



至诚至坚

博学笃行

EMRRG: Efficient Fine-Tuning Pre-trained X-ray Mamba Networks for Radiology Report Generation

Mingzheng Zhang, Jinfeng Gao, Dan Xu, Jiangrui Yu
Yuhan Qiao, Lan Chen, Jin Tang, Xiao Wang

讲者：徐丹
指导老师：王逍
日期：2025.11.30

安徽大学

计算机科学与技术学院

Introduction to X-ray Medical Report Generation

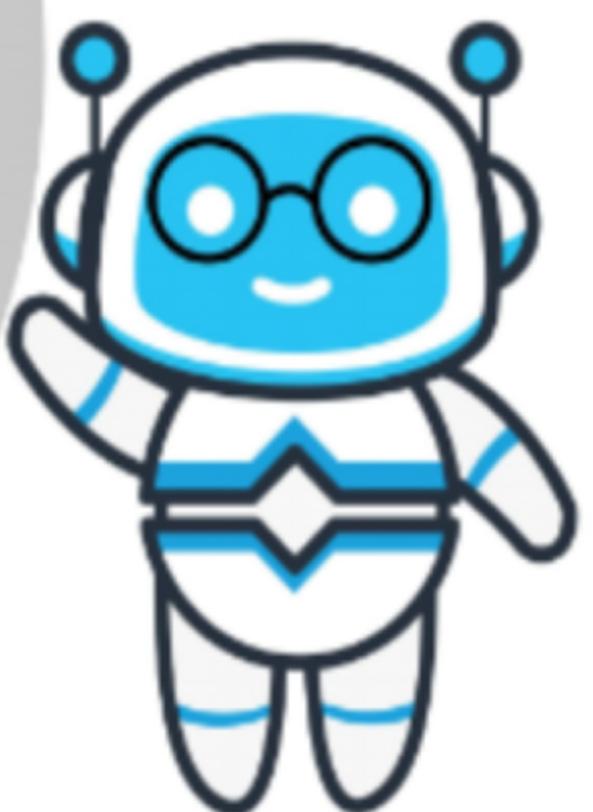


What is Radiology Report Generation?

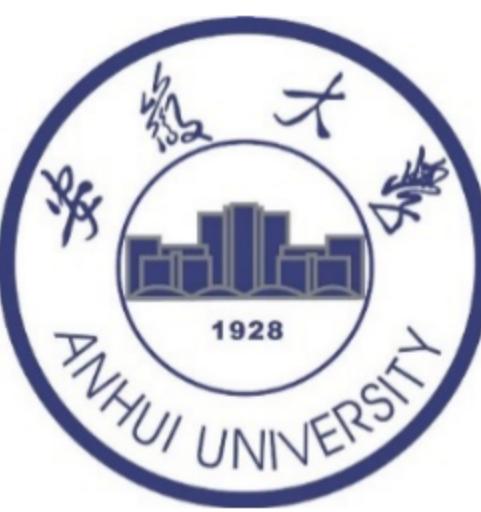
- An AI task that automatically translates medical images into coherent, structured textual reports
- It bridges Computer Vision and Natural Language Processing, mimicking the workflow of a radiologist
- Core Process: Image → Visual Features → Textual Description

Task Introduction

- Automatically generate diagnostic reports from radiographs using AI
- Goal: Assist clinicians by creating preliminary, descriptive text reports



Existing Work: R2Gen

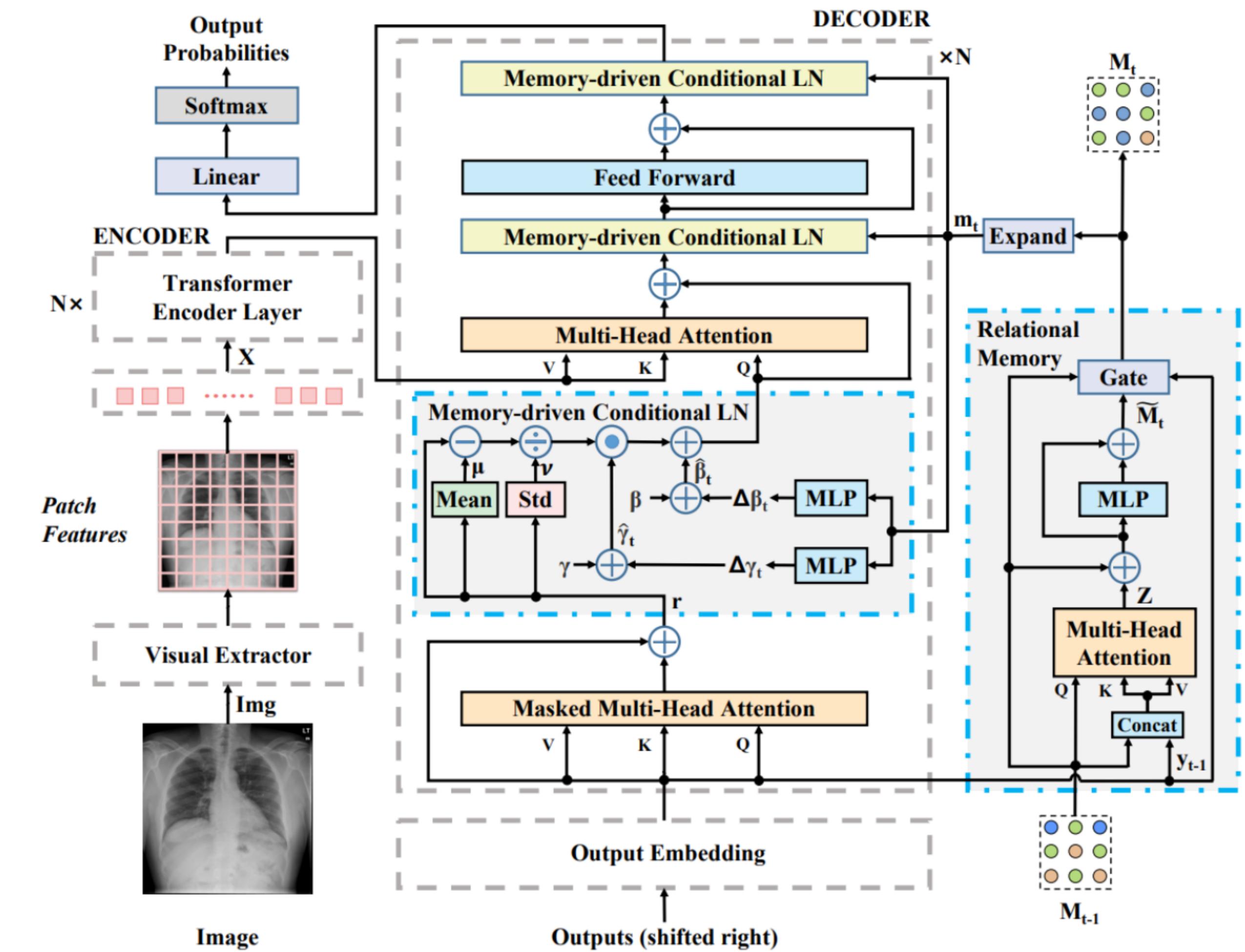


➤ Core Idea

- Uses a Relational Memory (RM) to capture recurrent patterns in reports.
- Integrates memory into the Transformer decoder via Memory-driven Conditional Layer Normalization (MCLN).

➤ Key Contributions

- Achieved state-of-the-art (SOTA) performance on IU X-Ray and MIMIC-CXR.
- First work to comprehensively benchmark report generation on the MIMIC-CXR dataset.



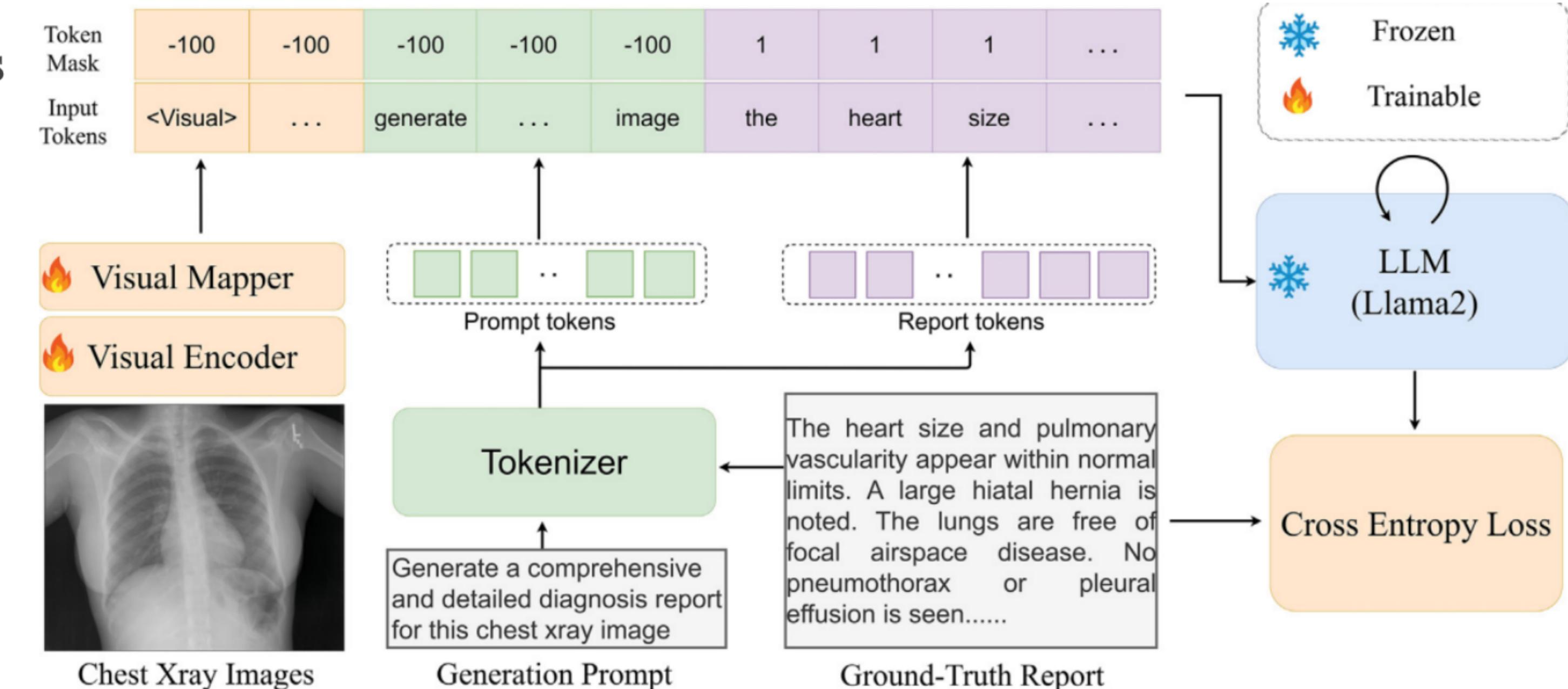
Existing Work: R2GenGPT

➤ Core Idea

- Align & Freeze: Maps visual features into a frozen LLM's (Llama2) embedding space.
- Parameter-Efficient: Explores 3 tuning strategies with minimal trainable parameters.

➤ Key Contributions

- Pioneering Use of LLM: First to leverage frozen LLMs for radiology report generation.
- Rigid LLM: High Efficiency: Achieves strong performance by training only 0.07% parameters.
- SOTA Performance: Reaches or surpasses state-of-the-art on IU X-Ray and MIMIC-CXR.



Existing Work: MambaXray-VL: Pre-training for X-ray Report Generation

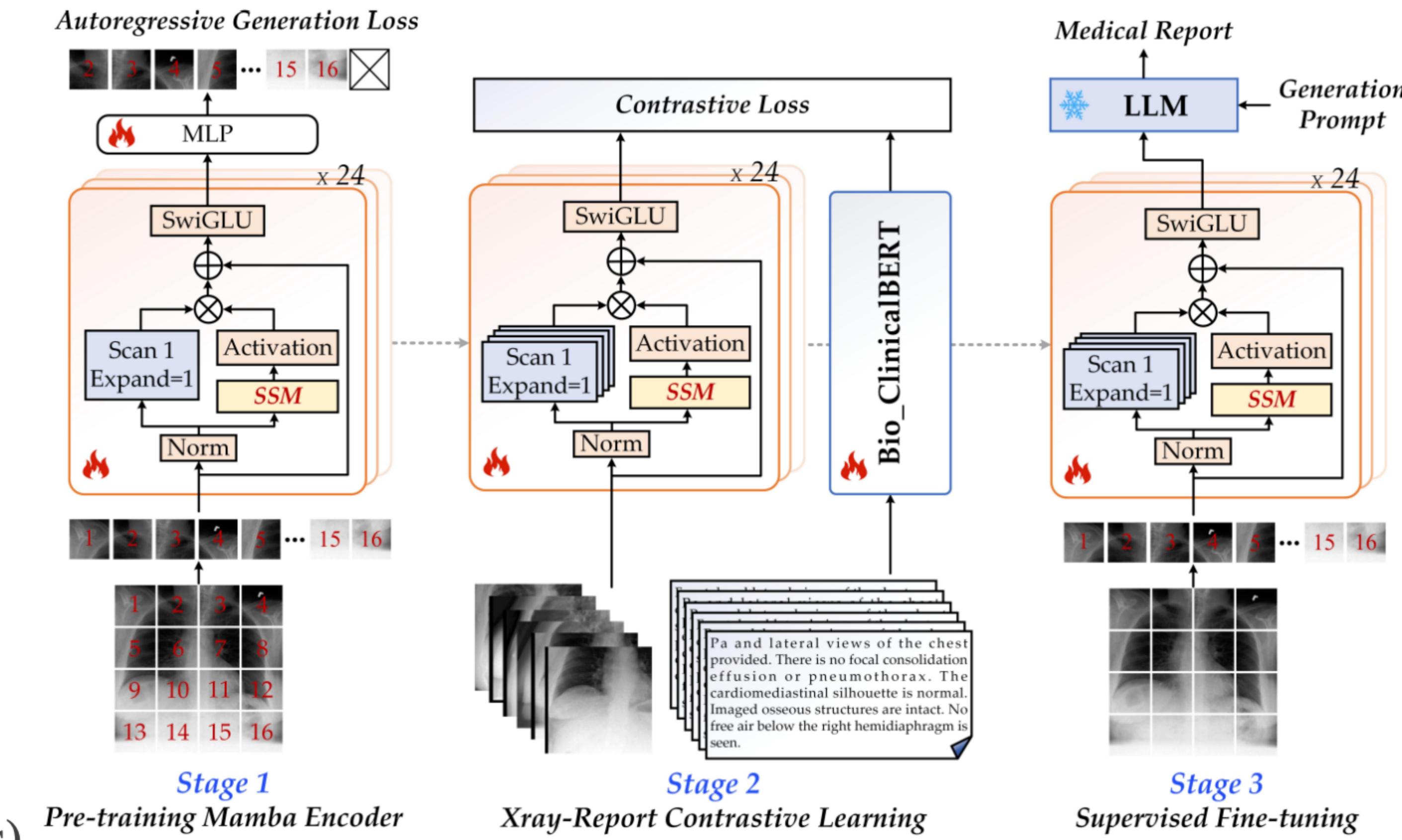


➤ Core Idea

- Multi-stage pre-training: autoregressive + contrastive + fine-tuning
- Mamba vision encoder with $O(N)$ complexity
- Replaces traditional Transformer-based encoders

➤ Key Contributions

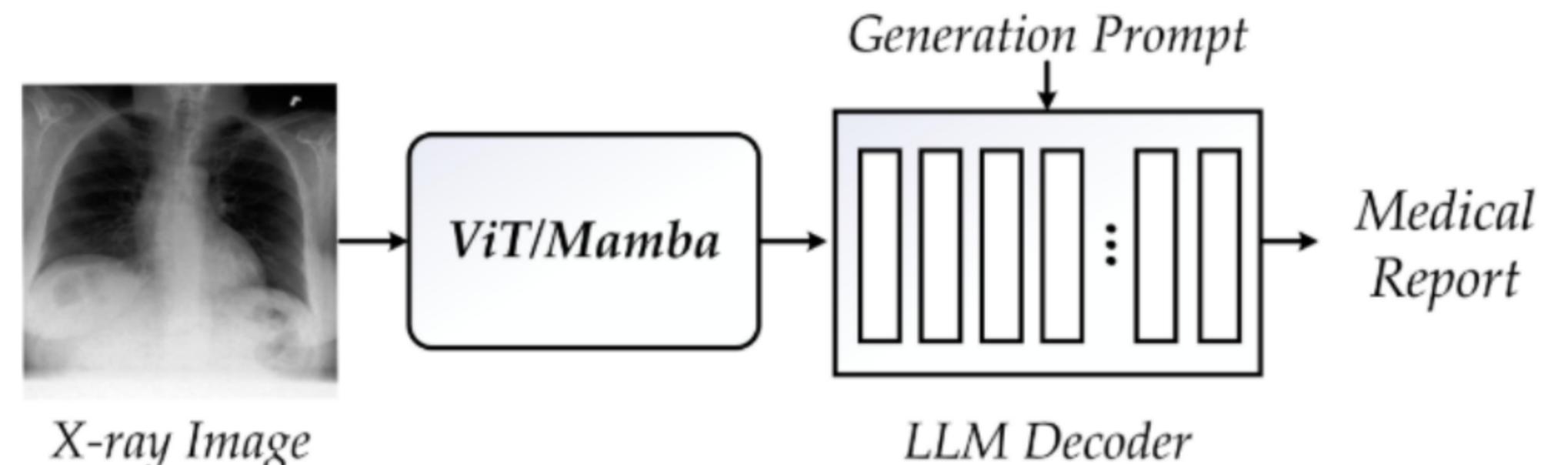
- Built CXPMRG-Bench - first large benchmark on CheXpert Plus
- Achieves SOTA across multiple datasets
- ARG pre-training outperforms MAE on X-rays
- Evaluates 35+ models (19 MRG + 16 LLMs/VLMs)



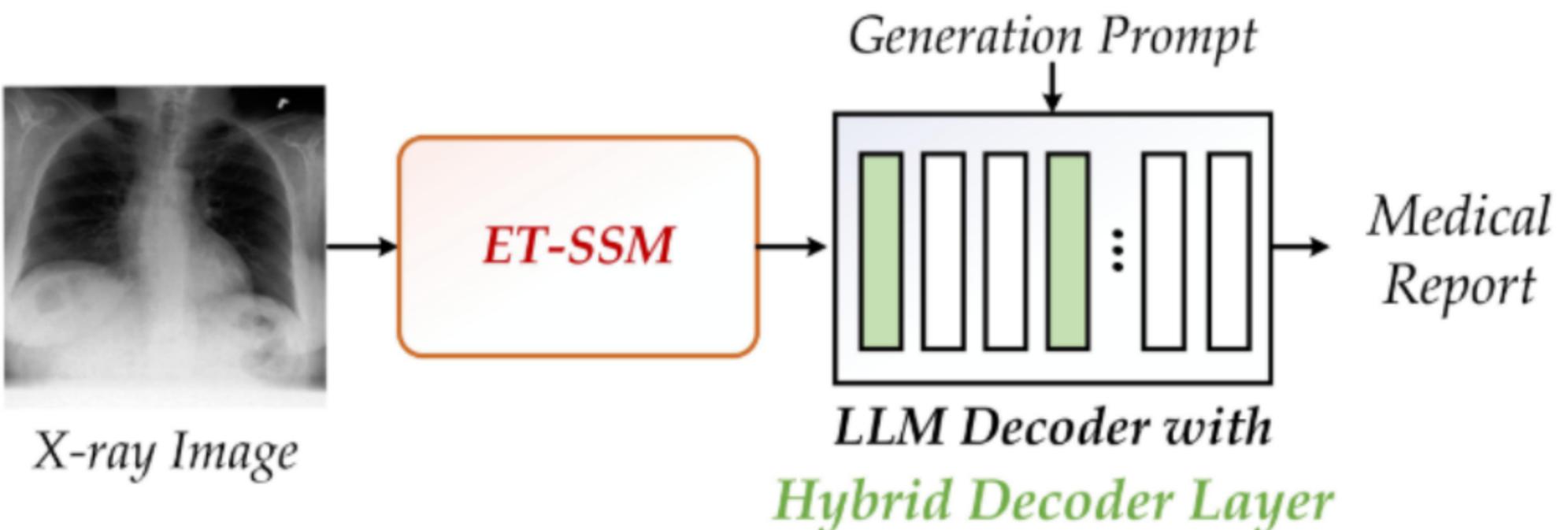
Limitations of Current Approaches and Our Contributions



Existing MRG systems face critical limitations that our EMRRG framework addresses through novel architectural innovations and efficient tuning strategies



(a). Existing ViT/Mamba-LLM framework for X-ray MRG



(b). Our newly proposed Efficient-Tuning SSM for MRG

① Key Limitations of Current Methods

- Insufficient exploration of pre-trained vision models with lightweight fine-tuning
- Limited cross-attention capabilities in LLM-based generation frameworks
- Underutilization of non-Transformer architectures like Mamba networks

↗ Our EMRRG Framework Contributions

- Novel Mamba-based architecture with parameter-efficient fine-tuning strategy
- Hybrid decoder layer augmented LLM for enhanced cross-attention
- Strong performance across IU X-ray, MIMIC-CXR, and CheXpert Plus datasets

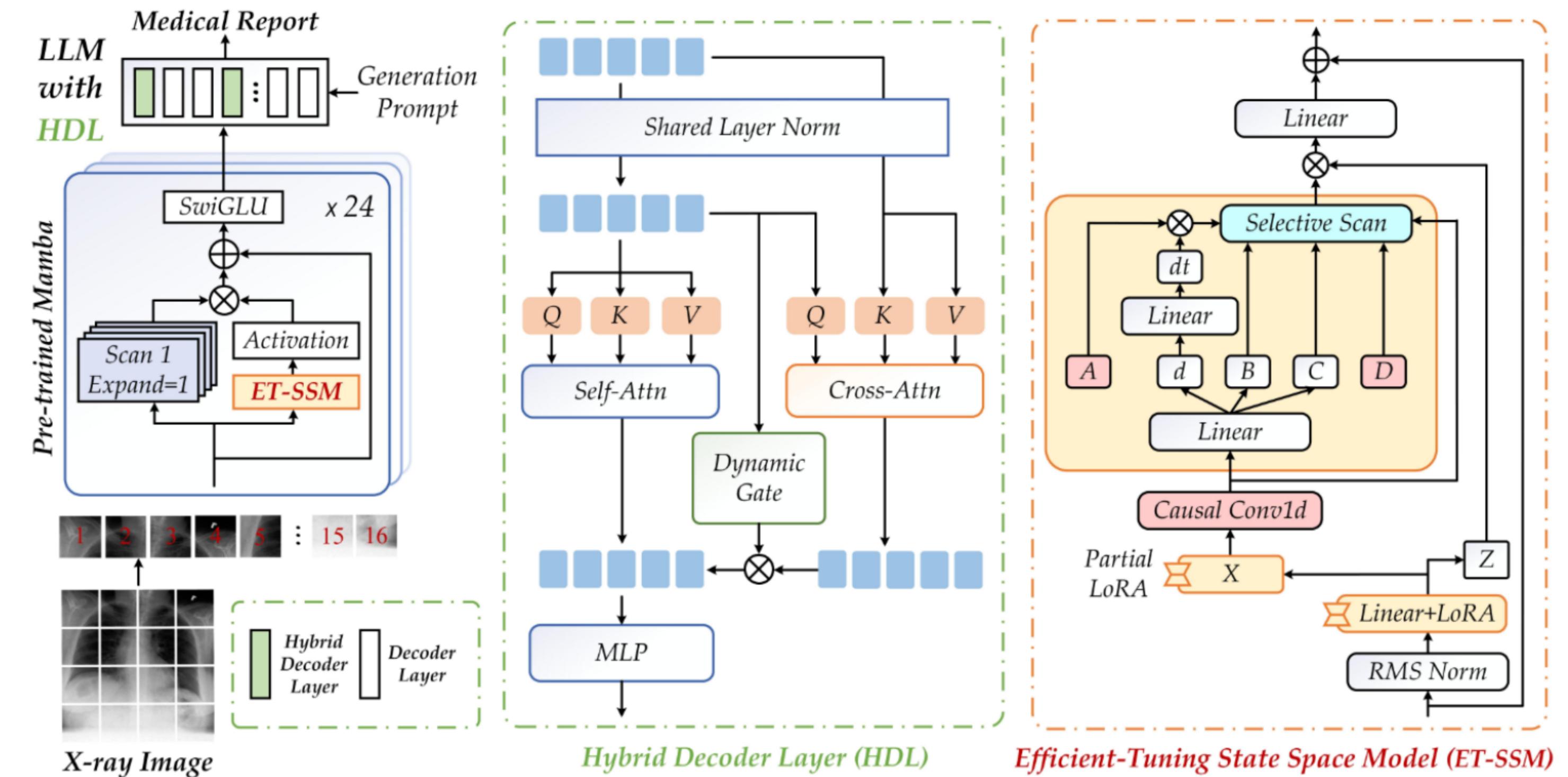


Overview of EMRRG Framework



❖ Framework Components

- Combines **efficient-tuning Mamba networks** with hybrid decoder-augmented LLMs
- Utilizes **ET-SSM** (Efficient-Tuning based SSM) for visual feature extraction
- Incorporates **hybrid decoder layers** in LLM for cross-modal integration

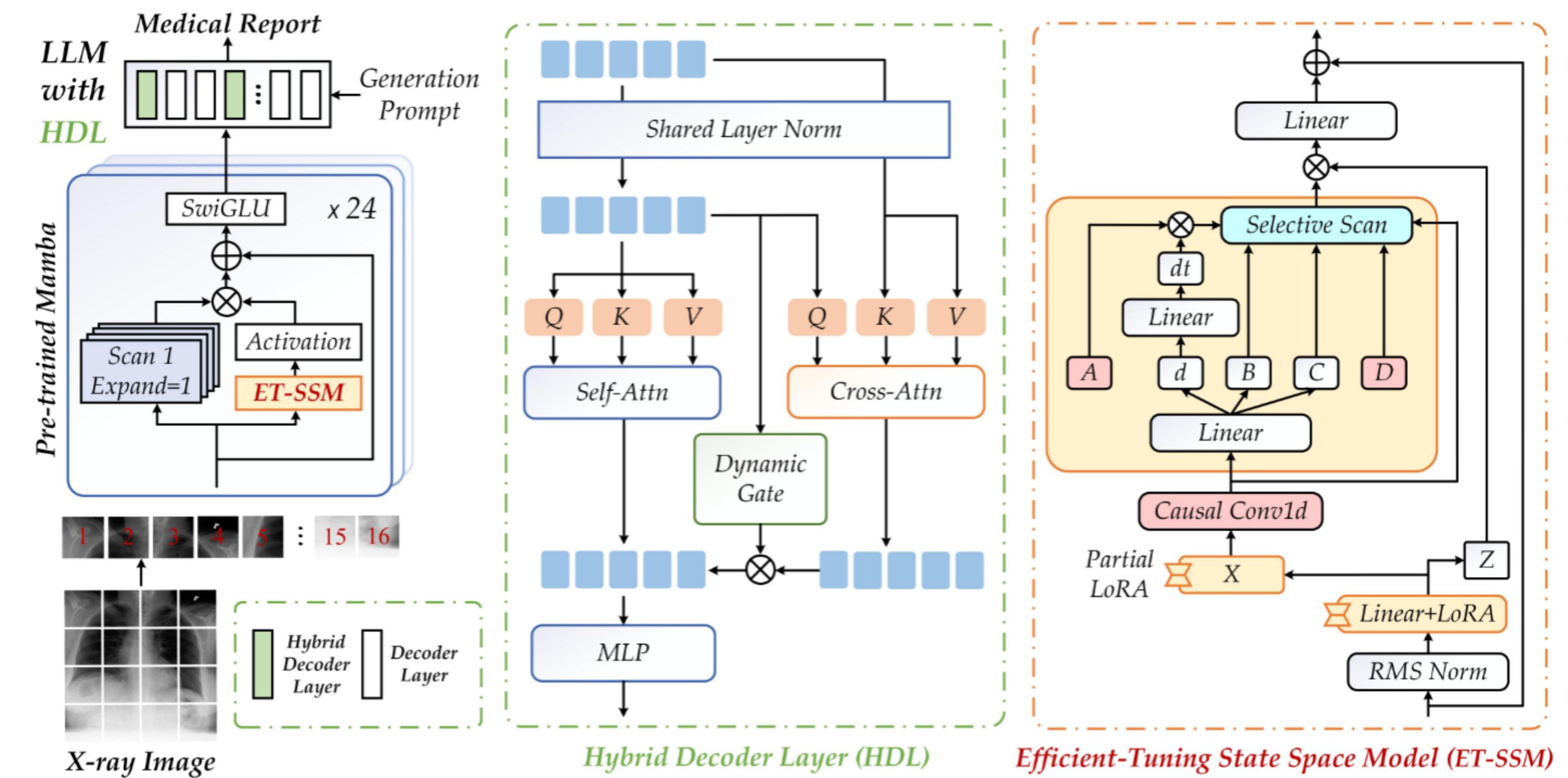


Overview of EMRRG Framework



Key Innovations

- **Partial LoRA (LoRA_P(X))** selectively adapts Mamba intermediate features
- **Cross-attention mechanism** enables dynamic visual-textual information fusion
- **Dynamic gating** modulates information flow between modalities



Efficient-Tuning SSM (ET-SSM)



➤ Partial LoRA ($\text{LoRA}_P(X)$):

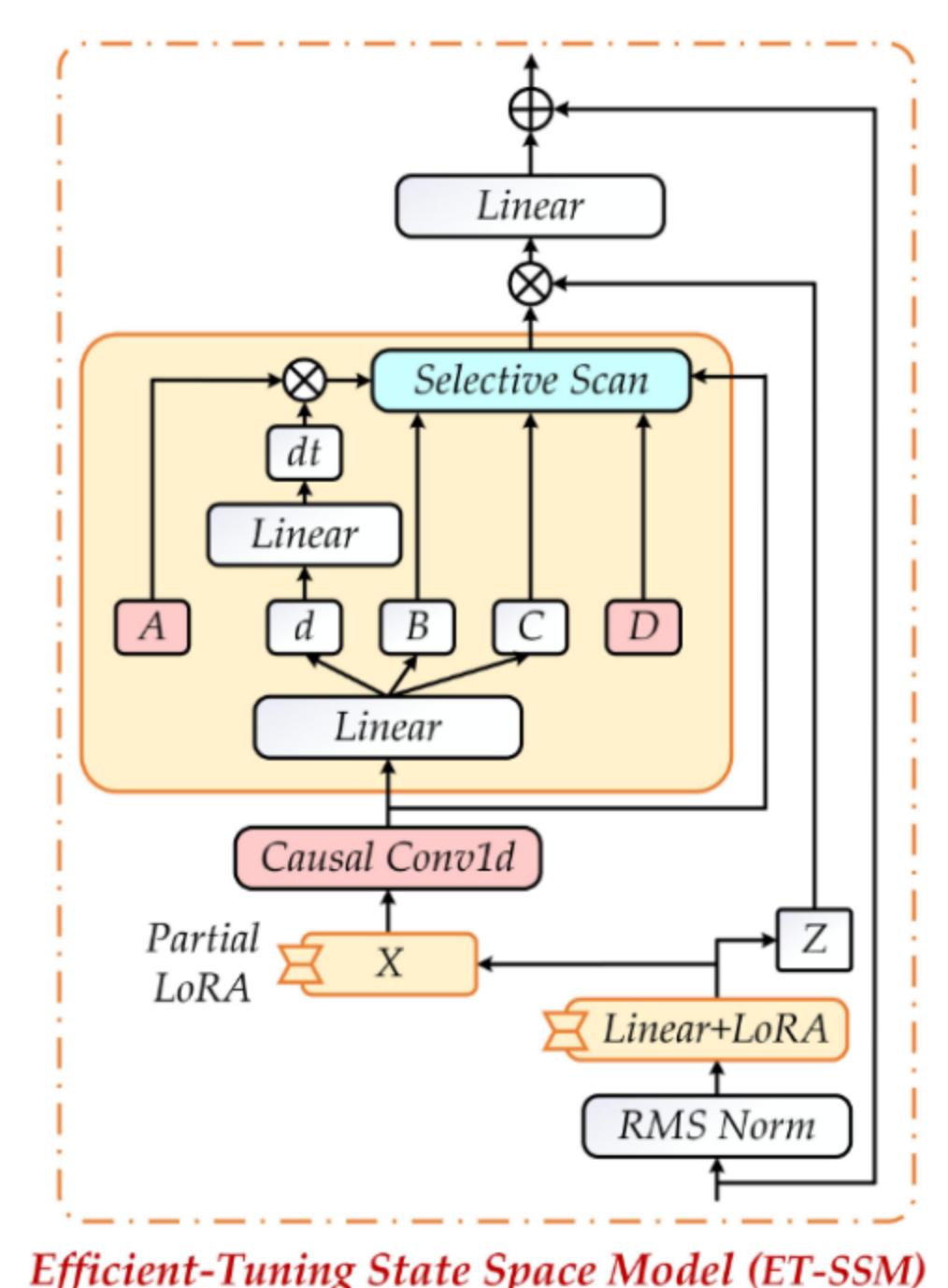
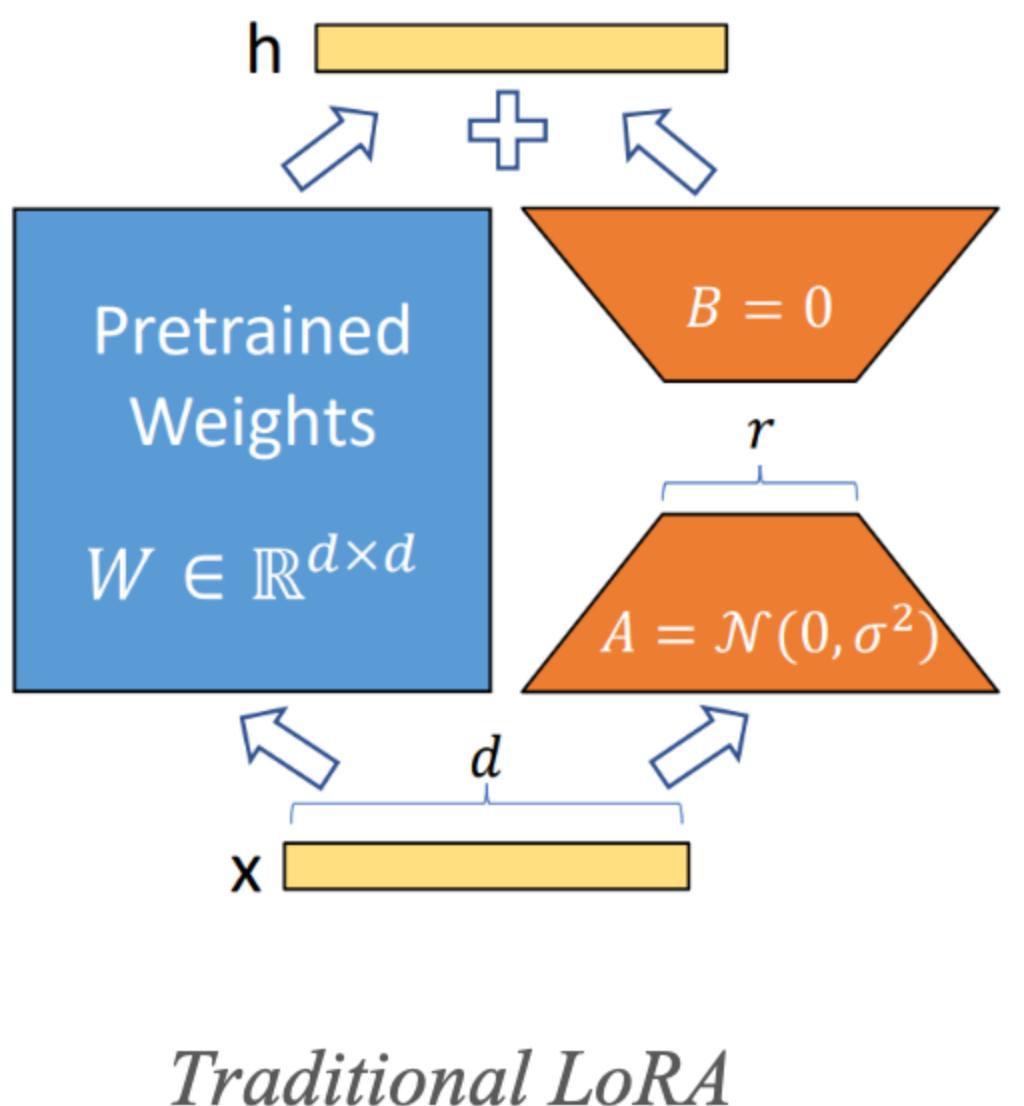
- Applies LoRA only to intermediate feature X in Mamba blocks
- Preserves other features (Z, dt, B, C) to maintain model stability

➤ Traditional LoRA:

- Fine-tunes input projection layer for better source representation

➤ Advantage:

- Combines targeted adaptation with global optimization, improving training efficiency without sacrificing capacity



Hybrid Decoder Layer Design

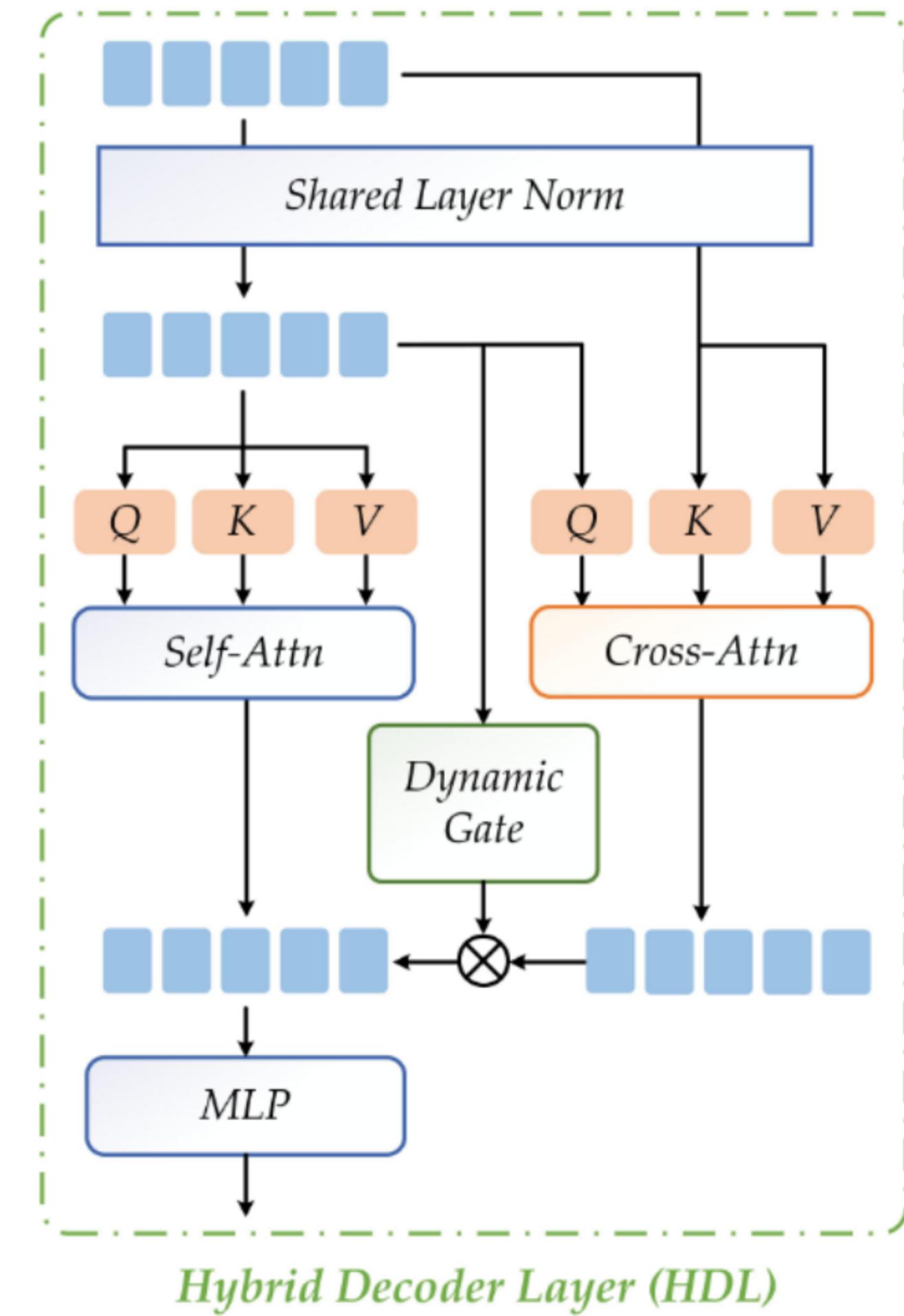


➤ Architecture:

- Parallel Self-Attention (text-text) and Cross-Attention (text-visual)
- Skip connections for training stability

➤ Dynamic Gating Mechanism:

- Token-wise gating values via linear layer + tanh
- Warm-up initialization for stable training
- Adaptive fusion of visual features into text embeddings



Hybrid Decoder Layer (HDL)

Experimental Setup and Datasets



Overview of datasets, evaluation metrics, and implementation details for EMRRG validation

Datasets



- Three public benchmarks: **IU-Xray**, **MIMIC-CXR**, and **CheXpert Plus**
- Standard datasets for evaluating medical report generation performance

Evaluation Metrics



- **Natural Language Metrics:** BLEU, ROUGE-L, METEOR, CIDEr
- **Clinical Metrics:** Precision, Recall, and F1-measure

Implementation Details



- **IU-Xray:** 30 epochs, batch size 20, Qwen-1.5-1.8B LLM
- **MIMIC-CXR/CheXpert Plus:** 6 epochs, batch size 18, Llama2-7B LLM
- **LoRA_{P(X)}** fine-tuning with rank=32 on in_proj layer weights



Performance Comparison on Public Benchmarks



🕒 IU X-ray Results

- Achieved **SOTA performance** on BLEU-2 (0.325) and ROUGE-L (0.385)
- Second-best on BLEU-1 (0.487) and METEOR (0.226) metrics
- Performance comparable to other SOTA report generation approaches

+ MIMIC-CXR & CheXpert Plus Results

- Strong performance on MIMIC-CXR with **0.311 ROUGE-L** and **0.239 CIDEr**
- CheXpert Plus SOTA results: BLEU-4 (0.104), ROUGE-L (0.273), METEOR (0.152)
- Consistent performance across all three benchmark datasets

| Dataset | Method | Publication | BLEU-1 | BLEU-2 | BLEU-3 | BLEU-4 | ROUGE-L | METEOR | CIDEr |
|---------------|----------------------------|---------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| IU X-Ray | R2Gen [4] | EMNLP 2020 | 0.470 | 0.304 | 0.219 | 0.165 | 0.371 | 0.187 | - |
| | R2GenCMN [5] | ACL-IJCNLP 2021 | 0.475 | 0.309 | 0.222 | 0.170 | 0.375 | 0.191 | - |
| | METransformer [34] | CVPR 2023 | 0.483 | 0.322 | 0.228 | 0.172 | 0.380 | 0.192 | 0.435 |
| | DCL [15] | CVPR 2023 | - | - | - | 0.163 | 0.383 | 0.193 | 0.586 |
| | R2GenGPT [†] [35] | Meta Radiology 2023 | 0.465 | 0.299 | 0.214 | 0.161 | 0.376 | 0.219 | 0.542 |
| | PromptMRG [13] | AAAI 2024 | 0.401 | - | - | 0.098 | 0.160 | 0.281 | - |
| | Med-LMM [20] | ACM MM 2024 | - | - | - | 0.168 | 0.381 | 0.209 | 0.427 |
| | SILC [19] | IEEE TMI 2024 | 0.472 | <u>0.321</u> | <u>0.234</u> | <u>0.175</u> | 0.379 | 0.192 | 0.368 |
| | KIA [40] | COLING 2025 | 0.501 | 0.325 | 0.240 | 0.183 | 0.375 | 0.207 | <u>0.559</u> |
| | EMRRG | Ours | <u>0.487</u> | 0.325 | 0.222 | 0.167 | 0.385 | <u>0.226</u> | 0.476 |
| MIMIC-CXR | R2Gen [4] | EMNLP 2020 | 0.353 | 0.218 | 0.145 | 0.103 | 0.277 | 0.142 | - |
| | R2GenCMN [5] | ACL-IJCNLP 2021 | 0.353 | 0.218 | 0.148 | 0.106 | 0.278 | 0.142 | - |
| | METransformer [34] | CVPR 2023 | 0.386 | 0.250 | 0.169 | 0.124 | 0.291 | 0.152 | 0.362 |
| | DCL [15] | CVPR 2023 | - | - | - | 0.109 | 0.284 | 0.150 | <u>0.281</u> |
| | R2GenGPT [†] [35] | Meta Radiology 2023 | 0.408 | 0.256 | 0.174 | 0.125 | 0.285 | 0.167 | 0.244 |
| | PromptMRG [13] | AAAI 2024 | 0.398 | - | - | 0.112 | 0.268 | 0.157 | - |
| | Med-LMM [20] | ACM MM 2024 | - | - | - | <u>0.128</u> | 0.289 | 0.161 | 0.265 |
| | AdaMatch-Cyclic [3] | ACL 2024 | 0.379 | 0.235 | 0.154 | 0.106 | 0.286 | 0.163 | - |
| | GIT-CXR [26] | arXiv 2025 | 0.403 | <u>0.254</u> | 0.215 | 0.136 | 0.311 | 0.161 | - |
| | EMRRG | Ours | <u>0.407</u> | 0.256 | <u>0.175</u> | 0.125 | 0.288 | <u>0.164</u> | 0.239 |
| CheXpert Plus | R2Gen [4] | EMNLP 2020 | 0.301 | 0.179 | 0.118 | 0.081 | 0.246 | 0.113 | 0.077 |
| | R2GenCMN [5] | ACL-IJCNLP 2021 | 0.321 | 0.195 | 0.128 | 0.087 | 0.256 | 0.127 | 0.102 |
| | XProNet [30] | ECCV 2022 | 0.364 | 0.225 | 0.148 | 0.100 | 0.265 | <u>0.146</u> | 0.121 |
| | ORGan [11] | ACL 2023 | 0.320 | 0.196 | 0.128 | 0.086 | 0.261 | 0.135 | 0.107 |
| | R2GenGPT [35] | Meta Radiology 2023 | 0.361 | 0.224 | <u>0.149</u> | <u>0.101</u> | <u>0.266</u> | 0.145 | <u>0.123</u> |
| | ASGMD [38] | ESWA 2024 | 0.267 | 0.149 | 0.094 | 0.063 | 0.220 | 0.094 | 0.044 |
| | Token-Mixer [39] | IEEE TMI 2024 | 0.378 | <u>0.231</u> | 0.153 | 0.091 | 0.262 | 0.135 | 0.098 |
| | PromptMRG [13] | AAAI 2024 | 0.326 | 0.174 | - | 0.095 | 0.222 | 0.121 | 0.044 |
| | R2GenCSR [32] | arXiv 2024 | 0.364 | 0.225 | 0.148 | 0.100 | 0.265 | <u>0.146</u> | 0.121 |
| | EMRRG | Ours | <u>0.375</u> | 0.232 | 0.153 | 0.104 | 0.273 | 0.152 | 0.167 |

NLG Metrics Comparison



Clinical Evaluation Metrics Results



MIMIC-CXR Clinical Metrics

- Precision: 0.421, Recall: 0.372, F1-score: 0.395
- Performance comparable to state-of-the-art clinical evaluation results
- Balanced performance between precision and recall metrics

| Method | Publication | MIMIC-CXR | | |
|--------------------|---------------|--------------|--------------|--------------|
| | | Precision | Recall | F1 |
| R2Gen [4] | EMNLP 2020 | 0.333 | 0.273 | 0.276 |
| METransformer [34] | CVPR 2023 | 0.364 | 0.309 | 0.311 |
| KiUT [12] | CVPR 2023 | 0.371 | 0.318 | 0.321 |
| DCL [15] | CVPR 2023 | <u>0.471</u> | 0.352 | 0.373 |
| CoFE [16] | ECCV 2024 | 0.489 | 0.370 | 0.405 |
| HERGen [29] | ECCV 2024 | 0.415 | 0.301 | 0.317 |
| SILC [19] | IEEE TMI 2024 | 0.457 | 0.337 | 0.330 |
| OaD [17] | IEEE TMI 2024 | 0.364 | <u>0.382</u> | 0.372 |
| GIT-CXR [26] | arXiv 2025 | 0.349 | 0.403 | 0.336 |
| EMRRG | Ours | 0.421 | 0.372 | <u>0.395</u> |

Clinical Metrics Comparison



Clinical Evaluation Metrics Results



❖ CheXpert Plus Clinical Metrics

- Precision: 0.341, Recall: 0.273, F1-score: 0.272
- Consistent clinical relevance across different datasets
- Demonstrates EMRRG's ability to capture clinically meaningful findings

| Method | Publication | CheXpert Plus | | |
|------------------|----------------|---------------|--------------|--------------|
| | | Precision | Recall | F1 |
| R2Gen [4] | EMNLP 2020 | 0.318 | 0.200 | 0.181 |
| R2GenCMN [5] | ACL 2021 | 0.329 | 0.241 | 0.231 |
| XProNet [30] | ECCV 2022 | 0.314 | 0.247 | 0.259 |
| R2GenGPT [35] | Meta-Rad. 2023 | 0.315 | 0.244 | 0.260 |
| Zhu et al. [44] | MICCAI 2023 | 0.217 | 0.308 | 0.205 |
| PromptMRG [13] | AAAI 2024 | 0.258 | 0.265 | <u>0.281</u> |
| Token-Mixer [39] | IEEE TMI 2024 | 0.309 | 0.270 | 0.288 |
| EMRRG | Ours | 0.341 | <u>0.273</u> | 0.272 |

Clinical Metrics Comparison



Runtime Efficiency Analysis



| Index | Algorithm | Publish | Encoder | Decoder | Time (min) | Param (M) |
|-------|---------------------|-------------|------------------|-------------|------------|-----------|
| #01 | TIMER [36] | CHIL23 | Transformer | Transformer | 26.5 | 79.28 |
| #02 | CvT2DistilGPT2 [22] | AIM23 | Transformer | GPT2 | 13.93 | 128.00 |
| #03 | ORGan [11] | ACL23 | CNN | Transformer | 46.66 | 67.50 |
| #04 | Zhu et al. [44] | MICCAI23 | Transformer | Transformer | 10.03 | 85.95 |
| #05 | CAMANet [31] | IEEE JBH23 | Swin-Former | Transformer | 23.08 | 43.22 |
| #06 | ASGMD [38] | ESWA24 | ResNet-101 | Transformer | 87.37 | 277.41 |
| #07 | Token-Mixer [39] | IEEE TMI23 | ResNet-50 | Transformer | 17.54 | 104.34 |
| #08 | PromptMRG [13] | AAAI24 | ResNet-101 | Bert | 108.45 | 219.92 |
| #09 | R2GenGPT [35] | Meta-Rad.23 | Swin-Transformer | Llama2 | 77.80 | 90.90 |
| #10 | R2GenCSR [32] | arXiv24 | VMamba | Llama2 | 31.20 | 91.70 |
| #11 | Wang et al. [33] | arXiv24 | ViT | Llama2 | 10.82 | 358.80 |
| #12 | MambaXray-VL [33] | CVPR2025 | MambaXray-VL | Llama2 | 50.66 | 57.31 |
| #13 | EMRRG | Ours | MambaXray-VL | Llama2 | 26.84 | 1.32 |

Runtime Efficiency Comparison

▣ Computational Efficiency

- Training time: **26.84 minutes** on CheXpert Plus dataset
- Only **1.32M parameters** required for fine-tuning
- Significantly faster than PromptMRG (108.45 min) and ASGMD (87.37 min)

▣ Efficiency Advantages

- **2.3% of parameters** compared to full fine-tuning methods
- Maintains performance close to full parameter fine-tuning
- MambaXray-VL encoder + Llama2 decoder architecture optimizes speed



Ablation Study and Component Analysis



| Index | SSM | | LLM | | | IU X-ray | | | | | | |
|-------|-----|-----------------------|-----|---------|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | FT | LoRA _P (X) | L2 | L2+LoRA | L2+HDL | B1 | B2 | B3 | B4 | R-L | M | C |
| #01 | ✓ | ✗ | ✓ | ✗ | ✗ | 0.480 | 0.322 | 0.226 | 0.175 | 0.383 | 0.215 | 0.478 |
| #02 | ✓ | ✗ | ✗ | ✓ | ✗ | 0.473 | 0.311 | 0.216 | 0.171 | 0.384 | 0.210 | 0.483 |
| #03 | ✓ | ✗ | ✗ | ✗ | ✓ | 0.489 | <u>0.324</u> | 0.231 | 0.182 | 0.391 | <u>0.219</u> | 0.490 |
| #04 | ✗ | ✓ | ✓ | ✗ | ✗ | 0.475 | 0.309 | 0.211 | 0.161 | 0.372 | 0.208 | 0.464 |
| #05 | ✗ | ✓ | ✗ | ✓ | ✗ | 0.471 | 0.313 | 0.216 | 0.158 | 0.374 | 0.210 | 0.461 |
| #06 | ✗ | ✓ | ✗ | ✗ | ✓ | <u>0.487</u> | 0.325 | 0.222 | 0.167 | <u>0.385</u> | 0.226 | 0.476 |

Component Analysis Results

Component Analysis

- LoRA_P(X) method achieves performance close to full fine-tuning
- Llama2+HDL outperforms standard Llama2 and Llama2+LoRA
- Combination of FT + Llama2+HDL achieves optimal results on most metrics

Parameter Efficiency

- EMRRG requires **2.3% of parameters** compared to full fine-tuning
- Maintains performance while significantly improving training efficiency
- LoRA_P(X)+Llama2+HDL achieves optimal BLEU-2 and METEOR scores



Ablation Study – Tuning Strategies



| Setting | BLEU-1 | BLEU-2 | BLEU-3 | BLEU-4 | ROUGE-L | METEOR | CIDEr |
|------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| LoRA(Llama2) | 0.479 | 0.320 | 0.233 | 0.177 | 0.386 | <u>0.212</u> | <u>0.489</u> |
| LoRA(embedding) | 0.463 | 0.302 | 0.218 | 0.167 | 0.379 | 0.207 | 0.530 |
| LoRA(x_proj) | 0.457 | 0.301 | <u>0.226</u> | 0.162 | 0.383 | 0.205 | 0.473 |
| LoRA(dt_proj) | 0.458 | 0.300 | 0.212 | 0.161 | 0.378 | 0.203 | 0.485 |
| LoRA(in_proj) | 0.444 | 0.288 | 0.205 | 0.152 | 0.368 | 0.194 | 0.431 |
| LoRA(out_proj) | 0.456 | 0.297 | 0.211 | 0.156 | 0.374 | 0.199 | 0.451 |
| LoRA _p (Z) | 0.466 | 0.302 | 0.213 | 0.163 | 0.381 | 0.204 | 0.467 |
| LoRA _p (dt) | 0.458 | 0.299 | 0.221 | 0.157 | <u>0.387</u> | 0.197 | 0.484 |
| LoRA _p (B) | 0.462 | 0.307 | 0.208 | 0.165 | 0.384 | 0.201 | 0.472 |
| LoRA _p (C) | 0.473 | 0.303 | 0.217 | 0.154 | 0.379 | 0.209 | 0.466 |
| LoRA _p (X) | 0.485 | <u>0.311</u> | 0.223 | <u>0.169</u> | 0.388 | 0.216 | 0.474 |

Ablation on Fine-tuning Strategies

- **LoRA_P(X) Achieves Best Overall Performance:**
 - BLEU-1: 0.485, ROUGE-L: 0.388
 - Outperforms LoRA_P(Z), LoRA_P(dt), LoRA_P(B), LoRA_P(C)
- **Interpretation:** Fine-tuning feature X most effectively adapts visual representations for report generation.

Qualitative Results & Visualization



| X-ray Image | Ground Truth | Mamba-Xray-VL | EMRRG (Ours) |
|-------------|---|---|---|
| | Cardiomediastinal silhouette is within normal limits. The lungs are clear without evidence of focal consolidation. There are degenerative changes of the shoulder bilaterally with surgical rotator cuff fixation pins in place. | Cardiomediastinal silhouette is within normal limits .The lungs are clear without evidence of focal consolidation or pleural effusion the visualized osseous structures are unremarkable . degenerative changes are seen in the thoracic spine and shoulder bilaterally with degenerative changes . | Cardiomediastinal silhouette is within normal limits . The lungs are clear without evidence of focal consolidation pleural effusion or pneumothorax . Degenerative changes are seen in the thoracic spine and shoulder bilaterally with degenerative changes . |
| | A single portable AP view of the chest demonstrates a right-sided subclavian vein catheter with tip at the cavoatrial junction. The lung volumes are low. No definite areas of consolidation are evident. The soft tissues and bones are unremarkable. | A single portable ap view of the chest demonstrates a right-sided subclavian vein catheter with tip at the cavoatrial junction. No evidence of pneumothorax pulmonary edema or pleural effusions . The bones and soft tissues are unremarkable. | A single portable AP view of the chest demonstrates a right-sided subclavian vein catheter with tip at the cavoatrial junction. Low lung volumes are demonstrated. There is no evidence of focal consolidation pulmonary edema or pneumothorax. The soft tissues and bones are unremarkable . |
| | There is redemonstration of a right internal jugular tunneled central venous catheter, terminating in the distal superior vena cava. The lungs remain clear. There is no evidence of pleural effusion. The cardiomediastinal silhouette is within normal limits. Regional osseous structures appear unremarkable. | The cardiomediastinal silhouette and pulmonary vascularity are within normal limits . there is no evidence of focal consolidation pleural effusion or pneumothorax . Soft tissues and osseous structures are unremarkable. | There has been redemonstration of a right internal jugular tunneled central venous catheter with the tip in the distal superior vena cava . The cardiomediastinal silhouette is within normal limits . The lungs are clear. There is no evidence of pleural effusion. The visualized osseous structures are unremarkable. |
| | Interval intubation with the endotracheal tube tip 3.7 cm above the carina. Interval placement of a left internal jugular sheath and Swan-Ganz catheter with the tip in the main pulmonary artery. Interval removal of a left internal jugular central venous catheter. Decreased bilateral trace pleural effusions and mild pulmonary edema. Persistent bibasilar opacities, likely atelectasis versus consolidation. No pneumothorax. | Interval placement of a right internal jugular central venous catheter with the tip in the superior vena cava . Persistent bibasilar opacities likely atelectasis versus consolidation. Decreased bilateral pleural effusions. Mild pulmonary edema . Cardiomediastinal silhouette is unchanged . | Interval intubation with the endotracheal tube tip approximately 3 cm above the carina.. Interval placement of a left internal jugular Swan-Ganz catheter with the tip in the main pulmonary artery . stable appearance of a small left pleural effusion and bibasilar opacities which may represent atelectasis or consolidation . No pneumothorax . |

- **Case Studies:** Compare Ground Truth, MambaXray-VL, and EMRRG reports
- **Highlights:** Yellow: MambaXray-VL matches; Cyan: EMRRG matches; Pink: Both match
- **Observation:** EMRRG generates more accurate and detailed descriptions of pathologies.

Conclusion and Future Work



This section summarizes EMRRG's contributions and outlines potential research directions for X-ray medical report generation.

Key Contributions



- Proposed **EMRRG** architecture for efficient X-ray medical report generation
- Developed **LoRA_{P(X)}** method for parameter-efficient Mamba structure fine-tuning
- Introduced **Hybrid Decoder Layers** to replace conventional decoder layers in LLMs

Experimental Validation



- Achieved **SOTA performance** on IU X-ray, MIMIC-CXR, and CheXpert Plus datasets
- Demonstrated **0.325 BLEU-2** and **0.385 ROUGE-L** scores on IU X-ray dataset
- Validated efficiency with only **2.3% parameter training** compared to full fine-tuning

Future Research Directions



- Improve generalization across diverse NLG metrics like **CIDEr score**
- Enhance clinical accuracy for rare disease descriptions
- Optimize hybrid decoder layer distribution for better performance



Code Availability & Contributions



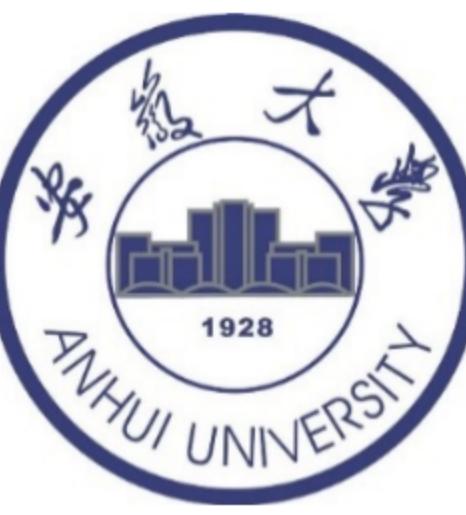
➤ **Open-Source Code:** https://github.com/Event-AHU/Medical_Image_Analysis

➤ **Our Related Projects:**

- [2025.10.27] EMRRG is accepted by BICS 2025 [<https://bics-ai.org/>]
- [2025.08.20] AM-MRG is accepted by IEEE Transactions on Medical Imaging (TMI)!
- [2025.08.06] [[R2GenKG](#)] is released on arXiv:2508.03426.
- [2025.02.27] [[CXPMRG-Bench](#)] is accepted by CVPR 2025!
- [2025.01.08] Activating Associative Disease-Aware Vision Token Memory for LLM-Based X-ray Report Generation, arXiv:2501.03458 is released on [[arXiv](#)]
- [2024.10.23] Pre-training on High Definition X-ray Images: An Experimental Study is accepted by [Visual Intelligence \(VI\)](#) Journal.
- [2024.10.01] A Pre-trained Large Model MambaXray-VL and the benchmark for the CheXpert Plus dataset is released [[arXiv:2410.00379](#)]
- [2024.08.20] R2GenCSR: Retrieving Context Samples for Large Language Model based X-ray Medical Report Generation is released [[arXiv:2408.09743](#)]
- [2024.04.28] Pre-training on High Definition X-ray Images: An Experimental Study, Xiao Wang, Yuehang Li, Wentao Wu, Jiandong Jin, Yao Rong, Bo Jiang, Chuanfu Li, Jin Tang arXiv pre-print, arXiv 2024, is released on arXiv [[Paper](#)]



About the Authors



汤进



王道



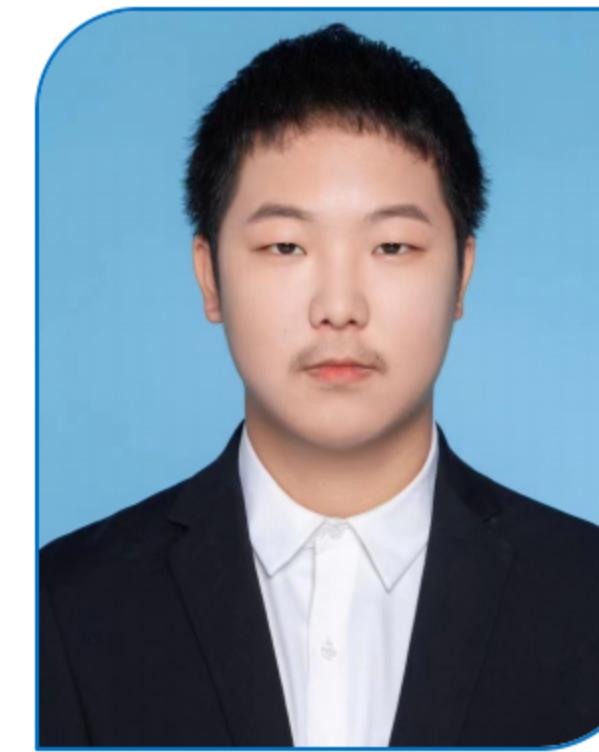
陈岚



张明政



高金锋



余江睿



乔雨晗



徐丹





Thank You!