LUS Image Classification of Covid-19 and Pneumonia

Abhijith C
242CS003
Under the guidance of **Dr. Jeny Rajan**

March 24, 2025

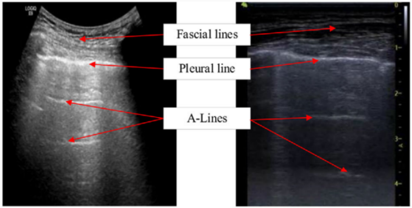
Introduction

LUS Images

Lung ultrasound (LUS) images provide key diagnostic information by capturing different artifacts and structures in the lungs. Key elements in LUS images include:

- **Pleural Line** – The bright, horizontal line at the top of the image, representing the lung's surface.
- **A-Lines** – Repetitive horizontal artifacts indicating normal aerated lungs.
- **B-Lines** – Vertical, hyperechoic (bright) lines extending to the bottom, associated with lung pathologies like pneumonia and pulmonary edema.

LUS Images



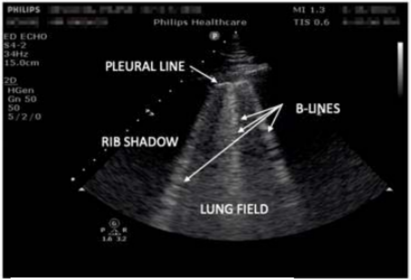


Image Sources: [1], [2]

LUS Differences: COVID-19 vs. Pneumonia

Feature	**COVID-19**	**Pneumonia**
B-Lines	Scattered, widespread	Localized in one area
Pleural Line	Thick and uneven	Mostly normal
Lung Involvement	Both lungs (bilateral)	One lung (unilateral)

Dataset

Dataset Overview

Label	Count
COVID-19	524
Pneumonia	463

Table 1: Dataset Label Distribution

Data Loading and Preprocessing

Data Loading Process:

- Resize images to uniform dimensions (224×224).
- Convert labels into numerical format (COVID $= 0$, Pneumonia $= 1$).

Data Splitting:

- Train (80%) + Validation (10%) + Test (10%) split.

Data Augmentation Steps

- **Geometric Transformations:**
  - Rotation: Randomly rotates images by up to $\pm 20^\circ$.
  - Width Shift: Shifts images horizontally by $\pm 10\%$.
  - Height Shift: Shifts images vertically by $\pm 10\%$.
  - Shear: Applies shearing transformation up to 0.2 radians.
  - Zoom: Random zooming in/out up to $\pm 20\%$.
  - Horizontal Flip: Randomly flips images left-right.
- **Fill Mode:** Uses `nearest` pixel values to fill missing areas.

Experiments Performed

ResNet50

Model Architecture:

- **Base Model:** Pretrained ResNet50 model with weights trained on ImageNet
- **Custom Layers:**
  - Flatten Layer
  - Fully Connected Layer (512 units, ReLU activation)
  - Output Layer (1 unit, Sigmoid activation)
- **Loss Function:** Binary Cross-Entropy
- **Optimizer:** Adam (Learning Rate = 0.0001)

ResNet50

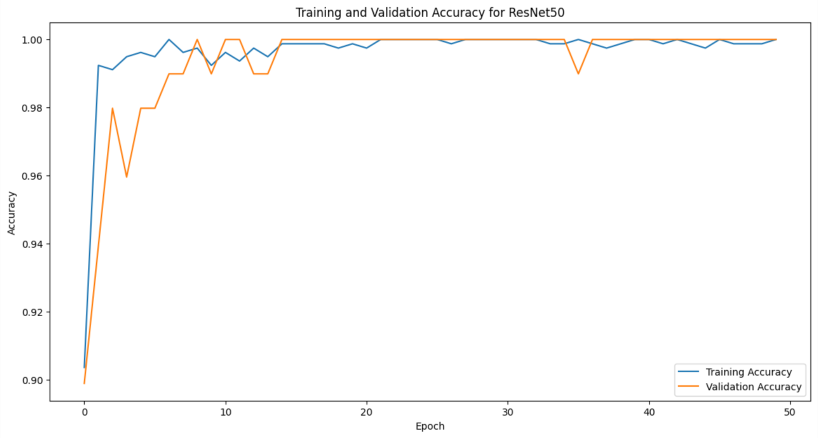


Figure: Accuracy Curve

ResNet50

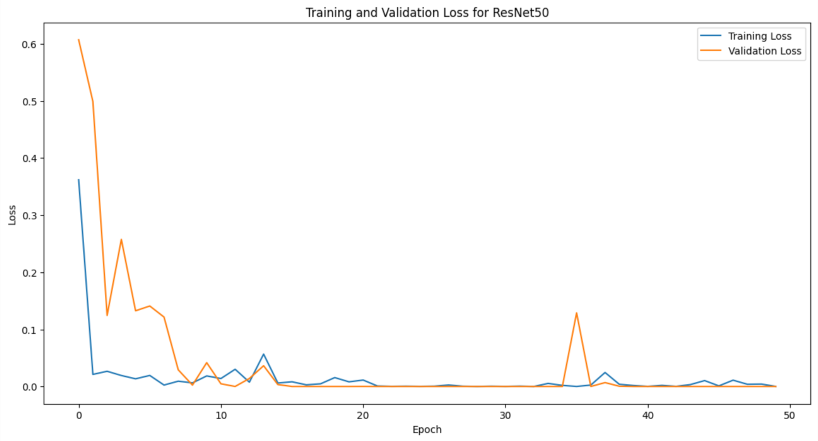


Figure: Loss Curve

ResNet50

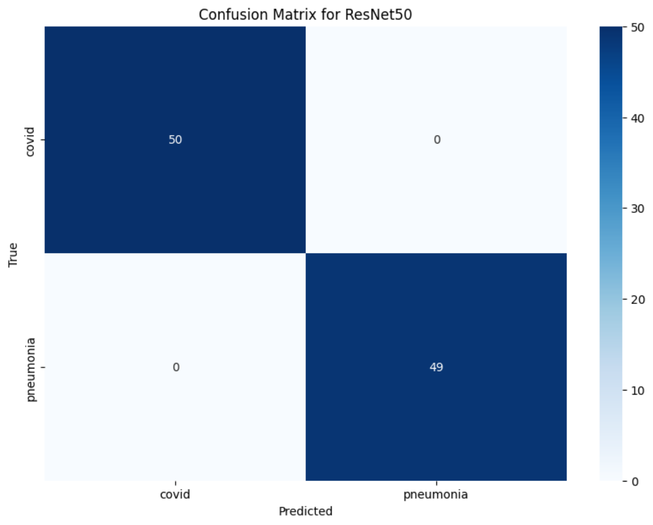


Figure: Confusion Matrix

ResNet50

```
Final Test Accuracy: 1.0000
Final Test Loss: 0.0003
```

```
Classification Report for ResNet50:
              precision    recall  f1-score   support

       covid       1.00      1.00      1.00        50
   pneumonia       1.00      1.00      1.00        49

    accuracy                           1.00        99
   macro avg       1.00      1.00      1.00        99
weighted avg       1.00      1.00      1.00        99
```

VGG16

Model Architecture:

- **Base Model:** Pretrained VGG16 model with weights trained on ImageNet
- **Custom Layers:**
  - Flatten Layer
  - Fully Connected Layer (512 units, ReLU activation)
  - Output Layer (1 unit, Sigmoid activation)
- **Loss Function:** Binary Cross-Entropy
- **Optimizer:** Adam (Learning Rate = 0.0001)

VGG16

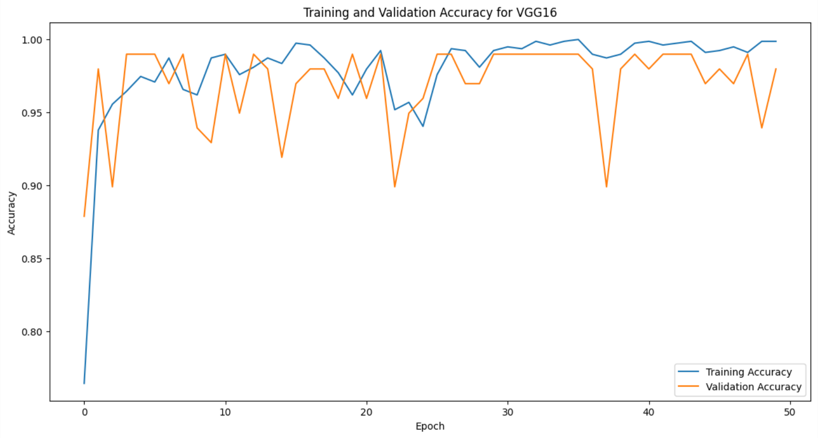


Figure: Accuracy Curve

VGG16

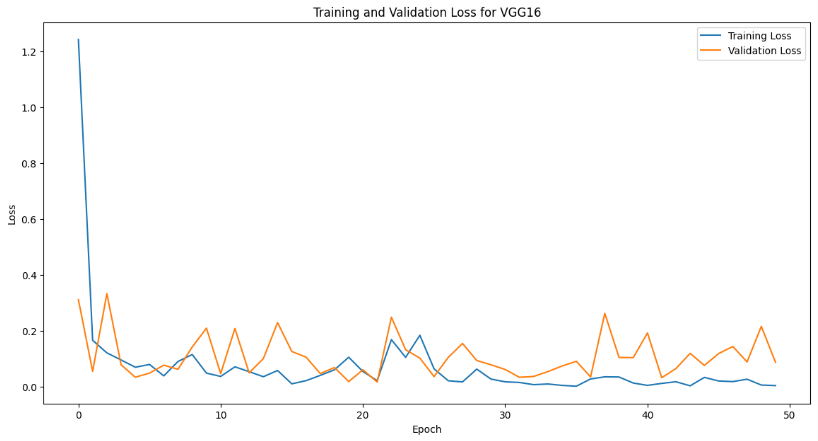


Figure: Loss Curve

VGG16

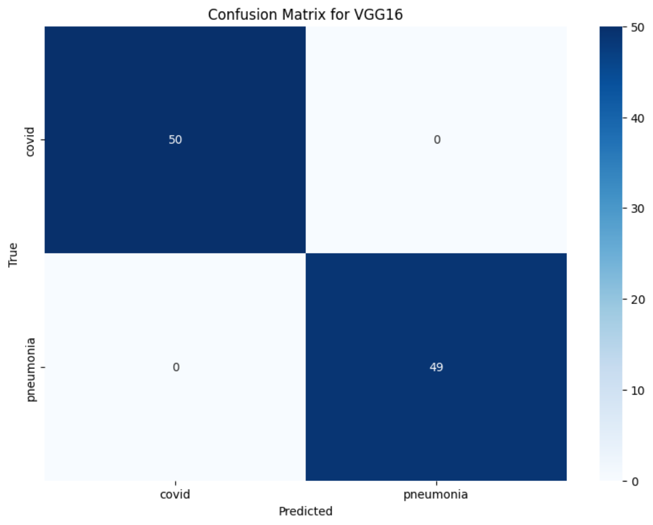


Figure: Confusion Matrix

VGG16

```
Final Test Accuracy: 1.0000
Final Test Loss: 0.0033
```

Classification Report for VGG16:

	precision	recall	f1-score	support
covid	1.00	1.00	1.00	50
pneumonia	1.00	1.00	1.00	49
accuracy			1.00	99
macro avg	1.00	1.00	1.00	99
weighted avg	1.00	1.00	1.00	99

K-Fold Cross-Validation for ResNet50

Key Steps:

- **Splitting Data:** The dataset is divided into 5 folds.
- **Training & Validation:** Each fold is used once as validation while the remaining 4 folds are used for training.
- **Performance Evaluation:** Validation accuracy and loss are computed for each fold.
- **Final Model Training:** After 5 iterations, a final model is trained on the full dataset.
- **Test Set Evaluation:** The final model is evaluated on the test set.

K-Fold Cross-Validation for ResNet50

Fold	Validation Accuracy	Validation Loss
Fold 1	0.9430	0.6063
Fold 2	0.9937	0.0074
Fold 3	0.9937	0.0267
Fold 4	0.9494	1.8617
Fold 5	0.8280	40.4298

```
Average Validation Accuracy across 5 folds: 0.9416
Average Validation Loss across 5 folds: 8.5864
Average Training Time per fold: 117.22 seconds
```

K-Fold Cross-Validation for ResNet50

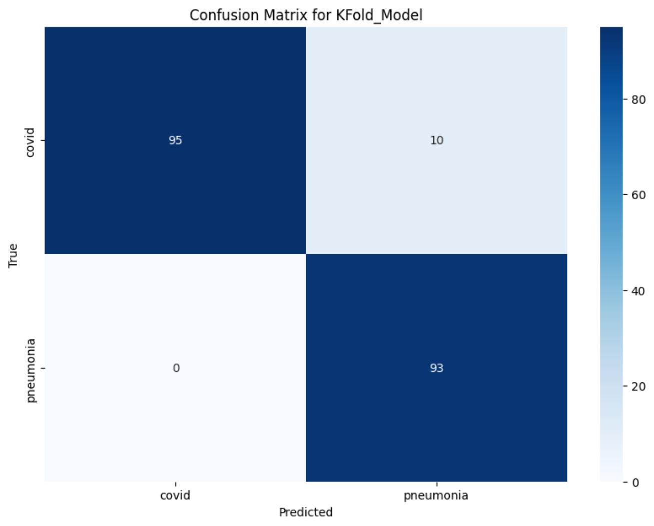


Figure: Confusion Matrix

K-Fold Cross-Validation for ResNet50

```
Final Test Accuracy: 0.9495
Final Test Loss: 6.3086
```

Classification Report for KFold_Model:

	precision	recall	f1-score	support
covid	1.00	0.90	0.95	105
pneumonia	0.90	1.00	0.95	93
accuracy			0.95	198
macro avg	0.95	0.95	0.95	198
weighted avg	0.95	0.95	0.95	198

ResNet50 with Imbalanced Dataset

- **Original Class Distribution:**
  - Class 0 (Covid): 524 samples
  - Class 1 (Pneumonia): 463 samples
- **Imbalanced Class Distribution:**
  - Class 0 (Covid): 524 samples
  - Class 1 (Pneumonia): 46 samples

ResNet50 with Imbalanced Dataset

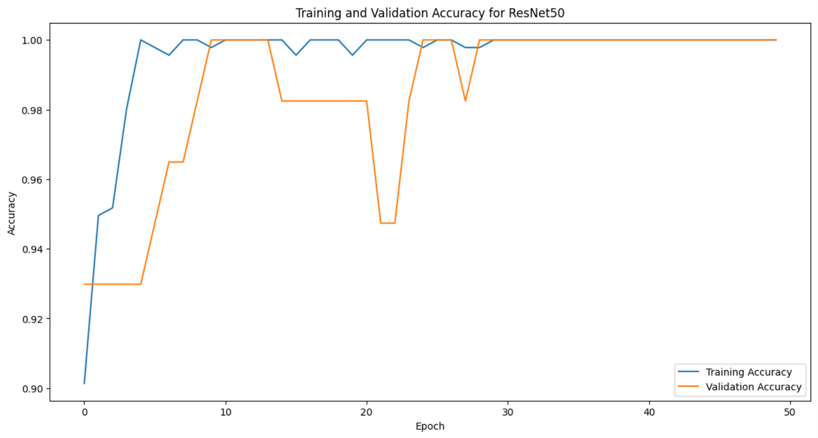


Figure: Accuracy Curve

ResNet50 with Imbalanced Dataset

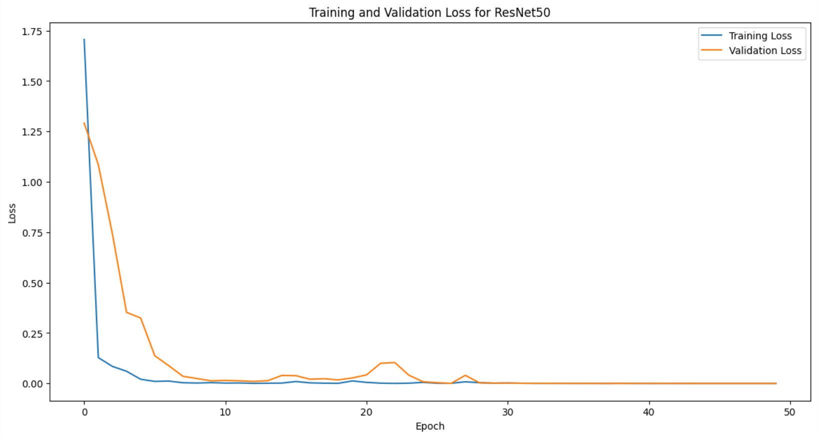


Figure: Loss Curve

ResNet50 with Imbalanced Dataset

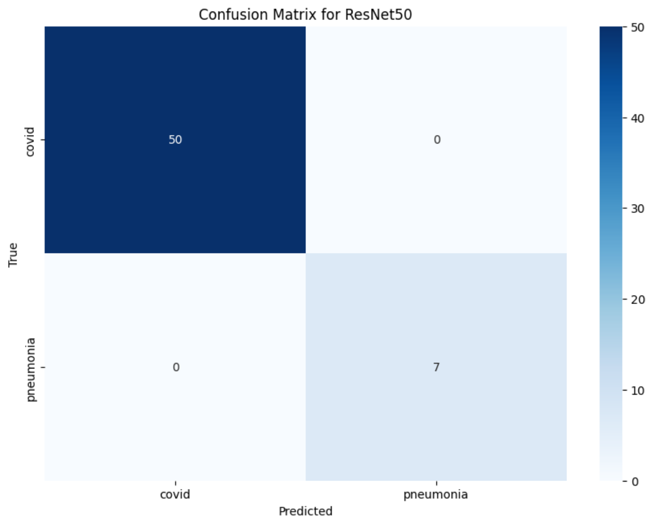


Figure: Confusion Matrix

ResNet50 with Imbalanced Dataset

```
Final Test Accuracy: 1.0000
Final Test Loss: 0.0057
```

```
Classification Report for ResNet50:
              precision    recall  f1-score   support

       covid       1.00      1.00      1.00        50
   pneumonia       1.00      1.00      1.00         7

    accuracy                           1.00        57
   macro avg       1.00      1.00      1.00        57
weighted avg       1.00      1.00      1.00        57
```

Introduction to Vision Transformers (ViTs)

- Originally designed for **Natural Language Processing (NLP)**, used in models like **GPT** and **BERT**.
- **First applied to computer vision** in 2020 for **image classification and segmentation** by Dosovitskiy et al. [3].
- The **core idea**:
  - Instead of convolutional filters, **images are divided into patches**.
  - Each patch is treated as a **token**, similar to words in NLP.
  - A standard **Transformer encoder** is used to process these tokens.
- Great potential in Ultrasound Image Analysis [2].

ViT Architecture

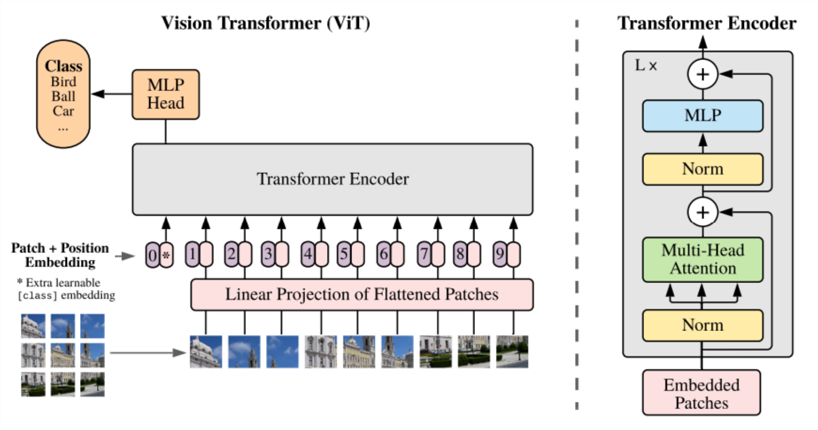


Figure: Model Overview of Vision Transformers

Image Source: [3]

Introduction to Swin Transformer

- Swin Transformer was proposed by Microsoft Research in 2021 [1].
- The **core idea**:
  - Uses a **shifted window attention mechanism** instead of global self-attention.
  - Reduces computational complexity compared to ViT.
- Captures Fine Details with Local & Global Features.
- Scales well and processes high resolution images faster.
- Used in COVID-19 Pneumonia Assessment in Lung Ultrasound Images [3].

Swin Tiny Transformer

- **Base Model:** Swin Tiny Transformer pretrained on ImageNet-1K
- **Input Size:** $224 \times 224 \times 3$
- **Frozen Layers:** All except classification head
- **New Classification Head:**
  - Linear (num_features $\rightarrow$ 256)
  - ReLU Activation
  - Dropout (0.5)
  - Linear (256 $\rightarrow$ 2 classes)
- **Loss Function:** CrossEntropyLoss
- **Optimizer:** Adam (lr$=1 \times 10^{-4}$, weight decay$=1 \times 10^{-5}$)
- **Scheduler:** StepLR (Step size = 5, Gamma = 0.5)

Swin Tiny Transformer

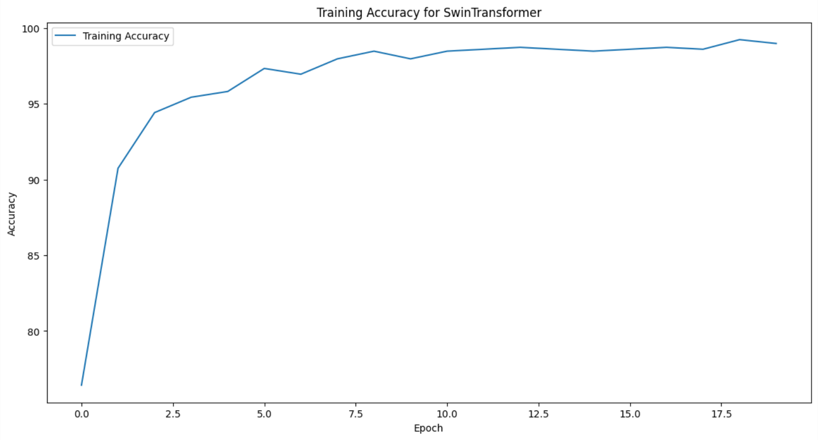


Figure: Accuracy Curve

Swin Tiny Transformer

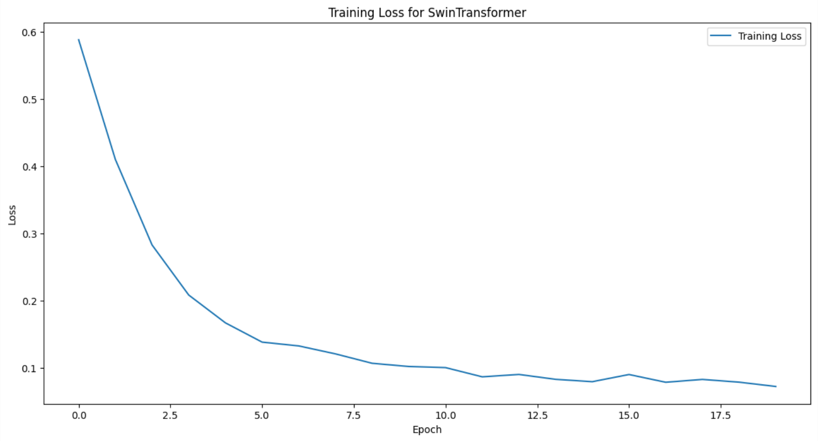
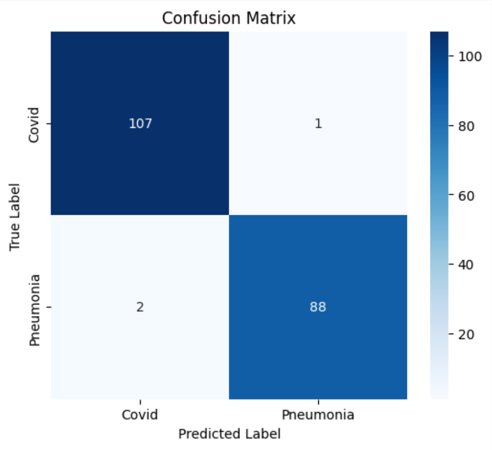


Figure: Loss Curve

Swin Tiny Transformer



Test Accuracy: 98.48%
Test Loss: 0.0783

Conclusion

Summary

Performance Metrics for Different Models

Model	Train Accuracy	Test Accuracy	Train Loss	Test Loss
ResNet50	99.69	100	0.0413	0.003
VGG16	97.59	100	0.1192	0.033
ResNet50 K-Fold	99.60	94.95	0.0123	6.3086
ResNet50 Imbalanced	97.79	100	0.2008	0.0057
Swin Transformer	96.60	98.48	0.1477	0.0783

References

- Xing, W., et al. "Automatic detection of A-line in lung ultrasound using deep learning." Medical Physics 50.1 (2023): 330-343.
- Baloescu, C., et al. "Automated lung ultrasound B-line assessment using deep learning." IEEE T-UFFC 67.11 (2020): 2312-2320.
- Dosovitskiy, A., et al. "An image is worth 16x16 words: Transformers for image recognition at scale." arXiv:2010.11929 (2020).

References

- Liu, Z., et al. "Swin Transformer: Hierarchical ViT using shifted windows." ICCV (2021).
- Vafaeezadeh, M., Behnam, H., & Gifani, P. "Ultrasound image analysis with vision transformers." Diagnostics 14.5 (2024): 542.
- Fiorentino, M. C., et al. "ViT approaches for COVID-19 pneumonia assessment in lung ultrasound." IEEE MetroXRAINE (2024).

Thank You!