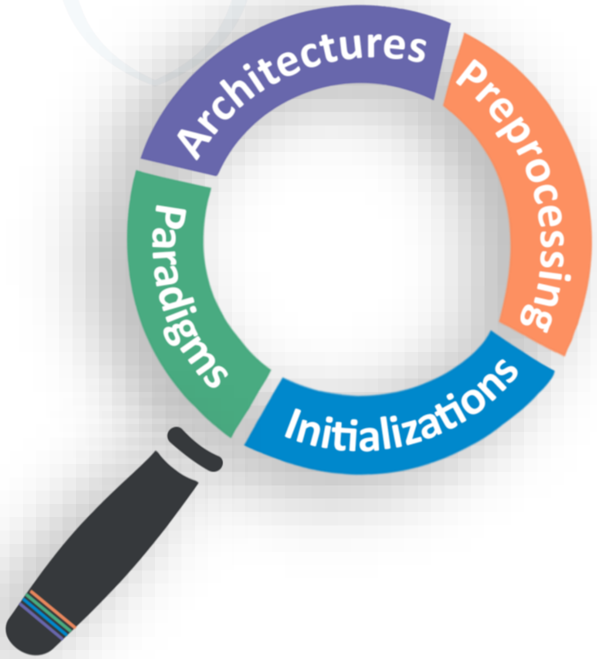
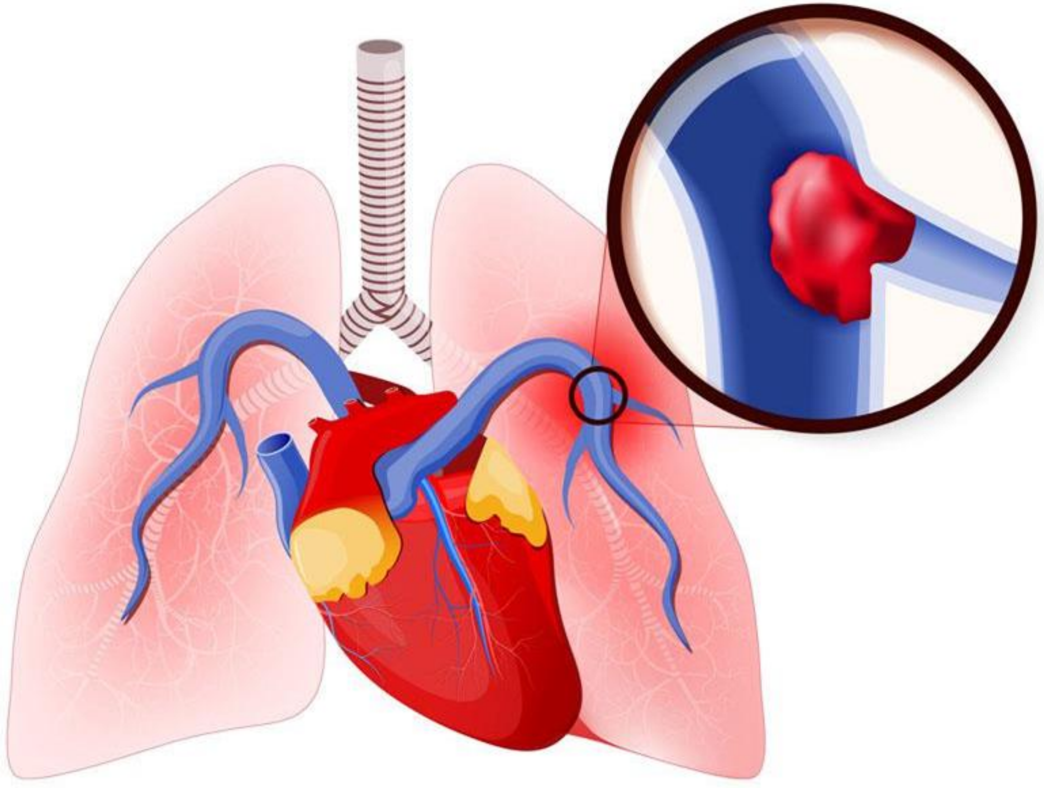
Seeking an Optimal Approach for Computer-aided Diagnosis of Pulmonary Embolism



Nahid Ul Islam[1], Zongwei Zhou[1], Shiv Gehlot[1], Michael B Gotway[2], and Jianming Liang[1]

[1]Arizona State University [2]Mayo Clinic

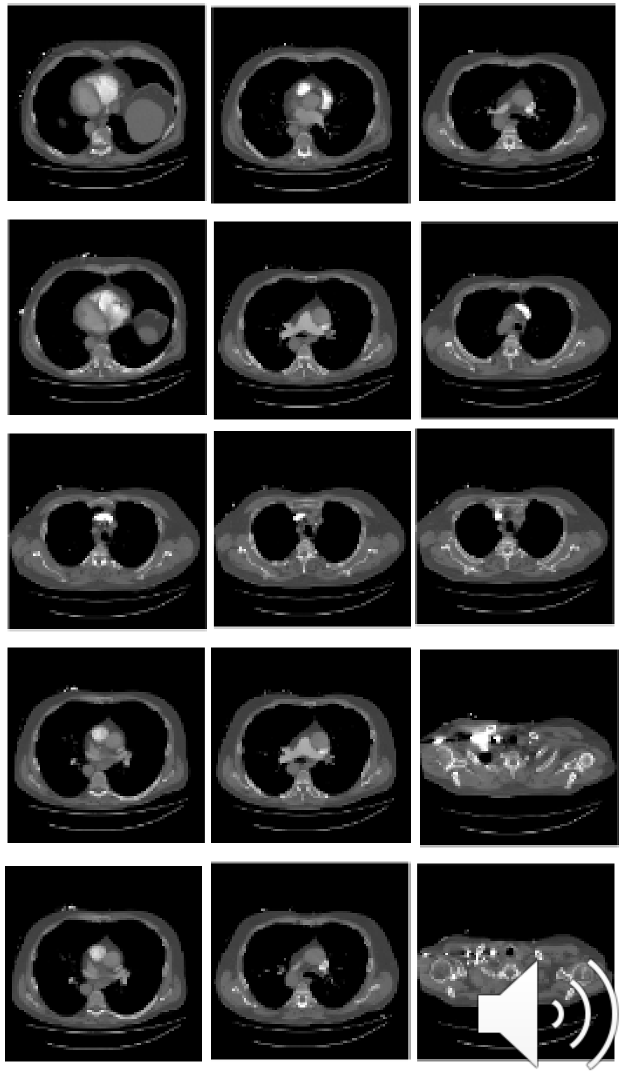
Pulmonary Embolism



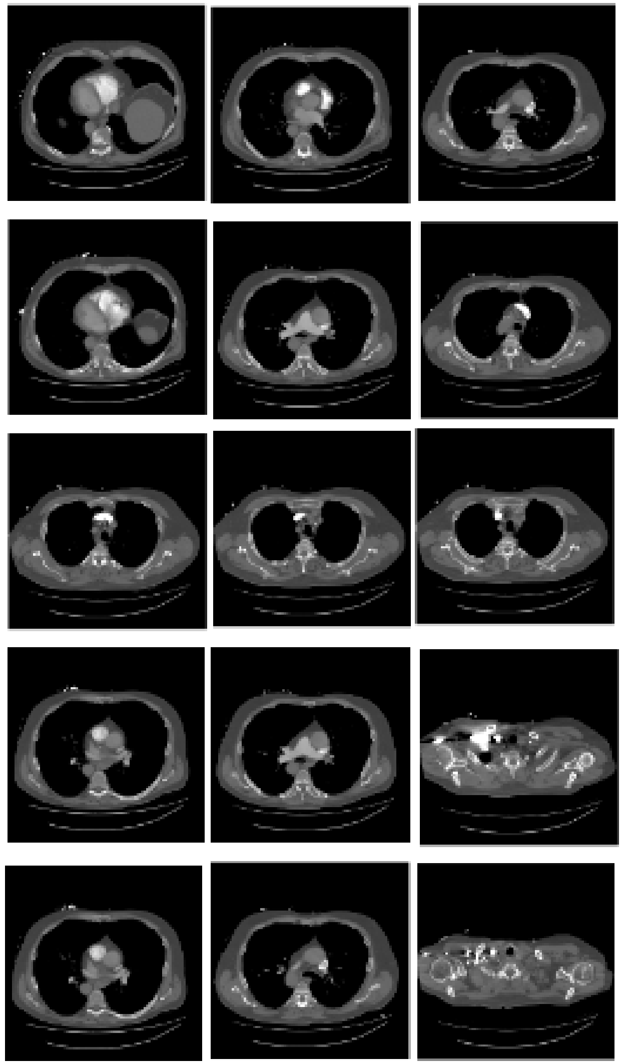
CT Pulmonary Angiography

CT Pulmonary Angiography

CT Pulmonary Angiography

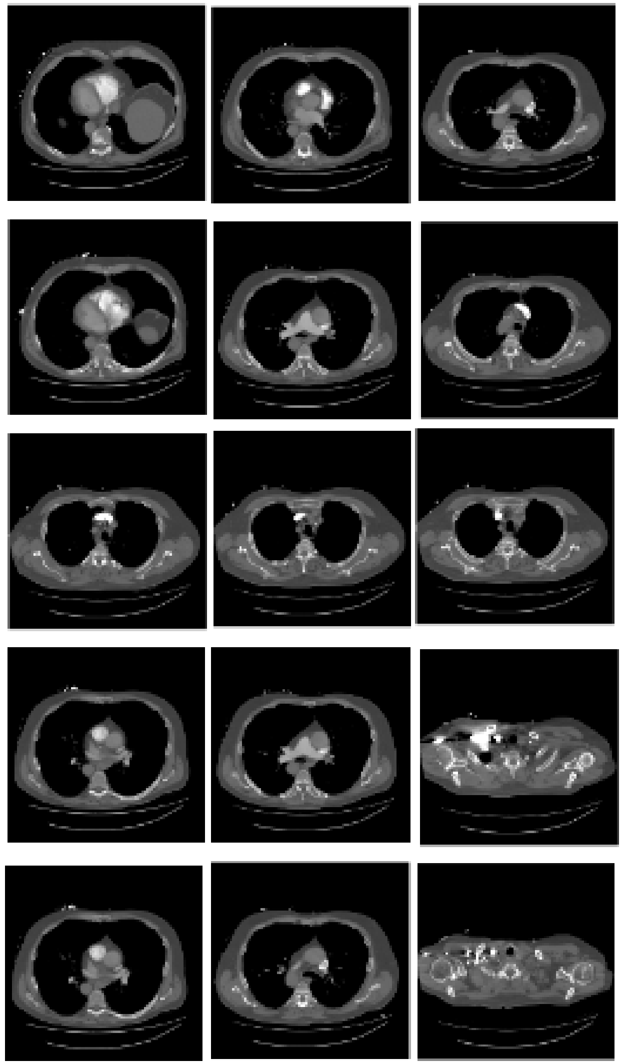


CT Pulmonary Angiography



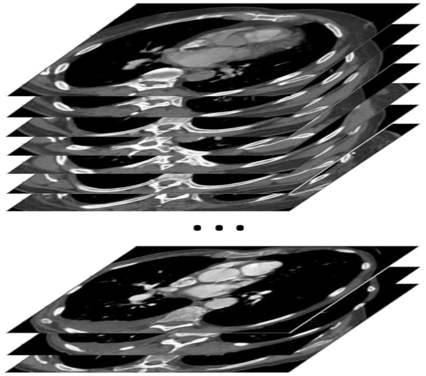
. . .

CT Pulmonary Angiography

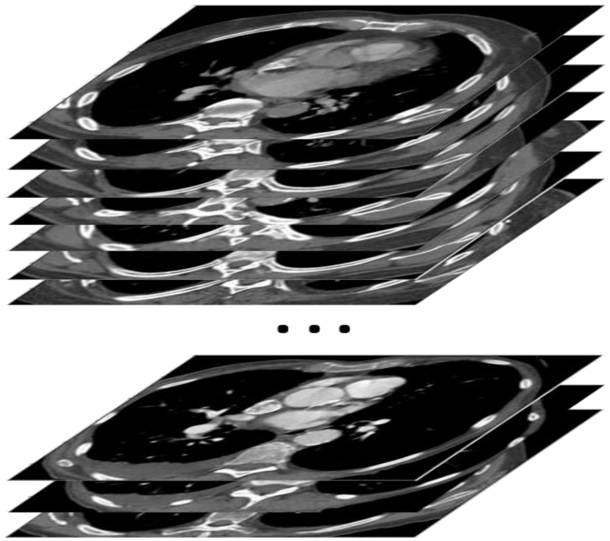


. . .

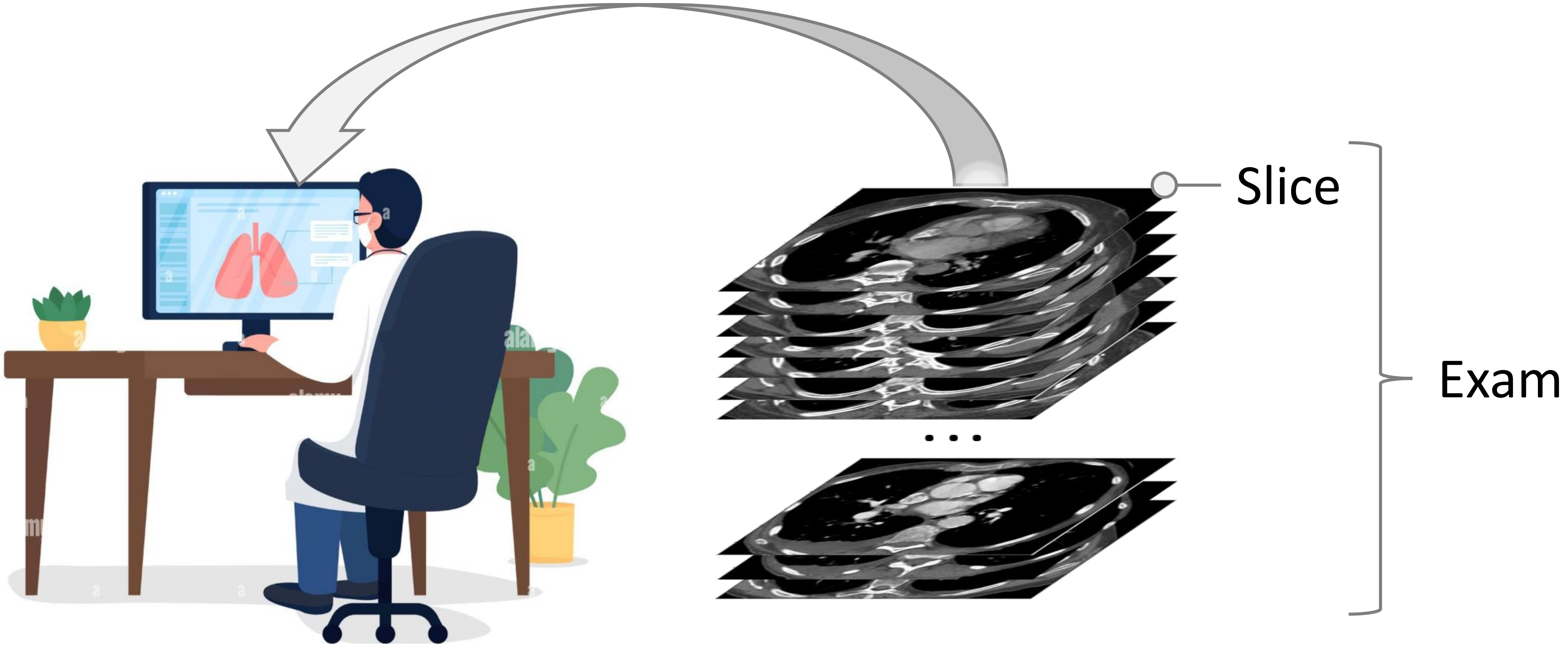
CT Pulmonary Angiography



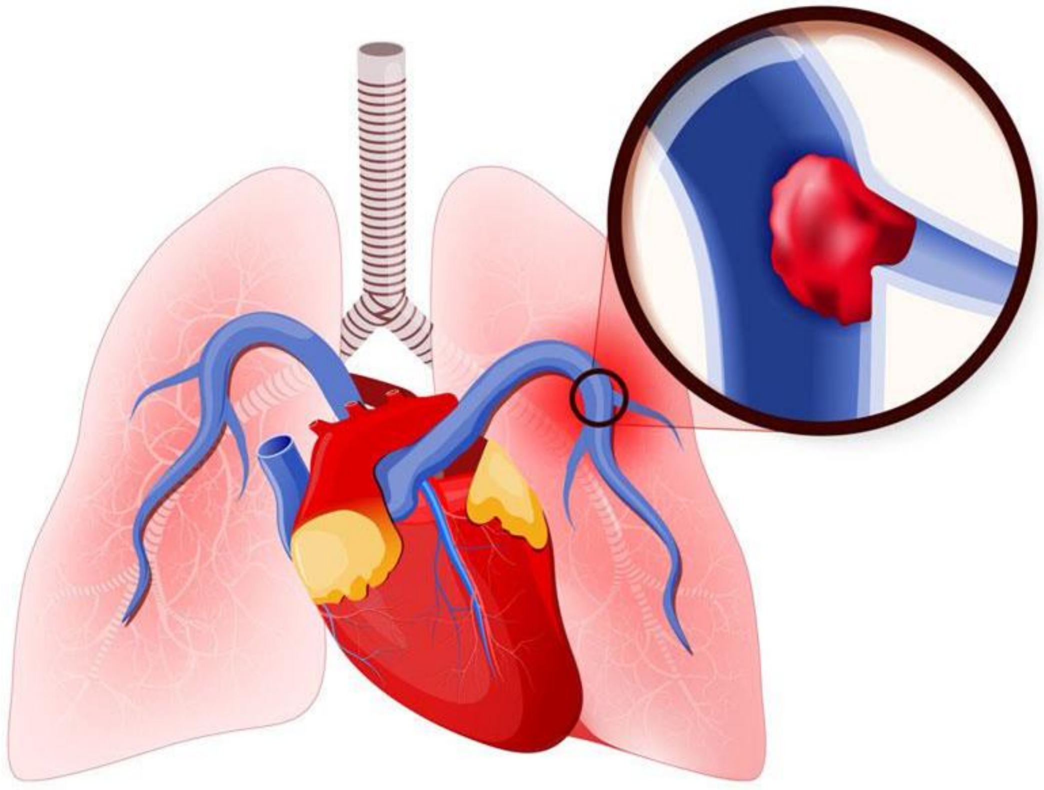
CT Pulmonary Angiography

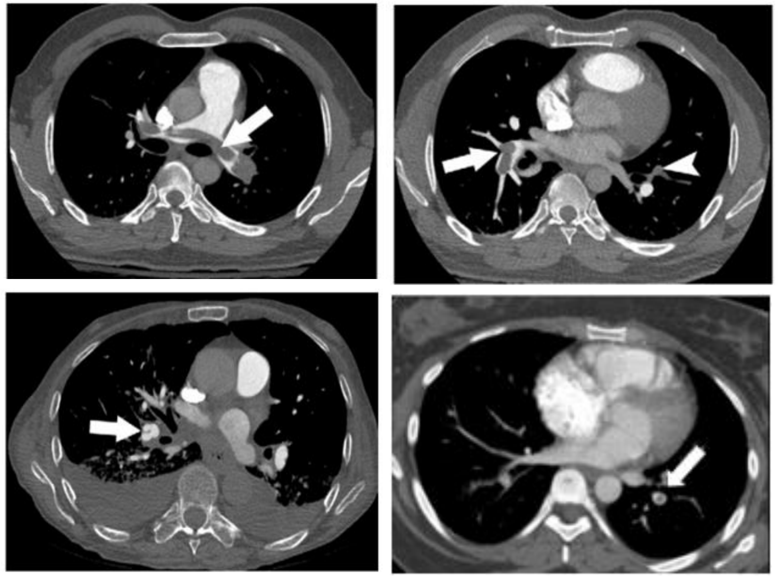


CT Pulmonary Angiography

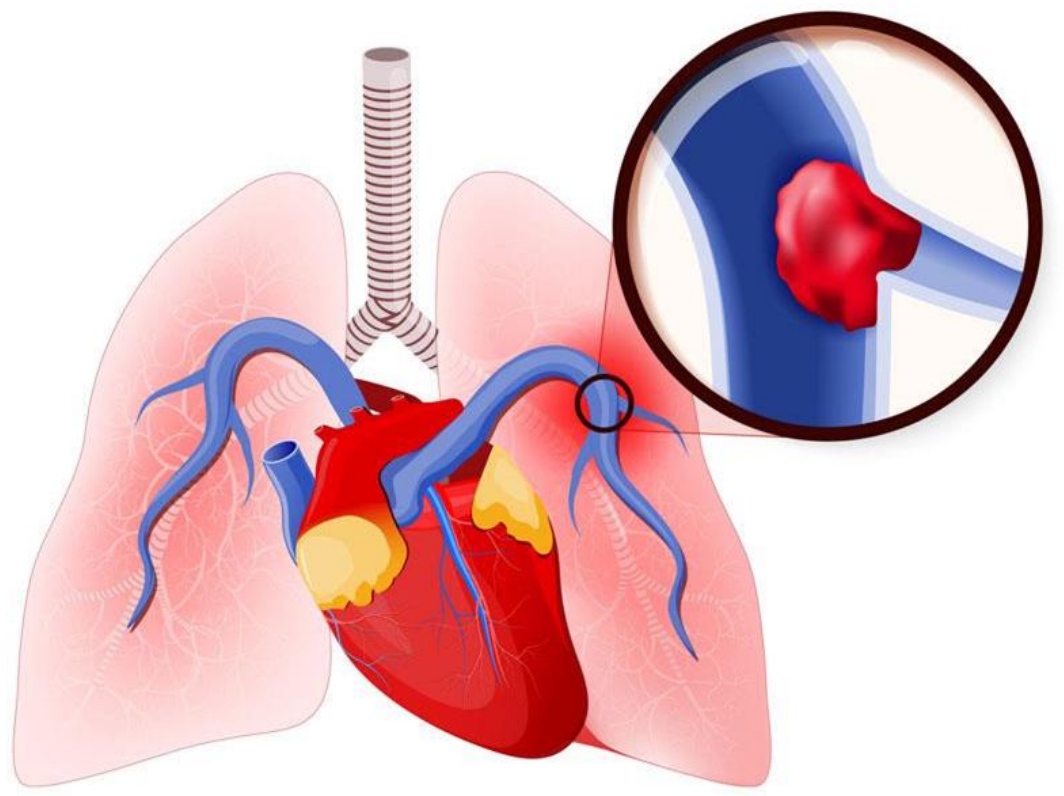


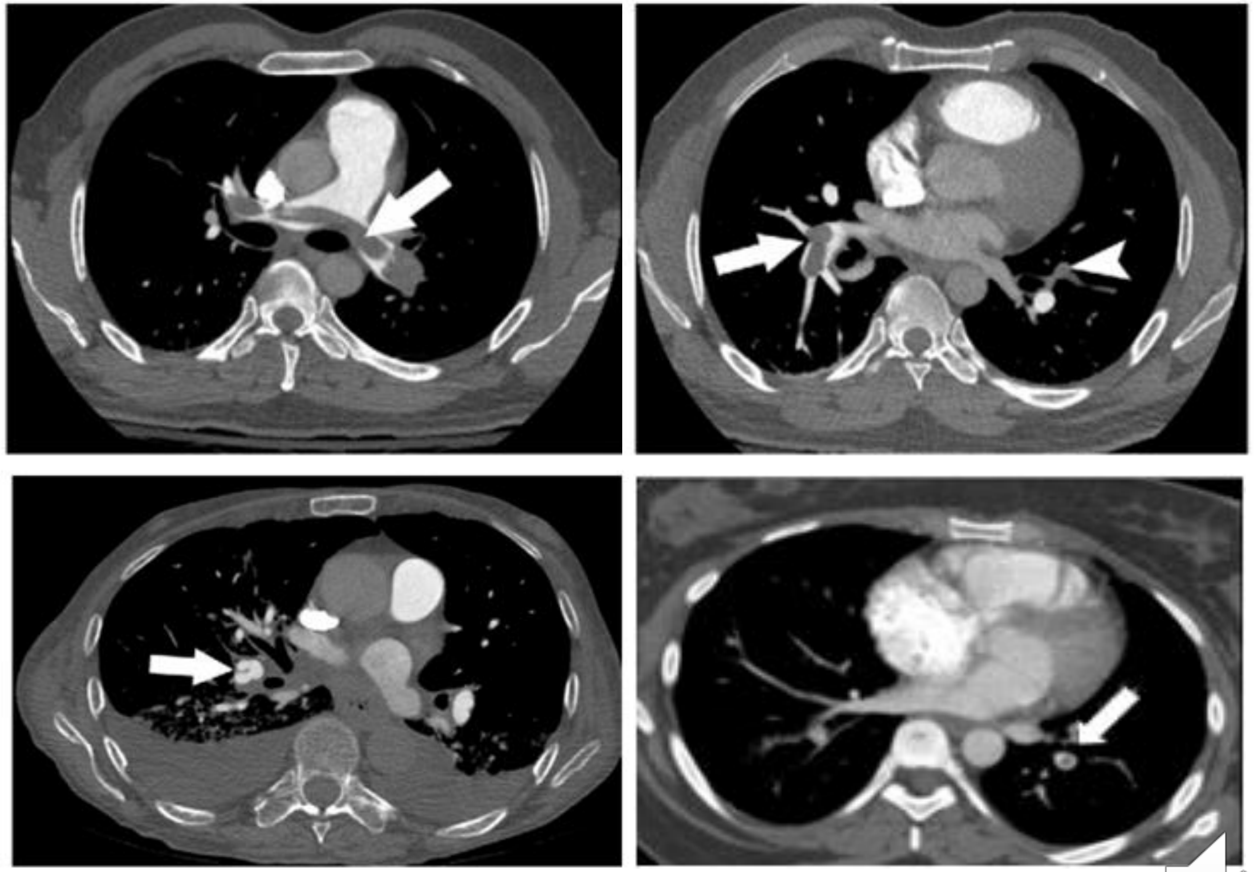
Slices Containing Pulmonary Embolism



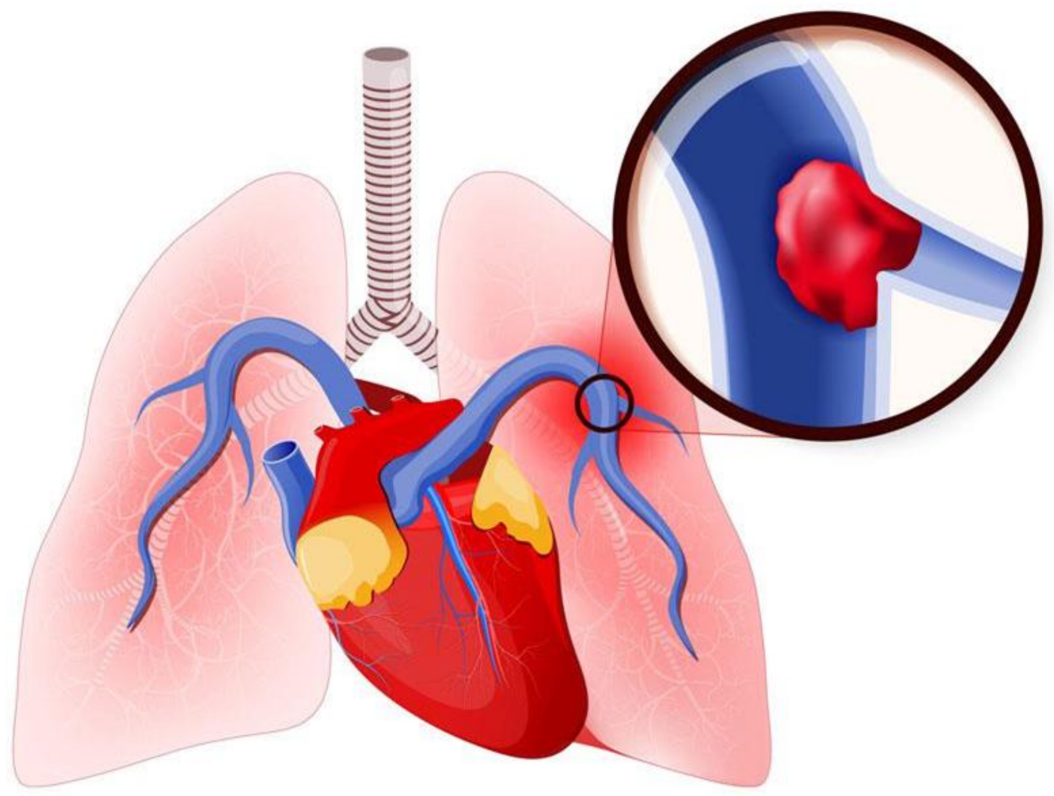


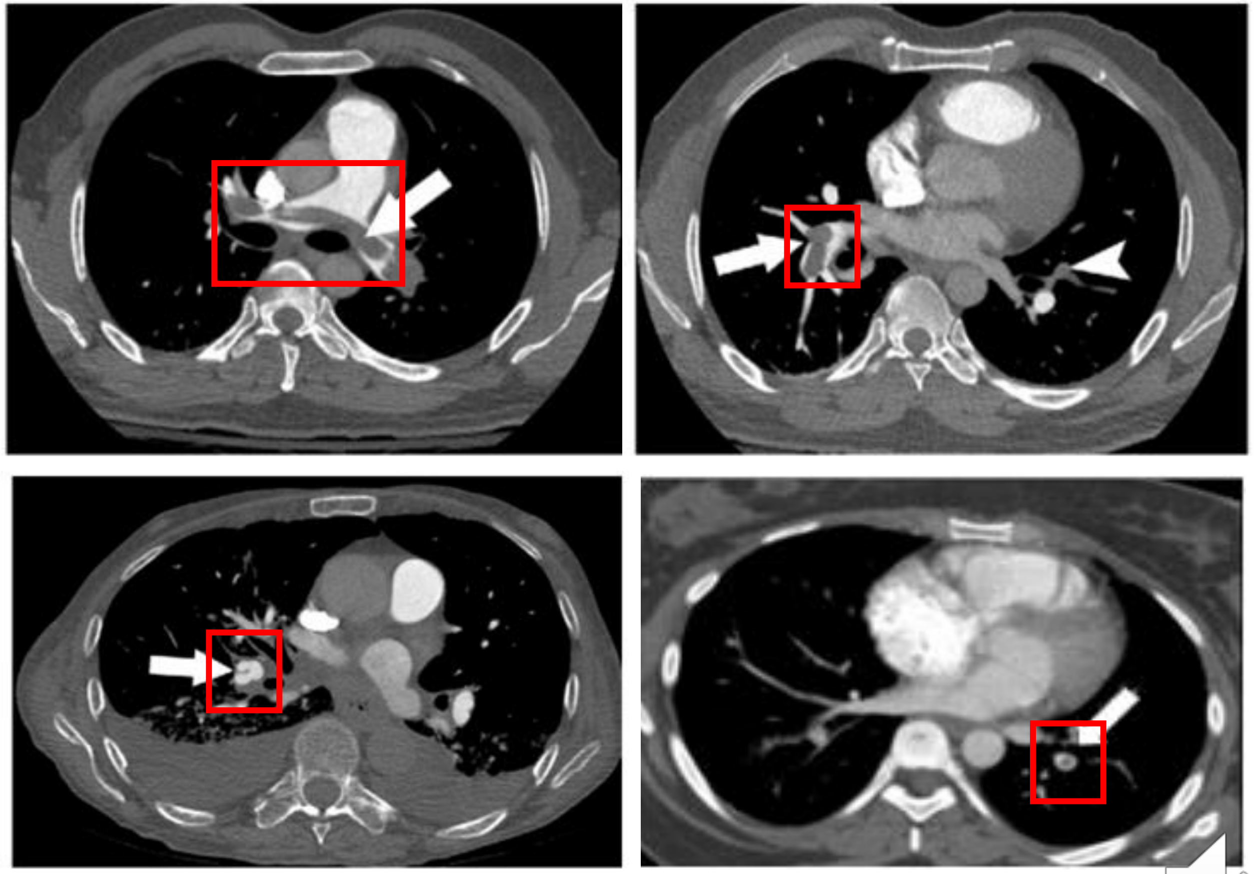
Slices Containing Pulmonary Embolism



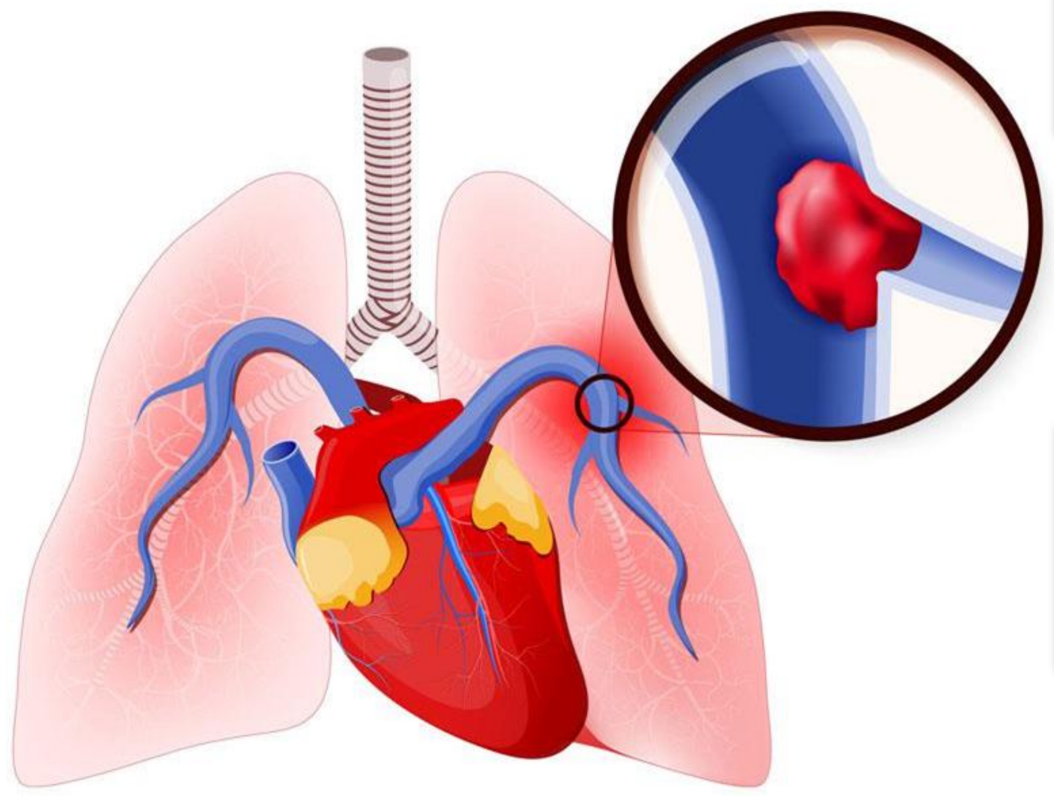


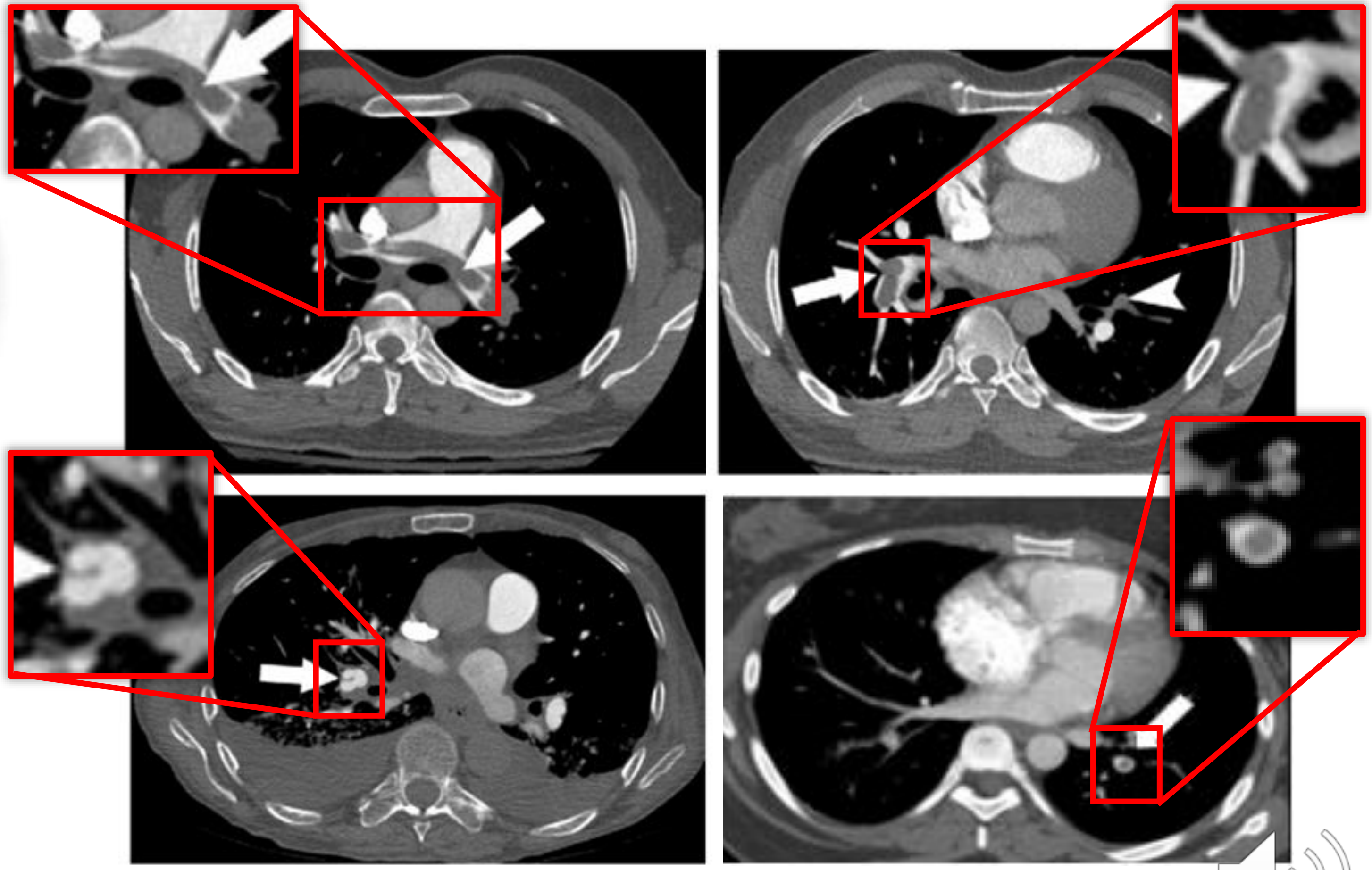
Slices Containing Pulmonary Embolism





Slices Containing Pulmonary Embolism





Pulmonary Embolism Dataset

RSNA Pulmonary Embolism Dataset

FUMPE Dataset

CAD-PE Challenge Dataset

In-house Dataset

Pulmonary Embolism Dataset

RSNA Pulmonary Embolism Dataset

FUMPE Dataset

CAD-PE Challenge Dataset

In-house Dataset

Exams: 7,279

Slices: 1,790,624

Task:
Slice-level Classification
Exam-level Diagnosis

Ground Truth:

Slice-level: PE present or not

Exam-level:

1. Negative Exam for PE
2. Indeterminate
3. Chronic PE
4. Acute & Chronic PE
5. Central PE
6. Left PE
7. Right PE
8. RV LV Ratio > 1
9. RV LV Ratio < 1

Colak, E., Kitamura, F. C., Hobbs, S. B., Wu, C. C., Lungren, M. P., Prevedello, L. M., … & RSNA-STR Annotators and Dataset Curation Contributors. (2021). The RSNA pulmonary embolism CT dataset. Radiology: Artificial Intelligence, 3(2), e200254.

Pulmonary Embolism Dataset

RSNA Pulmonary Embolism Dataset

FUMPE Dataset

CAD-PE Challenge Dataset

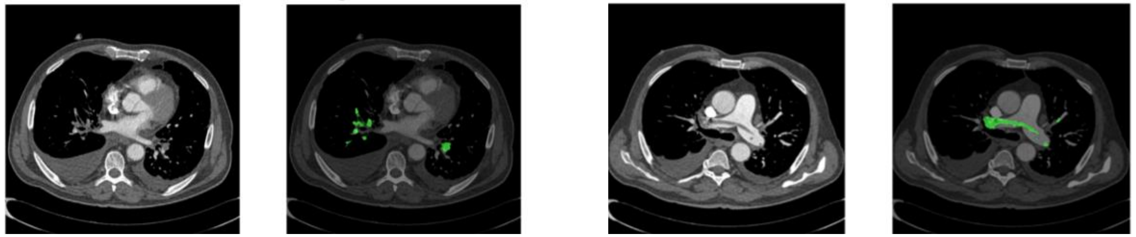
In-house Dataset

Exams: 35

Slices: 8,792

Task:
Slice-level Classification

Ground Truth:



Masoudi, M., Pourreza, H. R., Saadatmand-Tarzjan, M., Eftekhari, N., Zargar, F. S., & Rad, M. P. (2018). A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. Scientific data, 5(1), 1-9.

Pulmonary Embolism Dataset

RSNA Pulmonary Embolism Dataset

FUMPE Dataset

CAD-PE Challenge Dataset

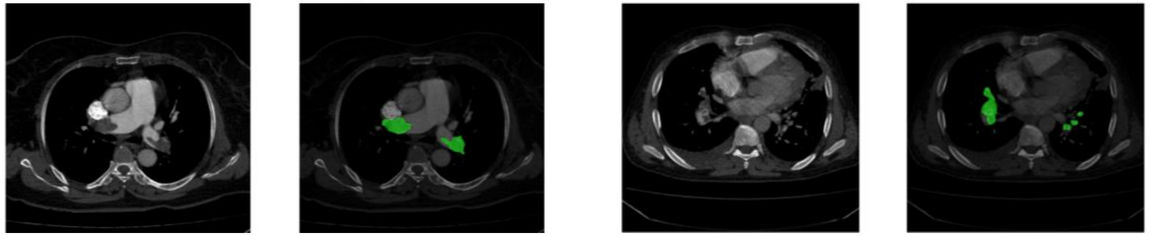
In-house Dataset

Exams: 91

Slices: 41,256

Task:
Slice-level Classification

Ground Truth:



González, G., Jimenez-Carretero, D., Rodríguez-López, S., Cano-Espinosa, C., Cazorla, M., Agarwal, T., ... & Ledesma-Carbayo, M. J. (2020). Computer aided detection for pulmonary embolism challenge (CAD-PE). arXiv preprint arXiv:2003.13440.

Pulmonary Embolism Dataset

RSNA Pulmonary Embolism Dataset

FUMPE Dataset

CAD-PE Challenge Dataset

In-house Dataset

Exams: 121

Emboli: 326

Task:
False Positive Reduction

Ground Truth:
Clot level

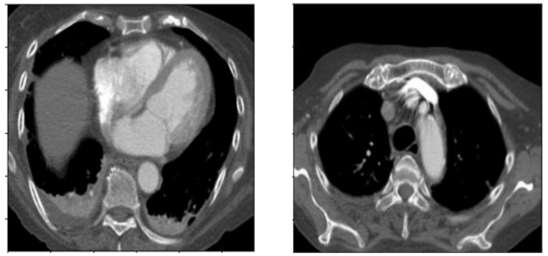
Tajbakhsh, N., Shin, J. Y., Gotway, M. B., & Liang, J. (2019). Computer-aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation. Medical image analysis, 58, 101541.

Classification and Diagnosis Task

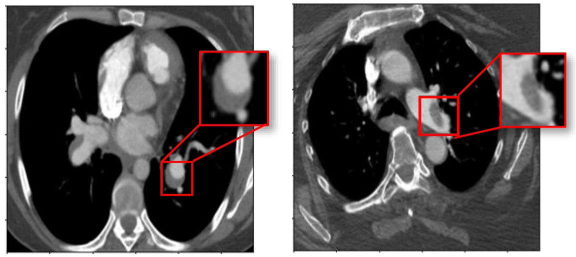
Exams: 7,279 (with 1,790,624 slices)

Ground Truth

1. PE present or not



PE absent



PE present

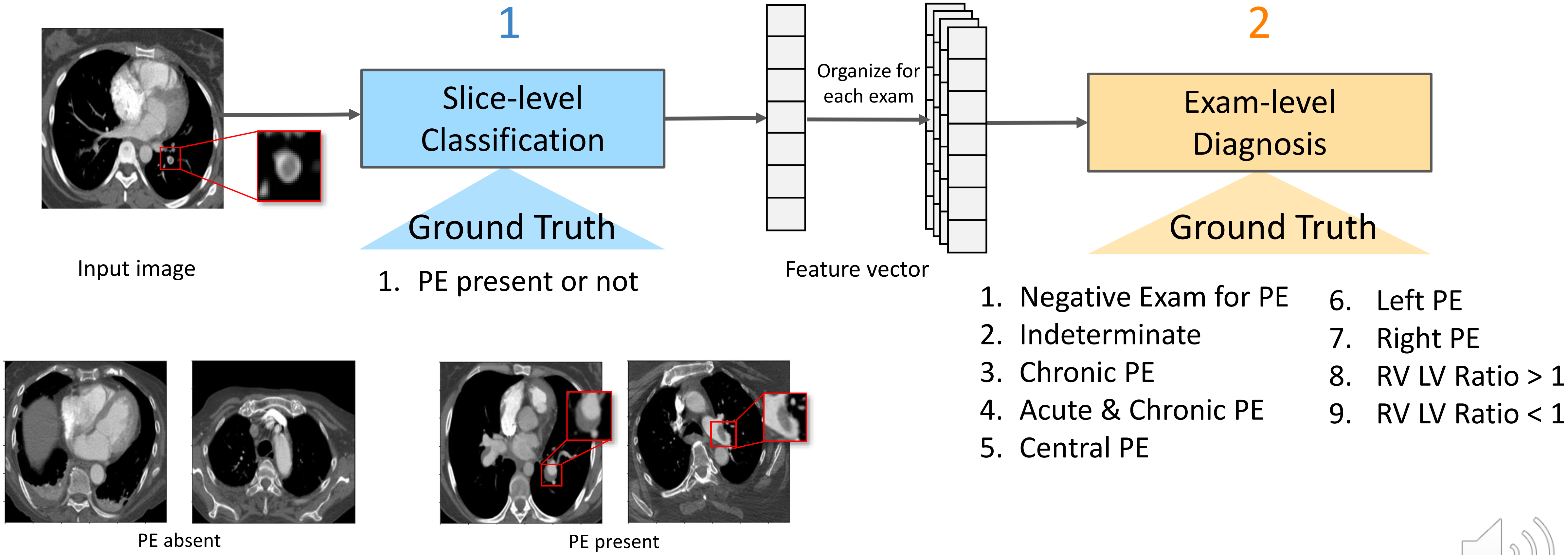
Ground Truth

1. Negative Exam for PE
2. Indeterminate
3. Chronic PE
4. Acute & Chronic PE
5. Central PE
6. Left PE
7. Right PE
8. RV LV Ratio > 1
9. RV LV Ratio < 1

Colak, E., Kitamura, F. C., Hobbs, S. B., Wu, C. C., Lungren, M. P., Prevedello, L. M., Kalpathy-Cramer, J., Ball, R. L., Shih, G., Stein, A., Halabi, S. S., Altinmakas, E., Law, M., Kumar, P., Manzalawi, K. A., Nelson Rubio, D. C., Sechrist, J. W., Germaine, P., Lopez, E. C., Amerio, T., ... RSNA-STR Annotators and Dataset Curation Contributors (2021). The RSNA Pulmonary Embolism CT Dataset. *Radiology. Artificial intelligence*, *3*(2), e200254. https://doi.org/10.1148/ryai.2021200254

Classification and Diagnosis Task

Exams: 7,279 (with 1,790,624 slices)

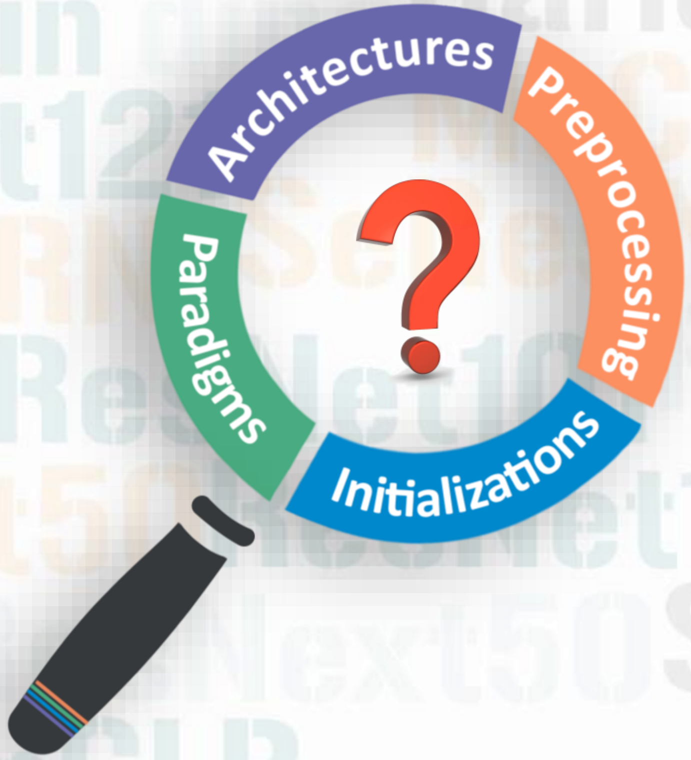


Colak, E., Kitamura, F. C., Hobbs, S. B., Wu, C. C., Lungren, M. P., Prevedello, L. M., Kalpathy-Cramer, J., Ball, R. L., Shih, G., Stein, A., Halabi, S. S., Altinmakas, E., Law, M., Kumar, P., Manzalawi, K. A., Nelson Rubio, D. C., Sechrist, J. W., Germaine, P., Lopez, E. C., Amerio, T., … RSNA-STR Annotators and Dataset Curation Contributors (2021). The RSNA Pulmonary Embolism CT Dataset. *Radiology. Artificial intelligence*, *3*(2), e200254. https://doi.org/10.1148/ryai.2021200254

InfoMin
dino
BarlowTwins
DenseNet121
MoCo
SNet154
clsa
PIRL
DRN
SeResNet50
obow
simsiam
ResNet101
ResNet50
BYOL
ResNext50
ResNet18
Xception
PCL
SeResNext50
SeResNet101
SWAV
SimCLR
InsDis
Sela
DeepCluster

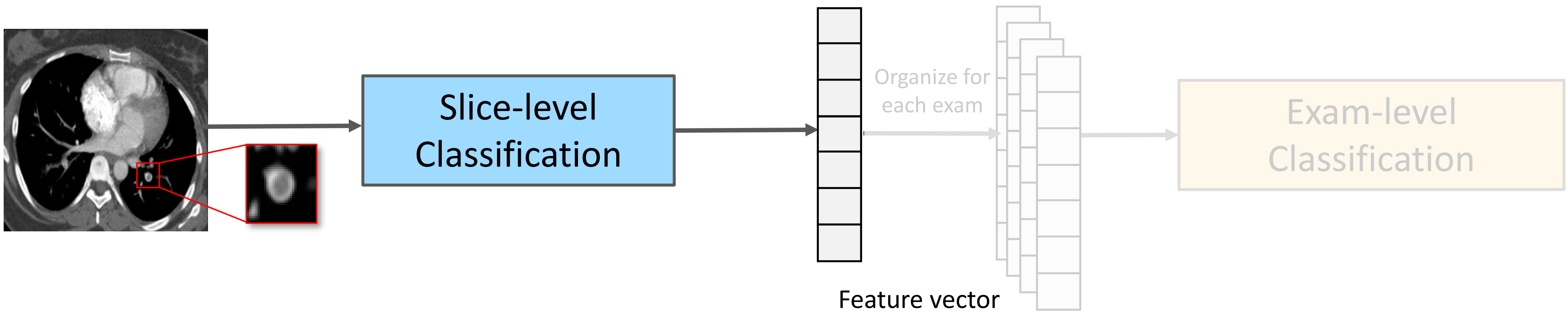
InfoMin
dino
BarlowTwins
MoCo
SNet154
DenseNet121
SeResNet50
obow
clsa
PIRL
DRN
ResNet101
ResNet50
simsiam
Xception
BYOL
ResNext50
ResNet18
PCL
SeResNext50
SeResNet101
SWAV
SimCLR
InsDis
Sela
DeepCluster

InfoMin
dino
BarlowTwins
MoCo
SNet154
DenseNet121
DRN
Net50obow
clsa
PIRL
simsiam
ResNet101
ResNet50
BYOL
ResNext50
ResNet18
Xception
PCL
SeResNext50
SeResNet101
SWAV
SimCLR
InsDis
Sela
DeepCluster

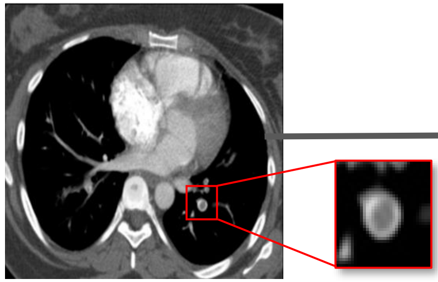


for computer-aided pulmonary embolism detection

First Stage


Slice-level Classification
Organize for each exam
Feature vector
Exam-level Classification

First Stage



Slice-level Classification

Feature vector

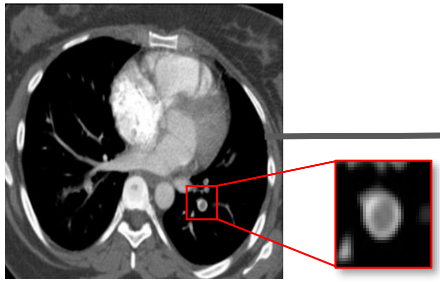
Organize for each exam

Exam-level Classification

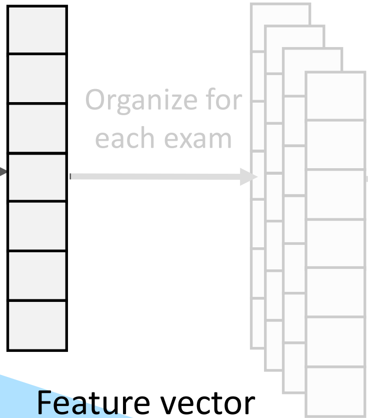
- Training from scratch
- Fine-tuning
  - Supervised pretrained models
  - Self-supervised pretrained models

https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection/discussion/194145

First Stage



Slice-level Classification



Exam-level Classification

Convolutional Neural Networks

1. ResNet18
2. ResNet50
3. ResNext50
4. SeResNet50
5. SeResNext50
6. DRN-A-50
7. Xception
8. SeXception
9. DenseNet121
10. SeNet154
11. ResNet101
12. SeResNet101

Transformers Based Models

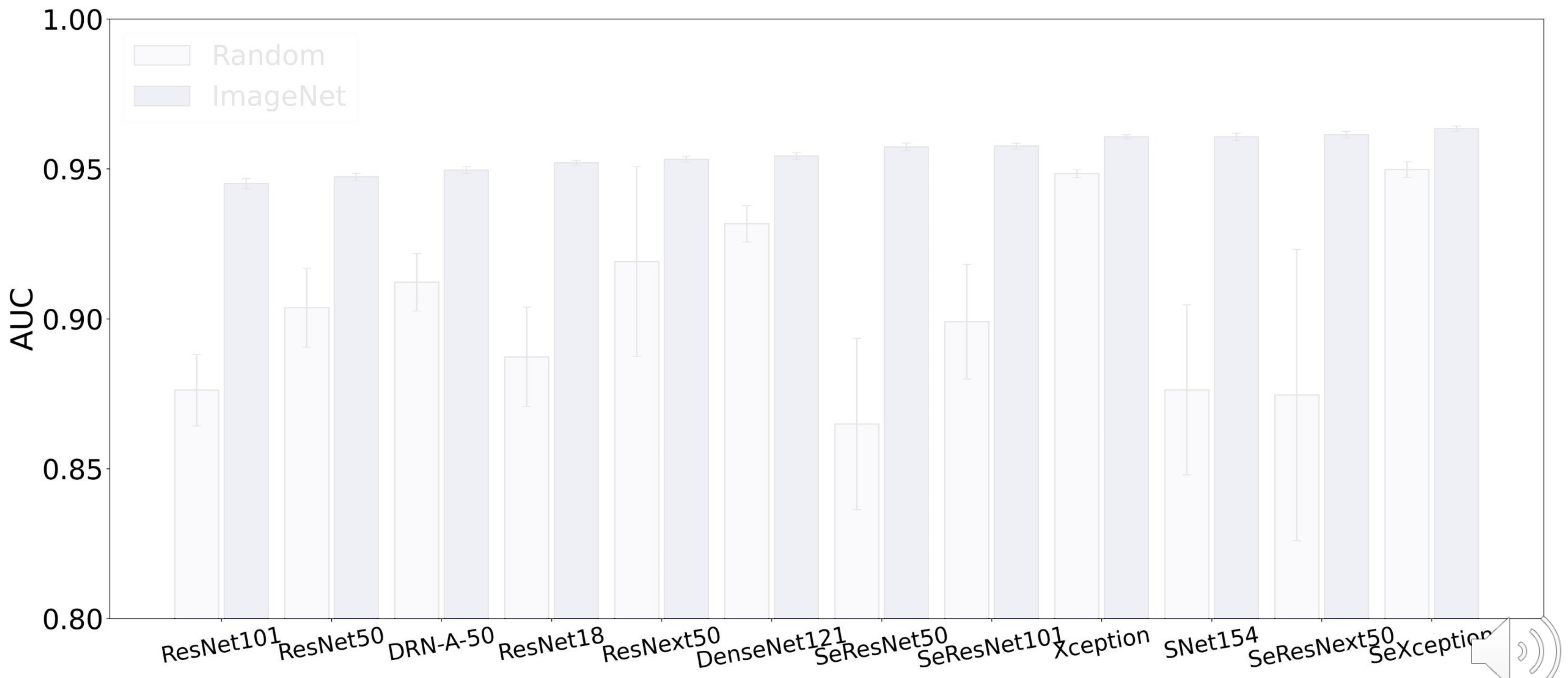
1. ViT
2. Swin

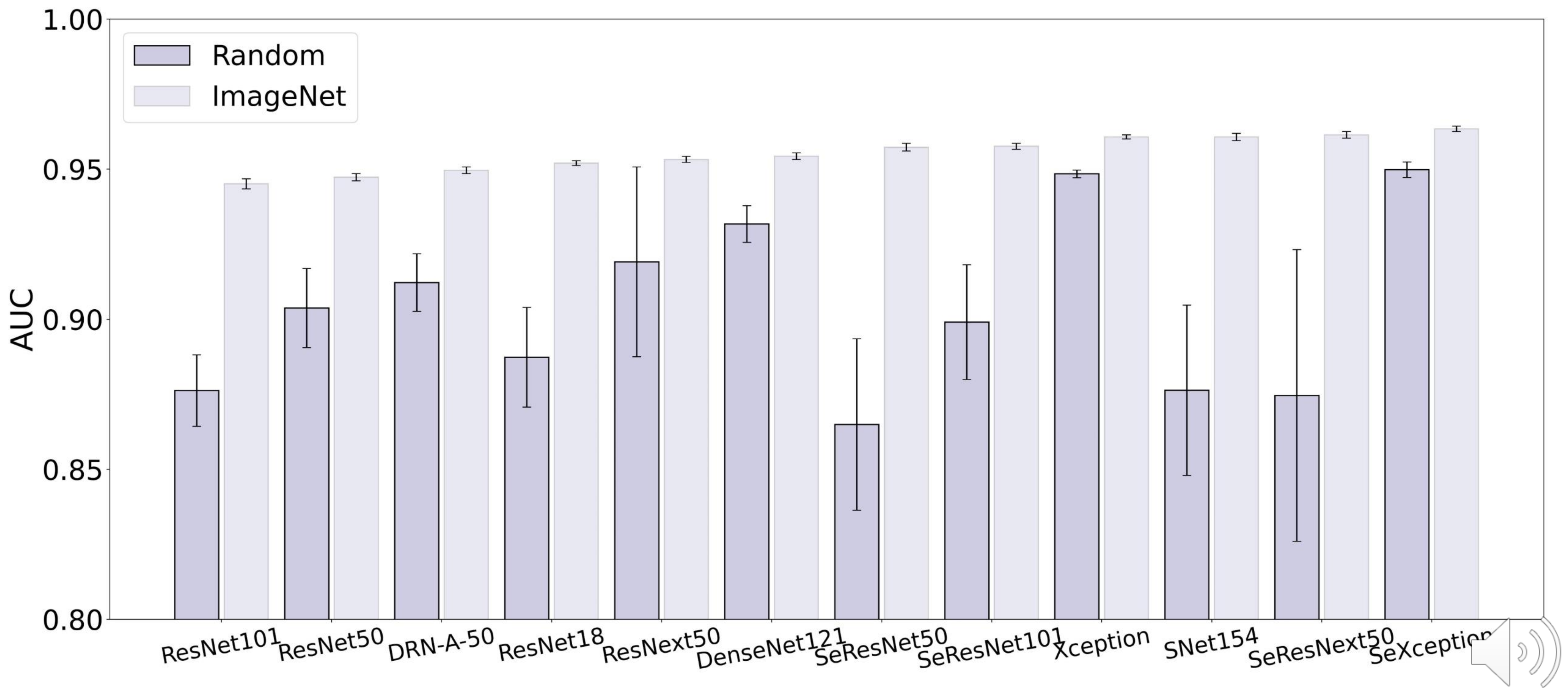
- Training from scratch
- Fine-tuning
  - Supervised pretrained models
  - Self-supervised pretrained models

Self-supervised Methods

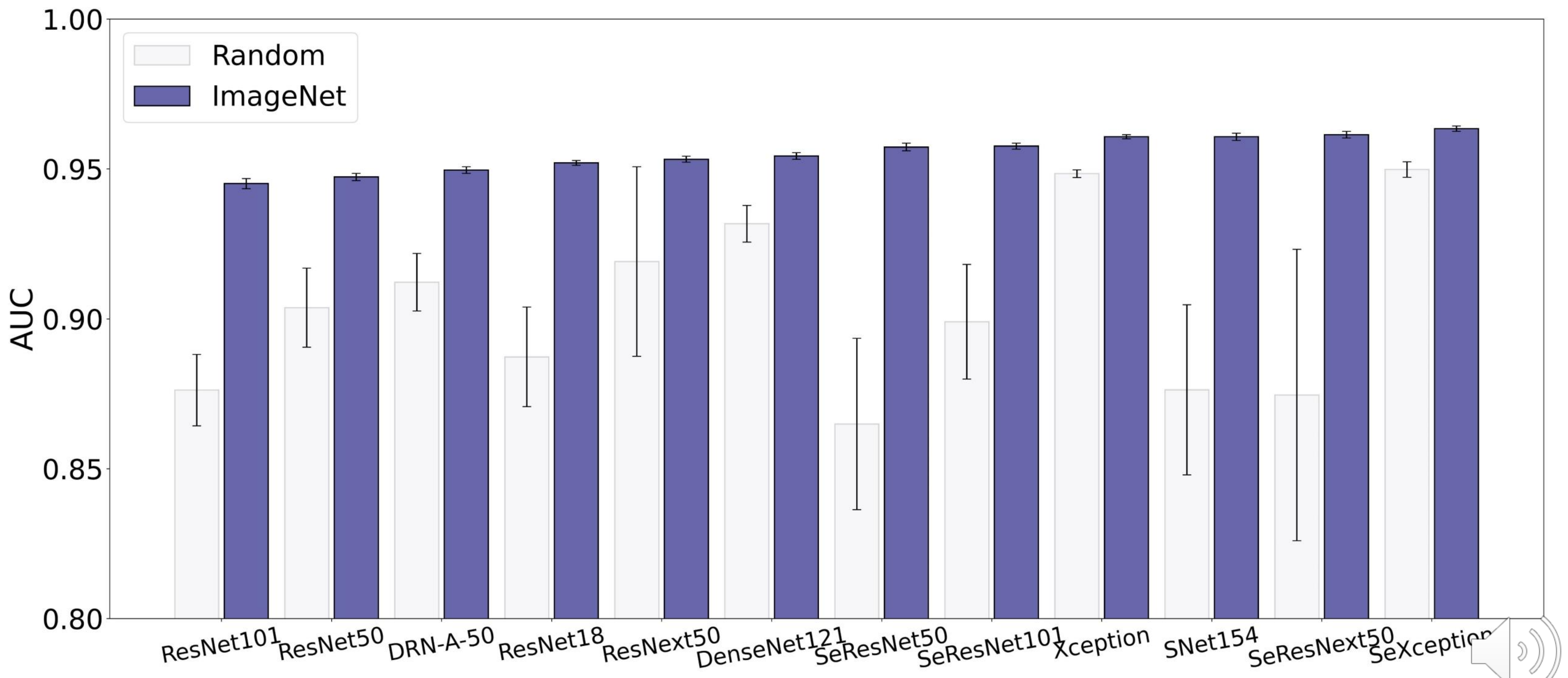
1. InsDis
2. InfoMin
3. MoCo-v1
4. BYOL
5. MoCo-v2
6. DeepCluster-v2
7. PCL-v1
8. SwAV
9. PCL-v2
10. SimCLR-v1
11. PIRL
12. SimCLR-v2
13. SeLa-v2
14. Barlow Twins
15. BYOL
16. SwAV
17. dino
18. clsa
19. obow

https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection/discussion/194145


1.00
0.95
0.90
0.85
0.80
AUC
Random
ImageNet
ResNet101
ResNet50
DRN-A-50
ResNet18
ResNext50
DenseNet121
SeResNet50
SeResNet101
Xception
SNet154
SeResNext50
SeXception

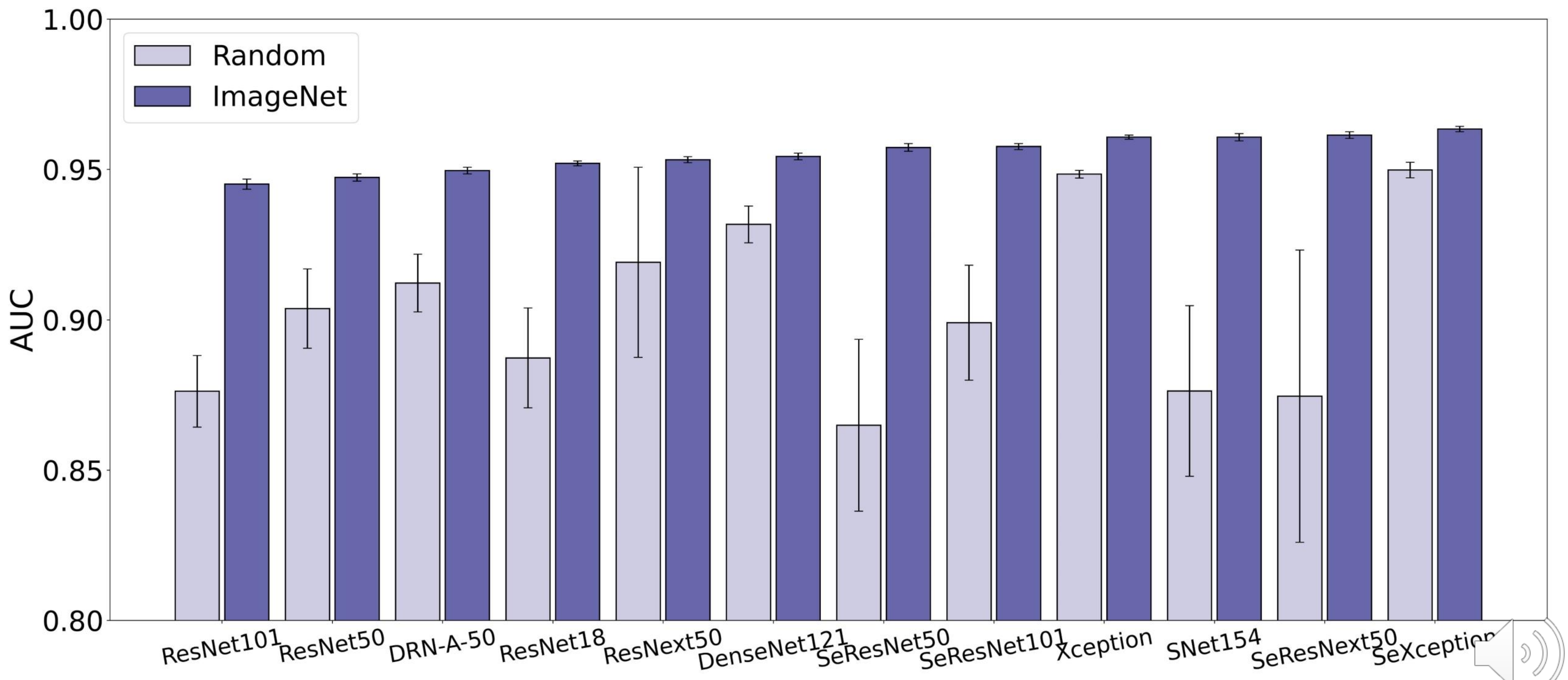


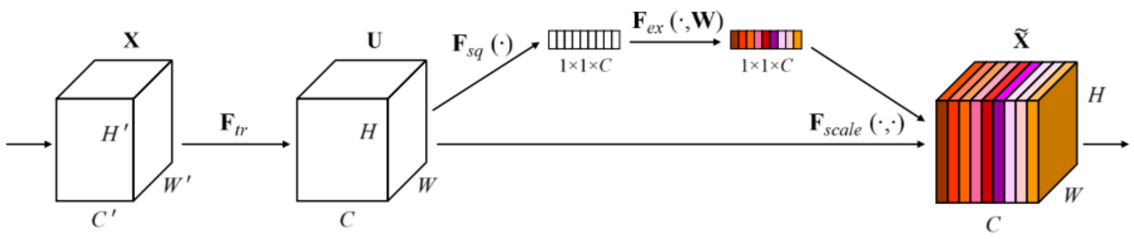
Random
ImageNet
AUC
1.00
0.95
0.90
0.85
0.80
ResNet101
ResNet50
DRN-A-50
ResNet18
ResNext50
DenseNet121
SeResNet50
SeResNet101
Xception
SNet154
SeResNext50

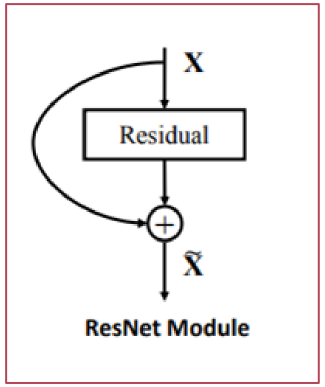


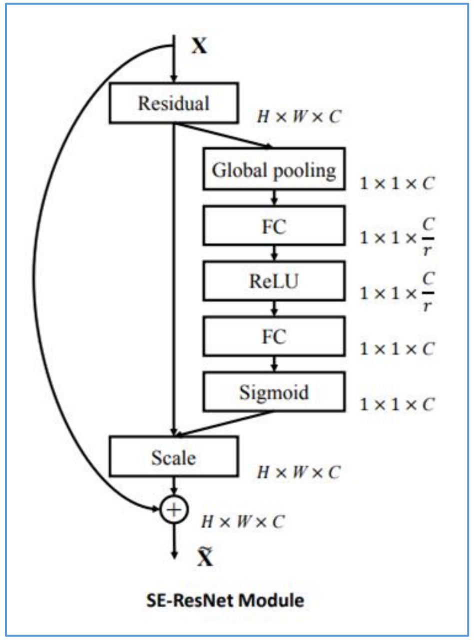
1.00
0.95
0.90
0.85
0.80
AUC
Random
ImageNet
ResNet101
ResNet50
DRN-A-50
ResNet18
ResNext50
DenseNet121
SeResNet50
SeResNet101
Xception
SNet154
SeResNext50
SeXception

Transfer learning consistently improves performance across the 12 different CNN architectures

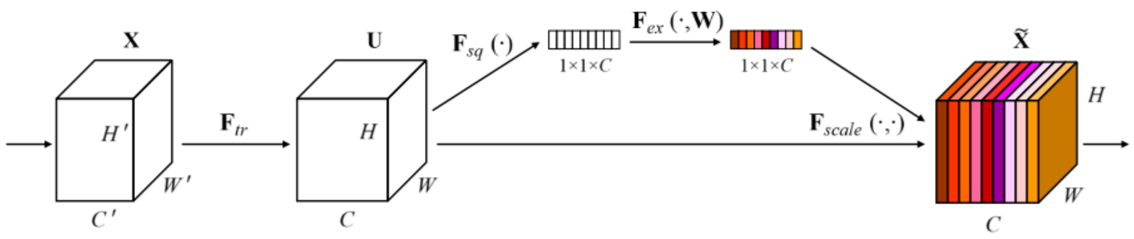


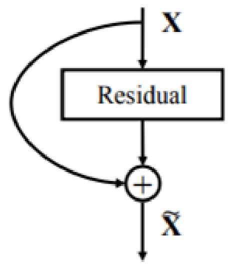




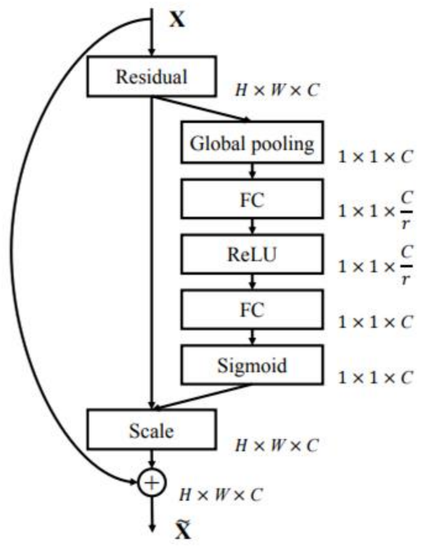


Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).





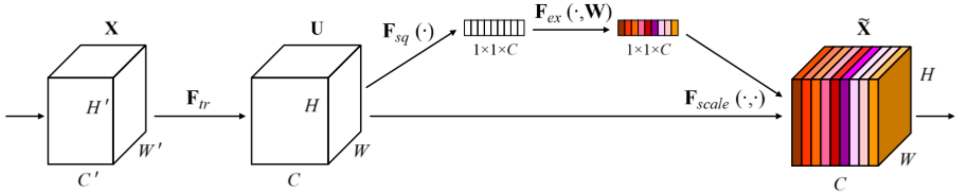
ResNet Module

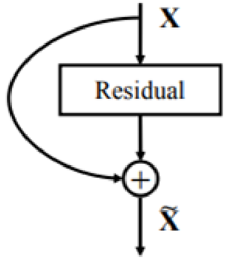


SE-ResNet Module

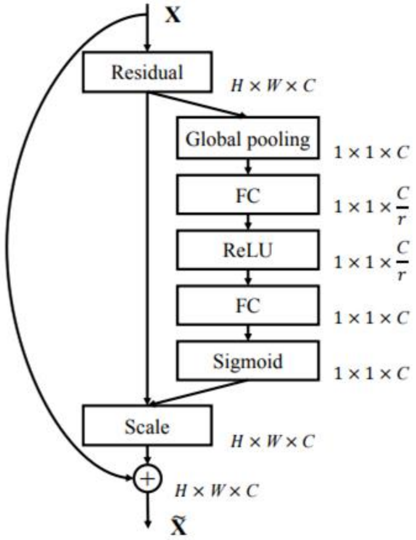
```
def se_block(in_block, ch, ratio=16):
    x = GlobalAveragePooling2D()(in_block)
    x = Dense(ch//ratio, activation='relu')(x)
    x = Dense(ch, activation='sigmoid')(x)
    return multiply()([in_block, x])
```

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

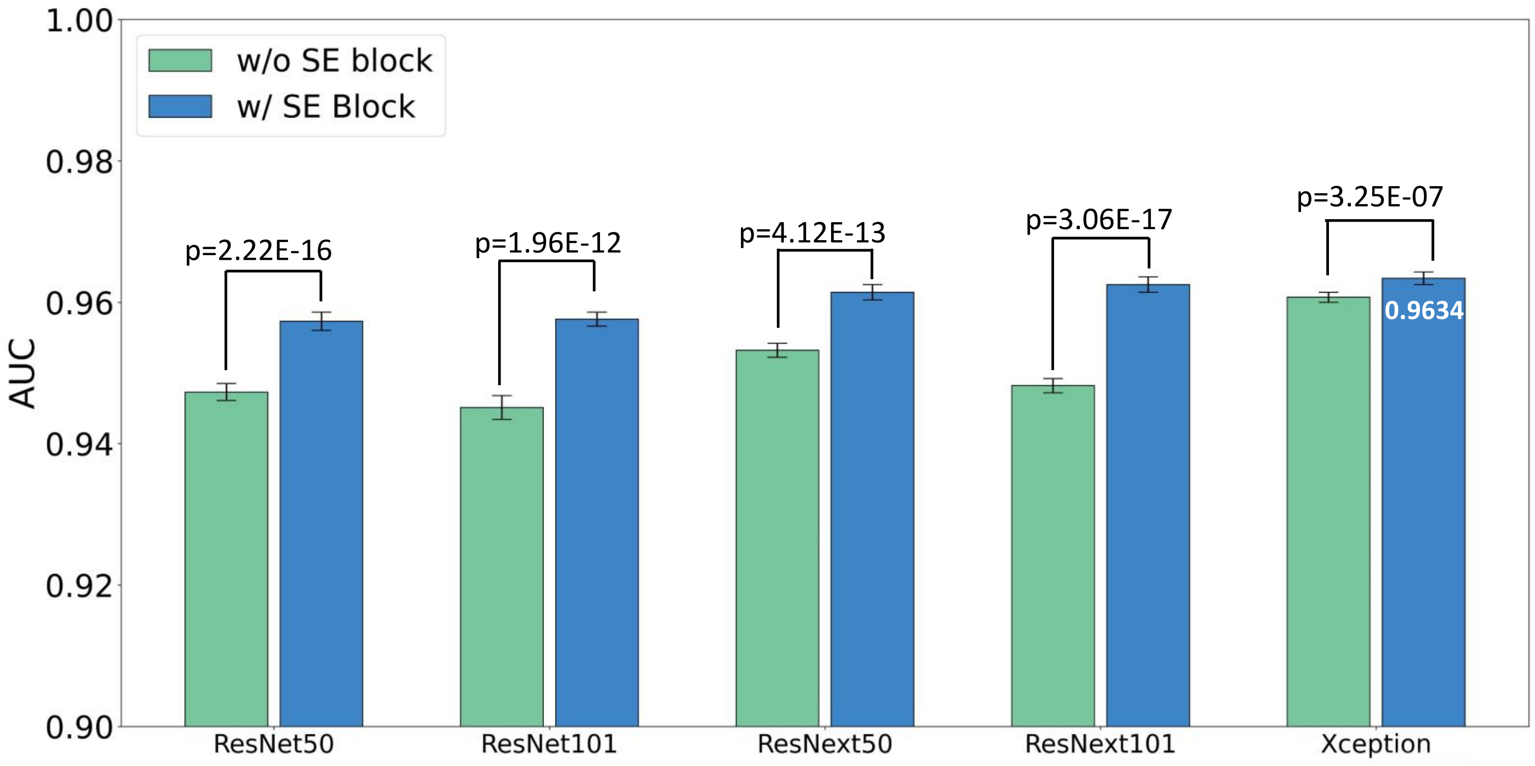




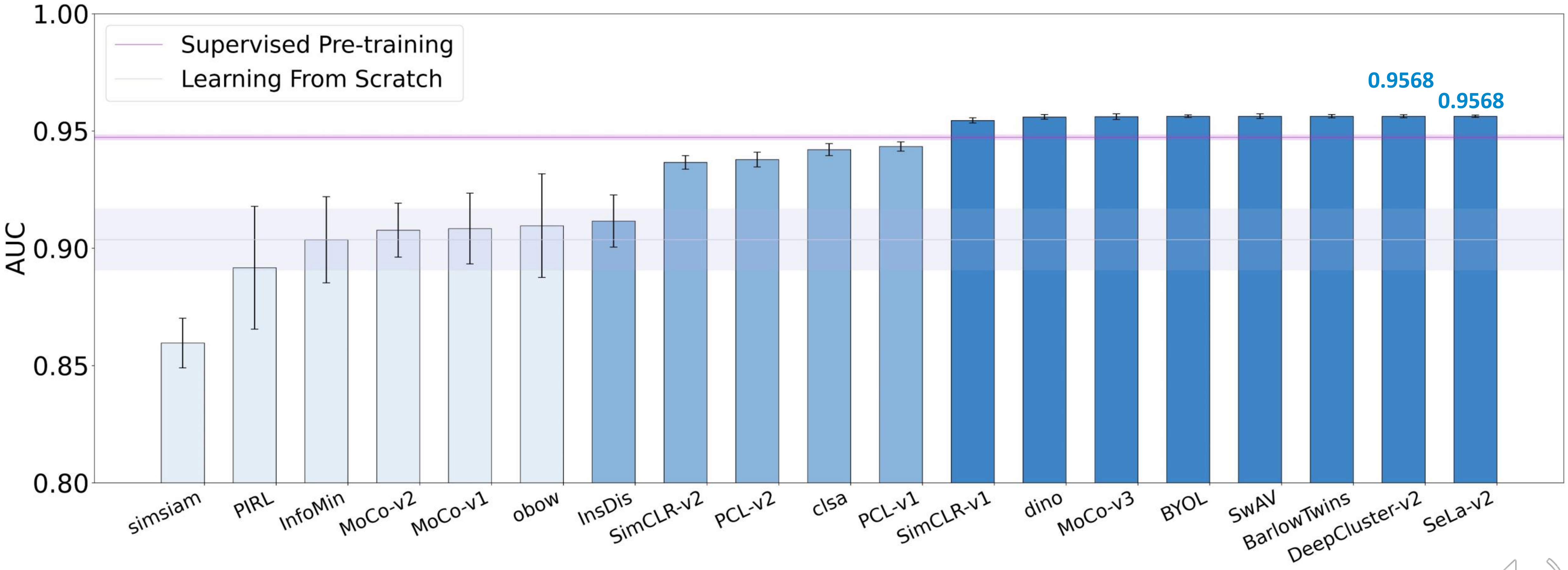
ResNet Module



SE-ResNet Module

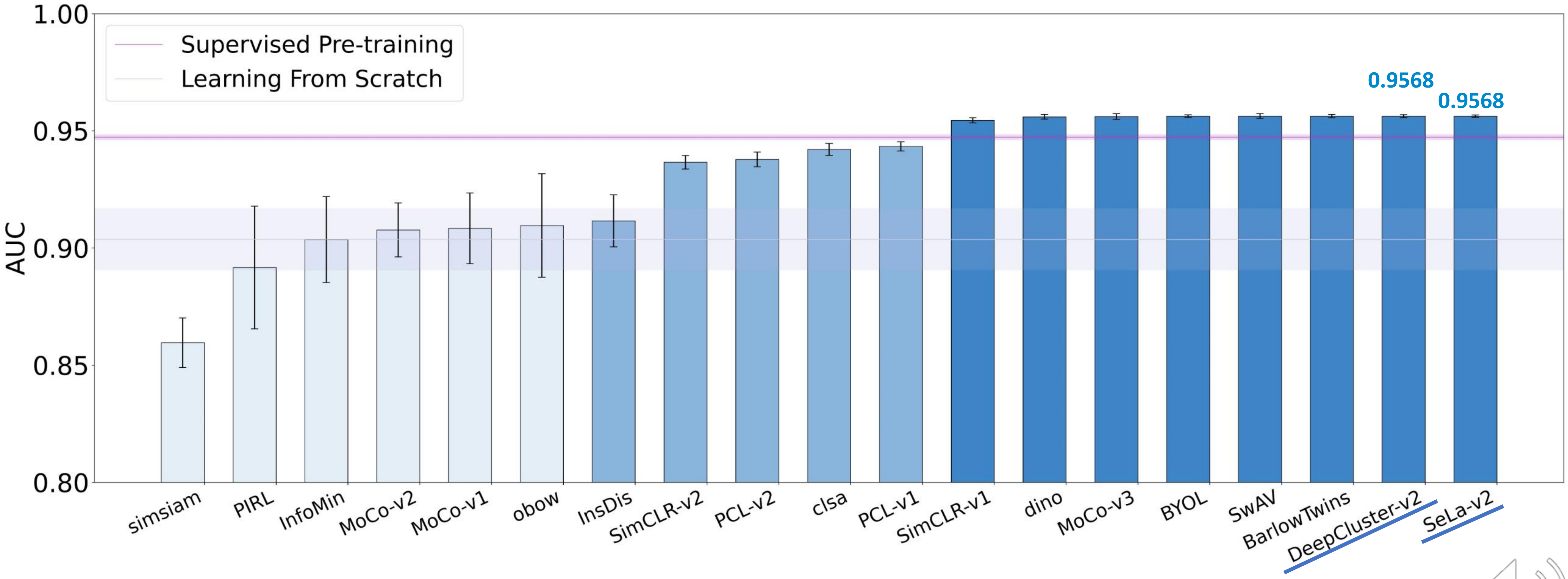


Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

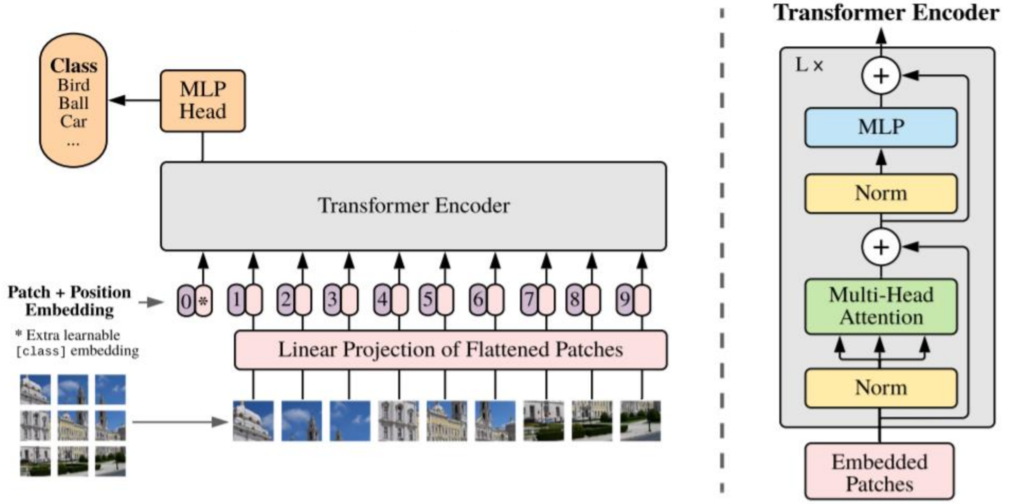


Ericsson, L., Gouk, H., & Hospedales, T. M. (2021). How Well Do Self-Supervised Models Transfer?. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5414-5423).

Self-supervised pre-training overtakes (fully) supervised pre-training



Ericsson, L., Gouk, H., & Hospedales, T. M. (2021). How Well Do Self-Supervised Models Transfer?. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5414-5423).

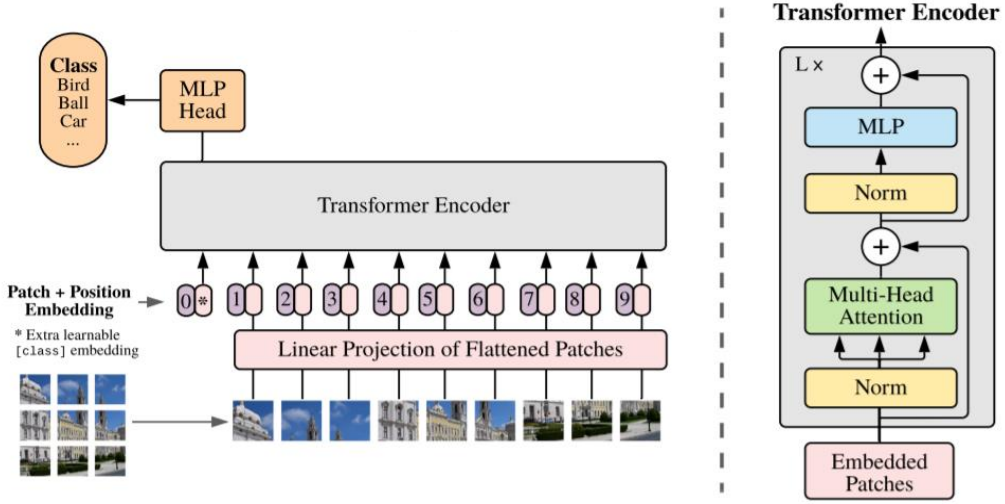


Vision Transformer

Backbone	Method	Pretrained Data	Slice-level AUC
SeXception	Supervised	ImageNet-1k	**0.9634**
ViT-B	Random	--	0.8468
	Supervised	ImageNet-1k	0.9131
	Supervised	ImageNet-21k	0.9119
	SimMIM	ImageNet-1k	0.9139
	MoBY	ImageNet-1k	0.9071

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

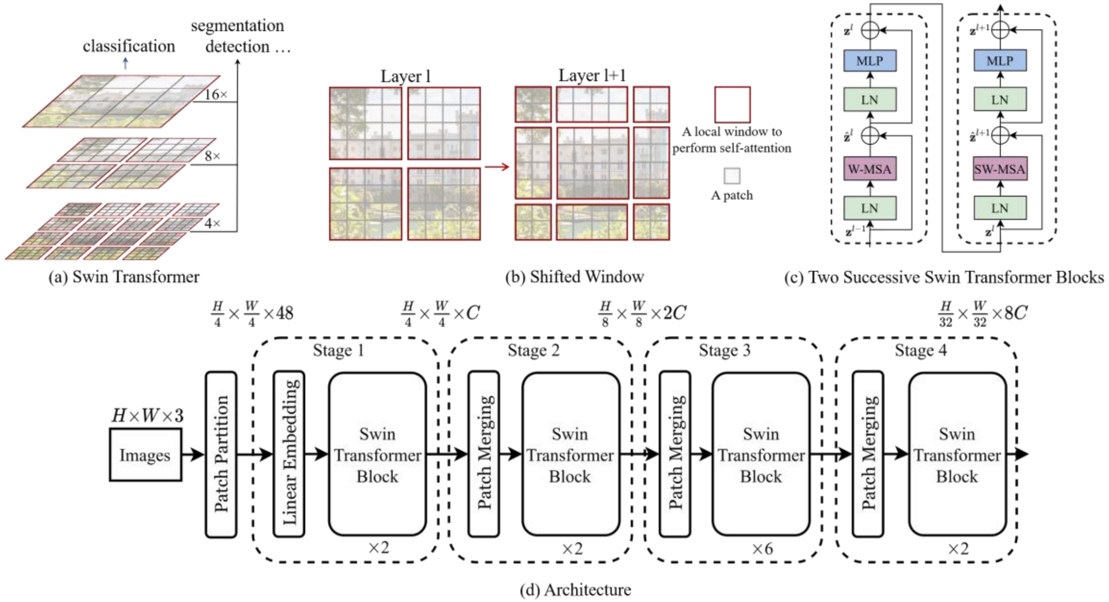
Vision transformer (ViT) underperform CNNs



Vision Transformer

Backbone	Method	Pretrained Data	Slice-level AUC
SeXception	Supervised	ImageNet-1k	**0.9634**
ViT-B	Random	--	0.8468
	Supervised	ImageNet-1k	0.9131
	Supervised	ImageNet-21k	0.9119
	SimMIM	ImageNet-1k	0.9139
	MoBY	ImageNet-1k	0.9071

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.



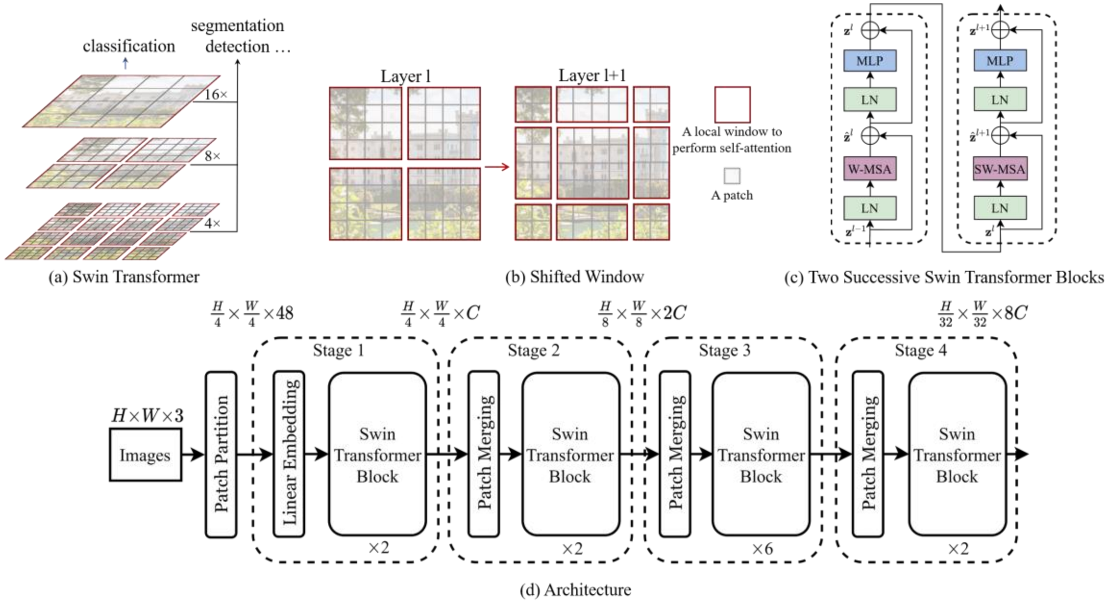
Swin Transformer

Backbone	Method	Pretrained Data	Slice-level AUC
SeXception	Supervised	ImageNet-1k	**0.9634**
Swin-B	Random	--	0.9063
	Supervised	ImageNet-1k	0.9485
	Supervised	ImageNet-21k	0.9458
	SimMIM	ImageNet-1k	**0.9527**
	MoBY	ImageNet-1k	0.9456

p=2.24E-14

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

Swin Transformer demonstrates a slower convergence rate compared to CNN.



Swin Transformer

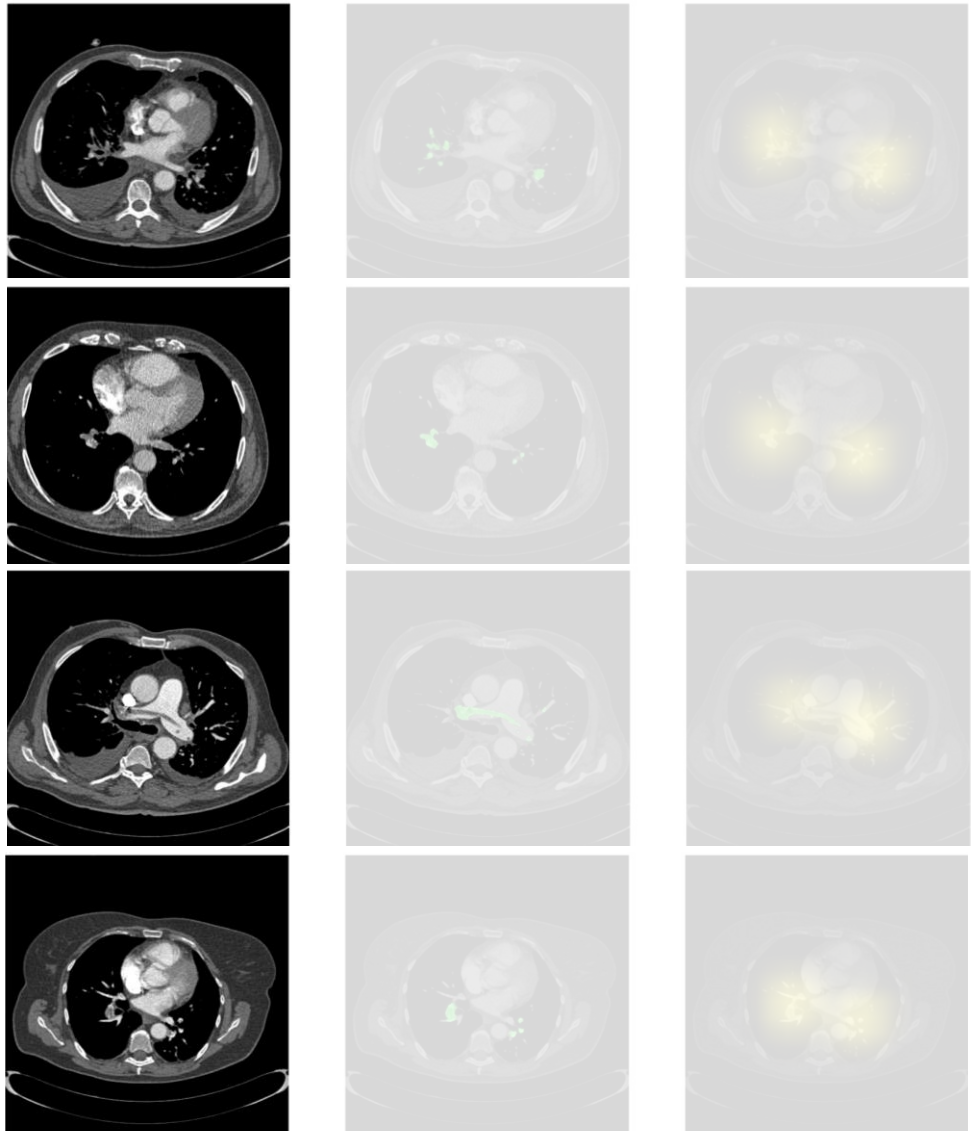
Backbone	Method	Pretrained Data	Slice-level AUC
SeXception	Supervised	ImageNet-1k	**0.9634**
Swin-B	Random	--	0.9063
	Supervised	ImageNet-1k	0.9485
	Supervised	ImageNet-21k	0.9458
	SimMIM	ImageNet-1k	**0.9527**
	MoBY	ImageNet-1k	0.9456

p=2.24E-14

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

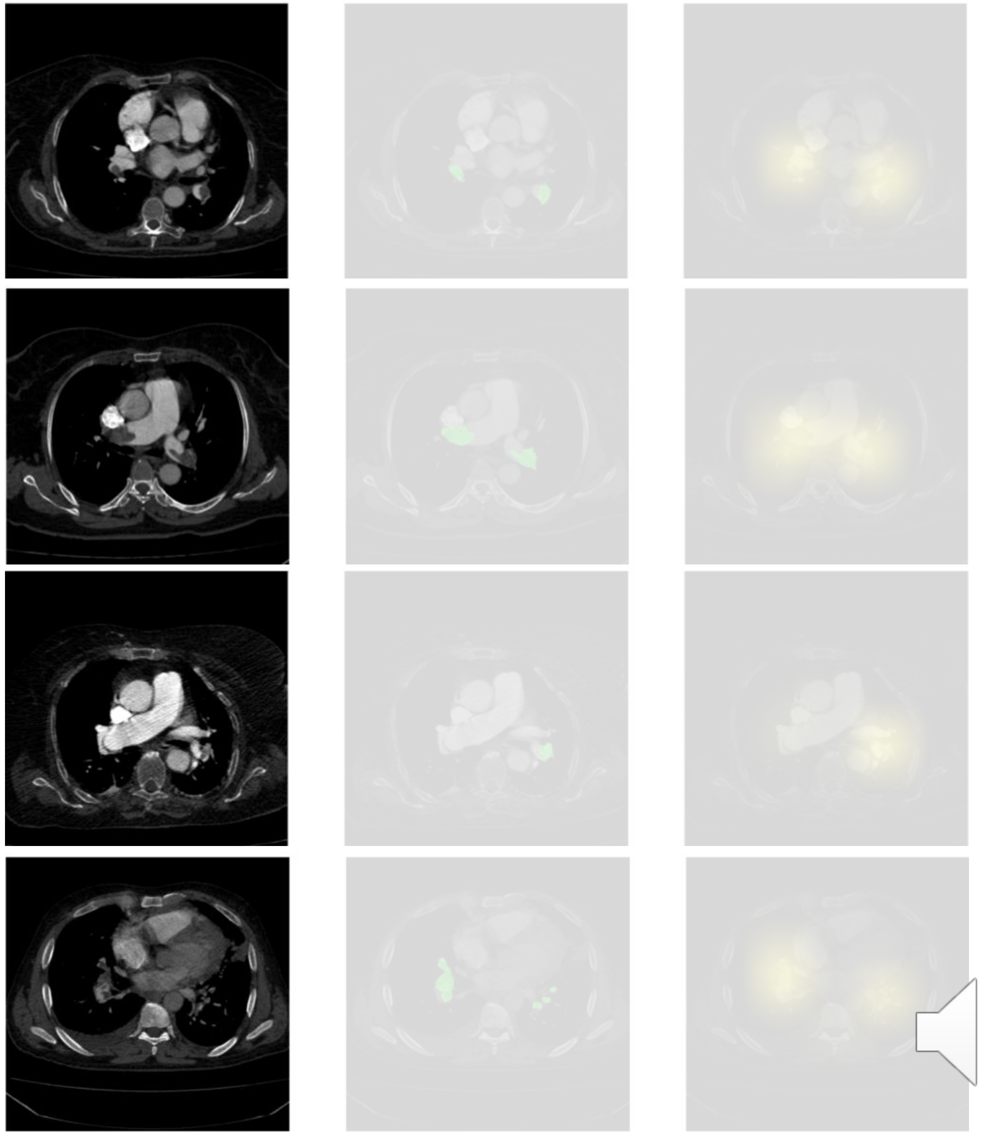
CAD-PE Challenge Dataset

Original | Ground Truth | Predicted



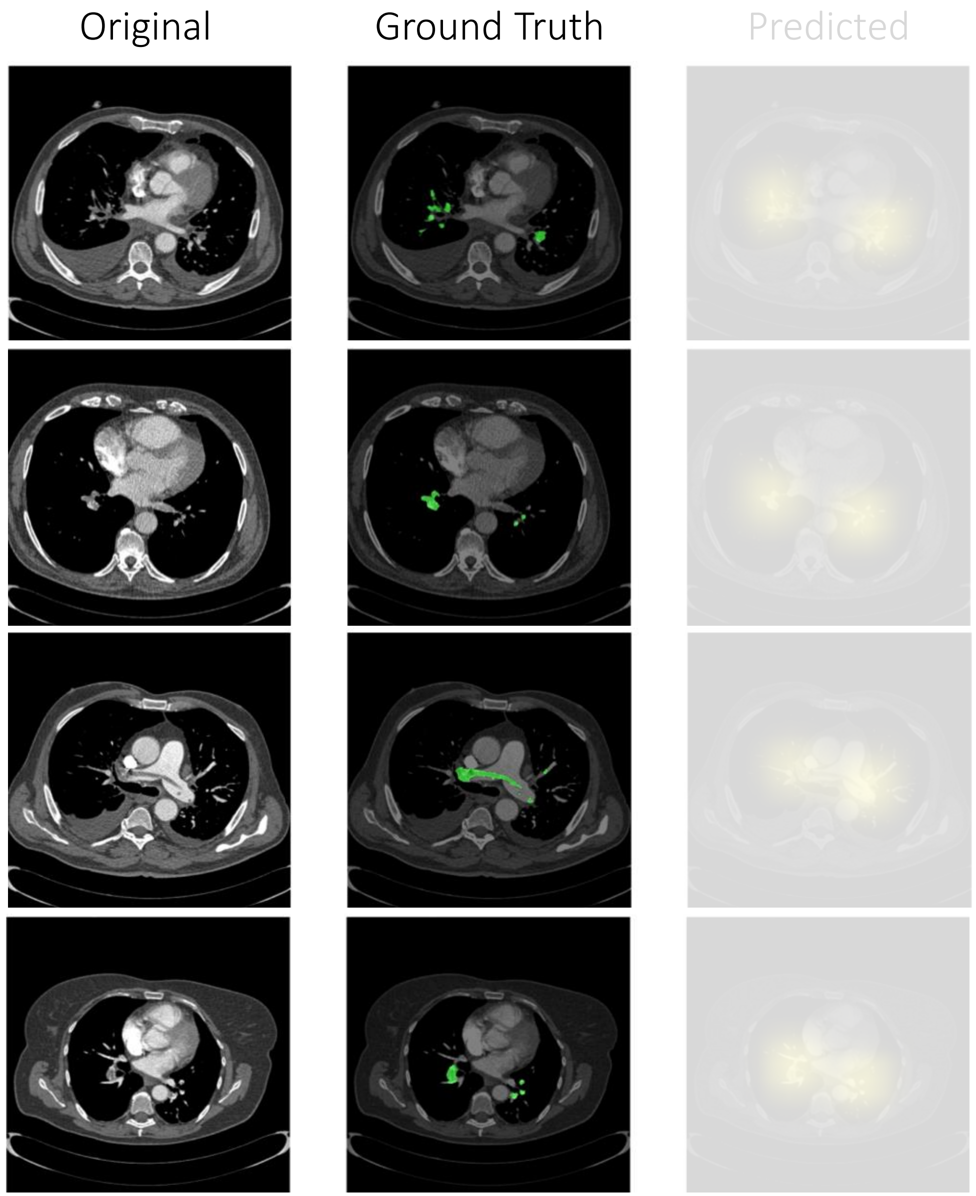
FUMPE Dataset

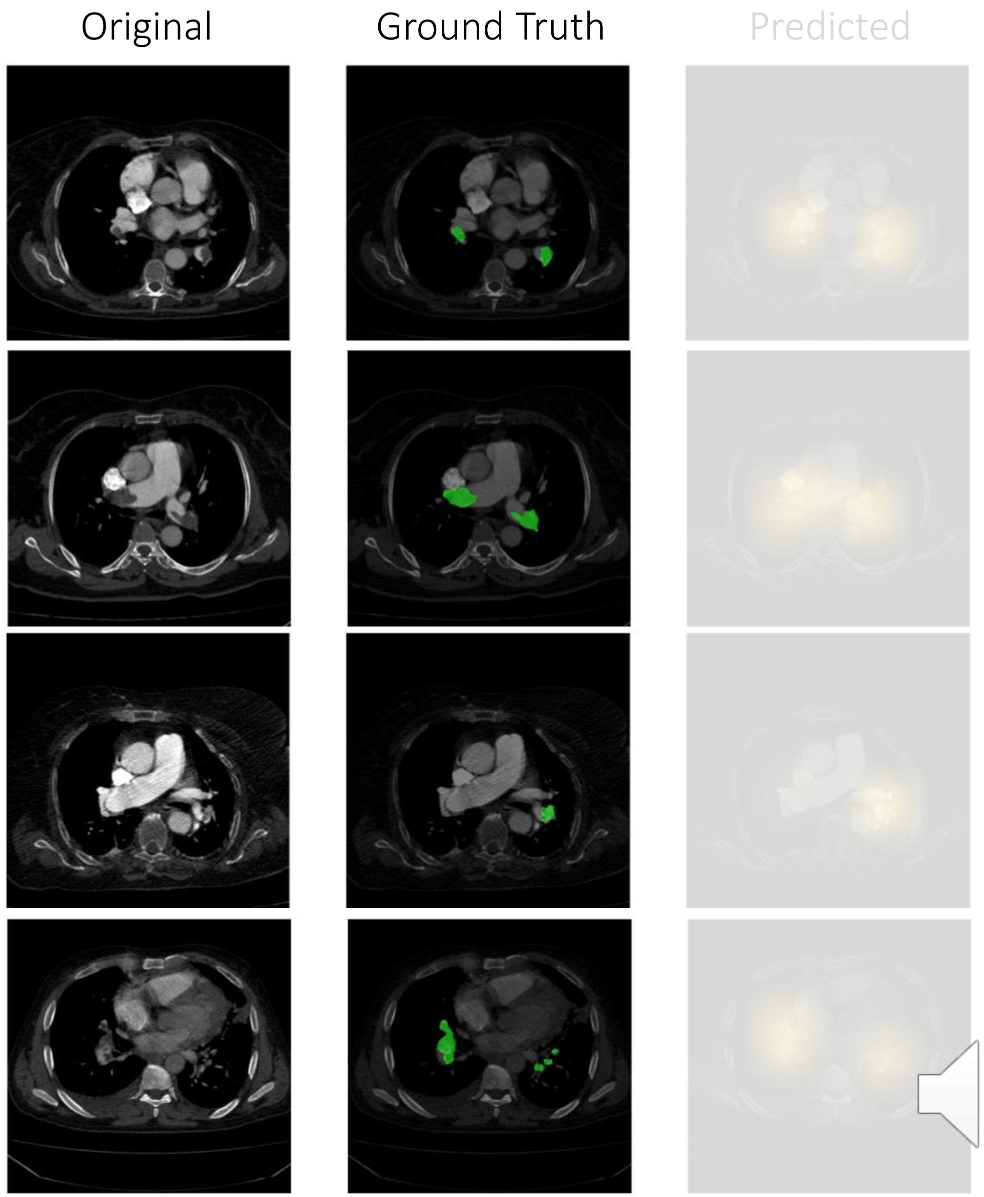
Original | Ground Truth | Predicted



CAD-PE Challenge Dataset

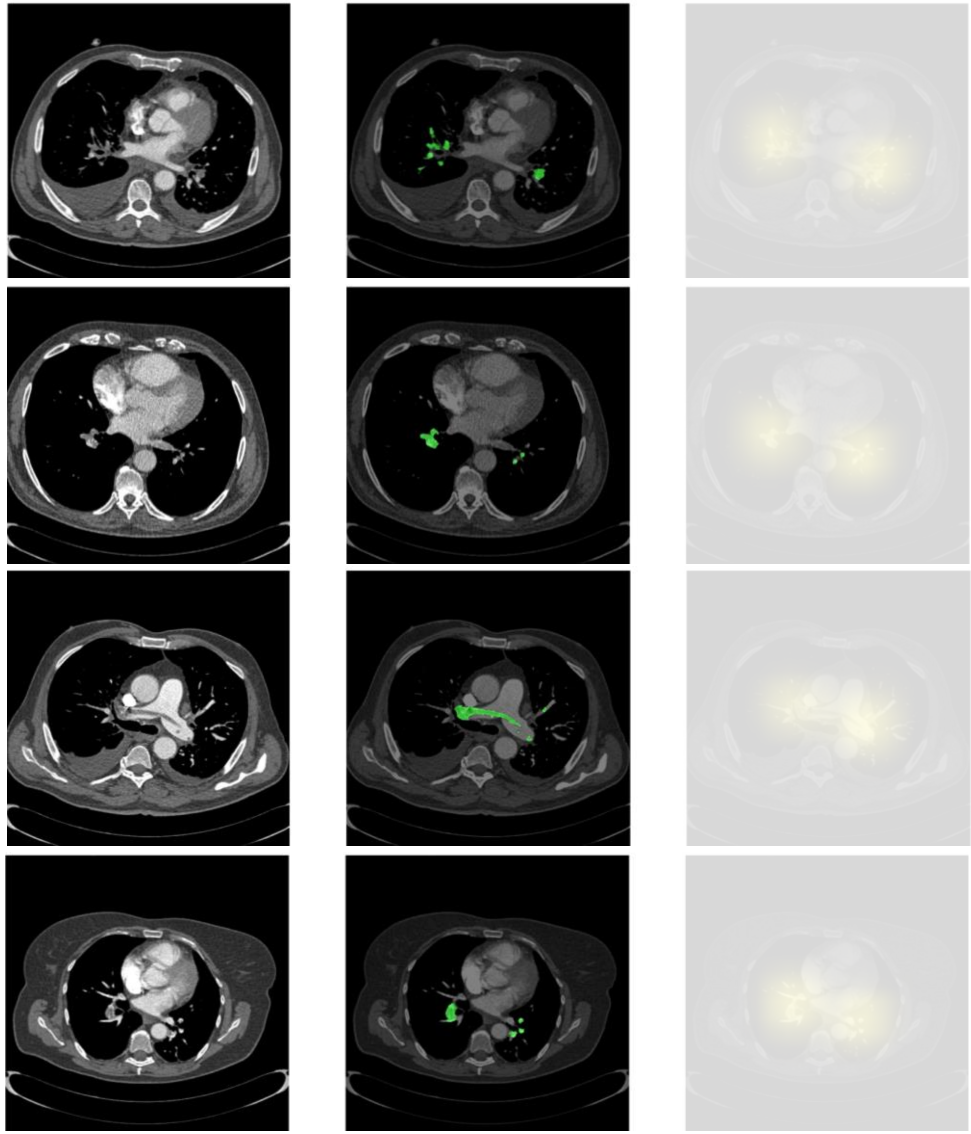
FUMPE Dataset





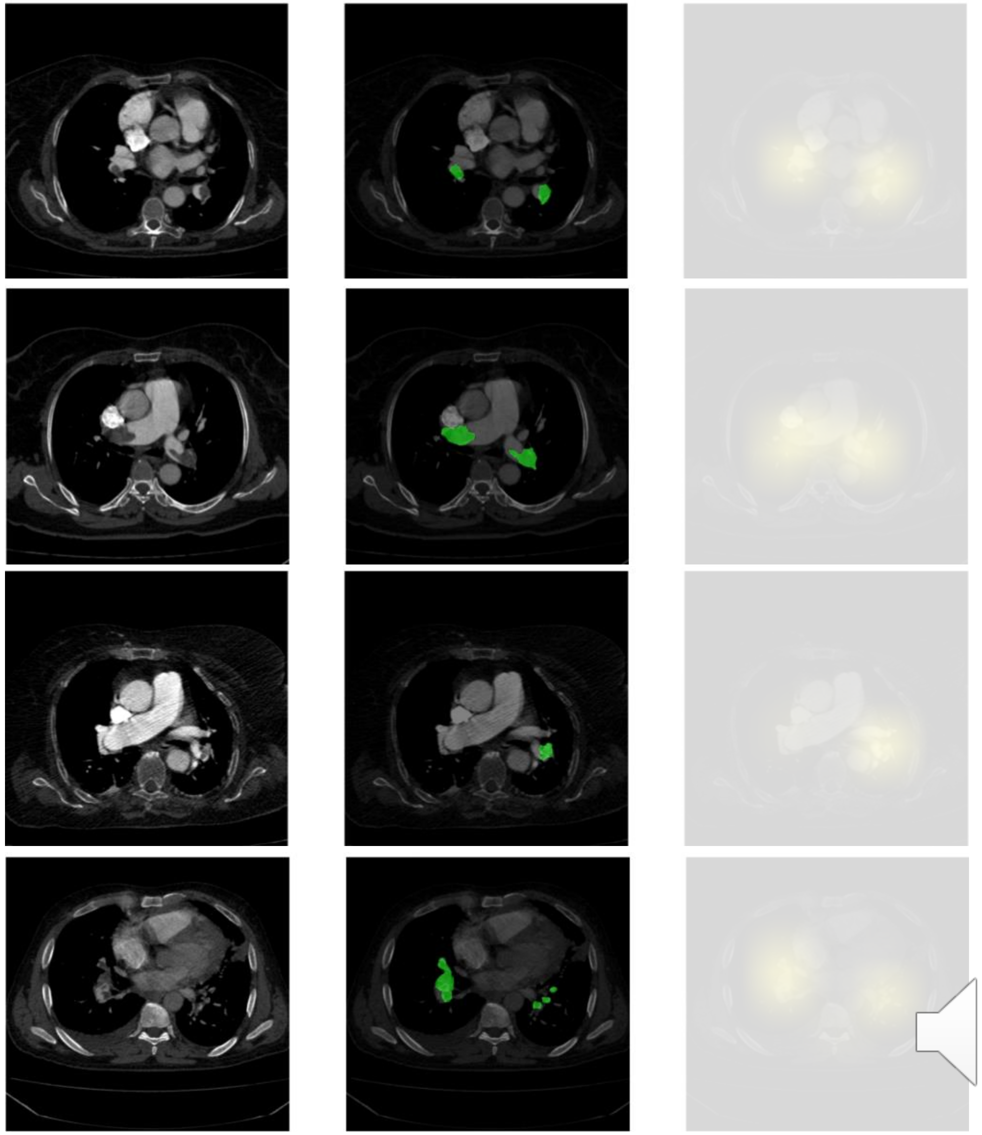
CAD-PE Challenge Dataset

Original Ground Truth Predicted



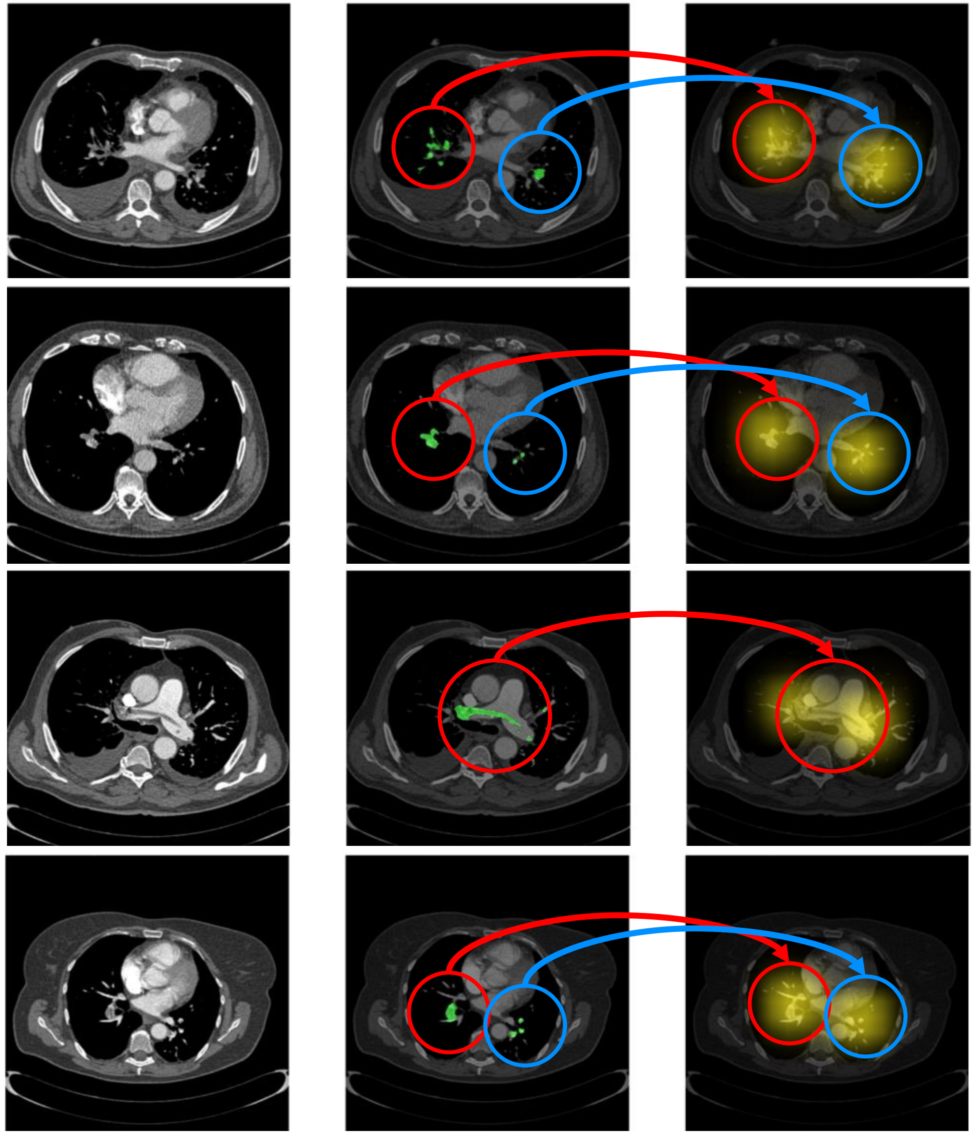
FUMPE Dataset

Original Ground Truth Predicted



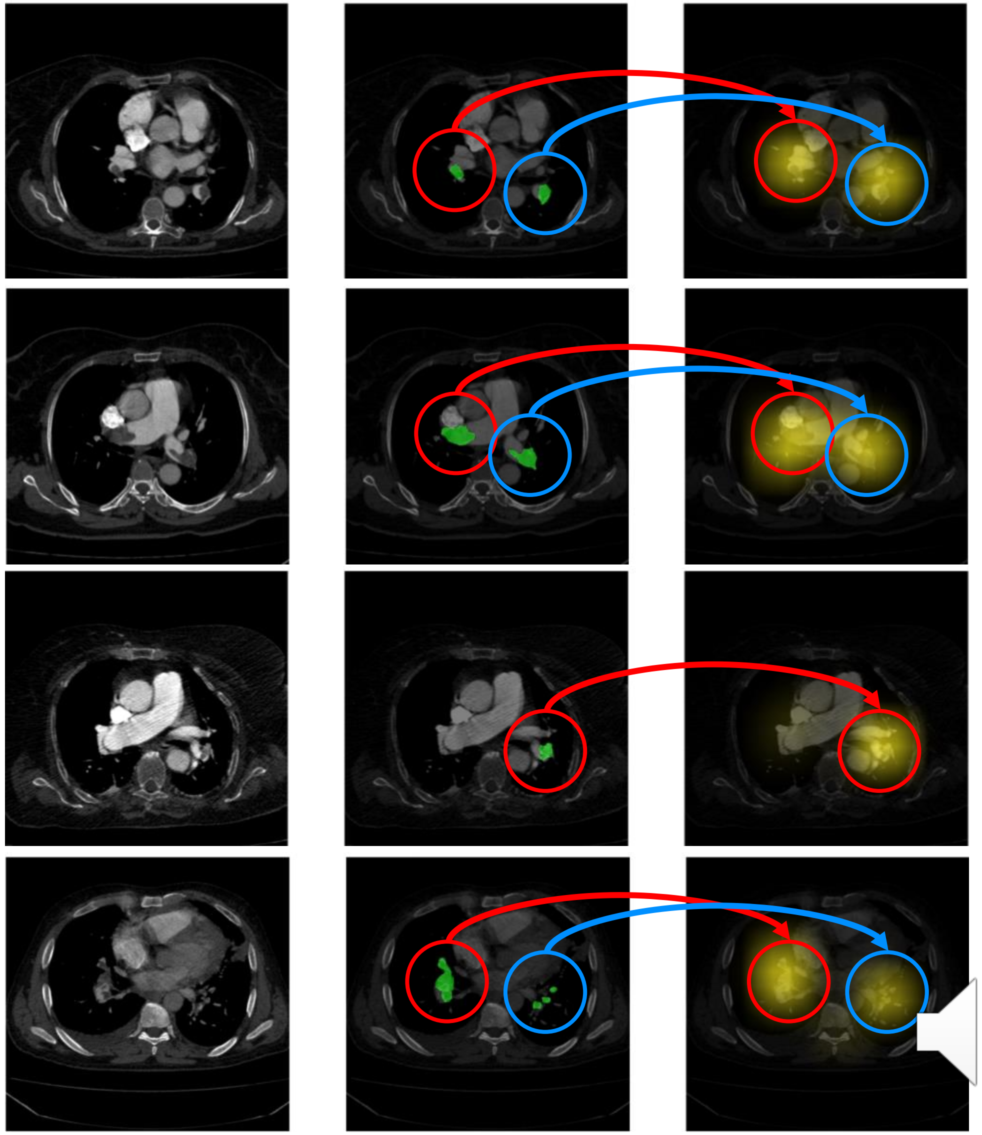
CAD-PE Challenge Dataset

Original | Ground Truth | Predicted

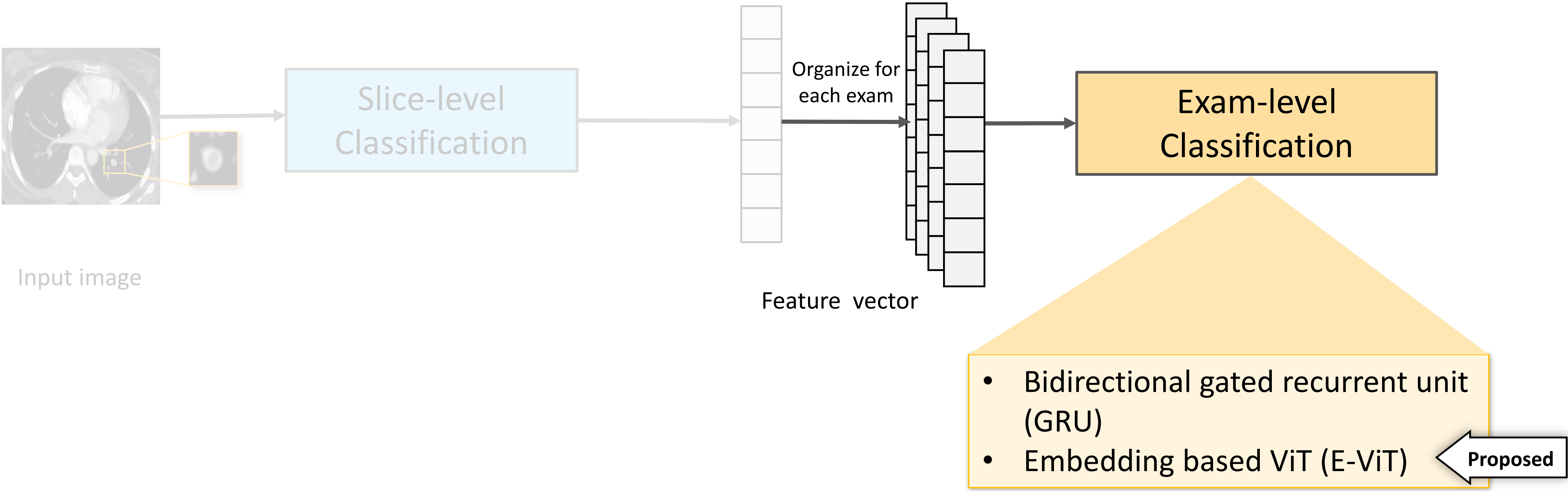


FUMPE Dataset

Original | Ground Truth | Predicted



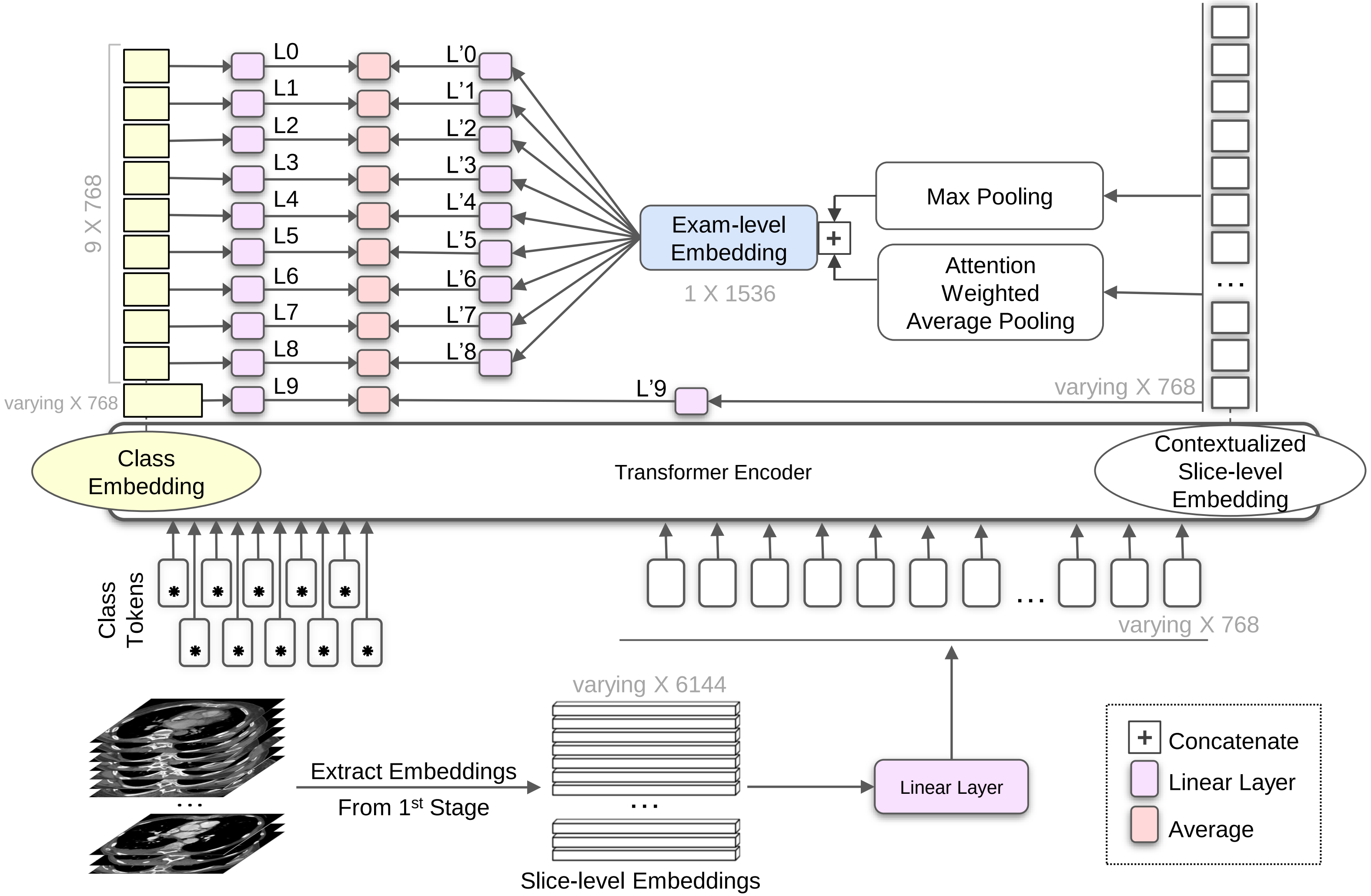
Second Stage



https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection/discussion/194145

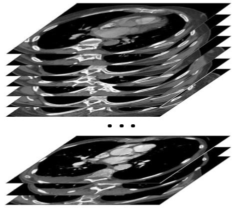
Second Stage

Proposed E-ViT



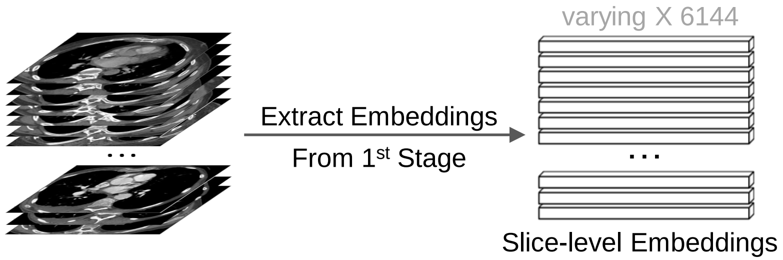
Second Stage

Proposed E-ViT



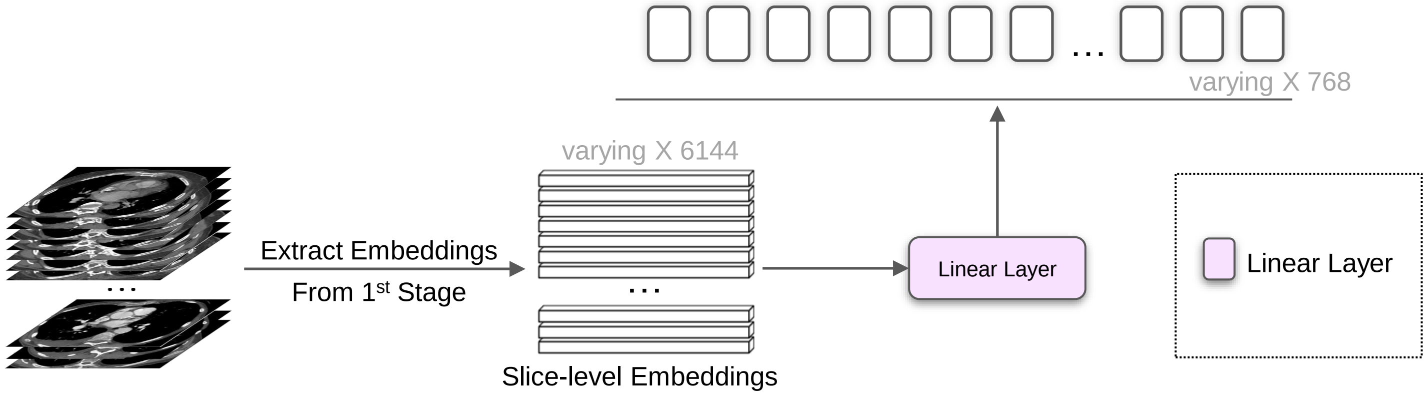
Second Stage

Proposed E-ViT



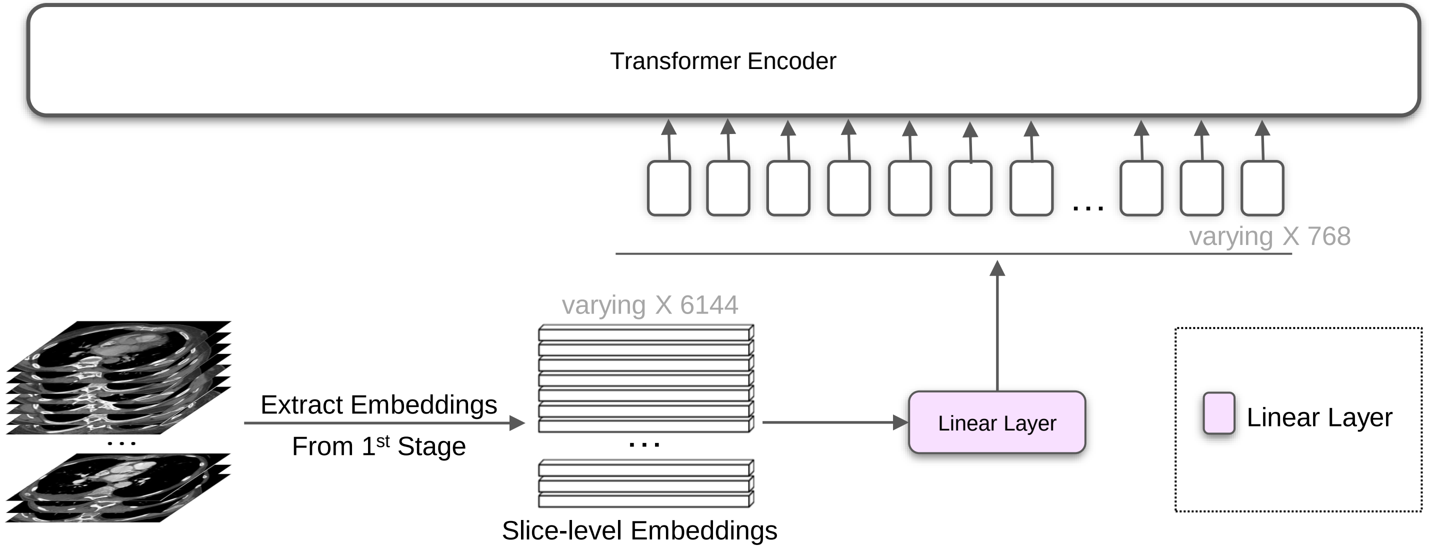
Second Stage

Proposed E-ViT



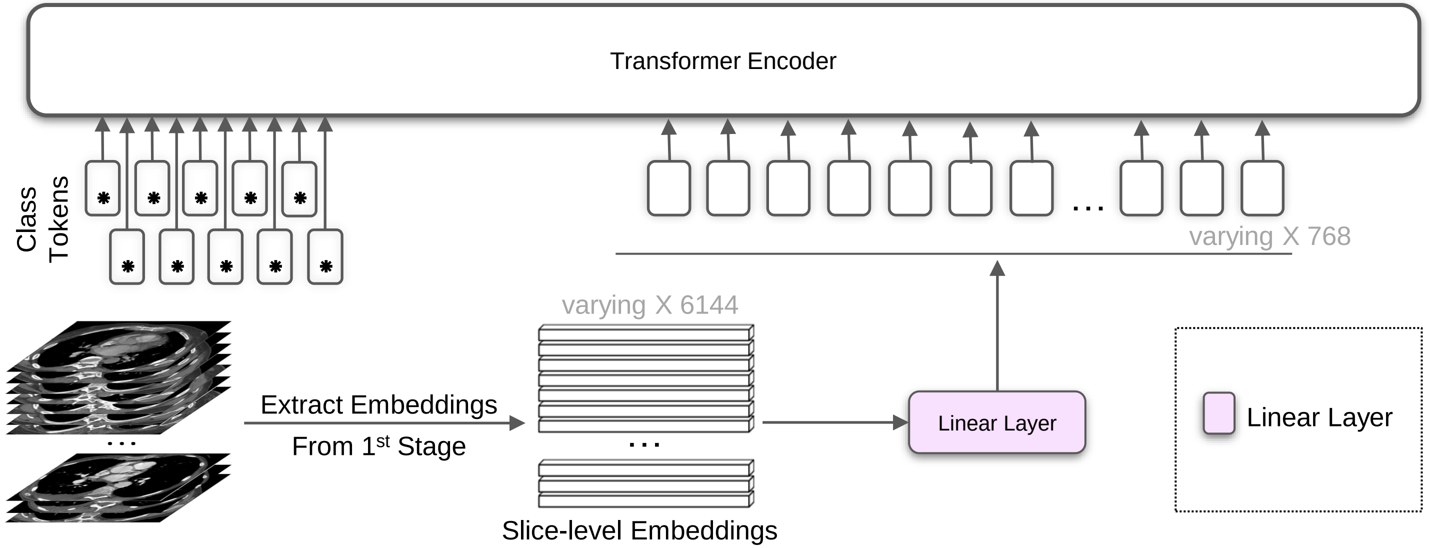
Second Stage

Proposed E-ViT



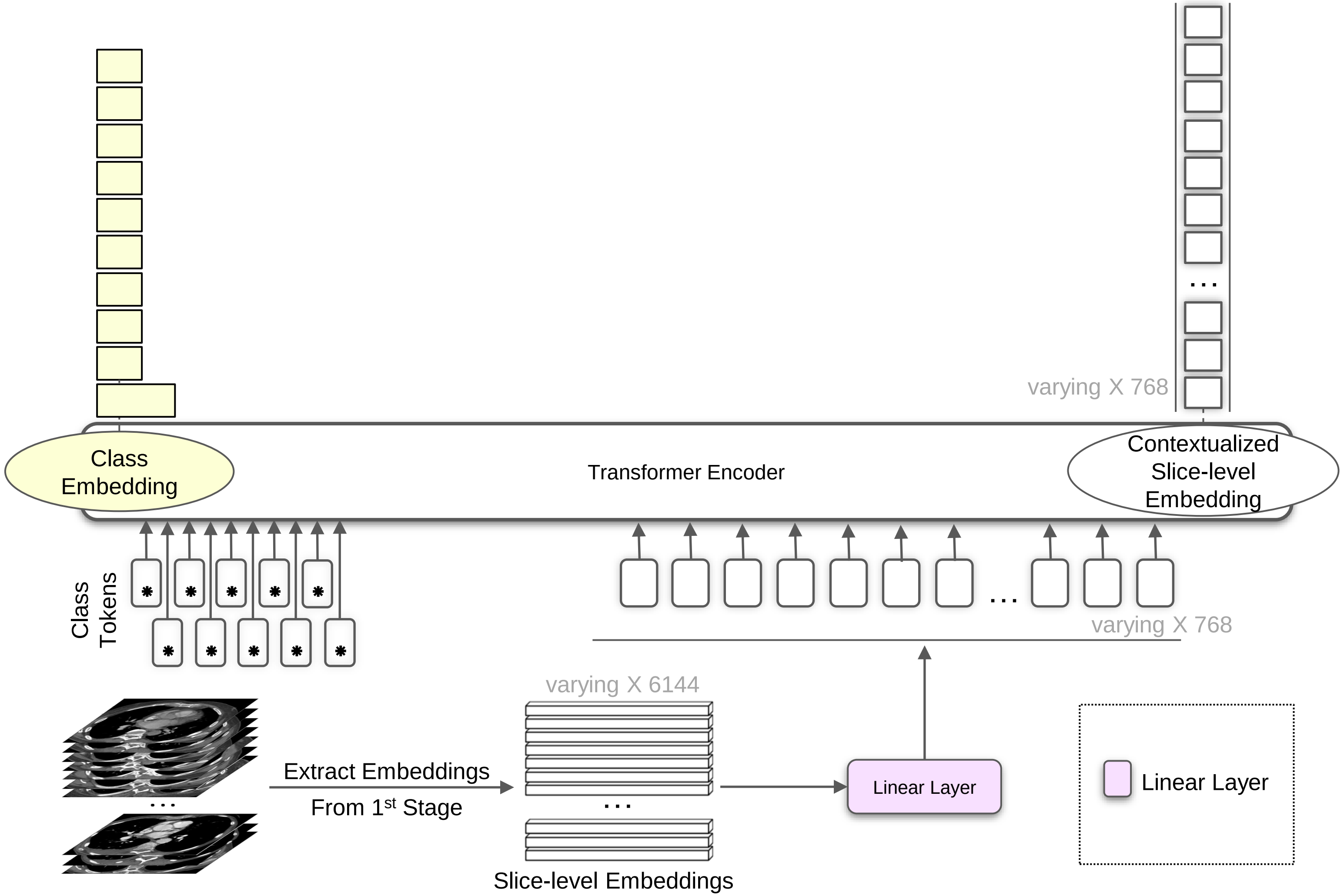
Second Stage

Proposed E-ViT



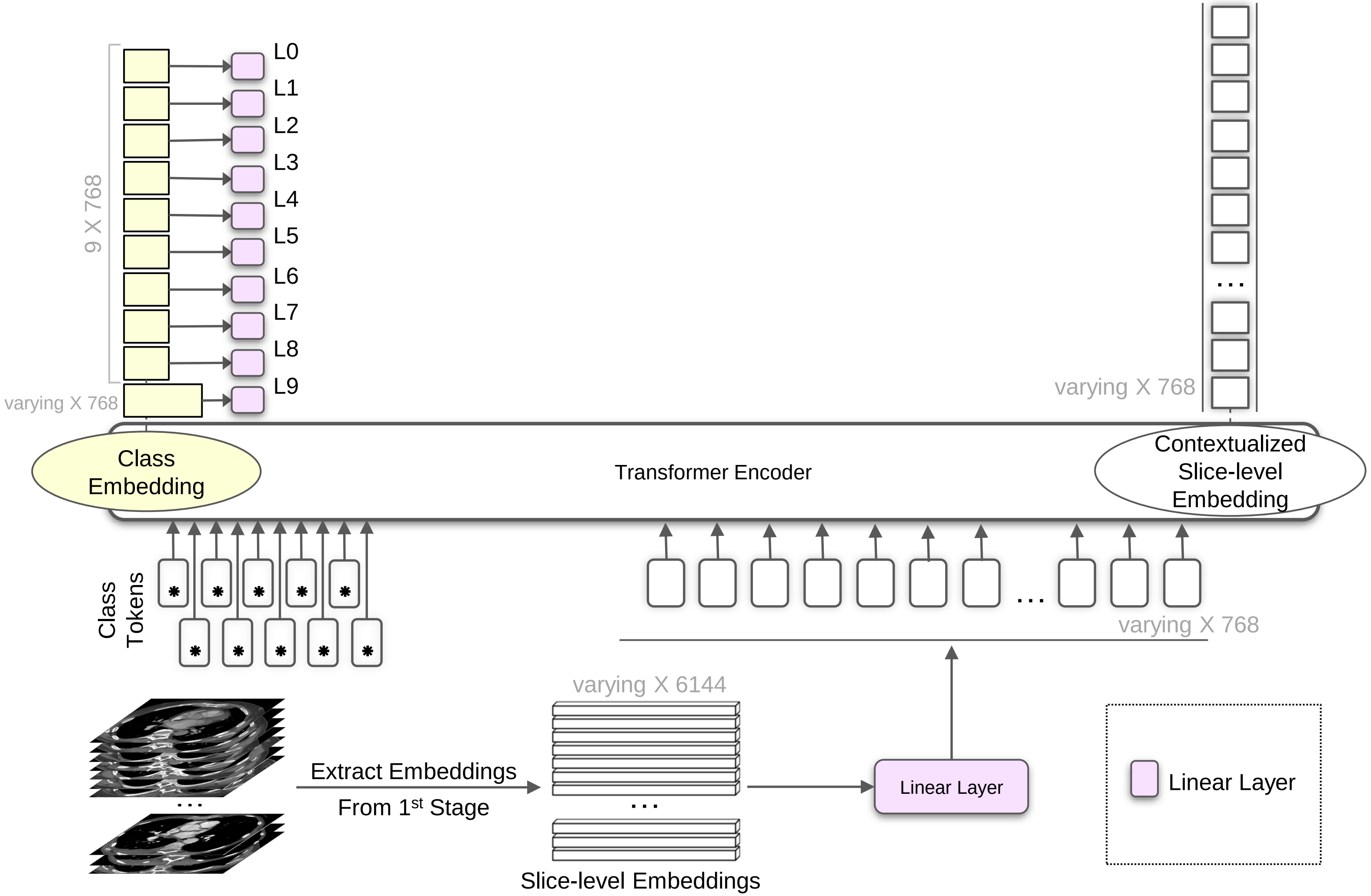
Second Stage

Proposed E-ViT



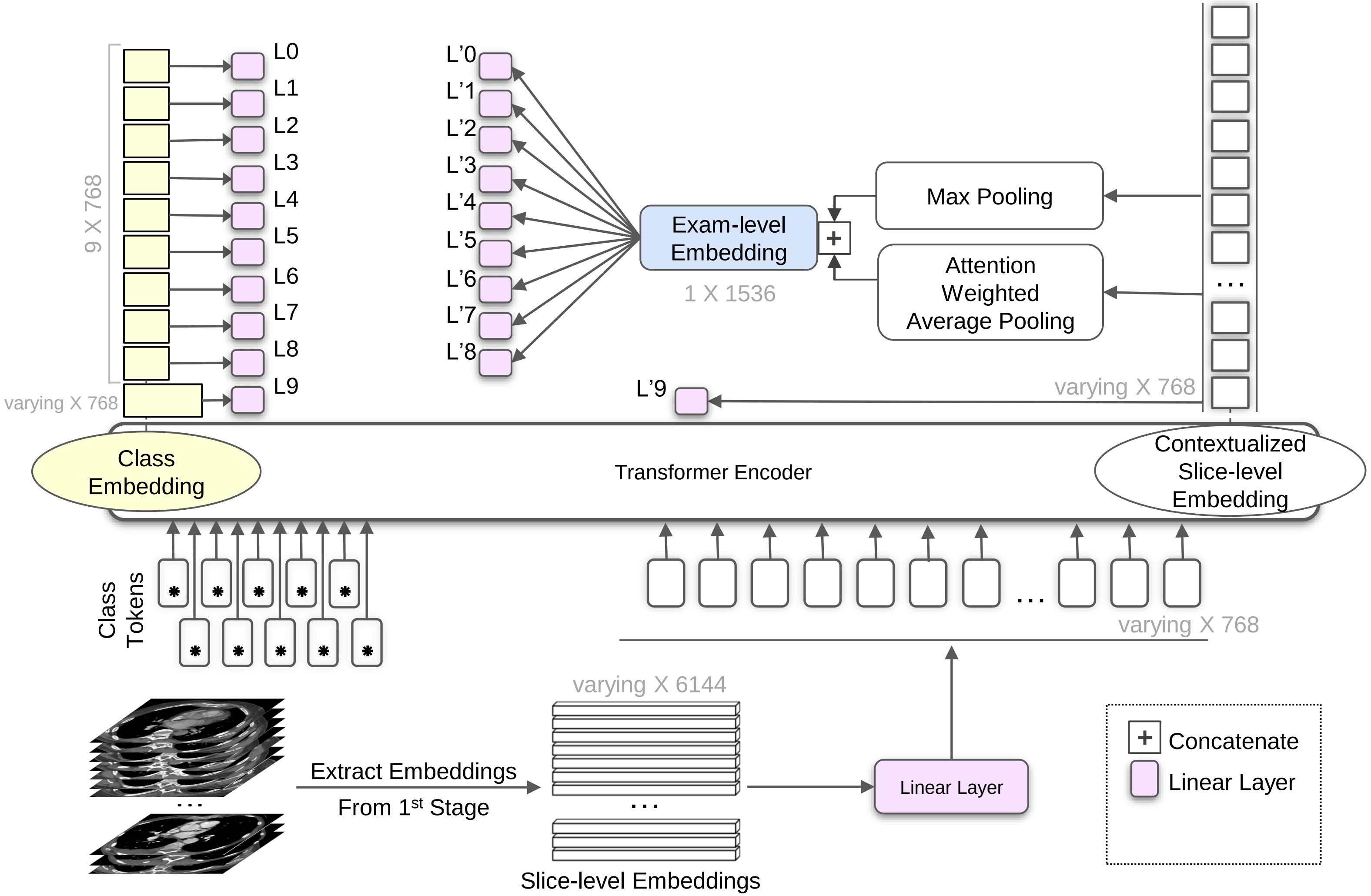
Second Stage

Proposed E-ViT



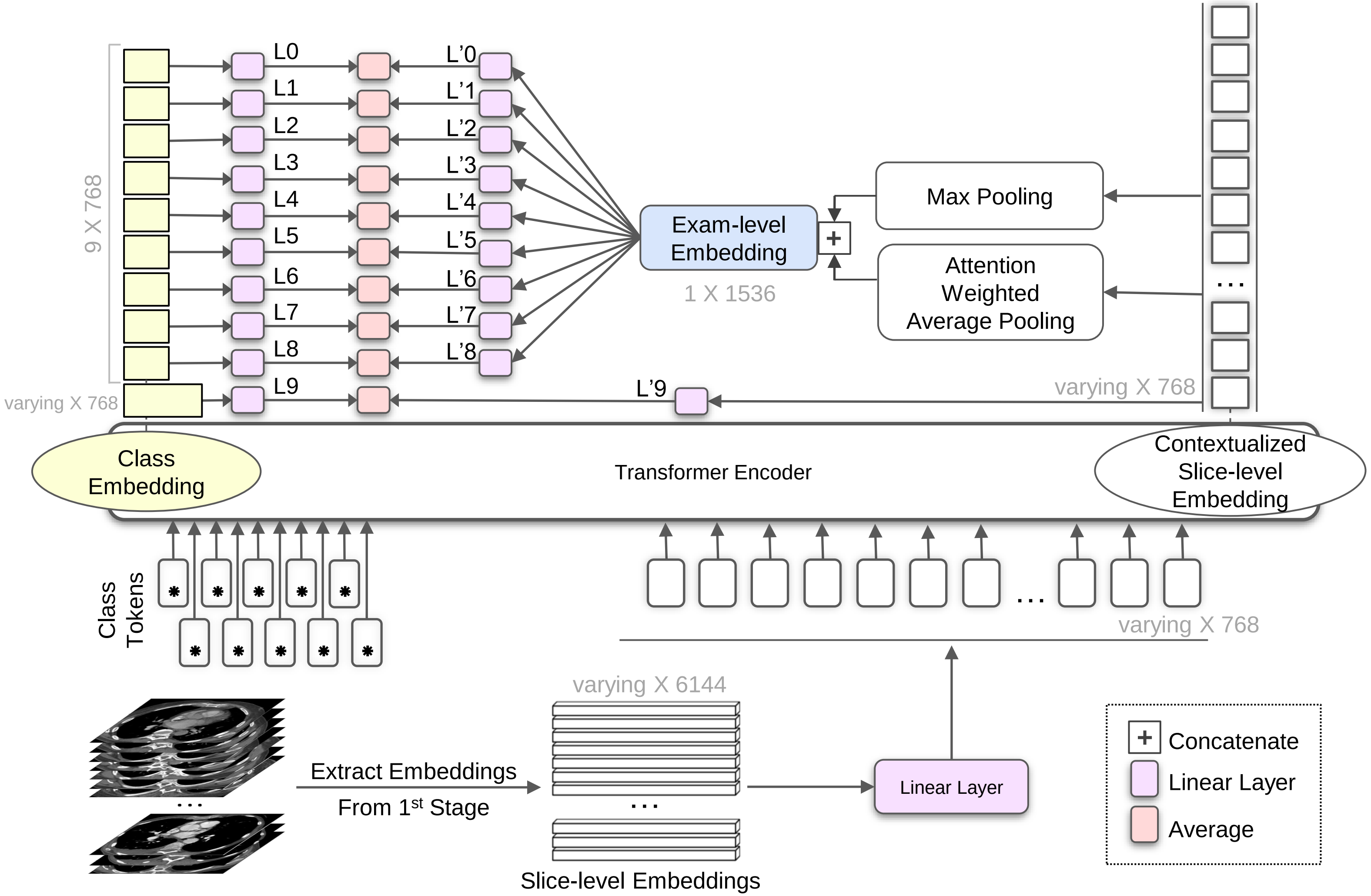
Second Stage

Proposed E-ViT

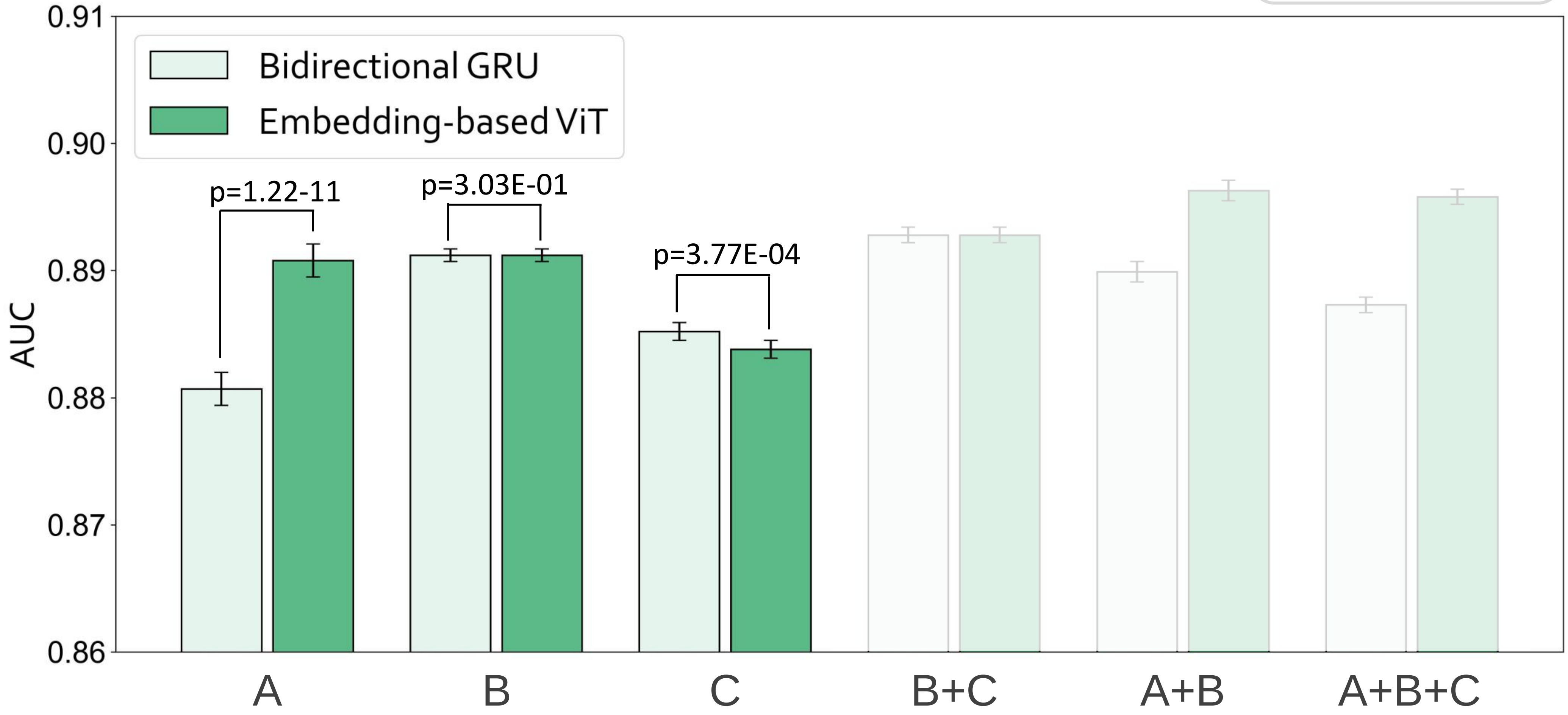


Second Stage

Proposed E-ViT

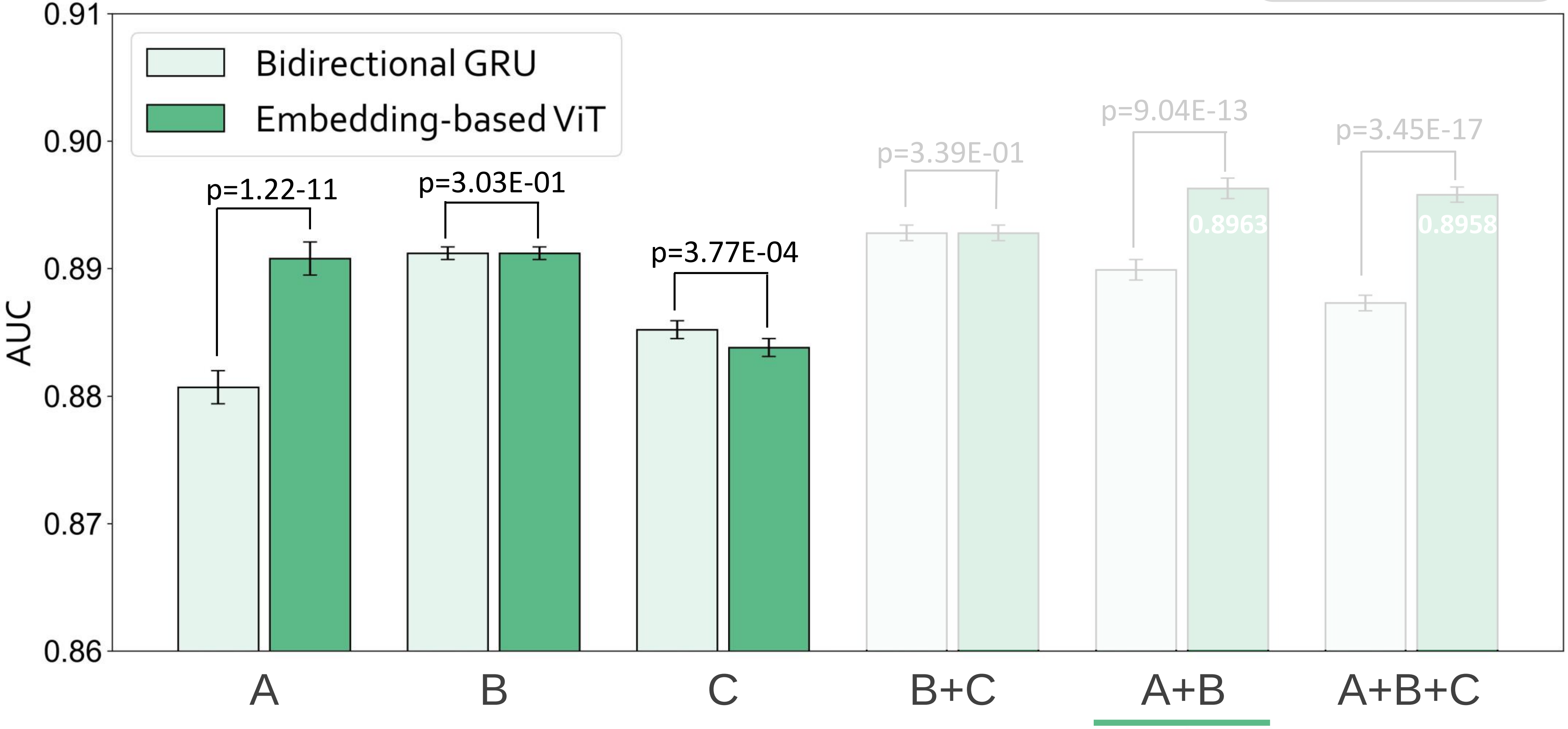


A = SeResNext50
B = Xception
C = SeXception

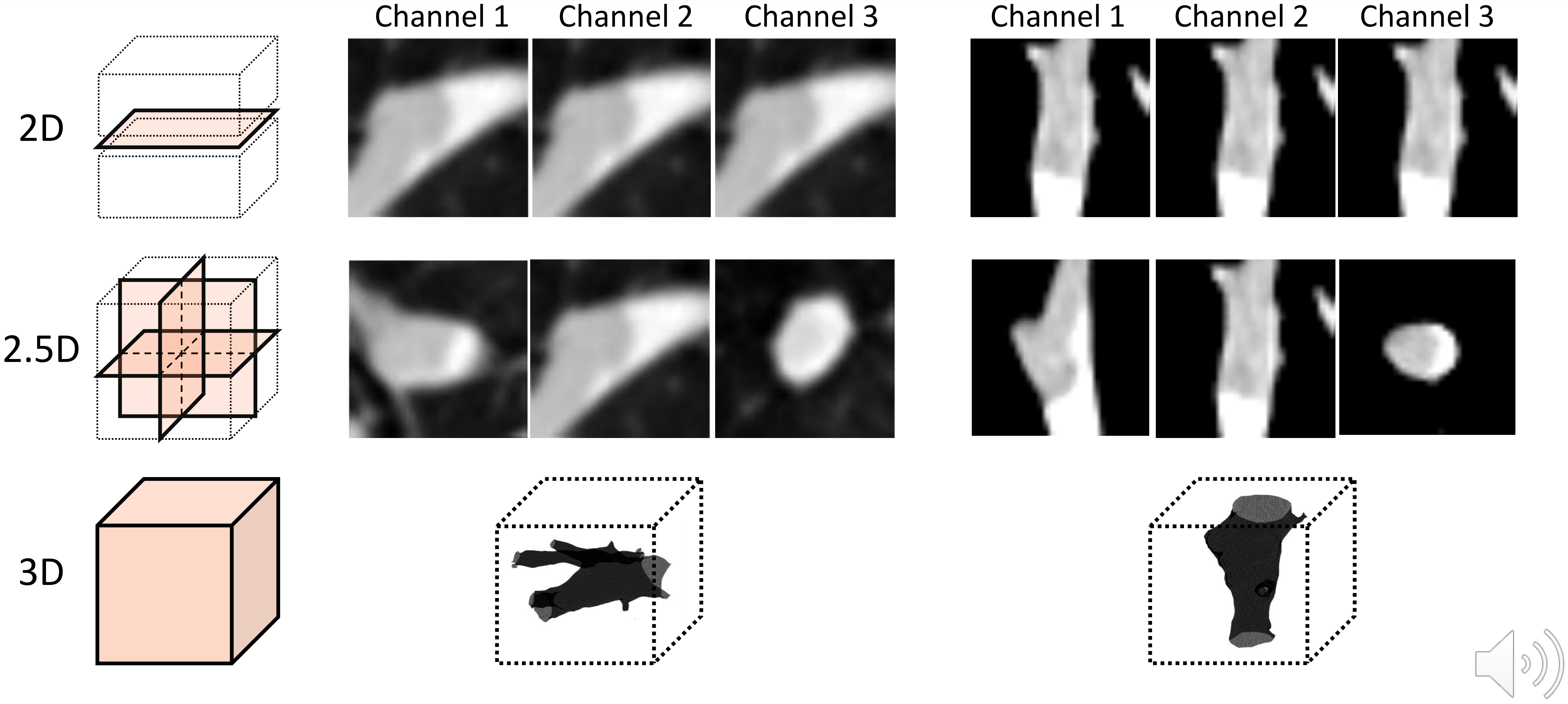


E-ViT outperforms BGRU at the exam level

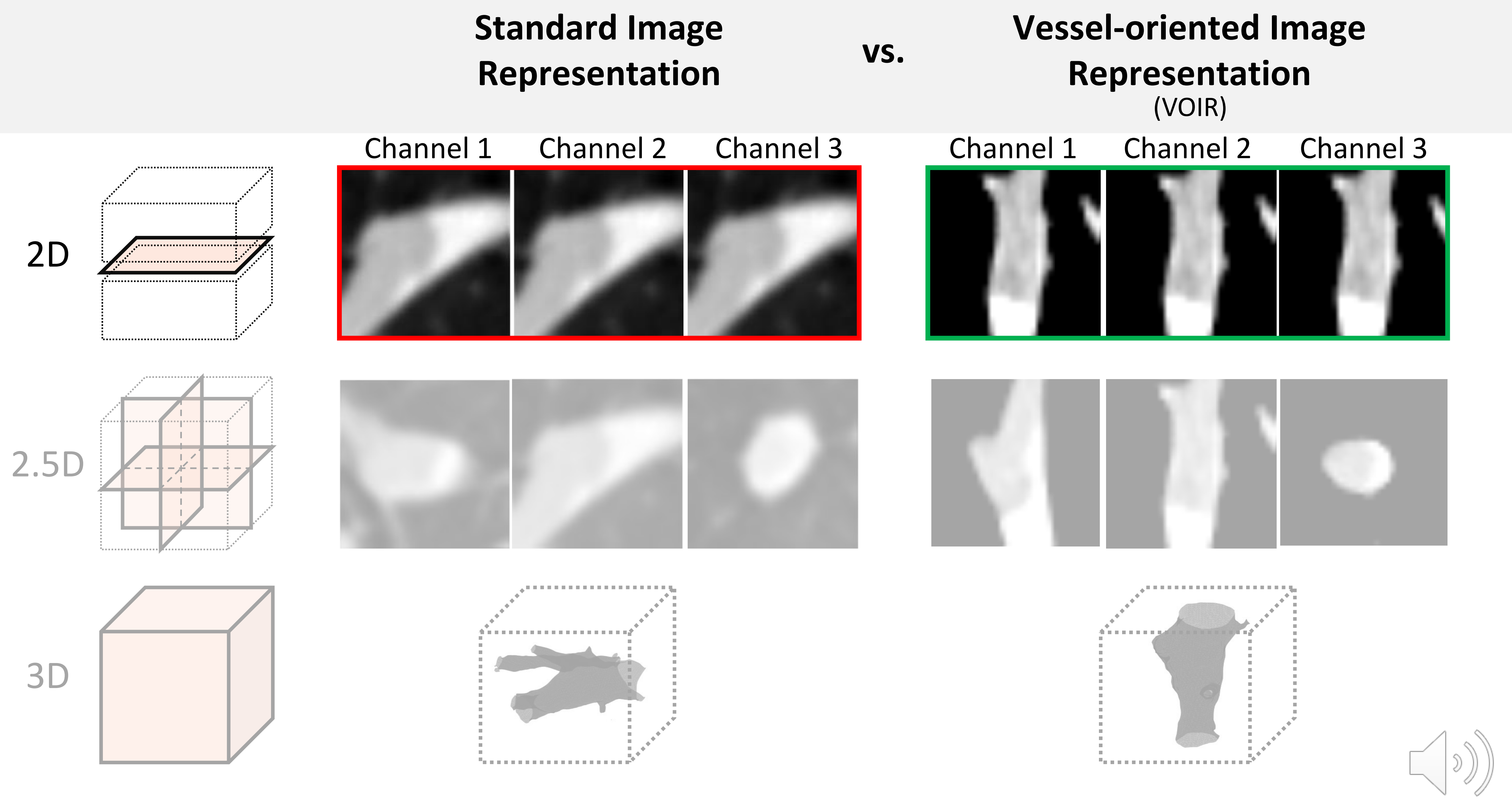
A = SeResNext50
B = Xception
C = SeXception



Standard Image Representation vs. Vessel-oriented Image Representation (VOIR)

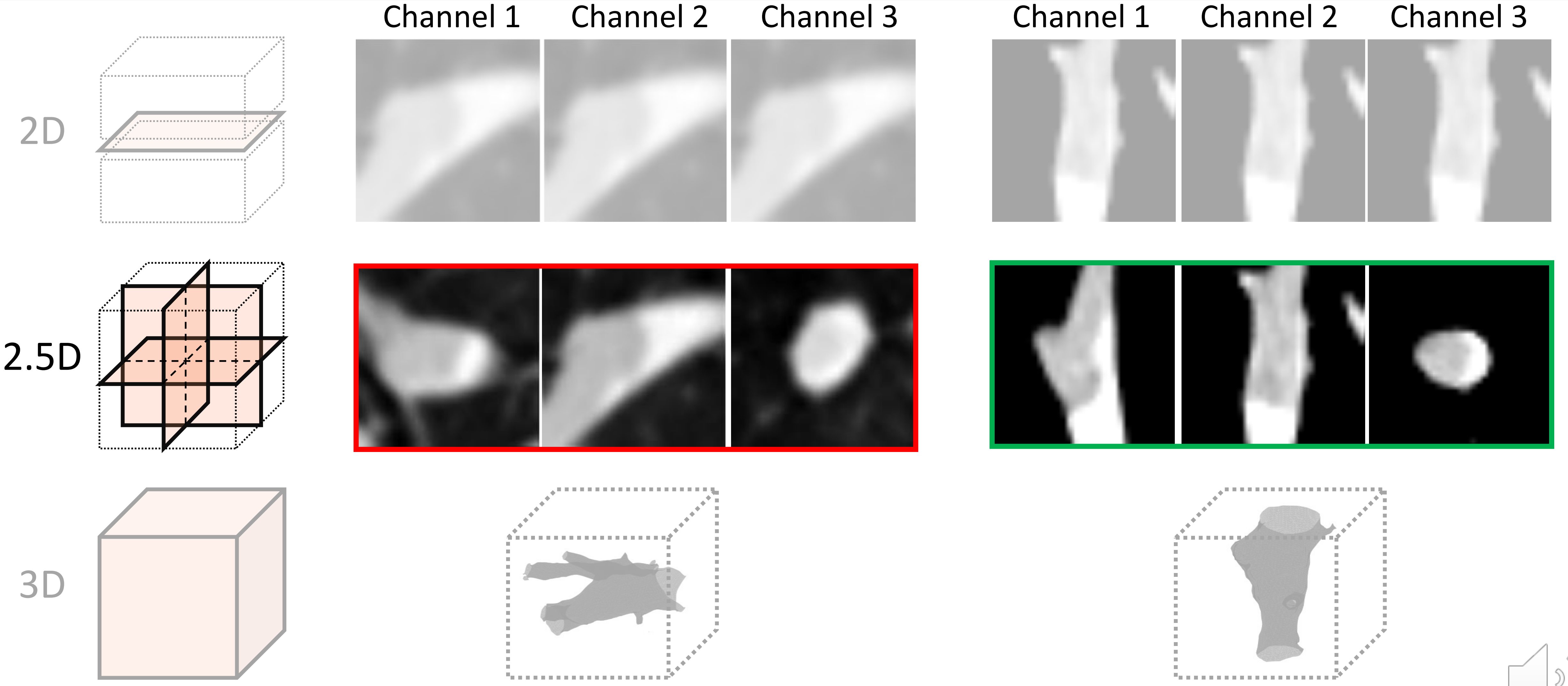


Zhou, Z. (2021). Towards annotation-efficient deep learning for computer-aided diagnosis (Doctoral dissertation, Arizona State University).



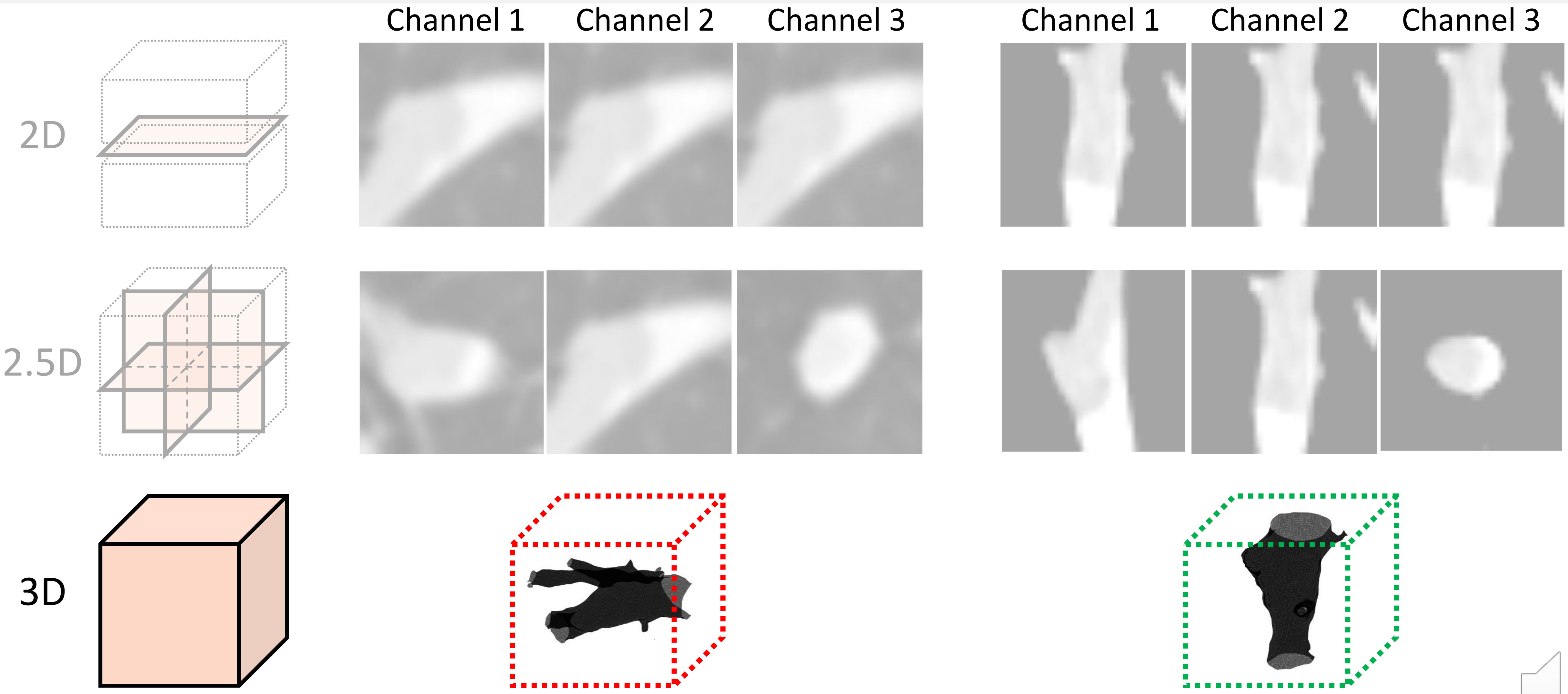
Zhou, Z. (2021). Towards annotation-efficient deep learning for computer-aided diagnosis (Doctoral dissertation, Arizona State University).

Standard Image Representation **vs.** **Vessel-oriented Image Representation** (VOIR)



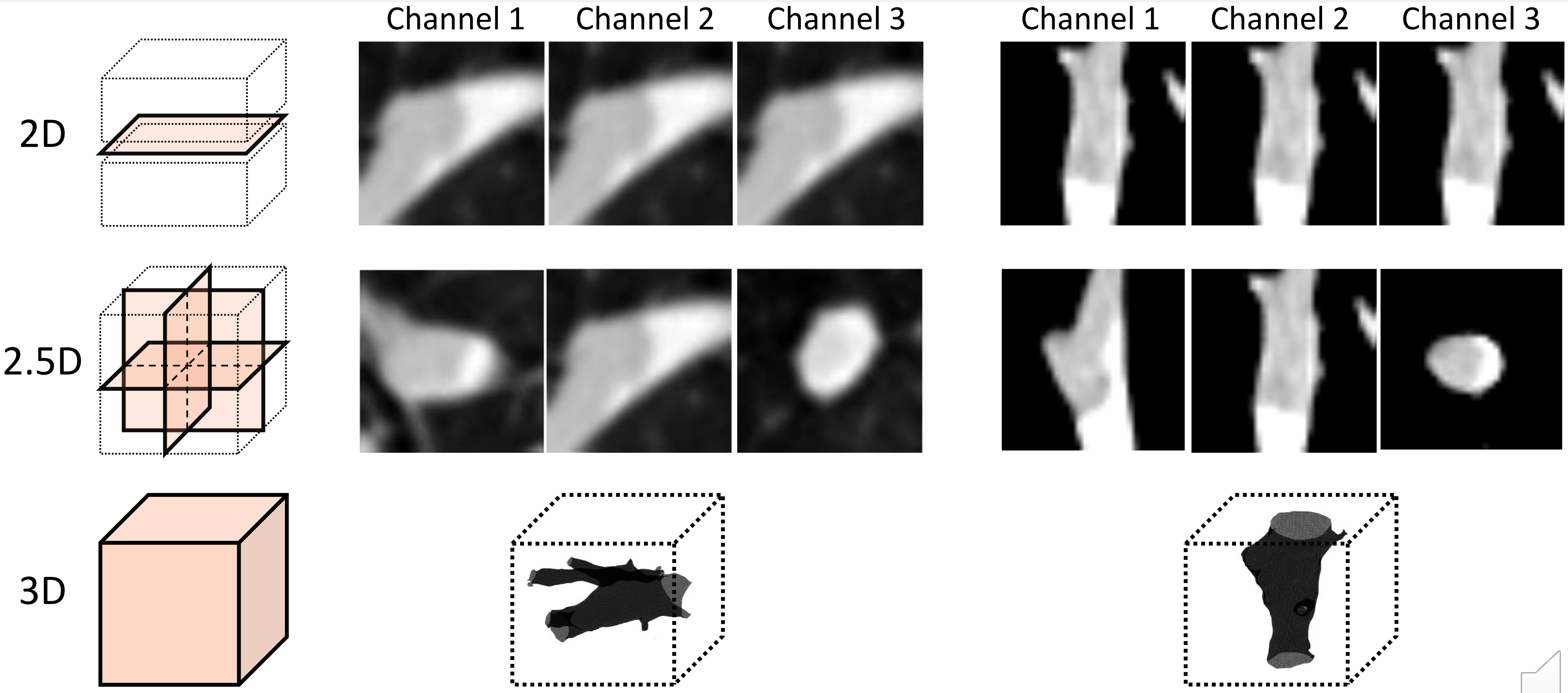
Zhou, Z. (2021). Towards annotation-efficient deep learning for computer-aided diagnosis (Doctoral dissertation, Arizona State University).

Standard Image Representation vs. Vessel-oriented Image Representation (VOIR)



Zhou, Z. (2021). Towards annotation-efficient deep learning for computer-aided diagnosis (Doctoral dissertation, Arizona State University).

Standard Image Representation vs. **Vessel-oriented Image Representation** (VOIR)



Zhou, Z. (2021). Towards annotation-efficient deep learning for computer-aided diagnosis (Doctoral dissertation, Arizona State University).

			Random	ImageNet	Models Genesis
Standard	2D	Slice-based Input	0.6303	0.6329	0.7211
	2.5D	Orthogonal Input	0.7881	0.8136	0.8534
	3D	Volume-based Input	0.8007	N/A	0.8766
VOIR	2D	Slice-based Input	0.8602	0.8581	0.8674
	2.5D	Orthogonal Input	0.8651	0.8729	0.8908
	3D	Volume-based Input	0.9135	N/A	0.9248

3D data offer higher performance than 2D and 2.5D data

			Random	ImageNet	Models Genesis
Standard	2D	Slice-based Input	0.6303	0.6329	0.7211
	2.5D	Orthogonal Input	0.7881	0.8136	0.8534
	3D	Volume-based Input	0.8007	N/A	0.8766
VOIR	2D	Slice-based Input	0.8602	0.8581	0.8674
	2.5D	Orthogonal Input	0.8651	0.8729	0.8908
	3D	Volume-based Input	0.9135	N/A	0.9248

VOIR is more informative than the standard image representation, bosting performance across image dimensions

			Random	ImageNet	Models Genesis
Standard	2D	Slice-based Input	0.6303	0.6329	0.7211
	2.5D	Orthogonal Input	0.7881	0.8136	0.8534
	3D	Volume-based Input	0.8007	N/A	0.8766
VOIR	2D	Slice-based Input	0.8602	0.8581	0.8674
	2.5D	Orthogonal Input	0.8651	0.8729	0.8908
	3D	Volume-based Input	0.9135	N/A	0.9248

			Random	ImageNet	Models Genesis
Standard	2D	Slice-based Input	0.6303	0.6329	0.7211
	2.5D	Orthogonal Input	0.7881	0.8136	0.8534
	3D	Volume-based Input	0.8007	N/A	0.8766
VOIR	2D	Slice-based Input	0.8602	0.8581	0.8674
	2.5D	Orthogonal Input	0.8651	0.8729	0.8908
	3D	Volume-based Input	0.9135	N/A	0.9248

Zhou, Z., Sodha, V., Pang, J., Gotway, M. B., & Liang, J. (2021). Models genesis. Medical image analysis, 67, 101840.

Hosseinzadeh Taher, M. R., Haghighi, F., Feng, R., Gotway, M. B., & Liang, J. (2021). A systematic benchmarking analysis of transfer learning for medical image analysis. In Domain Adaptation and Representation Transfer, and Affordable Healthcare and AI for Resource Diverse Global Health: Third MICCAI Workshop, DART 2021.

Same domain transfer learning with self-supervised pre-training enhances performance across image representations and dimensions

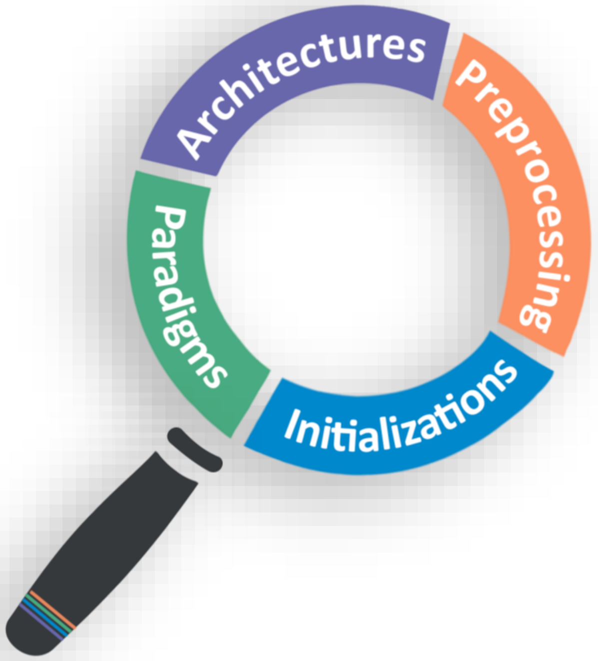
			Random	ImageNet	Models Genesis
Standard	2D	Slice-based Input	0.6303	0.6329	0.7211
	2.5D	Orthogonal Input	0.7881	0.8136	0.8534
	3D	Volume-based Input	0.8007	N/A	0.8766
VOIR	2D	Slice-based Input	0.8602	0.8581	0.8674
	2.5D	Orthogonal Input	0.8651	0.8729	0.8908
	3D	Volume-based Input	0.9135	N/A	0.9248

Zhou, Z., Sodha, V., Pang, J., Gotway, M. B., & Liang, J. (2021). Models genesis. Medical image analysis, 67, 101840.

Hosseinzadeh Taher, M. R., Haghighi, F., Feng, R., Gotway, M. B., & Liang, J. (2021). A systematic benchmarking analysis of transfer learning for medical image analysis. In Domain Adaptation and Representation Transfer, and Affordable Healthcare and AI for Resource Diverse Global Health: Third MICCAI Workshop, DART 2021.

Contributions

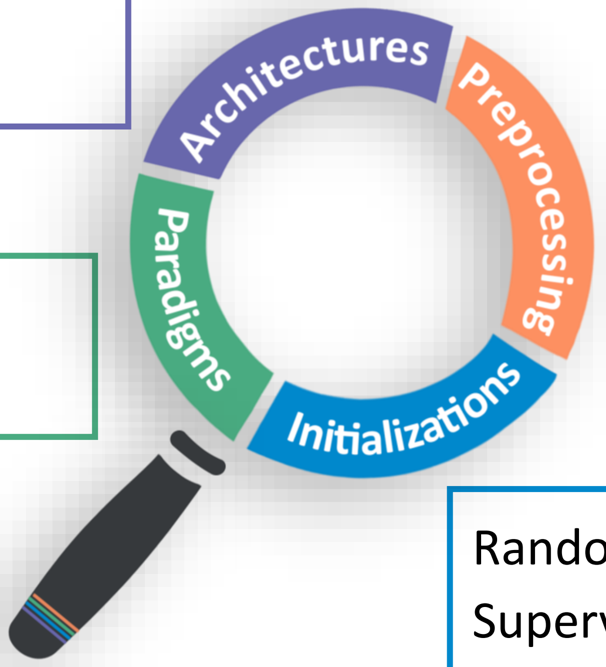
Contributions



Contributions

Convolutional neural networks
Vision transformer
Swin transformer

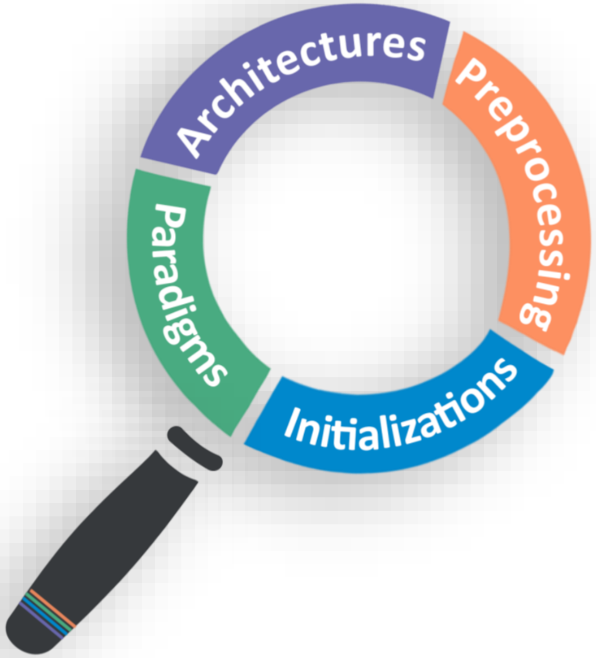
Standard image representation
Vessel oriented image representation



Sequential learning model
Proposed E-ViT

Random initialization
Supervised pre-training
Self-supervised pre-training

Conclusions



SeXception performs optimally for slice-level classification task

96.14% ➡ **96.34%**
0.20%

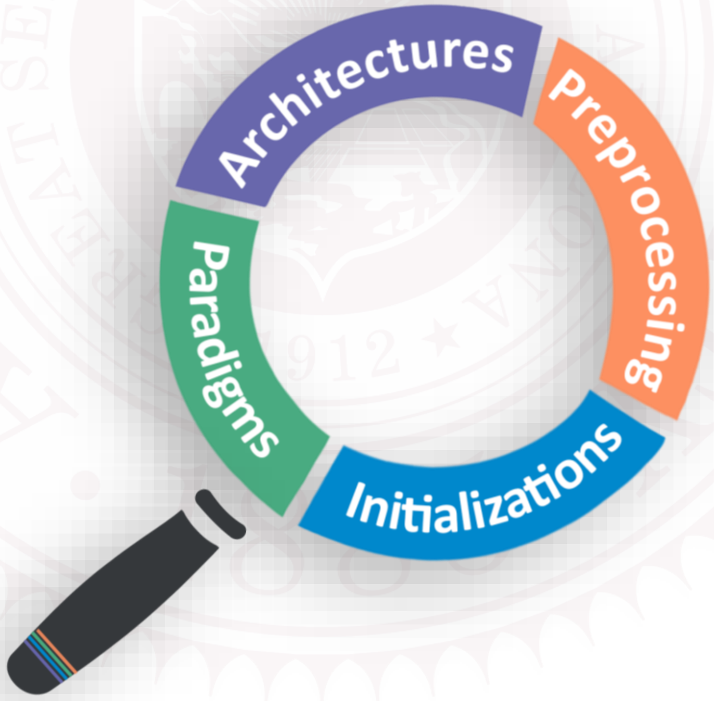
Our proposed E-ViT significantly improves performance for exam-level diagnosis

88.07% ➡ **89.63%**
1.53%

VOIR is more informative than the standard image representation, boosting performance across image dimensions

2D	2.5D	3D
63.03% ➡ **86.02%**	78.81% ➡ **86.51%**	80.07% ➡ **91.35 %**
22.99%	7.70%	11.28%

Seeking an Optimal Approach for Computer-aided Diagnosis of Pulmonary Embolism



View PDF

https://github.com/JLiangLab/CAD_PE