Deep Learning for Chest X-ray Report Generation

Presented By:-

Jaikrishna Bollampalli

Background

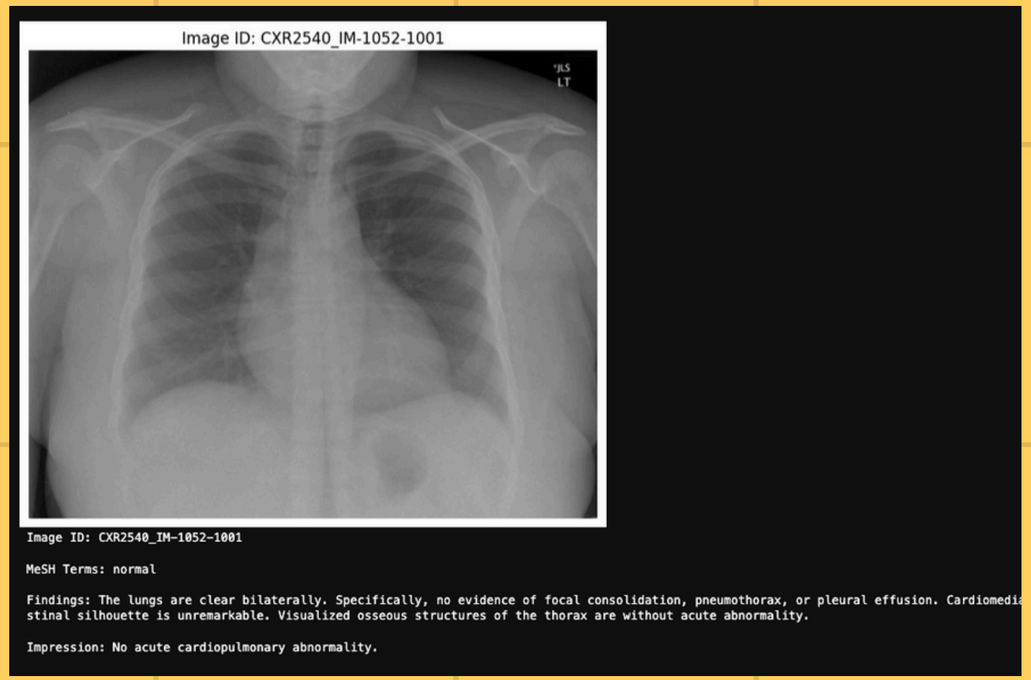
- Radiologists spend significant time generating reports for medical images
- High volume of chest X-rays in clinical settings creates workflow bottlenecks
- Automated report generation could improve efficiency and consistency

Motivation

- Assist radiologists by providing initial report drafts
- Reduce reporting time and backlog
- Standardize report language and structure
- Support clinical decision-making with AI assistance

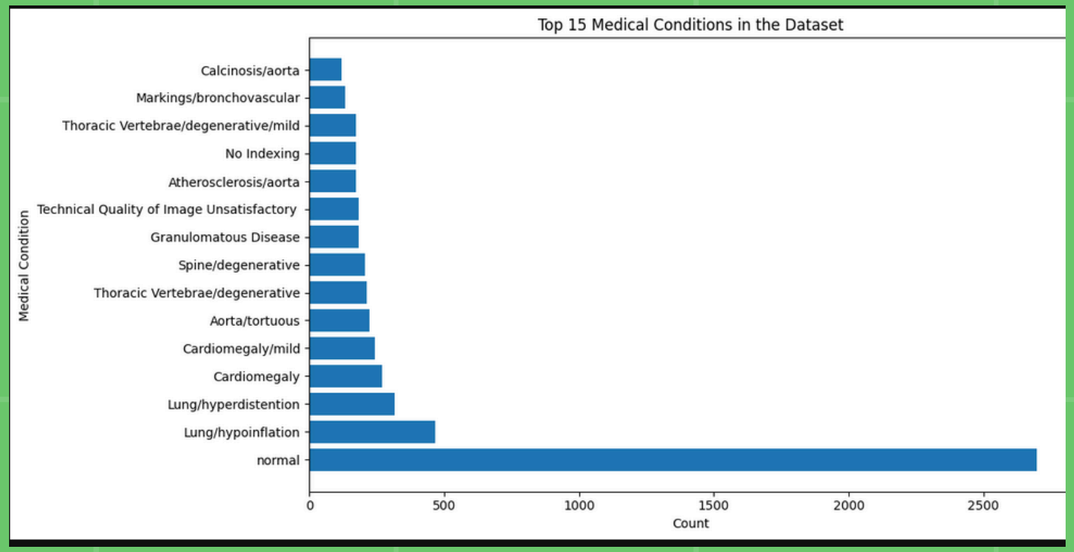
Dataset – IU X-Ray Overview

- Source: Indiana University Chest X-Ray Collection
- Total Images: ~7,000 chest X-rays (.png)
- Reports: Paired XML files containing two key sections:
- FINDINGS: Objective observations
- IMPRESSION: Diagnostic conclusions



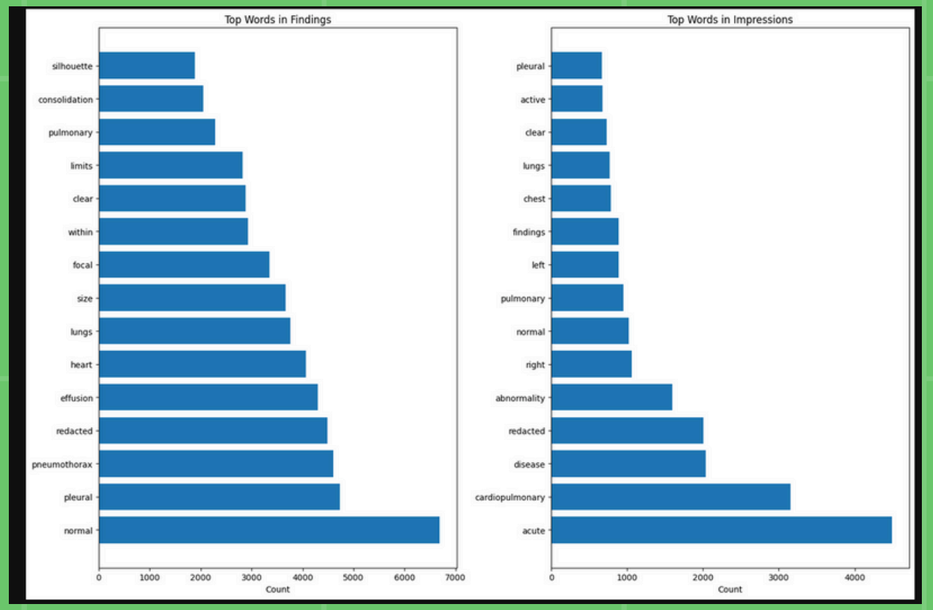
Data Exploration & Analysis

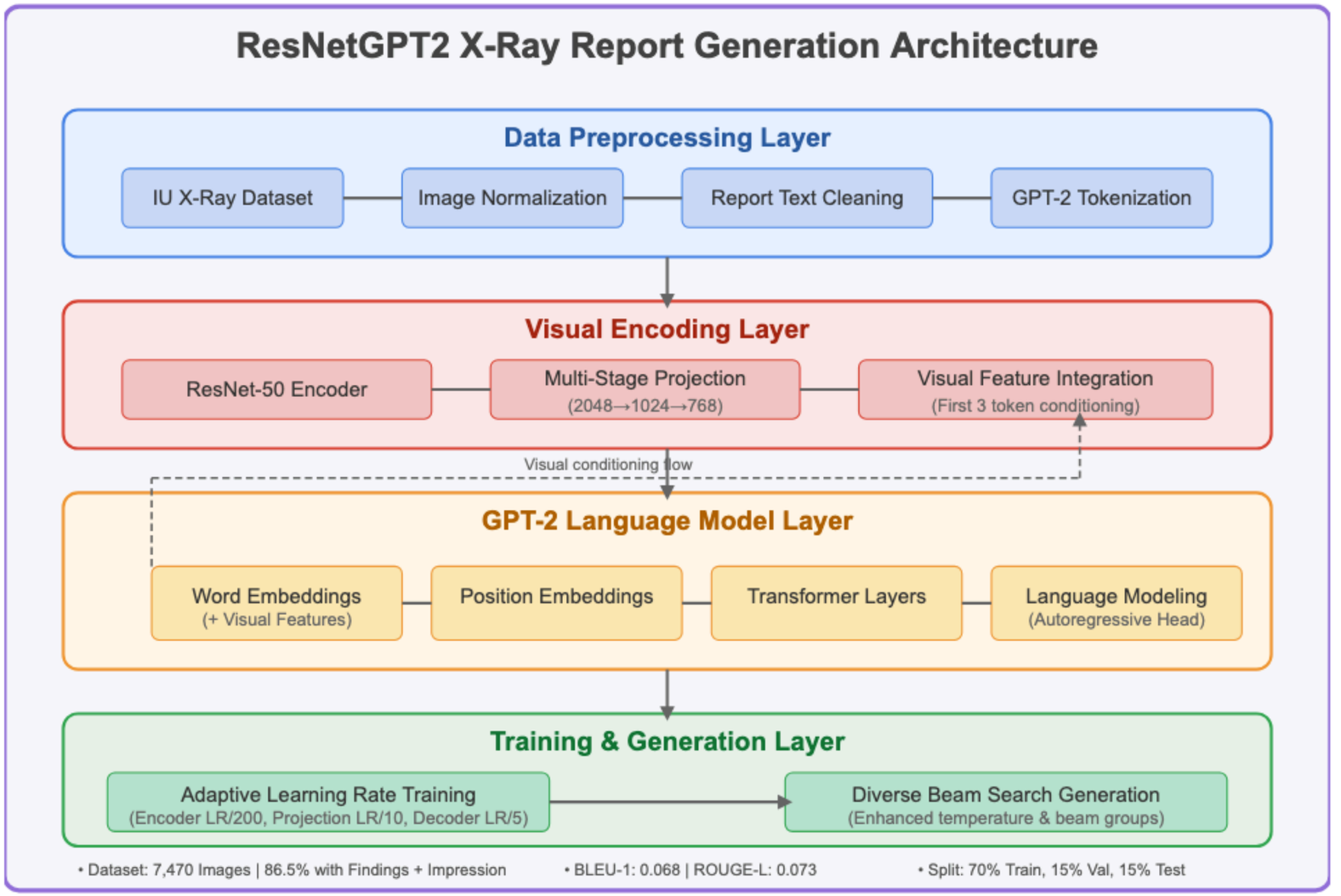
- Total Image-Report Pairs: ~3,800
- Average Word Count:
- Findings: ~25
- Impression: ~18
- Full Report: ~42
- MeSH Terms per Report: ~3.2 (avg)



Report Completeness:

- Both sections: 64.2%
- Only findings: 18.7%
- Only impression: 9.4%
- Neither: 7.7%



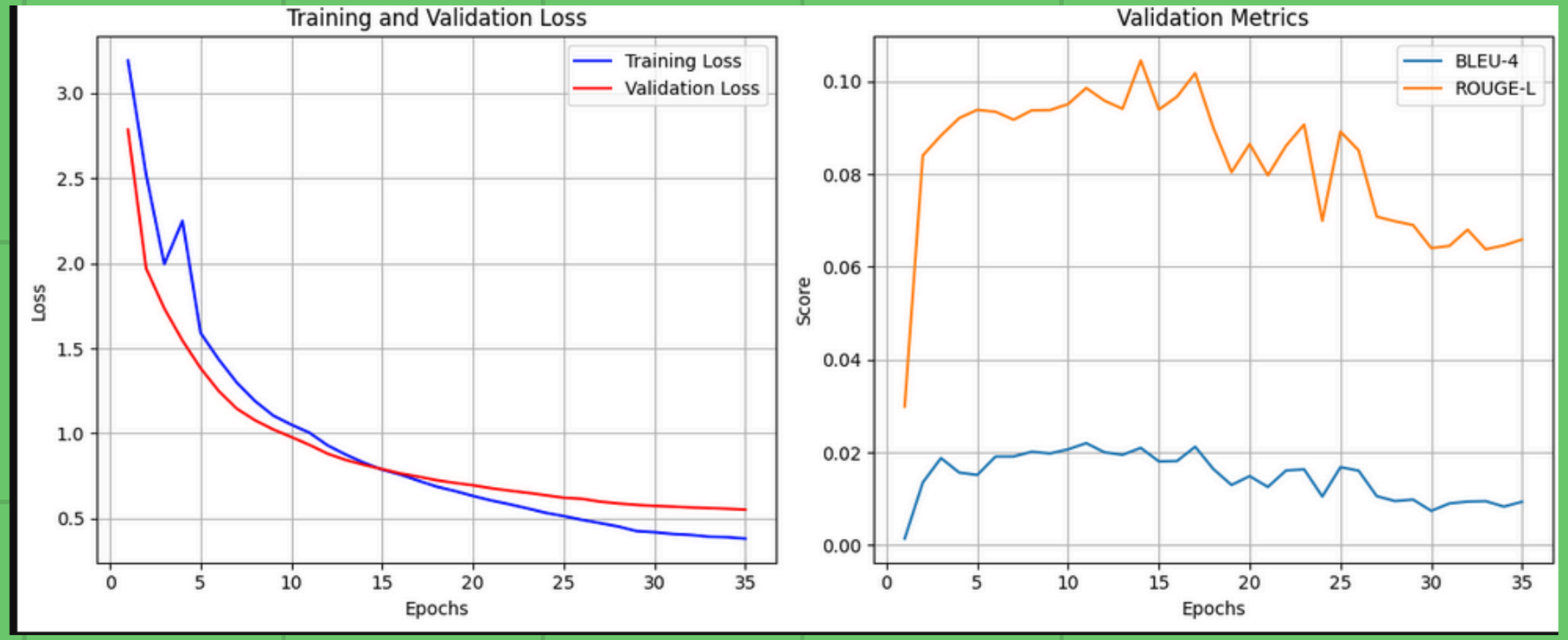

ResNetGPT2 X-Ray Report Generation Architecture
Data Preprocessing Layer
IU X-Ray Dataset
Image Normalization
Report Text Cleaning
GPT-2 Tokenization
Visual Encoding Layer
ResNet-50 Encoder
Multi-Stage Projection
(2048→1024→768)
Visual Feature Integration
(First 3 token conditioning)
Visual conditioning flow
GPT-2 Language Model Layer
Word Embeddings
(+ Visual Features)
Position Embeddings
Transformer Layers
Language Modeling
(Autoregressive Head)
Training & Generation Layer
Adaptive Learning Rate Training
(Encoder LR/200, Projection LR/10, Decoder LR/5)
Diverse Beam Search Generation
(Enhanced temperature & beam groups)
• Dataset: 7,470 Images | 86.5% with Findings + Impression
• BLEU-1: 0.068 | ROUGE-L: 0.073
• Split: 70% Train, 15% Val, 15% Test

Train-Test Pipeline Setup

```
Epoch 35/40
Epoch 35 [TRAIN]: 100% ████████████████████ 327/327 [01:16<00:00,  4.59it/s, loss=0.546]
[VALID]: 100% ████████████████████ 70/70 [23:05<00:00, 20.05s/it, loss=0.337]

Batch 0 samples:
Ref: Cardiomegaly with central pulmonary vascular prominence and coarsened interstitial markings, concern...
Gen: The examination reveals  1. The examination reveals: 2. Age-indeterminate wedging of several midthor...
Ref: 1. No acute abnormality. 2. Dextroscoliosis of thoracic spine unchanged. ....
Gen: The heart size and cardiomediastinal silhouette are within normal limits. Pulmonary vasculature appe...

Batch 1 samples:
Ref: No acute cardiopulmonary disease....
Gen: 1. No evidence of active cardiopulmonary disease. 2. Hyperexpansion and emphysematous changes 3. Age...
Ref: Heart size or great lungs are clear. Calcified 5 mm granuloma in the right upper lobe underneath the...
Gen: 1. The examination reveals: 2. Technically limited exam. 3. Incidental note  of large cervical spine...

Batch 2 samples:
Ref: Small streaky opacity lateral right lung, subsegmental atelectasis versus scarring....
Gen: The examination reveals  Mild cardiomegaly. No effusions or edema. Minimal bibasal subsegmental atel...
Ref: No acute cardiopulmonary abnormality.....
Gen: The heart size is normal. Mediastinal silhouette and pulmonary vascularity are within normal limits....

Medical condition detection:
  Normal: F1=0.5043, P=0.3720, R=0.7829
  Cardiomegaly: F1=0.0328, P=0.0345, R=0.0312
  Effusion: F1=0.1560, P=0.0872, R=0.7400
  Pneumothorax: F1=0.0749, P=0.0401, R=0.5610
  Opacity: F1=0.0000, P=0.0000, R=0.0000
  Edema: F1=0.0496, P=0.0283, R=0.2000
  Train Loss: 0.3819
  Val Loss: 0.5523
  BLEU-1: 0.0644
  BLEU-4: 0.0093
  ROUGE-L: 0.0659

Sample prediction:
Reference: Cardiomegaly with central pulmonary vascular prominence and coarsened interstitial markings, concerning for interstitial pulm
ry edema.
Prediction: The examination reveals  1. The examination reveals: 2. Age-indeterminate wedging of several midthoracic vertebral bodies. T
may be due to old injury or dislocations. If there is no known history, recommend further evaluation with dedicated views of the thoraci
egments. Dr.   was notified of this finding at / on ,  and accepted receipt of the results. There is no evidence of active tuberculous p
ess. 3. Mild osteopenia
Checkpoint saved to ./IUXRAY_OUTPUT_PIPELINE/checkpoints/checkpoint_epoch_35.pth
Removed old checkpoint: ./IUXRAY_OUTPUT_PIPELINE/checkpoints/checkpoint_epoch_34.pth
Checkpoint saved to ./IUXRAY_OUTPUT_PIPELINE/checkpoints/latest_checkpoint.pth
  No improvement in metrics for 8 epochs
```

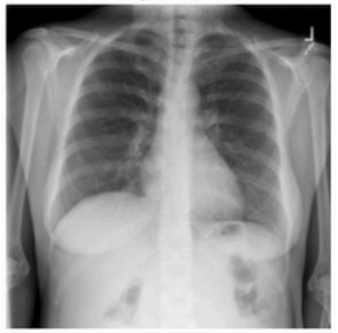
```
Epoch 37 [TRAIN]: 100% ████████████████████ 327/327 [01:16<00:00,  4.58it/s, loss=0.331]
[VALID]: 100% ████████████████████ 70/70 [23:03<00:00, 19.84s/it, loss=0.337]

Batch 0 samples:
Ref: Cardiomegaly with central pulmonary vascular prominence and coarsened interstitial markings, concern...
Gen: 1. No evidence of active cardiopulmonary disease. 2. Interval development of healing left sided rib ...
Ref: 1. No acute abnormality. 2. Dextroscoliosis of thoracic spine unchanged. ....
Gen: The examination reveals  Heart size is normal. Lungs are clear. No nodules, masses, or adenopathy. N...

Batch 1 samples:
Ref: No acute cardiopulmonary disease....
Gen: The examination reveals  Negative for acute cardiopulmonary disease. No evidence of active tuberculo...
Ref: Heart size or great lungs are clear. Calcified 5 mm granuloma in the right upper lobe underneath the...
Gen: The examination reveals  No evidence of acute cardiopulmonary process. Stable appearance of the ches...

Batch 2 samples:
Ref: Small streaky opacity lateral right lung, subsegmental atelectasis versus scarring....
Gen: The examination reveals  Heart size is normal. Lungs are clear. Calcified 5 mm right midlung granulo...
Ref: No acute cardiopulmonary abnormality.....
Gen: The heart size is normal and the mediastinal silhouette within normal limits. Pulmonary vasculature ...

Medical condition detection:
  Normal: F1=0.4545, P=0.3526, R=0.6390
  Cardiomegaly: F1=0.1326, P=0.1026, R=0.1875
  Effusion: F1=0.1680, P=0.0953, R=0.7100
  Pneumothorax: F1=0.0630, P=0.0343, R=0.3902
  Opacity: F1=0.0526, P=0.0488, R=0.0571
  Edema: F1=0.0738, P=0.0435, R=0.2444
  Train Loss: 0.3689
  Val Loss: 0.5453
  BLEU-1: 0.0649
  BLEU-4: 0.0091
  ROUGE-L: 0.0669

Sample prediction:
Reference: Cardiomegaly with central pulmonary vascular prominence and coarsened interstitial markings, concerning for interstitial pulmona
ry edema.
Prediction: 1. No evidence of active cardiopulmonary disease. 2. Interval development of healing left sided rib fractures, incompletely eva
luated. 3. Age-indeterminate  deformity of the midthoracic vertebral body. 4. Possible old T9 vertebral fracture is present. If there is no
known fracture, consider replacement with noncalcified tenderness from prior T11 exam. Otherwise, followup evaluation with dedicated rib se
ries is recommended to exclude an underlying malignancy. 5.Chronic emphysematous changes as
Checkpoint saved to ./IUXRAY_OUTPUT_PIPELINE/checkpoints/checkpoint_epoch_37.pth
Removed old checkpoint: ./IUXRAY_OUTPUT_PIPELINE/checkpoints/checkpoint_epoch_36.pth
Checkpoint saved to ./IUXRAY_OUTPUT_PIPELINE/checkpoints/latest_checkpoint.pth
  No improvement in metrics for 10 epochs
Early stopping after 37 epochs
```

Metrics


Training and Validation Loss
Training Loss
Validation Loss
Loss
Epochs
Validation Metrics
BLEU-4
ROUGE-L
Score
Epochs

Results

X-Ray Sample 3



Reference Report:
The cardiomediastinal silhouette and vasculature are within normal limits for size and contour. The lungs are normally inflated and clear. Osseous structures are within normal limits for patient age. IMPRESSION: 1. No acute radiographic cardiopulmonary process.

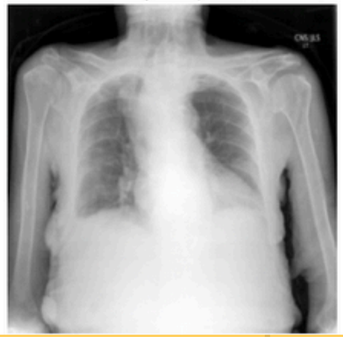
Generated Report:
The cardiac silhouette is normal. The lungs are clear. Stable degenerative changes of the thoracic spine. The examination reveals: 1. Cardiomegaly and large hiatal hernia. 2. Age-indeterminate deformity of the midthoracic vertebral body. This may be due to atelectasis or infiltrate in the osseous structures. If clinically indicated, further

X-Ray Sample 4

Reference Report:
Heart size and mediastinal contour within normal limits. Calcified granuloma in the left lung base. No focal airspace consolidation, pneumothorax, or large pleural effusion. No acute osseous abnormality. IMPRESSION: No acute cardiopulmonary abnormality.

Generated Report:
The examination reveals Heart size is normal. Lungs are clear. No evidence of acute cardiopulmonary process. Thoracic spondylosis. No effusion or pneumothorax. Mild degenerative changes of the thoracic spine. Elevation of the right hemidiaphragm. No acute bony abnormality. Minimal degenerative change of the midthoracic vertebral bodies. 1. Low lung

X-Ray Sample 5



Reference Report:
Low lung volumes. Stable ectasia of the thoracic aorta. Stable right upper mediastinal Bilateral small pleural effusions and bibasilar airspace opacities. The heart size and mediastinal silhouette are within normal limits for contour. No pneumothorax. Stable wedging of the anterior thoracic vertebral bodies. IMPRESSION: Bilateral small pleural effusions and associated atelectasis. Stable right upper mediastinal opacity consistent with <redacted> goiter.

Generated Report:
The heart is normal. The lungs are clear. There is no pneumothorax or pleural effusion. There are no active cardiopulmonary disease. No evidence of active tuberculosis. 1. No acute bony abnormality. 2. No active bony disease. 3. ilaterally calcified granulomas. 4. No visible active tuberculosis or active tuberculous process. 5. incidentally, no visible active or

Challenges

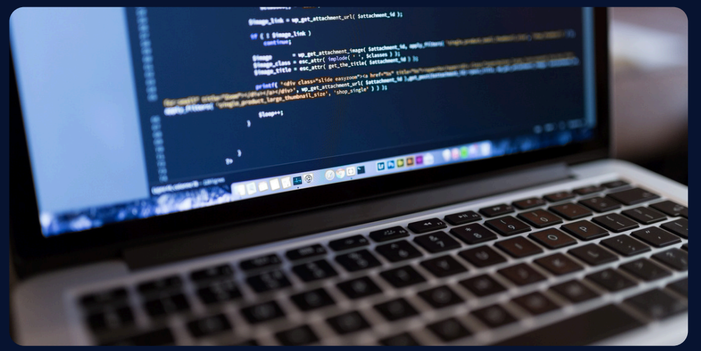
- Incomplete Reports: ~30% missing either Findings or Impression, limiting supervision.
- Imbalanced Conditions: Over-representation of terms like normal, under-representation of rare diseases.
- Ambiguous Language: Overlapping medical terms caused confusion for the model.
- Multi-Image Reports: Same text mapped to different images → reduced visual-text alignment.
- Short & Repetitive Texts: Limited vocabulary made generation less diverse and harder to evaluate.

Conclusion

- Built a ResNet-GPT2 encoder-decoder model to generate radiology reports from chest X-ray images.
- The model learned general radiological descriptions but struggled with specific clinical detail.
- Training and validation loss steadily decreased, indicating effective learning.
- BLEU-4 and ROUGE-L scores remained moderate due to dataset limitations and report complexity.

Future Directions

- Introduce attention mechanisms to enhance alignment between visual and textual features.
- Explore multi-modal pretraining strategies to improve contextual understanding.
- Integrate structured medical knowledge for more accurate and relevant outputs.
- Apply active learning to balance underrepresented findings and conditions.



Thank You