A deep learning-based all-in-one image analysis tool for cell quantification: Proof-of-concept using confocal images of immunostained cells on ECMO membranes

Z. Görmez[1], F. Hoeren[2,3], M. Richter[3], A. Krause[1], K.Troidl[1,2]

[1] Technische Hochschule Bingen, Bingen am Rhein, Germany; [2] Max Planck Institute for Heart and Lung Research / Dept. of Pharmacology / Bad Nauheim, Germany; [3]Justus-Liebig-Universität Giessen, Giessen, Germany

Introduction: In biomedical research, the segmentation and enumeration of immunostained cells imaged by confocal microscopy are challenging since clusters of overlapping cells and different sizes of nuclei aggravate the analysis. Due to the need of accurate cell counting in a large number of samples and to the time-consuming nature of quantification processes by hand, methods based on deep learning algorithms have become the state of the art. Nevertheless, a central problem with these methods is that they involve many sub-processes (pre-, main-, post-processing) that are difficult to apply and modulate by scientists that are inexperienced in programming. In our project, human cells that have adhered to oxygenator-membranes after Extracorporeal Membrane Oxygenation (ECMO) long-term support in patients have been stained and subjected to microscopy. A large data set of 3D images has been acquired. ECMO is a widely used treatment for patients with isolated or combined lung and/or heart failure. During ECMO treatment, venous blood is drained from the patient's body, oxygenated in the membrane oxygenator, and returned to the venous or arterial system with the help of a centrifugal pump. As a consequence of the blood-surface-interaction unwanted adhesion of cells to the foreign body material of the oxygenator occurs. The aim of this project is the quantification of the total number of cells at certain predilection points on the oxygenator membranes and it is the first step to gain information about the cell-surface-interaction during ECMO support.

Methods: To investigate the cell composition, membranes were examined for cell colonization after long-term patient support (>24h) using nuclear staining via DAPI (4′,6-diamidino-2-phenylindole). 240 image sequences were acquired by confocal laser microscopy, which automatically traverses a stacked volume in the Z-axis at different sites of the membranes, taking images of ten individual Z-planes at equally spaced intervals. This results in an image sequence that represents the spatial arrangement of the cells in the particular deposition site and is saved as a “.lif” file.

To quantify these cell depositions in certain areas and identify predilection sites for increased cell adherence, we established an automated deep learning-based all-in-one image processing pipeline and analyzed all image sequences of the examined ECMO-membranes as a proof of concept. The tool is written in Python, and since the targeted user group has no experience with command line usage, all necessary methods were combined in a ready-to-use graphical user interface (GUI).

Results: To achieve an optimal image analysis we added different features into the tool (Figure 1). They contain various adjustable parameters so that the user can adapt them to his own needs. Sub-processes of the tool produce not only the number of cells but also outlined images with cell nucleus borders (Figure 2) to provide that the users examine the results. All sub-processes were tested independently:

i) Batch-processing: This feature allows the selection of a directory or multiple images and starts the batch processing of all selected image sequences. In this step, the user can also divide the images into two groups to apply statistical analysis.

ii) Z-projection: Merging all ten images into one max-projection results in a crowded projection and overlapping of cells. To address this issue the feature tests different partial stacking options and provides optimal image stacking.

iii) Segmentation**:** Cellpose[1], a deep learning-based state-of-the-art method, was used as an established cellular segmentation method. It was chosen because of its strong ability to generalize which means that it can segment biological pictures out of the box without being trained on the specific dataset.

iv) Double Counting Analysis**:** Since partial stacking projections may result in unwanted double-counting of nuclei, we developed a method based on colocalization analysis to detect and eliminate those.

v) Elimination Processes: In order to avoid the detection of experimental artifacts as well as cells that can‘t be registered by the human eye because of their low intensities, a rule-based elimination process was performed.

vi) Cell Block Analysis: Detecting single nuclei borders by an algorithm is a common problem in tissue experiments due to accumulated cells, e.g. a cluster would be counted as one cell nucleus instead of the actual present amount of cell nuclei. To define this type of cell cluster, we developed a rule-based analysis.

vii) Statistics and Visualization: Finally, the tool allows statistical analysis and data visualization. Here, the Mann-Whitney-U-test was performed after the cell

quantification process to compare different predilection sites for increased cell deposition with each other. To visualize the statistical data box plots were performed (Figure 3). The new statistical tests and visualizations can easily be adapted to the needs of the user.

Conclusion: Our GUI-based tool addresses all issues by covering all processes of the analysis from segmentation to statistical testing. It is an easy way to provide the power of artificial neural networks to programming-untrained staff.

In future work, cell-type-specific immunostaining can be applied to characterize adherent cells in different predilection sites for increased cell deposition. The next goal after the characterization is to define the role of different cell types in systemic inflammation during ECMO therapy.

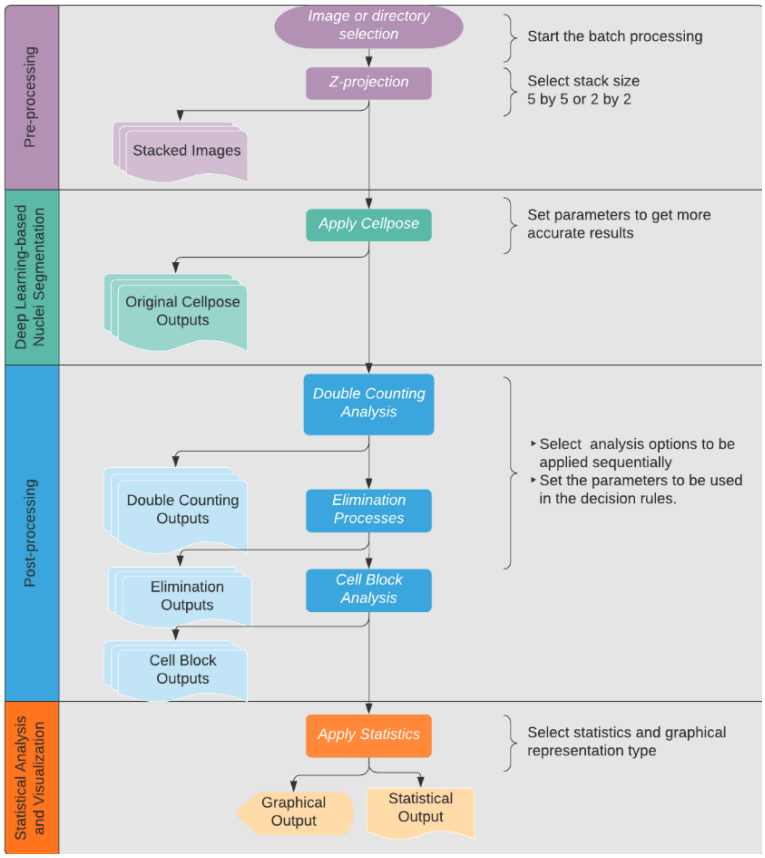

Figure 1: Overall flowchart representing the image analysis pipeline implemented in our tool. The first section (Pre-processing) includes batch processing and z-projection. The outputs of this section are partially stacked images used directly as input of Cellpose. The second section is the main segmentation process performed via Cellpose. The outputs of Cellpose are i) a mask list of an image which contains all detected cell nuclei, ii) the outlined image with cell nuclei borders. The third section (Post-processing) contains different sub-processes to achieve more accurate results. The outputs of each sub-process are similar to the outputs of the previous section: an edited mask list and the outlined image with nucleus borders colored in different colors (Figure 2). The last section is the statistical analysis that allows users to prove their hypothesis without using any other tool and transferring the data.

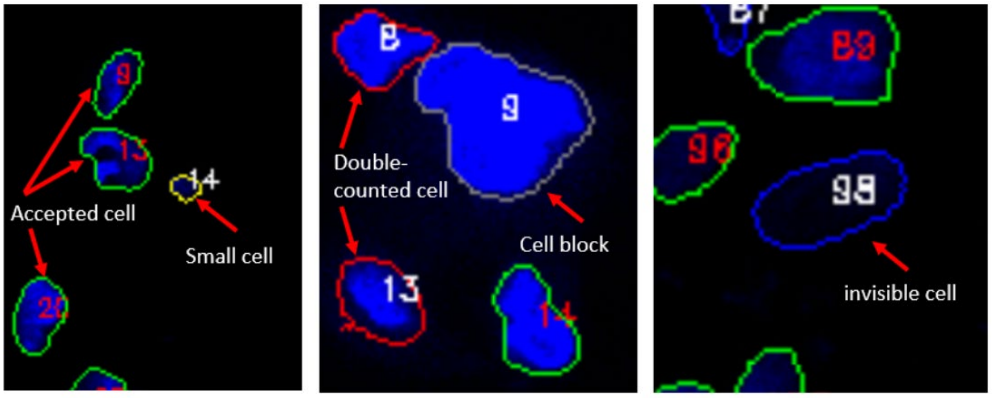

Figure 2: Outlined images with cell nucleus borders colored in different colors. Green outline indicates accepted cell nuclei that passed through all post-processing steps. Red outline represents double-counted cells with respect to the previous or next projection. Yellow represents a small cell nucleus. Blue was used for invisible cell nuclei and gray represents a cell nucleus block. Small, double-counted, and invisible cells are removed from the mask list of the related image and are not considered in the next step.

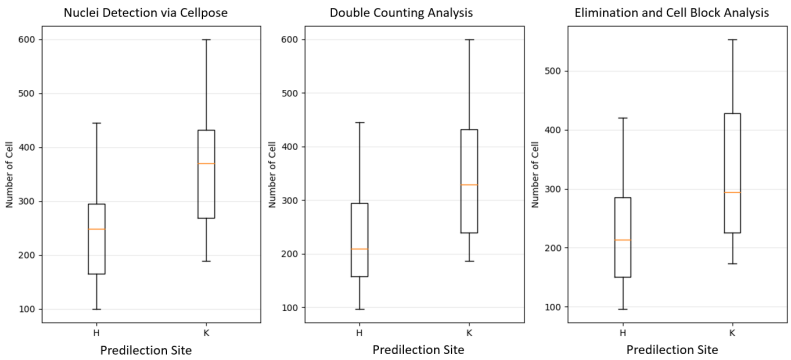

Figure 3: Box-plots showing the distribution of the total number of cells at certain predilection points after each process and the results of the Mann-Whitney-U-test indicate that they are significantly different.

References

1. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **18**, (2021).