1. APPENDIX

1.1. Implementation Details.

Datasets. Due to resource constraints, we randomly sample 30k studies from the training set while ensuring that all clinically meaningful subgroups are covered.

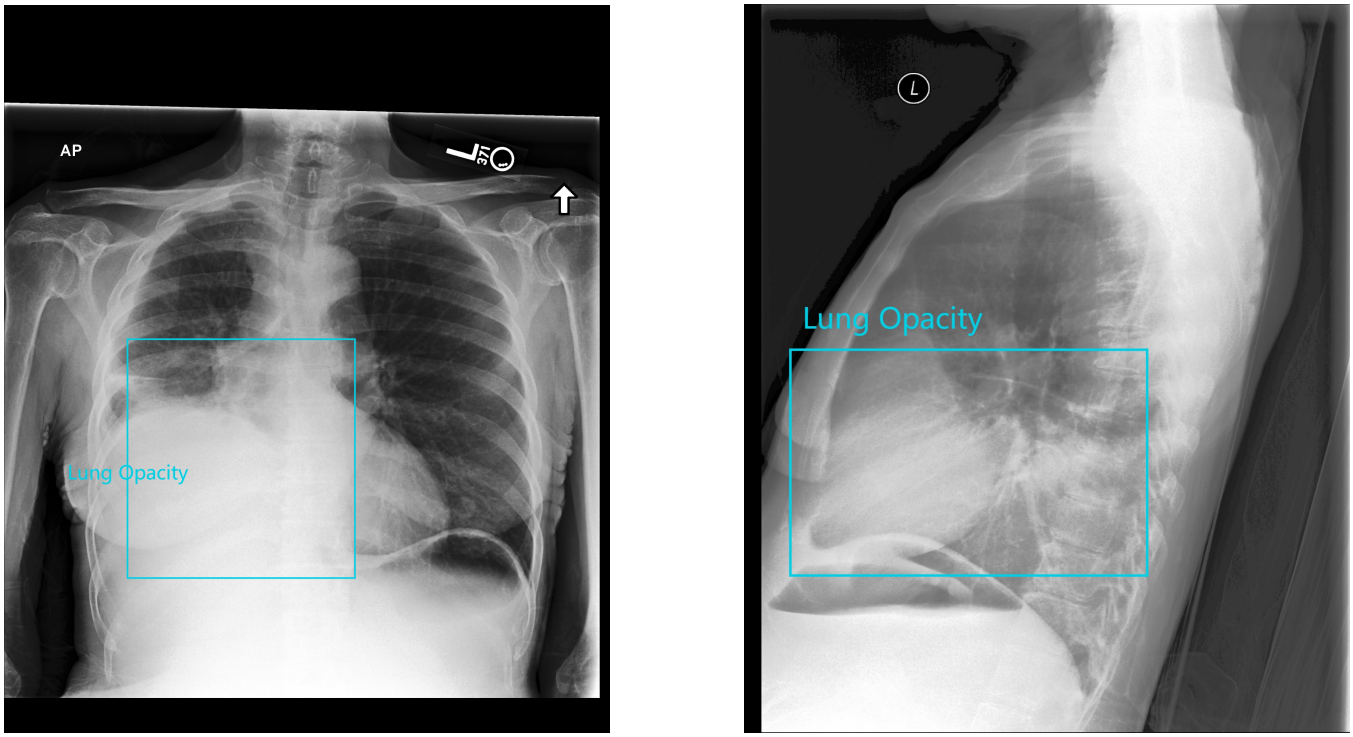
Model Implementation. We consider two baseline implementations. The first is a re-implementation of MARIA2, which adopts a multimodal training strategy similar to LLaVA, where image features extracted by the vision encoder are directly concatenated to form the image representation. The second follows an expert-fusion design, inspired by the idea of METransformer, but adapted to our setting by as signing subgroup-specific experts to different input configurations. For each study, only the experts corresponding to the available views pr ocess the images, and their outputs are fused to produce a single unified image representation.

For both approaches, the resulting image features are projected into the Vicuna-7B language space via an MLP projection module, and then concatenated with textual features to form the final multimodal input. We employ LoRA for parameter-efficient fine-tuning, AdamW as the optimizer, and a cosine learning rate scheduler. To stabilize training under limited GPU memory, gradient accumulation and mixed precision are used.

Model training. For each baseline, we conduct independent hyperparameter tuning to ensure comparable performance on the official test set in terms of standard metrics (e.g., BLEU-4, CheXbert F1). This setup eliminates discrepancies arising from architectural differences and allows us to focus on the effectiveness of ARCE in revealing subgroup-level disparities.

All experiments are conducted on two NVIDIA RTX 4090 GPUs.

1.2. Case study.

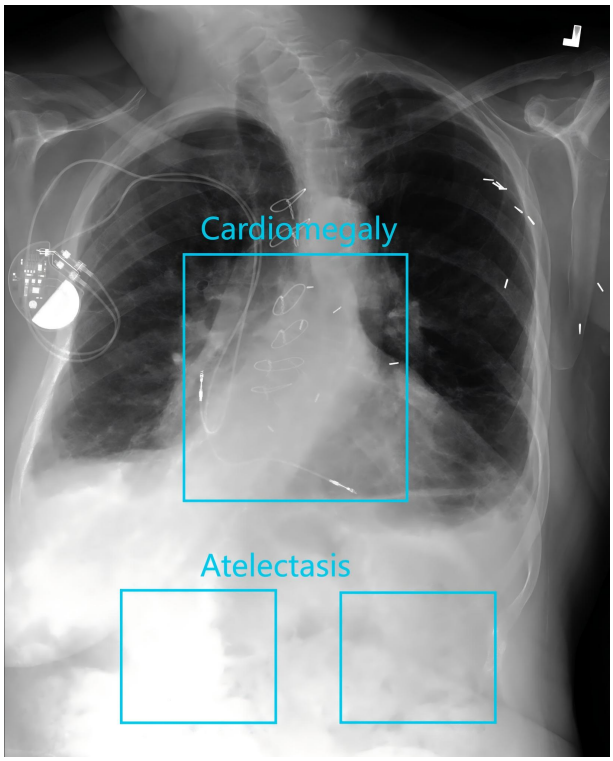


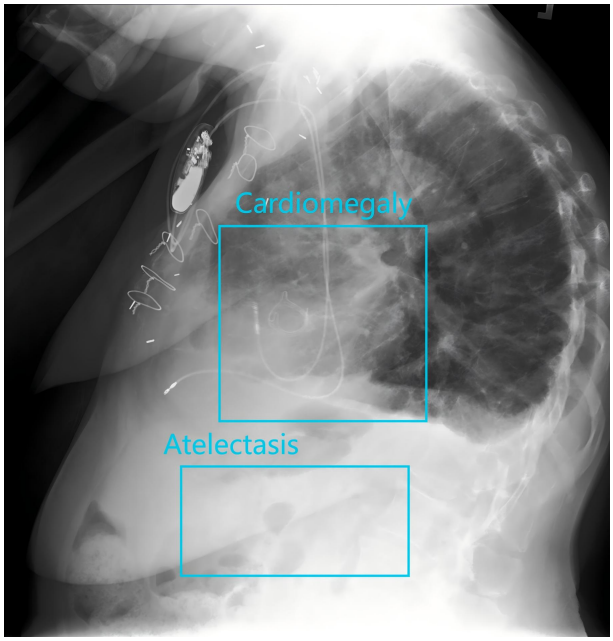
Ground Truth: Chest PA and lateral radiograph demonstrates a markedly elevated right hemidiaphragm with adjacent compressive atelectasis or consolidation. Minimal blunting of the posterior costophrenic angle may indicate a small right pleural effusion. Left lung is clear. Cardiomediastinal borders are unremarkable.

MODEL$_1$: A single frontal and lateral chest radiograph demonstrates a large right pleural effusion with adjacent air-fluid level. The left lung is clear, without focal consolidation or pneumothorax. There are no other concerning findings in the cardiomediastinal silhouette.

MODEL$_2$: Diffuse pulmonary opacities are seen with a left-sided predominance. The cardiomediastinal silhouette is within normal limits for technique. No pleural effusion or pneumothorax identified. There may be small bilateral apical scarring as well as an air bronchogram at the right lung base.

Fig. 1. Case study on the CFCL subgroup. Image: input frontal chest X-ray and Lateral chest X-ray. Text: generated reports from different models compared with the ground truth.



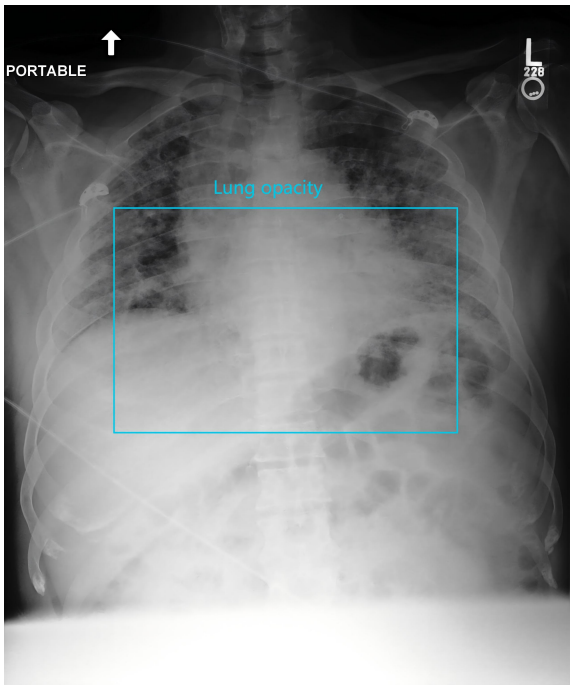


Ground Truth: There is enlargement of the cardiac silhouette with some evidence of elevated pulmonary venous pressure. Intact midline sternal wires are seen in a patient with previous CABG procedure and a dual-channel pacemaker in place. Axial clips are again seen. Some mild atelectatic changes and possible small effusions are seen at the bases.

MODEL$_1$: The patient has undergone CABG surgery as evidenced by intact sternal wires on both sides. There are small bilateral pleural effusion which have increased since prior study. Bibasilar opacities likely represent atelectasis or aspiration versus pneumonia. No focal consolidation to suggest pneumonia is present. Cardiomediastinal silhouette appears stable.

MODEL$_2$: Diffuse pulmonary opacities are seen with a left-sided predominance. The cardiomediastinal silhouette is within normal limits for technique. No pleural effusion or pneumothorax identified. There may be small bilateral apical scarring as well as an air bronchogram at the right lung base.

Fig. 2. Case study on the CFCL subgroup. Image: input frontal chest X-ray and Lateral chest X-ray. Text: generated reports from different models compared with the ground truth.



Ground Truth: Single, AP, upright, portable view of the chest was obtained. There are increased interstitial markings. The cardiac and mediastinal silhouettes are stable. There is slight blunting of both costophrenic angles, felt most likely due to overlying soft tissues, but a trace pleural effusions be difficult to exclude. No right pleural effusion is seen. There is no pneumothorax.

MODEL$_1$: There are diffuse bilateral interstitial opacities with bibasilar atelectasis and pulmonary edema, which could be mild to moderate in severity. Cardiomediastinal silhouette appears unchanged from the prior study. No pneumothorax or pleural effusion identified. Mild pulmonary edema is also noted. No large pleural effusion seen.

MODEL$_2$: There is increased opacification at the left base which may be due to aspiration or infection, but an infectious process cannot be excluded given prior pneumonia and recent surgery. The right lung appears clear.

Fig. 3. Case study on the CF subgroup. Image: input frontal chest X-ray. Text: generated reports from different models compared with the ground truth.

1.3. Details of some experimental results.

Table 1. Subgroup Results of MODELS

Subgroup	BLEU1	Rouge_L	neg_Macro_14_F1	neg_Micro_14_F1
Subgroup Results of MODEL$_1$				
CF	25.37	20.78	20.13	35.54
CF+CL	28.77	25.08	20.99	33.23
CF+CL	26.51	20.09	29.73	50.90
CF+PF+CL	28.80	24.09	25.78	38.97
Subgroup Results of MODEL$_2$				
CF	25.65	20.02	22.77	39.28
CF+CL	28.79	23.01	21.04	33.14
CF+CL	26.70	19.83	27.25	47.87
CF+PF+CL	27.83	21.86	27.79	38.76
Subgroup Results of MODEL$_3$				
CF	26.73	19.89	23.06	39.75
CF+CL	28.92	22.19	21.06	30.85
CF+CL	26.31	18.12	27.92	48.49
CF+PF+CL	28.66	20.92	26.48	38.43

Table 2. Subgroup Results of MODELS

Subgroup	BLEU1	Rouge_L	neg_Macro_14_F1	neg_Micro_14_F1
Subgroup Results of MODEL$_1$				
RARE	27.25	23.60	11.37	12.90
COMMON	27.03	21.22	24.51	48.50
Subgroup Results of MODEL$_2$				
RARE	29.01	22.00	13.46	15.97
COMMON	26.76	20.40	23.97	47.07
Subgroup Results of MODEL$_3$				
RARE	27.88	20.36	16.60	18.87
COMMON	28.66	20.92	24.43	47.46