Scoliosis Detection on X-ray Images via Transfer Learning

Arnel Malubay
Computational Systems Biology Laboratory
Supervisor : Prof. Naoaki ONO

Overview

I. Introduction

II. EfficientNet-based Model

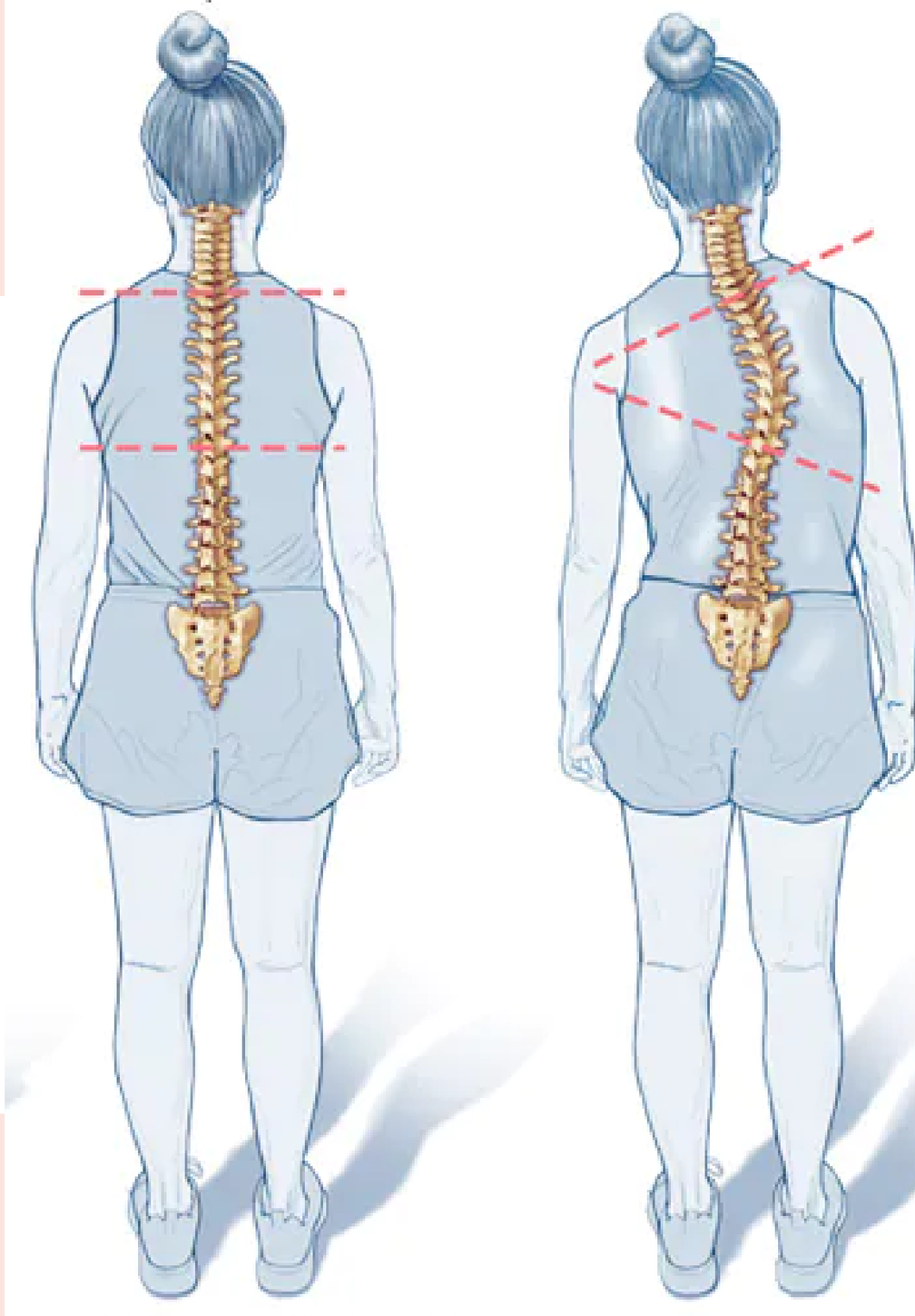
III. ChexNet-based Model

IV. Conclusion

V. Future Directions

Scoliosis

- a **sideways curvature of the spine** that most often is diagnosed in adolescents.
- most cases are **mild, but some curves worsen** as children grow
- computer-aided diagnosis through **X-ray images** are prevalent



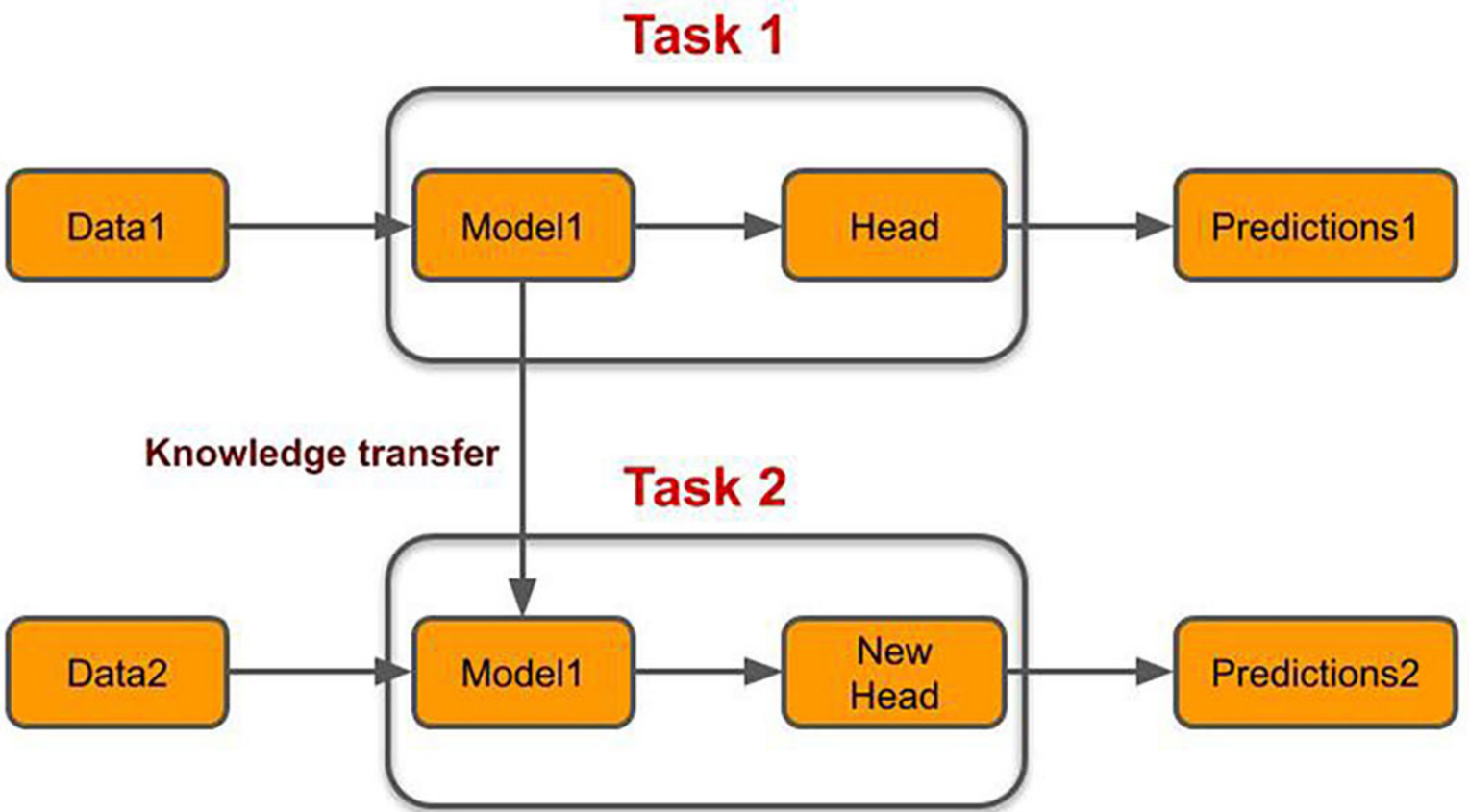
Task: Scoliosis Detection

Approach: **Supervised Machine Learning**
Input: **Chest X-ray Image**
Output: **Diagnosis (Binary Classification)**
Algorithm/Network: **Convolutional Neural Networks**

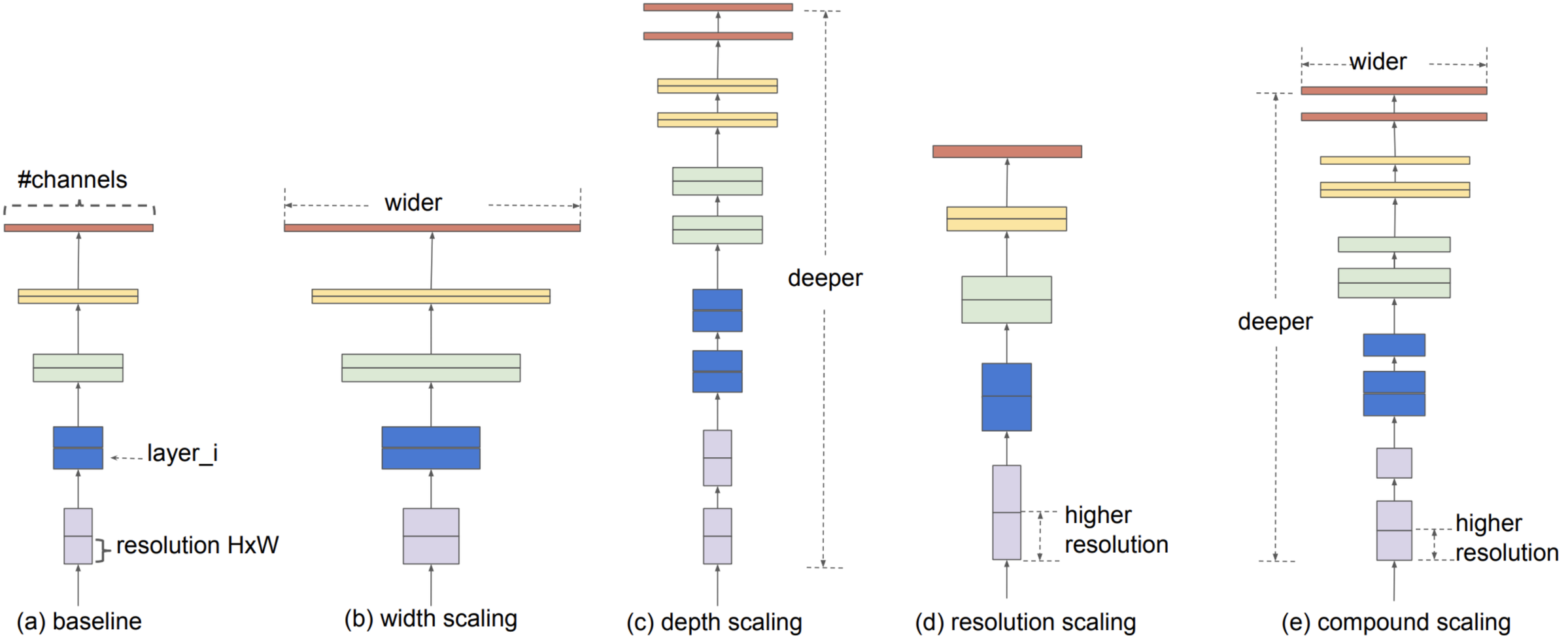
Dataset Overview
Source: **University of Tokyo**
Size: **82,997 (1308 positive, 81689 negative) [imbalanced]**
Resolution: **256 x 256**

Transfer Learning

- reusing a **pretrained model** to **speed up training** considerably and **utilize patterns learned previously**



EfficientNet: structured scaling



EfficientNet-based Model

EfficientNet's key characteristics:
- provides a **principled way to scale up** CNN's
- optimized using **neural architecture search**
- has implementations on **Pytorch and Tensorflow**

Transfer Learning using EfficientNet
Model Used: **EfficientNetB0**
Initial Weights: **Weights based on ImageNet**
Trainable Parameters: **4,729,342**

Training Process

Implementation: **PyTorch**
Number of Epochs: **30**
Optimizer: **AdaGrad**
Learning Rate: **0.001**
Batch Size: **20**

- Each batch is configured to have an **equal number** of positive and negative instances.
- **Random Rotation** was used for data augmentation.
- All **4,729,342** trainable parameters are trained (no layers were frozen)

Performance

	Positive Diagnosis	Negative Diagnosis	Total
Positive Prediction	**1,235**	**3,529**	**4,764**
Negative Prediction	**73**	**78,160**	**78,233**
Total	**1,308**	**81,689**	**82,997**

Accuracy: **95.66%**

Precision: **25.924%**

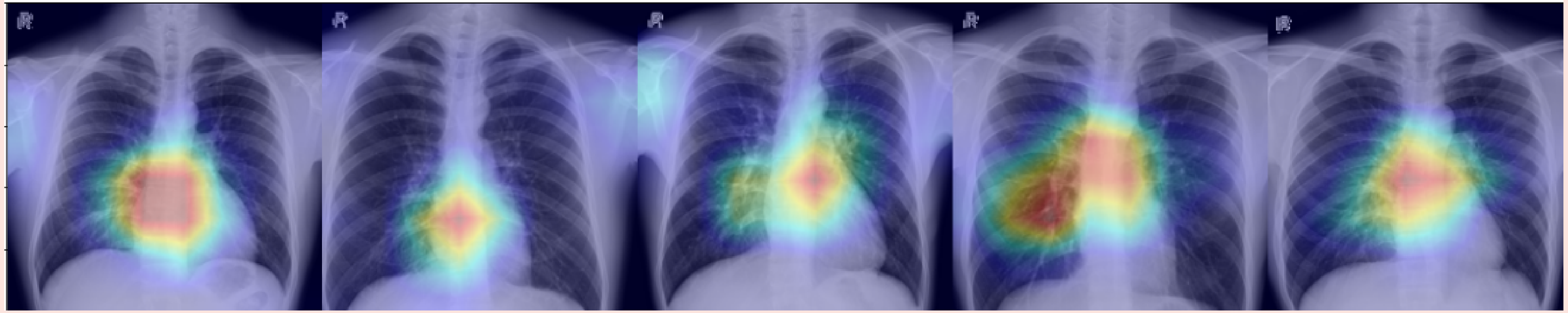
Recall: **94.419%**

F1 Score: **40.679%**

ROC-AUC Score: **98.725%**

Grad-CAM: visualizing region importance

True Positives



False Positives

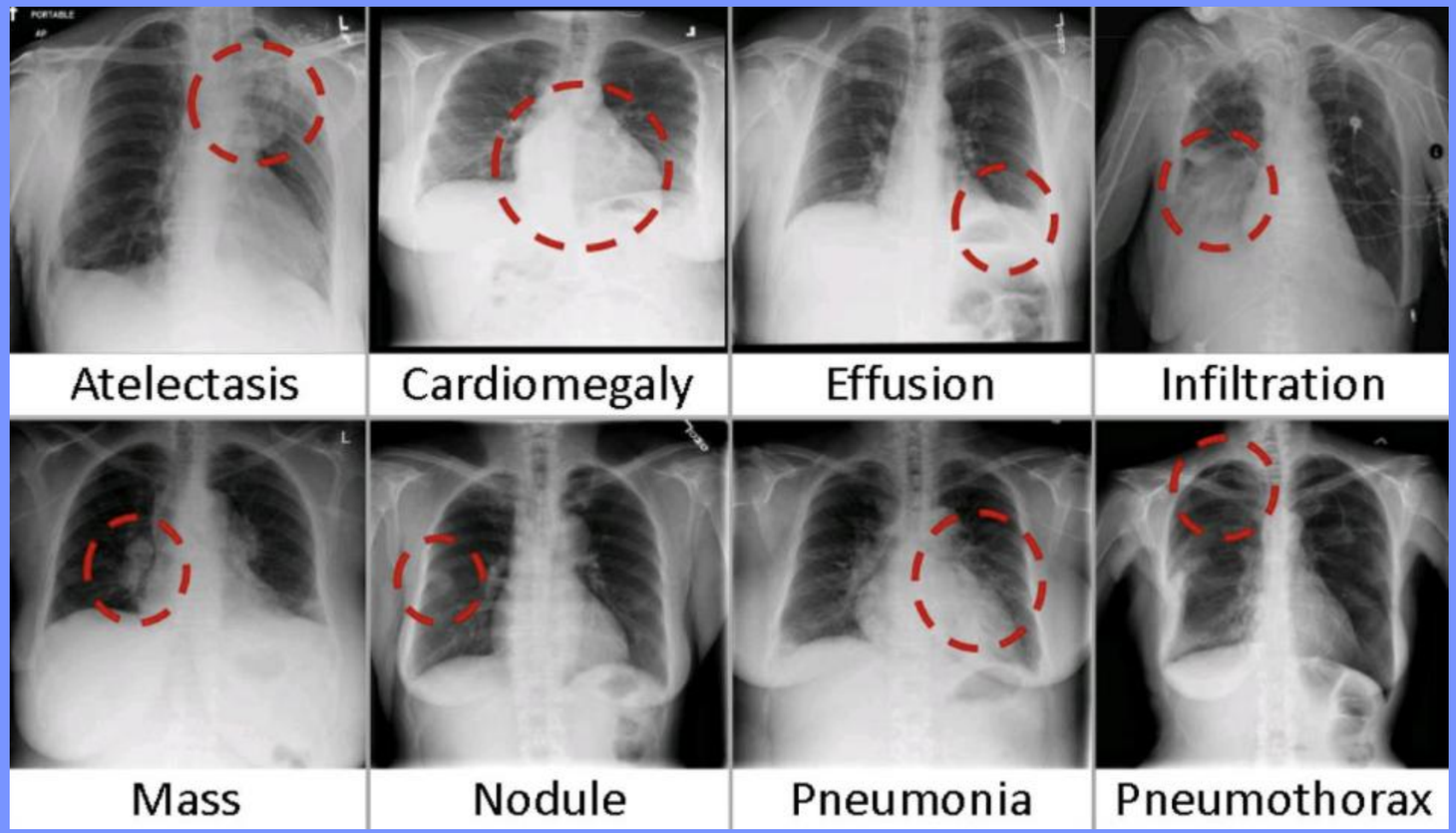
Grad-CAM: visualizing region importance

True Negatives

False Negatives

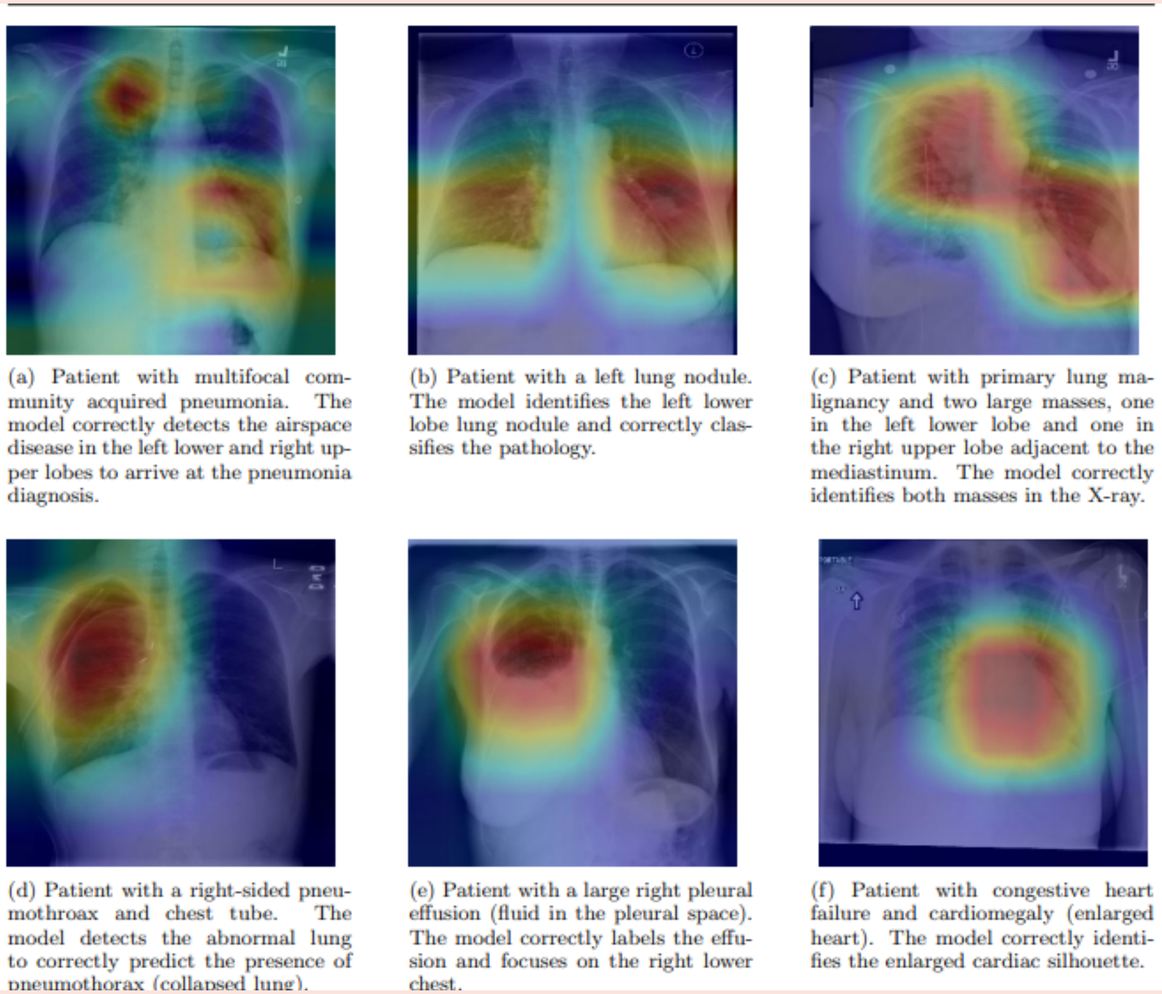
Contextualize further?

ImageNet Dataset



ChestX-ray14 Dataset

CheXNet: radiologist-level pneumonia detection



(a) Patient with multifocal community acquired pneumonia. The model correctly detects the airspace disease in the left lower and right upper lobes to arrive at the pneumonia diagnosis.

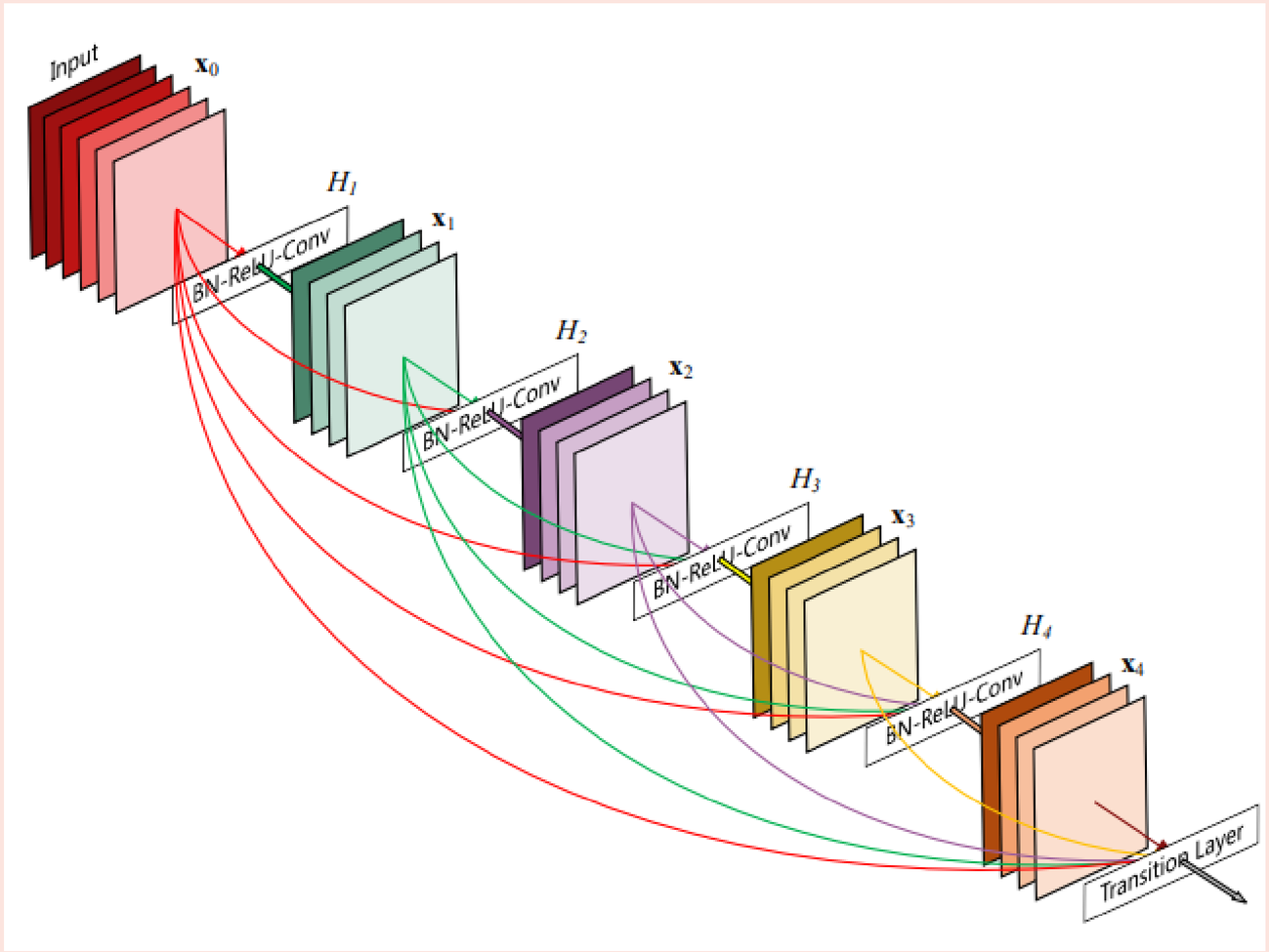
(b) Patient with a left lung nodule. The model identifies the left lower lobe lung nodule and correctly classifies the pathology.

(c) Patient with primary lung malignancy and two large masses, one in the left lower lobe and one in the right upper lobe adjacent to the mediastinum. The model correctly identifies both masses in the X-ray.

(d) Patient with a right-sided pneumothroax and chest tube. The model detects the abnormal lung to correctly predict the presence of pneumothorax (collapsed lung).

(e) Patient with a large right pleural effusion (fluid in the pleural space). The model correctly labels the effusion and focuses on the right lower chest.

(f) Patient with congestive heart failure and cardiomegaly (enlarged heart). The model correctly identifies the enlarged cardiac silhouette.



CheXNet-based Model

CheXNet's key characteristics:

- based on **DenseNet121**
- **outperformed radiologists** in pneumonia detection (based on F1 score)

Transfer Learning using CheXNet

Model Used: **CheXNet (only one version so far)**

Initial Weights: **Weights based on ChestX-ray14**

Trainable Parameters: **6,954,881**

Training Process: Round 1

Implementation: **Tensorflow**
Number of Epochs: **10 (7)**
Optimizer: **Adam**
Learning Rate: **0.001**
Batch Size: **16**

- Each batch is configured to have an **equal number** of positive and negative instances.
- **Random Rotation** was used for data augmentation.
- Only **1025** parameters were trained during this round. These parameters are the ones on the **final dense layer** (all other layers are frozen).

Training Process: Round 2

Implementation: **Tensorflow**
Number of Epochs: **10 (5)**
Optimizer: **Adam**
Learning Rate: **0.0001**
Batch Size: **16**

- Each batch is configured to have an **equal number** of positive and negative instances. **Random Rotation** was used for data augmentation.
- All **6,954,881** trainable parameters are trained (no layers were frozen). Same training, testing, and validation sets were used for both rounds to **prevent data leakage.**

Performance

	Positive Diagnosis	Negative Diagnosis	Total
Positive Prediction	**1,178**	**306**	**1,484**
Negative Prediction	**130**	**81,383**	**81,513**
Total	**1,308**	**81,689**	**82,997**

Accuracy: **99.475%**

Precision: **79.38%**

Recall: **90.061%**

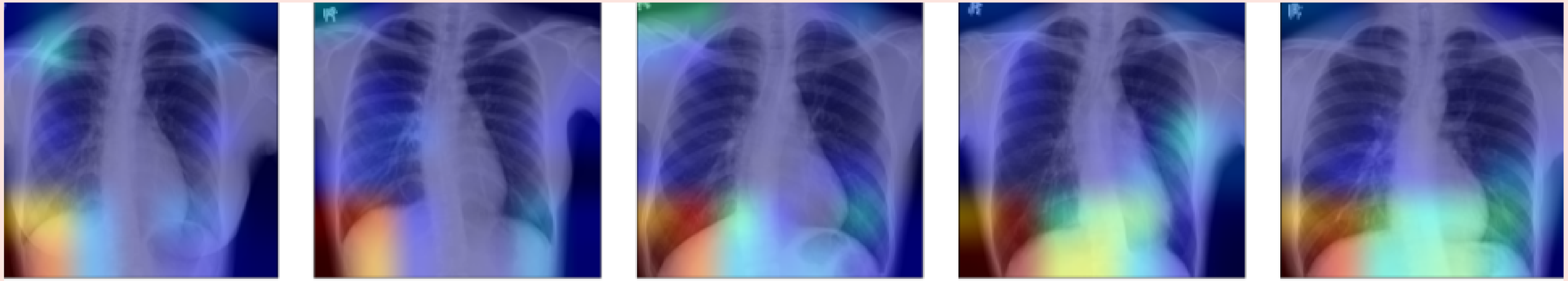
F1 Score: **84.384%**

ROC-AUC Score: **99.138%**

Performance Comparison

	Accuracy	Precision	Recall	F1 Score	ROC-AUC Score
EfficientNet-based Model	95.66%	25.924%	**94.419%**	40.679%	98.725%
CheXNet-based Model	**99.475%**	**79.38%**	90.061%	**84.384%**	**99.138%**

Grad-CAM: visualizing region importance

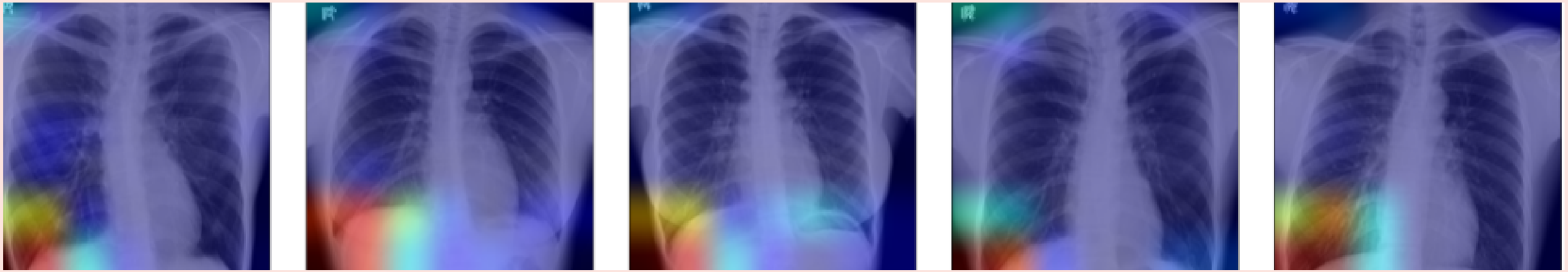


True Positives

False Positives

Grad-CAM: visualizing region importance

True Negatives



False Negatives

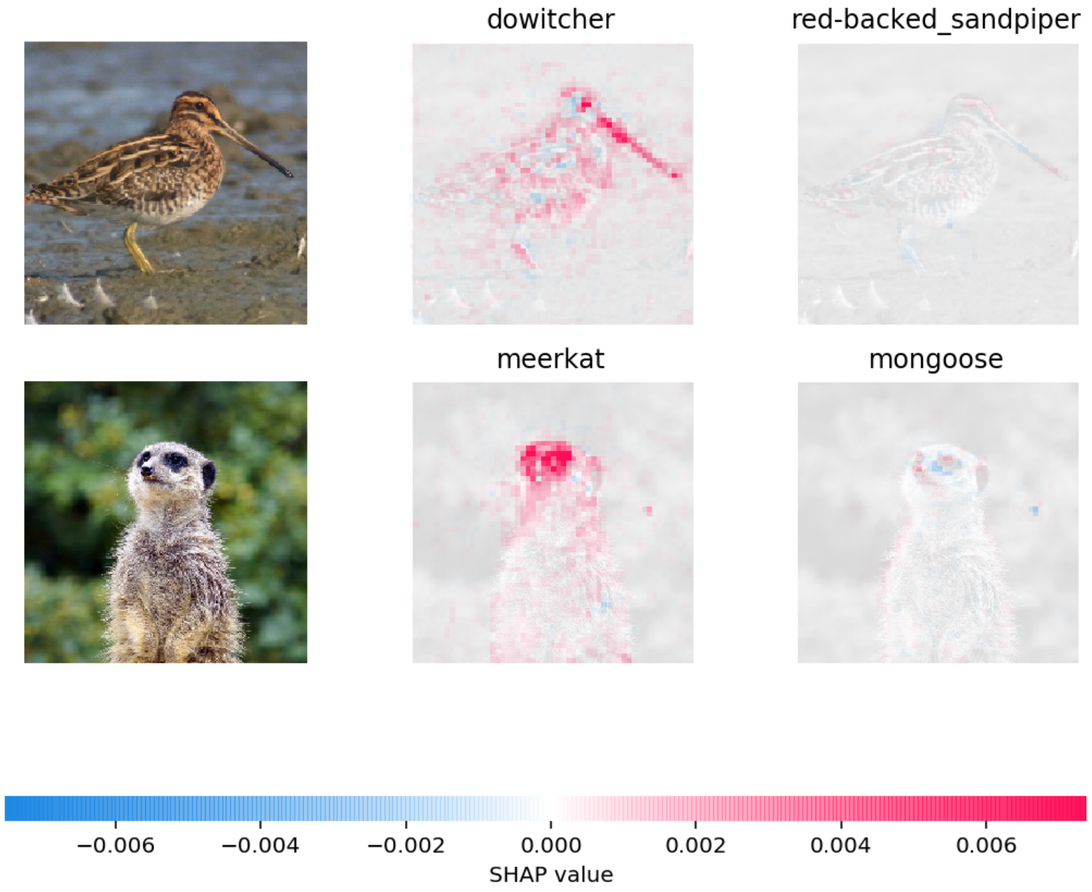
Conclusion

1. Dataset imbalance must be taken care of during both training (either through oversampling or through sample weights) and testing (model evaluation should not rely entirely on accuracy).
2. Considerably better performance and faster training convergence were observed when doing transfer learning on a pretrained model that is "closer" in context.

Future Directions

1. Investigate Grad-CAM results.
2. Experiment on different thresholds for classification in order to improve either precision or recall.

1. Explore the use of SHAP GradientExplainer on X-ray images.



Sources:

- https://www.mayoclinic.org/diseases-conditions/scoliosis/symptoms-causes/syc-20350716
- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow Second Edition by: Aurélien Géron
- https://www.topbots.com/transfer-learning-in-nlp/
- https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html
- https://paperswithcode.com/dataset/imagenet
- https://paperswithcode.com/dataset/chestx-ray14
- https://arxiv.org/pdf/1711.05225v3.pdf
- https://arxiv.org/pdf/1608.06993.pdf
- https://github.com/slundberg/shap