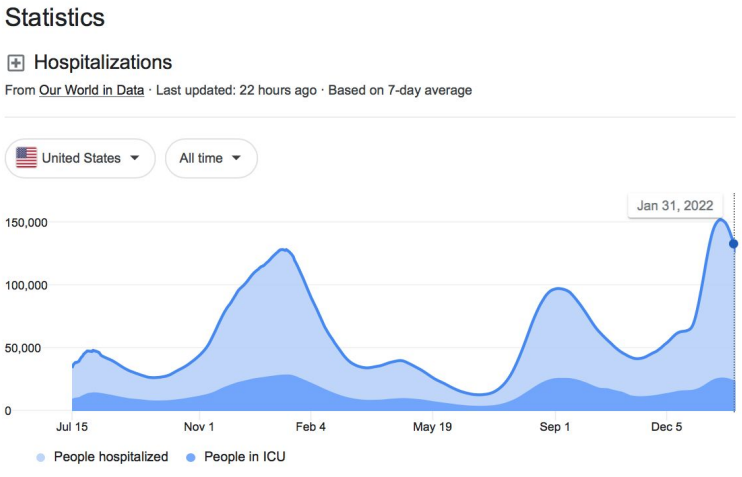
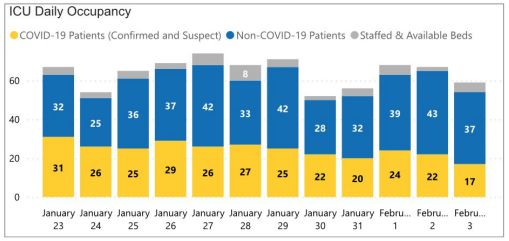
Classifying Pneumonia

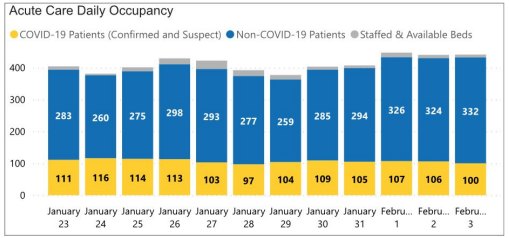
Springboard Capstone 3
Chloe Mai
February 2022

Problem

- Pneumonia is a possible complication of Covid-19
- During pandemic peaks, hospital utilization is maximized or overloaded, but medical staff has not grown
- Image classification can help filter X-rays to narrow the scope of radiology work





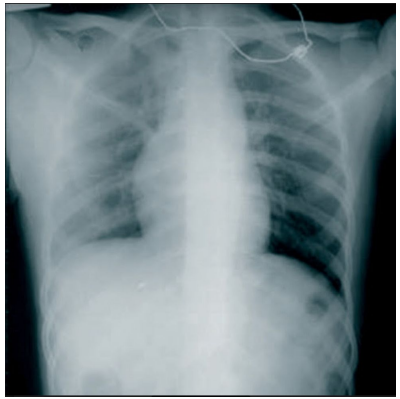


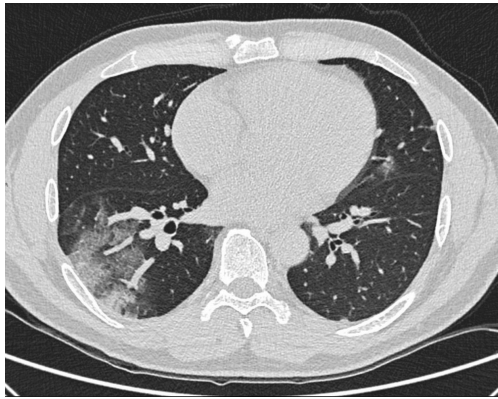
Source: ReddiNet daily poll of San Mateo County hospitals by SMC Health

Data

- 5935 images
  - 73% pneumonia vs 27% normal

	Normal	**Pneumonia**	**TOTAL**
Train	1196 *25%*	3551 *75%*	4747 *100%*
Test	234 *38%*	390 *62%*	624 *100%*





Data Wrangling

- This project focuses on Phase 1: pneumonia vs normal
- Data Wrangling for Phase 2 & 3: classifying causes of pneumonia
  - Phase 2: virus vs bacteria
  - Phase 3: if virus, covid vs not covid
- Actions Completed
  - Eliminated classifiers with < 10 data points
  - Relabeled NULL to Normal or Other
  - Created a validation dataset = 10% training dataset

Model

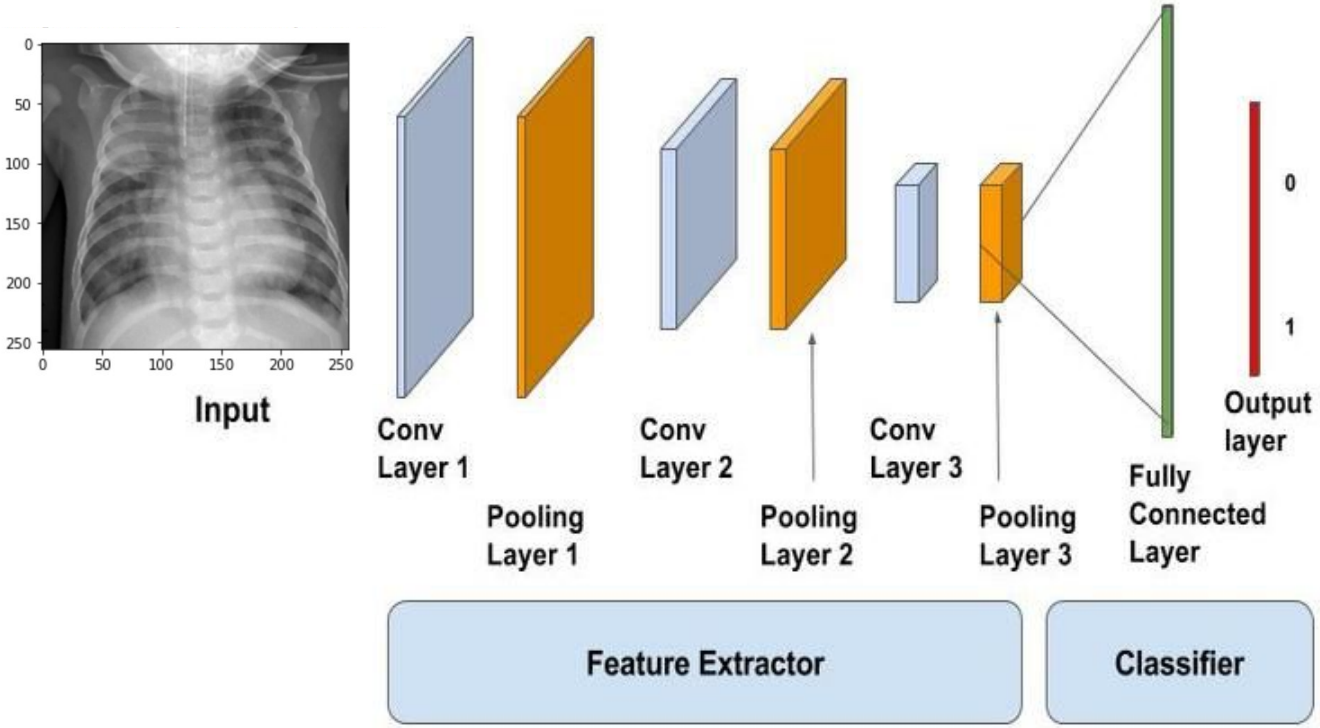


Diagram from Learn Open CV

Analysis

- Transfer Learning model produced the best recall
  - 93% of pneumonia patients were classified with pneumonia

Model #	Number of Convolution Layers	Number of Epoch	Filters	Training Accuracy	Test Recall
1	1	10	16	0.9031	81%
2	2	10	[32,16]	0.8599	74%
3	4	10	[32,64,32,16]	0.9248	66%
4	4	10	[128, 64, 32, 16]	0.9265	69%
5.1	3	20	[64, 32, 16]	0.9136	1st iteration - 79% 2nd iteration - 63.8%
5.2	3	+20	[64, 32, 16]	0.9328	66.9%
6	6	10	[32, 64, 32, 64, 32, 16]	0.9275	69%
7 - Transfer Learning	MobileNetV2	10	n/a	0.9130	93%

Recommendations

- Treat patients with pneumonia results
- Manually review X-ray results of patients who have pneumonia symptoms but model predicted Normal to verify False Negative

Future

- Improve model for higher recall and accuracy
- Create new models to predict virus and covid