Term Project:
Nuclei Detection

Hae-in Jo, Seyun Jeong, Min-jik Kim, Min-sik Sung

Programming for Data Science (IMEN574/IDSC574)

1. Introduction (DSB 2018)
2. Data info.
3. Approaches We Tried
4. Pushing Further:
Foundation Model-Based Approach

1. Introduction (DSB 2018)

The Challenge: Create an algorithm to automate nucleus detection



Spot Nuclei, Speed Cures

- Locate cells in varied conditions to enable faster cures
- Free biologists to focus on solutions
- Improve throughput for research and insight
- Reduce time-to-market for new drugs – currently 10 years
- Increase # of compounds for experiments
- Improve health and increase quality of life

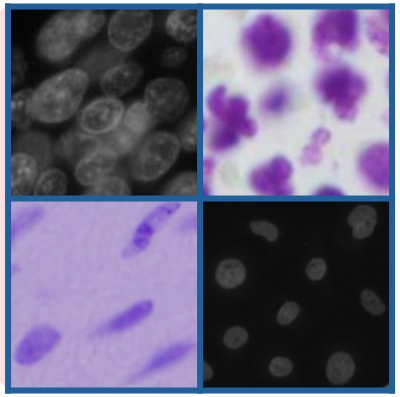
Introduction

Nuclei Detection



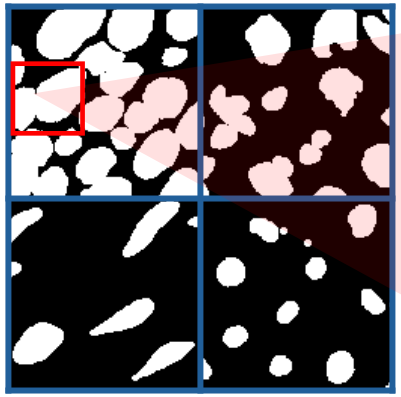
Microscope

Data acquisition



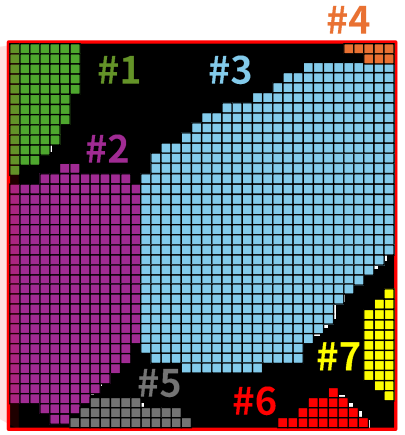
Nucleus

Segmentation



Masks

Instance separation



Magnified image

Run-length-encoding (RLE)

```
ID 1: 1464 10 1511 12 1556 17 1601 21 1643 3 1647 23
1690 29 1737 30 1784 31 1832 31 1879 32 1926 33 1974 33
2022 33 2070 33 2118 33 2165 34 2213 32 2261 31

ID 2: 1729 4 1777 5 1825 5 1873 5 1921 5 1969 5 2017 5
2065 5 2113 5 2161 4 2209 4 2257 4

ID 3: 1102 3 1148 5 1195 6 1242 7 1289 8 1336 9 1384 9
1431 10 1479 10 1526 11 1574 11 1623 10 1672 9 1720 9
1769 8 1817 8 1866 7 1915 5 1963 3 2012 2

ID 4: 398 6 445 9 483 20 529 23 577 24 625 25 673 29
721 30 769 30 817 30 865 30 913 31 961 31 1009 31 1057
31 1105 31 1153 31 1201 31 1249 30 1297 28 1345 27 1393
26 1441 23 1489 22 1537 19 1589 12

ID 5: 1 23 49 23 97 22 145 22 193 21 241 20 289 18 337
17 388 8

ID 6: 27 22 75 22 124 20 172 20 220 19 269 17 320 11
370 5
```

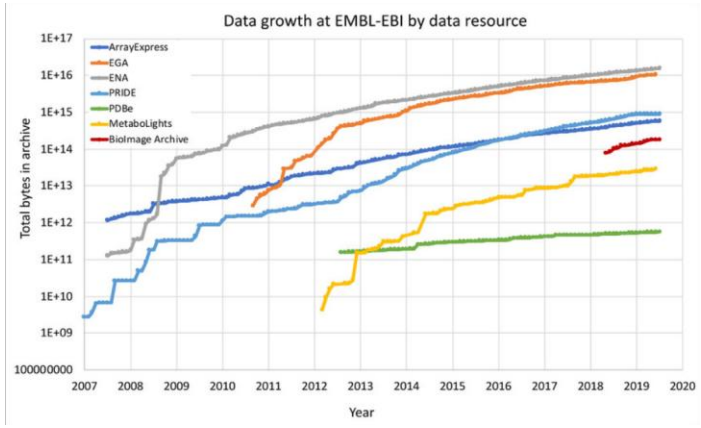
- **Objective & Scope**

Detect and segment individual cell nuclei in microscopy images, generating a separate mask for each nucleus.

Why Nuclei?: Spot Nuclei. Speed Cures.

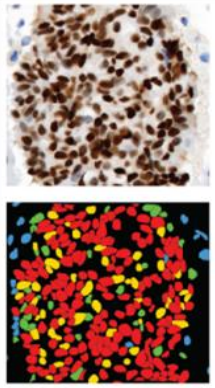
Technical Bottlenecks

- **Data Explosion**: HCS produces terabytes of images daily; manual boundary drawing by experts is physically impossible.
- **Extreme Heterogeneity**: Nuclei exhibit immense morphological variance in size and shape across distinct organisms.
- **Optical Generalization**: Algorithms must adapt to diverse noise, varied stains (Fluorescence vs. H&E), and densely clustered cells.



Clinical & Scientific Impact

- **Accelerated Discovery**: Rapidly quantifies experimental compound efficacy, drastically reducing the 10-year drug development cycle.
- **Precision Medicine**: Enables highly automated, objective disease grading in digital pathology (e.g., Allred scoring).
- **Resource Optimization**: Eliminates human bias and fatigue, freeing biologists to focus purely on finding cures.



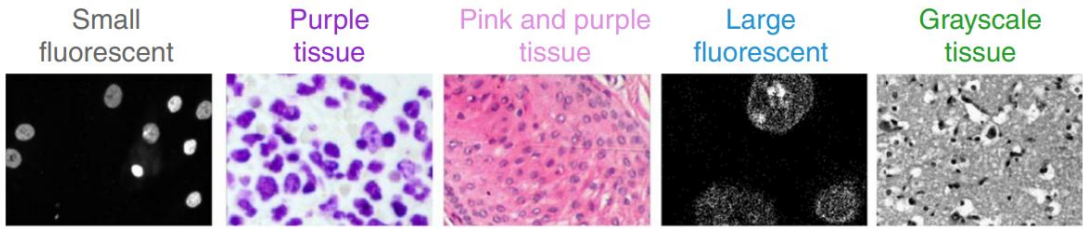
Total Nuclei	**279**
Negative	19
Weak	33
Moderate	51
Strong	176
Total Positive	**260 (93%)**
P Score	5
I Score	3
Allred Score	**8**

Feature of Data Science Bowl (DSB) 2018

Extreme Sample Diversity

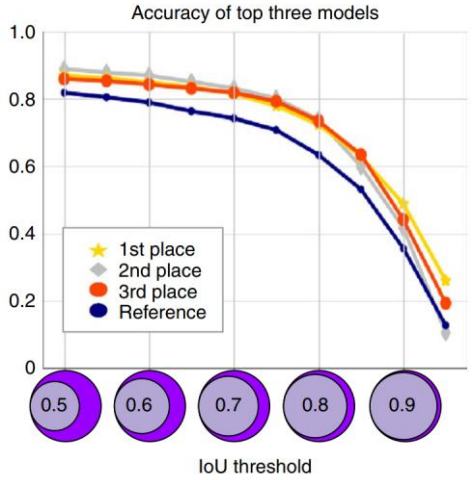
Requires models to generalize across highly heterogeneous conditions, preventing simple overfitting

- Organisms: Human, mouse, and fly cells.
- Modalities: Brightfield, fluorescence, and histology imaging.
- Conditions: Varied magnification and cell density.



Precision Evaluation (mAP)

Submissions evaluated on Mean Average Precision (mAP) via Intersection over Union (IoU)



$$IoU(A,B) = \frac{A \cap B}{A \cup B}$$

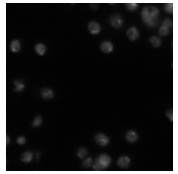
$$\frac{1}{|thresholds|} \sum_{t} \frac{TP(t)}{TP(t) + FP(t) + FN(t)}$$

Strict Metric: Evaluation sweeps IoU thresholds from 0.5 to 0.95 (step 0.05), penalizing poor localization.

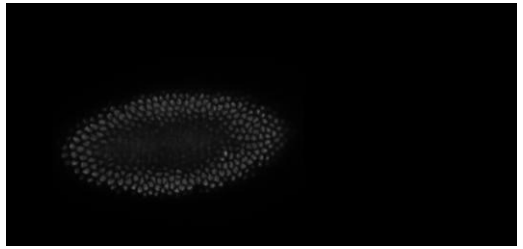
2. Data info.

Data Info.

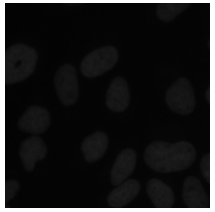
Dataset for Training & Validation



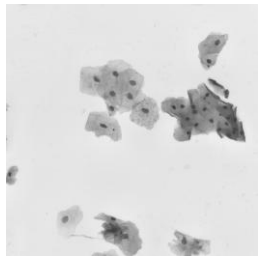
256x256



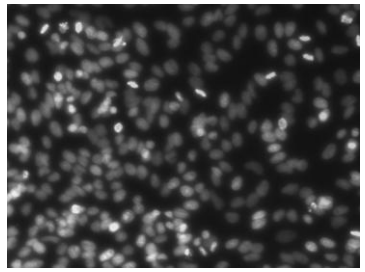
1272x603



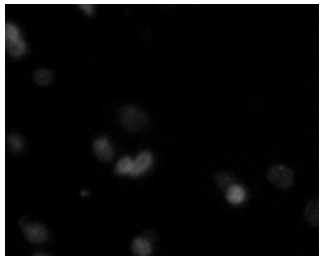
360x360



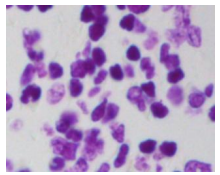
1024x1024



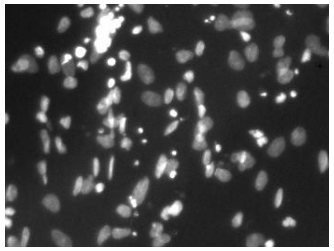
696x520



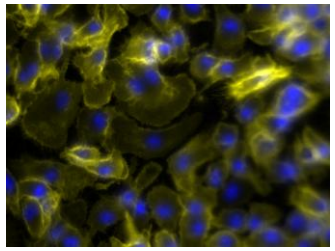
640x512



300x256



347x260



1388x1040

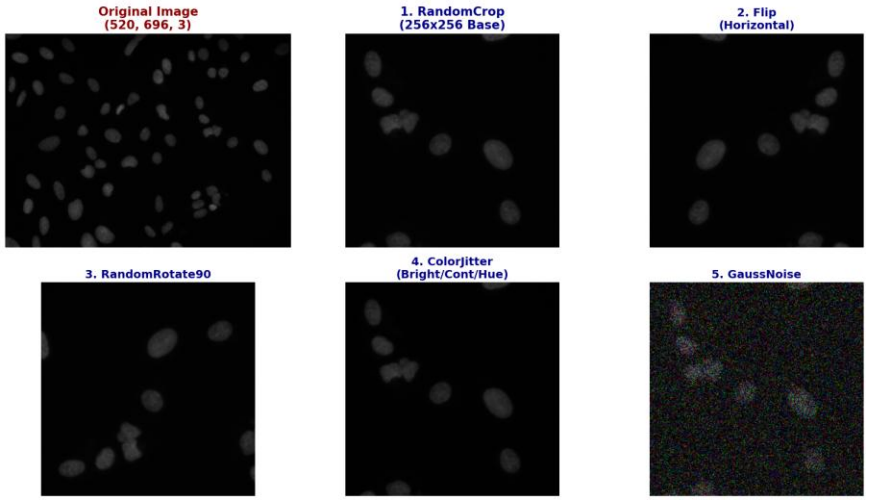
Training & Validation		Stage #1		Stage #2	
Size, (H*W)	# (ea)	Size, (H*W)	# (ea)	Size, (H*W)	# (ea)
256 x 256	340	256 x 256	24	256 x 256	1176
256 x 300	112	512 x 640	8	520 x 696	495
260 x 347	5	512 x 680	8	574 x 574	384
360 x 360	91	260 x 347	4	256 x 320	240
512 x 640	13	519 x 253	4	512 x 512	215
520 x 696	92	520 x 348	4	⋮	
603 x 1272	6	520 x 696	4	364 x 345	411
1024 x 1024	16	524 x 348	4	448 x 411	538
1040 x 1388	1	519 x 161	2	531 x 538	346
		519 x 162	2	553 x 346	556
		390 x 239	1	994 x 556	1360
Total	676	Total	65	Total	3019

→ The final score is evaluated based only on the results for Stage 2 dataset, which contains various size

3. Approaches We Tried

Method: U-Net

Augmentation



5 techniques randomly applied per epoch (p=0.5)

U-Net ++

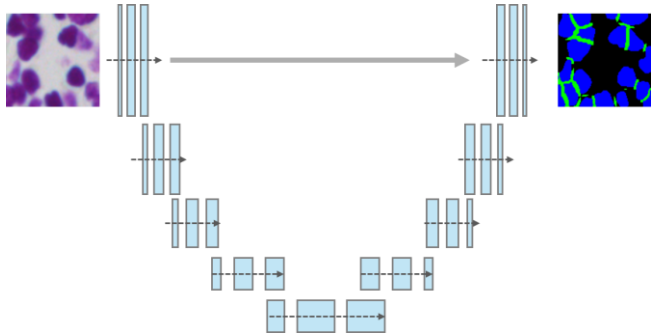
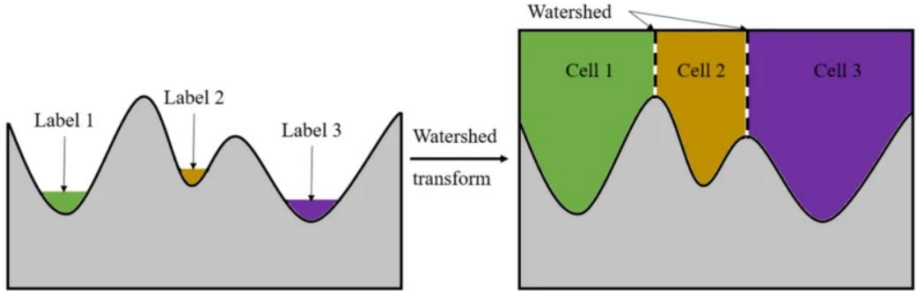


Table 3: Segmentation results (IoU: %) for U-Net, wide U-Net and our suggested architecture UNet++ with and without deep supervision (DS).

Architecture	Params	Dataset			
		cell nuclei	colon polyp	liver	lung nodule
U-Net [9]	7.76M	90.77	30.08	76.62	71.47
Wide U-Net	9.13M	90.92	30.14	76.58	73.38
UNet++ w/o DS	9.04M	**92.63**	**33.45**	79.70	76.44
UNet++ w/ DS	9.04M	92.52	32.12	**82.90**	**77.21**

- Encoder (ResNet-34)
Fast convergence & low-level feature extraction.
- Decoder (U-Net++)
Intermediate nodes bridge semantic gaps. High precision for small, overlapping nuclei.
- Output: sigmoid prediction

Watershed algorithm

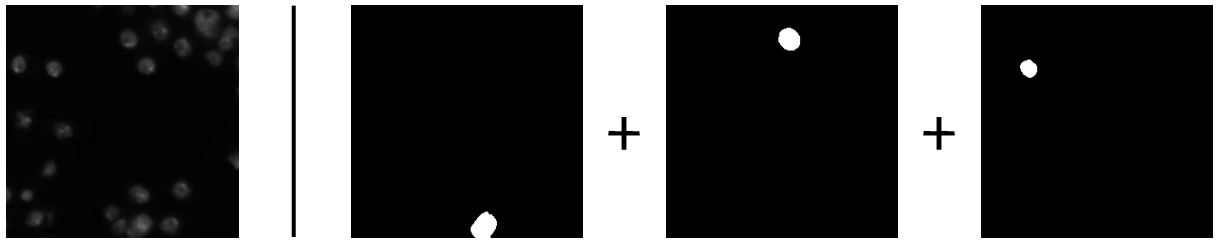


- **Topographic View**: Treats pixel values (e.g., distance transform) as elevation.
- **Flooding**: "Water" fills basins starting from local minima (Label 1, 2, 3).
- **Watershed Lines**: Boundaries are formed exactly where water from different basins meets.
- **Instance Separation**: Effectively splits touching or overlapping cells into individual instances.

Approach 1: Single-Channel UNet++ (baseline)

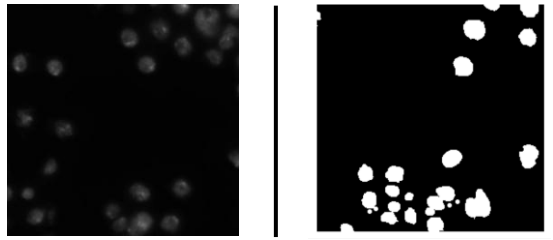
Data pipeline in training stage

Dataset given



Individual masks

Training Target

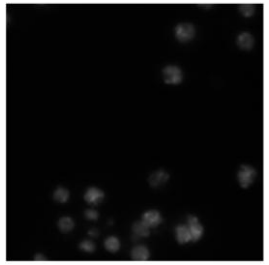


Single binary union

Pixels belonging to any nucleus → 1
Otherwise → 0

Approach 1: Single-Channel UNet++ (baseline)

Data pipeline in test stage

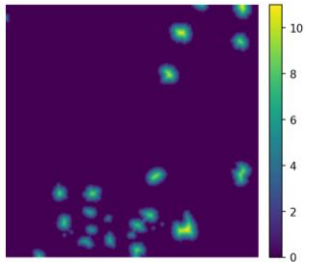


Raw Image



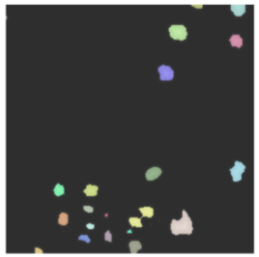
Model Prediction

- Pixels with nucleus probability >= 0.5 → set to 1
- Otherwise → set to 0



Distance Transform

- Distance from each pixel to the closest background pixel(0) → Peak at each nucleus center
- Touching nuclei produce separate peaks



Seeds

Pixels with distance > 30% of the maximum distance → extracted as seeds



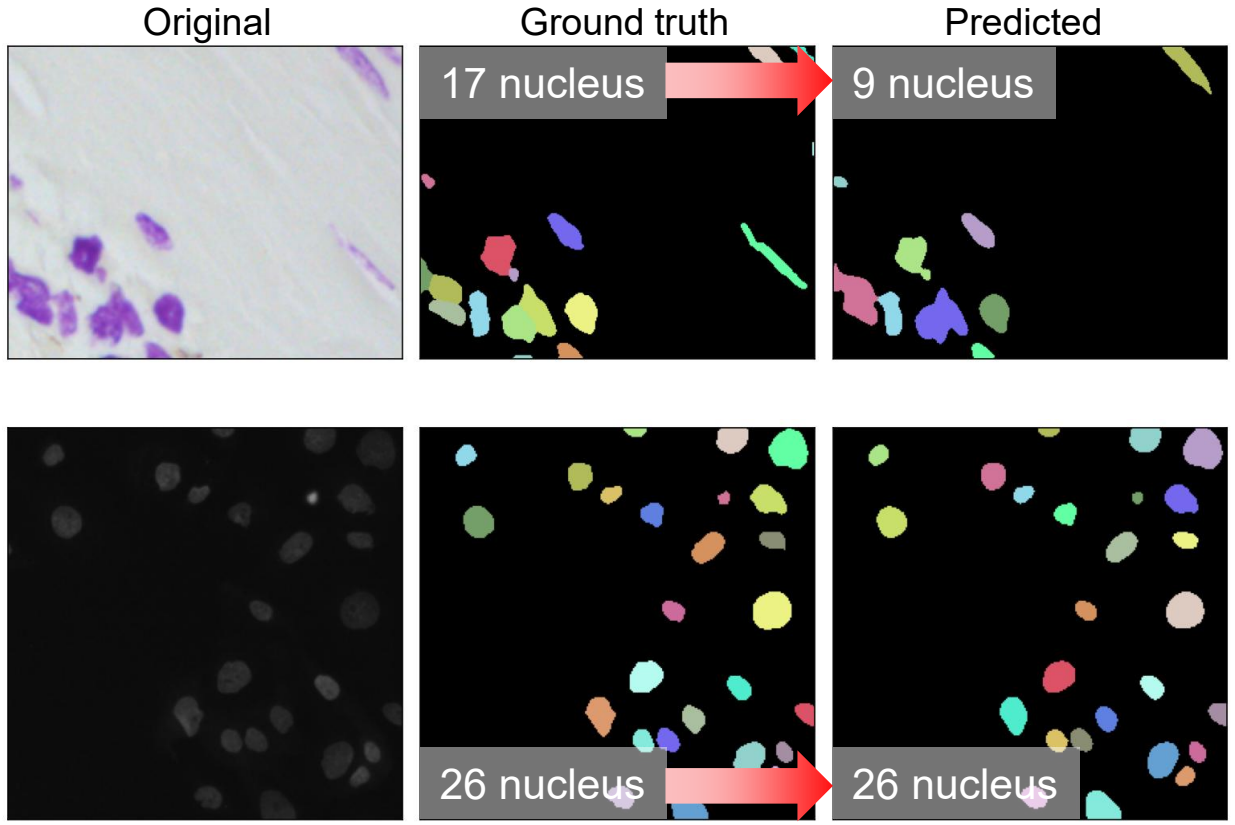
Instance Map Each nucleus assigned a unique integer ID

Starting from each seed, regions are expanded until two neighboring regions meet → the boundary between them is finalized

Approach 1: Result of Single-Channel UNet++ (baseline)

Quantitative Results

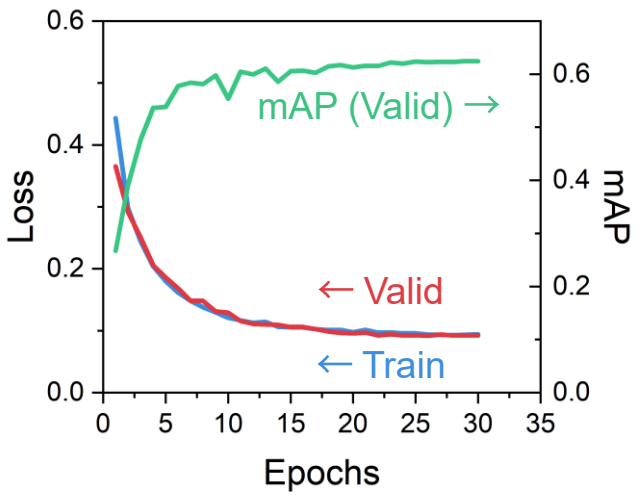
✓ Visualization of Original/GT/Pred images



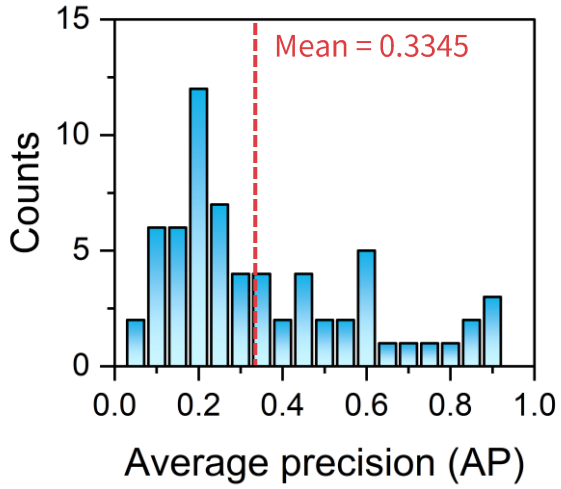
mAP: 0.46540

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

✓ Epoch-loss



✓ AP distribution (stage 1-test)

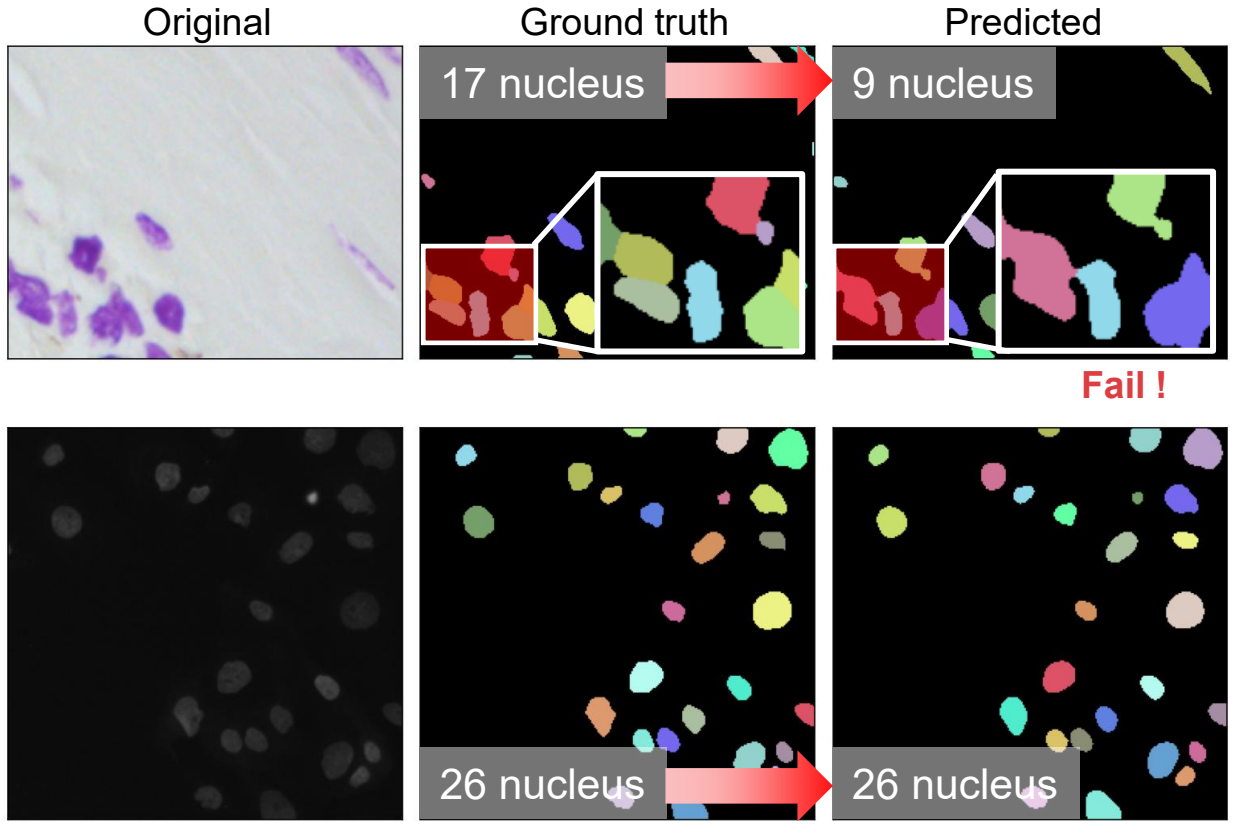


- When many cells are adjacent to one another, the model fails to accurately distinguish individual cells and detects multiple attached cells as a single instance, resulting in lower performance and accuracy.

Approach 1: Result of Single-Channel UNet++ (baseline)

Quantitative Results

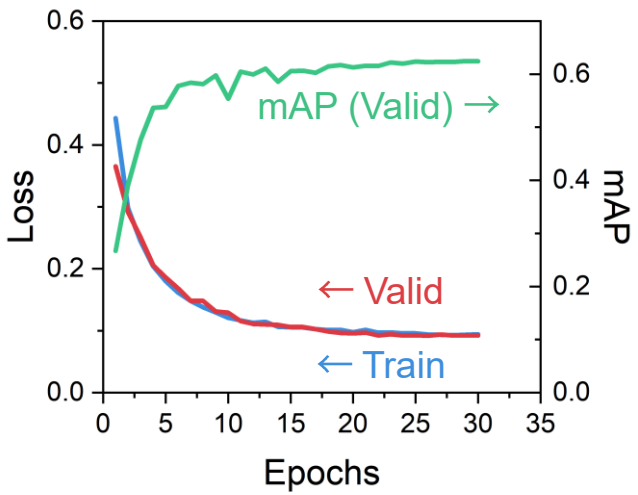
✓ Visualization of Original/GT/Pred images



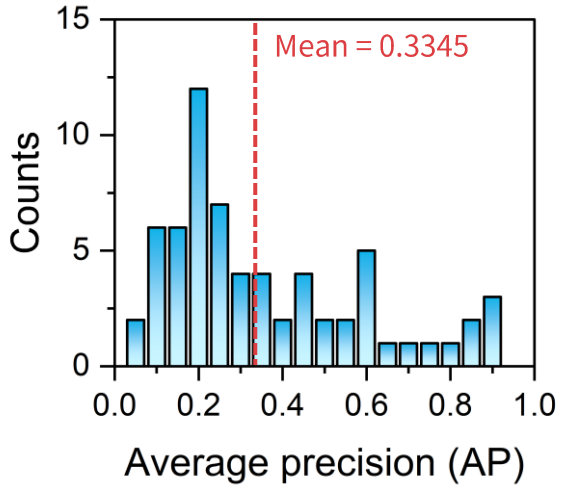
mAP: **0.46540**

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

✓ Epoch-loss



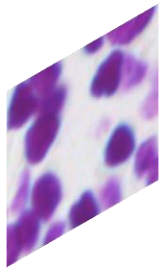
✓ AP distribution (stage 1-test)



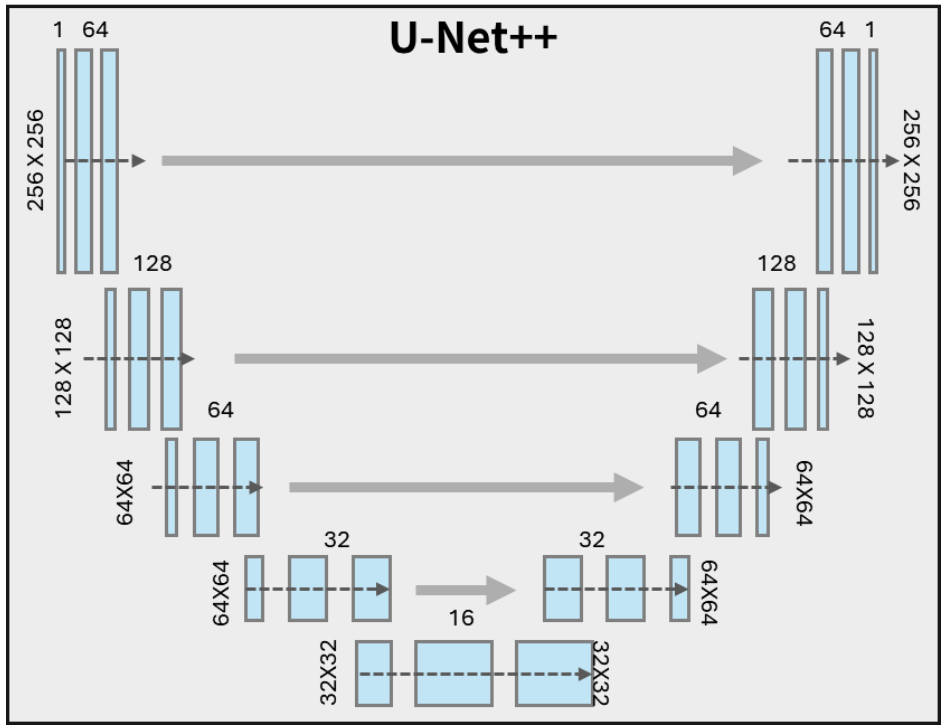
- When many cells are adjacent to one another, the model fails to accurately distinguish individual cells and detects multiple attached cells as a single instance, resulting in lower performance and accuracy.

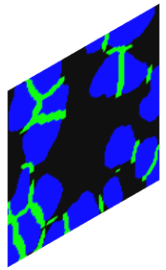
Approach 2: Two-channel UNet++

Why Use a 2-channel Mask ?



Original





2 ch. mask

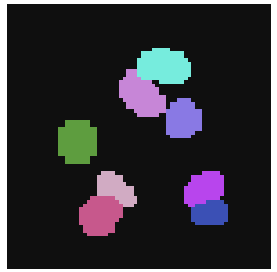
0th ch.: body

1th ch.: border

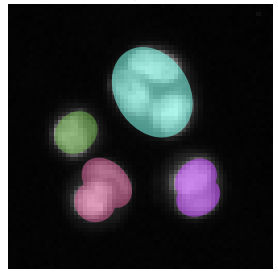
What is different ?



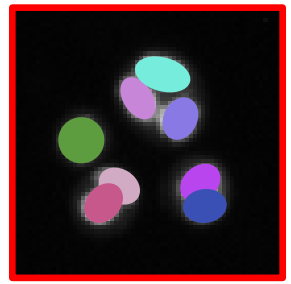
Watershed algorithm for separating each nuclei ...



Ground truth



Only body (1 ch.)



body + border (2 ch.)

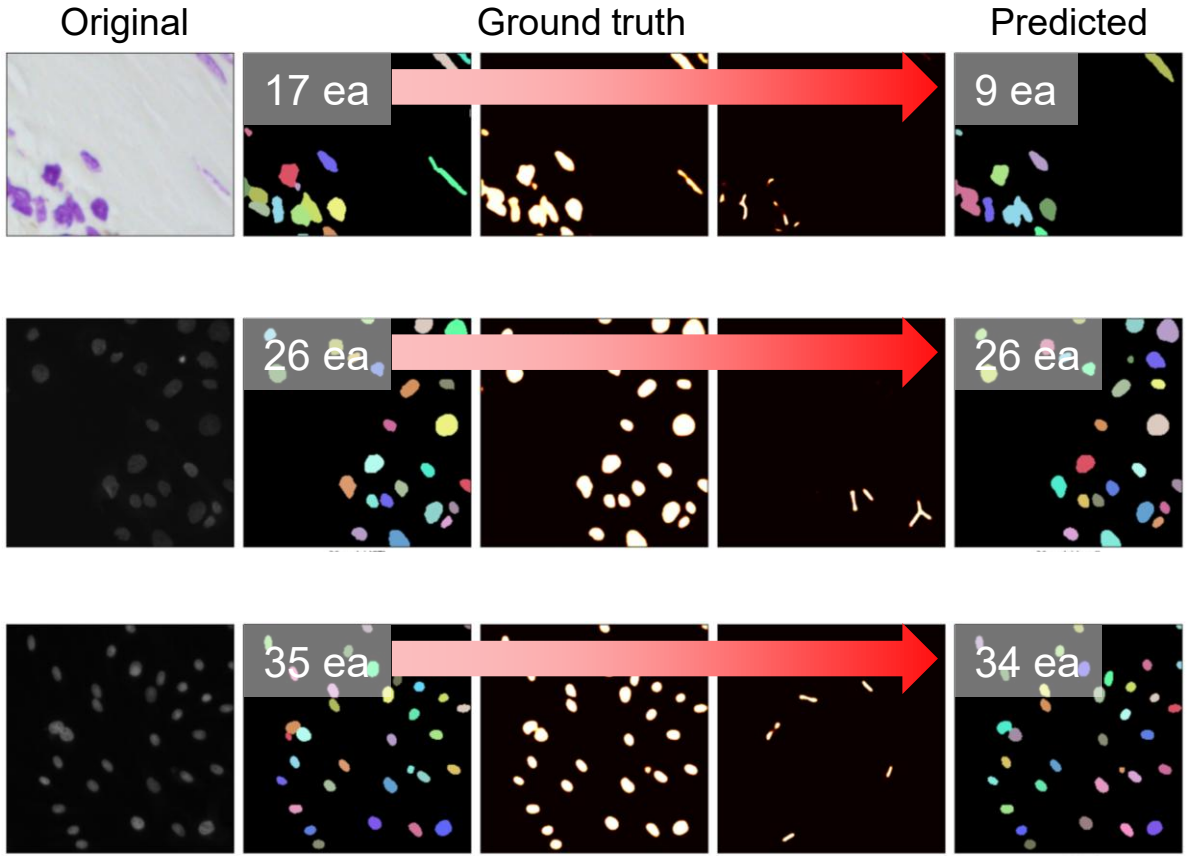
The border mask physically cuts the connection between two touching nuclei before the watershed even starts.

Approach 2: Result of Two-Channel UNet++

(previous) 0.46540

Quantitative Results

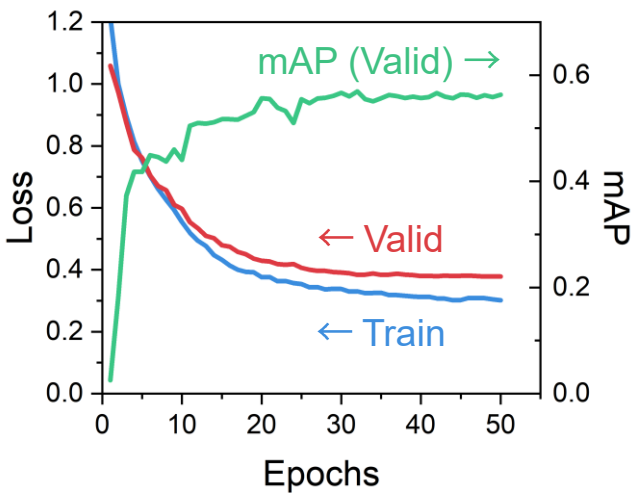
✓ Visualization of Original/GT/Pred images



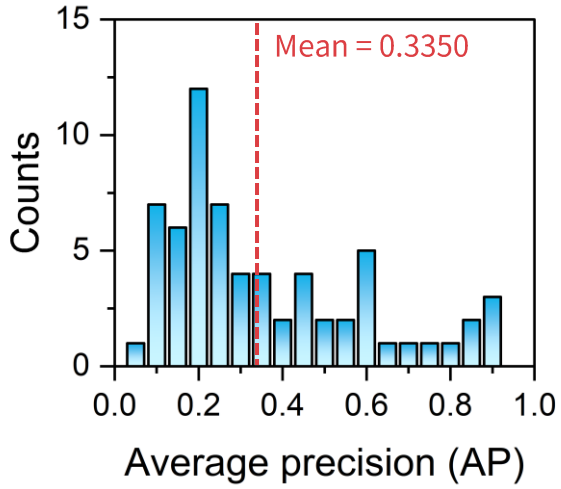
mAP: **0.49670 (↑6.3%)**

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

✓ Epoch-loss



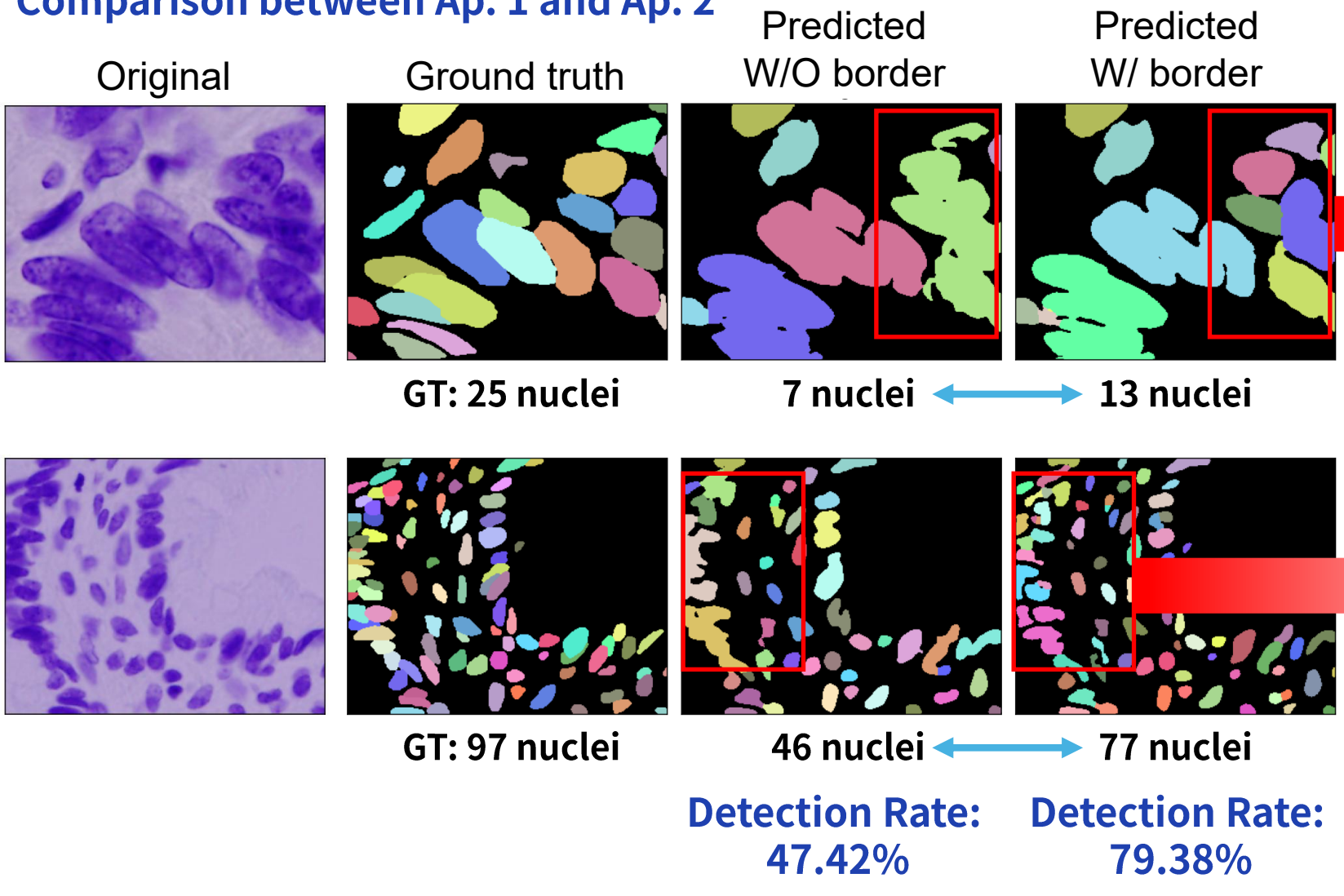
✓ AP distribution (stage 1-test)



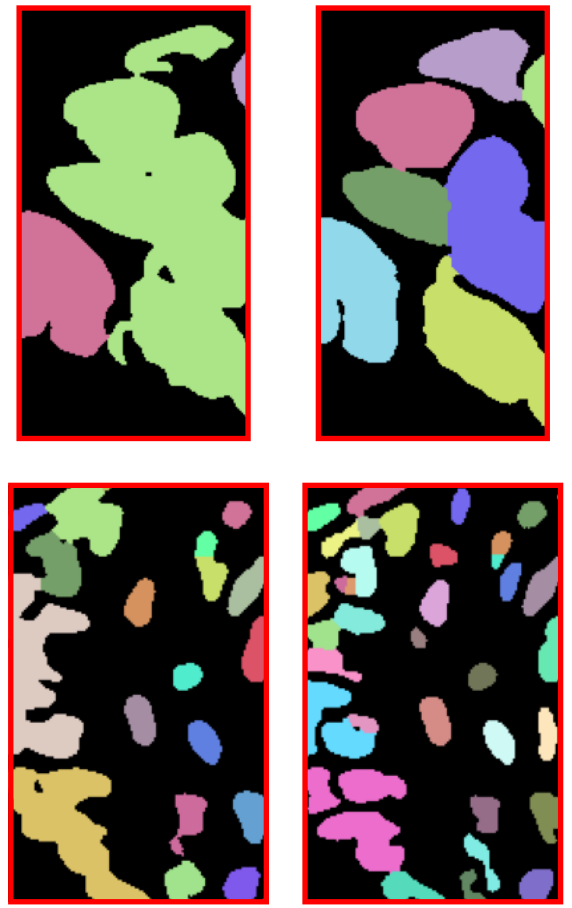
- Compared to the model trained without boundary masks, it shows slightly higher performance on the validation and Stage 1 test sets, but demonstrates a remarkable accuracy improvement of 6.3% on the Stage 2 test set of 3,019 images.

Approach 2: Result of Two-Channel UNet++

Comparison between Ap. 1 and Ap. 2

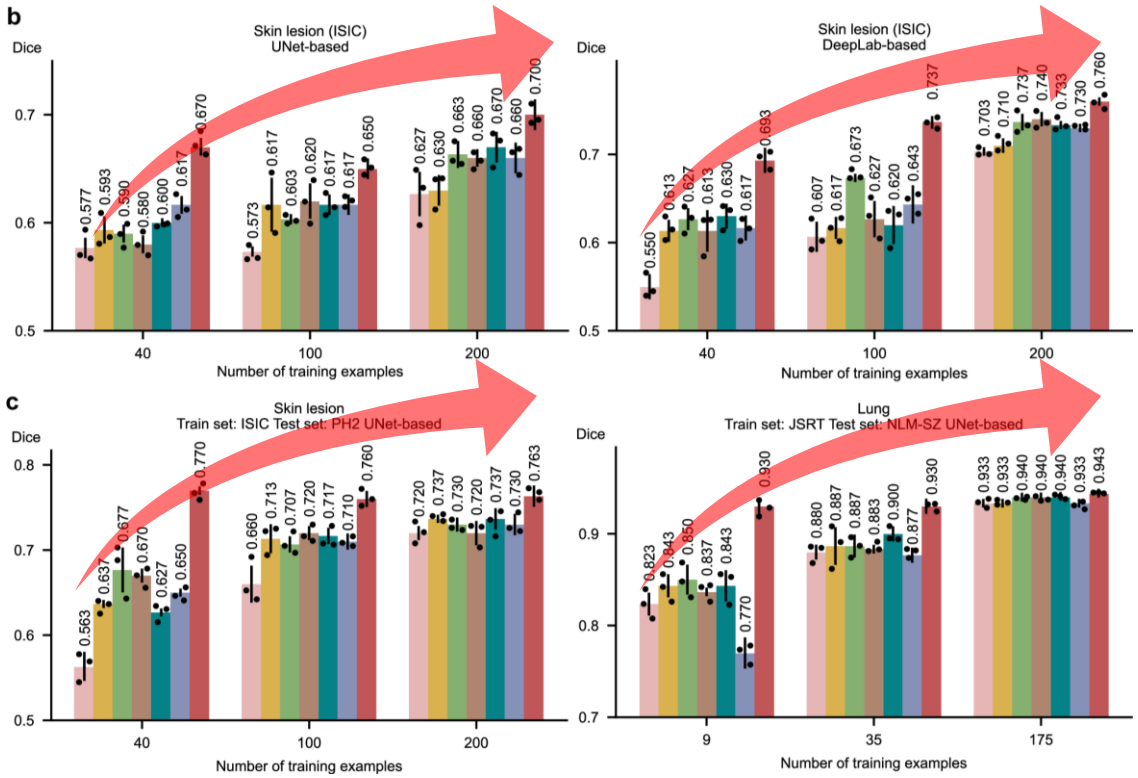


Magnified Results



Approach 3: Additional Datasets

Necessity of Additional Datasets



Augmenting training data has a clear and demonstrable impact on improving model performance.

✓ Additional Dataset Used

- **janowczyk** — a public dataset related to pathology images
- **nucleisegmentationbenchmark** — nucleus segmentation benchmark data
- **isbi2009** — ISBI 2009 challenge data (microscopy cell images)
- **BBBC020** — cell images from the Broad Bioimage Benchmark Collection
- **TNBC** — Triple Negative Breast Cancer tissue images

By incorporating diverse cell images from external sources, approximately 316 additional images were integrated into the training pipeline

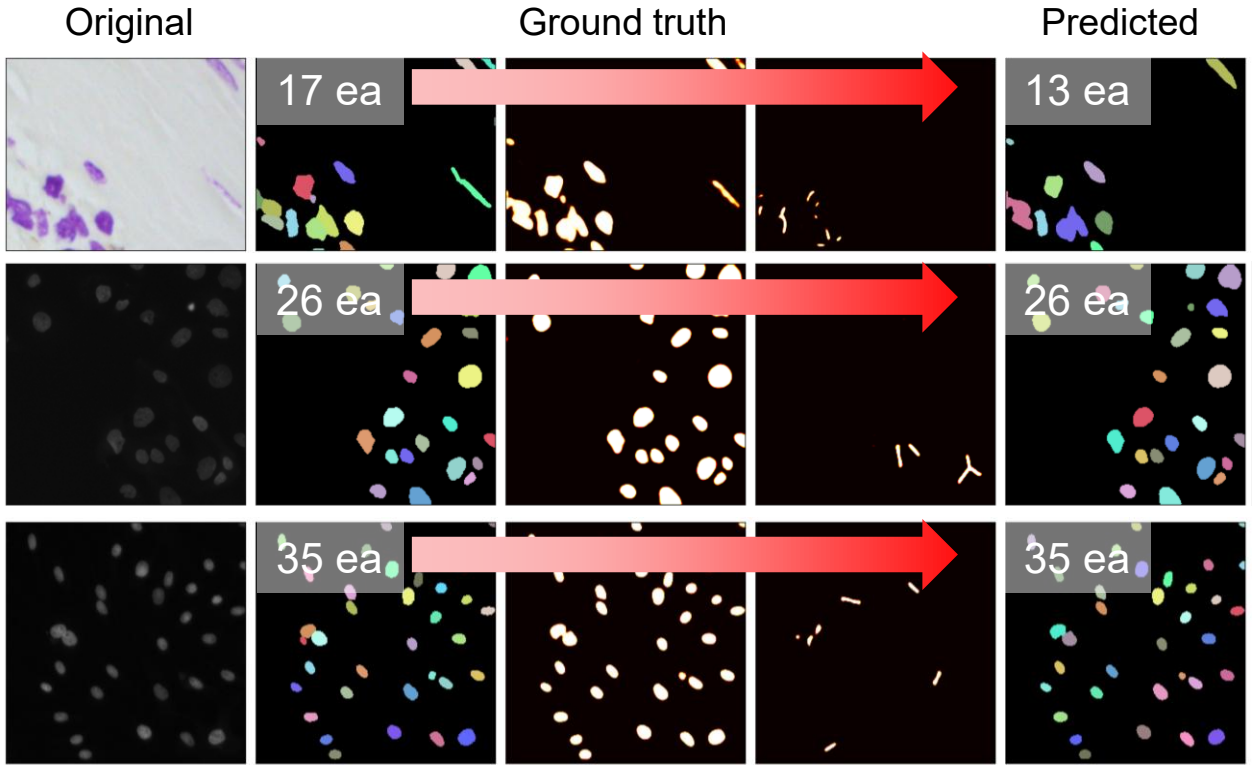
Zhang, L., Jindal, B., Alaa, A. *et al.* Generative AI enables medical image segmentation in ultra low-data regimes. *Nat Commun* **16**, 6486 (2025).

Approach 3: Results with Additional Dataset

(previous) 0.46540

Quantitative Results

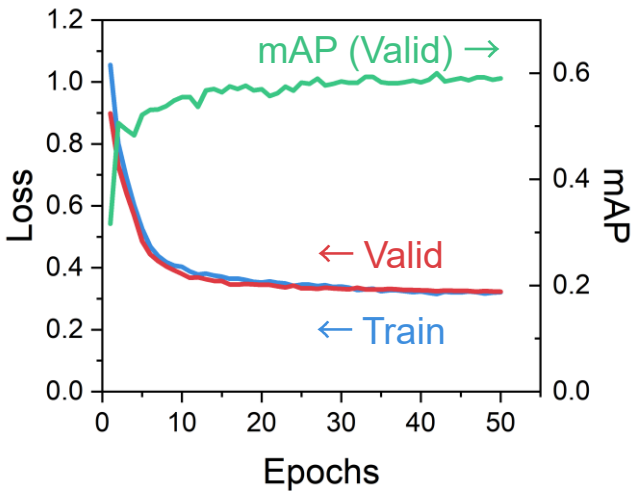
✓ Visualization of Original/GT/Pred images



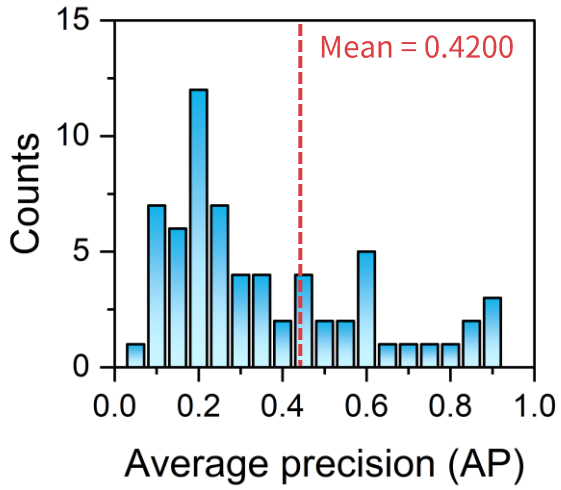
mAP: **0.54523** (↑ **14.64%**)

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

✓ Epoch-loss



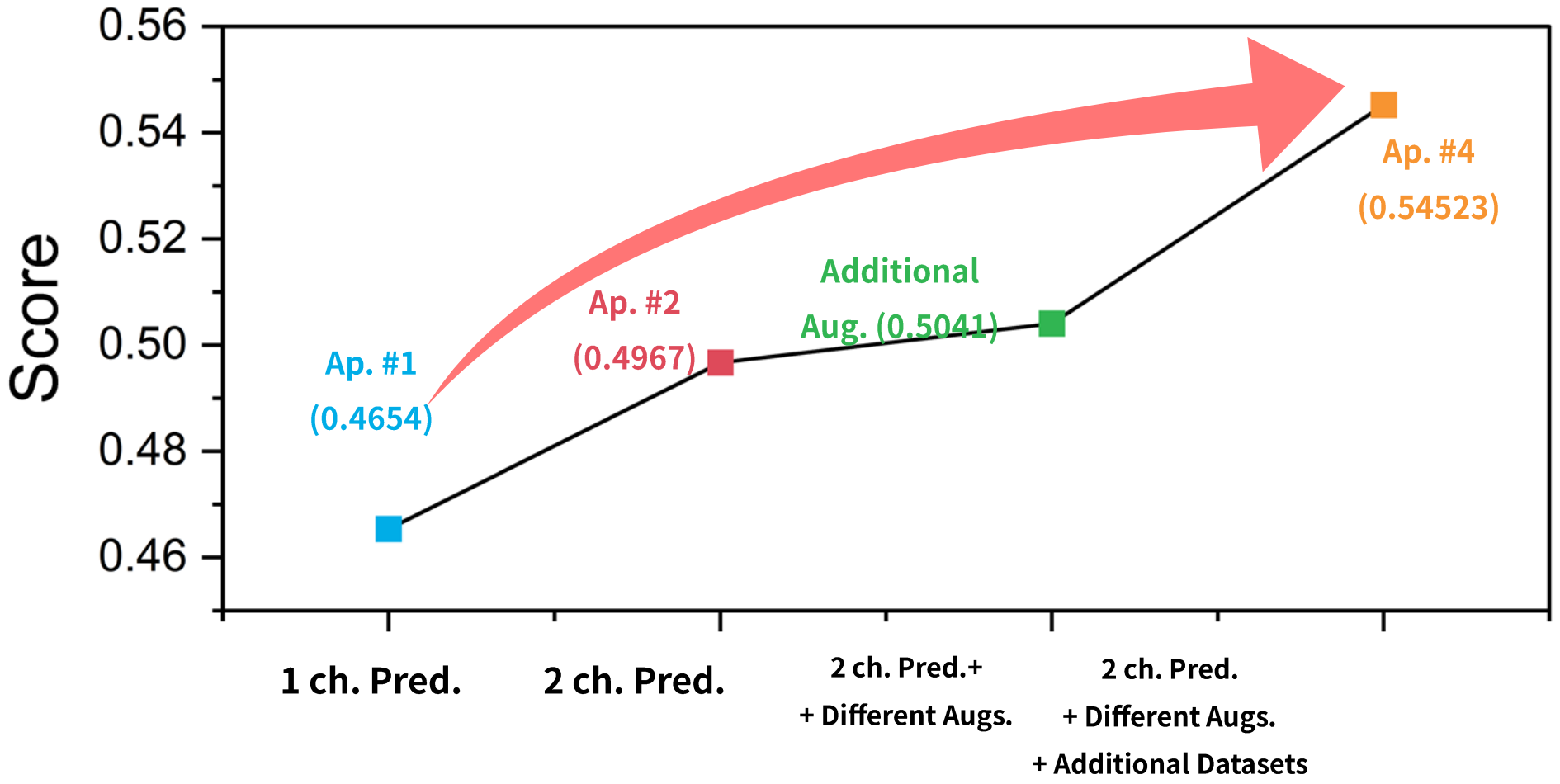
✓ AP distribution (stage 1-test)



- By training with an additional 316 cell images, the overall score demonstrated a 14.64% performance improvement over the baseline model.

Score Improvement

UNet++ Final Score and Performance Trend

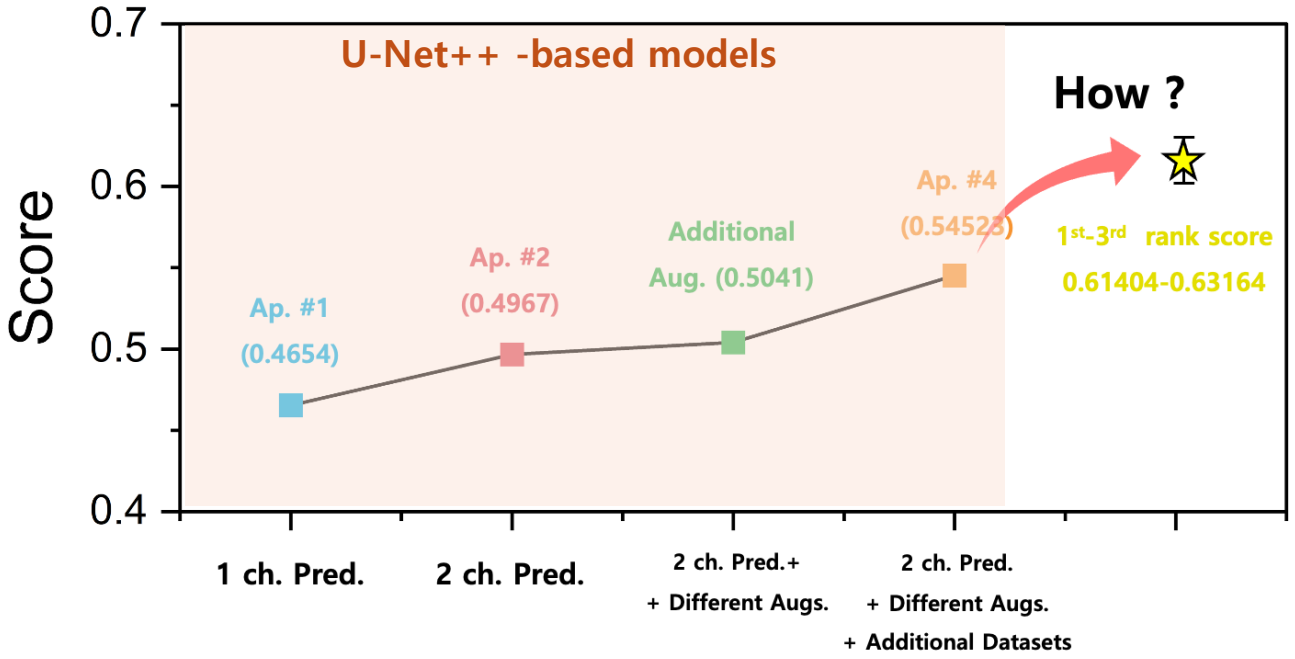


4. Pushing Further: Foundation Model-Based Approach

Why a Specialized Foundation Model?

Limitations of UNet++ for This Task

- Single supervised CNN with no prior knowledge of cellular morphology
- Multi-modality data (fluorescent + H&E + brightfield → imbalance) — one model struggles to cover all distributions
- Output is a binary/semantic mask — touching nuclei merged; separation depends on a extra post-processing step (watershed algorithm)



Implication

Need a model that already knows what cells look like — pre-trained on a large cellular corpus — and can be used without further training on our limited data.

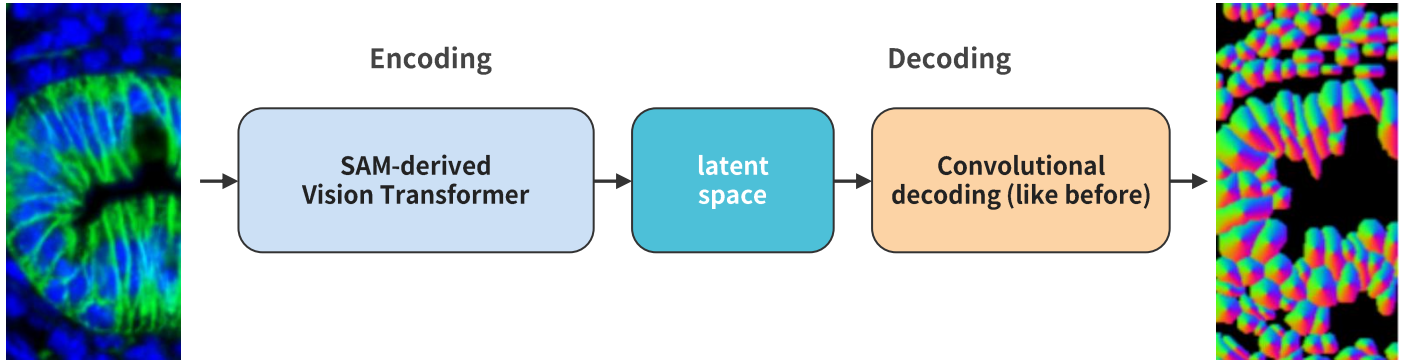
→ Utilizing SAM-based foundation models that are fine-tuned for cellular segmentation.

Approach 4: Cellpose 4 (cpsam)

Cellpose 4: Cellpose-SAM

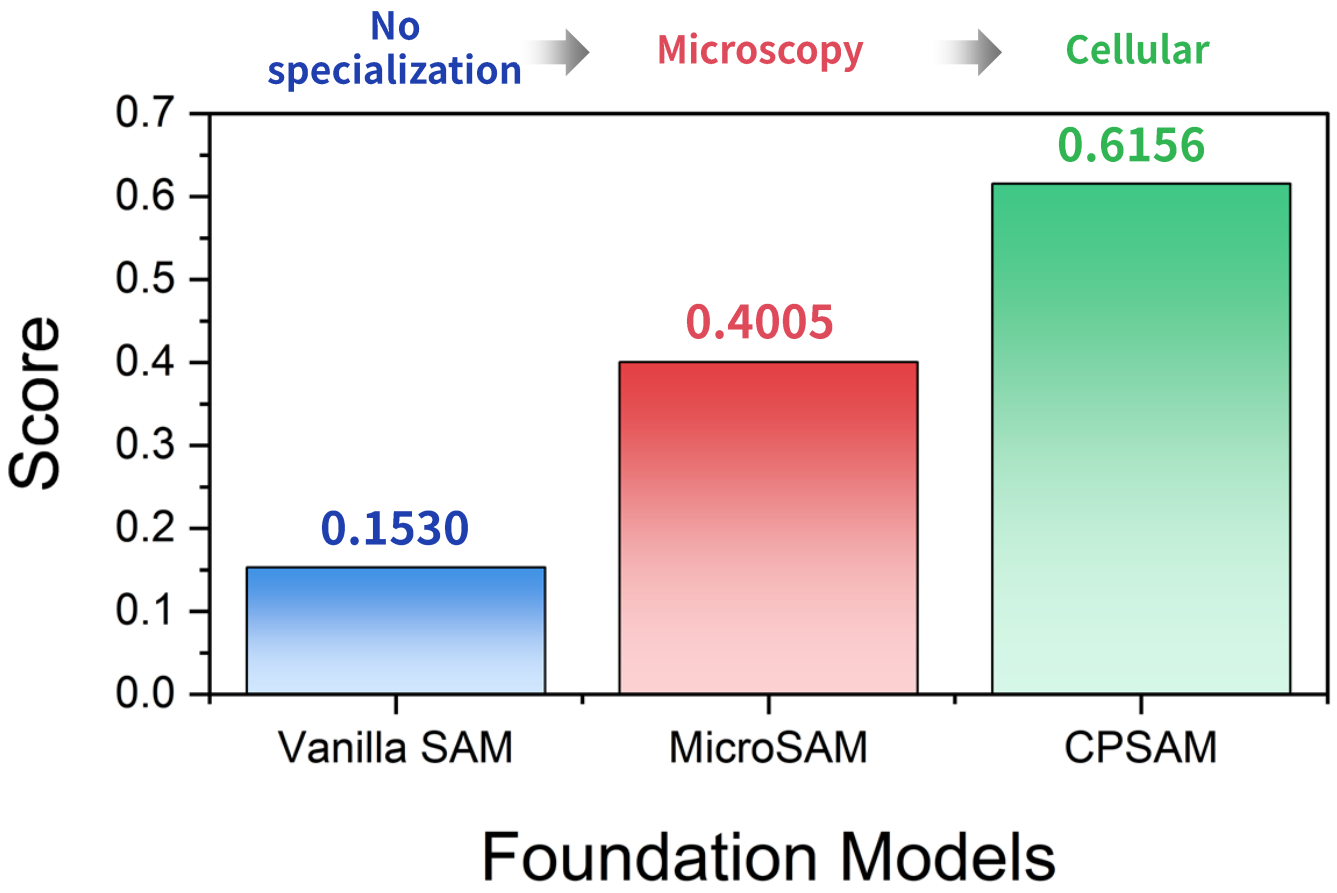
- ✓ Architecture based on Segment Anything Model (SAM) from meta
- ✓ Uses SAM weights & specialized them for bioimage segmentation
- ✓ Similar decoding architecture & the same flow prediction
- ✓ Generalized broader out-of-the-box & replaces Cellpose 3 image
- ✓ restoration approach

Cellpose 4 architecture



Approach 4: SAM Family Comparison — Specialization Ladder

Comparison for W/ and W/O domain specializations



Kaggle Scores

Vanilla SAM (Meta): 0.15305

— general-purpose, no cell fine-tune

MicroSAM: 0.40054

— broad microscopy fine-tune

cpsam (Cellpose 4): 0.61563 ★

— cellular-specific fine-tune

Same ViT-B backbone, different fine-tune data
→ very different downstream performance.

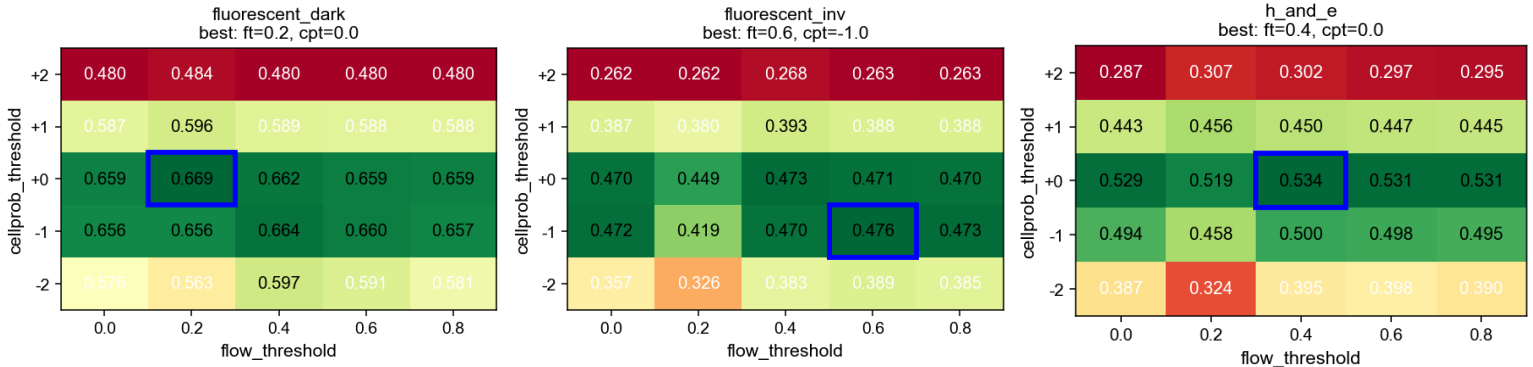
Takeaway

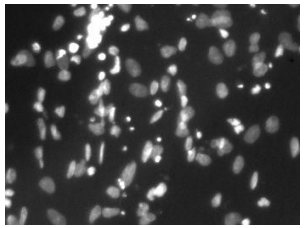
Closer fine-tune domain → higher score.

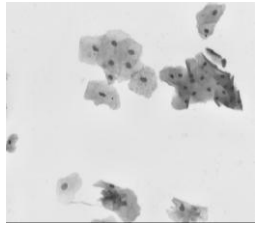
cpsam is the right model for nuclei segmentation.

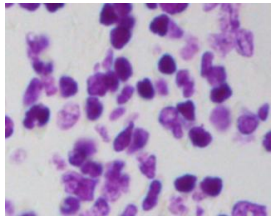
Per-Type Hyperparameter Grid Search

(previous) 0.46540









	fluorescent_dark	fluorescent_inv	h_and_e
ft	0.2	0.6	0.4
cpt	0.0	-1.0	0.0
Effect of these choices	stricter flow filter → removes noisy small false-positive masks	looser ft + lower → accepts darker / dimmer nuclei as cells	Default → already optimal typical nucleus size & contrast

flow_threshold(ft)

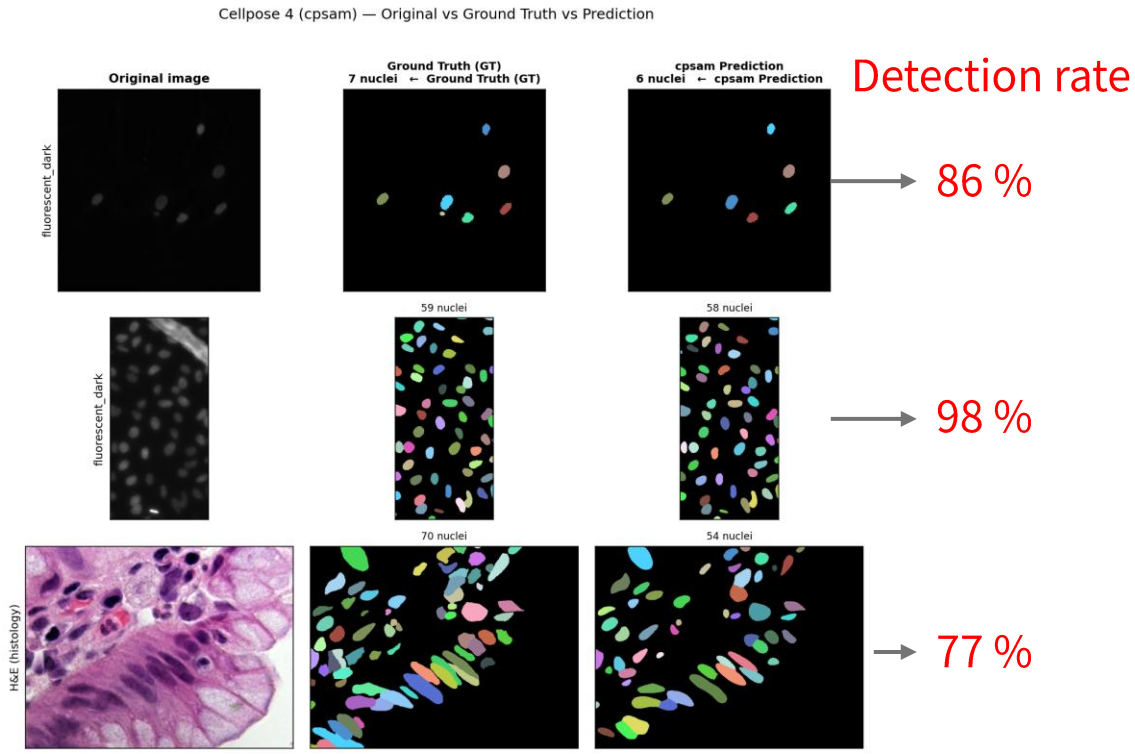
- Max allowed error of predicted flow inside a mask
- Lower → stricter, drops more inconsistent masks

cellprob_threshold(cpt)

- Min cell-probability for a pixel to be a cell
- Lower → includes dimmer pixels more inclusive

Result of cpsam

Visualization of HP-tuned Cellpose-SAM



Detection rate

86 %

98 %

77 %

mAP: **0.63132 (↑ 26.28%)**

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

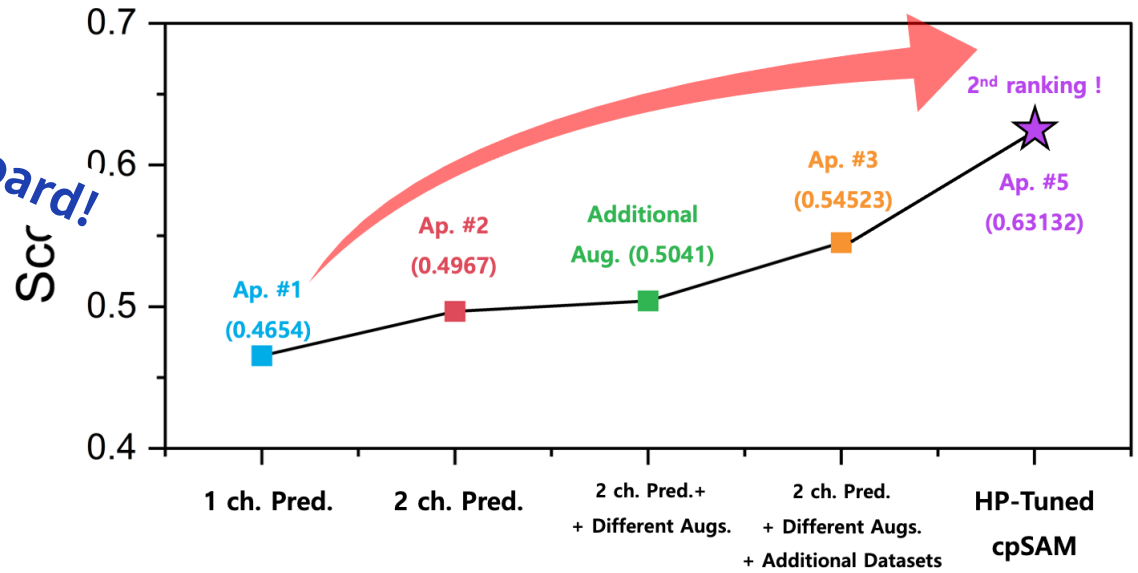
The domain-specific (cell characteristic) foundation model-based Cellpose 4 demonstrates high performance in distinguishing individual nuclei with high probability among numerous nuclei through hyperparameter tuning.

Final Result

Leaderboard Score (Private)

#	Team	Members		Entries	Last	Solution
1	[ods.ai] topcoders		0.63164			
2	Team 10 from IMEN574/ID574		0.63132			
2	jacobkie		0.61472	2	8y	
3	Deep Retina		0.61404	2	8y	
4	Nuclear Vision		0.61088	2	8y	
5	Inom Mirzaev		0.60929	4	8y	
6	ACS		0.59407	1	8y	

2nd rank in the leaderboard!



mAP: **0.63132**

For instance segmentation - A composite score measuring how accurately and completely nuclei are detected, evaluated across multiple levels of strictness

Thank you