Medical Image Analysis

Angshuman Paul

Assistant Professor

Department of Computer Science & Engineering

Annotation-efficient DL for MedIA

Learning a New Concept: Human vs Computer

- Human learning
  - Usually requires only a few instances
  - May efficiently learn visual appearance from semantic description
- Machine Learning (especially deep learning)
  - May require thousands (if not millions) of instances
  - Difficult to train using semantic information for visual understanding

What is Annotation-efficient Machine Learning

- Learning new concepts from only a few or no labeled training data
- Mimicking the human cognitive ability
- Overcoming the curse of large annotation
- Small training set: Difficult to train deep models

Why Annotation-efficient Machine Learning in Medical Images?

- Learn new diseases from a few labeled examples
  - Rare disease diagnosis

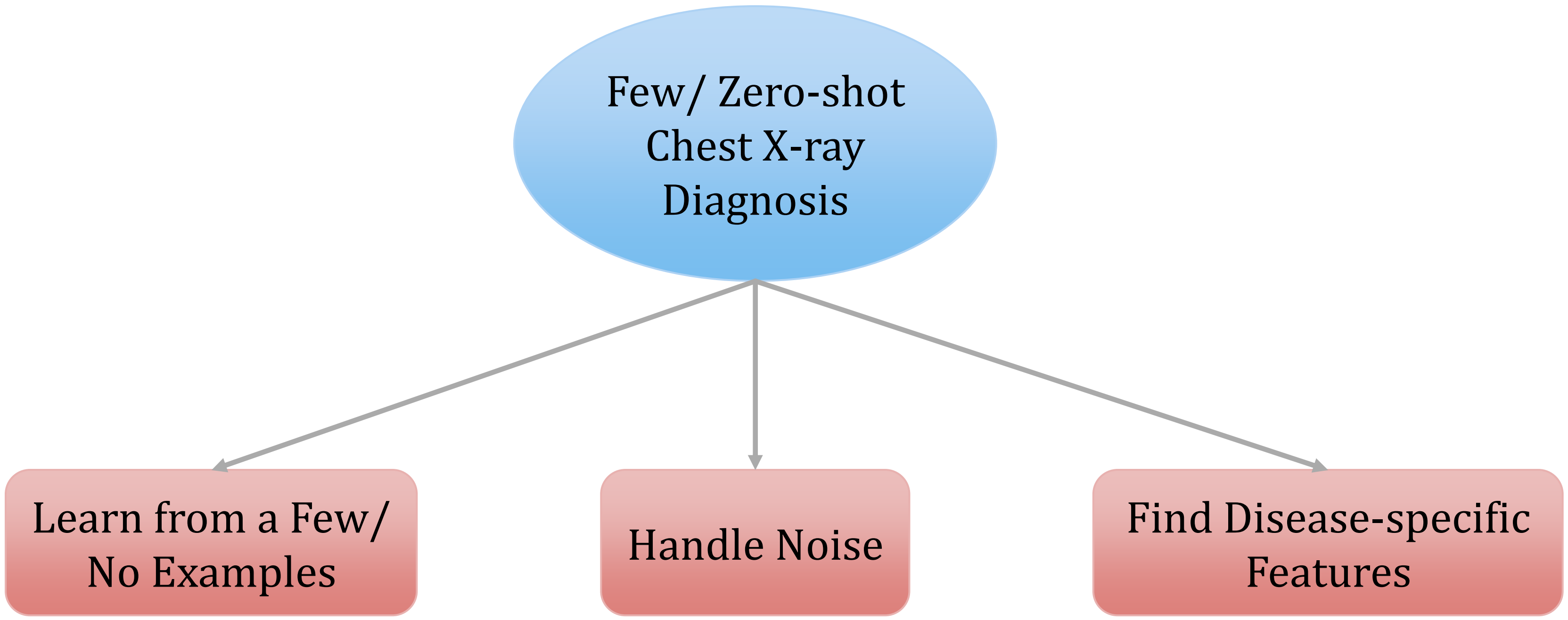
- Working solutions with small labeled dataset
  - Reducing the need of manual annotations
    - Expert-dependent
    - Time consuming

Annotation-efficient Machine Learning

Few-shot Learning

Zero-shot Learning

Annotation-efficient Machine Learning for Chest X-ray Diagnosis: Challenges

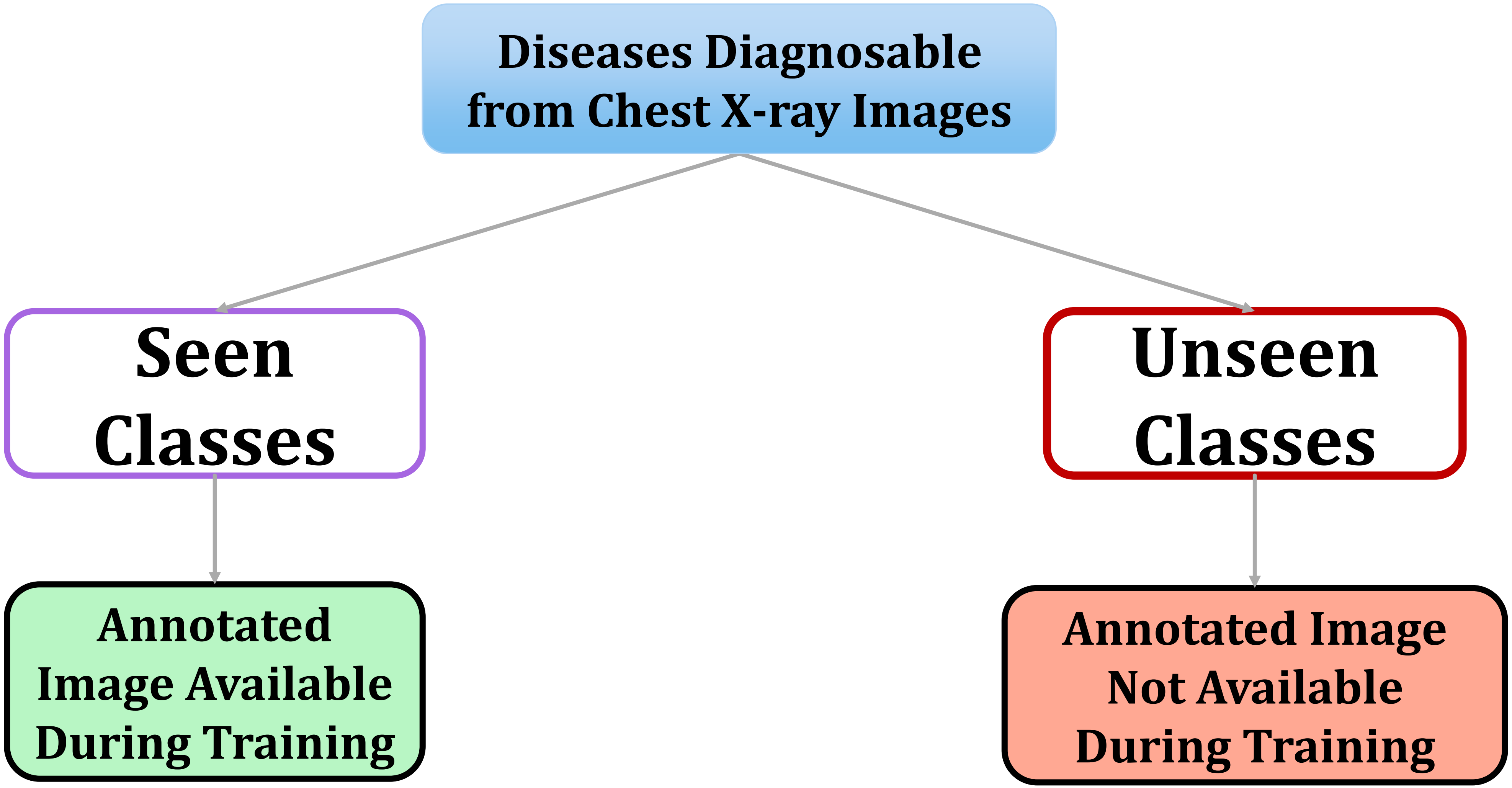


Zero-shot Chest X-ray Diagnosis

Zero-Shot Learning for Medical Images

- Common diseases: Large pool of annotated training images
- Rare diseases: few or no annotated training images
- Use annotated images of common diseases
- Learn new diseases from no examples: use **auxiliary information**
  - Rare disease diagnosis

Zero-shot Learning



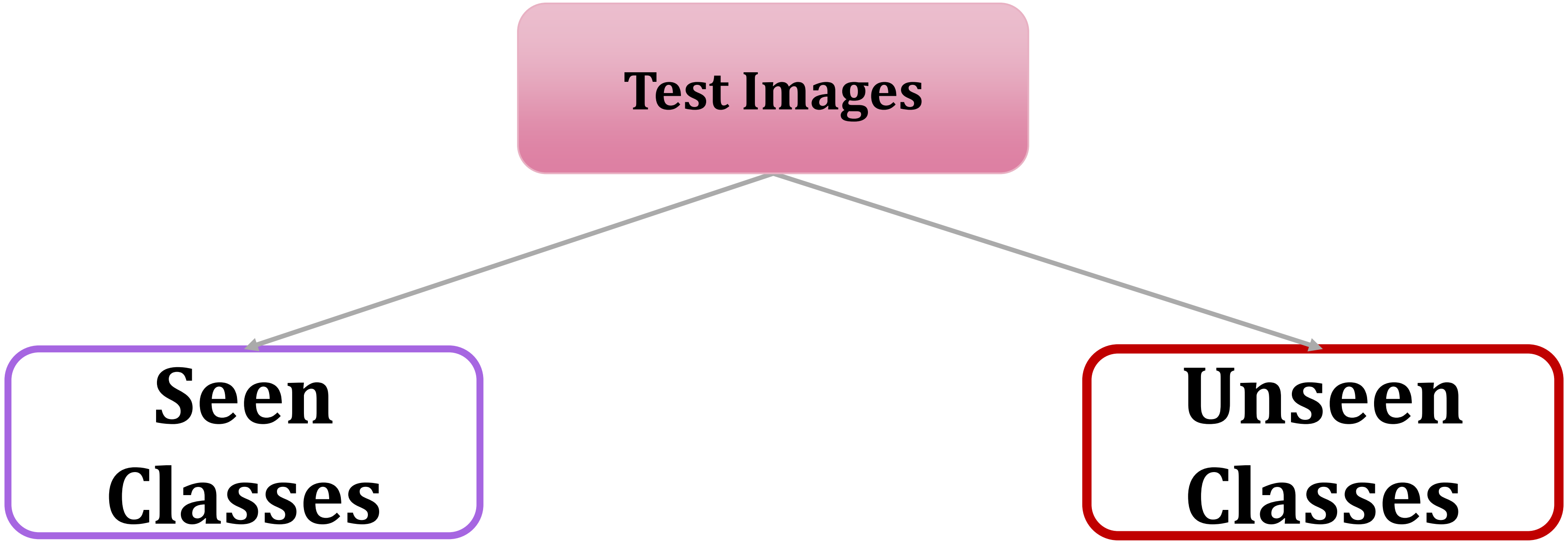
Zero-shot Learning: Training

Training Images

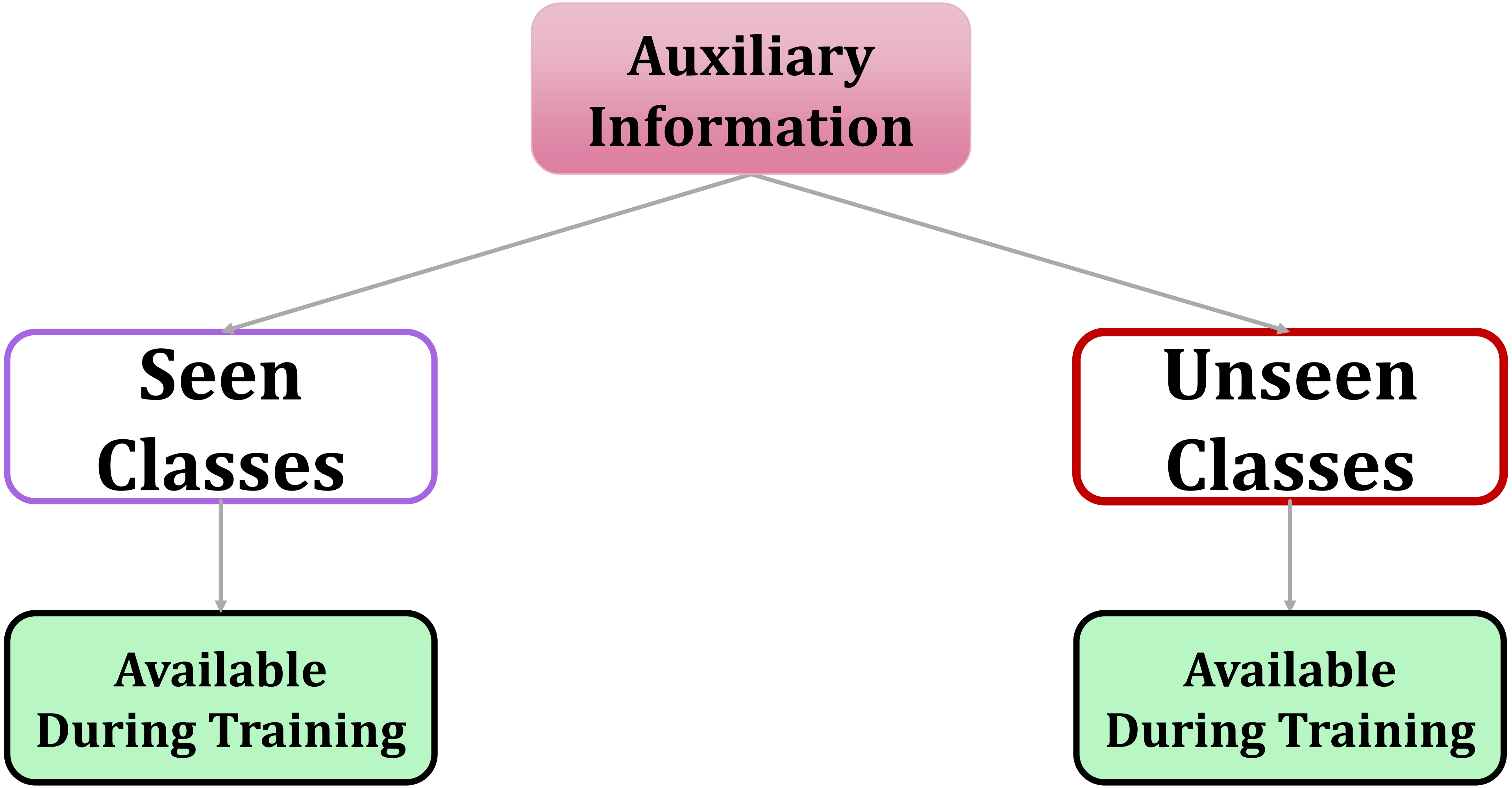
↓

Annotated Images of Seen Classes

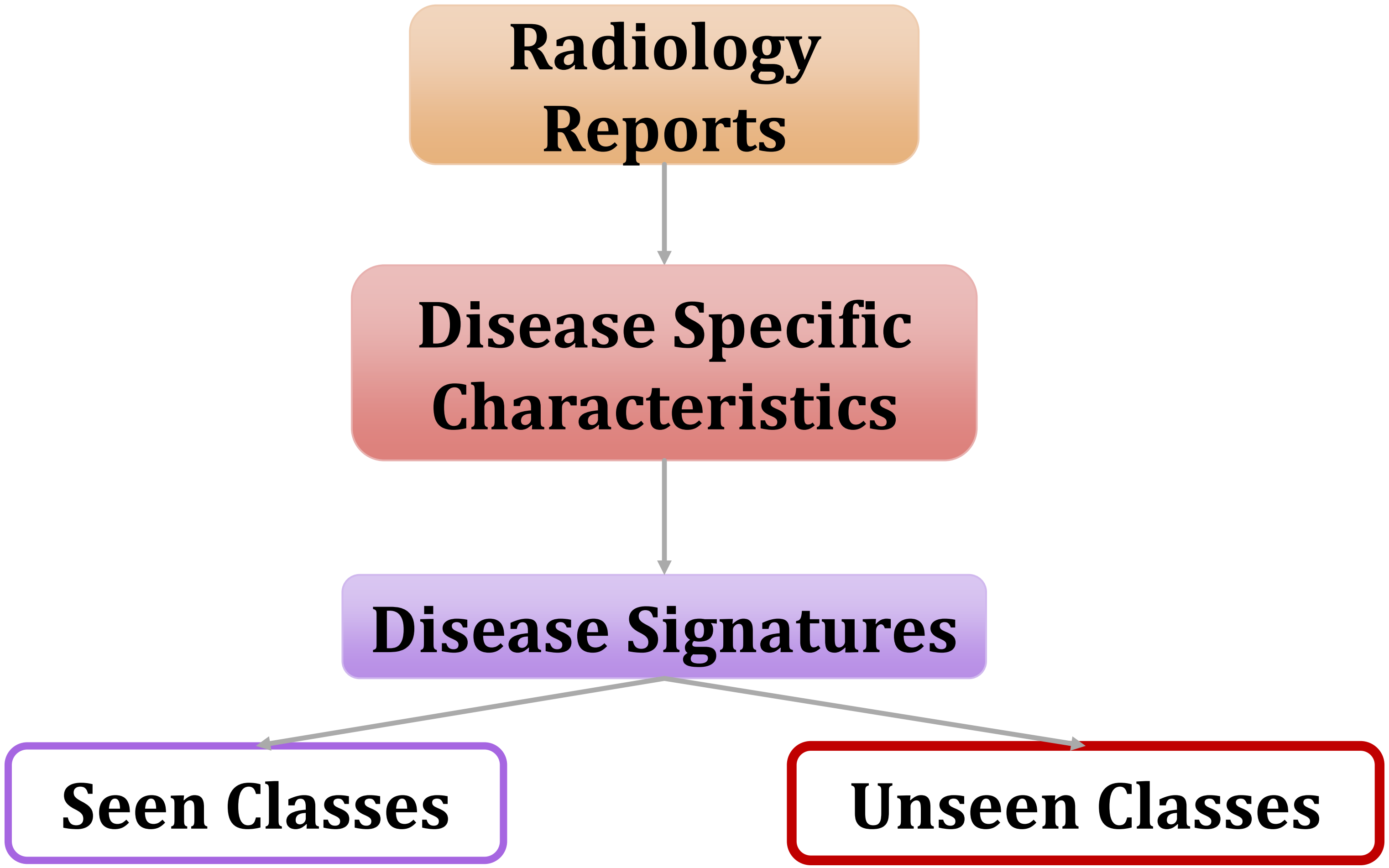
Generalized Zero-shot Learning: Testing



Zero-shot Learning

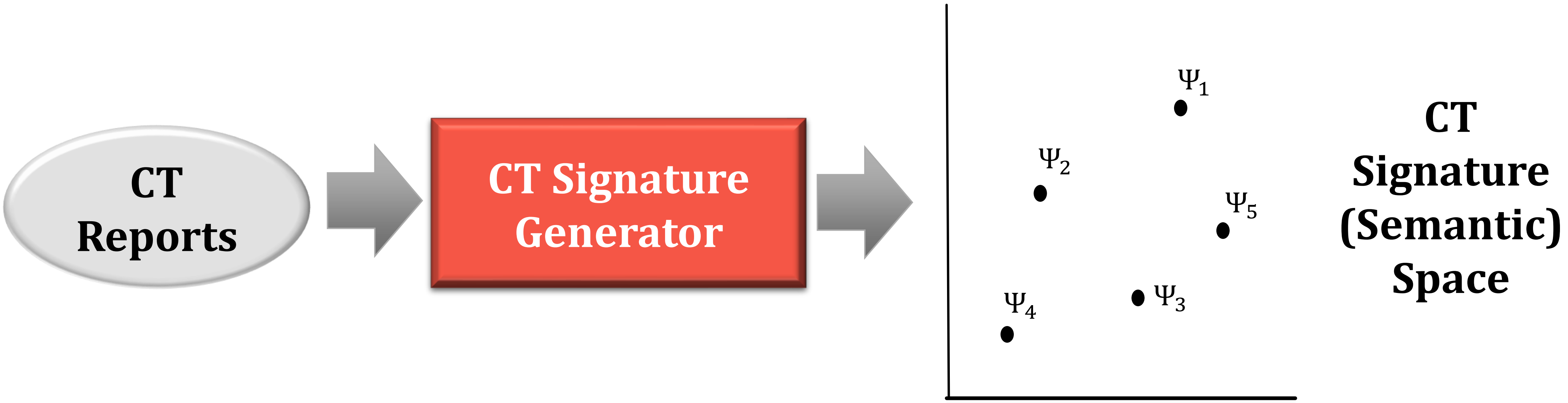


Auxiliary Information for Radiology Diagnosis



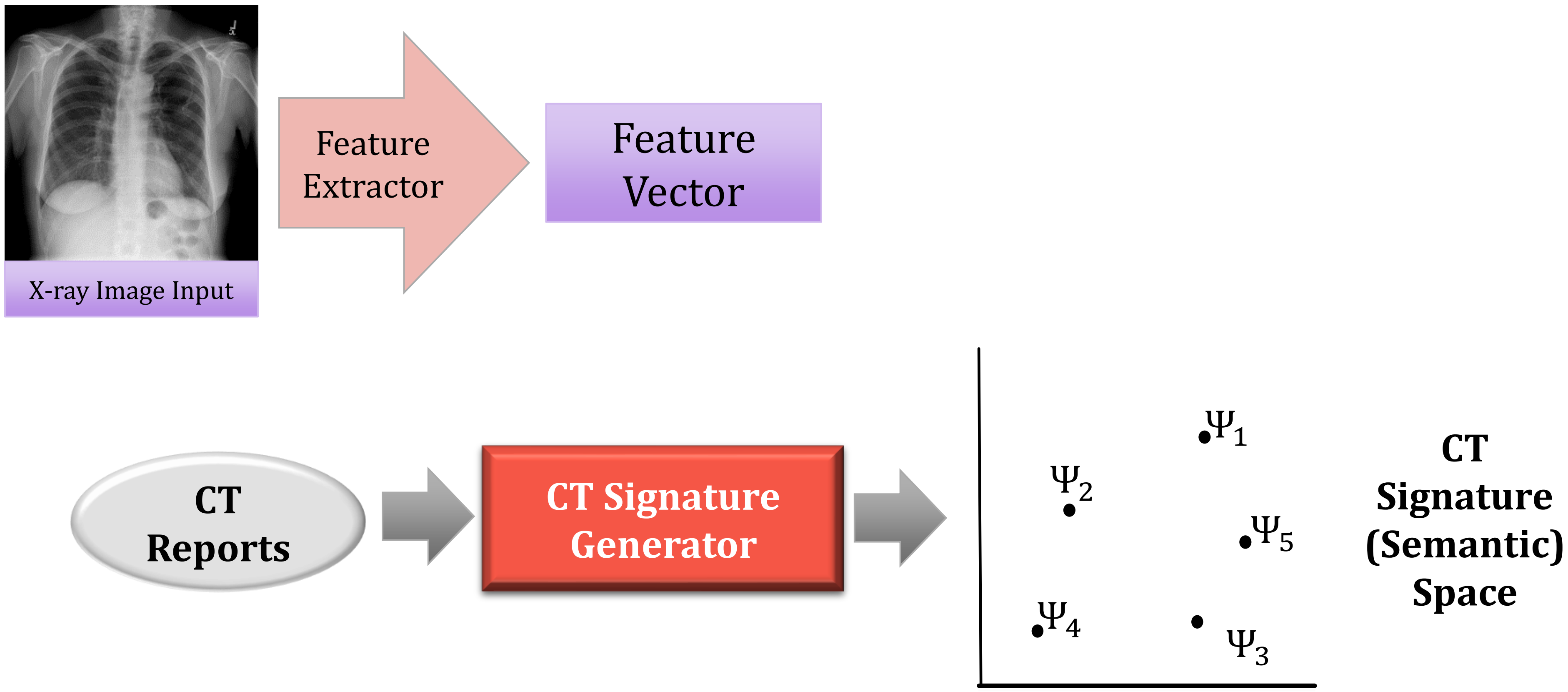
Disease Signatures

- Signature Generator: Intelligent Word Embedding (IWE)[1]

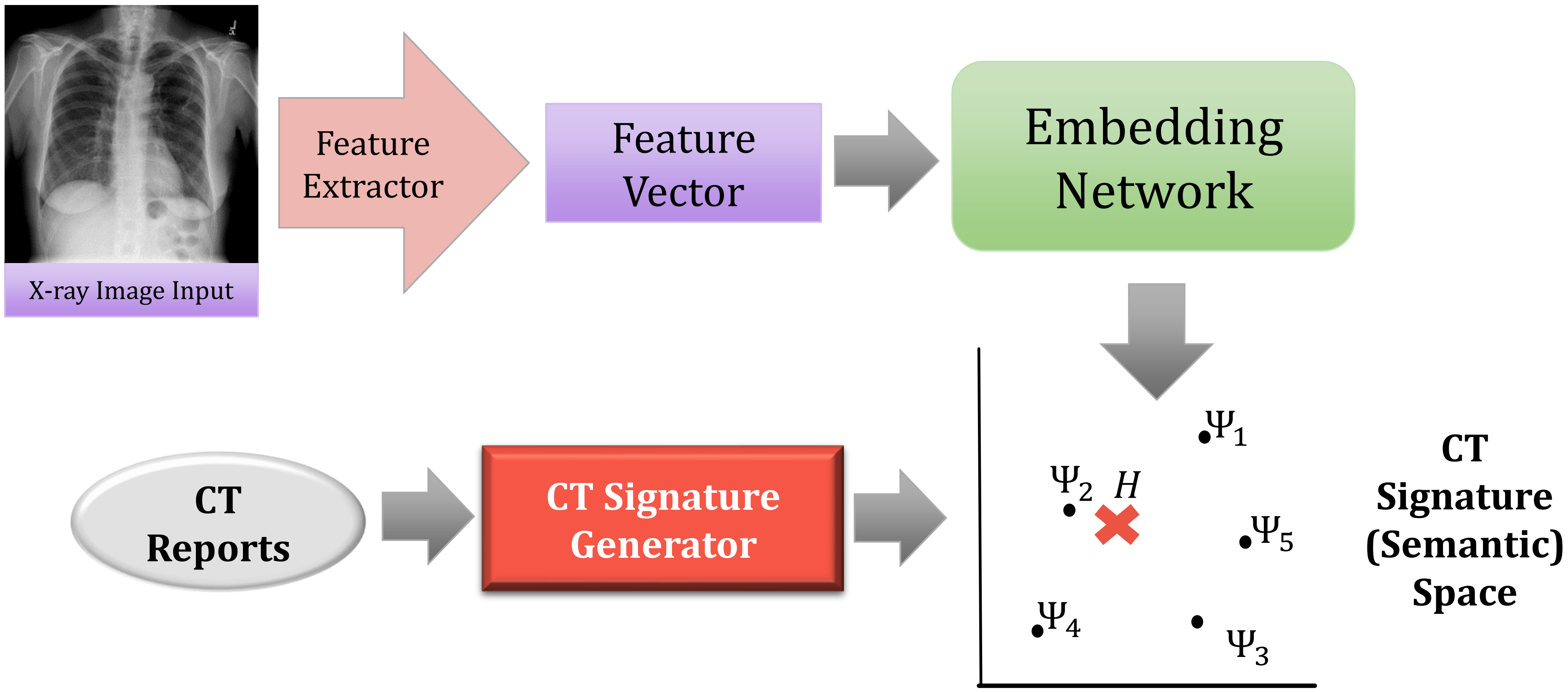


1. Banerjee *et al.*, “Intelligent word embeddings of free-text radiology reports”, (2017).

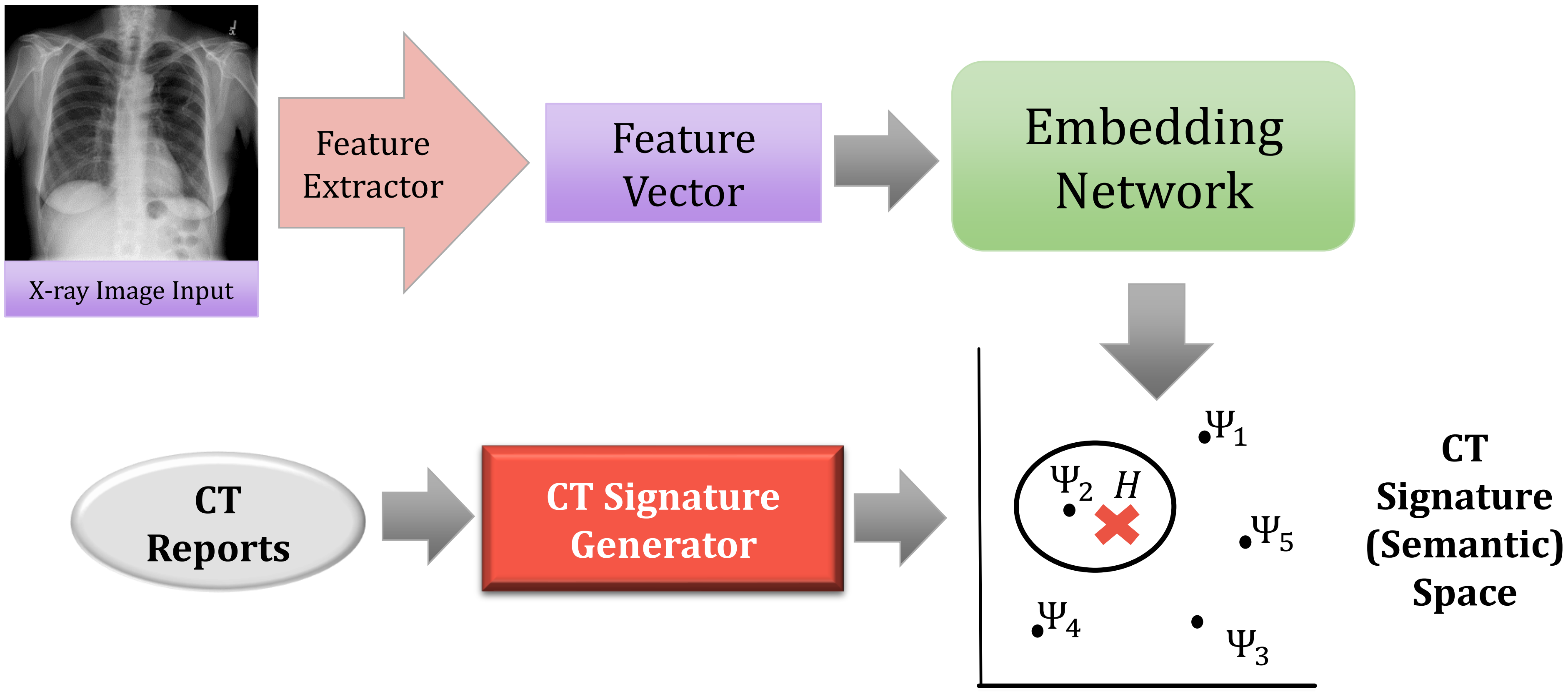
Zero-shot Diagnosis of Chest X-rays



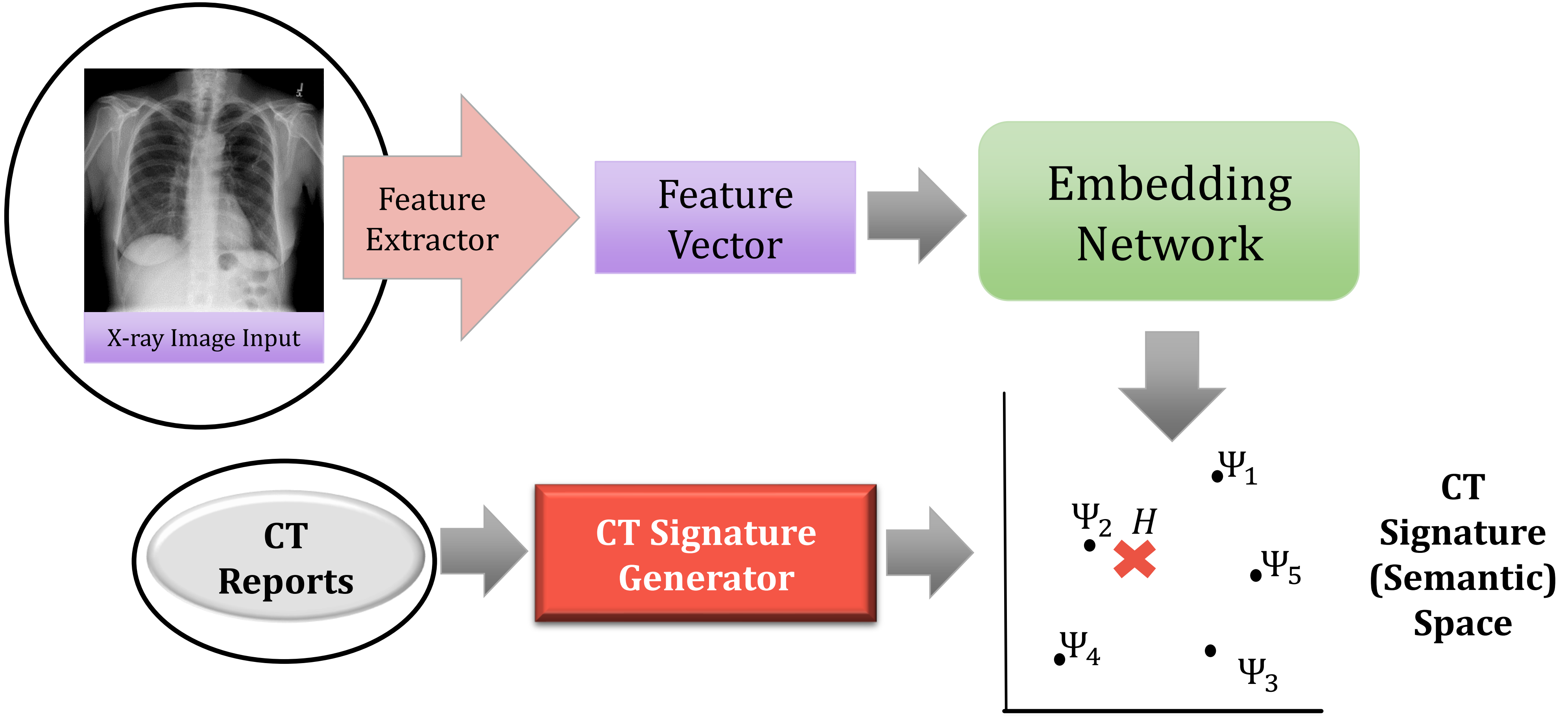
Zero-shot Diagnosis of Chest X-rays



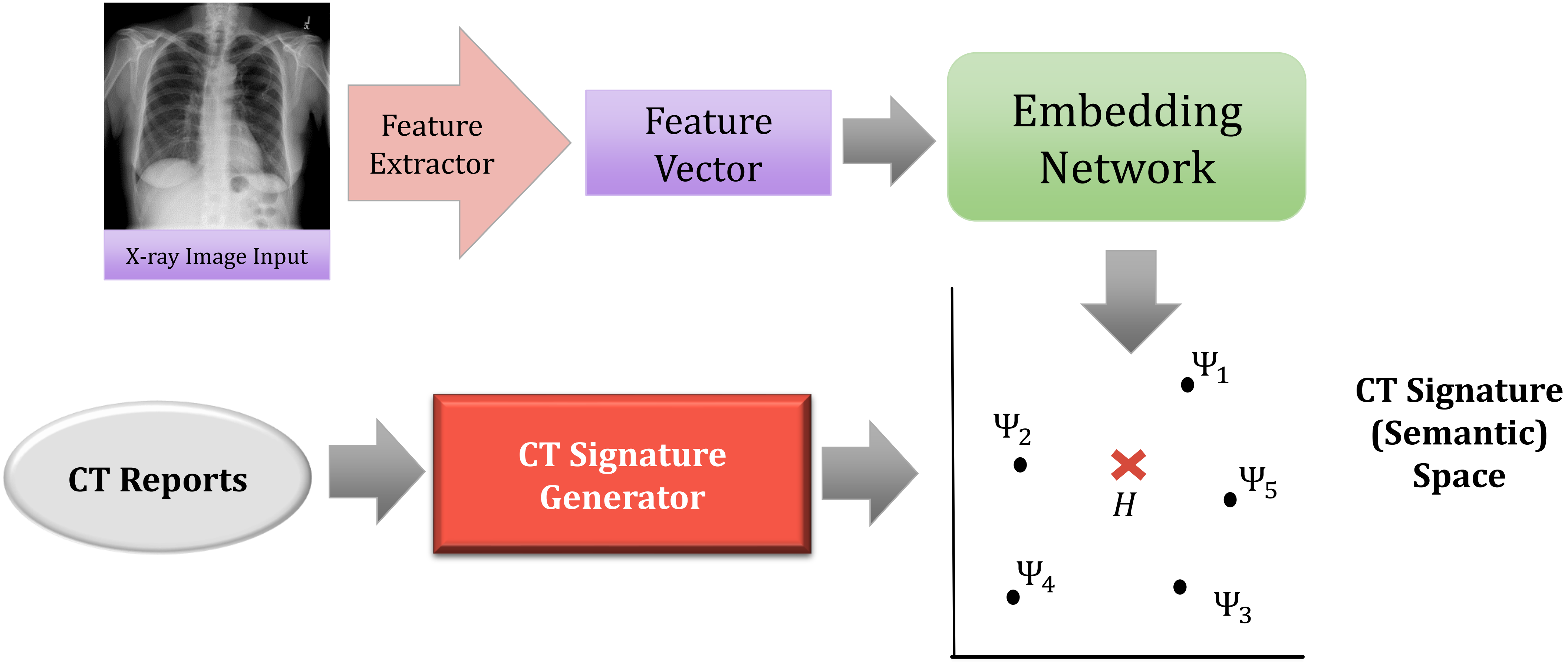
Zero-shot Diagnosis of Chest X-rays



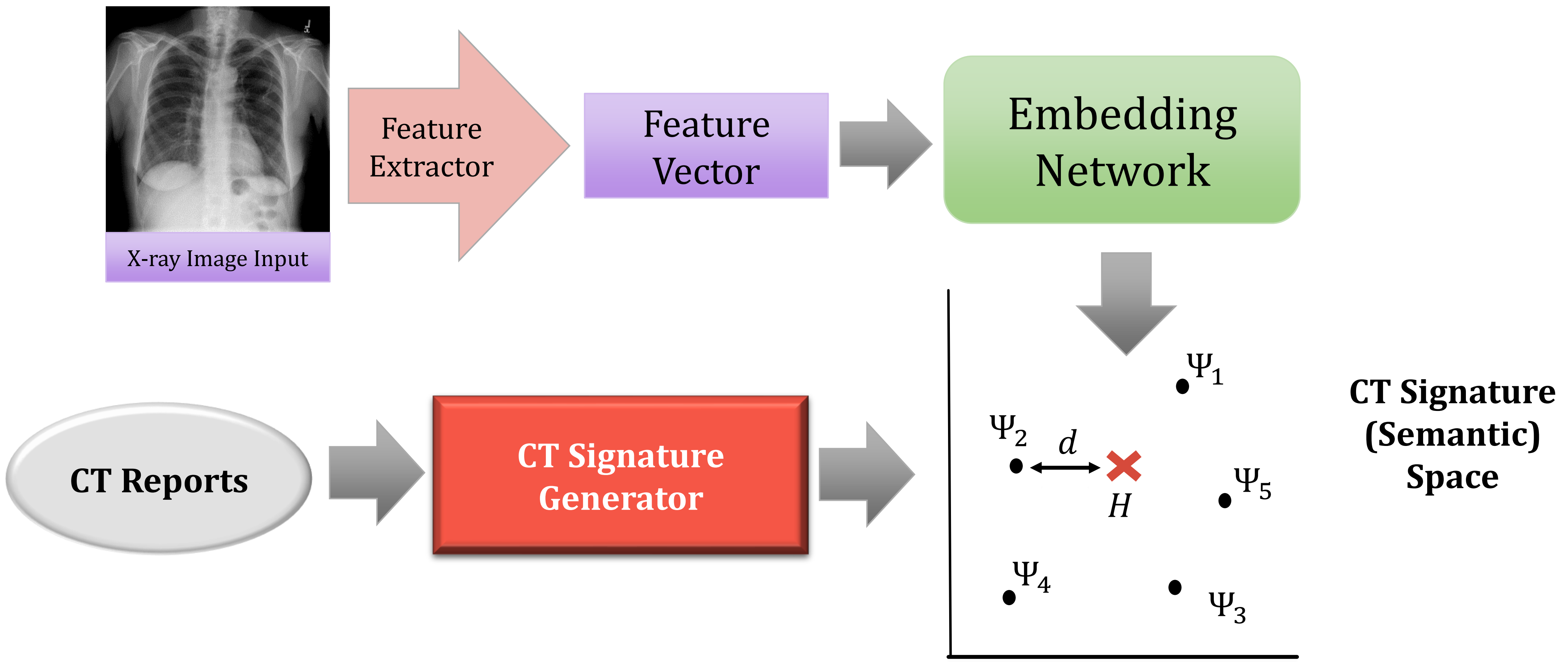
Cross-Modality Semantic Embedding



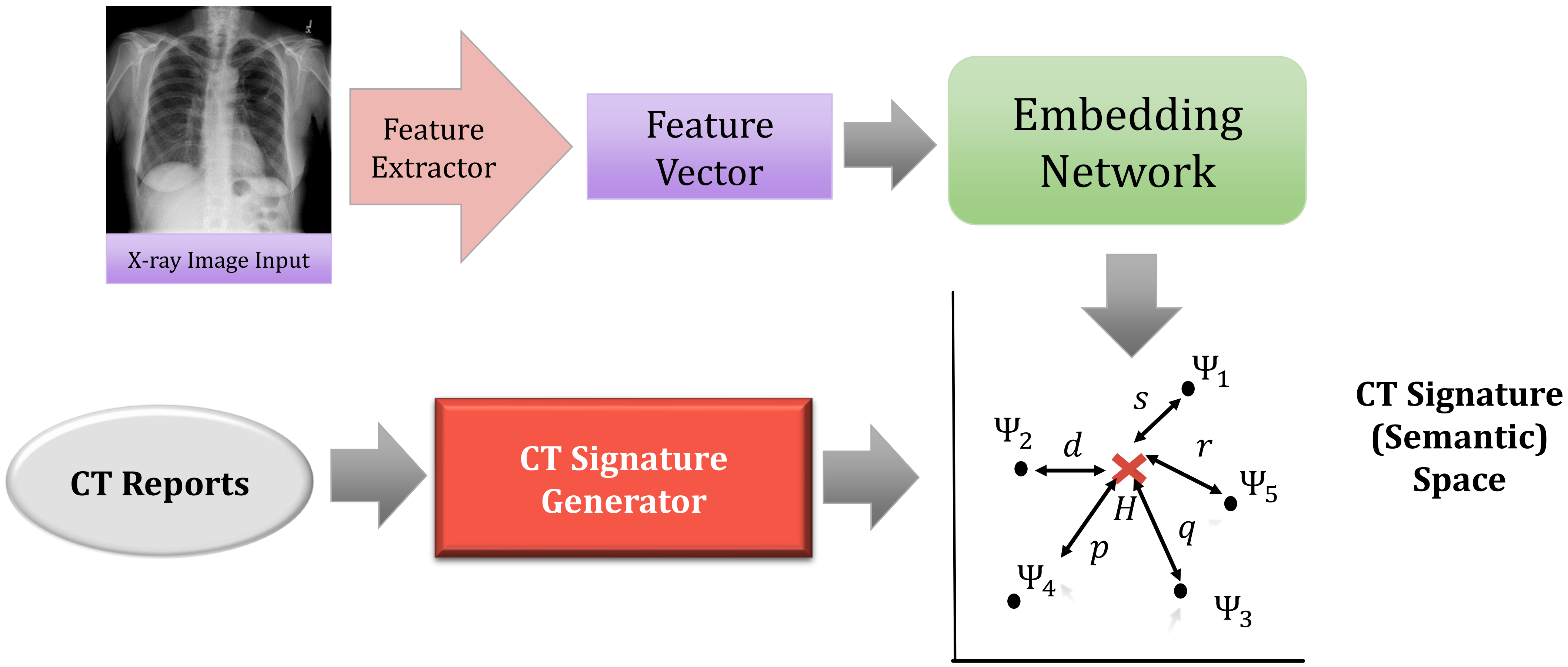
Meaningful Semantic Embedding



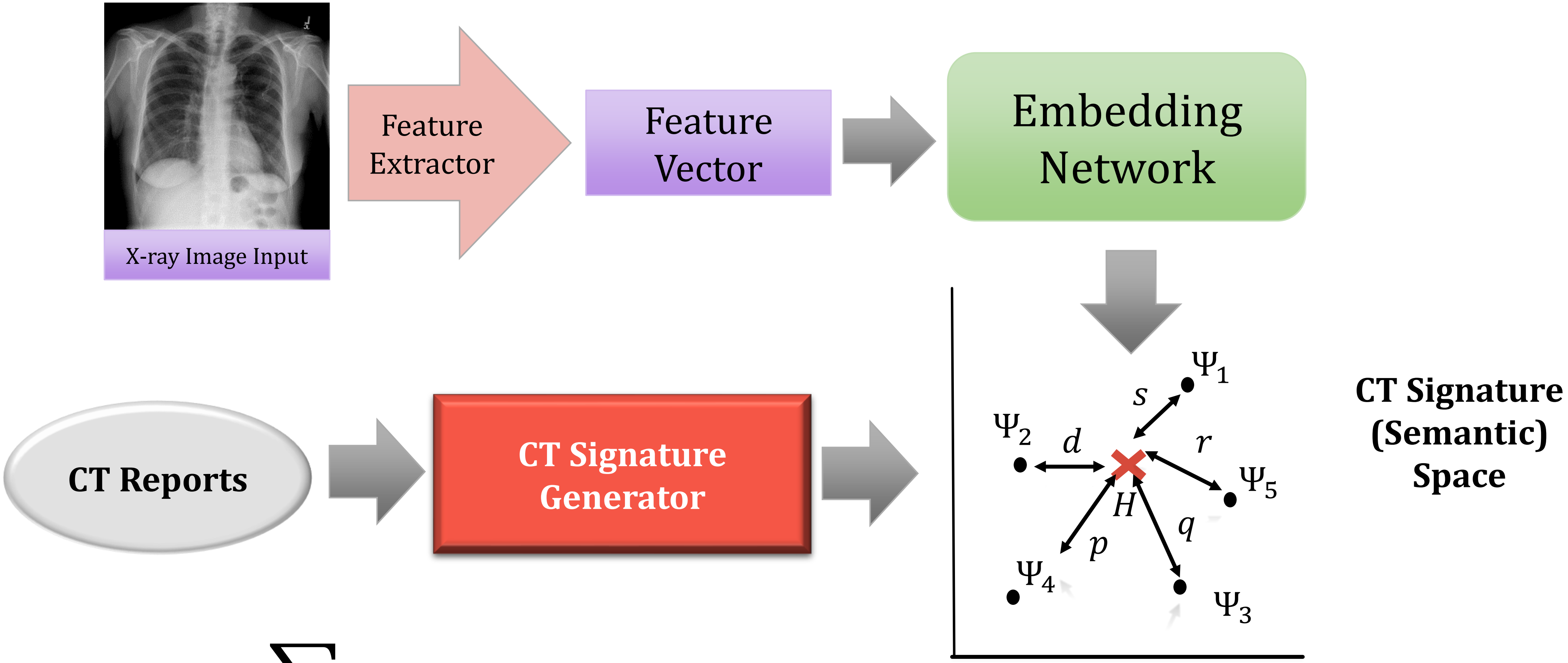
Meaningful Semantic Embedding



Meaningful Semantic Embedding

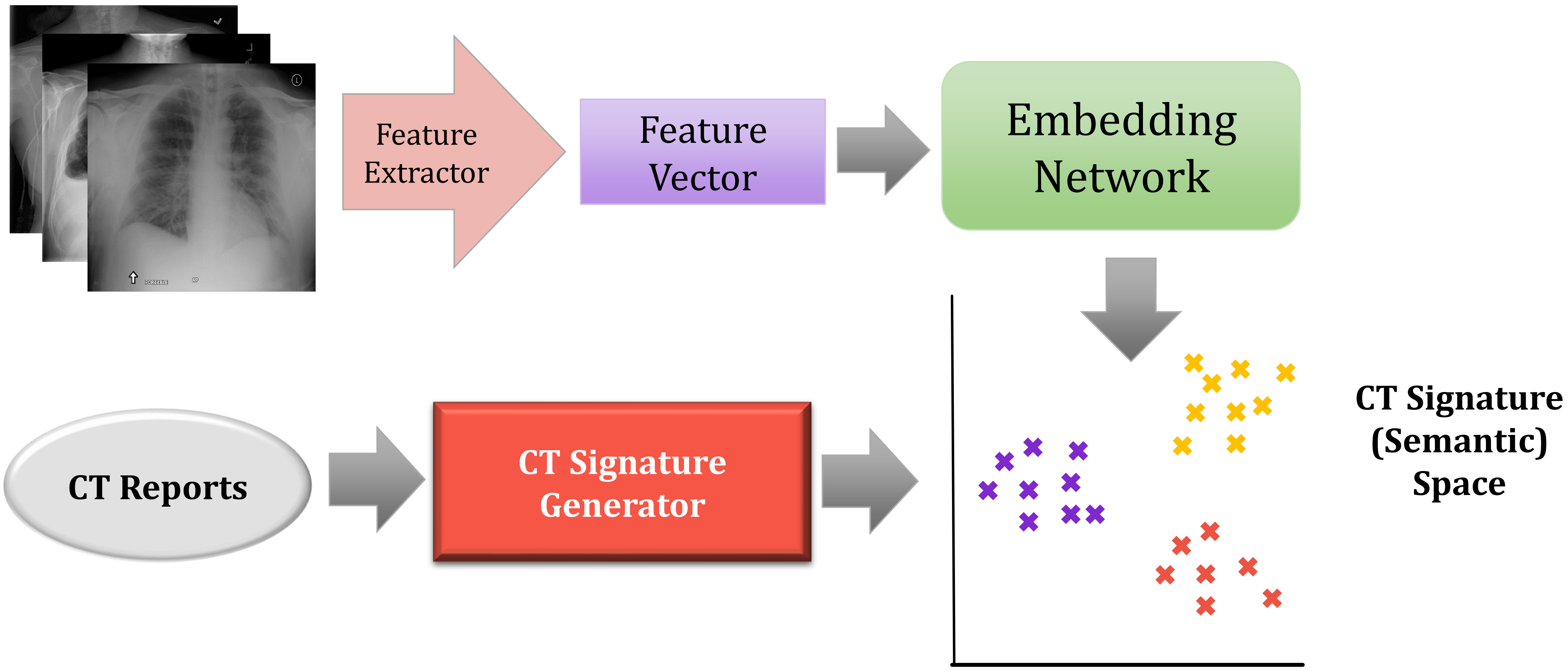


Meaningful Semantic Embedding

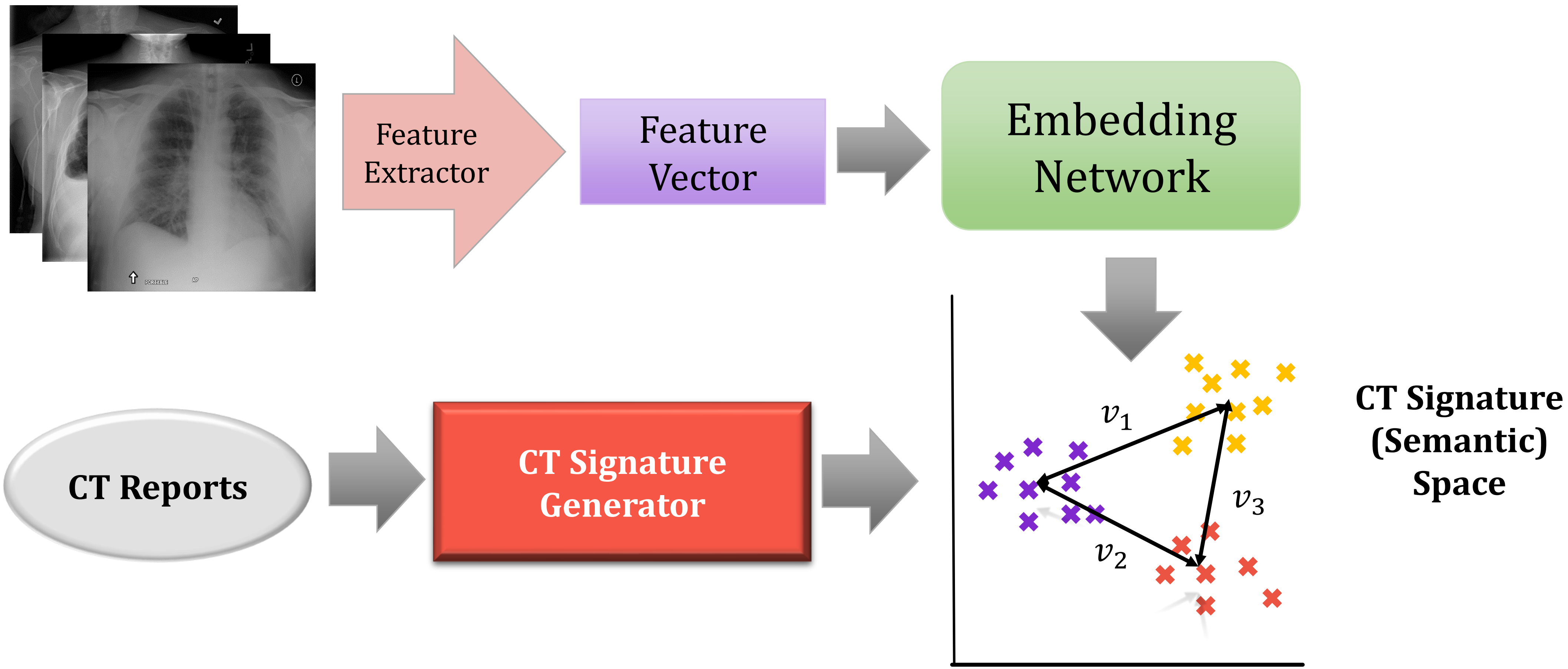


$$L_{se} = \sum d - (p + q + r + s)$$

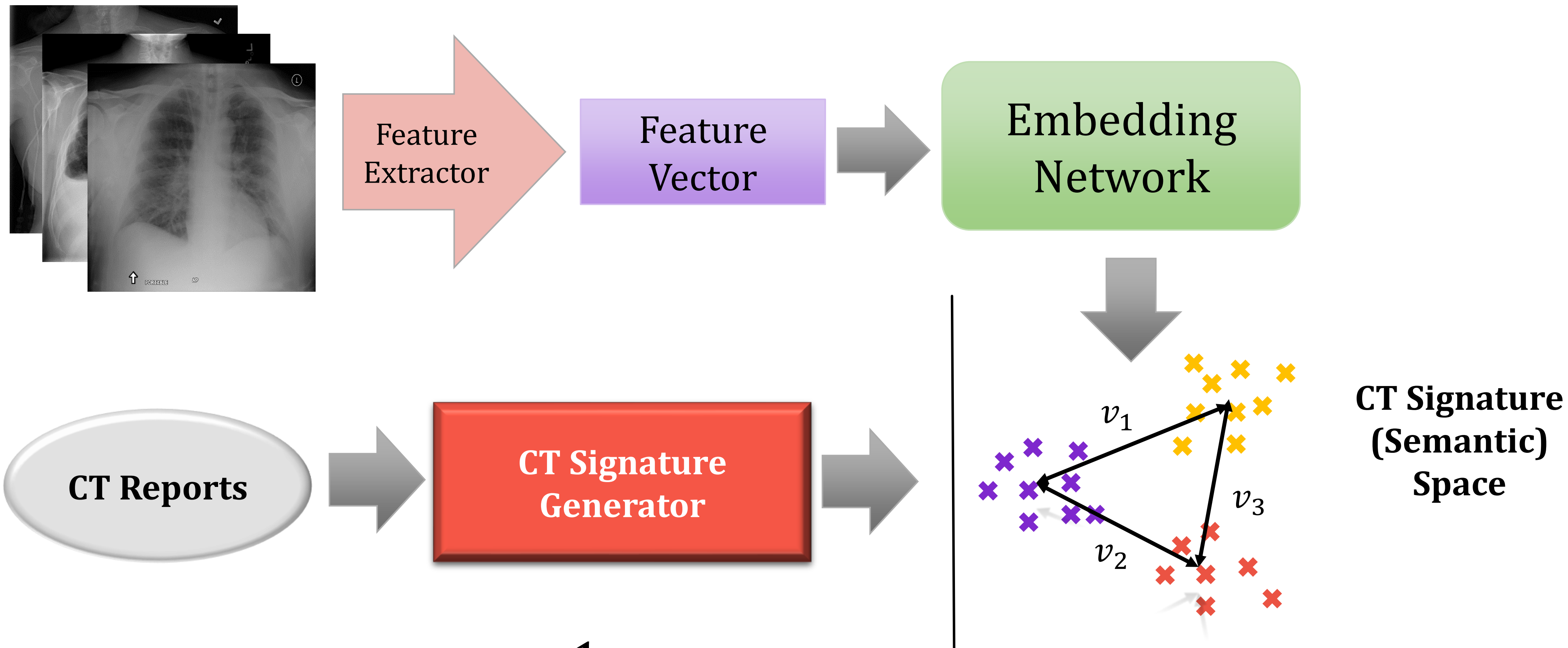
Meaningful Semantic Embedding



Meaningful Semantic Embedding

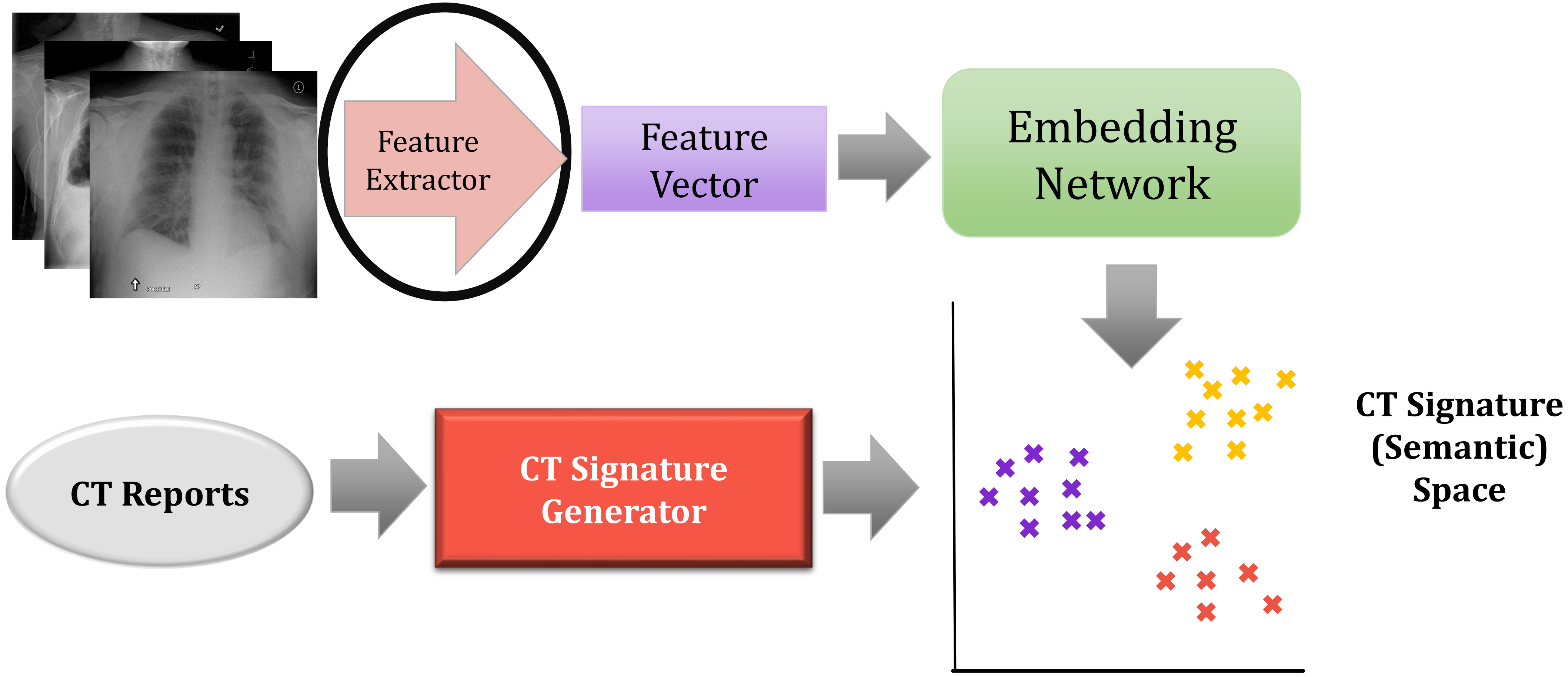


Meaningful Semantic Embedding



$$L_{sal} = \sum_{i} \left(\frac{1}{||v_i||^2 + \epsilon} \right)$$

Process Pipeline

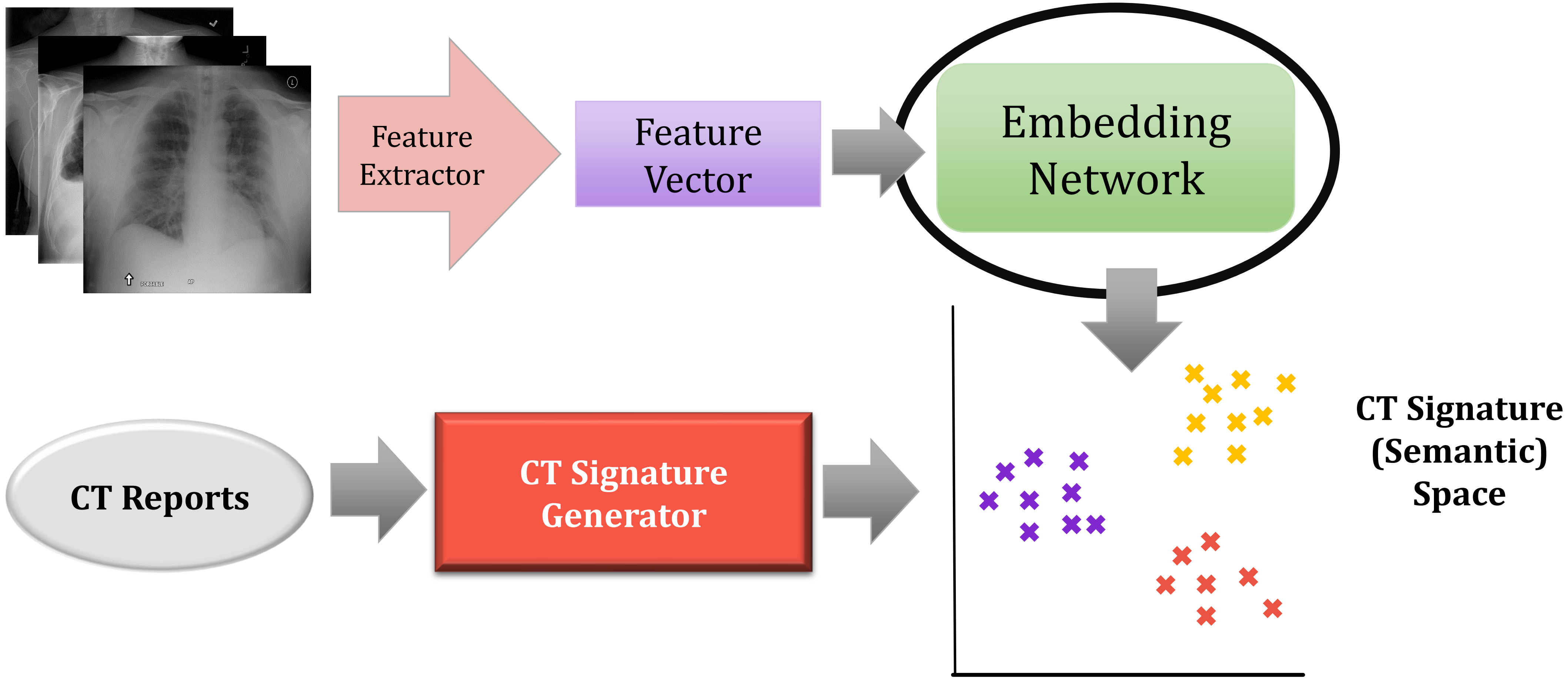


Feature Extractor

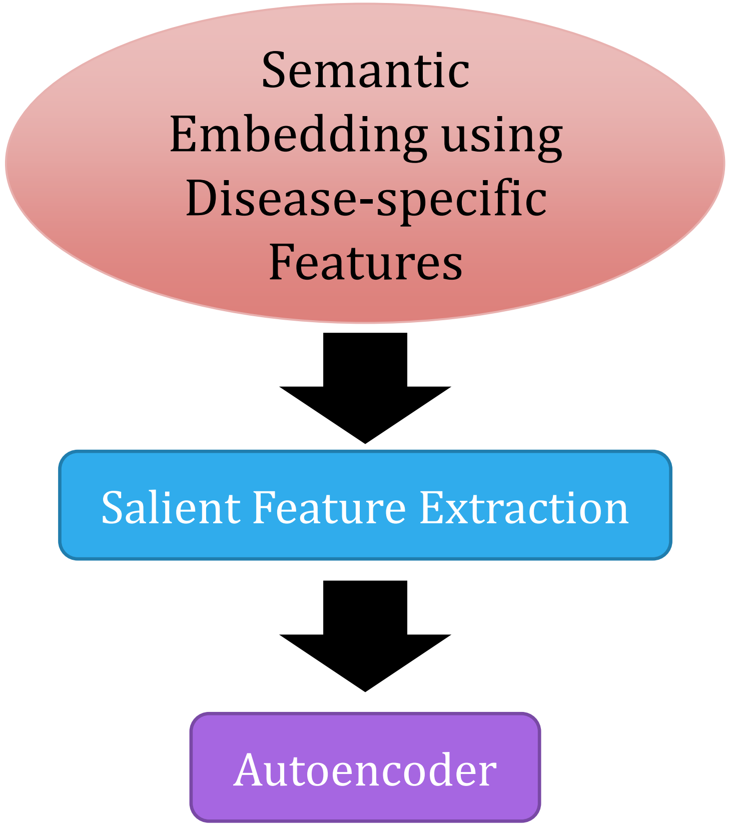
- DenseNet-121[2]
- Trained using seen classes
- Feature vectors from the penultimate layer
- Feature vectors: likely to be **noisy** for unseen classes

2. Rajpurkar *et al.*, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning", (2017).

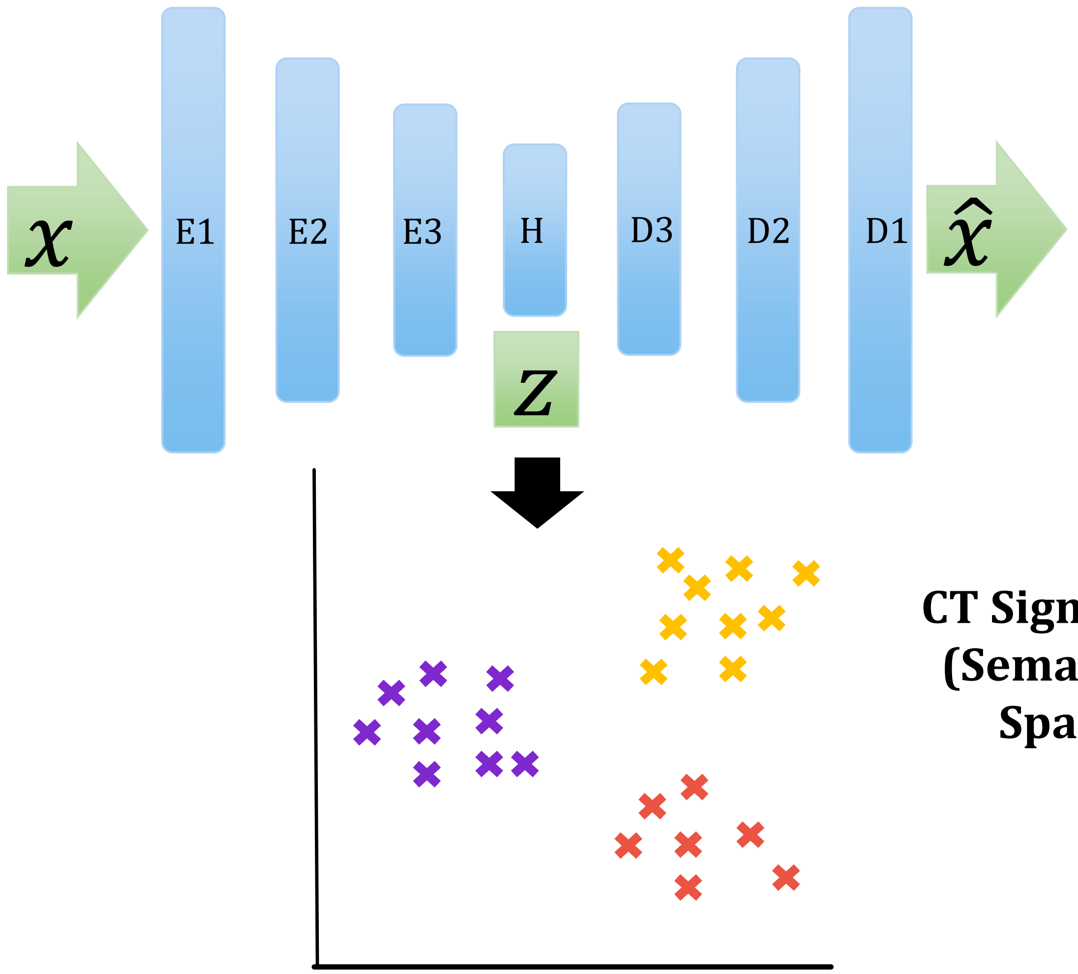
Process Pipeline



Choice of Embedding Network



Embedding Network: Autoencoders[3]



CT Signature (Semantic) Space

- Input data x
- Hidden state z
- Reconstructed output $\hat{x}$

3. Hinton *et al*. "Reducing the dimensionality of data with neural networks", (2006).

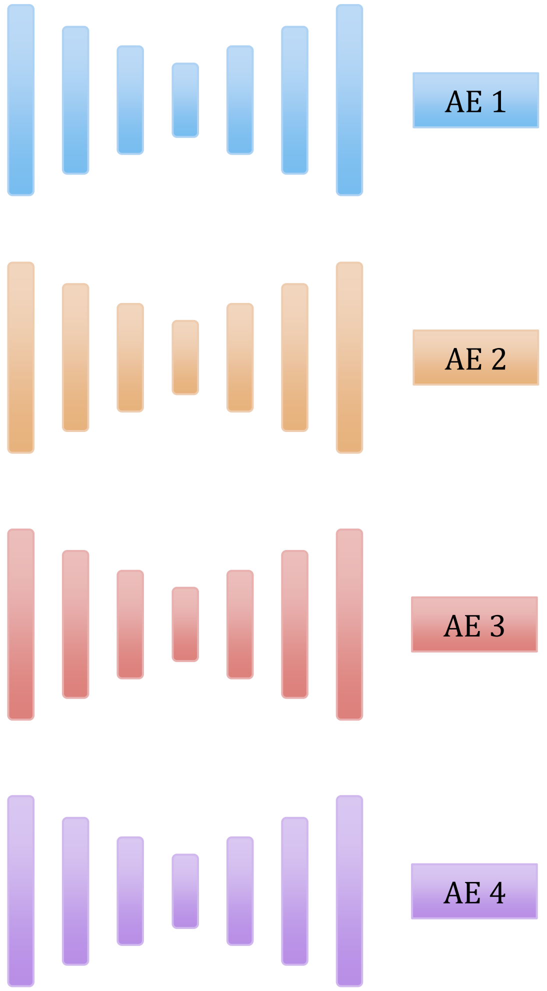
Embedding Network: Autoencoders

- Losses
  - Semantic embedding loss L_{se}
  - Semantic saliency loss L_{sal}
  - Reconstruction loss $L_{re} = ||x - \hat{x}||$

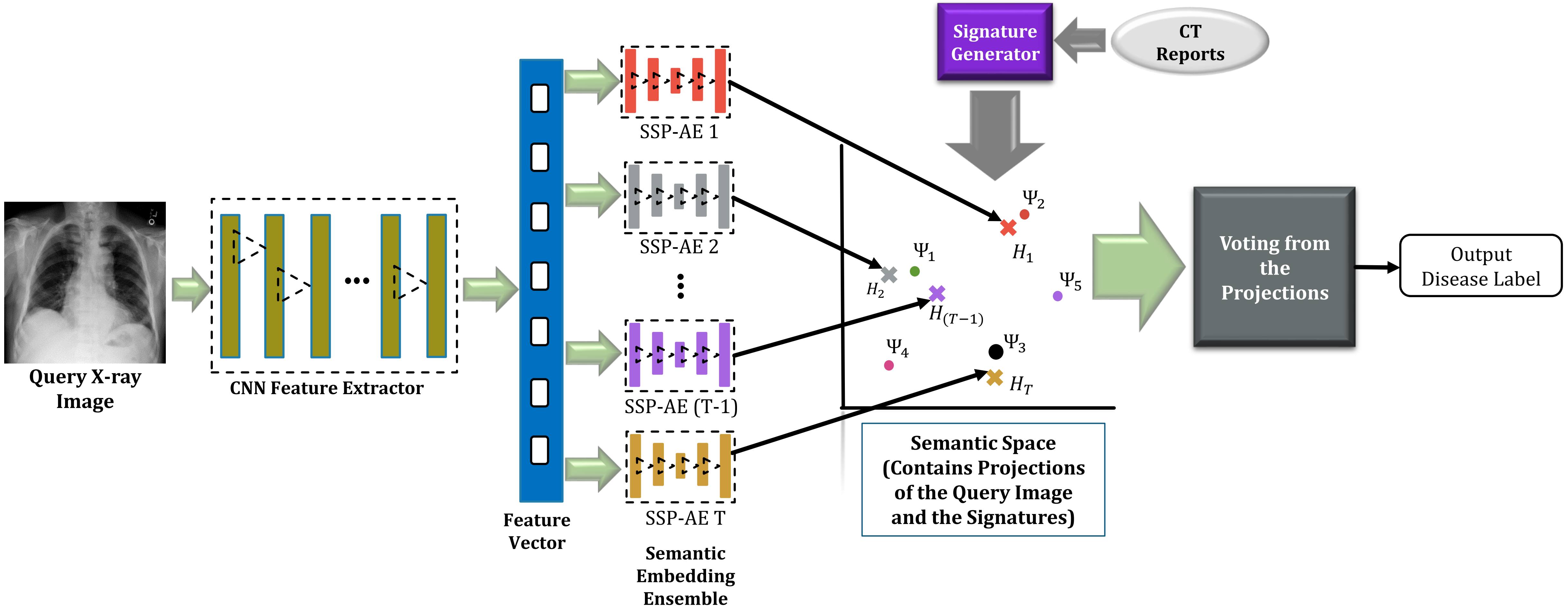
- Semantic Saliency Preserving Autoencoders (SSP-AE)

Dealing with Noisy Feature Vectors

- Ensemble[4] of Autoencoders
- Several autoencoders trained in parallel
- Each autoencoder
  - Explores different feature subspaces
    - Semi-deterministic selection of subspaces
  - Trained with different bootstrap samples[4]



4. Breiman, L., "Random forests", (2001).



Query X-ray Image
CNN Feature Extractor
Feature Vector
SSP-AE 1
SSP-AE 2
SSP-AE (T-1)
SSP-AE T
Semantic Embedding Ensemble
Signature Generator
CT Reports
Ψ_1
Ψ_2
Ψ_3
Ψ_4
Ψ_5
H_1
H_2
$H_{(T-1)}$
H_T
Semantic Space (Contains Projections of the Query Image and the Signatures)
Voting from the Projections
Output Disease Label

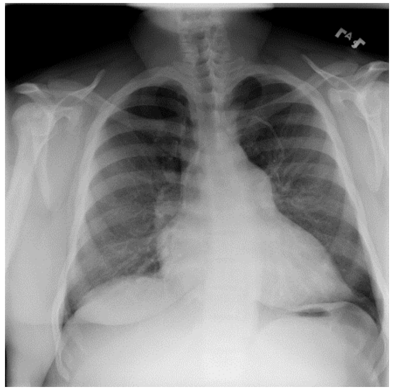
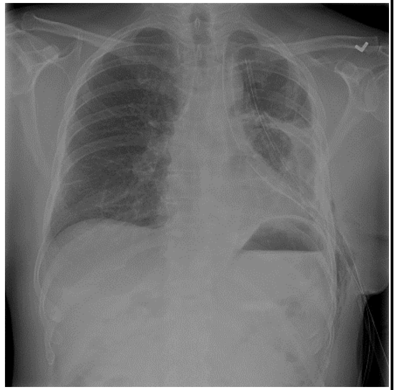
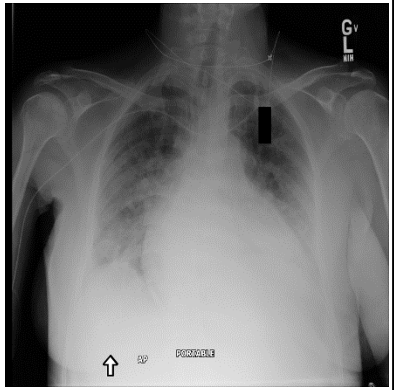
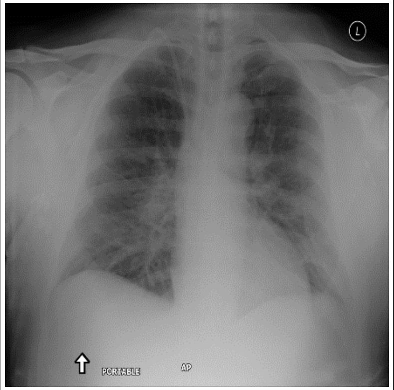
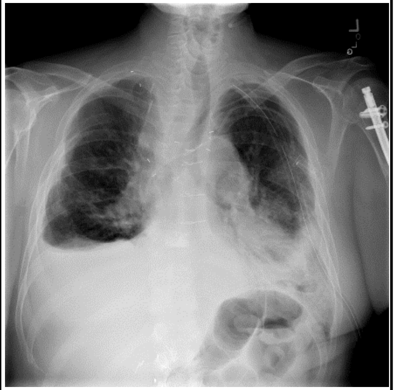
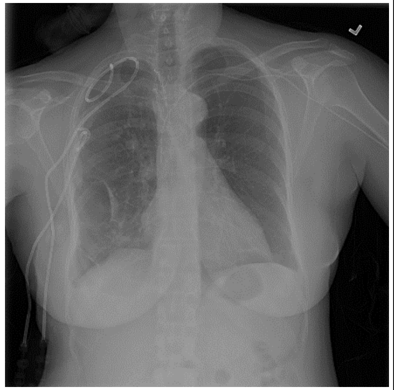
Seen & Unseen Classes

Class
Pneumonia
Nodule
Pneumothorax
Consolidation
Effusion
Infiltration
Edema
Emphysema
Cardiomegaly

Seen Classes

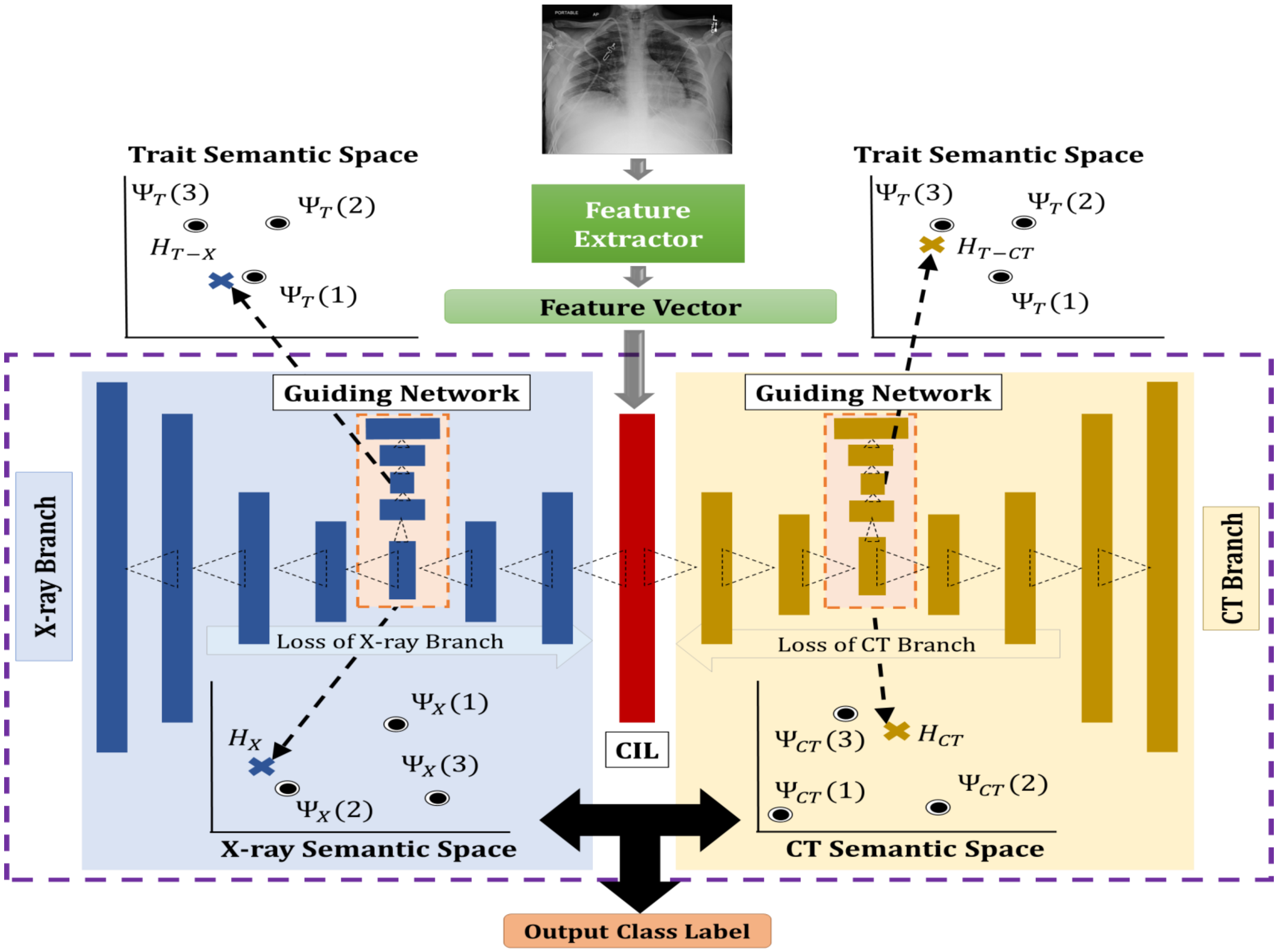
Unseen Classes

Visual Results

	Unseen Class			Seen Class		
CXR Examples						
GT	**Cardiomegaly**	Infiltration Pneumothorax **Emphysema**	Cardiomegaly Edema	**Nodule**	**Effusion** Infiltration	Pneumothorax
D	**Cardiomegaly**	**Emphysema**	Nodule	**Nodule**	**Effusion**	Pneumonia

Can We Use More Auxiliary Information?

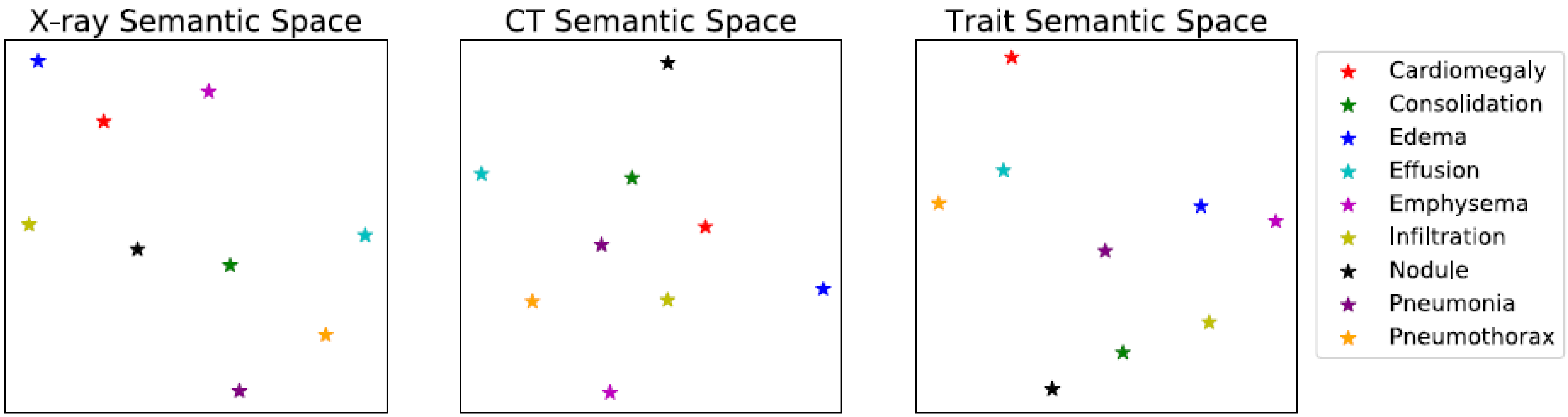
Trait-guided Multi-view Semantic Embedding for Zero-shot Chest X-ray Diagnosis



Visual Traits

- Location: Contains anatomical location information
  - lung, heart, etc.
- Position: Contains information of the pixel position corresponding to the abnormality
  - Upper portion of image, lower portion of image, etc.
- Opacity: high, low, medium
- Distribution: unilateral, bilateral
- Border sharpness: clear margin, indistinct margin
- Size (relative to lung volume)
- Aspect ratio

Semantic Spaces



Seen & Unseen Classes

Unseen Classes

Seen Classes

Class
Pneumonia
Nodule
Infiltration
Consolidation
Effusion
Pneumothorax
Edema
Emphysema
Cardiomegaly

Combination 1

Class
Pneumonia
Nodule
Infiltration
Consolidation
Effusion
Pneumothorax
Edema
Emphysema
Cardiomegaly

Combination 2

Class
Pneumonia
Nodule
Infiltration
Consolidation
Effusion
Pneumothorax
Edema
Emphysema
Cardiomegaly

Combination 3

Visual Results

Dataset	NIH-900		Open-i		PMC		
Image Examples							
Ground Truth	Cardiomegaly	Infiltration	Edema Cardiomegaly	Pneumonia	Consolidation	Nodule	Pneumothorax
Detected	Cardiomegaly (S)	Infiltration (U)	Edema (S)	Pneumonia (U)	Consolidation (S)	Nodule (U)	Nodule (U)

VAE for Zero-shot CXR Diagnosis

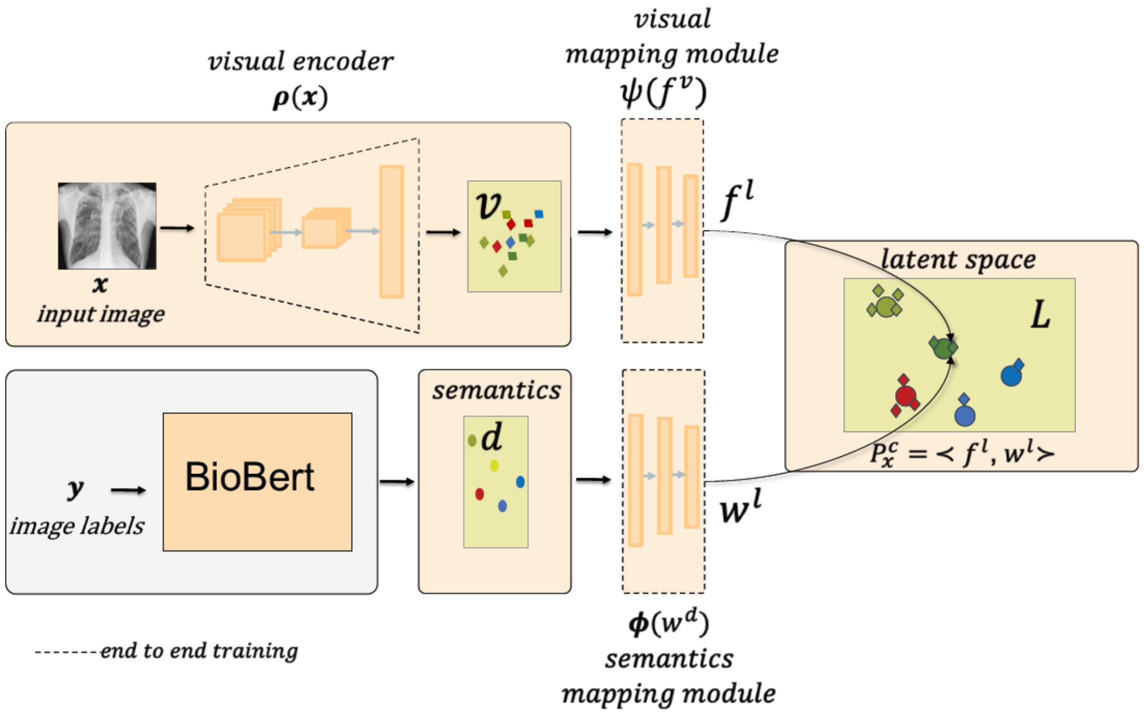
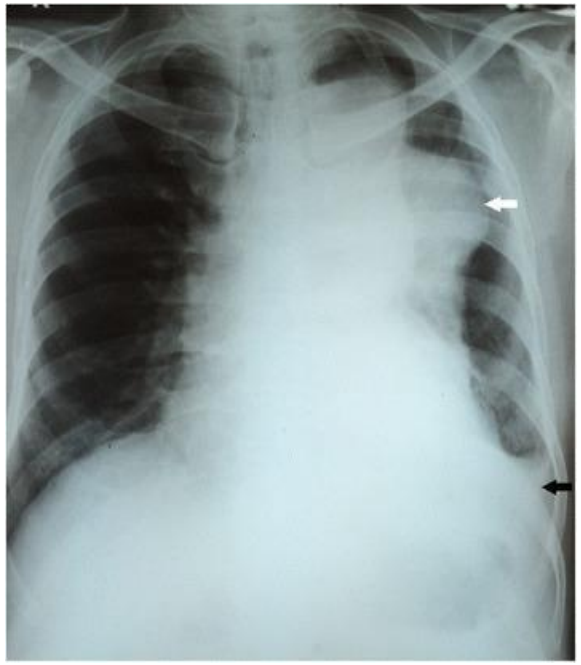
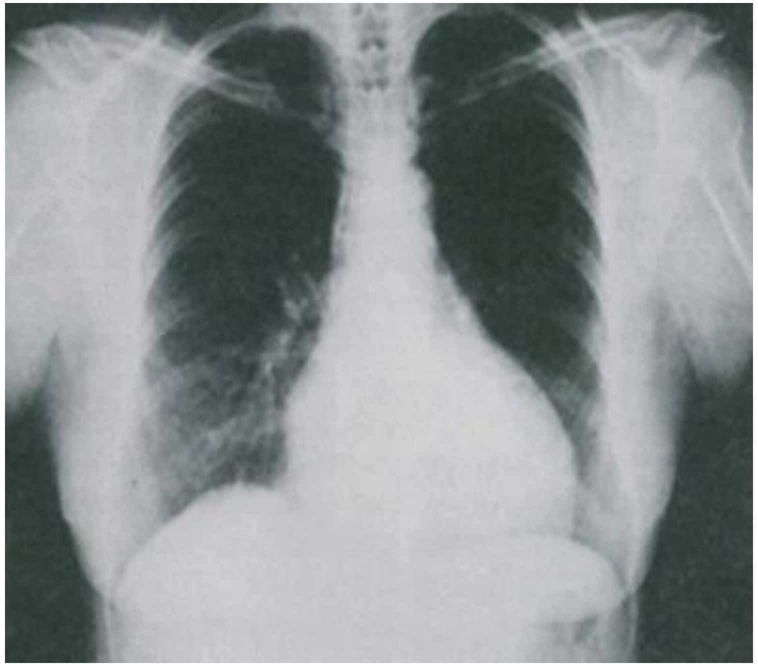


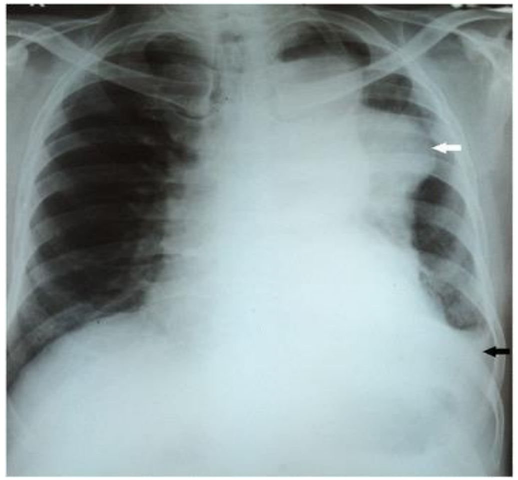
Figure 1: **Overview of our CXR-ML-GZSL model for learning visual representations of chest X-rays.** An overview of our network for chest X-ray images. It includes a trainable visual encoder and v- and d-dimensional visual and semantics spaces, respectively. For an input image x and its labels y, the network learns a visual representation guided by semantics extracted by BioBert. We perform end-to-end training of the visual encoder, visual mapping module, and the semantics mapping module, as indicated by the black dashed line.

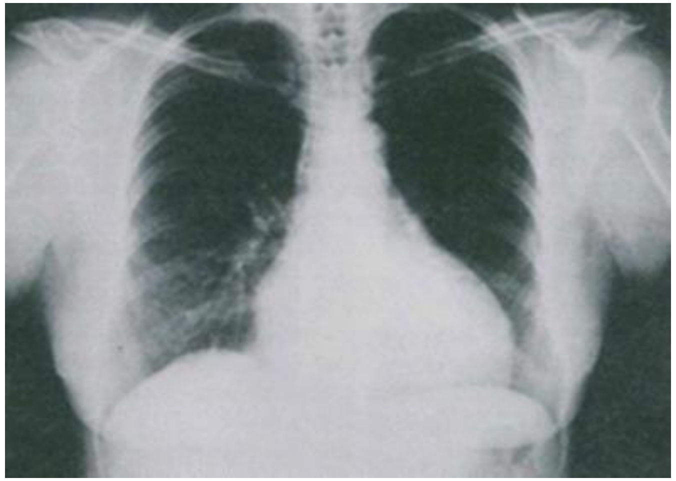
Few-shot Chest X-ray Diagnosis Using Images from the Published Scientific Literature

Sample X-ray Images in Published Literature

Challenges in Few-shot Learning from Published Literature





- Learning from few images
- Artifacts, low-resolution

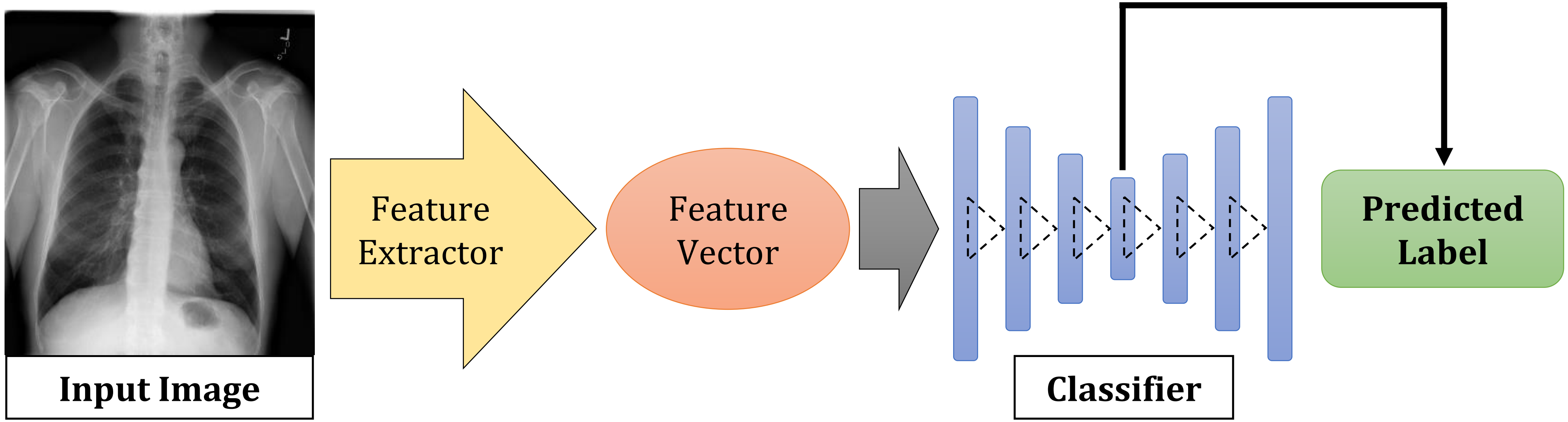
Solution: Use of Labeled and Unlabeled Images

- Labeled images from published literature: PubMed Central (PMC)
  - Initial training of few-shot learning model

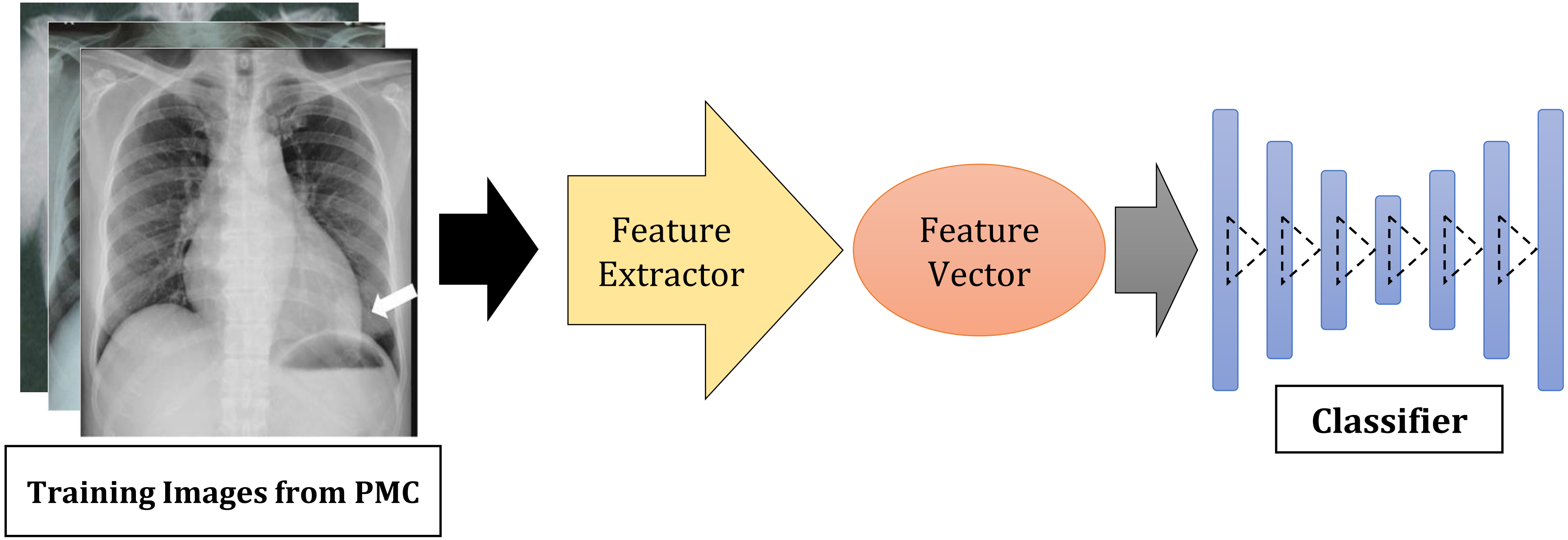
- Unlabeled images: from NIH CXR dataset
  - High-resolution, less artifacts

- Re-training with pseudo labels for the NIH CXR dataset
  - Dealing with the problem of noisy labels

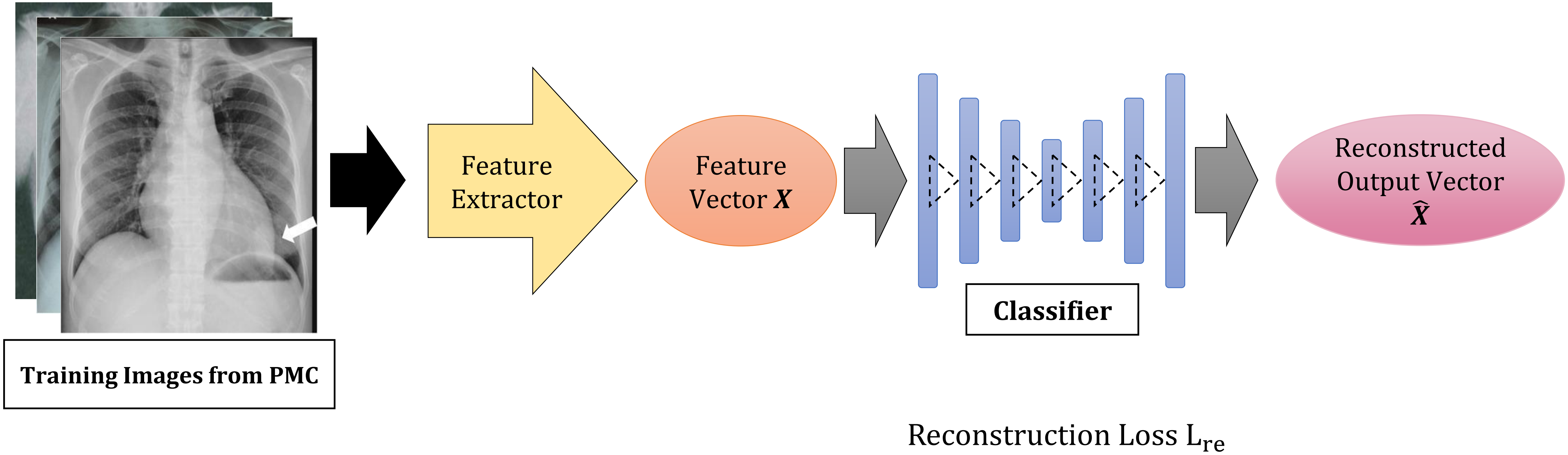
Process Pipeline



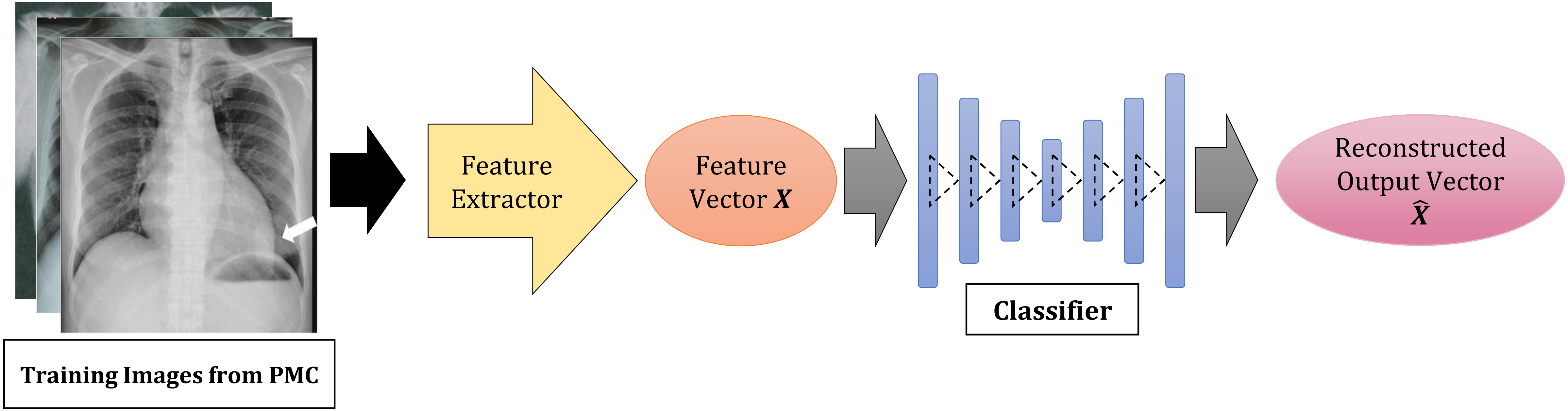
Process Pipeline: Initial Training of Classifier



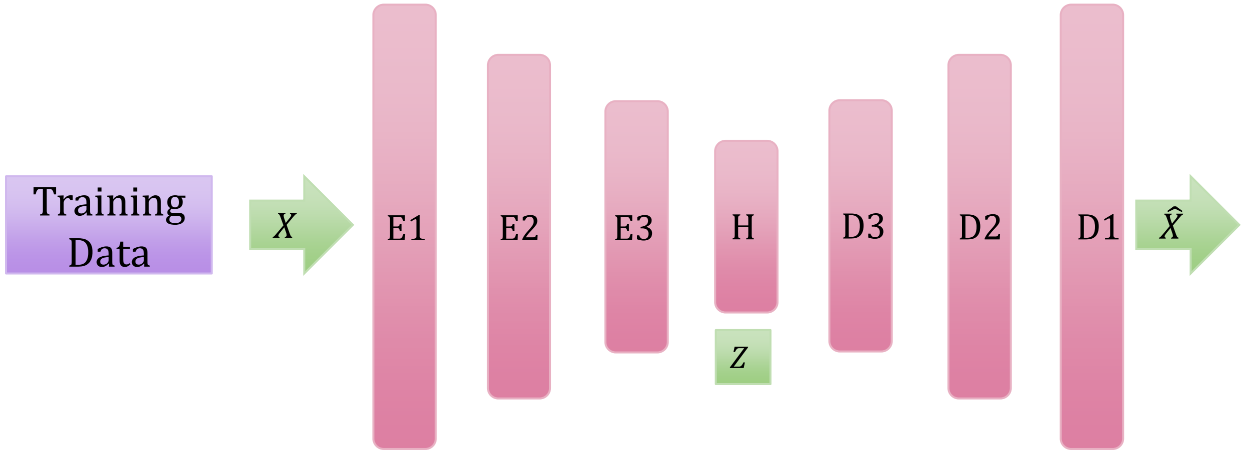
The Loss Function



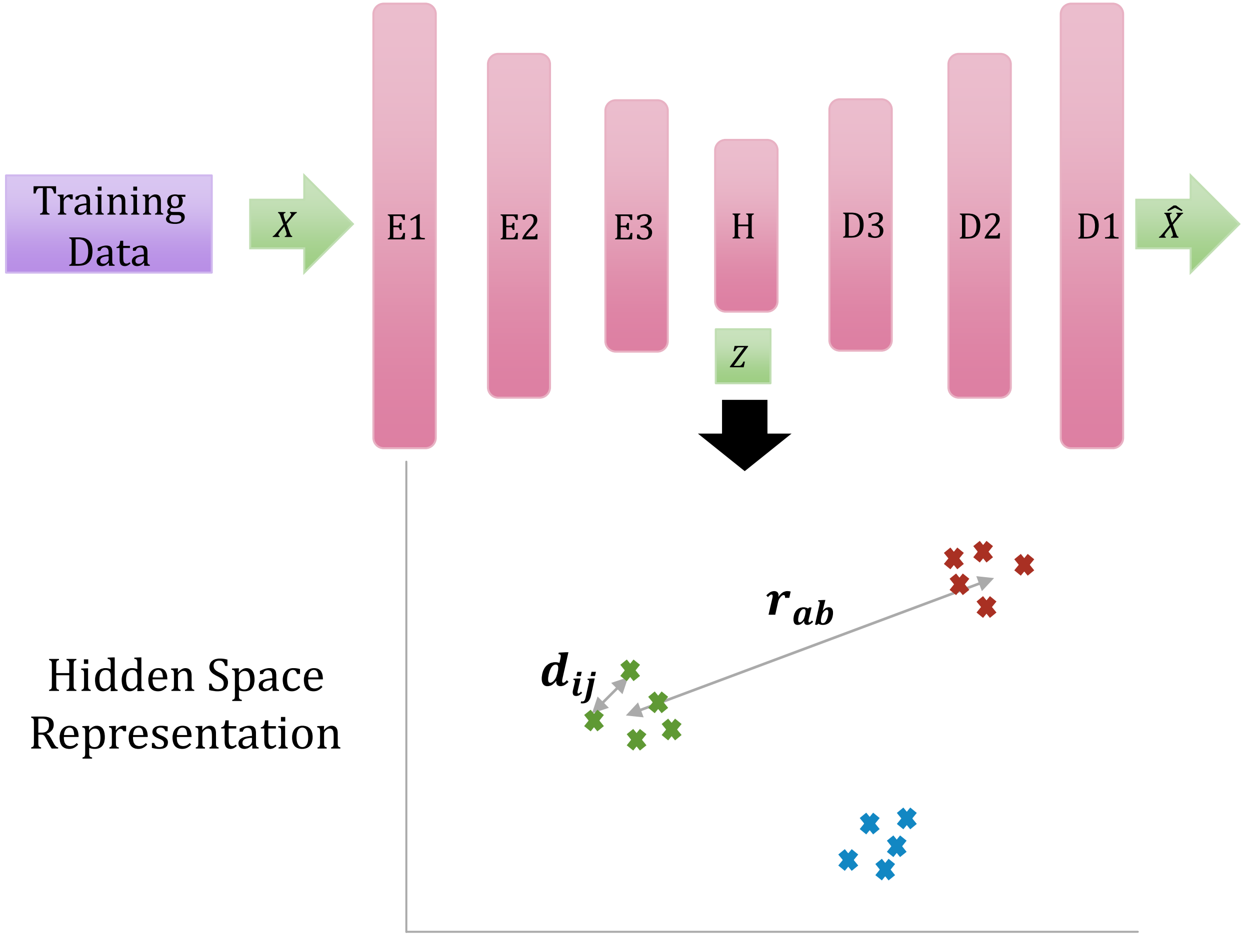
Loss Computed From Clusters



Loss Computed From Clusters



Loss Computed From Clusters



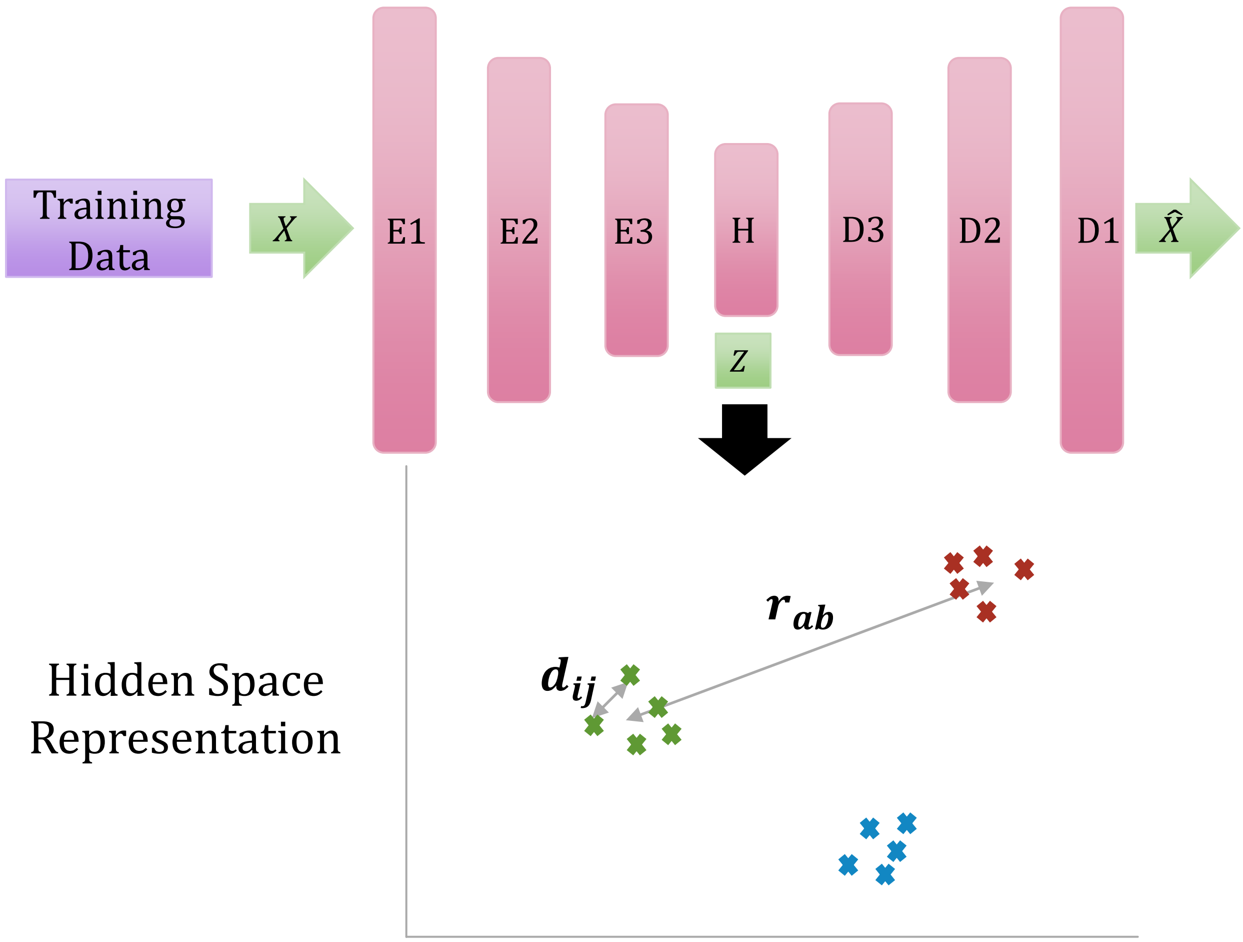
- Minimize every $\boldsymbol{d_{ij}}$ inside each cluster

$$L_{con} = \sum_{\forall c \in C} \Big(\sum_{\forall\, x_i, x_j \in S_c} d_{ij} \Big)$$

- Maximize $\boldsymbol{r_{ab}}$ for every pair of clusters

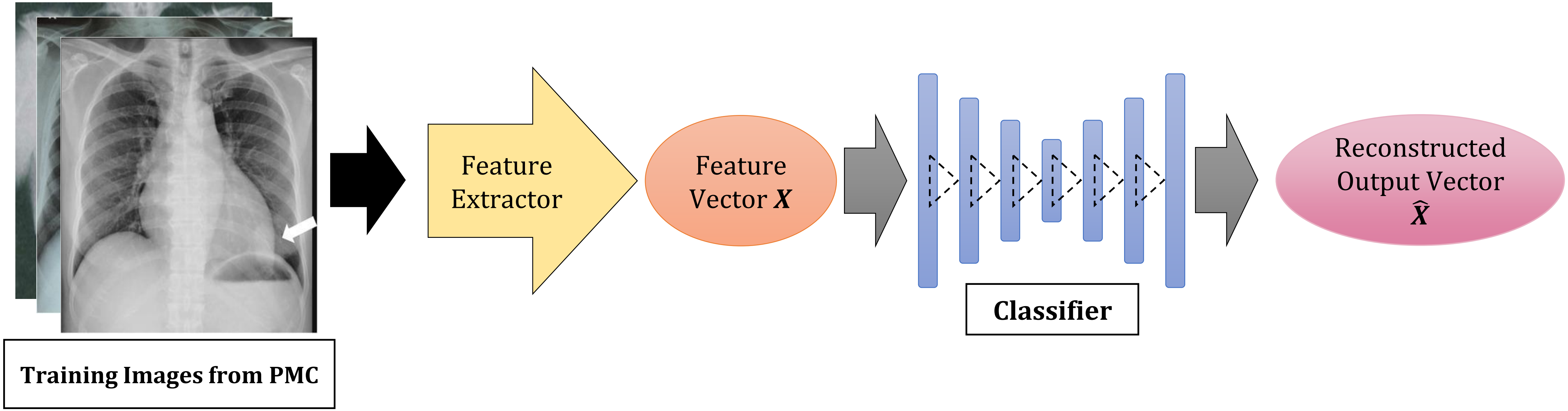
$$L_{sep} = -\sum_{a,b \in C} r_{ab}$$

Loss Computed From Clusters

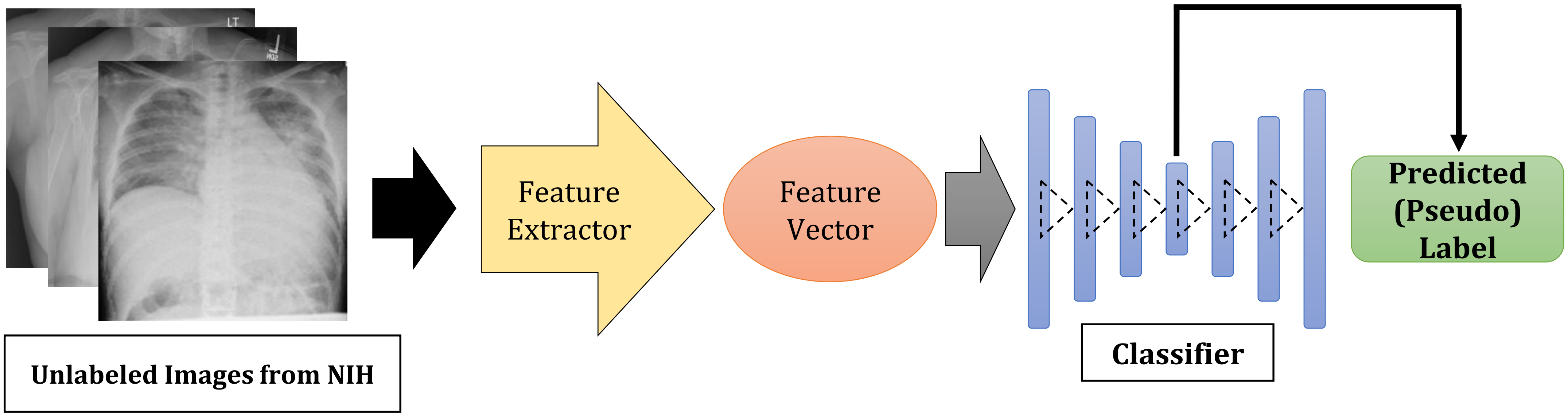


- Minimize every $\boldsymbol{d_{ij}}$ inside each cluster (L_{con})
- Maximize $\boldsymbol{r_{ab}}$ for every pair of clusters (L_{sep})
- $L = L_{re} + \lambda_1 L_{con} + \lambda_2 L_{sep}$

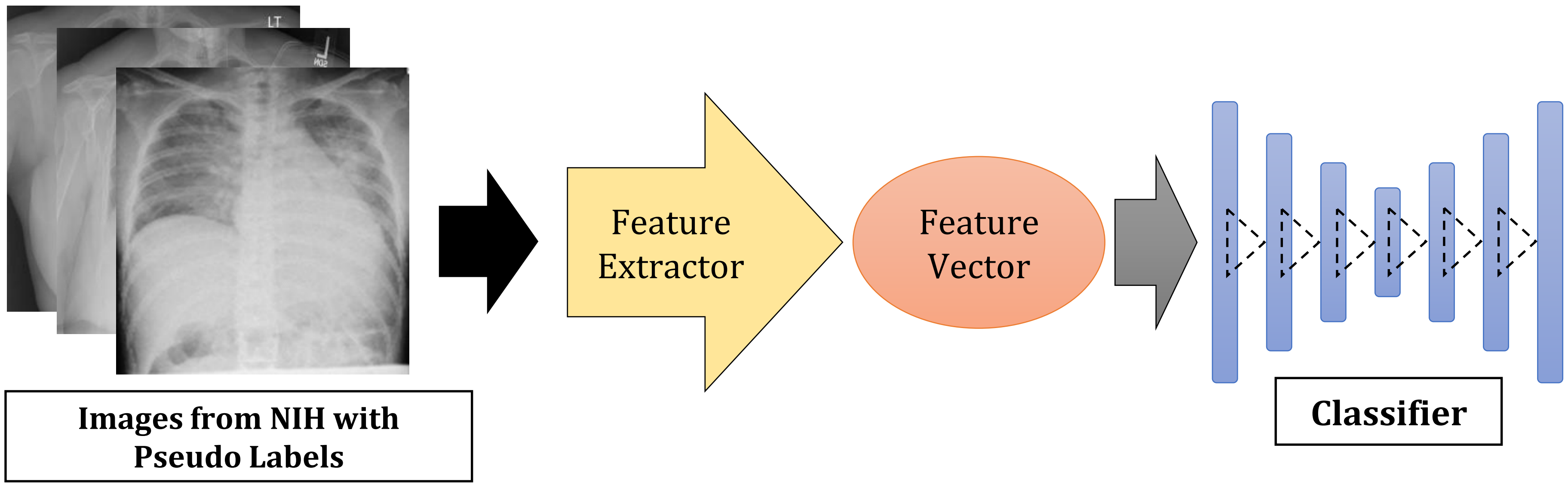
Loss Computed From Clusters



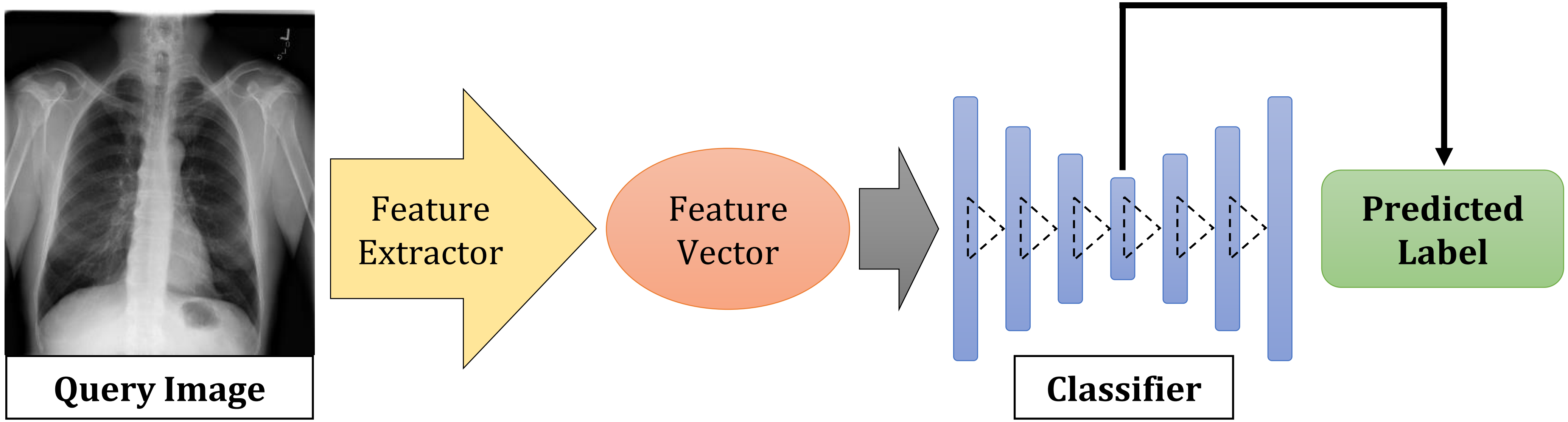
Inference for Unlabeled Images: Pseudo Labels



Re-training


Feature
Extractor
Feature
Vector
Images from NIH with
Pseudo Labels
Classifier

Inference

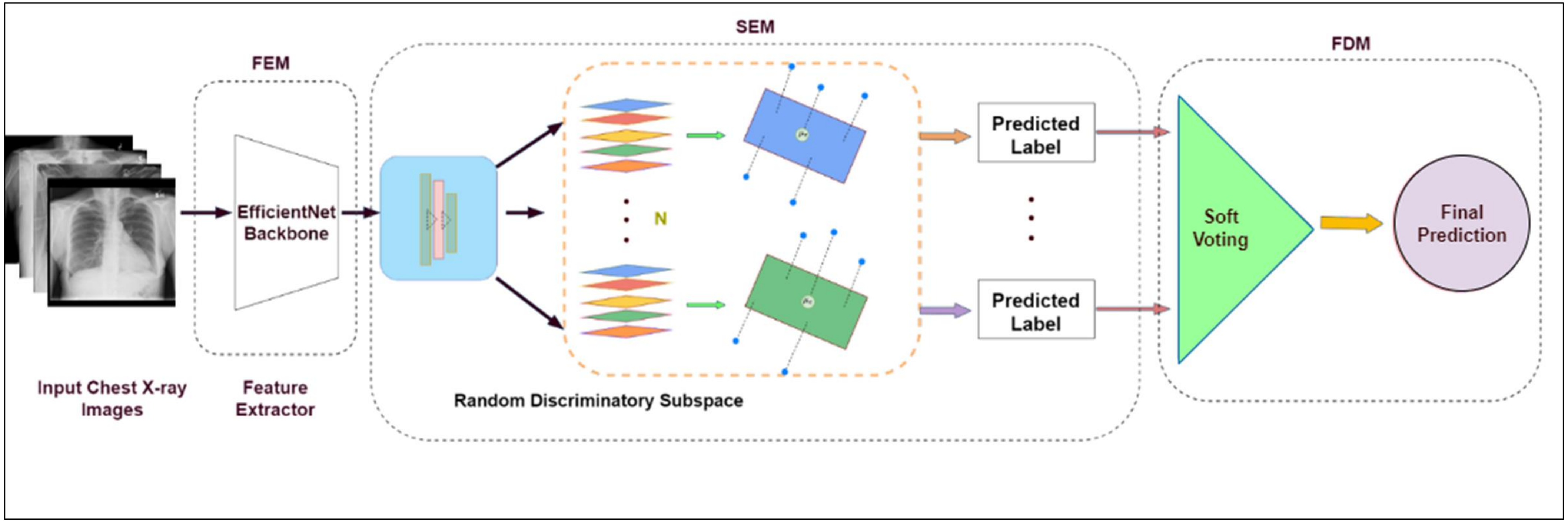


Feature Extractor
Feature Vector
Predicted Label
Query Image
Classifier

Image Results

Image Examples			
Ground Truth	Cardiomegaly	Edema	Edema
Detected	**Cardiomegaly**	**Edema**	**Cardiomegaly**

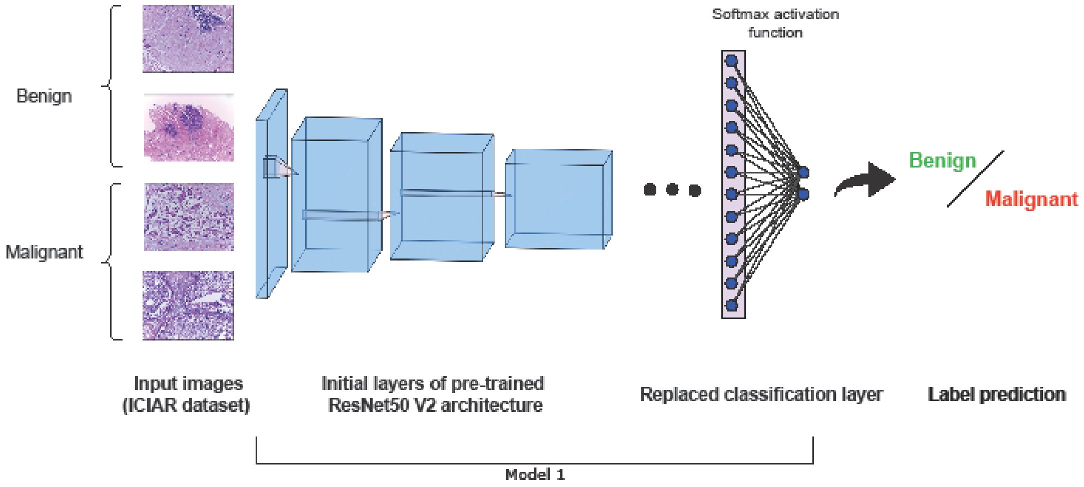
FSL Using an Ensemble of Subspaces



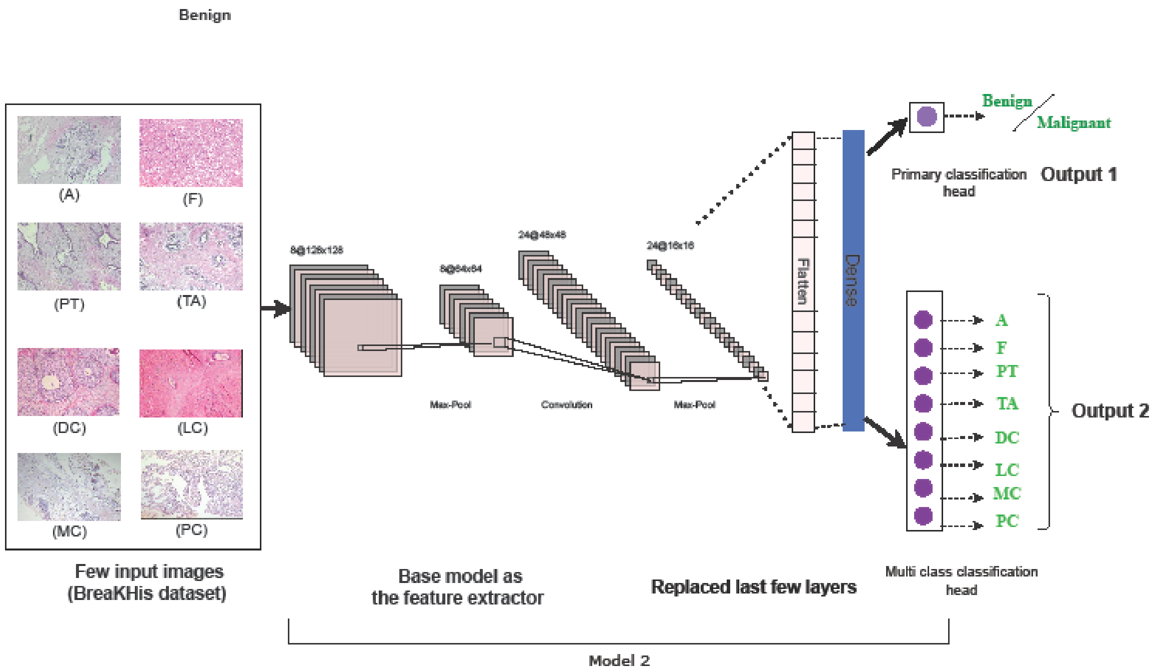
FSL Using an Ensemble of Subspaces

Images					
GT	Hernia	Fibrosis	Emphysema	Edema	Pneumonia
P	**Hernia**	**Fibrosis**	**Emphysema**	**Cardiomegaly**	**Pneumonia**

Few-shot Learning for Breast Cancer Detection

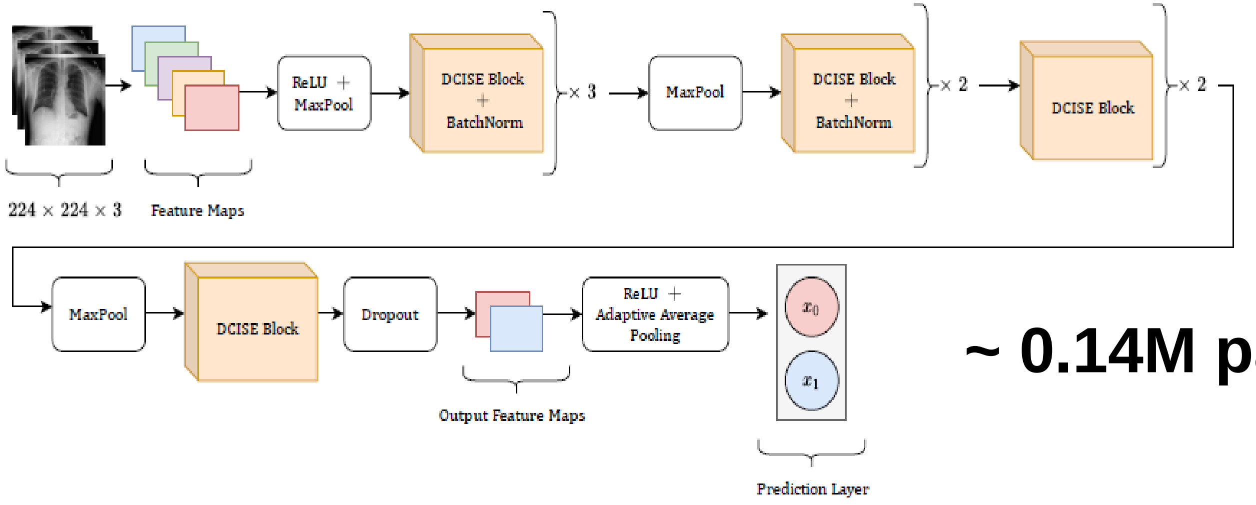


a) Binary classification on the ICIAR dataset by fine-tuning with ResNet50 V2 pre-trained architecture



b) Multi output model for primary and fine-grain classification trained on very few histopathology images from the BreaKHis dataset. Model 1 has been used as the base model.

Lightweight CNN Model for Chest X-ray Diagnosis



~ 0.14M parameters and ~ 550 KB size

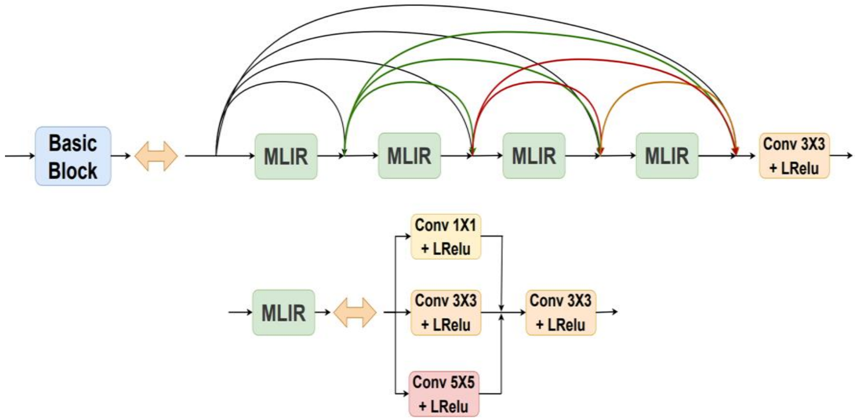
SI						
GT	Pleural Effusion	Cardiomegaly	Atelectasis	No Finding	Pneumothorax	Consolidation
D	**Pleural Effusion**	**Cardiomegaly**	**Atelectasis**	**No Finding**	**Pneumothorax**	**Atelectasis**

Single Image Super-resolution for Chest X-rays

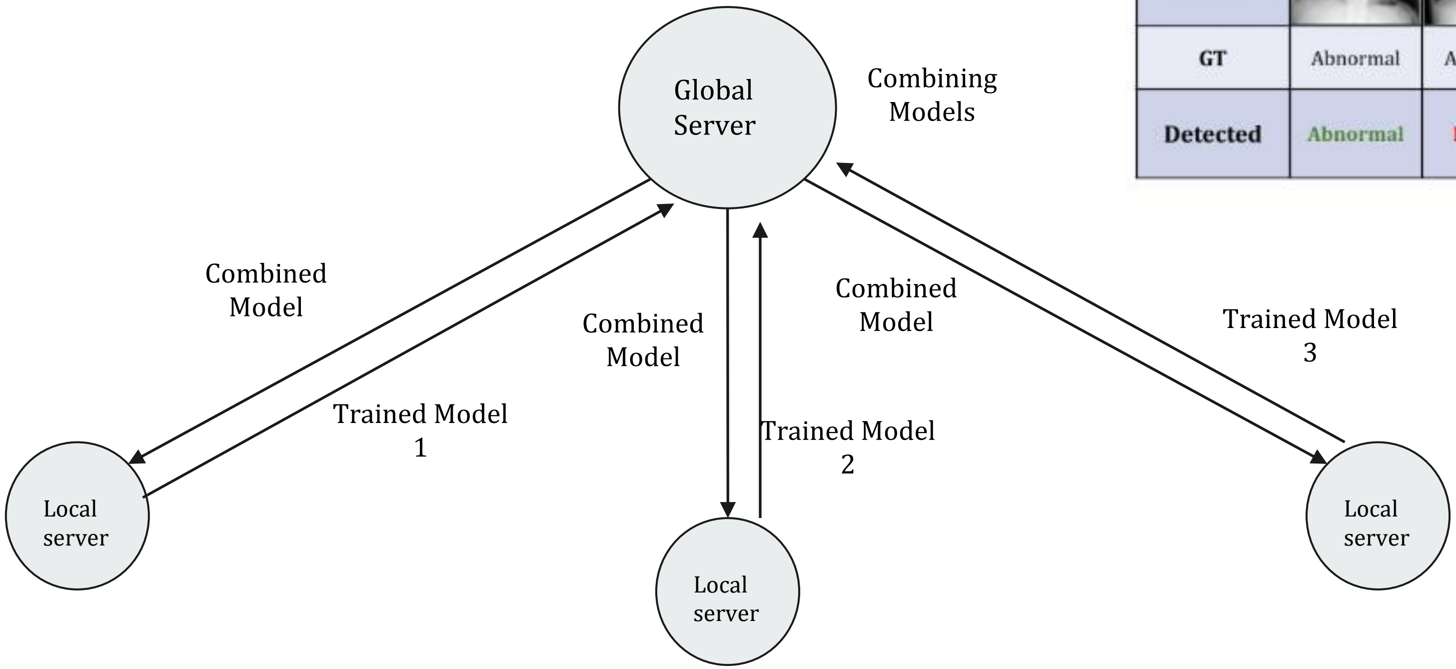
Low Res Image

Our Model

Ground Truth



Federated Learning for Radiology Diagnosis



Dataset	NIH			CheXpert			VinBig		
Image Examples									
GT	Abnormal	Abnormal	Abnormal	Normal	Normal	Normal	Normal	Normal	Normal
Detected	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal	Normal

Type	Method	F1 score NIH	F1 score CX	F1 score VB	Accuracy NIH	Accuracy CX	Accuracy VB
Local	NIH-L	0.93	0.89	0.64	0.93	0.81	0.68
	CX-L	0.90	0.90	0.70	0.90	0.84	0.75
	VB-L	0.80	0.87	0.88	0.83	0.79	0.90
Global	DN121-SA	0.90	0.89	0.70	0.91	0.82	0.79
	DN121-TL	0.83	0.88	0.65	0.85	0.81	0.78
	DN121-VL	0.88	0.88	0.69	0.89	0.81	0.80
	[6]	0.89	0.88	0.49	0.88	0.80	0.61
	Proposed	0.90	0.90	0.71	0.90	0.84	0.76