RWTH-GE (2)

Authors: Dennis Eschweiler, Johannes Stegmaier

Email: Johannes.stegmaier@lfb.rwth-aachen.de

Platform: Windows 10 (x64)

Prerequisites: MATLAB 2018b (x64)

RWTH-GE (2): SUMMARY

The presented algorithm for segmentation and tracking follows a three-step approach where we detect, track and finally segment nuclei. In the preprocessing phase, we detect centroids of the cell nuclei using a Laplacian-of-Gaussian Scale Space Maximum Projection approach. Tracking was performed in a backward fashion on the predicted seed points, i.e., starting at the last frame and sequentially connecting corresponding objects until the first frame was reached. Correspondences were identified by propagating detections of a frame t to its preceding frame t - 1 and by combining redundant detections using a hierarchical clustering approach. The tracked centroids were then used as input to a seeded watershed algorithm to obtain the final segmentation.

RWTH-GE (2): PREPROCESSING

For **Fluo-N3DH-CE**, **Fluo-N3DL-TRIC**, and **Fluo-N3DL-DRO**, we used our nucleus detection algorithm [1]. In brief, the original images were filtered using differently scaled Laplacian-of-Gaussian filters, where the selected scales, σ_{min} and σ_{max}, were matched to the observed cell sizes. The 4D scale-space was reduced to a 3D image by a maximum intensity projection of the individual LoG-filtered images. Local maxima were then identified in the 3D LoG ScaleSpace Maximum Projection (LoGSSMP) and we additionally allowed to detect intensity plateaus to reduce the number of false negative detections in cases where no single maximum pixel was present in the center of a nucleus. To reduce false positive detections in background regions, only detections with an intensity larger than the global mean plus two standard deviations of the LoGSSMP image intensities were considered. Moreover, the raw images were preprocessed using a 3D Gaussian filter with the standard deviation σ_{gauss}, a 3D median filter with the neighborhood radius r_{median} (only the first 50 frames of **Fluo-N3DH-CE**), or a grayscale 3D morphological erosion with a disc structuring element of a radius $r_{erosion}$. Redundant detections on intensity plateaus were combined in the subsequent tracking phase.

RWTH-GE (2): SEGMENTATION

Segmentation was performed as the final step of the pipeline. See *RWTH-GE (2): POST-PROCESSING*.

RWTH-GE (2): TRACKING

Tracking was performed in a backward fashion by sequentially linking corresponding objects until the first frame was reached. Correspondences were identified by assigning a new tracking label to all unlabeled detections at frame t and by copying the labeled detections to the preceding frame t - 1. At frame t - 1 we then performed a hierarchical clustering using Ward's linkage criterion and identify clusters that contain detections from time point t and t - 1. The distance-based cut-off is determined based on the average spatial distance of each detection to a set of nearest neighbors (large in less dense frames, smaller in densely occupied frames). We empirically chose to use half the average distance to the 3rd to 8th nearest neighbors of each detection, i.e., excluding potential redundant detections or division events from the computations. There are multiple cases of label presence in the clusters: (1) if a cluster contains no labeled detection, a new track label is introduced, (2) if one of the detections has a valid tracking label assigned, the label is copied for the cluster, (3) if two or more detections in the cluster have a valid tracking label assigned, a merge (i.e., a cell division) is introduced by adding a new tracking label and by correctly assigning the predecessor and successor links, and (4) if only one labeled detection is contained, the track should end. The detections contained in a cluster are averaged and form the set of tracked objects at frame t - 1. The steps above are repeated until the first frame of the sequence is reached. In principle, this tracking approach is similar to a nearest neighbor tracking but additionally handles redundant detections by clustering nearby seed points. Moreover, cell divisions are naturally included in the algorithm design, as daughter cells are clustered together if their spatial distance falls below the cluster cut-off value. A drawback of the clustering-based approach, however, is the required global cut-off value that is used for the entire dataset. This unavoidably leads to fragmented tracks if the datasets exhibit large density variations (e.g., **Fluo-N3DL-TRIC**). For **Fluo-N3DL-TRIC** and **Fluo-N3DL-DRO** with varying densities and where only a subset of the cells should be tracked, we thus performed an additional nearest neighbor tracking in the forward direction. Selected tracks that ended earlier than the last frame were linked to their closest neighbors in subsequent frames if the respective detections were not yet occupied by another selected trajectory.

RWTH-GE (2): POST-PROCESSING

The segmentation was largely based on a seeded watershed technique [2] with a few improvements to tune the results. The tracking labels were used to generate seed images with positions and labels identical to the tracked centroids. We analyzed small crops surrounding each of the detections in parallel fashion, similar to the TWANG segmentation approach described in [1]. Initially, we determine a binary threshold using Otsu's method separately for each of the cropped regions. However, instead of directly binarizing the small region with the identified threshold, we use a connected threshold image filter implemented in the Insight Toolkit (*itkConnectedThresholdImageFilter*). This filter is initialized with a seed point (the estimated centroid of the current nucleus) and then performs a region growing by extending to neighboring voxels if their intensity is above the threshold identified by the initial Otsu threshold. This step yields a single connected component that is connected to the seed point that was used for the initialization. As the binary connected component can potentially span over multiple neighboring nuclei, we perform an additional seeded watershed for clump-splitting on each of the regions by adding a seed in the center of the current nucleus of interest and by adding an artificial background marker on the image boundaries of the current region of interest. Thus, the center nucleus should be well separated from touching nuclei and background signal. The region size is set to the cut-off value used for the hierarchical clustering during tracking, i.e., it is variable and depends on the density of the cells contained in each frame. The obtained segmentations for each individual cell are then combined to the resulting full resolution segmentation image. For **Fluo-N3DL-TRIC**, we performed a maximum intensity projection of the 3D segmentation results and then resized the image in z to obtain the original number of slices. While this workaround does not necessarily provide good segmentation results in a biological sense, this was done to increase the chance of matching the slice-based annotations of the ground truth as 3D segments were systematically underestimated in the z direction and thus frequently ended prior to the slices of the ground truth annotations. The final segmentation images were double-checked with the tracking results and in cases where the segmentation algorithm erroneously missed a cell (e.g., if the background label flooded a cell region), we manually added the detections again to provide segmentation images that are consistent with the tracking results.

REFERENCES

1. Stegmaier J, Otte J, Kobitski A, Bartschat A, Garcia A, Nienhaus GU, Strähle U, Mikut R. Fast segmentation of stained nuclei in terabyte-scale, time resolved 3D microscopy image stacks. *PLoS ONE* **9,** e90036 (2014).

2. Beare R, Lehmann G. The watershed transform in ITK – discussion and new developments. *The Insight Journal* **6** (2006).