Learning Low-Rank Feature for Thorax Disease Classification

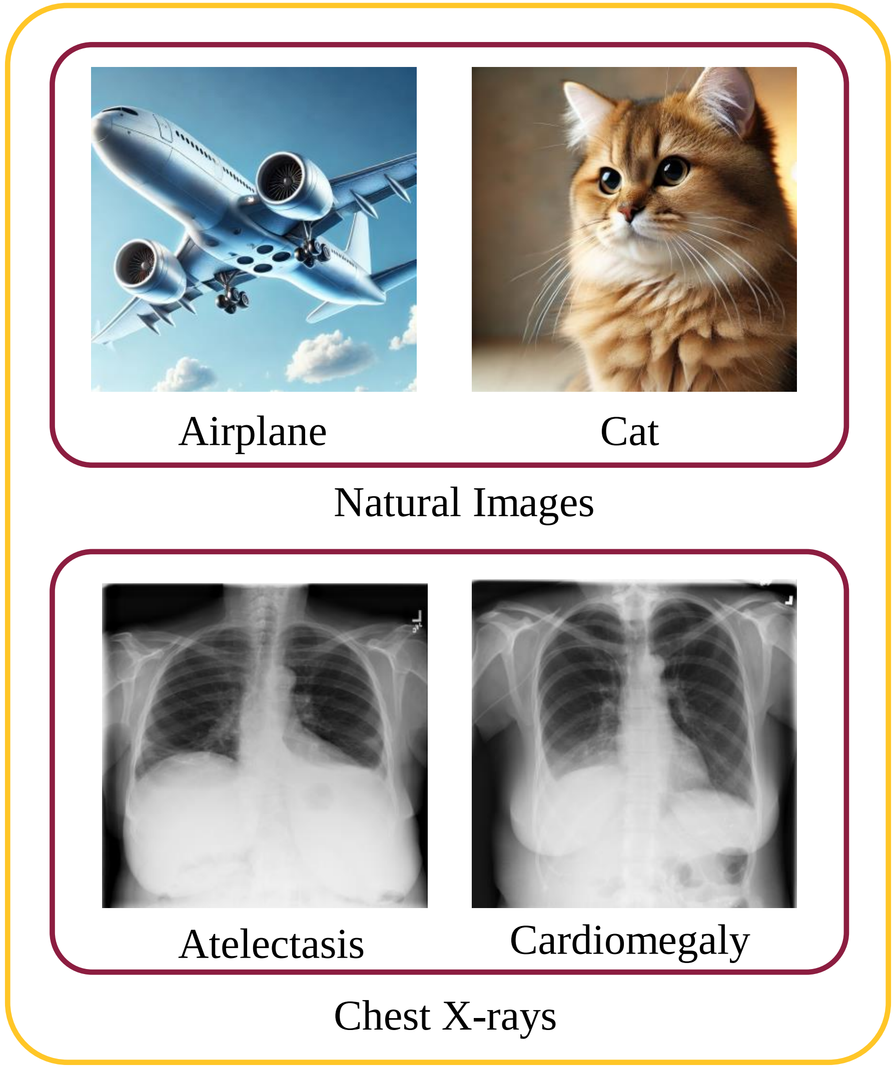
Authors: Rajeev Goel*[1], Yancheng Wang*[1], Utkarsh Nath[1], Alvin C. Silva[2], Teresa Wu[1], Yingzhen Yang[1]

Affiliations: Arizona State University[1], Mayo Clinic Arizona[2]

Copyright © 2023 Arizona Board of Regents

Introduction

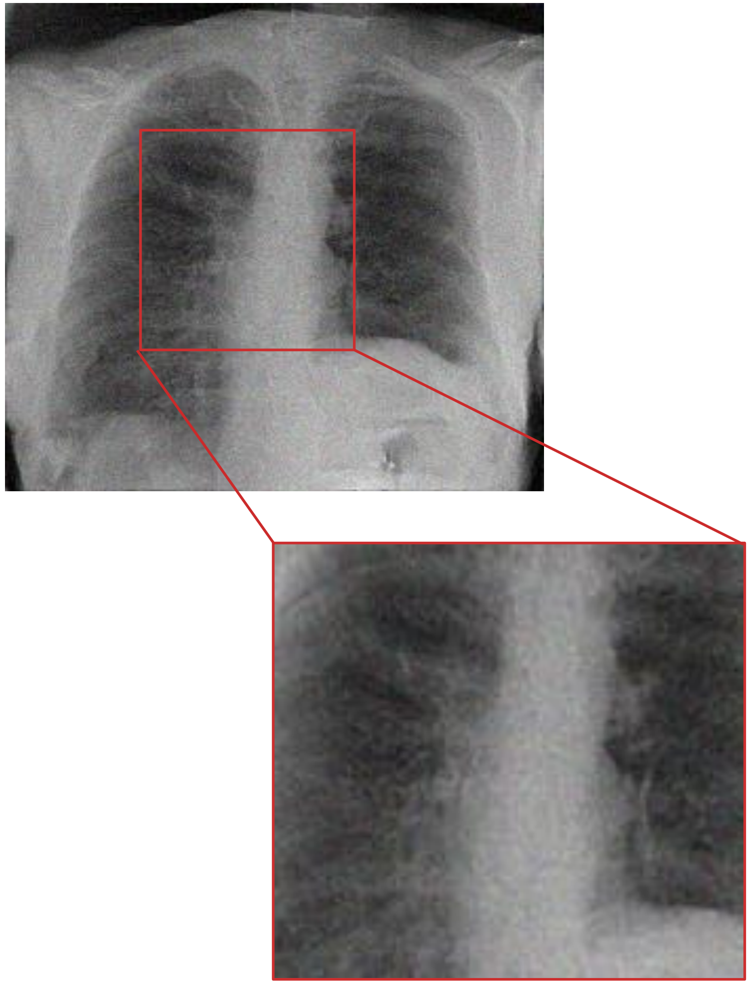
- Advancements in AI: Growing use of AI in medicine
- Challenges in Medical Imaging:
    - Medical images differ from natural images
    - Unique structures and features
    - Fewer labeled medical images
    - Noise complicates disease classification



Copyright © 2023 Arizona Board of Regents

Problem

- Key Challenges:
  - Limited availability of annotated data
  - Subtle disease features
  - Noise and non-disease interference
- Our Approach:
  - LRFL: Extract low-rank features
  - Minimize noise and enhance accuracy



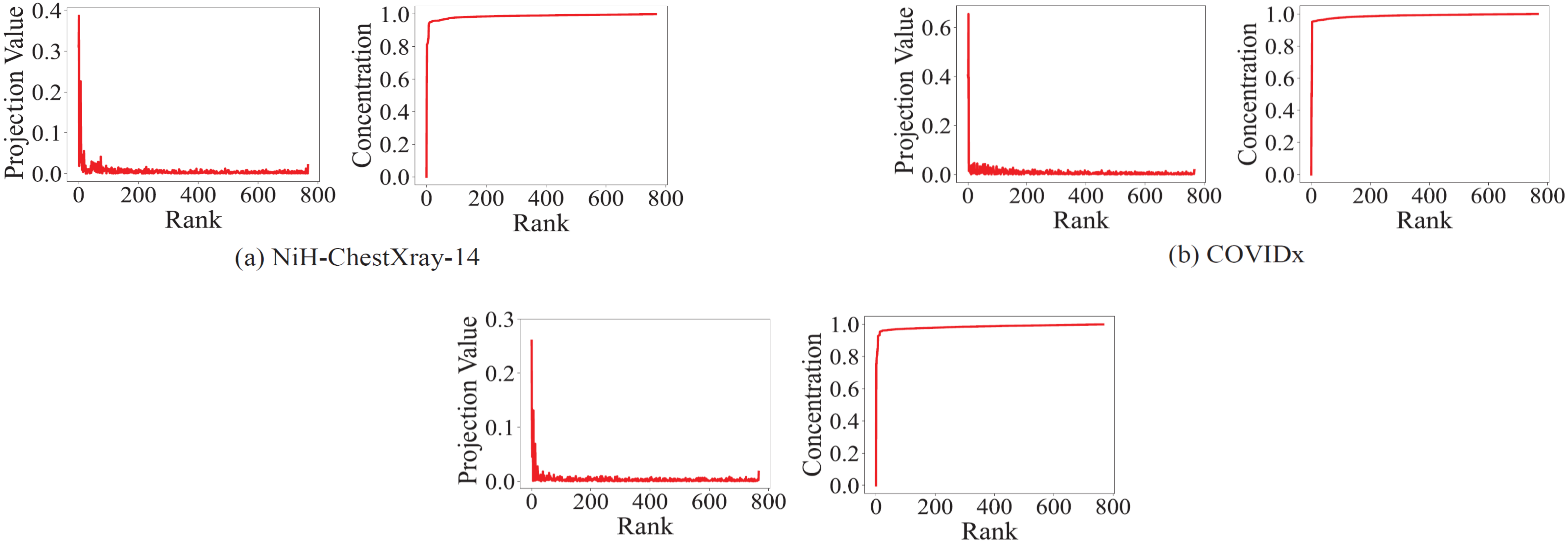
Noisy Chest X-ray

Copyright © 2023 Arizona Board of Regents

Methodology

Copyright © 2023 Arizona Board of Regents

Motivation



(a) NiH-ChestXray-14

(b) COVIDx

(c) CheXpert

- GT class labels retains key information.
- LRFL focuses on low-rank features.

Copyright © 2023 Arizona Board of Regents

Novel Approximation

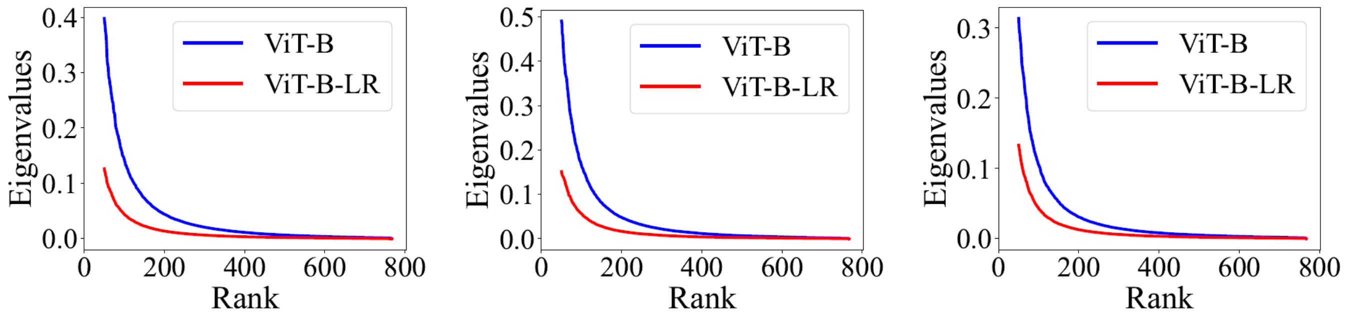
Algorithm 1 Training Algorithm with the Approximate Truncated Nuclear Norm by SGD

Initialize the weights $\mathbf{W}_1$ by $\mathbf{W}_1 = \mathbf{W}_1(0)$, and initialize $\mathbf{W}_2$ randomly
Compute feature $\mathbf{F}$ by the neural network, and its SVD as $\mathbf{F} = \mathbf{U\Sigma V}$
Update $\overline{\mathbf{U}} = \mathbf{U}$, $\overline{\mathbf{V}} = \mathbf{V}$
for $t = 1, 2, \ldots, t_{\max}$ **do**
    if $t \equiv 0 \pmod{t_0}$ **then**
        Compute feature $\mathbf{F}$ of the neural network, and its SVD $\mathbf{F} = \mathbf{U\Sigma V}$.
        Update $\overline{\mathbf{U}} = \mathbf{U}$, $\overline{\mathbf{V}} = \mathbf{V}$
    end if
    for $b = 1, 2, \ldots, B$ **do**
        Update $\mathbf{W}$ by applying gradient descent on batch $\mathcal{B}_j \subseteq [n]$ using the gradient of the loss $\mathcal{L}_j$ in Eq.(4)
    end for
end for
return The trained weights $\mathbf{W}$ of the network

- Novel Truncated Nuclear Norm (TNN) Approximation:
  - Unique low-rank approach.
  - Allows standard SGD optimization.
- Efficiency & Scalability:
  - SVD only after certain epochs to save computation
  - Scalable to large datasets, practical for real-world use.

Copyright © 2023 Arizona Board of Regents

Generalization Bound for LRFL



(a) ChestX-ray14 (b) COVIDx (c) CheXpert

Theorem 1. For every $x > 0$, with probability at least $1 - \exp(-x)$, after the t-th iteration of gradient descent for all $t \geq 1$, we have

$$L_{\mathcal{D}}(\mathrm{NN}_{\mathbf{W}}) \leq \left\|\mathbf{Y} - \bar{\mathbf{Y}}\right\|_{\mathrm{F}} + c_1 \left(1 - \eta\widehat{\lambda}_r\right)^{2t} \|\mathbf{Y}\|_{\mathrm{F}}^2 \quad + c_2 \min_{h \in [0,r]} \left(\frac{h}{n} + \sqrt{\frac{1}{n} \sum_{i=h+1}^{r} \widehat{\lambda}_i}\right) + \frac{c_3 x}{n},$$

where c_1, c_2, c_3 are positive constants.

- **Theoretical Guarantee**: Validates LRFL performance.
- **Reduced Overfitting**: Low-rank features effective with limited data.
- **Better Generalization**: Tighter error bounds vs. baselines.

Copyright © 2023 Arizona Board of Regents

Implementation Details and Results

Copyright © 2023 Arizona Board of Regents

Pipeline for Thorax Disease Classification

- **Pre-training**: Self-supervised learning with Masked Autoencoders (MAE)
- **Fine-tuning**: Cross-Entropy loss on target datasets
- **LRFL fine-tuning**: Using TNN regularization

Copyright © 2023 Arizona Board of Regents

Results

Table 1: Performance comparisons between LRFL models and SOTA baselines on CheXpert. The best result is highlighted in bold, and the second-best result is underlined. This convention is followed by all the tables in this paper. DN represents DenseNet.

Method	Architecture	Rank	Atelectasis	Cardiomegaly	Consolidation	Edema	Effusion	mAUC (%)
Irvin et al.[12]		-	81.8	82.8	<u>93.8</u>	93.4	92.8	88.9
Pham et al.[9]	DN121	-	82.5	85.5	93.7	93.0	92.3	89.4
Kang et al.[93]	DN121	-	82.1	85.9	**94.4**	89.2	93.6	89.0
MoCo v2 [2]	DN121	-	78.5	77.9	92.5	92.8	92.7	88.7
ViT-S [2]	ViT-S/16	-	<u>83.5</u>	81.8	93.5	<u>94.0</u>	93.2	89.2
ViT-S-LR (Ours)	ViT-S/16	0.05r	**83.7**	<u>86.3</u>	90.9	93.7	93.1	<u>89.6</u>
ViT-B [2]	ViT-B/16	-	82.7	83.5	92.5	93.8	**94.1**	89.3
ViT-B-LR (Ours)	ViT-B/16	0.05r	81.6	85.4	93.4	**94.6**	<u>93.9</u>	**89.8**

Table 2: Performance comparisons between LRFL models and SOTA baselines on COVIDx (in accuracy). DN represents DenseNet.

Method	Architecture	Rank	Covid-19 Sensitivity	Accuracy
COVIDNet-CXR Small [8]	-	-	87.1	92.6
COVIDNet-CXR Large [8]	-	-	96.8	94.4
MoCo v2 [2]	DN121	-	94.5	94.0
DN121 [2]	DN121	-	97.0	93.5
ViT-S [2]	ViT-S/16	-	94.5	95.2
ViT-S-LR (Ours)	ViT-S/16	0.01r	<u>97.5</u>	<u>96.8</u>
ViT-B [2]	ViT-B/16	-	95.5	95.3
ViT-B-LR (Ours)	ViT-B/16	0.003r	**98.5**	**97.0**

Copyright © 2023 Arizona Board of Regents

Visualizations

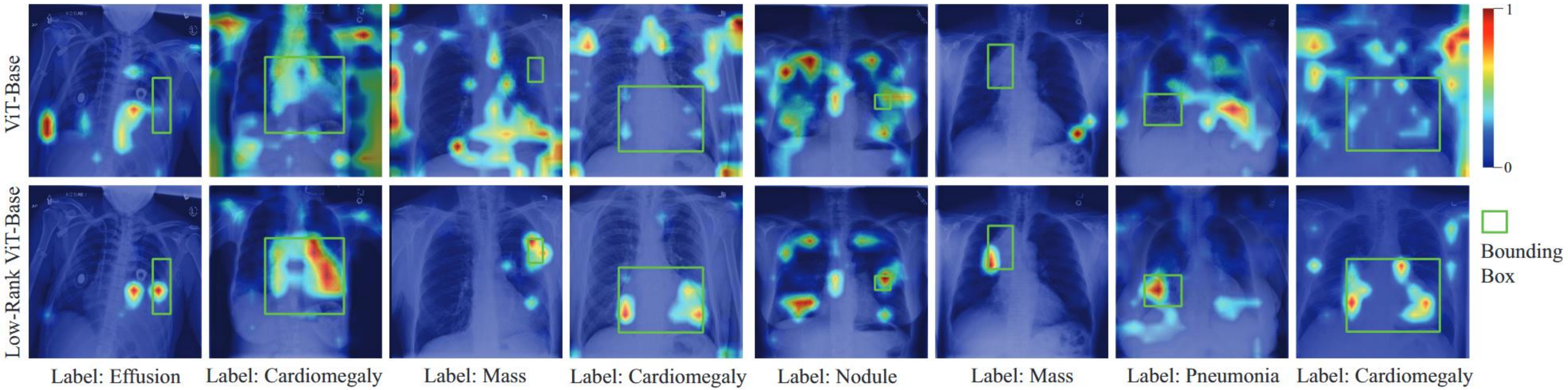


Figure 1: Grad-CAM visualization results on NIH ChestX-ray 14. The figures in the first row are the visualization results of ViT-Base, and the figures in the second row are the visualization results of Low-Rank ViT-Base.

Copyright © 2023 Arizona Board of Regents

Improved Results with Generative Data Augmentation

Table 3: Performance comparison of baseline models and LRFL models on the CheXpert and COVIDx datasets, with and without synthetic data. n denotes the number of training images in the respective dataset.

Method	Architecture	CheXpert			COVIDx		
		Rank	# Synthetic Images	mAUC (%)	Rank	# Synthetic Images	Accuracy (%)
ViT-S [2]	ViT-S/16	-	-	89.2	-	-	95.2
ViT-S-LR (Ours)	ViT-S/16	0.05r	-	89.6	0.01r	-	96.8
ViT-S (Ours)	ViT-S/16	-	$0.2n$	89.3	-	$1.0n$	97.0
ViT-S-LR (Ours)	ViT-S/16	0.05r	$0.2n$	89.7	0.01r	$1.0n$	<u>97.3</u>
ViT-B [2]	ViT-B/16	-	-	89.3	-	-	95.3
ViT-B-LR (Ours)	ViT-B/16	0.025r	-	89.8	0.003r	-	97.0
ViT-B (Ours)	ViT-B/16	-	$0.25n$	<u>89.9</u>	-	$1.0n$	97.0
ViT-B-LR (Ours)	ViT-B/16	0.025r	$0.25n$	**90.4**	0.003r	$1.0n$	**97.5**

Copyright © 2023 Arizona Board of Regents

Contributions

- **Novelty:** Introduction of a separable approximation for the TNN
- **Theoretical Foundation:** Sharp generalization bound
- **Performance**: Improvements in classification accuracy and mAUC
- **Impact**: Effective noise reduction and improved disease localization

Our paper is available at
https://openreview.net/pdf?id=GkzrVxs9LS

Copyright © 2023 Arizona Board of Regents