This work was based on my Ph.D. dissertation at Arizona State University, supervised by Dr. Jianming Liang.
I thank N. Tajbakhsh, J. Shin, V. Sodha, M. M. Rahman Siddiquee, J. Pang, R. Feng, L. Zhang, S. Bajpai, F. Haghighi, M. R. Hosseinzadeh Taher, Z. Guo, P. Zhang for their helps in paper writing and experiments.

Towards Annotation-Efficient Deep Learning for Computer-Aided Diagnosis

Zongwei Zhou, PhD

Postdoc, Department of Computer Science
Johns Hopkins University, Baltimore, MD
P: 1-(480)738-2575 | E: zzhou82@jh.edu

Introduction

Objective

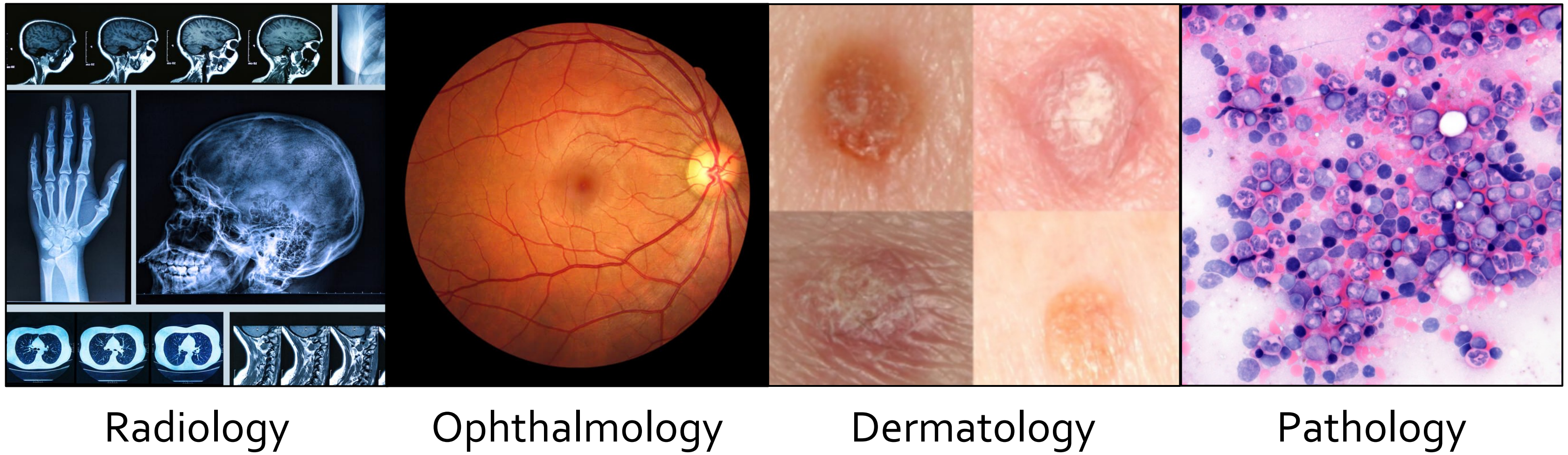
Aim 1

Aim 2

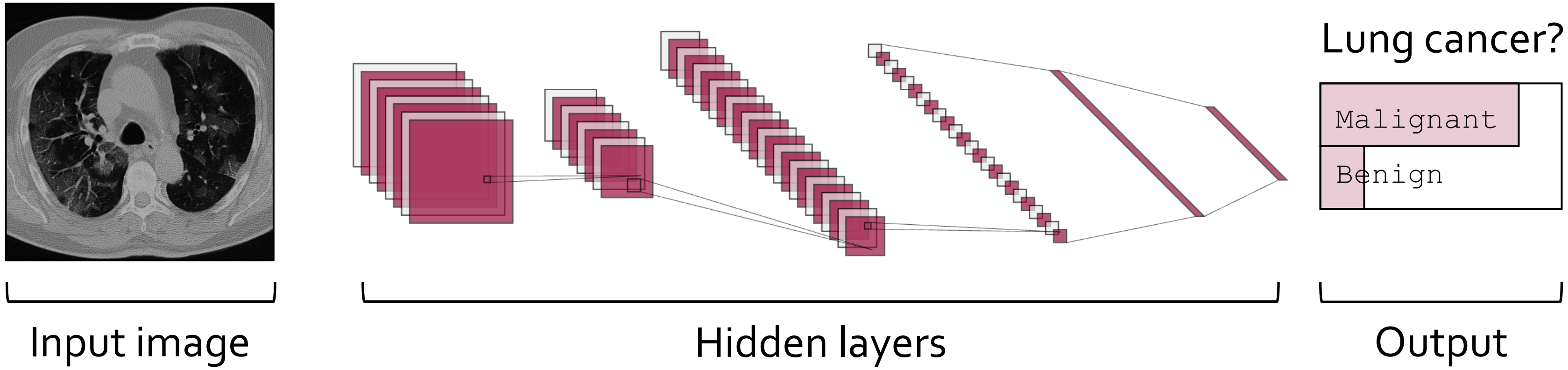
Aim 3

Summary

Imaging data account for about 90% of all healthcare data



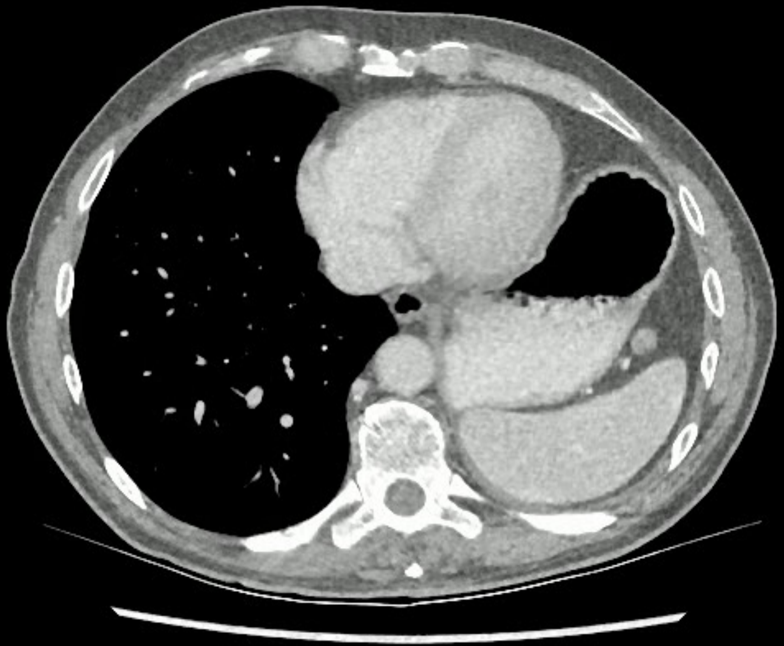
Deep Learning has ushered in a revolution in medical imaging



1. "The Digital Universe Driving Data Growth in Healthcare." published by EMC with research and analysis from IDC (12/13)
2. LeCun, Yann, et al. "Deep learning." Nature, 2015.

Radiologists hate annotation, but computer scientists love annotation.

CT



Ground Truth

annotated by human experts

Deep Learning

Liver

Liver tumor

- I thank S. Bajpai for this experiment; the code can be found at https://github.com/MrGiovanni/UNetPlusPlus/tree/master/pytorch

1. Bajpai, Shivam. "Pre-Trained Models for nnUNet." Master diss., Arizona State University, 2021.
2. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
3. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

Introduction

Objective

Aim 1

Aim 2

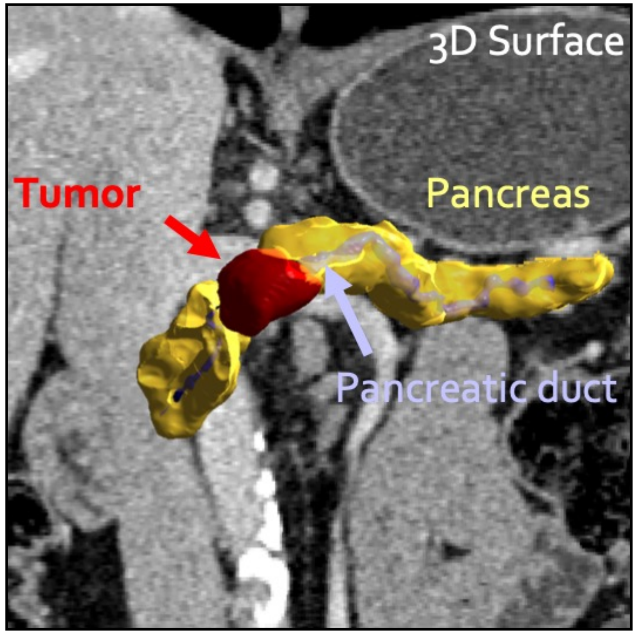
Aim 3

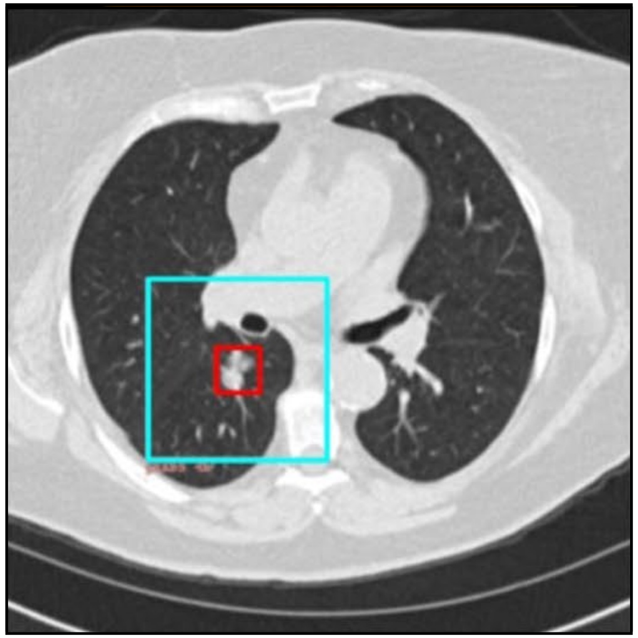
Summary

To match human diagnostic precision, deep learning requires enormous annotation cost.

- **1,511,400** radiologist-annotated CT images for pancreatic cancer detection (*15 years to create*)
- **42,290** radiologist-annotated CT images for lung cancer diagnosis
- **129,450** dermatologist-annotated images for skin cancer classification

The FELIX Project at JHU





1. Xia, Yingda, et al. "The FELIX project: Deep networks to detect pancreatic neoplasms." medRxiv, 2022.
2. 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.
3. Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." Nature, 2017.

To match human diagnostic precision, deep learning requires enormous annotation cost.

- **1,511,400** radiologist-annotated CT images for pancreatic cancer detection (*15 years to create*)
- **42,290** radiologist-annotated CT images for lung cancer diagnosis
- **129,450** dermatologist-annotated images for skin cancer classification

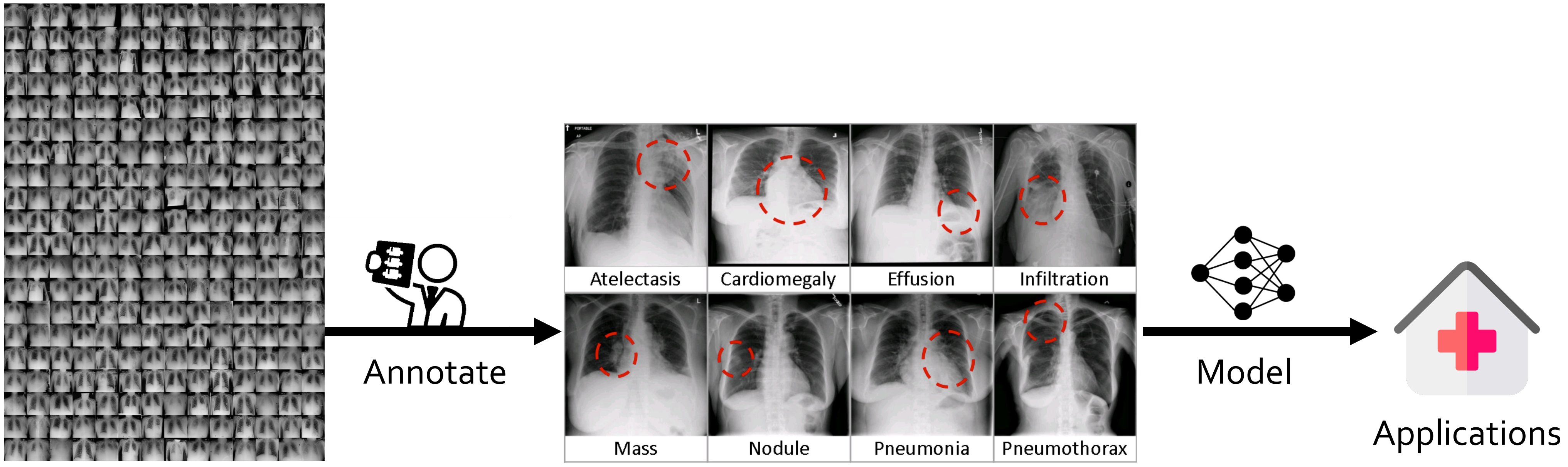
"How to develop annotation-efficient deep learning without such BIG annotated data?"

Significant, consider these scenarios:

- A flood of patients are waiting for imaging 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

Computer-*Aided* Diagnosis

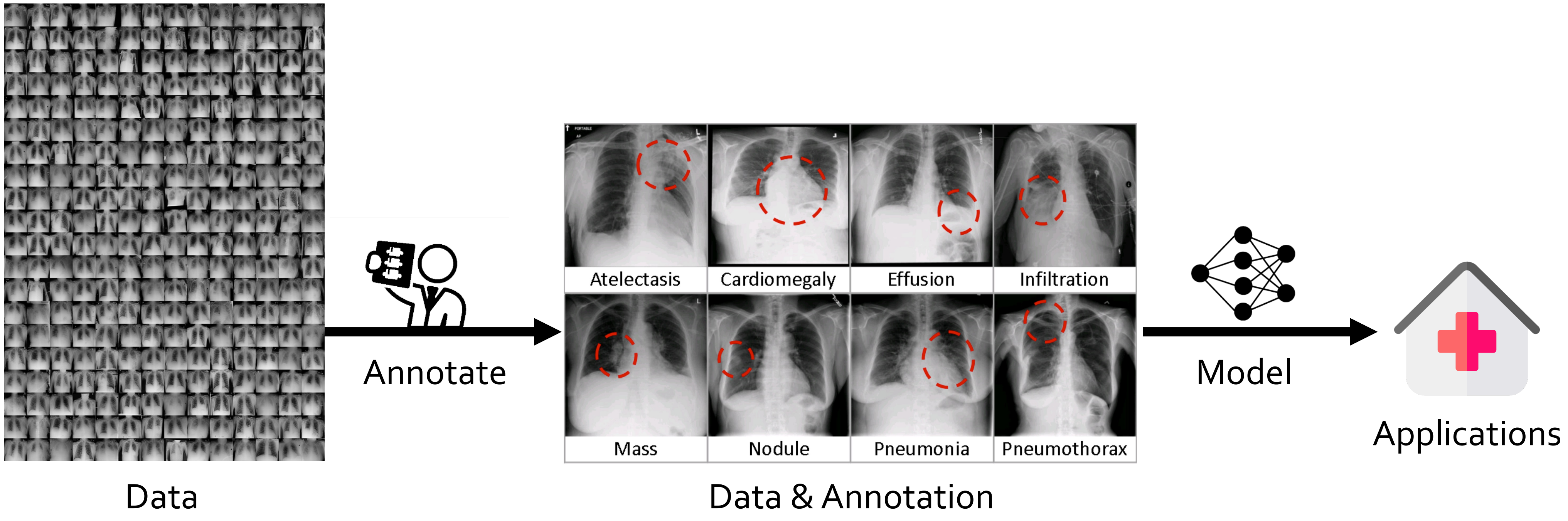
Assisting human experts to see more patients and
to deliver more accurate diagnosis (*beyond human eye*)



Data

Data & Annotation

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



Introduction

Objective

Aim 1

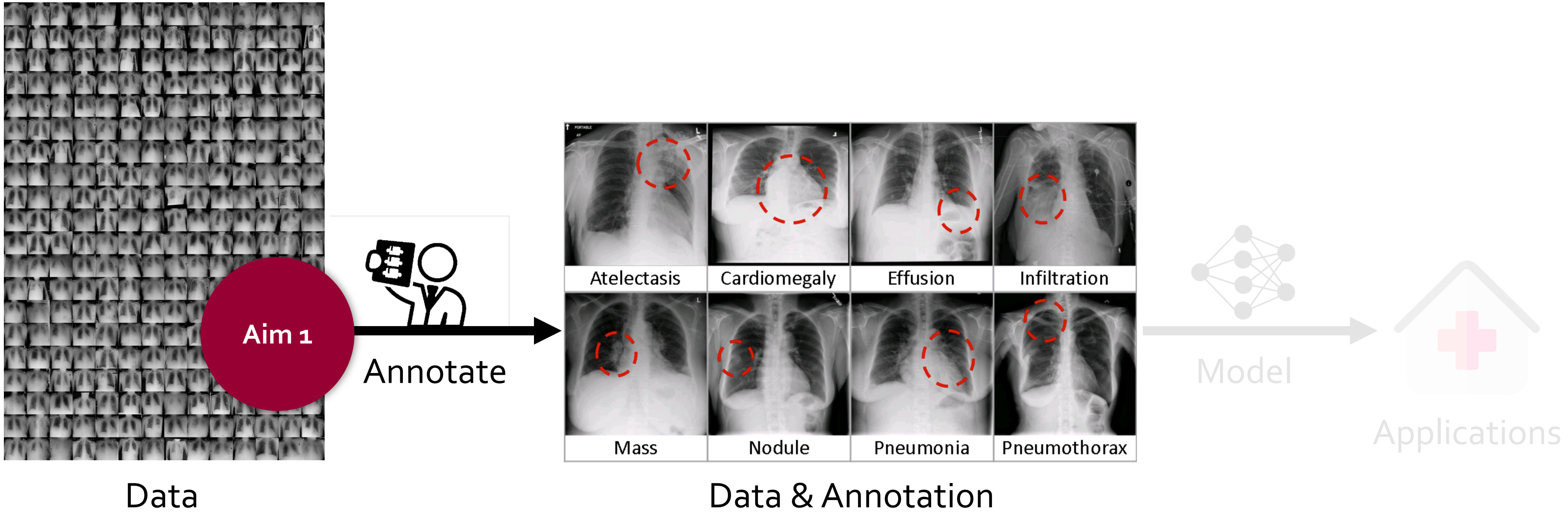
Aim 2

Aim 3

Summary

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

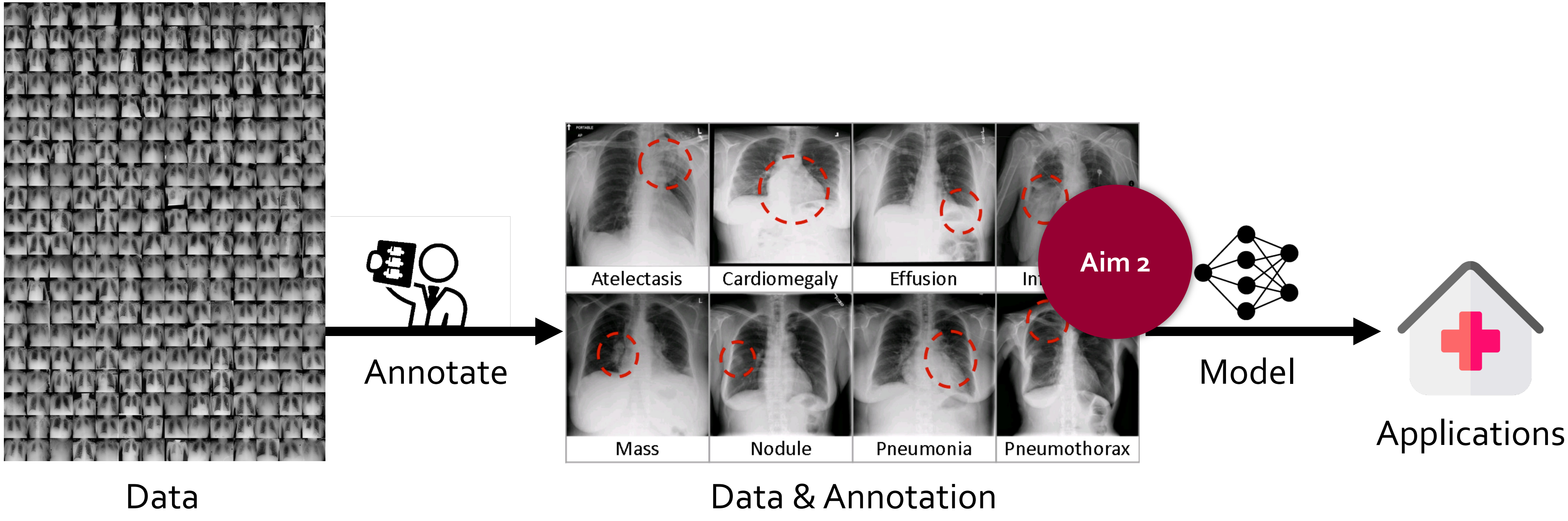
Aim 1: Acquiring necessary annotation efficiently from human experts



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

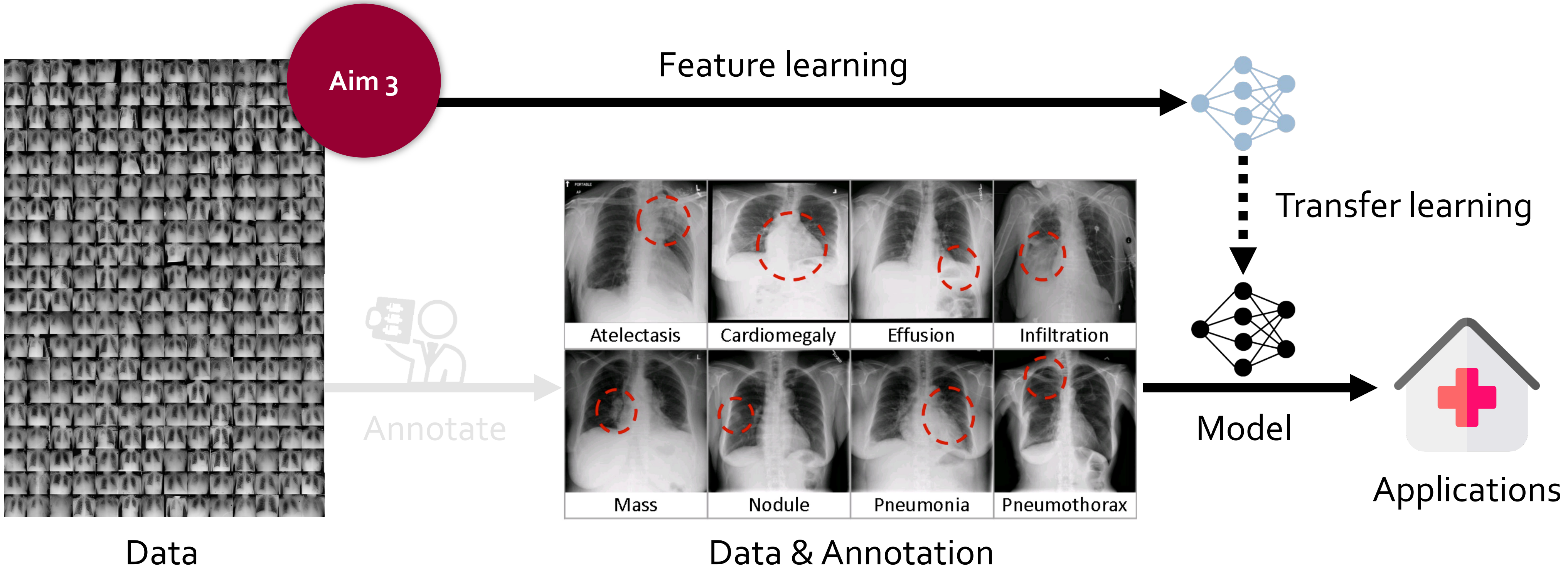


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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



Introduction

Objective

Aim 1

Aim 2

Aim 3

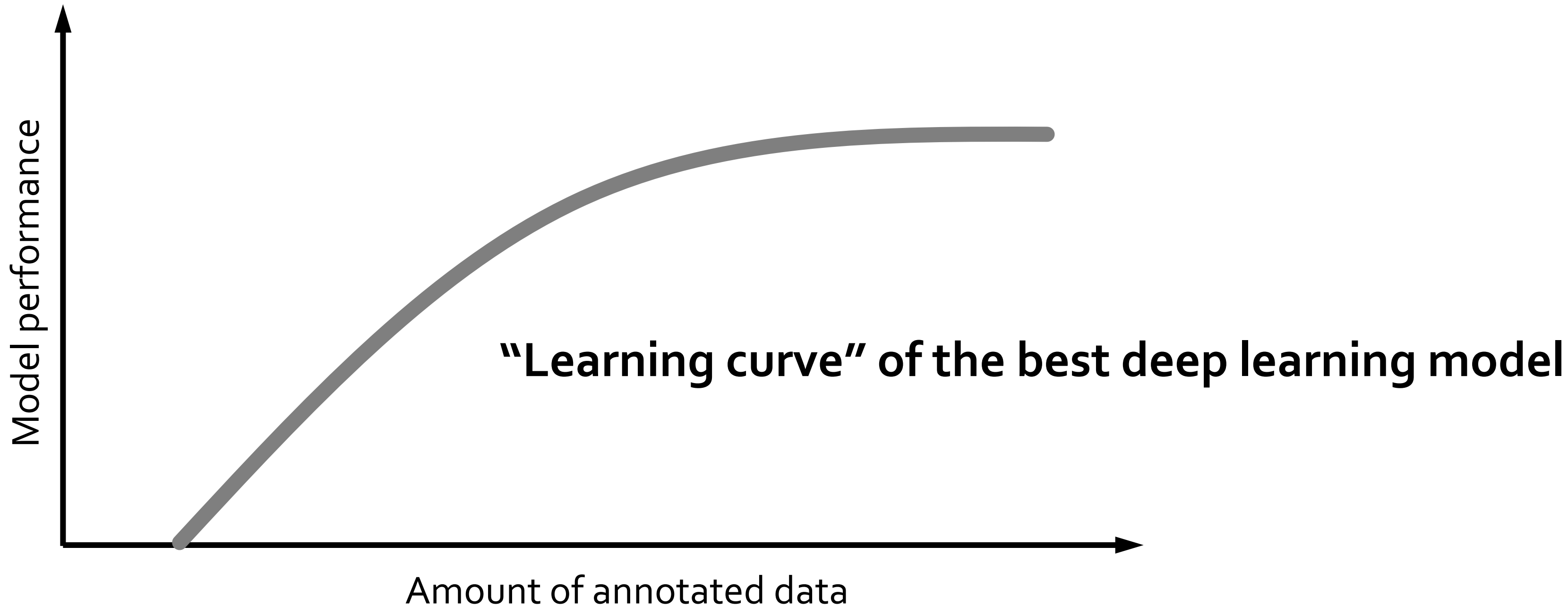
Summary

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

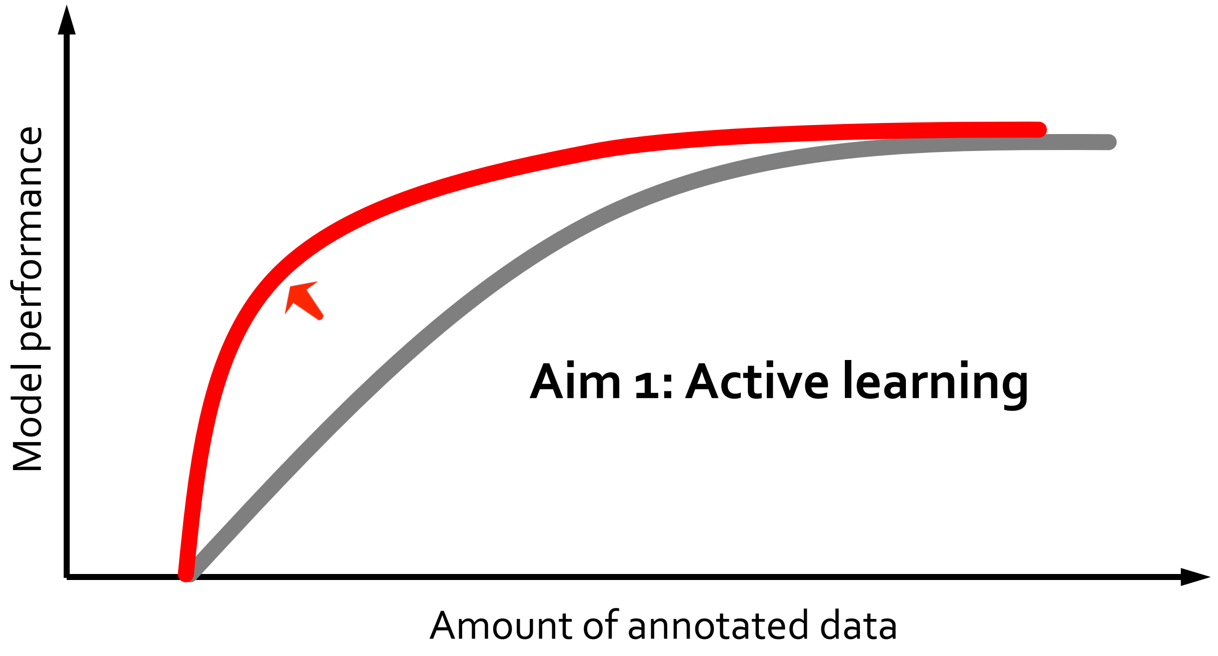


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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

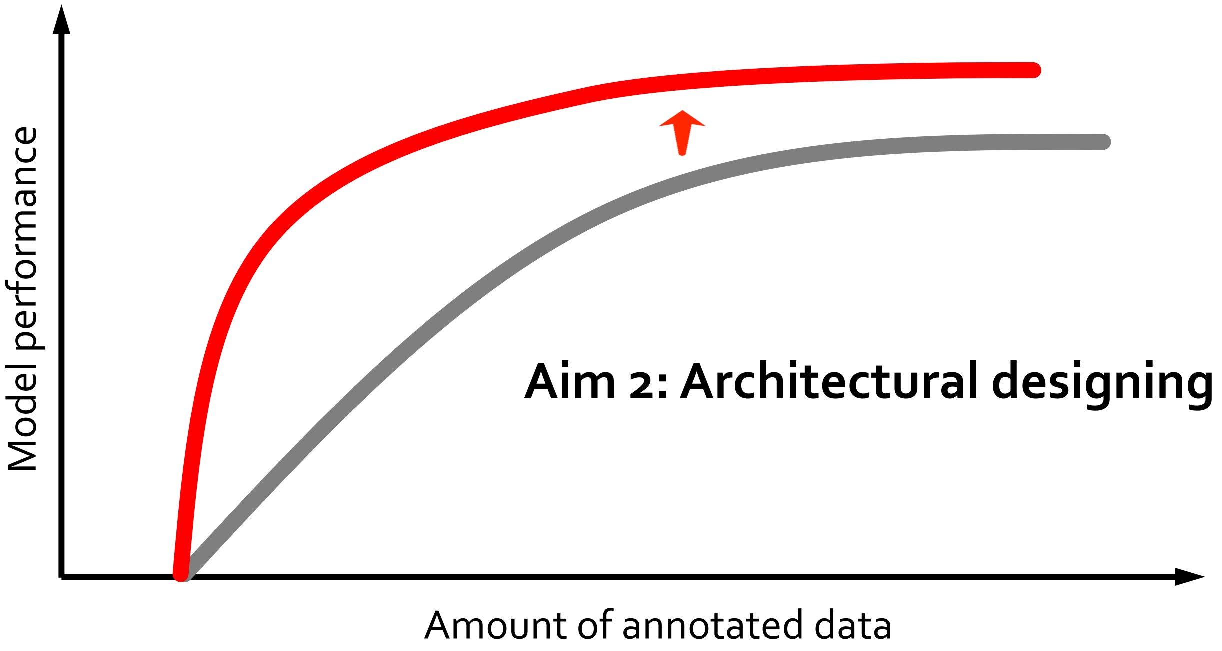


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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

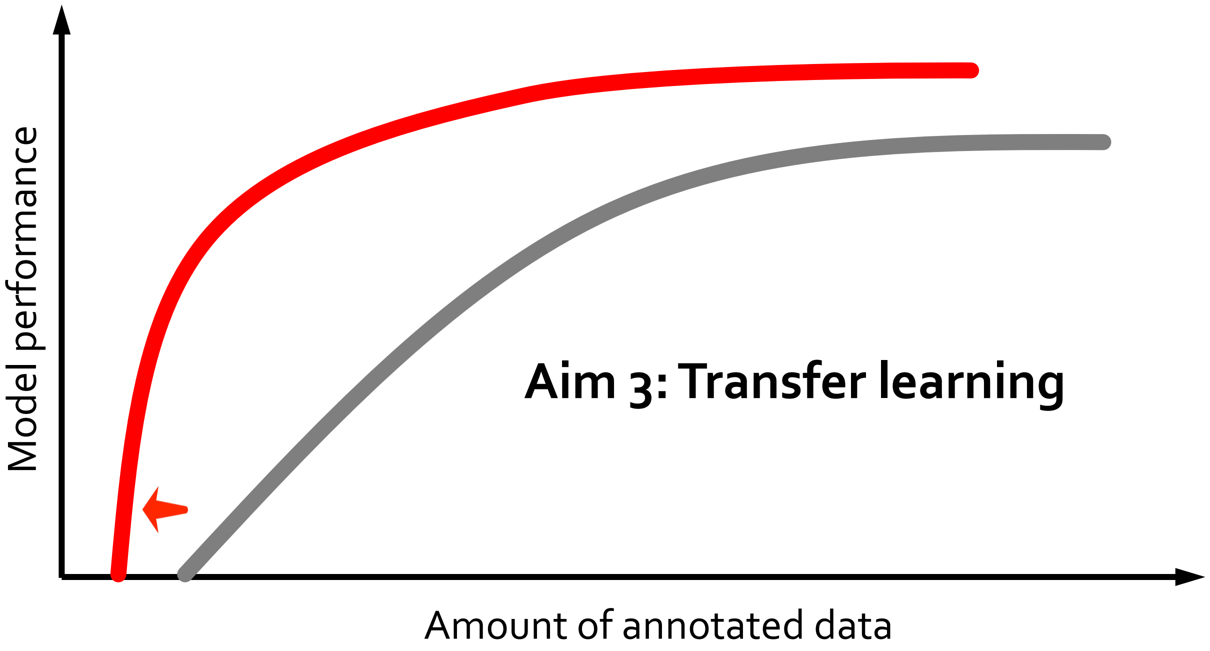


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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



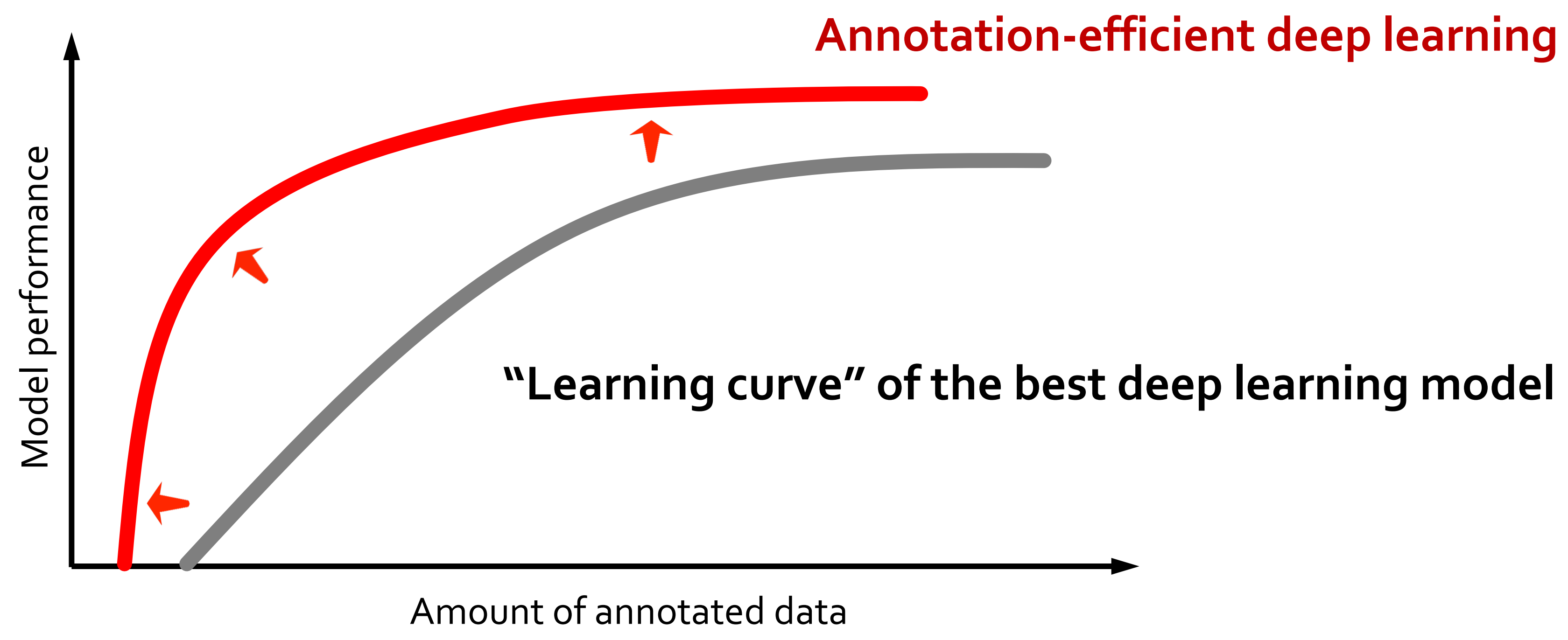
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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.



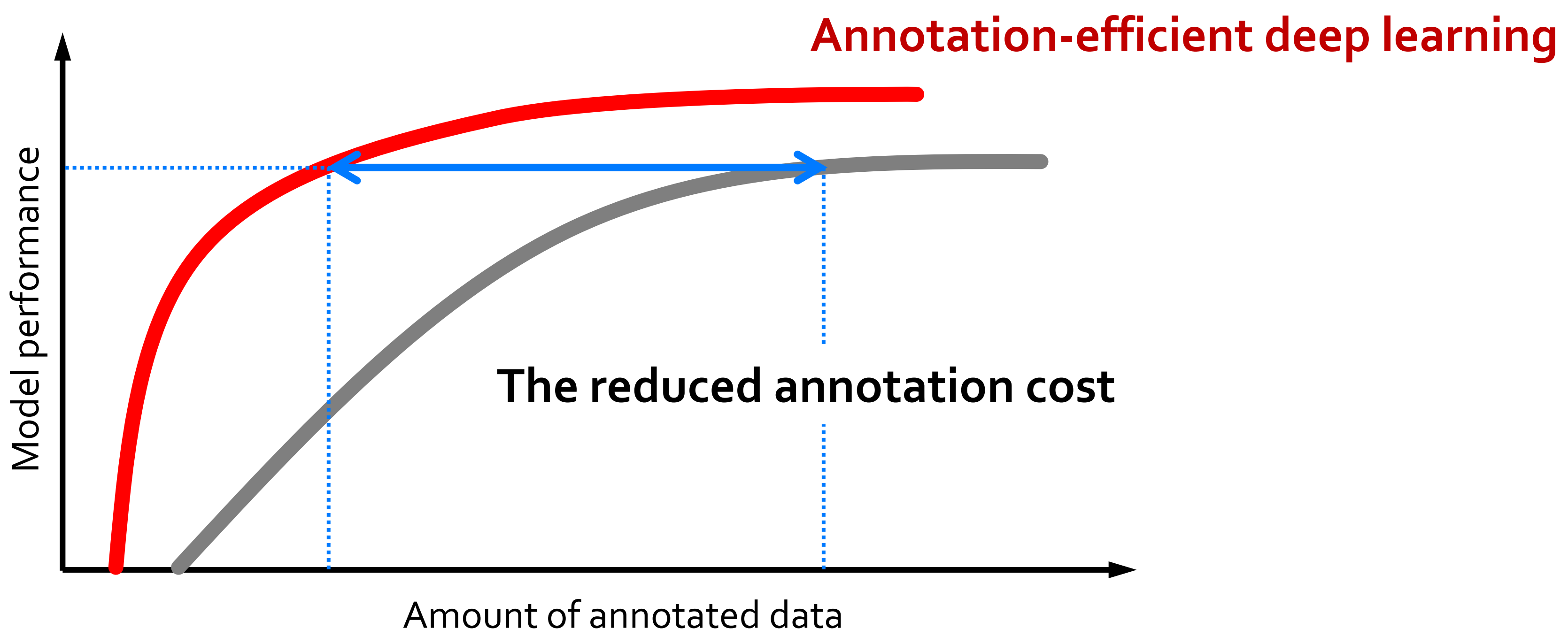
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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

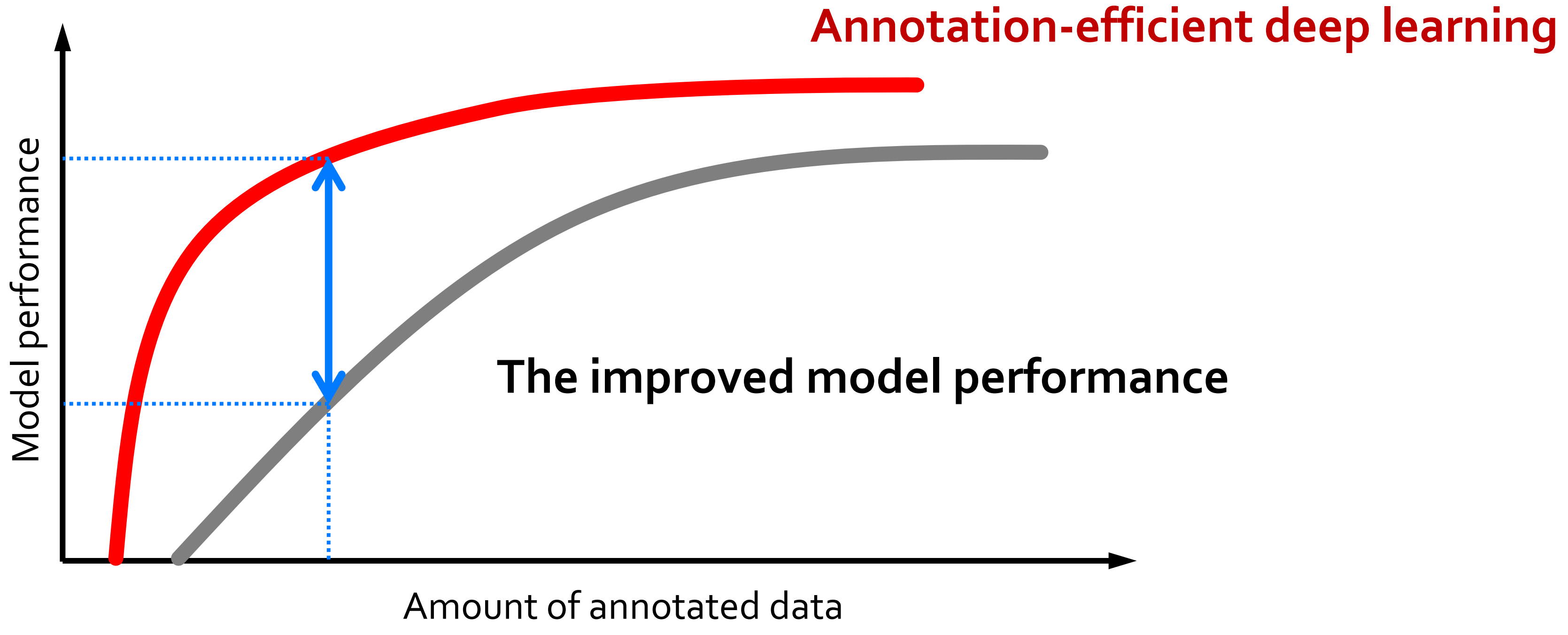
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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

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.

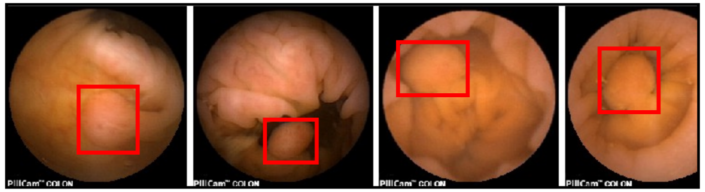


Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

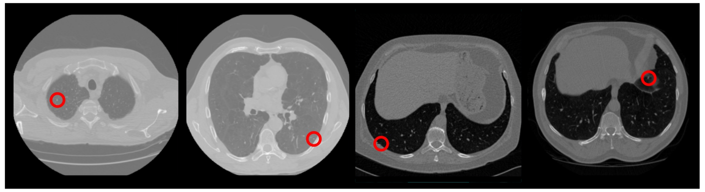
Aim 1: Acquiring necessary annotation efficiently from human experts

Aim 2: Utilizing existing annotation effectively from advanced architecture

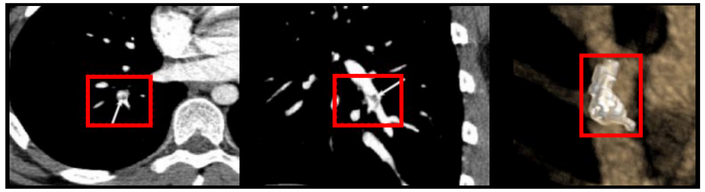
Aim 3: Extracting generic knowledge directly from unannotated images



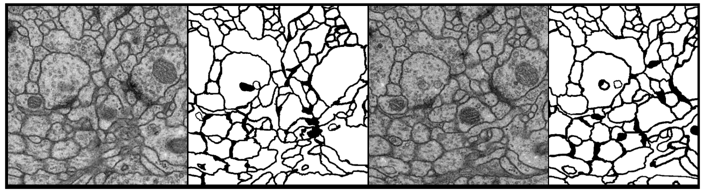
Polyp detection



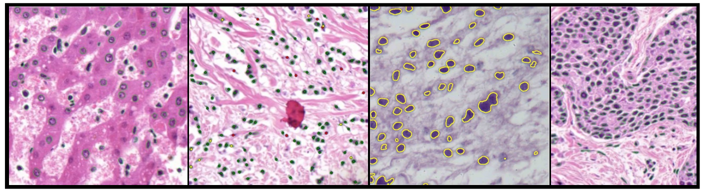
Lung nodule detection



Pulmonary embolism detection

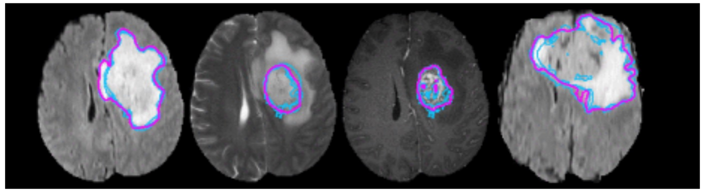


Neuronal structure segmentation

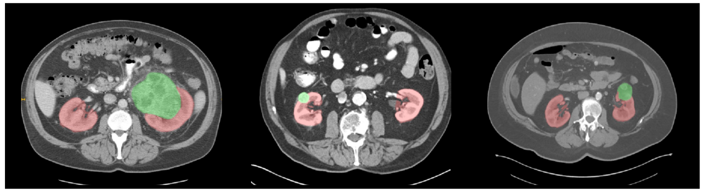


Cell/nuclei segmentation

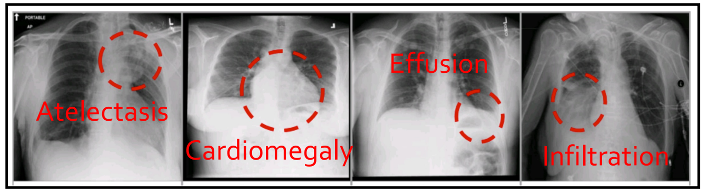
Liver/lesion segmentation



Brain/tumor segmentation



Kidney/lesion segmentation



Pulmonary diseases classification

Introduction

Objective

Aim 1

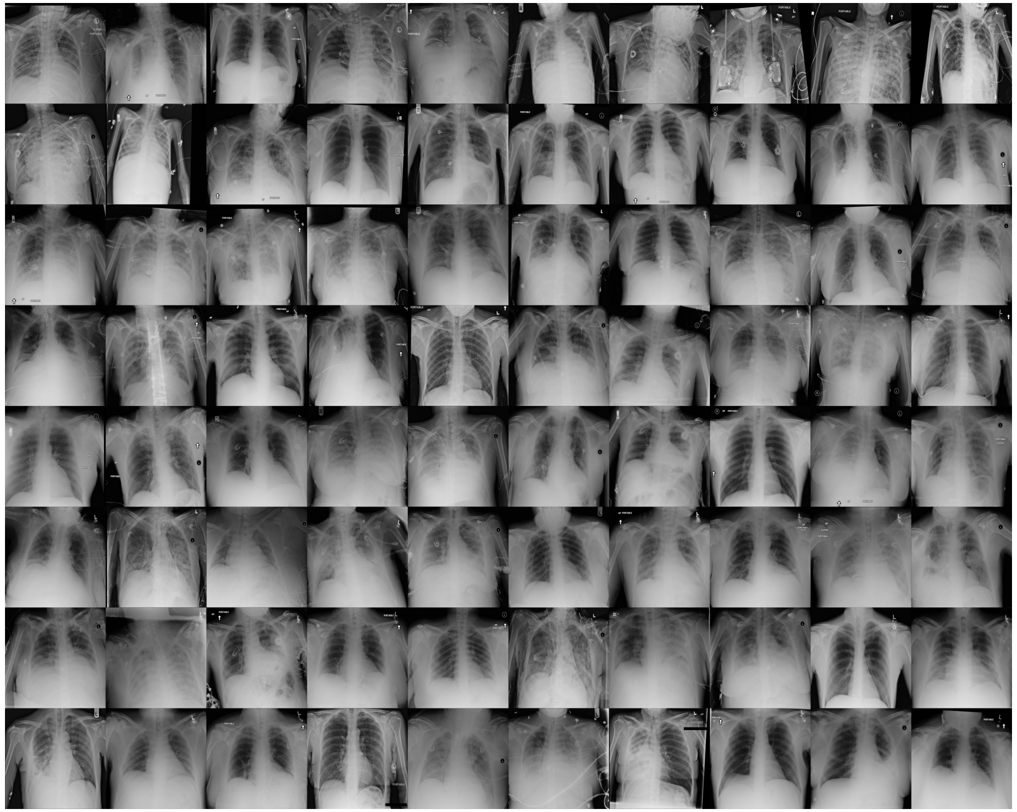
Aim 2

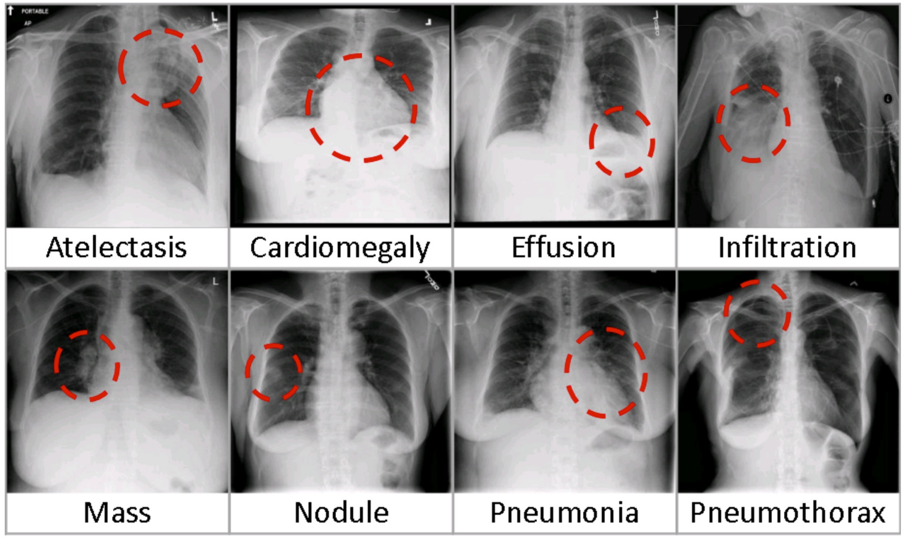
Aim 3

Summary

Aim 1: Acquiring necessary annotation efficiently from human experts

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





Introduction

Objective

Aim 1

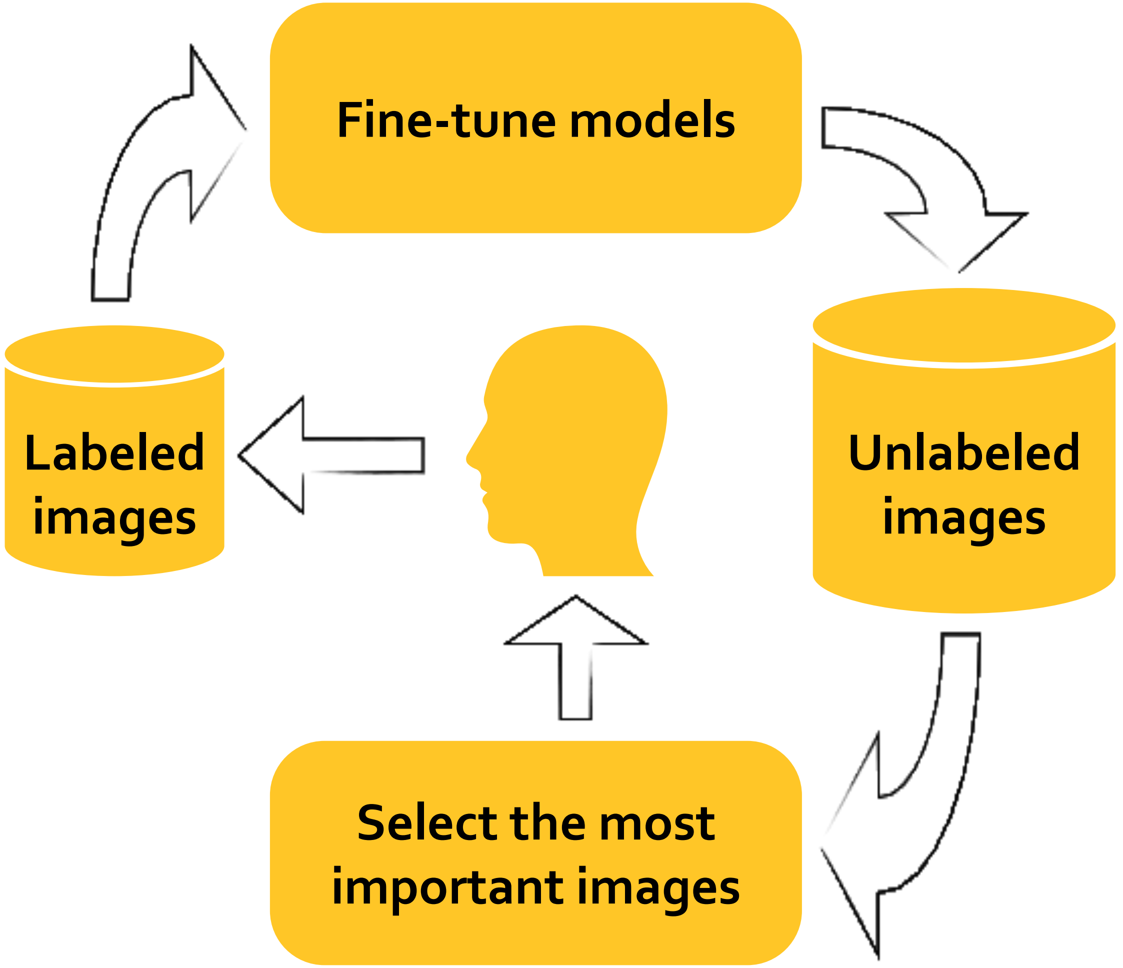
Aim 2

Aim 3

Summary

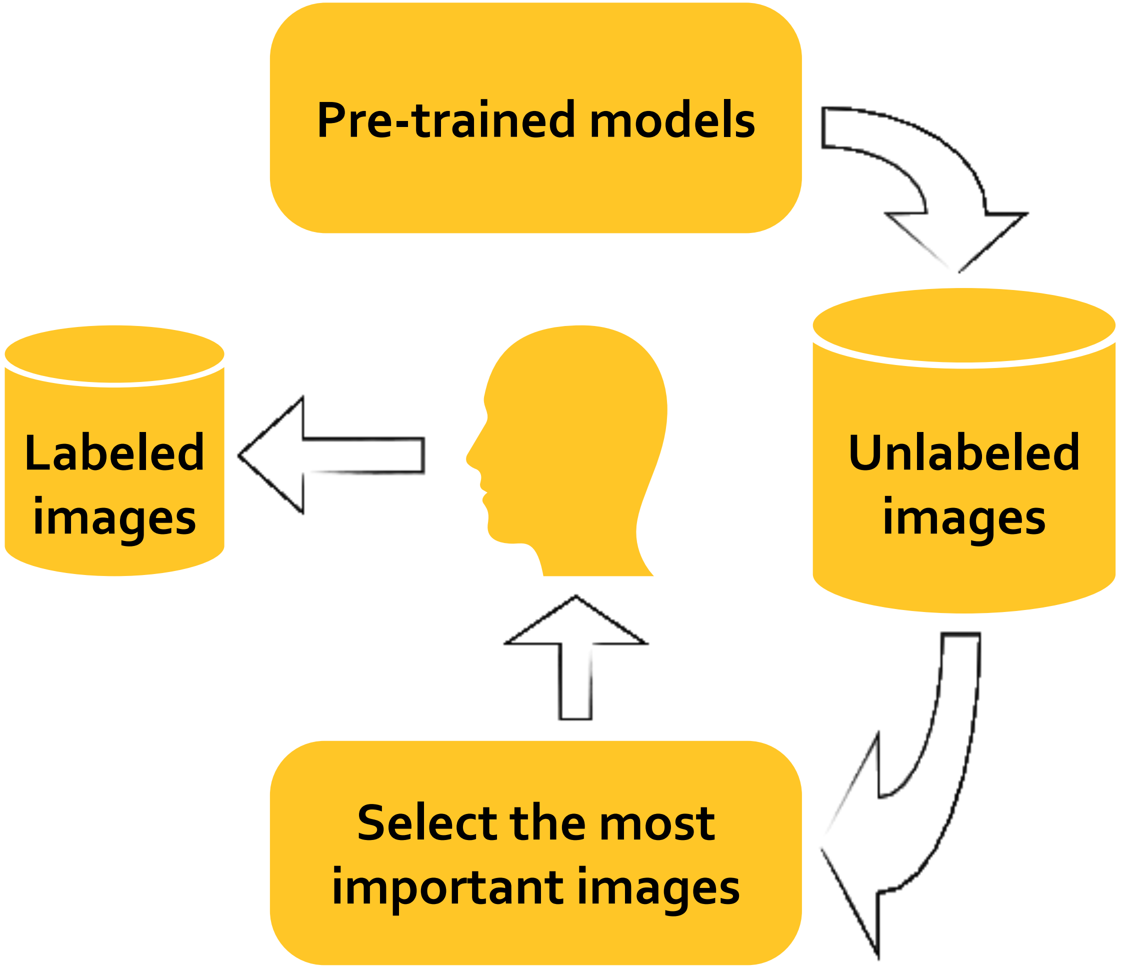
Aim 1: Acquiring necessary annotation efficiently from human experts

Approach: “Human-in-the-loop” active learning procedure



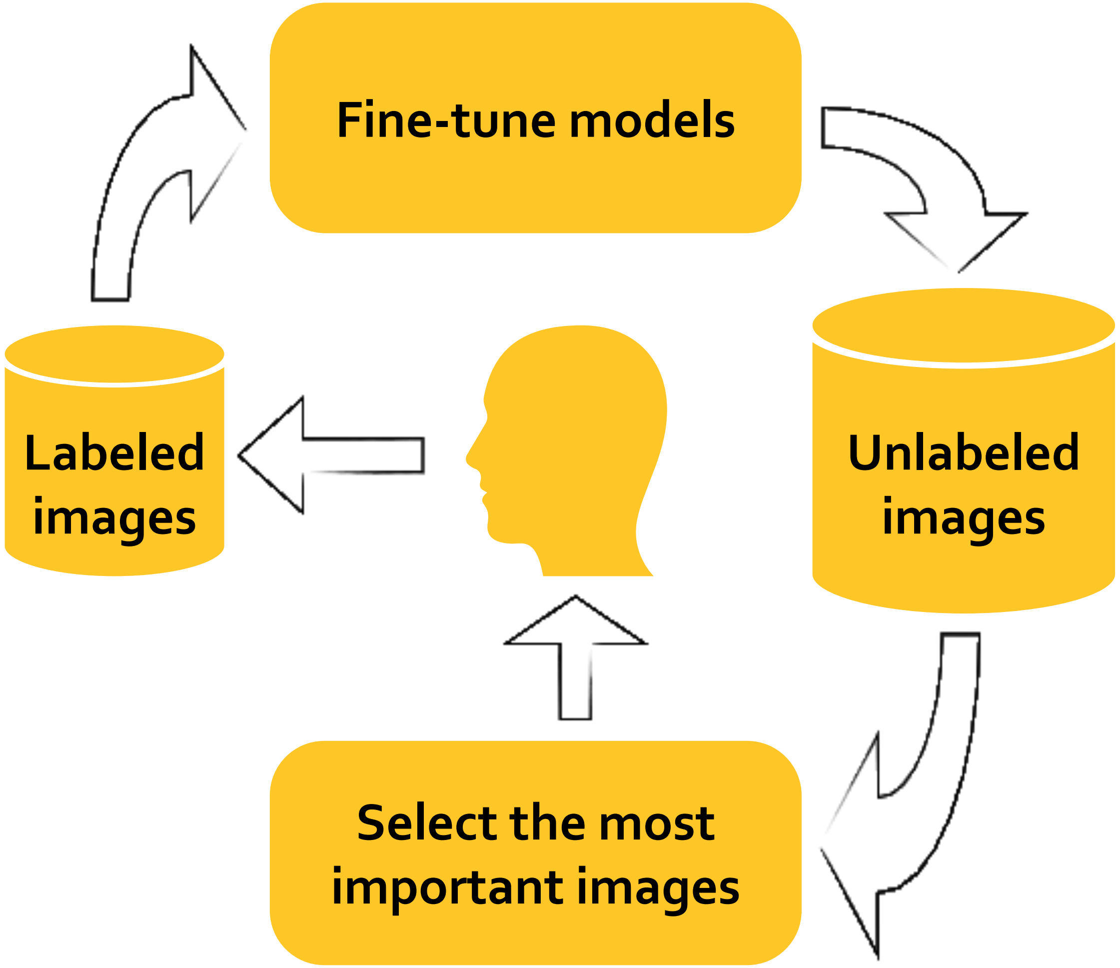
Aim 1: Acquiring necessary annotation efficiently from human experts

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



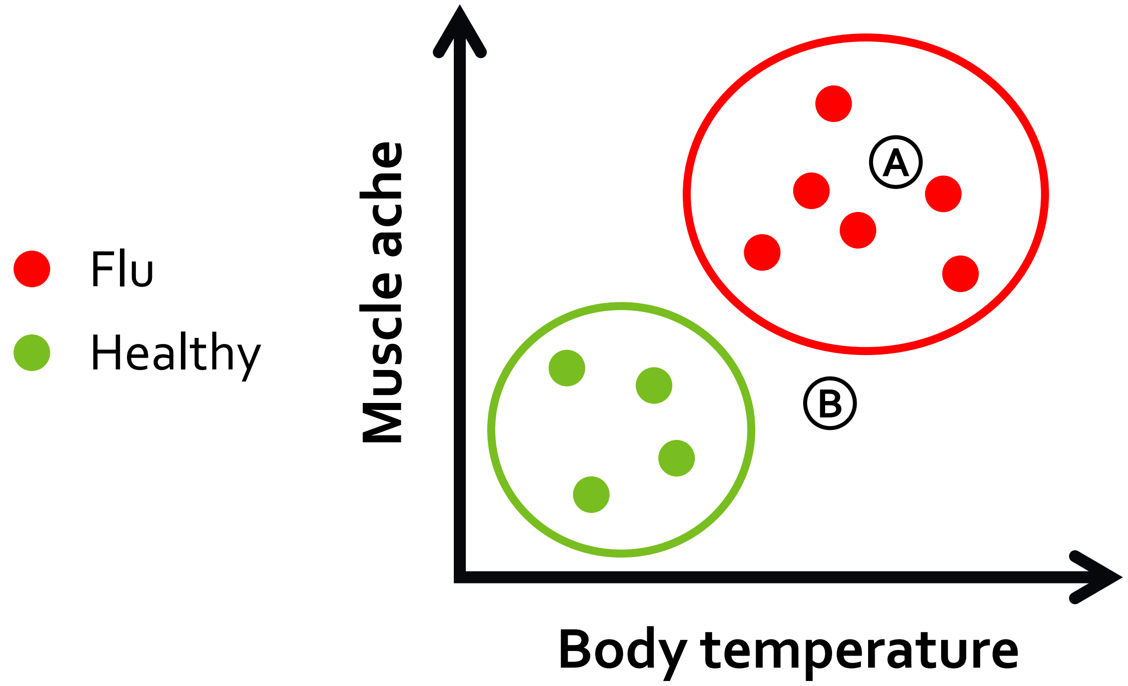
Aim 1: Acquiring necessary annotation efficiently from human experts

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

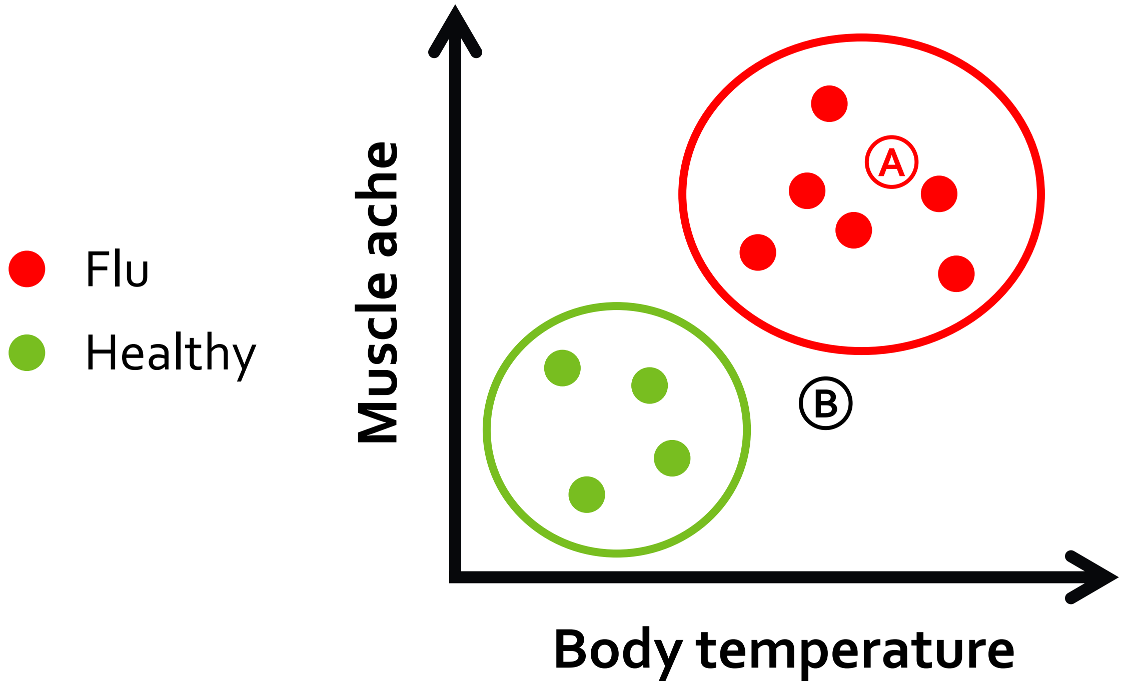


Select the most important images

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

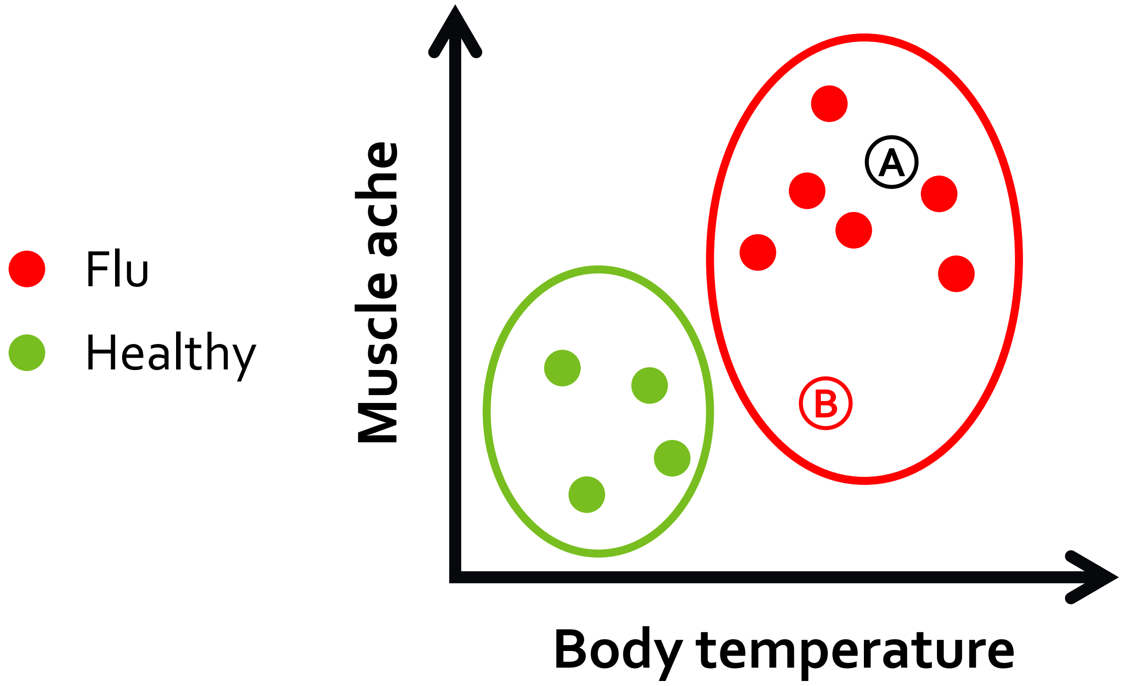


Select the most important images

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

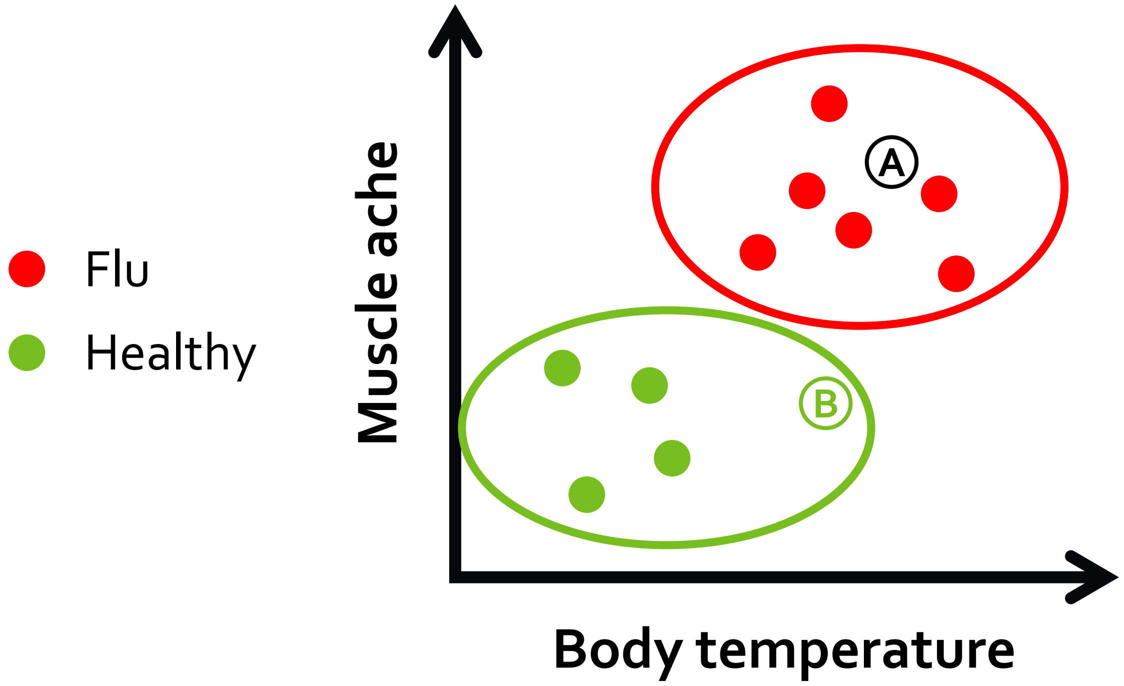


Select the most important images

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

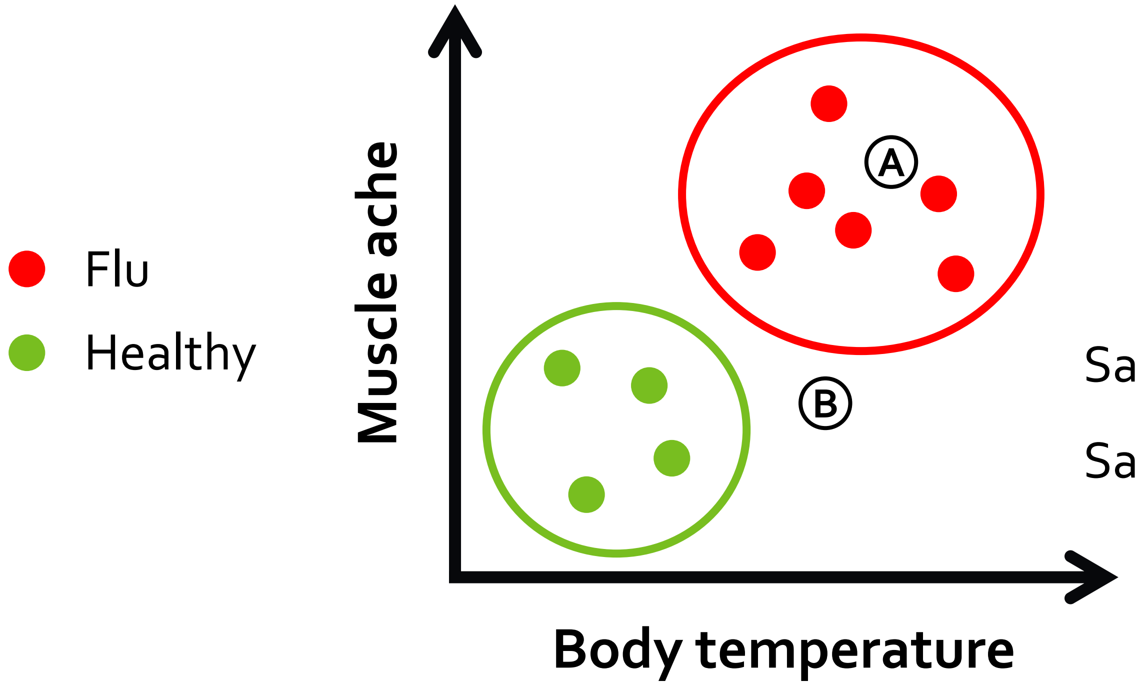


Select the most important images

Which sample would you annotate first, A or B?

Aim 1: Acquiring necessary annotation efficiently from human experts

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



Sample A: *low* uncertainty

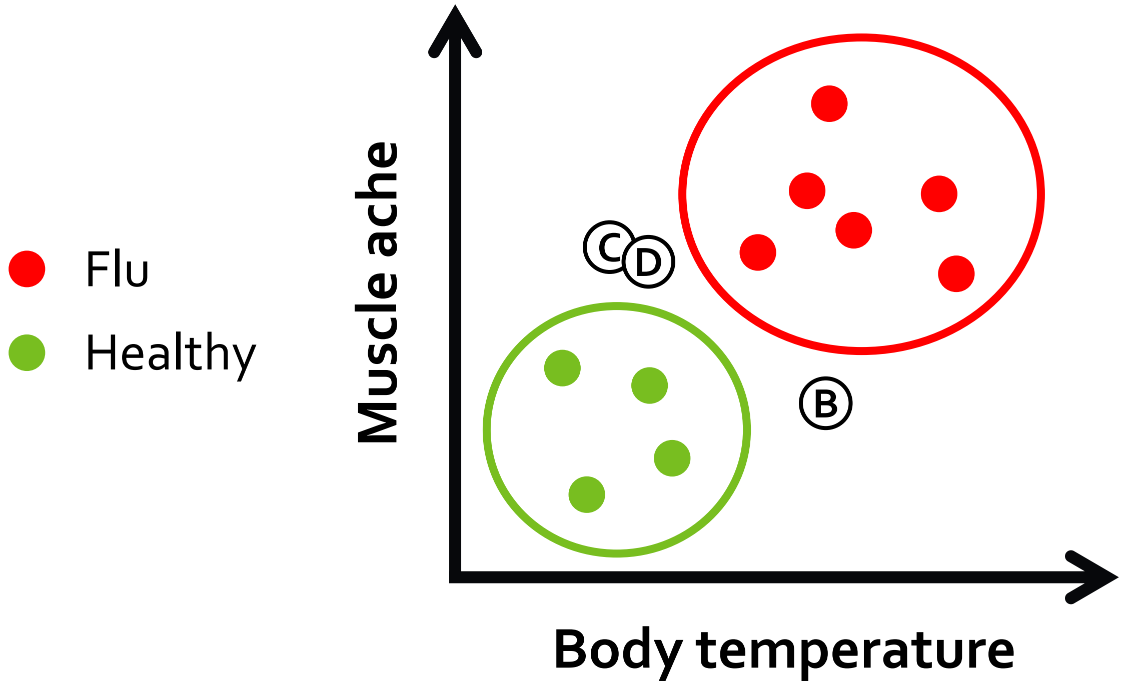
Sample B: *high* uncertainty

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*

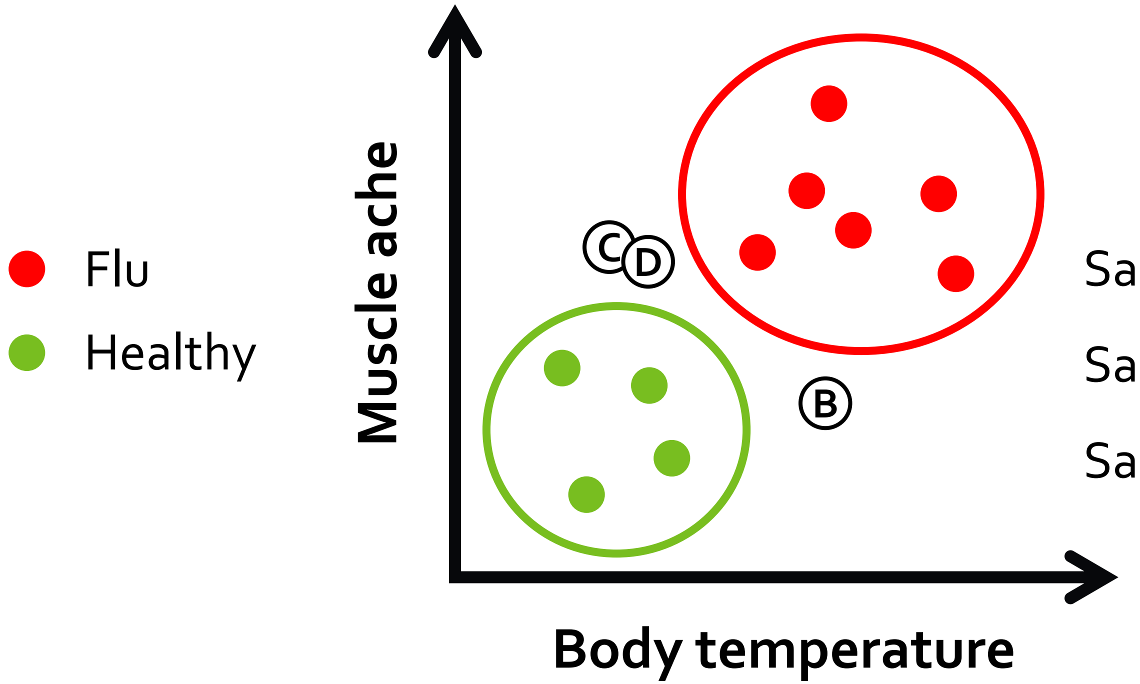


Select the most important images

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

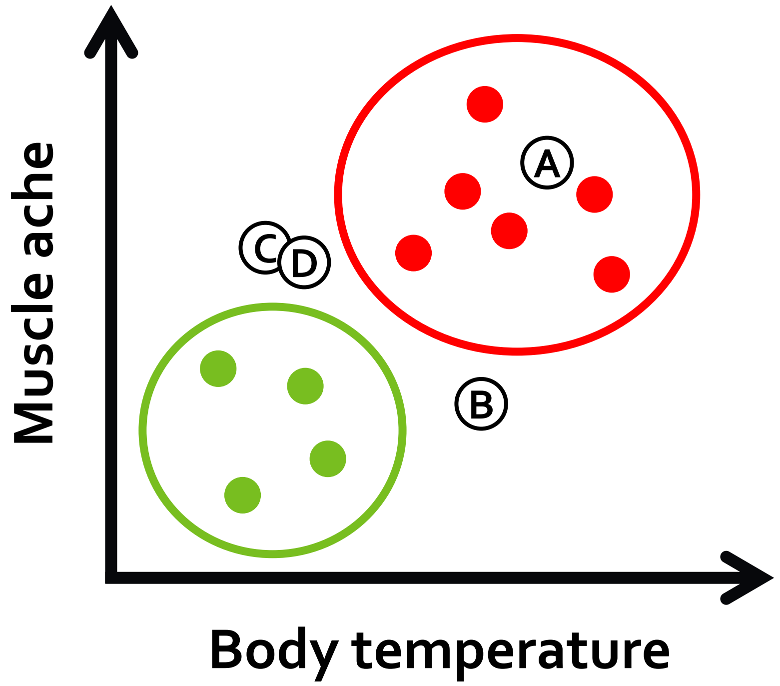
Samples C, D: *low* diversity

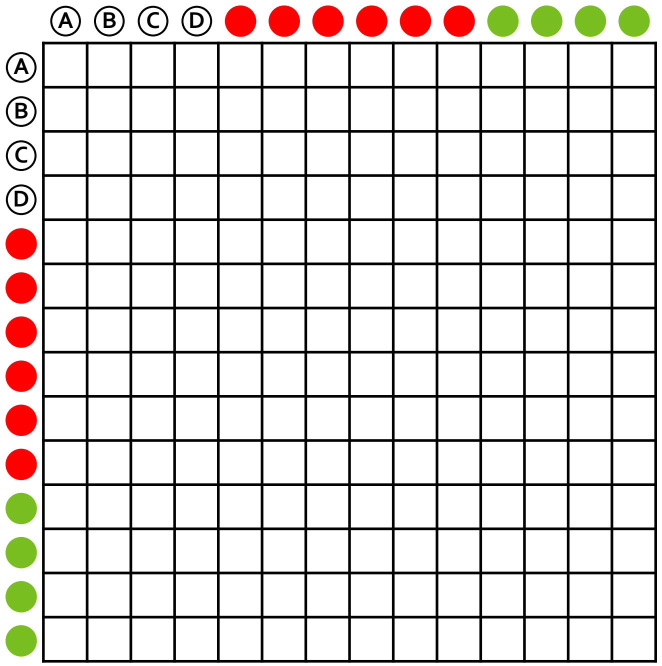
Select the most important images

Which two samples would you annotate first?

Aim 1: Acquiring necessary annotation efficiently from human experts

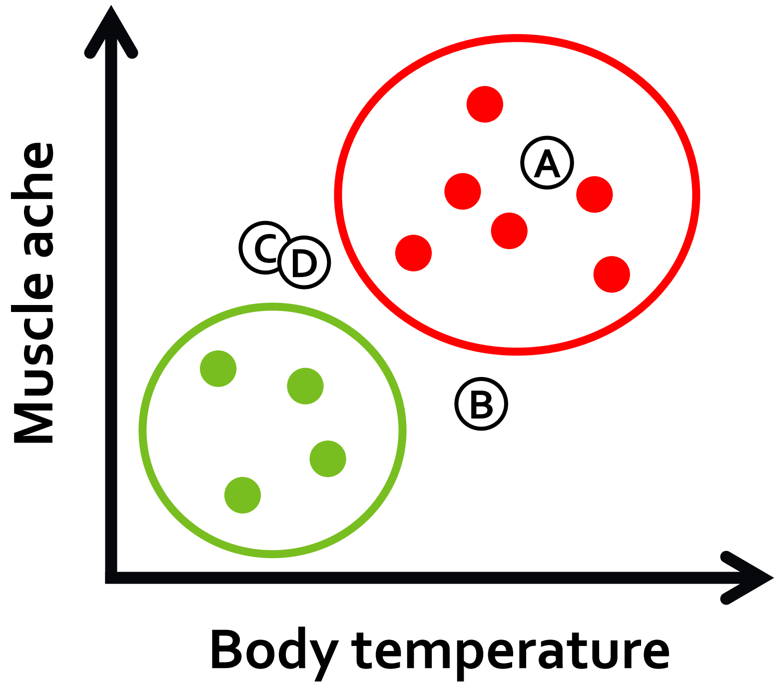
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*

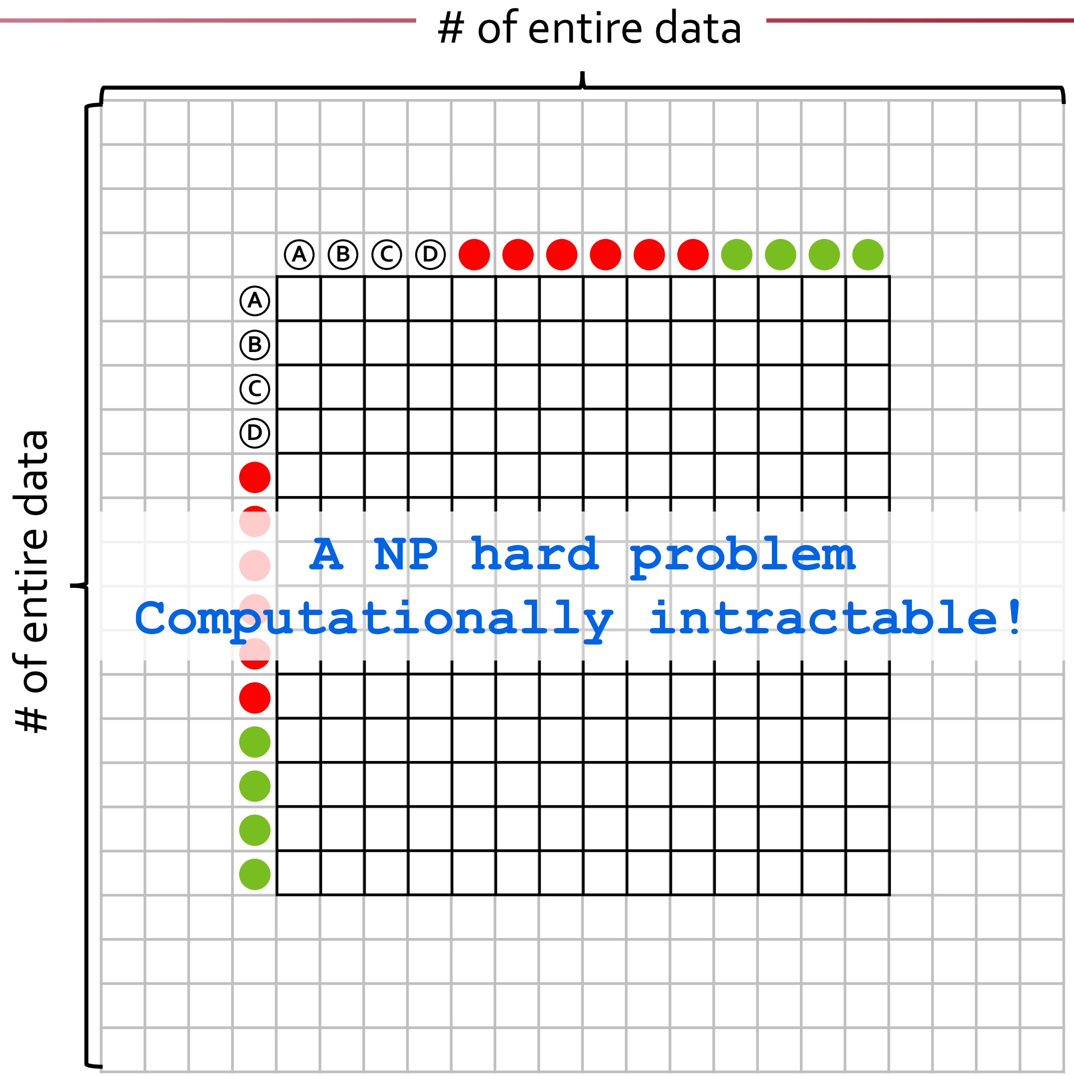




Aim 1: Acquiring necessary annotation efficiently from human experts

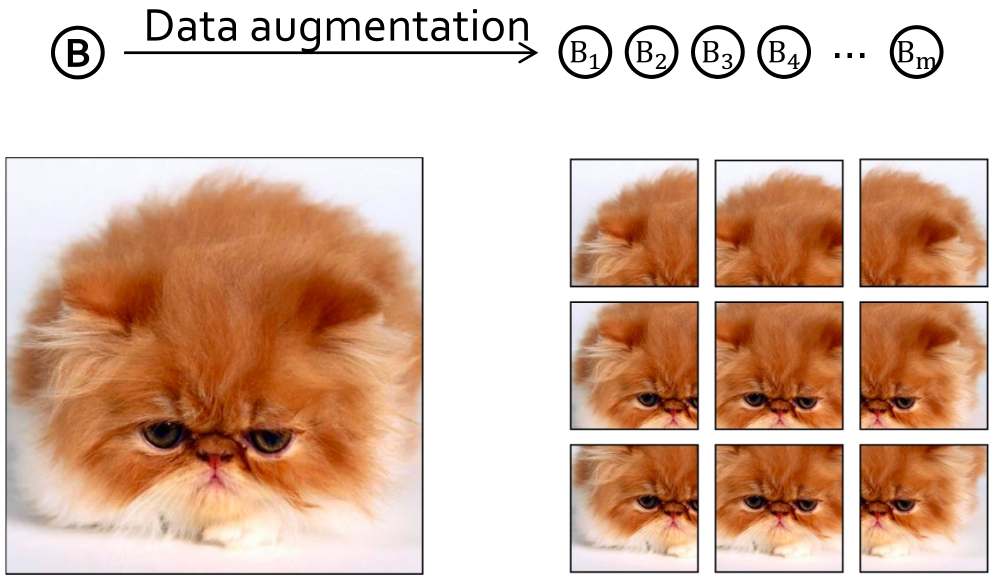
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*





Aim 1: Acquiring necessary annotation efficiently from human experts

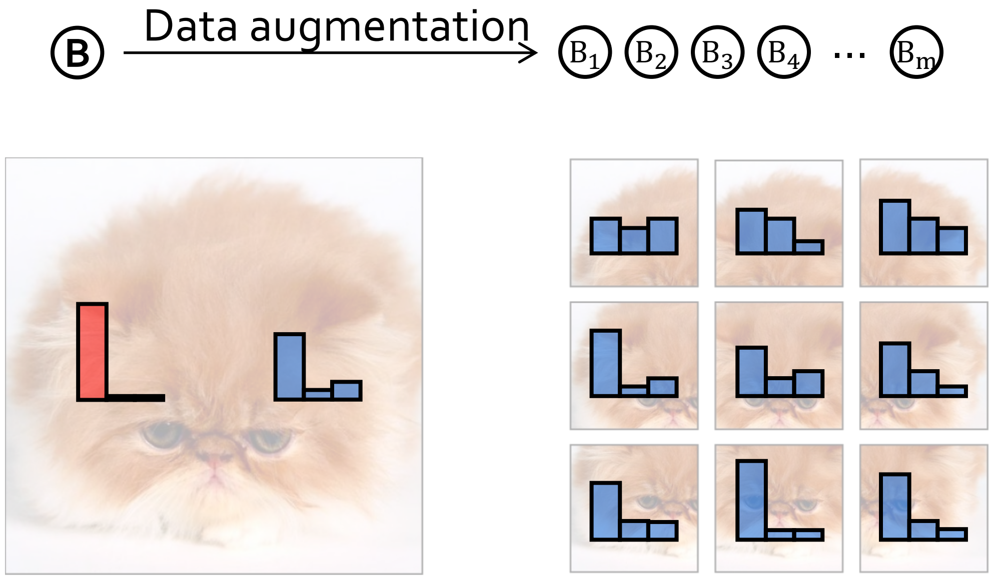
Approach: "Human-in-the-loop" active learning procedure, *diversity-based*



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*

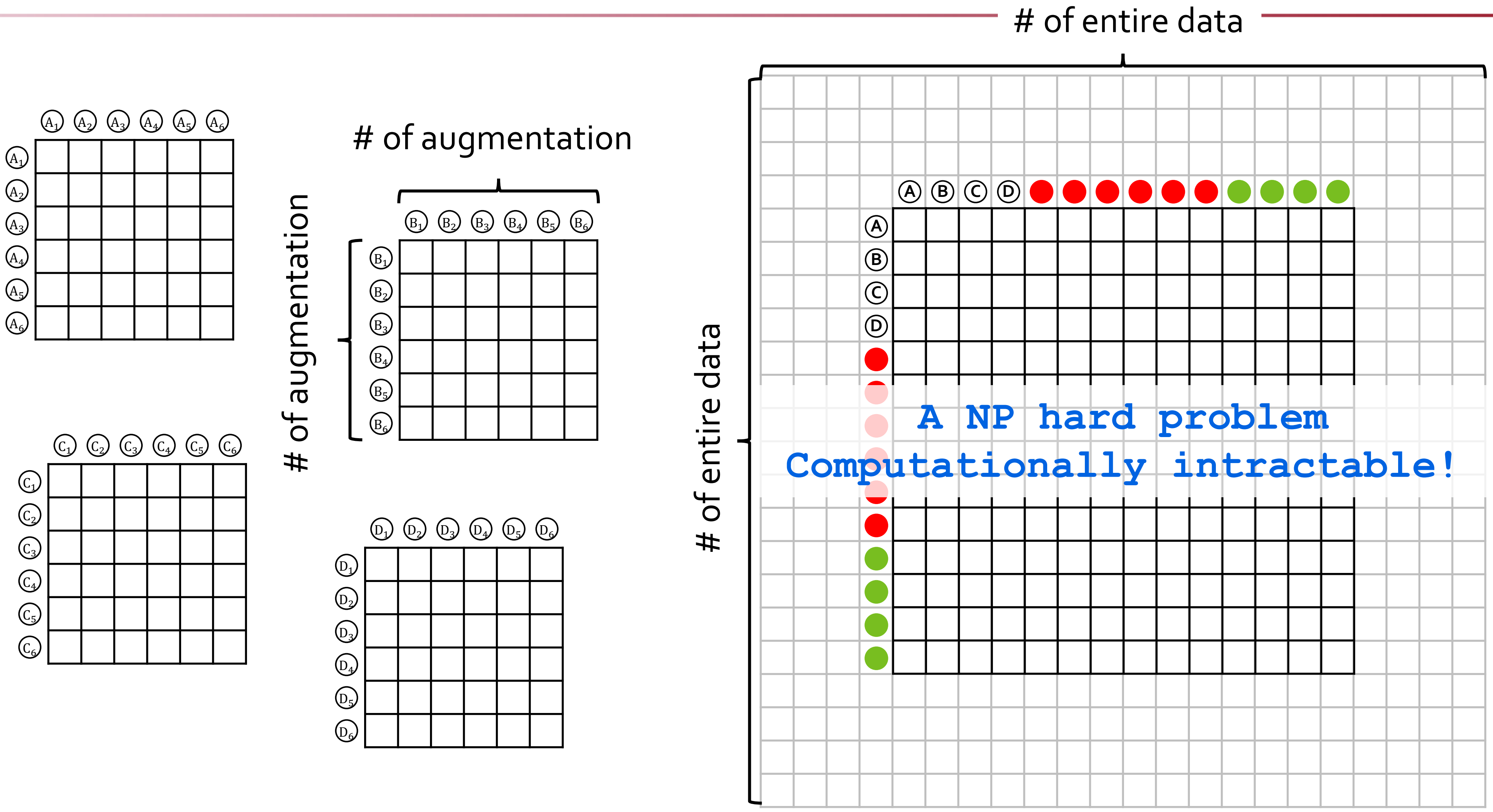


Cross-entropy between human labels and model predictions

Consistency among model predictions for the same sample

Aim 1: Acquiring necessary annotation efficiently from human experts

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



Introduction

Objective

Aim 1

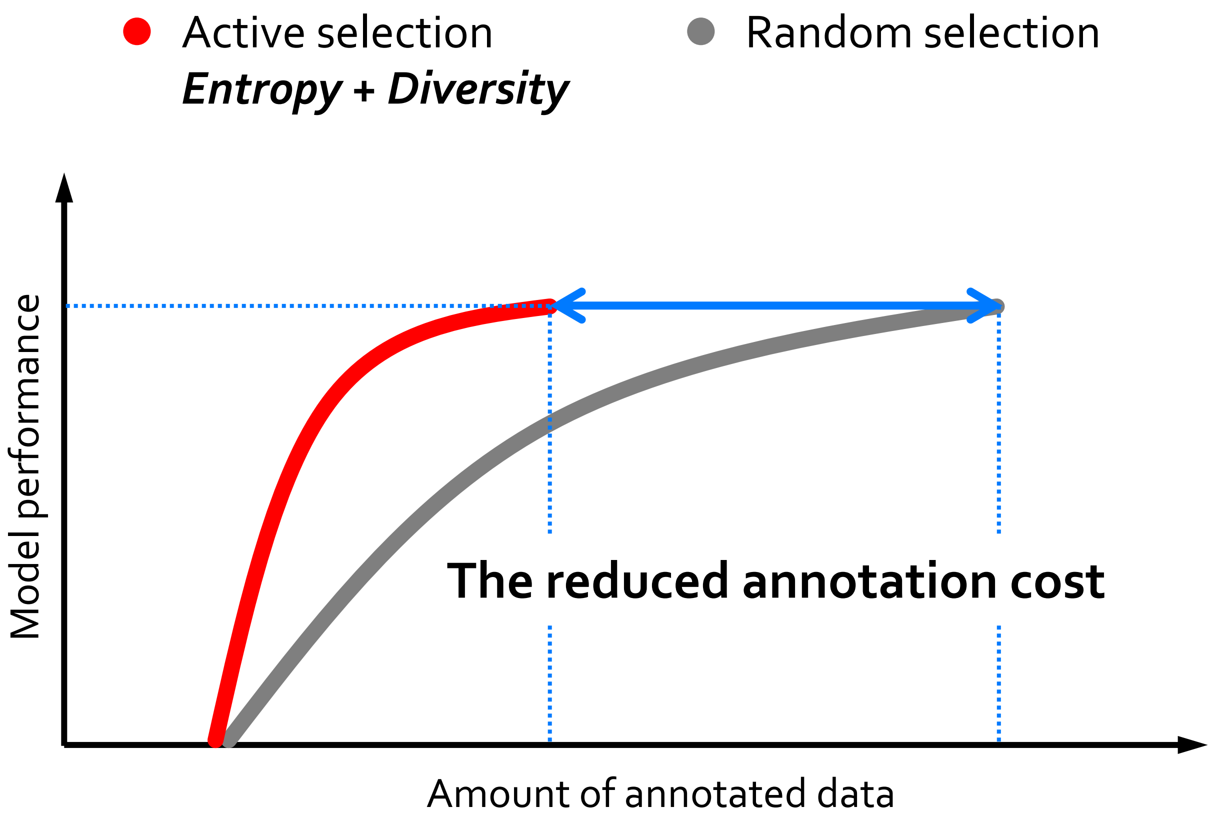
Aim 2

Aim 3

Summary

Aim 1: Acquiring necessary annotation efficiently from human experts

Hypothesis: Wisely selecting important samples can reduce annotation cost



Introduction

Objective

Aim 1

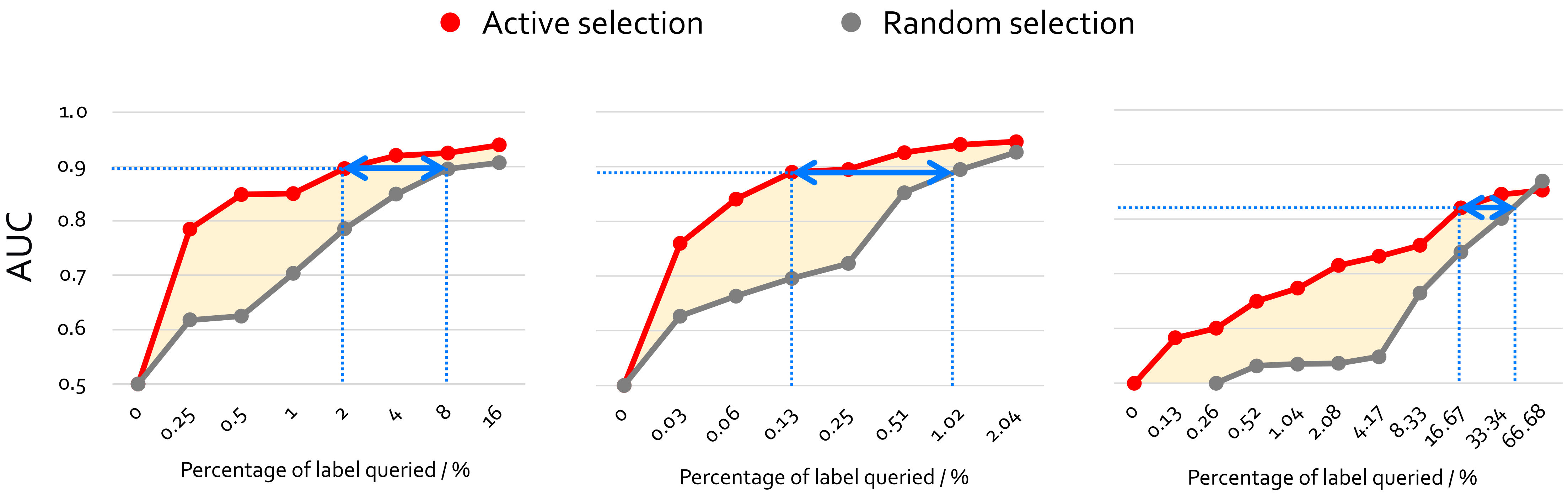
Aim 2

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, 2019.
2. Zhou, Zongwei, et al. "Active, Continual Fine Tuning of Convolutional Neural Networks for Reducing Annotation Efforts." Medical Image Analysis , 2021.
3. Zhou, Zongwei, et al. "Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally." CVPR, 2017.

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*.

- L. Chen, Y. Bai, S. Huang, Y. Lu, B. Wen, A. Yuille, Z. Zhou. 2022. A Guide to Your First Choice: Addressing Cold Start Problem in Vision Active Learning. *NeurIPS HILL*.

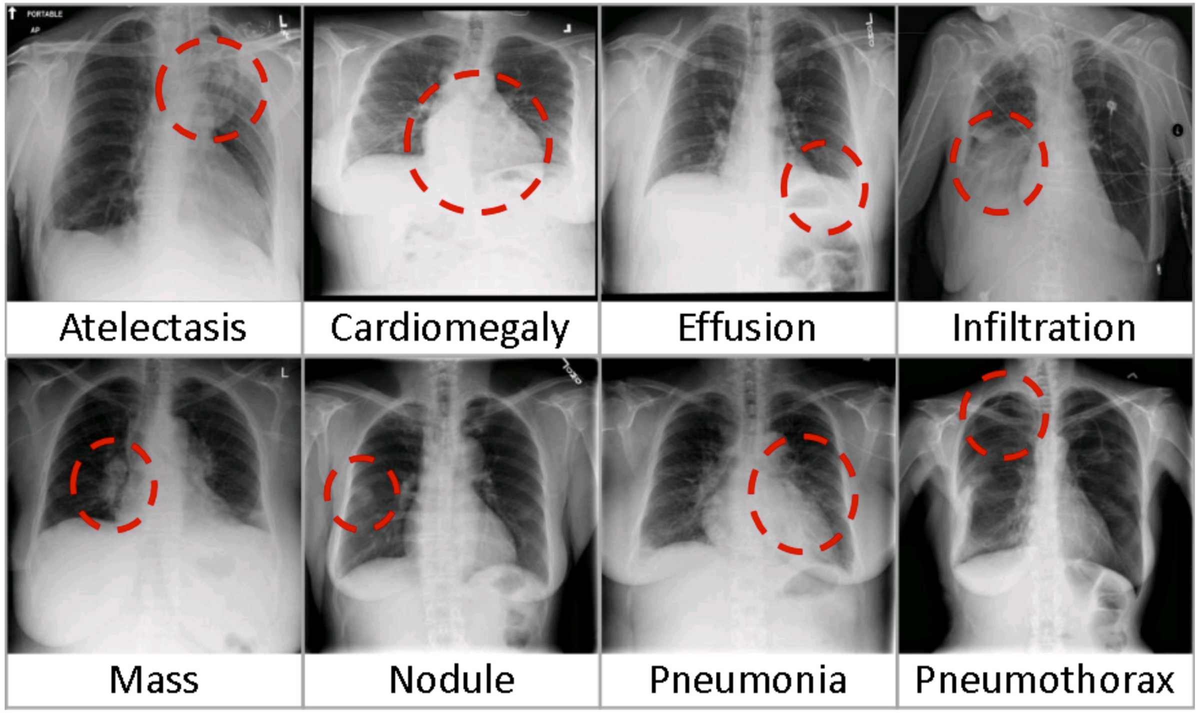
Not All Data Is Created Equal

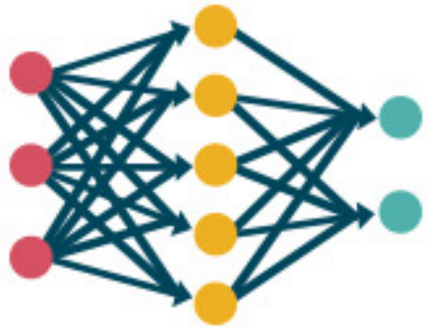
Clinical Impacts of Aim 1:

- An efficient "human-in-the-loop" procedure helps radiologists identify the *most important data*, therefore dramatically reducing the burden of annotation.
- The *continual learning capability* of deep models encourages data, label, and model reuse, significantly improving the training efficiency.

Aim 2: Utilizing existing annotation effectively from advanced architecture

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

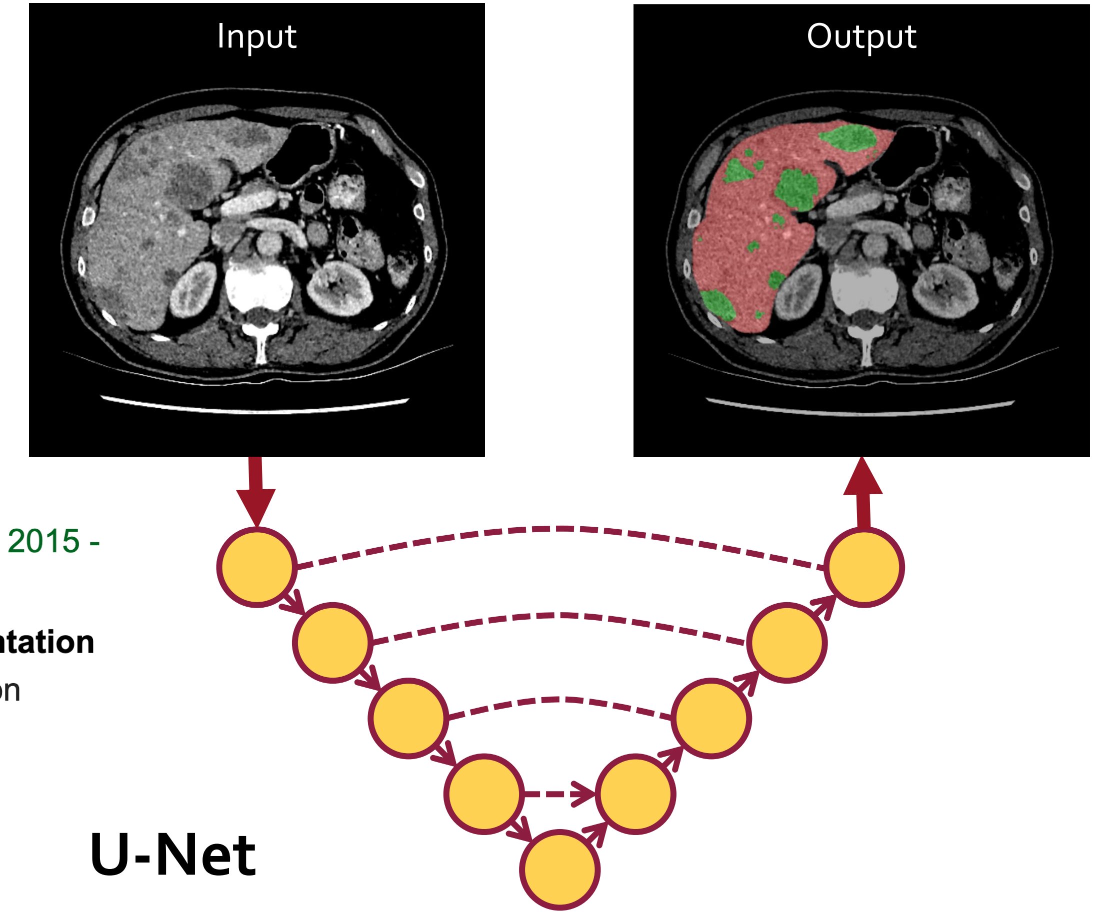




Aim 2: Utilizing existing annotation effectively from advanced architecture

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

e.g., liver & lesion segmentation



U-net: **Convolutional networks** for **biomedical image segmentation**

O Ronneberger, P Fischer, T Brox - … Conference on Medical **image** …, 2015 - Springer

… We demonstrate the application of the **u-net** to three different **segmentation** tasks. The first task is the **segmentation** of neuronal structures in electron microscopic recordings. An …

☆ Save ❞ Cite Cited by 51327 Related articles All 30 versions

Introduction

Objective

Aim 1

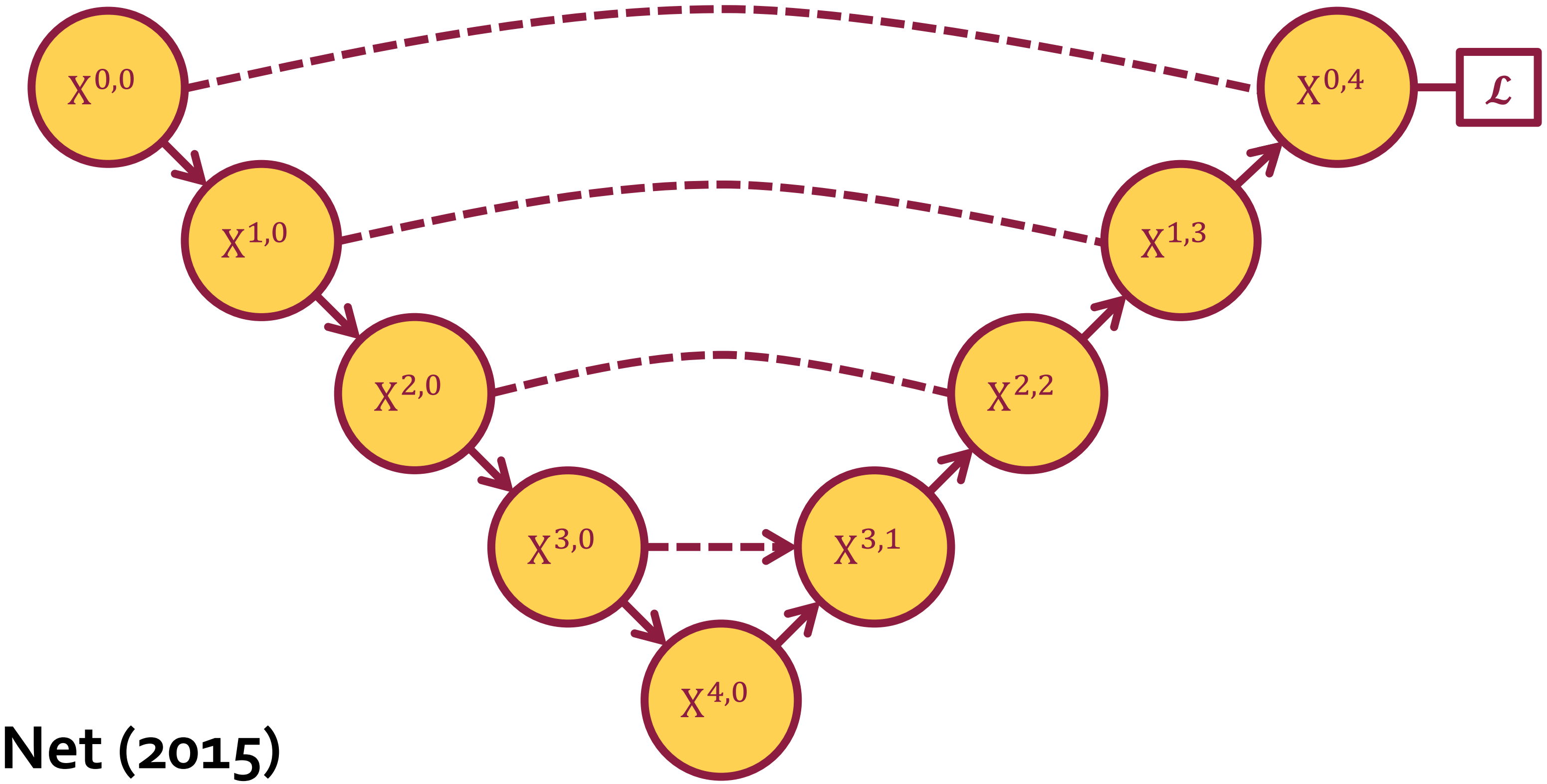
Aim 2

Aim 3

Summary

Aim 2: Utilizing existing annotation effectively from advanced architecture

Hypothesis: Multi-scale feature aggregation leads to powerful models



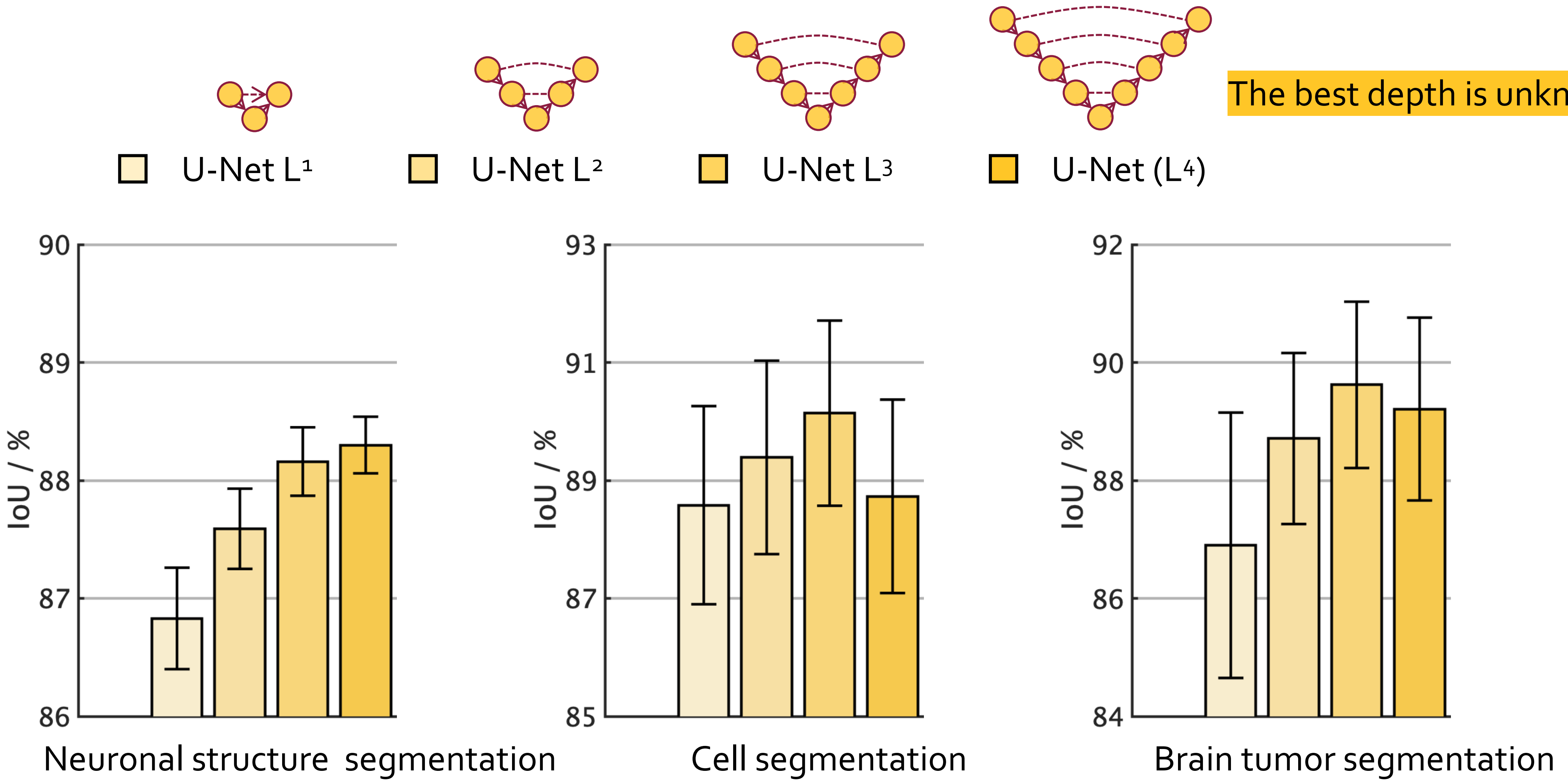
U-Net (2015)

1. Ronneberger, Olaf, et al. "U-net: Convolutional networks for biomedical image segmentation." MICCAI, 2015.

Aim 2: Utilizing existing annotation effectively from advanced architecture

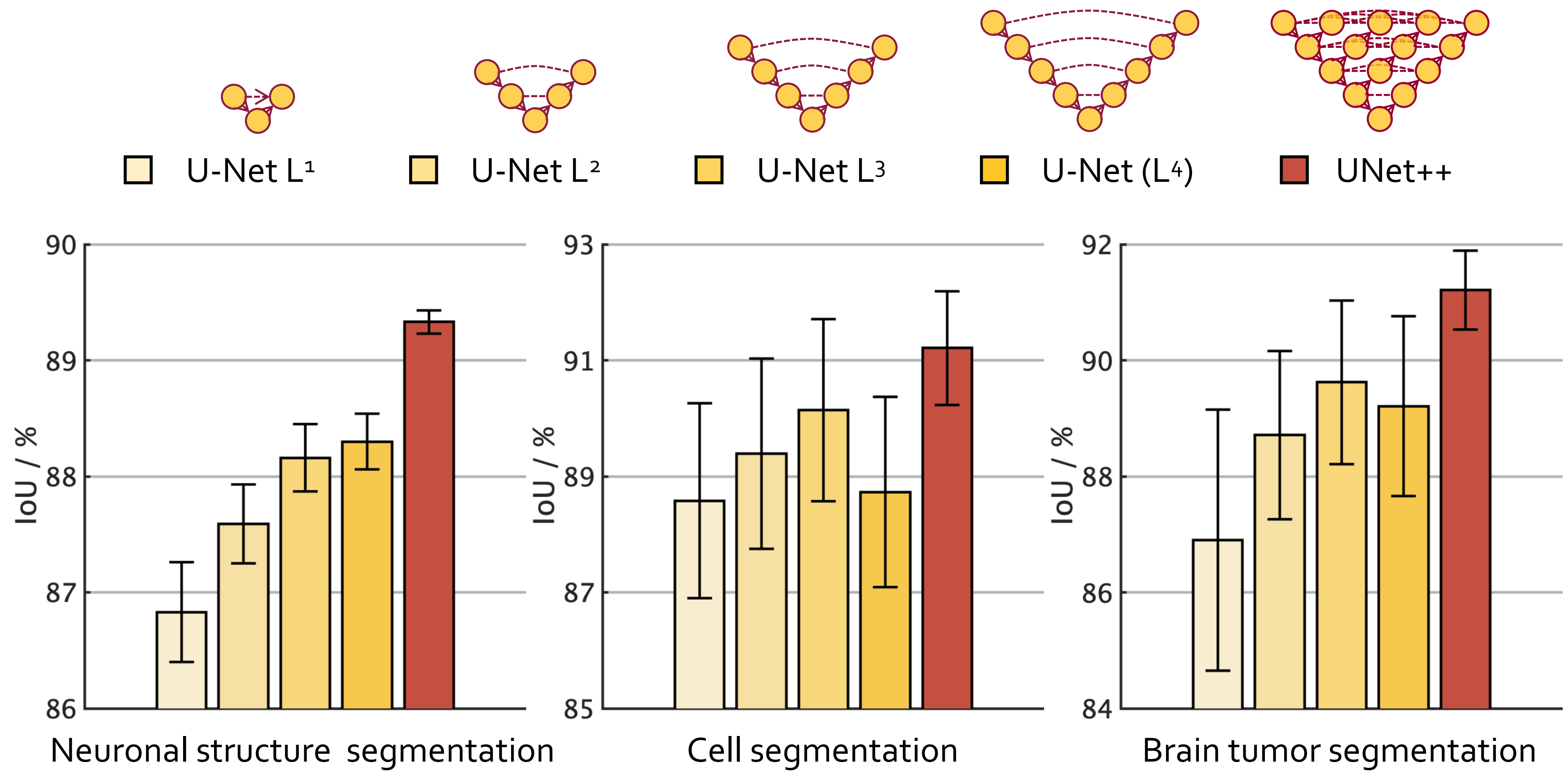
Approach: Redesigned skip connections aggregate multi-scale features

The best depth is unknown



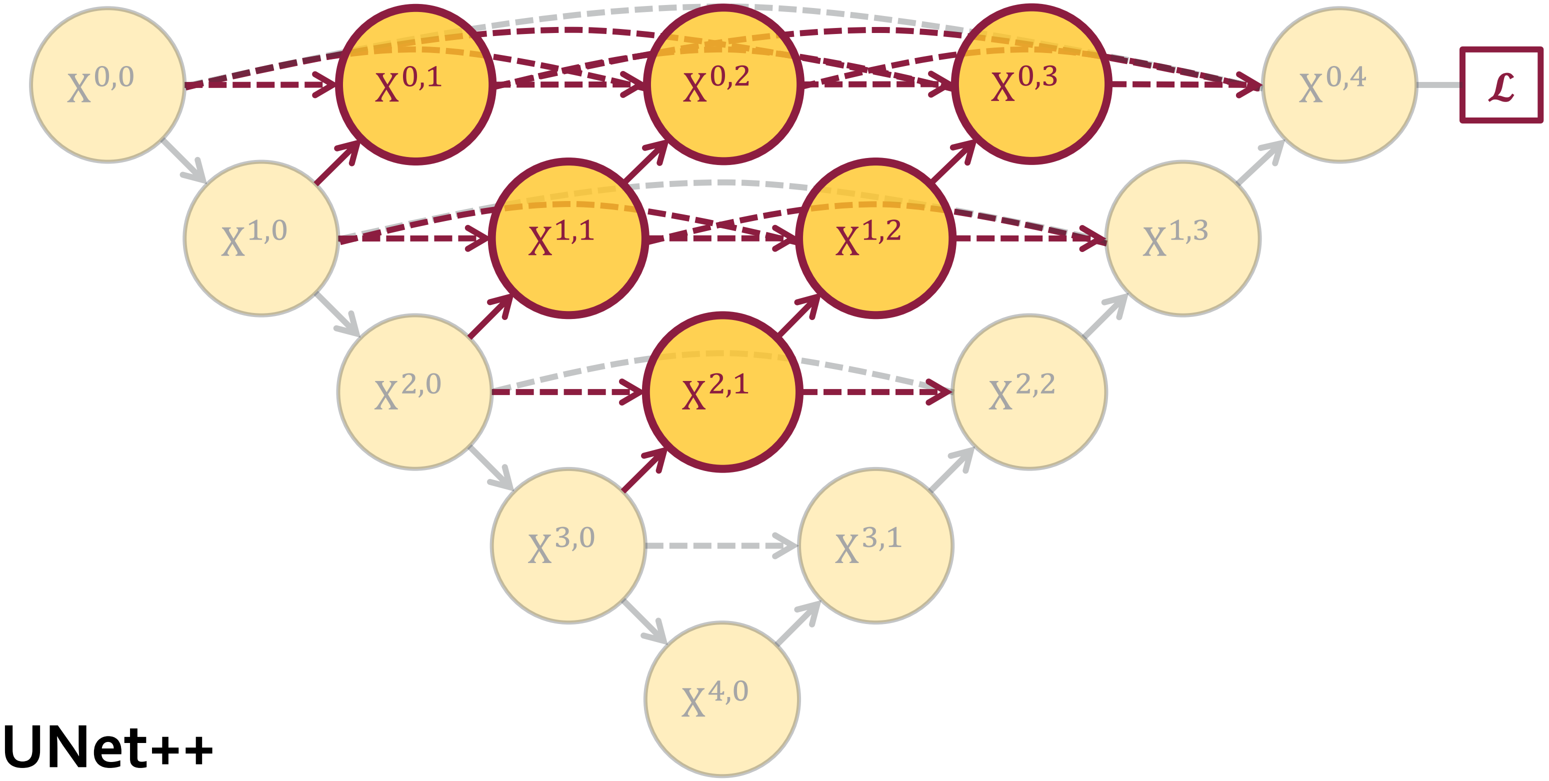
Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Redesigned skip connections aggregate multi-scale features



Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Redesigned skip connections aggregate multi-scale features

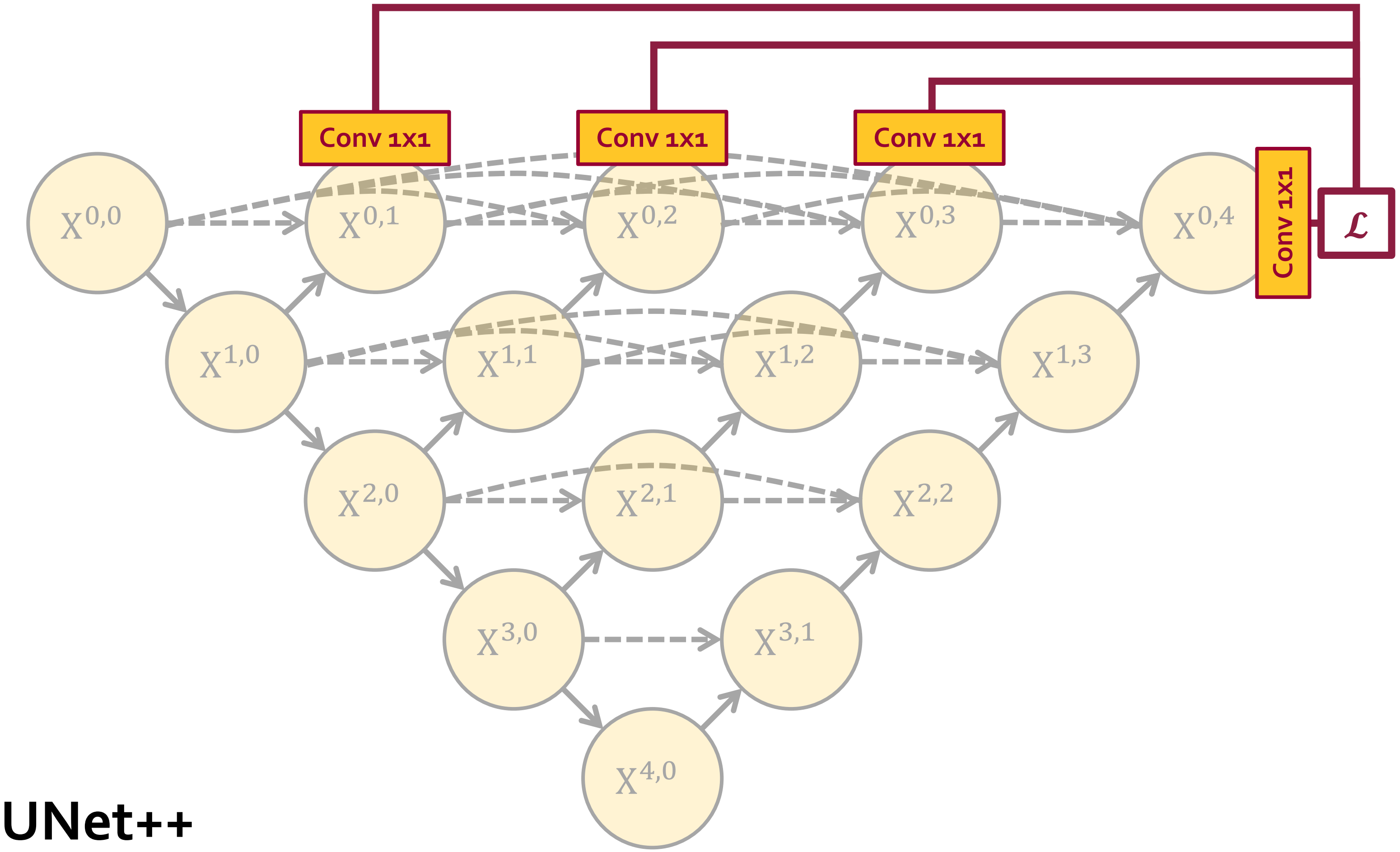


UNet++

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Deep supervision stabilizes model training and enables model pruning

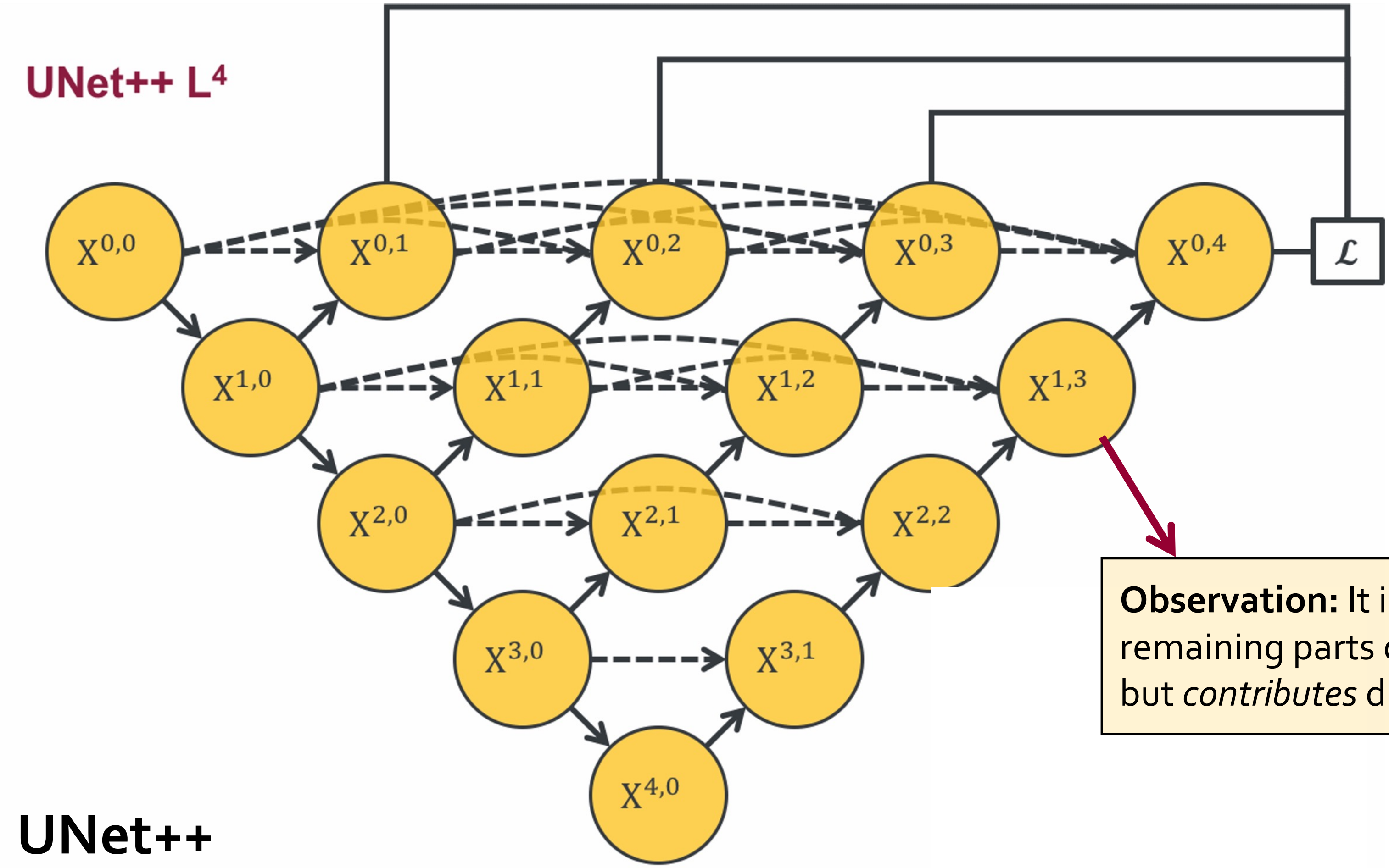


UNet++

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Approach: Deep supervision stabilizes model training and enables model pruning



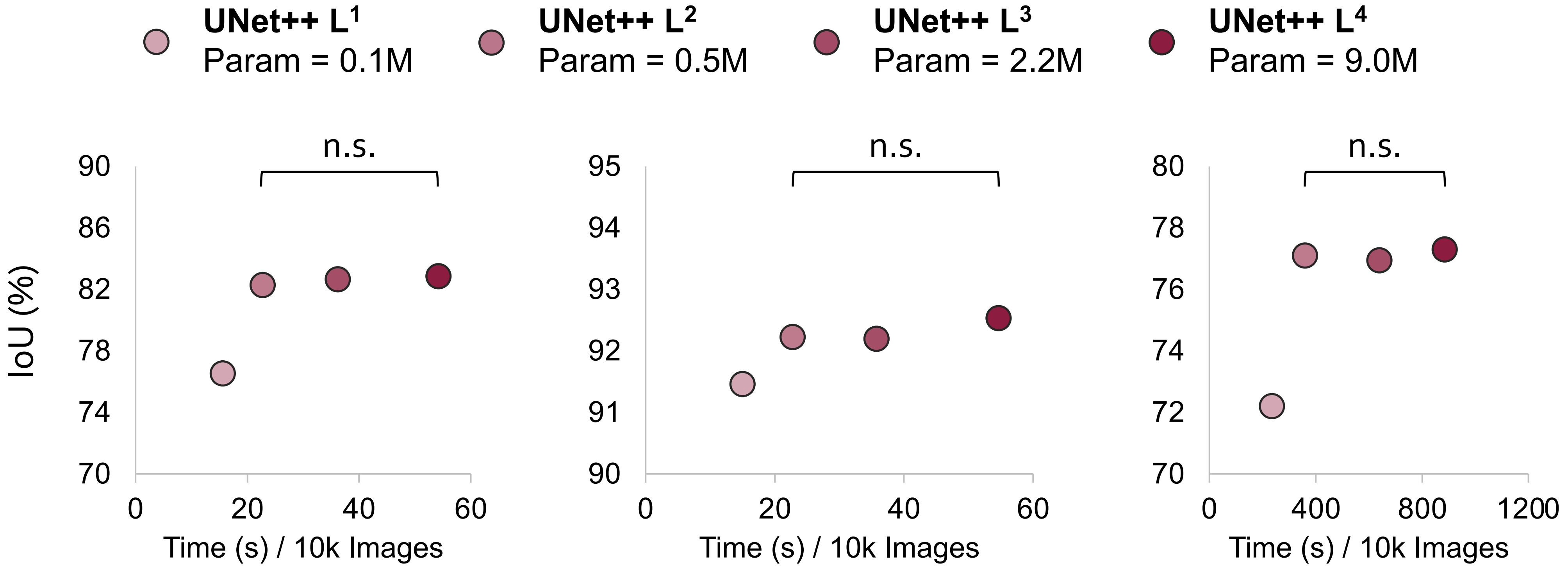
Observation: It is *independent* with the remaining parts during testing phase, but *contributes* during training phase.

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

Aim 2: Utilizing existing annotation effectively from advanced architecture

Contribution: UNet++ L^2 can process 2.4x more images than U-Net with similar accuracy.

Reducing the memory footprint by 18 times!



Liver segmentation

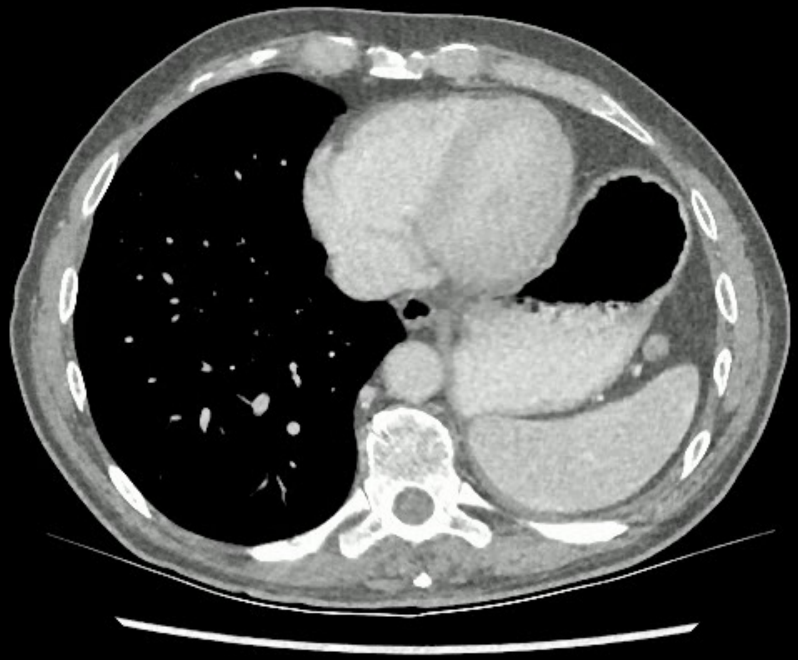
Cell nuclei segmentation

Lung nodule segmentation

1. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
2. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

UNet++ can precisely detect tumors in CT scans

CT



Ground Truth

UNet++ Prediction

Liver

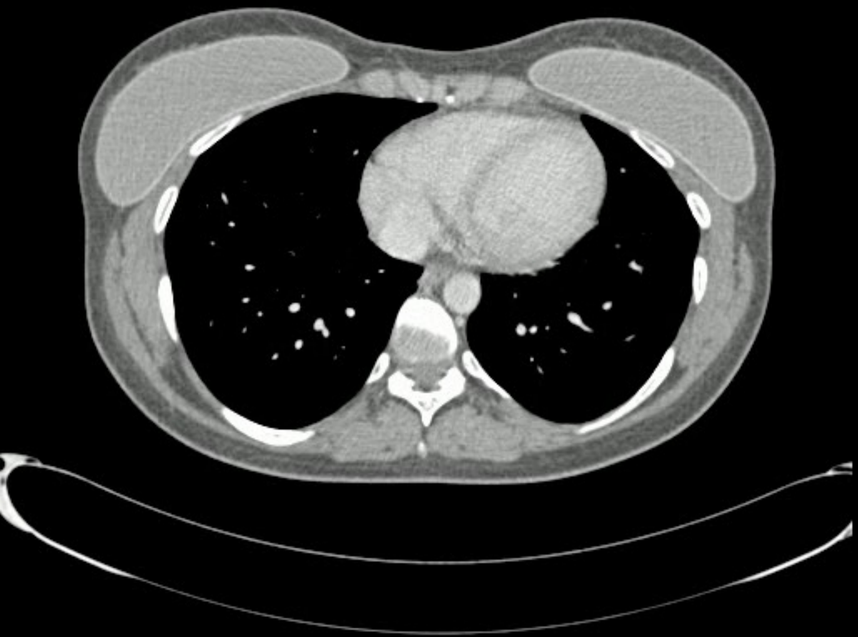
Liver tumor

- I thank S. Bajpai for this experiment; the code can be found at https://github.com/MrGiovanni/UNetPlusPlus/tree/master/pytorch

1. Bajpai, Shivam. "Pre-Trained Models for nnUNet." Master diss., Arizona State University, 2021.
2. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
3. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

UNet++ can detect very small tumors without predicting too many false positives

CT | Ground Truth | UNet++ Prediction



Liver

Liver tumor

• I thank S. Bajpai for this experiment; the code can be found at https://github.com/MrGiovanni/UNetPlusPlus/tree/master/pytorch

1. Bajpai, Shivam. "Pre-Trained Models for nnUNet." Master diss., Arizona State University, 2021.
2. Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." DLMIA, 2018.
3. Zhou, Zongwei, et al. "Unet++: Redesigning skip connections to exploit multiscale features in image segmentation." TMI, 2019.

43.9% → 58.1% (U-Net → UNet++)
Covid-19 segmentation (CT)
[Fan et al., IEEE TMI]

78.6% → 82.9% (U-Net →UNet++)
Fiber tracing (corneal confocal microscopy)
[Mou et al., MICCAI]

86.5% → 89.5% (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.

86.6% → 87.2% (U-Net → UNet++)
SegTHOR 2019 Challenge (CT)
[Zhang et al., IEEE TMI]

90.2% → 92.0% (U-Net → UNet++)
Optic Disc & Cup Segmentation (fundus image)
[Meng et al., MICCAI]

60.3% → 71.6% (U-Net → UNet++)
Ground-glass opacity segmentation (CT)
[Zheng et al., IEEE Access]

51.2% → 58.6% (U-Net → UNet++)
Esophagus segmentation (CT)
[Huang et al., IEEE Access]

63.7% → 66.3% (U-Net → UNet++)
Liver tumor segmentation (CT)
[Bajpai et al., Master Thesis]

90.7% → 91.6% (U-Net → UNet++)
Heart segmentation (MRI)
[Ji et al., MICCAI]

Unet++: A nested u-net architecture for medical image segmentation
Z Zhou, MM Rahman Siddiquee, N Tajbakhsh… - Deep learning in …, 2018 - Springer
In this paper, we present UNet++, a new, more powerful architecture for medical image
☆ Save 99 Cite Cited by 2807 Related articles All 15 versions

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. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*.

U.S. Patent

- US Patent 11,164,067, Systems, Methods, and Apparatuses for Implementing a Multi-resolution Neural Network for Use with Imaging Intensive Applications Including Medical Imaging

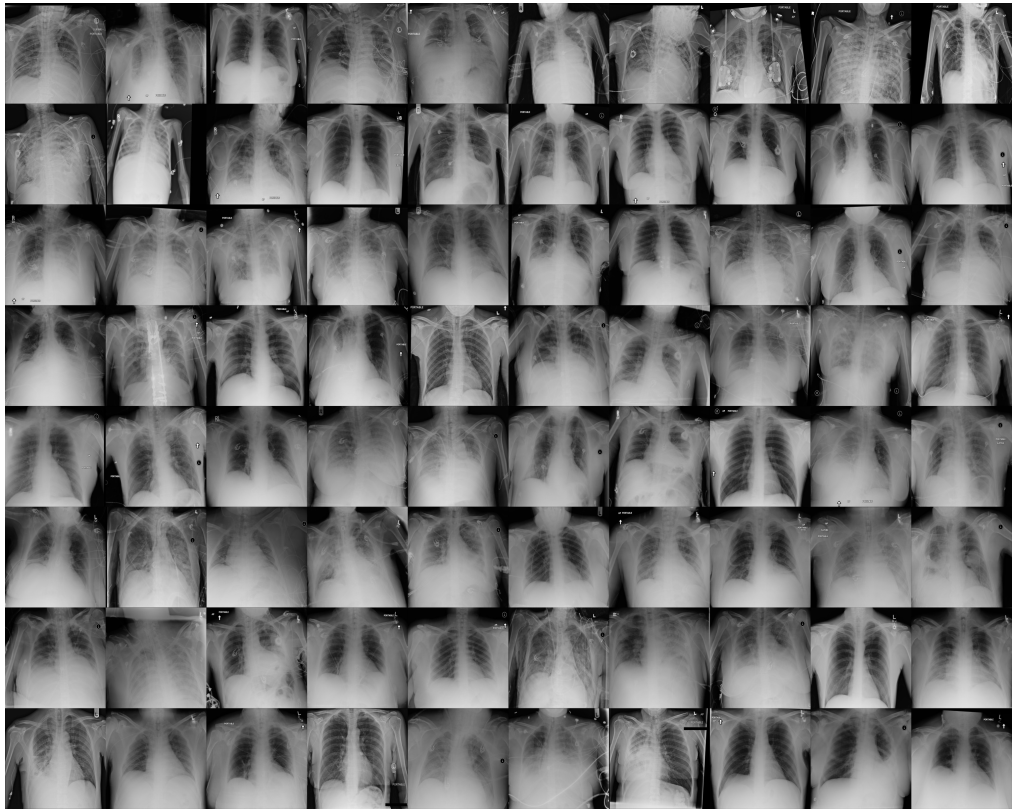
Intertwine the visual representation

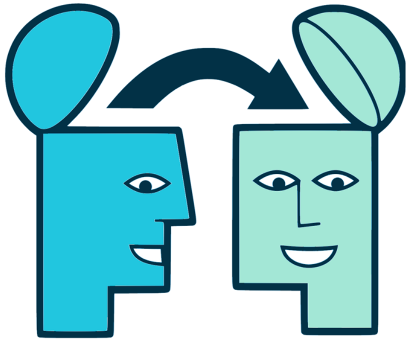
Clinical Impacts of Aim 2:

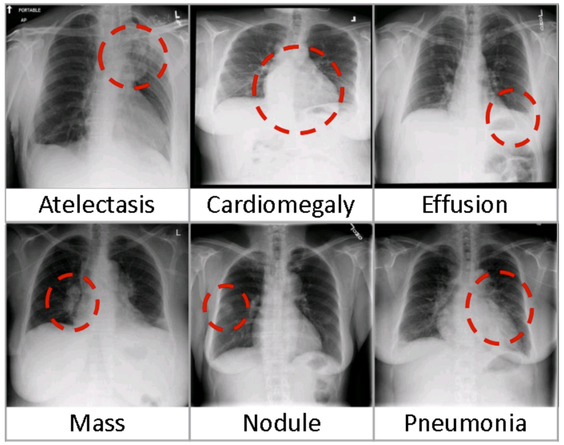
- Image segmentation can help compute clinically more accurate and desirable *imaging bio-markers* or *precision measurement*.
- Model pruning has the potential to exert important impact on deploying CAD systems to *mobile devices* and *ordinary desktop/laptop PCs* in clinical practice.

Aim 3: Extracting generic knowledge directly from unannotated images

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







Introduction

Objective

Aim 1

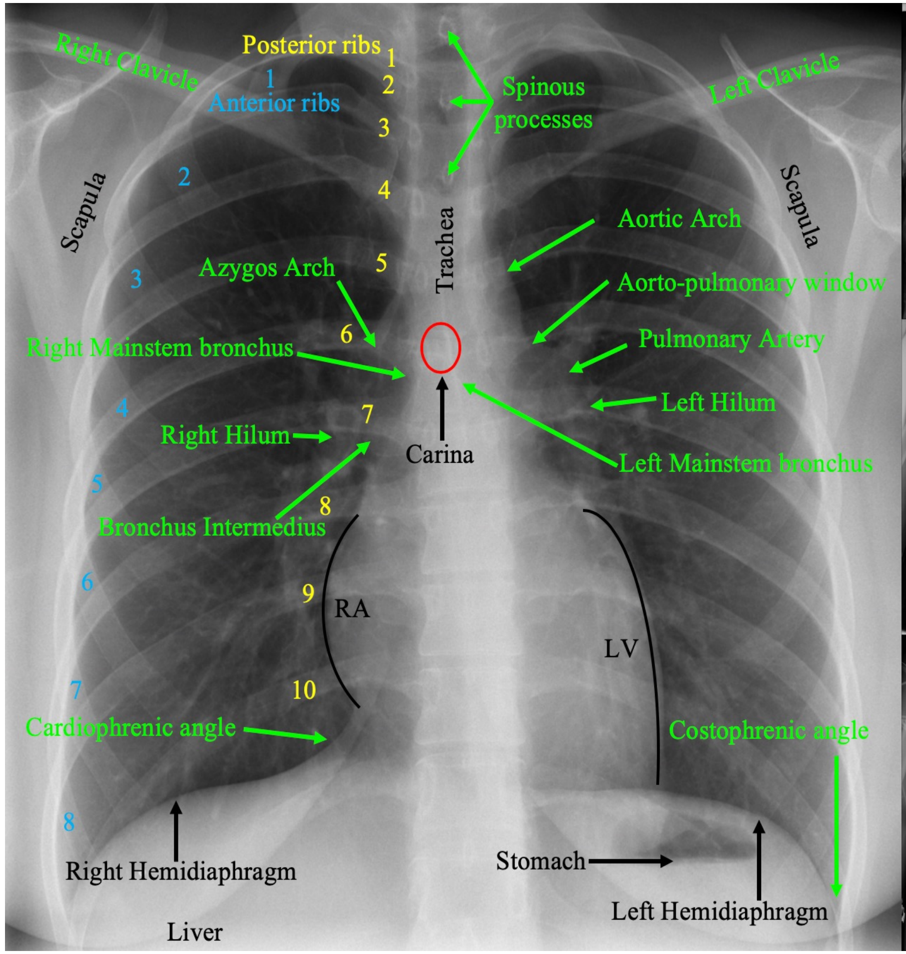
Aim 2

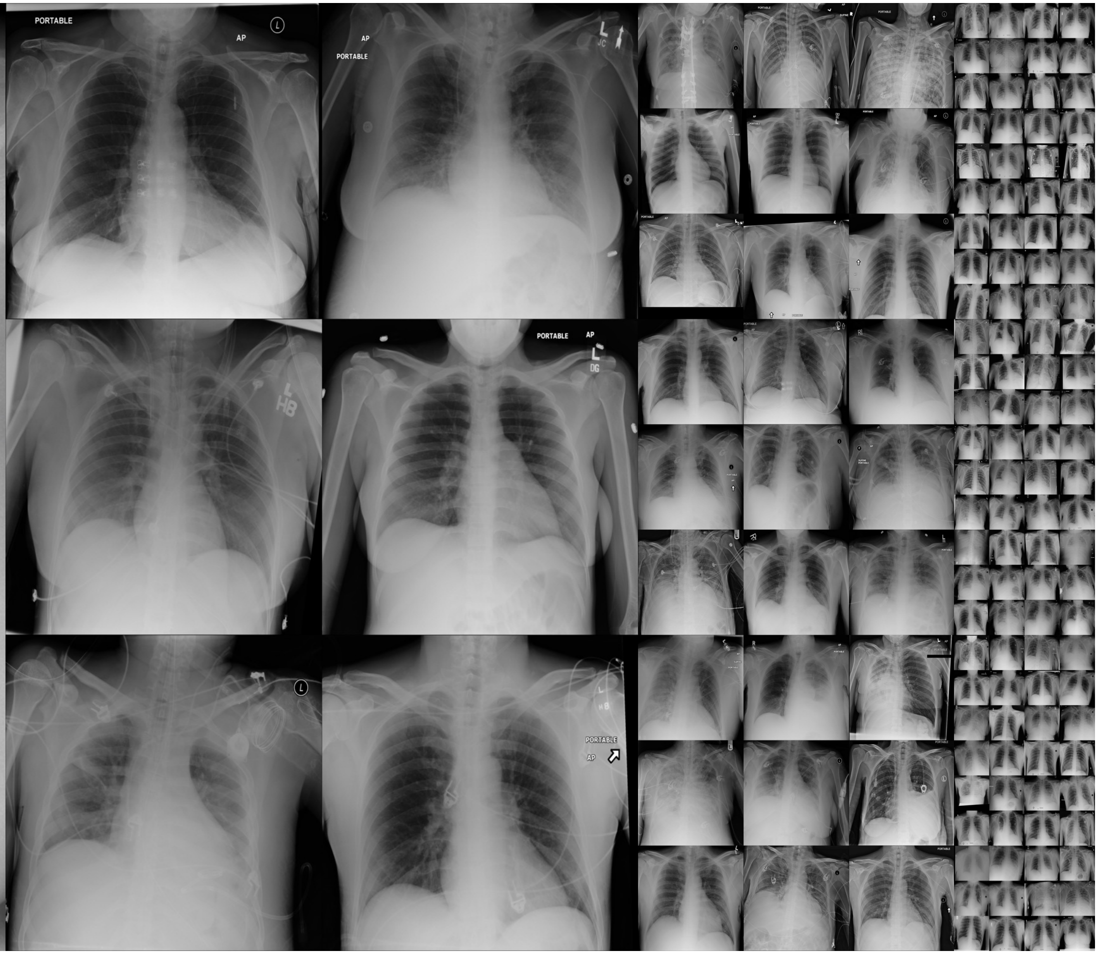
Aim 3

Summary

Aim 3: Extracting generic knowledge directly from unannotated images

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

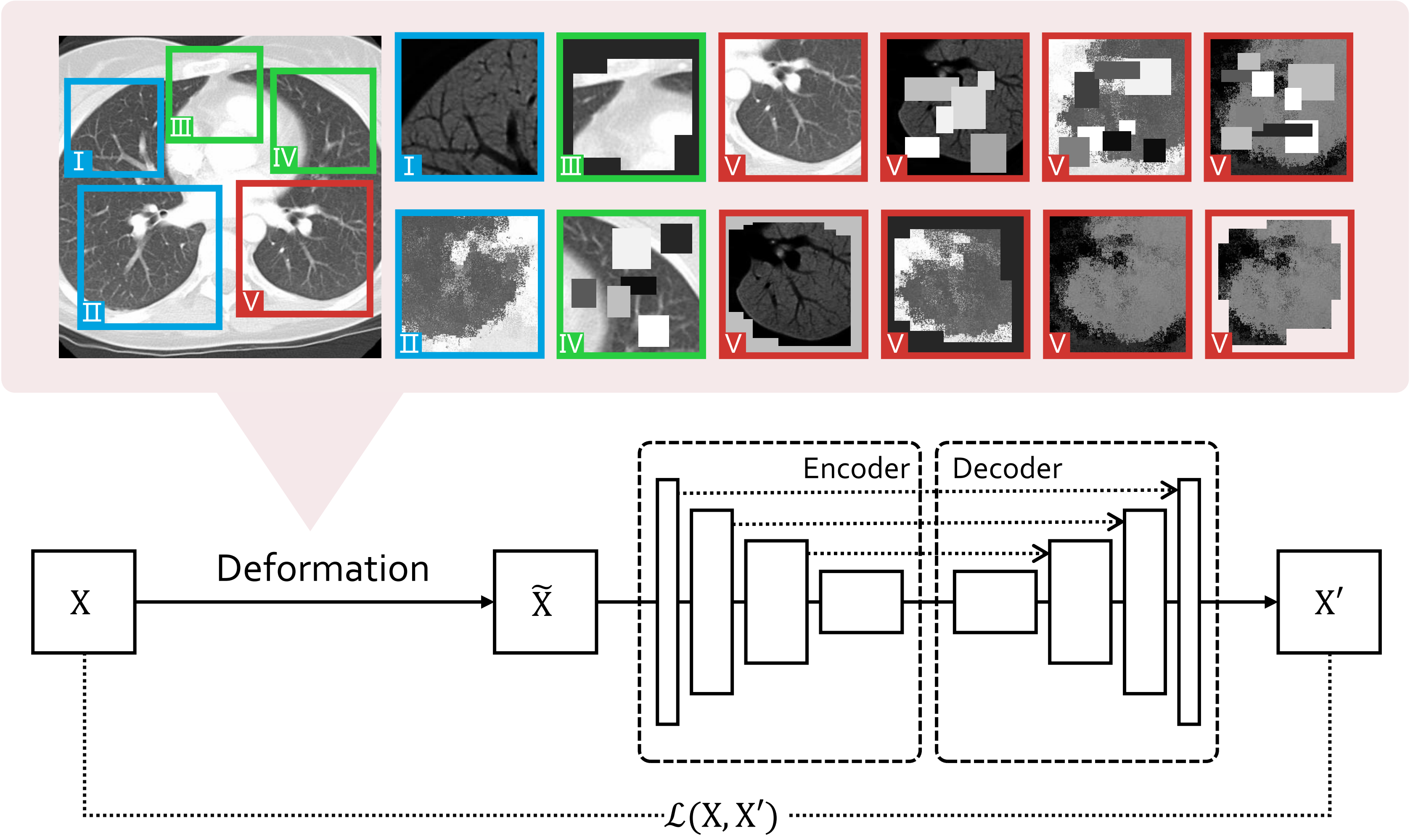




- I thank M. R. Hosseinzadeh Taher and M. Gotway for producing the left figure.

Aim 3: Extracting generic knowledge directly from unannotated images

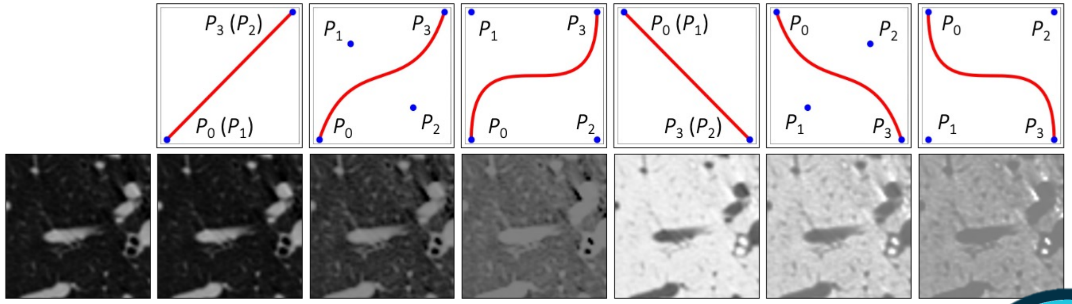
Approach: Image restoration task helps model learn image representation

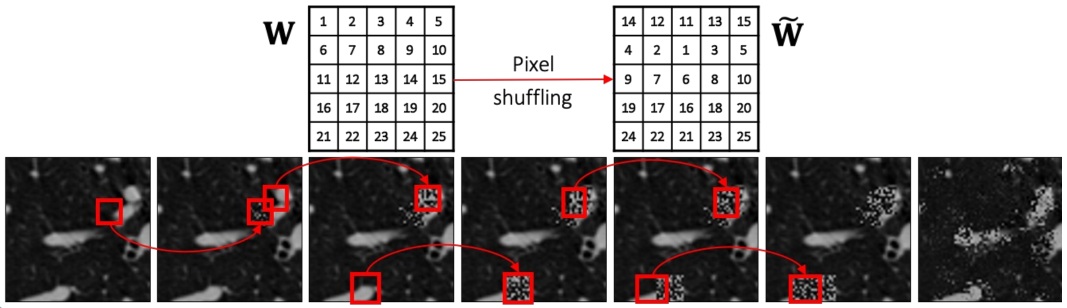


Introduction
Objective
Aim 1
Aim 2
Aim 3
Summary

Aim 3: Extracting generic knowledge directly from unannotated images

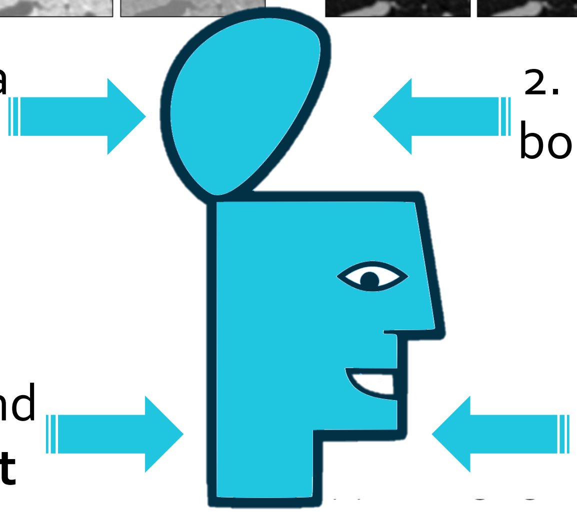
Approach: Learning from multiple perspectives leads to robust models





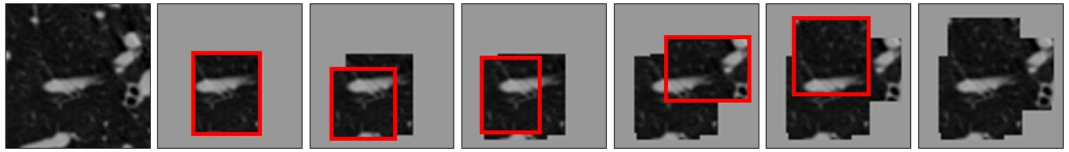
1. Learning organ appearance via **non-linear transformation**

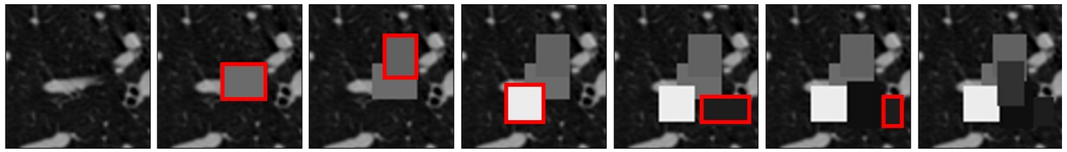
2. Learning organ texture and local boundaries via **local pixel shuffling**



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

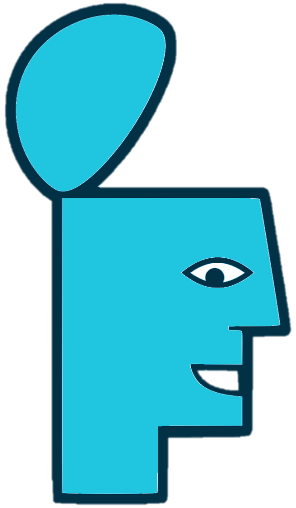
4. Learning local continuities of organs via **inner-cutout**

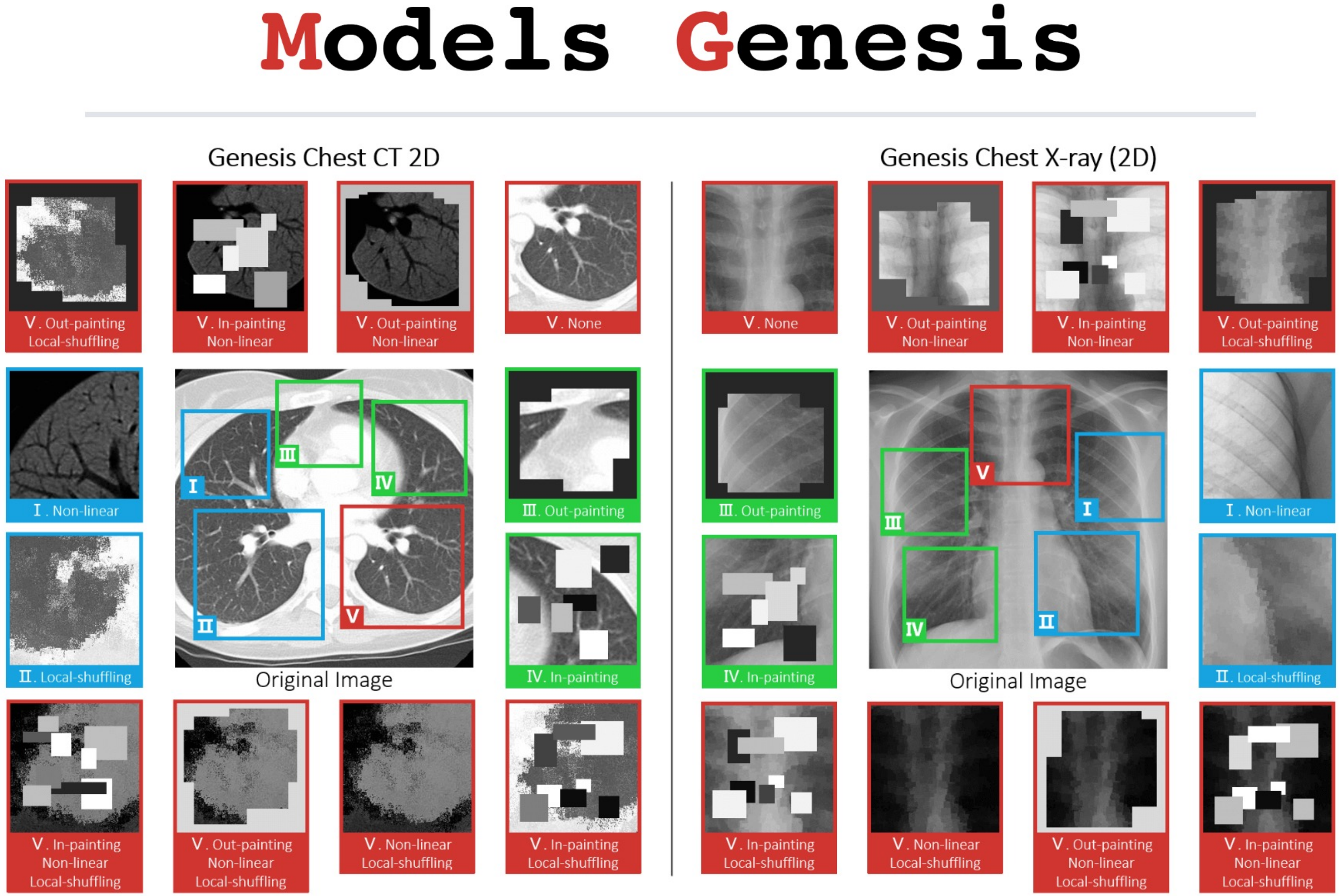




Aim 3: Extracting generic knowledge directly from unannotated images

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





1. Zhou, Zongwei, et al. "Models genesis: Generic autodidactic models for 3d medical image analysis." MICCAI, 2019.
2. Zhou, Zongwei, et al. "Models genesis." Medical image analysis, 2021.

Aim 3: Extracting generic knowledge directly from unannotated images

Contribution: Among the *first* endeavor for publicly available pre-trained 3D models

deepmind / **kinetics-i3d** 240,000 annotated videos

Watch 52 | Unstar 1.1k | Fork 355

Code | Issues 68 | Pull requests 4 | Actions | Projects 0 | Wiki | Security | Insights

Convolutional neural network model for video classification trained on the Kinetics dataset.

NifTK / **NiftyNet** 90 annotated subjects

Used by 19 | Watch 89 | Unstar 1.1k | Fork 365

Code | Issues 95 | Pull requests 3 | Actions | Projects 3 | Wiki | Security | Insights

An open-source convolutional neural networks platform for research in medical image analysis and image-guided therapy
http://niftynet.io

Tencent / **MedicalNet** 1,638 annotated subjects

Unwatch 43 | Unstar 790 | Fork 210

Code | Issues 28 | Pull requests 2 | Actions | Projects 0 | Wiki | Security | Insights

Many studies have shown that the performance on deep learning is significantly affected by volume of training data. The MedicalNet project provides a series of 3D-ResNet pre-trained models and relative code.

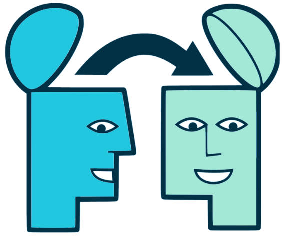
1. Carreira, Joao, et al. "Quo vadis, action recognition? a new model and the kinetics dataset." CVPR, 2017.
2. Gibson, Eli, et al. "NiftyNet: a deep-learning platform for medical imaging." Computer methods and programs in biomedicine, 2018.
3. Chen, Sihong, et al. "Med3d: Transfer learning for 3d medical image analysis." arXiv:1904.00625, 2019.

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

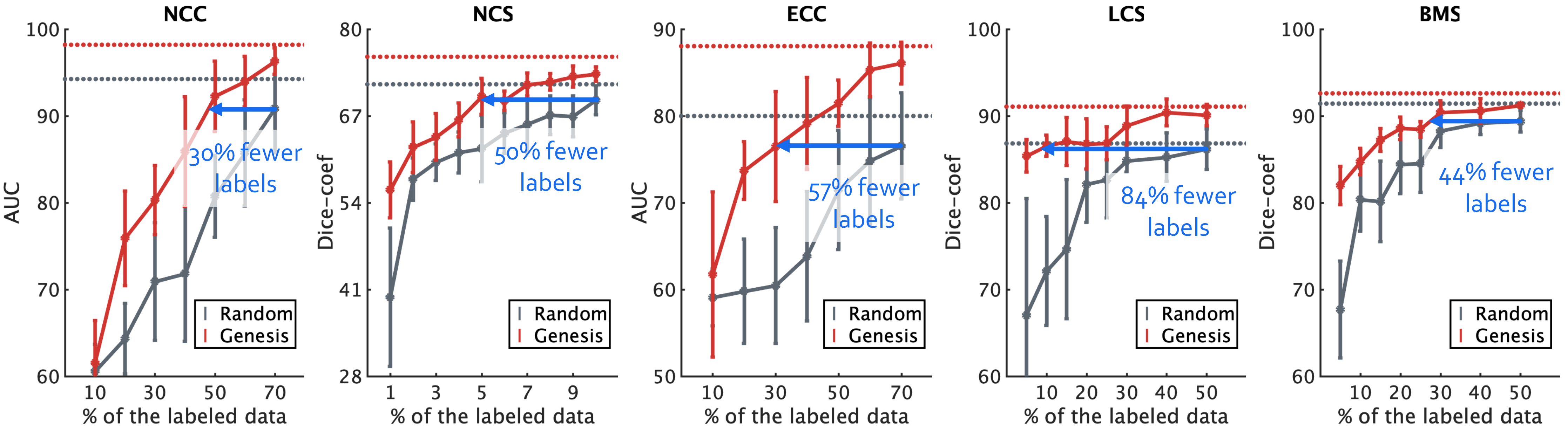


• I thank Z. Guo, F. Haghighi, M. R. Hosseinzadeh Taher, M. M. Rahman Siddiquee, and P. Zhang for their supports in experiments.

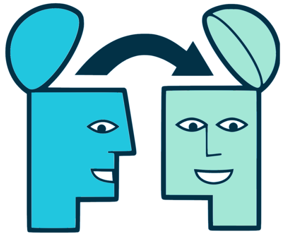
1. Zhou, Zongwei, et al. "Models genesis: Generic autodidactic models for 3d medical image analysis." MICCAI, 2019.
2. Zhou, Zongwei, et al. "Models genesis." Medical image analysis, 2021.

Aim 3: Extracting generic knowledge directly from unannotated images

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



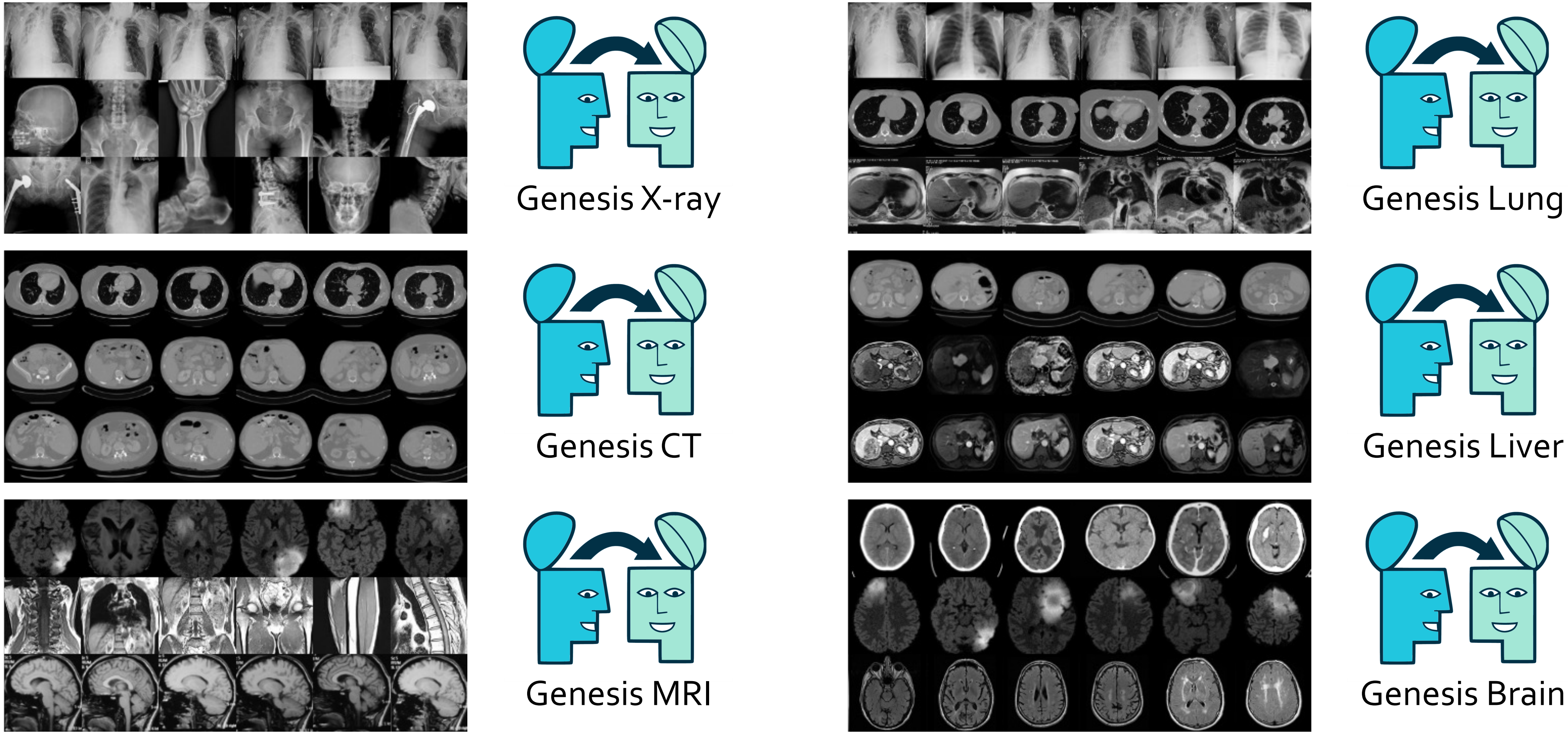
[1]NCC	Lung nodule false positive reduction in CT images
[2]NCS	Lung nodule segmentation in CT images
[3]ECC	Pulmonary embolism false positive reduction in CT images
[4]LCS	Liver segmentation in CT images
[5]BMS	Brain tumor segmentation in MR images

Genesis Chest CT  Target models

1. Zhou, Zongwei, et al. "Models genesis: Generic autodidactic models for 3d medical image analysis." MICCAI, 2019.
2. Zhou, Zongwei, et al. "Models genesis." Medical image analysis, 2021.

Aim 3: Extracting generic knowledge directly from unannotated images

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



69.0% → 73.9% (Scratch → MG) Prostate segmentation (MRI) [Taleb et al., arXiv:1912.05396, 2019]	**83.1% → 88.3% (Scratch → MG)** Lymph node classification (histology) [Xu et al., BIBM, 2020]	**72.3% → 85.8% (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.

67.0% → 74.5% (Scratch → MG) Blood cavity segmentation (MRI) [Zhang et al., arXiv:2010.06107, 2020]	**67.8% → 69.3% (Scratch → MG)** 13 organ segmentation (CT) [Xie et al., arXiv:2011.12640, 2020]	**90.0% → 95.0% (Scratch → MG)** Liver segmentation (CT&MRI) [Taleb et al., arXiv:1912.05396, 2019]
77.5% → 92.5% (Scratch → MG) COVID-19 classification (CT) [Sun et al., arXiv:2012.06457, 2020]	**76.0% → 77.5% (Scratch → MG)** Liver tumor segmentation (CT) [Bajpai et al., Master Thesis, 2021]	**74.0% → 79.3% (Scratch → MG)** Alzheimer's disease classification (MRI) [Zhang et al., arXiv:2010.06107, 2020]

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, Young Scientist Publication Impact Award Finalist*
- 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:

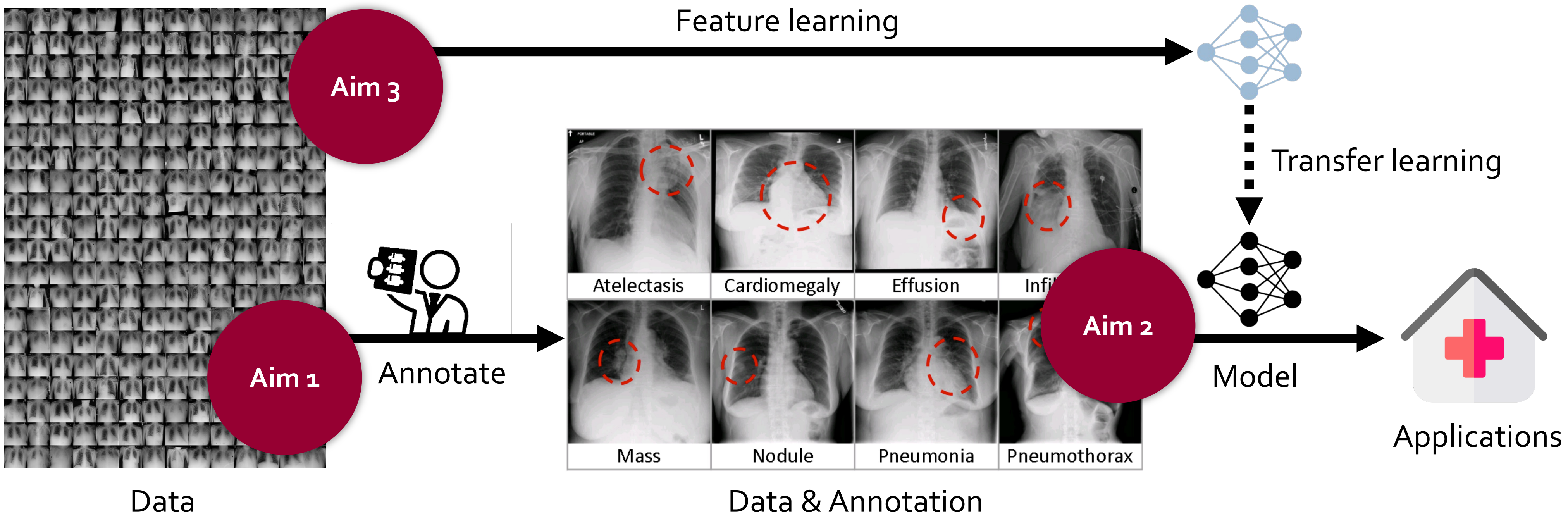
- Instead of building a model from scratch (demanding numerous data and label acquisition), a *smaller dataset* can be used to efficiently fine-tune the existing model.
- Generic pre-trained models can serve as a *primary source of transfer learning* for many medical imaging applications, leading to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

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



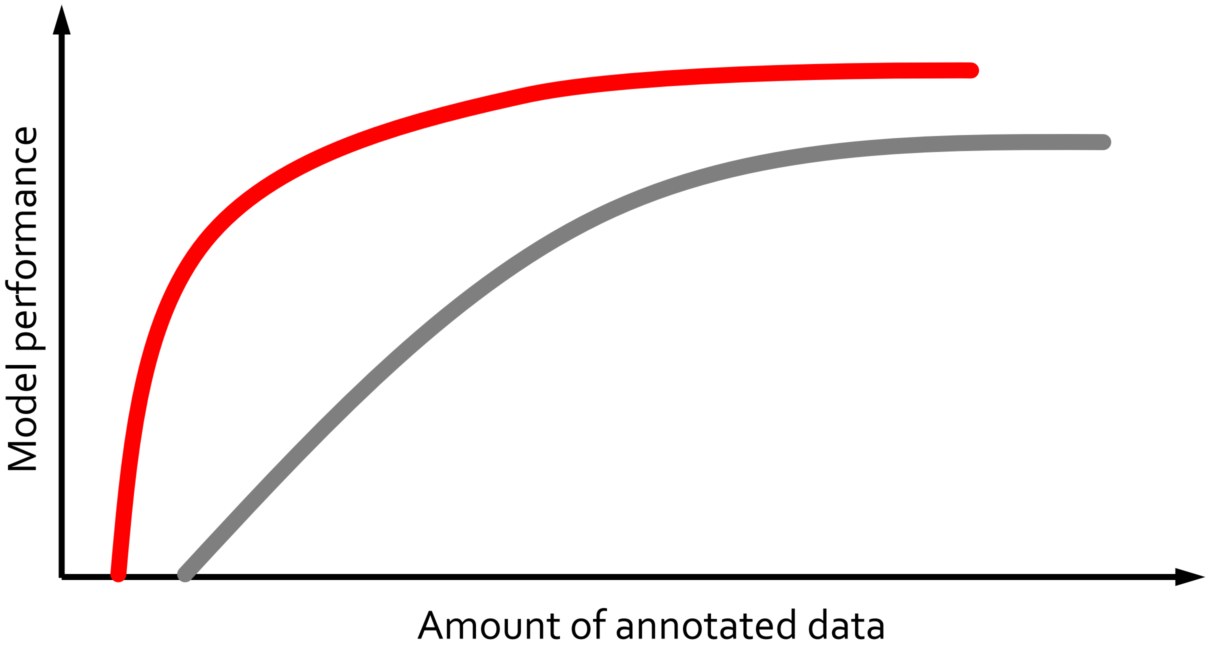
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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.



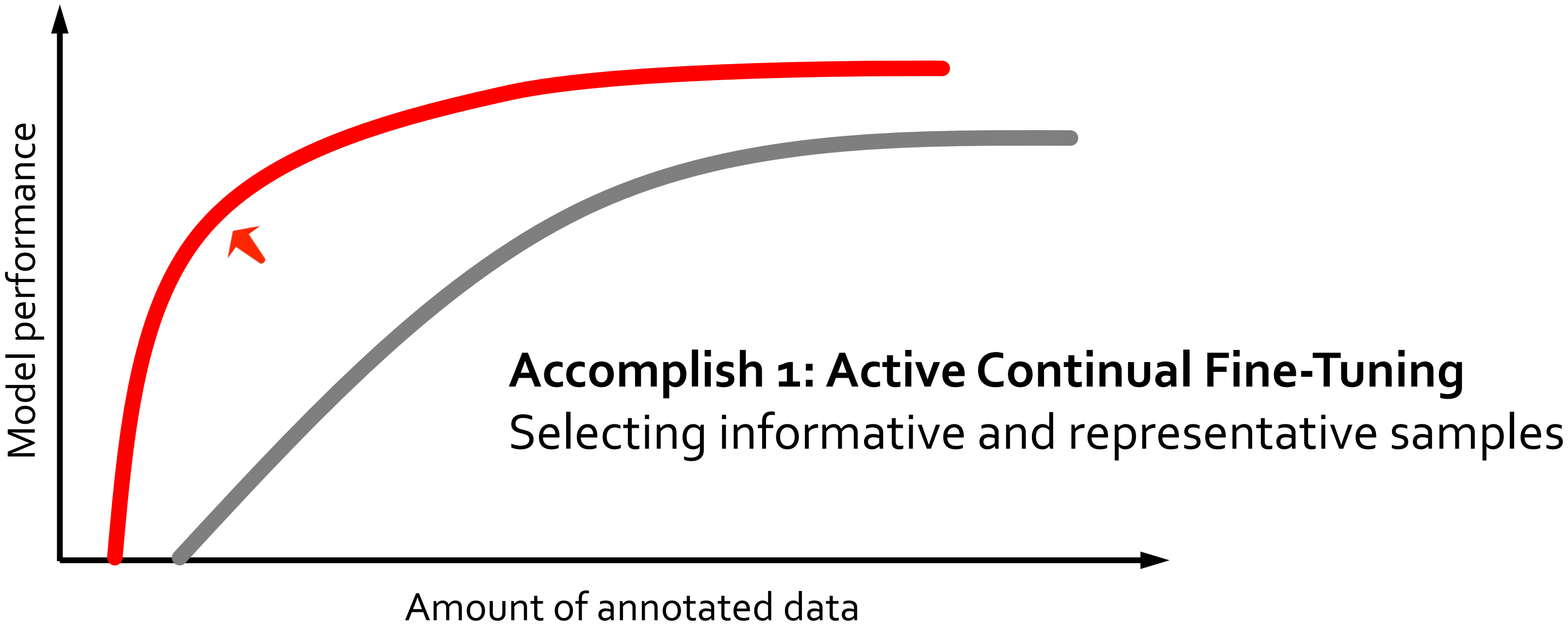
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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!**



Introduction
Objective
Aim 1
Aim 2
Aim 3
Summary

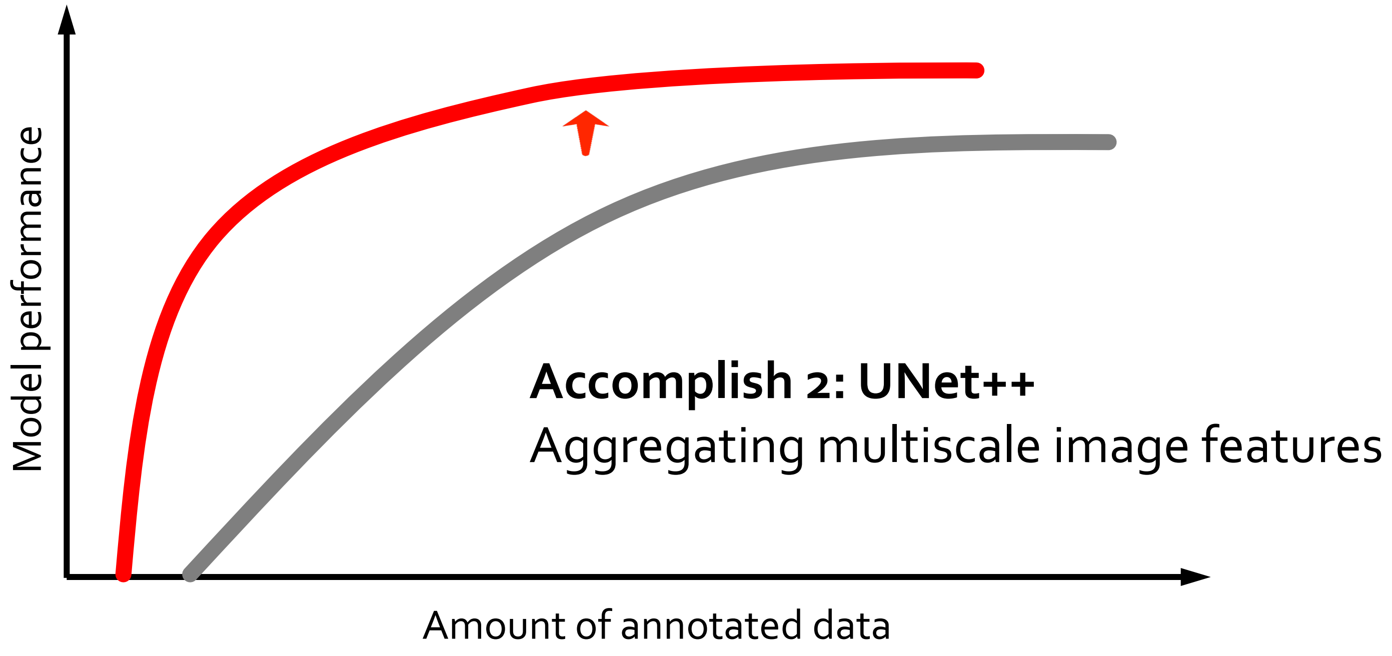
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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!**



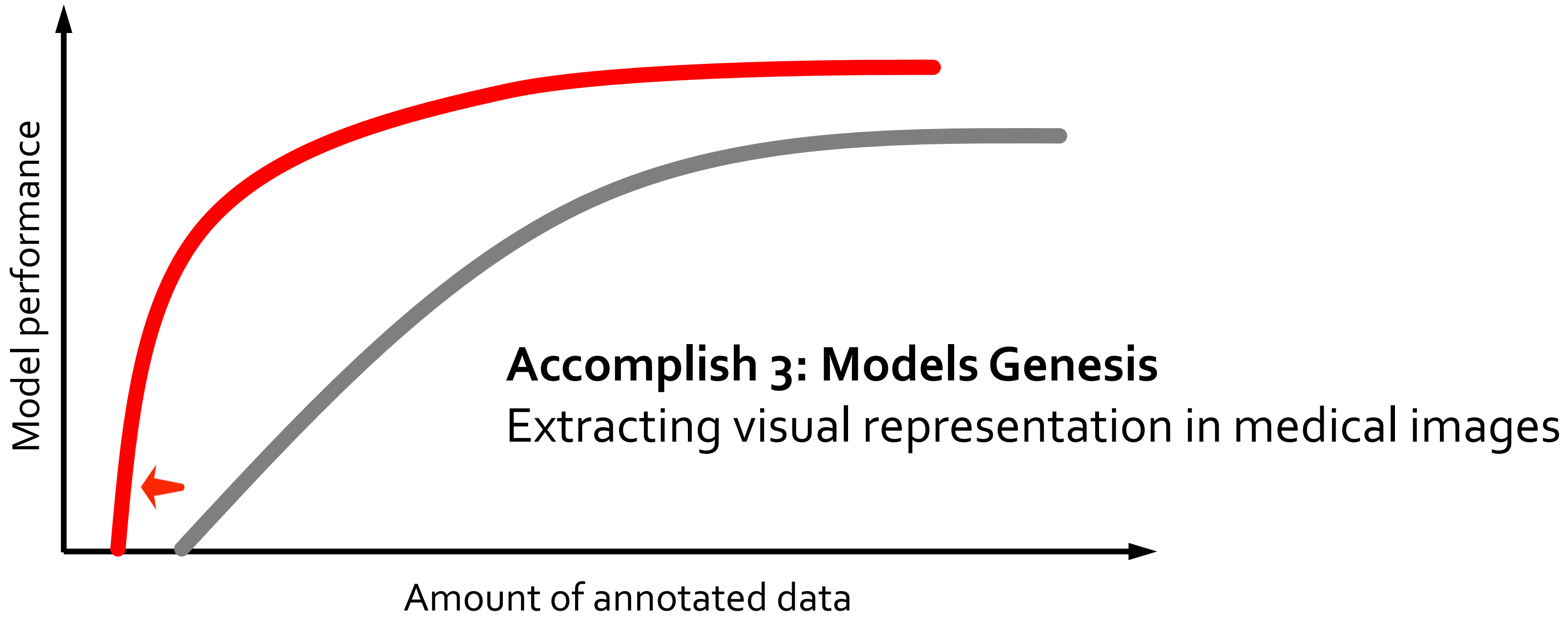
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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!**



Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

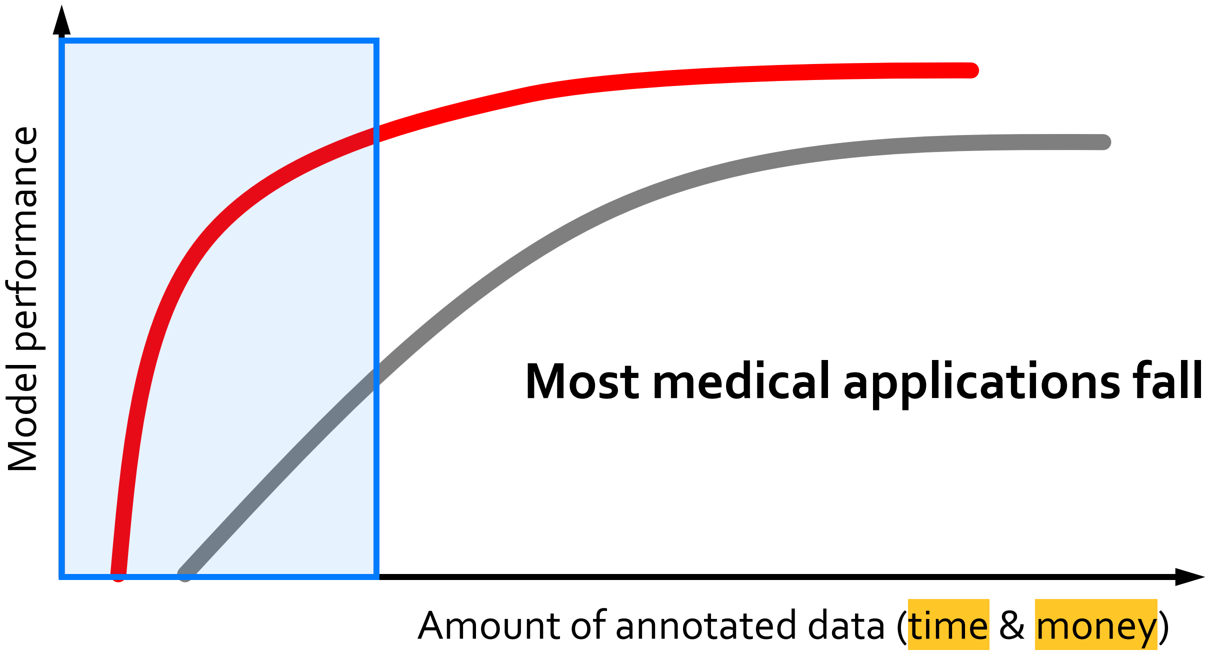
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

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

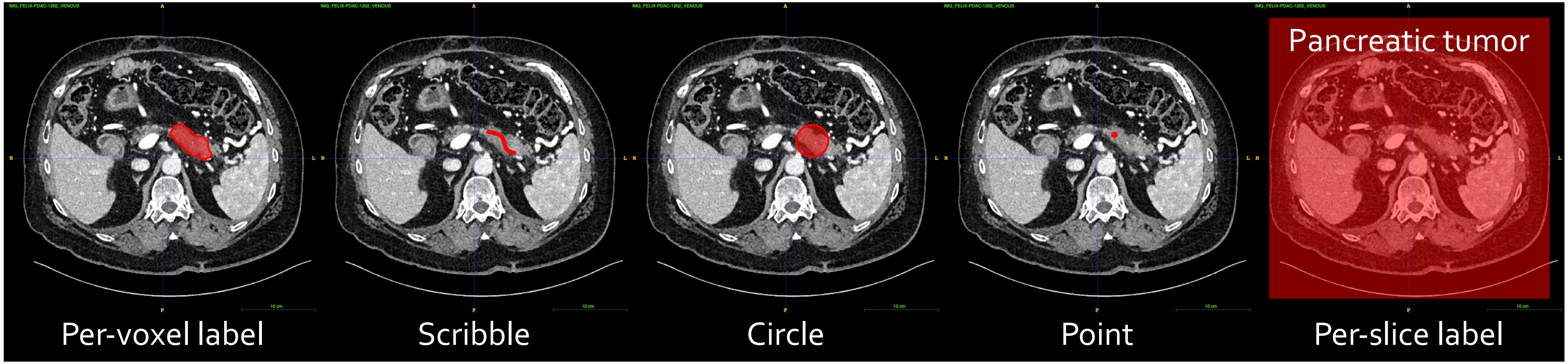
Conclusion: With a small part of the dataset annotated, we can deliver deep models that approximate or even outperform those that require annotating the entire dataset. **Yes, we can!**



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

Aims 1-3 (PhD): Reducing *detailed* annotation (per-voxel labels)

Aim Pro: Investigating *weak* annotation (scribble, circle, point, per-slice labels, etc.)



15 human-year to create
Sensitivity = 92.4%
Specificity = 93.0%

→

x10 faster to create
Sensitivity = 92.5%
Specificity = 91.3%

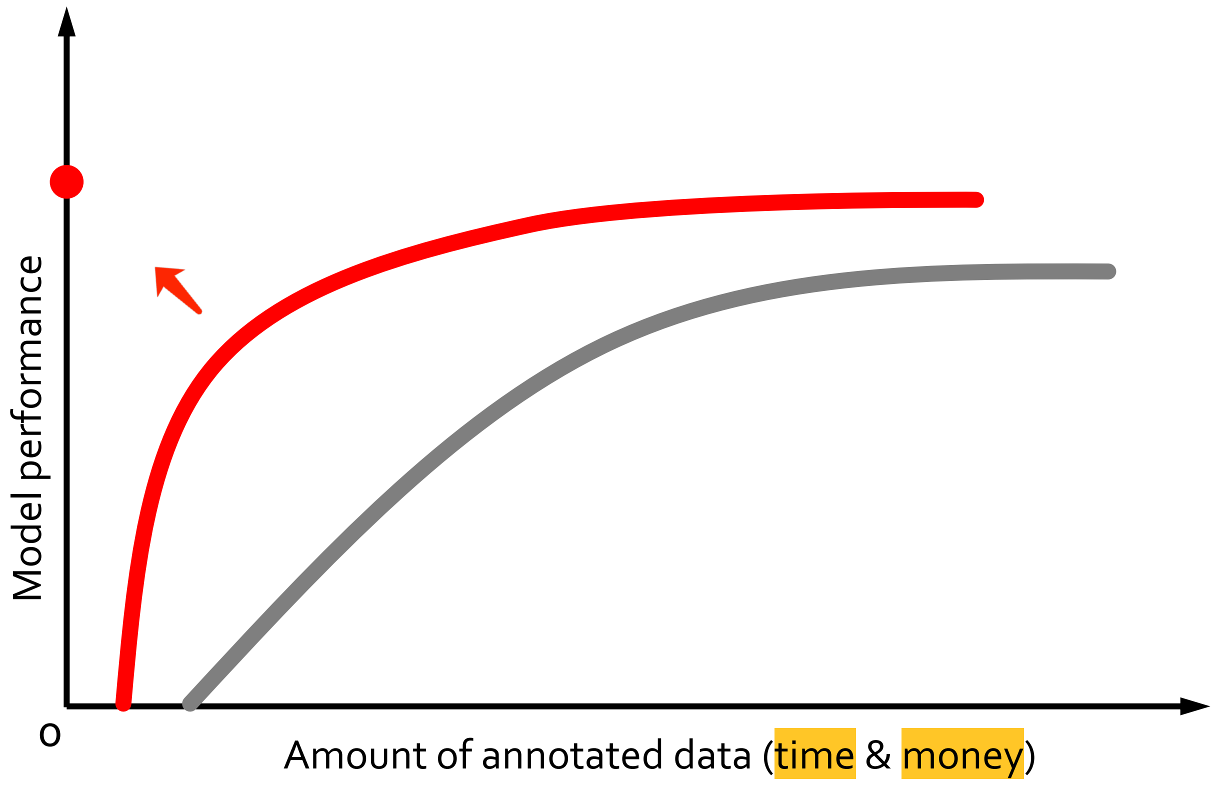
1. Yang, Shuojue et al. "Pancreatic Ductal Adenocarcinoma (PDAC) Detection Using Per-Slice Annotation." Radiological Society of North America (RSNA), 2022. (Oral Presentation)

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

Aims 1-3 (PhD): Reducing *detailed* annotation (per-voxel labels)

Aim Pro: Investigating *weak* annotation (per-slice, points, scribbles, boxes, etc.)

Aim Pro Max: Exploring *ultra-weak* annotation (radiology reports, synthetic data, etc.)



Medical professionals with over 6-year experience cannot tell which are real and which are synthetic tumors

Can you?

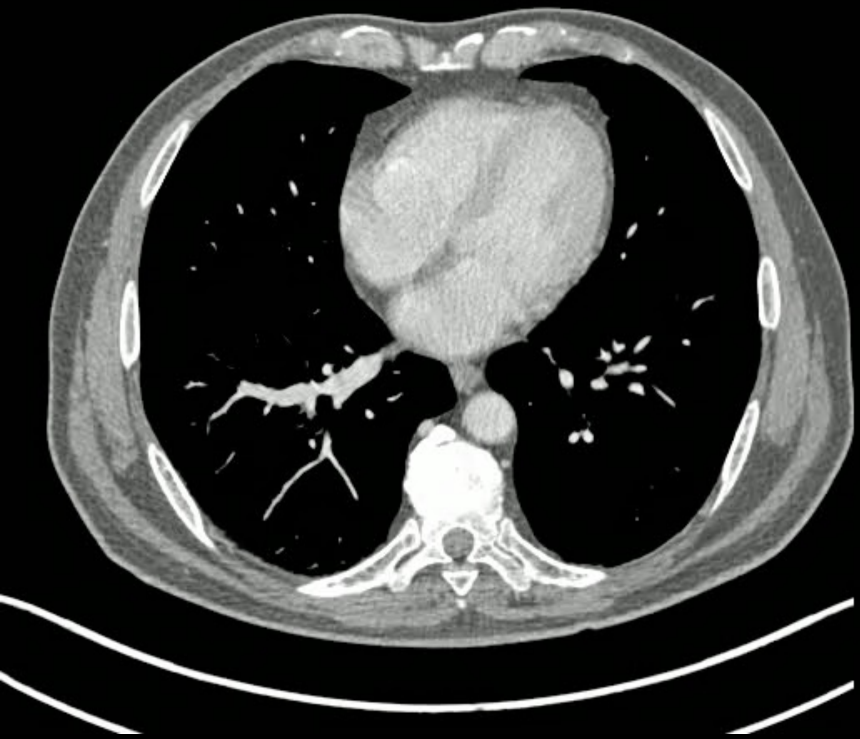
1. Hu, Qixin et al. "Synthetic Tumors Make AI Segment Tumors Better". NeurIPS Workshop on Medical Imaging Meets NeurIPS, 2022.

Training AI on synthetic tumors performs almost as well as training it on real tumors.

CT

UNet++ prediction
trained on real tumors
with per-voxel annotation

UNet++ prediction
trained on synthetic tumors
with no annotation



Liver

Liver tumor

1. Hu, Qixin et al. "Synthetic Tumors Make AI Segment Tumors Better". NeurIPS Workshop on Medical Imaging Meets NeurIPS, 2022.

Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

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

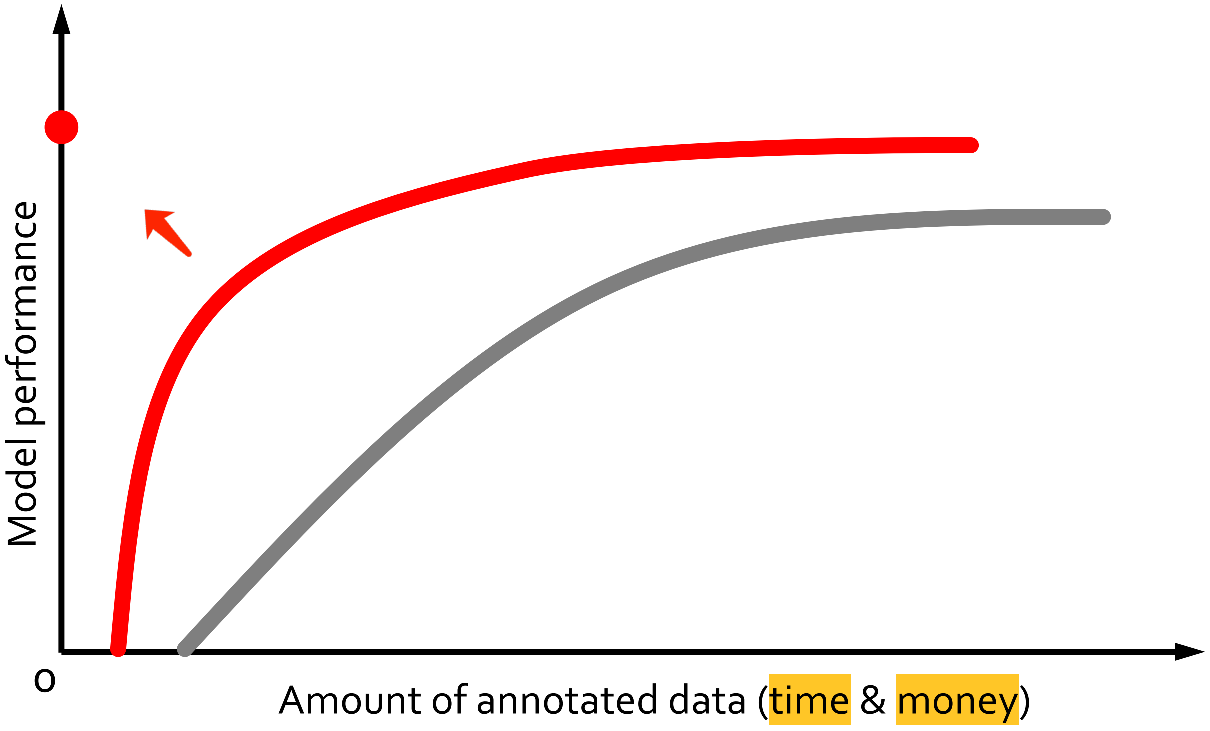
Aims 1-3 (PhD): Reducing *detailed* annotation (per-voxel labels)

Aim Pro: Investigating *weak* annotation (per-slice, points, scribbles, boxes, etc.)

Aim Pro Max: Exploring *ultra-weak* annotation (radiology reports, synthetic data, etc.)

I am actively looking for a faculty position to make this happen :)

- *Imaging Informatics, Computer Vision, Healthcare*



Introduction

Objective

Aim 1

Aim 2

Aim 3

Summary

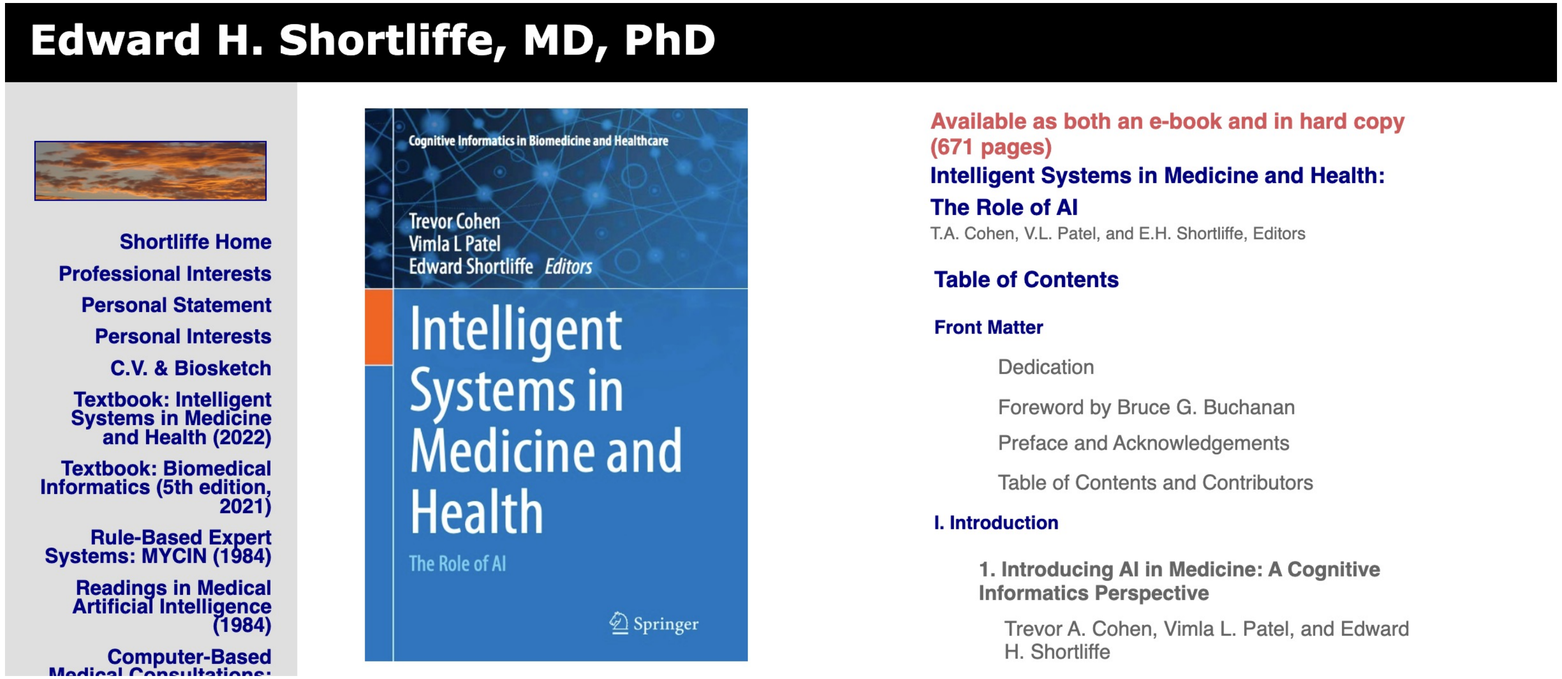
Goal: Minimize manual annotation efforts for rapid, precise computer-aided diagnosis systems

Aims 1-3 (PhD): Reducing *detailed* annotation (per-voxel labels)

Aim Pro: Investigating *weak* annotation (per-slice, points, scribbles, boxes, etc.)

Aim Pro Max: Exploring *ultra-weak* annotation (radiology reports, synthetic data, etc.)

Interpreting Medical Images: A book chapter that overviews AI in medical image interpretation



1. Zhou, Zongwei et al. "Interpreting Medical Images." In Cognitive Informatics in Biomedicine and Healthcare. Intelligent Systems in Medicine and Health: The Role of AI. T. Cohen, V. Patel and E. Shortliffe (eds.). Springer Nature, 2022.

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.
- Alan L. Yuille, Ph.D.

Funding for research program supported by

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

Credit: Baymax!

Towards Annotation-Efficient Deep Learning for Computer-Aided Diagnosis

Zongwei Zhou

Postdoc, Department of Computer Science

Johns Hopkins University, Baltimore, MD

P: 1-(480)738-2575 | E: zzhou82@jh.edu

People
Who
Don't
Give Up