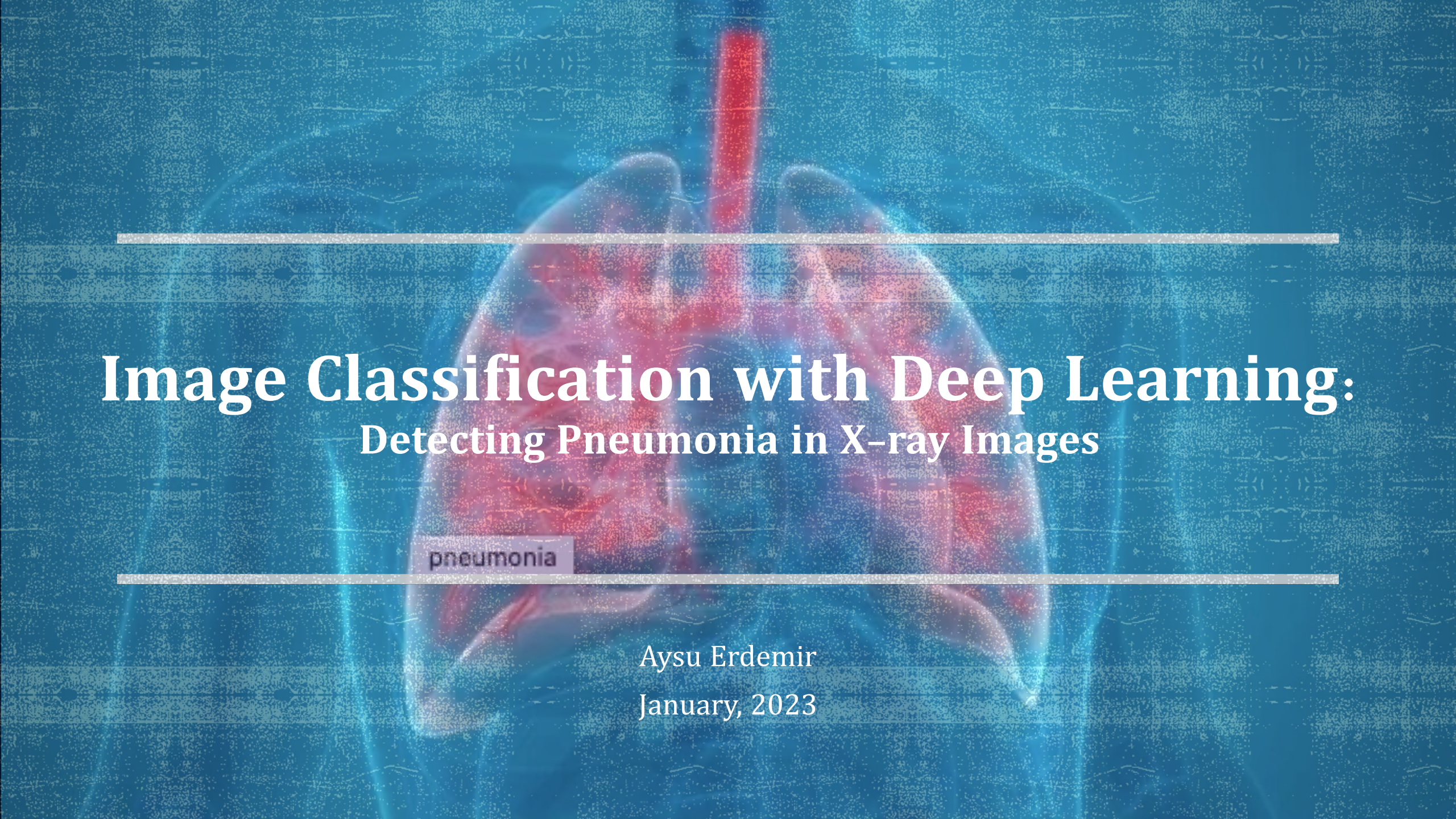


Image Classification with Deep Learning:

Detecting Pneumonia in X-ray Images

Aysu Erdemir

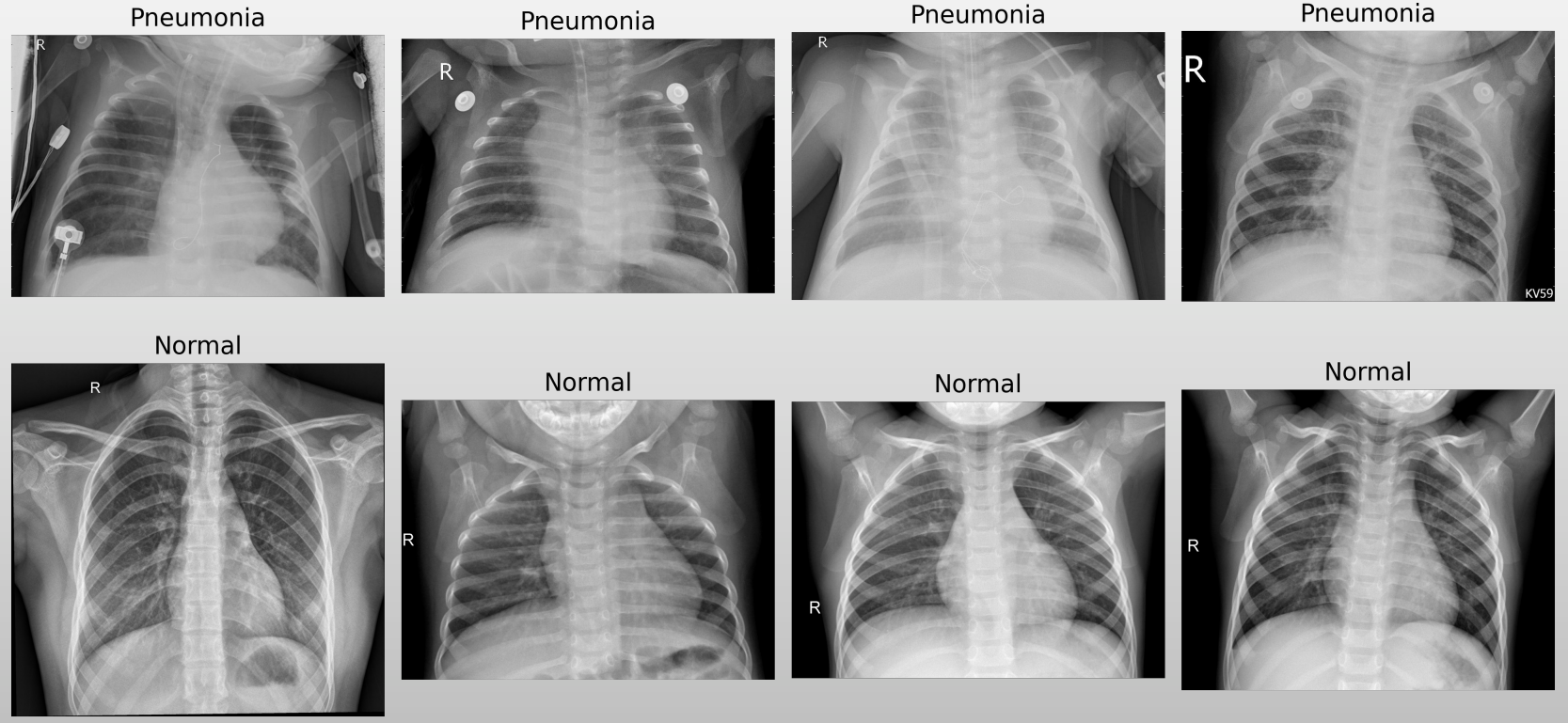
January, 2023

Stakeholder:	Be Well Healthcare Center, Radiology Department
Business Problem:	Looking for an efficient way to screen chest x-ray images for pneumonia detection in pediatric patients.
Project Value:	○ Decrease the workload of its radiologists. ○ Minimize diagnosis time and allows for faster treatment. ○ Lower the risk of human errors and improve patient safety.

Data：

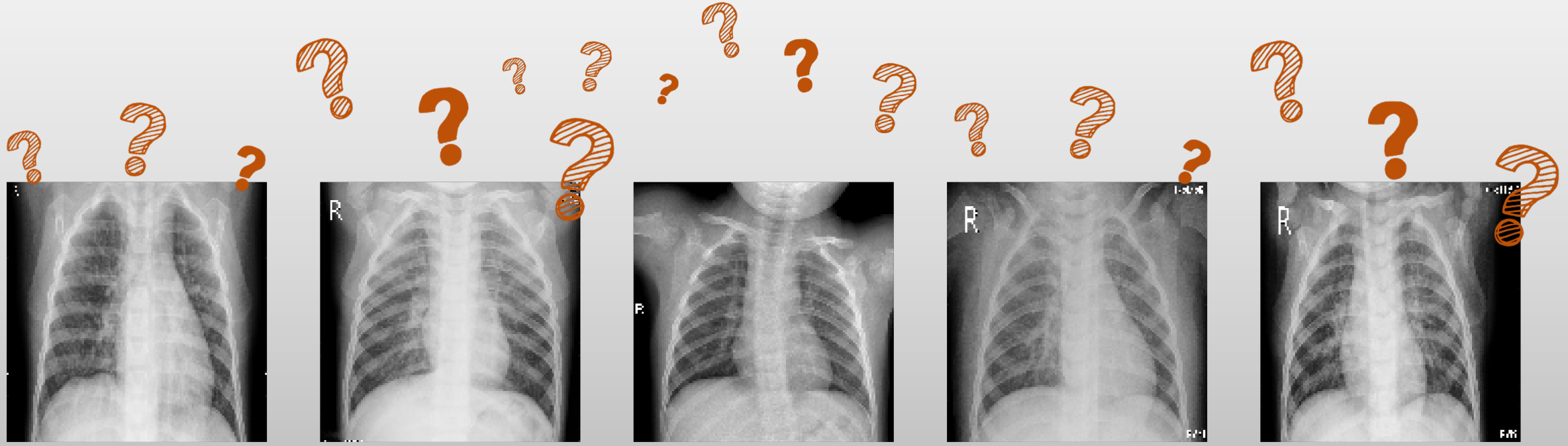
Kaggle Dataset of Chest X-Ray Images：

✓ Chest X-ray images from 5,863 pediatric patients separated by the image categories that they belong to（Pneumonia/Normal）.

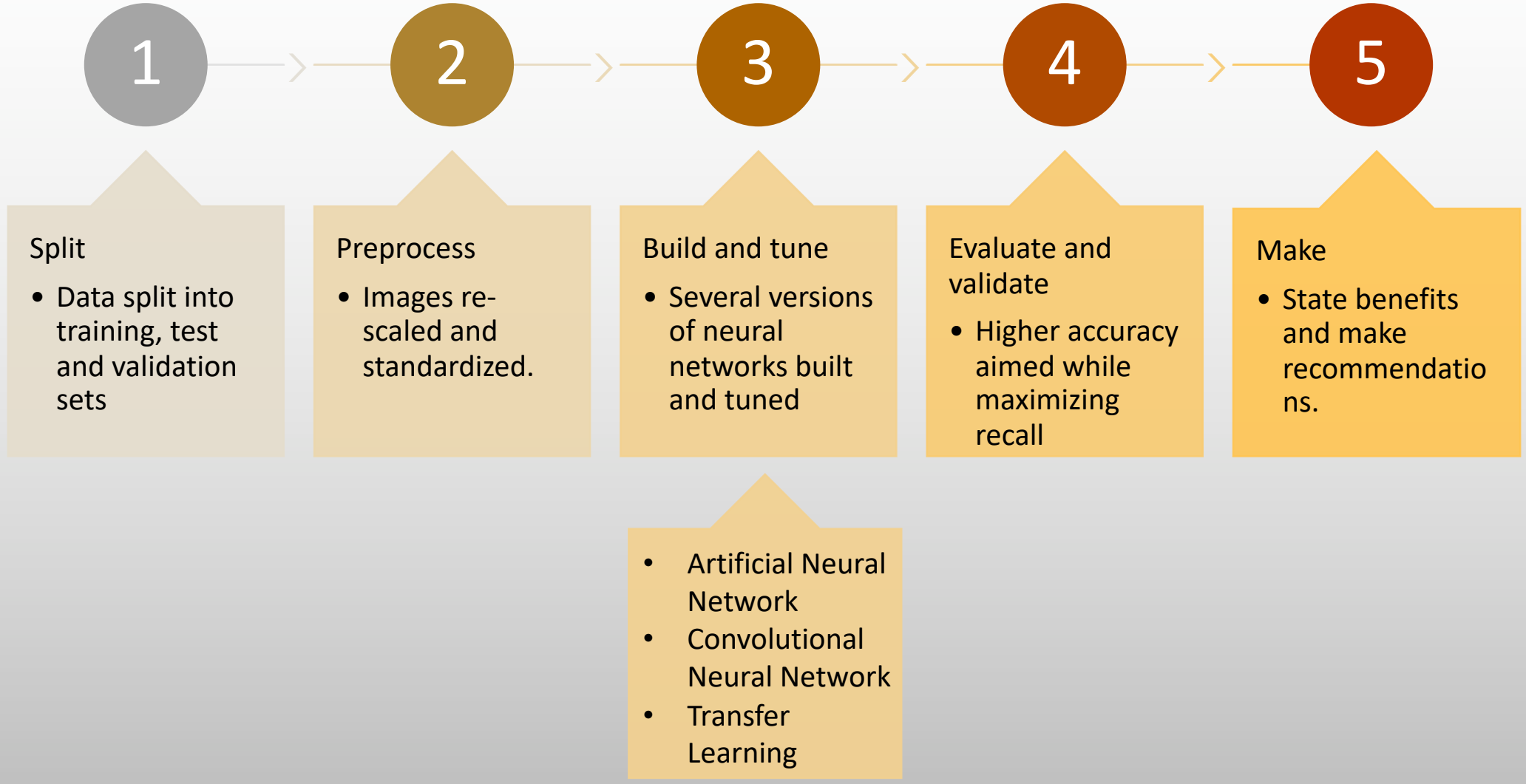


Goal:

- Generate a **Neural Network** that detects the presence of pneumonia in X-ray images.
- **Target Accuracy**: Accurately predict the status of the lungs (**Normal vs pneumonia**) (on the test set).
- **Maximize Recall**: Identify majority of the **True Positive** cases correctly so that we catch as many kids with pneumonia as possible.

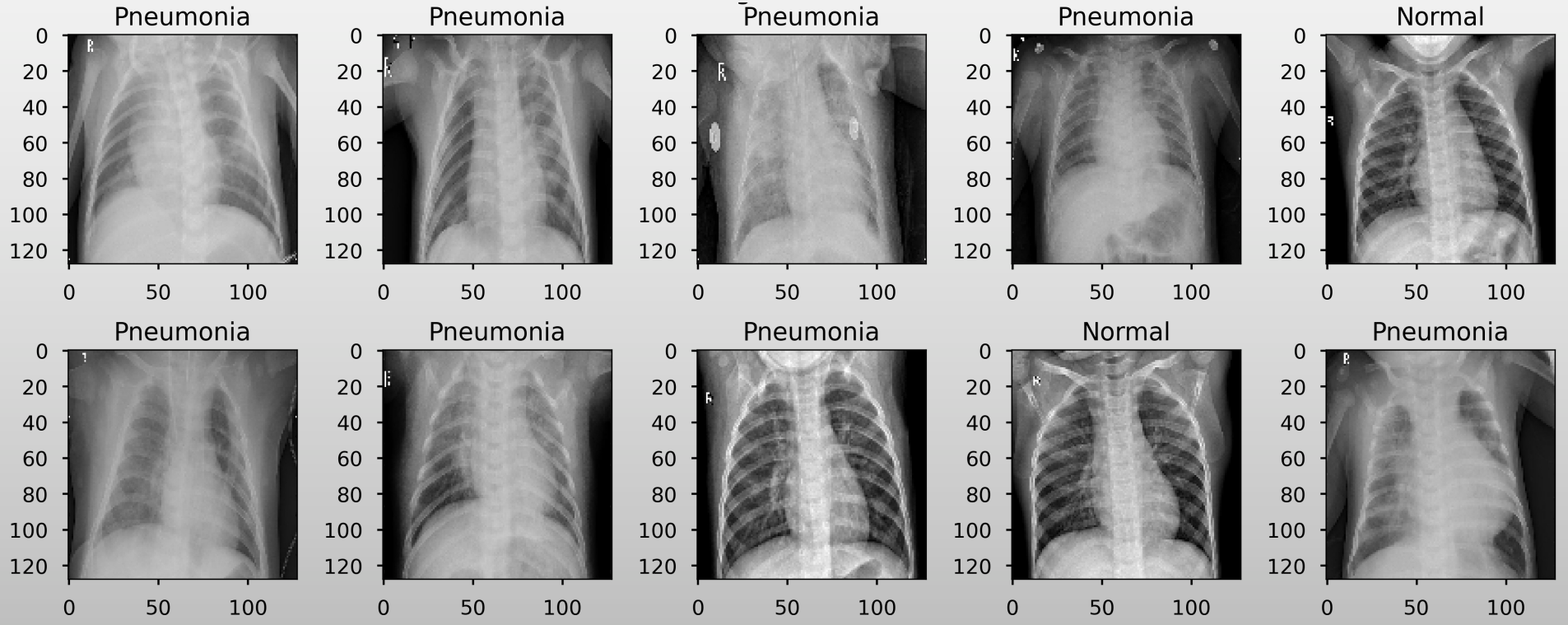


Methodology:



Split and Preprocess:

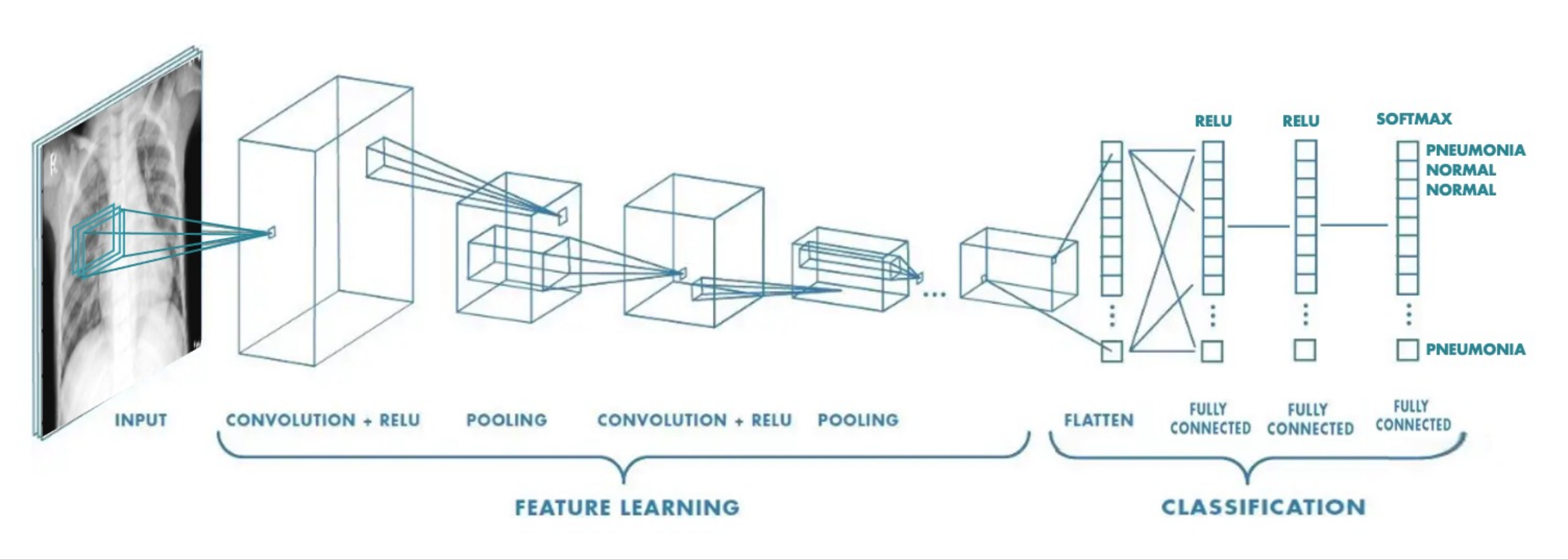
- Train (80%), test (10%), validation (10%) sets formed
- All images are downsized to a size of 128 x 128 pixels
- Pixels values (0–255) normalized to 0–1



Build and Tune:

Top Performing Model: Convolutional Neural Network

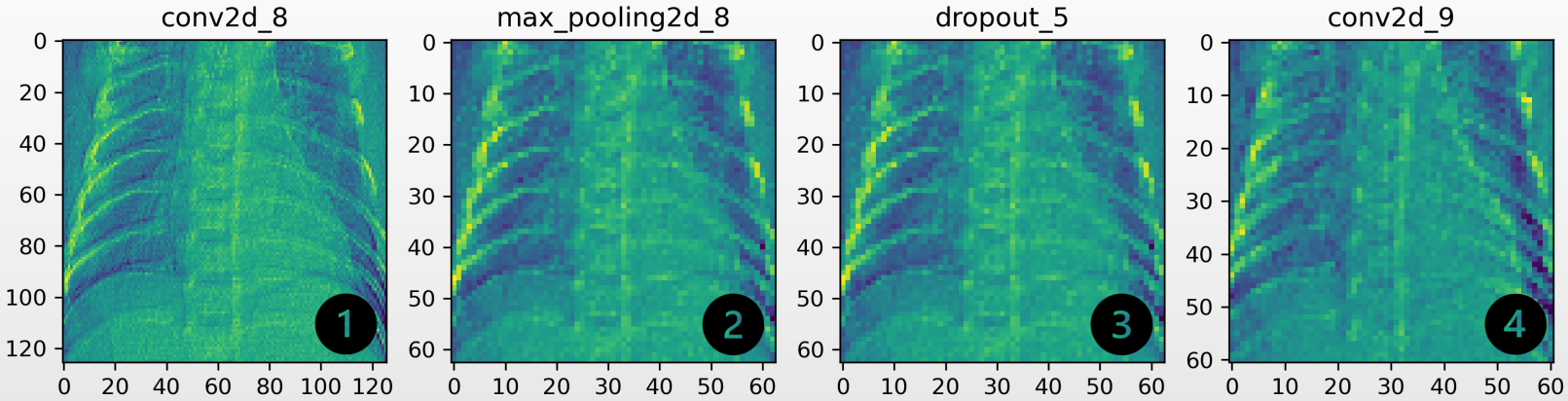
with dropout regularization and lower learning rate

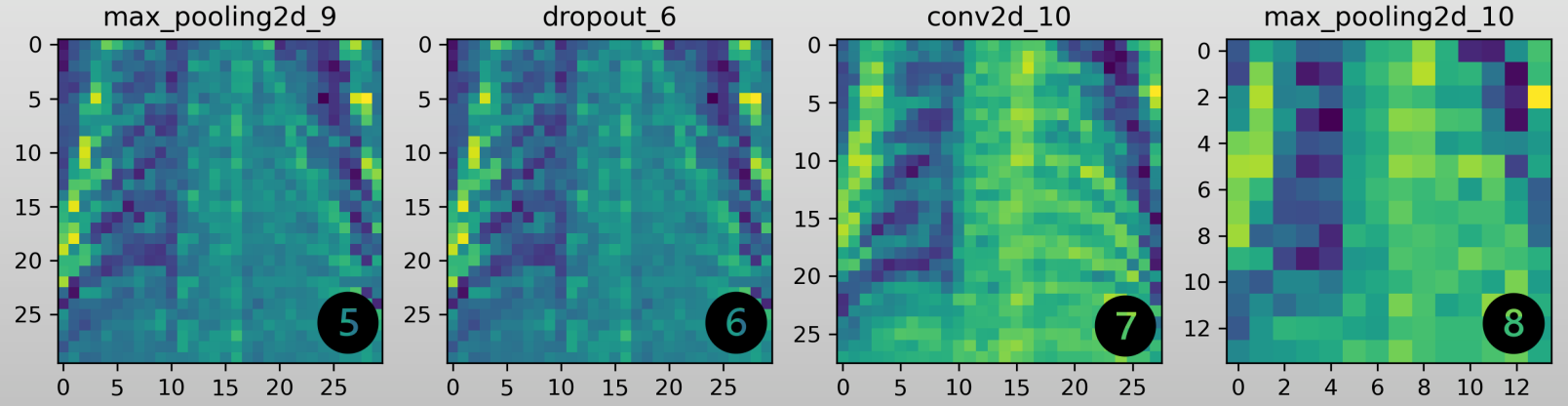


Visualize Features:

The visualization of a single feature map/channel for each of the convolutional layers as they go deeper into the network

Low-Level features such as edges, color





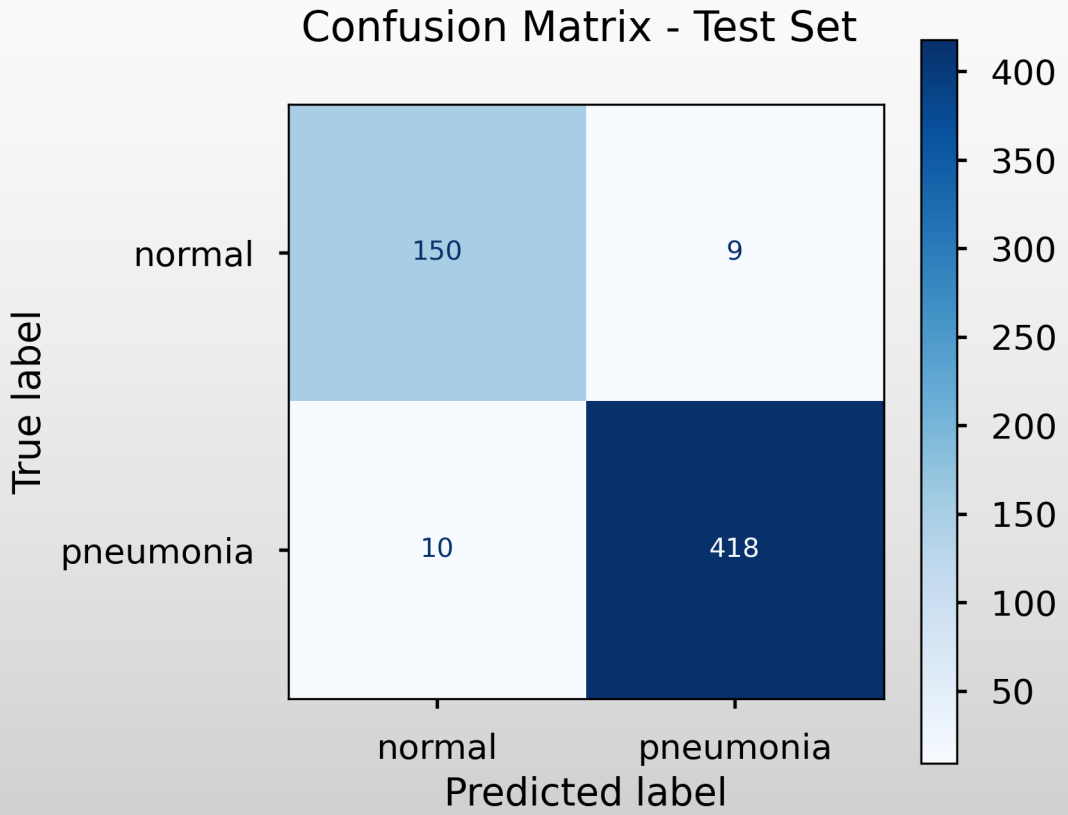
High-Level more abstract features

Validate and Evaluate:

Model Performance

	Pneumonia	Normal
Accuracy	**97%**	
Precision	98%	94%
Recall *(True positive rate)*	**98%**	94%
f1-score	**98%**	**94%**

"**Very Good**" performance in all metrics.



Recommendations:

Use of CNN to classify x-ray images

- Stream-line the diagnosing process → quicker return time and greater patient satisfaction.
- Begin the treatment right away for patients classified as high-risk.
- Doctors can allocate more time to go over the images that fall into the grey zone more rigorously, and for more demanding and complex procedures in general.

Limitations and Improvements

- We can use data augmentation methods to increase the size of the training set.
- We can crop the images to exclude the electrodes and letter R from the display which might be negatively affecting the image processing algorithm.
- We could address the class imbalance issue using oversampling techniques.

Thank You

Email: erdemiraysu@gmail.com

GitHub: @erdemiraysu

LinkedIn: linkedin.com/in/aysuerdemir/