Cell Reports Methods, Volume *1*

Supplemental information

A deep learning-based segmentation pipeline for profiling cellular morphodynamics using multiple types of live cell microscopy

Junbong Jang, Chuangqi Wang, Xitong Zhang, Hee June Choi, Xiang Pan, Bolun Lin, Yudong Yu, Carly Whittle, Madison Ryan, Yenyu Chen, and Kwonmoo Lee

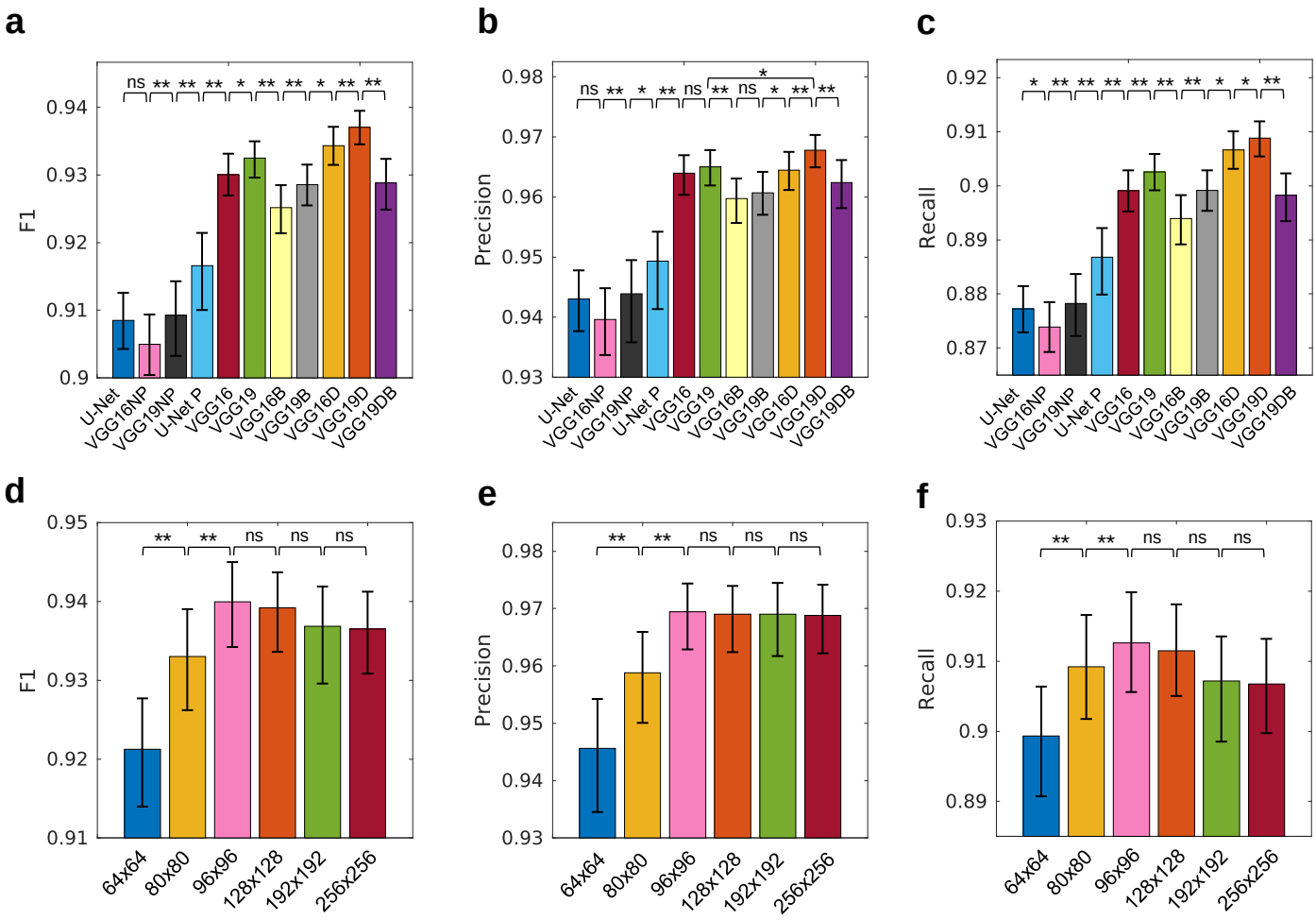

Figure S1. Effects of VGG variants and patch size on the models trained on phase contrast datasets, related to Figure 3. (**a-c**) Average performance of U-Net and its variants with VGG encoders, pretraining, batch normalization or dropout layers. Model names suffixed by “NP” or U-Net are not pretrained, but other models including U-Net P are ImageNet pretrained. Suffix B denotes batch normalization and D denotes Dropout. (**d-f**) Average performance of ImageNet pretrained VGG19D-U-Net trained on different sizes of two-dimensional cropped image patches, ranging from 64x64 to 256x256 pixels. (**a-f**) The tests of significance by two-sided Wilcoxon signed-rank test with p >= 0.05 are indicated by ns, p < 0.05 are indicated by * and p < 0.0001 are indicated by **. Only the statistical tests of differences between adjacent bars are shown except for (**b**) to compare VGG19D-U-Net with the next best model, VGG19-U-Net. Error bars: 95% confidence intervals of the bootstrap mean. The number of evaluated frames is n=202, which is roughly 40 frames from each phase contrast live cell movie.

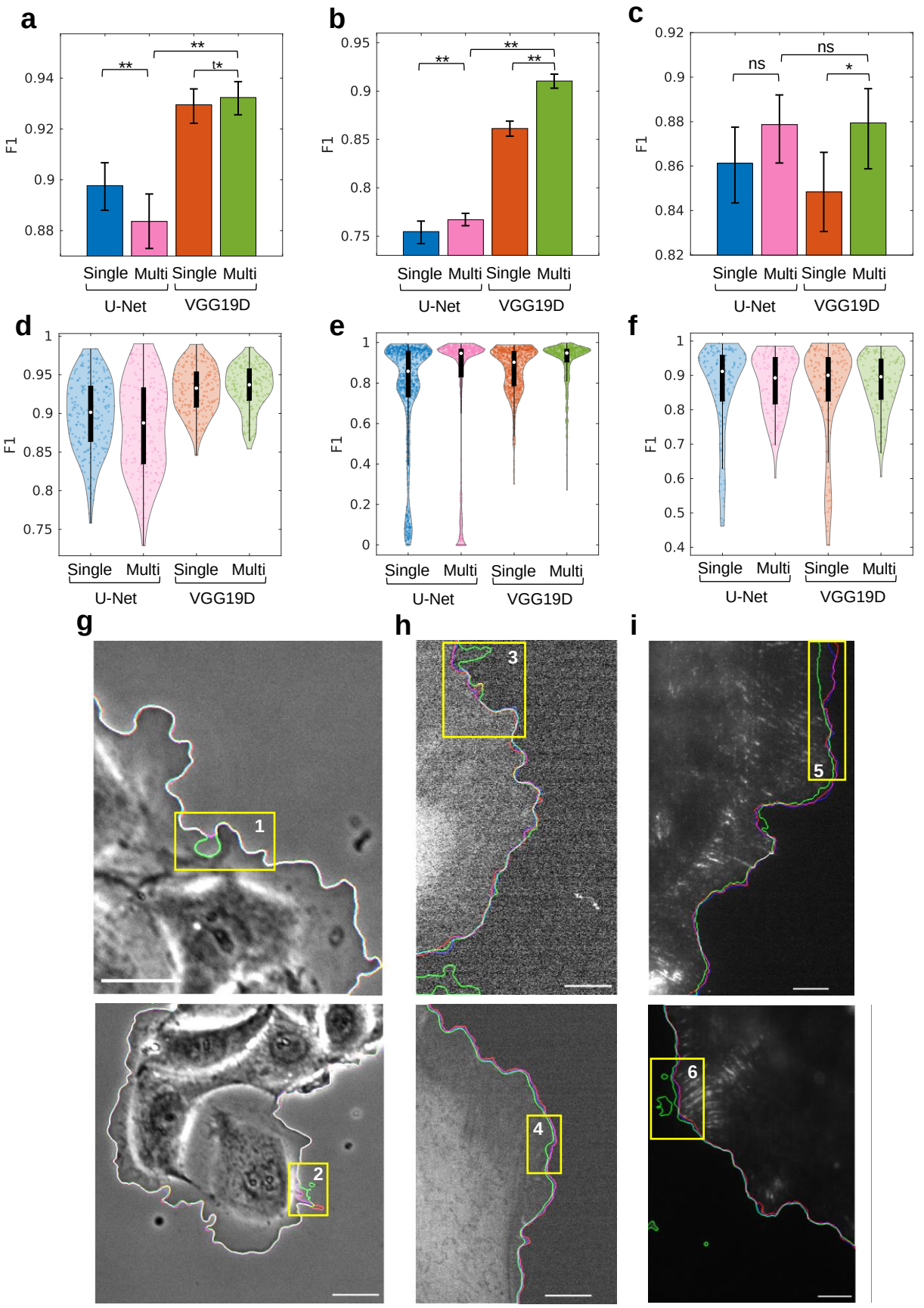
a
b
c
d
e
f
g
h
i
F1
Single
Multi
U-Net
VGG19D
ns
1
2
3
4
5
6
Ground Truth (GT)
Single-Microscopy (single)
Multiple-Microscopy (multi)
GT & multi overlap
GT & single overlap
single & multi overlap
GT & single & multi overlap

Figure S2. Comparisons of single-microscopy-type and multiple-microscopy-type training using U-Net and VGG19D-U-Net models, related to Figure 6. (**a-c**) Average F1 scores of models evaluated on phase contrast (**a**) , SDC (**b**) , and TIRF (**c**). (**d-f**) The distribution of F1 score in violin plot and box plot in black with a median indicated by the white circle for phase contrast (**d**) , SDC (**e**) , and TIRF (**f**) . (**a-f**) The numbers of evaluated frames are n=202 for phase contrast, n=1000 for SDC, and n=132 for TIRF. The statistical significance by two-sided Wilcoxon signed-rank test with p >= 0.05 are indicated by ns, p <0.05 are indicated by *, and p < 0.001 are indicated by **. The statistical significance by paired t-test with p <0.05 is indicated by t*. Error bars: 95% confidence intervals of the bootstrap mean. (**g-i**) Visualization of segmentation results on three microscopies: phase contrast (**g**) , SDC (**h**) , and TIRF (**i**) . Each yellow box with the inset number corresponds to the close-up views shown in **Fig. 6**. Edges extracted from ground truth masks and predictions are overlaid on the original image. Each edge is represented by one of three primary colors: red, green and blue. Overlap of two or more edges is represented by the combination of those colors. (**g**) Bar: 32.5 μm. (**h**) Bar: 7.2 μm. (**i**) Bar: 6.5 μm,

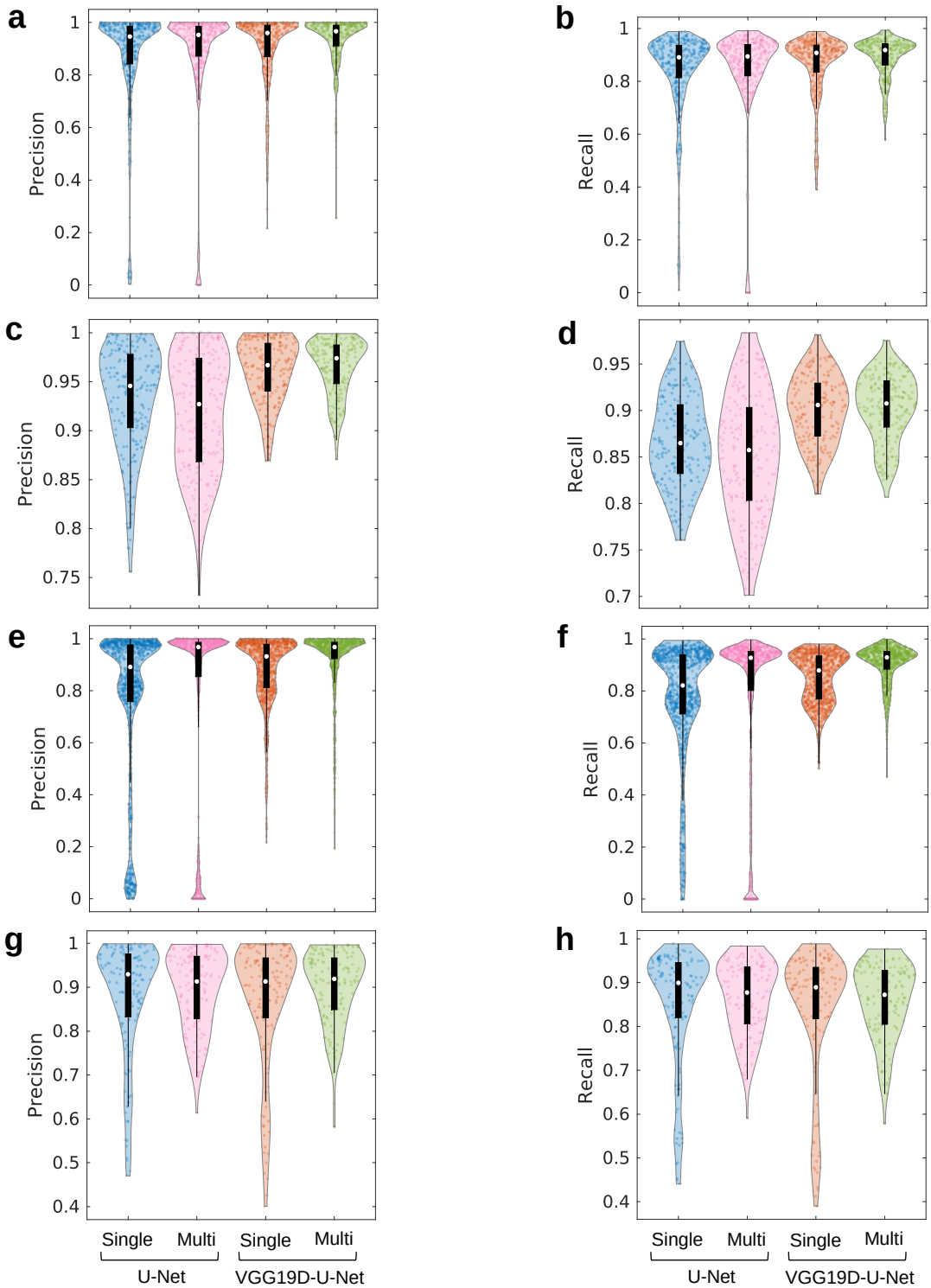

Figure S3. Comparisons of U-Net and VGG19D-U-Net models trained on single or multiple-microscopy-type dataset, related to Figure 6 (**a-h**) The distribution of precision and recall of U-Net and VGG19D-U-Net either trained on single-microscopy-type or multiple-microscopy-type datasets. The number of evaluated frames are (**a-b**) n=335 for dataset comprised of all microscopy types, (**c-d**) n=202 for Phase Contrast dataset, (**e-f**) n=1000 for SDC dataset, and (**g-h**) n=132 for TIRF dataset. Within each violin plot, there is a box plot in black with a median indicated by the white circle.

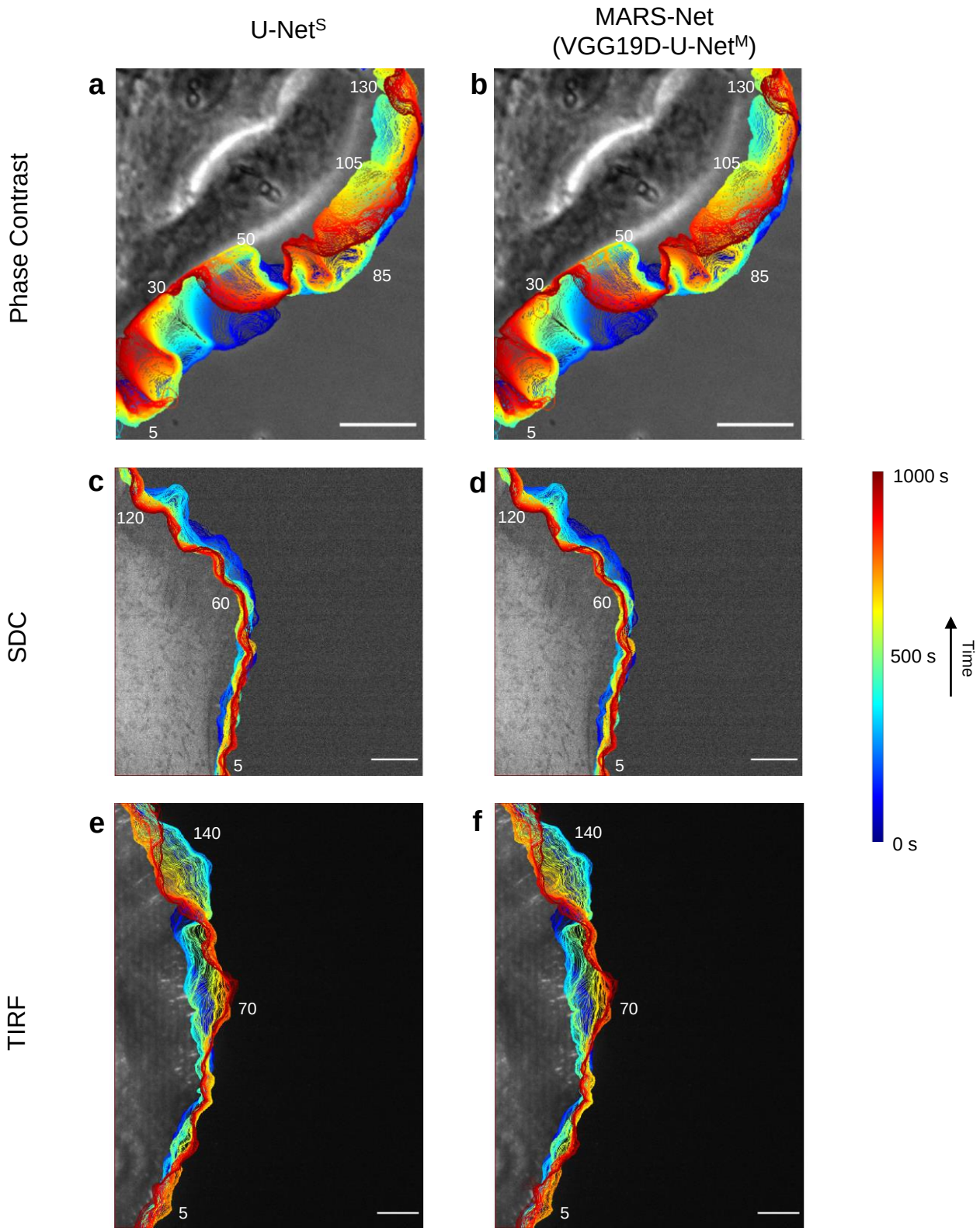

Figure S4. Visualization of single-microscopy-type U-Net and MARS-Net, related to Figure 6. (**a-f**) Progression of the cell edges overlaid on the first frame of the movie using segmentation results from U-Net^S (**a,c,e**) and MARS-Net (**b,d,f**). (**a-b**) Phase contrast live cell movie, Bars: 32.5 μm. (**c-d**) SDC live cell movie, Bars: 7.2 μm. (**e-f**) TIRF live cell movie, Bars: 6.5 μm. The numbers corresponding to some of the window numbers in protrusion maps in **Fig. 7** are written in white text color on the image. (blue, 0s; red, 1000s time points).

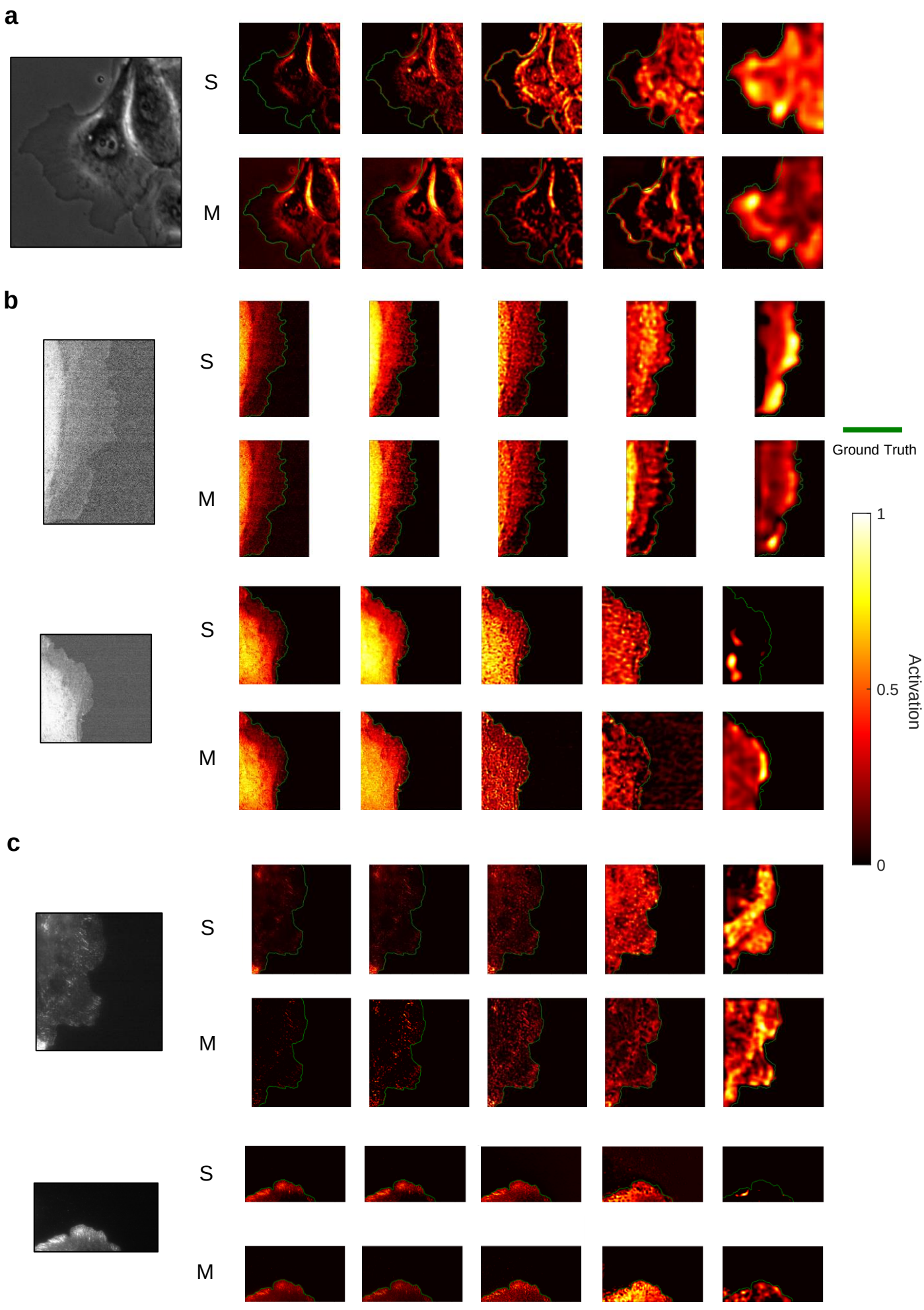
a
S
M
b
S
M
S
M
Ground Truth
1
0.5
0
Activation
c
S
M
S
M

Figure S5. Class activation map of convolutional layers in the single-microscopy-type and multiple-microscopy-type VGG19D-U-Net with respect to the ground truth edges, related to Figure 6. One frame was randomly chosen from all movies of each microscopy type: **(a)** Phase contrast, **(b)** SDC and **(c)** TIRF**.** On the left, the original image of the corresponding frame is shown. The word 'S' represents the single-microscopy-type dataset, and 'M' represents the multiple-microscopy-type dataset. The class activation maps from the first to the fifth block in the encoder are organized in order from left to right. In the heatmap, the value 0 means no activation, and 1 means the highest activation. The green lines represent the ground truth edges.

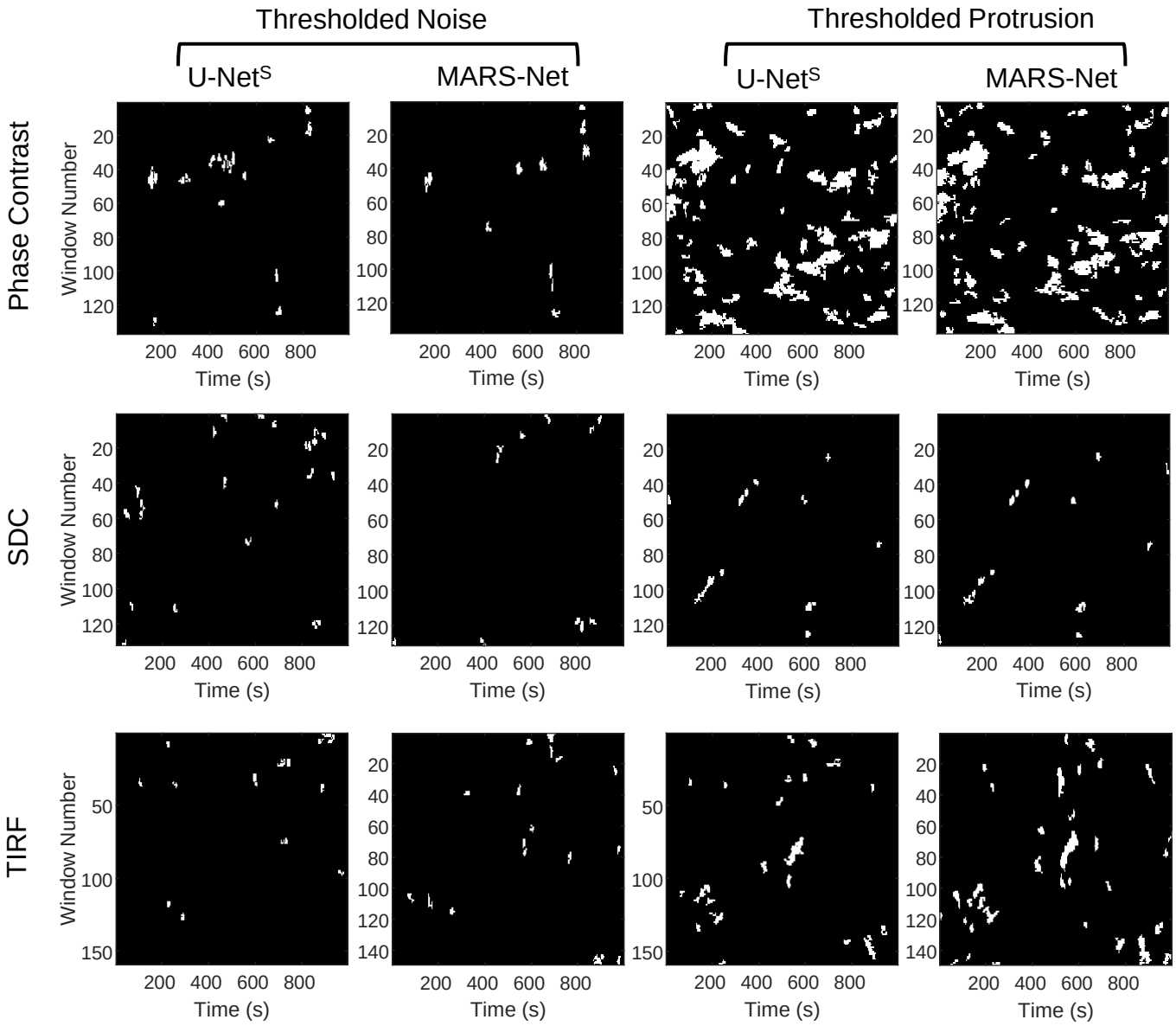

Figure S6. Threshold noise and protrusion in the protrusion maps of Phase Contrast, SDC, and TIRF live cell movies segmented by single-microscopy-type U-Net and MARS-Net, related to Figure 7. These are obtained by filtering and thresholding protrusion maps in **Fig. 7**. In each row, the same cropped region in the movie is segmented by both models and their segmentation are profiled for cellular morphodynamics.