SAFE AND DATA-EFFICIENT MEDICAL IMAGE ANALYSIS IN THE AGE OF HIGHLY-PERFORMANT AI

Dr. Christian Baumgartner

OVERARCHING VISION OF OUR RESEARCH GROUP

OUR PHD STUDENTS

PhD Students

Jaivardhan Kapoor

PhD student

Probabilistic inference in spatio-temporal models

Nikolas Morshuis

PhD student

Robust MRI analysis and reconstruction using physics-informed networks

Paul Fischer

PhD student

Uncertainty quantification in medical prediction systems

Stefano Woerner

PhD student

Few-shot and meta-learning for learning from few data

Susu Sun

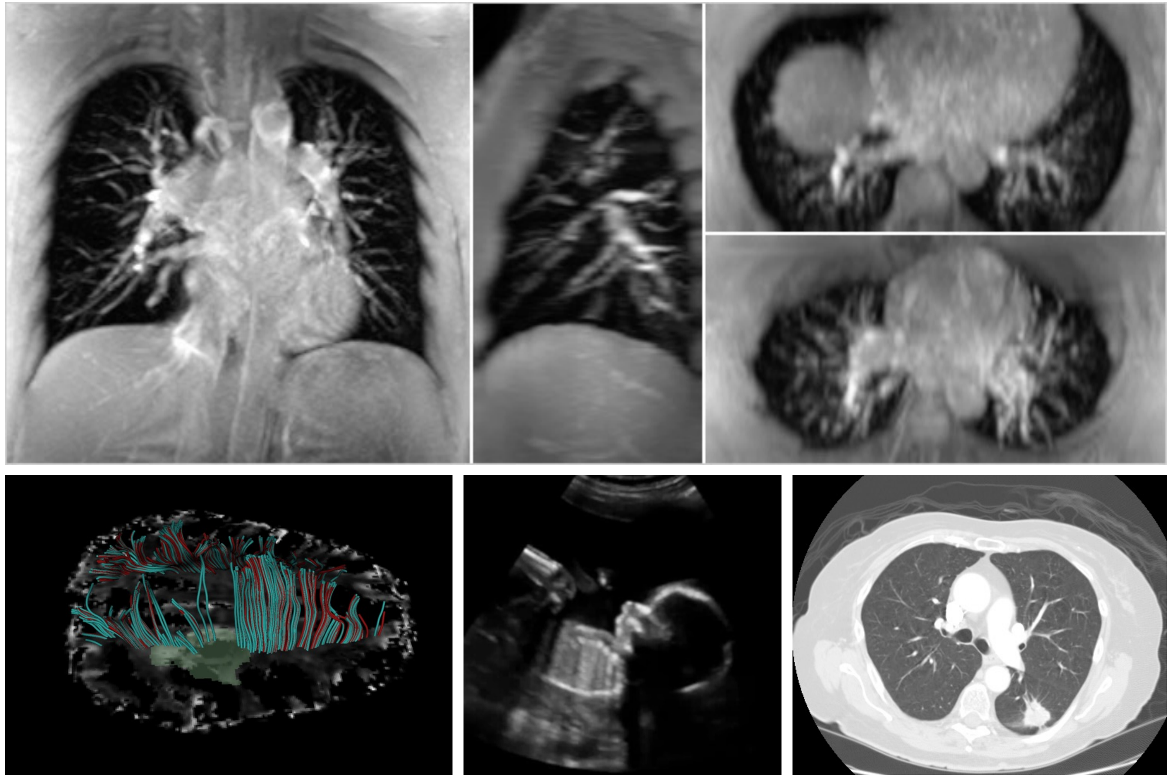
PhD student

Interpretable Machine Learning, Incorporation of Prior Domain Knowledge into Deep Neural Networks

MEDICAL IMAGING

Other Diagnostic Information

Genetic Information



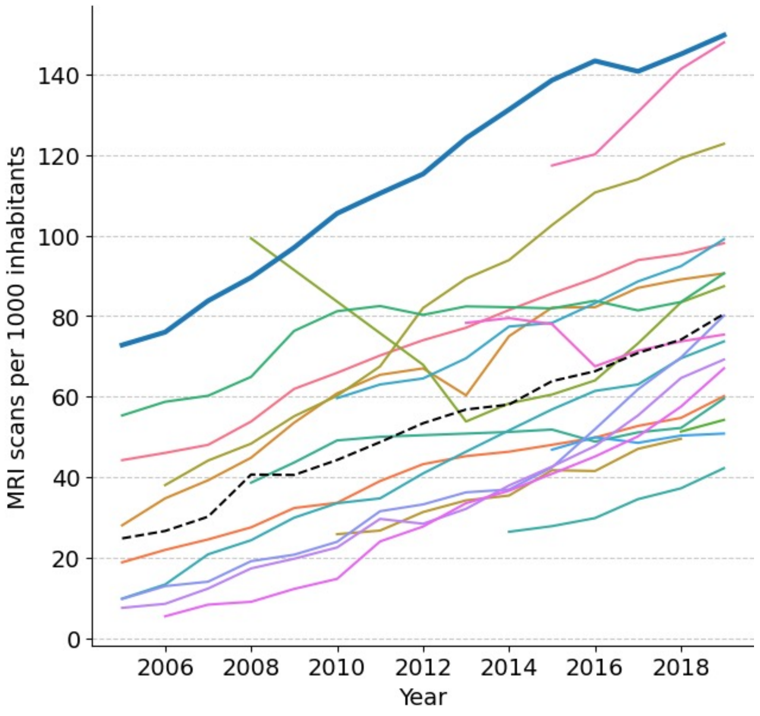
Diagnoses

Treatment Planning & Guidance

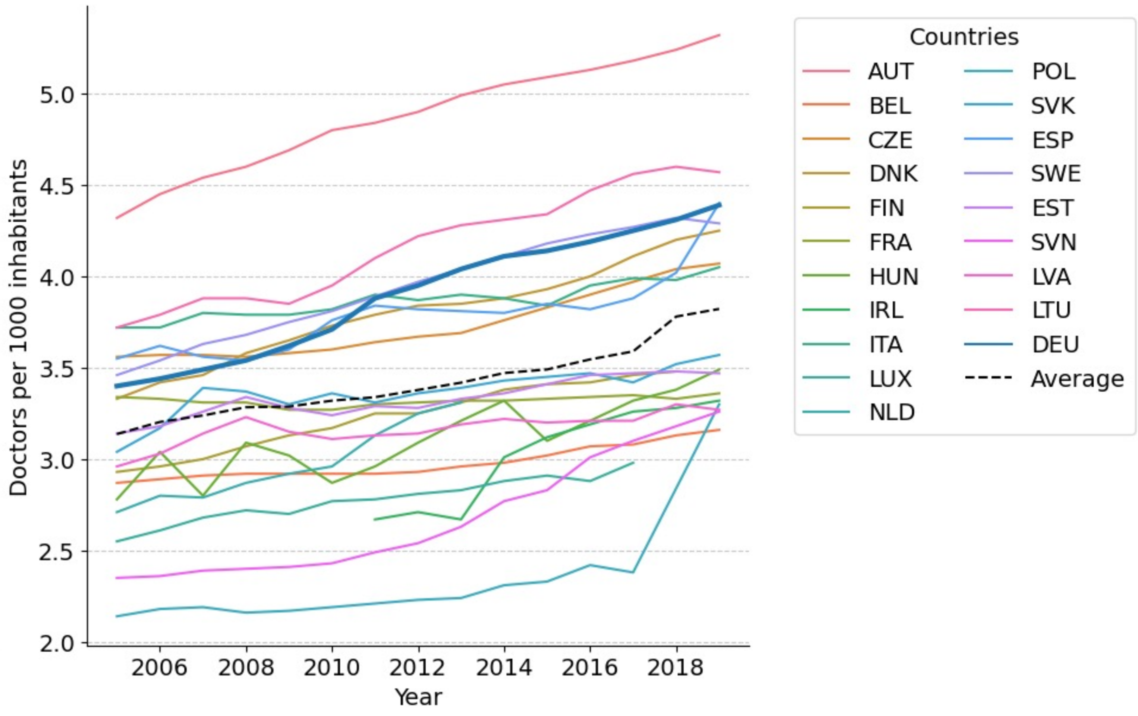
Biomarkers Extraction

SOLUTION TO MANY PROBLEMS
BUT ALSO CAUSE OF MANY PROBLEMS

Increase in magnetic resonance (MR) scans



Increase in doctors



2009-2019 on average: **+85% MR scans** **+17% doctors**

The volume of information doctors need to process is growing year by year!

Source: OECD

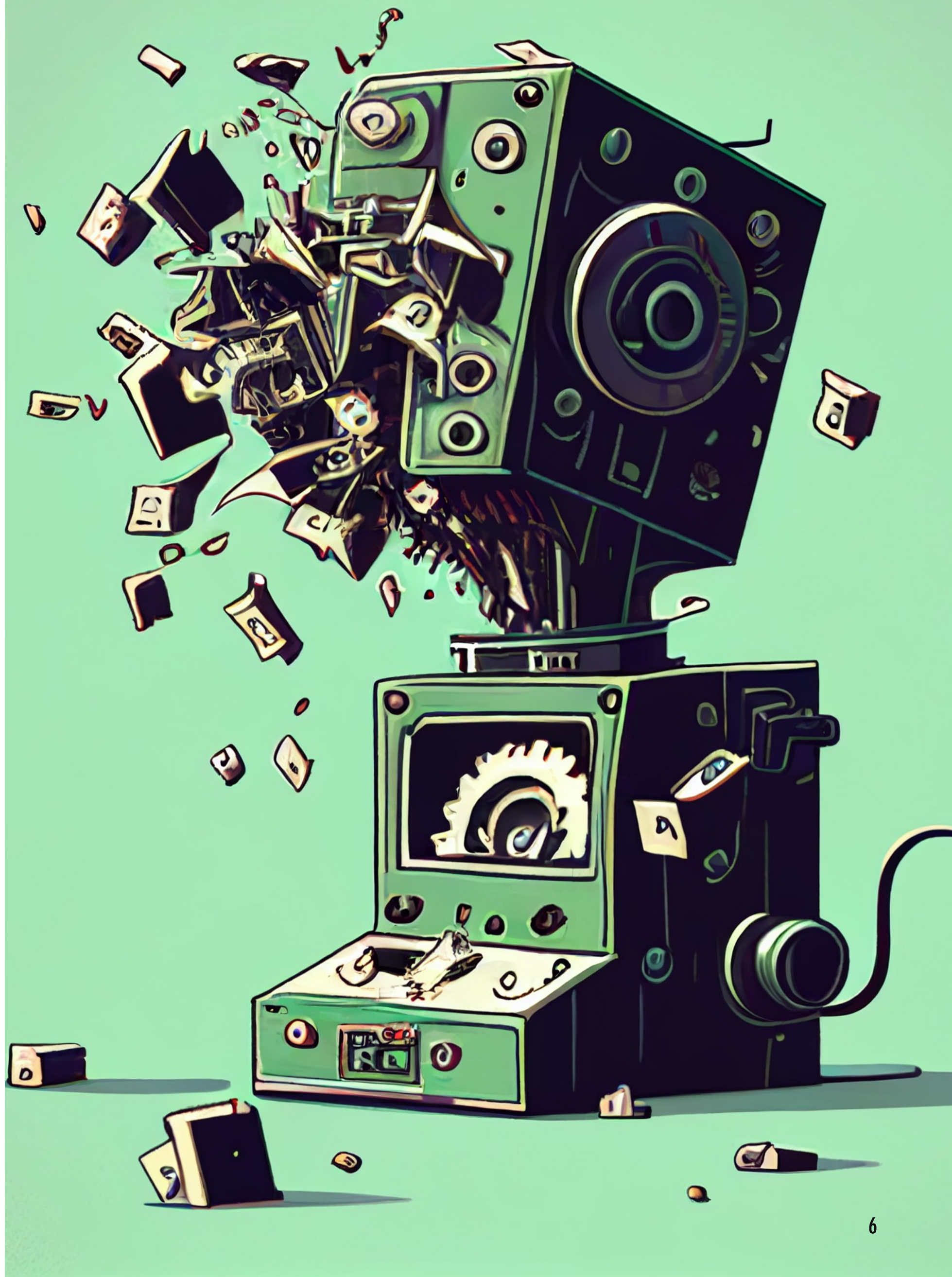
IMPACT ON CLINICAL WORKFORCE

UK radiology workforce census 2021 [1]:

- 98% of clinical directors in the UK are worried about workforce **morale, stress** and **burnout** in their departments
- 81% of clinical directors cited **worries about patient safety**

ACR Survey 2016 [2]:

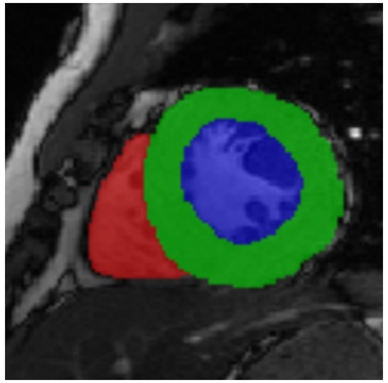
- 77% of practice leaders say **burnout is a significant problem.**



[1] UK Radiology workforce 2021 census
[2] Harolds et al., "Burnout of Radiologists: Frequency, Risk Factors, and Remedies: A Report of the ACR Commission on Human Resources", Journal of the American College of Radiology (2016)

HOW CAN ML-BASED CLINICAL DECISION SUPPORT SYSTEMS HELP?

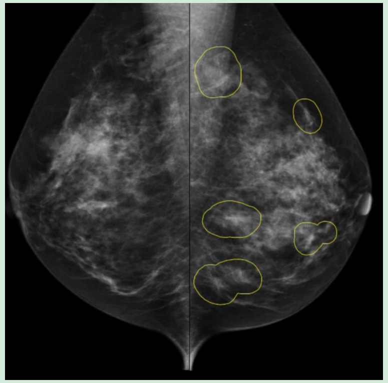
Workflow support



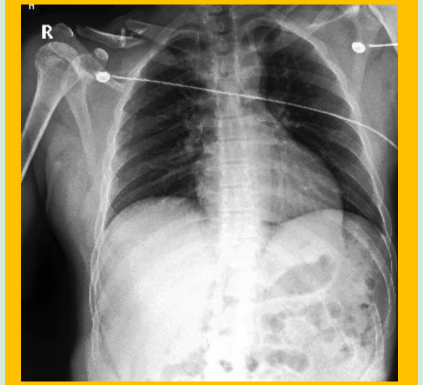
Contouring [1]
(e.g. volumetry)

Scan plane
Detection [2]
(e.g. guidance in US)

Expert system



Computer-aided
Detection
(e.g. screening)



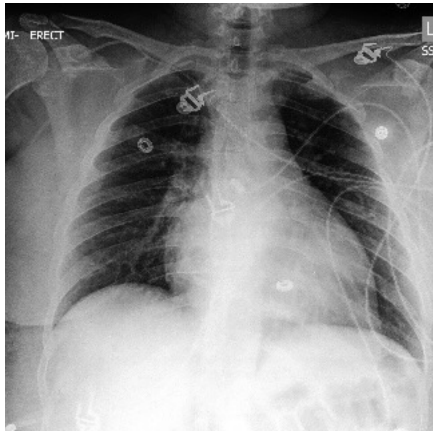
Computer-aided
Diagnosis
(e.g. prioritization,
second reader)

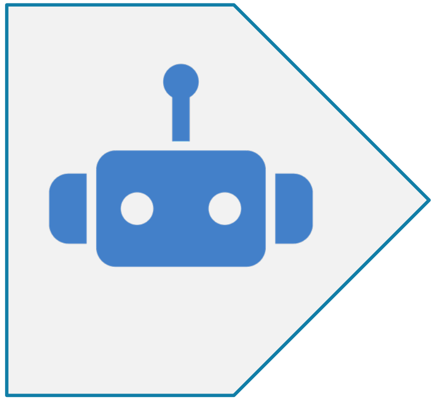
ML tools based on neural networks have been shown to perform on a par with clinicians

[1] Baumgartner et al., “An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation”, Proc. STACOM (2017)

[2] Baumgartner et al., “SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound”, IEEE Transactions in Medical Imaging (2017)

HOW WILL ML-BASED EXPERT SYSTEMS FIT INTO CLINICAL WORKFLOWS?





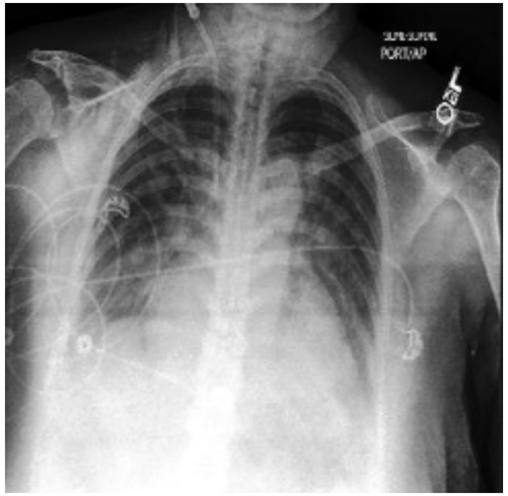
"This patient has early stage lung cancer"

We need explainable ML!

We need robust ML!

We need uncertainty-aware ML!

WILL MACHINES EVER BE ABLE TO PERFORM TASKS HUMAN CANNOT?



Black

Asian

	Area under the receiver operating characteristics curve
Race detection in radiology imaging	
Chest x-ray (internal validation)*	
MXR (Resnet34, Densenet121)	0·97, 0·94
CXP (Resnet 34)	0·98
EMX (Resnet34, Densenet121, EfficientNet-B0)	0·98, 0·97, 0·99

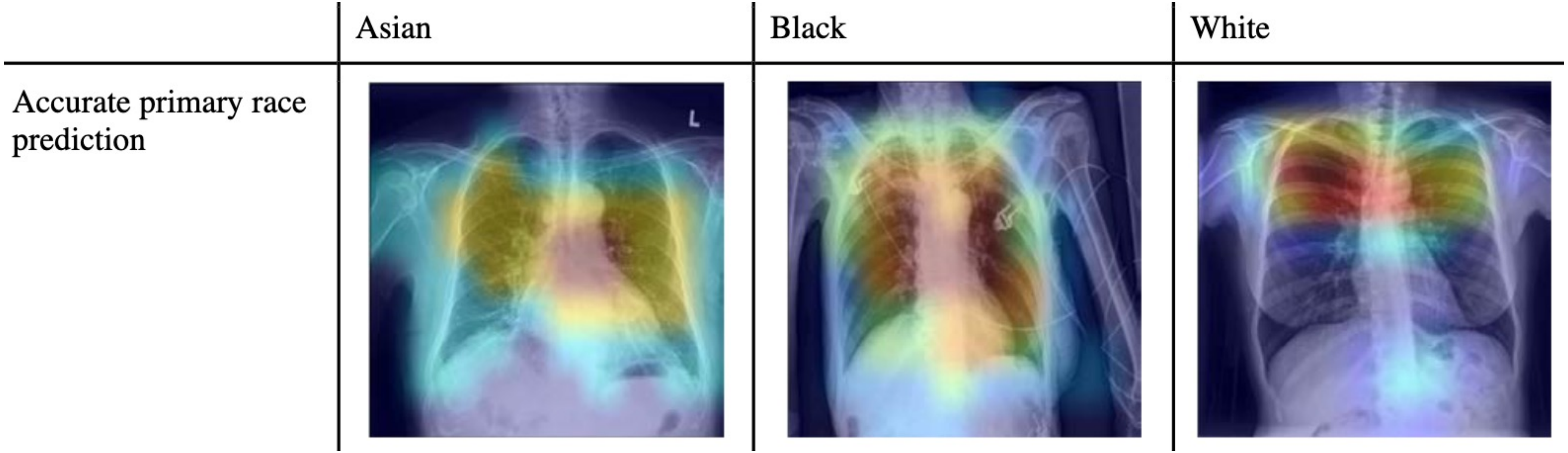
Doctors cannot do this!

The authors were not able to figure out how the ML algorithm does it

These developments underline the ethical and legal requirements for explainable ML

Gichoya et al., "AI recognition of patient race in medical imaging: a modelling study", Lancet Digital Health (2022)

PIXEL ATTRIBUTION FOR RACE PREDICTION



Pixel attribution techniques were not useful for figuring out how network predicts race

Gichoya et al., "AI recognition of patient race in medical imaging: a modelling study", Lancet Digital Health (2022)

AN EXPLAINABLE CHEST X-RAY EXPERT SYSTEM FOR X-RAY DIAGNOSIS

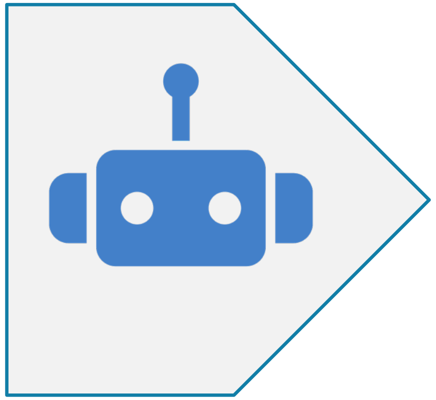
Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals

Susu Sun[1] SUSU.SUN@UNI-TUEBINGEN.DE
[1] *Cluster of Excellence – Machine Learning for Science, University of Tübingen, Germany*
Stefano Woerner[1] STEFANO.WOERNER@UNI-TUEBINGEN.DE
Andreas Maier[2] ANDREAS.MAIER@FAU.DE
[2] *Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*
Lisa M. Koch[3,4] LISA.KOCH@UNI-TUEBINGEN.DE
[3] *Hertie Institute for Artificial Intelligence in Brain Health, University of Tübingen, Germany*
[4] *Institute of Ophthalmic Research, University of Tübingen, Germany*
Christian F. Baumgartner[1] CHRISTIAN.BAUMGARTNER@UNI-TUEBINGEN.DE

Accepted to the Conference on Medical Imaging with Deep Learning (MIDL) 2023

PROBLEM SETTING



Atelectasis

Cardiomegaly

Consolidation

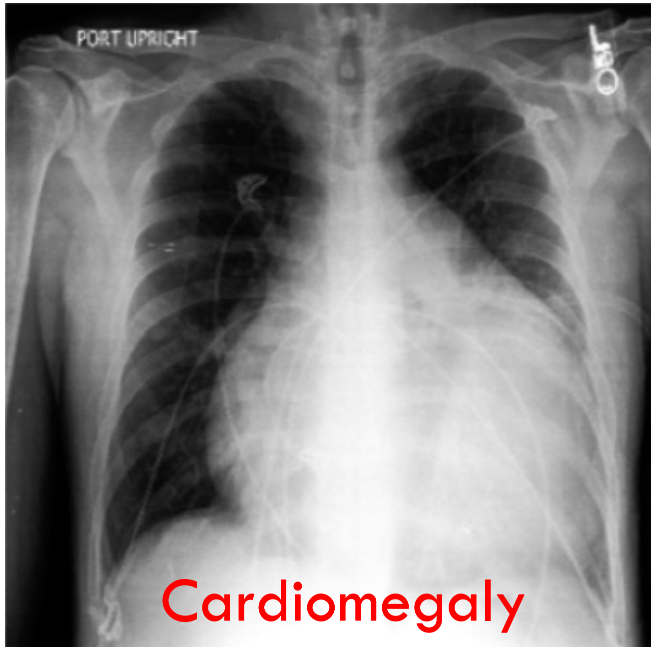
Edema

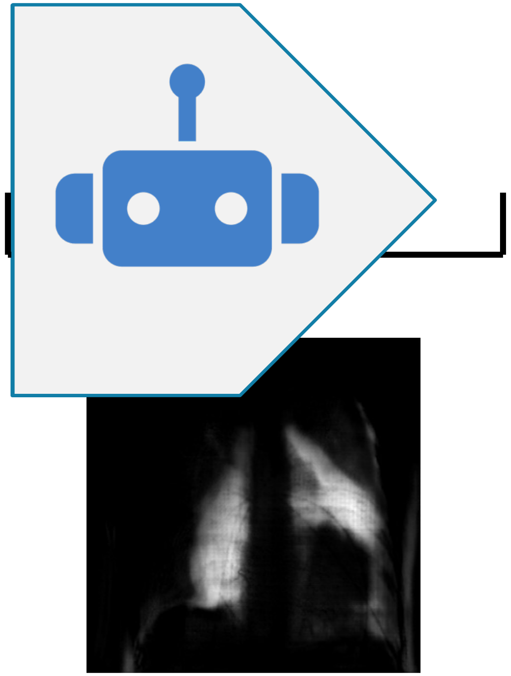
Effusion

Sun et al., “Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals”, Proc. MIDL 2023

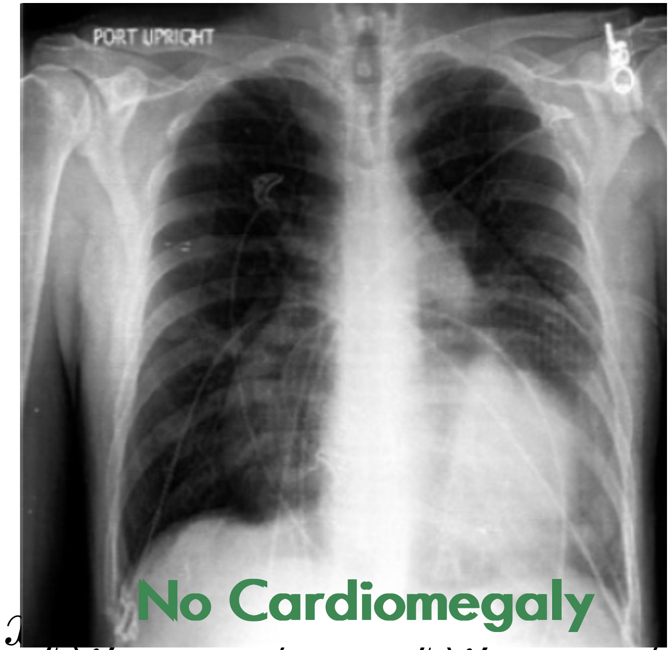
OUR APPROACH: WHAT WOULD AN IMAGE LOOK IF IT DID NOT HAVE A PARTICULAR DISEASE?

$x_i(y = 1)$





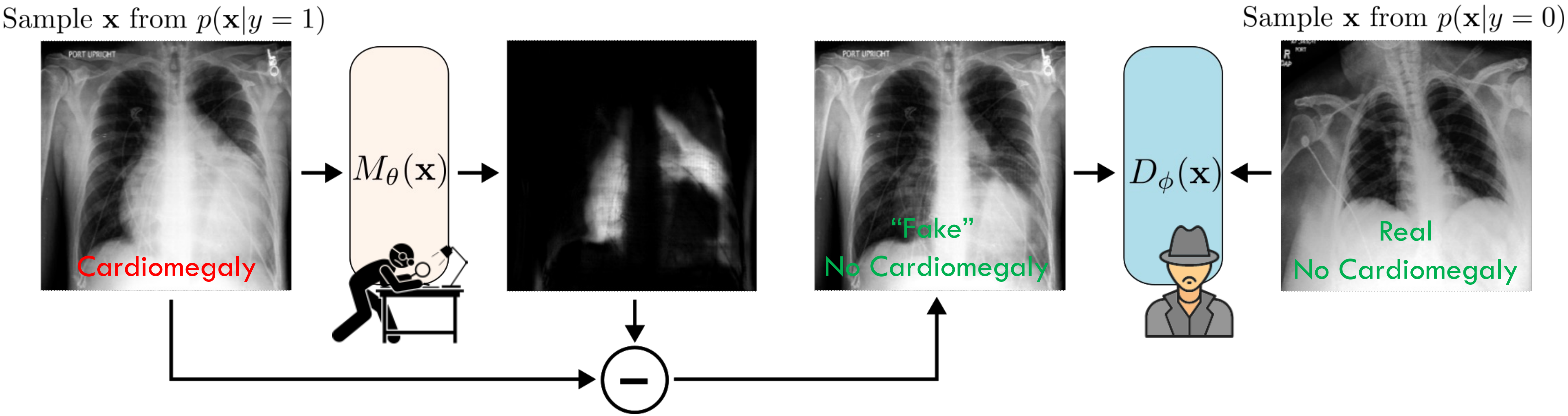
$x_i(y = 0)$



"Individual Disease Effect"

Sun et al., "Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals", Proc. MIDL 2023

COUNTERFACTUAL GENERATION USING GENERATIVE ADVERSARIAL NETWORKS (GAN)



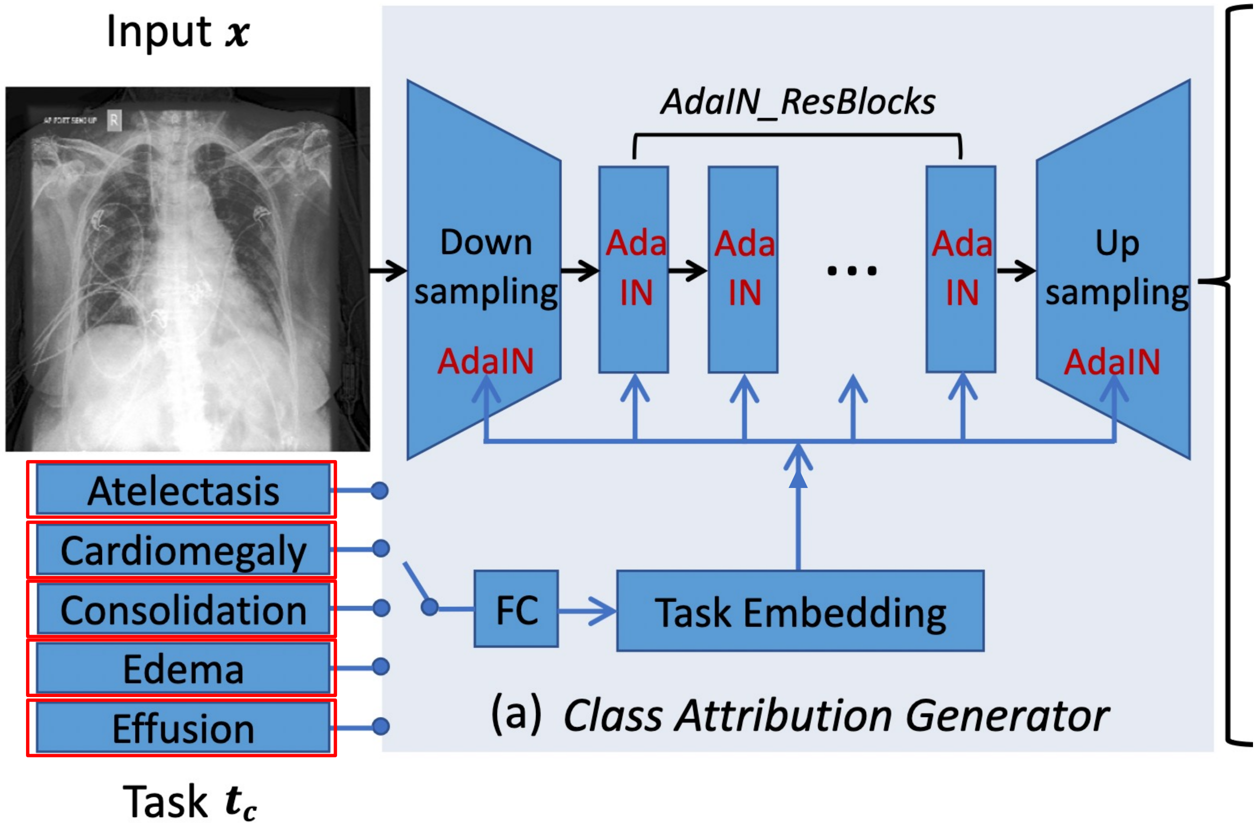
$$\mathcal{L}_{\mathrm{GAN}}(\theta,\phi) = -\mathbb{E}_{\mathbf{x}\sim p(\mathbf{x}|y=1)}[D_\phi(\mathbf{x}-M_\theta(\mathbf{x}))] + \mathbb{E}_{\mathbf{x}\sim p(\mathbf{x}|y=0)}[D_\phi(\mathbf{x})]$$

$$\mathcal{L}_{\mathrm{reg}}(\theta) = \mathbb{E}_{\mathbf{x}\sim p(\mathbf{x})}[||M_\theta(\mathbf{x})||_1]$$

$$\theta^* = \arg\min_\theta \max_\phi \mathcal{L}_{\mathrm{GAN}}(\theta,\phi)$$

Sun et al., "Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals", Proc. MIDL 2023

BUILDING A CLASSIFIER BASED ON COUNTERFACTUALS: ATTRI-NET



Sun et al., "Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals", Proc. MIDL 2023

DISEASE CLASSIFICATION PERFORMANCE

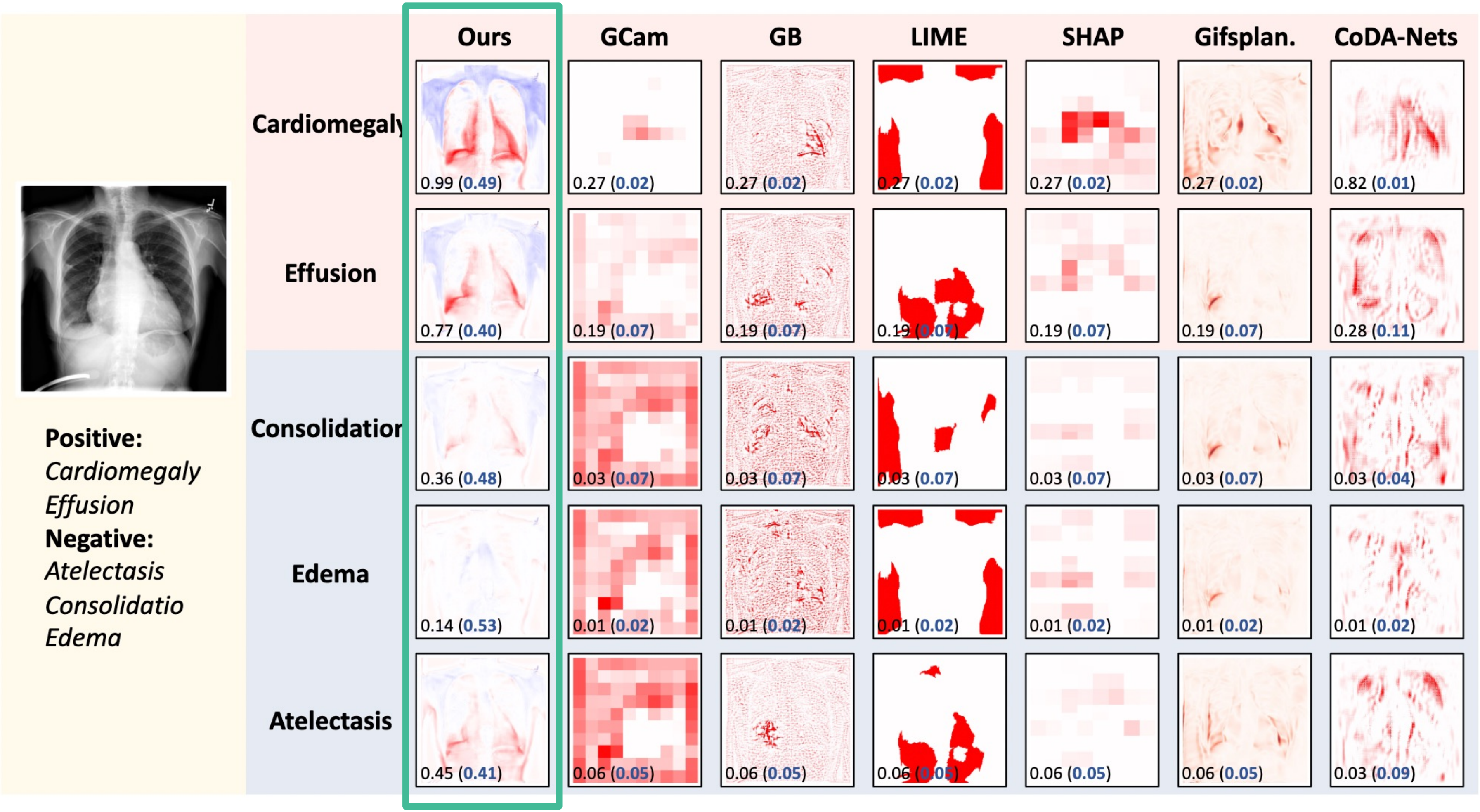
Area under ROC curve for different methods

Model	CheXpert	ChestX-ray8	VindrCXR
ResNet50 (Azizi et al., 2021)	0.7687	-	-
SimCLR (Azizi et al., 2021)	0.7702	-	-
LSE (Ye et al., 2020)	-	0.7554	-
ChestNet (Ye et al., 2020)	-	0.7896	-
ResNet50	0.7727	0.7445	0.8986
CoDA-Nets	0.7659	0.7727	0.9322
ours	0.7405	0.7762	0.9405

→Attri-Net performs on a par with state-of-the-art classification approaches

Sun et al., "Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals", Proc. MIDL 2023

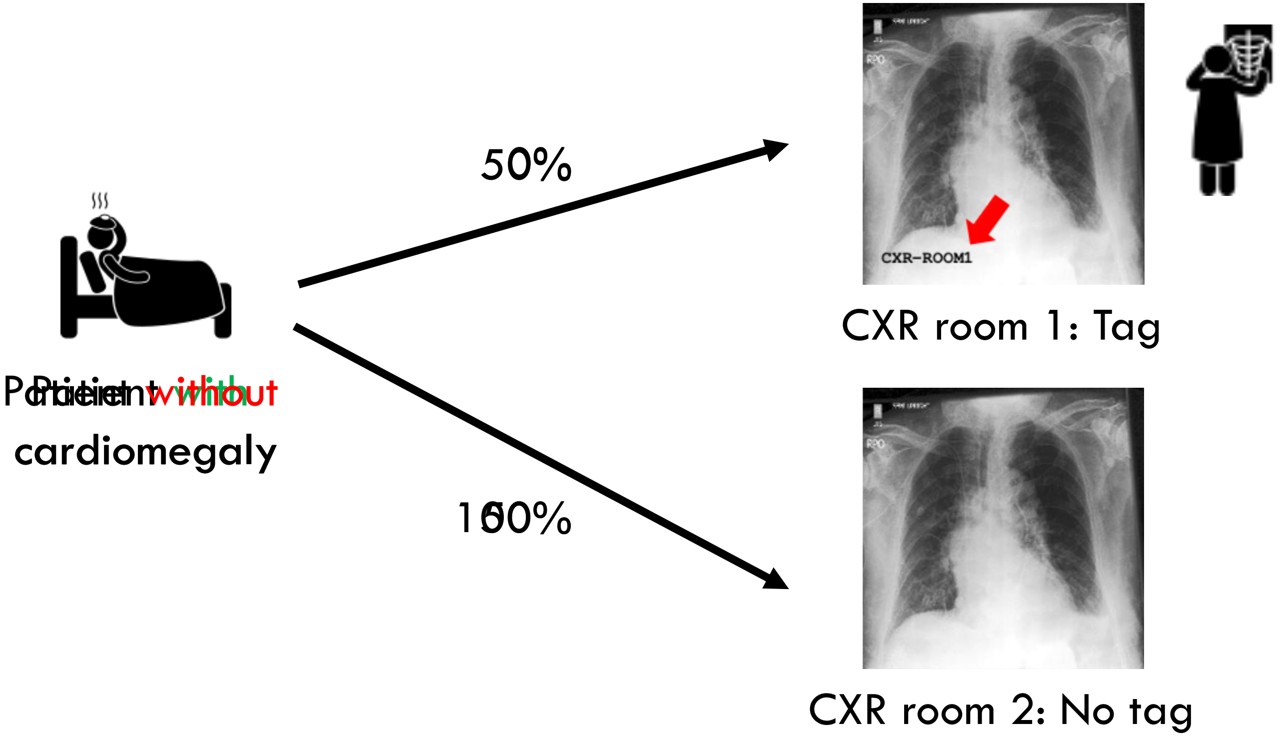
COMPARISON OF EXPLANATION QUALITY TO BASELINE TECHNIQUES



Sun et al., "Inherently Interpretable Multi-Label Classification Using Class-Specific Counterfactuals", Proc. MIDL 2023

ANOTHER EVALUATION APPROACH: DETECTING CONFOUNDING FEATURES

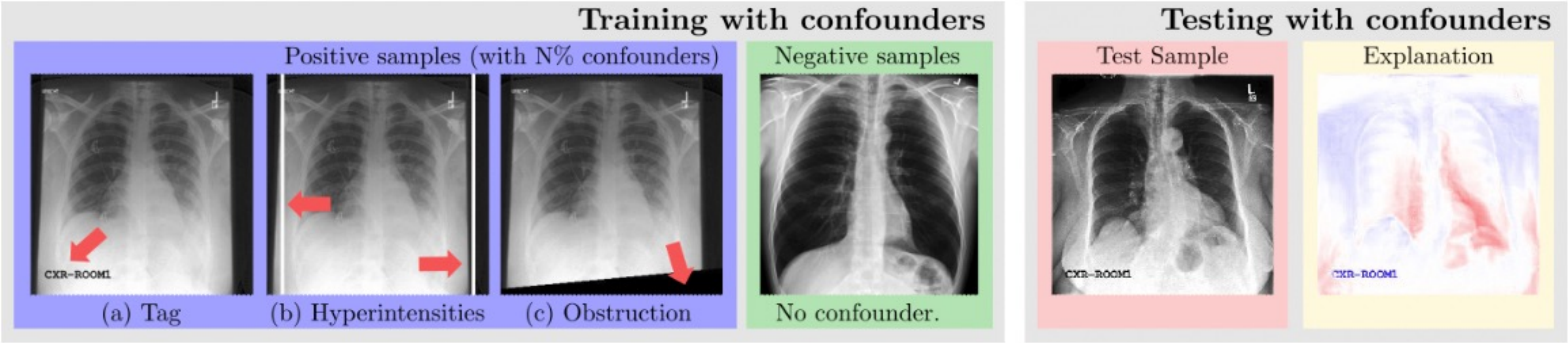
There may be signals that are *correlated* with a disease but *unrelated* to it



Sun et al., “Right for the wrong reason: can interpretable ML techniques detect spurious correlations?”, Proc. MICCAI 2023 (under review)

EVALUATION WITH SHORTCUT LEARNING

N% of cardiomegaly cases are modified to include **a confounding signal**

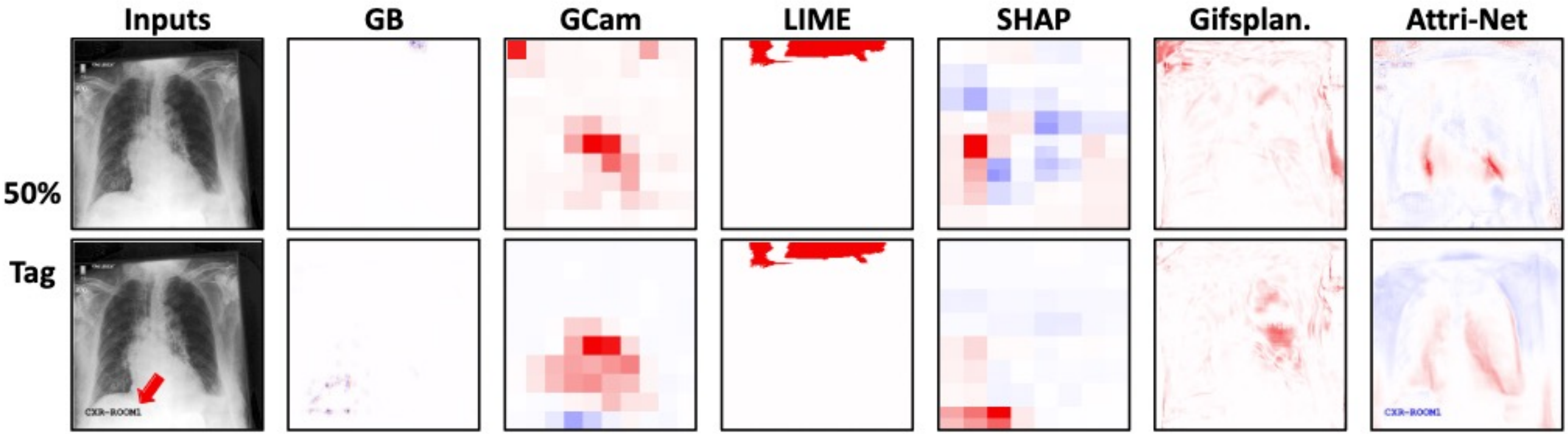


→ Finding 1: All investigated techniques learn a shortcut (pay attention to confounder)

→ Question: Can explainable ML techniques help us uncover this unwanted dependence?

Source: Sun et al., "Right for the wrong reason: can interpretable ML techniques detect spurious correlations?", Proc. MICCAI 2023 (under review)

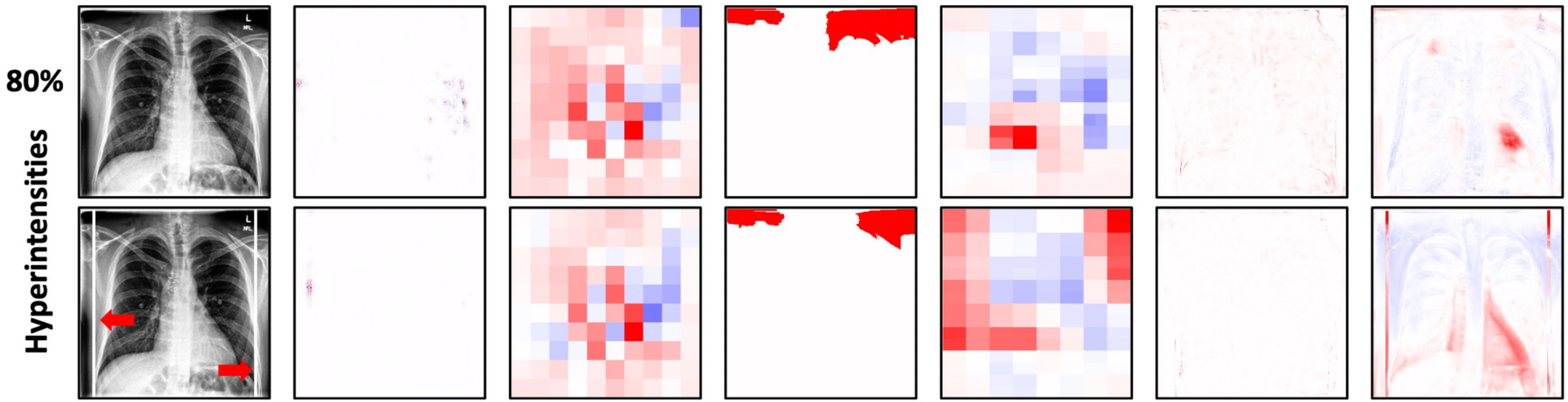
EVALUATION WITH SHORTCUT LEARNING



1) Explanations should highlight confounded pixels *if prediction switches*
2) Explanations with and without tag should be different *if prediction switches*

Source: Sun et al., "Right for the wrong reason: can interpretable ML techniques detect spurious correlations?", Proc. MICCAI 2023 (under review)

EVALUATION WITH SHORTCUT LEARNING



1) Explanations should highlight confounded pixels *if prediction switches*
2) Explanations with and without tag should be different *if prediction switches*

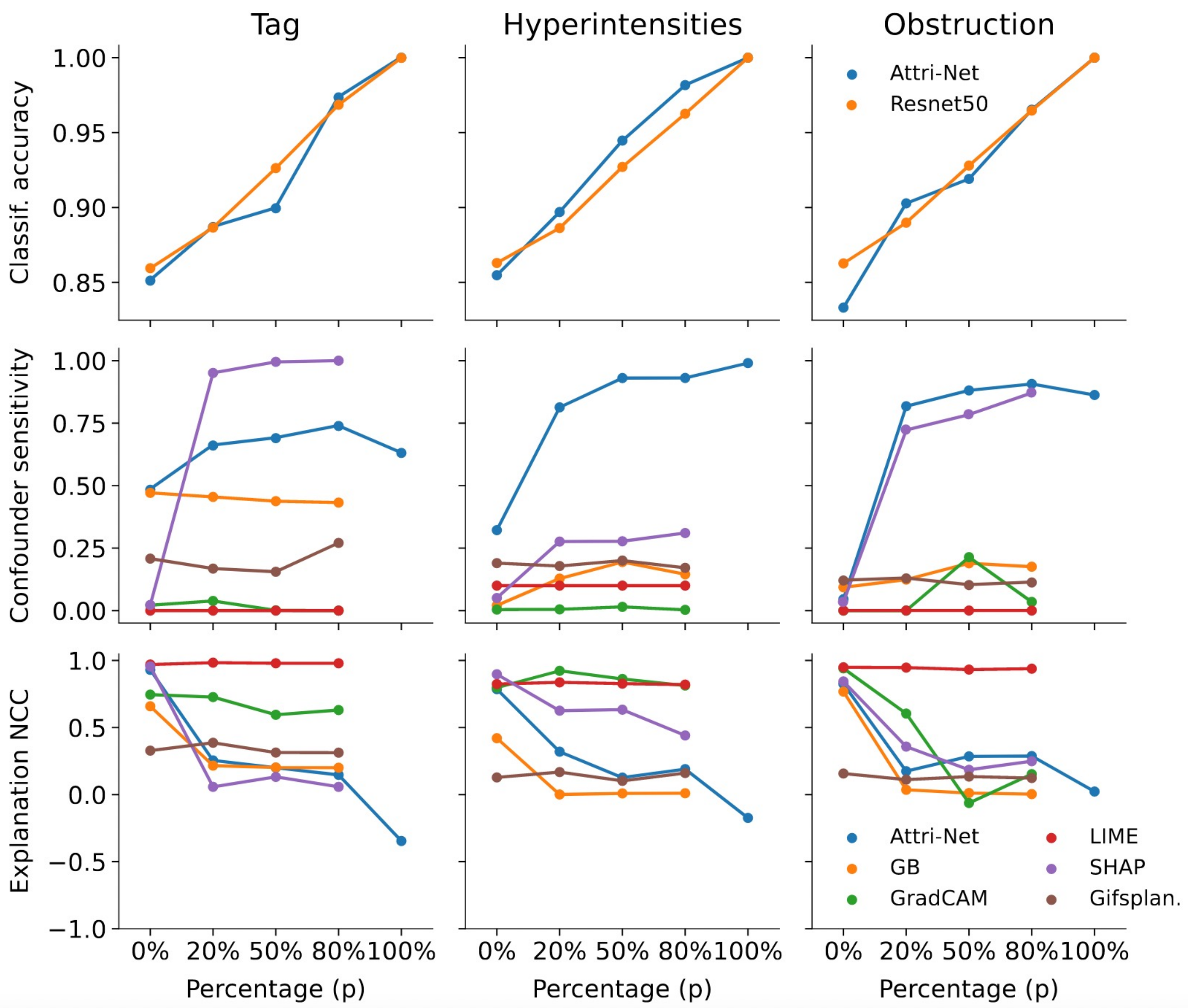
Source: Sun et al., "Right for the wrong reason: can interpretable ML techniques detect spurious correlations?", Proc. MICCAI 2023 (under review)

QUANTITATIVE EVALUATION

Accuracy increases with more % confounded in training and test data

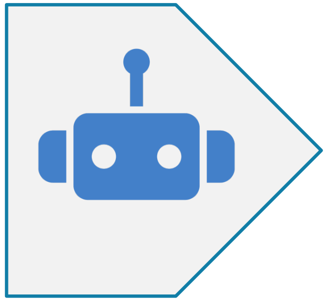
1) Explanations should highlight confounded pixels *if prediction switches*

2) Explanations with and without tag should be different *if prediction switches*

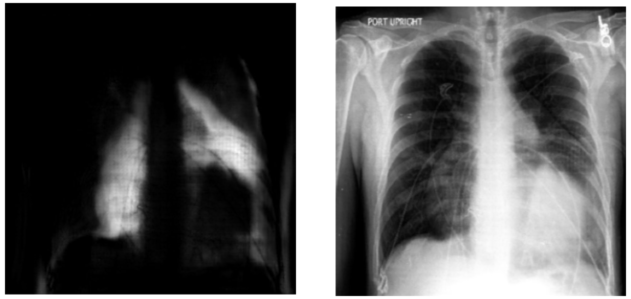


Source: Sun et al., "Right for the wrong reason: can interpretable ML techniques detect spurious correlations?", Proc. MICCAI 2023 (under review)

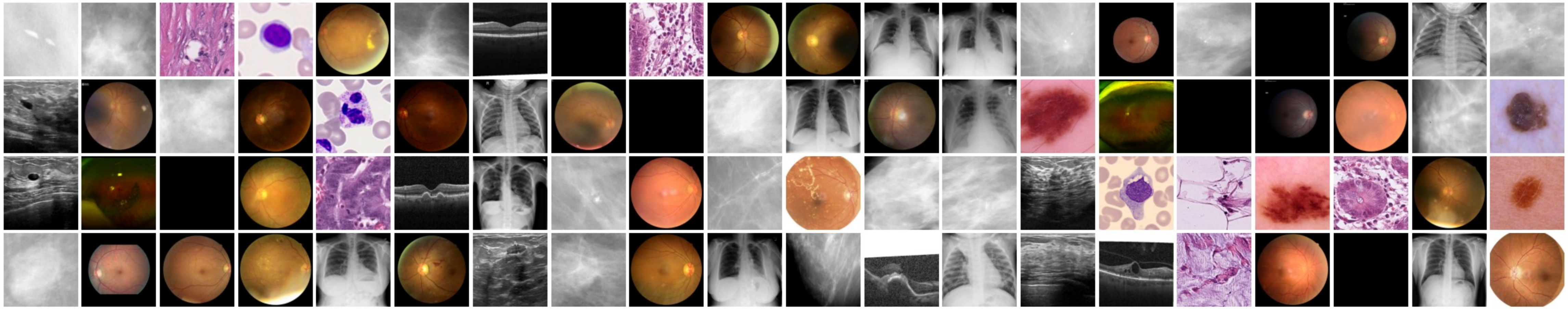
FUTURE WORK: WILL IT IMPROVE CLINICAL WORKFLOWS?



“This patient has cardiomegaly, **because…**”



Ongoing work: evaluation of this approach in a **doctor-in-the-loop setting**



THE MICCAI LEARN2LEARN CHALLENGE: LEARNING FROM FEW DATA POINTS

Project lead:
Stefano Woerner

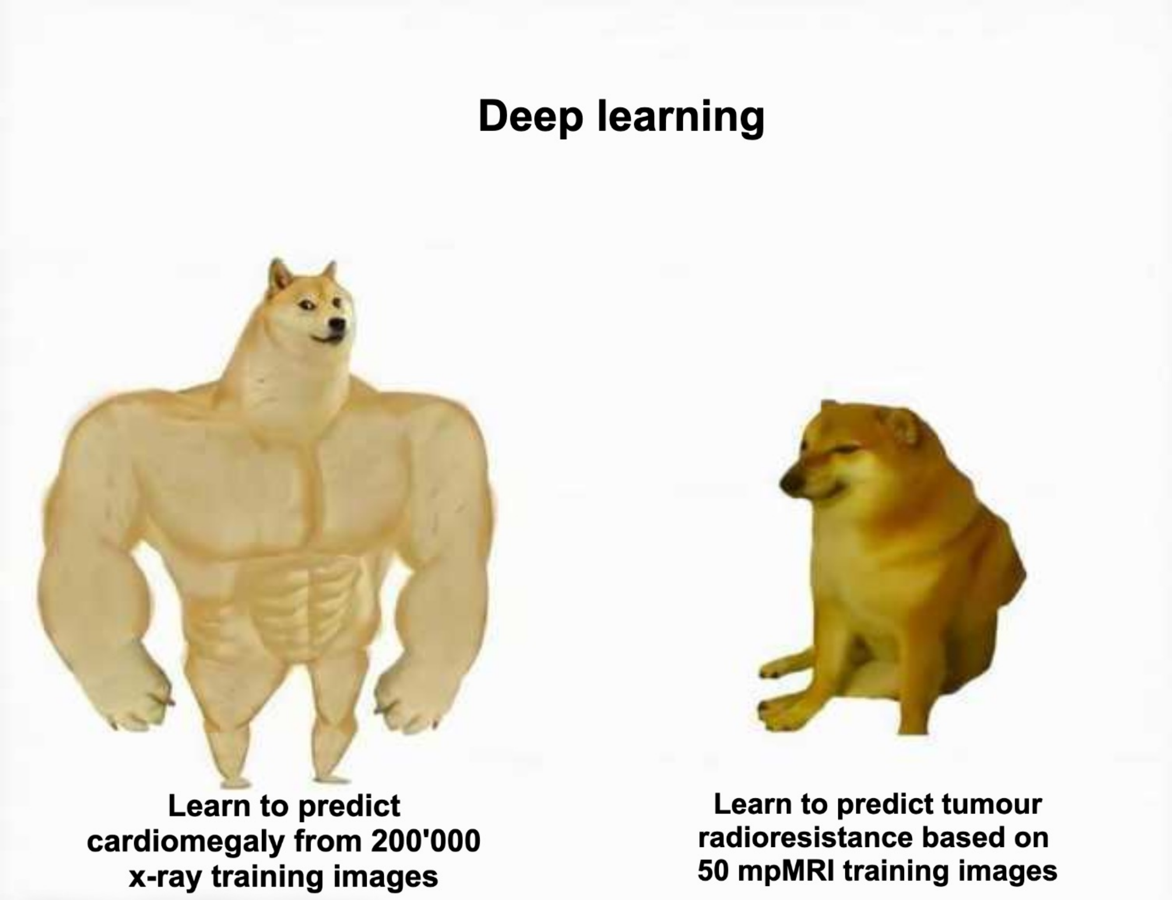
MOTIVATION

Deep learning gets applied to classification only in a handful of tasks with very good data availability

Many interesting tasks remain unstudied due to lack of data

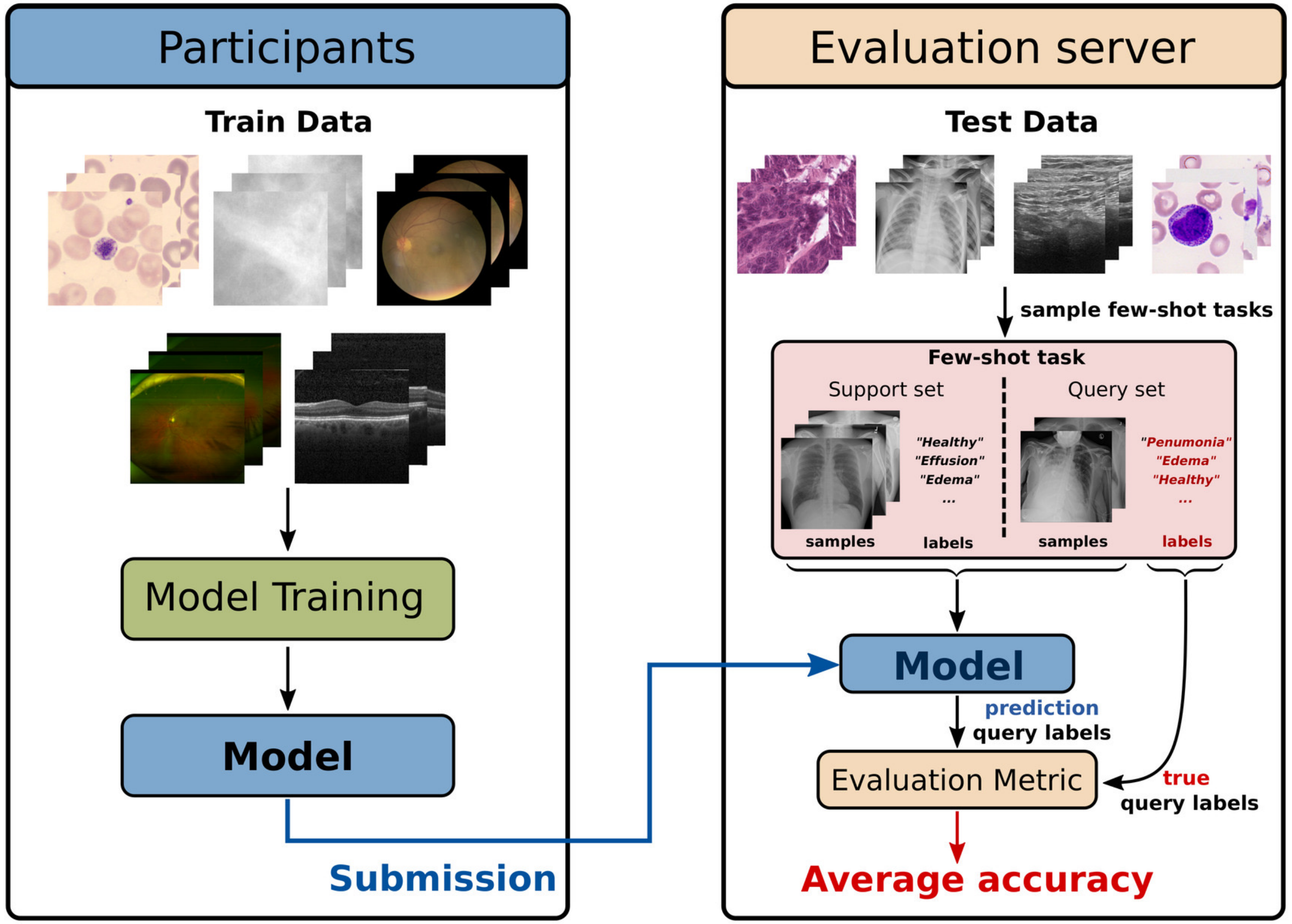
Transfer learning may be part of the solution

Many tasks are very related, which data/task do you transfer from?



Main objective: A general purpose learning algorithm that can be applied to any small dataset

BASIC IDEA OF THE L2L CHALLENGE



DATASET AND CHALLENGE TIMELINE

Extended benchmark to **28 tasks** derived from **17 publicly available datasets**

We also publish easy-to-use and unified pytorch data loaders and FSL toolbox

- **02 May 2023** - Data release
- **25 May 2023** - Submis
- **29 June 2023** - Regist
- **10 August 2023** - Submission deadline
- **31 August 2023** - Top-performing teams contacted
- **8-12 October 2023** - Results announced at MICCAI 2023

1000 CAD cash prize!

ImFusion

www.l2l-challenge.org

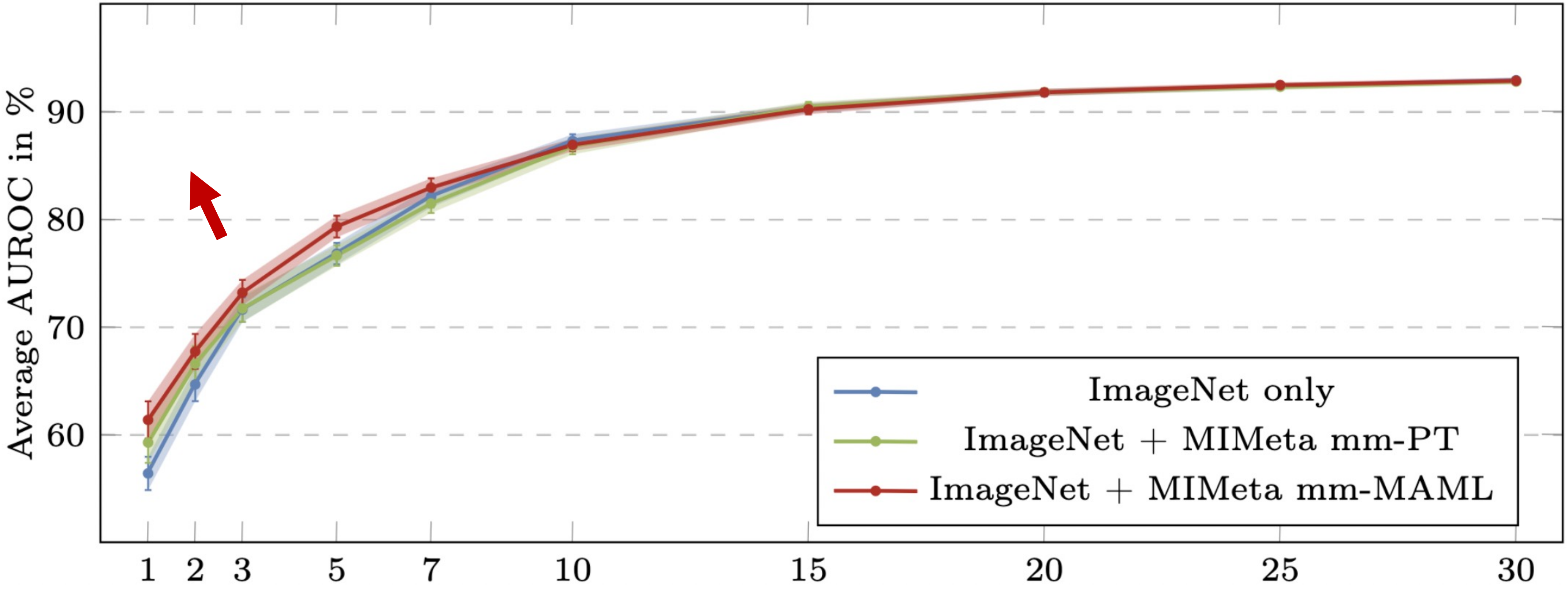
HOW CAN YOU LEARN TO LEARN?
ONE POSSIBLE APPROACH: META-LEARNING

Strategies for Meta-Learning with Diverse Tasks

Stefano Woerner STEFANO.WOERNER@UNI-TUEBINGEN.DE
Christian F. Baumgartner CHRISTIAN.BAUMGARTNER@UNI-TUEBINGEN.DE
Cluster of Excellence Machine Learning: New Perspectives for Science, University of Tübingen

Presented as short paper at MIDL 2022

APPLICATION TO OUR CHALLENGE DATASET



Source: Woerner et al., "The MIMeta Benchmark: Towards Cross-Domain Few-shot Learning in Medical Image Analysis", Proc. MICCAI 2023 (under review)

THANKS

www.l2l-challenge.org

www.mlmia-unitue.de

christian.baumgartner@uni-tuebingen.de

machine learning
new perspectives
for science

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Carl Zeiss
Stiftung