Marker and Motion Guided Deep Networks for Cell Segmentation and Detection Using Weakly Supervised Microscopy Data

Gani Rahmon, Kannappan Palaniappan, Imad Eddine Toubal, DDW Cornelison

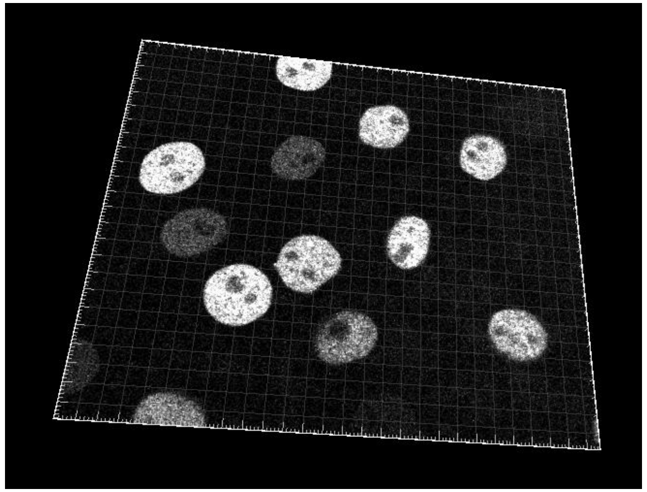
University of Missouri-Columbia, MO, USA

AIPR 2023

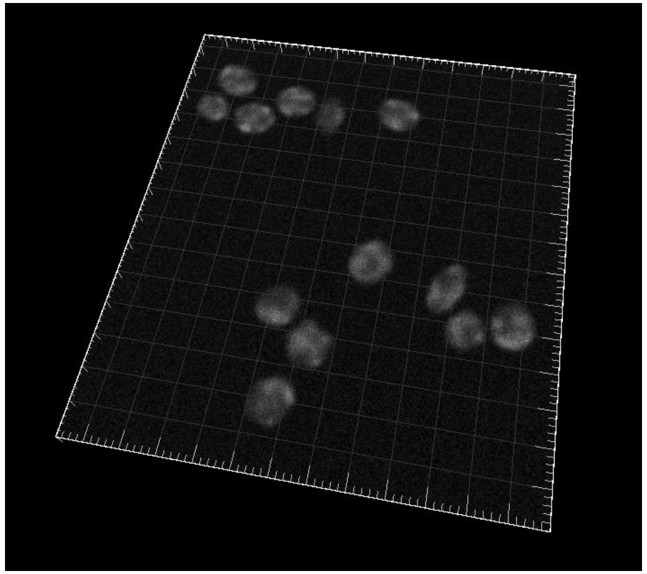
Introduction

Cell segmentation:

- Segmenting, detecting, and tracking moving cells in time-lapse sequences is a challenging task
- Required for many applications in both scientific and industrial settings
- Properly characterizing how cells
  - change their shapes
  - move as they interact with their surrounding environment
  - key to understanding the mechanobiology of cell migration
  - its multiple implications
    - normal tissue development
    - many diseases



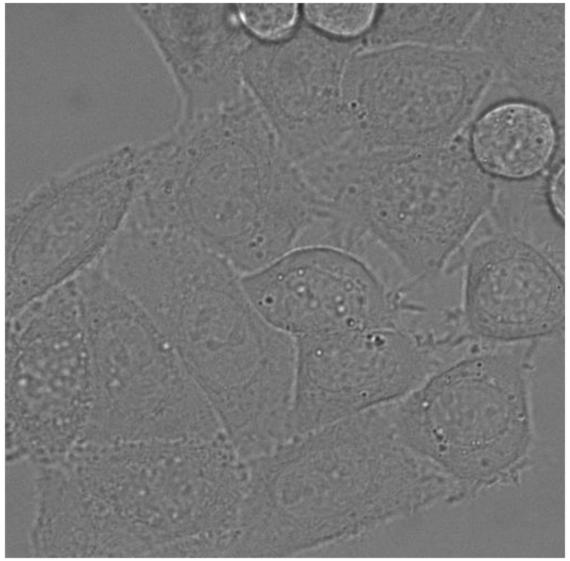
Fluo-N2DH-GOWT1 (REAL)



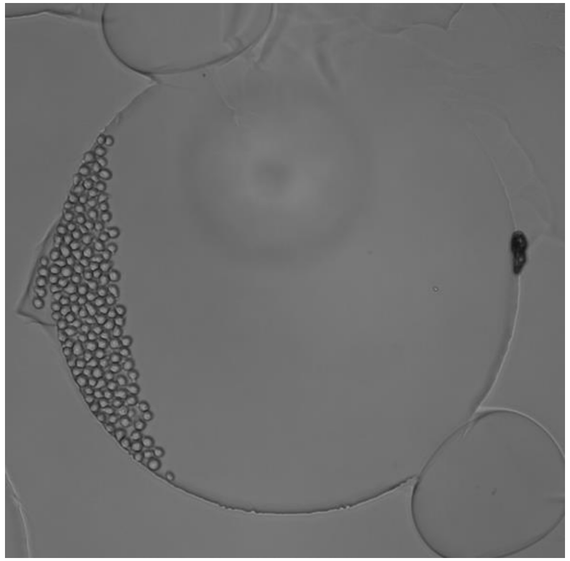
Fluo-N2DH-SIM+ (Simulated)

Cell Segmentation and Detection Challenges

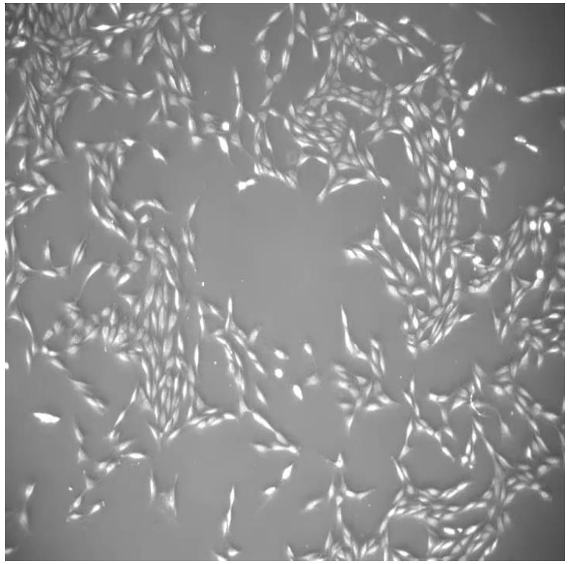
- Difficult to separate clustered cells
- Variation in cell shape
- High cell density
- Difference in imaging modalities
- Similarity with background
- Low contrast
- Mitosis



DIC-C2DH-HeLa



BF-C2DL-HSC

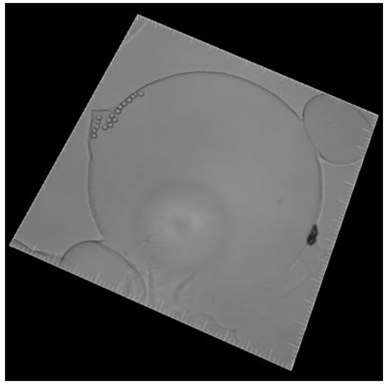


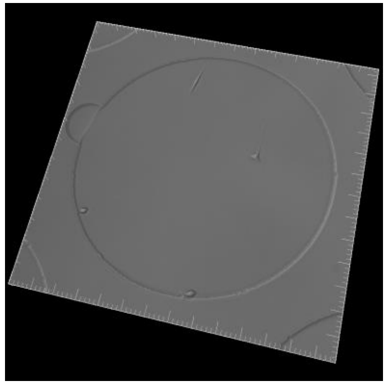
PhC-C2DL-PSC

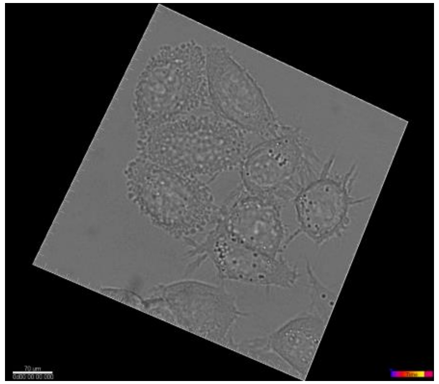
Fluo-N2DL-HeLa

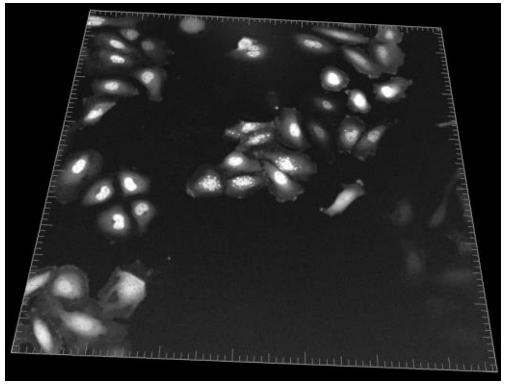
Dataset: 2D Subset of Cell Tracking Challenge [1]

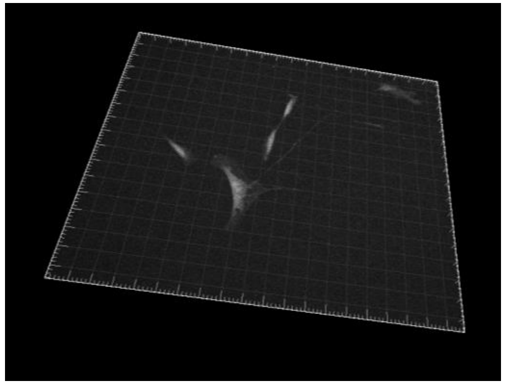
- Number of images with **Weakly Supervised Labels** provided is shown in brackets for each dataset

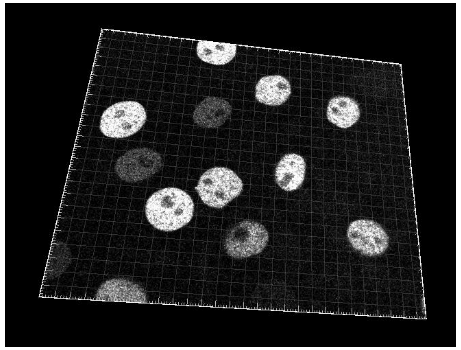

BF-C2DL-HSC (3528)

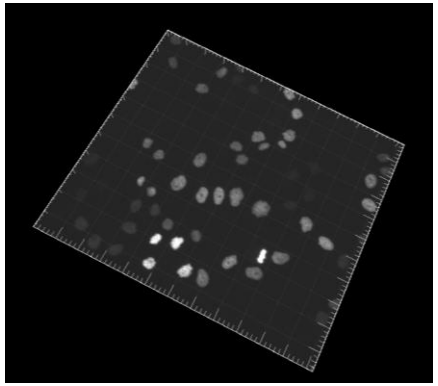

BF-C2DL-MuSC (2752)

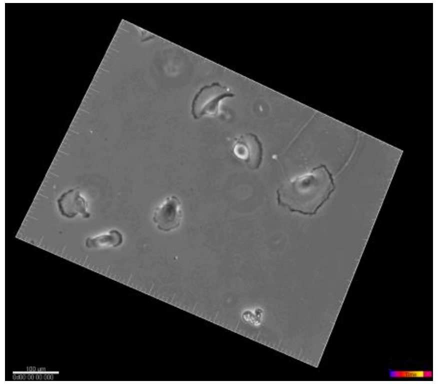

DIC-C2DH-HeLa (168)

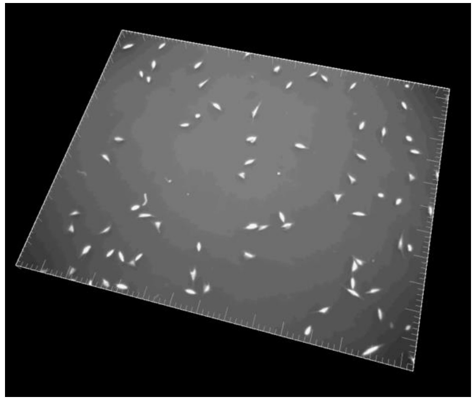

Fluo-C2DL-Huh7 (13)

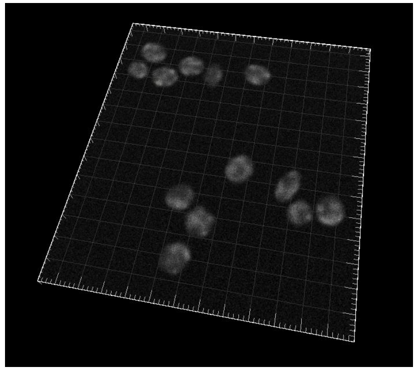

Fluo-C2DL-MSC (96)


Fluo-N2DH-GOWT1 (184)


Fluo-N2DL-HeLa (184)


PhC-C2DH-U373 (230)


PhC-C2DL-PSC (600)

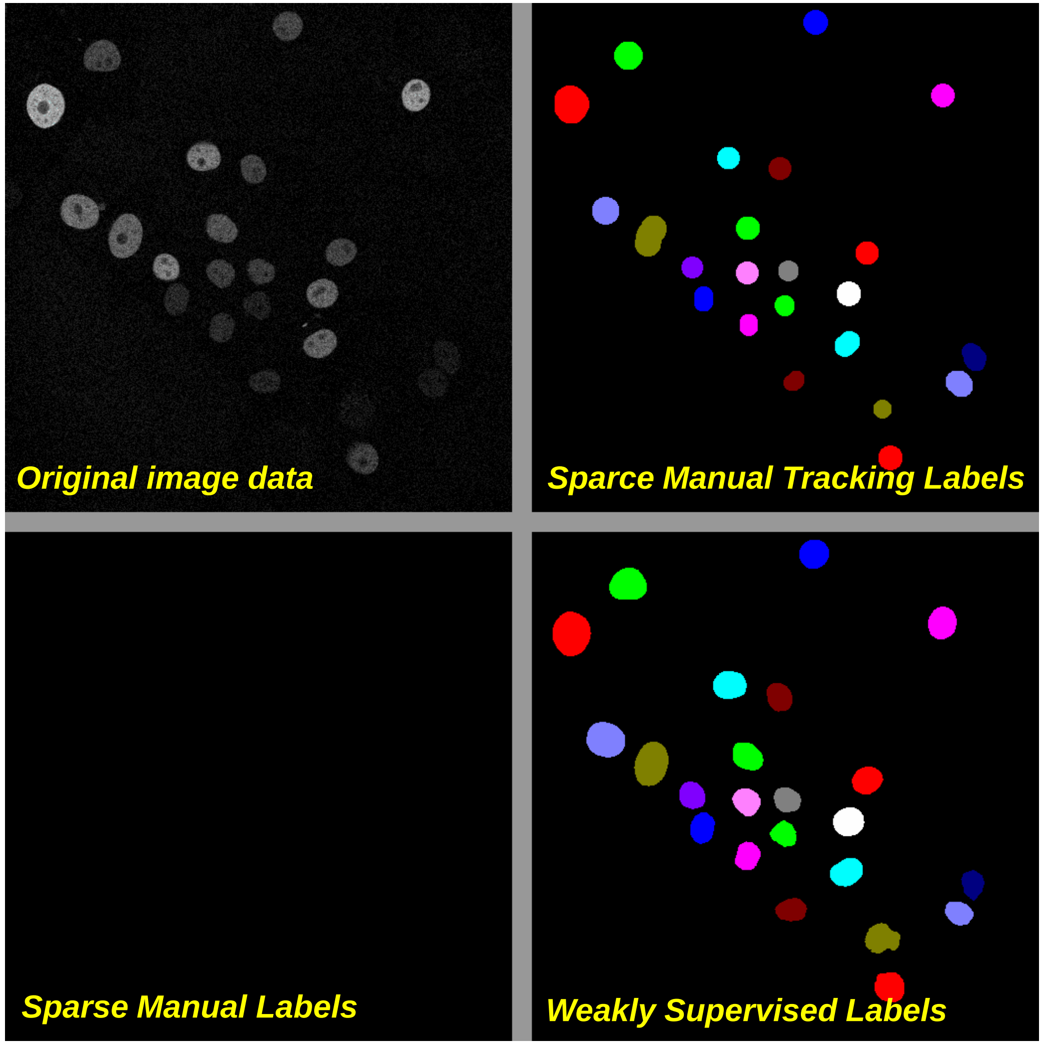

Fluo-N2DH-SIM+ (215)

[1] http://celltrackingchallenge.net/

Training Dataset Using Weakly Supervised Labels

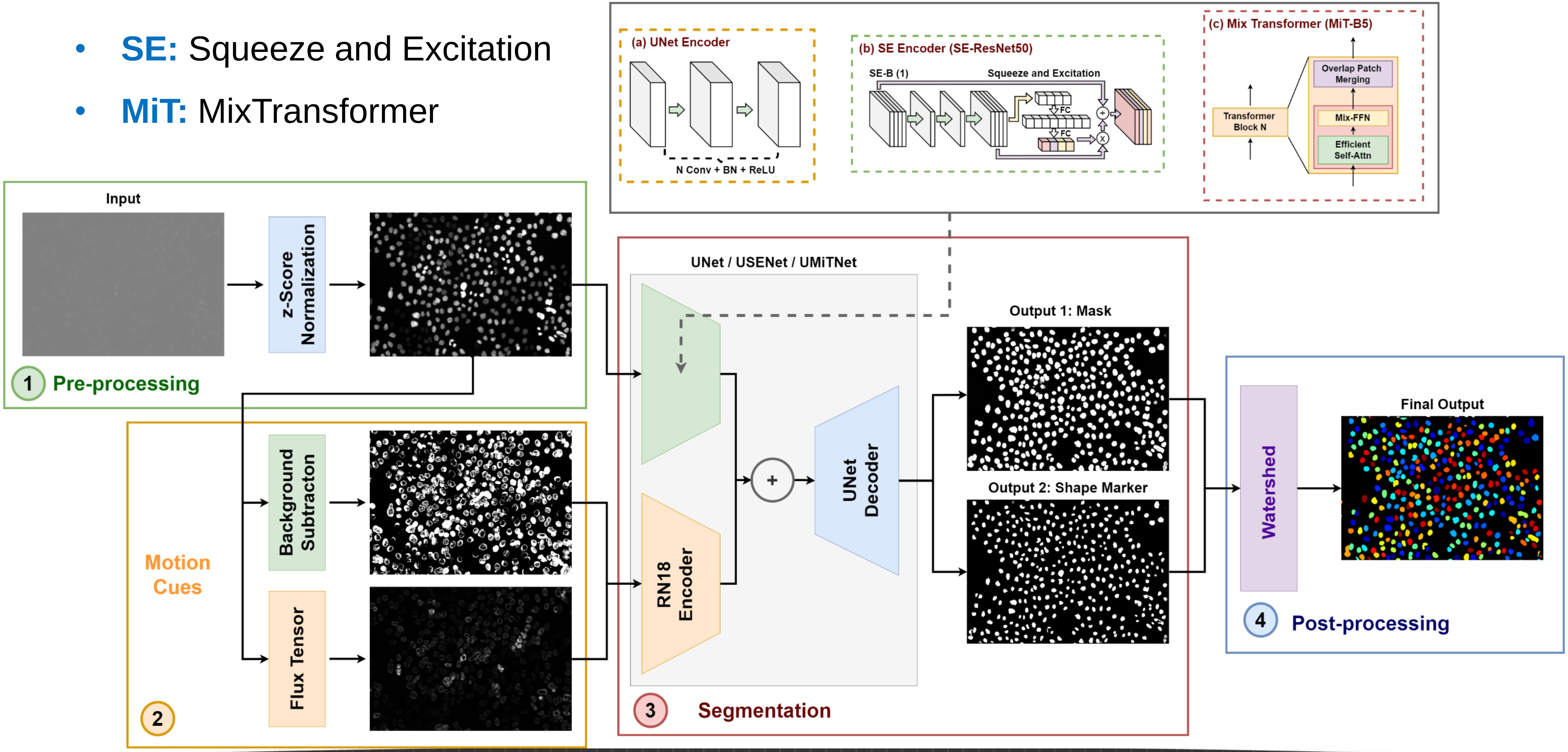
- Annotations based on their origin:
  - **Ground Truth** containing exact reference annotations, available only for simulated datasets;
  - **Sparse Manual Ground Truth (gold truth)** containing human-made reference annotations, obtained as a consensual or majority opinion of several human experts;
  - **Weak Supervision (Silver Truth)** containing computer-generated reference annotations **(noisy labels),** obtained as the majority opinion over the results of several automatic analysis methods submitted by former challenge participants.

- **Question:** How well does learning with weak supervision work for cell video segmentation?

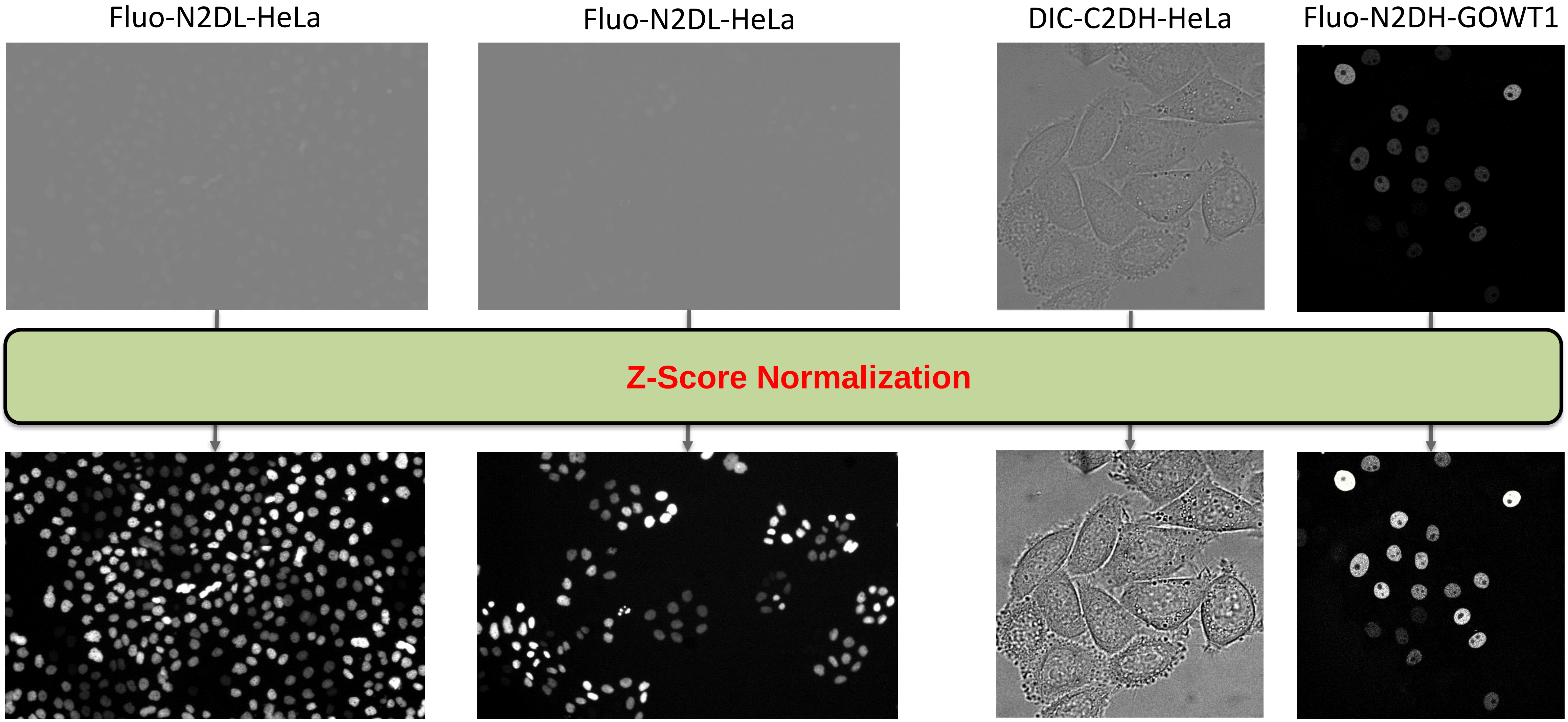


Pipeline: Deep Segmentation using UNet, USENet, UMiTNet

- **SE:** Squeeze and Excitation
- **MiT:** MixTransformer



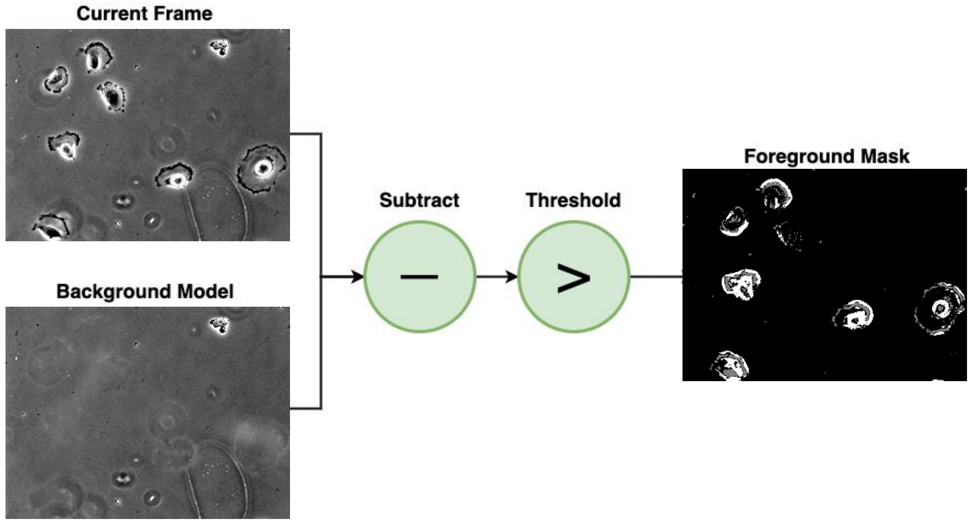
Pre-processing: Z-Score Normalization



Motion Cues: Background Subtraction & Flux Tensor

1. Multi-modal background subtraction (BGS) for change estimation

- <u>Change</u>: estimated using an <u>adaptive multi-modal background subtraction approach</u> (OpenCV: BackgroundSubtractorMOG2 [2]).

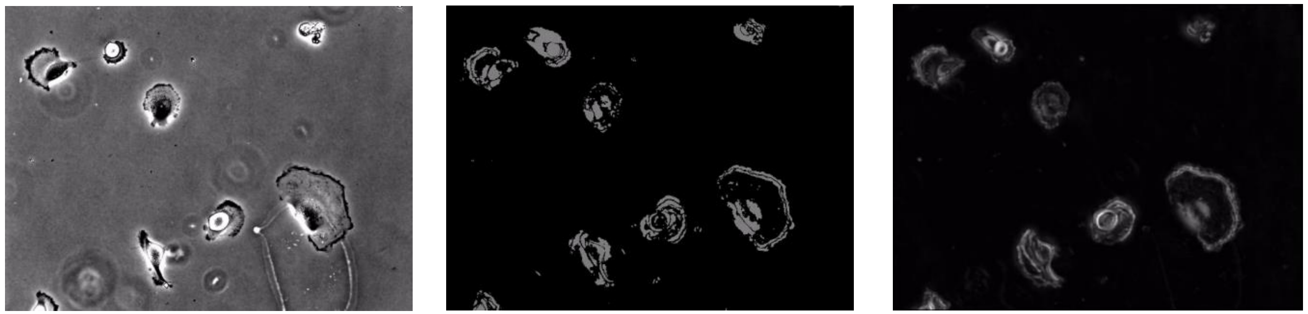


2. Flux tensor motion estimation

- **Flux tensor**: temporal variation of the optical flow field within local 3D spatio-temporal volume

$$J_F = \begin{bmatrix} \int_\Omega \left\{ \frac{d^2 I}{dx dt} \right\}^2 dy & \int_\Omega \frac{d^2 I}{dx dt}\frac{d^2 I}{dy dt} dy & \int_\Omega \frac{d^2 I}{dx dt}\frac{d^2 I}{dt^2} dy \\ \int_\Omega \frac{d^2 I}{dy dt}\frac{d^2 I}{dx dt} dy & \int_\Omega \left\{ \frac{d^2 I}{dy dt} \right\}^2 dy & \int_\Omega \frac{d^2 I}{dy dt}\frac{d^2 I}{dt^2} dy \\ \int_\Omega \frac{d^2 I}{dt^2}\frac{d^2 I}{dx dt} dy & \int_\Omega \frac{d^2 I}{dt^2}\frac{d^2 I}{dy dt} dy & \int_\Omega \left\{ \frac{d^2 I}{dt^2} \right\}^2 dy \end{bmatrix}$$

$$trace(J_F) = \int_\Omega || \frac{d}{dt} \nabla I ||^2 dy$$



Input Image BGS Flux

[2] *Opencv: How to use background subtraction methods*. [Online]. Available: https://docs.opencv.org/3.4/d1/dc5/ tutorial background subtraction.html.

Motion U-Net: Motion Cue

Flux tensor motion estimation

- **Flux tensor**: temporal variation of the optical flow field within local 3D spatio-temporal volume
- Minimization results in:

$$J_F(x,W) = \int_{\Omega} W(x,y)\,\frac{\partial}{\partial t}\nabla I(x).\frac{\partial}{\partial t}\nabla I^T(x)\,dy$$

$$\frac{\partial}{\partial t}\left(\frac{dI(x)}{dt}\right) = \frac{\partial^2 I(x)}{\partial x \partial t} v_x + \frac{\partial^2 I(x)}{\partial y \partial t} v_y + \frac{\partial^2 I(x)}{\partial t^2} v_t + \frac{\partial I(x)}{\partial x} a_x + \frac{\partial I(x)}{\partial y} a_y + \frac{\partial I(x)}{\partial t} a_t$$

$$J_F = \begin{bmatrix} \int_\Omega \left\{\frac{d^2 I}{dxdt}\right\}^2 dy & \int_\Omega \frac{d^2 I}{dxdt}\frac{d^2 I}{dydt} dy & \int_\Omega \frac{d^2 I}{dxdt}\frac{d^2 I}{dt^2} dy \\ \int_\Omega \frac{d^2 I}{dydt}\frac{d^2 I}{dxdt} dy & \int_\Omega \left\{\frac{d^2 I}{dydt}\right\}^2 dy & \int_\Omega \frac{d^2 I}{dydt}\frac{d^2 I}{dt^2} dy \\ \int_\Omega \frac{d^2 I}{dt^2}\frac{d^2 I}{dxdt} dy & \int_\Omega \frac{d^2 I}{dt^2}\frac{d^2 I}{dydt} dy & \int_\Omega \left\{\frac{d^2 I}{dt^2}\right\}^2 dy \end{bmatrix}$$

Advantages:

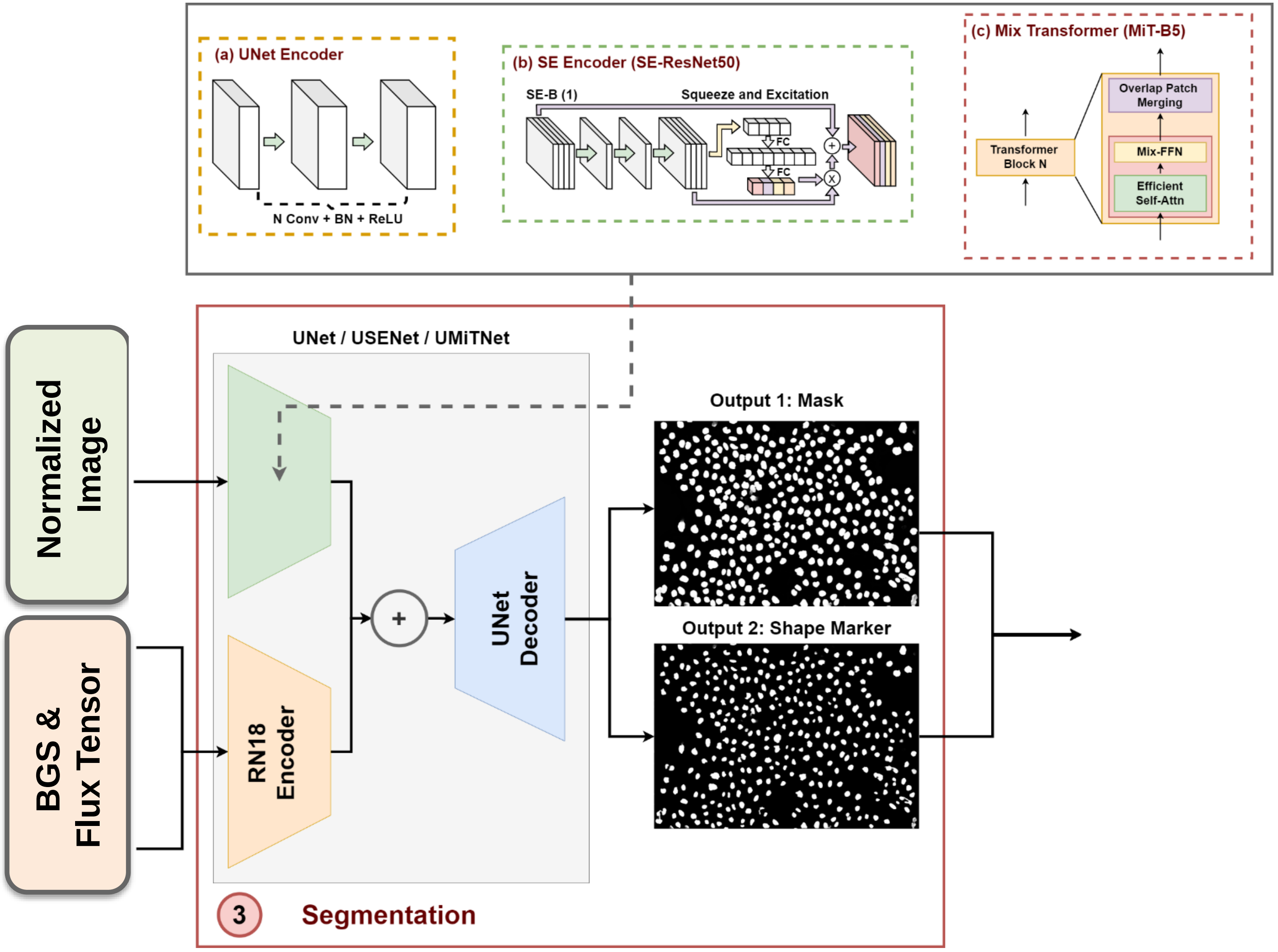
- Elements of the flux tensor incorporate information about temporal gradient changes: <u>efficient discrimination between stationary and moving image features</u>
- The trace of the flux tensor matrix can be directly used to classify moving and non-moving regions <u>without the need for expensive eigenvalue decompositions</u>.

$$trace(J_F) = \int_\Omega \left|\left|\frac{d}{dt}\nabla I\right|\right|^2 dy$$

F. Bunyak, K. Palaniappan, S. Nath, and G. Seetharaman, "Flux tensor constrained geodesic active contours with sensor fusion for persistent object tracking," *Journal of multimedia*, vol. 2, no. 4, p. 20, 2007.

Segmentation: UNet, USENet, UMiTNet

- **Appearance Feature Only (Single Stream):**
  - **UNet:**
    - Backbone: CONV+BN+ReLU
  - **USENet:**
    - Backbone: SE-ResNet-50
  - **UMiTNet:**
    - Backbone: MiT-B5
- **Appearance and Motion Features (Two Streams):**
  - **MUNet (Motion UNet):**
    - Backbone: CONV+BN+ReLU,
    - Backbone2: ResNet-18
  - **MUSENet (Motion USENet):**
    - Backbone: SE-ResNet-50,
    - Backbone2: ResNet-18
  - **MUMiTNet (Motion UMiTNet):**
    - Backbone: MiT-B5,
    - Backbone2: ResNet-18

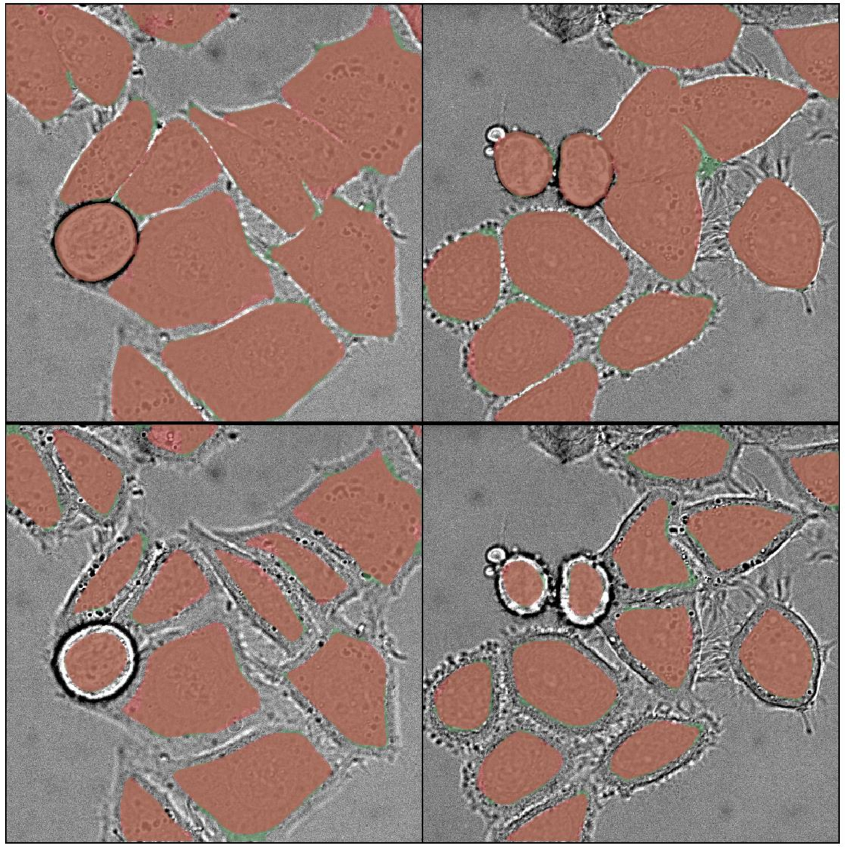


Masks and Markers

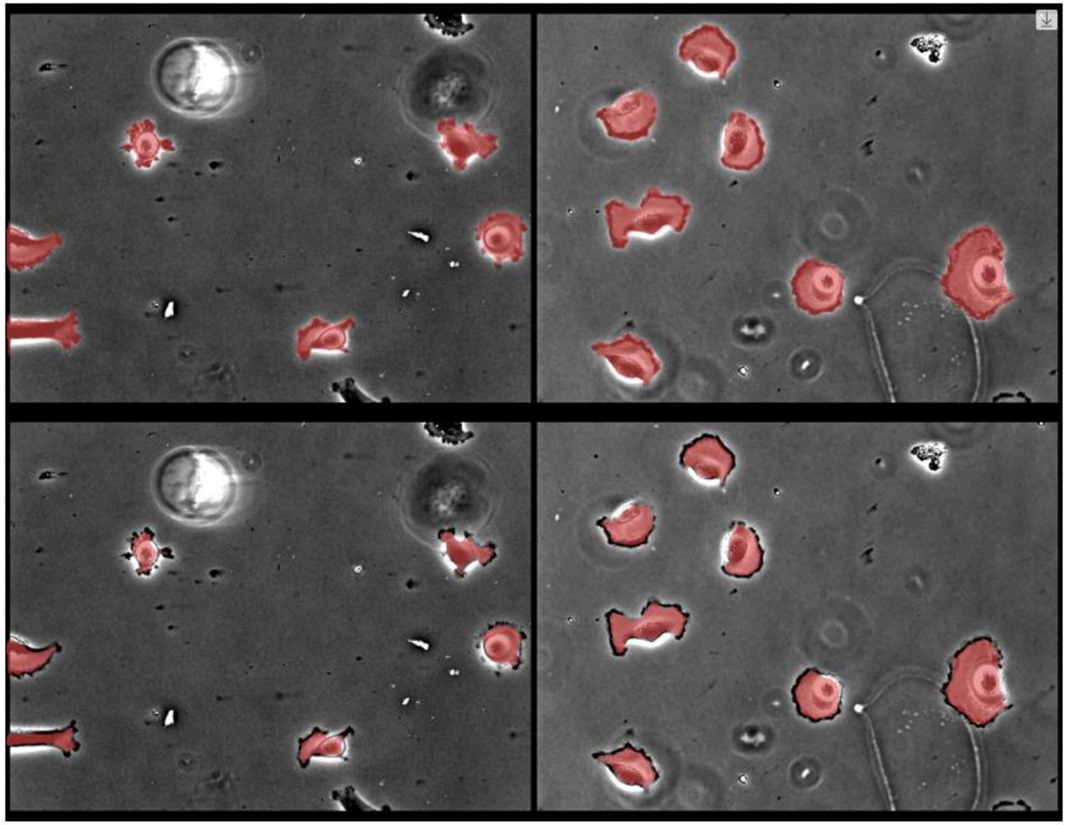
Different erosion level is applied to different datasets based on the size and shape of the cells.

Mask

Markers



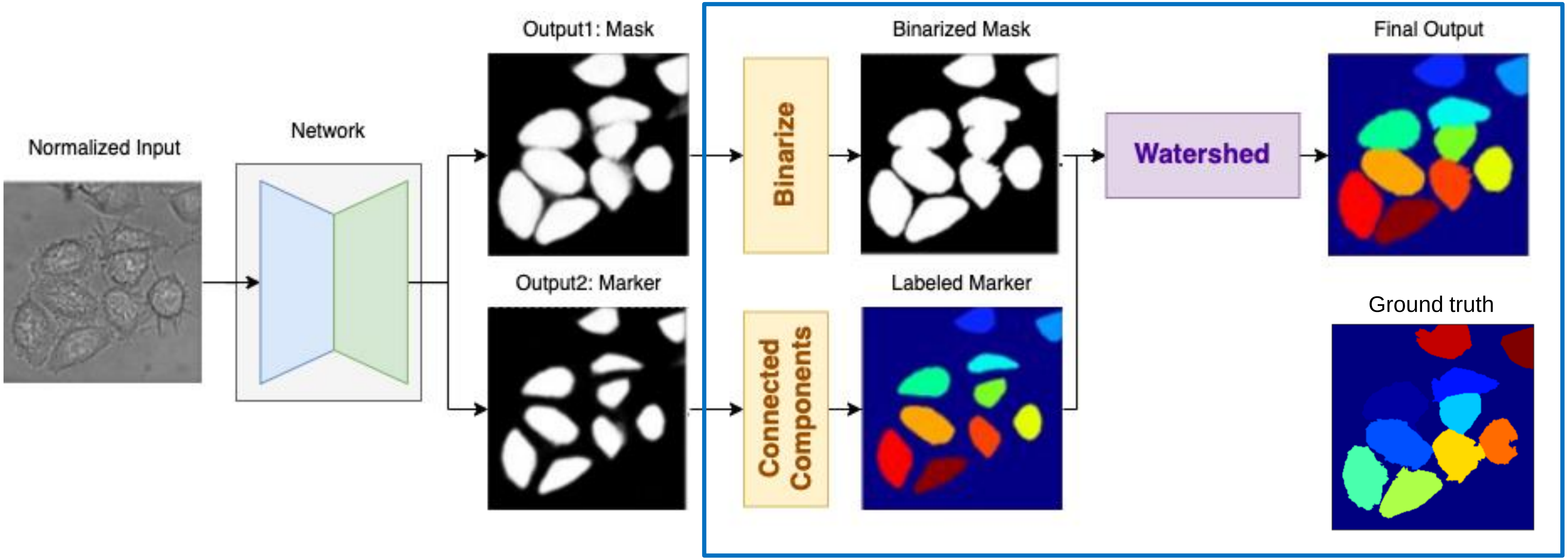
DIC-C2DH-HeLa (Erosion: 20px)



PhC-C2DH-U373 (Erosion: 10px)

Post-processing Across Microscopy Cell Types and Datasets

- Thresholding to get foreground/background mask and shape marker
- Connected components on the shape marker
- Binarized mask and labeled shape marker map used as input to watershed algorithm



Evaluation Metrics

- Segmentation accuracy is measured using OP_{CSB} :

$$\boxed{OP_{CSB} = 0.5 \times (DET + SEG)}$$

  - **SEG** measurement based on the Jaccard similarity index J of the sets of pixels of matching objects

$$\boxed{SEG = J(GT, R) = \frac{GT \cap R}{GT \cup R} \xrightarrow{\text{If}} |GT \cap R| > 0.5.|GT|}$$

    - where GT is Ground Truth and R is the segmentation result

  - **DET** measurement is computed as:

$$\boxed{DET = 1 - min(AOGM_D, AOGM_{D_0})/AOGM_{D_0}}$$

    - where Acyclic Oriented Graph Matching (AOGM) – D is the cost of transforming a set of nodes provided by the participant into the set of GT nodes
    - AOGM-D_0 is the cost of creating the set of GT nodes from scratch (i.e., it is AOGM-D for empty detection results)

Training Details for Unet, USENet, UMiTNet

- **Training:**
  - Data: Weakly supervision is used as labels, 90% / 10% split training/validation
  - Train: training model per dataset (10 models)
  - Train*: training a single model using all data from 6 different datasets (1 model) (**train:1310, validation: 152**)
- **Augmentation:**
  - Horizontal / Vertical Flip, rotation, crop, noise

	# Params
UNet	31M
USENet	35M
UMiTNet	88M

- **Network parameter:**

  input size = 512x512, Adam optimizer, Learning Rate = 1e-4, 250 epochs, mini-batch size = 8, trained / tested on Tesla V100-PCIE-32GB

- **Loss Function:**

$$L_{BCE} = \frac{1}{N}\sum_{i=1}^{N}[-g_i \log p_i + (1-g_i)\log(1-p_i)]$$

$$L_{Dice} = 1 - \frac{2\sum_{i=1}^{N} p_i\, g_i}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2 + \epsilon}$$

- where **pi** prediction, **gi** ground-truth, **N** is the total number of pixels in the mini-batch set of images, **λ** is the weight parameter and it is empirically chosen as **0.5**

$$\boxed{Loss = \lambda * L_{BCE} + (1-\lambda) * L_{Dice}}$$

Experiment Results: Train Separate Model For Each Dataset
(Training Data – *Weakly Supervised*, Evaluating on Training Data – *Sparse Manual*)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7	Fluo-N2DH-SIM+	Average
UNet	**OPCSB**	**0.894**	**0.821**	**0.915**	**0.704**	**0.911**	**0.889**	**0.931**	**0.837**	**0.634**	**0.857**	**0.839**
	SEG	0.803	0.696	0.869	0.605	0.865	0.809	0.885	0.691	0.378	0.778	0.738
	DET	0.985	0.946	0.962	0.802	0.957	0.968	0.978	0.983	0.890	0.937	0.941
USENet	**OPCSB**	**0.915**	**0.827**	**0.949**	**0.771**	**0.937**	**0.916**	**0.927**	**0.837**	**0.881**	**0.911**	**0.887**
	SEG	0.832	0.699	0.923	0.636	0.923	0.850	0.880	0.690	0.820	0.844	0.810
	DET	0.999	0.955	0.974	0.907	0.951	0.983	0.974	0.984	0.942	0.977	0.965
UMiTNet	**OPCSB**	**0.911**	**0.816**	**0.949**	**0.788**	**0.937**	**0.911**	**0.943**	**0.835**	**0.828**	**0.917**	**0.884**
	SEG	0.824	0.687	0.923	0.665	0.926	0.842	0.903	0.694	0.746	0.853	0.806
	DET	0.997	0.946	0.975	0.910	0.948	0.979	0.983	0.977	0.910	0.982	0.961
MUNet	**OPCSB**	**0.914**	**0.802**	**0.905**	**0.666**	**0.902**	**0.901**	**0.933**	**0.832**	**0.796**	**0.823**	**0.847**
	SEG	0.831	0.664	0.861	0.565	0.849	0.829	0.890	0.688	0.670	0.735	0.758
	DET	0.997	0.940	0.948	0.767	0.955	0.974	0.975	0.976	0.923	0.912	0.937
MUSENet	**OPCSB**	**0.917**	**0.832**	**0.949**	**0.773**	**0.931**	**0.914**	**0.937**	**0.838**	**0.856**	**0.911**	**0.886**
	SEG	0.835	0.705	0.928	0.643	0.906	0.845	0.896	0.693	0.790	0.845	0.809
	DET	0.999	0.959	0.971	0.902	0.956	0.983	0.979	0.983	0.923	0.978	0.963
MUMiTNet	**OPCSB**	**0.908**	**0.806**	**0.939**	**0.712**	**0.937**	**0.914**	**0.940**	**0.831**	**0.845**	**0.918**	**0.875**
	SEG	0.822	0.672	0.912	0.566	0.925	0.843	0.897	0.687	0.727	0.858	0.791
	DET	0.995	0.941	0.965	0.858	0.949	0.984	0.984	0.975	0.963	0.977	0.959

Experiment Results: Train Single Model
(Training Data – *Weakly Supervised*, Evaluating on Training Data – *Sparse Manual*)

		DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 (Unseen)	Fluo-N2DH-SIM+ (Unseen)	Average
UNet	**OPCSB**	**0.912**	**0.608**	**0.732**	**0.834**	**0.888**	**0.807**	**0.571**	**0.680**	**0.754**
	SEG	0.868	0.515	0.671	0.727	0.828	0.646	0.249	0.597	0.638
	DET	0.955	0.701	0.792	0.941	0.948	0.968	0.894	0.764	0.870
USENet	**OPCSB**	**0.920**	**0.732**	**0.779**	**0.877**	**0.922**	**0.803**	**0.622**	**0.770**	**0.803**
	SEG	0.878	0.609	0.712	0.783	0.865	0.644	0.359	0.665	0.689
	DET	0.961	0.855	0.845	0.971	0.979	0.962	0.884	0.874	0.917
UMiTNet	**OPCSB**	**0.920**	**0.733**	**0.818**	**0.843**	**0.929**	**0.807**	**0.558**	**0.735**	**0.793**
	SEG	0.879	0.630	0.762	0.729	0.878	0.646	0.240	0.636	0.675
	DET	0.961	0.836	0.875	0.957	0.980	0.967	0.877	0.834	0.911
MUNet	**OPCSB**	**0.905**	**0.599**	**0.733**	**0.827**	**0.913**	**0.807**	**0.576**	**0.700**	**0.757**
	SEG	0.862	0.481	0.670	0.710	0.852	0.645	0.256	0.616	0.637
	DET	0.948	0.716	0.796	0.945	0.974	0.968	0.897	0.784	0.878
MUSENet	**OPCSB**	**0.917**	**0.691**	**0.829**	**0.875**	**0.923**	**0.799**	**0.608**	**0.735**	**0.797**
	SEG	0.877	0.562	0.778	0.780	0.875	0.639	0.338	0.646	0.687
	DET	0.957	0.819	0.880	0.971	0.972	0.958	0.878	0.825	0.907
MUMiTNet	**OPCSB**	**0.924**	**0.724**	**0.841**	**0.840**	**0.929**	**0.797**	**0.567**	**0.749**	**0.796**
	SEG	0.889	0.602	0.789	0.720	0.879	0.629	0.260	0.646	0.677
	DET	0.960	0.845	0.893	0.959	0.980	0.966	0.874	0.851	0.916

Experiment Results: Train Separate Model For Each Dataset (Test Data – Unseen)

- Manual sparse labels of unseen testing data are not publicly available by Cell Tracking Challenge
- Test data was annotated by our lab members using **CVAT** [3] tool

		DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	Average
UNet	**OPCSB**	**0.791**	**0.579**	**0.905**	**0.724**	**0.910**	**0.782**
	SEG	0.736	0.534	0.882	0.488	0.866	0.701
	DET	0.847	0.624	0.928	0.961	0.953	0.862
USENet	**OPCSB**	**0.847**	**0.712**	**0.898**	**0.711**	**0.910**	**0.815**
	SEG	0.792	0.646	0.875	0.487	0.875	0.735
	DET	0.901	0.778	0.921	0.935	0.945	0.896
UMiTNet	**OPCSB**	**0.823**	**0.704**	**0.893**	**0.717**	**0.923**	**0.812**
	SEG	0.769	0.640	0.866	0.491	0.886	0.730
	DET	0.876	0.768	0.919	0.944	0.960	0.893
MUNet	**OPCSB**	**0.795**	**0.399**	**0.902**	**0.731**	**0.911**	**0.748**
	SEG	0.743	0.368	0.876	0.494	0.873	0.671
	DET	0.846	0.431	0.928	0.969	0.950	0.825
MUSENet	**OPCSB**	**0.856**	**0.690**	**0.908**	**0.711**	**0.930**	**0.819**
	SEG	0.800	0.610	0.883	0.487	0.891	0.734
	DET	0.911	0.770	0.933	0.935	0.968	0.903
MUMiTNet	**OPCSB**	**0.835**	**0.643**	**0.890**	**0.711**	**0.920**	**0.800**
	SEG	0.780	0.572	0.866	0.486	0.883	0.718
	DET	0.890	0.713	0.914	0.936	0.957	0.882

[3] https://www.cvat.ai/

Experiment Results: Train Single Model (Test Data-Unseen)

		DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	Average
UNet	**OPCSB**	**0.864**	**0.467**	**0.583**	**0.712**	**0.838**	**0.693**
	SEG	0.797	0.448	0.571	0.465	0.787	0.614
	DET	0.930	0.486	0.595	0.958	0.888	0.772
USENet	**OPCSB**	**0.880**	**0.628**	**0.819**	**0.726**	**0.909**	**0.792**
	SEG	0.814	0.584	0.787	0.481	0.858	0.705
	DET	0.946	0.673	0.850	0.971	0.960	0.880
UMiTNet	**OPCSB**	**0.876**	**0.522**	**0.848**	**0.708**	**0.912**	**0.773**
	SEG	0.809	0.487	0.815	0.469	0.858	0.688
	DET	0.943	0.558	0.881	0.947	0.966	0.859
MUNet	**OPCSB**	**0.867**	**0.449**	**0.612**	**0.721**	**0.892**	**0.708**
	SEG	0.801	0.424	0.591	0.468	0.836	0.624
	DET	0.934	0.474	0.633	0.975	0.948	0.792
MUSENet	**OPCSB**	**0.887**	**0.536**	**0.816**	**0.717**	**0.907**	**0.772**
	SEG	0.820	0.505	0.784	0.474	0.852	0.687
	DET	0.954	0.566	0.847	0.959	0.963	0.858
MUMiTNet	**OPCSB**	**0.881**	**0.518**	**0.864**	**0.698**	**0.906**	**0.773**
	SEG	0.815	0.475	0.827	0.455	0.850	0.685
	DET	0.947	0.561	0.900	0.940	0.962	0.862

Experiment Results: All, Metric: OPCSB (1st best, 2nd best, 3rd best)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 *	Fluo-N2DH-SIM+*	Average
Train Separate Model	UNet	0.8940	0.8210	0.9150	0.7040	0.9110	0.8890	0.9310	0.8370	0.6340	0.8570	**0.8393**
	USENet	0.9150	0.8270	0.9490	0.7710	0.9370	0.9160	0.9270	0.8370	0.8810	0.9110	**0.8871**
	UMiTNet	0.9110	0.8160	0.9490	0.7880	0.9370	0.9110	0.9430	0.8350	0.8280	0.9170	**0.8835**
	MUNet	0.9140	0.8020	0.9050	0.6660	0.9020	0.9010	0.9330	0.8320	0.7960	0.8230	**0.8474**
	MUSENet	0.9170	0.8320	0.9490	0.7730	0.9310	0.9140	0.9370	0.8380	0.8560	0.9110	**0.8858**
	MUMiTNet	0.9080	0.8060	0.9390	0.7120	0.9370	0.9140	0.9400	0.8310	0.8450	0.9180	**0.8750**
Train Single Model	UNet	x	x	0.9120	0.6080	0.7320	0.8340	0.8880	0.8070	0.5710	0.6800	**0.7540**
	USENet	x	x	0.9200	0.7320	0.7790	0.8770	0.9220	0.8030	0.6220	0.7700	**0.8031**
	UMiTNet	x	x	0.9200	0.7330	0.8180	0.8430	0.9290	0.8070	0.5580	0.7350	0.7929
	MUNet	x	x	0.9050	0.5990	0.7330	0.8270	0.9130	0.8070	0.5760	0.7000	0.7575
	MUSENet	x	x	0.9170	0.6910	0.8290	0.8750	0.9230	0.7990	0.6080	0.7350	**0.7971**
	MUMiTNet	x	x	0.9240	0.7240	0.8410	0.8400	0.9290	0.7970	0.5670	0.7490	**0.7964**
Test Separate Model	UNet	x	x	0.7910	0.5790	0.9050	0.7240	0.9100	x	x	x	**0.7818**
	USENet	x	x	0.8470	0.7120	0.8980	0.7110	0.9100	x	x	x	**0.8156**
	UMiTNet	x	x	0.8230	0.7040	0.8930	0.7170	0.9230	x	x	x	**0.8120**
	MUNet	x	x	0.7950	0.3990	0.9020	0.7310	0.9110	x	x	x	**0.7476**
	MUSENet	x	x	0.8560	0.6900	0.9080	0.7110	0.9300	x	x	x	**0.8190**
	MUMiTNet	x	x	0.8350	0.6430	0.8900	0.7110	0.9200	x	x	x	**0.7998**
Test Single Model	UNet	x	x	0.8640	0.4670	0.5830	0.7120	0.8380	x	x	x	**0.6928**
	USENet	x	x	0.8800	0.6280	0.8190	0.7260	0.9090	x	x	x	**0.7924**
	UMiTNet	x	x	0.8760	0.5220	0.8480	0.7080	0.9120	x	x	x	**0.7732**
	MUNet	x	x	0.8670	0.4490	0.6120	0.7210	0.8920	x	x	x	**0.7082**
	MUSENet	x	x	0.8870	0.5360	0.8160	0.7170	0.9070	x	x	x	**0.7726**
	MUMiTNet	x	x	0.8810	0.5180	0.8640	0.6980	0.9060	x	x	x	**0.7734**

*** means unseen data for a single model case**

Experiment Results: All, Metric: OPCSB (1st best, 2nd best, 3rd best)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 *	Fluo-N2DH-SIM+*	Average
Train Separate Model	UNet	0.8940	0.8210	0.9150	0.7040	0.9110	0.8890	0.9310	0.8370	0.6340	0.8570	**0.8393**
	USENet	0.9150	0.8270	0.9490	0.7710	0.9370	0.9160	0.9270	0.8370	0.8810	0.9110	**0.8871**
	UMiTNet	0.9110	0.8160	0.9490	0.7880	0.9370	0.9110	0.9430	0.8350	0.8280	0.917[illegible]	**0.8835**
	MUNet	0.9140	0.8020	0.9050	0.6660	0.9020	0.9010	0.9330	0.8320	0.7960	0.8230	**0.8474**
	MUSENet	0.9170	0.8320	0.9490	0.7730	0.9310	0.9140	0.9370	0.8380	0.8560	0.9110	**0.8858**
	MUMiTNet	0.9080	0.8060	0.9390	0.7120	0.9370	0.9140	0.9400	0.8310	0.8450	0.9180	**0.8750**
Train Single Model	UNet	x	x	0.9120	0.6080	0.7320	0.8340	0.8880	0.8070	0.5710	0.6800	**0.7540**
	USENet	x	x	0.9200	0.7320	0.7790	0.8770	0.9220	0.8030	0.6220	0.7700	0.8031
	UMiTNet	x	x	0.9200	0.7330	0.8180	0.8430	0.9290	0.8070	0.5580	0.7350	0.7929
	MUNet	x	x	0.9050	0.5990	0.7330	0.8270	0.9130	0.8070	0.5760	0.7000	0.7575
	MUSENet	x	x	0.9170	0.6910	0.8290	0.8750	0.9230	0.7990	0.6080	0.7350	0.7971
	MUMiTNet	x	x	0.9240	0.7240	0.8410	0.8400	0.9290	0.7970	0.5670	0.7490	0.7964
Test Separate Model	UNet	x	x	0.7910	0.5790	0.9050	0.7240	0.9100	x	x	x	**0.7818**
	USENet	x	x	0.8470	0.7120	0.8980	0.7110	0.9100	x	x	x	**0.8156**
	UMiTNet	x	x	0.8230	0.7040	0.8930	0.7170	0.9230	x	x	x	**0.8120**
	MUNet	x	x	0.7950	0.3990	0.9020	0.7310	0.9110	x	x	x	**0.7476**
	MUSENet	x	x	0.8560	0.6900	0.9080	0.7110	0.9300	x	x	x	**0.8190**
	MUMiTNet	x	x	0.8350	0.6430	0.8900	0.7110	0.9200	x	x	x	**0.7998**
Test Single Model	UNet	x	x	0.8640	0.4670	0.5830	0.7120	0.8380	x	x	x	**0.6928**
	USENet	x	x	0.8800	0.6280	0.8190	0.7260	0.9090	x	x	x	**0.7924**
	UMiTNet	x	x	0.8760	0.5220	0.8480	0.7080	0.9120	x	x	x	**0.7732**
	MUNet	x	x	0.8670	0.4490	0.6120	0.7210	0.8920	x	x	x	**0.7082**
	MUSENet	x	x	0.8870	0.5360	0.8160	0.7170	0.9070	x	x	x	**0.7726**
	MUMiTNet	x	x	0.8810	0.5180	0.8640	0.6980	0.9060	x	x	x	**0.7734**

1
3
2

*** means unseen data for a single model case**

Experiment Results: All, Metric: OPCSB (1st best, 2nd best, 3rd best)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 *	Fluo-N2DH-SIM+*	Average
Train Separate Model	UNet	0.8940	0.8210	0.9150	0.7040	0.9110	0.8890	0.9310	0.8370	0.6340	0.8570	**0.8393**
	USENet	0.9150	0.8270	0.9490	0.7710	0.9370	0.9160	0.9270	0.8370	0.8810	0.9110	**0.8871**
	UMiTNet	0.9110	0.8160	0.9490	0.7880	0.9370	0.9110	0.9430	0.8350	0.8280	0.9170	**0.8835**
	MUNet	0.9140	0.8020	0.9050	0.6660	0.9020	0.9010	0.9330	0.8320	0.7960	0.8230	**0.8474**
	MUSENet	0.9170	0.8320	0.9490	0.7730	0.9310	0.9140	0.9370	0.8380	0.8560	0.9110	**0.8858**
	MUMiTNet	0.9080	0.8060	0.9390	0.7120	0.9370	0.9140	0.9400	0.8310	0.8450	0.9180	**0.8750**
Train Single Model	UNet	x	x	0.9120	0.6080	0.7320	0.8340	0.8880	0.8070	0.5710	0.6800	**0.7540**
	USENet	x	x	0.9200	0.7320	0.7790	0.8770	0.9220	0.8030	0.6220	0.7700	**0.8031**
	UMiTNet	x	x	0.9200	0.7330	0.8180	0.8430	0.9290	0.8070	0.5580	0.7350	0.7929
	MUNet	x	x	0.9050	0.5990	0.7330	0.8270	0.9130	0.8070	0.5760	0.7000	0.7575
	MUSENet	x	x	0.9170	0.6910	0.8290	0.8750	0.9230	0.7990	0.6080	0.7350	**0.7971**
	MUMiTNet	x	x	0.9240	0.7240	0.8410	0.8400	0.9290	0.7970	0.5670	0.749[illegible]	**0.7964**
Test Separate Model	UNet	x	x	0.7910	0.5790	0.9050	0.7240	0.9100	x	x	x	**0.7818**
	USENet	x	x	0.8470	0.7120	0.8980	0.7110	0.9100	x	x	x	**0.8156**
	UMiTNet	x	x	0.8230	0.7040	0.8930	0.7170	0.9230	x	x	x	**0.8120**
	MUNet	x	x	0.7950	0.3990	0.9020	0.7310	0.9110	x	x	x	**0.7476**
	MUSENet	x	x	0.8560	0.6900	0.9080	0.7110	0.9300	x	x	x	**0.8190**
	MUMiTNet	x	x	0.8350	0.6430	0.8900	0.7110	0.9200	x	x	x	**0.7998**
Test Single Model	UNet	x	x	0.8640	0.4670	0.5830	0.7120	0.8380	x	x	x	**0.6928**
	USENet	x	x	0.8800	0.6280	0.8190	0.7260	0.9090	x	x	x	**0.7924**
	UMiTNet	x	x	0.8760	0.5220	0.8480	0.7080	0.9120	x	x	x	**0.7732**
	MUNet	x	x	0.8670	0.4490	0.6120	0.7210	0.8920	x	x	x	**0.7082**
	MUSENet	x	x	0.8870	0.5360	0.8160	0.7170	0.9070	x	x	x	**0.7726**
	MUMiTNet	x	x	0.8810	0.5180	0.8640	0.6980	0.9060	x	x	x	**0.7734**

* means unseen data for a single model case

Experiment Results: All, Metric: OPCSB (1st best, 2nd best, 3rd best)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 *	Fluo-N2DH-SIM+*	Average	
Train Separate Model	UNet	0.8940	0.8210	0.9150	0.7040	0.9110	0.8890	0.9310	0.8370	0.6340	0.8570	**0.8393**	
	USENet	0.9150	0.8270	0.9490	0.7710	0.9370	0.9160	0.9270	0.8370	0.8810	0.9110	**0.8871**	
	UMiTNet	0.9110	0.8160	0.9490	0.7880	0.9370	0.9110	0.9430	0.8350	0.8280	0.9170	**0.8835**	
	MUNet	0.9140	0.8020	0.9050	0.6660	0.9020	0.9010	0.9330	0.8320	0.7960	0.8230	**0.8474**	
	MUSENet	0.9170	0.8320	0.9490	0.7730	0.9310	0.9140	0.9370	0.8380	0.8560	0.9110	**0.8858**	
	MUMiTNet	0.9080	0.8060	0.9390	0.7120	0.9370	0.9140	0.9400	0.8310	0.8450	0.9180	**0.8750**	
Train Single Model	UNet	x	x	0.9120	0.6080	0.7320	0.8340	0.8880	0.8070	0.5710	0.6800	**0.7540**	
	USENet	x	x	0.9200	0.7320	0.7790	0.8770	0.9220	0.8030	0.6220	0.7700	**0.8031**	
	UMiTNet	x	x	0.9200	0.7330	0.8180	0.8430	0.9290	0.8070	0.5580	0.7350	0.7929	
	MUNet	x	x	0.9050	0.5990	0.7330	0.8270	0.9130	0.8070	0.5760	0.7000	0.7575	
	MUSENet	x	x	0.9170	0.6910	0.8290	0.8750	0.9230	0.7990	0.6080	0.7350	**0.7971**	
	MUMiTNet	x	x	0.9240	0.7240	0.8410	0.8400	0.9290	0.7970	0.5670	0.7490	**0.7964**	
Test Separate Model	UNet	x	x	0.7910	0.5790	0.9050	0.7240	0.9100	x	x	x	**0.7818**	
	USENet	x	x	0.8470	0.7120	0.8980	0.7110	0.9100	x	x	x	**0.8156**	2
	UMiTNet	x	x	0.8230	0.7040	0.8930	0.7170	0.9230	x	x	x	**0.8120**	3
	MUNet	x	x	0.7950	0.3990	0.9020	0.7310	0.9110	x	x	x	**0.7476**	
	MUSENet	x	x	0.8560	0.6900	0.9080	0.7110	0.9300	x	x	x	**0.8190**	1
	MUMiTNet	x	x	0.8350	0.6430	0.8900	0.7110	0.9200	x	x	x	**0.7998**	
Test Single Model	UNet	x	x	0.8640	0.4670	0.5830	0.7120	0.8380	x	x	x	**0.6928**	
	USENet	x	x	0.8800	0.6280	0.8190	0.7260	0.9090	x	x	x	**0.7924**	
	UMiTNet	x	x	0.8760	0.5220	0.8480	0.7080	0.9120	x	x	x	**0.7732**	
	MUNet	x	x	0.8670	0.4490	0.6120	0.7210	0.8920	x	x	x	**0.7082**	
	MUSENet	x	x	0.8870	0.5360	0.8160	0.7170	0.9070	x	x	x	**0.7726**	
	MUMiTNet	x	x	0.8810	0.5180	0.8640	0.6980	0.9060	x	x	x	**0.7734**	

* means unseen data for a single model case

Experiment Results: All, Metric: OPCSB (1st best, 2nd best, 3rd best)

		BF-C2DL-HSC	BF-C2DL-MuSC	DIC-C2DH-HeLa	Fluo-C2DL-MSC	Fluo-N2DH-GOWT1	Fluo-N2DL-HeLa	PhC-C2DH-U373	PhC-C2DH-PSC	Fluo-C2DL-Huh7 *	Fluo-N2DH-SIM+*	Average	
Train Separate Model	UNet	0.8940	0.8210	0.9150	0.7040	0.9110	0.8890	0.9310	0.8370	0.6340	0.8570	**0.8393**	
	USENet	0.9150	0.8270	0.9490	0.7710	0.9370	0.9160	0.9270	0.8370	0.8810	0.9110	**0.8871**	
	UMiTNet	0.9110	0.8160	0.9490	0.7880	0.9370	0.9110	0.9430	0.8350	0.8280	0.9170	**0.8835**	
	MUNet	0.9140	0.8020	0.9050	0.6660	0.9020	0.9010	0.9330	0.8320	0.7960	0.8230	**0.8474**	
	MUSENet	0.9170	0.8320	0.9490	0.7730	0.9310	0.9140	0.9370	0.8380	0.8560	0.9110	**0.8858**	
	MUMiTNet	0.9080	0.8060	0.9390	0.7120	0.9370	0.9140	0.9400	0.8310	0.8450	0.9180	**0.8750**	
Train Single Model	UNet	x	x	0.9120	0.6080	0.7320	0.8340	0.8880	0.8070	0.5710	0.6800	**0.7540**	
	USENet	x	x	0.9200	0.7320	0.7790	0.8770	0.9220	0.8030	0.6220	0.7700	**0.8031**	
	UMiTNet	x	x	0.9200	0.7330	0.8180	0.8430	0.9290	0.8070	0.5580	0.7350	0.7929	
	MUNet	x	x	0.9050	0.5990	0.7330	0.8270	0.9130	0.8070	0.5760	0.7000	0.7575	
	MUSENet	x	x	0.9170	0.6910	0.8290	0.8750	0.9230	0.7990	0.6080	0.7350	**0.7971**	
	MUMiTNet	x	x	0.9240	0.7240	0.8410	0.8400	0.9290	0.7970	0.5670	0.7490	**0.7964**	
Test Separate Model	UNet	x	x	0.7910	0.5790	0.9050	0.7240	0.9100	x	x	x	**0.7818**	
	USENet	x	x	0.8470	0.7120	0.8980	0.7110	0.9100	x	x	x	**0.8156**	
	UMiTNet	x	x	0.8230	0.7040	0.8930	0.7170	0.9230	x	x	x	**0.8120**	
	MUNet	x	x	0.7950	0.3990	0.9020	0.7310	0.9110	x	x	x	**0.7476**	
	MUSENet	x	x	0.8560	0.6900	0.9080	0.7110	0.9300	x	x	x	**0.8190**	
	MUMiTNet	x	x	0.8350	0.6430	0.8900	0.7110	0.9200	x	x	x	**0.7998**	
Test Single Model	UNet	x	x	0.8640	0.4670	0.5830	0.7120	0.8380	x	x	x	**0.6928**	
	USENet	x	x	0.8800	0.6280	0.8190	0.7260	0.9090	x	x	x	**0.7924**	1
	UMiTNet	x	x	0.8760	0.5220	0.8480	0.7080	0.9120	x	x	x	**0.7732**	3
	MUNet	x	x	0.8670	0.4490	0.6120	0.7210	0.8920	x	x	x	**0.7082**	
	MUSENet	x	x	0.8870	0.5360	0.8160	0.7170	0.9070	x	x	x	**0.7726**	
	MUMiTNet	x	x	0.8810	0.5180	0.8640	0.6980	0.9060	x	x	x	**0.7734**	2

* means unseen data for a single model case

Experiment Results: Simulated / Weakly Supervised and Real Data (1st best, 2nd best, 3rd best)

		Trained on Fluo-N2DH-SIM+ (Simulated Data)					
		UNet	USENet	UMiTNet	MUNet	MUSENet	MUMiTNet
Fluo-N2DH-GOWT1	**OPCSB**	**0.2728**	**0.5236**	**0.8450**	**0.6098**	**0.5784**	**0.9383**
	SEG	0.2844	0.4749	0.7918	0.5473	0.5356	0.9280
	DET	0.2612	0.5722	0.8982	0.6723	0.6213	0.9485

		Trained on Fluo-N2DH-GOWT1 (Real Data)					
		UNet	USENet	UMiTNet	MUNet	MUSENet	MUMiTNet
Fluo-N2DH-SIM+	**OPCSB**	**0.6086**	**0.4474**	**0.6697**	**0.6103**	**0.6729**	**0.9180**
	SEG	0.4955	0.4617	0.6140	0.5312	0.6030	0.8581
	DET	0.7217	0.4331	0.7253	0.6894	0.7428	0.9778

Experiment Results: Simulated / Weakly Supervised and Real Data
(1st best, 2nd best, 3rd best)

		Trained on Fluo-N2DH-SIM+ (Simulated Data)					
		UNet	USENet	UMiTNet	MUNet	MUSENet	MUMiTNet
Fluo-N2DH-GOWT1	**OPCSB**	**0.2728**	**0.5236**	**0.8450** 2	**0.6098** 3	**0.5784**	**0.9383** 1
	SEG	0.2844	0.4749	0.7918	0.5473	0.5356	0.9280
	DET	0.2612	0.5722	0.8982	0.6723	0.6213	0.9485

33%

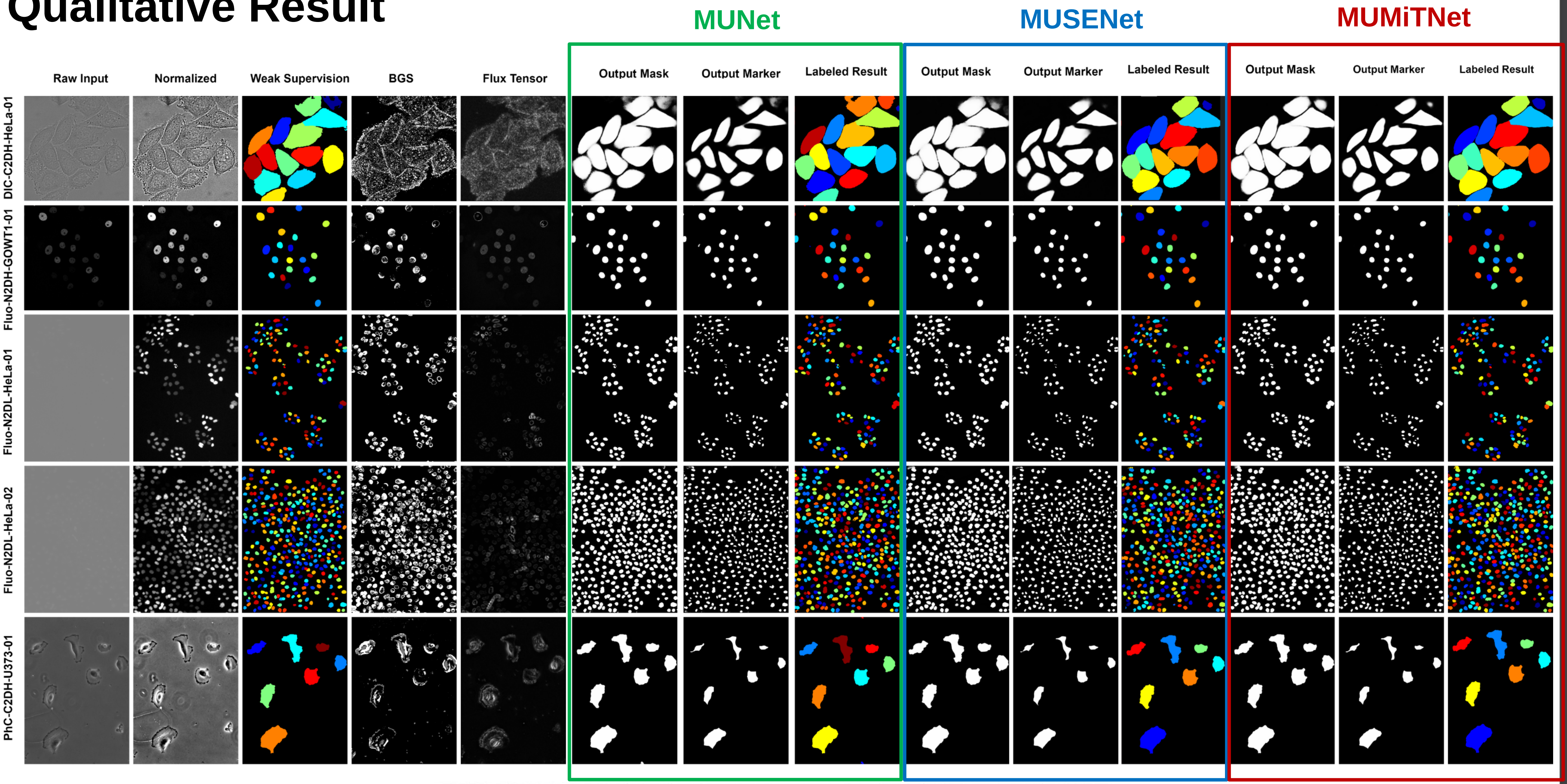
9%

		Trained on Fluo-N2DH-GOWT1 (Real Data)					
		UNet	USENet	UMiTNet	MUNet	MUSENet	MUMiTNet
Fluo-N2DH-SIM+	**OPCSB**	**0.6086**	**0.4474**	**0.6697** 3	**0.6103**	**0.6729** 2	**0.9180** 1
	SEG	0.4955	0.4617	0.6140	0.5312	0.6030	0.8581
	DET	0.7217	0.4331	0.7253	0.6894	0.7428	0.9778

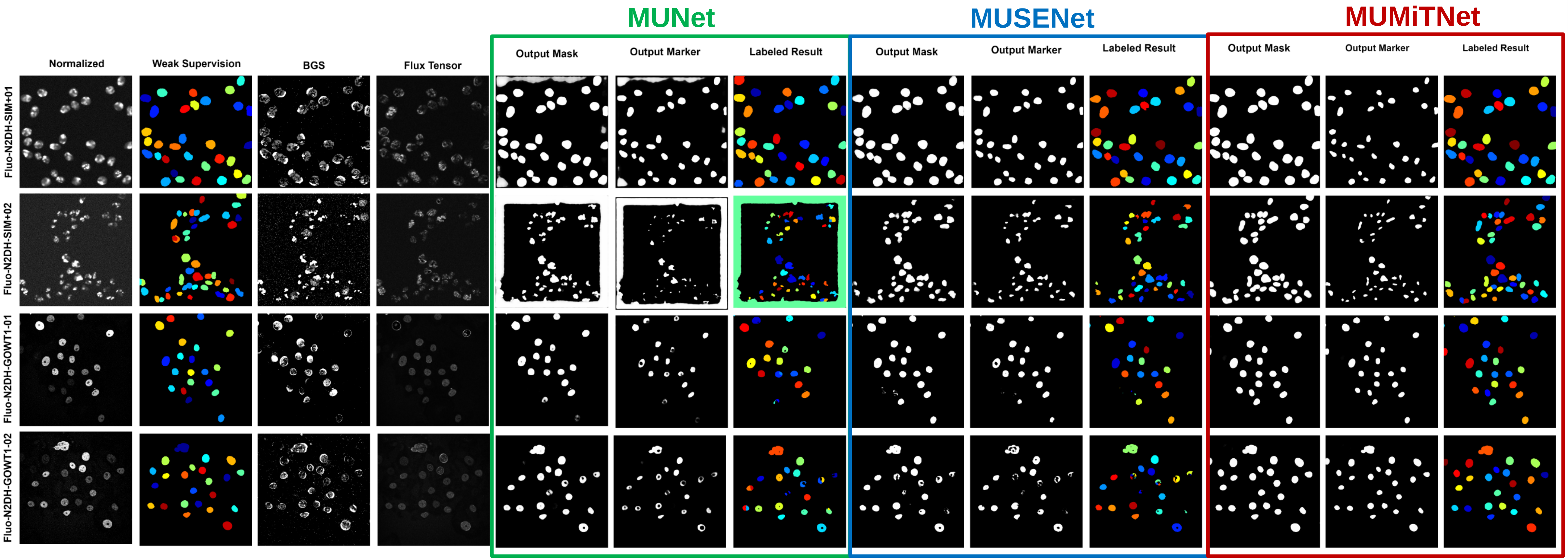
24%

25%

Qualitative Result



Qualitative Result for Simulated and Real Data



Conclusion

- UNet, USENet, and UMiTNet have been evaluated on the Cell Tracking Challenge Dataset.
- Motion cues have been added as an additional stream to each deep-learning architecture.
- From the experiments, it can be concluded that **USENet** with only input image outperformed UNet and UMiTNet in most cases.
- UMiTNet has almost similar accuracy as USENet with less than **1% accuracy**.
- **MUMiTNet** outperforms all other methods:
  - by **33%** (MUNet -> MUMiTNet) when trained on **simulated** data and tested on **real**
  - by **24%** (MUSENet -> MUMiTNet) when trained on **real** data and tested on **simulated**
- Adding **motion cues improved** the accuracy of the networks:
  - by **9%** (UMiTNet -> MUMiTNet) when trained on **simulated** data and tested on **real**
  - by **25%** (UMiTNet -> MUMiTNet) when trained on **real** data and tested on **simulated**
- In other cases, adding motion cues did not make a significant impact

Future Work

- Training and Testing on 3D subset of Cell Tracking Challenge
- Improving pre & post-processing
- Using Sparse Manual Labels (Gold Truth) with a combination of Weakly Supervised Labels (Silver Truth)