RWTH Cell Linking Challenge Submission 30.11.2024

Author: Zhu Chen, Johannes Stegmaier
Email: zhu.chen@lfb.rwth-aachen.de
Platform: Linux
Prerequisites: python 3.10

Summary

The proposed algorithm is designed to link pre-segmented masks across adjacent time frames and generate new masks when a cell's mask is missing in an intermediate time frame.
For 2D datasets, the algorithm works as a zero-shot method, using the pretrained Segment-Anything 2 (SAM2) [1] model without requiring additional training or fine-tuning.
For 3D datasets, there are two operational modes:

1. Linking-only mode: ideal for datasets with a nearly complete set of masks, this mode is training-free and computationally efficient, only focusing on linking existing masks across time frames.

2. Mask generation mode: designed for datasets with a significant number of missing masks, this mode uses a fine-tuned SAM-Med3D [2] model, utilizing few-shot learning to generate the missing 3D masks for intermediate time frames.

Preprocessing

The algorithm begins from the last time frame and iteratively moves backward to the first time frame. It assumes that the set of pre-segmented masks is complete in the last time frame, which is defined as the initialization of the tracking lineages.
For each lineage tracked, the algorithm extracts a local patch surrounding the known mask in time frame t and the corresponding region in time frame $t-1$ with a side length of L_{window}, where the linked mask is to be identified. As preprocessing, the intensity values of the extracted image patches are normalized to the range of (0,255) using min-max normalization.
Additional preprocessing steps are applied to specific datasets:

1. Gamma transformation: Applied to the datasets Fluo-N2DL-HeLa, Fluo-N2DH-GOWT1, Fluo-N2DH-SIM+, and Fluo-N3DH-SIM+ to improve contrast. The γ value can be found in the parameter table.

2. Median Filtering: To reduce noise, 2D median filters are applied to Fluo-N2DH-GOWT1 and Fluo-N2DH-SIM+, while 3D median filters are used for Fluo-N3DH-SIM+. The filter size k_{median} is listed in the parameter table.

Linking

Once the masks of all cells are detected in a given time frame t, the algorithm begins linking by identifying the corresponding masks in the previous time frame $t-1$. For each mask in t, an image patch surrounding the mask is cropped, along with a corresponding patch at the same coordinates from time frame $t-1$.
The pair of image patches, I_t and I_{t-1}, is provided as input to SAM2 as a video sequence. Using the known mask in I_t, the algorithm generates prompts for SAM2, which include:

1. Bounding box prompts: defined based on the cell boundary.

2. Point prompts: a set of randomly selected points from the foreground and background areas of the mask.

With the prompts applied to I_t, SAM2 predicts the mask of the same cell within the corresponding patch I_{t-1}. This process effectively identifies the movement and potential mask of the cell in time frame $t-1$.
Once the predicted mask in I_{t-1} is obtained, the algorithm calculates its centroid coordinates and searches for a matching mask in the existing set of masks for $t-1$. If a matching mask is found, it is added to the lineage, and the masks are linked across the two time frames. For 3D masks, since SAM2 only works with 2D images, the algorithm selects the slice along the z-axis with the largest mask area to crop the image patches, so that the linking process is applied to the most representative slice of the mask.
Once SAM2 predicts the mask in time frame $t-1$, there are two different cases:

- Existing mask found: if a pre-segmented mask in $t-1$ matches the predicted mask, it is directly linked to the known mask in t.
- No existing mask: If no matching mask exists in the pre-segmented mask set of $t-1$, two possibilities are considered:
  - Disappearing cell: if the predicted mask is smaller than a predefined threshold $A_{\text{disappear}}$ or its centroid is located very close to the image boundary, the cell is classified as "disappearing." This indicates that t is the **first time frame** of the lineage due to the backward tracking approach. Consequently, tracking for this lineage is terminated in earlier time frames.
  - Missing mask: if the predicted mask is valid, it is assumed that the mask is missing from the pre-segmented mask set. In this case, the mask predicted by SAM2 is added to the lineage.

For 3D datasets, pre-segmented masks are linked based on the SAM2 prediction of the slice along the z-axis with the largest mask area and a predefined threshold d_{neighbor}. In linking-only mode, if no existing mask from the pre-segmented set can be linked to the predicted 2D mask, the tracking of the current lineage is terminated. In mask-generation mode, when no existing mask is linked, a new 3D mask is generated using SAM-Med3D. The validity of the generated mask is checked using the same criteria applied for 2D datasets.

Post-Processing

After processing all lineages from t to $t-1$:

- Mitosis Detection: If two masks in t link to the same mask in $t-1$, this is identified as a mitosis event. If more than two masks link to the same mask (a scenario that would validate biological prior knowledge), only two links with shortest distances are kept, and the third is discarded.
- Unlinked Masks in $t-1$: For the remaining masks in the pre-segmented mask set of $t-1$ that are not linked to any mask in t, the mask value at their centroid in time frame t is checked to attempt a re-linking. The following scenarios are considered:
  - If a mask previously marked as "disappeared" is detected, it is re-linked, and the "disappeared" status is removed.
  - If the mask is re-linked to a merged cell, the mitosis detection is identified as erroneous. One of the merged cells is re-linked to the unlinked mask, and the incorrect mitosis detection is removed.

– If no re-linking is feasible, a new lineage is initiated starting from the unlinked mask in $t-1$.

References

[1] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, K. V. Alwala, N. Carion, C.-Y. Wu, R. Girshick, P. Dollár, and C. Feichtenhofer. Sam 2: Segment anything in images and videos. *arXiv:2408.00714*, 2024.

[2] H. Wang, S. Guo, J. Ye, Z. Deng, J. Cheng, T. Li, J. Chen, Y. Su, Z. Huang, Y. Shen, B. Fu, S. Zhang, J. He, and Y. Qiao. Sam-med3d. *arXiv:2310.15161*, 2023.