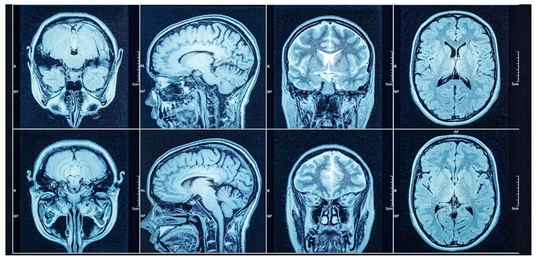
RegionGate-CT: Region Adaptive Fusion for Anatomy-Grounded CT Report Generation

Md Mustafizur Rahman and Dr. Mylène C.Q. Farias
Intelligent Multimedia Processing and Computing Technologies (IMPACT) Lab, Department of Computer Science, Texas State University

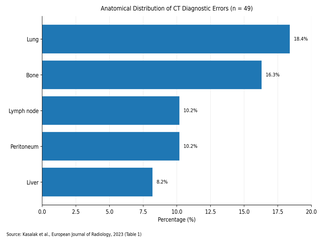
Motivation

- CT scans contain hundreds of slices, requiring reasoning across multiple anatomical regions.
- Radiologists operate under high workload and limited availability. Diagnostic errors still occur and can impact patient outcomes.



- Accurate report generation requires precise, region-specific descriptions grounded in anatomy.
- Recent AI-based methods attempt to assist relying on static fusion and uniform global context, leading to weak anatomical grounding and irrelevant contextual information.

To address this, we develop a **framework for anatomically grounded CT report generation**.

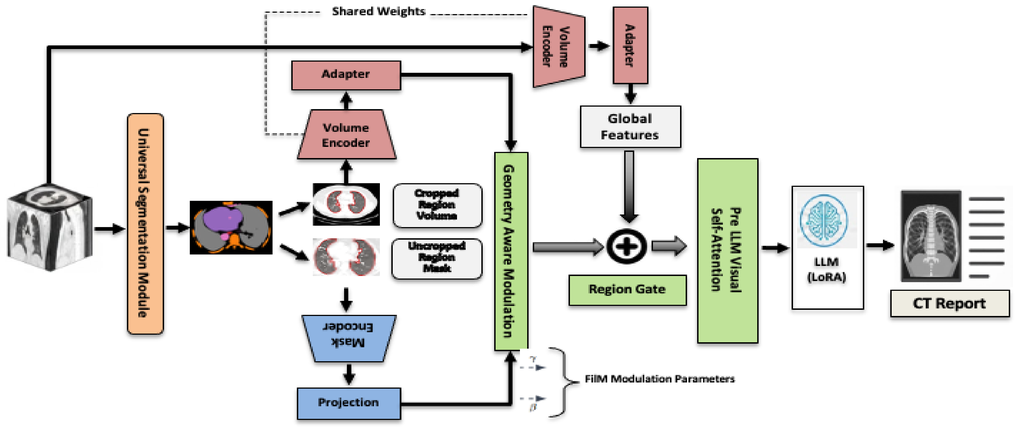


Existing CT Report Generation Methods

- Automated CT report generation aims to map volumetric imaging data to structured clinical text, reducing reliance on manual interpretation.
- Recent vision–language models improve performance by linking textual descriptions with anatomical regions, enhancing interpretability.
- However, many existing methods rely on static fusion of global and regional features, often using simple concatenation.
- They also apply the same global context across all regions, which can introduce irrelevant information and weaken region-specific reasoning.
- These limitations highlight the need for region-adaptive modeling in CT report generation.

System Architecture

- The input CT volume is decomposed into multiple anatomical regions using segmentation masks.
- Each region is encoded to obtain regional features, while the full volume provides a global feature.



- Geometry-aware modulation injects spatial priors into regional features to preserve anatomical structure.
- A region gate selectively integrates global context into each region, enabling region-specific conditioning.
- The combined visual tokens are refined via self-attention and passed to a large language model (LLM) to generate the CT report.

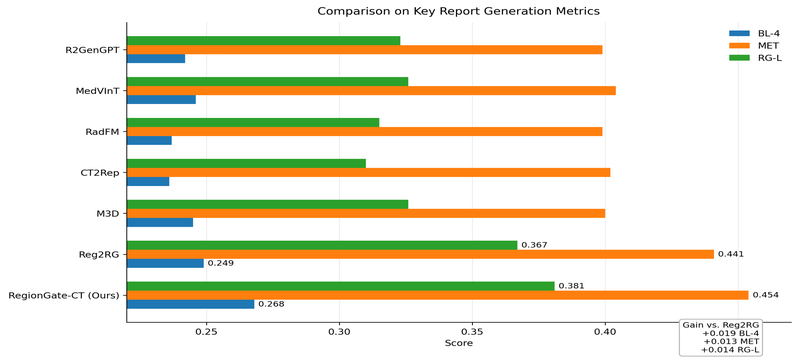
Dataset

We use the RadGenome–Chest CT dataset with region-level annotations for anatomically grounded report generation.

Scans	∼25K CT volumes
Patients	∼21K
Regions	197 anatomical structures
Annotations	Grounded reports + VQA

Experimental Results

- RegionGate-CT achieves consistent improvements across all evaluation metrics compared to prior methods.



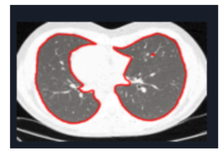
Analysis

- The proposed method outperforms the strongest baseline (Reg2RG) by +1.9 BLEU-4, +1.3 METEOR, and +1.4 ROUGE-L.
- These gains indicate improved linguistic quality and semantic alignment in generated reports.
- The results validate the effectiveness of region-adaptive global–regional fusion.

Qualitative Example of the Generated Report

RegionGate-CT generates globally consistent report.

Region: Lung



Ground Truth	Aeration of both lung parenchyma is normal and no infiltrative lesion is detected. A few millimetric nonspecific nodules and mild recessions are observed in the right lung upper and lower lobes.
Generated	Aeration of both lung parenchyma is normal and no nodular or infiltrative lesion is detected in the lung parenchyma.

Conclusion

- We propose RegionGate-CT, a region-adaptive framework for CT report generation.
- Region-specific integration of global context improves anatomical grounding and report quality.
- Experimental results show consistent gains over previous state of the art methods

Future Works

- Extend to more diverse datasets.
- Incorporate clinical priors for improved diagnostic specificity.

Md Mustafizur Rahman, yks9@txstate.edu
Dr. Mylène C.Q. Farias, mylene@txstate.edu

TEXAS STATE UNIVERSITY