NODE21

Chest Nodule Generation

Abu_Shahid(B20CS003) Atharva_Pandey(B20CS007)

WHY? NEED?

- Lung cancer is the leading cause of cancer deaths globally in both men and women.

Why? need?

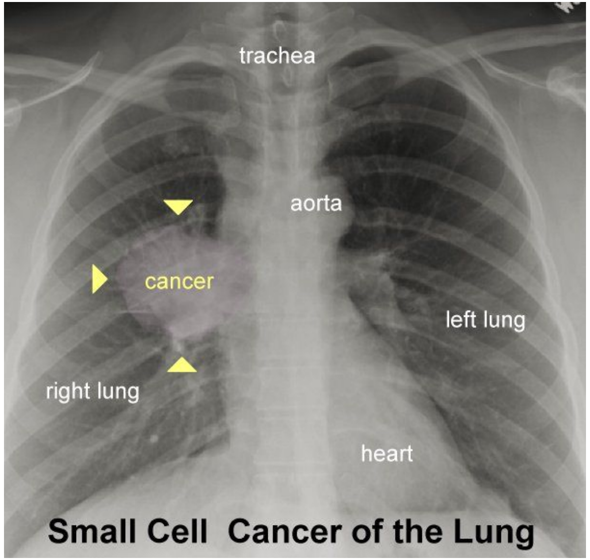
- Lung cancer is the leading cause of cancer deaths globally in both men and women.
- Symptoms of lung cancer often appear in later stages, reducing the effectiveness of treatment.

Why? need?

- Lung cancer is the leading cause of cancer deaths globally in both men and women.
- Symptoms of lung cancer often appear in later stages, reducing the effectiveness of treatment.
- Early detection is crucial for lowering mortality rates.

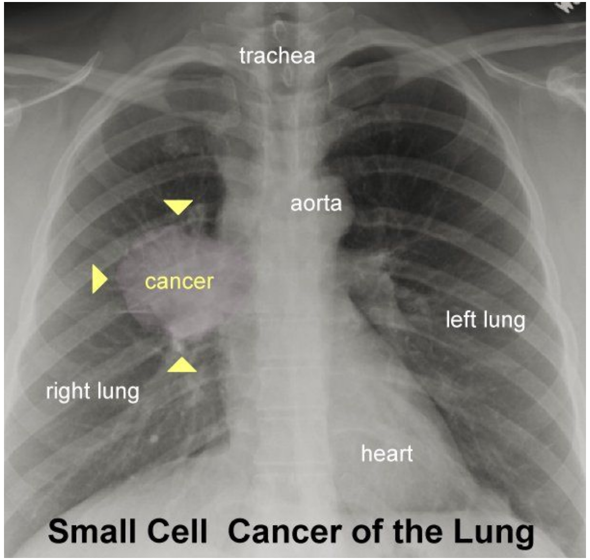
Why? need?

- Pulmonary nodules are the first signs of lung cancer, often visible before symptoms appear.



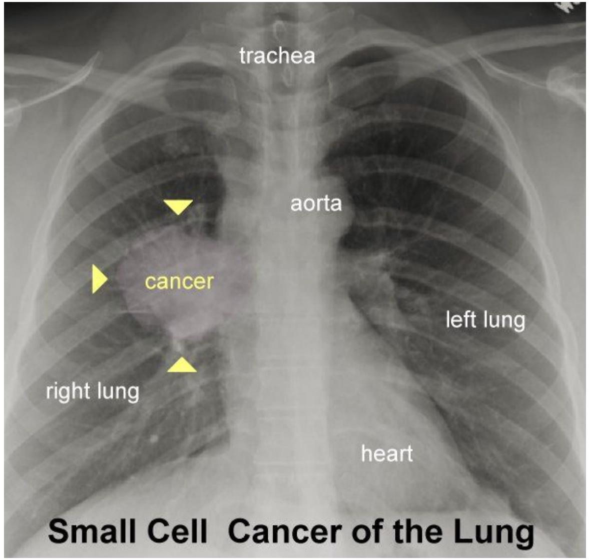
Why? need?

- Pulmonary nodules are the first signs of lung cancer, often visible before symptoms appear.
- Chest radiography (CXR) is the most common imaging method globally.



WHY? NEED?

- Pulmonary nodules are the first signs of lung cancer, often visible before symptoms appear.
- Chest radiography (CXR) is the most common imaging method globally.
- Pulmonary nodules are commonly found incidentally during routine examinations or CXR imaging for other health concerns.



Difficulty

- Detecting lung nodules on CXR can be challenging due to factors such as size, density, and location.

Difficulty

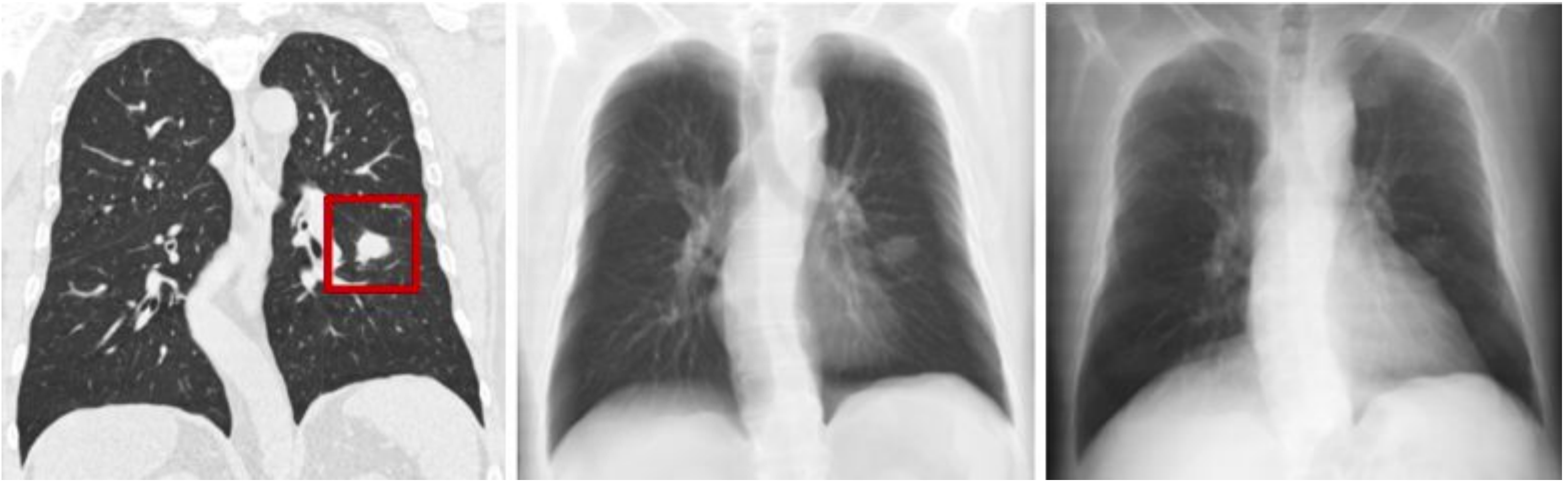
- Detecting lung nodules on CXR can be challenging due to factors such as size, density, and location.
- **Projection images in CXR can cause nodules to overlap with other anatomical structures like the heart, hilum, or diaphragm.**

Difficulty

- Detecting lung nodules on CXR can be challenging due to factors such as size, density, and location.
- **Projection images in CXR can cause nodules to overlap with other anatomical structures like the heart, hilum, or diaphragm.**
- This overlap can make nodules difficult to discern and potentially impossible to detect.

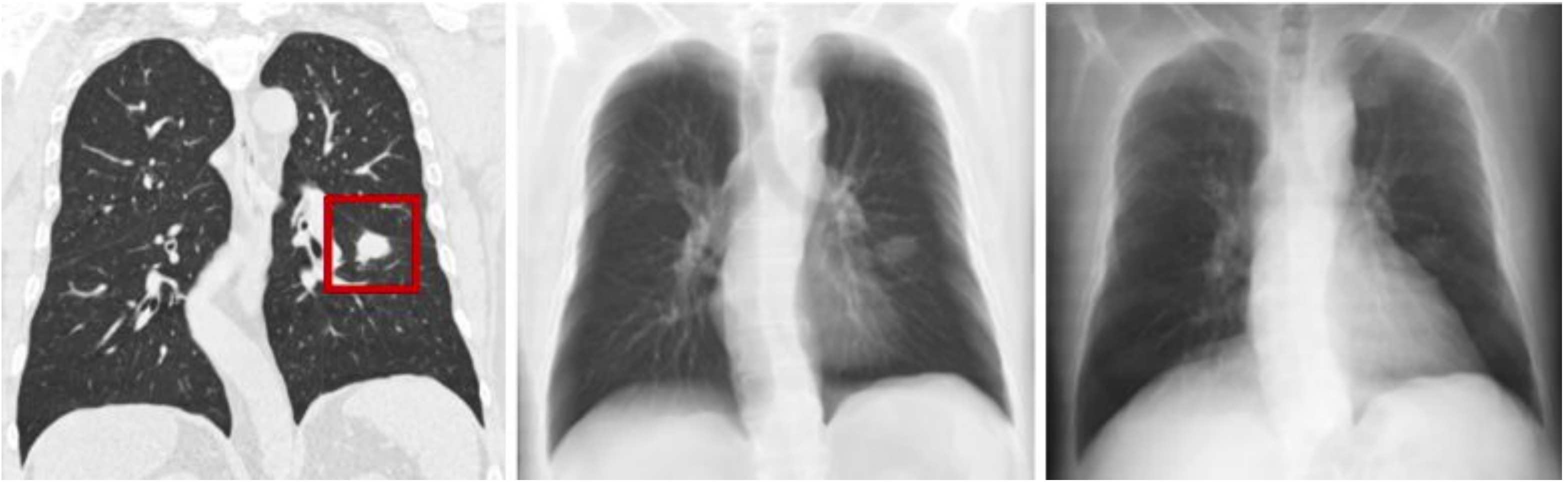
BIggest Underlying issue??

- Lack of data



BIggest Underlying issue??

- Lack of data
- Uneconomical costs of annotating such a data



NODE21 public CXR training dataset

- It contains **4882** frontal chest radiographs.

NODE21 public CXR training dataset

- It contains **4882** frontal chest radiographs.
- **1134** images are annotated with bounding boxes around nodules, totaling **1476** nodules.

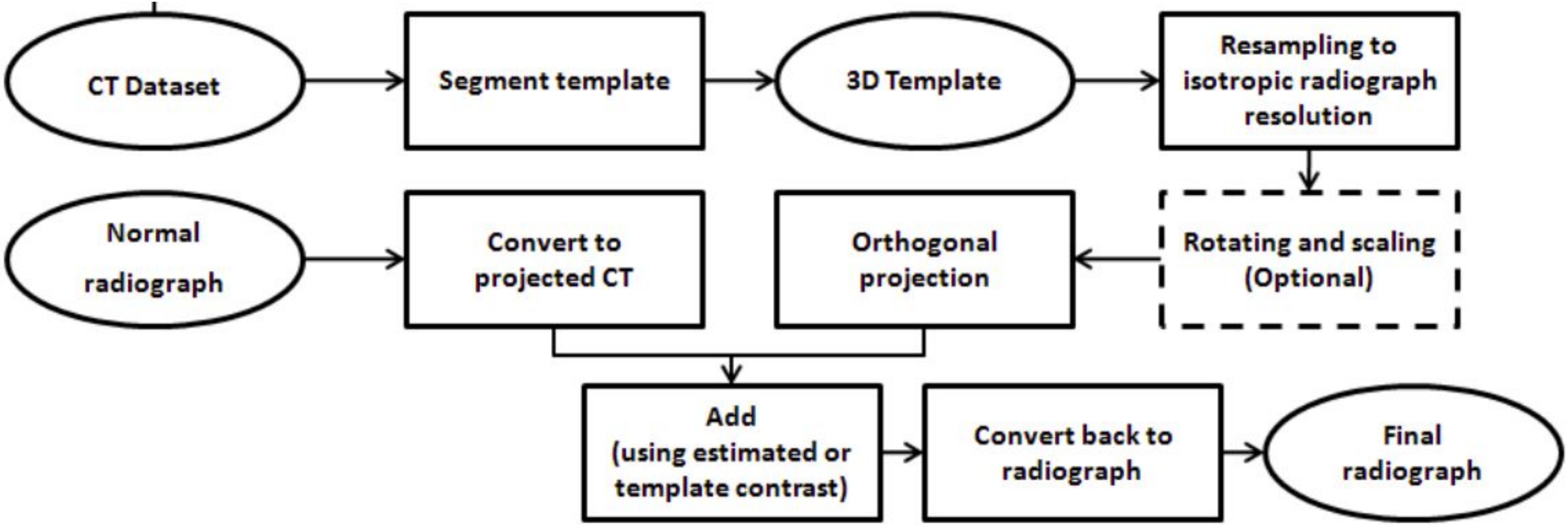
NODE21 public CXR training dataset

- It contains 4882 frontal chest radiographs.
- 1134 images are annotated with bounding boxes around nodules, totaling 1476 nodules.
- **3748** images represent the negative class, as they do not contain nodules.

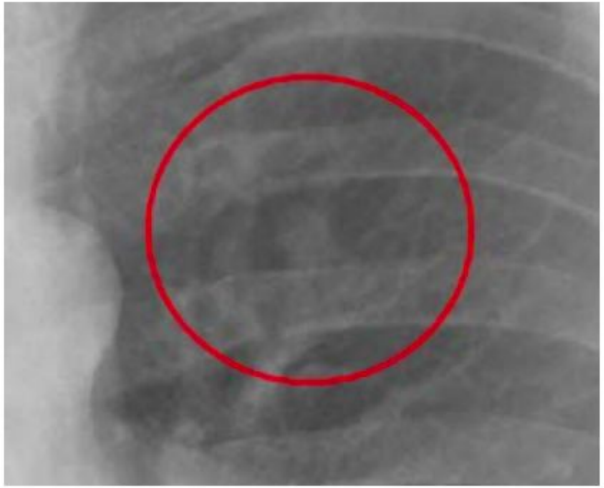
NODE21 public CXR training dataset

- It contains 4882 frontal chest radiographs.
- 1134 images are annotated with bounding boxes around nodules, totaling 1476 nodules.
- 3748 images represent the negative class, as they do not contain nodules.
- Nodules should be generated for each CXR image marked with label==0, indicating the absence of nodules.

Approach 0: node21_generation_baseline



Approach 0: node21_generation _baseline: results


Approach 0:node21_generation _baseline: results

Approach 1: Eigen Nodules

- Extract nodules from dataset, resize to **64*64**, and store

Approach 1: Eigen Nodules

- Extract nodules from dataset, resize to 64*64, and store
- Extract vector subspace from the images and create a convex hull parameterized eigenvectors

Approach 1: Eigen Nodules

- Extract nodules from dataset, resize to 64*64, and store
- Extract vector subspace from the images and create a convex hull parameterized eigenvectors
- Randomly sample vectors from the hull and reshape to 64*64 for nodule reconstruction

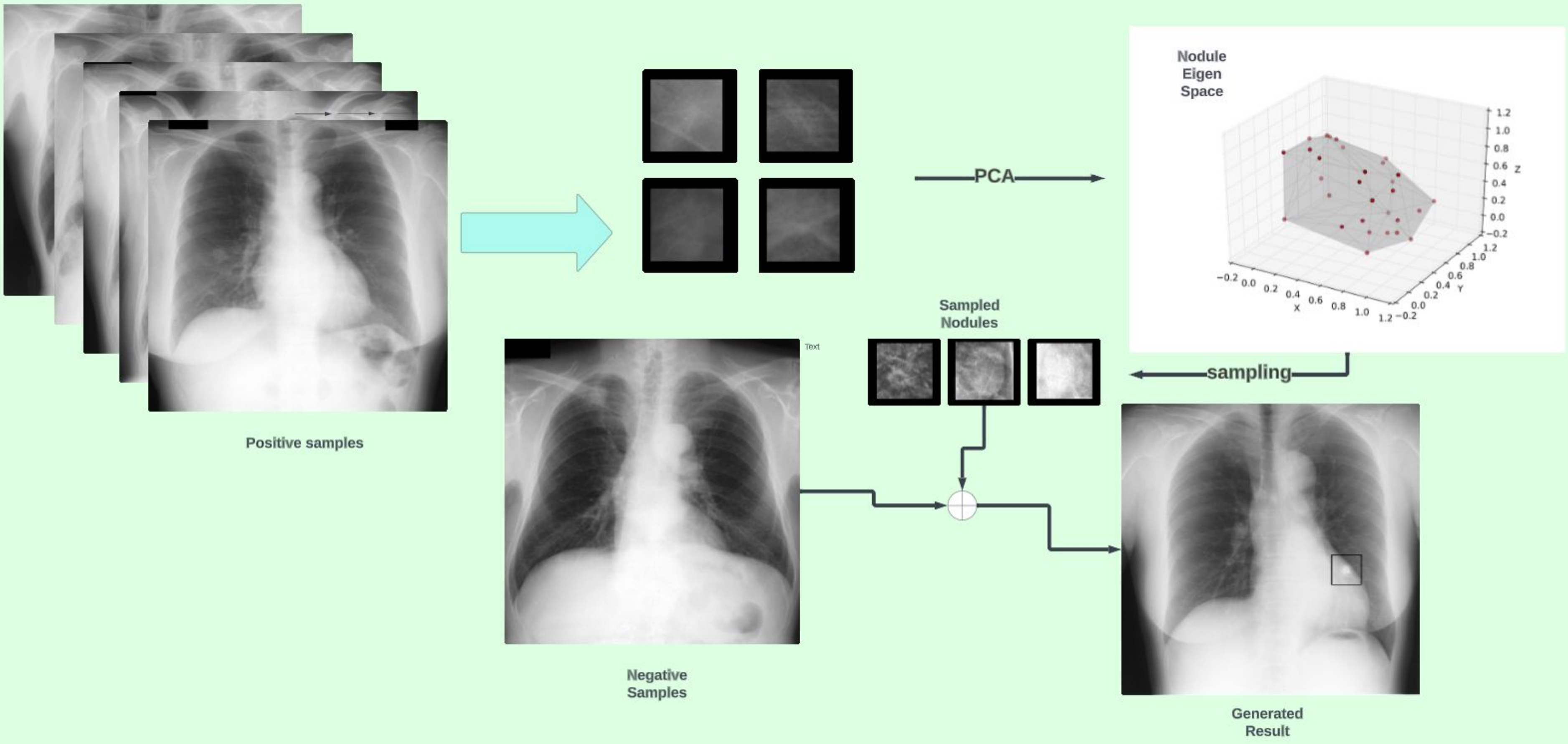
Approach 1: Eigen Nodules

- Extract nodules from dataset, resize to 64*64, and store
- Extract vector subspace from the images and create a convex hull parameterized eigenvectors
- Randomly sample vectors from the hull and reshape to 64*64 for nodule reconstruction
- Use Gaussian mixture model to create a region proposal network

Approach 1: Eigen Nodules

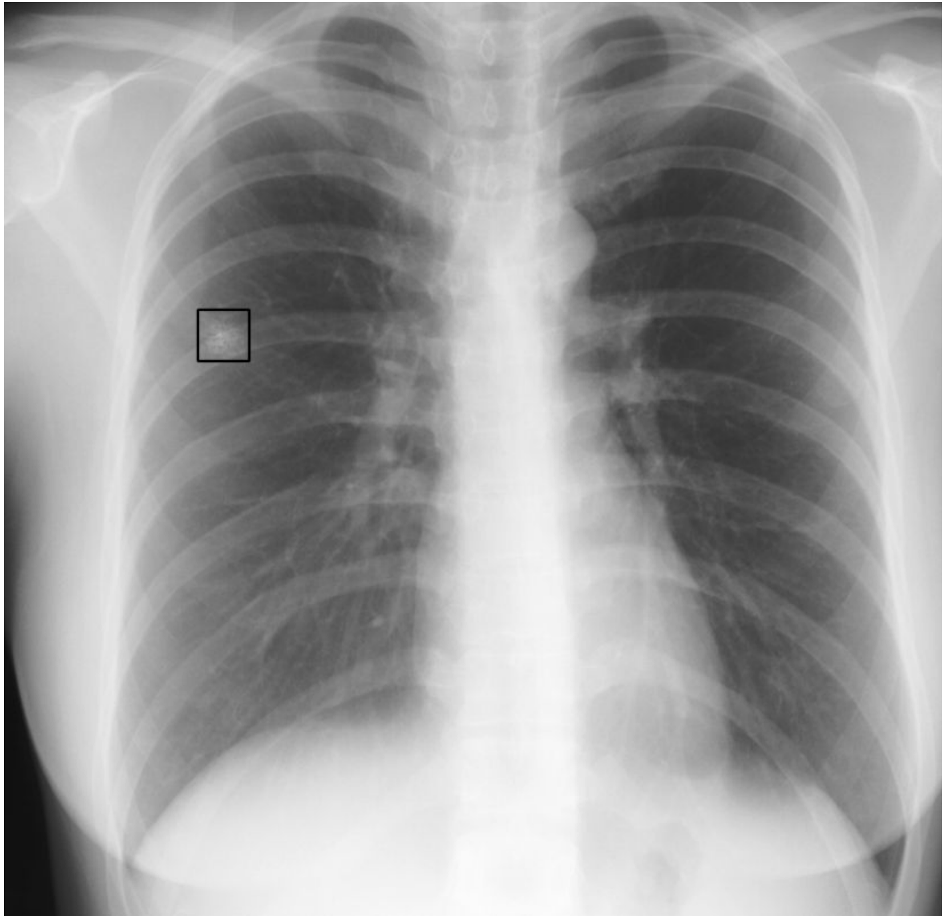
- Extract nodules from dataset, resize to 64*64, and store
- Extract vector subspace from the images and create a convex hull parameterized eigenvectors
- Randomly sample vectors from the hull and reshape to 64*64 for nodule reconstruction
- Use Gaussian mixture model to create a region proposal network
- Blend the generated nodule on a CT scan image using a Gaussian mask

Approach 1: Eigen Nodules

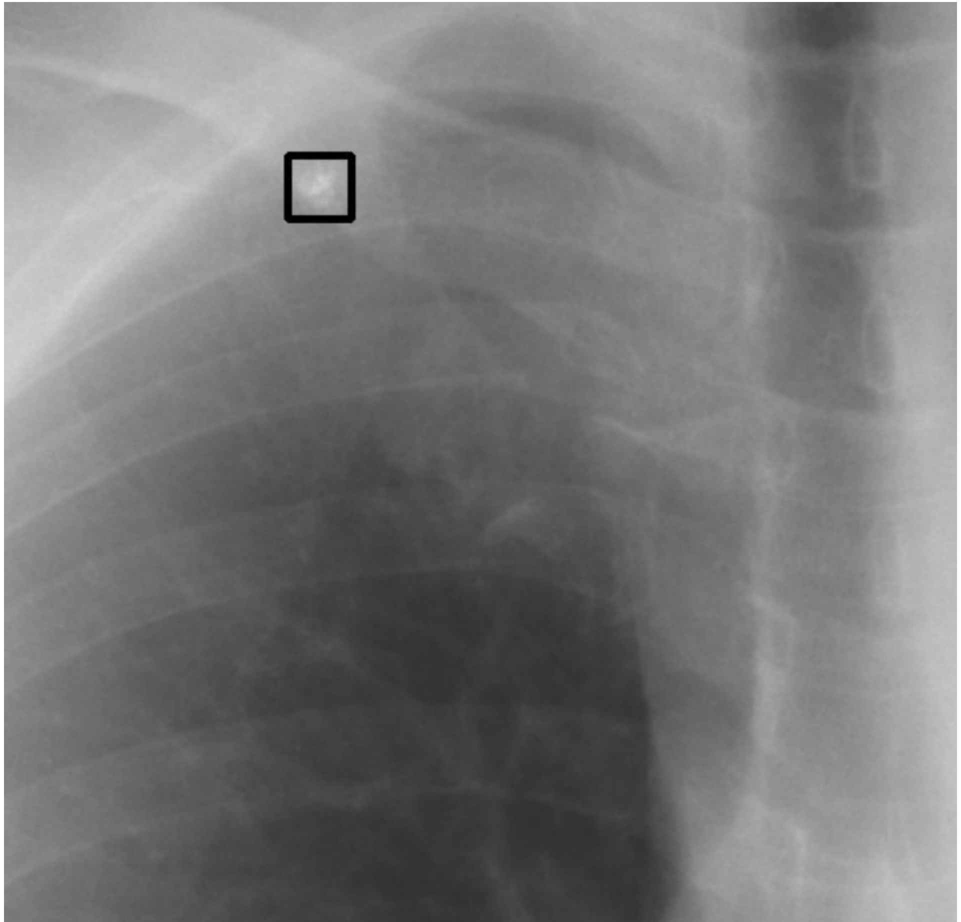


Approach 1: Eigen Nodules; results

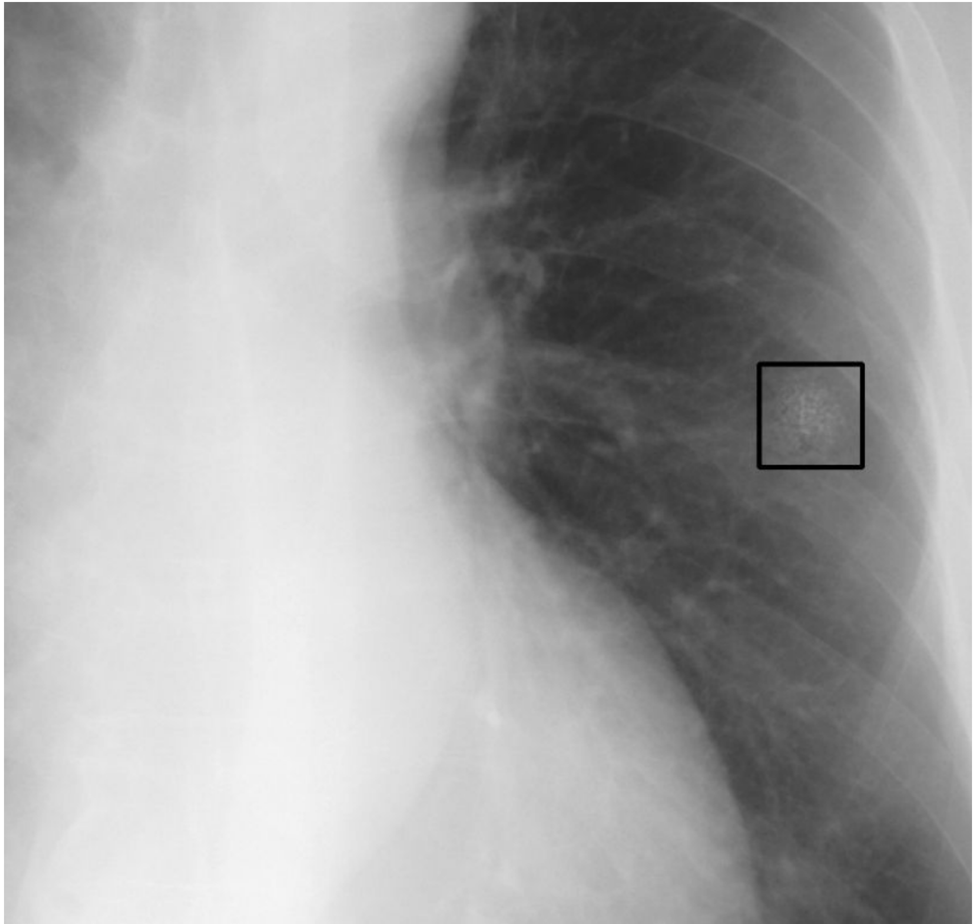
Approach 1: Eigen Nodules; results



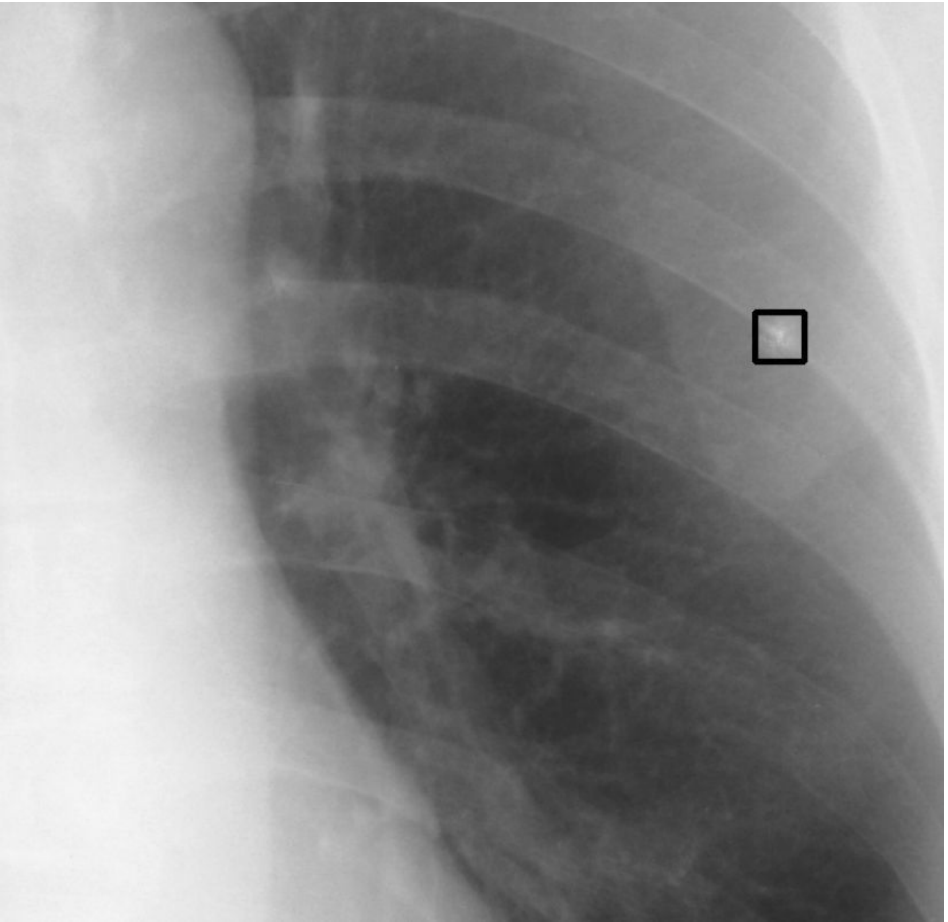
Approach 1: Eigen Nodules; results



Approach 1: Eigen Nodules; results

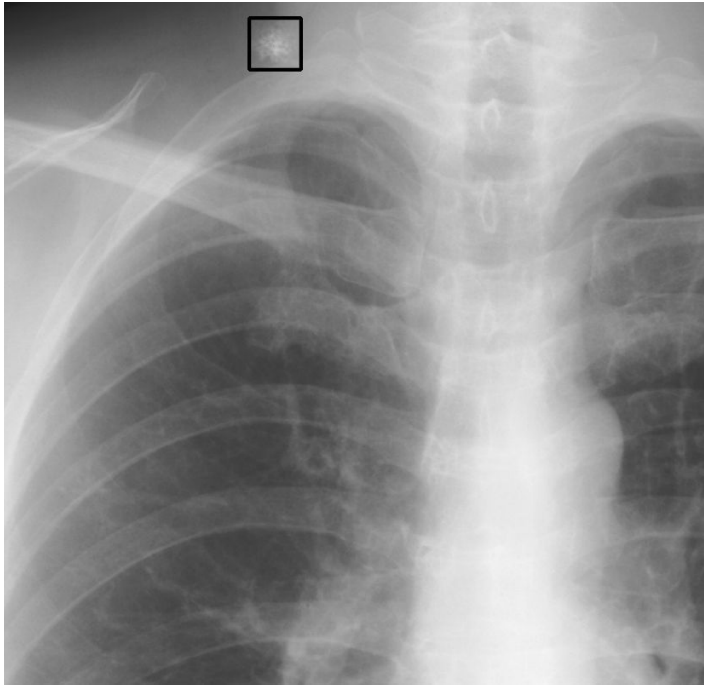


Approach 1: Eigen Nodules; results



Approach 1: Eigen Nodules; FLAWS

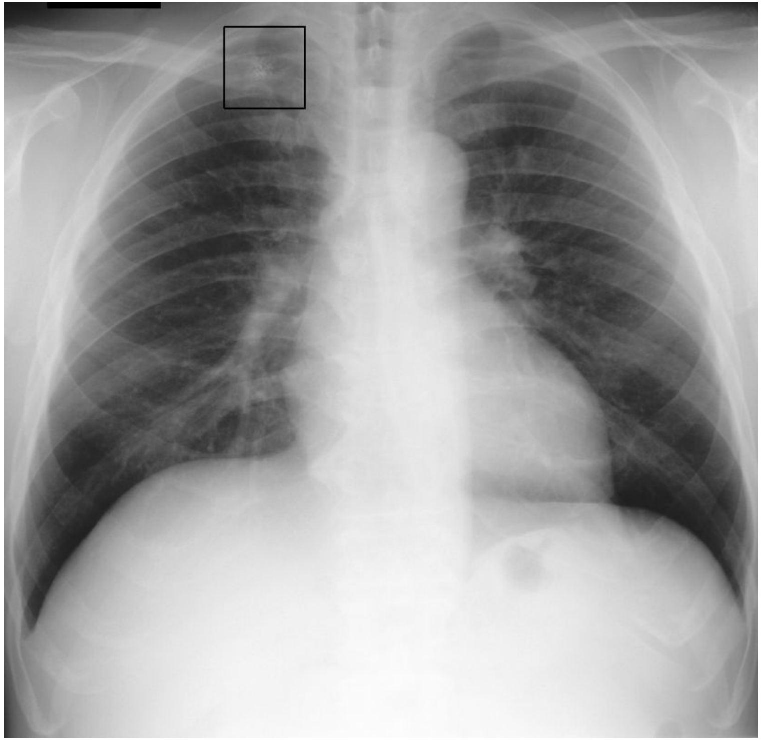
- Generated nodule may not always lie in the lung region.
- Eigenvectors may not form a convex hull
- Nodule generated on bone; cannot be differentiated
- Nodules of small size



Node generated outside the lung

Approach 1: Eigen Nodules; Flaws

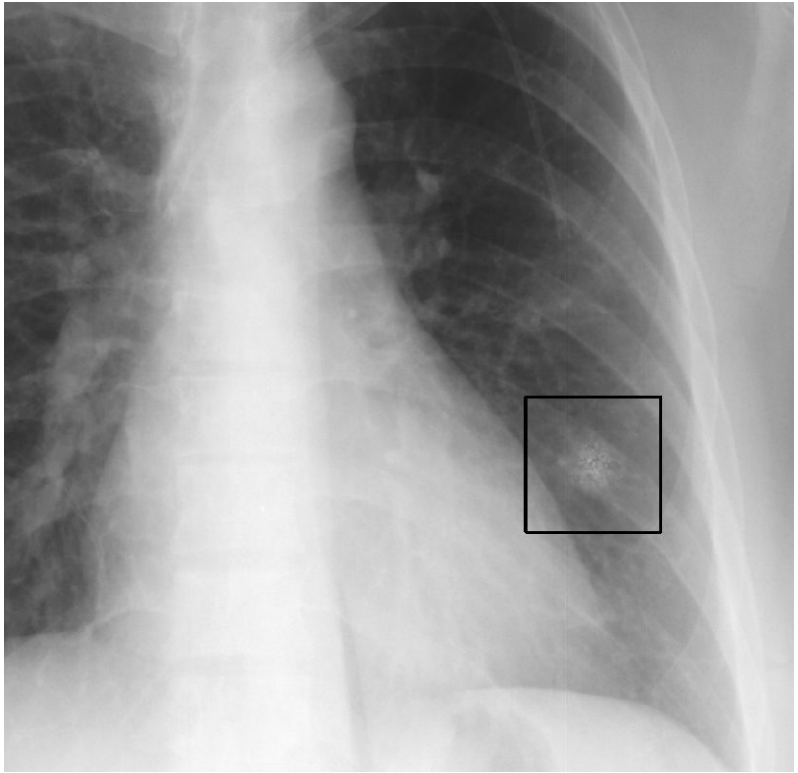
- Generated nodule may not always lie in the lung region.
- Eigenvectors may not form a convex hull
- Nodule generated on bone; cannot be differentiated
- Nodules of small size



Nodule generated on the bone

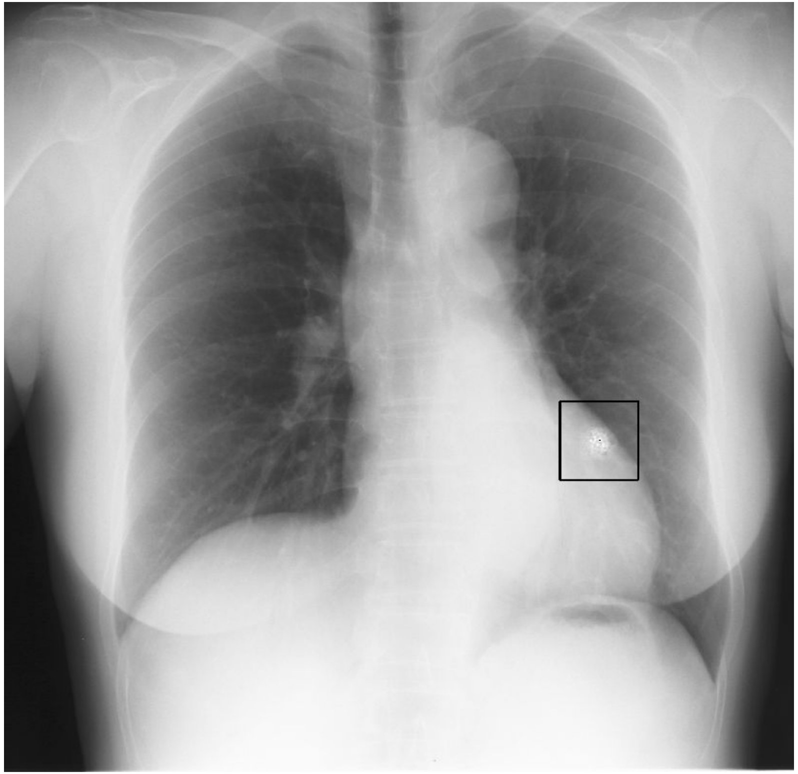
Approach 1.2: Eigen Nodules w/ Clahe;

- Extracted nodule patches are enhanced using CLAHE before PCA



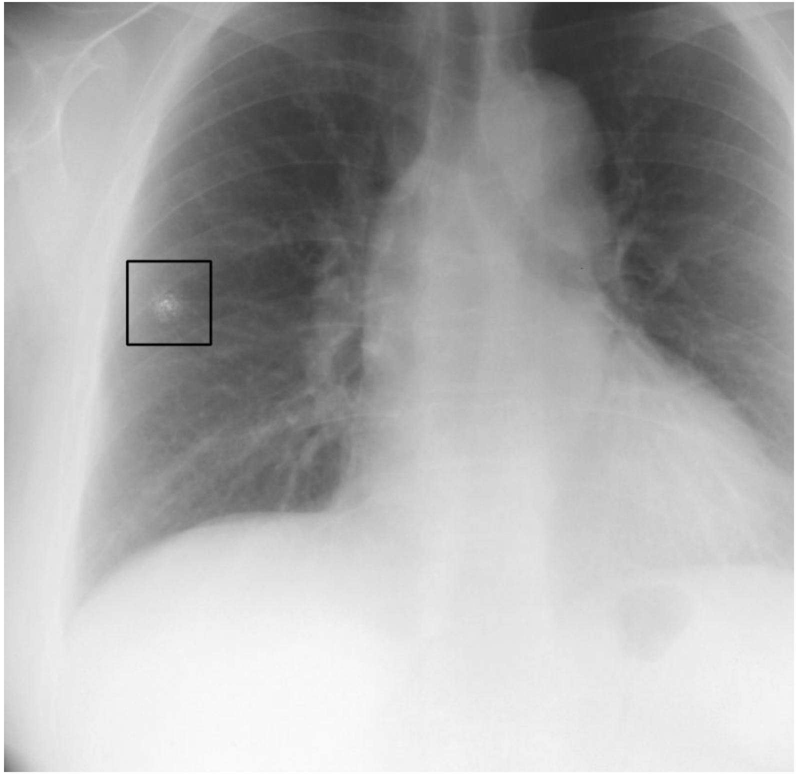
Approach 1.2: Eigen Nodules w/ Clahe;

- Extracted nodule patches are enhanced using CLAHE before PCA



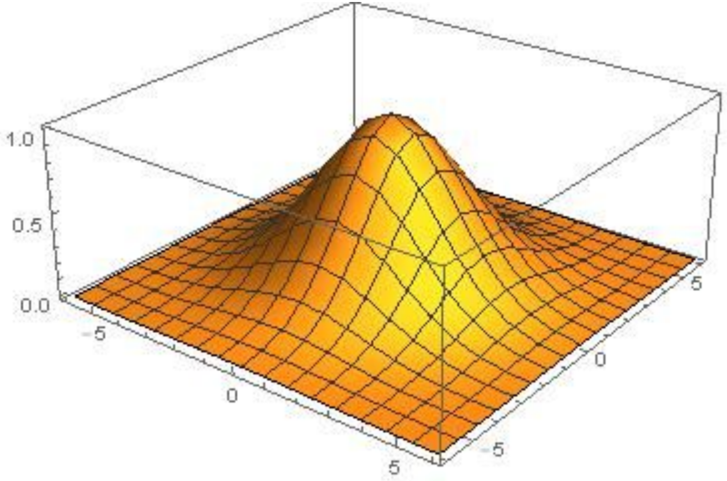
Approach 1.2: Eigen Nodules w/ Clahe;

- Extracted nodule patches are enhanced using CLAHE before PCA



How was blending done?

- Gaussian mask
- Local intensity informed alpha blending



Approach 2: GANs

- The generator network synthesizes chest X-ray images with nodules, attempting to mimic real data,
- Discriminator evaluates these generated images against authentic ones.
- This adversarial process refines the generator's ability to produce realistic images.
- GANs contribute to data augmentation strategies by generating diverse X-ray images

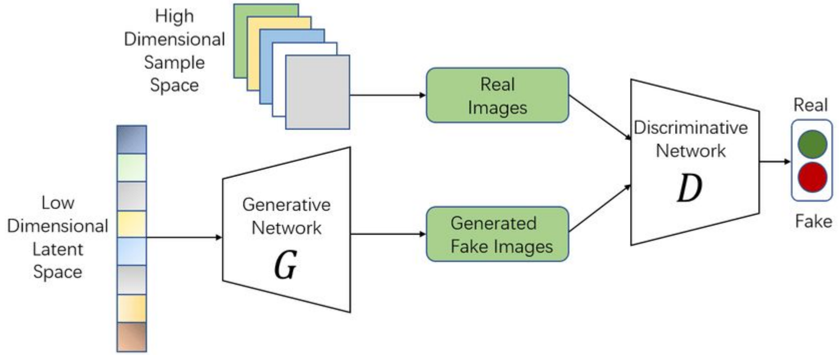
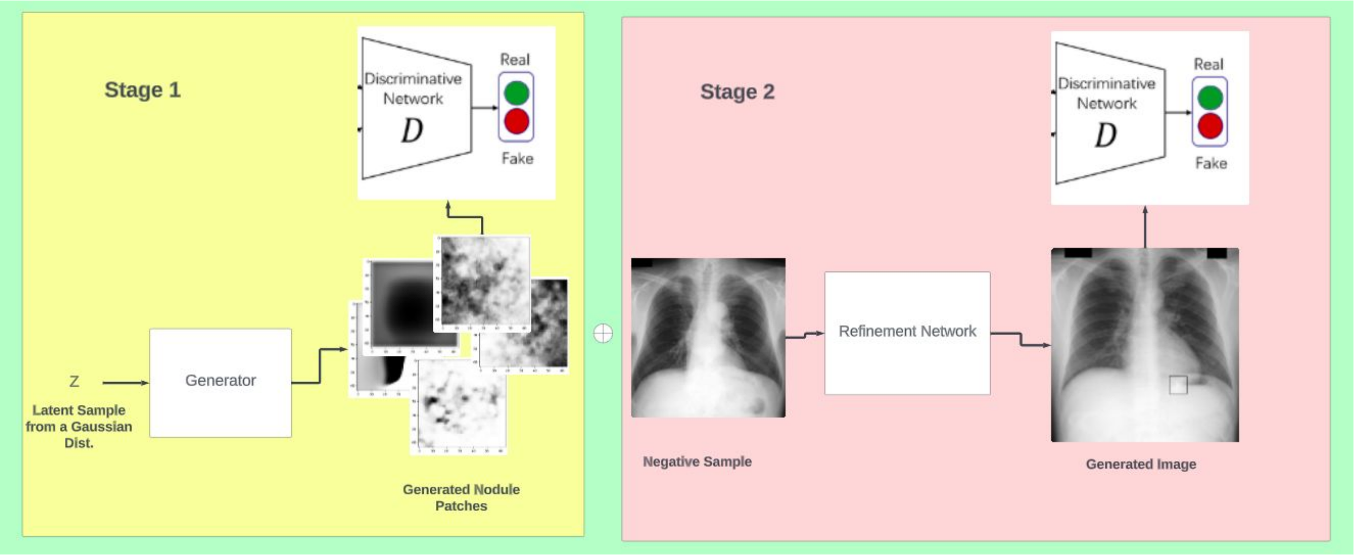


Figure illustrating GANs

Approach 2: GANs

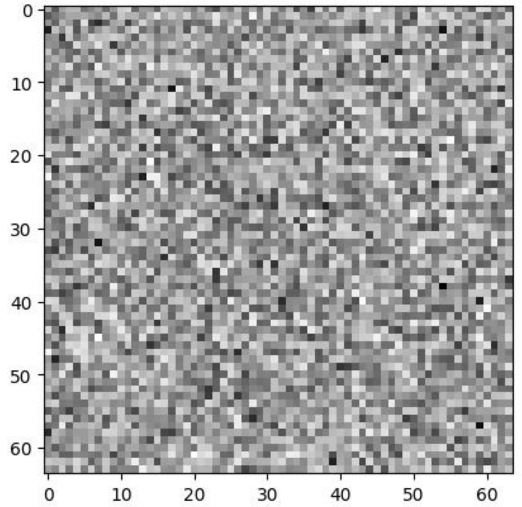
- Modelled distribution of nodules
- Once we have a nodule generated, overlaid it on CXR by regressing location, size
- Decided to blend nodule using another refinement network.
  - Proposed an RPN



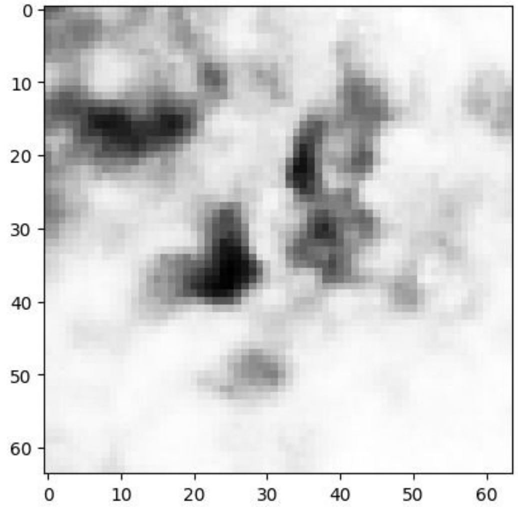
Our Pipeline

Approach 2.1: Stage 1; training

- ANN based generator
  - Noisy generation
  - Unable to capture pattern generation
- Solution?
  - Gradient Minimization loss included;



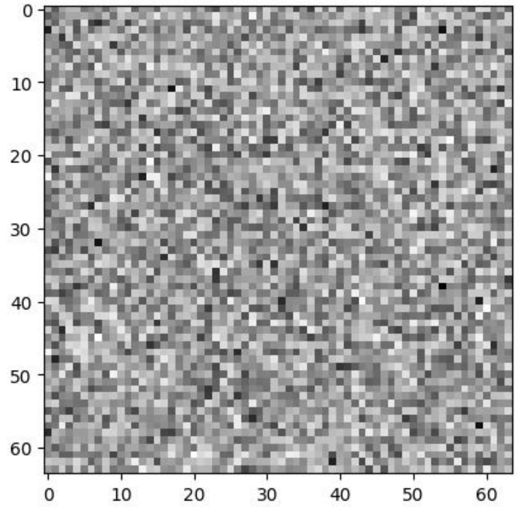
before



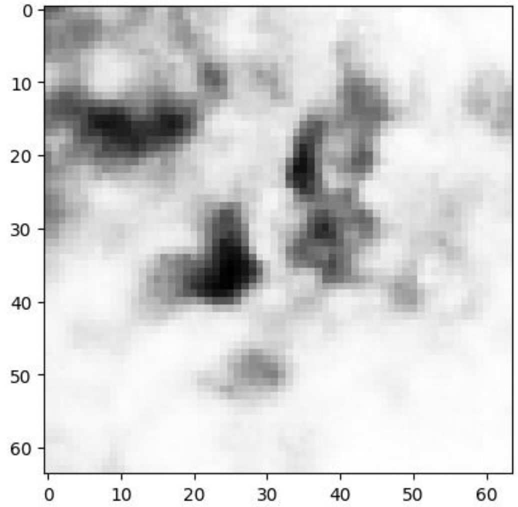
after

Approach 2.1: Stage 1; training

- ANN based generator
  - Noisy generation
  - Unable to capture pattern generation
- Solution?
  - Gradient Minimization loss included; **can give even better results**



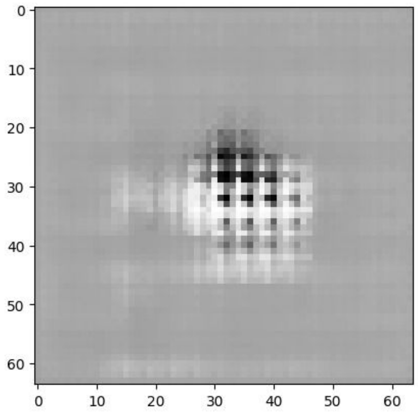
before

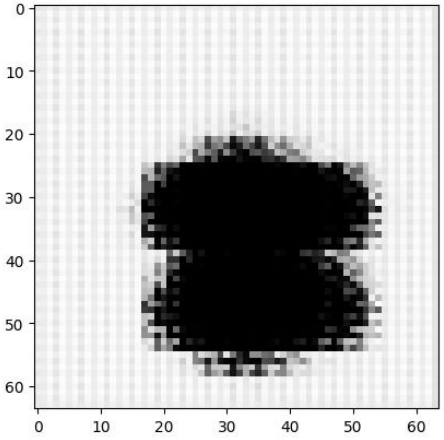


after

Approach 2.2: Stage 1; training

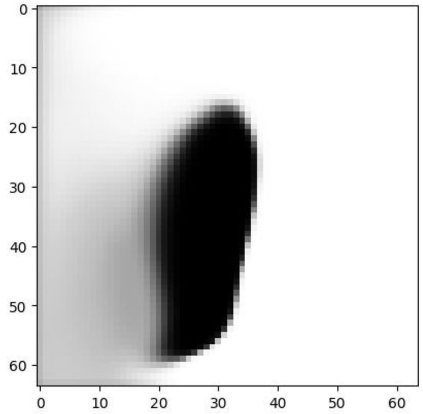
- Result still noisy; only diffused structures captured
- Solution?
  - Transpose Convolutional based generator
- Issue*
  - Checked artefacts

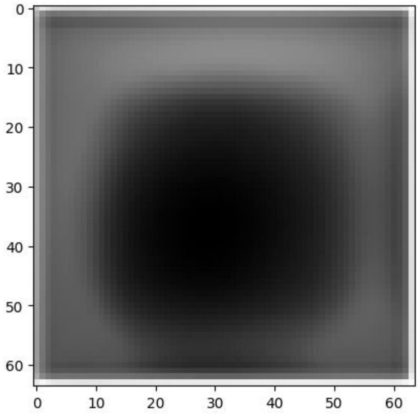




Approach 2.3: Stage 1; training

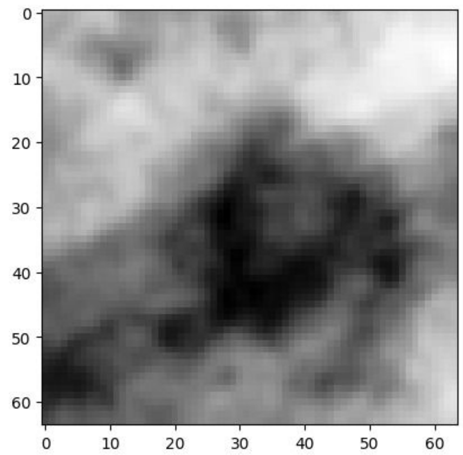
- Solution?
  - Replace **Transpose Convolutional based generator w/ Upsampling & convolution layer**
- Bilinear interpolation used for upsampling
- Issues*
  - Learned noises too smooth
  - Diffused patterns not learned

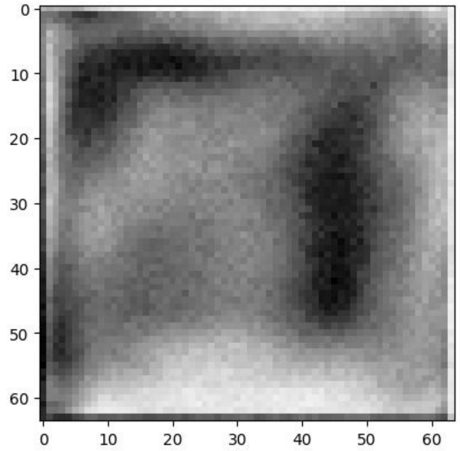




Approach 2.4: Stage 1; training

- Solution?
  - Weighted average of both the generators in **v2.1&v2.3**
- Domain knowledge needed for further explainability
  - All the approaches were trained in GAN setup

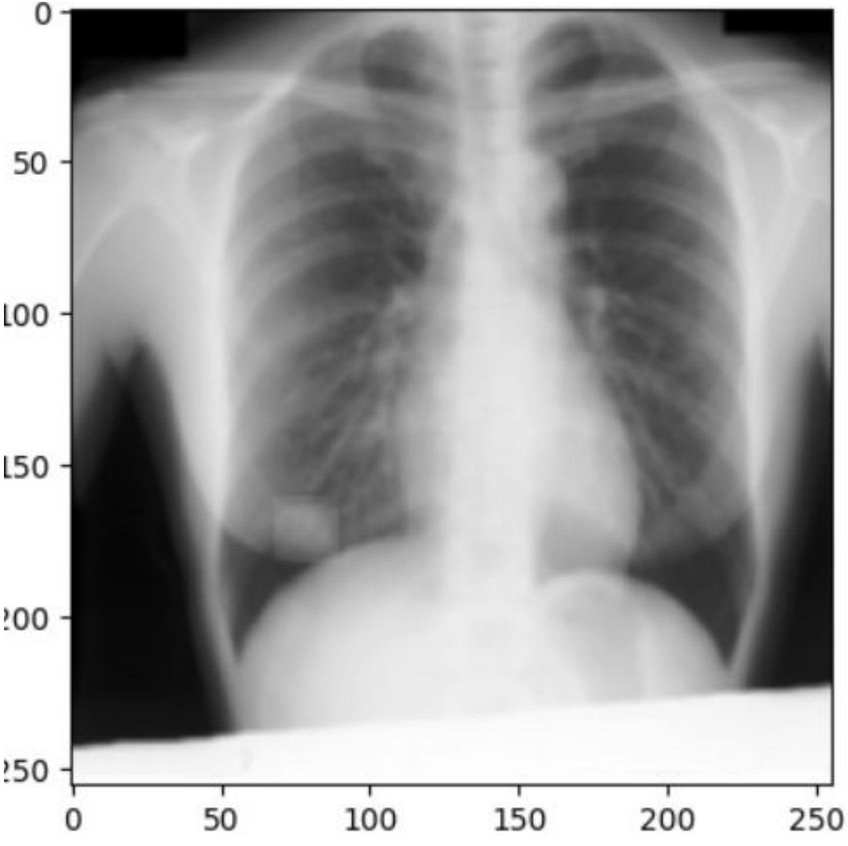




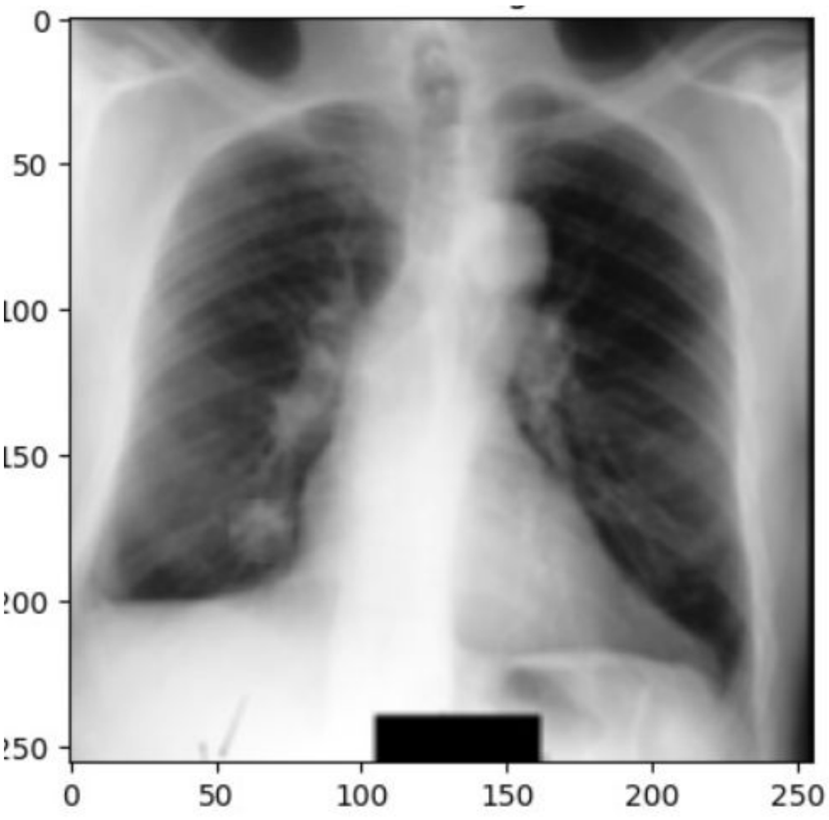
Approach 2.4: Stage 2; training

- Refinement network for blending
  - Nodule patch generated was reshaped using size regressed via Gaussian Mixture Model
  - Patch & Neg. sample sent via 2 channels into a shallow CNN, trained in GAN setup

Approach 2.4: Results

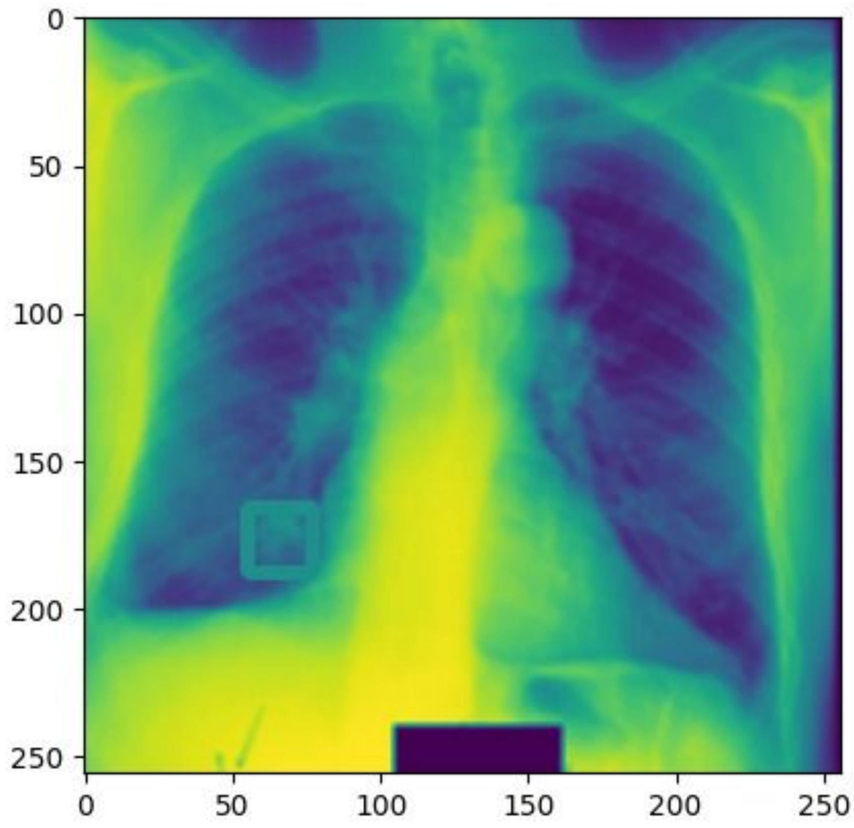


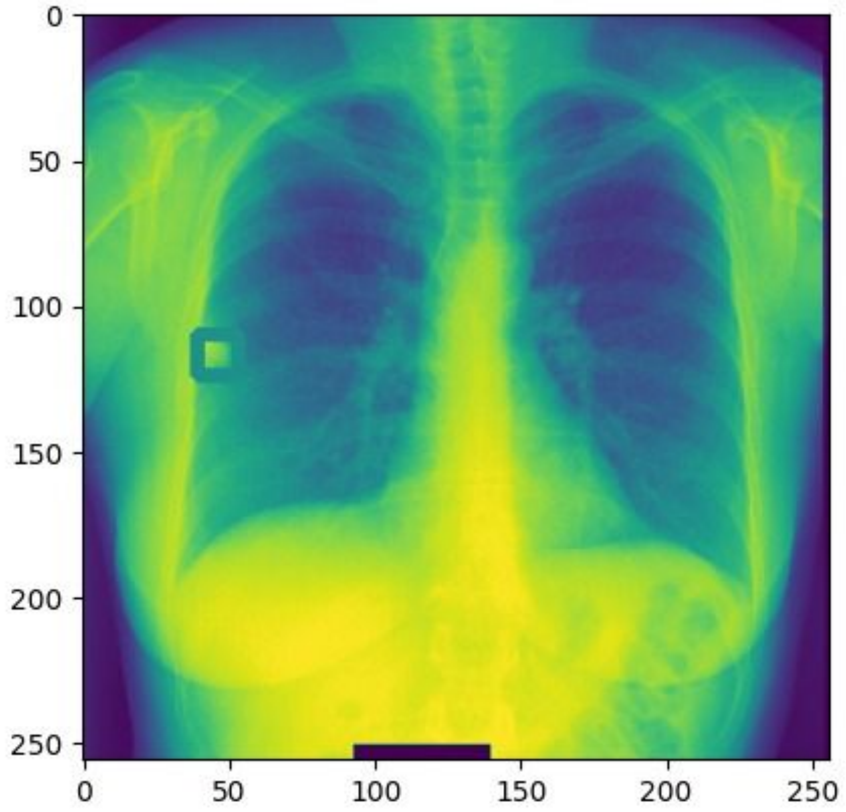
Without refinement



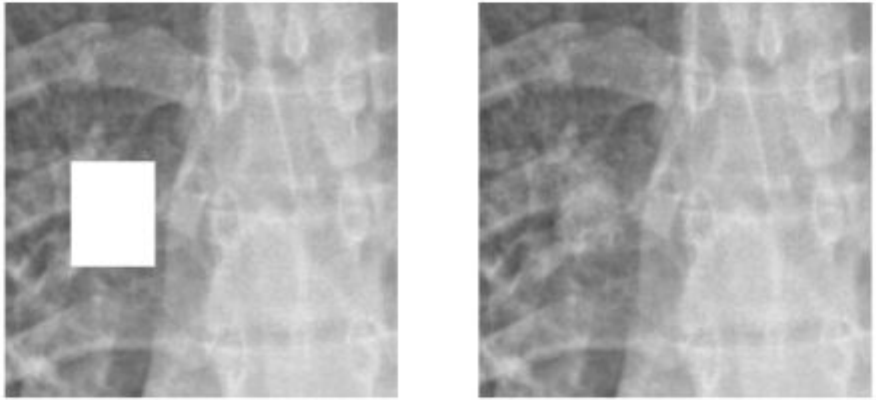
With refinement

Approach 2.4: Results

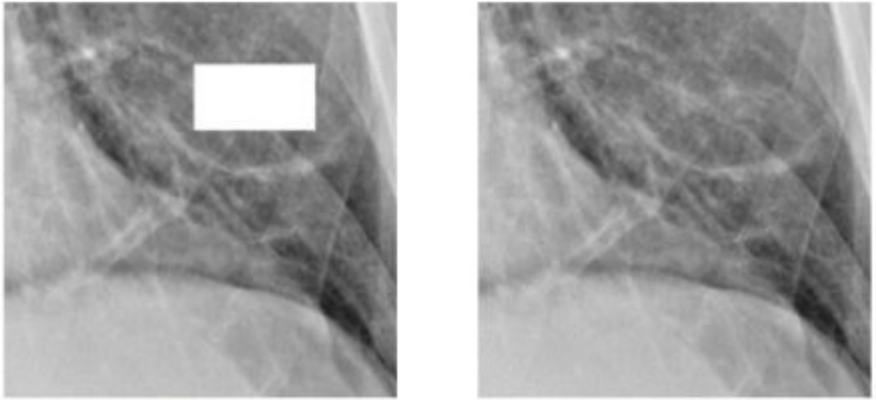




Vs SOTA



(c) (d)



Related works;

- Attempted bone suppression to extract better nodules for Eigen-nodules
  - Outdated repo; could not make work
- Disentanglement for CXR
  - Was trained on MNIST; no pretrained model
  - Was taking huge compute, time to get trained
- Ideas that could not come to fruition
  - CRFill- wanted to tinker around
  - Diffusion Models- Compute issues

Thank you :)