

# InspectorCell: Finding Ground Truth in Multiplexed Microscopy Images

Andre Gosselink<sup>1,2\*</sup>, Tatsiana Hofer<sup>3\*</sup>, Elvira Criado-Moronati<sup>2</sup>, Arndt von Haeseler<sup>3,4</sup>, and Jutta Kollet<sup>2</sup>

<sup>1</sup> Institute of Medical Statistics and Computational Biology, University of Cologne, Bachemer Str. 86, 50931 Cologne, Germany <sup>2</sup> R&D Department, Miltenyi Biotec B.V. & Co. KG, Friedrich Ebert Straße 68, 51429 Bergisch Gladbach, Germany <sup>3</sup> Center for Integrative Bioinformatics Vienna, Max Perutz Labs, University of Vienna, Medical University of Vienna, Dr. Bohr Gasse 9, 1030 Vienna, Austria <sup>4</sup> Faculty of Computer Science, University of Vienna, Währinger Str. 29, 1090 Vienna, Austria \* These authors contributed equally.

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

## Summary

Multiplexed immunofluorescence microscopy gives deep insights into multiple characteristics of cells in biological samples relevant to fields like cancer research. These characteristics are often image intensities that reflect the expression of a marker on cell surfaces. Typical tasks in analyzing multiplexed microscopy images are the segmentation and quantification markers on cell surfaces in the biological sample. These tasks are only feasible with automation, often using trained machine learning models. However, training these models requires high-quality labeled datasets. With recent increases in image stack sizes, the generation of labeled datasets for supervised machine learning has become a significant bottleneck. InspectorCell alleviates this bottleneck by providing an intuitive, graphical interface for synchronized manual segmentation and annotation of cells in highly multiplexed microscopy images. The modular implementation of InspectorCell in Python enables tight integration into existing applications such as Orange3 or CellProfiler. An image dataset with exemplary annotations is available at: <https://doi.org/10.7303/syn37910913.2>

## Statement of Need

The cellular composition of tumors is of great scientific interest as the presence of tumor-infiltrating lymphocytes correlates with the survival of cancer patients (Idos et al., 2020; Santoiemma & Powell, 2015). Multiplexed tissue imaging methods like CODEX (Goltsev et al., 2018) or MACSima<sup>TM</sup> (Kinkhabwala et al., 2022) are used to investigate tumor samples' cellular characteristics. These techniques generate large image stacks of the same tumor sample slice, where each image covers the intensity profile of a specific cell surface marker. All cells in a tissue sample slice are categorized – for example, into cancer and immune cells – based on the cell surface marker intensities on each cell. This cell-based analysis for ever-growing large datasets seems to be only feasible with machine learning. However, a significant bottleneck when working with machine learning models is the lack of suitable multiplexed image datasets with segmented and quantified cells for model training.

Generating these datasets is time-consuming and cumbersome because current software does not integrate synchronized viewing, editing, and annotation of multiple images simultaneously. While it is possible to edit cell segments in CellProfiler (Carpenter et al., 2006), only a single channel can be evaluated simultaneously. FIJI (Schindelin et al., 2012) displays several immune images in parallel, but the different views on the images are not synchronized. An annotation

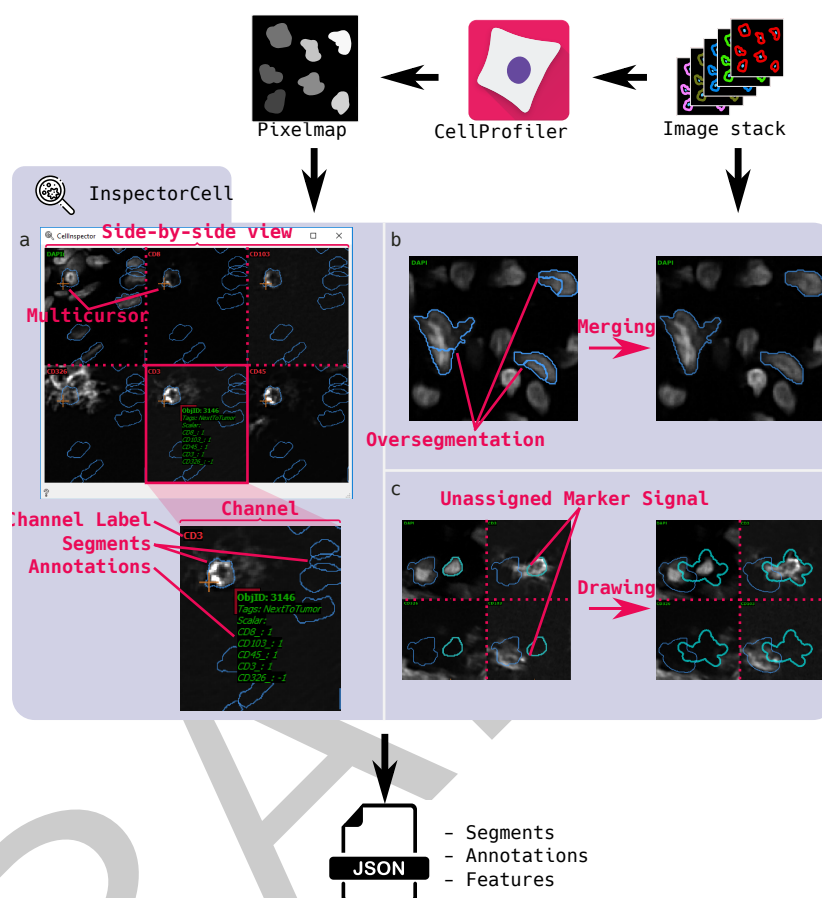
or change of a cell segment must be performed in each window individually, but finding the exact location of a single cell in all images is cumbersome. Applications like ilastik (Sommer et al., 2011) or CellPose (Stringer et al., 2021) fuse the analysis of images with the training of a machine learning model. However, a user can evaluate only one image at a time, and the generated data is not viable for subsequent model training.

We used InspectorCell to manually analyze cells in multiplexed immunofluorescence microscopy images of an ovarian cancer tissue slice obtained with the MACSima™ imaging platform. The tissue was stained with 99 cell markers, and a subset of the dataset is freely available. An initial cell segmentation was generated with CellProfiler and imported into InspectorCell. The cell segmentations were refined, and the cell marker expressions were quantified (fig. Figure 1a). In the initial CellProfiler segments were 1960 cells, and after refinement with InspectorCell, only 1750 cells remained. Among the immunological analysis results, we obtained a ground truth training dataset for subsequent model training, e.g., for continuous training of the CellPose model.

A synchronized overview over multiple images with editing capabilities for single-cell segmentation and quantification made this manual analysis possible. With InspectorCell, we provide a solution for efficient manual segmentation and annotation of large image stacks. The primary benefit of InspectorCell is the ability to view cells within the context of multiple cell markers, which accelerates manual analysis of cells in large image stacks. Expert immunologists can use InspectorCell to evaluate various cell characteristics at a glance and simultaneously generate high-quality training data.

DRAFT

## Figures



**Figure 1:** Exemplary use cases of InspectorCell. Segmentation of image stacks are generated with CellProfiler and opened with InspectorCell. (a) Six immunofluorescence images of the image stack are displayed side-by-side in a 3×2 grid. Cell segments are displayed as blue polygons (top). A synchronized multi cursor (orange) is a visual anchor in all channels. The CD3 channel (lower middle) is enlarged below the main window. The cell segment annotation is editable and displayed for the active cell segment in green font. (b) Over-segmentation can be merged with a single keystroke after mouse selection. (c) A cell segment (light blue) can be edited to embrace the marker signal area attributable to a distinct cell. Multiple segments can enclose the same areas to reflect cell overlaps and interactions. The manual edits of cell segmentations and annotations are saved in a single JSON file. Additionally, the JSON file can store extracted cell features, for example, mean pixel intensities. The JSON file is in plain text and can be readily used in downstream analysis.

## Acknowledgements

We thank Ali Kinkhabwala for his help during development, Bianca Heemskerk for releasing the data, Paurush Praveen, Werner Müller and Achim Tresch for reviewing the manuscript. This project was partially funded by Miltenyi Biotec B.V. & Co. KG.

## References

Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., Golland, P., & Sabatini, D. M.

- (2006). CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10), R100. <https://doi.org/fqfn3j>
- Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S., & Nolan, G. P. (2018). Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell*, 174(4), 968–981.e15. <https://doi.org/gdwtdq>
- Idos, G. E., Kwok, J., Bonthala, N., Kysh, L., Gruber, S. B., & Qu, C. (2020). The Prognostic Implications of Tumor Infiltrating Lymphocytes in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Scientific Reports*, 10(1), 3360. <https://doi.org/gkspmt>
- Kinkhabwala, A., Herbel, C., Pankratz, J., Yushchenko, D. A., Rüberg, S., Praveen, P., Reiß, S., Rodriguez, F. C., Schäfer, D., Kollet, J., Dittmer, V., Martinez-Osuna, M., Minnerup, L., Reinhard, C., Dzionek, A., Rockel, T. D., Borbe, S., Büscher, M., Krieg, J., ... Bosio, A. (2022). MACSima imaging cyclic staining (MICS) technology reveals combinatorial target pairs for CAR T cell treatment of solid tumors. *Scientific Reports*, 12(1), 1911. <https://doi.org/gqxxsr>
- Santoemma, P. P., & Powell, D. J. (2015). Tumor infiltrating lymphocytes in ovarian cancer. *Cancer Biology & Therapy*, 16(6), 807–820. <https://doi.org/gp65td>
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. <https://doi.org/f34d7c>
- Sommer, C., Strähle, C., Köthe, U., & Hamprecht, F. A. (2011). Ilastik: Interactive learning and segmentation toolkit. *Eighth IEEE International Symposium on Biomedical Imaging (ISBI 2011). Proceedings*, 230–233. <https://doi.org/c44gd9>
- Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods*, 18(1), 100–106. <https://doi.org/ghrgms>