Towards Annotation-Efficient Deep Learning for Computer-Aided Diagnosis

Zongwei Zhou

Ph.D. Candidate, Biomedical Informatics

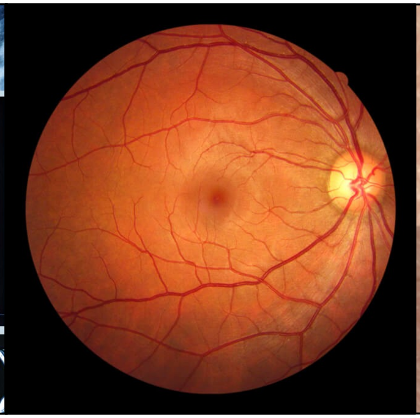
College of Health Solutions, Arizona State University

P: 1-(480)738-2575 | E: zongweiz@asu.edu

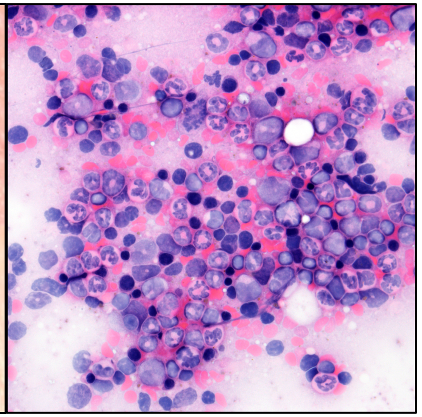
People
Who
Don't
Give Up

Imaging data account for about 90% of all healthcare data

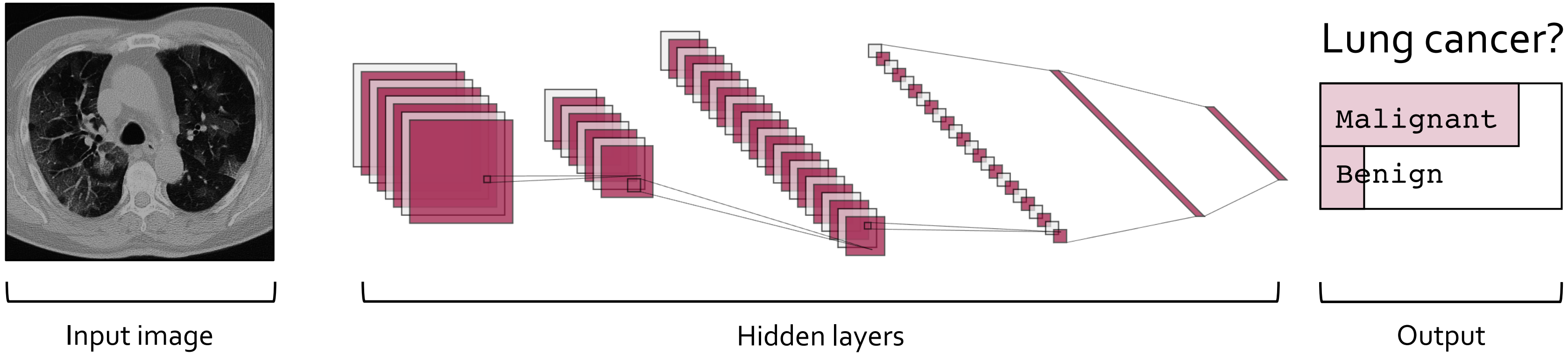

Radiology


Ophthalmology


Dermatology


Pathology

Deep Learning has ushered in a revolution in medical imaging



1. "The Digital Universe Driving Data Growth in Healthcare." published by EMC with research and analysis from IDC (12/13)
2. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444

Introduction

Objective

Aim 1

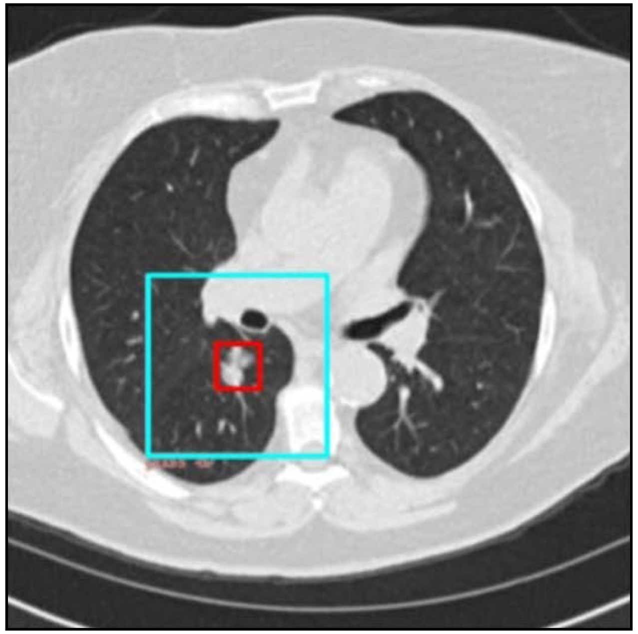
Aim 2

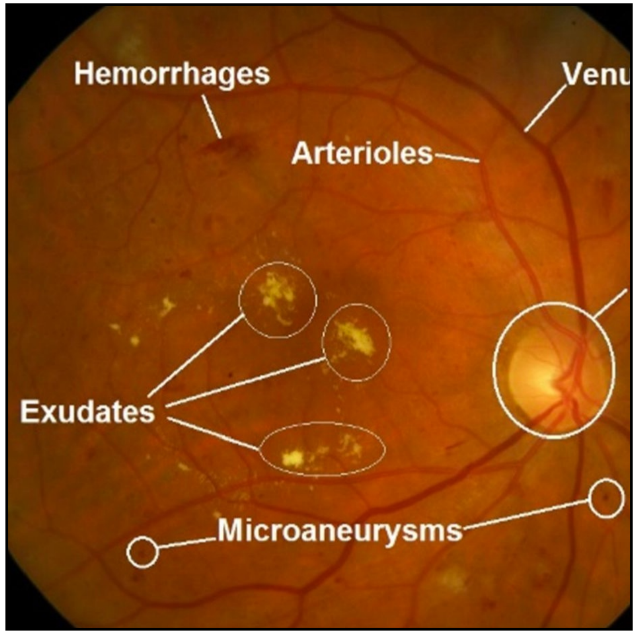
Aim 3

Summary

To match human diagnostic precision, deep learning requires a lot of annotation cost.

- **42,290** radiologist-annotated CT images for lung cancer diagnosis
- **128,175** ophthalmologist-annotated retinal images for diabetic retinopathy detection
- **129,450** dermatologist-annotated images for skin cancer classification





1. Ardila, Diego, et al. "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography." Nature medicine 25.6 (2019): 954-961.
2. Gulshan, Varun, et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." Jama 316.22 (2016): 2402-2410.
3. Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." nature 542.7639 (2017): 115-118.

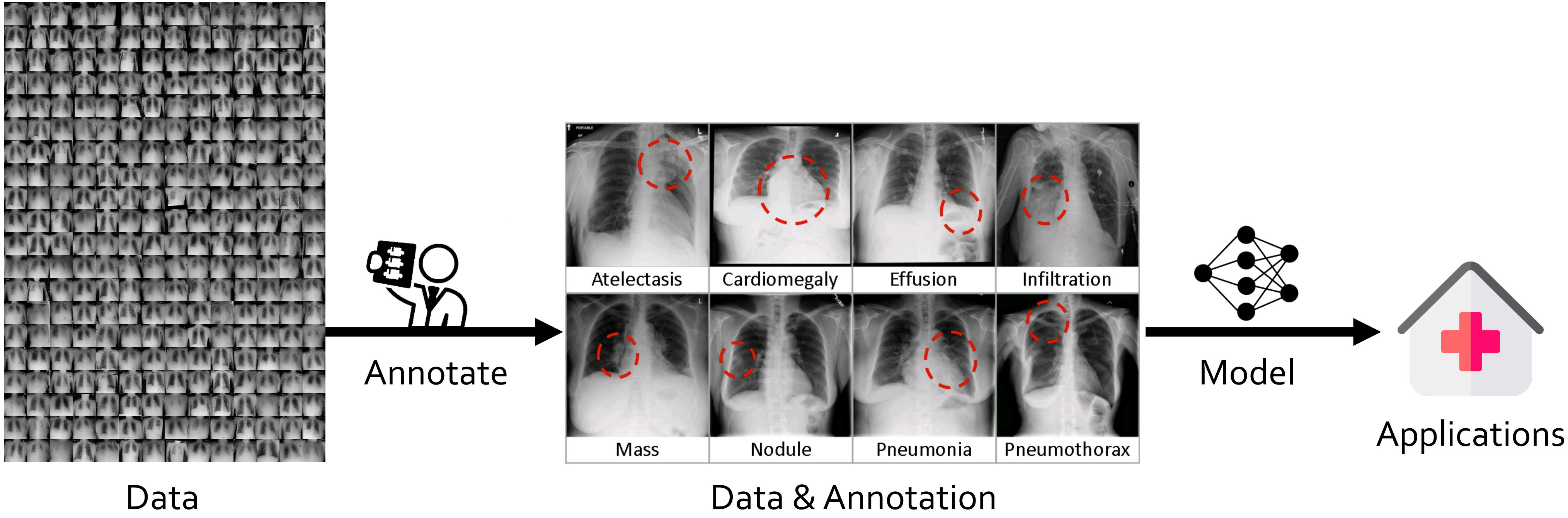
To match human diagnostic precision, deep learning requires a lot of annotation cost.

- **42,290** radiologist-annotated CT images for lung cancer diagnosis
- **128,175** ophthalmologist-annotated retinal images for diabetic retinopathy detection
- **129,450** dermatologist-annotated images for skin cancer classification

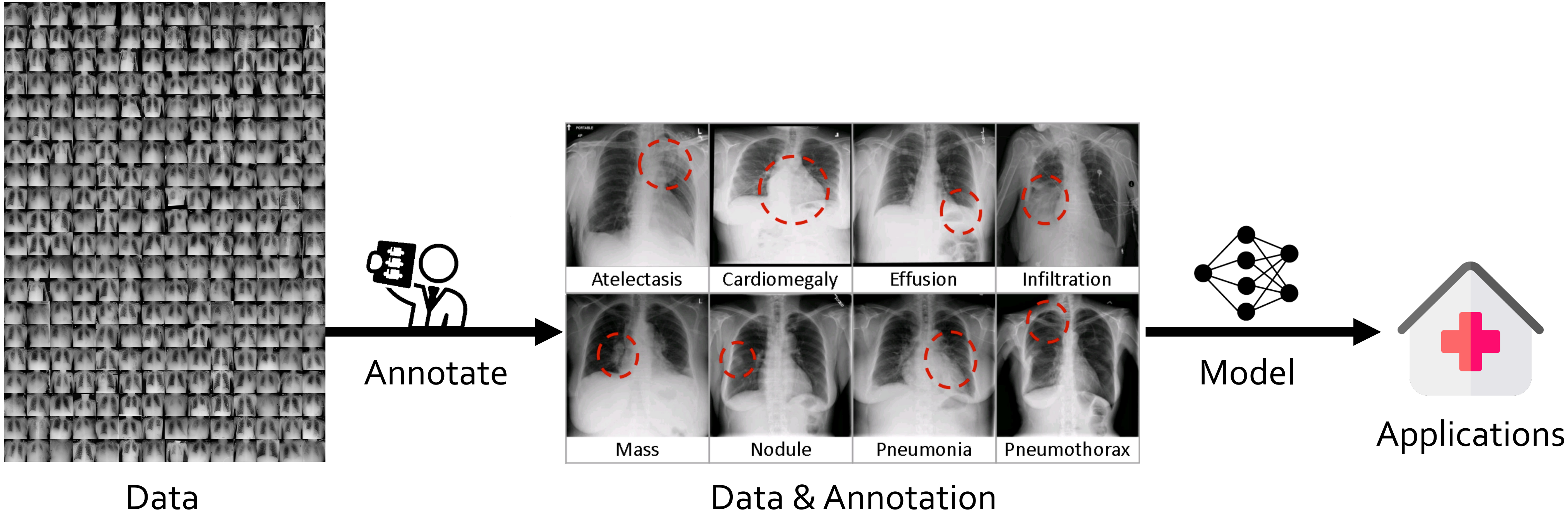
How to develop annotation-efficient deep learning without such BIG annotated data?

Significant, especially for these scenarios:

- A flood of patients are waiting for results during an outbreak
- Doctors do not have time to annotate every case for algorithm development
- Not many doctors have expertise for novel/rare diseases

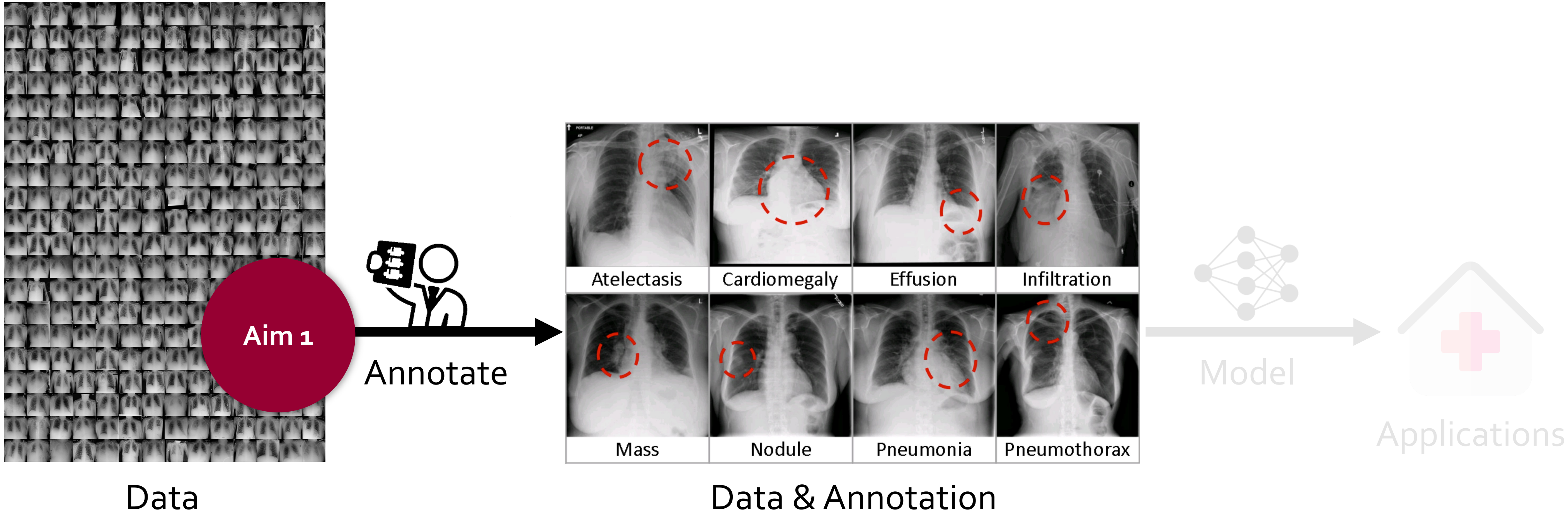

Annotate
Atelectasis
Cardiomegaly
Effusion
Infiltration
Mass
Nodule
Pneumonia
Pneumothorax
Model
Applications
Data
Data & Annotation

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems



Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts



Introduction

Objective

Aim 1

Aim 2

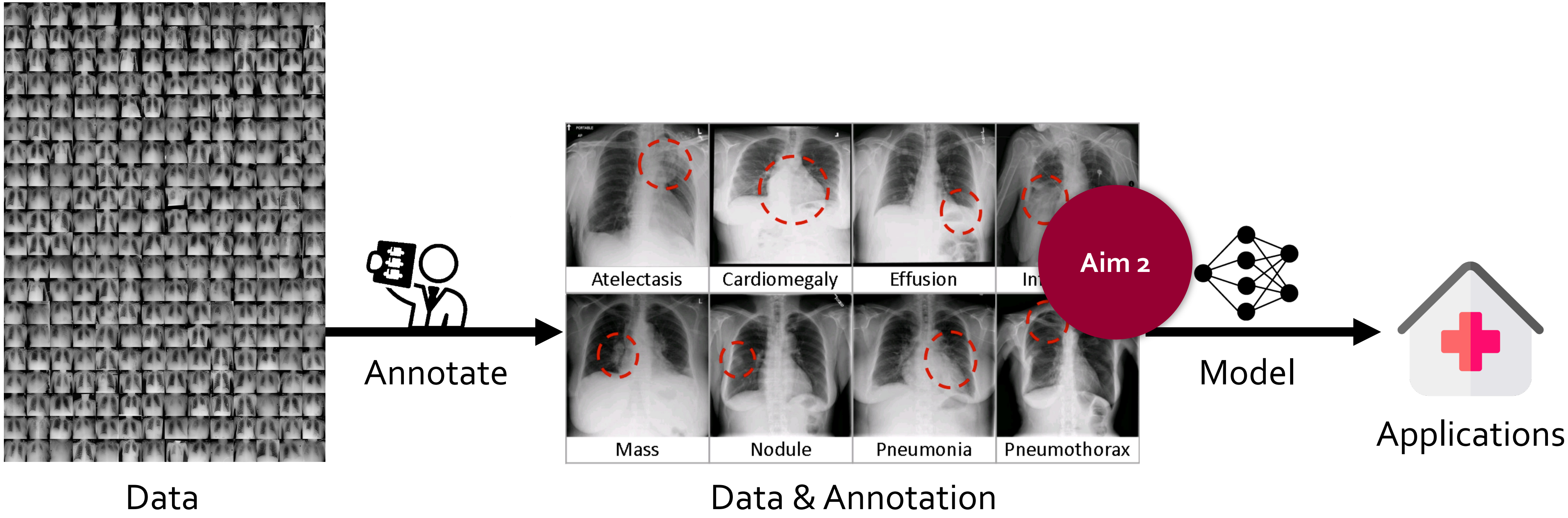
Aim 3

Summary

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

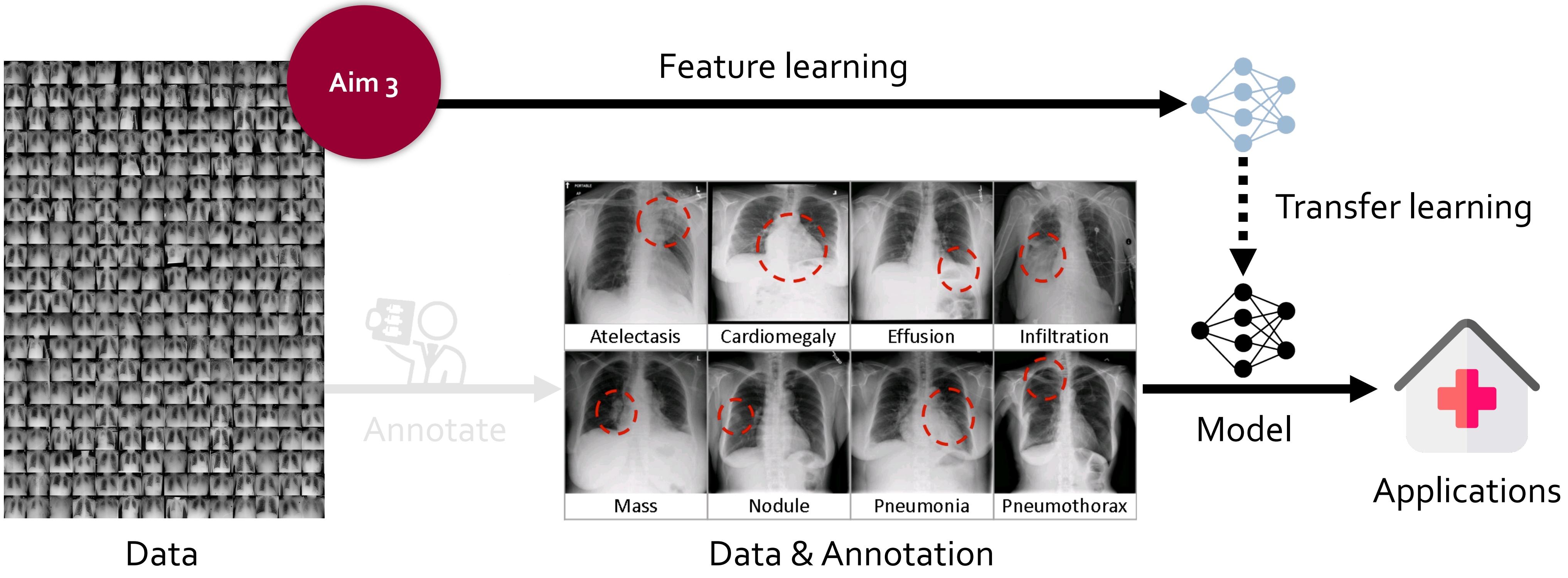


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

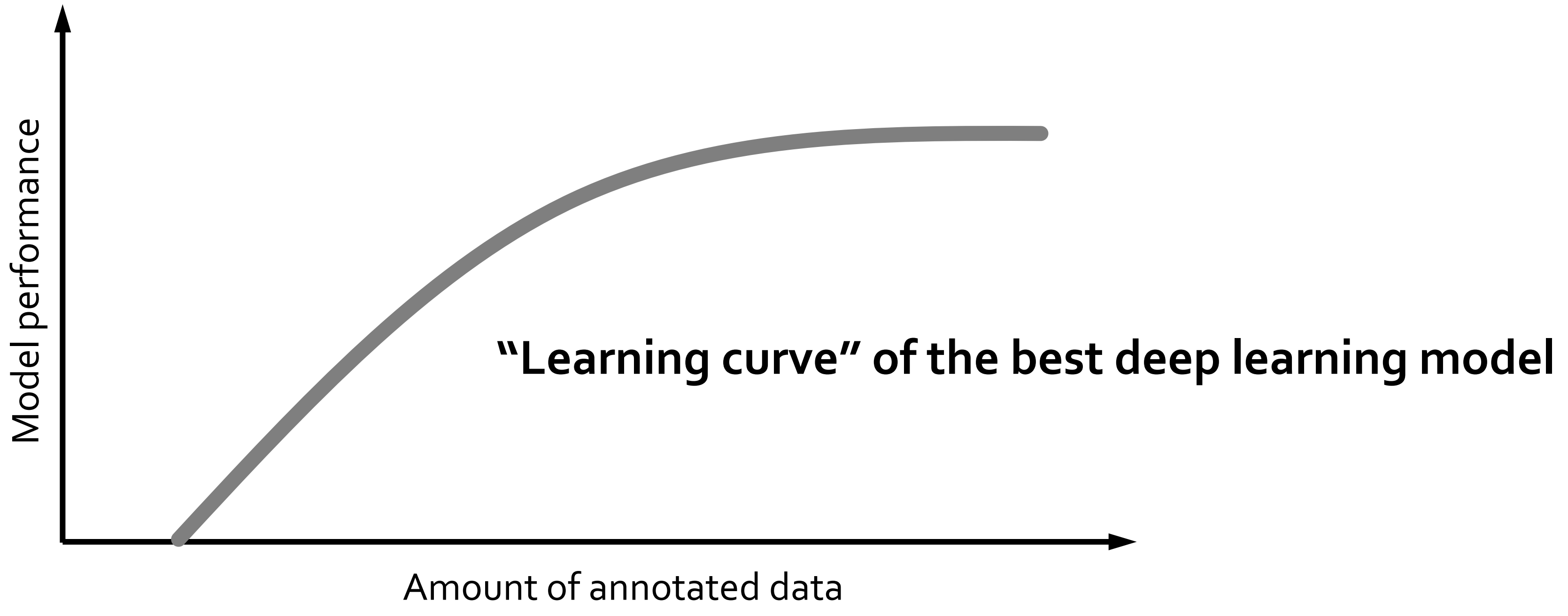


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images



Introduction

Objective

Aim 1

Aim 2

Aim 3

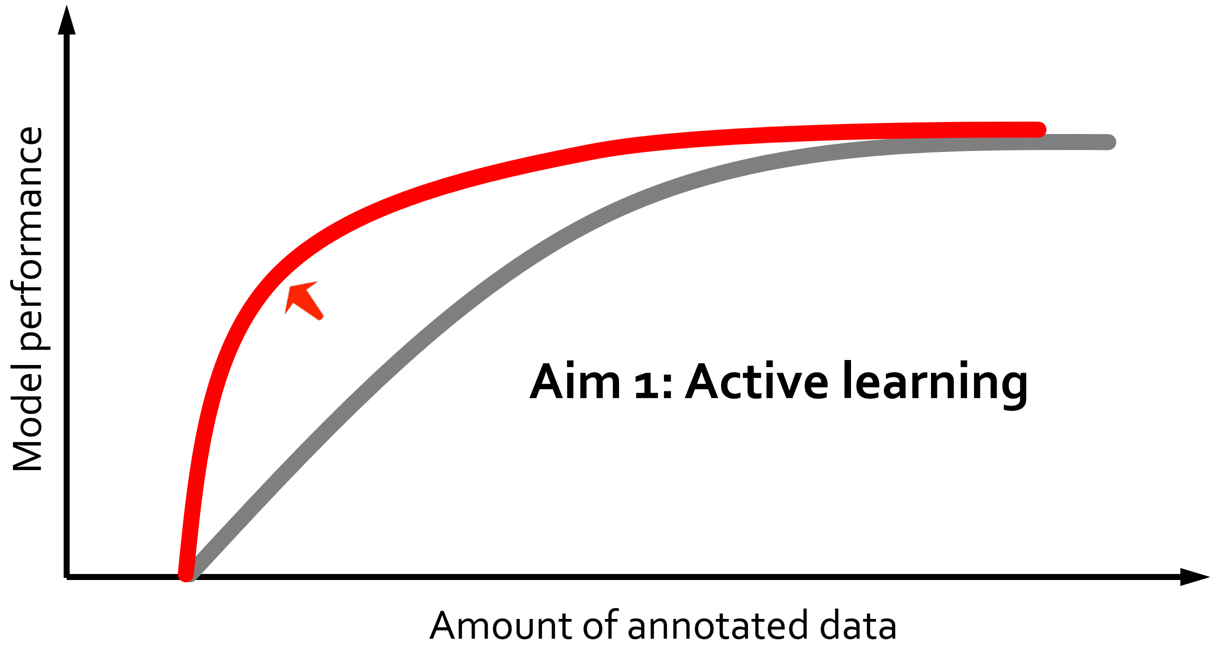
Summary

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

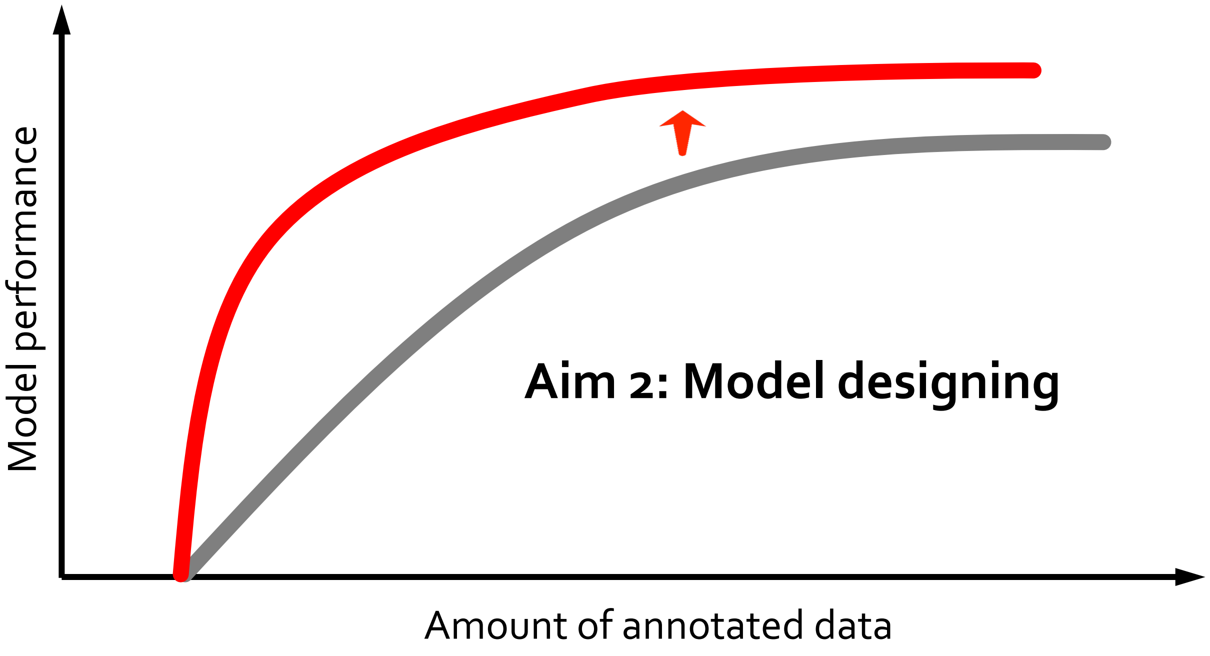


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images



Introduction

Objective

Aim 1

Aim 2

Aim 3

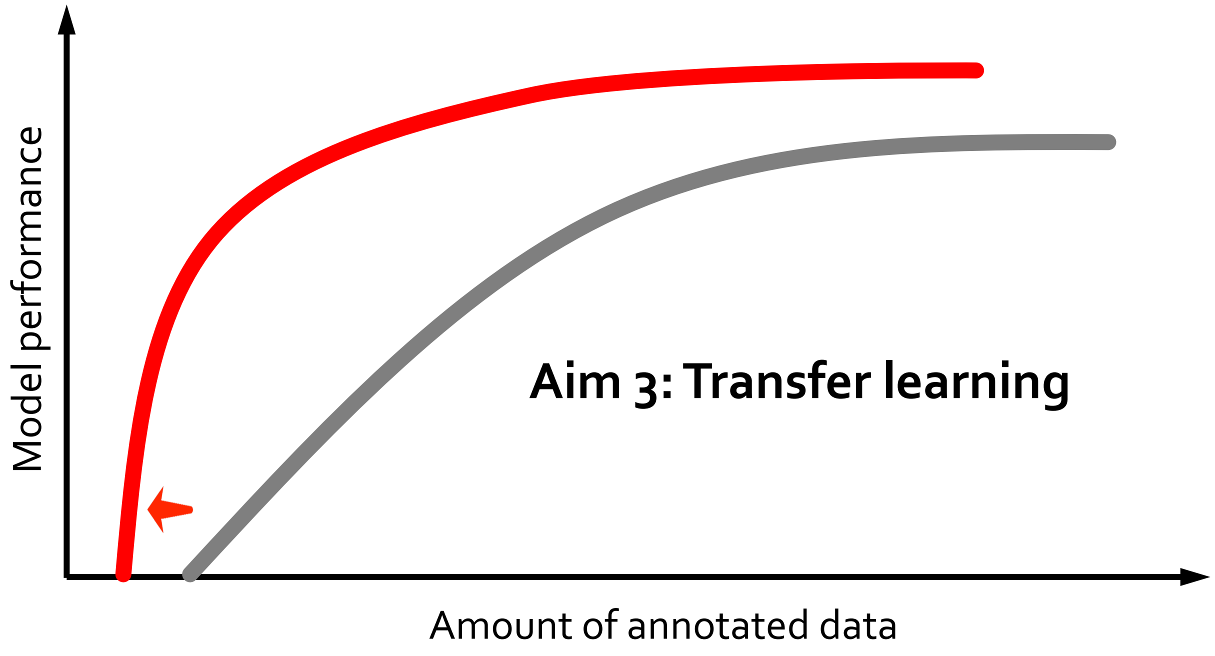
Summary

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images



Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

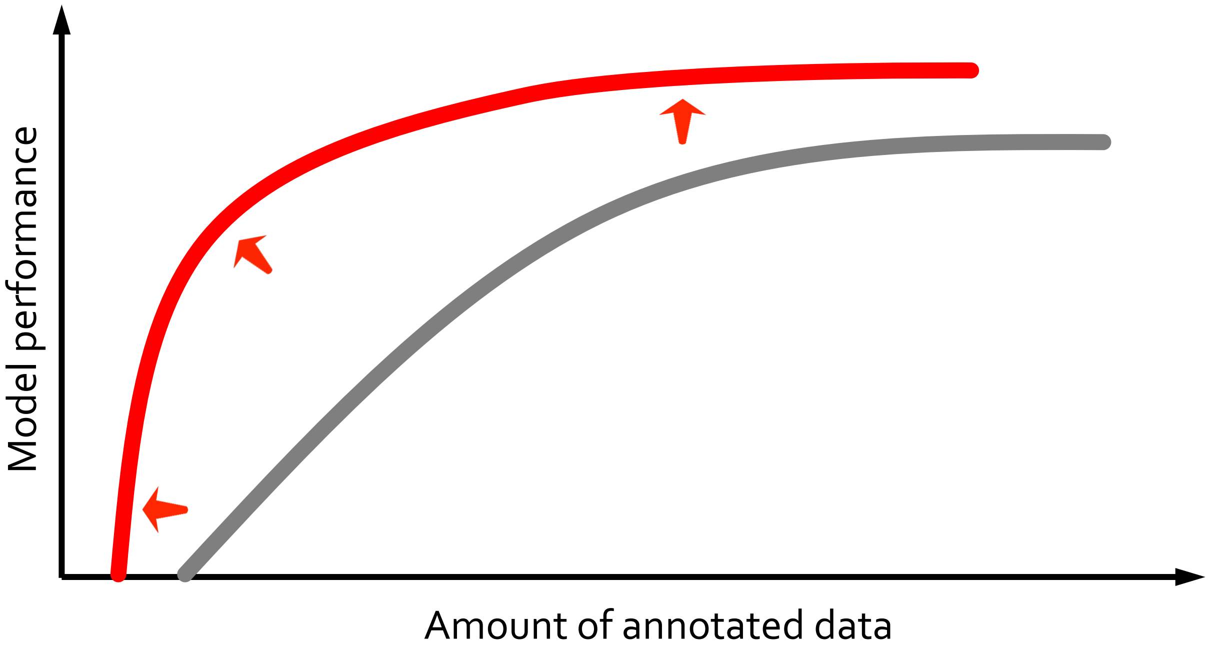
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Hypothesis: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset.



Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

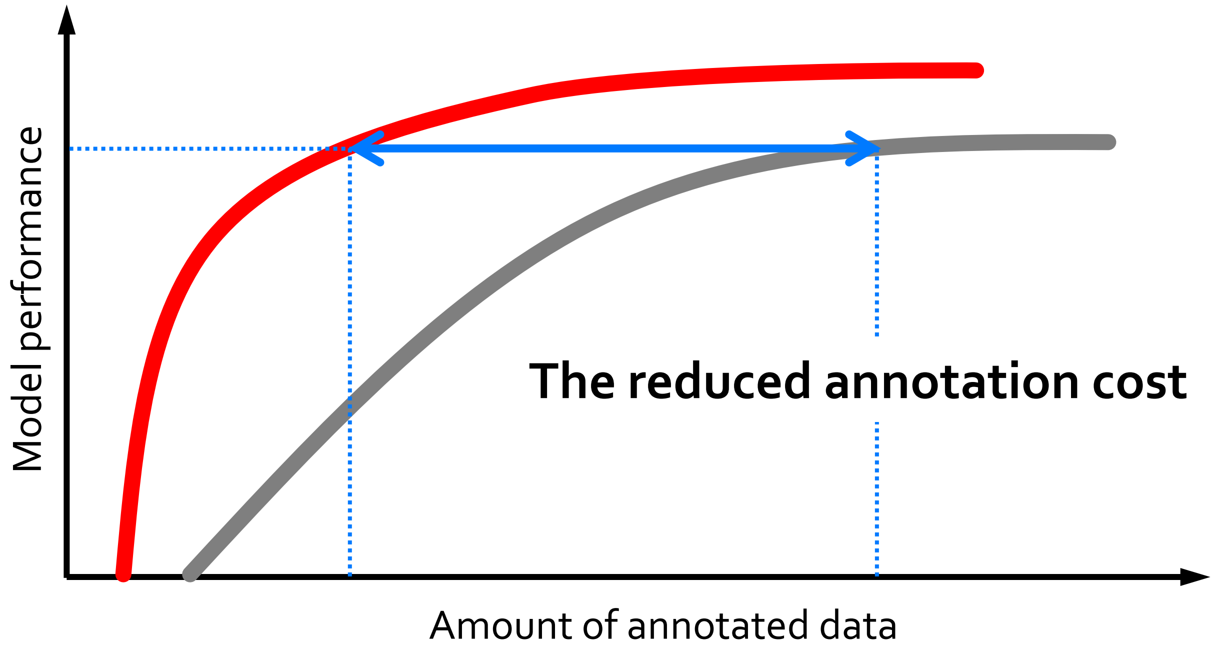
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Hypothesis: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset.



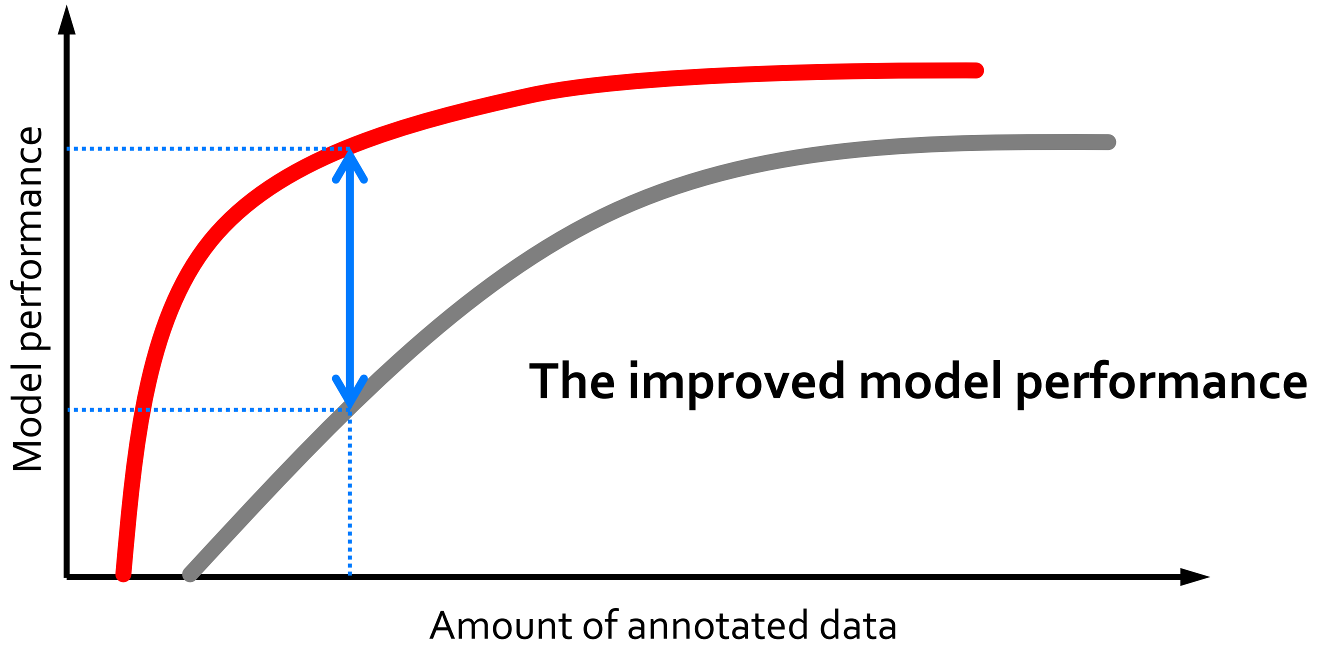
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Hypothesis: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset.



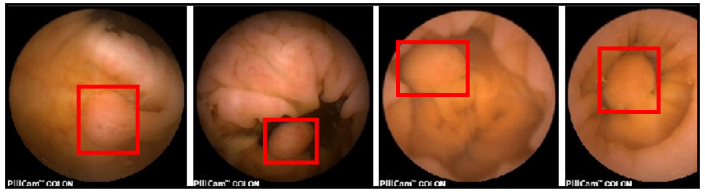
Introduction
Objective
Aim 1
Aim 2
Aim 3
Summary

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

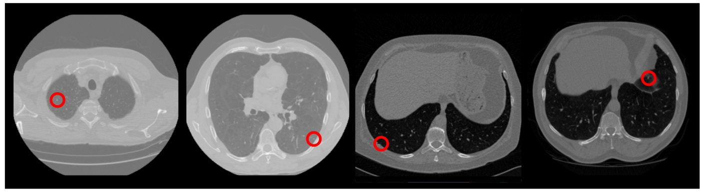
Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

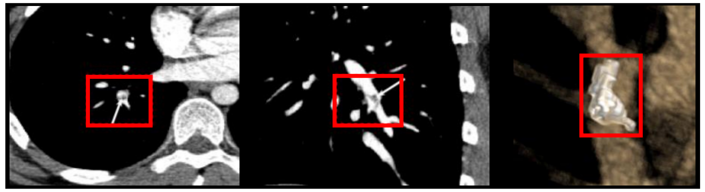
Aim 3: Extracting generic knowledge directly from unannotated images



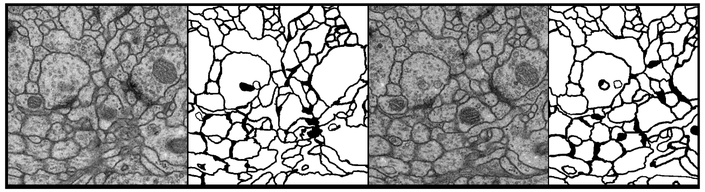
Polyp detection



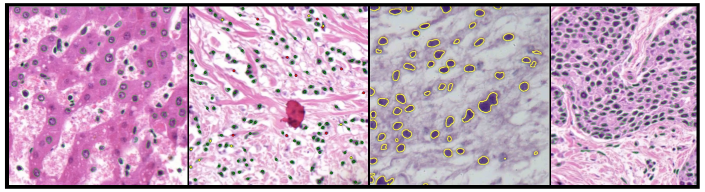
Lung nodule detection



Pulmonary embolism detection

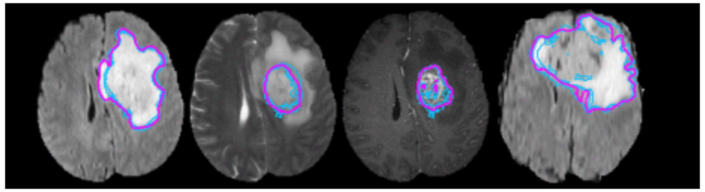


Neuronal structure segmentation

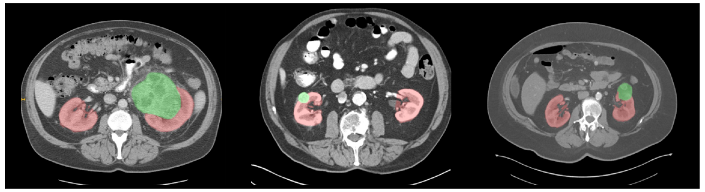


Cell/nuclei segmentation

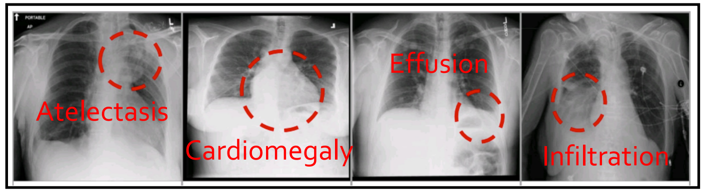
Liver/lesion segmentation



Brain/tumor segmentation



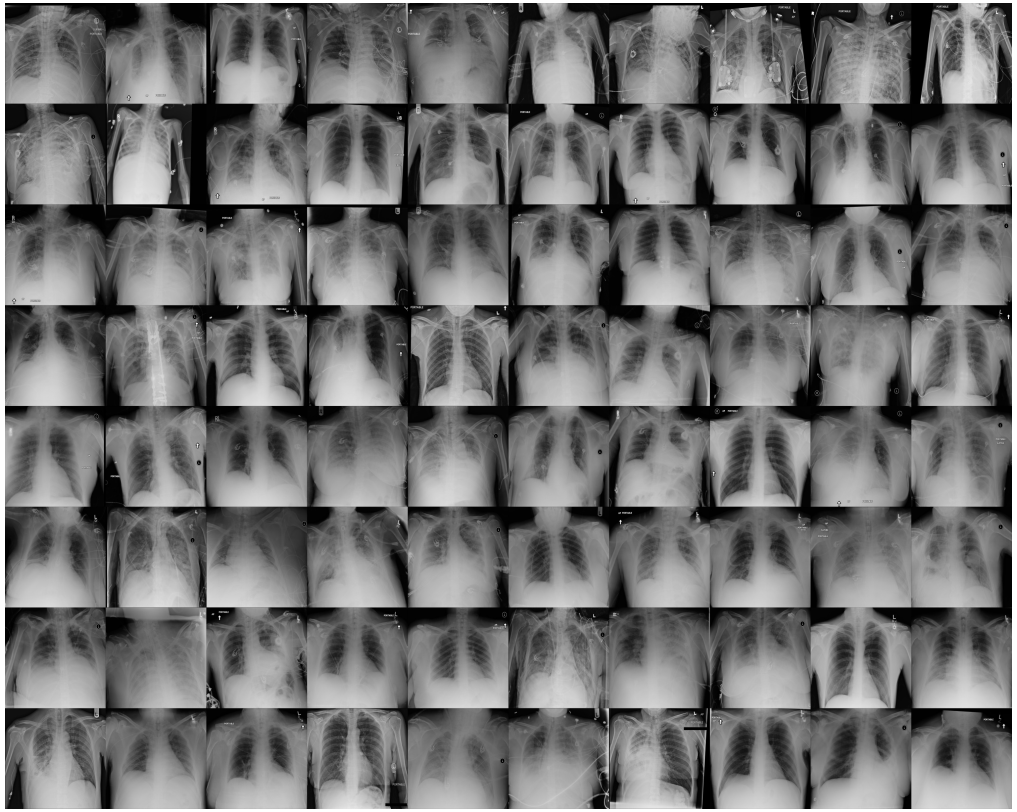
Kidney/lesion segmentation

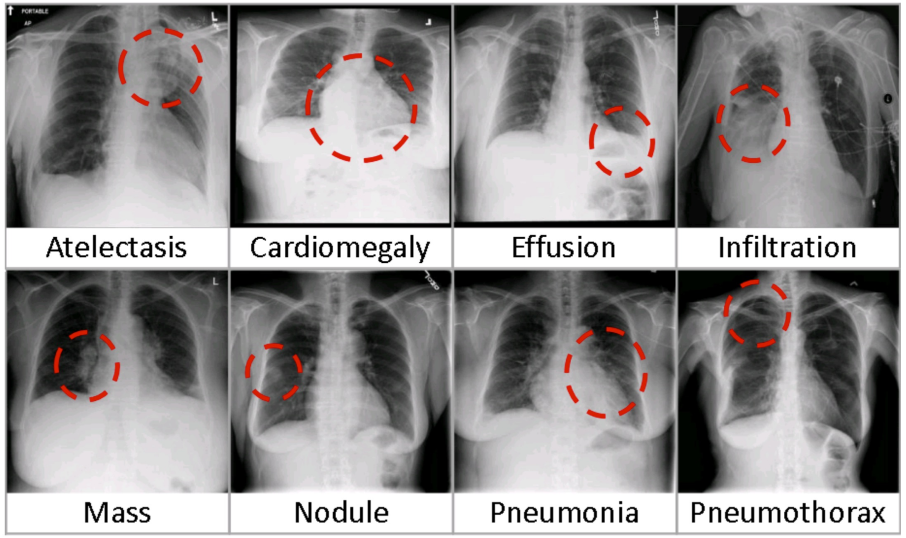


Pulmonary diseases classification

Aim 1: Acquiring necessary annotation efficiently from human experts

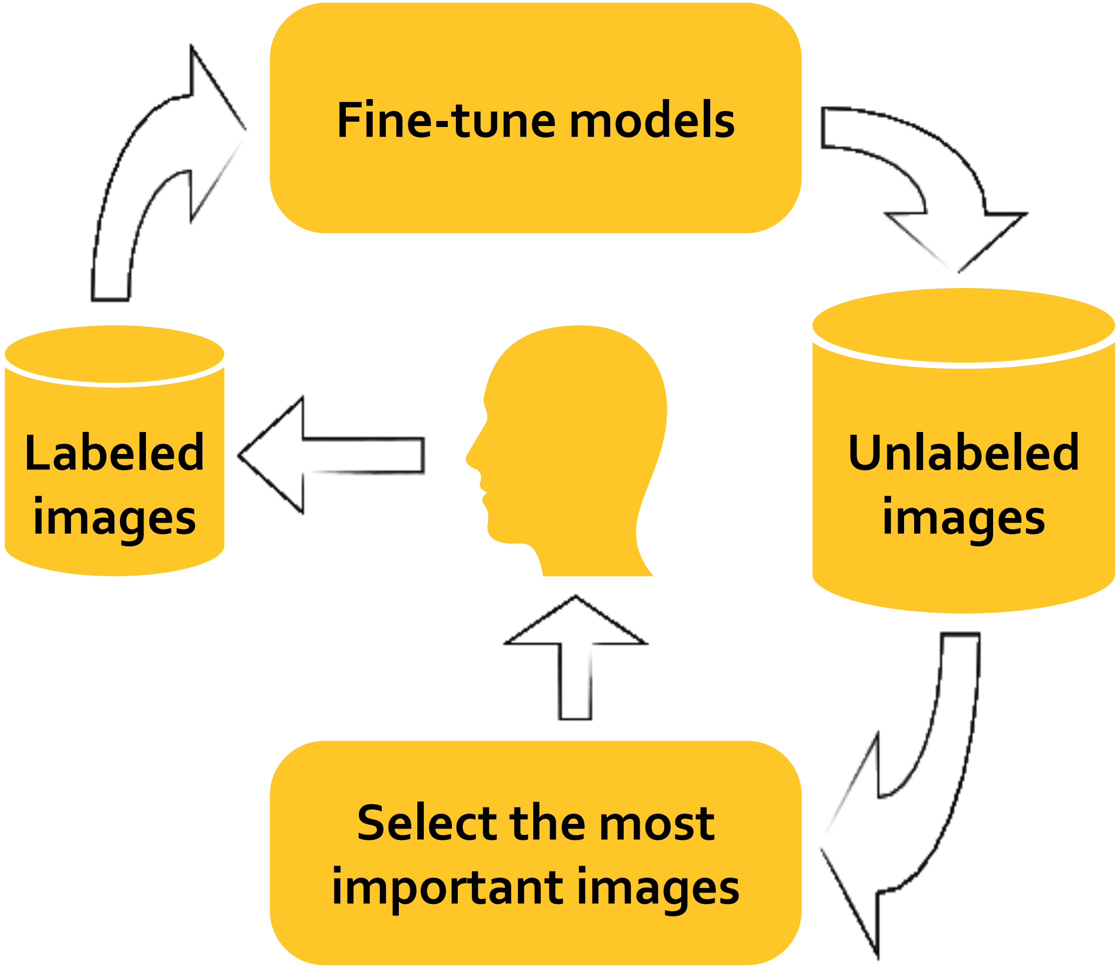
Task: Find the most important 1,000 images from 1,000,000 images





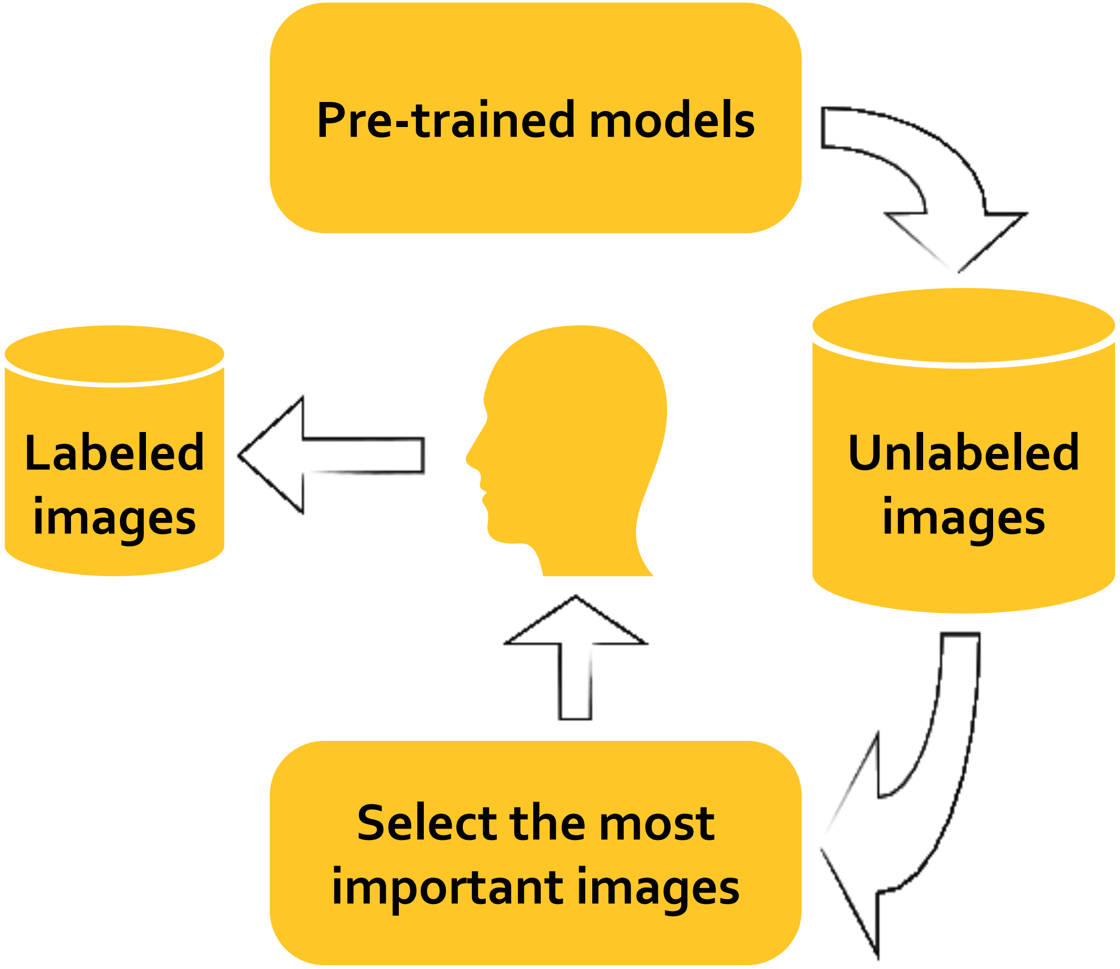
Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure



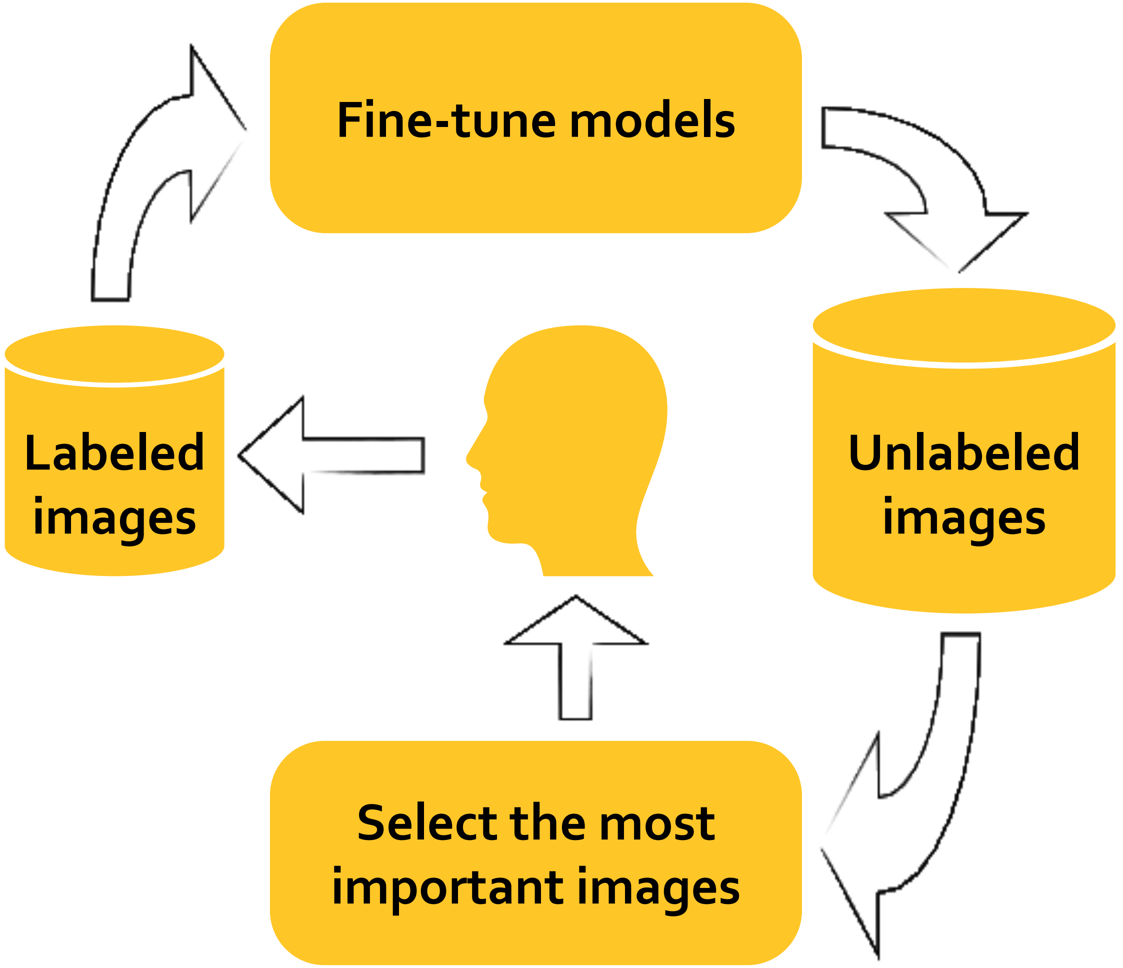
Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure



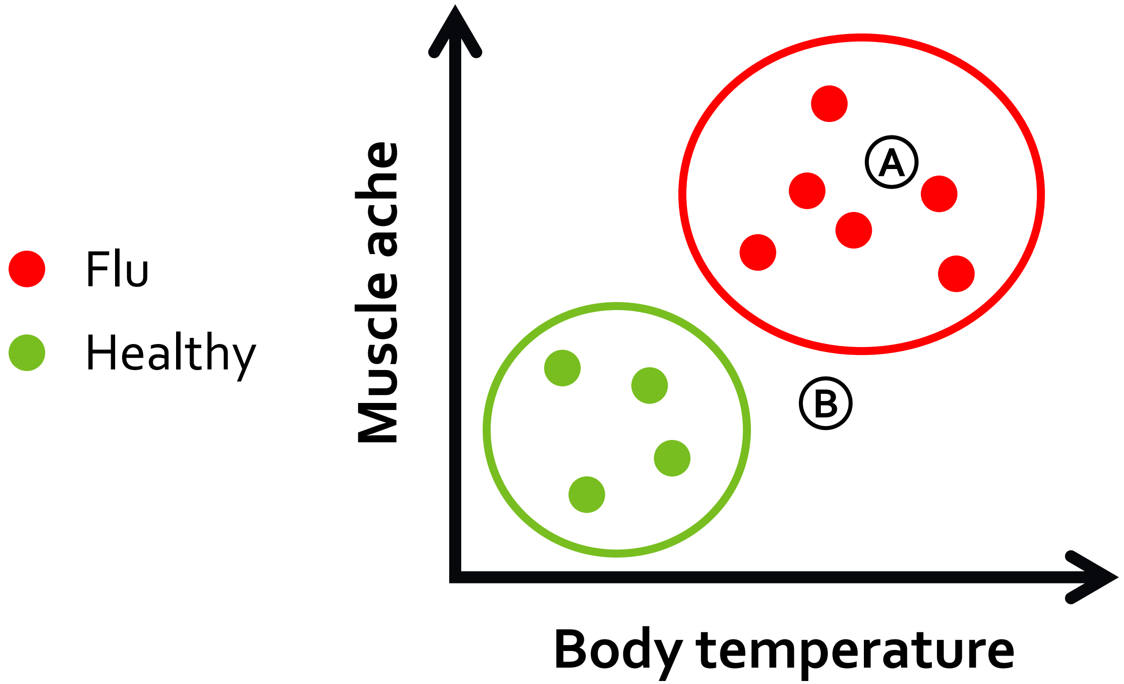
Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure



Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure

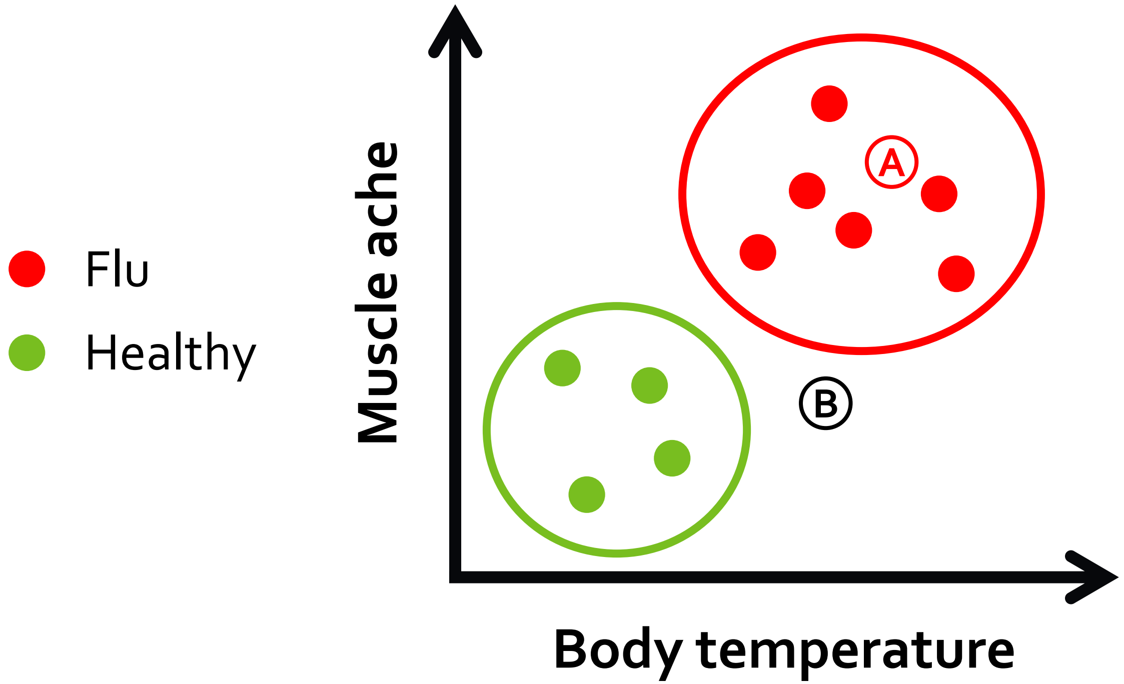


Select the most important images

Which sample would you annotate first, A or B?

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure

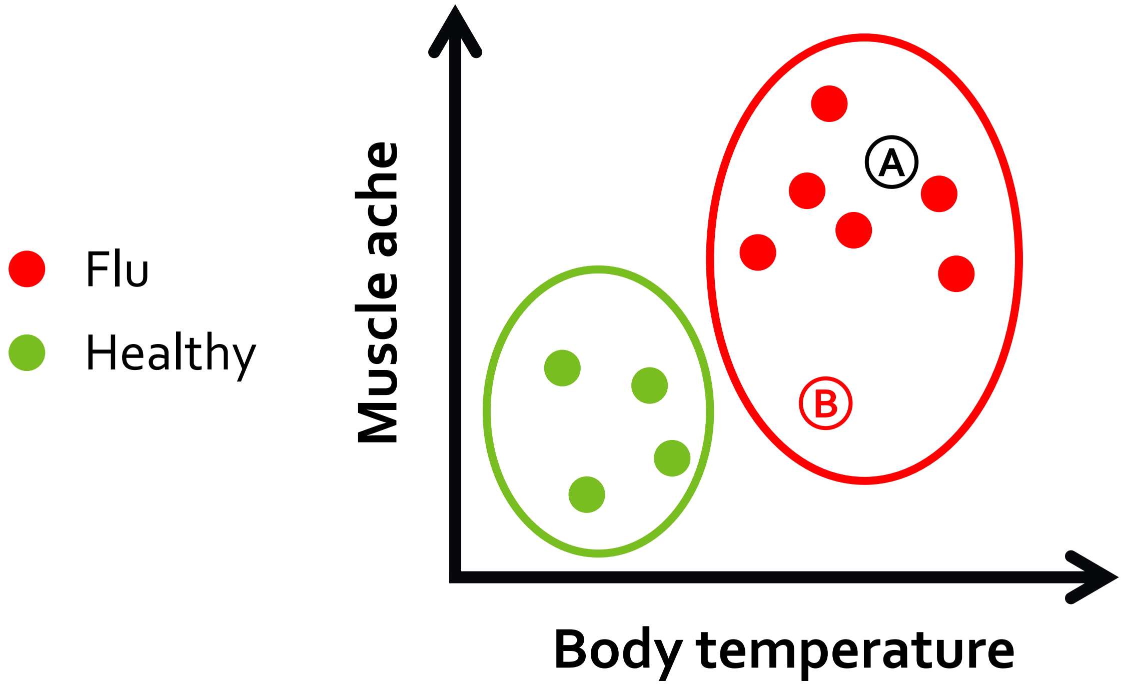


Select the most important images

Which sample would you annotate first, A or B?

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure

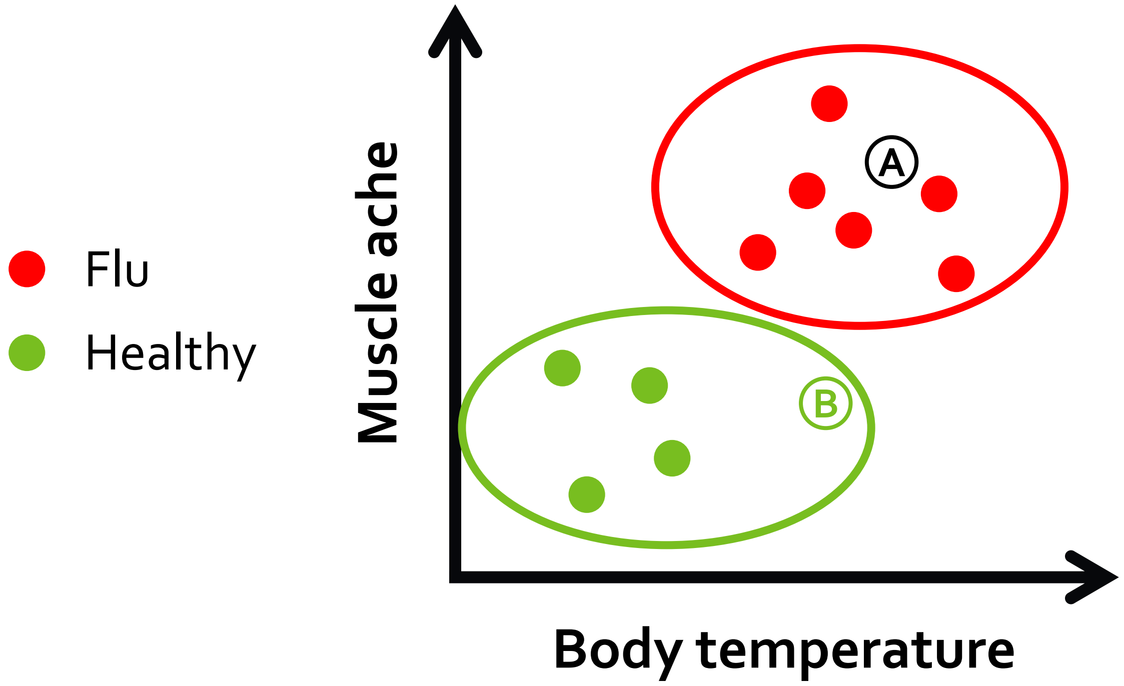


Select the most important images

Which sample would you annotate first, A or B?

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure



Select the most important images

Which sample would you annotate first, A or B?

Introduction

Objective

Aim 1

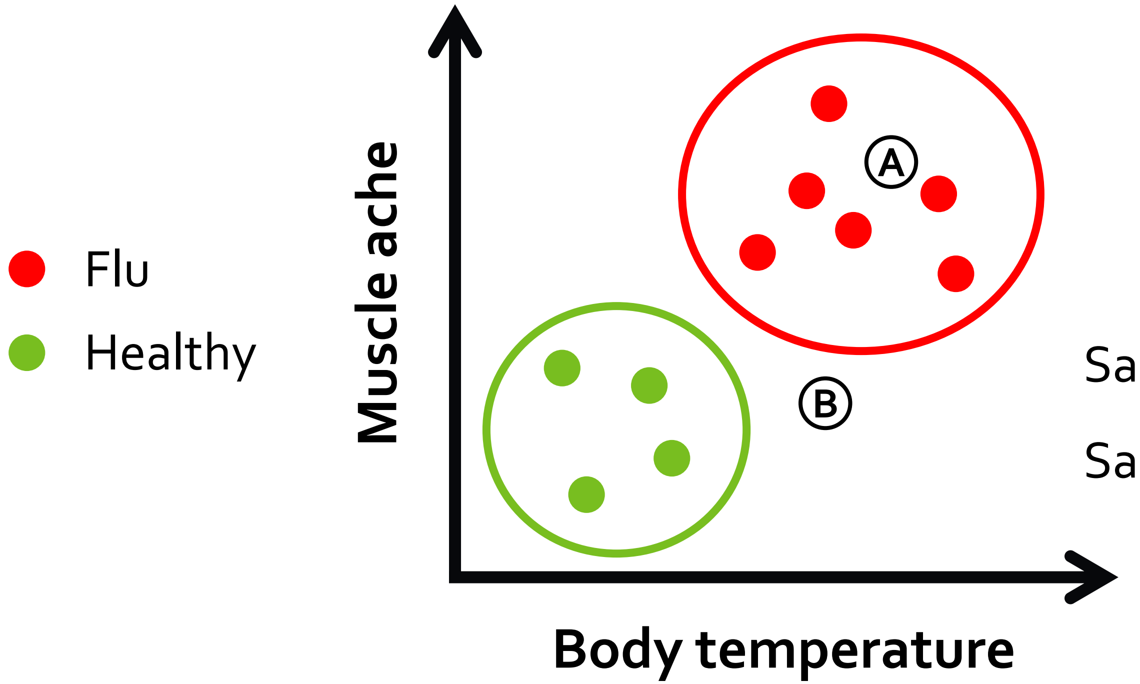
Aim 2

Aim 3

Summary

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure, *uncertainty-based*



Sample A: *low* uncertainty

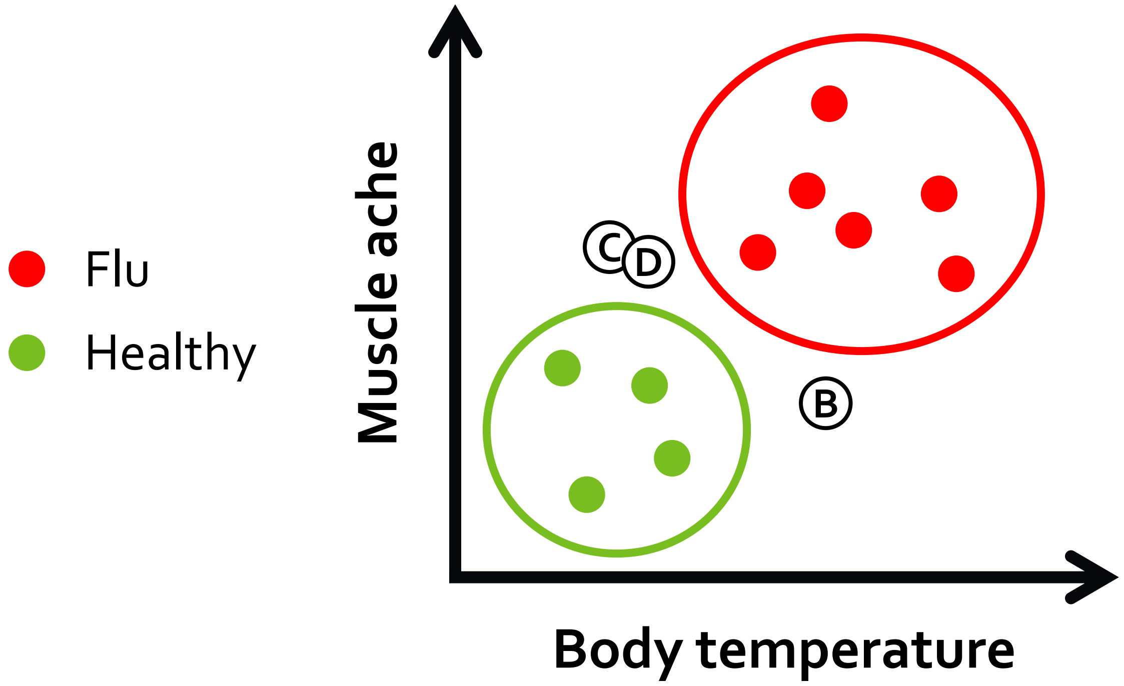
Sample B: *high* uncertainty

Select the most important images

Which sample would you annotate first, A or B?

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure, *diversity-based*

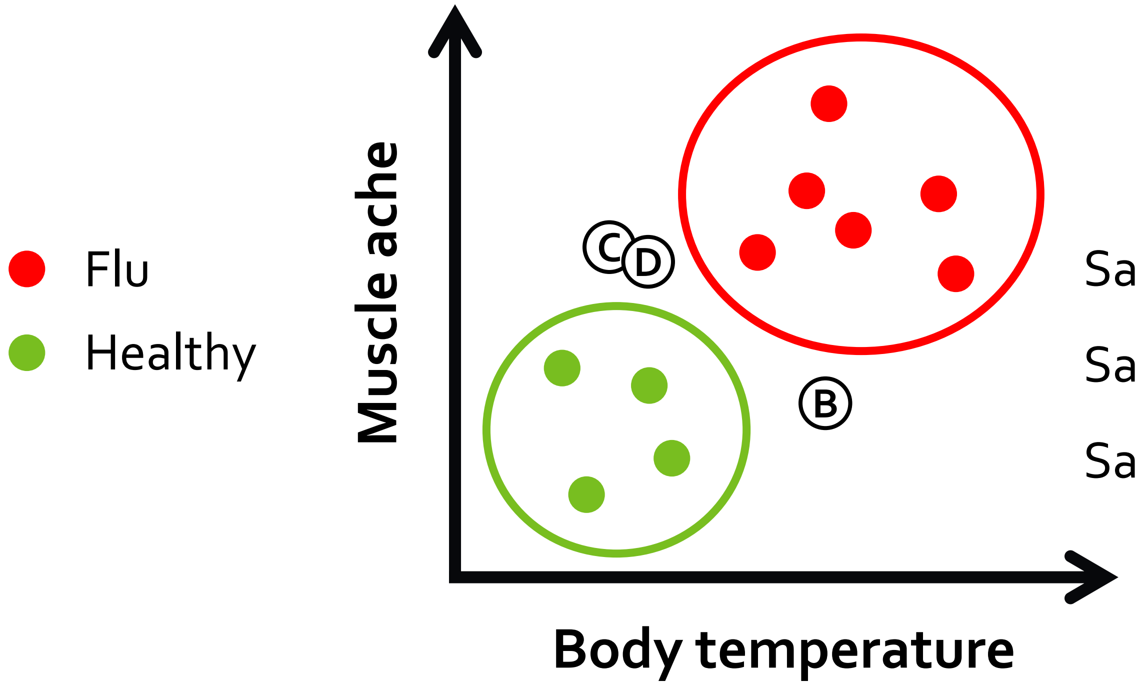


Select the most important images

Which two samples would you annotate first?

Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: "Human-in-the-loop" active learning procedure, *diversity-based*



Samples B, C: *highest* diversity

Samples B, D: *high* diversity

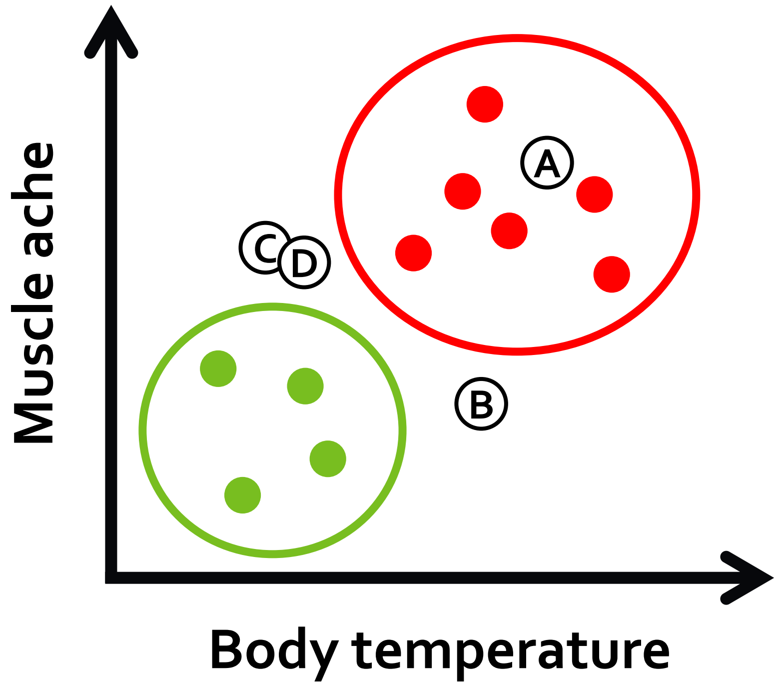
Samples C, D: *low* diversity

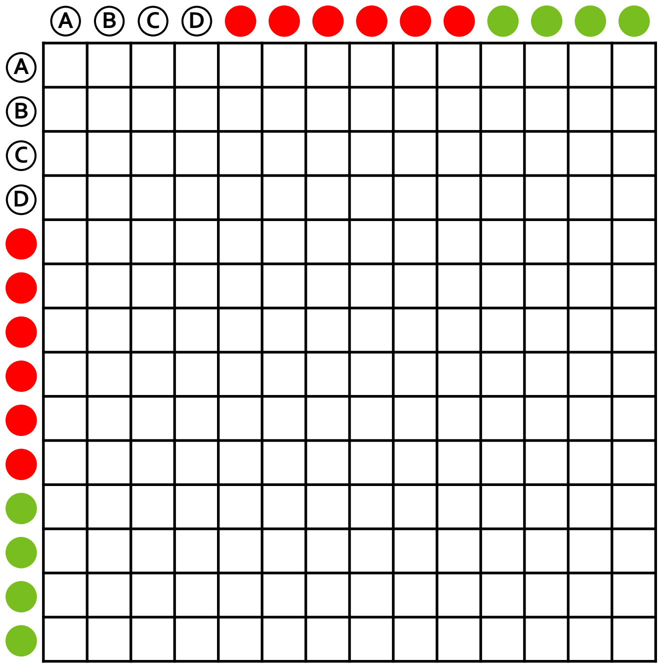
Select the most important images

`Which two samples would you annotate first?`

Aim 1: Acquiring necessary annotation efficiently from human experts

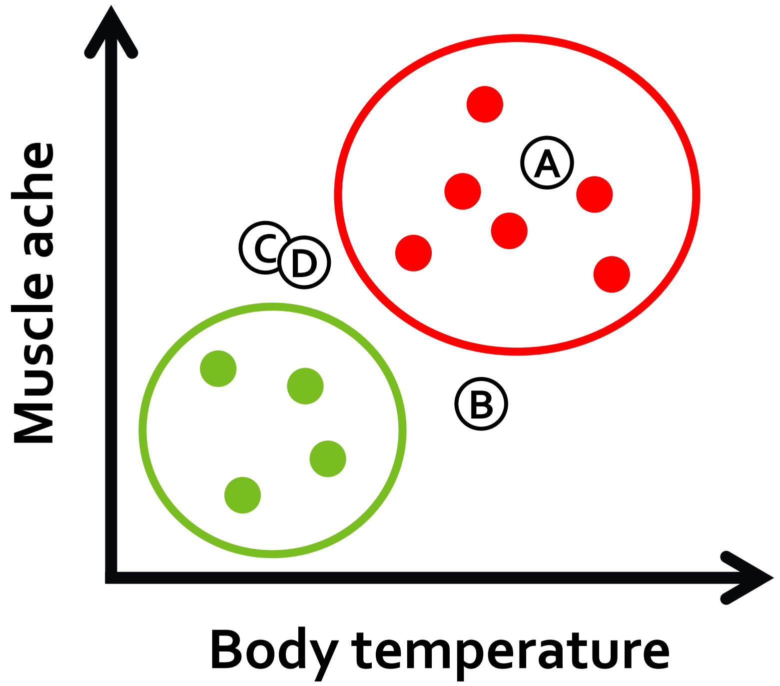
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*

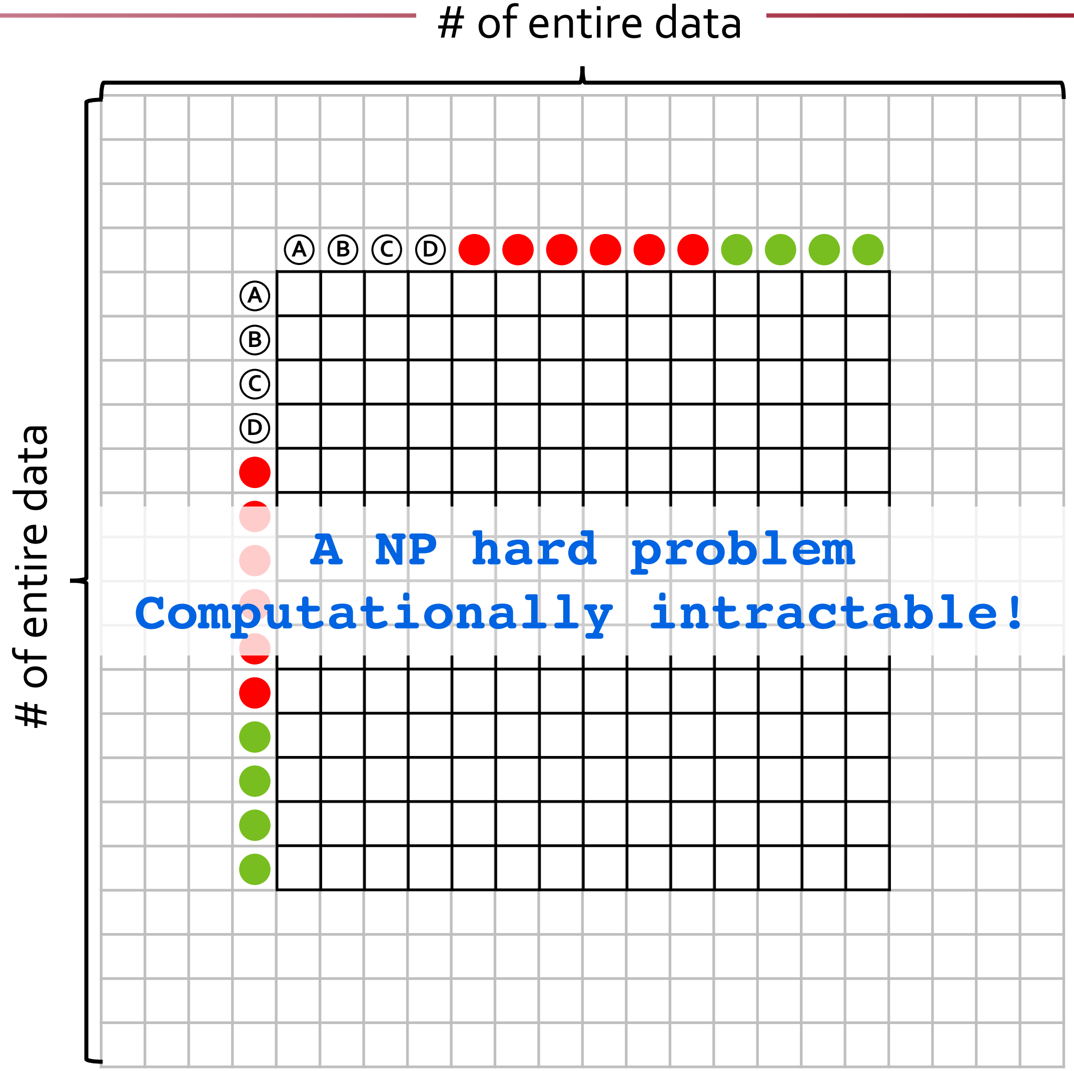




Aim 1: Acquiring necessary annotation efficiently from human experts

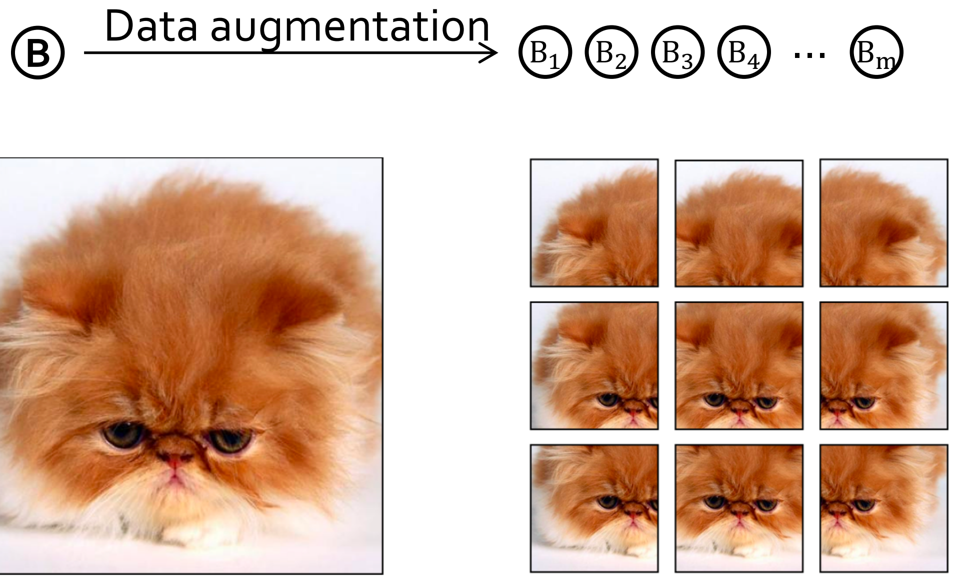
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*





Aim 1: Acquiring necessary annotation efficiently from human experts

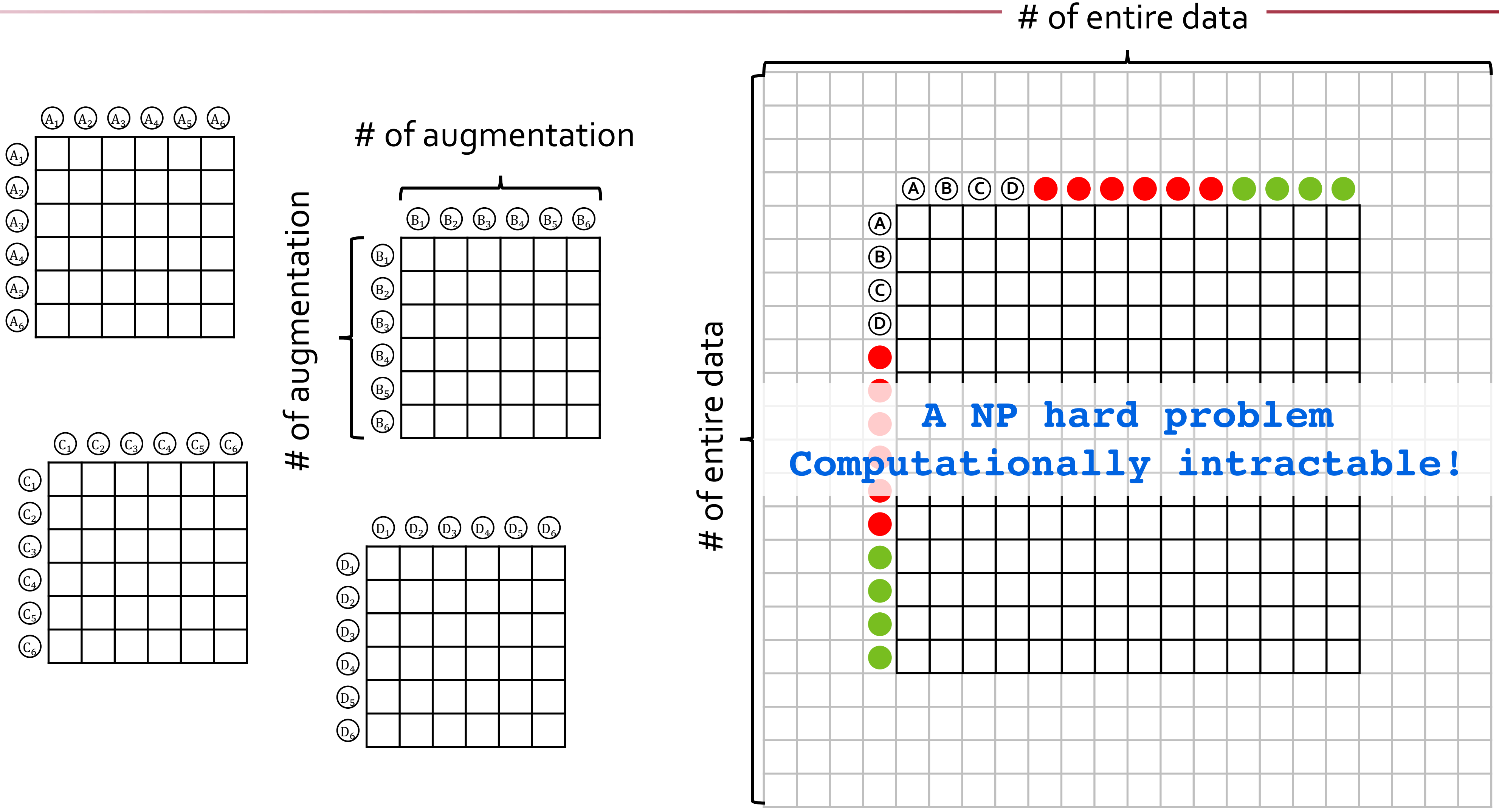
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*



To train the deep model, many patches are usually generated via data augmentation; these patches generated from the same image share the ***same label*** *(cat)*, and they are expected to have ***similar predictions*** by the current model.

Aim 1: Acquiring necessary annotation efficiently from human experts

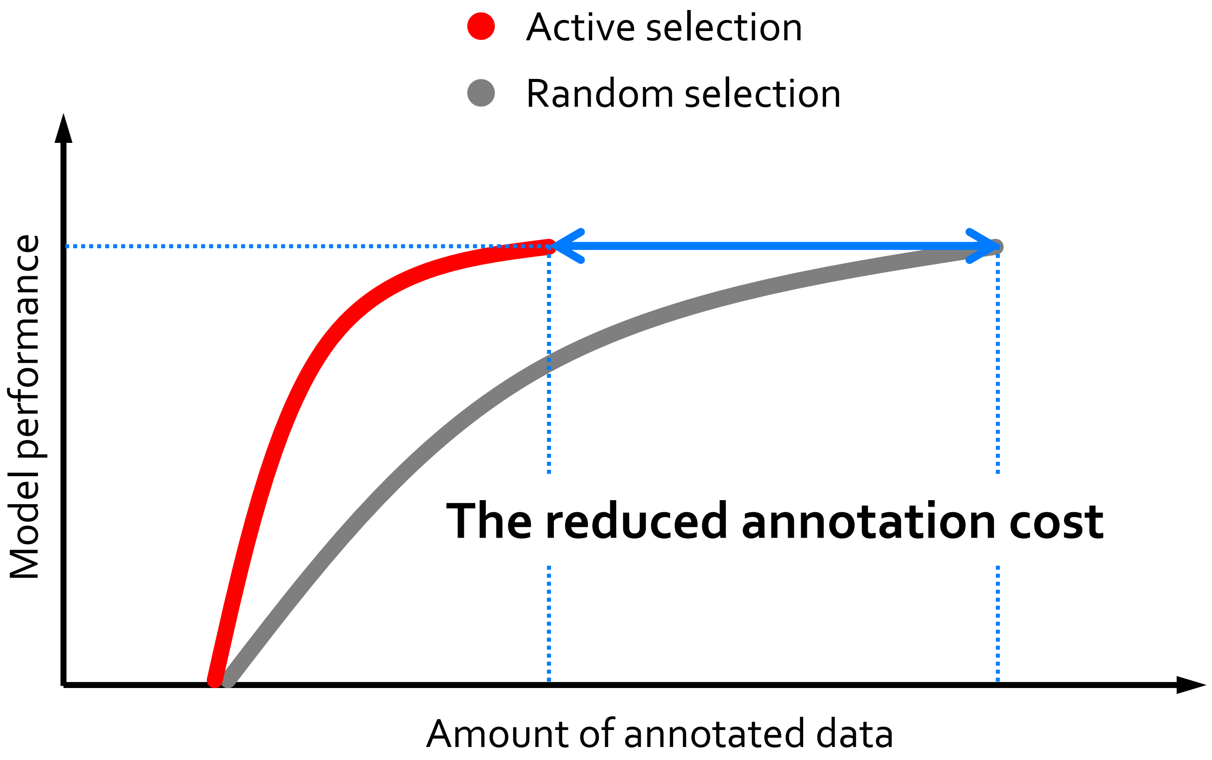
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*



Introduction
Objective
Aim 1
Aim 2
Aim 3
Summary

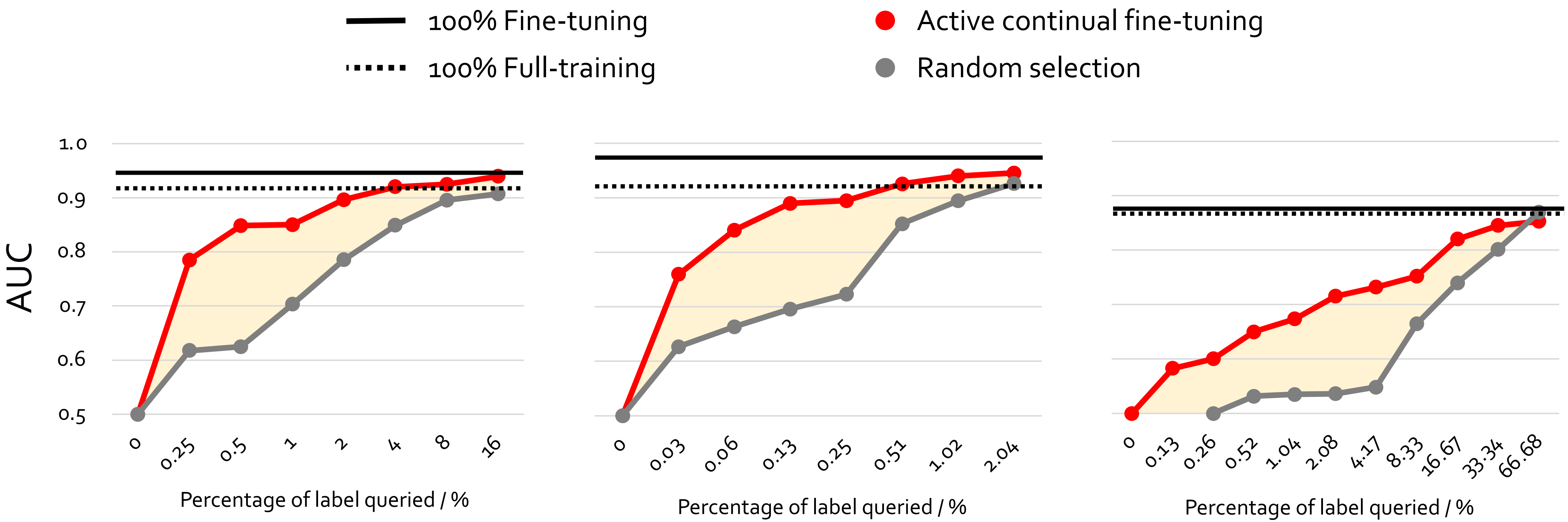
Aim 1: Acquiring necessary annotation efficiently from human experts

Hypothesis: Wisely selecting important samples can reduce annotation cost



Aim 1: Acquiring necessary annotation efficiently from human experts

Contribution: Reduce annotation cost by over 80% compared with random selection



Colonoscopy Frame Classification

Polyp Detection

Pulmonary Embolism Detection

1. Zhou, Zongwei, *et al.* "Integrating active learning and transfer learning for carotid intima-media thickness video interpretation." Journal of digital imaging 32.2 (2019): 290-299.
2. Zhou, Zongwei, *et al.* "Active, Continual Fine Tuning of Convolutional Neural Networks for Reducing Annotation Efforts." Medical Image Analysis (2021): 101997.
3. Zhou, Zongwei, *et al.* "Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7340-7351. 2017.

Introduction

Objective

Aim 1

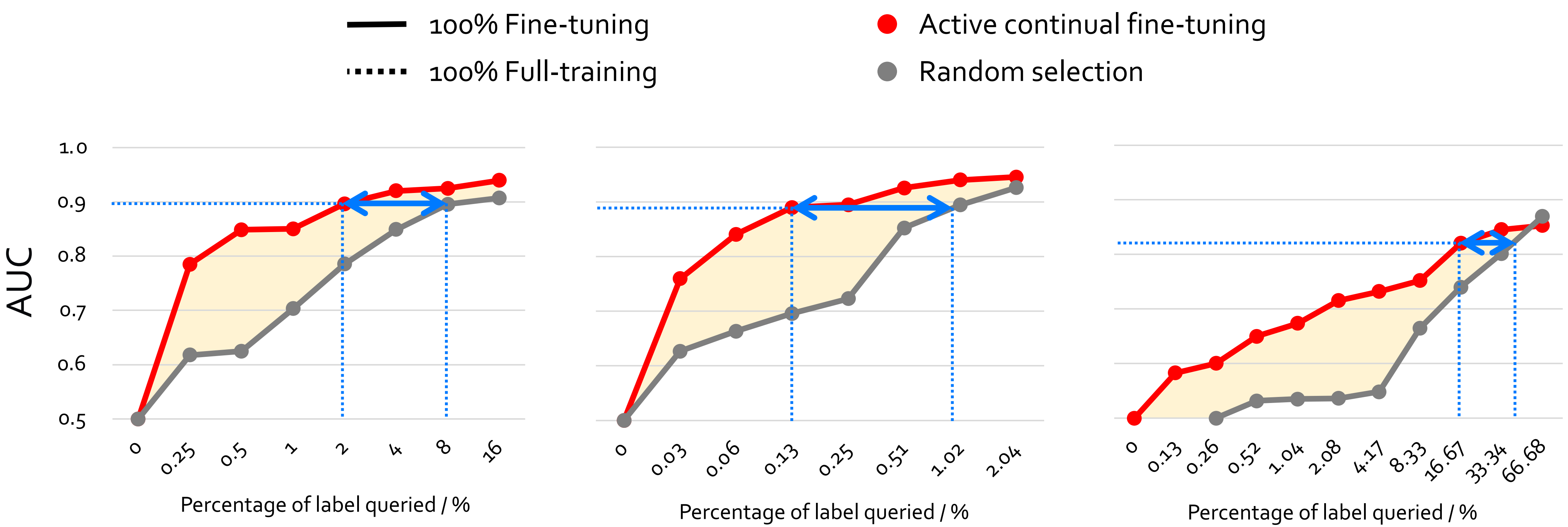
Aim 2

Aim 3

Summary

Aim 1: Acquiring necessary annotation efficiently from human experts

Contribution: Reduce annotation cost by over 80% compared with random selection



Colonoscopy Frame Classification (-81.5%)

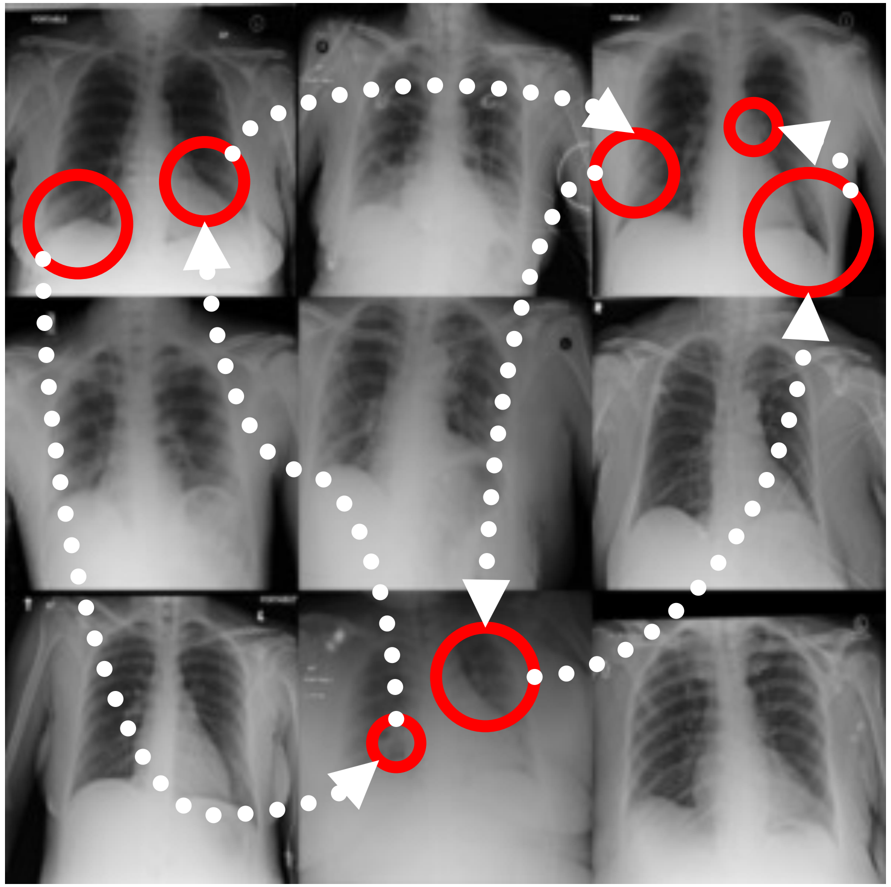
Polyp Detection (-86.3%)

Pulmonary Embolism Detection (-80.3%)

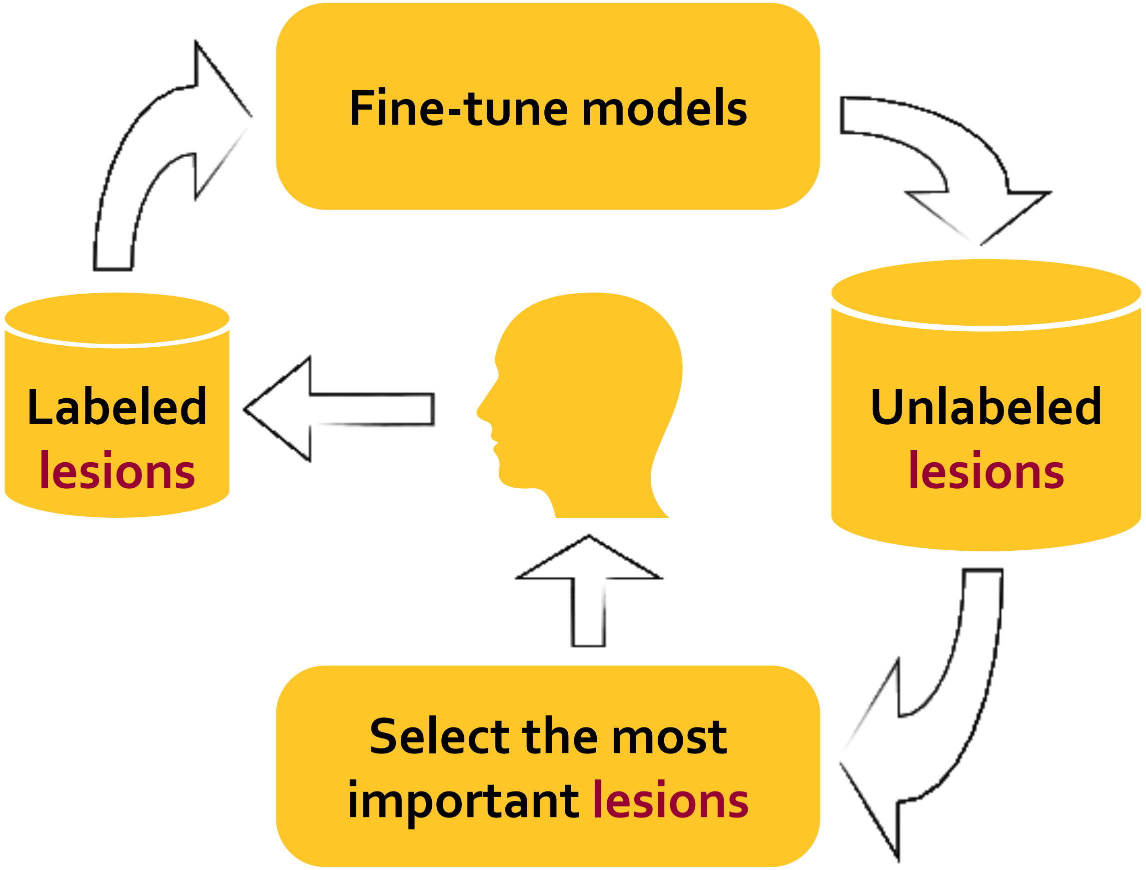
1. Zhou, Zongwei, *et al.* "Integrating active learning and transfer learning for carotid intima-media thickness video interpretation." Journal of digital imaging 32.2 (2019): 290-299.
2. Zhou, Zongwei, *et al.* "Active, Continual Fine Tuning of Convolutional Neural Networks for Reducing Annotation Efforts." Medical Image Analysis (2021): 101997.
3. Zhou, Zongwei, *et al.* "Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7340-7351. 2017.

Aim 1: Acquiring necessary annotation efficiently from human experts

Discussion: Iteratively suggest important samples at the patient-level

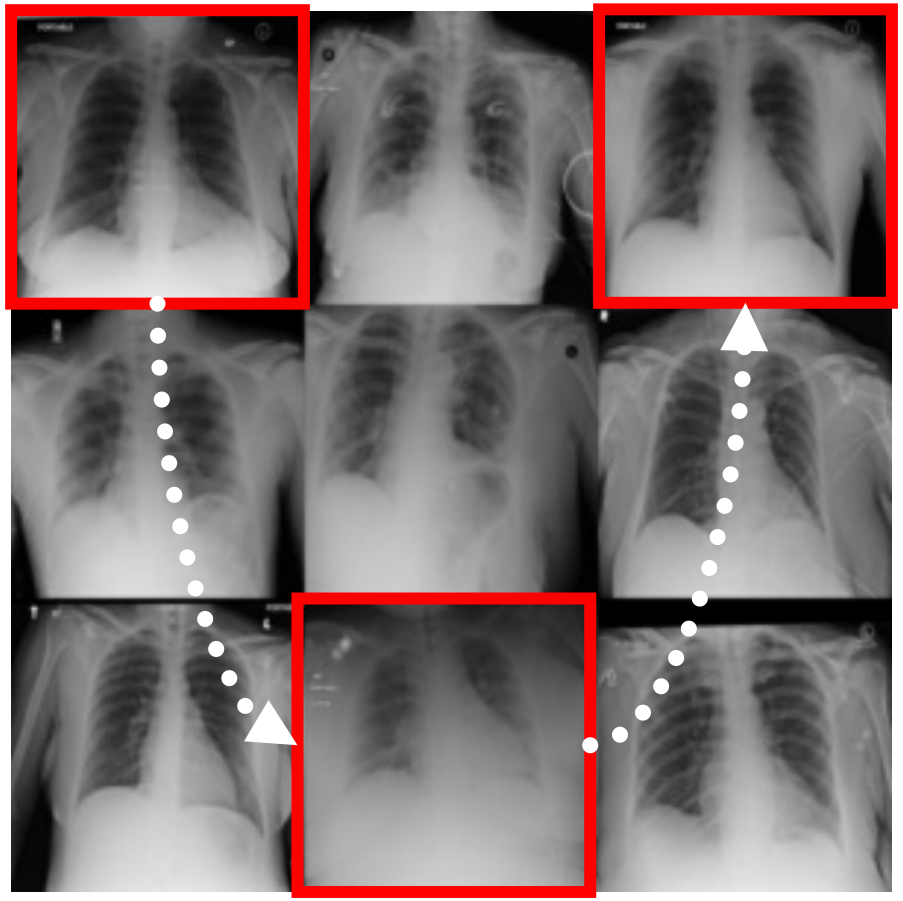


Lesion-level annotation

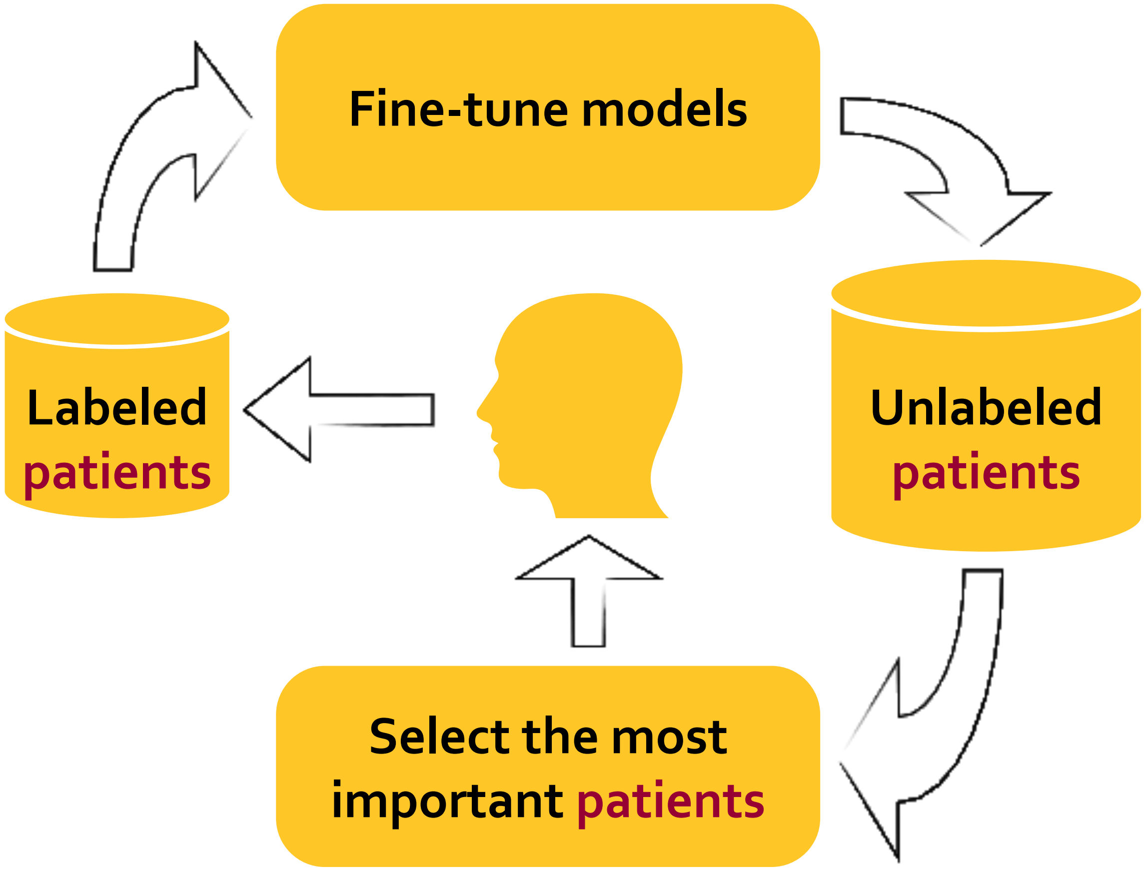


Aim 1: Acquiring necessary annotation efficiently from human experts

Discussion: Iteratively suggest important samples at the patient-level



Patient-level annotation



Not All Data Is Created Equal

Publications for Aim 1:

- Z. Zhou, J. Shin, L. Zhang, S. Gurudu, M. Gotway, J. Liang, 2017. Fine-tuning Convolutional Neural Networks for Biomedical Image Analysis: Actively and Incrementally. *CVPR'17, one of only five papers in biomedical imaging accepted by CVPR'17.*
- Z. Zhou, J. Shin, S. Gurudu, M. Gotway, J. Liang, 2021. Active, Continual Fine Tuning of Convolutional Neural Networks for Reducing Annotation Efforts. *Medical Image Analysis.*
- Z. Zhou, J. Shin, R. Feng, R. Hurst, C. Kendall, J. Liang, 2019. Integrating Active Learning and Transfer Learning for Carotid Intima-Media Thickness Video Interpretation. *Journal of Digital Imaging.*

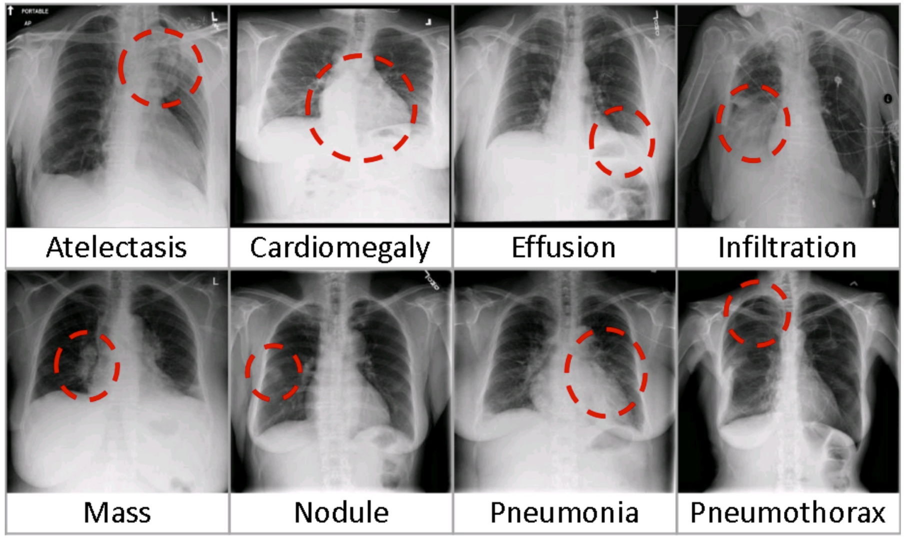
Not All Data Is Created Equal

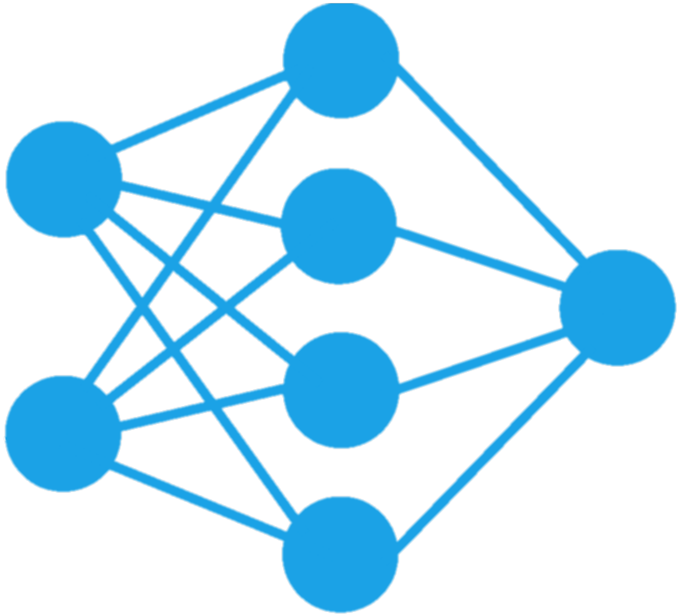
Clinical Impacts of Aim 1:

- The *continual learning capability* of deep models encourages data, label, and model reuse, significantly improving the training efficiency.
- An efficient “human-in-the-loop” procedure assists radiologists in *quickly dismissing patients with negative results*, therefore dramatically reducing the burden of annotation.
- An instant on-line feedback process makes it possible for CAD systems to be *self-learning* and *self-improving* via continual fine-tuning.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Task: Enhance the architecture for modeling 1,000 annotated images





Introduction

Objective

Aim 1

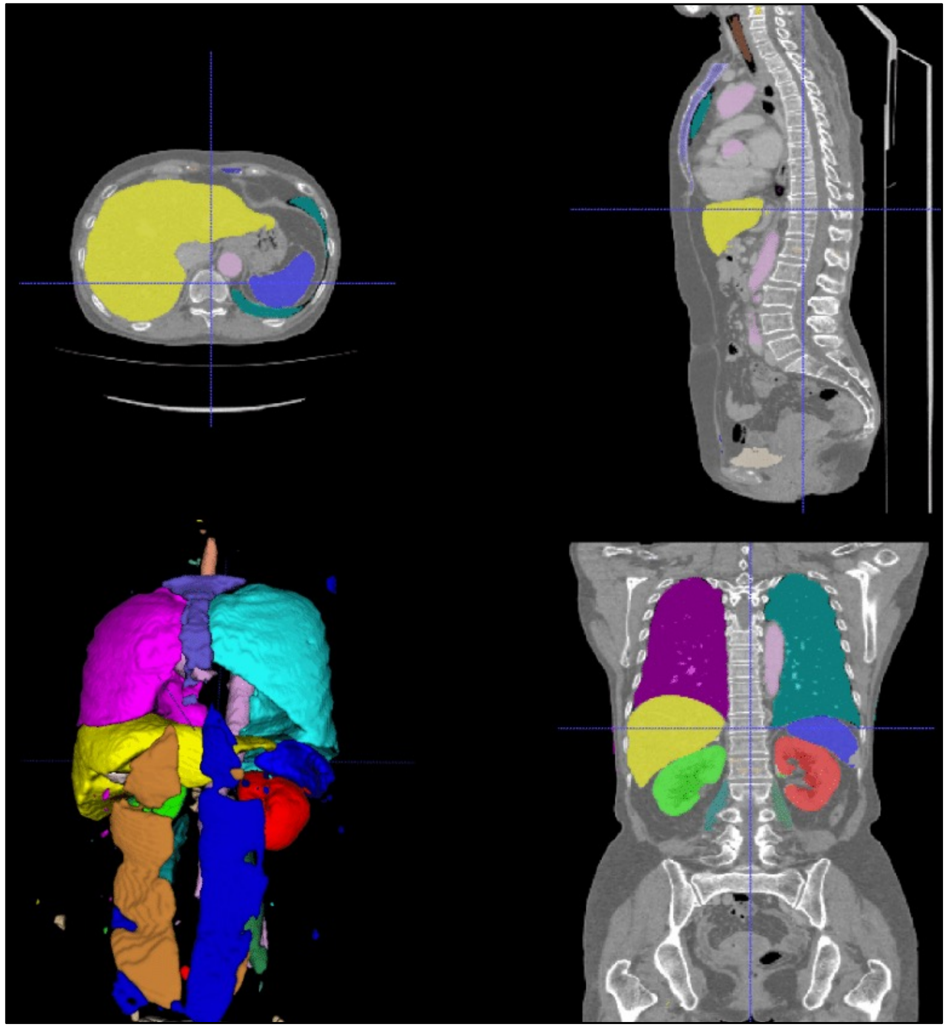
Aim 2

Aim 3

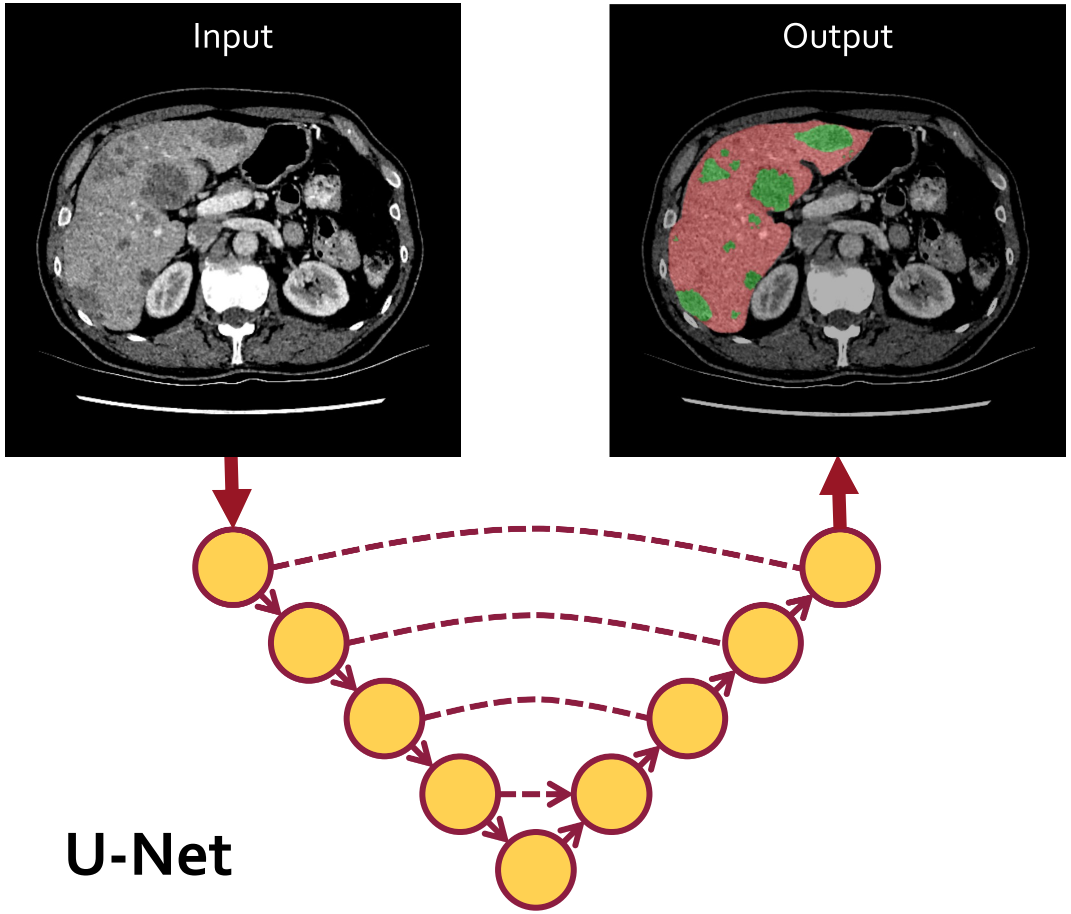
Summary

Aim 2: Utilizing existing annotation effectively from advanced architecture

Segmentation: Partition an image into multiple segments to ease the analysis



e.g., liver & lesion segmentation



Introduction

Objective

Aim 1

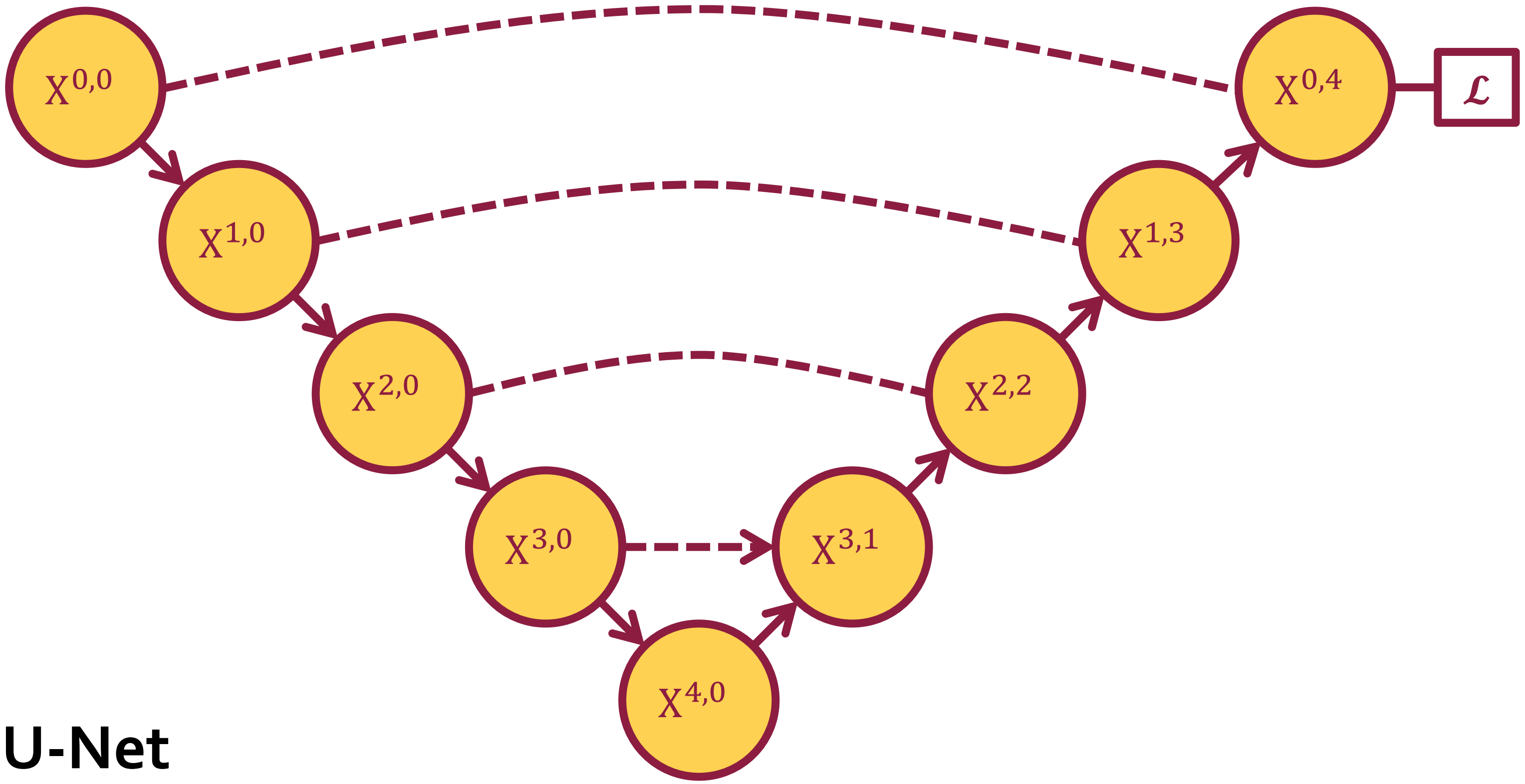
Aim 2

Aim 3

Summary

Aim 2: Utilizing existing annotation effectively from advanced architecture

Hypothesis: Multi-scale feature aggregation leads to powerful models

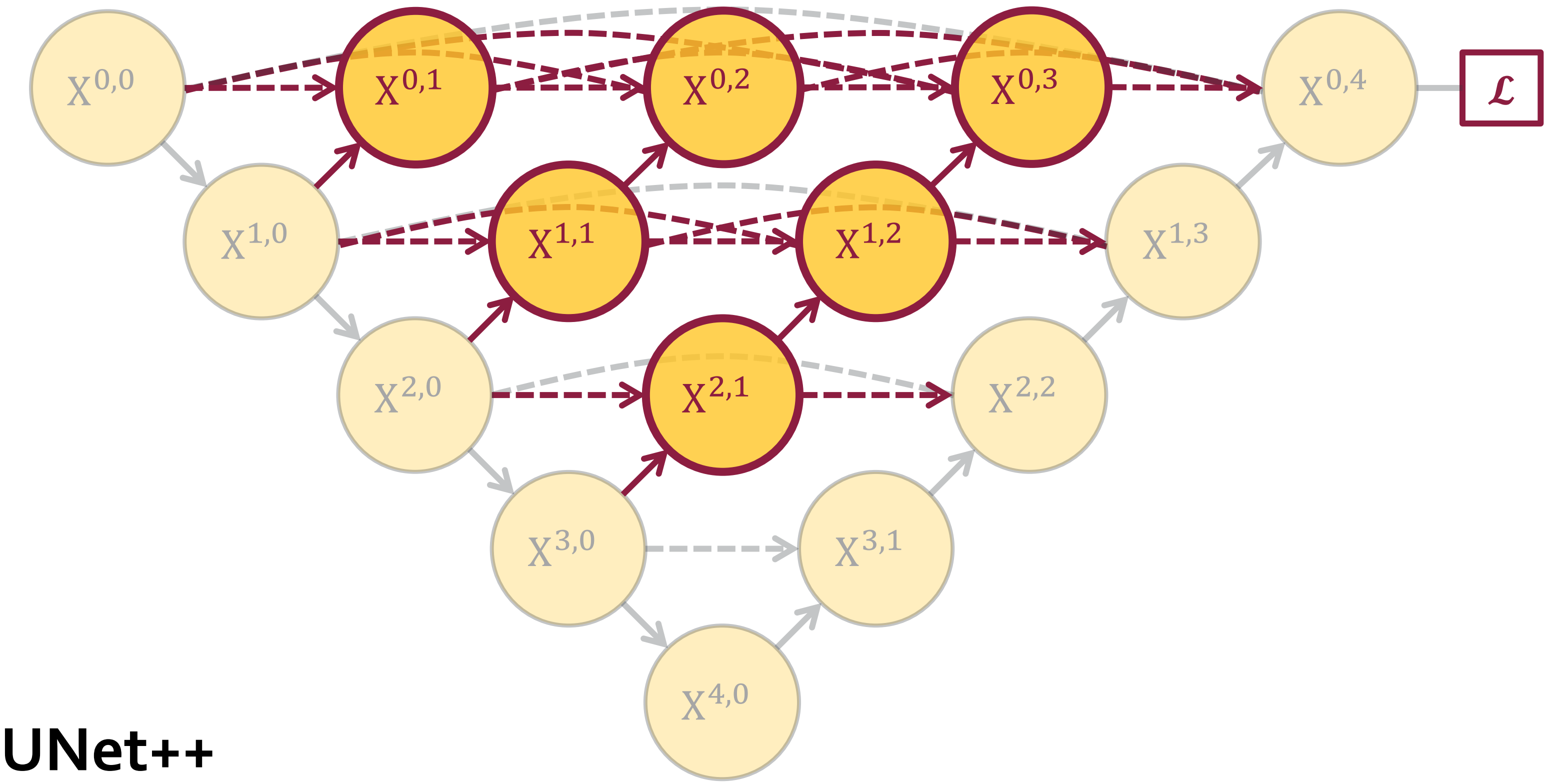


U-Net

1. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Redesigned skip connections aggregate multi-scale features

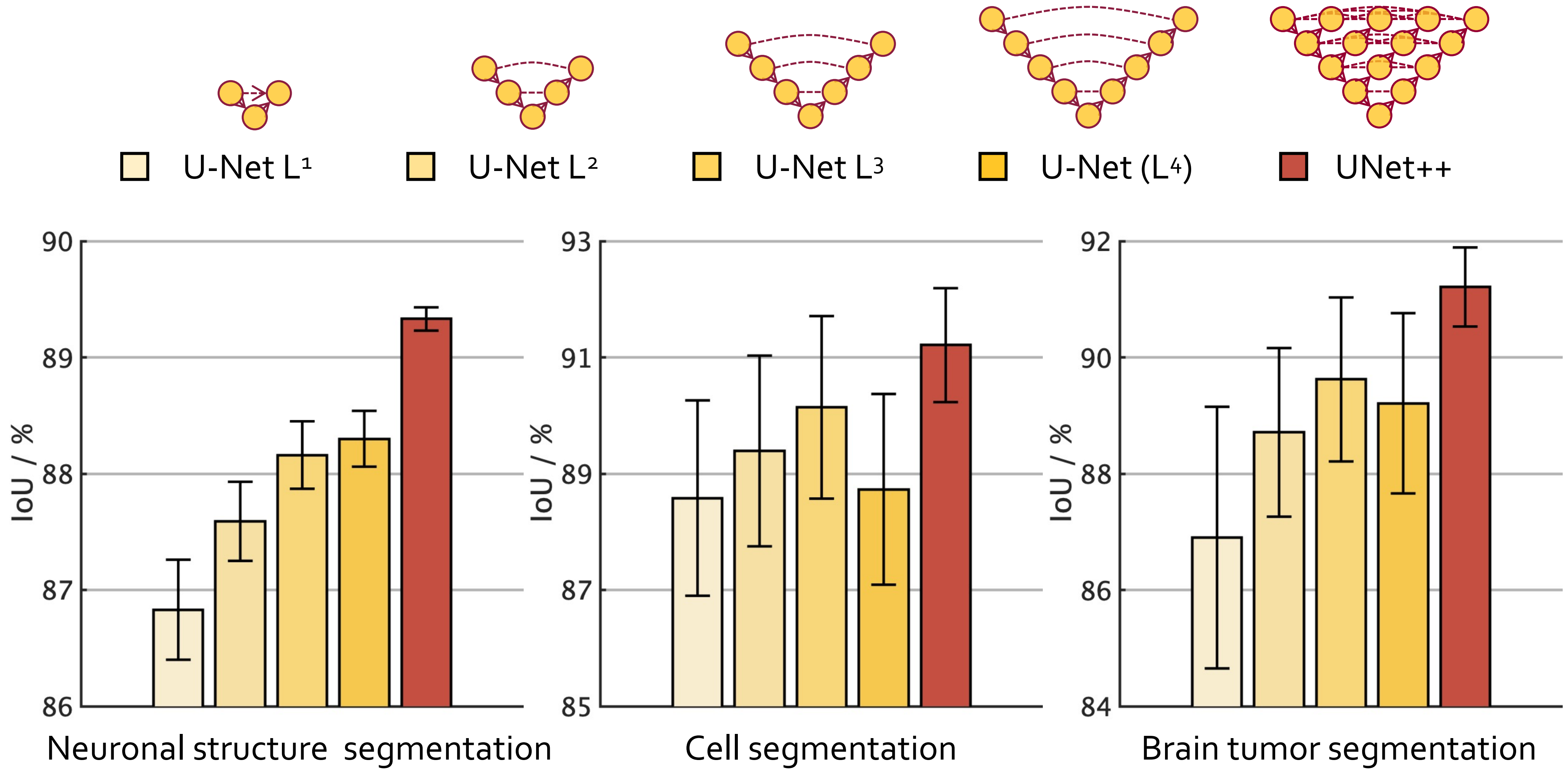


UNet++

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11.

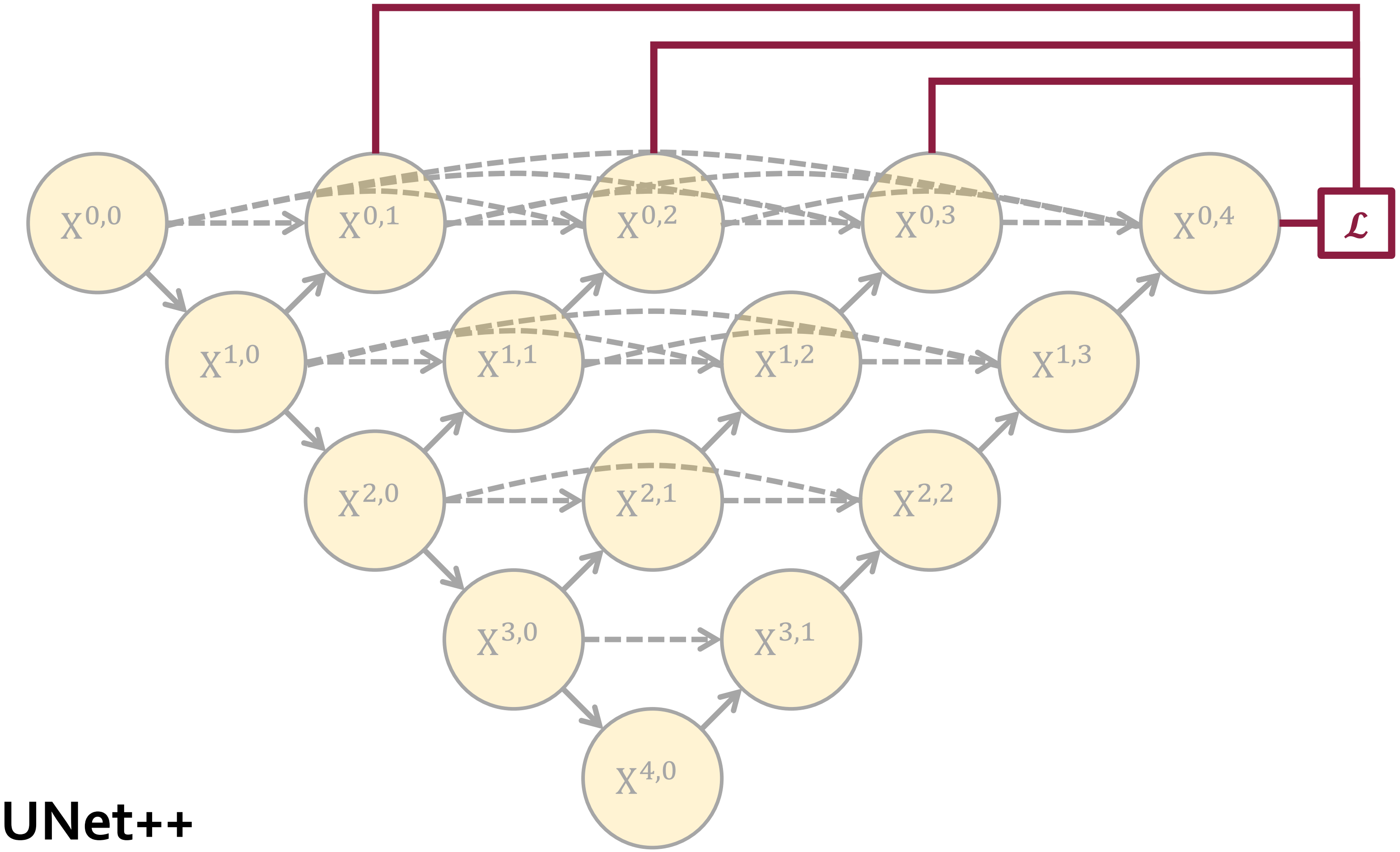
Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Redesigned skip connections aggregate multi-scale features



Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Deep supervision stabilizes model training and enables model pruning



UNet++

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11.

Introduction

Objective

Aim 1

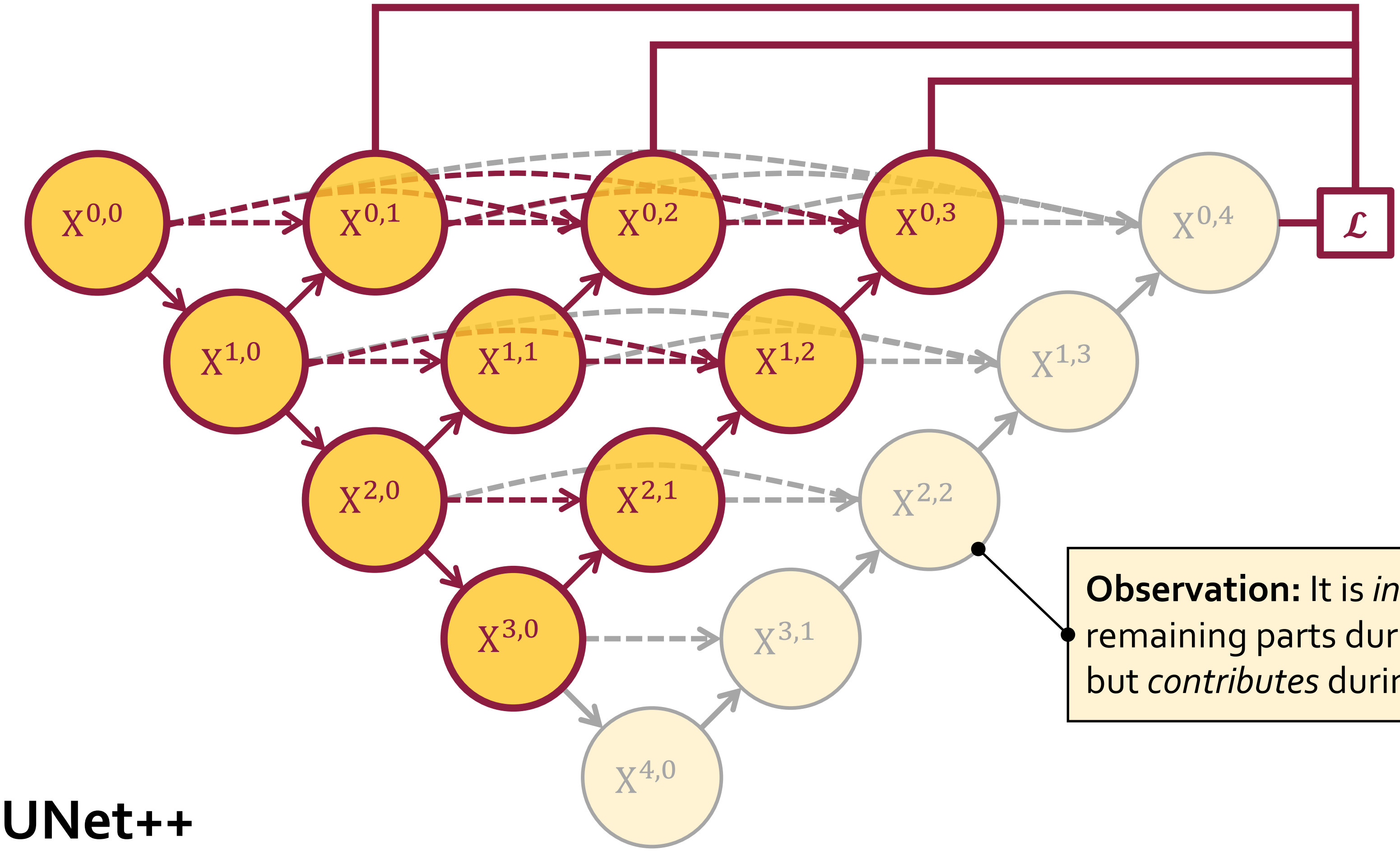
Aim 2

Aim 3

Summary

Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Deep supervision stabilizes model training and enables model pruning



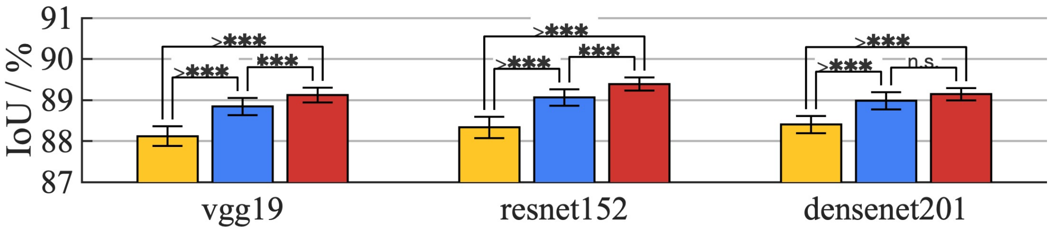
Observation: It is *independent* with the remaining parts during testing phase, but *contributes* during training phase.

UNet++

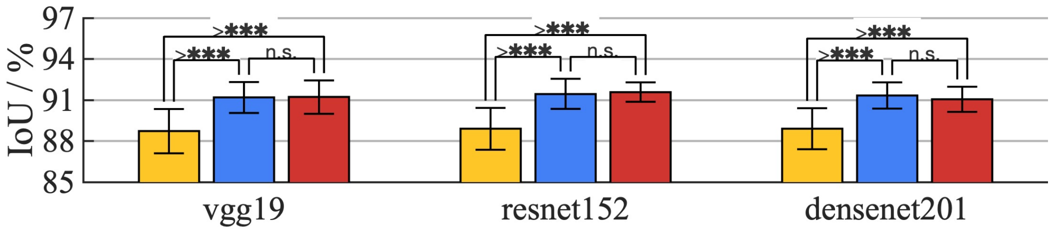
1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11.

Aim 2: Utilizing existing annotation effectively from advanced architecture

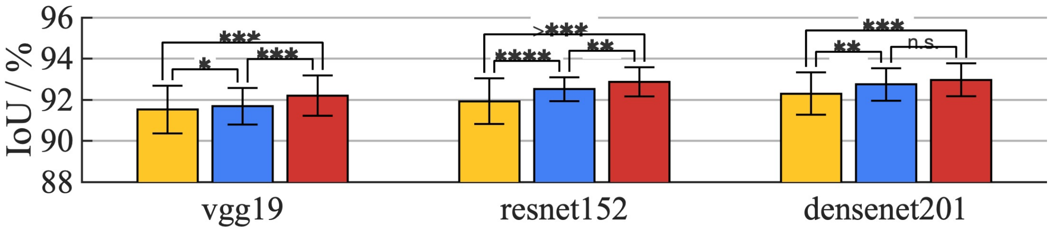
Contribution: UNet++ significantly improves disease/organ segmentation



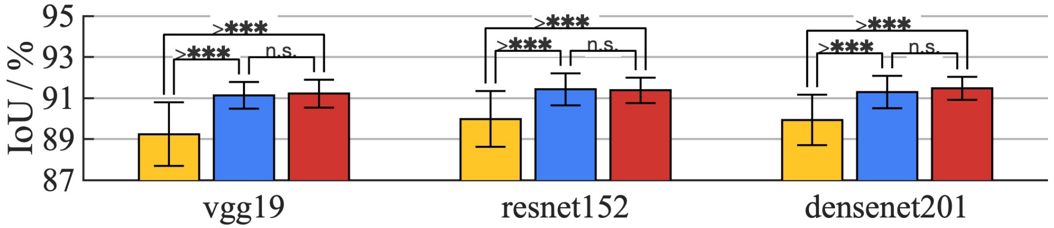
Neuronal structure segmentation



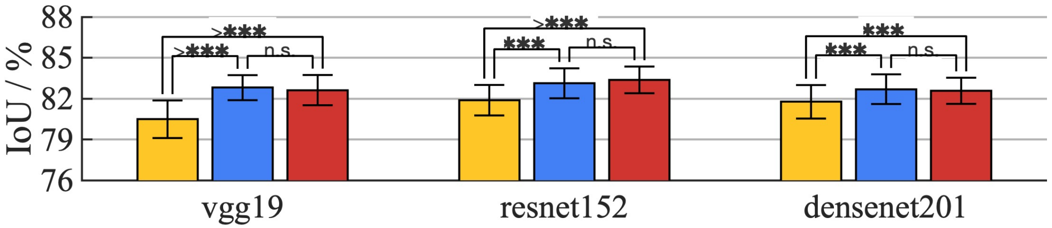
Cell segmentation



Nuclei segmentation



Brain tumor segmentation

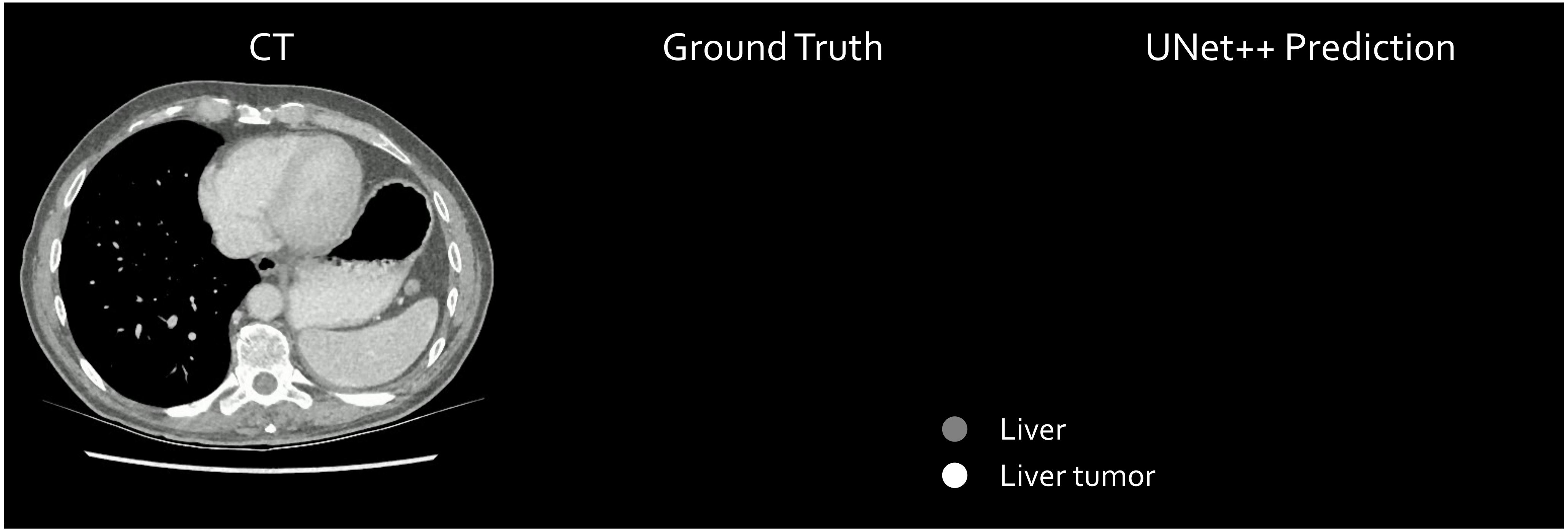


Liver segmentation

1. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE transactions on medical imaging 39.6 (2019): 1856-1867.

Aim 2: Utilizing existing annotation effectively from advanced architecture

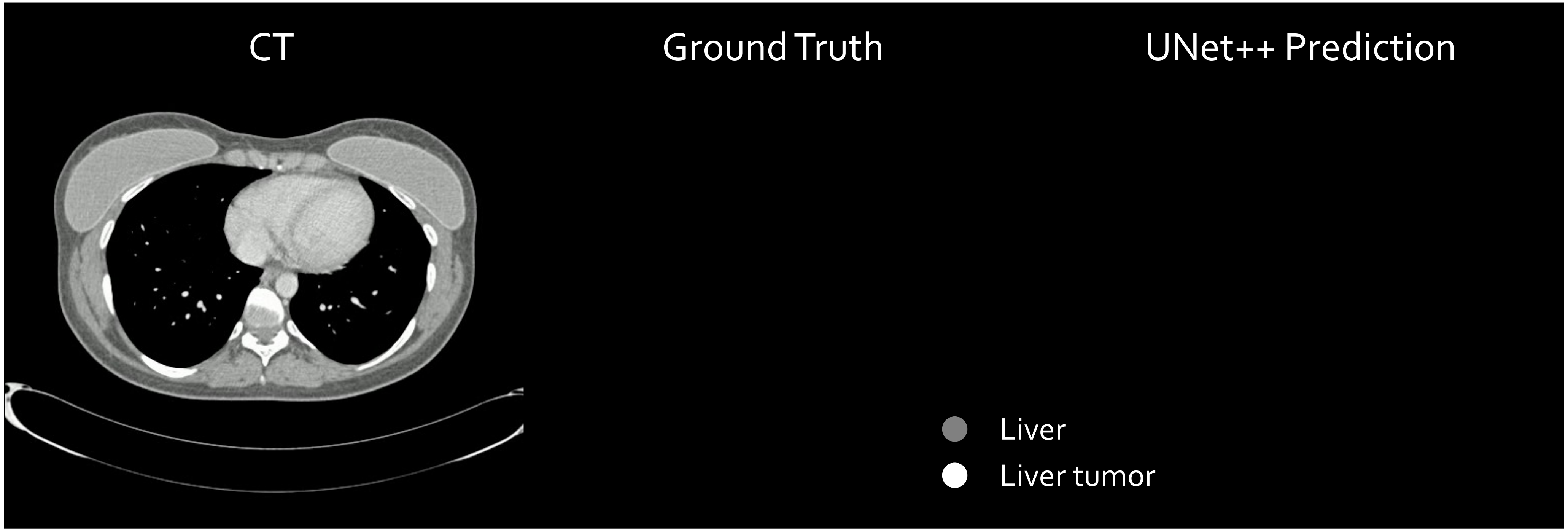
Contribution: UNet++ significantly improves disease/organ segmentation



1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE transactions on medical imaging 39.6 (2019): 1856-1867.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Contribution: UNet++ significantly improves disease/organ segmentation



1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer, Cham, 2018. 3-11.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." IEEE transactions on medical imaging 39.6 (2019): 1856-1867.

Intertwine the visual representation

Publications for Aim 2:

- Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, 2019. UNet++: Redesigning Skip Connections to Exploit Multi-Resolution Features in Image Segmentation. *IEEE Transactions on Medical Imaging, ranked among the most popular articles in IEEE TMI.*
- Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, 2018. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *DLMIA'18.*

Intertwine the visual representation

Clinical Impacts of Aim 2:

- Image segmentation can help compute clinically more accurate and desirable *imaging bio-markers* or *precision measurement*.
- Model pruning has the potential to exert important impact on deploying CAD systems to *mobile devices* and *ordinary desktop/laptop PCs* in clinical practice.

43.90% → 58.10% (U-Net → UNet++)
Covid-19 segmentation (CT)
[Fan et al., IEEE TMI]

78.56% → 82.90% (U-Net →UNet++)
Fiber tracing (corneal confocal microscopy)
[Mou et al., MICCAI]

86.48% → 89.53% (U-Net → UNet++)
Spleen segmentation (MRI)
[Li et al., Computers & Graphics]

Intertwine the visual representation

Research Impacts of Aim 2: *https://github.com/MrGiovanni/UNetPlusPlus*

- We have made UNet++ open science to stimulate collaborations among the research community and to help translate the technology to clinical practice.

86.59% → 87.22% (U-Net → UNet++)
SegTHOR 2019 Challenge (CT)
[Zhang et al., IEEE TMI]

90.16% → 91.98% (U-Net → UNet++)
Optic Disc & Cup Segmentation (fundus image)
[Meng et al., MICCAI]

60.34% → 71.60% (U-Net → UNet++)
Ground-glass opacity segmentation (CT)
[Zheng et al., IEEE Access]

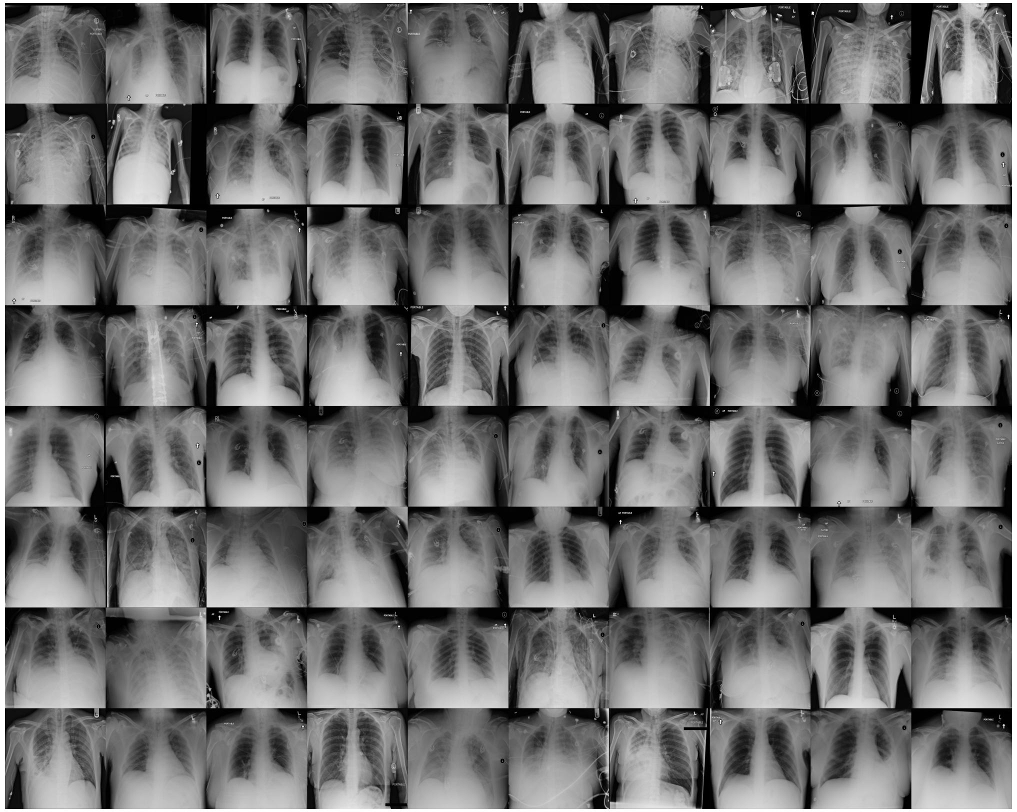
51.20% → 58.60% (U-Net → UNet++)
Esophagus segmentation (CT)
[Huang et al., IEEE Access]

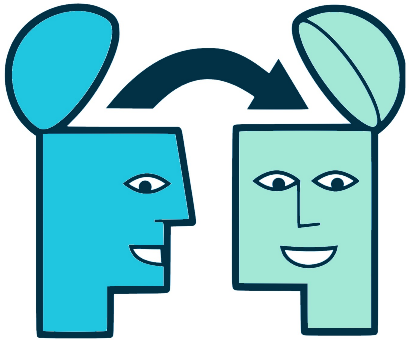
63.72% → 66.25% (U-Net → UNet++)
Liver tumor segmentation (CT)
[Bajpai et al., Master Thesis]

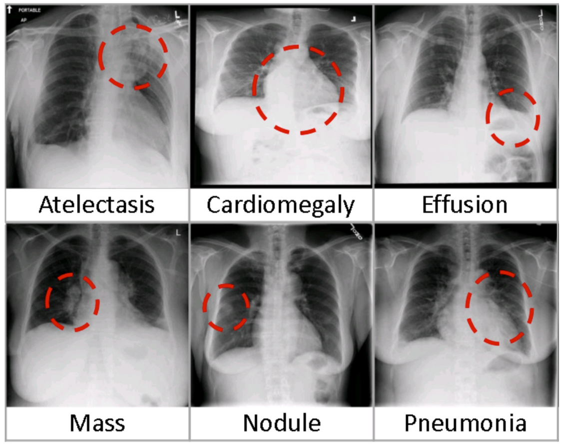
90.70% → 91.56% (U-Net → UNet++)
Heart segmentation (MRI)
[Ji et al., MICCAI]

Aim 3: Extracting generic knowledge directly from unannotated images

Task: Utilize 1,000,000 images without systematic annotation

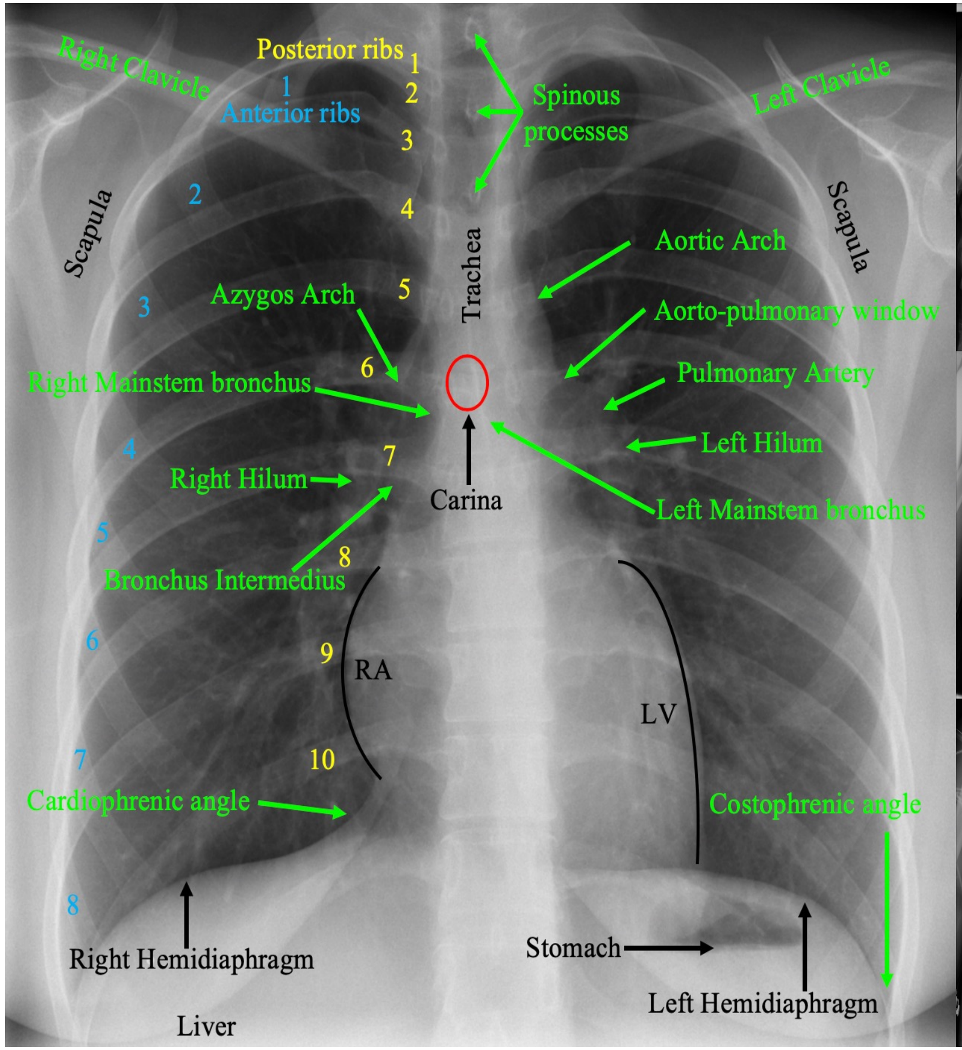


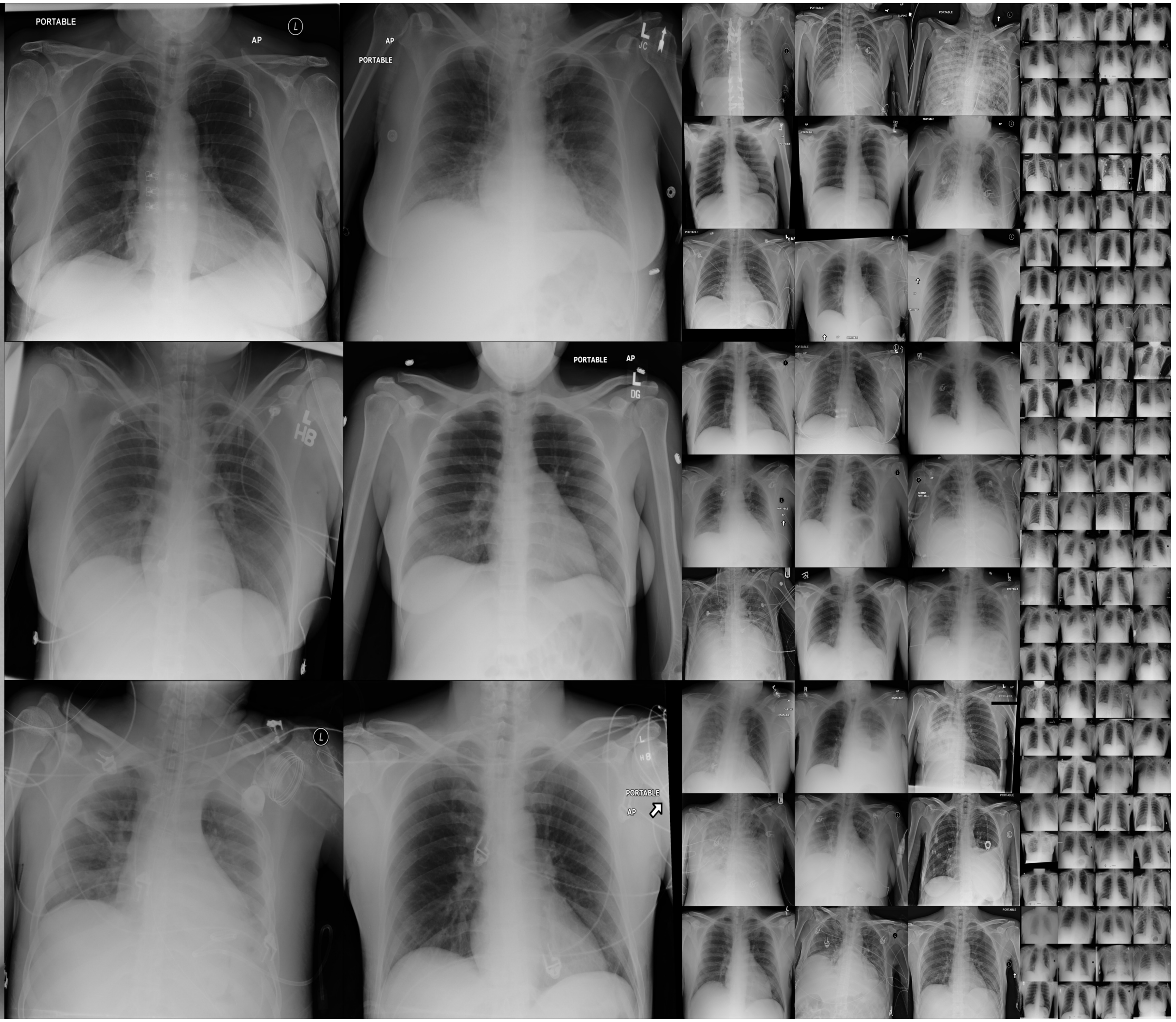




Aim 3: Extracting generic knowledge directly from unannotated images

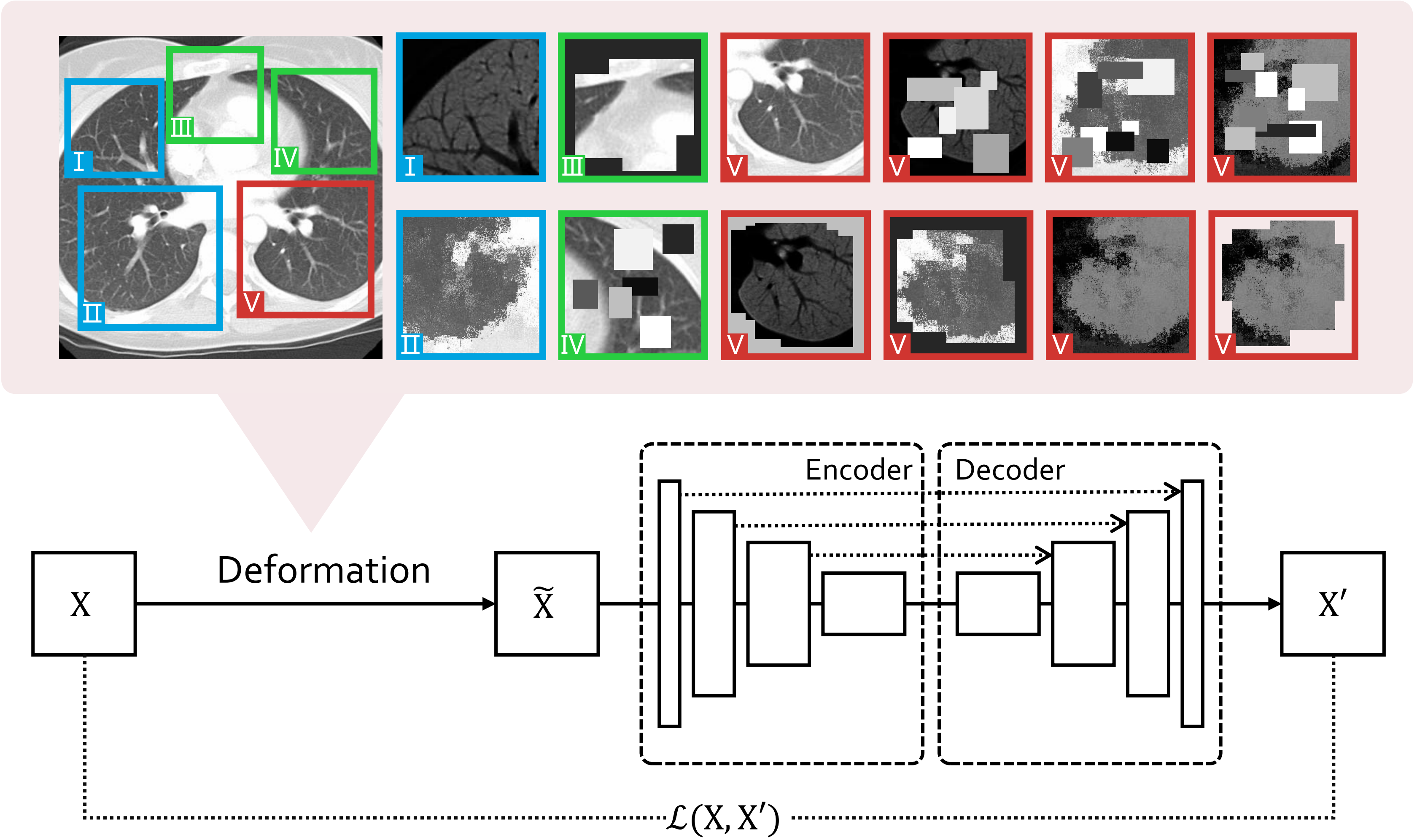
Hypothesis: Generic models can be built upon consistent, recurrent anatomy





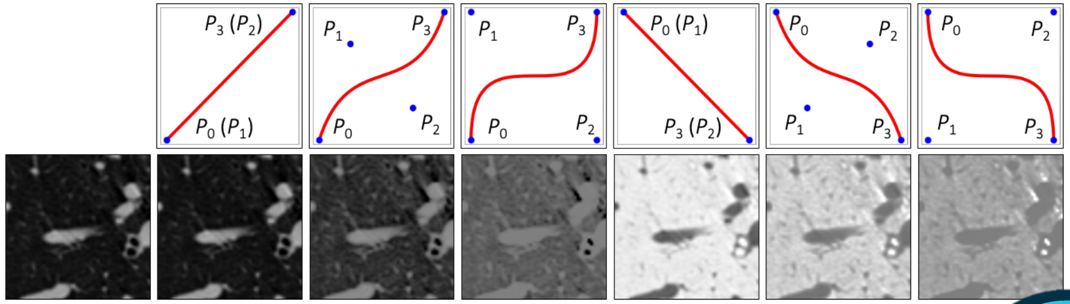
Aim 3: Extracting generic knowledge directly from unannotated images

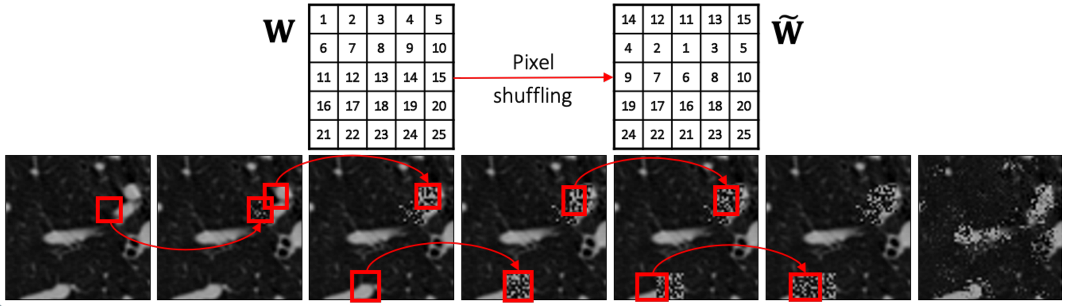
Approach: Image restoration task helps model learn image representation



Aim 3: Extracting generic knowledge directly from unannotated images

Approach: Learning from multiple perspectives leads to robust models



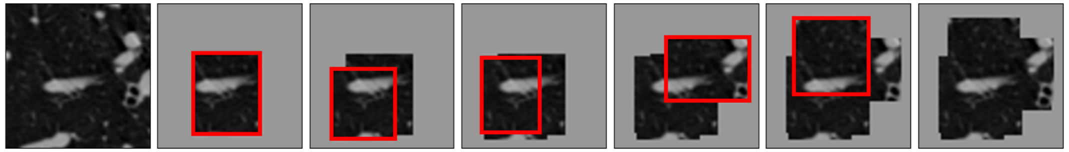


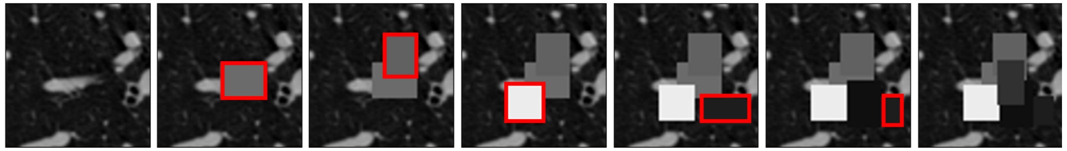
1. Learning organ appearance via **non-linear transformation**

2. Learning organ texture and local boundaries via **local pixel shuffling**

3. Learning organ spatial layout and global geometry via **outer-cutout**

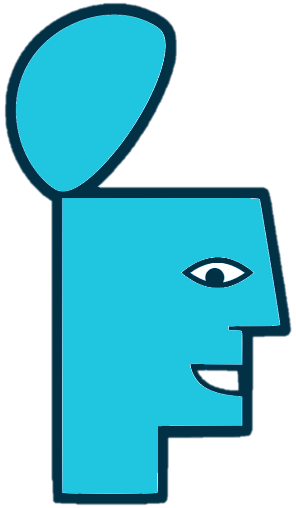
4. Learning local continuities of organs via **inner-cutout**



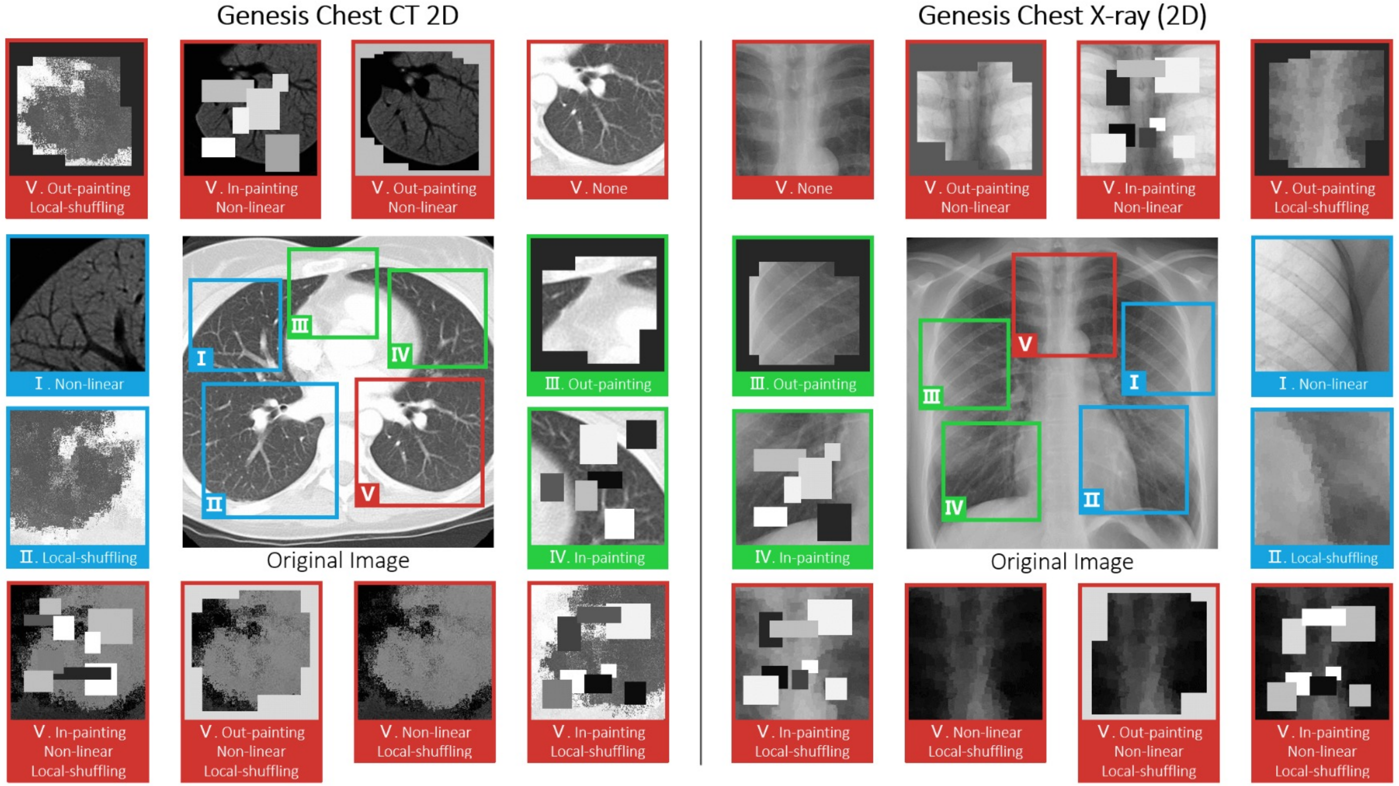


Aim 3: Extracting generic knowledge directly from unannotated images

Contribution: Build generic pre-trained 3D models, named "Models Genesis"



Models Genesis



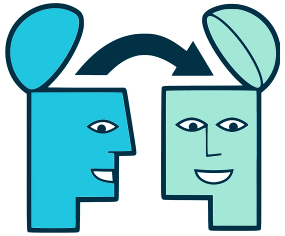
1. Zhou, Zongwei, *et al.* "Models genesis: Generic autodidactic models for 3d medical image analysis." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019.
2. Zhou, Zongwei, *et al.* "Models genesis." Medical image analysis 67 (2021): 101840.

Aim 3: Extracting generic knowledge directly from unannotated images

Contribution: Models Genesis exceed publicly available pre-trained 3D models

Pre-training	Approach	Target tasks				
		NCC[1] (%)	NCS[2] (%)	ECC[3] (%)	LCS[4] (%)	BMS[5] (%)
No	Random with Uniform Init	94.74±1.97	75.48±0.43	80.36±3.58	78.68±4.23	60.79±1.60
	Random with Xavier Init (Glorot and Bengio, 2010)	94.25±5.07	74.05±1.97	79.99±8.06	77.82±3.87	58.52±2.61
	Random with MSRA Init (He et al., 2015)	96.03±1.82	76.44±0.45	78.24±3.60	79.76±5.43	63.00±1.73
(Fully) supervised	I3D (Carreira and Zisserman, 2017)	98.26±0.27	71.58±0.55	80.55±1.11	70.65±4.26	67.83±0.75
	NiftyNet (Gibson et al., 2018b)	94.14±4.57	52.98±2.05	77.33±8.05	83.23±1.05	60.78±1.60
	MedicalNet (Chen et al., 2019b)	95.80±0.49	75.68±0.32	86.43±1.44	**85.52±0.58**†	66.09±1.35
Self-supervised	De-noising (Vincent et al., 2010)	95.92±1.83	73.99±0.62	85.14±3.02	84.36±0.96	57.83±1.57
	In-painting (Pathak et al., 2016)	91.46±2.97	76.02±0.55	79.79±3.55	81.36±4.83	61.38±3.84
	Jigsaw (Noroozi and Favaro, 2016)	95.47±1.24	70.90±1.55	81.79±1.04	82.04±1.26	63.33±1.11
	DeepCluster (Caron et al., 2018)	97.22±0.55	74.95±0.46	84.82±0.62	82.66±1.00	65.96±0.85
	Patch shuffling (Chen et al., 2019a)	91.93±2.32	75.74±0.51	82.15±3.30	82.82±2.35	52.95±6.92
	Rubiks Cube (Zhuang et al., 2019)	96.24±1.27	72.87±0.16	80.49±4.64	75.59±0.20	62.75±1.93
	Genesis Chest CT (ours)	**98.34±0.44**	**77.62±0.64**	**87.20±2.87**	85.10±2.15	**67.96±1.29**

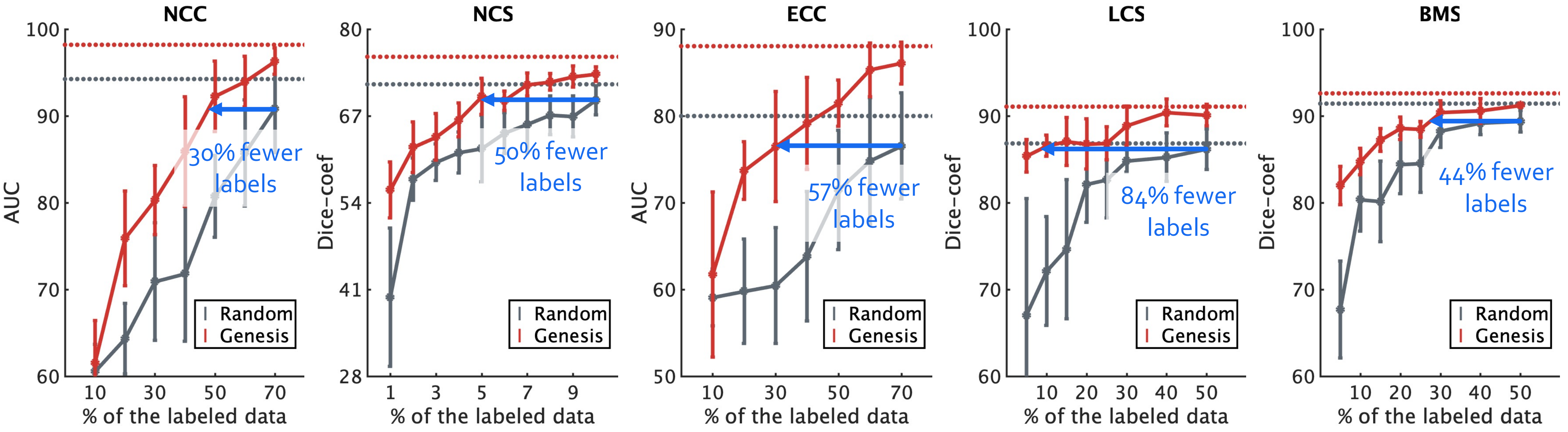
[1]NCC Lung nodule false positive reduction in CT images
[2]NCS Lung nodule segmentation in CT images
[3]ECC Pulmonary embolism false positive reduction in CT images
[4]LCS Liver segmentation in CT images
[5]BMS Brain tumor segmentation in MR images

Genesis Chest CT

Target models

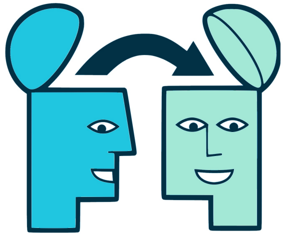
1. Zhou, Zongwei, *et al.* "Models genesis: Generic autodidactic models for 3d medical image analysis." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019.
2. Zhou, Zongwei, *et al.* "Models genesis." Medical image analysis 67 (2021): 101840.

Aim 3: Extracting generic knowledge directly from unannotated images

Contribution: Models Genesis reduce annotation efforts by at least 30%



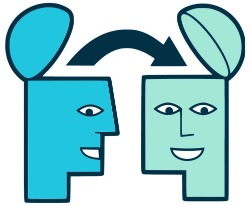
[1]NCC Lung nodule false positive reduction in CT images
[2]NCS Lung nodule segmentation in CT images
[3]ECC Pulmonary embolism false positive reduction in CT images
[4]LCS Liver segmentation in CT images
[5]BMS Brain tumor segmentation in MR images



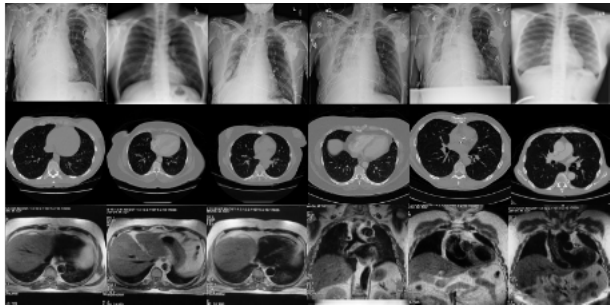
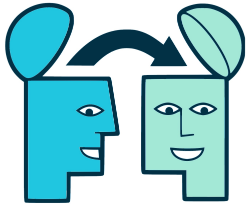
1. Zhou, Zongwei, *et al.* "Models genesis: Generic autodidactic models for 3d medical image analysis." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019.
2. Zhou, Zongwei, *et al.* "Models genesis." Medical image analysis 67 (2021): 101840.

Aim 3: Extracting generic knowledge directly from unannotated images

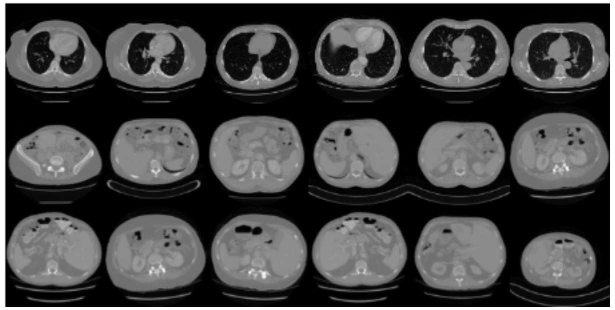
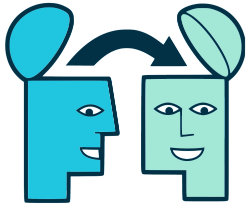
Discussion: Extend to modality-oriented and organ-oriented models

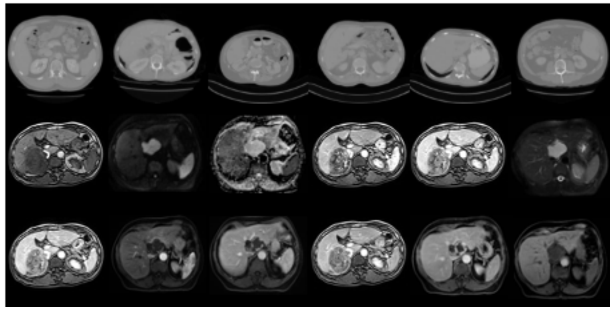
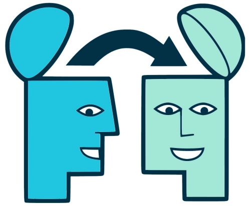
Genesis X-ray

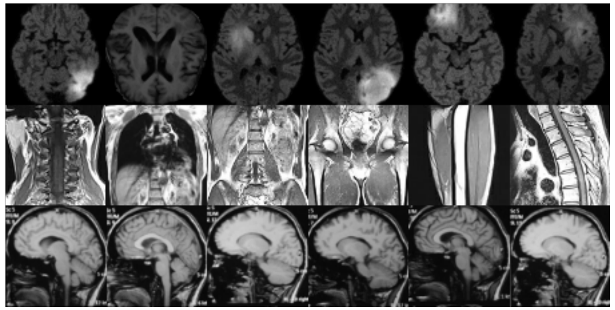
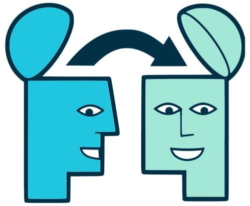
Genesis Lung

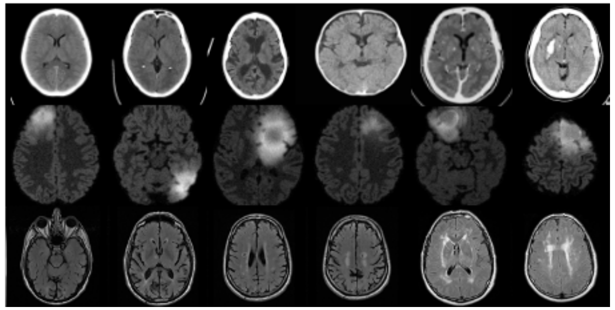
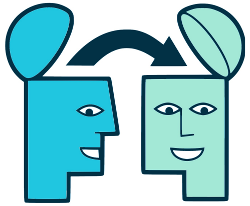
Genesis CT

Genesis Liver

Genesis MRI

Genesis Brain

Effective image features across diseases, organs, and modalities.

Publications for Aim 3:

- Z. Zhou, V. Sodha, M. M. Rahman Siddiquee, R. Feng, N. Tajbakhsh, M. Gotway, J. Liang, 2019. Models Genesis: Generic Autodidactic Models for 3D Medical Image Analysis. *MICCAI'19, Young Scientist Award.*
- Z. Zhou, V. Sodha, J. Pang, M. Gotway, J. Liang, 2020. Models Genesis. *Medical Image Analysis, MedIA Best Paper Award.*

Effective image features across diseases, organs, and modalities.

Clinical Impacts of Aim 3:

- Instead of building a model from scratch (demanding numerous data and label acquisition), a *smaller dataset* can be used to efficiently fine-tune the existing model.
- Generic pre-trained models can serve as a *primary source of transfer learning* for many medical imaging applications, leading to accelerated training and improved performance.

68.98% → 73.85% (Scratch → MG) Prostate segmentation (MRI) [Taleb et al., arXiv:1912.05396, 2019]	**83.14% → 88.30% (Scratch → MG)** Lymph node classification (histology) [Xu et al., BIBM, 2020]	**72.30% → 85.81% (Scratch → MG)** Brain hemorrhage classification (CT) [Zhu et al., arXiv:2012.07477, 2020]

Effective image features across diseases, organs, and modalities.

Research Impacts of Aim 3: *https://github.com/MrGiovanni/ModelsGenesis*

- We have made Models Genesis open science to stimulate collaborations among the research community and to help translate the technology to clinical practice.

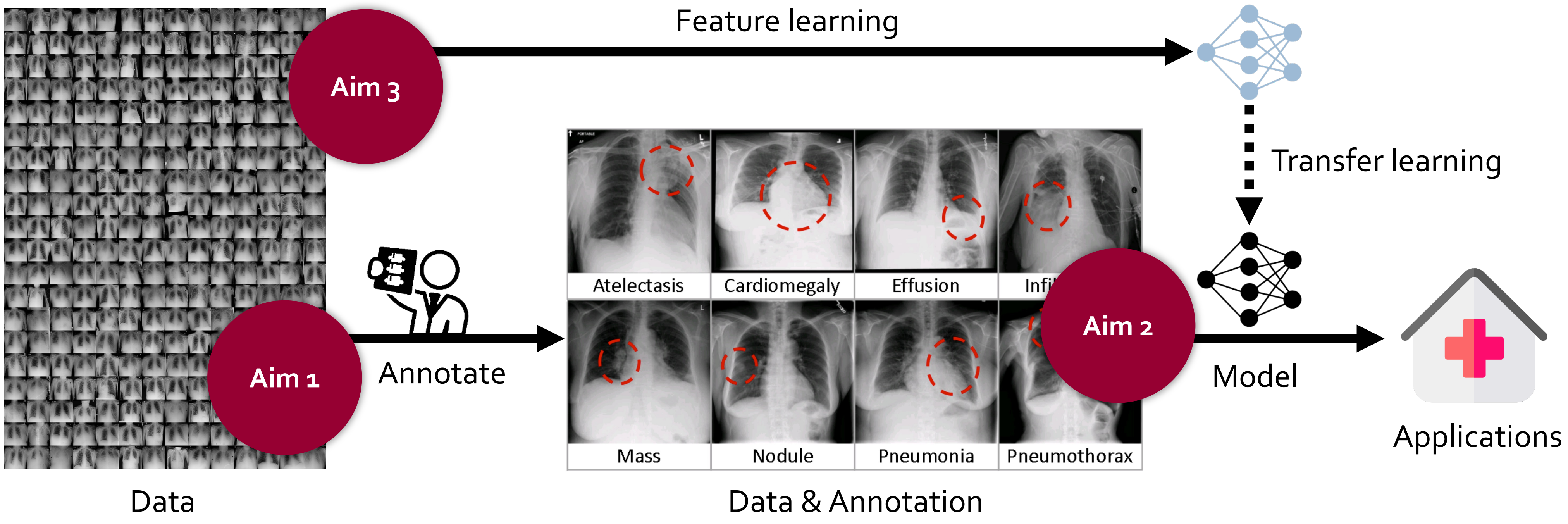
67.04% → 74.53% (Scratch → MG) Blood cavity segmentation (MRI) [Zhang et al., arXiv:2010.06107, 2020]	**67.84% → 69.27% (Scratch → MG)** 13 organ segmentation (CT) [Xie et al., arXiv:2011.12640, 2020]	**89.98% → 95.01% (Scratch → MG)** Liver segmentation (CT&MRI) [Taleb et al., arXiv:1912.05396, 2019]
77.50% → 92.50% (Scratch → MG) COVID-19 classification (CT) [Sun et al., arXiv:2012.06457, 2020]	**75.97% → 77.50% (Scratch → MG)** Liver tumor segmentation (CT) [Bajpai et al., Master Thesis, 2021]	**74.00% → 79.33% (Scratch → MG)** Alzheimer's disease classification (MRI) [Zhang et al., arXiv:2010.06107, 2020]

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images



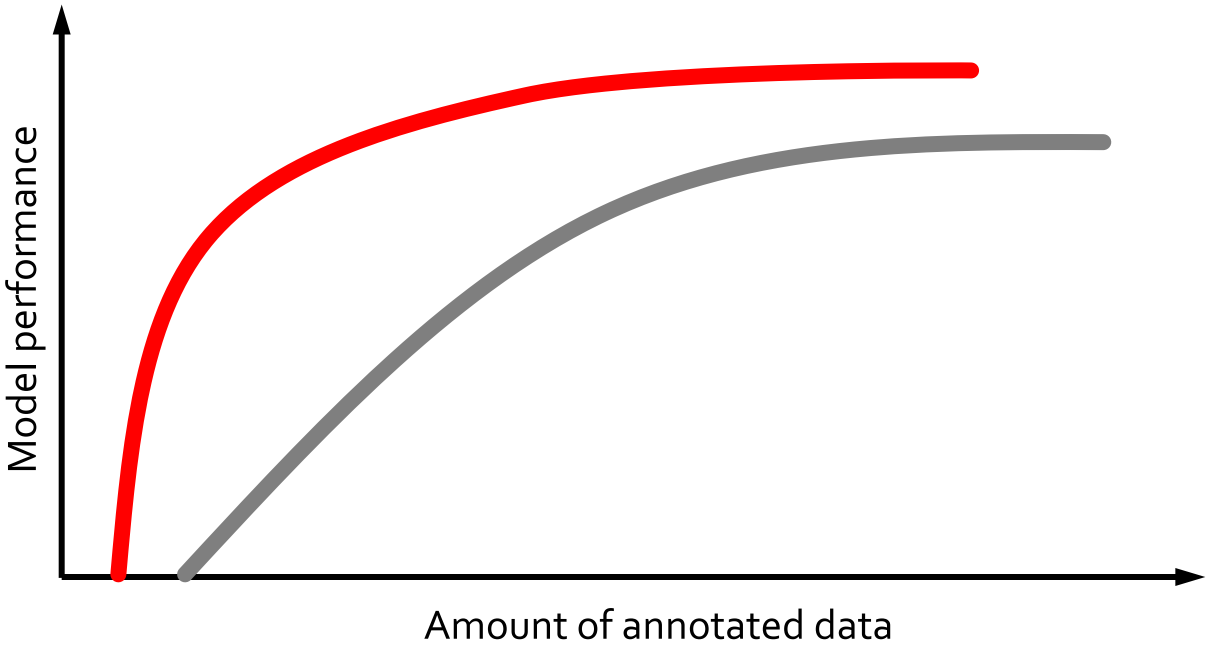
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Hypothesis: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset.



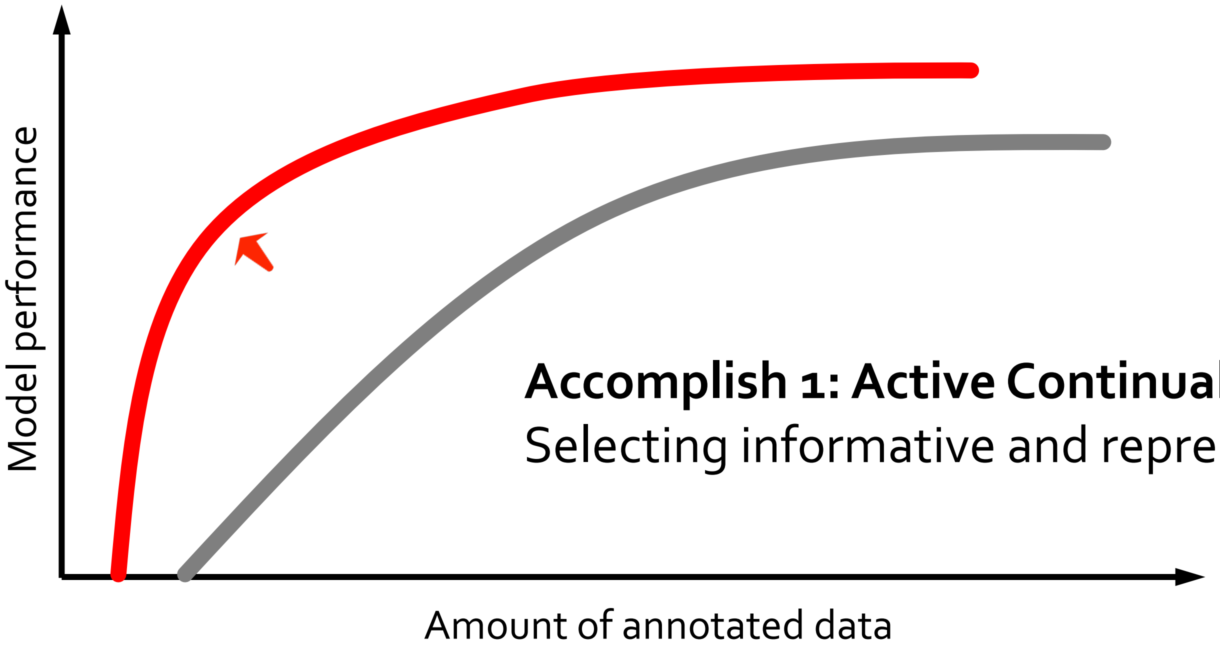
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**



Accomplish 1: Active Continual Fine-Tuning
Selecting informative and representative samples

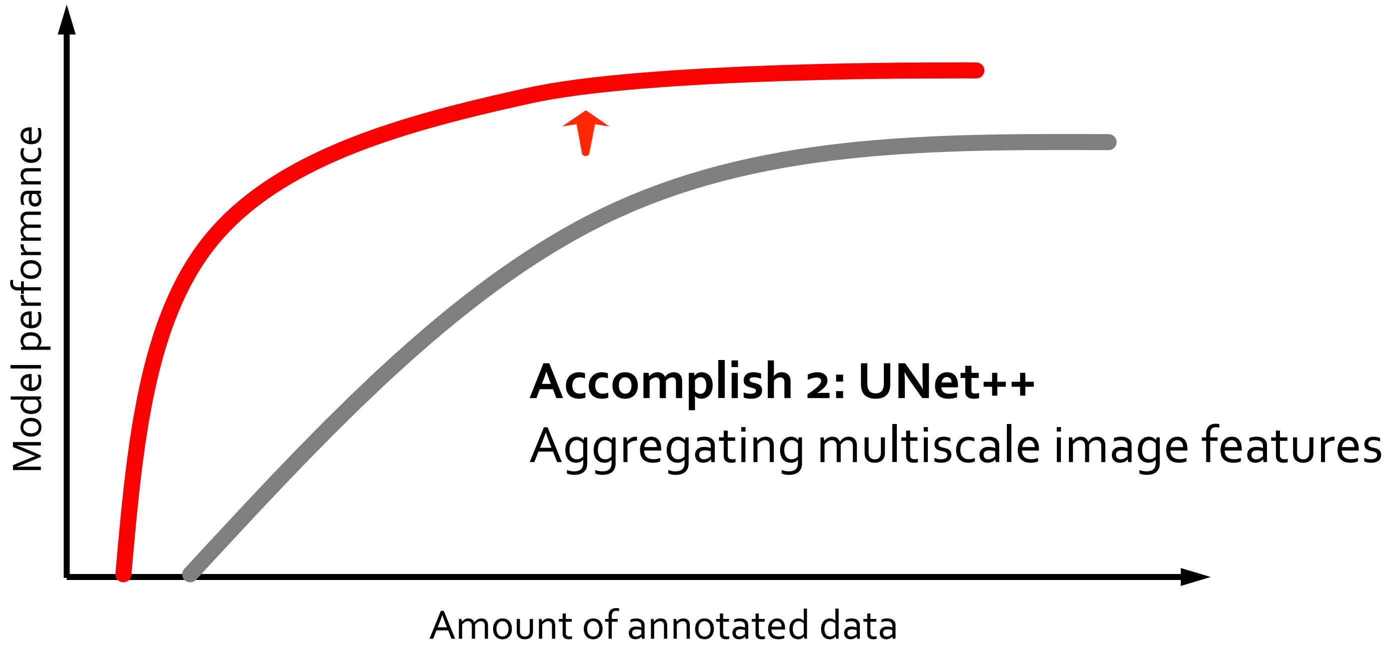
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**



Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

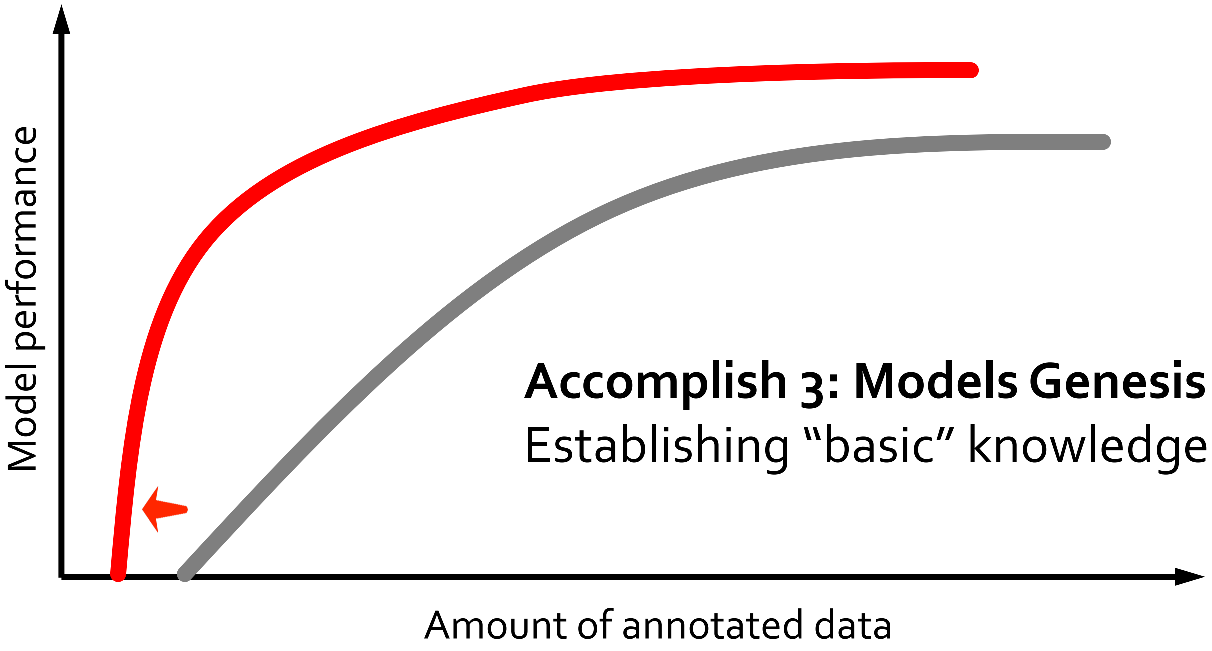
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**



Accomplish 3: Models Genesis

Establishing "basic" knowledge in medical images

Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

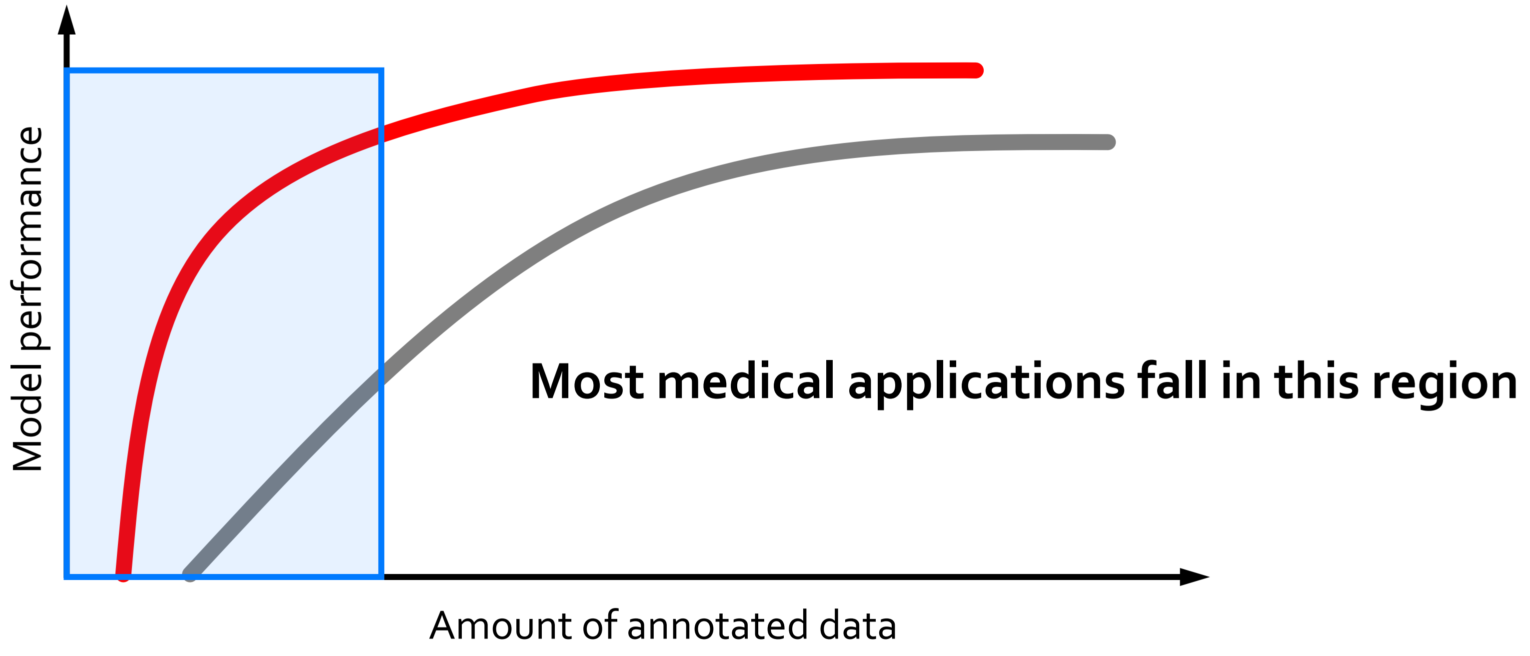
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**



Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**

Annotation-efficiency: Applications of pulmonary embolism detection (*rank #3*) and liver tumor segmentation (*rank #1*, Shivam 2021)

Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

Aim 3: Extracting generic knowledge directly from unannotated images

Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**

Annotation-efficiency: Applications of pulmonary embolism detection (*rank #3*) and liver tumor segmentation (*rank #1*, Shivam 2021)

Interpreting medical images: A book chapter overviewing AI in medical image interpretation

THANK YOU

- Jianming Liang, Ph.D.
- Edward H. Shortliffe, M.D., Ph.D.
- Robert A. Greenes, Ph.D.
- Baoxin Li, Ph.D.
- Michael B. Gotway, M.D.
- Murthy Devarakonda, Ph.D.

Funding for research program supported by

- NIH R01 (R01HL128785)
- ASU-Mayo Grant
- Mayo Innovation Grant

Towards Annotation-Efficient Deep Learning for Computer-Aided Diagnosis

Zongwei Zhou

Ph.D. Candidate, Biomedical Informatics

College of Health Solutions, Arizona State University

P: 1-(480)738-2575 | E: zongweiz@asu.edu

People
Who
Don't
Give Up