MultiOrg: A Multi-Rater Organoid-Detection Dataset

NeurIPS 2024: Datasets & Benchmarks

Christina Bukas[1], Harshavardhan Subramanian[1], Fenja See[2], Carina Steinchen[2], Ivan Ezhov[3], Gowtham Boosarpu[4], Sara Asgharpour[2], Gerald Burgstaller[2], Mareike Lehmann[2,4], Florian Kofler[1,5,6], Marie Piraud[1]

[1]Helmholtz AI, Helmholtz Munich, Germany
[2]Institute of Lung Health and Immunity (LHI), Helmholtz Munich, Germany
[3]Department of Computer Science, Technical University of Munich, Germany
[4]Institute for Lung Research, Philipps-University Marburg, Germany
[5]Department of Neuroradiology, Technical University of Munich, Munich, Germany.
[6]Department of Quantitative Biomedicine, University of Zurich, Switzerland.

Motivation & Background

What are organoids?

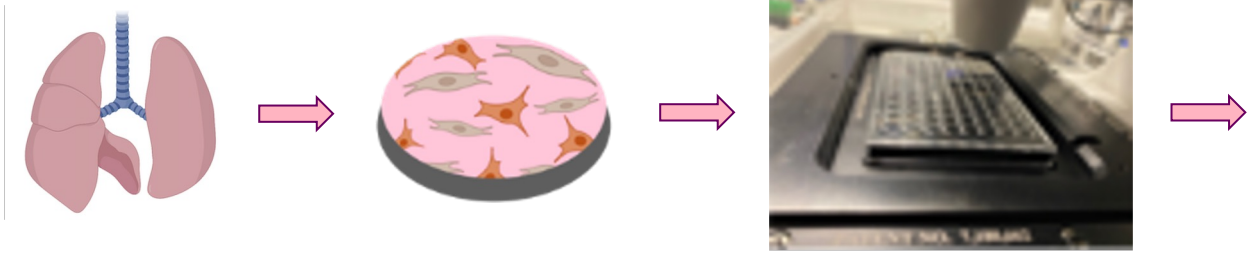
Image generated using https://deepai.org/

Organoids are used for:

- Disease understanding
- Drug development
- Personalised medicine

How Organoids are made

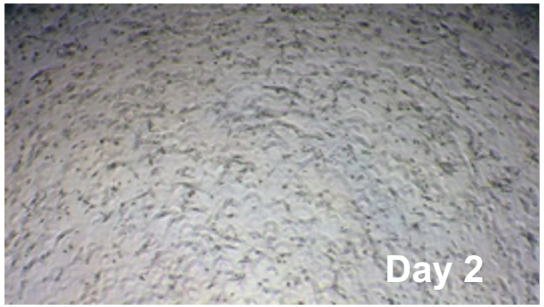
Growing tiny organs in the lab

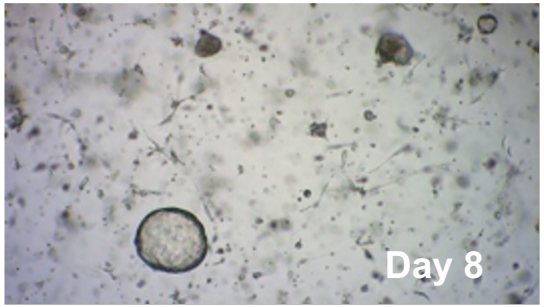


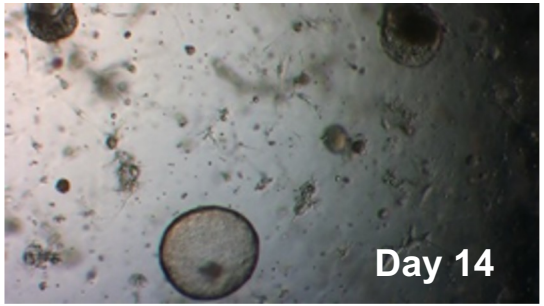
Lung Stem Cells

Organoid culture

Image acquisition

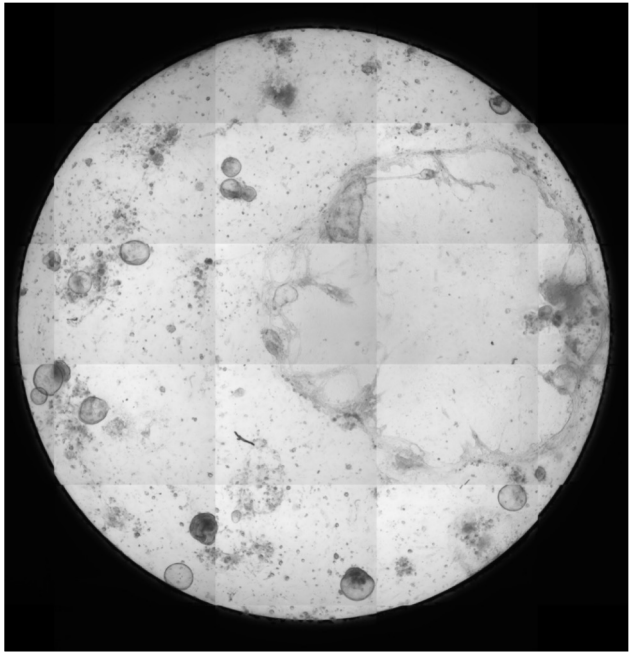






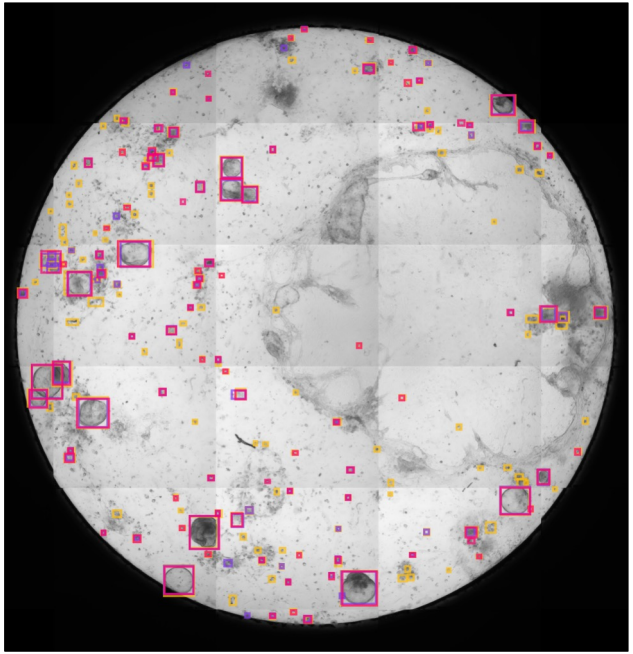
Challenges in Organoid Detection

Labeling uncertainty and noise



Challenges in Organoid Detection

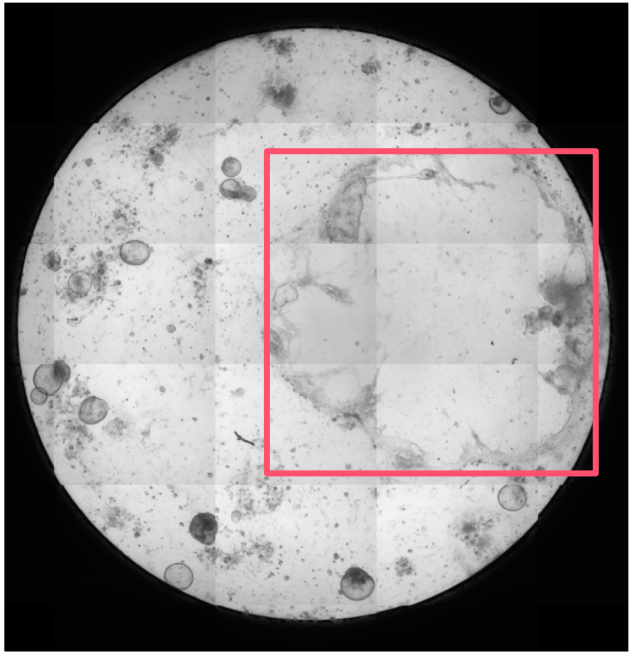
Labeling uncertainty and noise



- Time-consuming manual annotations

Challenges in Organoid Detection

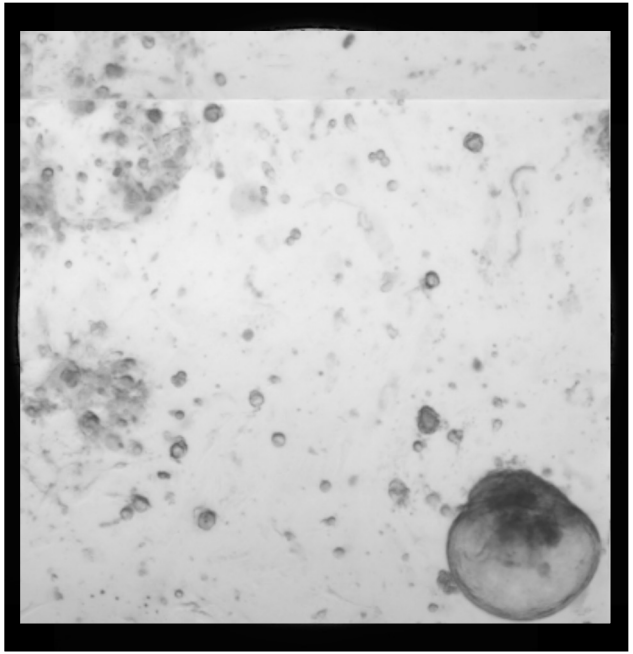
Labeling uncertainty and noise



- Time-consuming manual annotations
- Noise due to artifacts, dead cells and debris

Challenges in Organoid Detection

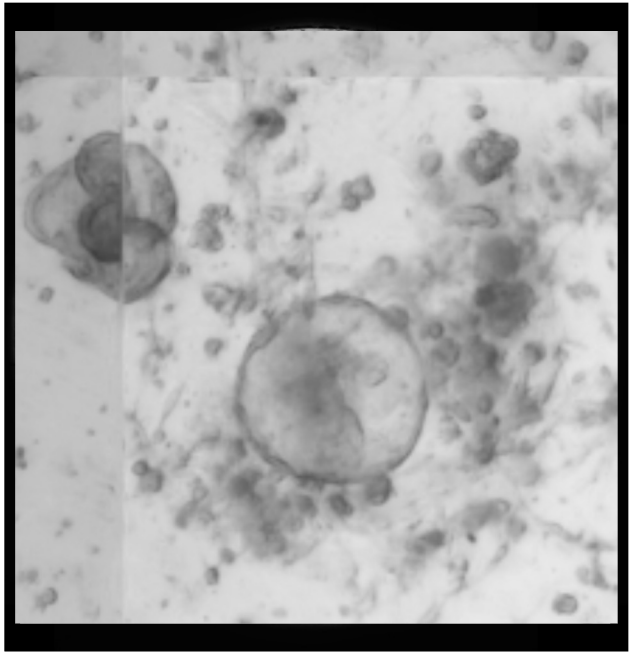
Labeling uncertainty and noise



- Time-consuming manual annotations
- Noise due to artifacts, dead cells and debris
- Large range of organoid sizes

Challenges in Organoid Detection

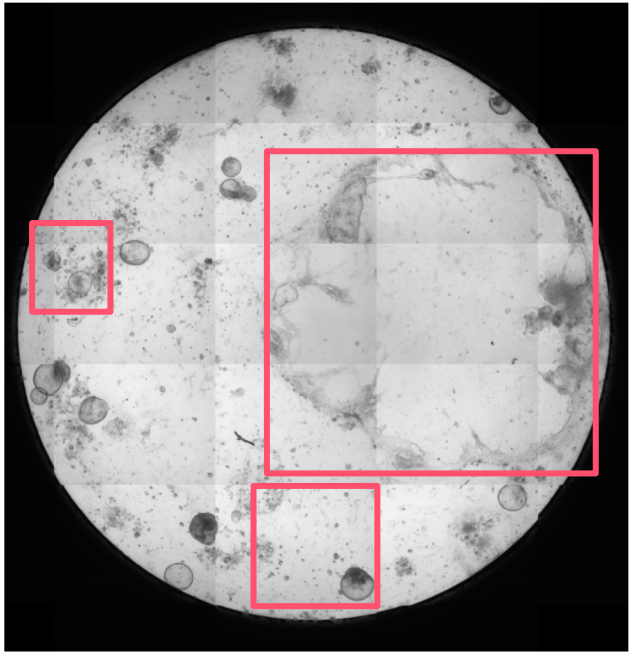
Labeling uncertainty and noise



- Time-consuming manual annotations
- Noise due to artifacts, dead cells and debris
- Large range of organoid sizes
- Organoid clusters

Challenges in Organoid Detection

Labeling uncertainty and noise



- Time-consuming manual annotations
- Noise due to artifacts, dead cells and debris
- Large range of organoid sizes
- Organoid clusters

No datasets in organoid detection addressing label uncertainty!

The MultiOrg dataset

Overview

- 400+ 2D microscopy images

Study Type	Normal		Macros		Combined	
	# Images	# Organoids	# Images	# Organoids	# Images	# Organoids
Train set						
$train_A$	181	30,710	15	2,669	196	33,379
$train_B$	135	20,263	25	1,781	160	22,044
Total	**316**	**50,973**	**40**	**4,450**	**356**	**55,423**
Test set						
$test^0_A$	8	1,145	14	1,865	22	3,010
$test^0_B$	20	3,020	13	1,493	33	4,513
Total (Label set $test^0$)	**28**	**4,165**	**27**	**3,358**	**55**	**7,523**

The MultiOrg dataset

Overview

- 400+ 2D microscopy images
- 60,000+ organoid annotations

Study Type	Normal		Macros		Combined	
	# Images	# Organoids	# Images	# Organoids	# Images	# Organoids
Train set						
$train_A$	181	30,710	15	2,669	196	33,379
$train_B$	135	20,263	25	1,781	160	22,044
Total	**316**	**50,973**	**40**	**4,450**	**356**	**55,423**
Test set						
$test^0_A$	8	1,145	14	1,865	22	3,010
$test^0_B$	20	3,020	13	1,493	33	4,513
Total (Label set $test^0$)	**28**	**4,165**	**27**	**3,358**	**55**	**7,523**

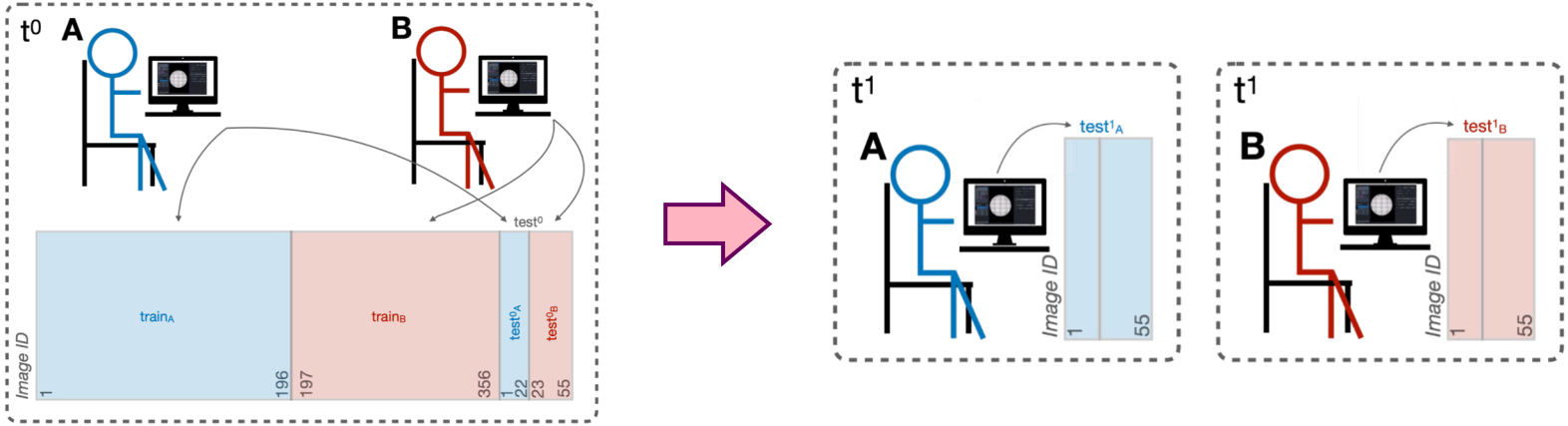
The MultiOrg dataset

Overview

- 400+ 2D microscopy images
- 60,000+ organoid annotations
- Three expert labels on the test set

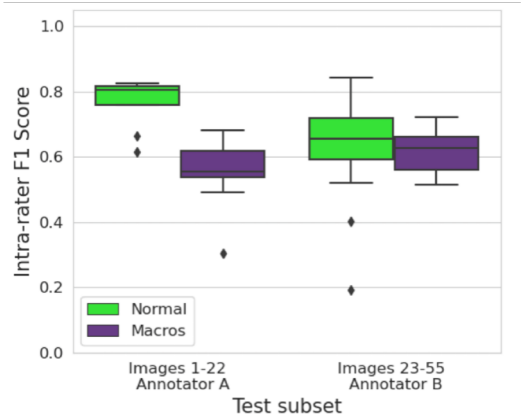
Study Type	Normal		Macros		Combined	
	# Images	# Organoids	# Images	# Organoids	# Images	# Organoids
Train set						
$train_A$	181	30,710	15	2,669	196	33,379
$train_B$	135	20,263	25	1,781	160	22,044
Total	**316**	**50,973**	**40**	**4,450**	**356**	**55,423**
Test set						
$test^0_A$	8	1,145	14	1,865	22	3,010
$test^0_B$	20	3,020	13	1,493	33	4,513
Total (Label set $test^0$)	**28**	**4,165**	**27**	**3,358**	**55**	**7,523**
Label set $test^1_A$	**28**	**2,748**	**27**	**1,981**	**55**	**4,729**
Label set $test^1_B$	**28**	**2,655**	**27**	**2,301**	**55**	**4,956**

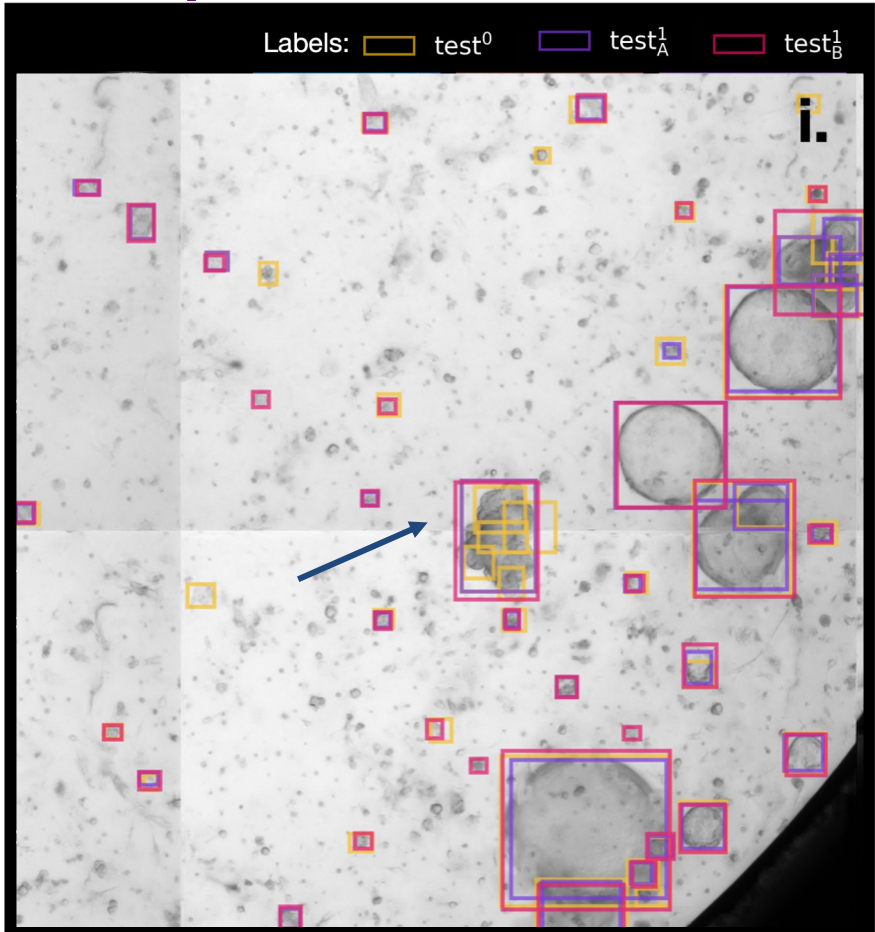
Multi-rater setting



Multi-rater setting at two distinct timepoints

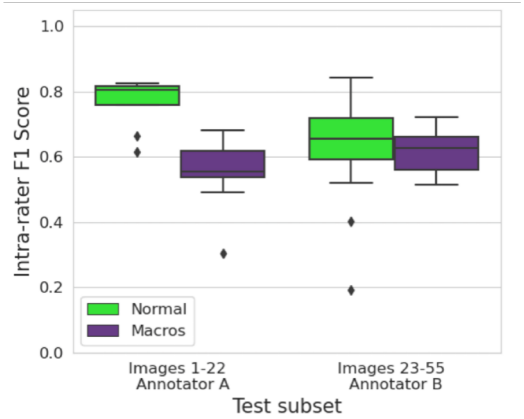
Allows for intra- and inter-rater analyses

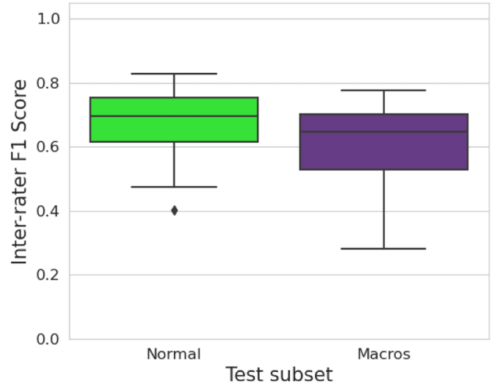


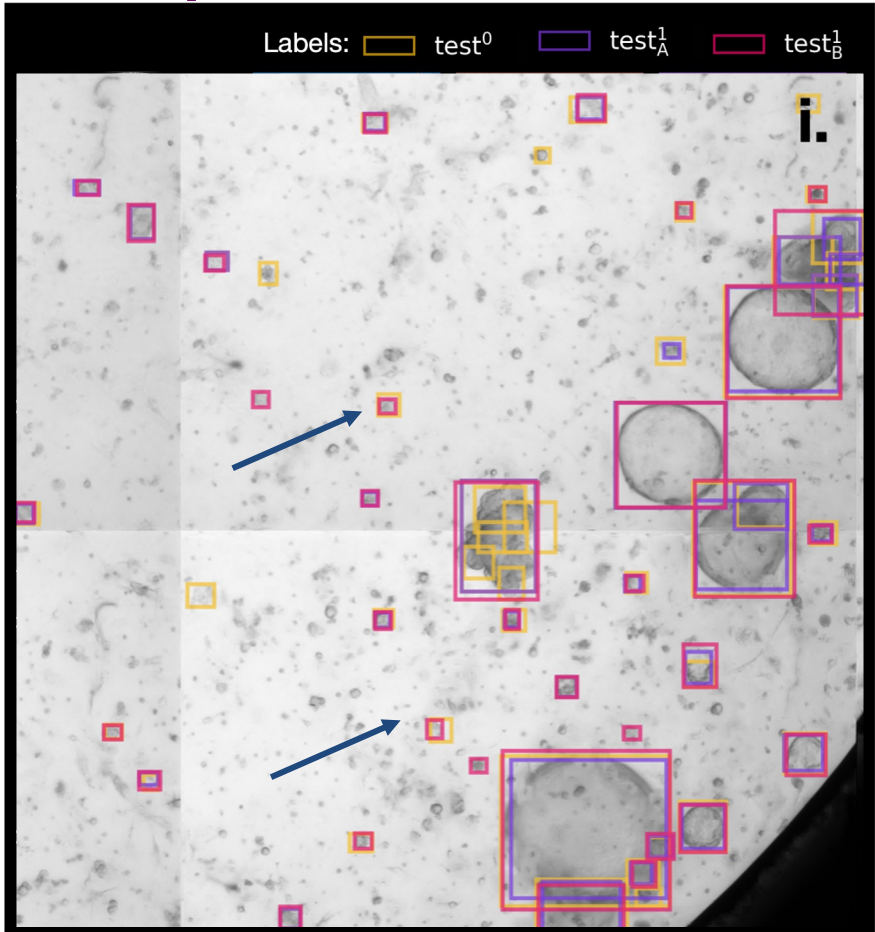


Multi-rater setting at two distinct timepoints

Allows for intra- and inter-rater analyses

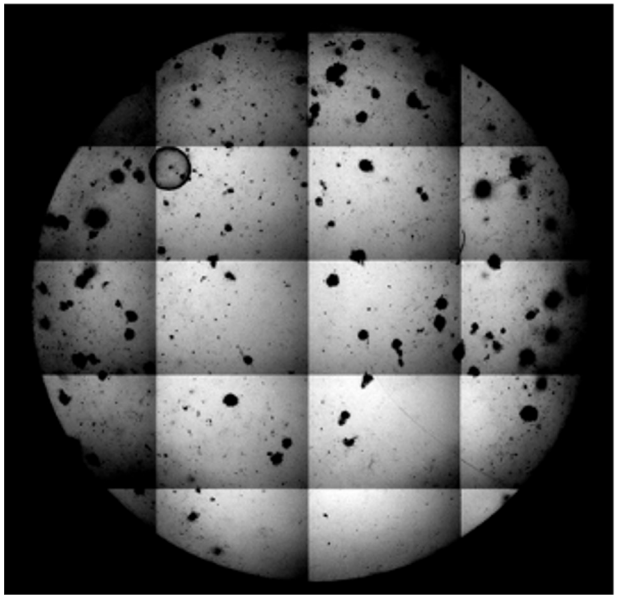




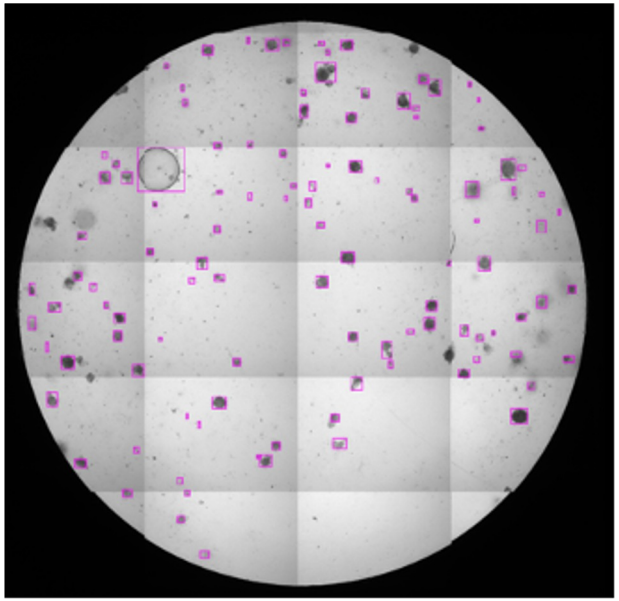


MultiOrg: Model Benchmark

Object-detection models



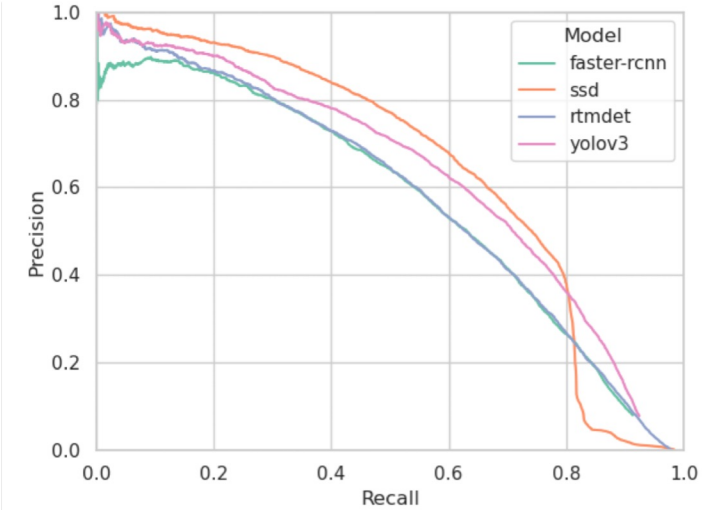
- Faster R-CNN
- SSD
- YOLOv3
- RTMDet



https://www.kaggle.com/code/christinabukas/multiorg-detection-benchmark

MultiOrg: Model Benchmark

Object Detection Scores

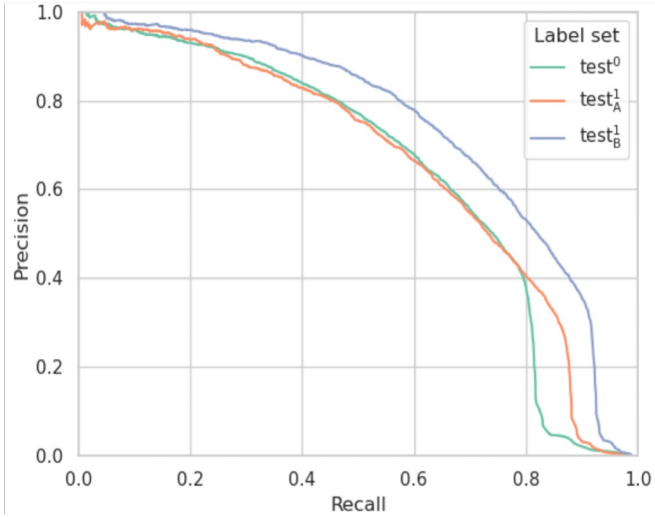


PR curves of all models on $test^0$

Metric	Label set	Faster R-CNN	SSD	YOLOv3	RTMDet
Precision	$test^0$	0.23	0.61	**0.73**	0.64
	$test^1_A$	0.16	0.44	0.58	0.54
	$test^1_B$	0.18	0.50	0.67	0.56
	mean	**0.19**	**0.52**	**0.66**	**0.58**
Recall	$test^0$	0.84	0.67	0.48	0.51
	$test^1_A$	0.92	0.78	0.62	0.69
	$test^1_B$	**0.97**	0.83	0.67	0.68
	mean	**0.91**	**0.76**	**0.59**	**0.63**
F1-score	$test^0$	0.36	0.64	0.58	0.57
	$test^1_A$	0.27	0.57	0.60	0.61
	$test^1_B$	0.30	0.62	**0.67**	0.62
	mean	**0.31**	**0.61**	**0.62**	**0.60**
mAP@0.5IoU (%)	$test^0$	56.56	64.40	62.55	57.71
	$test^1_A$	57.09	65.79	61.11	63.87
	$test^1_B$	68.36	**73.88**	70.25	63.23
	mean	**60.67**	**68.09**	**64.64**	**61.60**
mAP@0.75IoU (%)	$test^0$	17.48	21.81	19.15	22.56
	$test^1_A$	23.53	23.42	19.13	30.13
	$test^1_B$	**46.98**	**46.48**	39.01	32.85
	mean	**29.33**	**30.57**	**25.76**	**28.51**

MultiOrg: Model Benchmark

Object Detection Scores



PR curves of SSD model on all label sets

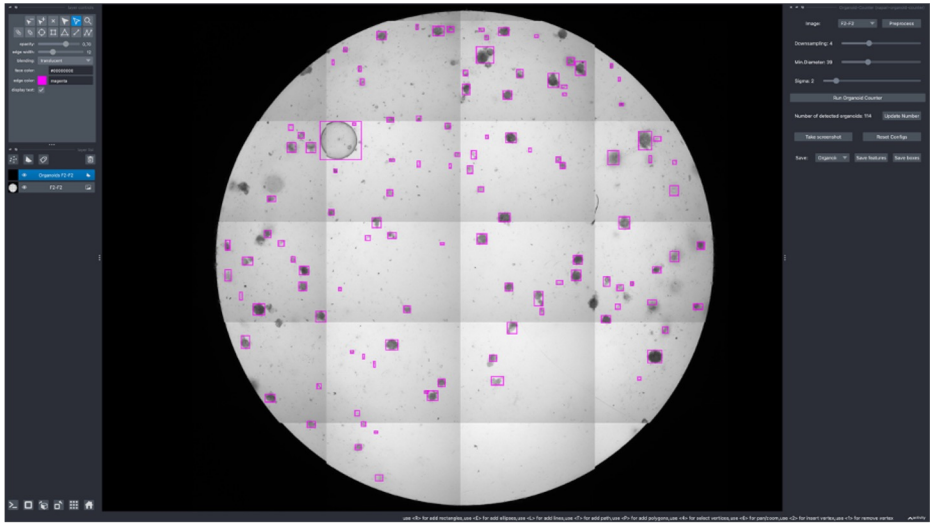
Metric	Label set	Faster R-CNN	SSD	YOLOv3	RTMDet
Precision	$test^0$	0.23	0.61	**0.73**	0.64
	$test^1_A$	0.16	0.44	0.58	0.54
	$test^1_B$	0.18	0.50	0.67	0.56
	mean	**0.19**	**0.52**	**0.66**	**0.58**
Recall	$test^0$	0.84	0.67	0.48	0.51
	$test^1_A$	0.92	0.78	0.62	0.69
	$test^1_B$	**0.97**	0.83	0.67	0.68
	mean	**0.91**	**0.76**	**0.59**	**0.63**
F1-score	$test^0$	0.36	0.64	0.58	0.57
	$test^1_A$	0.27	0.57	0.60	0.61
	$test^1_B$	0.30	0.62	**0.67**	0.62
	mean	**0.31**	**0.61**	**0.62**	**0.60**
mAP@0.5IoU (%)	$test^0$	56.56	64.40	62.55	57.71
	$test^1_A$	57.09	65.79	61.11	63.87
	$test^1_B$	68.36	**73.88**	70.25	63.23
	mean	**60.67**	**68.09**	**64.64**	**61.60**
mAP@0.75IoU (%)	$test^0$	17.48	21.81	19.15	22.56
	$test^1_A$	23.53	23.42	19.13	30.13
	$test^1_B$	**46.98**	**46.48**	39.01	32.85
	mean	**29.33**	**30.57**	**25.76**	**28.51**

MultiOrg: Key contributions

Open-source resources

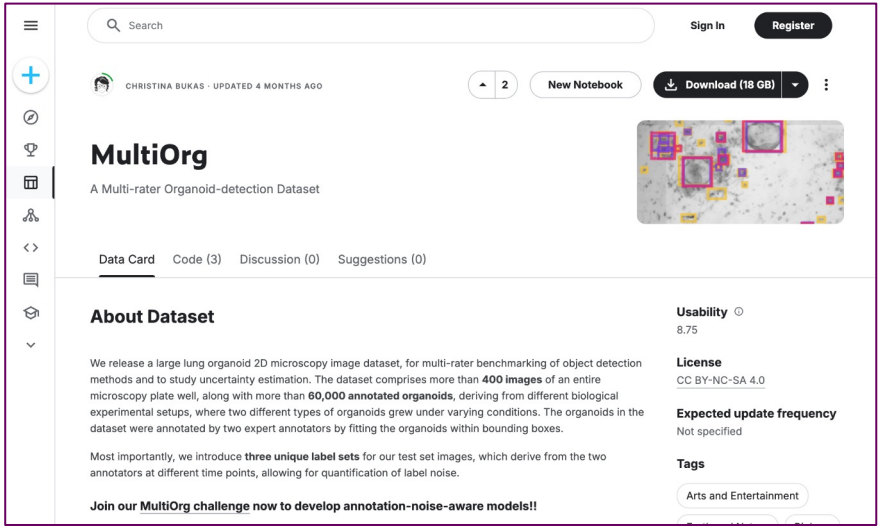
SCAN ME

Napari plugin for the biologist community



https://github.com/HelmholtzAI-Consultants-Munich/napari-organoid-counter

Dataset, benchmark & challenge



NeurIPS 2024: Datasets & Benchmarks

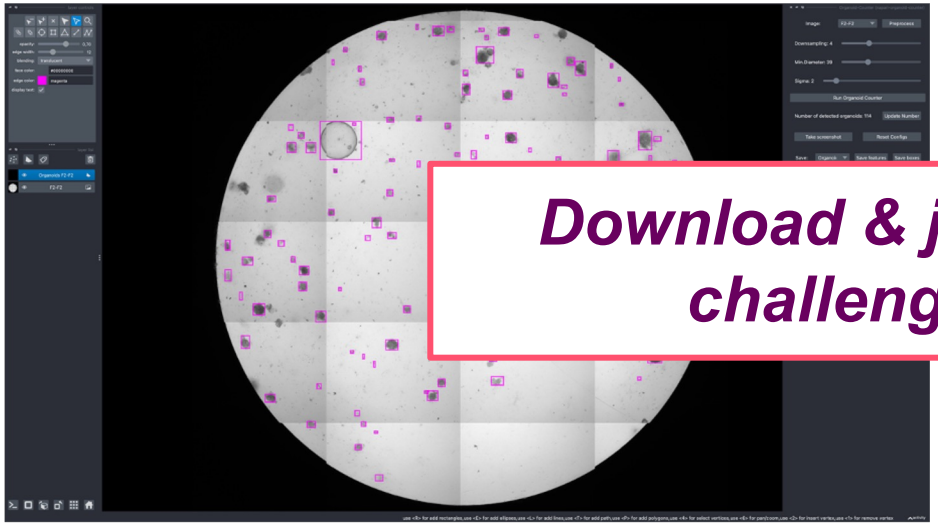
MultiOrg: Key contributions

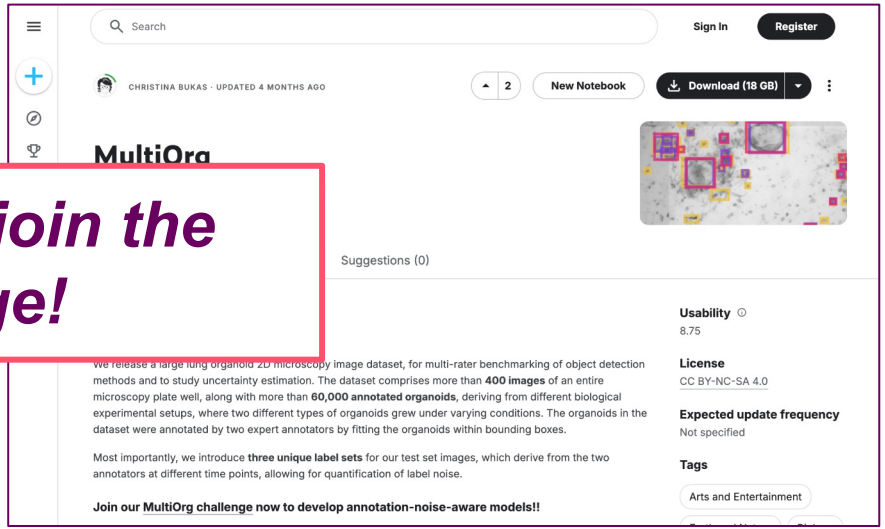
Open-source resources

SCAN ME

Napari plugin for the biologist community

Dataset, benchmark & challenge





Download & join the challenge!

https://github.com/HelmholtzAI-Consultants-Munich/napari-organoid-counter

NeurIPS 2024: Datasets & Benchmarks