STAT 453 PROJECT

Detection of Lung Diseases through image classification

Group 7
Erika Yeonseo Park
Jina Yang
Ching-Wen Wang

Contents

1. Introduction

- Motivation
- Objective
- Method Description
- Roadmap

Current Situation

COVID-19 Pandemic

Severity

- Easy Spread
- Fatal to Vulnerable & Elderly

Early Diagnosis

Confusion in Diagnosis

- similar signs & symptoms to other respiratory disease

Accurate Diagnosis

Lacking Resources

- medical resources vs. amount of patients

Efficient Diagnosis

Objective

- **Quickly and Accurately** identify lung diseases
- Guide patients for **appropriate medical treatment**
- Efficient **Digital Assistants for Hospitals**

Hospital wait rooms

Our Project Goal:

- **Accuracy:** 90% +
- high recall in **COVID** & **Viral Pneumonia** class

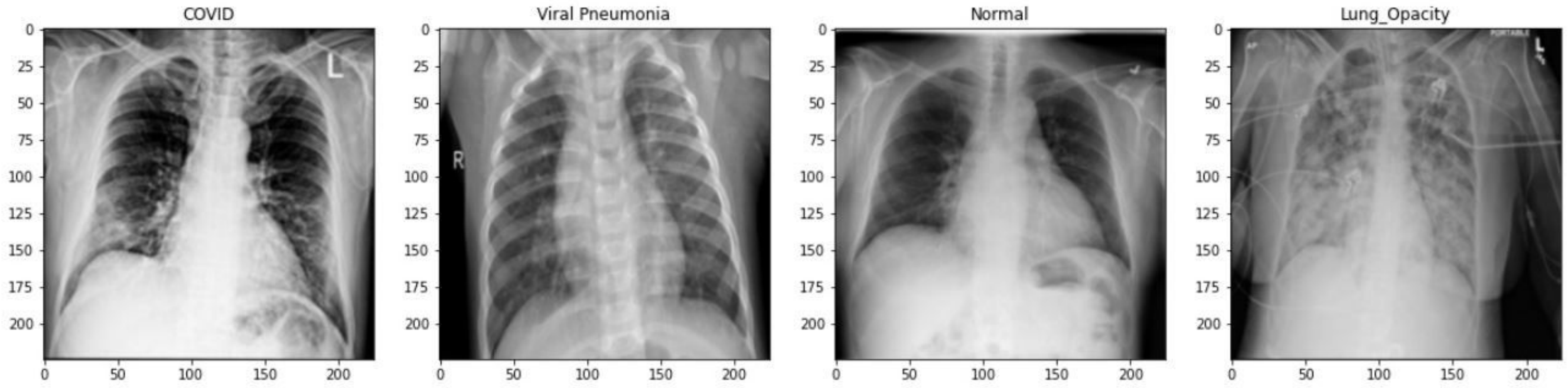
Treatment process

Why X-radiation Scans (X-Ray) ?

- Detect not only COVID, but **“Lung Diseases”**
- CT vs. X-ray
    - X-Ray is more **available** & **affordable** & **faster** & **high mobility**

X-ray is preferable than other imaging modalities for general public!

Images from Dataset

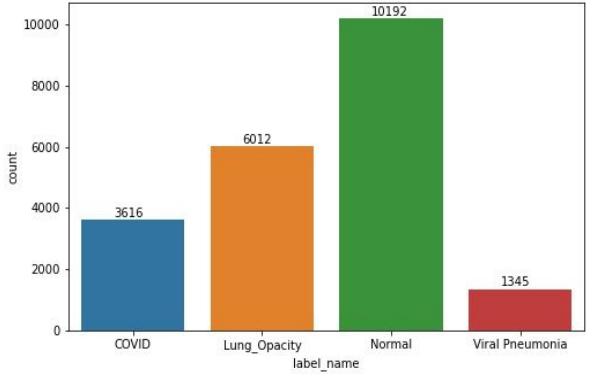


Dataset

Radiography Database

- Gathered from Different Sources + 43 publications
- Contains 21,165 chest X-ray images with 4 classes

Distribution



Class	COVID	Lung Opacity	Viral Pneumonia	Normal
Number of Images	3,616	6,012	1,345	10,192
Total	21,165			

Convolutional Neural Network

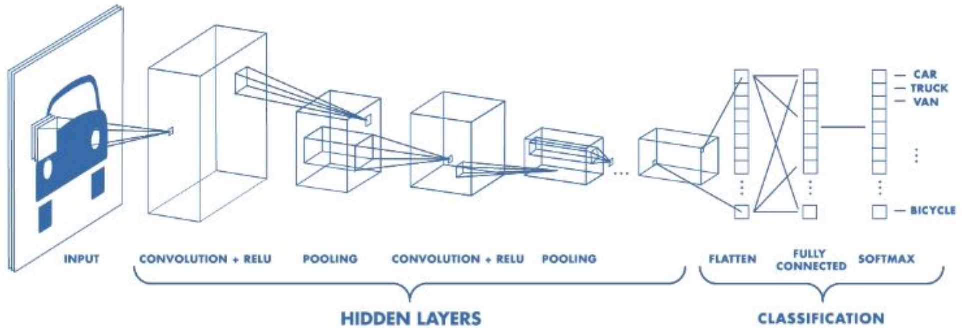
CNN: Basic & reliable architecture for Image classification

Speciality

- "Feature extraction" layers
- pooling layers

Activation functions between layers

- ReLu: Introduce non-linearity



CNN variants

AlexNet

- 8 layers
  - 5 convolution layers
  - 3 fully connected linear layers

Shortcoming:
- overfitting
- less computational efficiency

VGG (Visual Geometry Group)

- Layers more than AlexNet
  - small sized conv filters
  - 3 fully connected linear layers

Shortcoming:
- less computational efficiency
- vanishing gradients

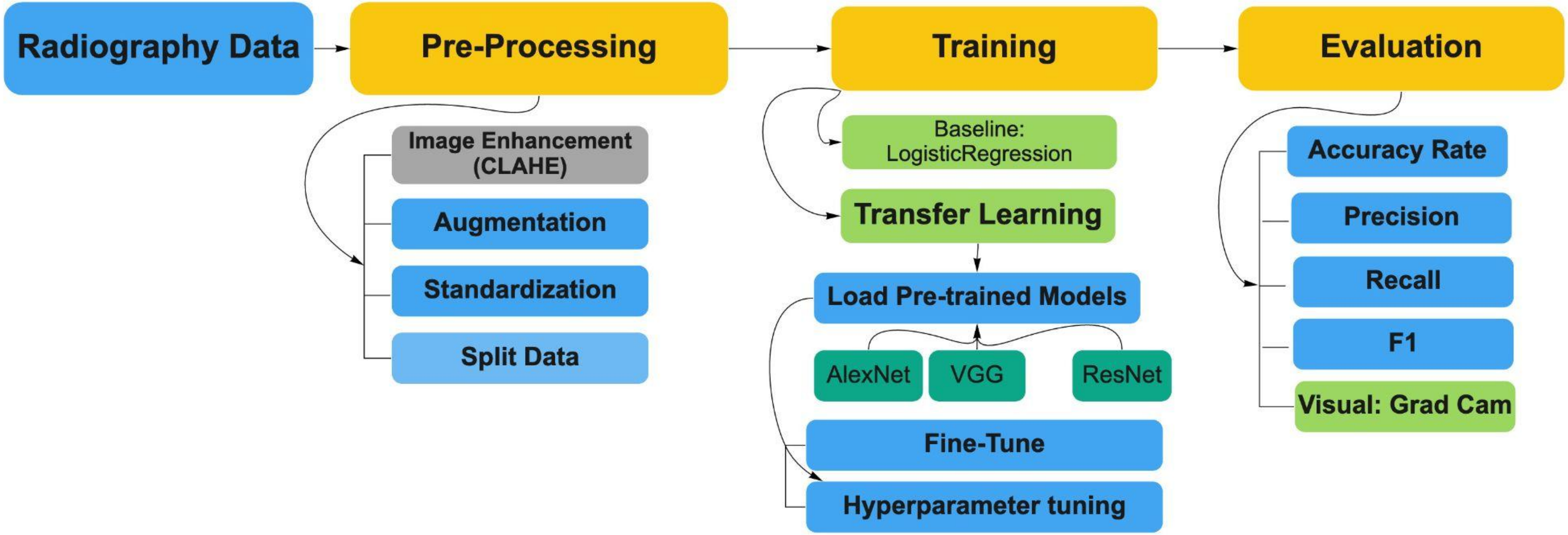
ResNet (Residual Network)

- Simplifies Network
  - small sized conv filters
  - skip layers
  - solve vanishing gradient problem
  - lessen degradation (accuracy saturation)

Shortcoming:
- high training error

Activation: **ReLu** ⇒ Nonlinearity & Amplification of learning effects

General Procedure



2. Train

- Pre-processing
- Train
- Evaluation

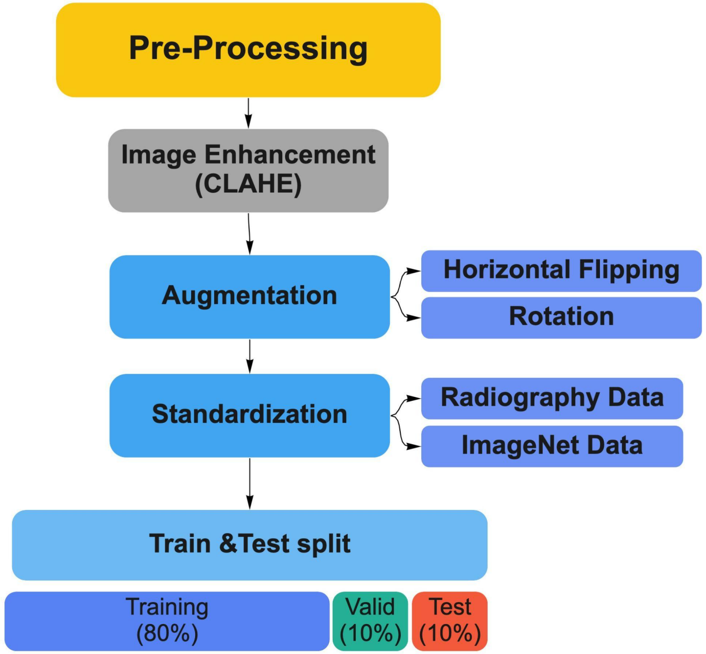
Pre-processing

1. Augmentation

- random rotation
- random horizontal flipping
- Resizing Images:
  - Matching pre-trained models

2. Standardization

- Mean & std. from:
  - Radiography dataset
  - ImageNet dataset



Pre-processing

3. Splitting the Dataset

- Train (80%) → 16, 932 Images
- Validation (10%) → 2,117 Images
- Test (10%) → 2, 116 Image

Equal portion from each class

Before Transformation

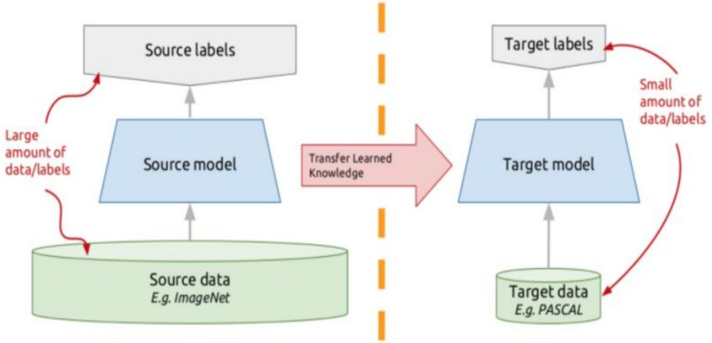
After Transformation

Training: Transfer learning

- Benefit:
  - **generalized better** with small, imbalance dataset
  - relatively **faster training** time
  - available for low spec devices

- Pytorch pre-trained CNN models on ImageNet

  Dataset : 1000 class, 1.2 million RGB images
  - AlexNet
  - VGG
  - ResNet



Hyperparameter tuning

	# Layers	Batch size	# Epoch	Optimizer	Fine-tune
AlexNet	5	32, 64, 128, 256	20 ~ 50	Adam / SGD	Freeze/unfreeze gradients for convolution & linear layers
VGG	17, 19				
ResNet	18, 34, 50				

Evaluation

Evaluation Metrics:

1. Accuracy Rates
2. Precision, Recall, and F1 scores

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Criterion for Selecting good models:

- High Overall Accuracy on test set
- High Recall rates (specifically for COVID & Viral Pneumonia class)

3. Evaluation & Discussion

- Result tables
- findings and Limitations
- Future works

Results I

(FN)TRUE LABEL is the number of images that are misclassified
as "**Normal**" when the real label is **TRUE LABEL** in test set prediction

Model	Hyper Parameters	Train Acc(%)	Test Acc(%)	(FN) COVID	(FN) Viral Pneumonia	(FN) Lung Opacity
AlexNet	**FT:** 5th conv layr & 3 fc Batch size: 64, epoch: 50	97.87	93.86	5	3	63
VGG19_bn	**FT:** last 3 conv layers & 6 fc layrs Batch size: 64, epoch: 20	98.12	94.33	6	5	51
ResNet 50	**FT:** 3rd + 4th layrs (3 blocks each) & last linear layer Batch size: 128, epoch: 20	**98.59**	**95.46**	**3**	**6**	54
Logistic Regression	**FT:** None Batch size: 64, epoch: 50	60.84	62.24	214	25	197

Results II

Precision/ Recall / F1

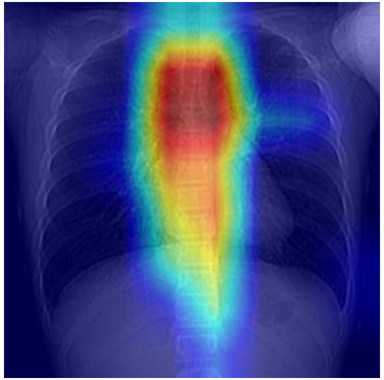
Class label \ Model Name	Logistic Regression	AlexNet	VGG19_bn	Resnet50
COVID	25.86 / 28.75 / 27.23	96.55 / 94.65 / 95.59	97.7 / 94.44 / 96.05	98.28 / 97.71 / 97.99
Lung Opacity	53.85 / 70.59/ 61.09	88.78 / 94.22 / 91.42	90.38 / 94.31 / 92.31	90.71 / 96.42 / 93.48
Normal	88.04 / 67.14 / 76.19	95.65 / 93.17 / 94.39	95.36 / 93.96 /94.65	97.43 / 93.99 / 95.68
Viral Pneumonia	0.0 / NA / NA	96.97 / 95.52 / 96.24	96.21 / 96.95 / 96.58	95.45 / 96.92 / 96.18

Model Visualization

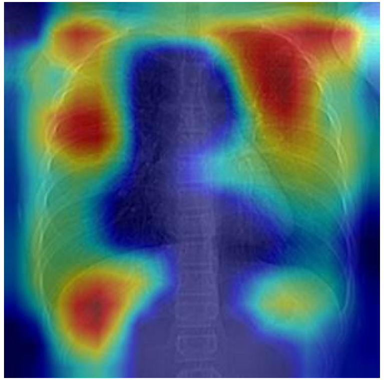
GradCam: visualizes for focus region of model's layer when predicting.

Original
COVID image

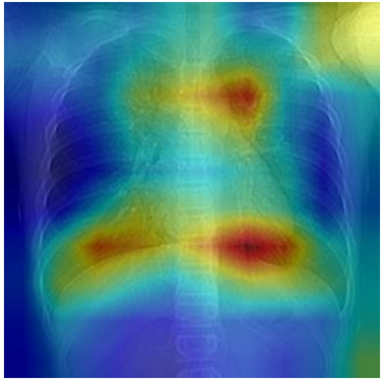
AlexNet



VGG19 batchNorm

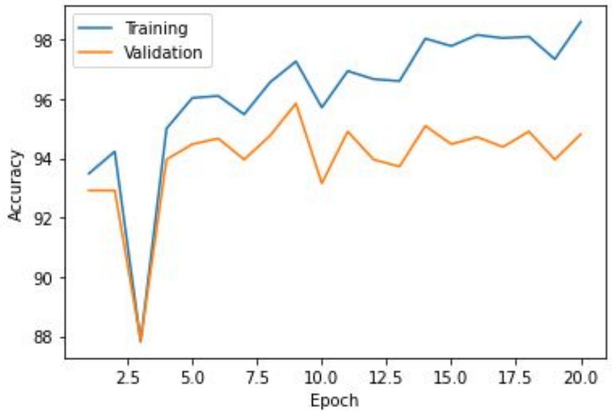


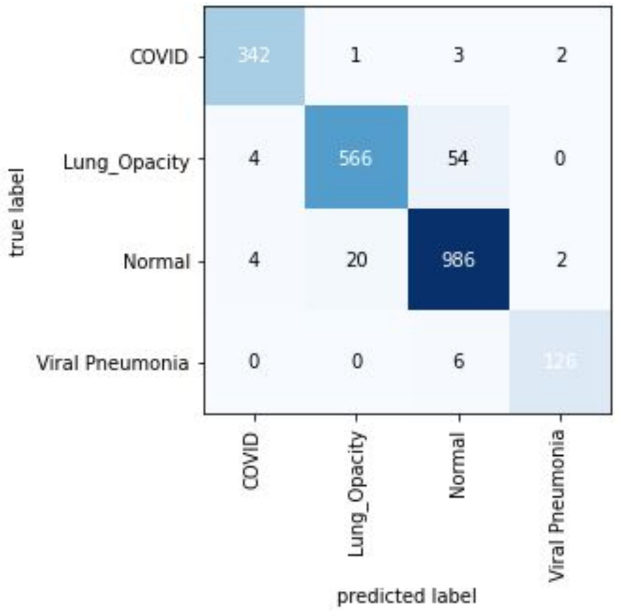
ResNet50



Conclusion

- **Best Model**
  - **ResNet 50** 🏆





Findings & Limitations

- fine tuning
  - freezing + unfreezing gradients in different layers
- limitations of evaluation
- limitations of hardware devices

Future Work

- Image preprocessing methods
  - lung segmentation
  - rib cage shadow removal

- Transfer learning
  - fine tuning more layers
  - models that are pre-trained on X-ray images

Thank You!

Any questions?

ypark258@wisc.edu

cwang553@wisc.edu

jyang633@wisc.edu