

KiUT: Knowledge-injected U-Transformer for Radiology Report Generation

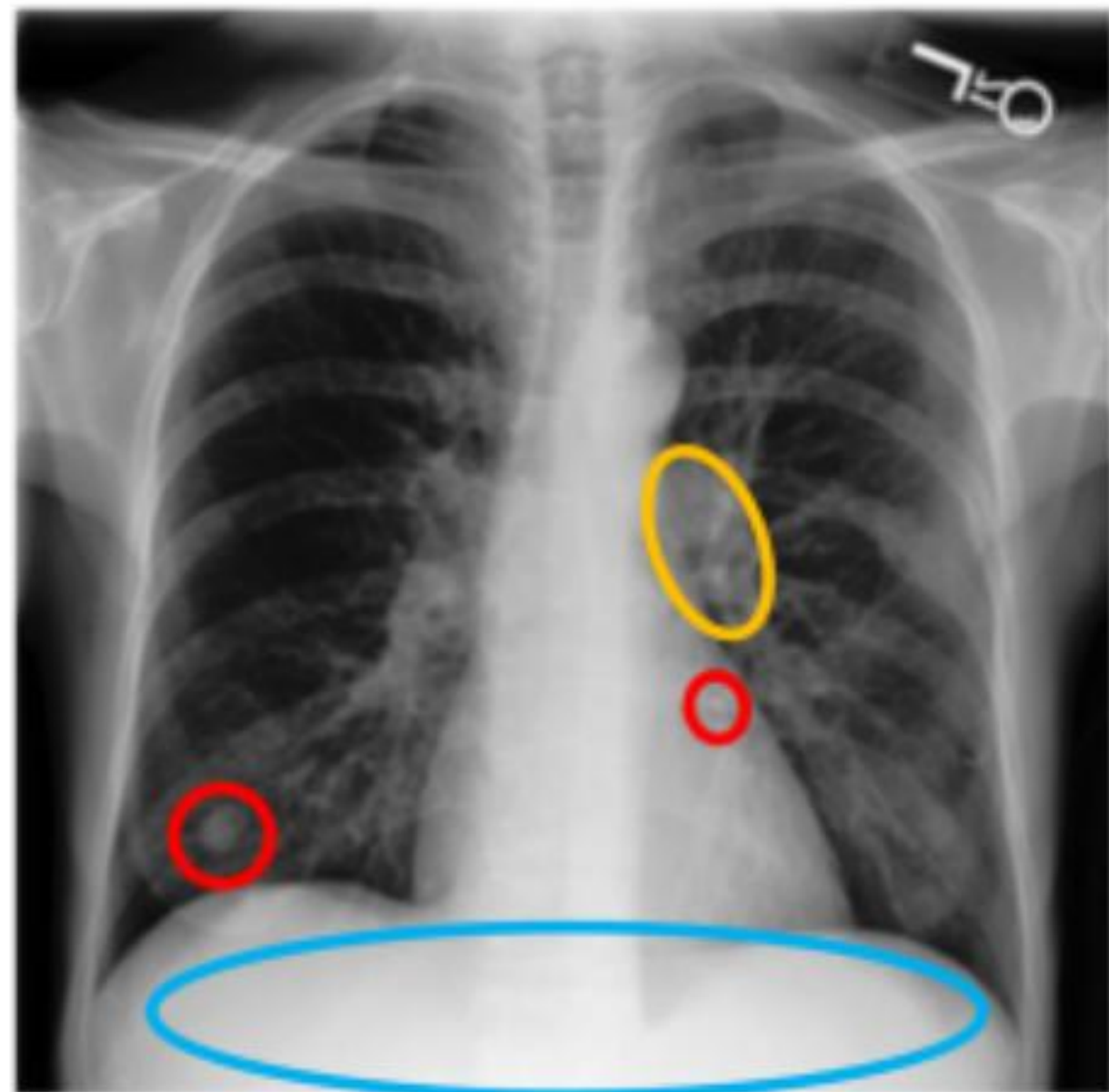
Zhongzhen Huang^{1,2}, Xiaofan Zhang^{1,2}, Shaoting Zhang^{2,3}

¹Shanghai Jiao Tong University ²Shanghai AI Laboratory ³SenseTime Research



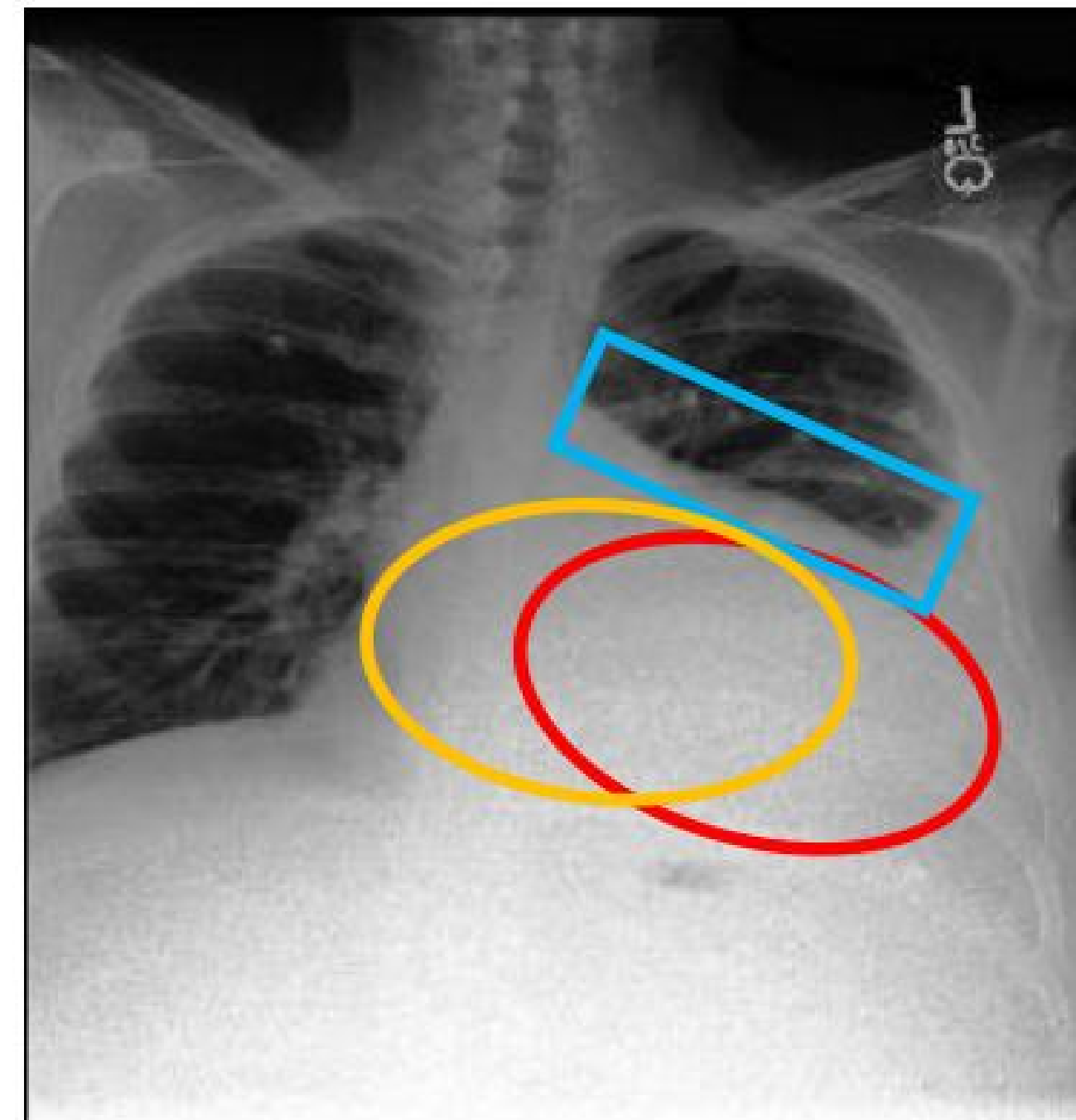
Clinical Radiology Report

Findings: The cardiomeastinal silhouette is normal in size and contour. Masslike opacification of right apex. No pneumothorax or large pleural effusion. The osseous structures are grossly normal.

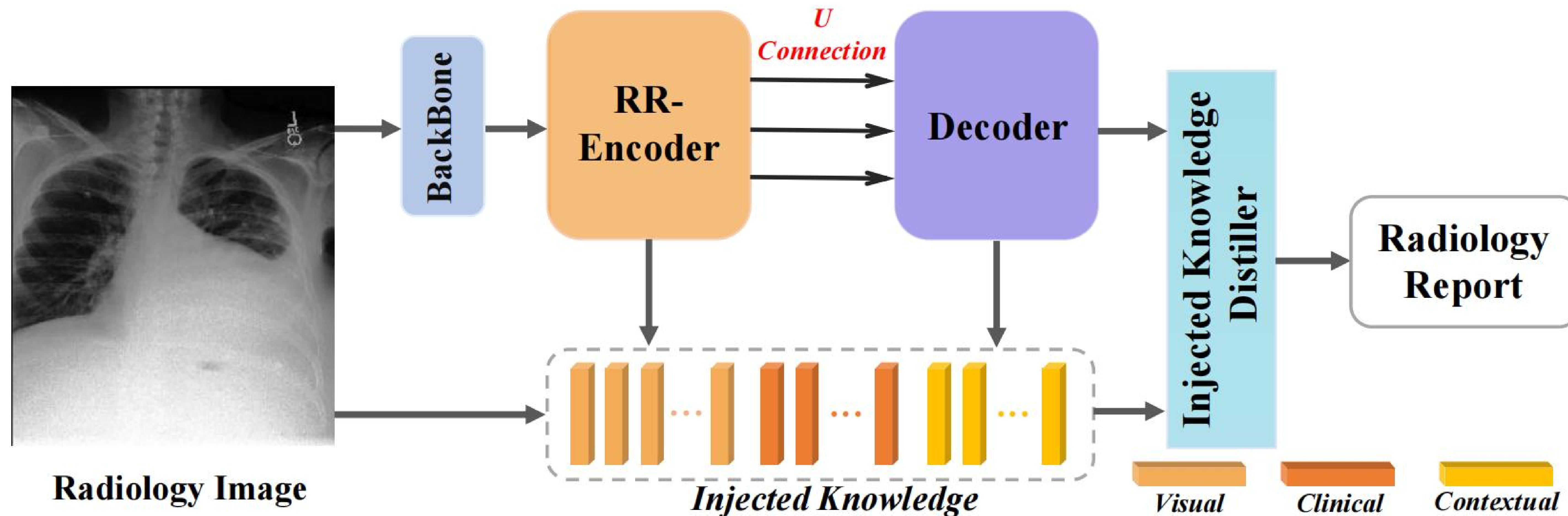


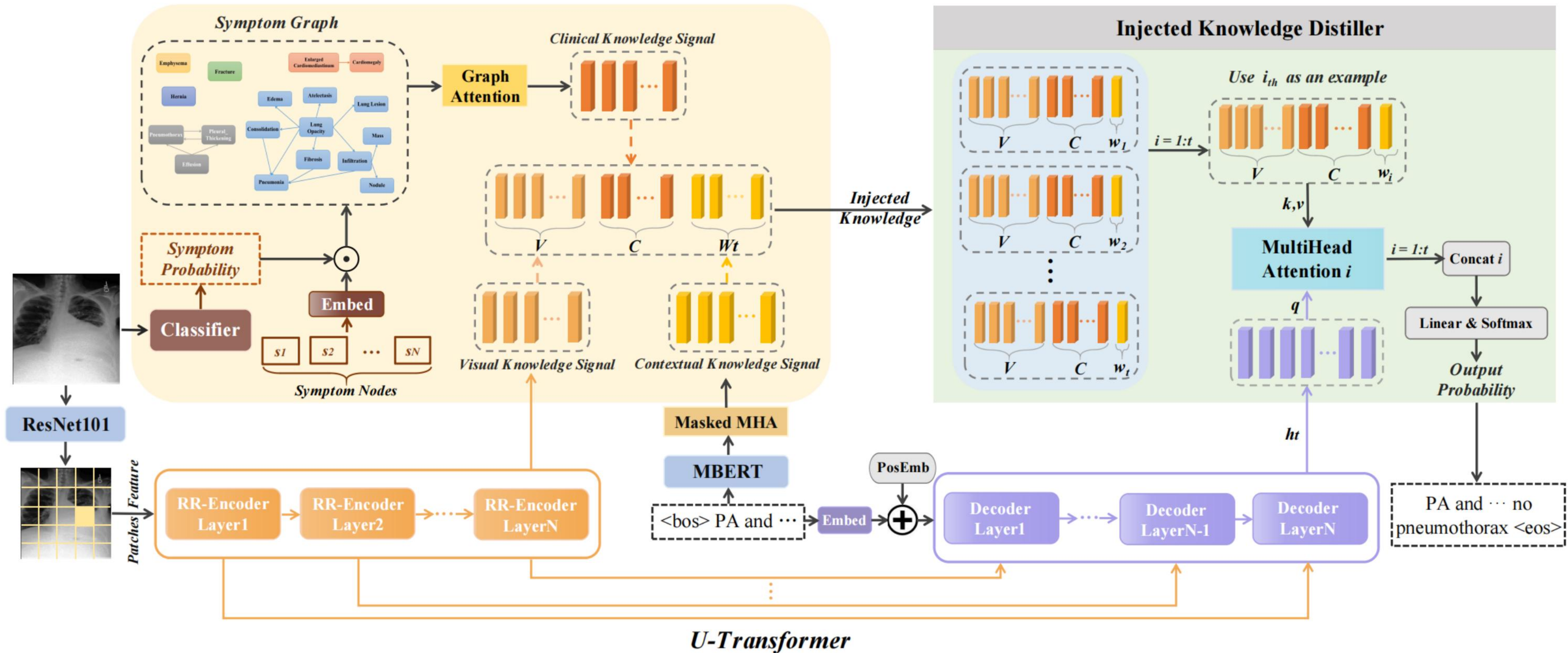
Clinical Radiology Report

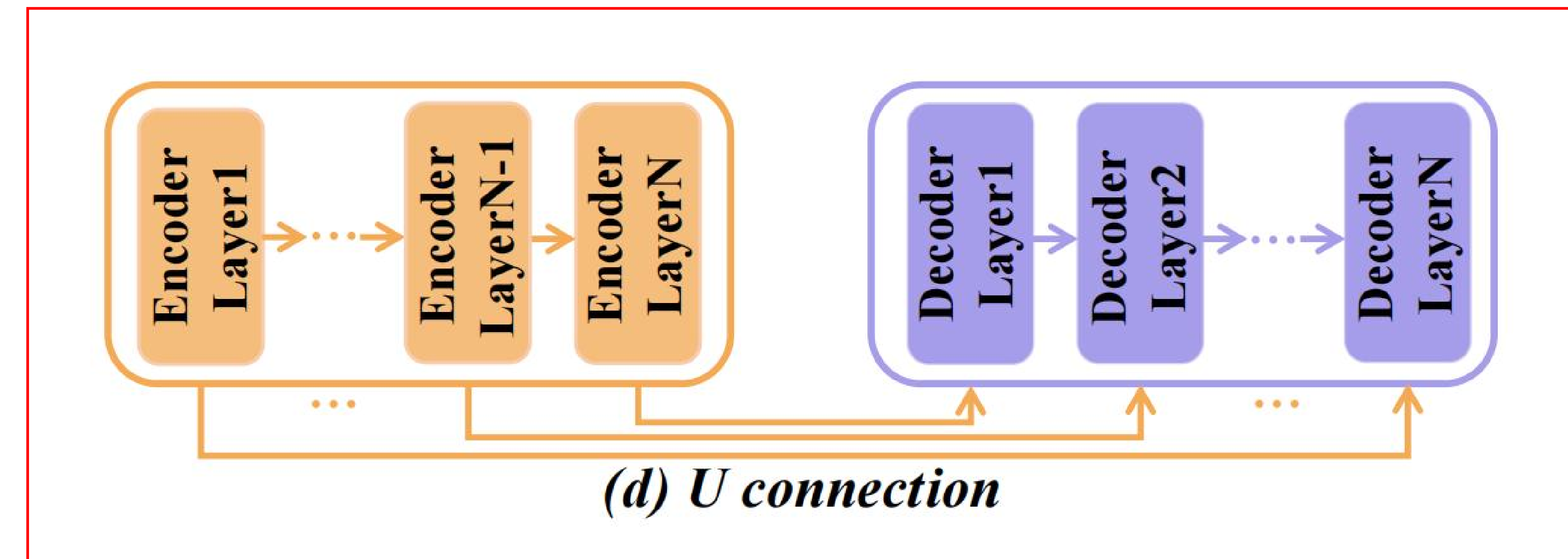
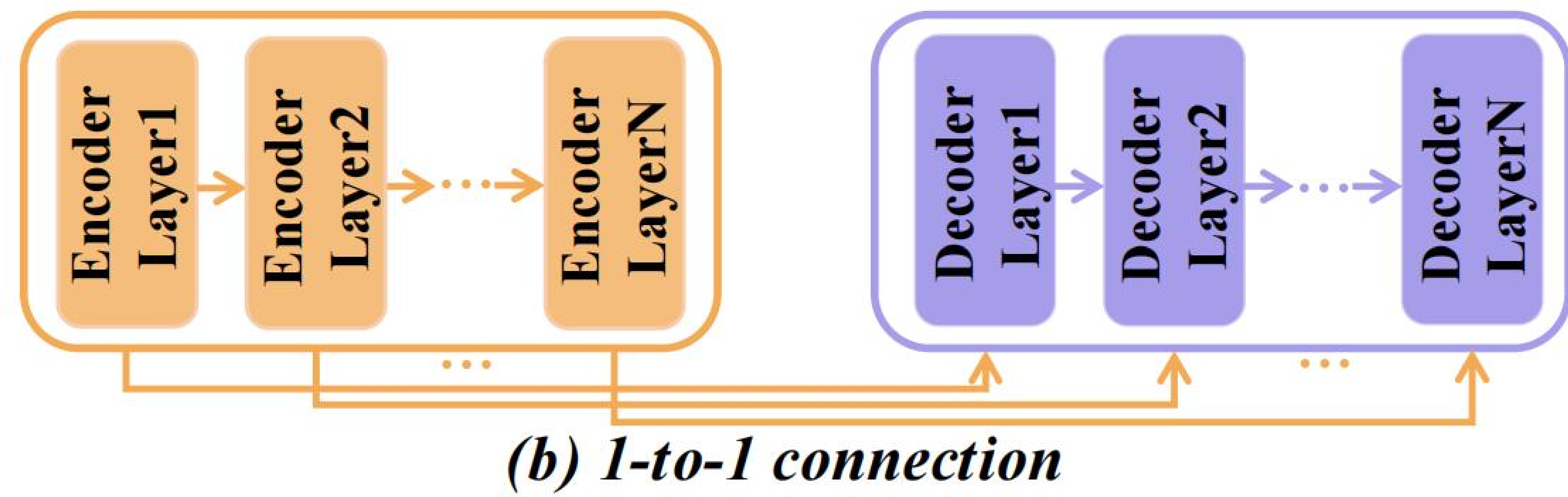
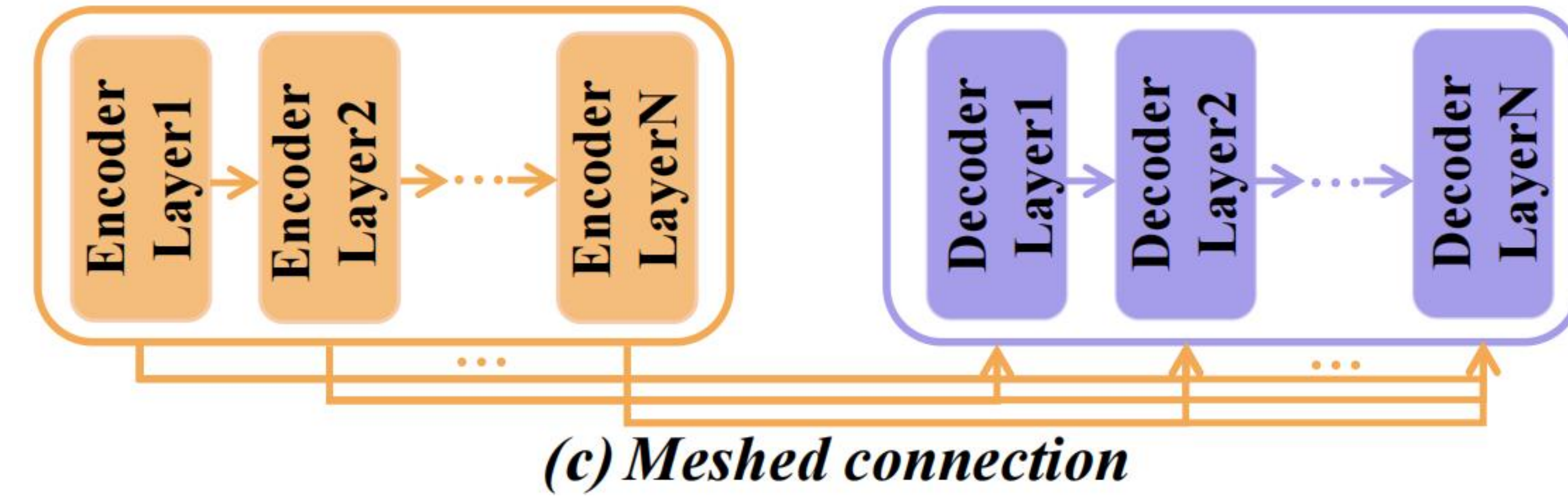
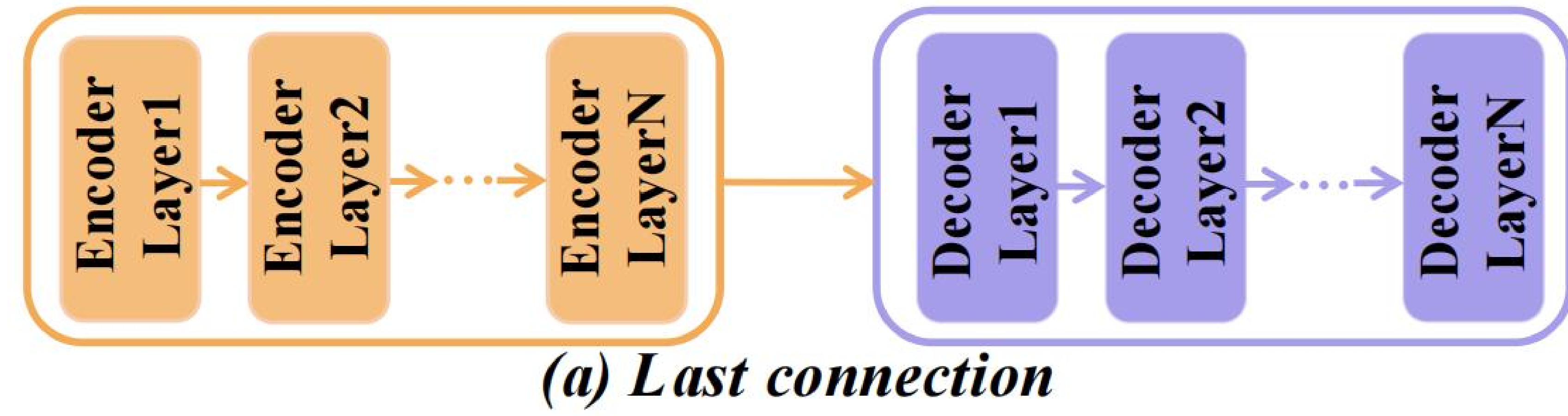
Findings: There is no focal consolidation, pleural effusion or pneumothorax. **Bilateral nodular** opacities that most likely represent nipple shadows. The cardiomeastinal silhouette is normal. Clips project over the **left lung**, potentially within the breast. The imaged **upper abdomen** is unremarkable.

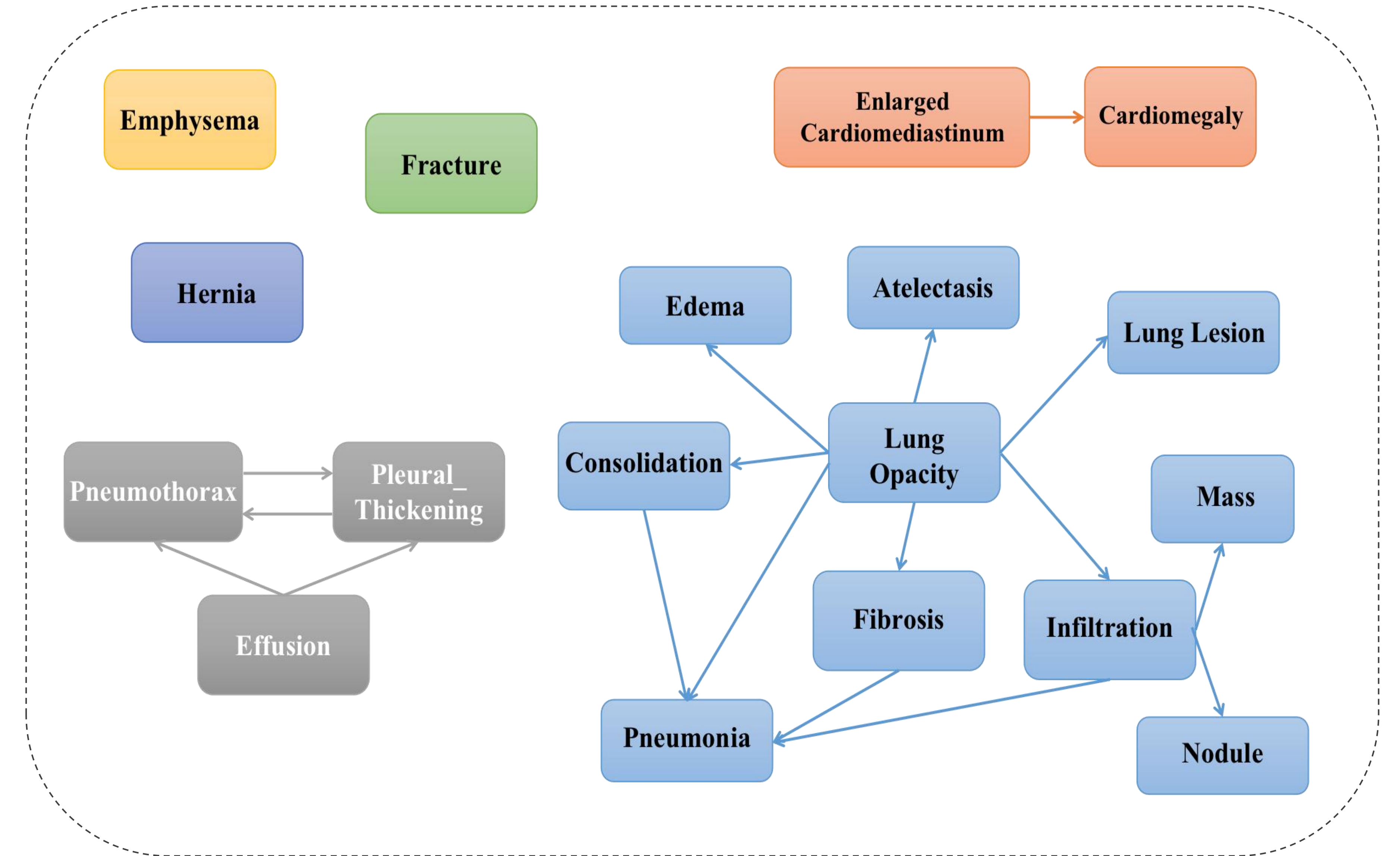
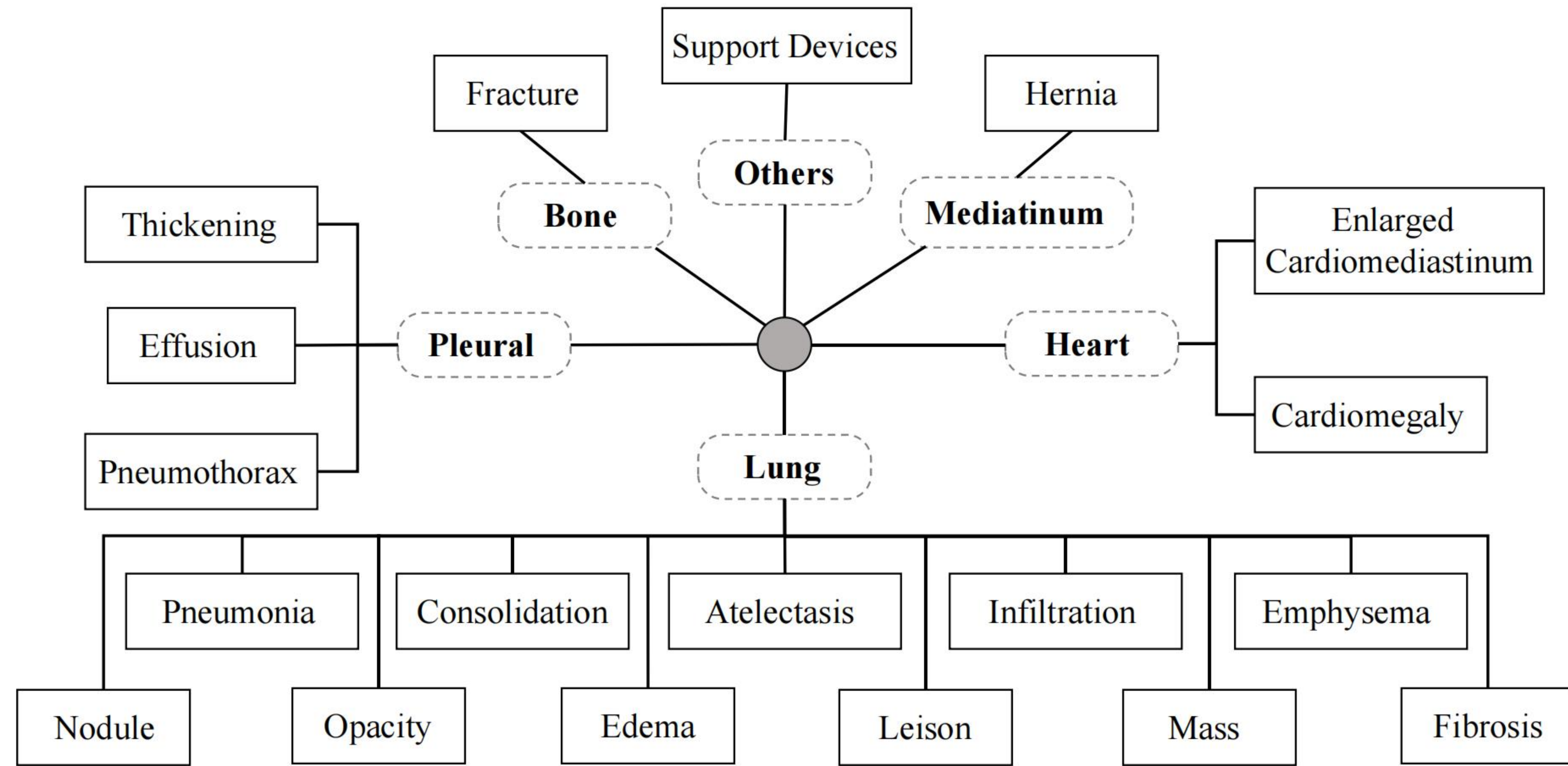


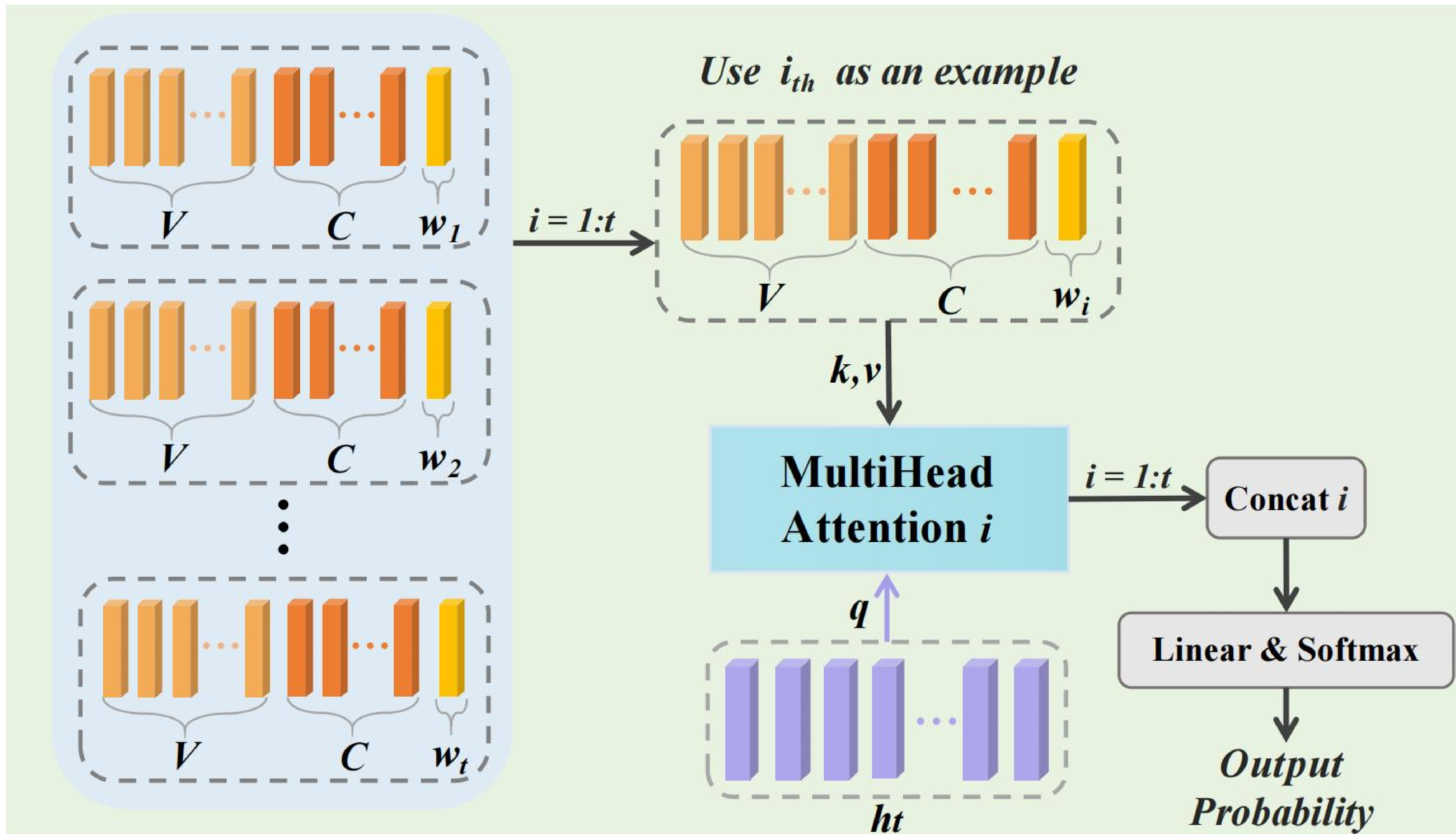
PA and lateral views of chest demonstrate **an extensive left-sided pleural effusion** with **compressive atelectasis**. An underlying pneumonia cannot be excluded. A tiny right pleural effusion may also be present. **The cardiac silhouette also appears enlarged**, but it is difficult to completely assess the left border given the large pleural effusion. The right lung is clear of focal opacities worrisome for pneumonia. There is no pneumothorax.












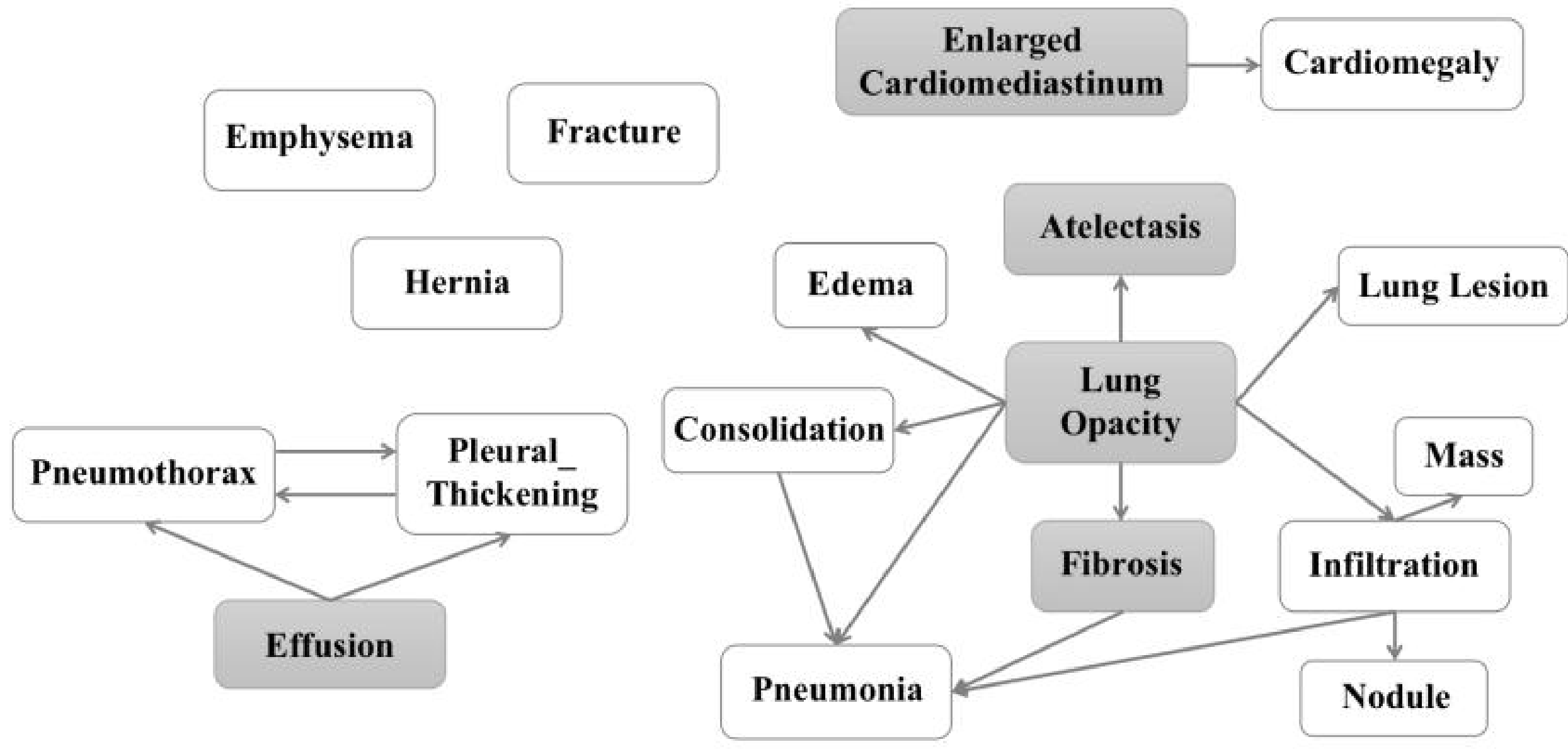
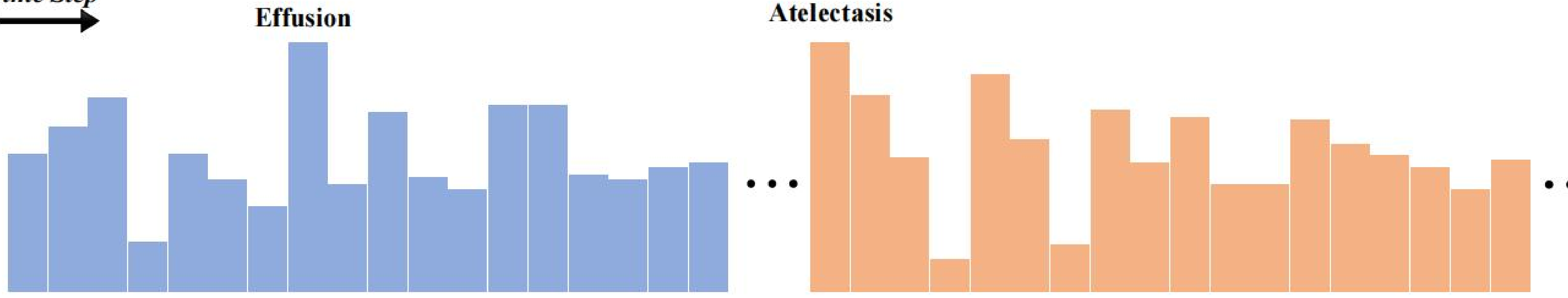
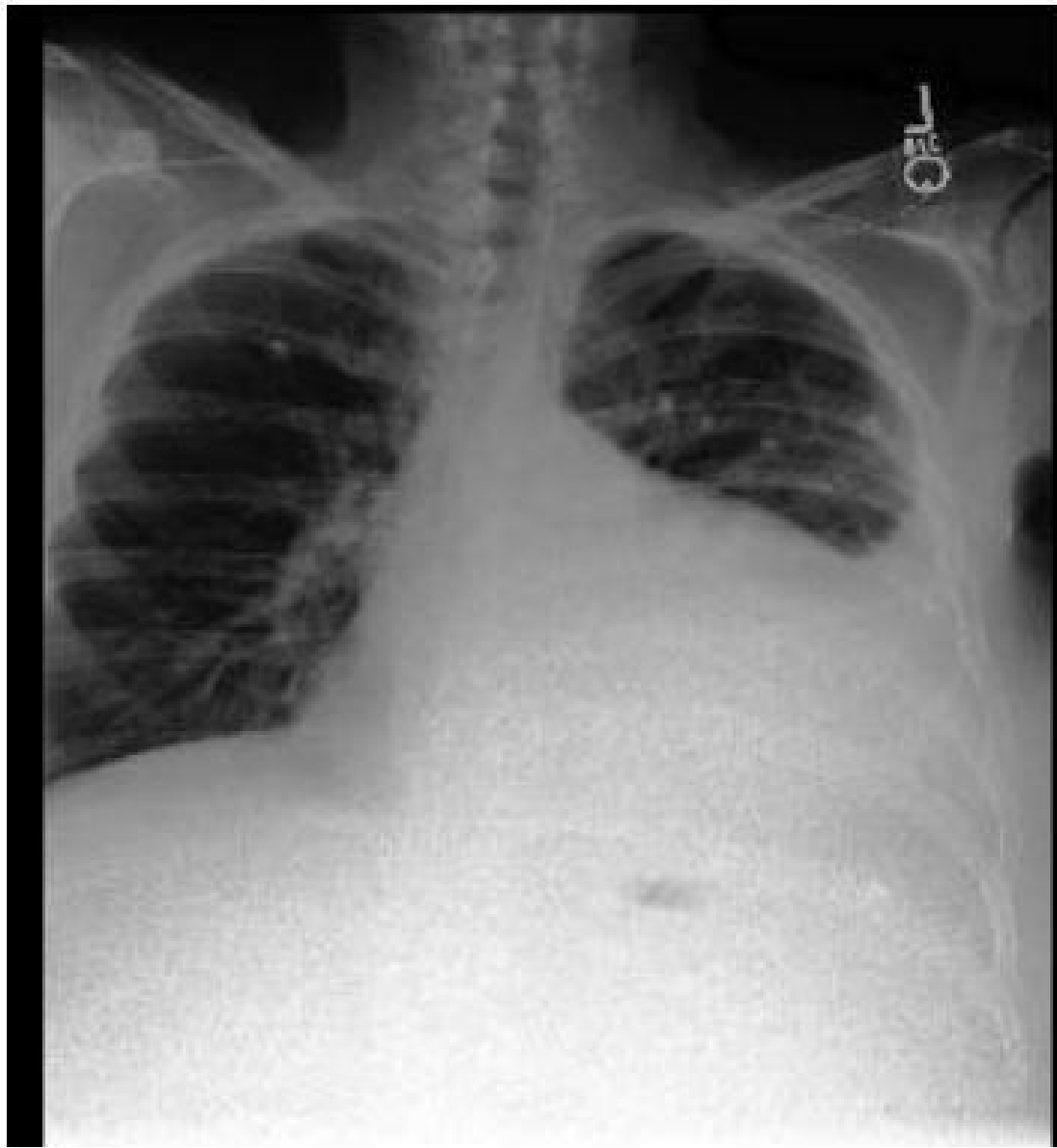
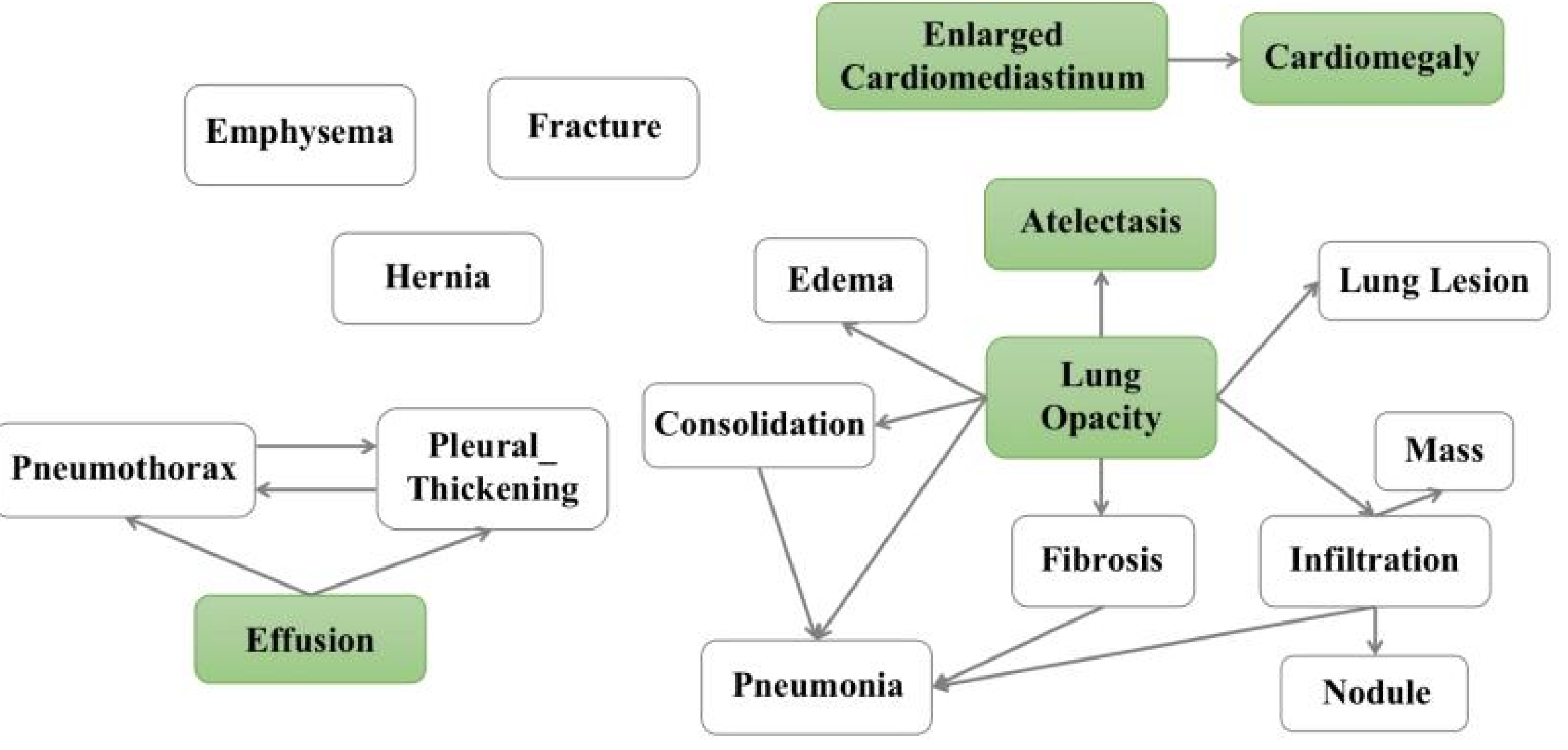

$$\tilde{F}_i = [\tilde{X}, \tilde{C}, \tilde{w}_i] \quad i \in [1, t], \quad \tilde{w}_i \in \tilde{W}_t$$

$$\text{MHA}_i(h_t, \tilde{F}_i) = \text{Softmax} \left(\frac{QK_i^T}{\sqrt{d_n}} \right) V_i$$


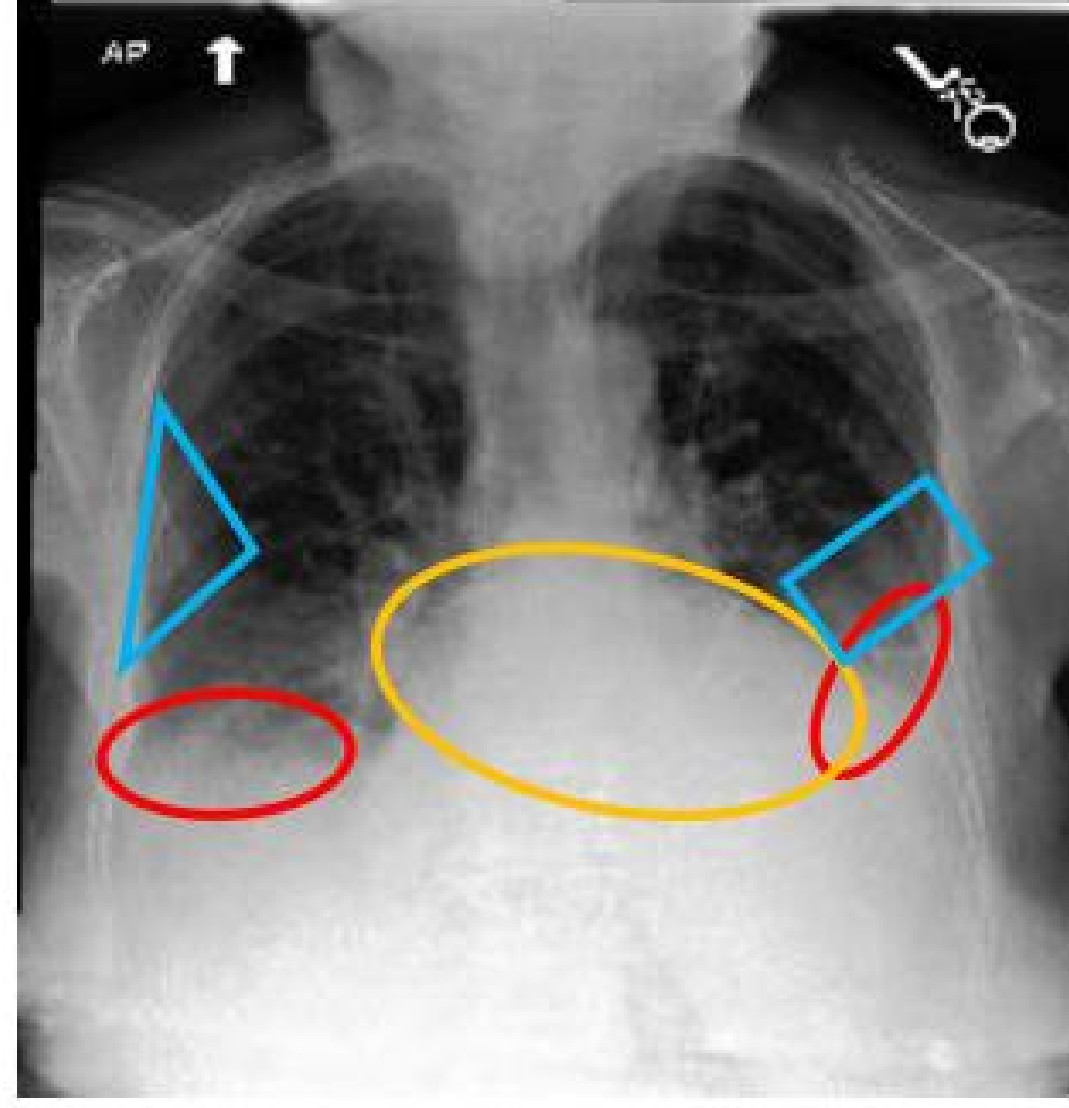
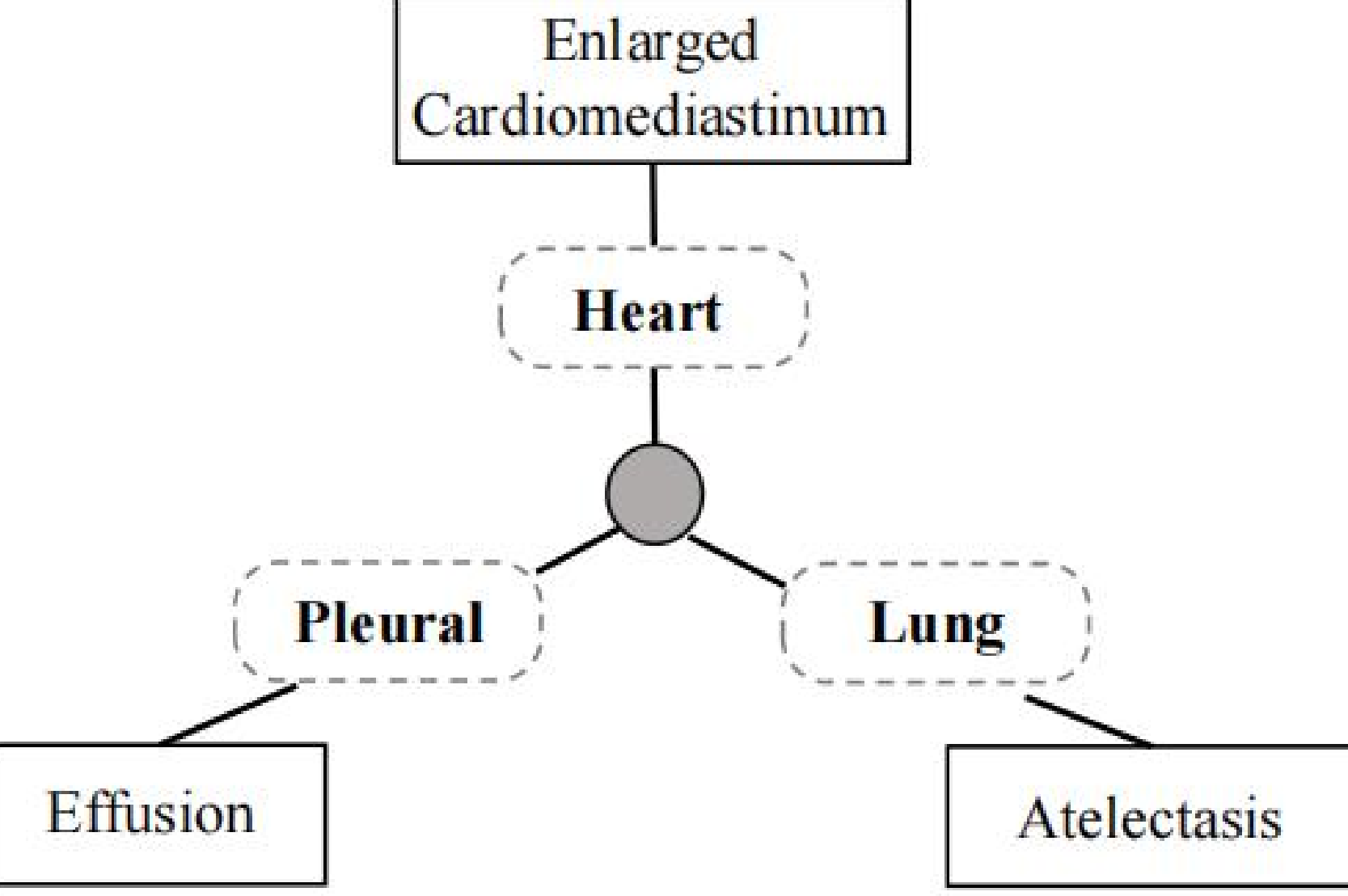

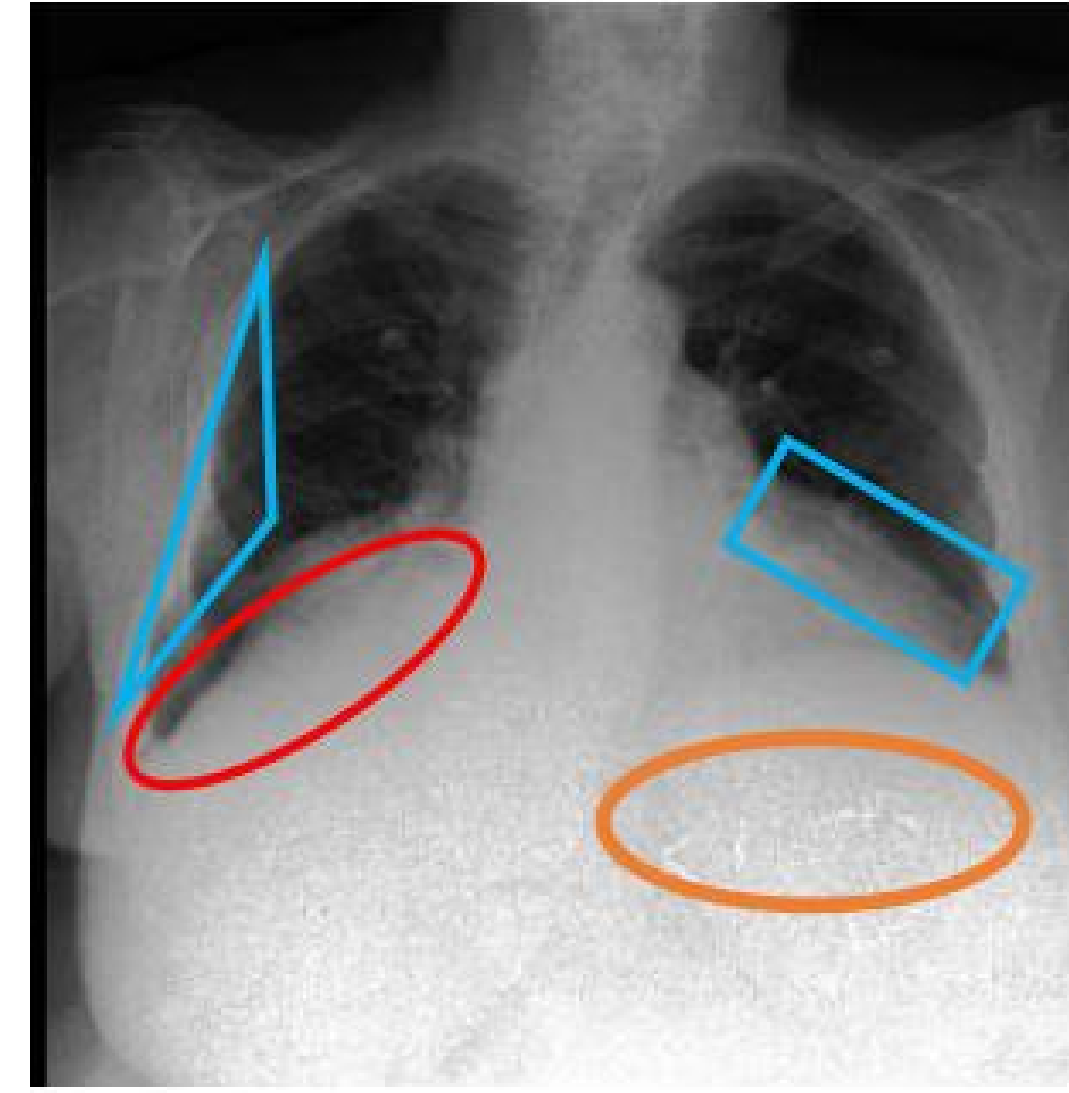
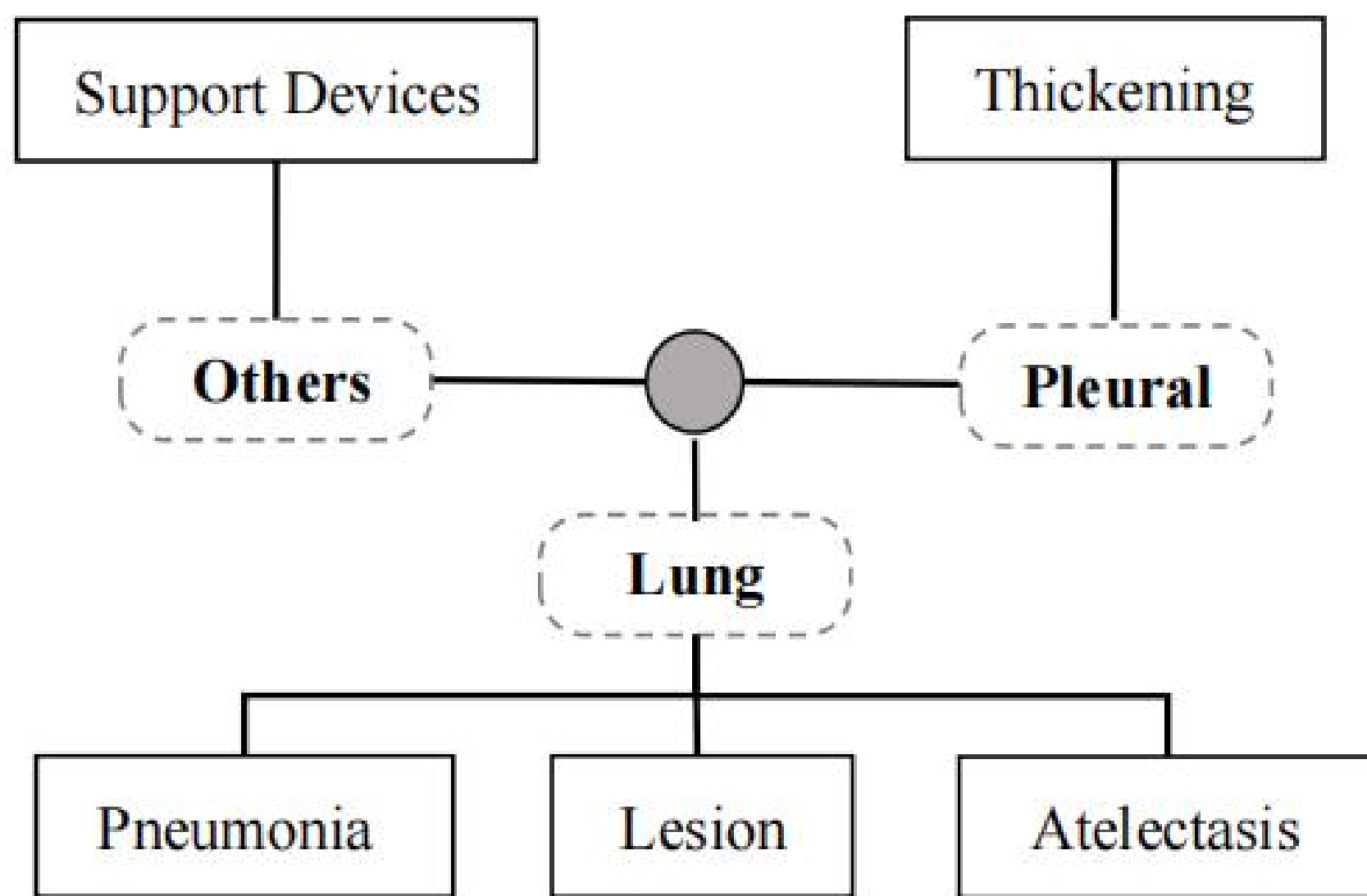
$$Q = h_t W^Q, \quad K_i = \tilde{F}_i W^K, \quad V_i = \tilde{F}_i W^V$$

$$\tilde{h}_t = \text{Concat}(\text{MHA}_1(h_t, \tilde{F}_1), \dots, \text{MHA}_t(h_t, \tilde{F}_t))$$

$$y_{t+1} \sim p_{t+1} = \text{Softmax} \left(\tilde{h}_t W_p + b_p \right)$$

Radiology Image	Initial Symptom Graph	Contribution Degree of Each Symptom
		<p>Time Step →</p>  <p>Frontal and Lateral views of the chest were obtained. There are moderate bilateral pleural effusions</p> <p>Frontal and Lateral views of the chest were obtained. There are moderate bilateral pleural effusions with overlying atelectasis</p>
		<p>Time Step →</p>  <p>Frontal and lateral views of the chest were obtained . There is a large left pleural effusion</p> <p>Frontal and lateral views of the chest were obtained. There is a large left pleural effusion ... There is no evidence of pneumothorax. The cardiac</p>

Dataset	Method	NLG Metric						CE Metric		
		Bleu1	Bleu2	Bleu3	Bleu4	Meteor	Rouge_L	Precision	Recall	F1
IU- Xray	SentSAT+KG [44]	0.441	0.291	0.203	0.147	—	0.367	—	—	—
	R2GenCMN [3]	0.470	0.304	0.219	0.165	0.187	0.371	—	—	—
	PPKED [27]	0.483	0.315	0.224	0.168	0.190	0.376	—	—	—
	AlignTrans [41]	0.484	0.313	0.225	0.173	0.204	0.379	—	—	—
	Contrastive [28]	0.492	0.314	0.222	0.169	0.193	0.381	—	—	—
	XPRONET [38]	0.525	0.357	0.262	0.199	0.220	0.411	—	—	—
	Ours	0.525	0.360	0.251	0.185	0.242	0.409	—	—	—
MIMIC CXR	Up-Down [1]	0.317	0.195	0.130	0.092	0.128	0.267	0.320	0.231	0.238
	Att2in [33]	0.325	0.203	0.136	0.096	0.134	0.276	0.322	0.239	0.249
	R2GenCMN [3]	0.353	0.218	0.145	0.103	0.142	0.277	0.333	0.273	0.276
	PPKED [27]	0.360	0.224	0.149	0.106	0.149	0.284	—	—	—
	Contrastive [28]	0.350	0.219	0.152	0.109	0.151	0.283	0.352	0.298	0.303
	AlignTrans [41]	0.378	0.235	0.156	0.112	0.158	0.283	—	—	—
	XPRONET [38]	0.344	0.215	0.146	0.105	0.138	0.279	—	—	—
	Ours	0.393	0.243	0.159	0.113	0.160	0.285	0.371	0.318	0.321

Radiology Image	Labeled Image	Related Symptom Graph	Ground Truth and Generated Report
			<p style="text-align: center;">Ground Truth</p> <p>There are moderate bilateral pleural effusions with overlying atelectasis, underlying consolidation not excluded. Mild prominence of the interstitial markings suggests mild pulmonary edema. The cardiac silhouette is mildly enlarged. The mediastinal contours are unremarkable. There is no evidence of pneumothorax.</p> <p style="text-align: center;">KIUT</p> <p>Frontal and Lateral views of the chest were obtained. There are moderate bilateral pleural effusions with overlying atelectasis underlying consolidation is not excluded. The cardiac silhouette remains enlarged. The aorta is calcified and tortuous. No overt pulmonary edema is seen.</p>
			<p style="text-align: center;">Ground Truth</p> <p>PA and lateral views of the chest. There is mild bibasilar atelectasis. Persistent slight elevation of the right hemidiaphragm. There is no focal parenchymal opacities concerning for pneumonia. There is no pleural effusion or pneumothorax. The cardiac and mediastinal silhouettes are stable. Multiple surgical clips in the left upper quadrant.</p> <p style="text-align: center;">KiUT</p> <p>Frontal and lateral views of the chest were obtained. There are low lung volumes which accentuate the bronchovascular markings. Given this there is mild elevation of the right hemidiaphragm. No definite focal consolidation is seen. There is no pleural effusion or pneumothorax. The cardiac and mediastinal silhouettes are stable. Surgical clips are seen in the upper abdomen.</p>

- A novel framework for radiology report generation that focuses on extracting and distilling multi-level information and multiple injected knowledge.
- KiUT encodes images with the extrinsic and intrinsic relationships among image regions and decodes words by an injected knowledge distiller.
- A novel U-connection schema to exploit the interaction between the encoder and decoder, which is unprecedented for other architectures in such a cross-modal scenario.
- Future work: sophisticated knowledge constructing and structured report template filling solutions.

Thank You