ObiWan-Microbi: OMERO-based integrated workflow for annotating microbes in the cloud

Johannes Seiffarth[1,2], Tim Scherr[3], Bastian Wollenhaupt[1], Oliver Neumann[3], Hanno Scharr[4], Dietrich Kohlheyer[1,5], Ralf Mikut[3], Katharina Nöh[1,*]

[1] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

[2] Computational Systems Biotechnology (AVT.CSB), RWTH Aachen University, Aachen, Germany

[3] Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany

[4] Institute for Advanced Simulation, IAS-8, Forschungszentrum Jülich GmbH, Jülich, Germany

[5] Microscale Bioengineering (AVT-MSB), RWTH Aachen University, Aachen, German

Abstract

Summary: Reliable deep learning segmentation for microfluidic live-cell imaging requires comprehensive ground truth data. *ObiWan-Microbi* is a microservice platform combining the strength of state-of-the-art technologies into a unique integrated workflow for data management and efficient ground truth generation for instance segmentation, empowering collaborative semi-automated image annotation in the cloud.

Availability and Implementation: *ObiWan-Microbi* is open-source and available under the MIT license at https://github.com/hip-satomi/ObiWan-Microbi, along documentation and usage examples.

Contact: k.noeh@fz-juelich.de

Supplementary information: Supplementary data are available online.

Introduction

The combination of automated live-cell microscopy and microfluidic lab-on-chip technology unlocks unique spatio-temporal insights into the morphology of individual cells, the development of microbial populations, and community interactions (Burmeister and Grünberger, 2020). Time-lapse imaging produces large amounts of image data, calling for reliable high-throughput image processing pipelines. Deep learning approaches have become state-of-the-art for many bioimage applications (Moen *et al.*, 2019), especially when adaptability to various morphologies and imaging modalities are required (Caicedo *et al.*, 2019). Still, extracting microbial single-cell properties like cell count or area from microscopic image sequences remains challenging (Jeckel and Drescher, 2021).

The practical application of deep learning for segmentation (DLS) is often hampered by the lack of sufficient high-quality ground truth (GT) annotations for training and validation (Laine *et al.*, 2021), as well as a limited accessibility for users due to complex software dependencies or special hardware requirements (Moen *et al.*, 2019). Recently, solutions addressing individual aspects have been developed, such as Omnipose providing a user interface for cell segmentation (Cutler *et al.*, 2021), or ImJoy that simplifies DLS usage utilizing browser-based user interfaces, a flexible plugin architecture and remote computation

resources (Ouyang *et al.*, 2019). Still, there is no integrated end-to-end workflow that provides state-of-the-art DLS methods for microbial live-cell images, is extensible, allows automated instance segmentation and convenient manual correction, connects to remote computational resources, is collaboration-ready, and integrates image and meta-data management.

To leverage integrated, efficient and reproducible workflows in microbial DLS with all the mentioned features optimized for fast GT creation, we developed *ObiWan-Microbi* (Fig. 1). *ObiWan-Microbi* comprises an extensible collection of DLS methods to be executed on a remote server, is customizable by end-users, comes with a user-friendly web app interface providing semi-automated DLS for image data stored in an OMERO backend (Allan *et al.*, 2012), and facilitates team and project-based collaboration using cloud technology.

Approach and Implementation

The open *ObiWan-Microbi* platform consists of three major parts (Fig. 1): *SegServe* for convenient DLS execution, the graphical user interface *SegUI*, and the microservice architecture *SegMSA* for deployment and integration of OMERO data management. DLS execution is often difficult in practice due to hardware and software requirements. *SegServe* is a convenient REST API designed to execute DLS methods, bundled in git repositories providing reproducible code execution (SI S.1). The execution is performed on a single server, accelerated by GPU hardware and batch processing, with no need for local installation or software dependency handling on the user-side. The REST API functionality is automatically documented using FastAPI.

To unlock the power of DLS for GT generation, we developed the user-friendly graphical user interface *SegUI*, an Angular and Ionic based web app that connects to the OMERO data backend. *SegUI* allows browsing image data and performing fully manual or fast semi-automated image annotation of time-lapse sequences. Results from automated DLS by *SegServe* are available in the web app *SegUI* for manual correction using efficient drawing tools, a data model to undo/redo annotation actions, and multi-label annotations for complex imaging data. Custom DLS approaches are supported by *SegUI* to readily include new or optimized segmentation approaches and parameters settings. Thus, *SegUI* provides a convenient and interactive DLS experience for collaborative microfluidic image annotation.

SegUI and *SegServe* are combined with OMERO to form the microservice architecture *SegMSA* (SI S.2). OMERO takes care of image data handling and segmentation storage including user and access management, facilitating FAIR data principles. This way, OMERO-compatible modules, such as Fiji can be utilized to extract single-cell properties. *SegMSA* utilizes docker to manage services and their connections providing a convenient installation procedure for the microservice architecture deployment.
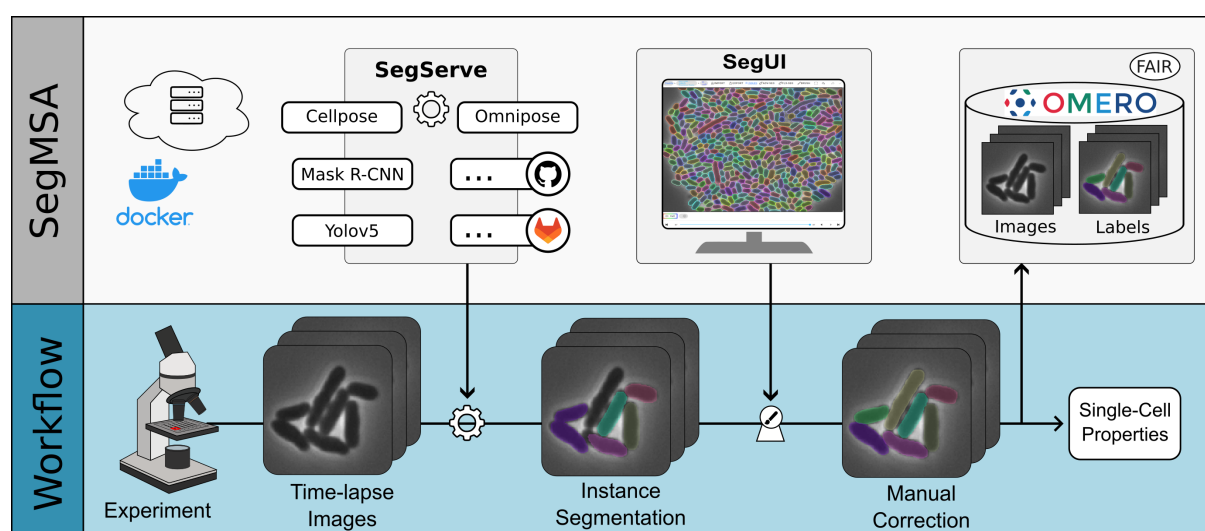
Figure 1. *ObiWan-Microbi* platform overview. Fully integrated end-to-end workflow for semi-automated image annotation and segmentation of time-lapse microscopy in the browser. *ObiWan-Microbi* consists of *SegServe*, *SegUI*, and OMERO jointly distributed in the microservice architecture *SegMSA*.

Results

The combination of the three components forms the platform *ObiWan-Microbi*, used to segment microbial organisms including *E. coli, C. glutamicum* and *B. subtilis* (SI S.3). For semi-automated segmentation, four pre-trained DLS algorithms are provided out-of-the-box and accessible in *SegUI* including Cellpose (Stringer *et al.*, 2021), Omnipose (Cutler *et al.*, 2021), Mask R-CNN trained on simulated image data (Sachs *et al.*, 2022) and Yolov5 for general object detection in images. Using semi-automated segmentation, an annotation speed of more than 200 cells/minute was achieved (SI S.4).

Conclusion

ObiWan-Microbi is a collaborative annotation platform for the efficient creation of high-precision GT data that debottlenecks the broad usage of DLS in microbial image processing, especially for distributed teams. Based on growing GT data sets, and in combination with integrated training workflows (Scherr *et al.*, 2022), we are looking forward to seeing increased usage of DLS methods and publicly available benchmarks for training and validation in the domain of live-cell imaging.

Acknowledgement

Funding

References

Allan,C. *et al.* (2012) OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods*, **9**, 245–253.

Burmeister,A. and Grünberger,A. (2020) Microfluidic cultivation and analysis tools for interaction studies of microbial co-cultures. *Curr. Opin. Biotechnol.*, **62**, 106–115.

Caicedo,J.C. *et al.* (2019) Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods*, **16**, 1247–1253.

Cutler,K.J. *et al.* (2021) Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *bioRxiv*.

Jeckel,H. and Drescher,K. (2021) Advances and opportunities in image analysis of bacterial cells and communities. *FEMS Microbiol. Rev.*, **45**, 1–14.

Laine,R.F. *et al.* (2021) Avoiding a replication crisis in deep-learning-based bioimage analysis. *Nat. Methods*, **18**, 1136–1144.

Moen,E. *et al.* (2019) Deep learning for cellular image analysis. *Nat. Methods*, **16**, 1233–1246.

Ouyang,W. *et al.* (2019) ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods*, **16**, 1199–1200.

Sachs,C.C. *et al.* (2022) CellSium – Versatile Cell Simulator for Microcolony Ground Truth Generation. *bioRxiv. doi 10.1101/2022.03.24.485611*.

Scherr,T. *et al.* (2022) microbeSEG: Accurate Cell Segmentation with OMERO Data Management. *bioRxiv. doi 10.1101/2022.04.29.489998*.

Stringer,C. *et al.* (2021) Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods*, **18**, 100–106.