CONVOLUTIONAL NEURAL NETWORKS FOR RECOGNIZING COVID-INDUCED PNEUMONIA IN LUNG X-RAY SCANS

PROJECT PROPOSAL

RYAN DIVIGALPITIYA

HEATHER HARTLEY

MOHAMMED ALANBARY

04.05.2021

CS 9542 Advanced AI II

Western
UNIVERSITY · CANADA

AGENDA

TOPIC OVERVIEW

RELATED WORK

DATA SET OVERVIEW

METHODOLOGIES

PRELIMINARY RESULTS

IMPROVEMENT STRATEGIES

EVALUATION METRICS

Healthy

Bacterial Pneumonia

Viral Pneumonia

COVID-19 Pneumonia

TOPIC OVERVIEW

- COVID-19 has plagued our world for over a year
- AI community continuously looking for better ways to detect pneumonia caused by COVID
- We seek to implement a CNN that can differentiate between a healthy patient, a patient that has COVID-induced pneumonia and a patient that has non-COVID pneumonia (3 classes in total)
- Practical use of a CNN that detects COVID-19 pneumonia cases extends to:
  - Treatment decision making
  - Outcome prediction
  - Severity of illness

RELATED WORK

- Why Convolutional Neural Networks?
- Studies show they are more effective at detecting chest diseases than other network types
- Abiyev, Rahib H, and Mohammad Khaleel Sallam Ma'aitah. "**Deep Convolutional Neural Networks for Chest Diseases Detection.**" *Journal of healthcare engineering* vol. 2018 4168538. 1 Aug. 2018, doi:10.1155/2018/4168538
  - "[…] *the CNN has achieved the highest recognition rate for [detecting chest disease type], compared to other employed networks.*"

TWO GOALS

1. First, see if we can build a model that can distinguish between a **healthy patient** versus a **patient with pneumonia** (any kind)

2. If we are successful, then we'll try to build a model that attempts to distinguish between **healthy patients** versus **COVID-induced pneumonia** versus **non-COVID-induced pneumonia**

DATASET

DATASET CHALLENGES

Different Anatomical Planes

Transverse

Coronal

Sagittal

Bad Cropping

Empty Space

Off-Centered

Differing Sizes

Aspect-Ratios

Dimensions

Pre-processing pipeline will have to remove all of these differences in the images

DATASET OVERVIEW

Healthy

(10,000 images)

COVID

(3,600 images)

Bacterial Pneumonia

(6,000 images)

Viral Pneumonia

(1,300 images)

Unhealthy

1

First, see if we can build a model that can distinguish between a **healthy patient** versus a **patient with pneumonia** (any kind)

Healthy

(10,000 images)

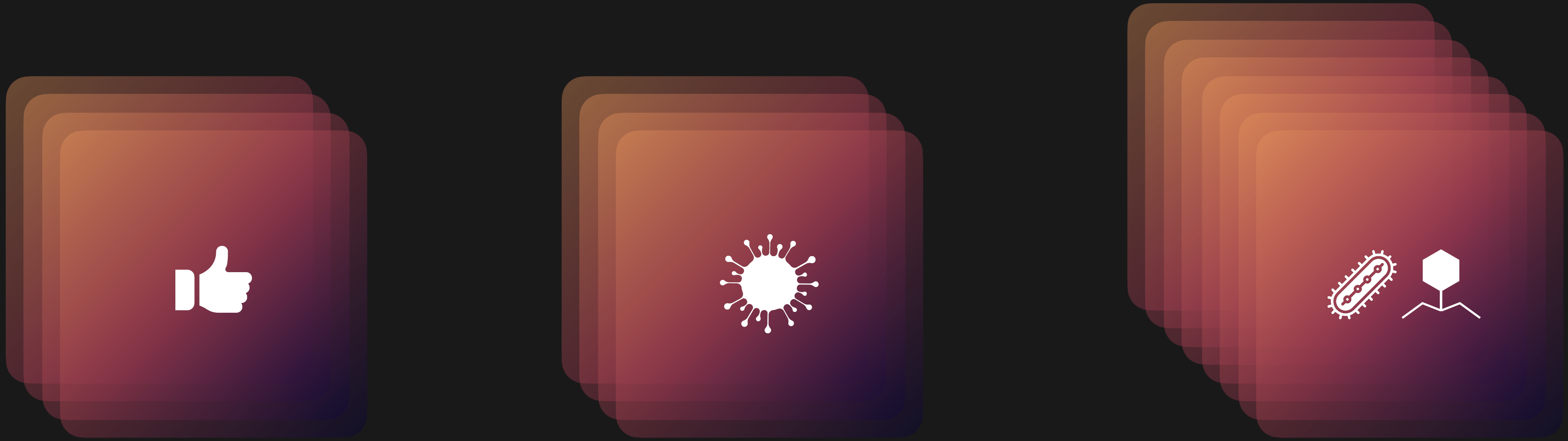
Unhealthy (Pneumonia)

(10,000 images)

2 Classes

2

If we are successful, then we'll try to build a model that attempts to distinguish between **healthy patients** versus **COVID-induced pneumonia** versus **non-COVID-induced pneumonia**



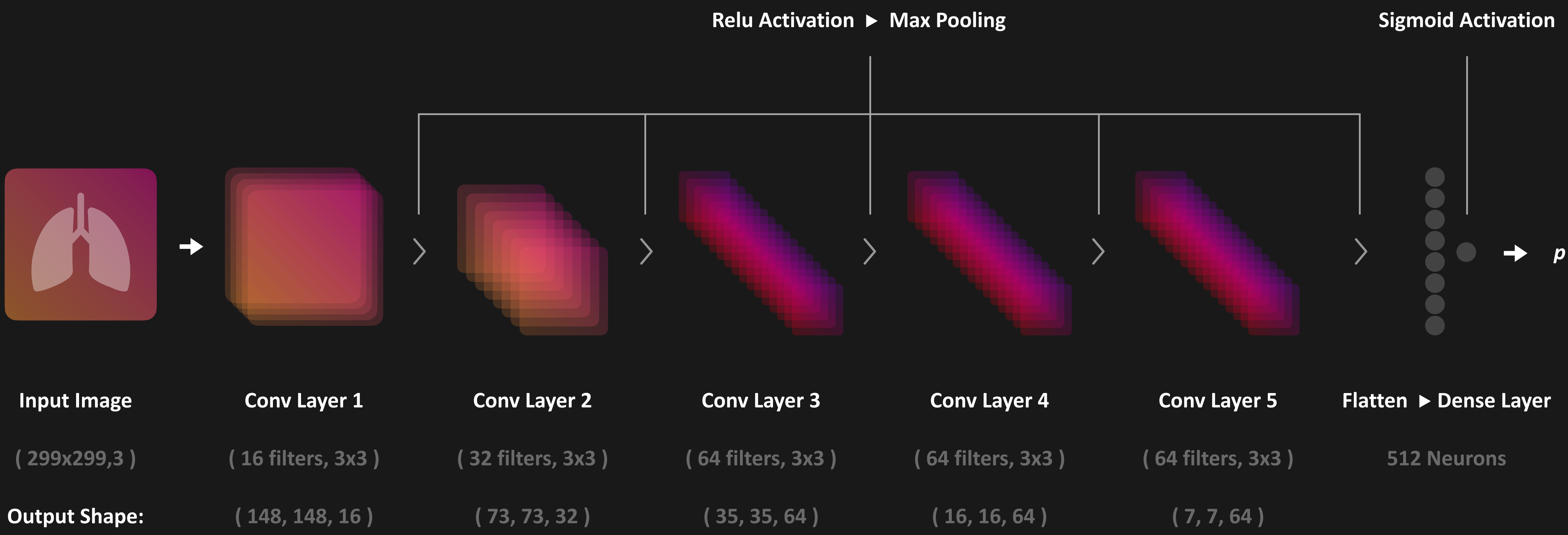
Healthy	COVID	Non-COVID
(10,000 images)	(3,600 images)	(7,300 images)

3 Classes

METHODOLOGIES

Preliminary Neural Net Architecture

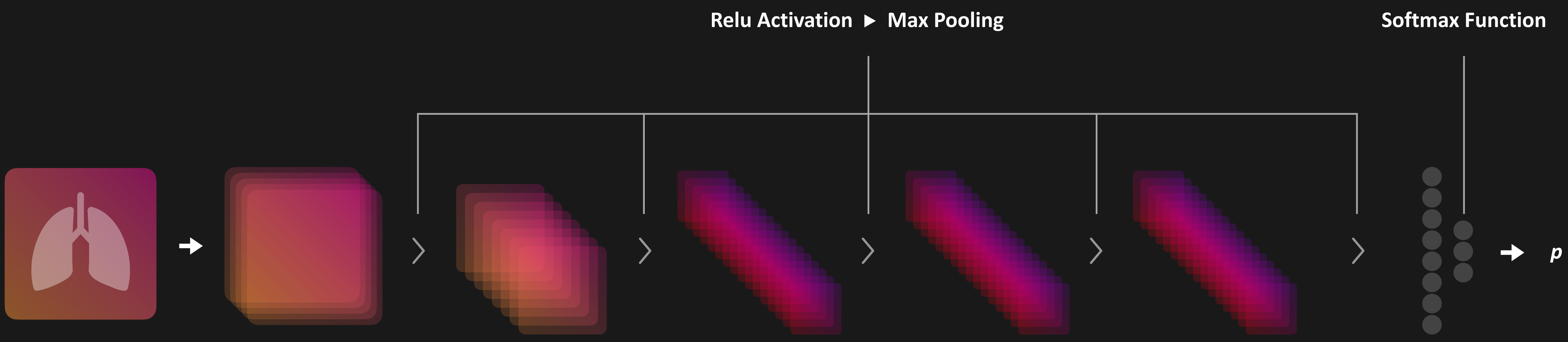
1 Binary Classification (Healthy Patient | Patient with Pneumonia (any kind))



1,704,097 Parameters

Preliminary Neural Net Architecture

2 Multi-Classification (3 Classes: Healthy Patient | Patient with COVID | Patient with Non-COVID Pneumonia)



Input Image	Conv Layer 1	Conv Layer 2	Conv Layer 3	Conv Layer 4	Conv Layer 5	Flatten ▶ Dense Layer
(299x299,3)	(16 filters, 3x3)	(32 filters, 3x3)	(64 filters, 3x3)	(64 filters, 3x3)	(64 filters, 3x3)	512 Neurons
Output Shape:	(148, 148, 16)	(73, 73, 32)	(35, 35, 64)	(16, 16, 64)	(7, 7, 64)	

1,705,123 Parameters

Preliminary Results

1 **Binary Classification (Healthy Patient | Patient with Pneumonia (any kind))**

Training Accuracy	Validation Accuracy
91%	88%

2 **Multi-Classification (3 Classes: Healthy Patient | Patient with COVID | Patient with Non-COVID Pneumonia)**

Training Accuracy	Validation Accuracy
88%	82%

Performance Boosting Strategies

- Pre-processing:
  - Same intensity/brightness range?
  - Did not normalize. We will do that next time to help converge faster

Performance Boosting Strategies

- Tuning Architecture of Network:
  - Tune number of convolution layers
  - Tune number of dense layers
  - **Will have to employ Cross-Validation + evaluation of a Test Set to guard against overfitting**

Performance Boosting Strategies

- Simply increase number of epochs (to an extent)
    - GPU Training

Future Evaluation Metrics

Lesser of two evils:
If we have to choose between higher false positive rate and higher false negative rate, we would want to choose the former: <u>for medical diagnosis, lower false-negative rate should be prioritized</u>

We will explore making use of the following evaluation metrics to properly assess model performance:

1. **Multi-Classification Confusion Matrix**
2. **Receiver Operator Curve**
3. **Precision**
4. **Recall**
5. **F1-Scores**
6. **Specificity**
7. **Sensitivity**
8. **Balanced Accuracy**

THANK YOU