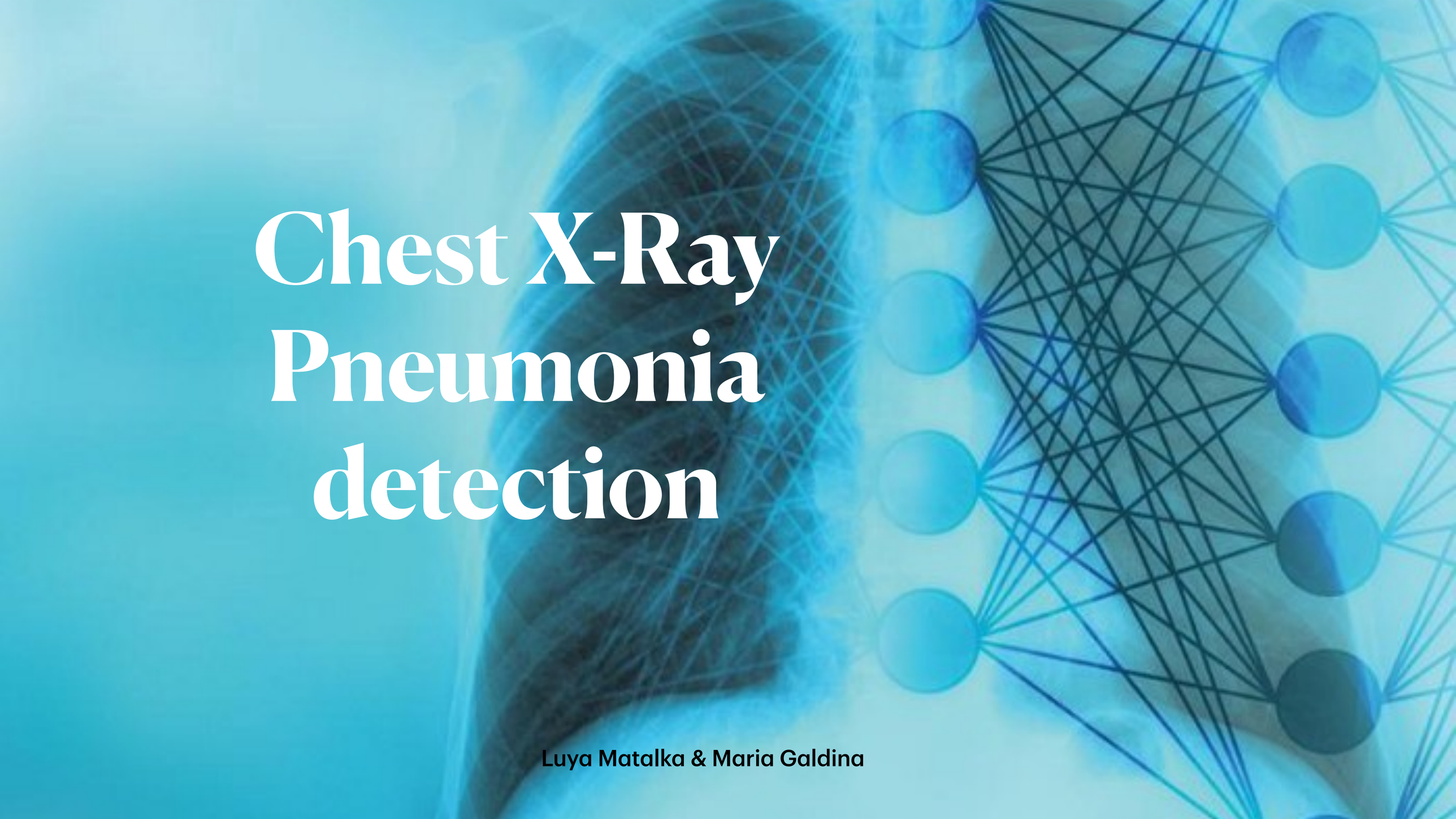
Chest X-Ray Pneumonia detection



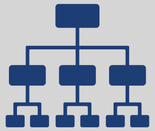
Luya Matalka & Maria Galdina

Contents

Introduction

Research Goal

Data Gathering


Data Pre-Processing

Results

Introduction

Radiology is a branch of medicine where the disease diagnosed by examine X-ray images

To reduce the human eye error and reduce time in diagnosing the disease medical centers implement computer aided systems for better

Machine learning techniques shown great results in image recognition

In this project we are trying to diagnose the Pneumonia from X-ray using MLT

Research Question

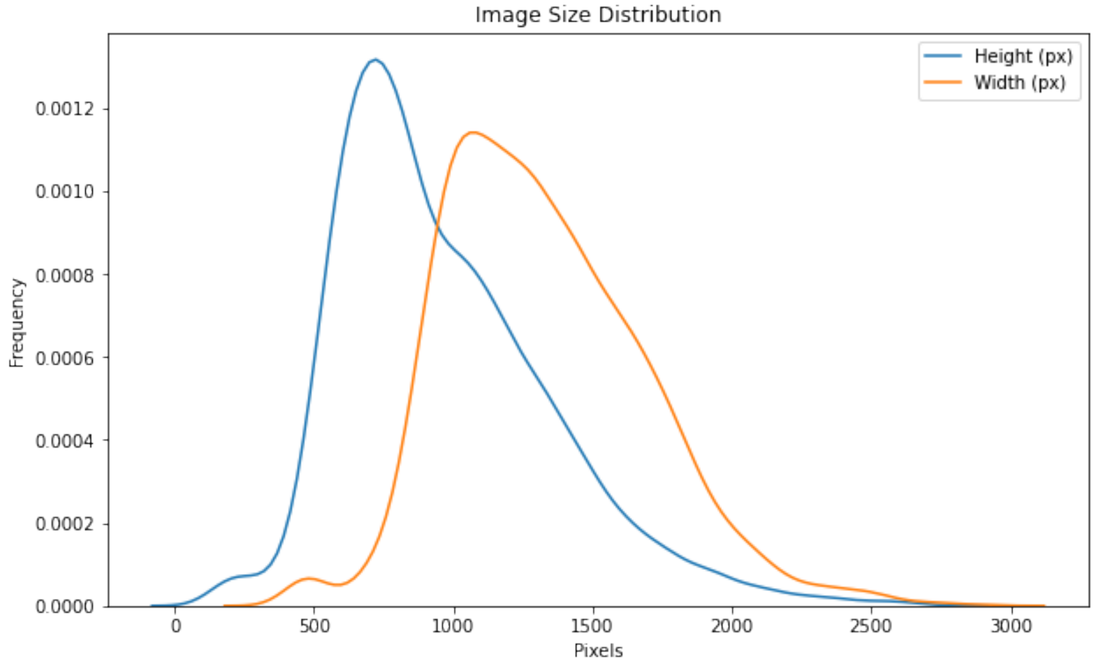
- ✓ Can we create a CNN model that can improve Pneumonia detection with recall score closer to as much as possible 100%?

Data Gathering

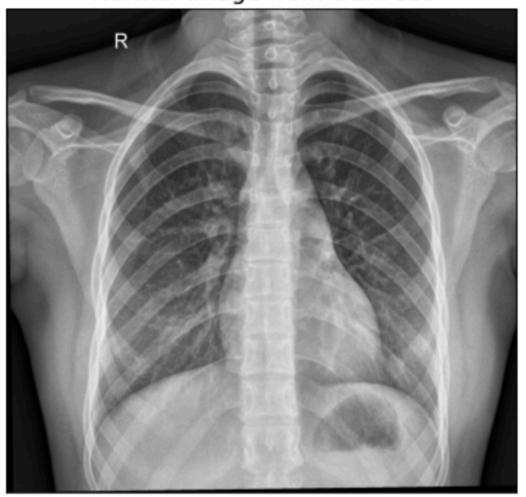
Data for this project is extracted from Kaggle prepared dataset

There are 5,856 X-Ray images in two categories (Pneumonia/Normal = 4273/1583)

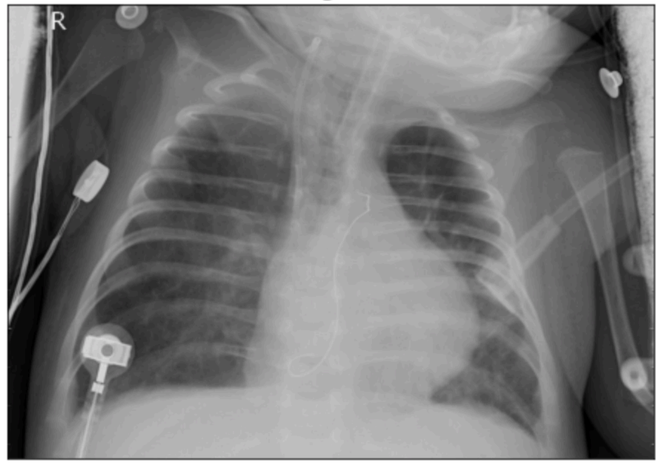
Images are in grayscale format, JPEG, different sizes



Normal Image from train set



Pneumonia Image from train set



Data Preprocessing

Rescaling: divided each image by 255

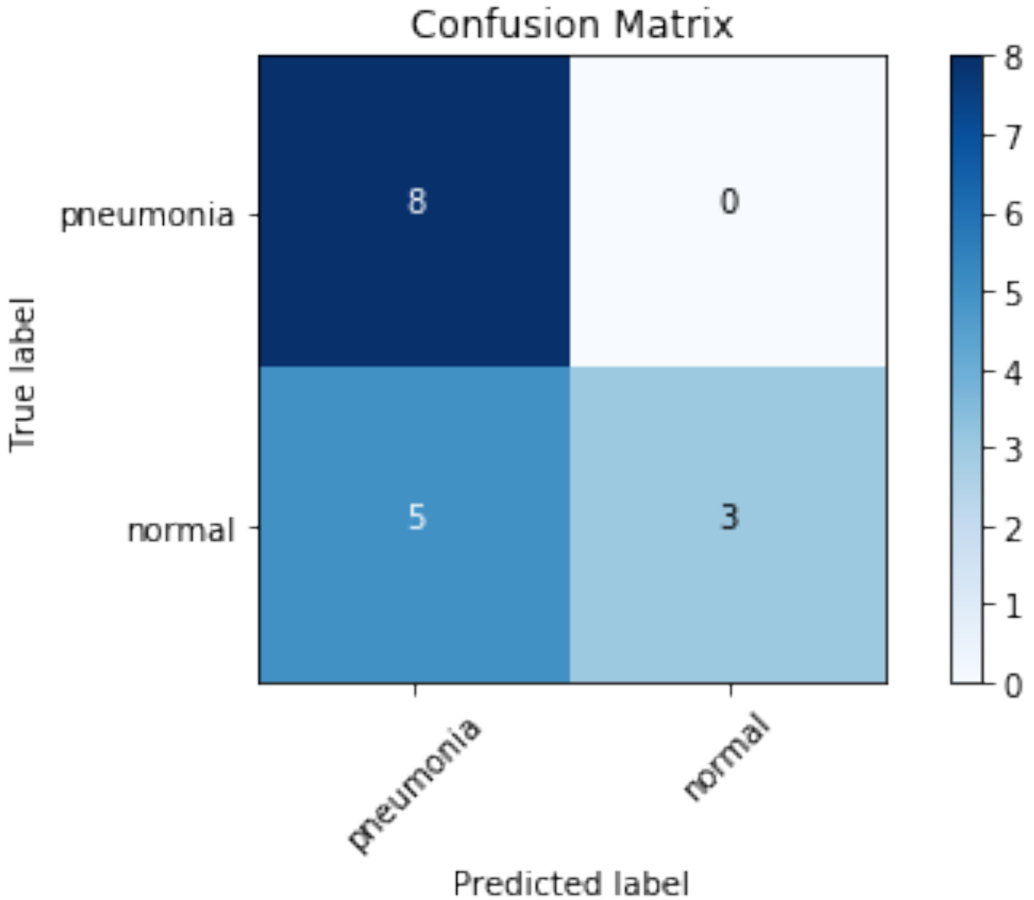
Resizing: all images have 224x224 sizes

Rotation: changed by 10 degrees

Also we used shearing transformation, randomly zooming inside picture and some horizontally flipping

Result

	precision	recall	f1-score	s
0	0.62	1.00	0.76	
1	1.00	0.38	0.55	
accuracy			0.69	
macro avg	0.81	0.69	0.65	
weighted avg	0.81	0.69	0.65	



PNEUMONIA
PNEUMONIA

NORMAL
NORMAL

PNEUMONIA
PNEUMONIA

PNEUMONIA
PNEUMONIA

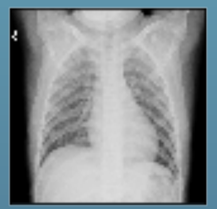
PNEUMONIA
PNEUMONIA

PNEUMONIA
PNEUMONIA

PNEUMONIA
PNEUMONIA

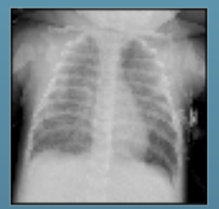
PNEUMONIA
NORMAL

PNEUMONIA
NORMAL



NORMAL
NORMAL

PNEUMONIA
PNEUMONIA



PNEUMONIA
NORMAL

NORMAL
NORMAL

PNEUMONIA
NORMAL

PNEUMONIA
PNEUMONIA

PNEUMONIA
NORMAL

Conclusion

- ✓ CNN model can be used under the supervision of a radiologist to enhance accuracy to improve hospitals' treatment

Recommendations

Integrate our model into medical apps

Use CNN-models can save radiologist time

Further Research

Identifying viral or bacterial pneumonia on the X-ray

Identifying other diseases using X-ray and MLT

Use GridSearch on GPU for discover better parameters for system

Try to improve data imbalance

THANK YOU!

Questions?

Appendix