MedScan AI : Transforming Disease Diagnosis

Group 4: BAAZIGAR

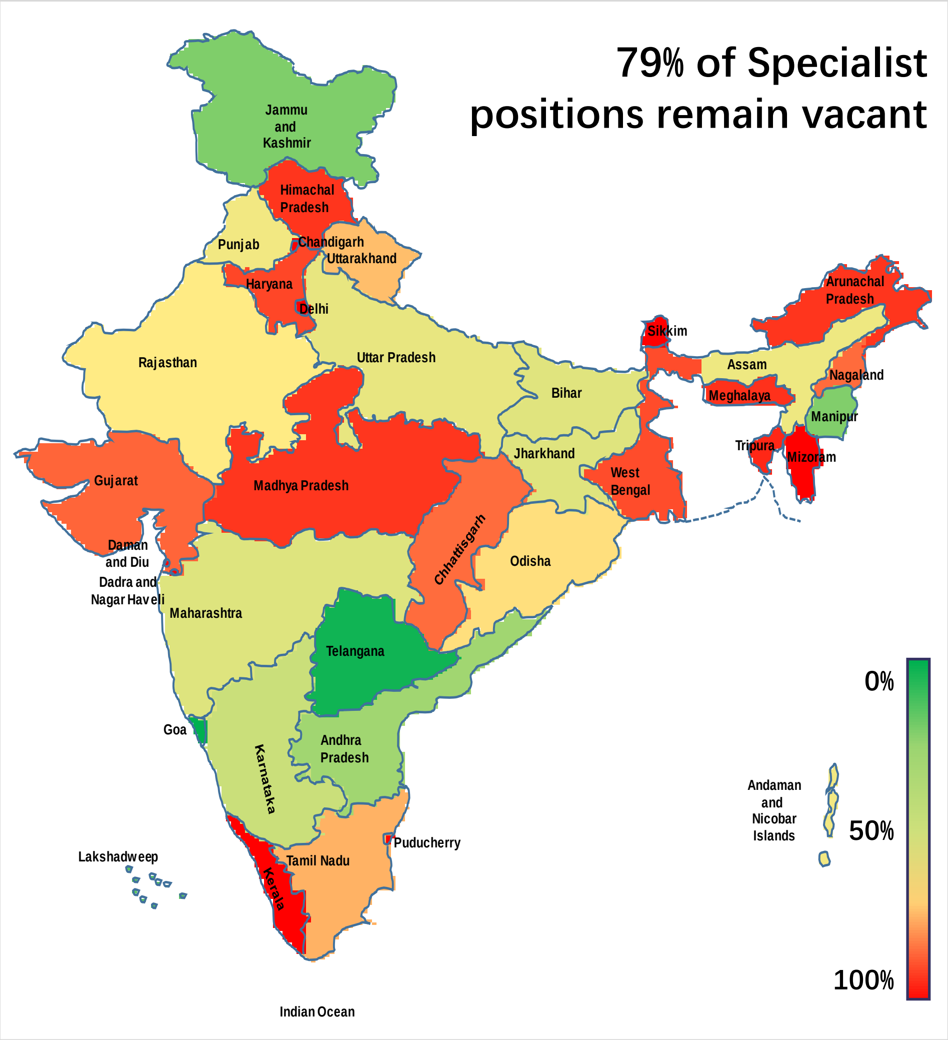
Abhinay Singh, Anubhav Bhatnagar, Pratik Parag, Vivek Mishra

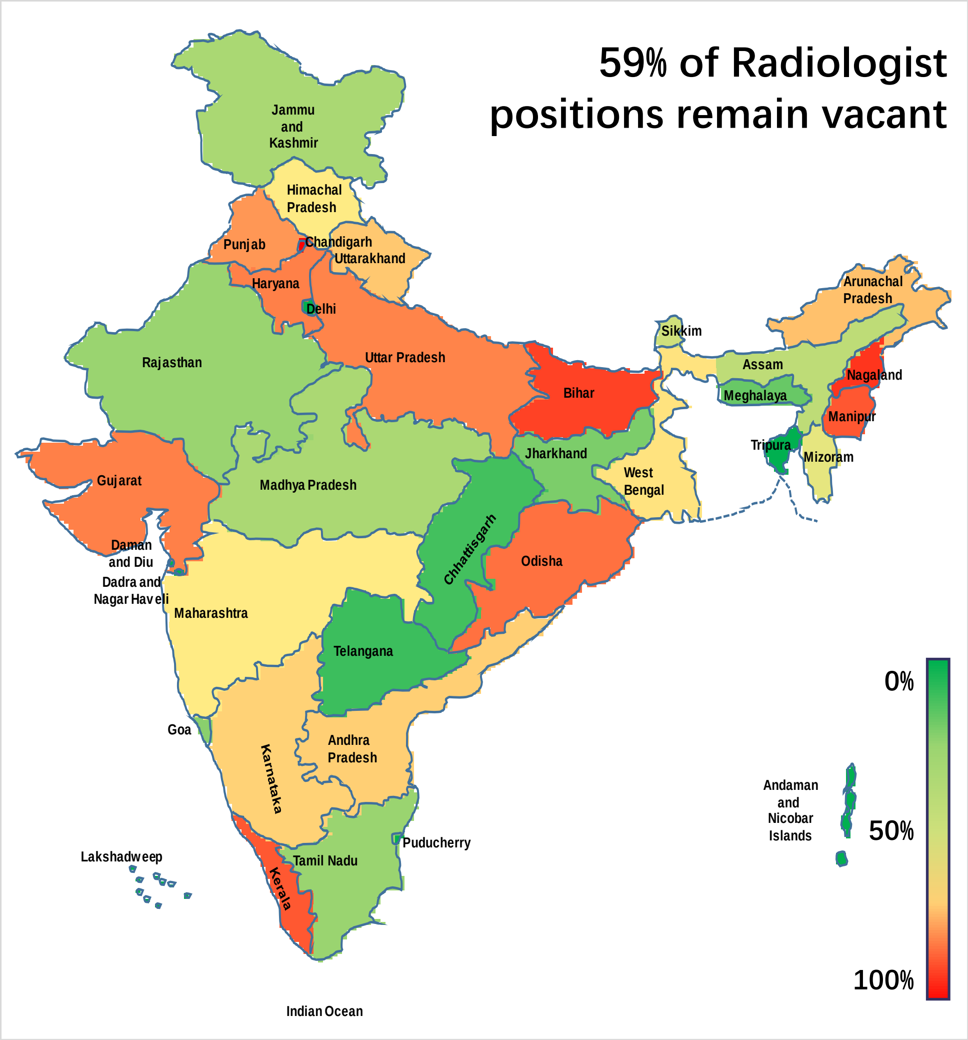
CONTENTS

01

Introduction

Skilled medical professionals in short supply





Source: MoHFW | Rural Health Statistics 2021-22

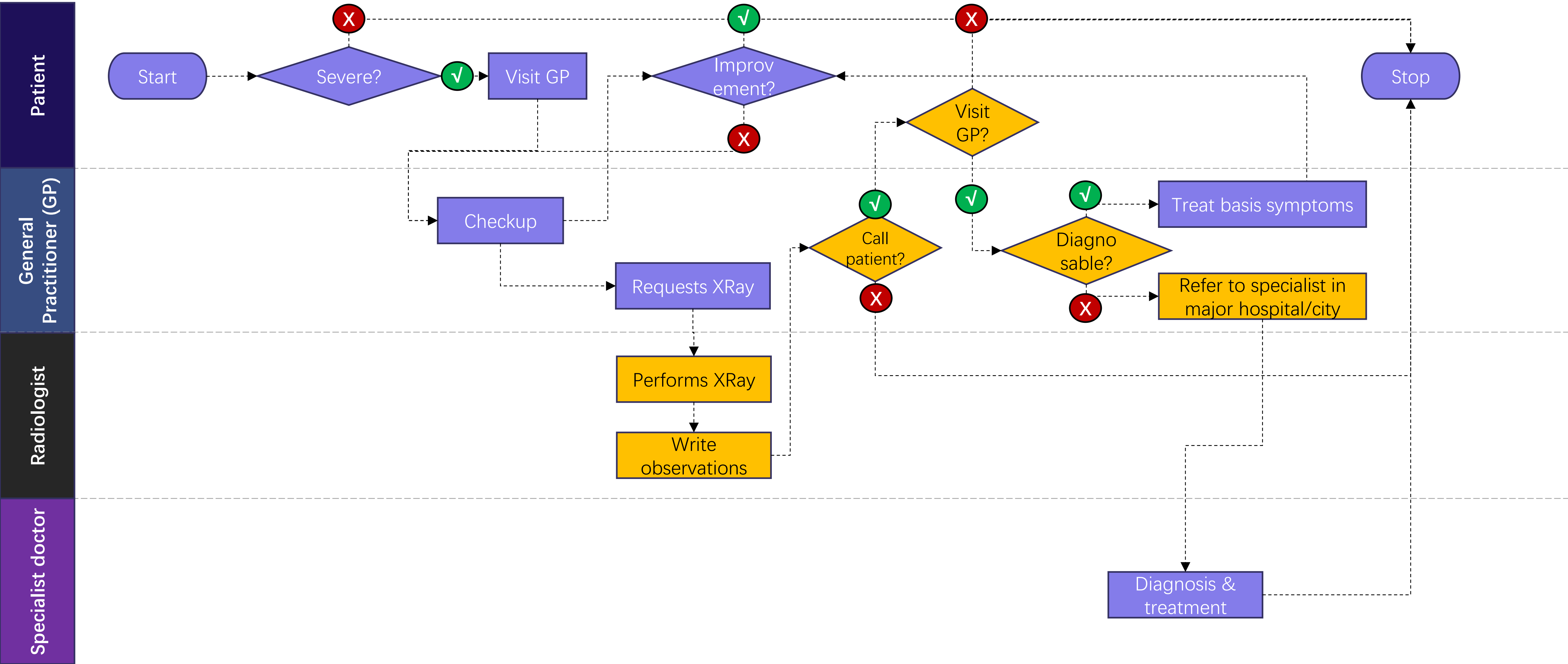
India leads the world in respiratory diseases

According to the latest Global Burden of Diseases Report, 2017, India has one of the highest burden of chronic respiratory diseases. India contributes to 15.69% of Global Chronic Respiratory Diseases but 30.28% of all global deaths due to Chronic Respiratory Disease occur in India. India has the highest number of COPD (Chronic Obstructive Pulmonary Disease) cases in the world, a whopping 55.23 million! The second-largest number of global deaths due to COPD, almost 0.85 million, occur in India. India also leads the world in deaths due to Asthma with 43% of global asthma deaths occurring in our country. Yet awareness in the healthcare community and the public remains low about these two chronic lung conditions which cause a huge health and economic burden on our health and health care system.

Source: IPCRG

How does this affect patients?

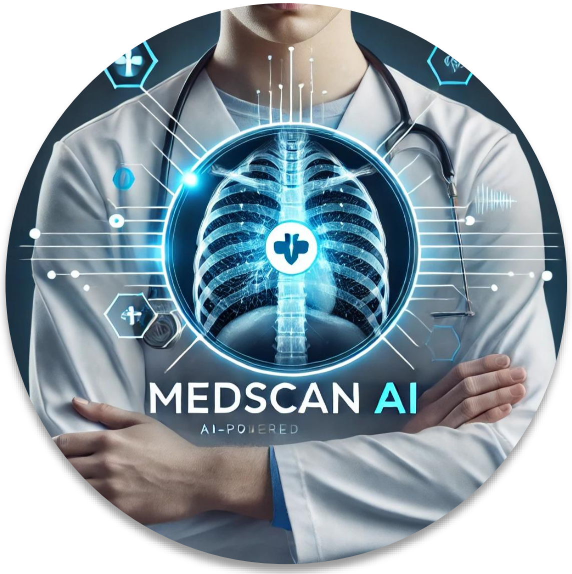
Patient journey | Current state



How thorough were the Radiologist's notes?

Is the GP in a position to diagnose/treat a patient?

Can we improve the chances of a patient going for a follow up?



Can we improve the chances of a patient going for a follow up?

The Opportunity

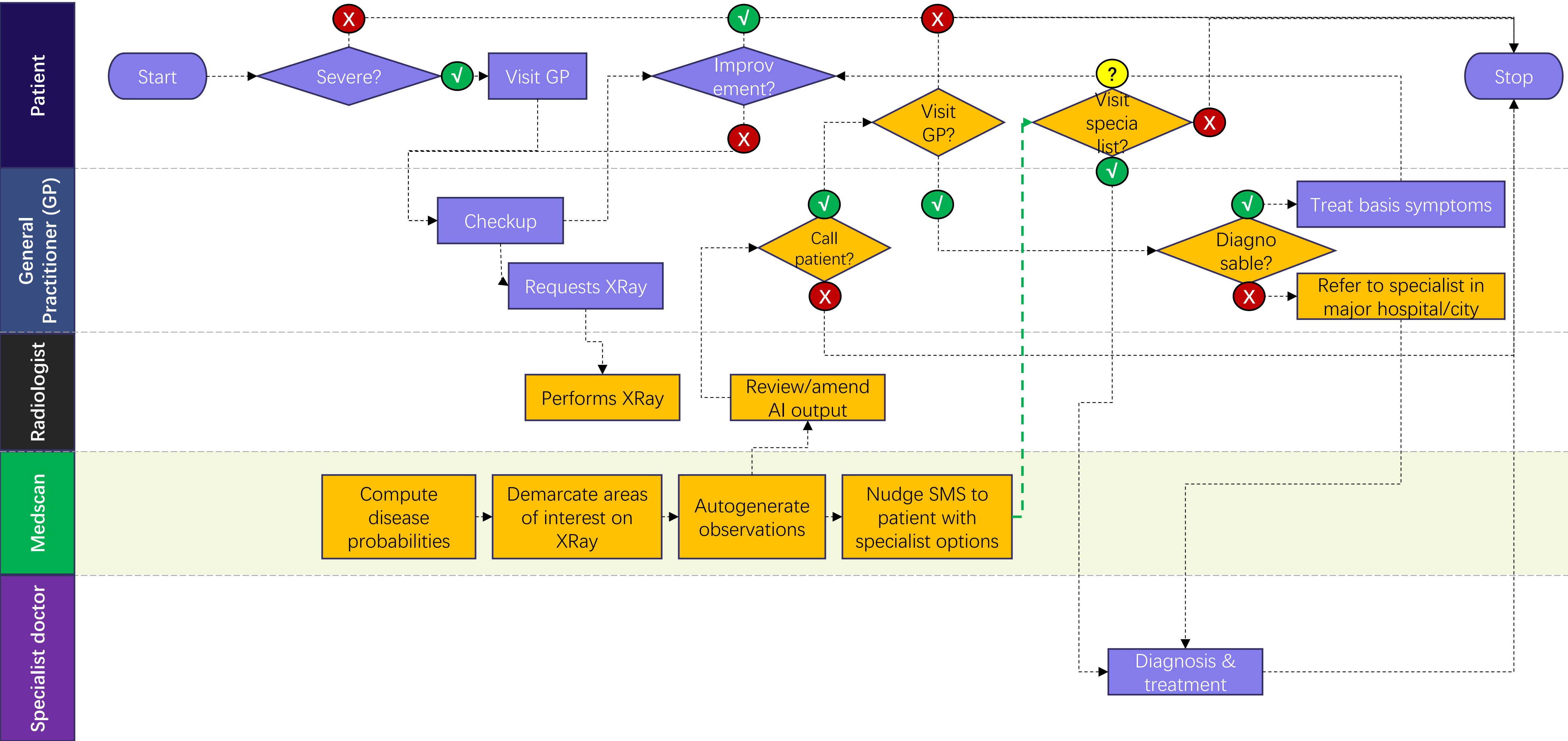
22%	Follow up rate amongst lung cancer patients in the US *(127,194 patients – McNulty, 2023)*
39%	Follow up rate amongst COPD patients in Sweden over 15 months *(19,857 patients - Sandelowsky et al., 2022)*
49%	Follow up rate amongst Pulmonary Nodule patients in the US over 4 years *(32,965 patients – Abrahams et al., 2024)*

Effectiveness of nudge notifications

Increase in follow ups amongst households with children needing immunization in Haryana *(14,760 patients – Banerjee et al., 2021)*	44% ▲
Increase in follow ups amongst US based cardiology patients *(39,209 patients – Telukuntla et al., 2021)*	19% ▲
Increase in follow ups for radiology session appointments amongst Saudi Arabian patients *(493 patients – Alturbag, 2024)*	5% ▲

Sources: *McNulty, 2023, Sandelowsky et al., 2022, Abrahams et al., 2024, Banerjee et al., 2021, Telukuntla et al., 2021, Alturbag, 2024*

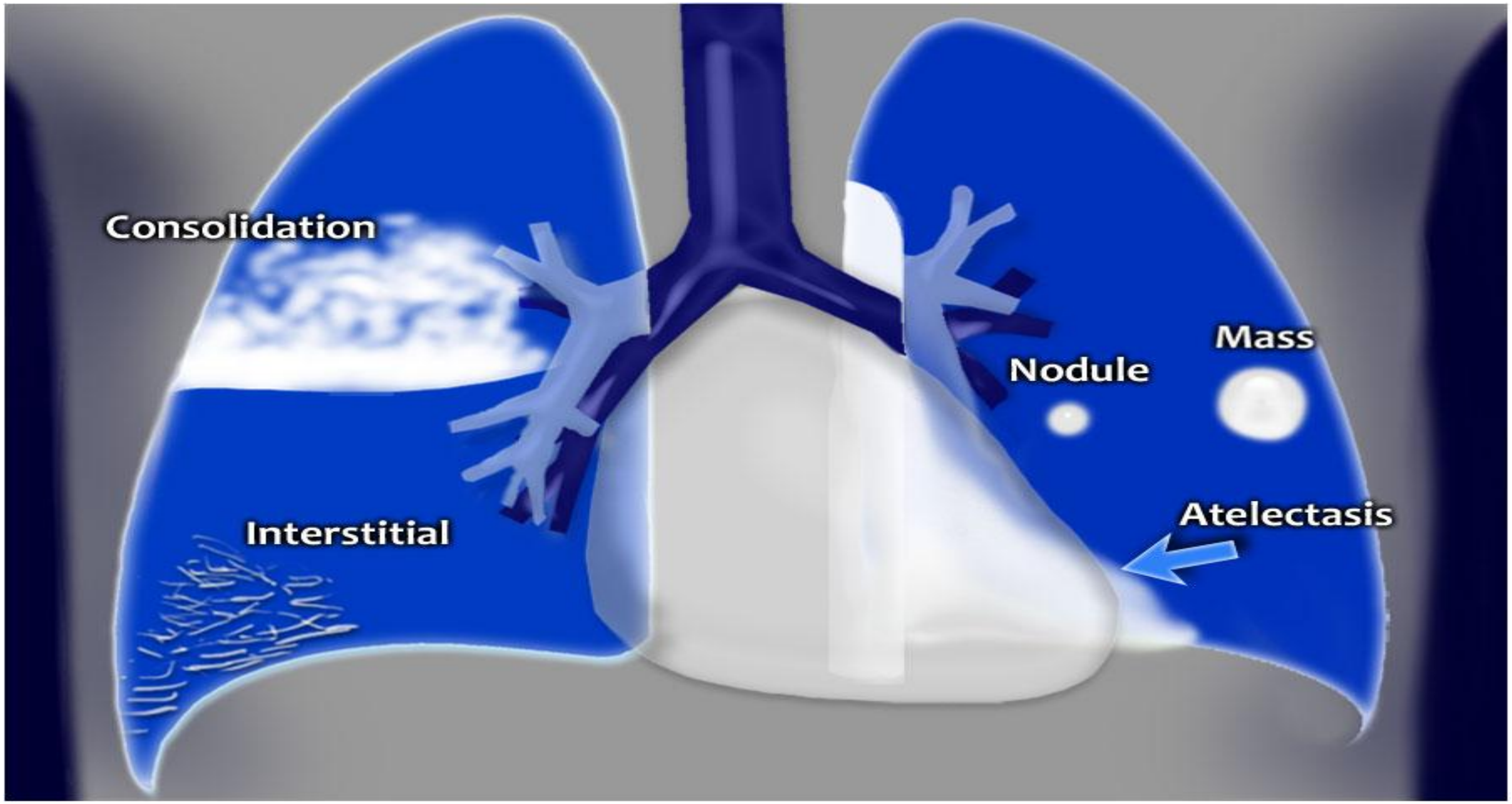
Patient journey | Current state



Family of Diseases Diagnosed by Medscan AI

Pneumonia	Lung cancer	Tuberculosis	COPD (Chronic obstructive pulmonary disease)	Others
Pneumonia *4 m cases/yr, 30% fatality*	Mass	Pleural Thickening	Emphysema	Edema
Infiltration	Nodules *5% malignancy rate, 81K deaths/yr*	Nodules *5% malignancy rate, 81K deaths/yr*		Fibrosis *0.4 m deaths/yr*
Atelectasis				Pneumothorax
Consolidation *42K deaths/yr*	Consolidation *42K deaths/yr*			Hernia
Effusion *0.7 m cases/yr, 32% mortality*	Effusion *0.7 m cases/yr, 32% mortality*			Cardiomegaly *50% mortality rate*

4 patterns of chest x-ray lung abnormalities



Source: Smithuis

02

Exploratory Data Analysis

Dataset Characteristics

- **Source:** National Institutes of Health, Bethesda, MD, USA (Wang et al., 2017)

- **Size and composition:**
    - **112,120 Lung X-Ray images** taken from 30,805 unique patients
    - Each image is **1024x1024** pixels
    - Metadata on patient ID, age and sex available

- **Nature of data:**
    - Images can belong to one or more of 14 lung conditions identified, or to a “no finding” class
    - Condition labels of the original dataset are derived through NLP, with 90% accuracy

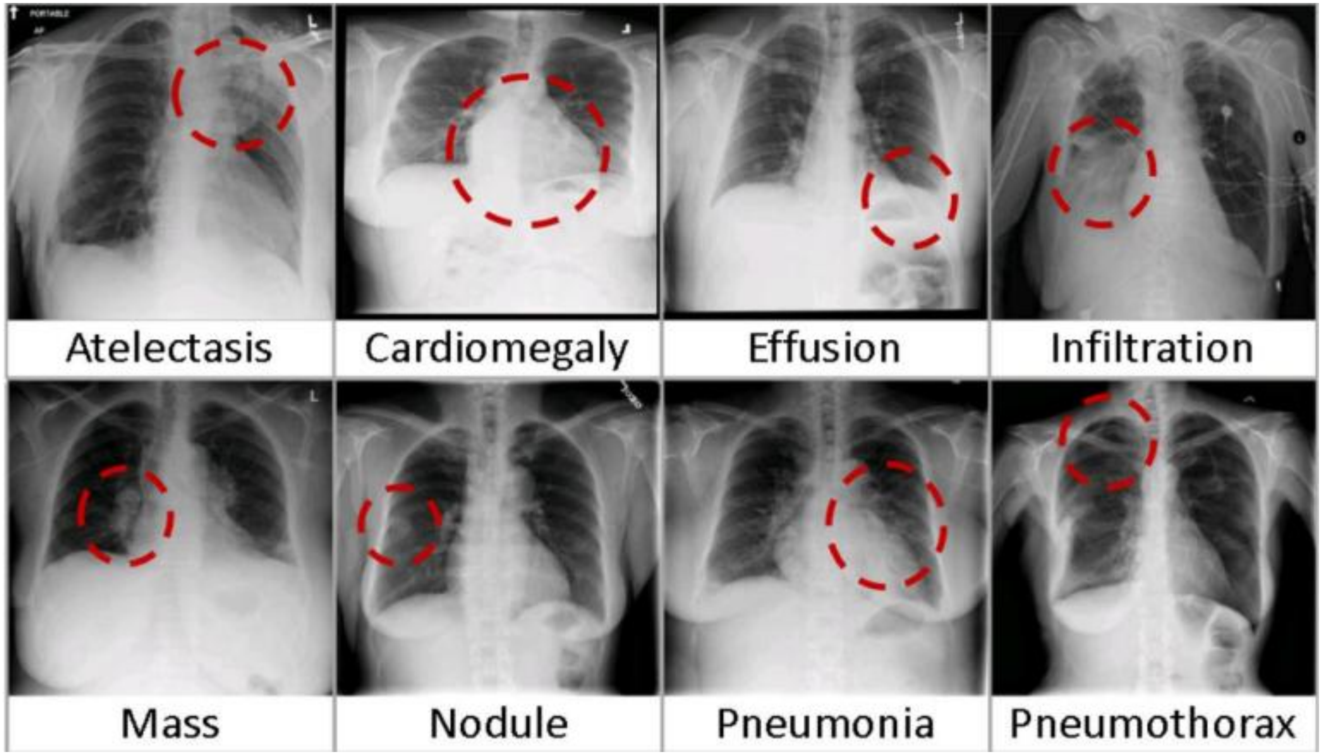


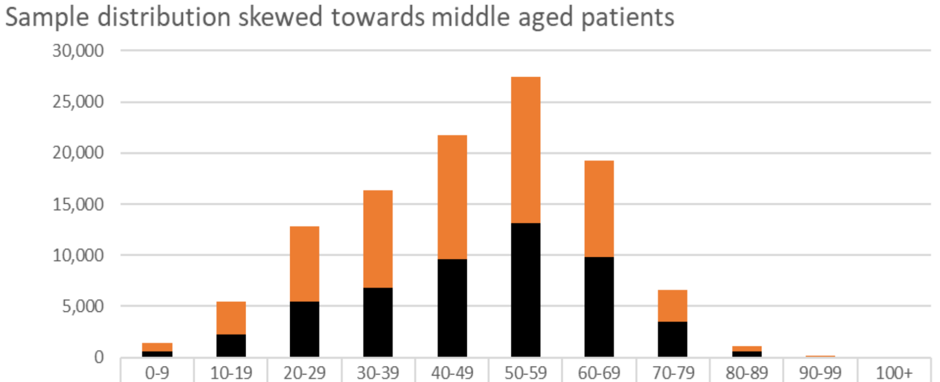
Image source: Wang et al., 2017

Distribution

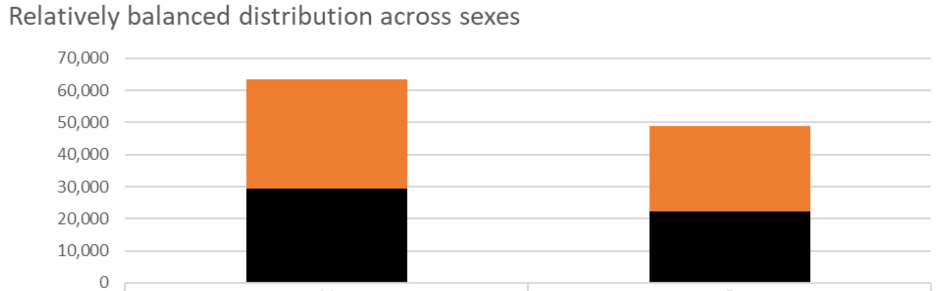
Age and sex distribution is relatively balanced

Certain limitations on use of this dataset do exist

- **Different biological makeup:** Scans in the dataset are from US based subjects, with no information on height, weight or biological origin of subjects
- **Pre existing conditions:** No information about comorbidity or preexisting conditions provided
- **Smoking and tobacco use** information not provided
- **Patient environment:** No information about exposure to high risk settings provided (eg. indoor cooking, masala factory, asbestos exposure, air pollution of surrounding area)



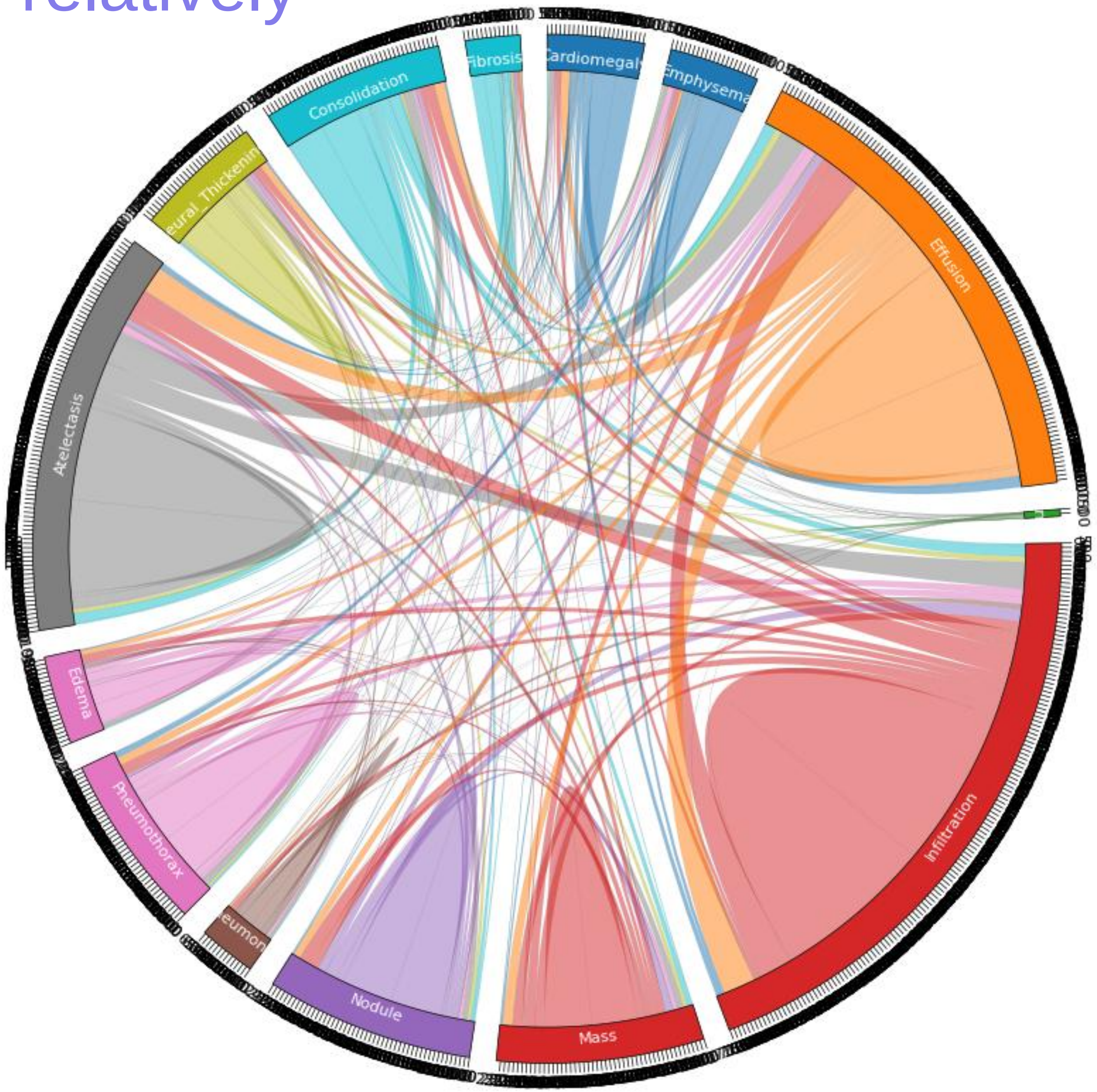
	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100+
No disease	840	3158	7354	9529	12137	14248	9441	3169	452	25	8
Diseased	563	2263	5444	6784	9594	13158	9831	3472	602	40	8

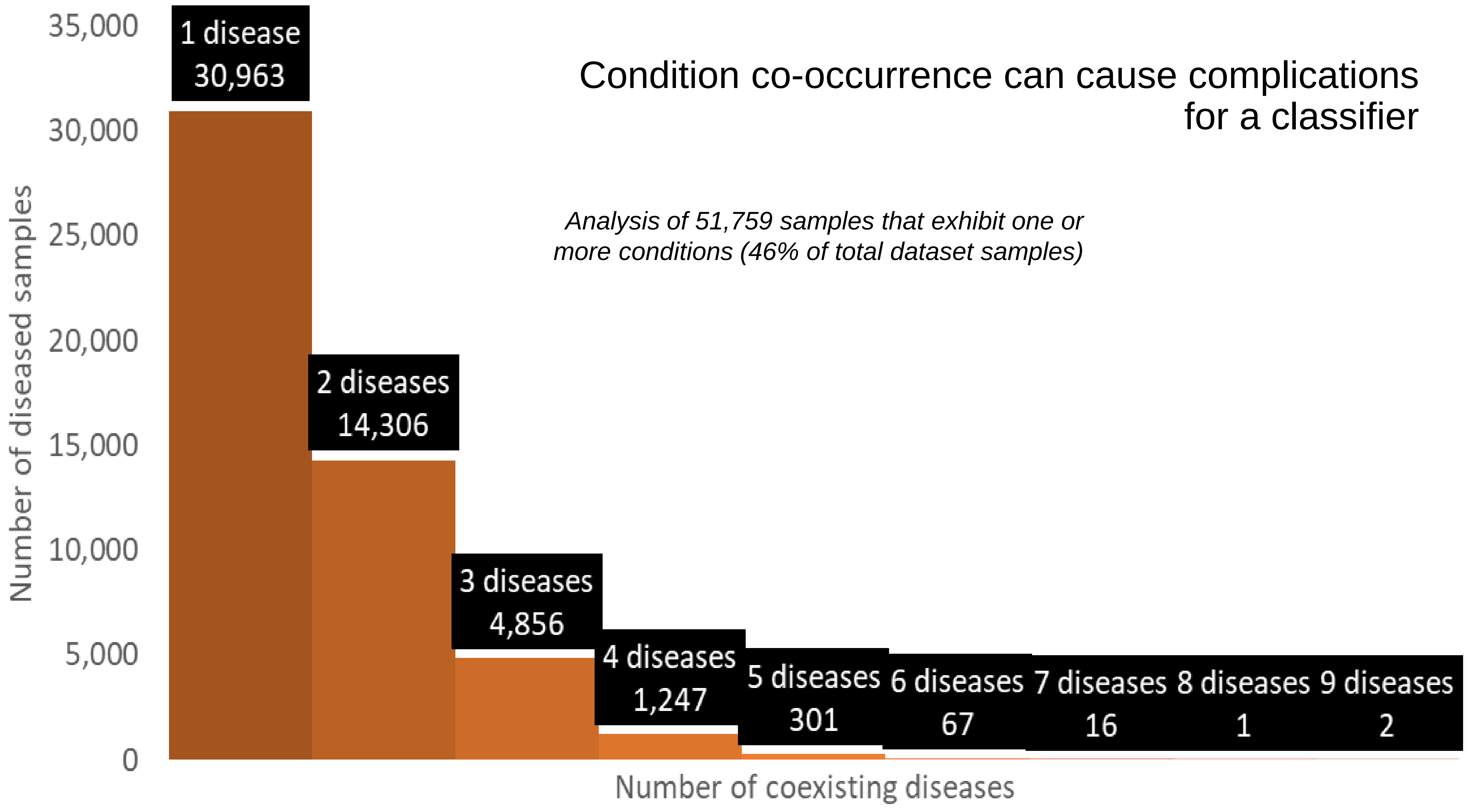


	M	F
No finding	33922	26439
Diseased	29418	22341

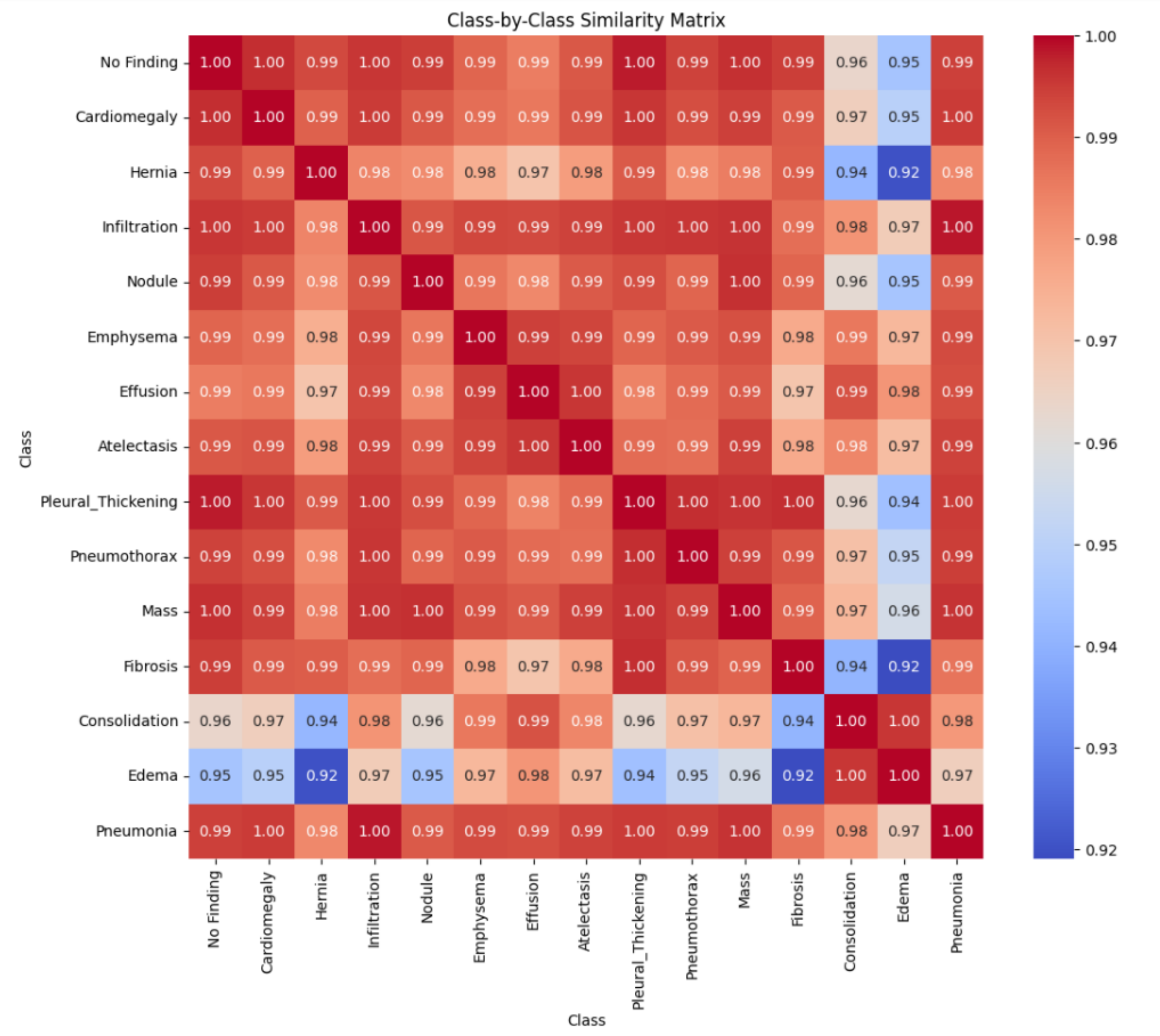
Condition co-occurrence is relatively common

Conditions like Pneumonia, lung cancer, Tuberculosis are *often characterized by multiple conditions existing at once in an X-Ray, depending on the advancement of the patient's condition*




Condition co-occurrence can cause complications for a classifier
Analysis of 51,759 samples that exhibit one or more conditions (46% of total dataset samples)
Number of diseased samples
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
1 disease 30,963
2 diseases 14,306
3 diseases 4,856
4 diseases 1,247
5 diseases 301
6 diseases 67
7 diseases 16
8 diseases 1
9 diseases 2
Number of coexisting diseases

Low class separation can hamper classifier discernment



Approach:

- **Select representative sample:**
  - 12,679 single condition images
  - Proportion of each class in dataset maintained
- **Generate image embeddings** for each image with EfficientNetV2-S (vector length:1280)
- **Reduce image embeddings for faster computation**, preserving 95% information using PCA
- **Create the mean embedding vector** for each class
- **Measure cosine distances between mean embedding vectors**
- **Compute similarity scores**

Disease categories do not naturally form distinct groups

0.0161

Adjusted Rand Index (ARI)

How well do the predicted clusters match the actual disease classes, accounting for randomness?

The clusters are random and don't match the real disease groups.

0.0444

Normalized Mutual Information (NMI)

How much information is shared between clusters and true labels?

No shared information (completely independent)

Approach:

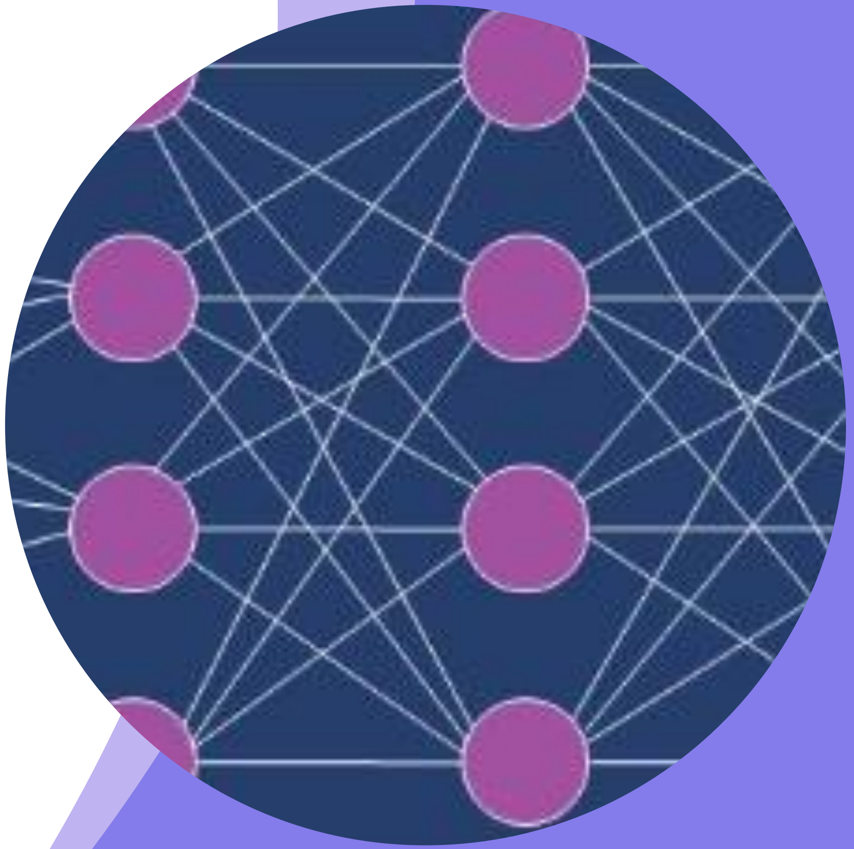
- **Perform K Means Clustering** (k= number of classes)
- **Perform Agglomerative Clustering** (number of clusters = number of classes) to compare against K Means

Insights for model design

Insight	Potential application
• Age and sex distribution is relatively balanced	• Random selection of samples within each class will not hamper demographic balance
• Careful class balancing and sample selection needed	• Choosing a subset with 500-1000 samples from each of the 15 classes will balance the training set
• Condition co-occurrence can cause complications for a classifier	• Limit or remove cases where more than one disease exists from the training set
• Low class separation can hamper classifier discernment	• Consider pretrained models like DenseNet201 or EfficientNetV2S for robust feature extraction
• Disease categories do not naturally form distinct groups	

03

Model Architecture & Training



Chest X-ray Image Validation - Preprocessing model for Chest X-Ray Classification

- **Image Preprocessing & Format Conversion**
  - Converts DICOM to PNG/JPG and applies resizing, normalization, and feature extraction (HOG + Canny) to prepare images.

- **Feature Extraction: HOG + Canny Edge Detection**

- Histogram of Oriented Gradients (HOG): Captures the shape and texture of the image, making it easier to distinguish chest X-rays from non-X-rays.Canny Edge Detection: Identifies key structural edges to differentiate medical X-rays from other images like normal photographs or unrelated scans.

- **Chest X-ray vs. Non-X-ray Classification**
  - A Random Forest model is trained on HOG + Canny features to detect whether an input image is a valid chest X-ray or not before disease classification.

- **Validation Before Disease Classification**
  - Ensures that only authentic chest X-ray images are passed to the disease classification model, reducing false predictions.

Chest X-ray Image Validation – Performance

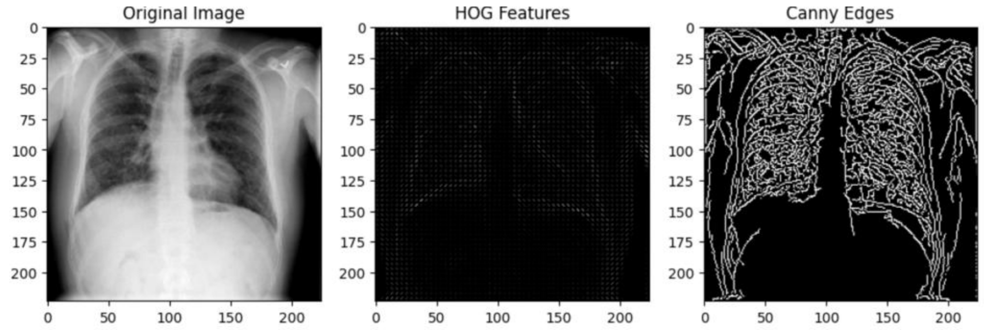
- Uses accuracy, precision, recall, and F1-score to measure the effectiveness of the validation model in correctly classifying chest X-rays.

✅ Accuracy: 0.9871
✅ Precision: 1.0000
✅ Recall: 0.9748
✅ F1-score: 0.9872

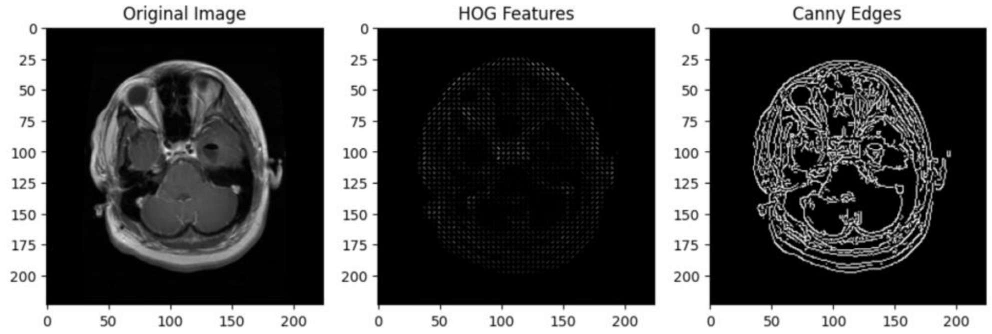
Classification Report:

	precision	recall	f1-score	support
0	0.97	1.00	0.99	266
1	1.00	0.97	0.99	278
accuracy			0.99	544
macro avg	0.99	0.99	0.99	544
weighted avg	0.99	0.99	0.99	544

Test result on a Sample Chest X-ray Image



Test result on a Sample Chest X-ray Image



Choosing a pretrained model: EfficientNetV2-S vs DenseNet201

Preprocessing

- **Patient-wise split**: Ensured no overlap by allocating 80% of unique patients to training and 20% to validation to prevent data leakage.
- **Preprocessing pipeline:** Applied image decoding, resizing, and augmentation only to training data for better generalization.
- **Fair evaluation**: Kept validation data unchanged to accurately assess model performance.
- **Optimized loading**: Used efficient batching and caching to speed up training.

Why these models?

- DenseNet201: Efficient feature reuse through dense connections helps in extracting intricate patterns from chest X-rays, improving disease detection while keeping the model lightweight.
- EfficientNetV2-S: Optimized for speed and accuracy, it reduces computational cost while maintaining high classification performance, making it ideal for real-time medical image analysis

Architecture

Input Layer: Accepts 600x600 RGB chest X-ray images with 3 channels.

Feature Extraction: Uses pretrained DenseNet201 (18.3M parameters) to extract deep image features.

Regularization: Includes Dropout layers to prevent overfitting and improve generalization.

Regularization: Feature Reduction: Global Average Pooling (GAP) converts feature maps into a compact 1920-element vector.

Output Layer: A Dense layer (14 neurons, Sigmoid activation) enables multi-label classification of chest diseases.

Model Performance Scores

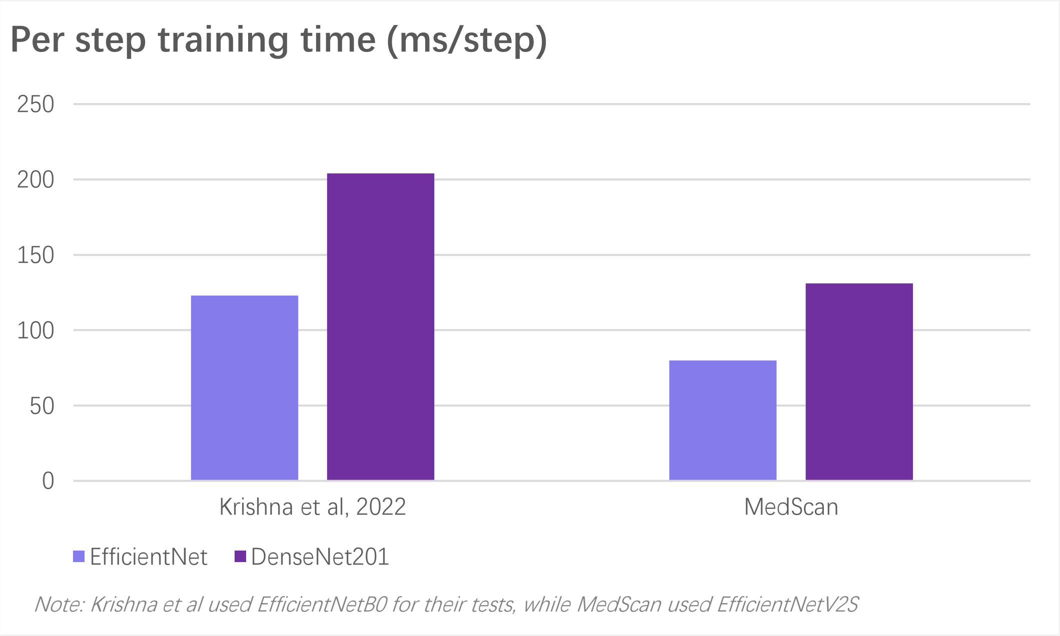
Reason 1: Better results

EfficientNetV2S outperformed DenseNet201 and other models on the majority of diseases (comparable performance on others)

Condition	AUC scores		Best performance
	DenseNet201	EfficientNetV2S	
Cardiomegaly	0.9219	0.9168	DenseNet
Infiltration	0.7404	0.7397	DenseNet
Effusion	0.8557	0.8551	DenseNet
Pneumothorax	0.8949	0.8940	DenseNet
Fibrosis	0.8444	0.8371	DenseNet
Edema	0.8629	0.8606	DenseNet
Hernia	0.9306	0.9376	EfficientNet
Nodule	0.8130	0.8164	EfficientNet
Emphysema	0.9351	0.9450	EfficientNet
Atelectasis	0.8002	0.8063	EfficientNet
Pleural_Thickening	0.7848	0.8010	EfficientNet
Mass	0.8473	0.8548	EfficientNet
Consolidation	0.7318	0.7394	EfficientNet
Pneumonia	0.6909	0.7291	EfficientNet

Reason 2: Faster processing and less compute intensive

EfficientNetV2S provided 39.7% time savings over DenseNet201 in similar medical image classification tasks (Krishna et al., 2022)



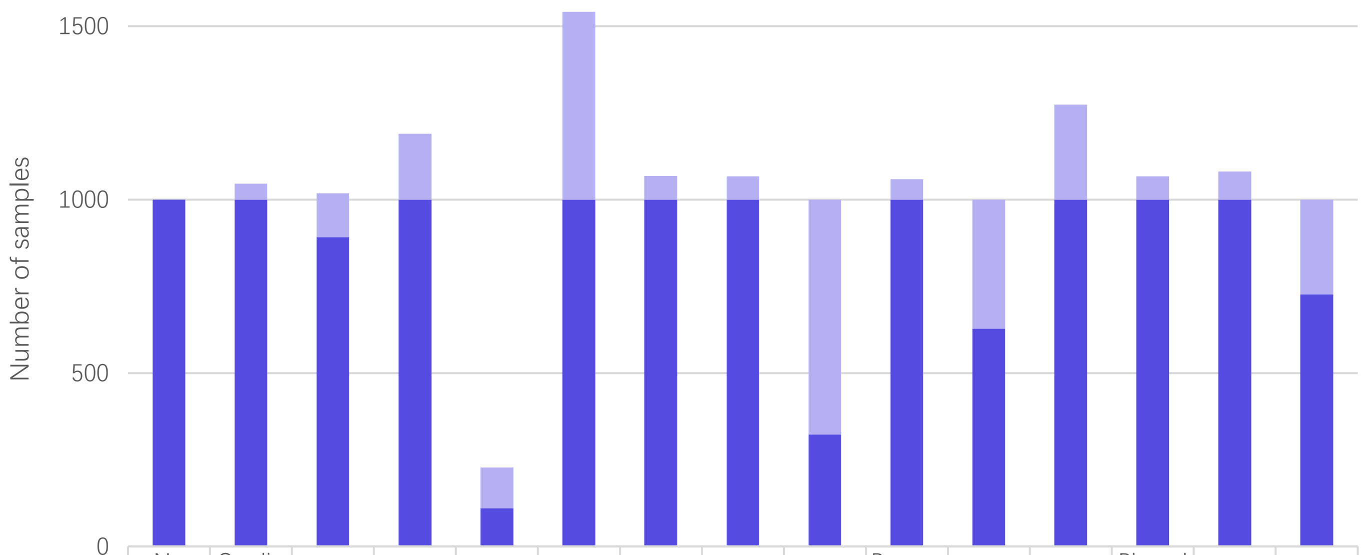
https://www.researchgate.net/figure/Performance-and-efficiency-comparison-of-DenseNet-and-EfficientNet_tbl4_357592591

EfficientNetV2S(14k images)|Pursuing balanced, unambiguous training

Need for single disease/condition samples and relatively balanced classes for improved classification power

We selected (without replacement) a minimum of **1000 samples** or the number of occurrences of each disease in the dataset, giving priority to scans showing a single disease. The validation dataset was a randomly selected 20% of this dataset

Relatively balanced classes



	No Finding	Cardiomegaly	Emphysema	Effusion	Hernia	Infiltration	Mass	Nodule	Pneumonia	Pneumothorax	Edema	Atelectasis	Pleural_Thickening	Consolidation	Fibrosis
Multiple	0	46	126	190	117	541	68	67	678	59	372	274	67	81	273
Single	1000	1000	892	1000	110	1000	1000	1000	322	1000	628	1000	1000	1000	727

Single Multiple

14,064

Samples chosen across 15 condition classes

Number of conditions/sample	Number of samples	Proportion of dataset
Single condition	12,679	90.15%
2 conditions	1,210	8.60%
3 conditions	166	1.18%
4 conditions	5	0.04%
5 conditions	3	0.02%
6 conditions	1	0.01%

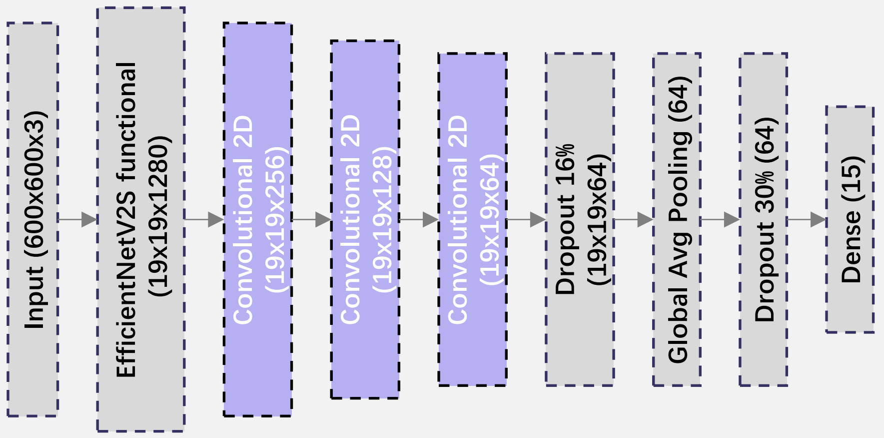
Finetuning the pretrained model | How much is needed?

Trade-offs considered when finetuning

- **Higher training time and costs** (in ascending order of complexity)
  - Option 1: Maintain standard EfficientNet (ie. Original)
  - Option 2**:** Add trainable convolutional layers, keep EfficientNet untrainable
  - Option 3: Make the last n blocks of EfficientNet trainable
- **Improved classification ability**
- **Better and faster GradCAM results**

Selected architecture (Option 2)

Adding convolutional layers post EfficientNet (layers highlighted in purple)



Effects:

Finetuning improved AUC performance across categories by four percentage points

Condition	AUC scores		Best performance
	Finetuned	Original	
Cardiomegaly	0.9564	0.9168	Finetuned
Hernia	0.8724	0.9376	Original
Infiltration	0.7733	0.7397	Finetuned
Nodule	0.8536	0.8164	Finetuned
Emphysema	0.9427	0.9450	Original
Effusion	0.8892	0.8551	Finetuned
Atelectasis	0.8682	0.8063	Finetuned
Pleural_Thickening	0.8720	0.8010	Finetuned
Pneumothorax	0.9232	0.8940	Finetuned
Mass	0.8982	0.8548	Finetuned
Fibrosis	0.8576	0.8371	Finetuned
Consolidation	0.8367	0.7394	Finetuned
Edema	0.9308	0.8606	Finetuned
Pneumonia	0.8394	0.7291	Finetuned

Other considerations

Model architecture

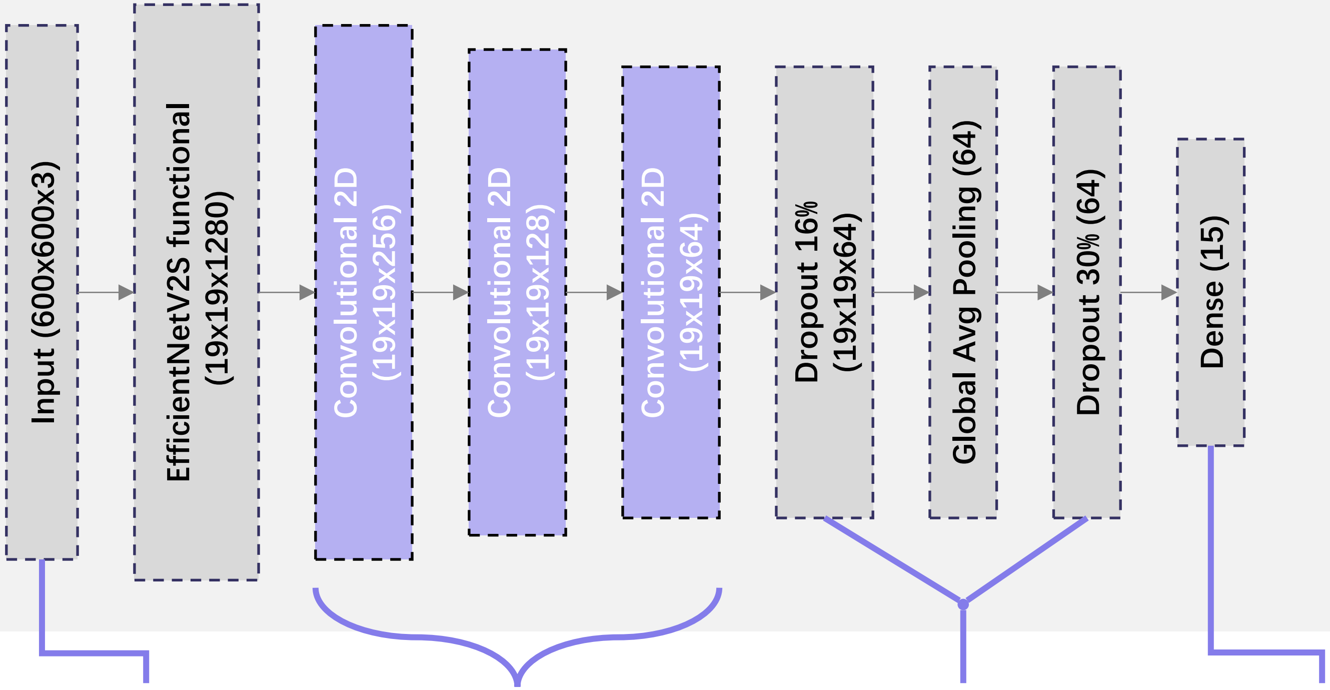


Image size reduced from 1024x1024 to 600x600 for faster, less expensive computations (Kufel et al use 512x512)

Kernel size (3x3) is consistent with EfficientNet kernels, and is computationally less demanding than larger kernels

ReLU activation is computationally more efficient and prevents overfitting by allowing for sparsity

Dropout layers allow generalization before Global Avg Pooling and Dense layers, preventing the model from overreliance on certain feature maps

Sigmoid activation allows for the model to assign probabilities to each disease independently (unlike softmax, which makes classes mutually exclusive)

Other training based considerations

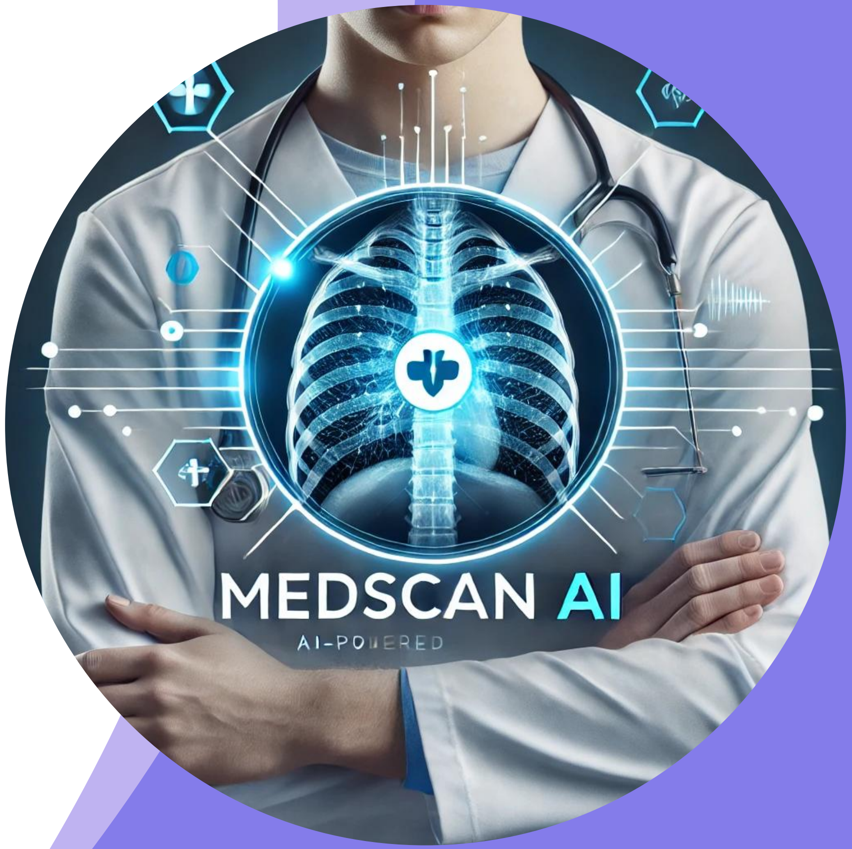
- **Image augmentation** in the form of image flips (left to right) and ±20% variations in brightness, contrast and saturation applied to reduce overfitting
- **Binary Crossentropy** used as a loss function as each choice is independent (as opposed to **Categorical Crossentropy** that is used when classes are mutually exclusive). Binary Crossentropy is also computationally cheaper and more resilient than functions like **Focal Loss**
- **Learning rate is reduced** by 10x (from 10^{-4} up to 10^{-6}) to allow for faster convergence initially but finer weight adjustment when validation loss has not changed for 2 epochs
- **Early Stopping** occurs after 5 epochs of no improvement in validation loss, and prevents overfitting
- **Model checkpointing** allows for a model with good weights to be saved and restored , even if the model gets overfit in subsequent epochs

Tools and hardware

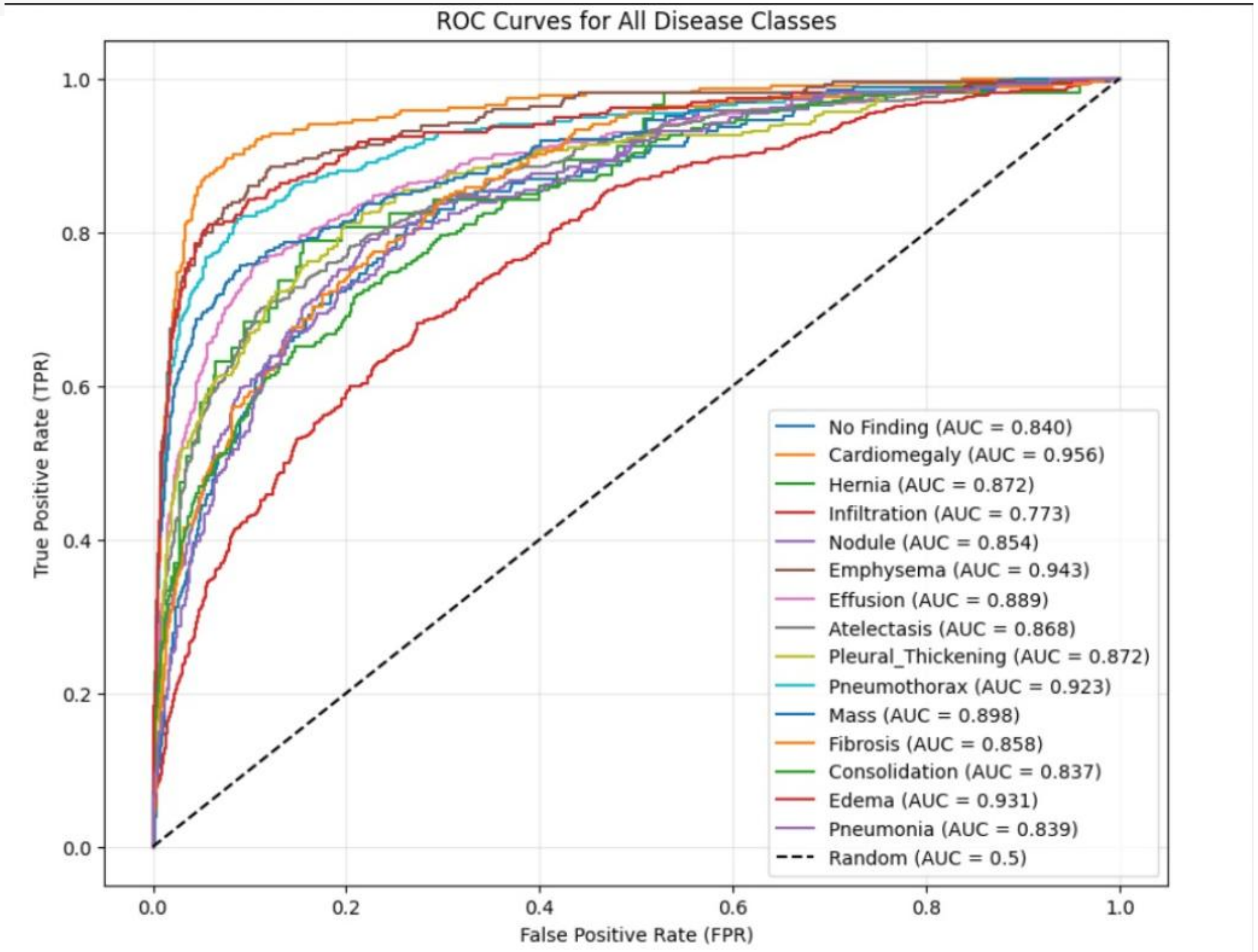
- **RunPod 1 x A100 PCIe** machine with 31 vCPU and 117 GB RAM (1 unit of GPU), 100 GB hot storage
- **Tensorflow 2.14.0**

04

Results and Analysis



ROC curve for all disease classes



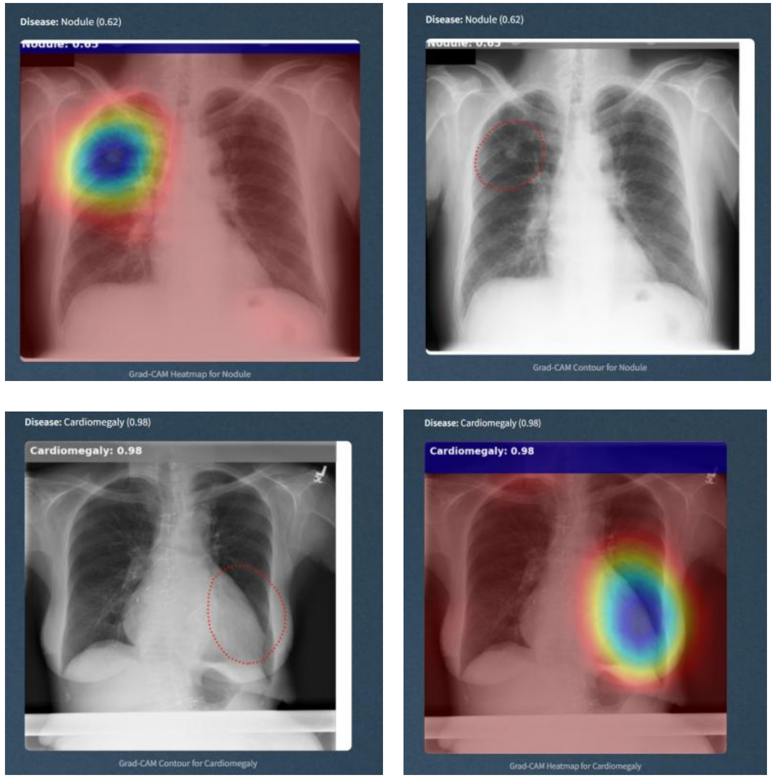
Performance vs similar studies

Medscan outperformed previous studies on almost all disease conditions

All figures below represent test AUC scores.

Pathology Label	Ğuendel et al.	Yan et al.	Baltruschat et al.	Kufel et al.	Medscan	Best performance
Official split	Yes	Yes	Yes	No	No	
Atelectasis	0.767	0.792	0.763	0.817	0.868	Medscan
Cardiomegaly	0.883	0.881	0.875	0.911	0.956	Medscan
Effusion	0.828	0.842	0.822	0.879	0.889	Medscan
Infiltration	0.709	0.710	0.694	0.716	0.773	Medscan
Mass	0.821	0.847	0.820	0.853	0.898	Medscan
Nodule	0.758	0.811	0.747	0.771	0.854	Medscan
Pneumonia	0.731	0.740	0.714	0.769	0.839	Medscan
Pneumothorax	0.846	0.876	0.840	0.898	0.923	Medscan
Consolidation	0.745	0.760	0.749	0.815	0.837	Medscan
Edema	0.835	0.848	0.846	0.908	0.931	Medscan
Emphysema	0.895	0.942	0.895	0.935	0.943	Medscan
Fibrosis	0.818	0.833	0.816	0.824	0.858	Medscan
Pleural_Thickening	0.761	0.808	0.763	0.812	0.872	Medscan
Hernia	0.896	0.934	0.937	0.890	0.872	Baltruschat et al.
Average	0.807	0.830	0.727	0.843	0.880	Medscan

Model Explainability



Grad-CAM

Using Grad- CAM for visual explanations of model decisions, showcasing which parts of the image contributed to the diagnosis.

Potential improvements to the model

- **Deeper training:**
  - Making more parts of the EfficientNet block trainable
  - Adding more convolutional layers
- **Different architecture:**
  - Experimenting with LSTM based models
  - Hybrid models that take image embeddings and HOG embeddings
  - Multi layer classifiers that consider families of diseases in the first step (lung cancer, TB, pneumonia) and further classify within those families in the second step
  - Utilizing self- attention techniques to focus on important regions within images for more precise diagnosis.
- **Combining datasets:**
  - Source more images on underrepresented classes (eg. Hernia, Pneumonia and Edema) for training from other datasets
  - Generate synthetic images of Hernia and Pneumonia to boost performance

05

Features of MedScan AI

Overview of MedScan AI Application

01.

Objectives

The MedScan AI project aims to leverage artificial intelligence to enhance medical imaging accuracy for lung X-Ray, streamline diagnostic procedures, and improve overall patient care.

02.

Scope and Impact

The project encompasses developing an innovative AI tools for analyzing medical images, potentially transforming diagnostic processes and significantly impacting healthcare quality and accessibility.

Key Features

- **Disease Detection**

  MedScan AI employs advanced algorithms to identify a wide range of diseases from medical images, significantly aiding in early diagnosis and treatment planning.

- **Accuracy Rates**

  The platform boasts impressive accuracy rates, ensuring that diagnoses are reliable and reducing the margin of error typically associated with manual interpretations.

- **User-Friendly Design**

  The user interface is intuitively designed, allowing healthcare professionals to navigate the system effortlessly, enhancing overall user experience and efficiency.

- **Support for Physicians**

  Decision support : To help physicians make informed decisions by providing evidence- based recommendations and critical data analysis.
  Reducing workloads.

- **Patient Experience**

  Early Diagnosis.
  Improved Outcomes.
  Cost saving

Importance of AI in Medical Imaging

Enhanced Diagnostic Accuracy

AI algorithms provide higher diagnostic precision by analyzing medical images meticulously, reducing the chances of human error.

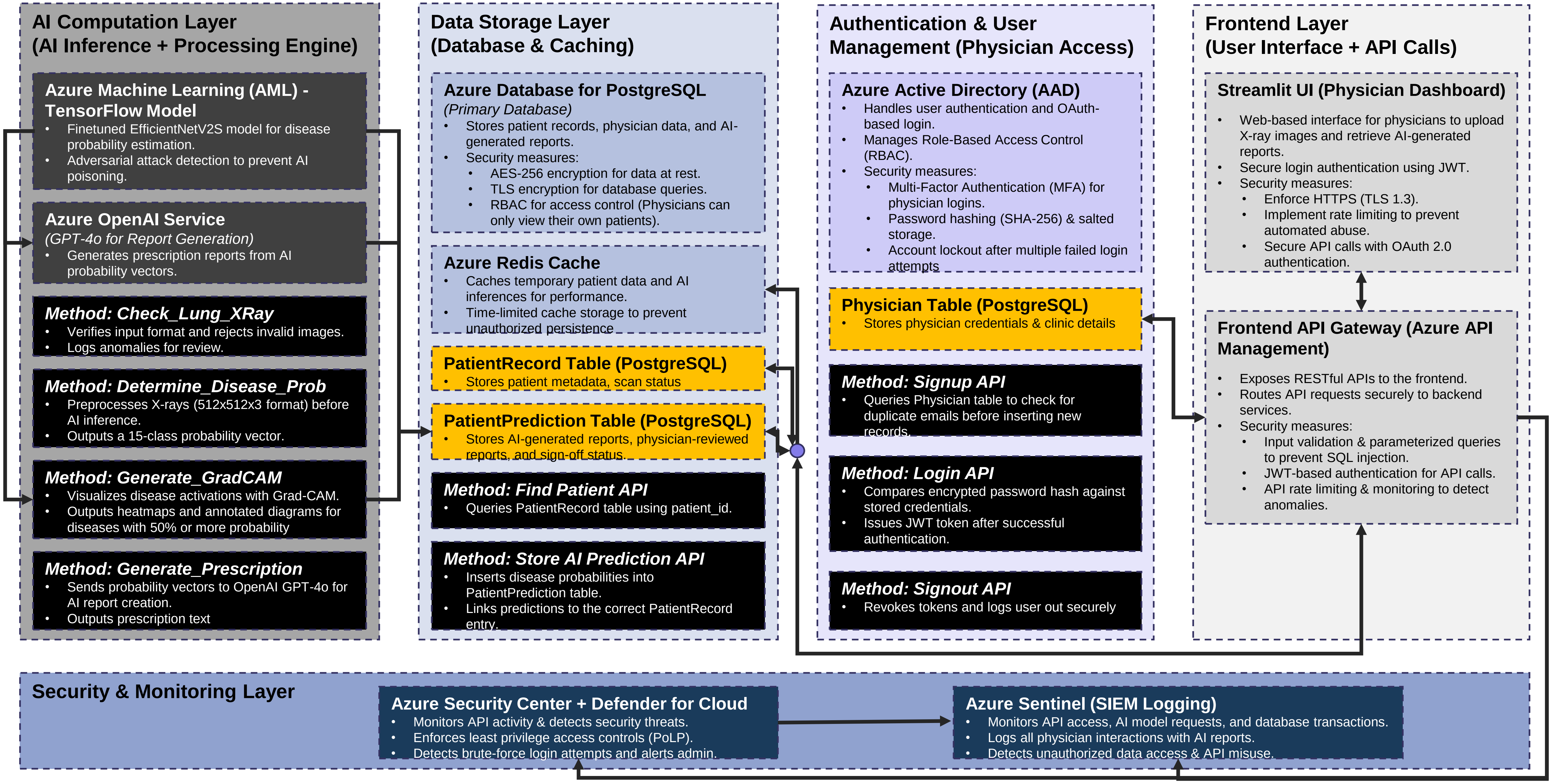
Speed and Efficiency

AI can process and interpret imaging data quickly, leading to faster diagnosis and treatment planning, enhancing patient outcomes.

Cost-Effectiveness

By improving accuracy and speed, AI reduces the need for additional tests and procedures, lowering overall healthcare costs.

Medscan AI App architecture



Challenges

Data Privacy Concerns

The use of AI in healthcare raises significant concerns about patient data privacy, necessitating stringent protocols to protect sensitive information.

Reliability and Accuracy

Ensuring the reliability and accuracy of algorithms is challenging, as they must be rigorously validated to meet medical standards and avoid misdiagnoses.

Data availability

Availability of dataset with different demography of patients etc is a challenge and potentially limiting the effectiveness of the diagnosis.

Need for Professional Training

Healthcare professionals need specialized training to effectively interpret and integrate AI- driven diagnostic tools into clinical practice, requiring ongoing education.

Ethical Use of AI & Data Security

Bias and Fairness

Ensuring AI systems are free from bias and treat all users fairly is vital to building trust and promoting equal opportunities.

Accountability

Developers and organizations must take responsibility for AI- driven decisions, ensuring transparent and traceable outcomes.

Data Protection Measures

Implementing robust data protection measures is essential to safeguard sensitive information from breaches and unauthorized access.

Compliance with Laws

Adhering to data security laws and regulations is critical for legal compliance and maintaining users' trust in AI applications.

06

Future Prospects

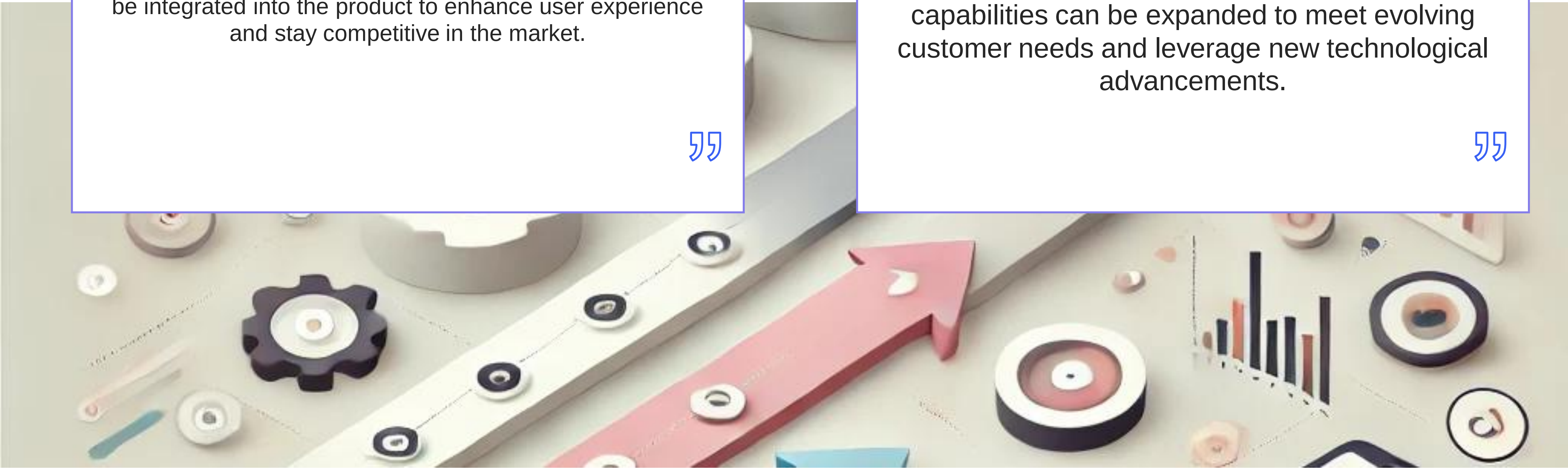
Potential Advances

New Features

This section explores the innovative new features that could be integrated into the product to enhance user experience and stay competitive in the market.

Expanding Capabilities

Discussing the ways in which the product's capabilities can be expanded to meet evolving customer needs and leverage new technological advancements.



Continuous Improvement

01

Research and Development

Highlighting the importance of ongoing research and development efforts to innovate and improve the product, ensuring long- term success.

02

Feedback and Iterations

Emphasizing the role of customer feedback in driving iterative improvements and adjustments to the product for better user satisfaction.


COLLEGE
Thank You
A COLLEGE PROJECT
THANK YOU