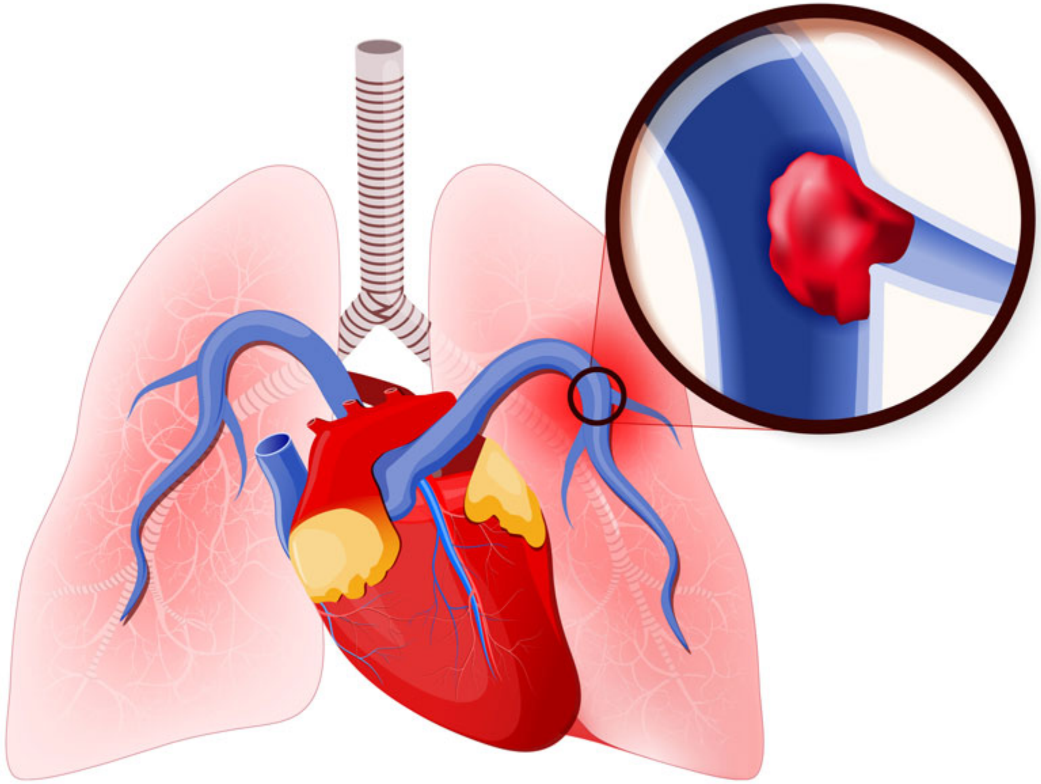
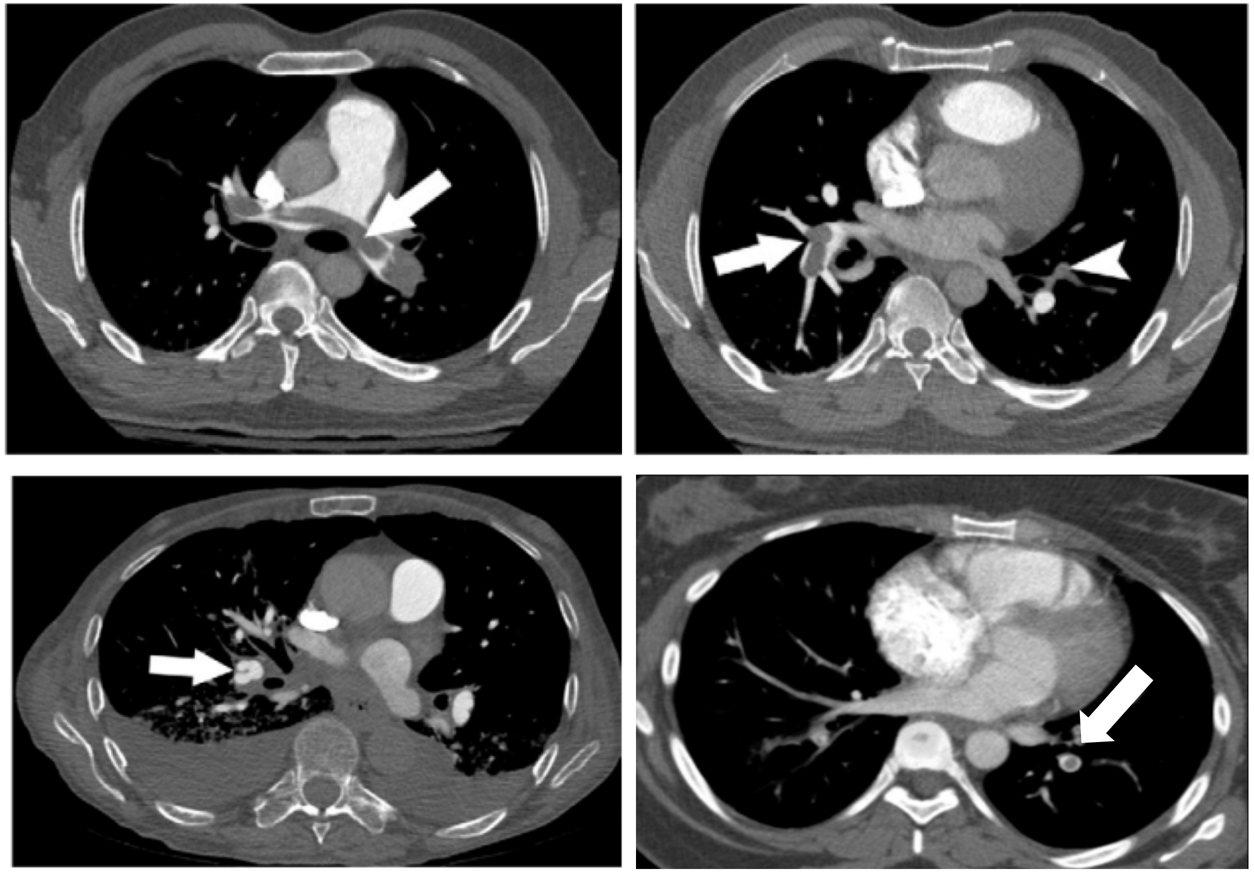
# Seeking an Optimal Approach for Computer-aided Pulmonary Embolism Detection


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

**[1]Arizona State University** **[2]Mayo Clinic**

# Pulmonary Embolism





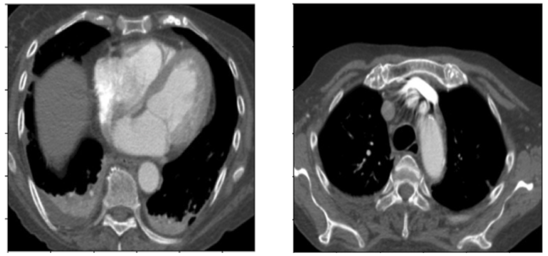
# RSNA Pulmonary Embolism Dataset

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

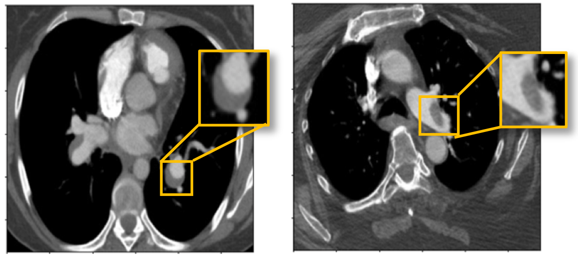
## Slice-level Classification

Ground Truth

1. PE present or not



PE absent



PE present

## Exam-level Classification

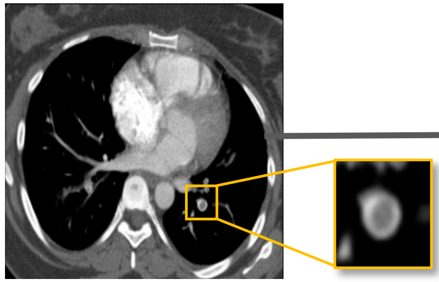
Ground Truth

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

# RSNA Pulmonary Embolism Dataset

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



Input image

**1**

Slice-level Classification

Ground Truth

1. PE present or not

Organize for each exam

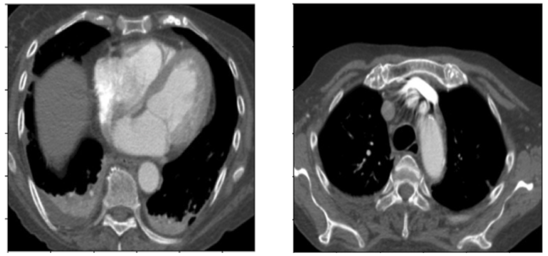
Feature vector

**2**

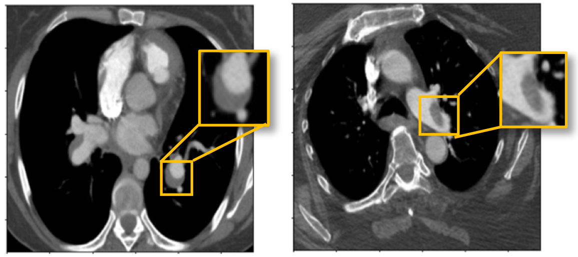
Exam-level Classification

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

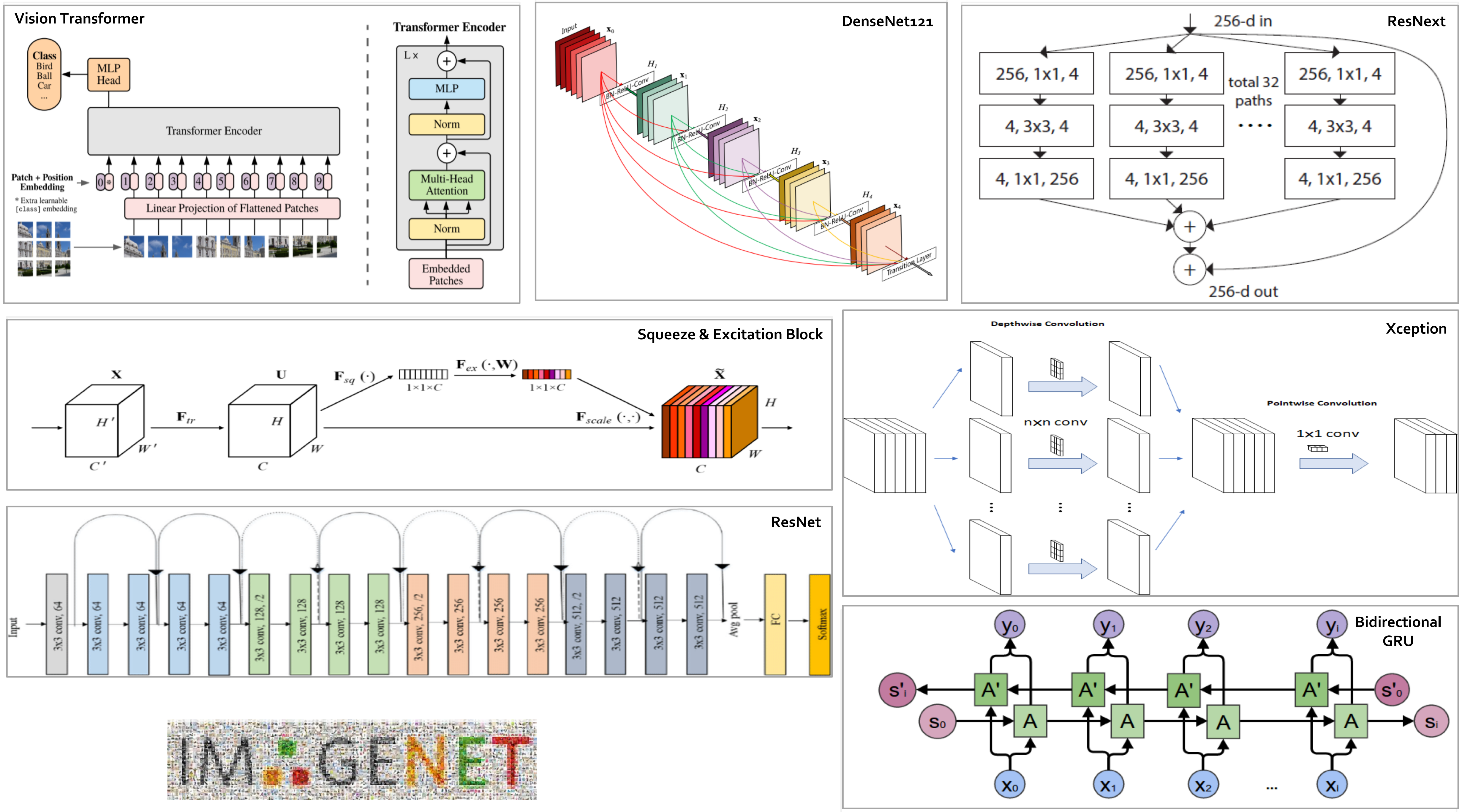


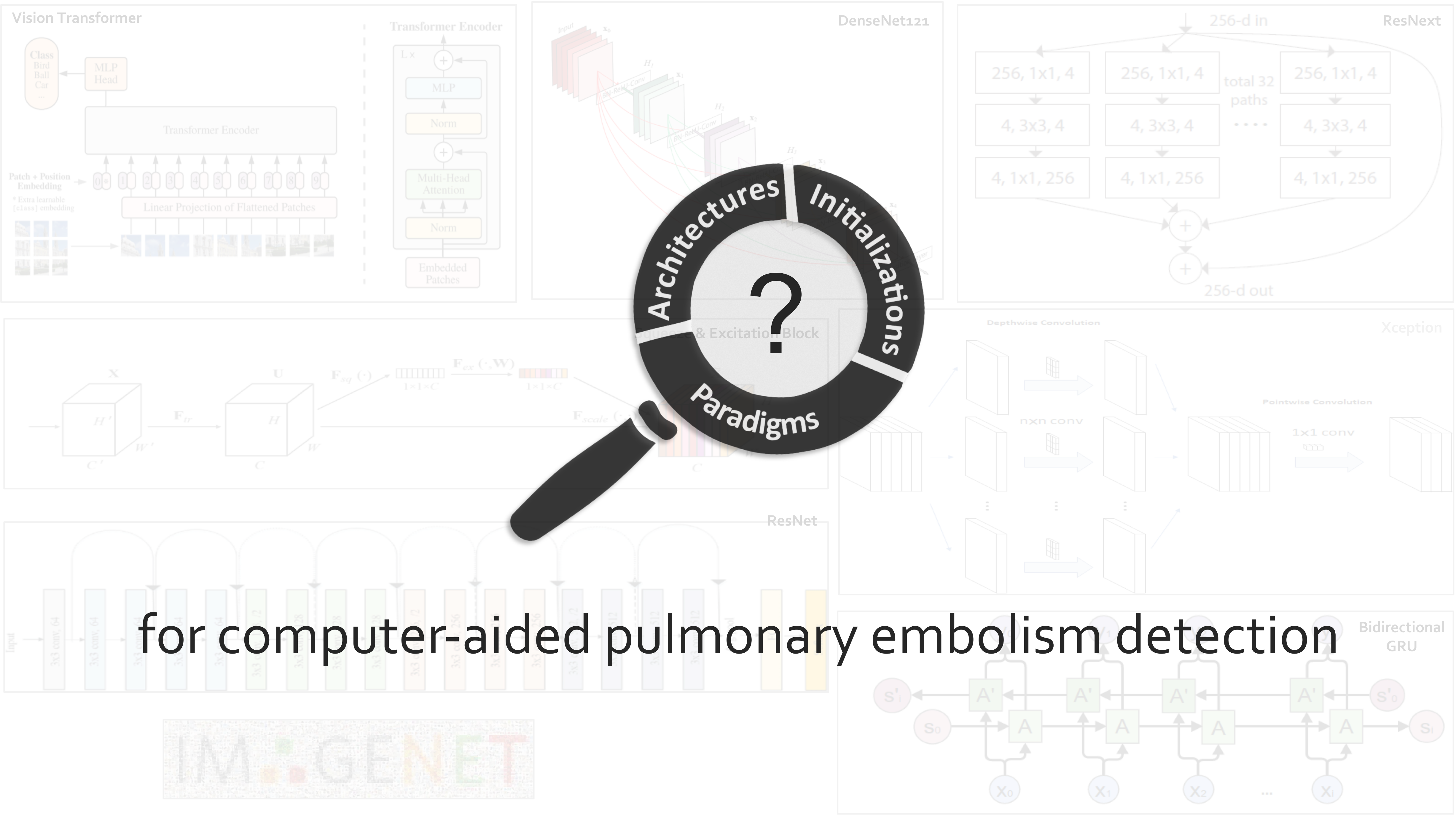
PE absent



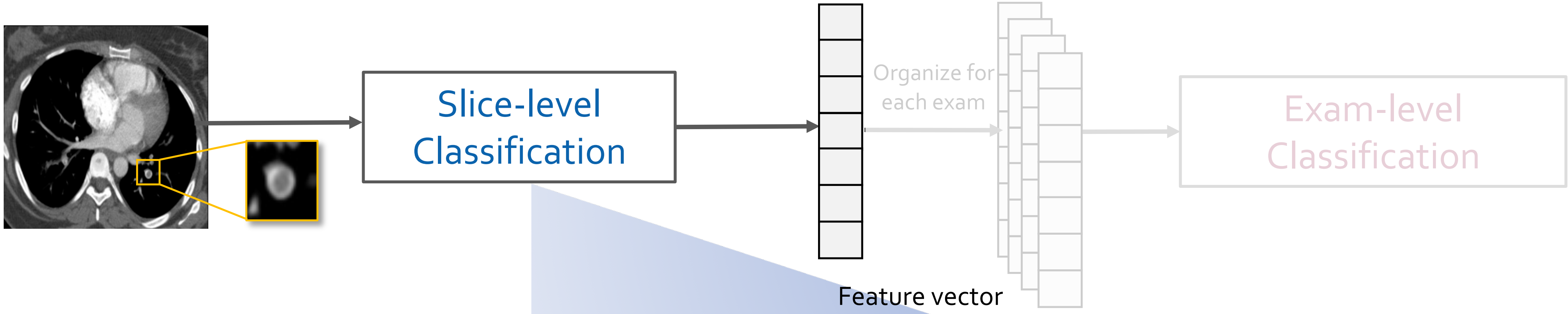
PE present

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


Vision Transformer
Class
Bird
Ball
Car
...
MLP Head
Transformer Encoder
Patch + Position Embedding
* Extra learnable [class] embedding
Linear Projection of Flattened Patches
Transformer Encoder
L x
MLP
Norm
Multi-Head Attention
Norm
Embedded Patches
DenseNet121
Input
BN-ReLU-Conv
Transition Layer
ResNext
256-d in
256, 1x1, 4
4, 3x3, 4
4, 1x1, 256
total 32 paths
256-d out
Squeeze & Excitation Block
X
U
H'
W'
C'
H
W
C
1×1×C
1×1×C
Xception
Depthwise Convolution
nxn conv
Pointwise Convolution
1x1 conv
ResNet
Input
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
Avg pool
FC
Softmax
Bidirectional GRU
IMAGENET


Vision Transformer
Transformer Encoder
DenseNet121
ResNext
256-d in
256, 1x1, 4
4, 3x3, 4
4, 1x1, 256
total 32 paths
256-d out
Xception
ResNet
Bidirectional GRU
Architectures
Initializations
Paradigms
?
for computer-aided pulmonary embolism detection

# First Stage




**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

**Vision Transformer**

1. ViT-B_16
2. ViT-B_32

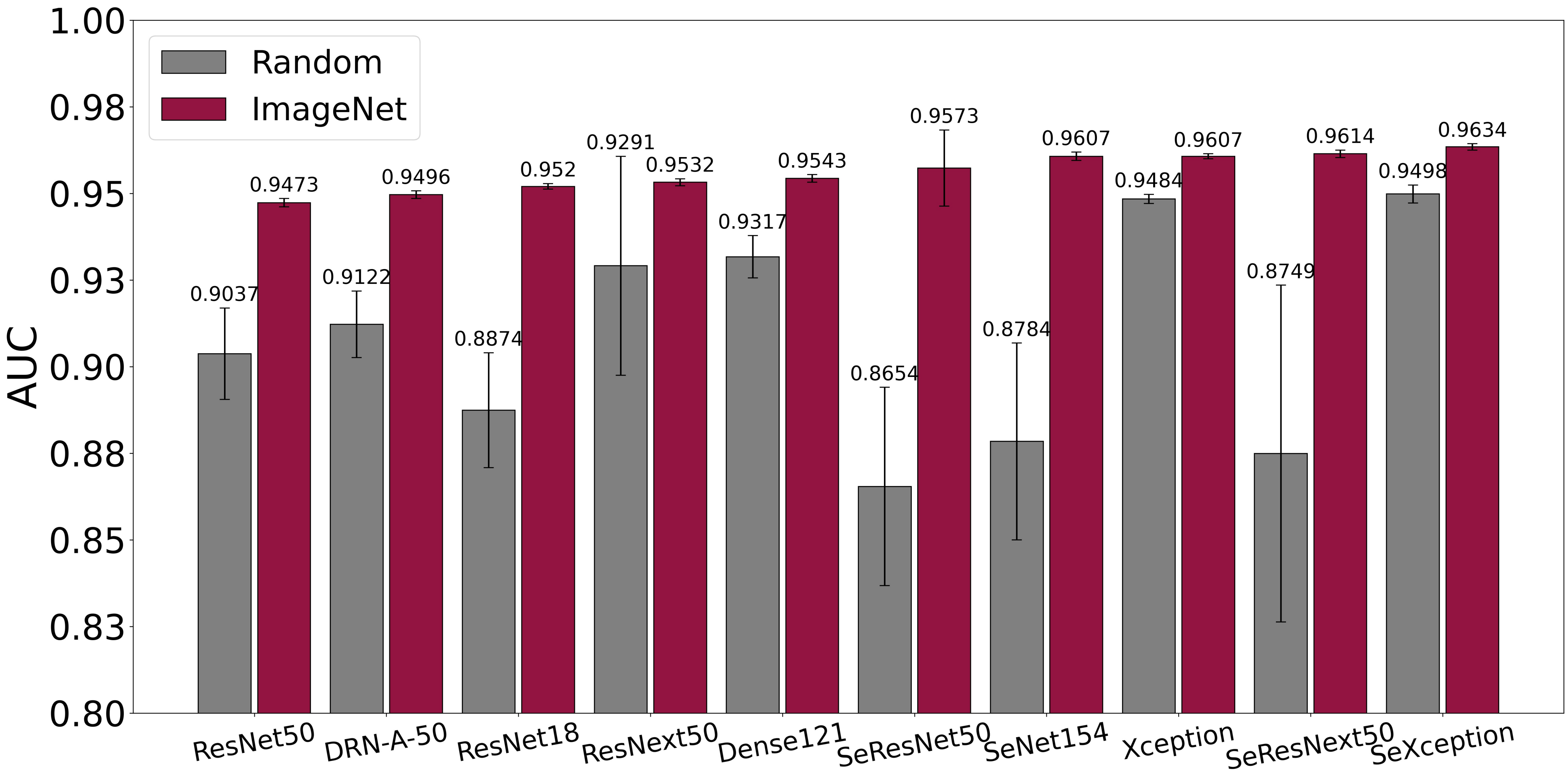
- Training from scratch
- Fine-tuning
  - Supervised pretrained models
  - Self-supervised pretrained models

**Self-supervised Methods**

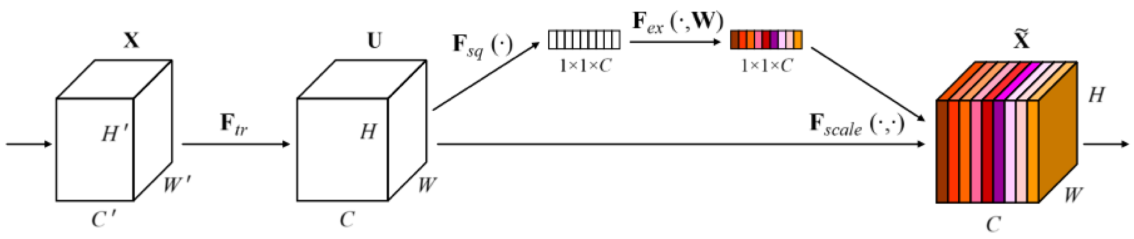
1. InsDis
2. InfoMin
3. MoCo-v1
4. BYOL
5. MoCo-v2
6. DeepCluster-v2
7. PCL-v1
8. SwAV
9. PCL-v2
10. SimCLR-v1
11. PIRL
12. SimCLR-v2
13. SeLa-v2
14. Barlow Twins

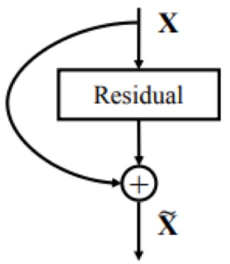
https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection/discussion/194145

# Transfer learning consistently improves performance across the 10 different CNN architectures

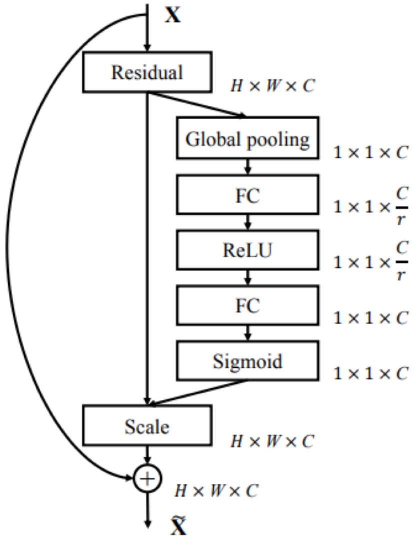


# Squeeze and excitation (SE) blocks enhance CNN performance

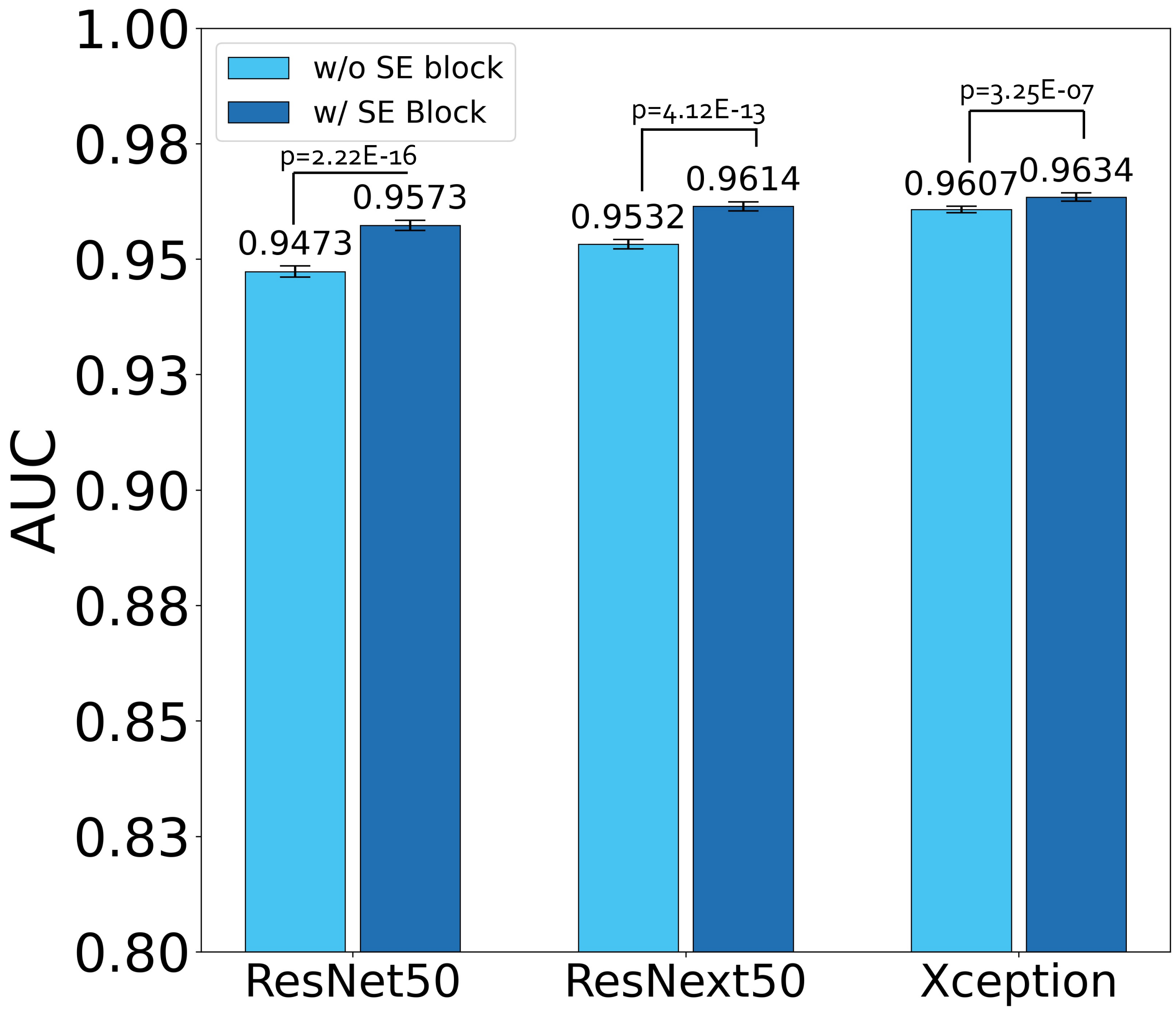


ResNet Module

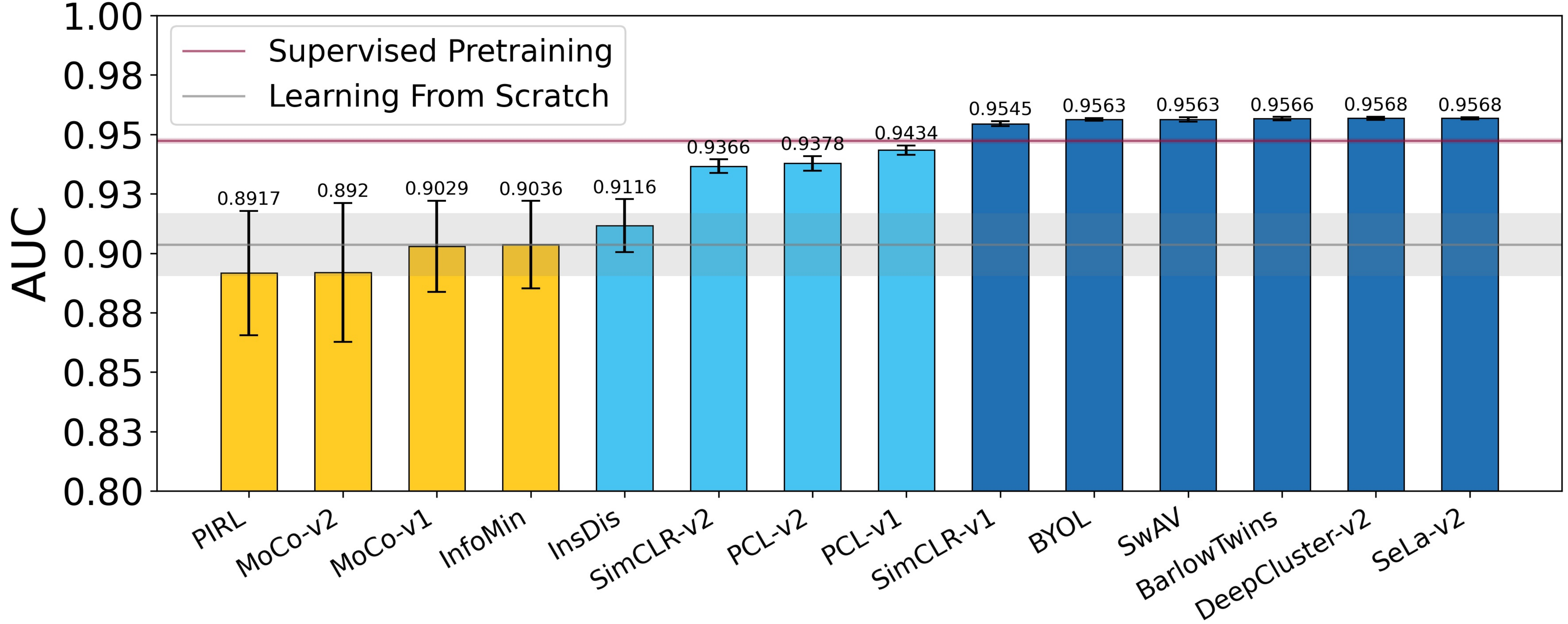


SE-ResNet Module



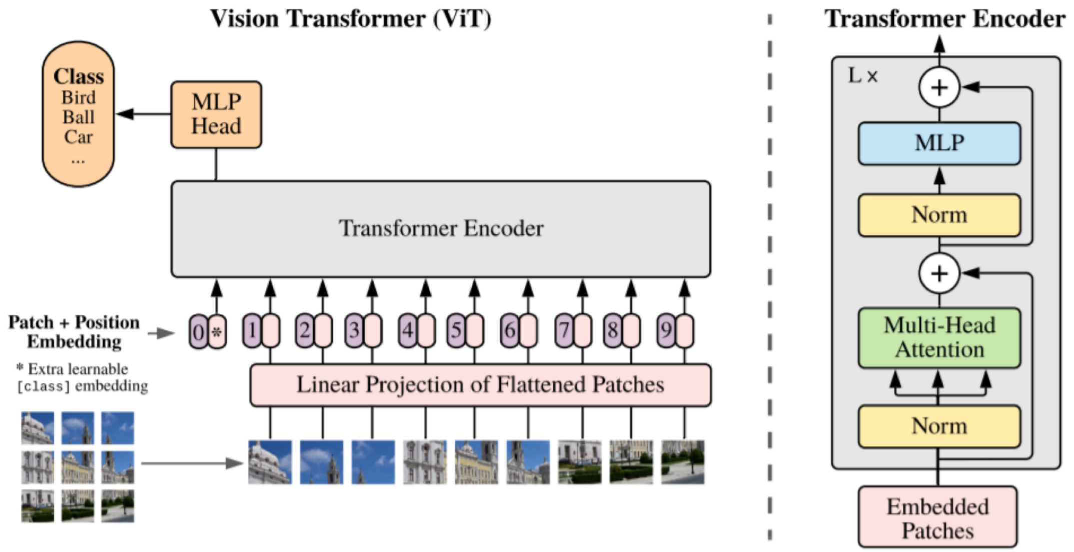

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

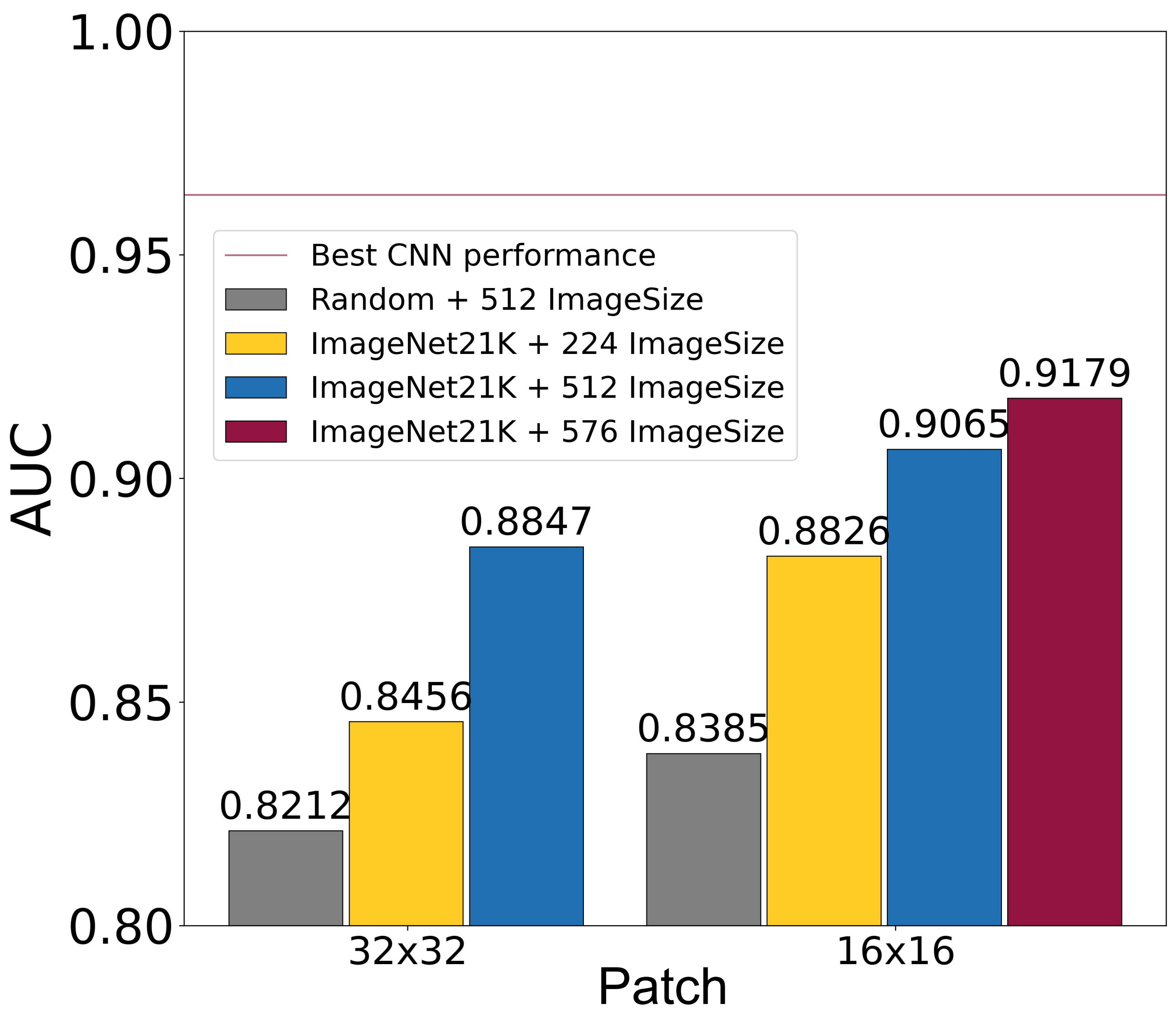
# Self-supervised pre-training overtakes (fully) supervised pre-training




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).

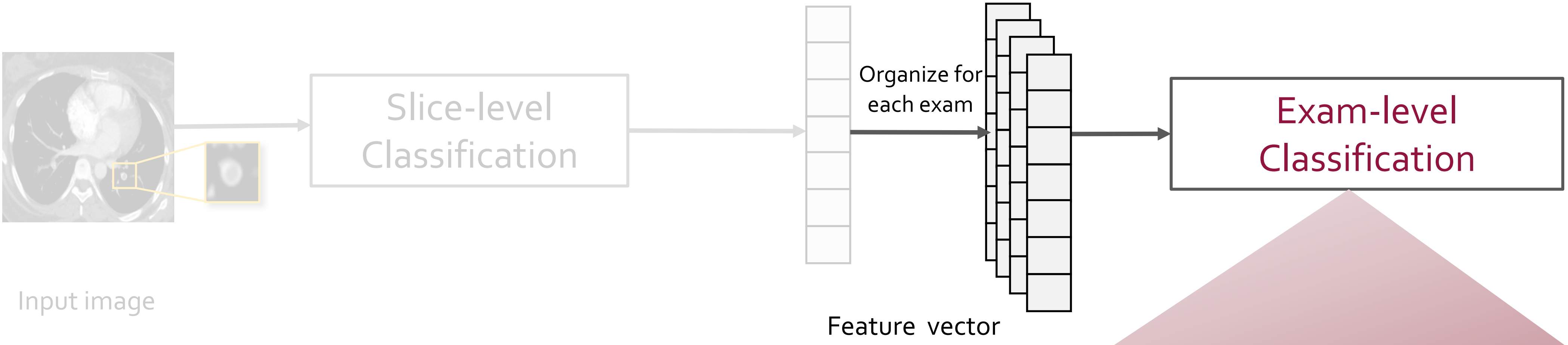
# Vision transformers (ViT) underperform CNNs




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*.

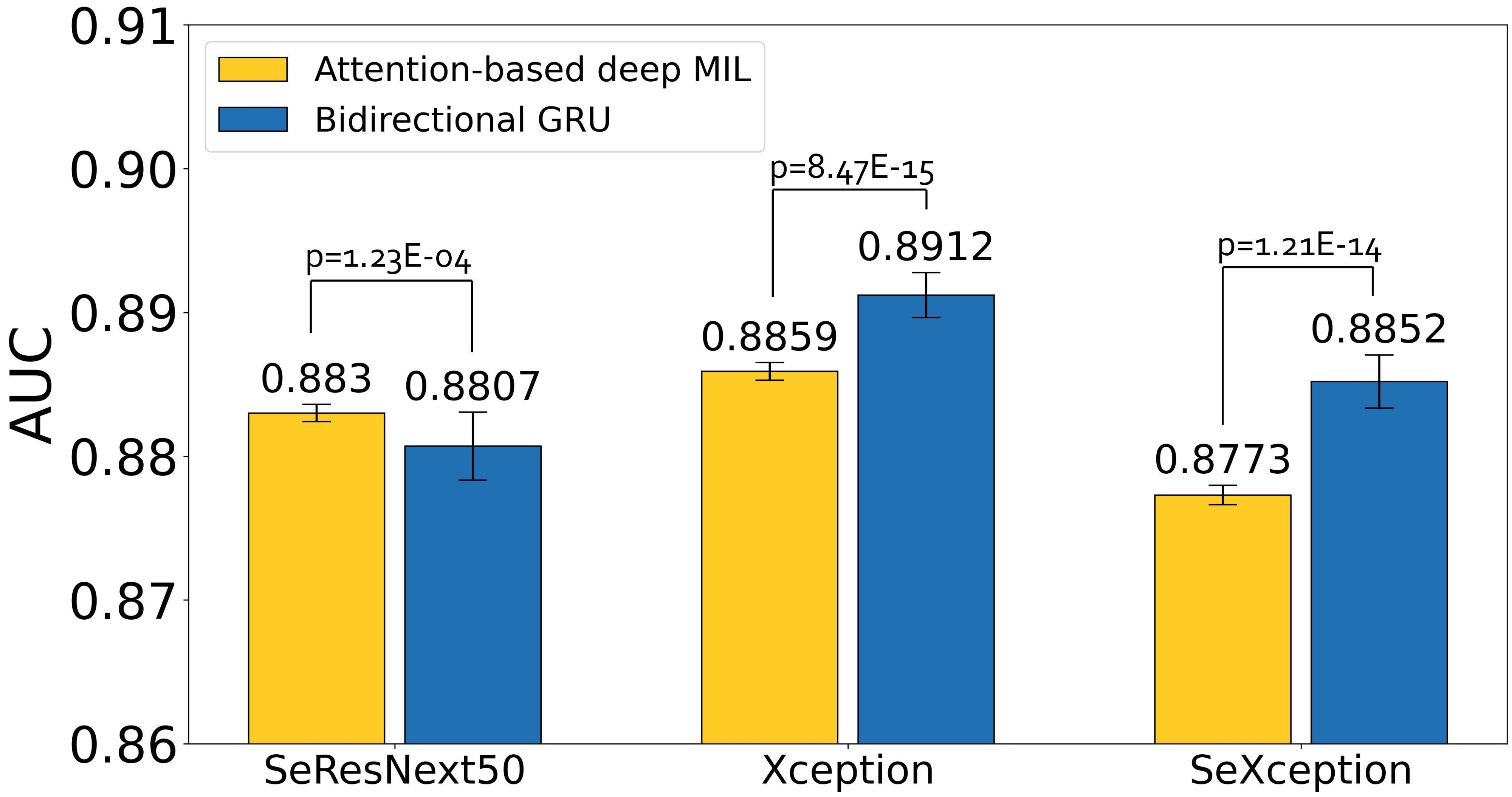
# Second Stage




- Bidirectional gated recurrent unit (GRU)
- Attention-based deep multiple instance learning (MIL)

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

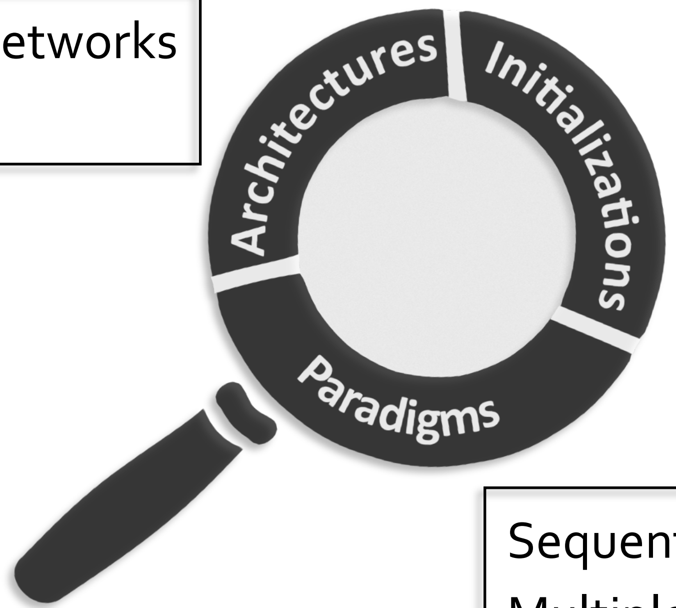
# Bidirectional GRU marginally outperforms attention-based MIL at the exam level




Ilse, M., Tomczak, J., & Welling, M. (2018, July). Attention-based deep multiple instance learning. In *International conference on machine learning* (pp. 2127-2136). PMLR.

# Contributions

Convolutional neural networks
Vision transformers

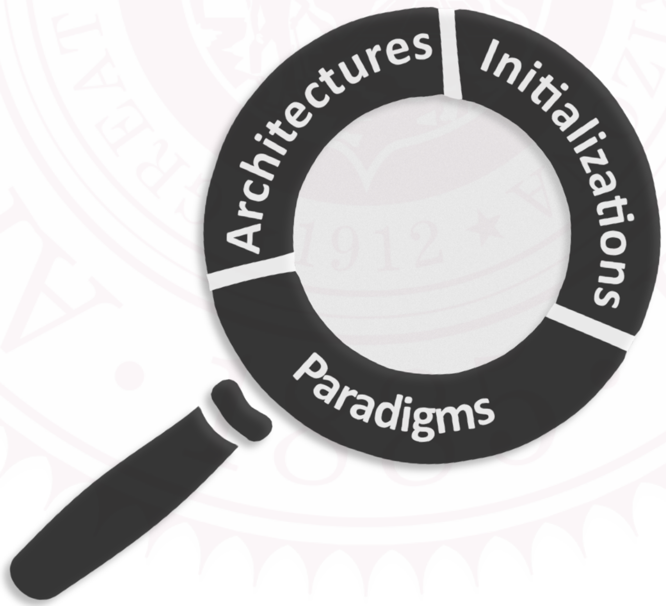



Random initialization
Supervised pre-training
Self-supervised pre-training

Sequential learning model
Multiple instance learning

Compared with the state-of-the-art, our optimal approach provides an AUC gain of 0.2% and 1.05% for slice-level and examination-level, respectively.

# Seeking an Optimal Approach for Computer-aided Pulmonary Embolism Detection




Poster Session, MLMI2021-P-46

Monday, 27th September 2021, 15:50 - 16:40

**https://github.com/JLiangLab/CAD_PE**