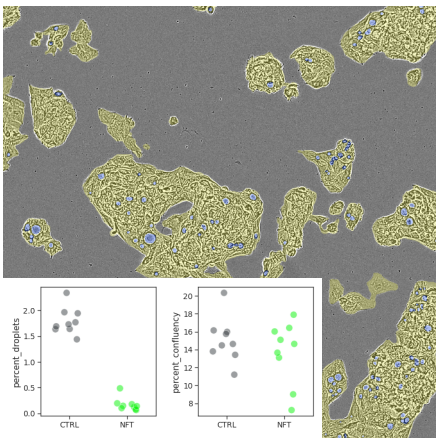

Fig. 2: HepG2 cells imaged analyzed with the VAIDR-System. Colored overlays show segmentation results: Yellow: Cells, blue: droplets. Inset shows diappearance of droplets under treatment with NFT, while confluency stays constant.

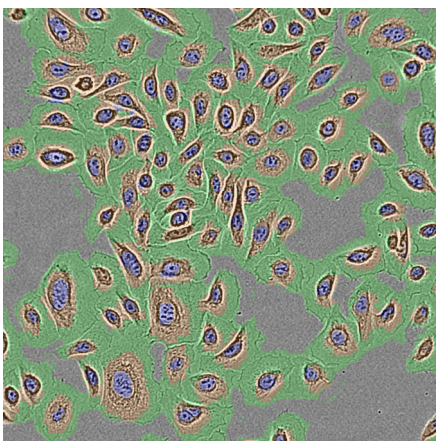

Fig. 3: Keratinocytes imaged analyzed with the VAIDR-System. Colored overlays show segmentation results: Blue: Nucleus, yellow: inner cytoplasm, green: outer cytoplasm.

Examples

With the method described above, precise and reliable image segmentation algorithms can be obtained. As shown in fig. 1, **very fine structures like neurites can be detected** and distinguished from the cell bodies. Such capabilities immediately lead to sensitive and robust assays, e.g. for the detection of subtle neuro-toxicological effects well below the threshold of actual aptotosis or necrosis .

As shown in fig. 2, treatment-induced effects, like the disappearance of sub-cellular structures, can be detected and quantified without the need to previously stain the structure. This opens the possibility of a **more un-biased approach to phenotypic assay design**, as compared to the total reliance on specific staining agents.

Finally, fig. 3 shows that segmentation of standard cell regions like nucleus and cytoplasm can be achieved with precision. Even the flat parts of adherent cells can be detected and quantified. Such a **standardized and generic approach of cell morphology quantification** opens the door to an un-biased, multiparametric characterization and clustering of complex physiological effects. This extends to sub-visual effects, i.e. morphological profiles that are characteristic of a particular condition, but which are not evident to a human observer. As an example, fig. 4 shows the successful classification of keratinocytes into the age-class of the cell donors.

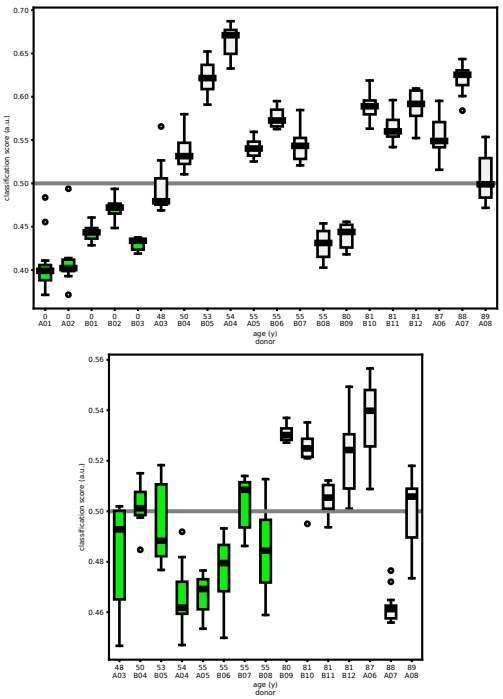

Fig. 4: Keratinocyte donor age prediction. ANN was trained to predict the donor age group (young, medium, old) in a two-step classification scheme:
(A) Classification results of young donors vs medium or old donors. (B) Results for medium vs old donors.

Acknowledgements

We thank Dr. Ken Raj, Public Health England, and Prof. Dr. Stefan Schildknecht, Universität Albstadt-Sigmaringen for providing the samples that were used for the analyses shown here.

TRI Thinking Research Instruments GmbH

info@vaidr.de

www.vaidr.de