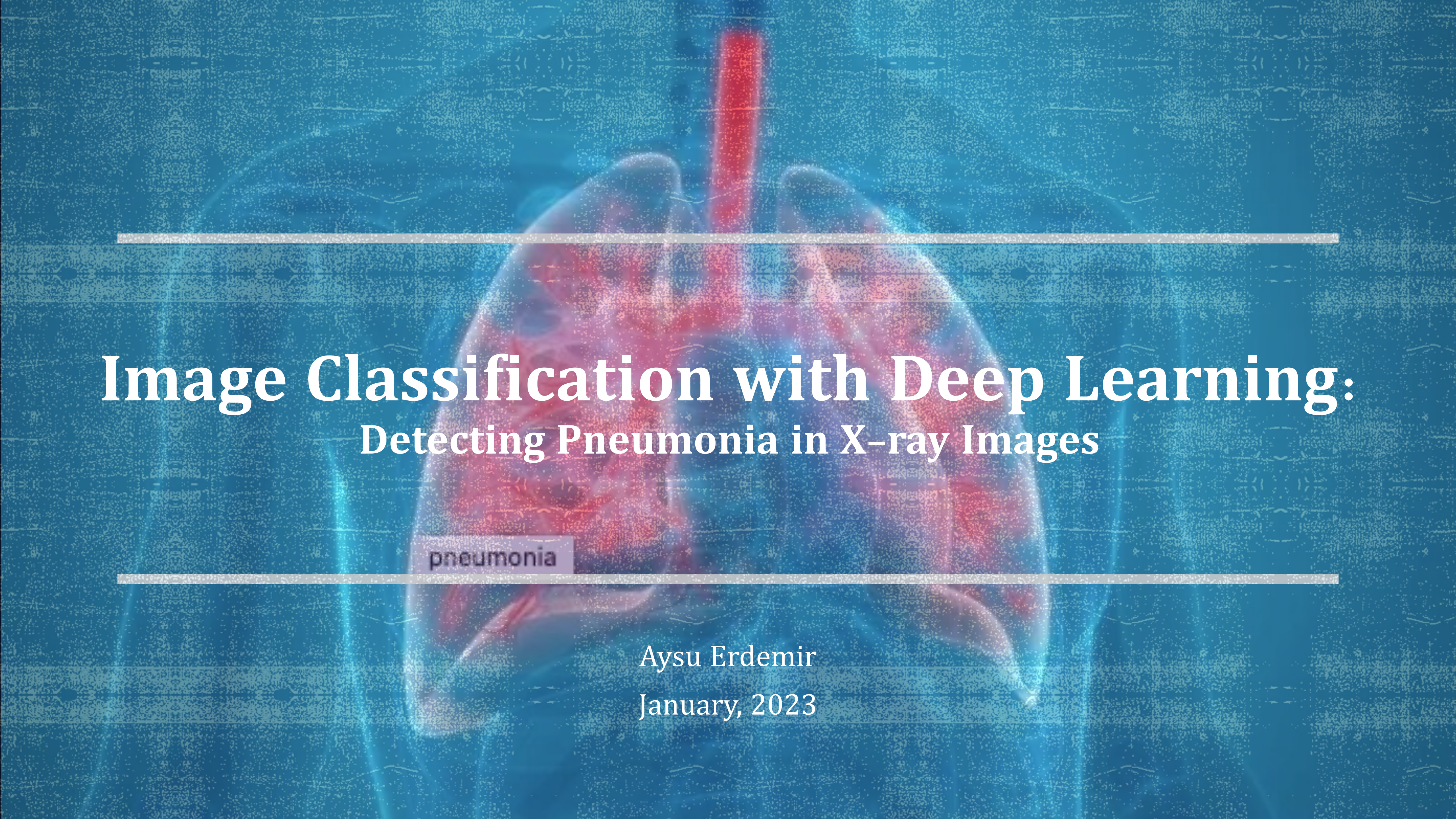


Image Classification with Deep Learning:
Detecting Pneumonia in X-ray Images

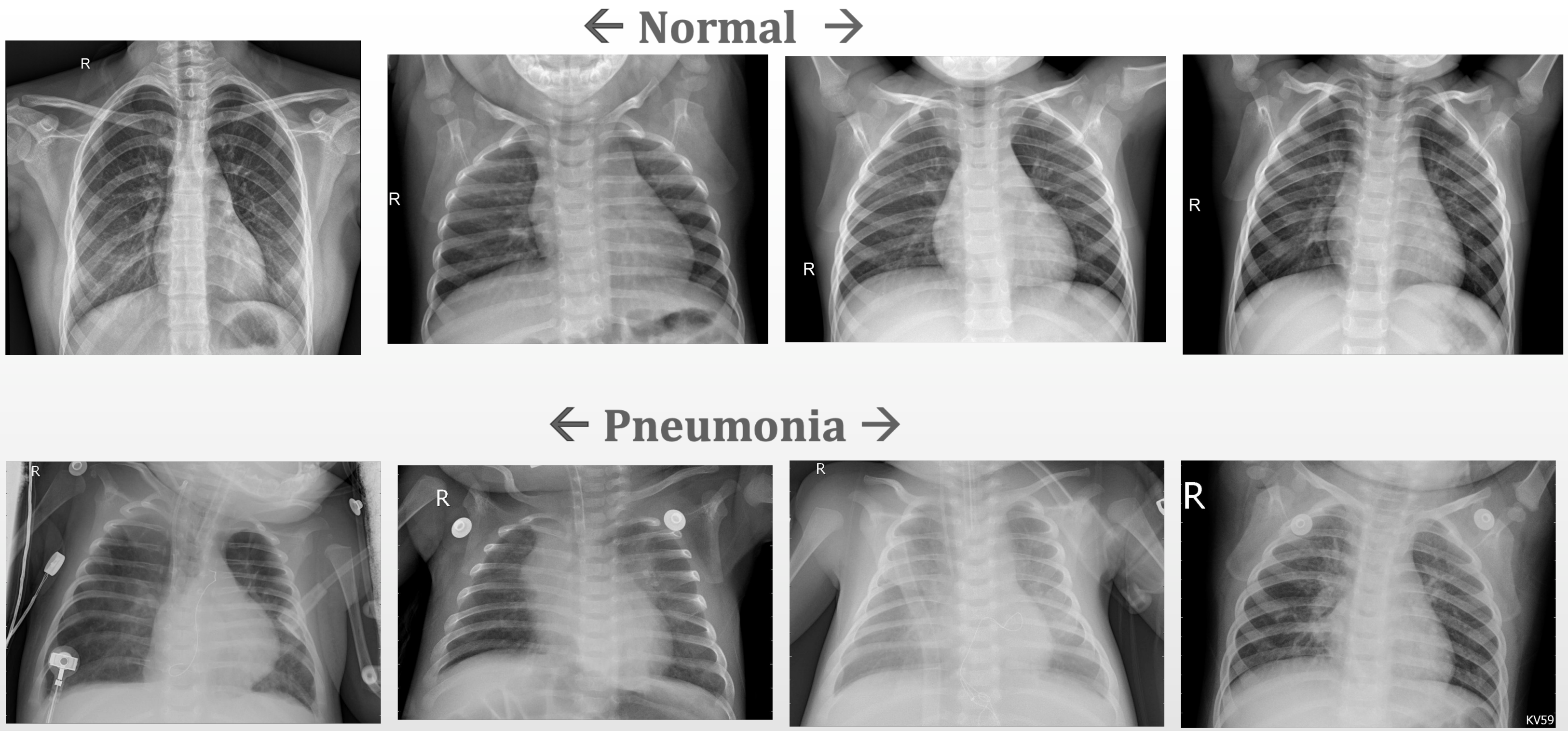
Aysu Erdemir

January, 2023

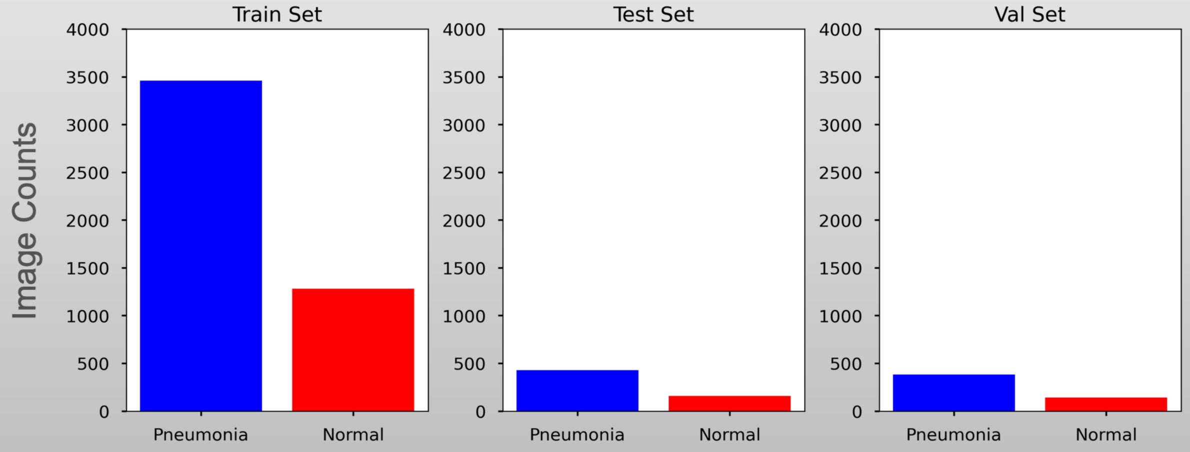
Stakeholder:	Be Well Healthcare Center, Radiology Department
Business Problem:	Looking for an efficient way to screen chest x–ray images for pneumonia detection in pediatric patients.
Project Value:	o Decrease the workload of its radiologists. o Minimize diagnosis time and allows for faster treatment.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

Data:

Kaggle Dataset of **Chest X-ray images** from **5,863 pediatric** patients.

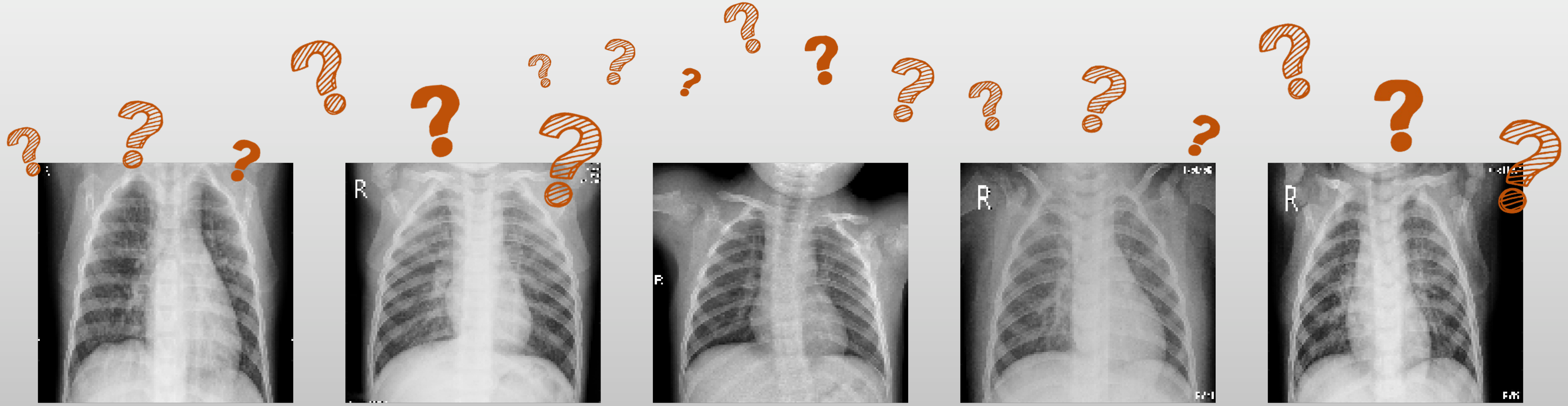


Data is separated by: **train** (80%), **test** (10%), **validation** (10%) sets and the categories they belong to (**Pneumonia/Normal**).



Goal:

- Generate a **Neural Network Model** that classifies the X-ray images by detecting pneumonia.
- **Target Accuracy**: Predict the status of the lungs (**Normal vs pneumonia**) as accurately as possible.
- **Maximize Recall**: Identify majority of the **True Positive** cases so that we catch as many kids with pneumonia as possible.



Methodology:

1 — Split

- Data split into training, test and validation sets

2 — Preprocess

- Images re-scaled and standardized.

3 — Build and tune

- Several versions of neural networks built and tuned
    - Artificial Neural Network
    - Convolutional Neural Network
    - Transfer Learning

4 — Evaluate and validate

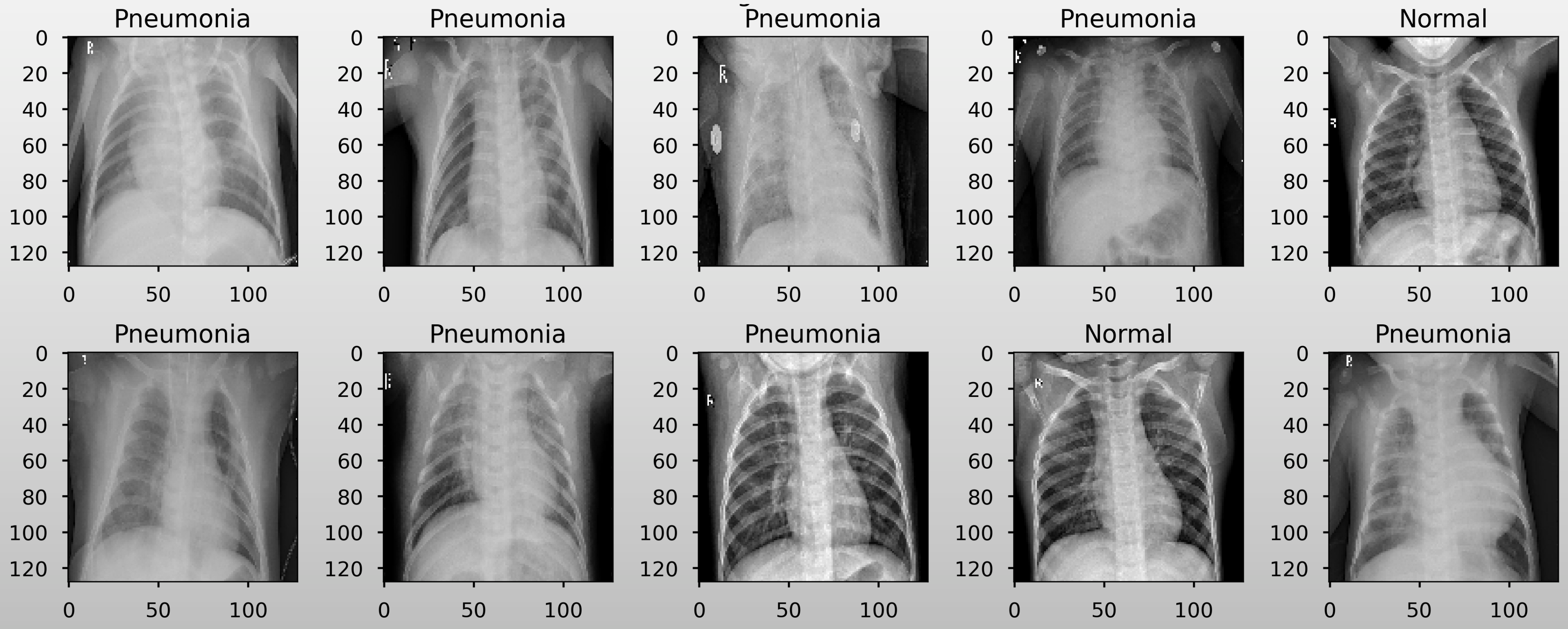
- Higher accuracy aimed while maximizing recall

5 — Conclusions

- State benefits and make recommendatio ns.

Preprocess:

- All images are downsized to a size of 128 x 128 pixels
- Pixels values（0–255）normalized to 0–1

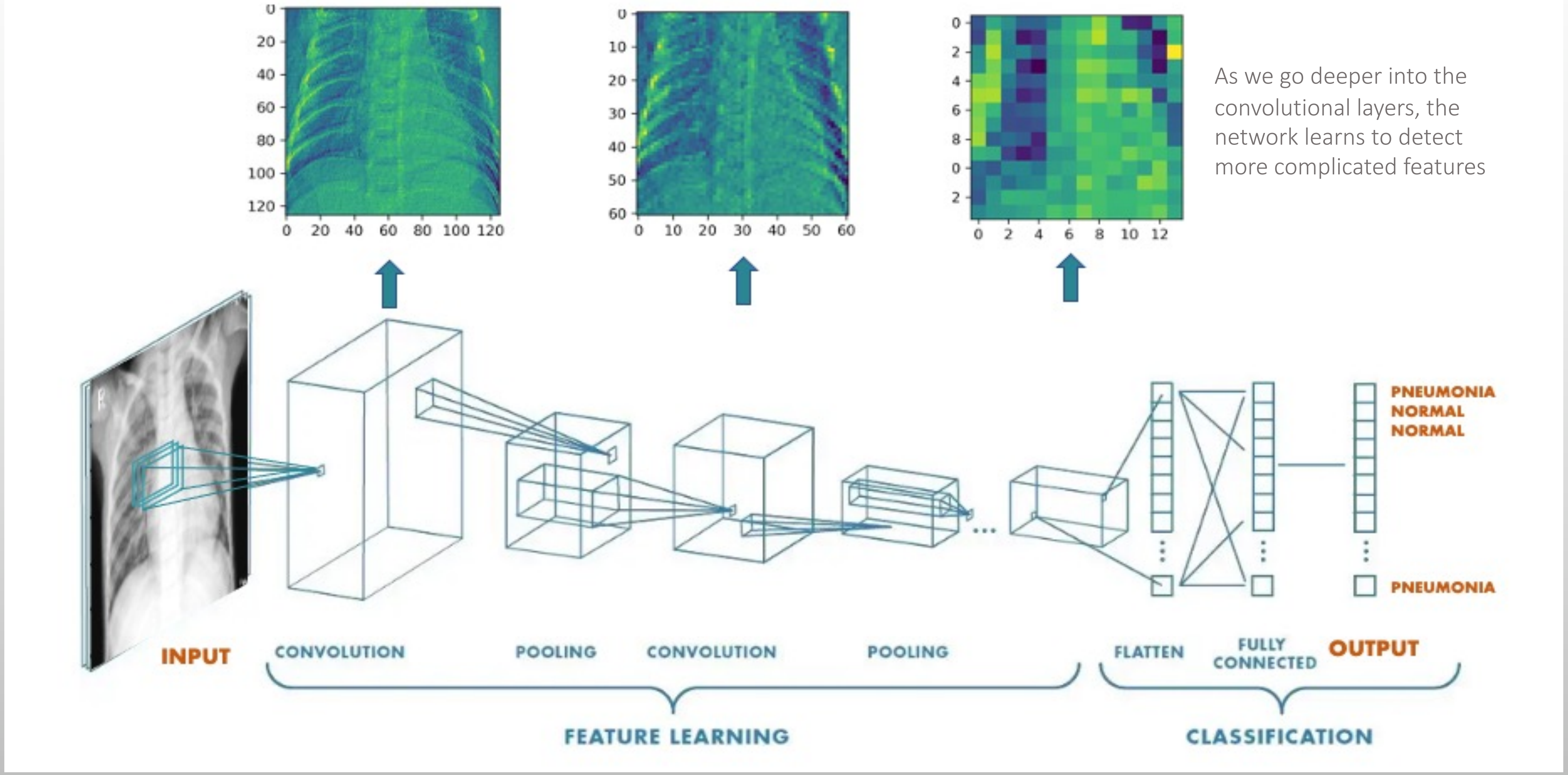


Build and Tune:

Top Performing Model: Convolutional Neural Network

Low-Level features such as edges, color

High-Level, more abstract features



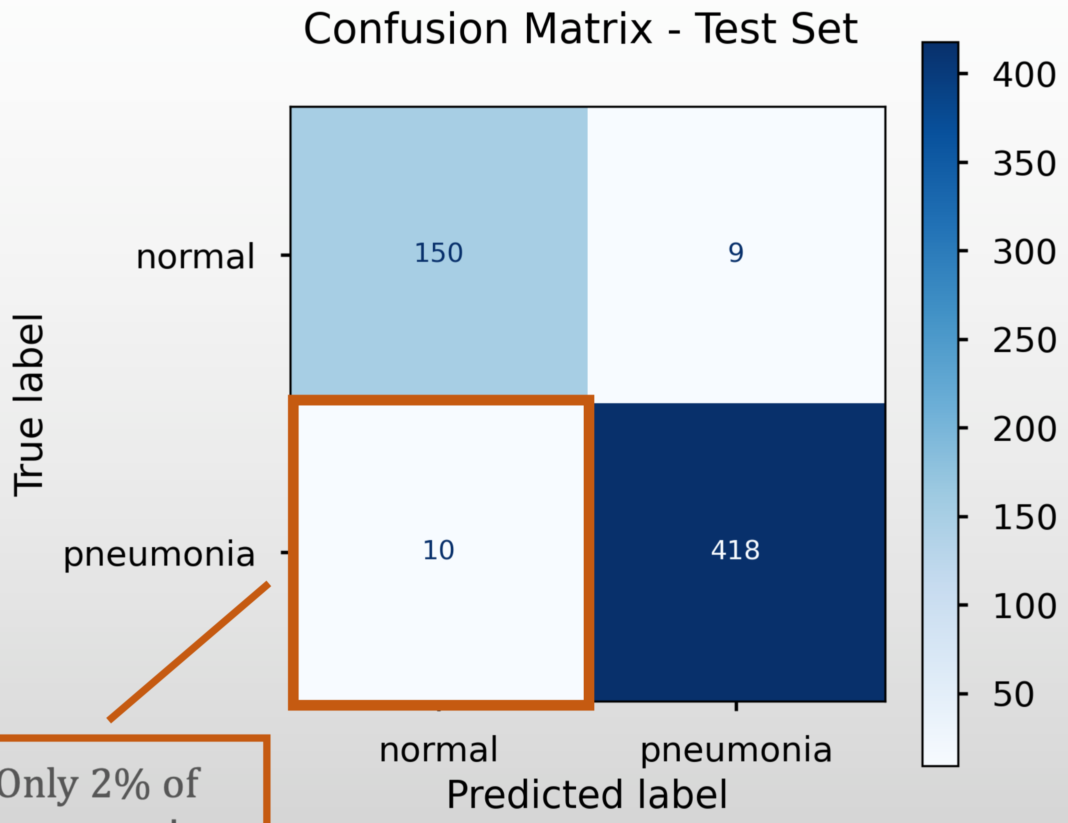
As we go deeper into the convolutional layers, the network learns to detect more complicated features

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.



Only 2% of pneumonia cases missed

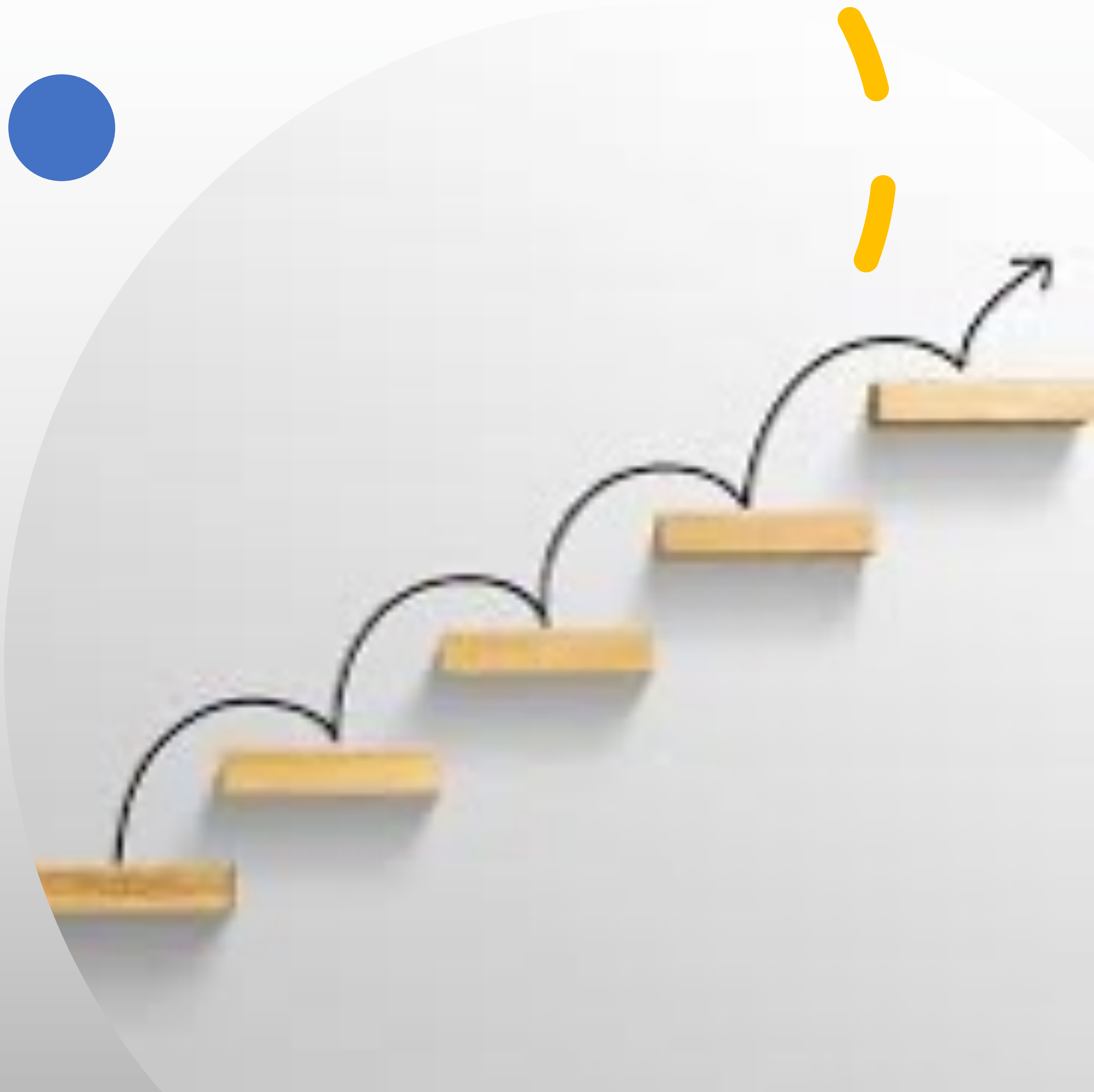
Recommendations:

Use of CNN to classify x-ray images

- **Decreased work-load** by stream-lining the diagnosing process → *quicker return time and greater patient satisfaction*.
    - **Faster initiation of treatment** for patients classified as high-risk.
    - Doctors / radiologists can allocate **more time** to go over the images in the grey zone more rigorously or **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.



Email: erdemiraysu@gmail.com

GitHub: @erdemiraysu

LinkedIn: linkedin.com/in/aysuerdemir/

Thank You