

OTrack

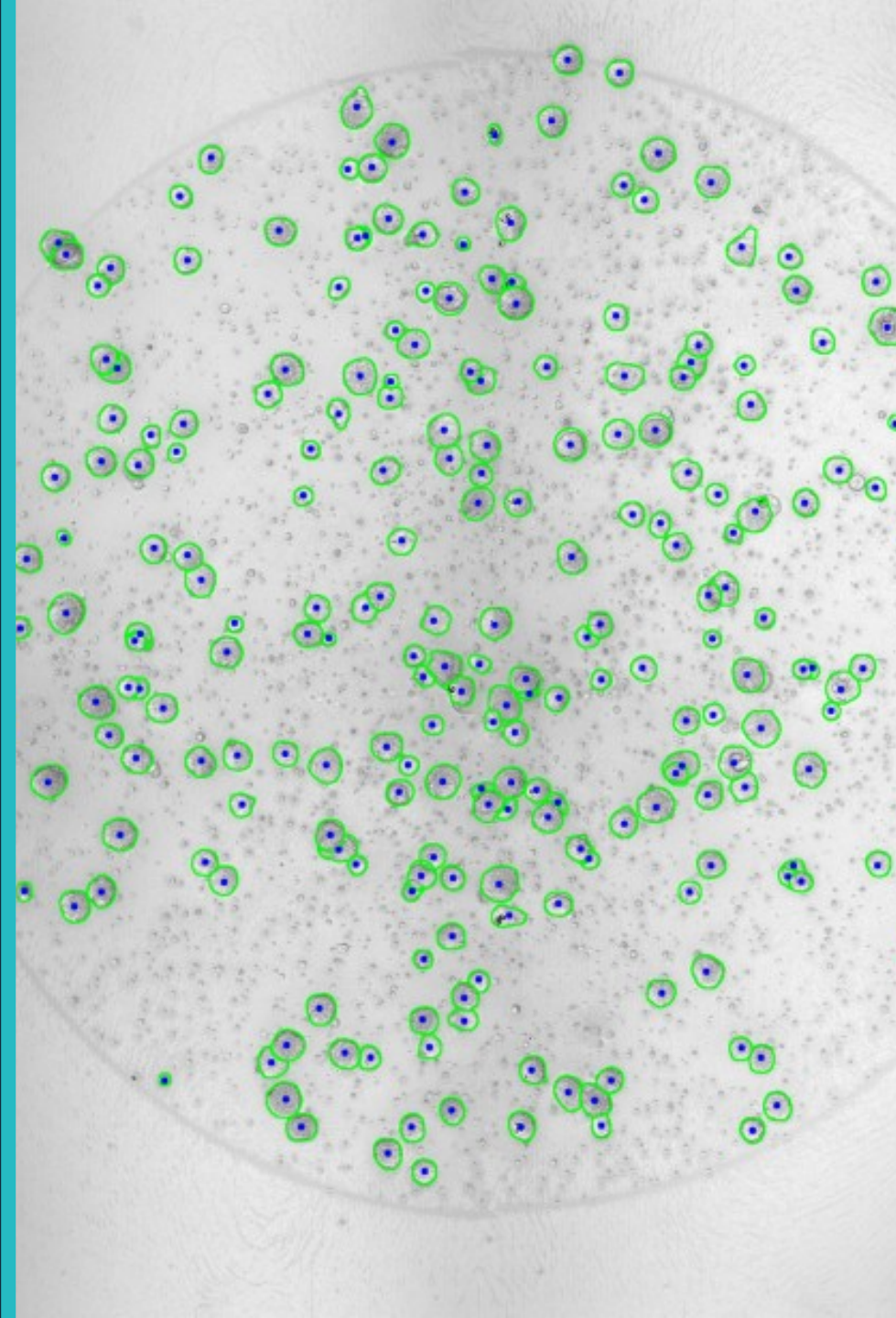
From Organoid Images to Growth Dynamics

A Training-Free Framework for
Robust Longitudinal Tracking

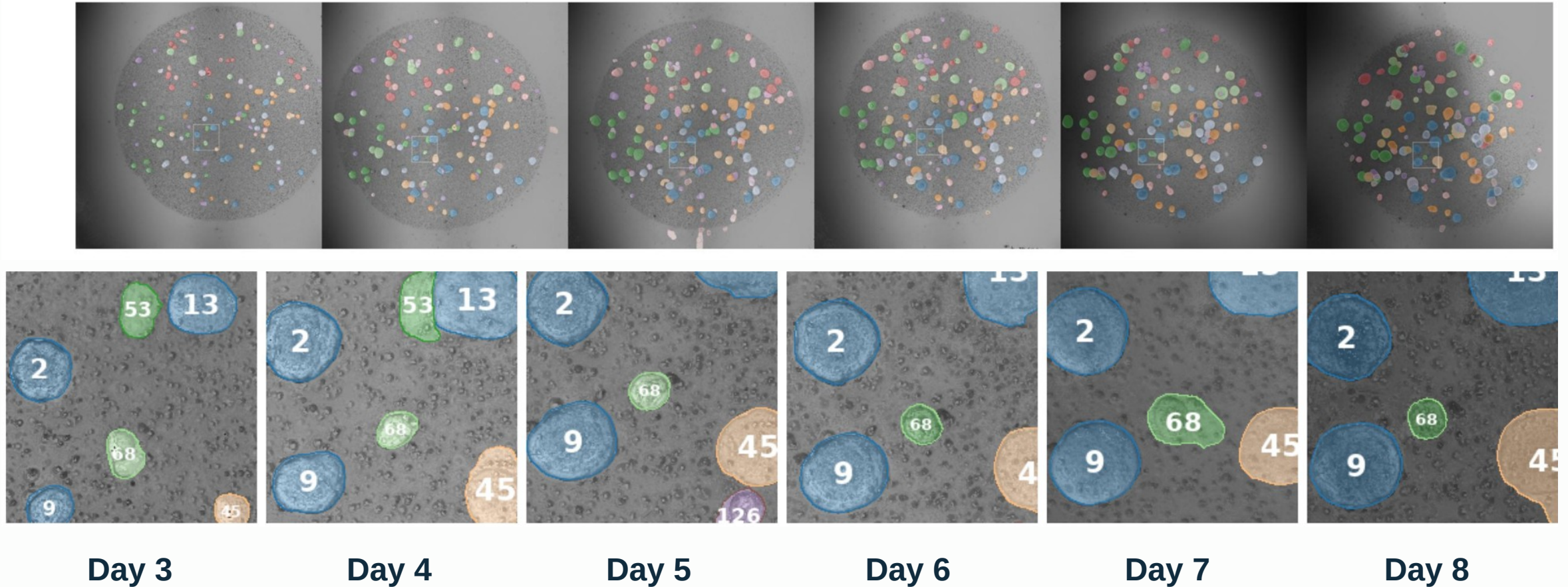
Liang Yang

College of Mathematics · Sichuan University

2 October 2026 | Tashkent, Uzbekistan



The scientific unit is a trajectory—not a frame



Four questions for the next

01

The imaging challenge

Why appearance alone is fragile

02

The geometric solution

From neighborhood cues to identities

03

The experimental evidence

Detection, continuity, and transfer

04

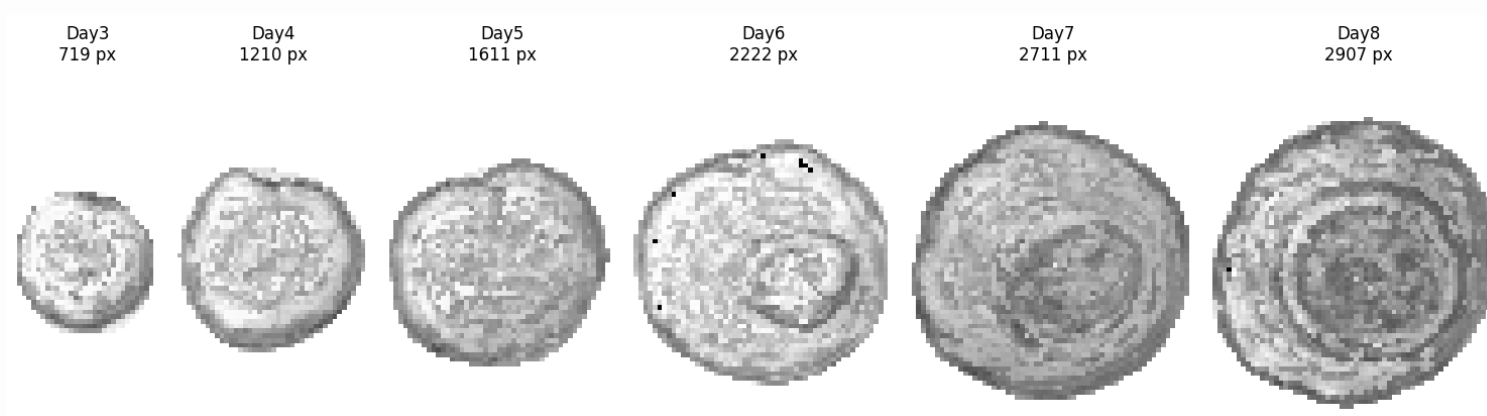
The biological readout

From identity to growth measurements

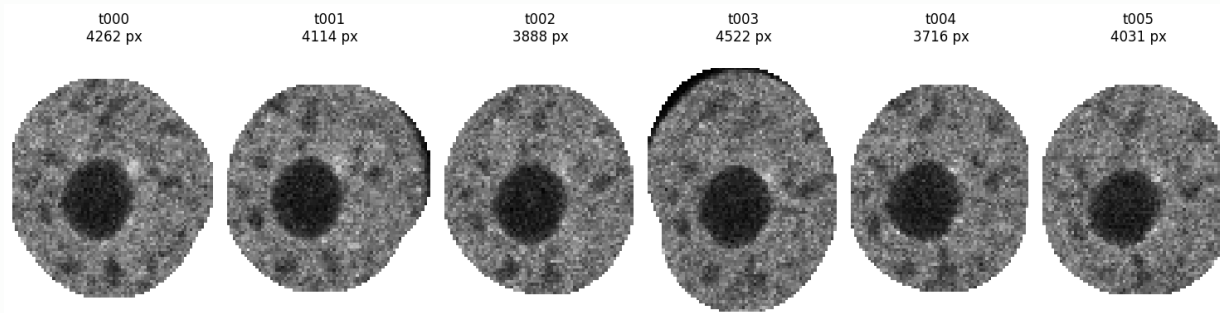
Observe → describe → align → assign → measure

Growth changes the object we are trying to recognize

DAILY ORGANOID IMAGING

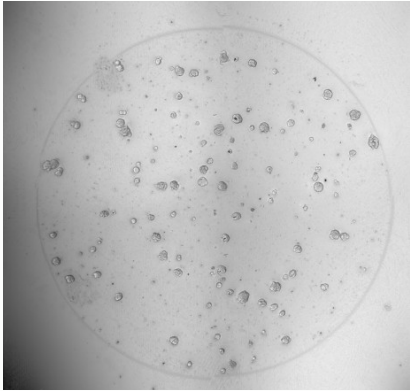


SHORT IMAGE SEQUENCE: MORE SIMILAR APPEARANCE

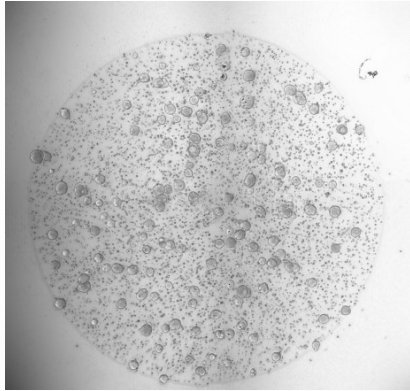


**The same identity
can have a different
size and appearance.**

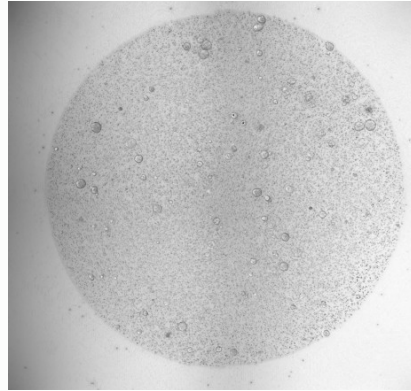
The observation process is already heterogeneous



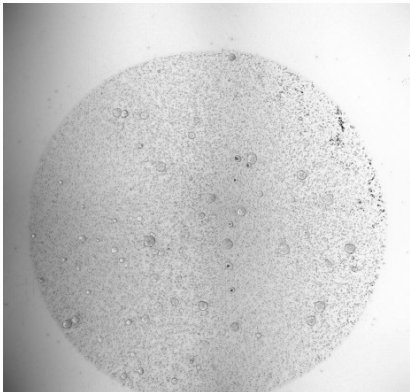
ESCC



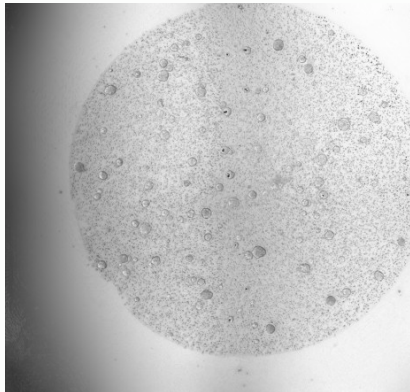
p53



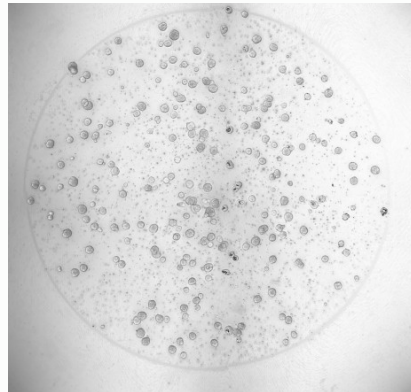
p53p16



p53p16sox2



p53p16sox2tpm4



wt

THREE SOURCES OF HETEROGENEITY

Density

Different spatial neighborhoods

Morphology

Different shapes and scales

Acquisition

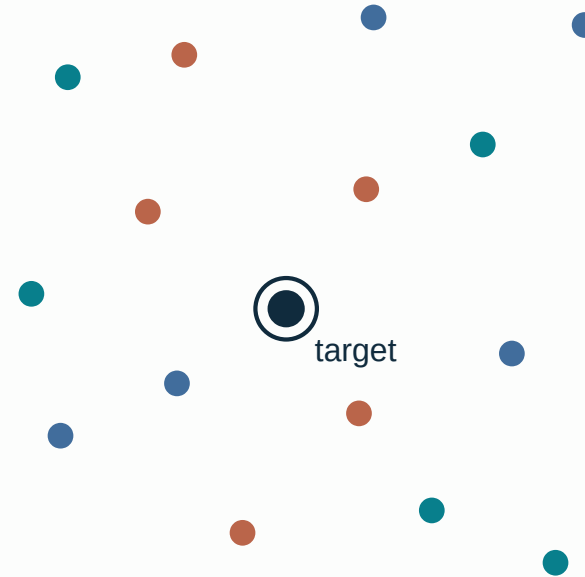
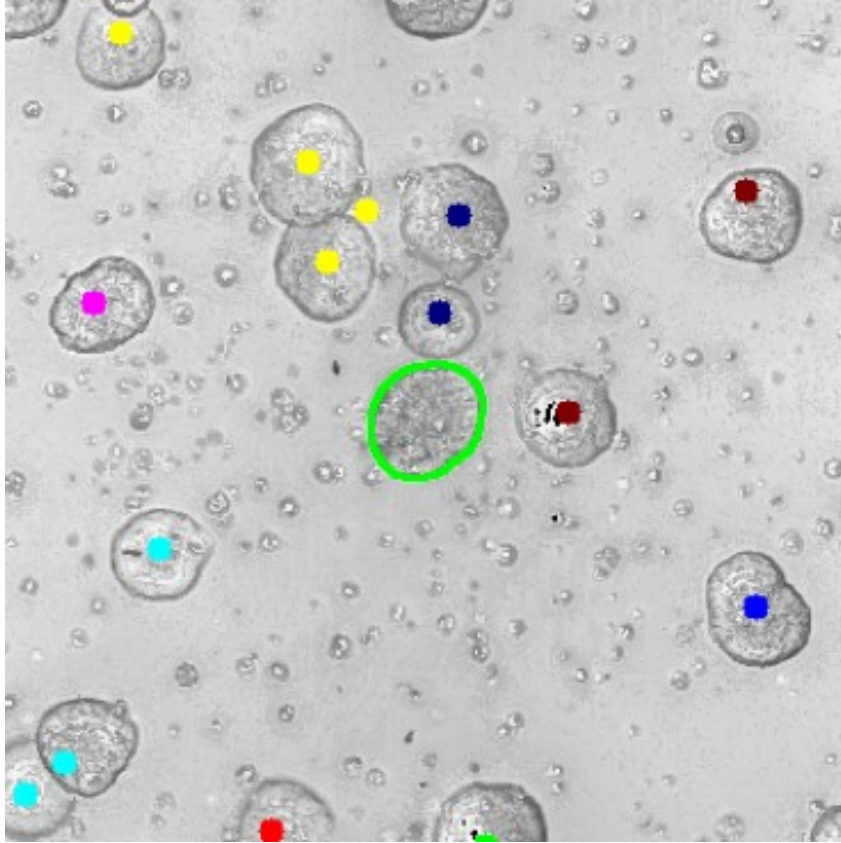
Focus, contrast, and field of view

A mathematical view: recover identity between two sets



Two unordered mask sets. One longitudinal identity problem.

Use the neighborhood as an identity cue



**An object
is more than
its own
pixels.**

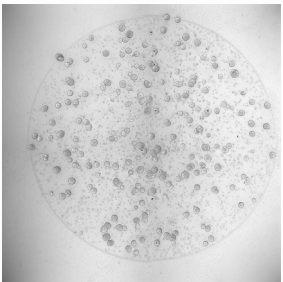
Nearby objects
provide a spatial
context for
matching.

OTrack separates perception, geometry, and assignment

01

Reconstruct

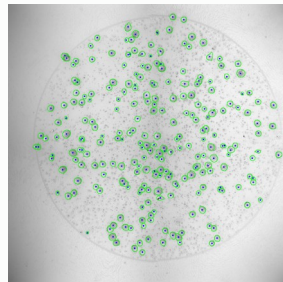
Focus stacking
Image stitching



02

Segment

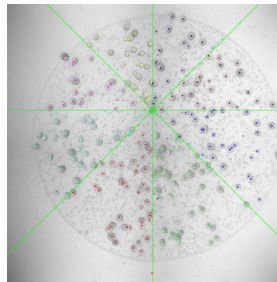
Pretrained SAM2
Cleaned masks



03

Describe

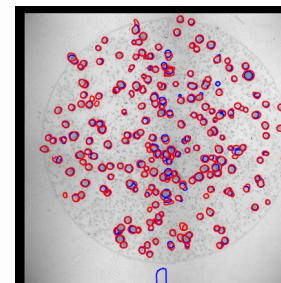
HoR + HoC
Morphology



04

Align

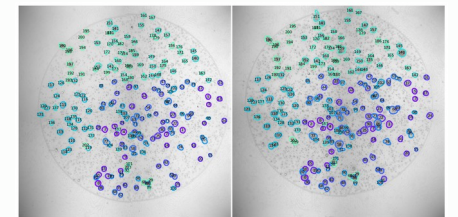
Candidate matches
GC-RANSAC



05

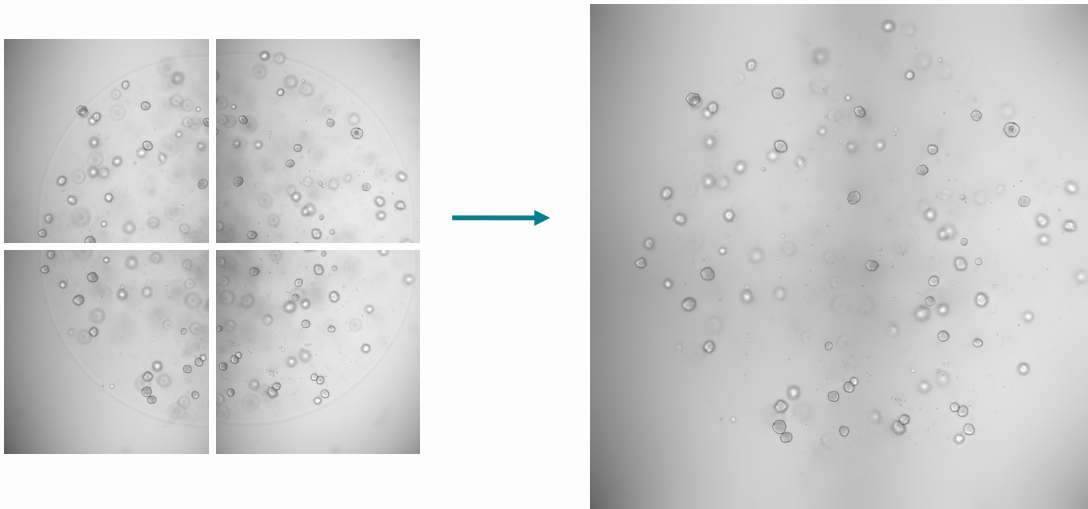
Associate

Aligned-mask IoU
Hungarian
assignment



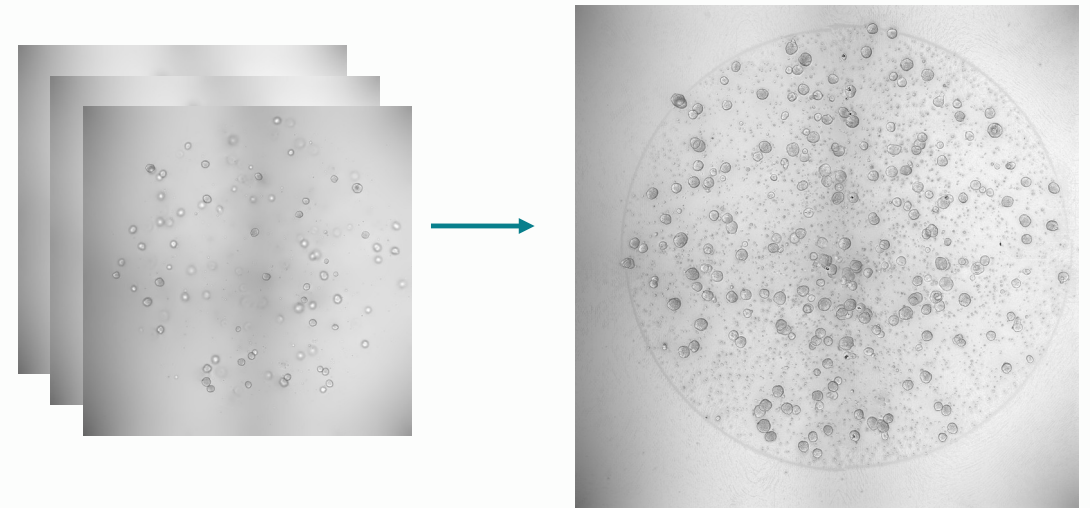
Reconstruct a consistent field of view

IMAGE STITCHING



Align tiles, optimize seams, blend

FOCUS STACKING



Fuse complementary focal planes

Training-free means no task-specific network training

PRETRAINED COMPONENT

SAM2 is reused

Inference with a model whose weights were learned before this task.

EXPLICIT COMPUTATION

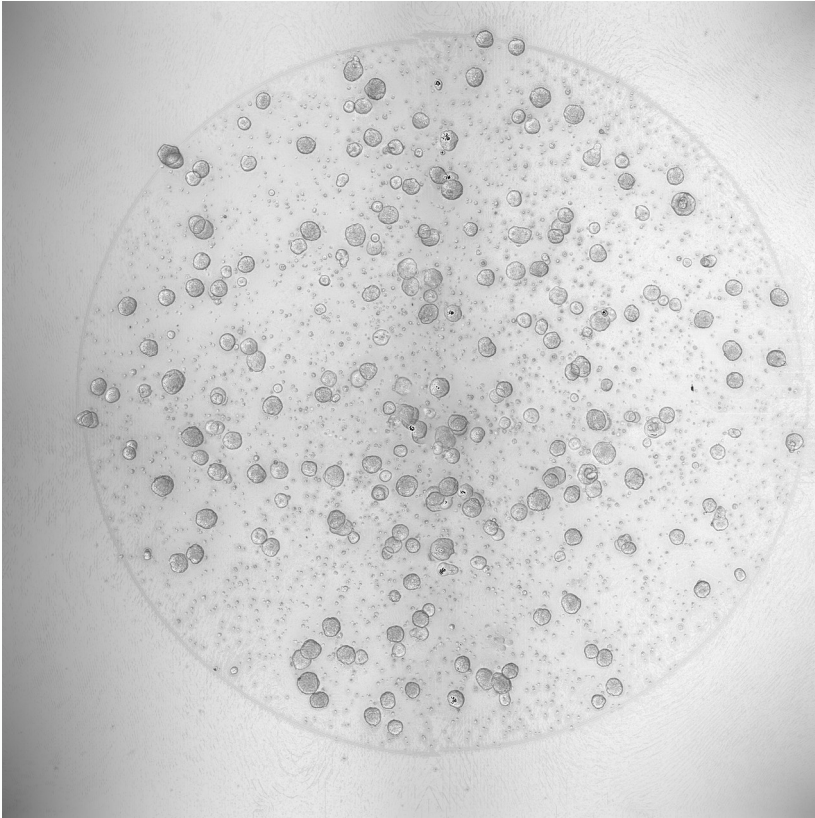
No new tracking network

Geometric descriptors, robust fitting, and combinatorial assignment.

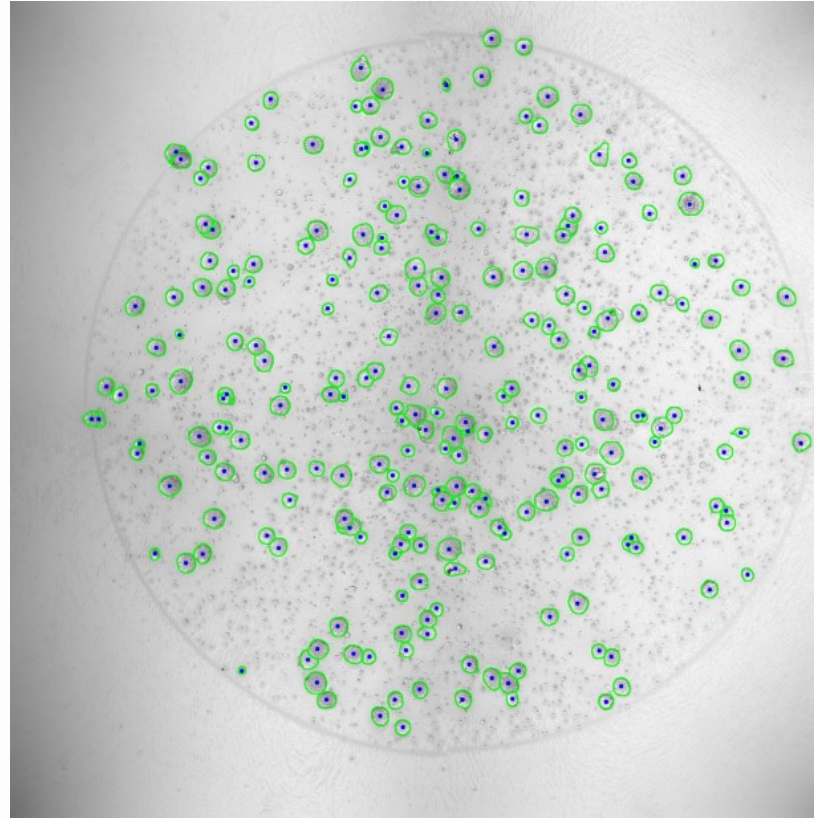
Training-free $\neq$ parameter-free

Prompt generation, filtering, descriptor weights, and algorithm settings still matter.

From an image to a cleaned set of instance masks



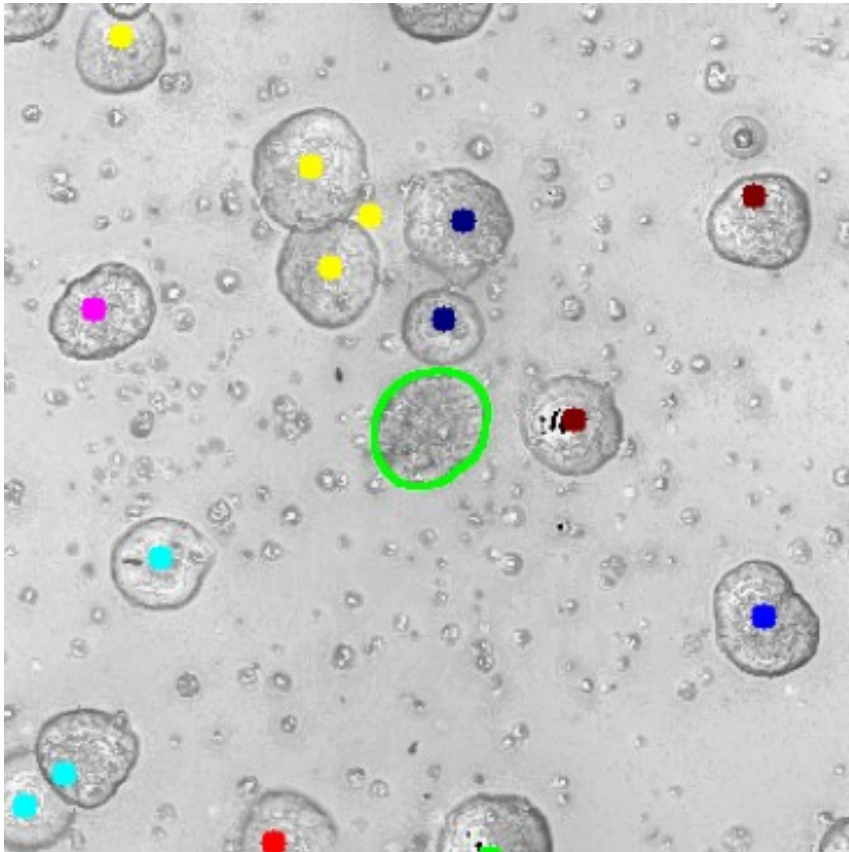
INPUT IMAGE



INSTANCES AND CENTROIDS

- Automatic prompts
- **SAM2 masks**
- Area filtering
- Morphological closing
- IoU-NMS

Masks provide both geometry and measurements



01

Mask $M_i(t)$

Which pixels belong to object i ?

02

Centroid $p_i(t)$

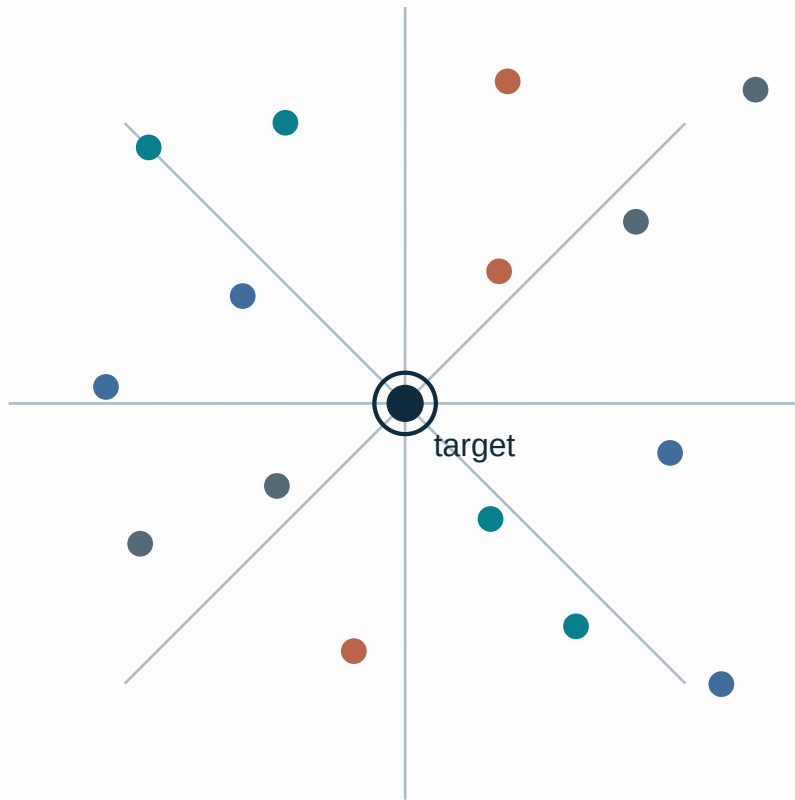
Where is the object located?

03

Morphology $m_i(t)$

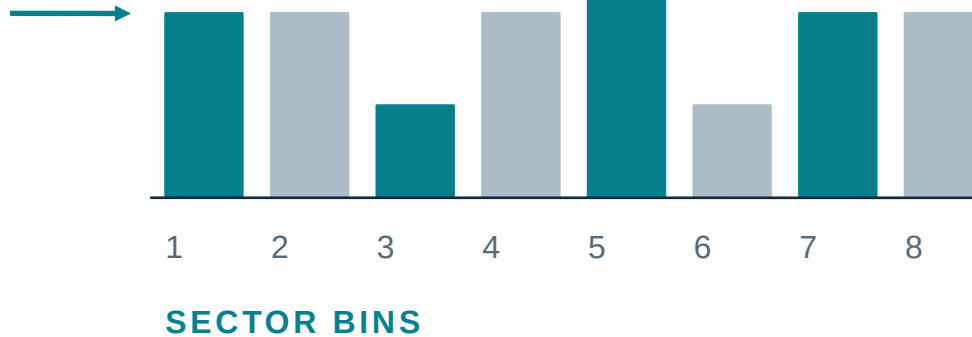
What can be measured from its mask?

HoR encodes where neighbors lie by direction

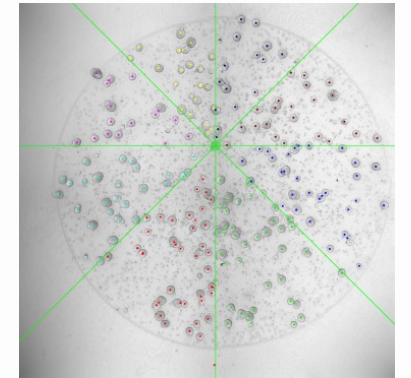


ANGULAR-SECTOR NEIGHBORHOOD

$$R = \text{Norm}(r_1, \dots, r_k)$$

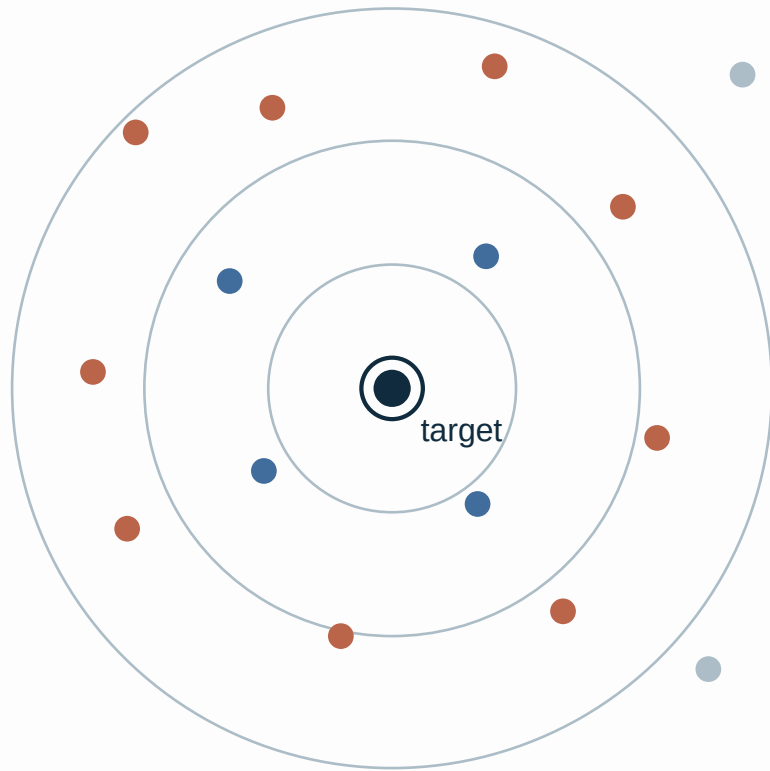


Count neighbors by direction.
Normalize the resulting histogram.



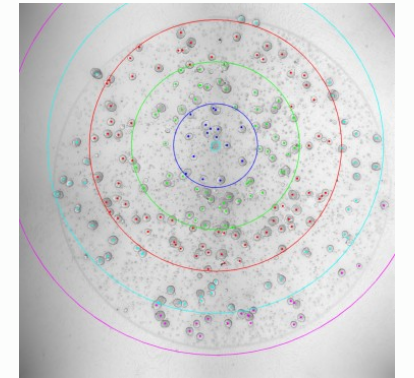
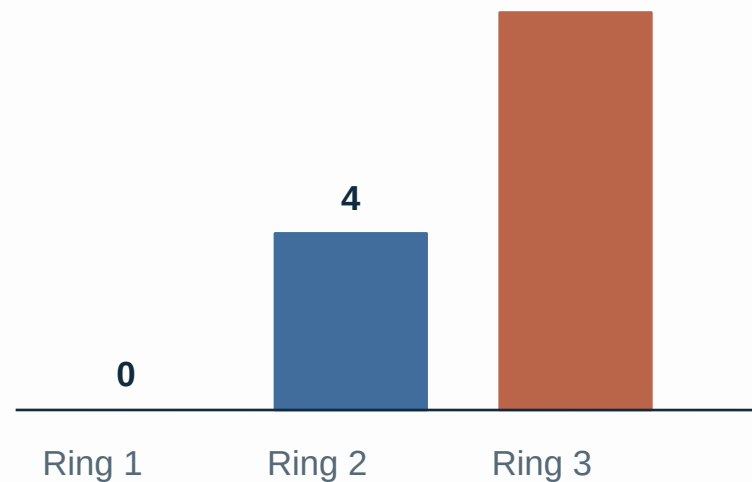
Original example

HoC encodes how far neighbors lie from the target



CONCENTRIC-RING NEIGHBORHOOD

$$\mathbf{C} = \text{Norm}(\mathbf{c}_1, \dots, \mathbf{c}_1)$$



Original example

Count neighbors by distance.
Complement the directional information.

Fuse neighborhood structure with morphology

$$F = [w_1 R_1, w_2 R_2, w_3 C_1, w_4 C_2, w_5 M]$$

$R_1 + R_2$

Directional context

$C_1 + C_2$

Distance context

M

Morphology / position

Complementary evidence is fused before geometric verification.

Relative geometry helps—but does not guarantee invariance

WHAT RELATIVE COORDINATES REMOVE

$$(p_j + b) - (p_i + b) = p_j - p_i$$

A shared translation cancels.

WHAT CAN STILL CHANGE THE DESCRIPTOR

Rotation

Moves angular-sector boundaries

Scale

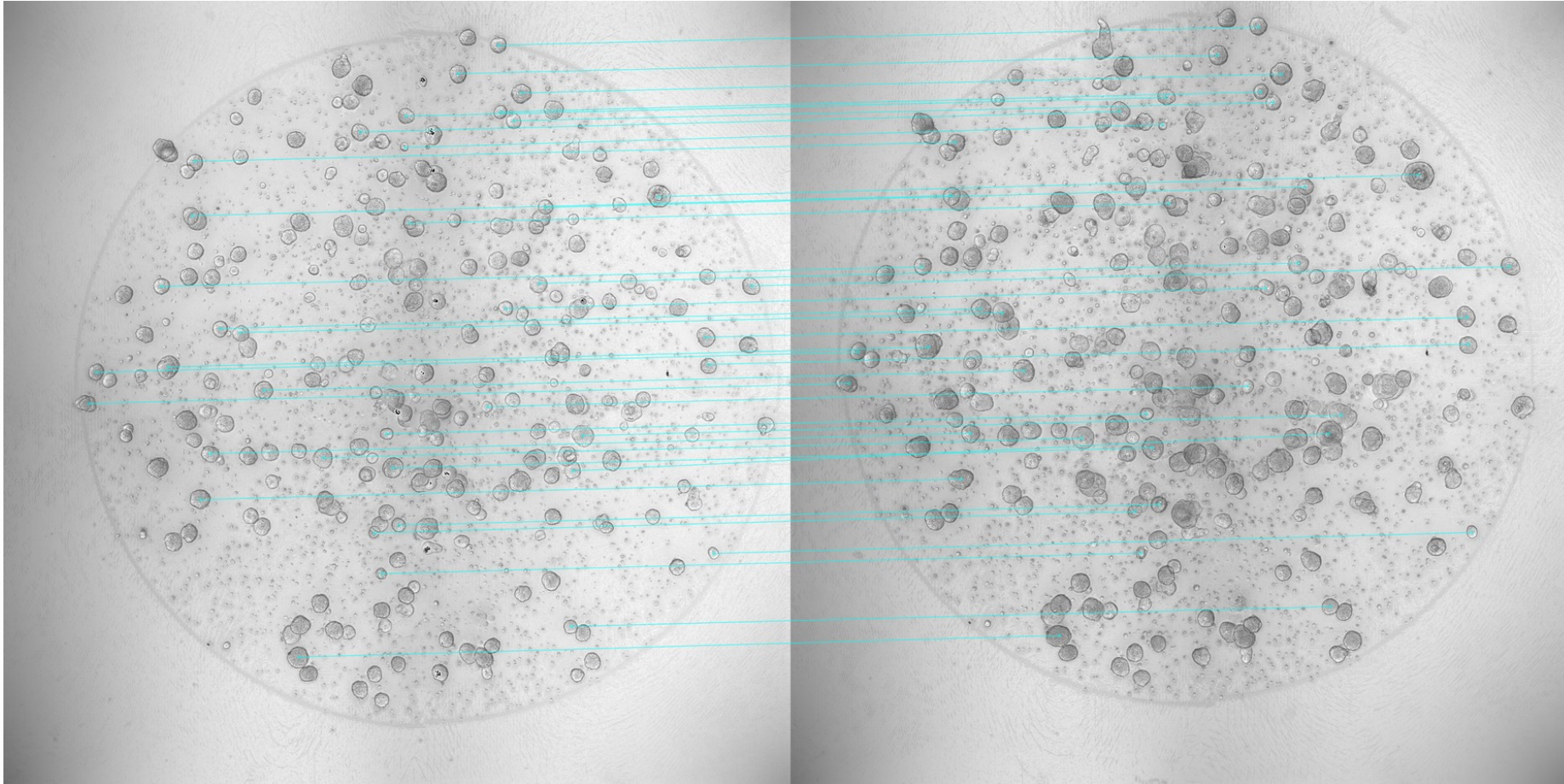
Moves concentric-ring boundaries

Missed objects

Changes neighborhood counts

Neighborhood geometry supplies a cue—not a universal identity certificate.

Candidate matches are hypotheses, not final identities



**Similar
descriptor
s
propose
matches.**

Geometry
must test
whether those
proposals
agree.

GC-RANSAC seeks a geometrically consistent subset

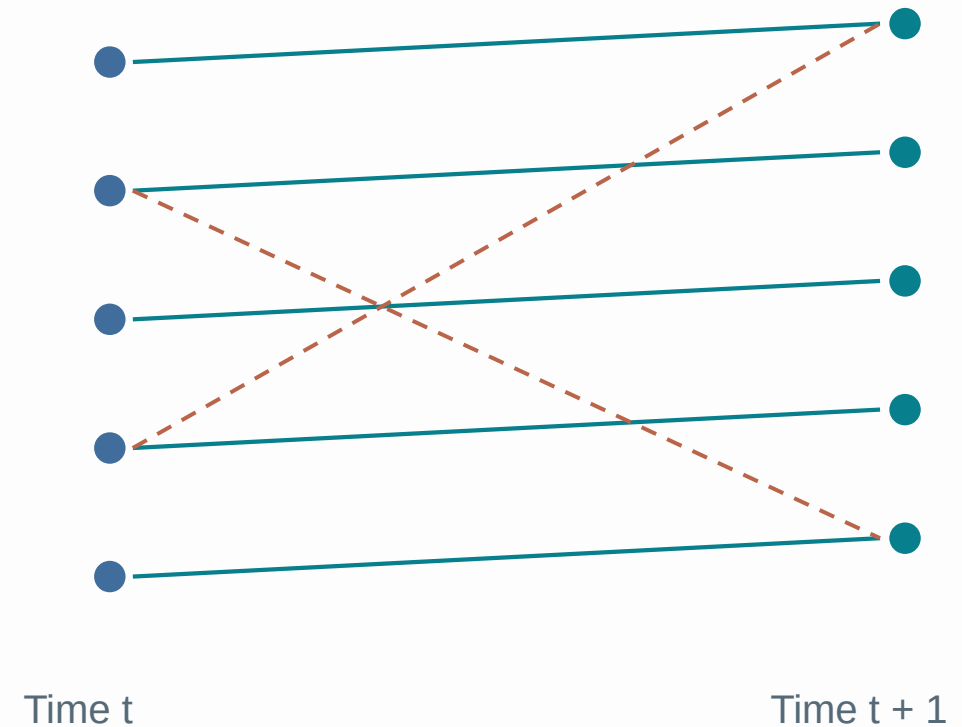
AFFINE MODEL

$$\mathbf{q} \approx \mathbf{A} \mathbf{p} + \mathbf{b}$$

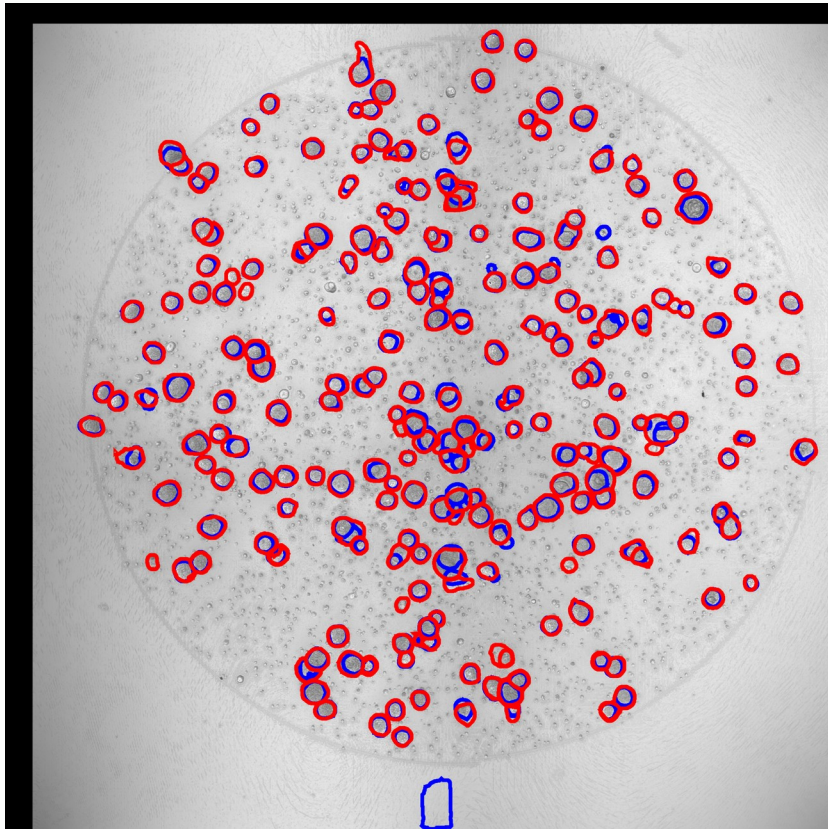
$$\text{residual} = \|\mathbf{A} \mathbf{p} + \mathbf{b} - \mathbf{q}\|$$

Seek a consistent inlier set.
Use spatial coherence in local refinement.

CONSISTENT VS. INCONSISTENT PROPOSALS



Align first; then solve the overlap assignment



ALIGNED MASKS

1 / MEASURE OVERLAP

$$\text{IoU}(M_i, M_j) = |M_i \cap M_j| / |M_i \cup M_j|$$

2 / FORM THE ASSIGNMENT COST

$$\text{cost}(i, j) = -\text{IoU}(M_i, M_j)$$

3 / ENFORCE ONE-TO-ONE CONSISTENCY

Hungarian assignment links objects across time.

Why a global assignment beats independent choices

ILLUSTRATIVE OVERLAP MATRIX

Time $t \rightarrow$	Object 1	Object 2
A	0.30	0.28
B	0.29	0.05

Both A and B prefer object 1 locally.
A consistent assignment must resolve that conflict.

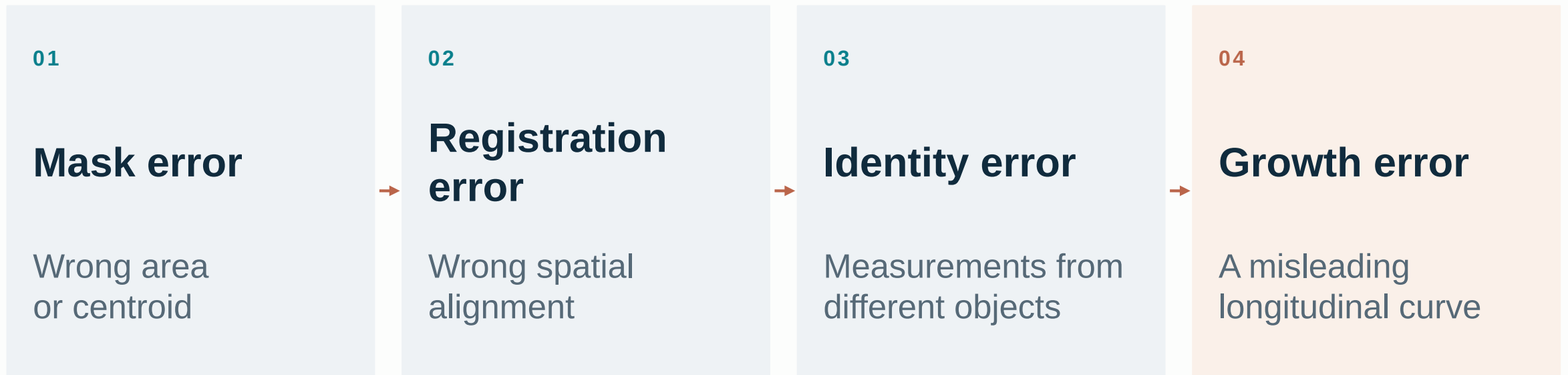
A $\rightarrow$ 1, B $\rightarrow$ 2

Total overlap = $0.30 + 0.05 = 0.35$

A $\rightarrow$ 2, B $\rightarrow$ 1

Total overlap = $0.28 + 0.29 = 0.57$

A correct curve requires more than a correct mask



Longitudinal validity depends on the entire chain.

Evaluate detection, association, and continuity separately

DETECTION + MASKS

DET / SEG

Were objects found?
Were their masks accurate?

LINKS + CONTINUITY

TRA / TF

Were identities linked correctly?
How much of each track stayed intact?

COMPLEMENTARY VIEWS

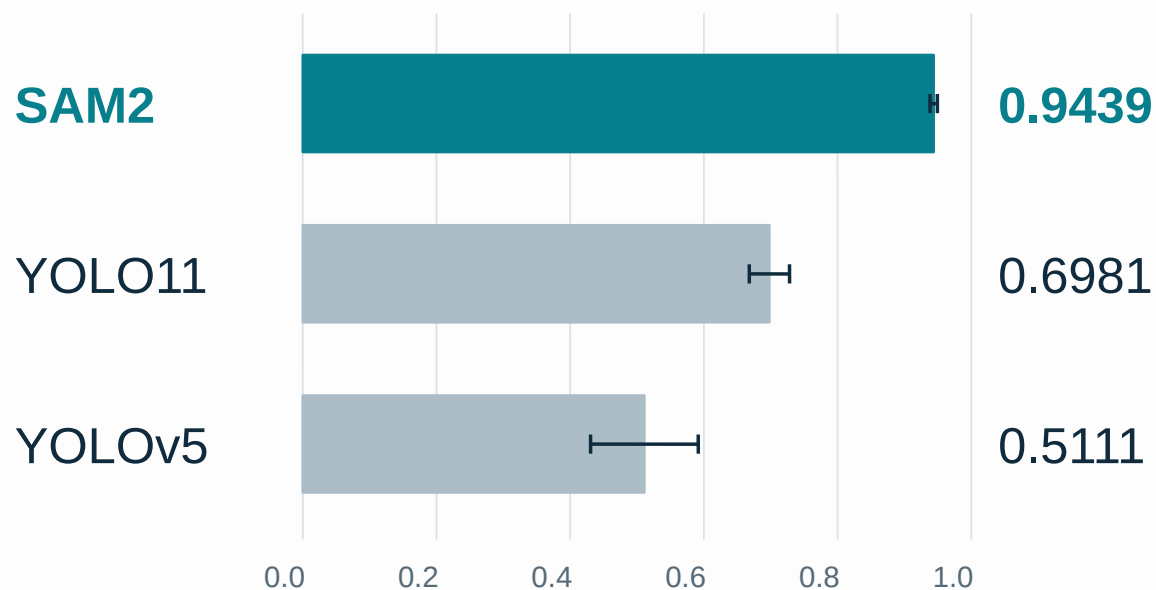
IDSW / MOTA

How often did identity switch?
What is the aggregate tracking score?

$$\text{OPCTB} = \frac{1}{2} (\text{SEG} + \text{TRA})$$

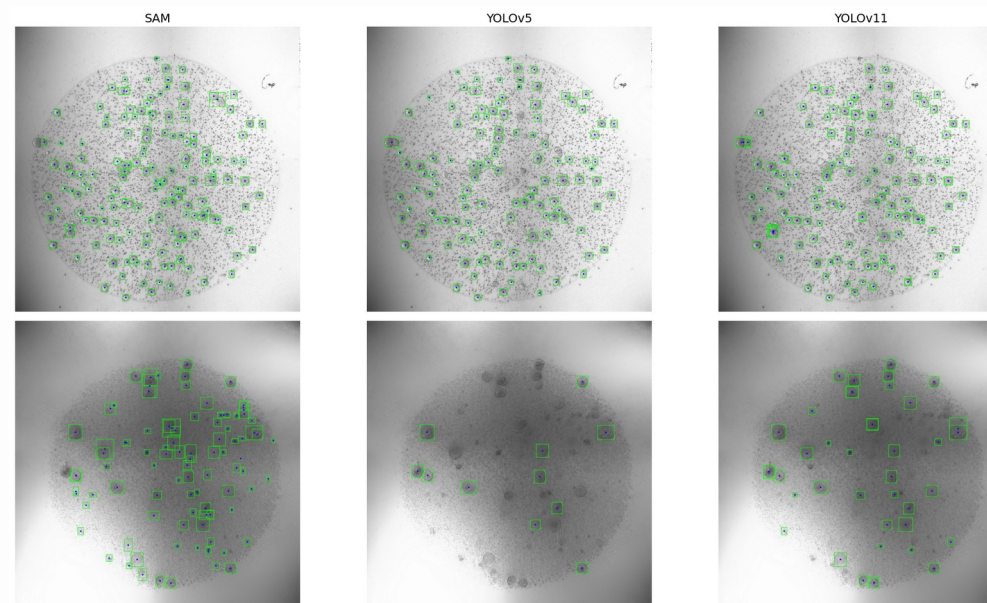
TF measures continuously recovered track fraction; it is not the complete-track score (CT).

SAM2 gives the highest reported DET in this comparison



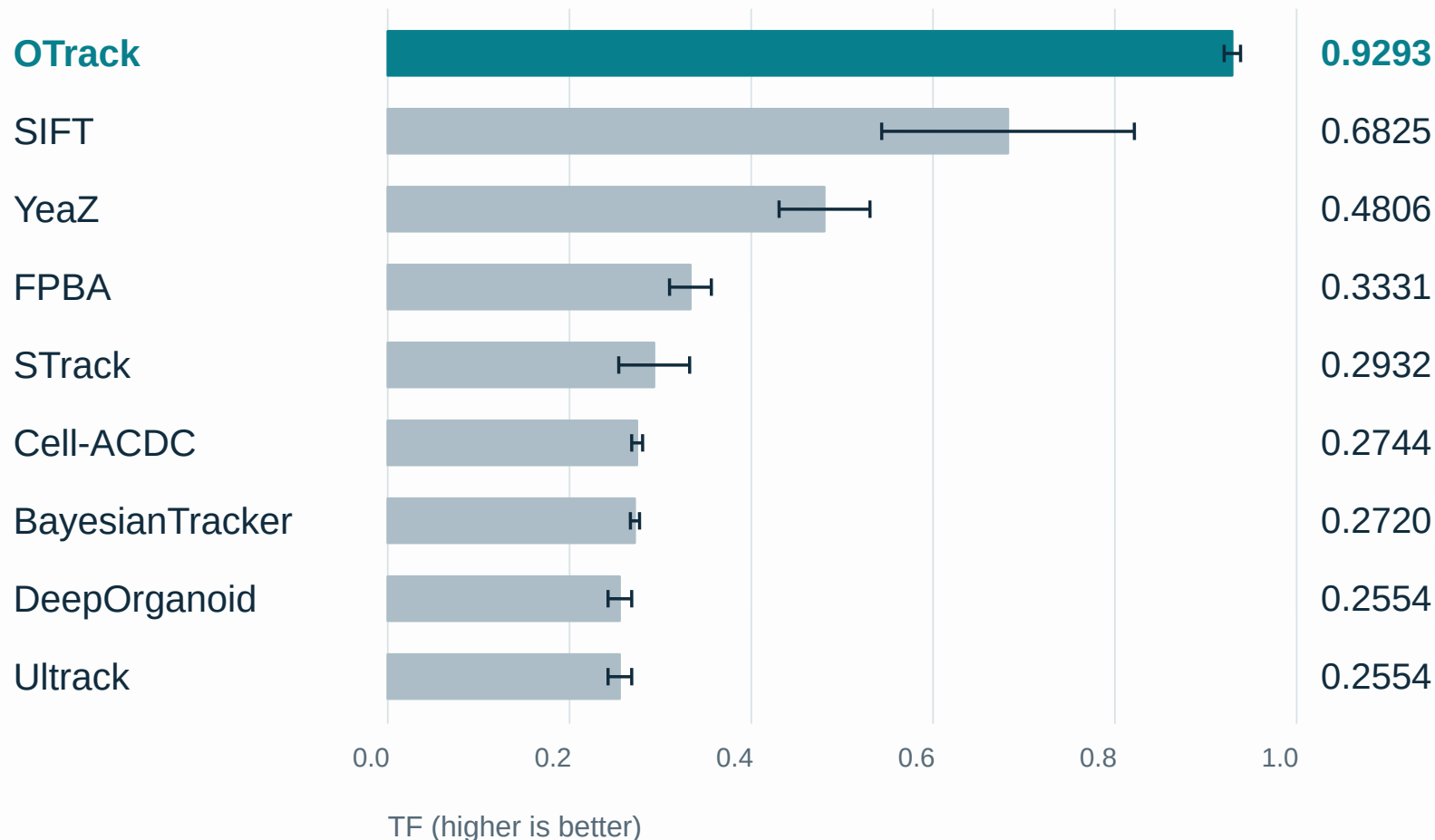
SAM2 DET · mean $\pm$ SD

0.9439 $\pm$ 0.0055



Original detection examples

The largest reported gain is in trajectory continuity

**0.9293**

OTrack TF

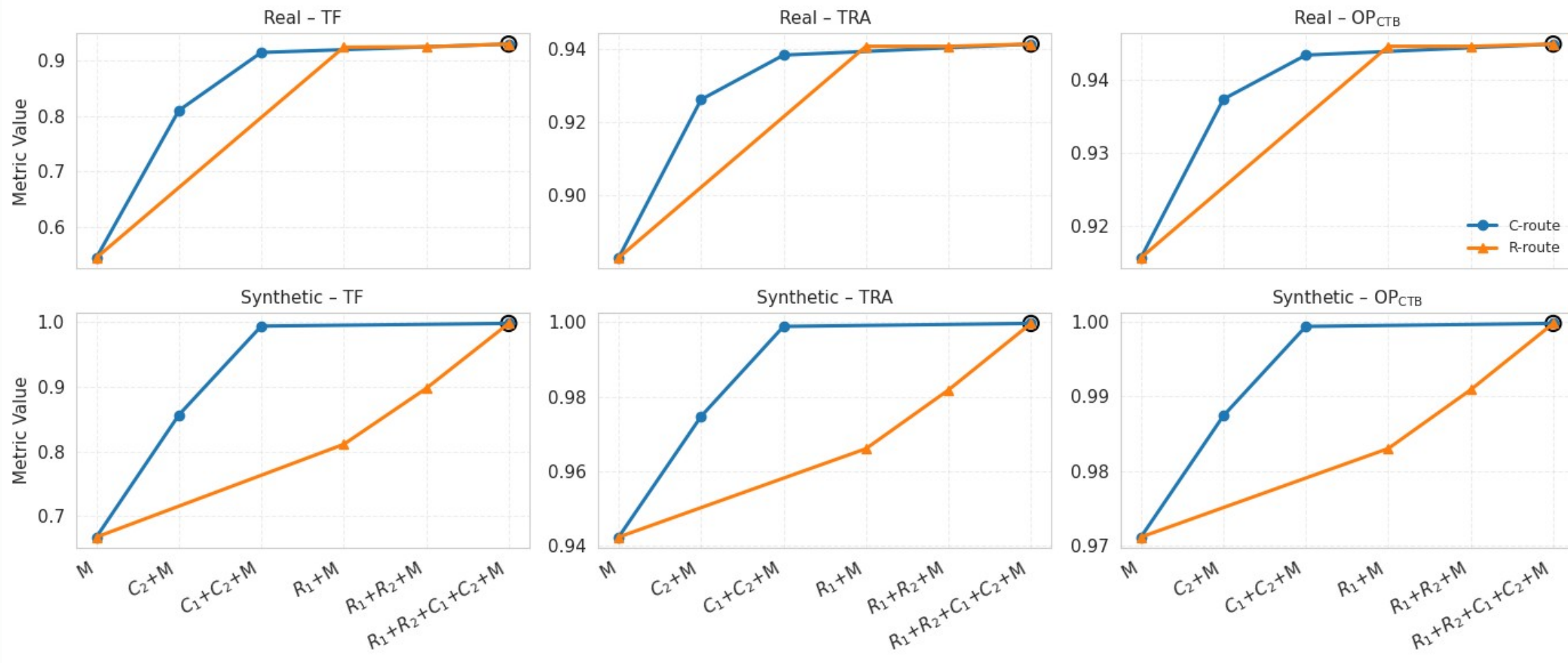
+24.68percentage points
vs. SIFT

Similar detection scores can accompany very different tracks

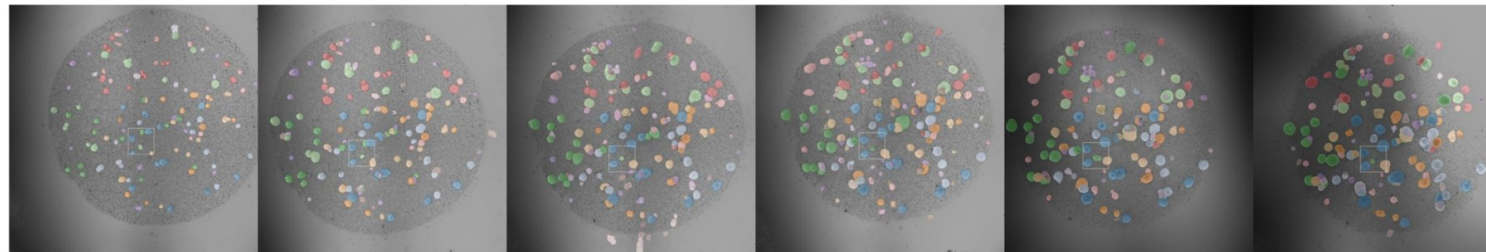
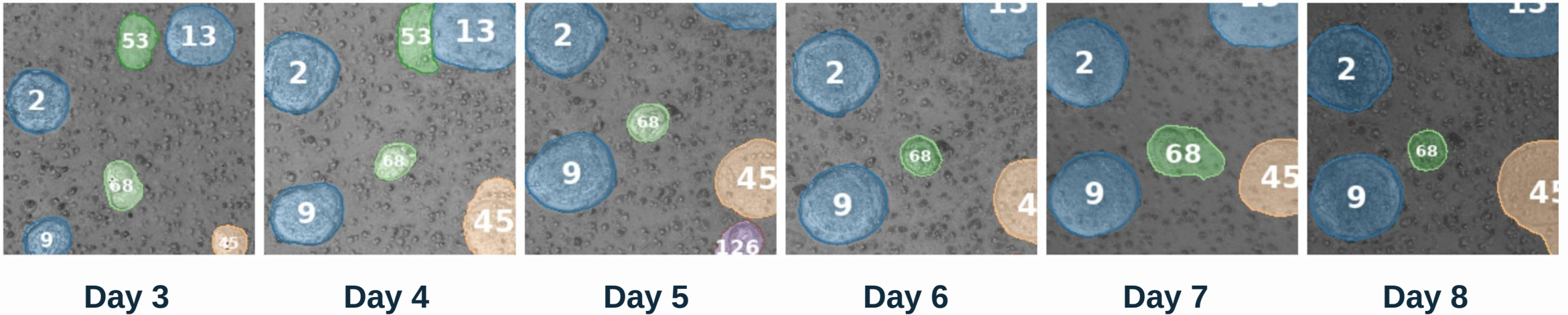
Metric	OTrack	SIFT	Difference
DET	0.9439	0.9439	0.0000
SEG	0.9485	0.9485	0.0000
TRA	0.9412	0.9061	+0.0351
TF	0.9293	0.6825	+0.2468
OPCTB	0.9448	0.9273	+0.0175

Frame-level scores alone do not describe longitudinal identity quality.

Ablation supports combining complementary descriptor blocks

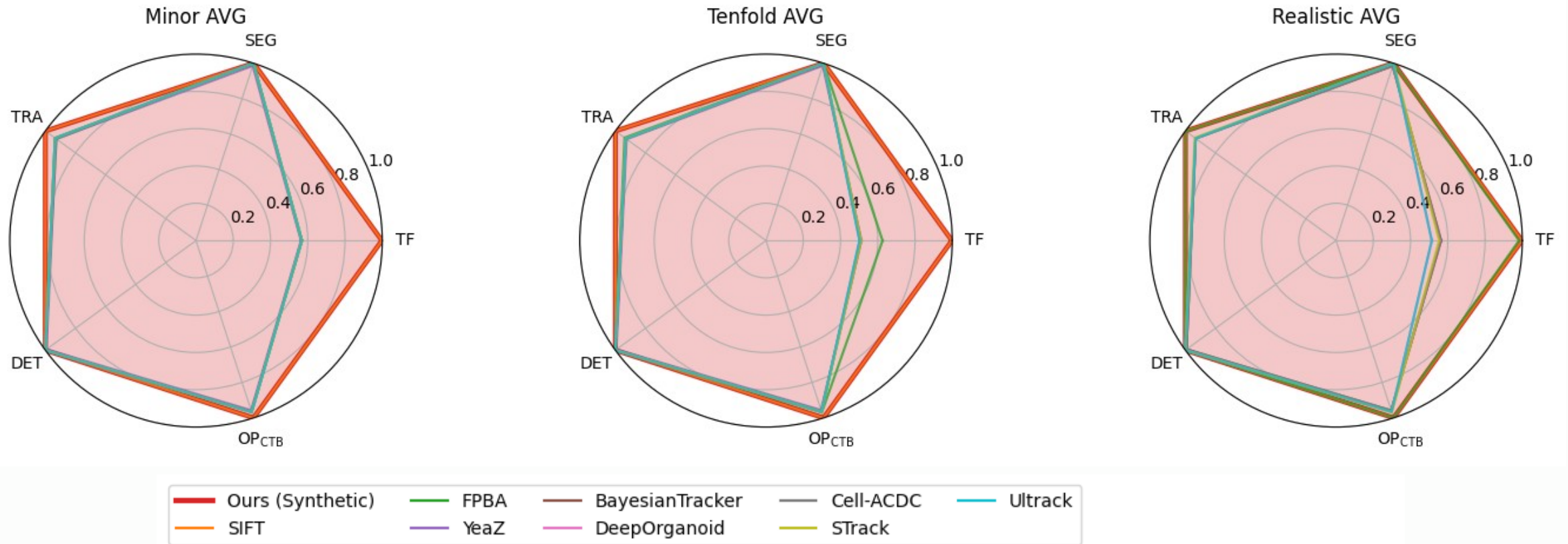


Qualitative trajectories connect the metrics to the images



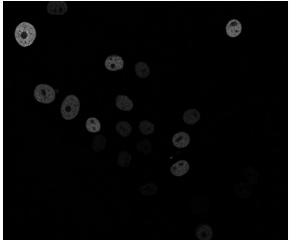
**Follow identity 68:
shape changes,
identity persists.**

Geometric stress tests probe one dimension of robustness

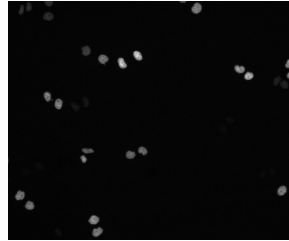


Geometric stress testing is not the same as biological or clinical validation.

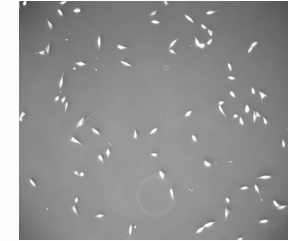
Transfer is tested on three public cell-image datasets



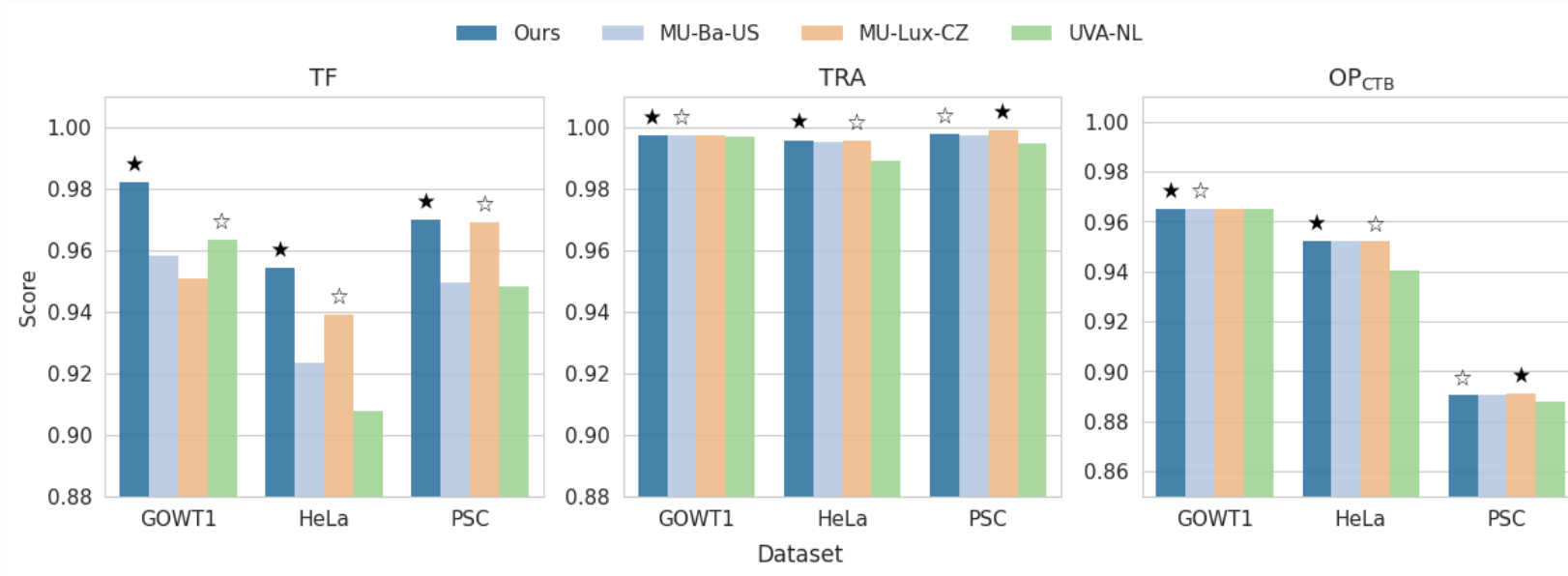
Fluo-N2DH-GOWT1



Fluo-N2DL-HeLa



PhC-C2DL-PSC

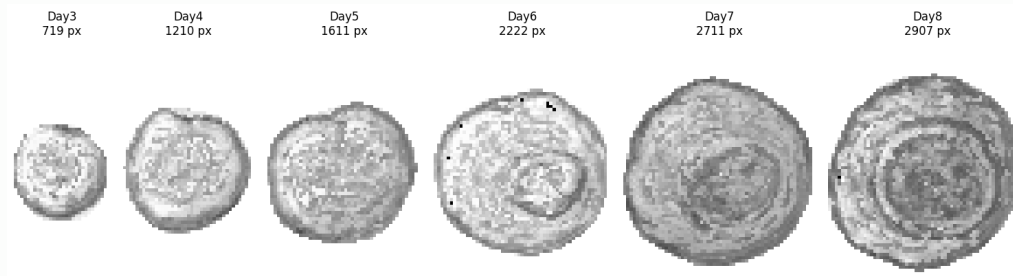


Tracking continuity transfers to these examples; rankings are metric-dependent.

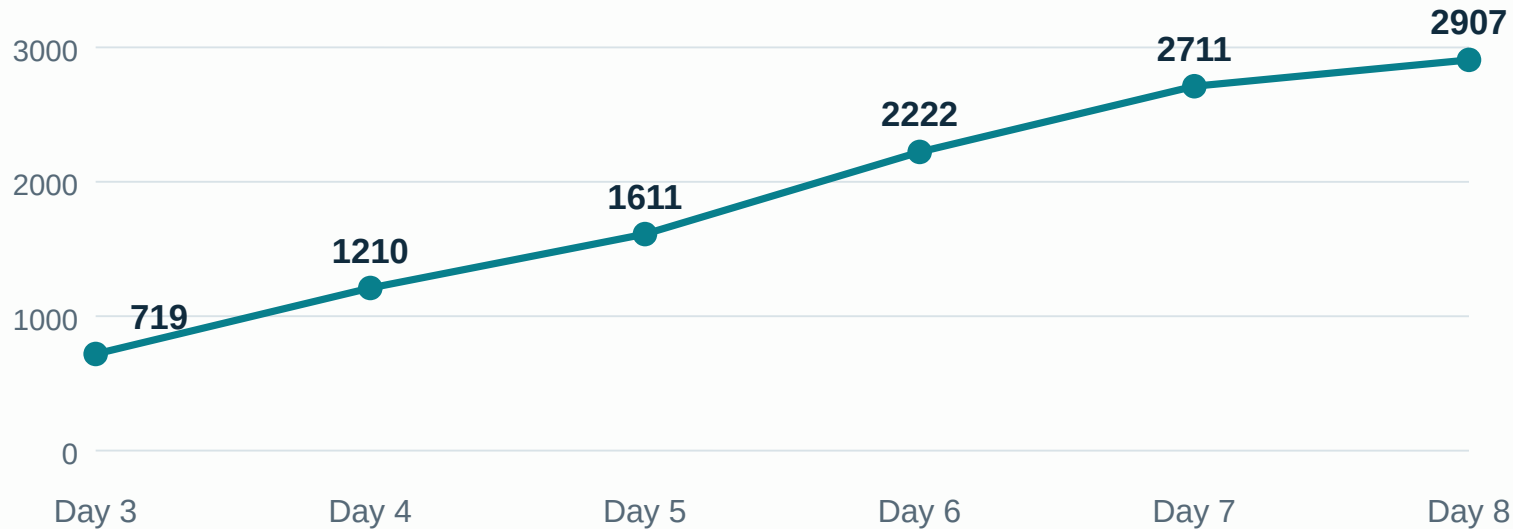
Identity switches reveal differences hidden by aggregate scores

Method	GOWT1	HeLa	PSC
OTrack	6	118	69
MU-Ba-US	7	236	126
MU-Lux-CZ	8	228	180
UVA-NL	6	379	369
GOWT1: OTrack and UVA-NL tie at 6 switches.			

A tracked example becomes a measurable growth history



Segmented area (pixels)



4.04x

Day 8 / Day 3

Segmented area
Not volume
Not cell count
Not a population effect

Identity unlocks interpretable longitudinal measurements

Area

$$A(t) = |M(t)|$$

Observed 2D footprint

Relative area

$$A(t) / A(t_0)$$

Change from the first observation

Interval change

$$[A(t_2) - A(t_1)] / (t_2 - t_1)$$

Change per unit time

Where the current approach is expected to struggle

01

Sparse neighborhoods

Less contextual information for matching.

02

Extreme local deformation

One affine model may be inadequate.

03

Missing or corrupted instances

Filtering and artifacts affect both cues and measurements.

04

2D observations

Area is not a direct measurement of 3D volume.

The next step is a stronger longitudinal measurement system

EXTEND THE OBSERVATION

3D and multimodal data

Move beyond the current 2D footprint.

STRENGTHEN RELIABILITY

Artifacts and uncertainty

Identify trajectories that need review.

SUPPORT SCALE

Efficiency and events

Measure runtime; clarify split/merge behavior.

A useful tracker should make its reliable operating region visible.

Thank you.

Questions and discussion

Liang Yang · Sichuan University

malyang@scu.edu.cn

JOINT WORK WITH

Binlin Chen · Jian Wang · Jiaqi He · Yuan Li · Jianfa Zhou
Linjie Guo · Haofeng Li · Bohui Chen