The *TYC Dataset* for Understanding Instance-Level Semantics and Motions of Cells in Microstructures

ICCV 2023 Workshop on BioImage Computing



Christoph Reich[1,2,‡], Tim Prangemeier[1], and Heinz Koeppl[1,†]

[1]Technische Universität Darmstadt, Germany
[2]NEC Laboratories America Inc., USA[§]

[‡]christoph.reich@bcs.tu-darmstadt.de
[†]heinz.koeppl@bcs.tu-darmstadt.de

[§]Work done at Technische Universität Darmstadt

Introduction

- Many biomedical applications require the segmentation of single cells in microscopy imagery

Introduction

- Many biomedical applications require the segmentation of single cells in microscopy imagery
- Multiple cell segmentation datasets have been proposed (*e.g.*, DSB2018 [Caicedo et al., 2019])

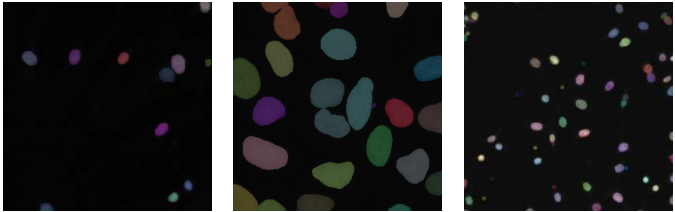

Figure: Sample of the DSB2018 dataset [Caicedo et al., 2019].

Introduction

- Many biomedical applications require the segmentation of single cells in microscopy imagery
- Multiple cell segmentation datasets have been proposed (*e.g.*, DSB2018 [Caicedo et al., 2019])

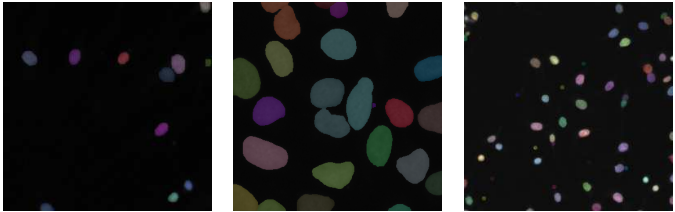

Figure: Sample of the DSB2018 dataset [Caicedo et al., 2019].

- Existing datasets only consider limited scenes (*e.g.*, no microstructures)

Introduction

- Many biomedical applications require the segmentation of single cells in microscopy imagery
- Multiple cell segmentation datasets have been proposed (*e.g.*, DSB2018 [Caicedo et al., 2019])

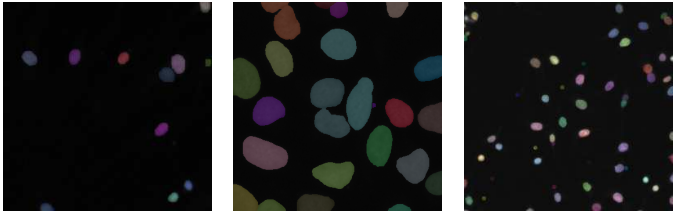

Figure: Sample of the DSB2018 dataset [Caicedo et al., 2019].

- Existing datasets only consider limited scenes (*e.g.*, no microstructures)
- **We present the first large-scale dataset including cells and microstructures**

Biological Application

Time-Lapse Fluorescence Microscopy

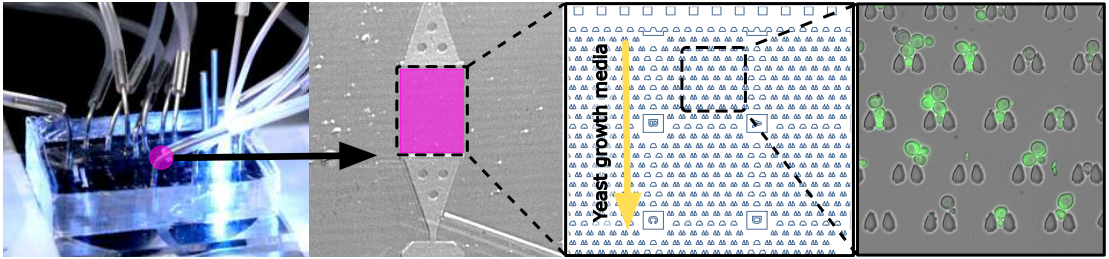

Figure: TLFM experiment setup [Prangemeier et al., 2020] for single-cell fluorescence measurement.

Biological Application

Time-Lapse Fluorescence Microscopy

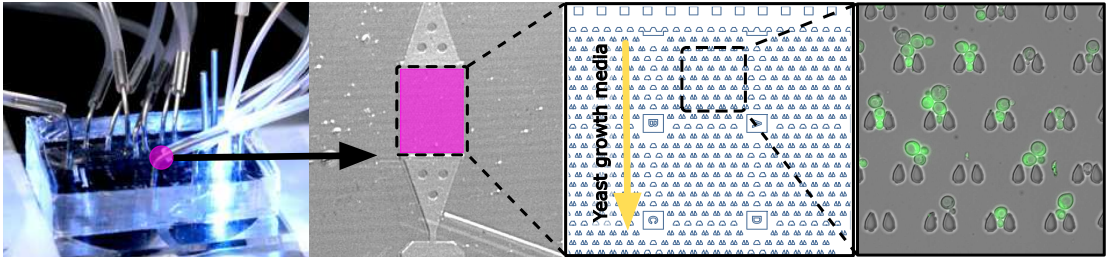

Figure: TLFM experiment setup [Prangemeier et al., 2020] for single-cell fluorescence measurement.

- We collect data from more than a dozen experiments conducted in our wet lab

Biological Application

Time-Lapse Fluorescence Microscopy

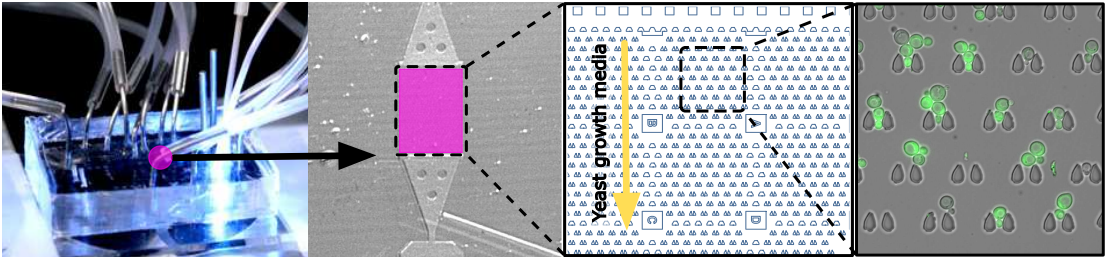

Figure: TLFM experiment setup [Prangemeier et al., 2020] for single-cell fluorescence measurement.

- We collect data from more than a dozen experiments conducted in our wet lab
- We also include data generated by the Swain lab [Bakker et al., 2018]

The TYC Dataset

Labeled Set

- We present ~ 19k instance masks for both cells and microstructures

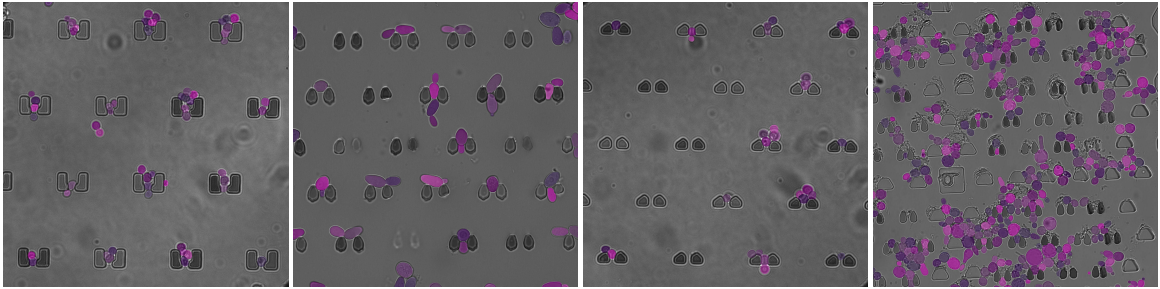

Figure: Labeled sample of our TYC dataset; shades of (dark) grey (■■■) indicate microstructures (traps) and shades of pink (■■■) indicate individual cell instances.

The TYC Dataset

Labeled Set Split

Split	# images	# ann. pixels [10^7]	# cells	# traps
Training	81	35.53	12296	3448
Validation	8	3.69	950	310
Test	8	3.59	753	346
OOD Test	8	3.58	542	301

Table: Split of our labeled set.

- Our out-of-domain does only include images from experiments not used in the other sets

The TYC Dataset

Unlabeled Set

- We also present 61 curated video clips to facilitate unsupervised learning

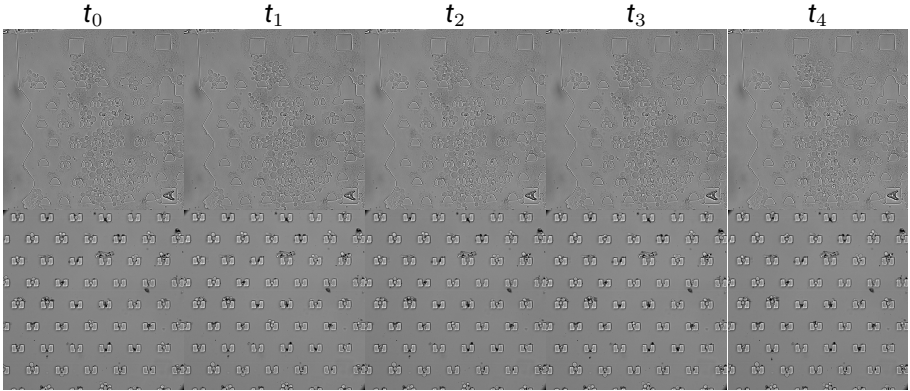

Figure: Unlabeled TLFM image sequences ($\Delta t = 10$min).

SAM Zero-Shot Results

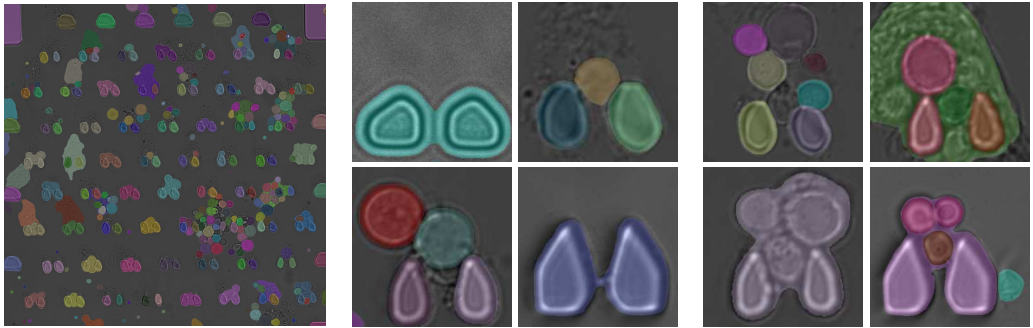

Figure: Zero-shot results of the Segment Anything Model [Kirillov et al., 2023] on our TYC dataset.

- SAM does not provide satisfactory segmentation results out of the box

Conclusion

- We proposed the **first large-scale instance segmentation dataset of cells in microstructures**

Conclusion

- We proposed the **first large-scale instance segmentation dataset of cells in microstructures**
- We additionally offer **unlabeled microscopy image sequences** for unsupervised learning

Conclusion

- We proposed the **first large-scale instance segmentation dataset of cells in microstructures**
- We additionally offer **unlabeled microscopy image sequences** for unsupervised learning
- We provide a **standardized instance and semantic-level evaluation strategy** for our dataset

Conclusion

- We proposed the **first large-scale instance segmentation dataset of cells in microstructures**
- We additionally offer **unlabeled microscopy image sequences** for unsupervised learning
- We provide a **standardized instance and semantic-level evaluation strategy** for our dataset

Project page | Code | Dataset | Paper

https://christophreich1996.github.io/tyc_dataset/

Work supported by: The Landesoffensive für wissenschaftliche Exzellenz, European Research Council, and NEC Laboratories America, Inc.

References

Bakker, E., Swain, P. S., and Crane, M. M. (2018).
Morphologically constrained and data informed cell segmentation of budding yeast.
Bioinformatics, 34(1):88–96.

Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., Heng, C., Becker, T., Doan, M., et al. (2019).
Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl.
Nature Methods, 16(12):1247–1253.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., et al. (2023).
Segment Anything.
arXiv preprint arXiv:2304.02643.

Prangemeier, T., Lehr, F.-X., Schoeman, R. M., and Koeppl, H. (2020).
Microfluidic platforms for the dynamic characterisation of synthetic circuitry.
Current Opinion in Biotechnology, 63:167–176.