Natural Language Processing for Automated Radiology Report Generation: A Paradigm Shift in Clinical Workflow

Rasit Dinc

Rasit Dinc Digital Health & AI Research

Published: October 26, 2025 | Medical Imaging AI

DOI: 10.5281/zenodo.17996505

Abstract

The exponential growth in medical imaging studies burdens radiologists, as traditional manual reporting causes inter-observer variability and workflow bottle...

Introduction

The exponential growth in medical imaging studies burdens radiologists, as traditional manual reporting causes **inter-observer variability** and workflow bottlenecks [1]. **Natural Language Processing (NLP)**, a core component of artificial intelligence, offers a transformative solution for automating and standardizing radiology report generation. This overview explores the technical foundations, benefits, and challenges of leveraging NLP.

The Technical Evolution of NLP in Radiology

Automated radiology report generation requires sophisticated NLP models to interpret medical images and translate findings into coherent, clinically accurate text. The field's evolution can be categorized into key technical approaches:

1. Traditional and Machine Learning Methods

Early approaches included **rules-based text-matching algorithms** and conventional machine learning models. Rules-based systems, while precise, require meticulous handling of complex linguistic features like **negation handling** (e.g., "no evidence of fracture") [2]. Conventional machine learning models, often using **Term Frequency-Inverse Document Frequency (TF-IDF)** vectorization, are limited by the need for extensive data preprocessing and their difficulty in capturing the nuanced context of medical language [2].

2. Deep Learning and Domain-Specific Models

Deep learning introduced powerful models like **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory (LSTM)** networks, which capture sequential dependencies. The adoption of the **Transformer** architecture, exemplified by **BERT**, marked a significant leap. Domain-specific variants, such as **ClinicalBERT**, fine-tuned on clinical notes, show superior performance in information extraction and automated report labeling due to their deep understanding of biomedical terminology [3].

3. The Rise of Large Language Models (LLMs) and Vision-Language Models (VLMs)

The most promising development is the application of **Large Language Models (LLMs)** and **Vision-Language Models (VLMs)**. VLMs are multimodal models that process both the radiological image and language, directly generating a report from the image data [4]. LLMs can refine, structure, or generate reports from radiologist impressions, promoting **structured reporting** which is critical for improving data consistency and facilitating downstream data analysis [5].

Key Benefits for Clinical Practice

The automation of report generation offers profound benefits for the radiologist and the healthcare system:

| Benefit | Description | Impact on Clinical Workflow | | :--- | :--- | :--- | | **Efficiency and Throughput** | Reduces the time required for report dictation and generation. | Decreases reporting turnaround time, increasing study throughput. | | **Consistency and Standardization** | Enforces standardized terminology and structure, minimizing variability. | Improves report clarity and reliability for clinicians and patient care. | | **Data Extraction and Research** | Converts unstructured free-text reports into structured, machine-readable data. | Facilitates large-scale data mining for research, quality assurance, and training of new AI models. | | **Decision Support** | Can highlight critical findings or suggest differential diagnoses based on image analysis. | Acts as a cognitive assistant, reducing the risk of oversight in complex cases. |

Studies have shown that keyword-based AI assistance can reduce reporting time by over 27% without a significant difference in report quality, underscoring the potential for efficiency gains [6].

Challenges and Future Directions

Despite rapid advancements, several challenges must be addressed before widespread clinical adoption:

1. **Data Quality and Availability:** Training robust models, especially VLMs, requires massive, high-quality, and ethically sourced datasets of paired images and reports. Publicly available datasets remain limited, particularly for modalities other than chest X-rays [2]. 2. **Clinical Validation and Trust:** Models must be rigorously validated in diverse clinical settings to ensure accuracy and safety across different patient populations and imaging protocols. Building trust among radiologists and referring physicians is paramount. 3. **Handling Medical Complexity:** The inherent complexity of medical language, including specialized jargon, acronyms, and the critical

importance of accurate negation handling, remains a technical hurdle that requires continuous model refinement [3].

The future lies in hybrid models that integrate the strengths of deep learning with the contextual reasoning of LLMs. As these models become more sophisticated and clinically validated, they are poised to become integral to the diagnostic process, enhancing patient care and redefining the role of the modern radiologist.

**

References

[1] Min, J. Y. (2025). Automating Radiology Report Generation: A Systematic Review*. Preprints.org. [https://www.preprints.org/manuscript/202503.2218/v1] (https://www.preprints.org/manuscript/202503.2218/v1) [2] Chng, S. Y., et al. (2023).* Automated labelling of radiology reports using natural language processing: Comparison of traditional and newer methods*. Health Care Science, 2(2), 120–128. [https://pmc.ncbi.nlm.nih.gov/articles/PMC11080679/] (https://pmc.ncbi.nlm.nih.gov/articles/PMC11080679/) [3] Hsu, E., et al. (2023).* Extraction of Radiological Characteristics From Free-Text Radiology Reports Using a Domain-Specific Natural Language Processing Pipeline*. JMIR AI, 2(1), e42884. [https://ai.jmir.org/2023/1/e42884/] (https://ai.jmir.org/2023/1/e42884/) [4] Artsi, Y., et al. (2025).* Large language models in radiology reporting*. Clinical Imaging, 111, 109968. [https://www.sciencedirect.com/science/article/pii/S2666521225000912] (https://www.sciencedirect.com/science/article/pii/S2666521225000912) [5] Busch, F., et al. (2024).* Large language models for structured reporting in radiology*. NPJ Digital Medicine, 7(1), 1-10. [https://pmc.ncbi.nlm.nih.gov/articles/PMC12021971/] (https://pmc.ncbi.nlm.nih.gov/articles/PMC12021971/) [6] Dong, F., et al. (2025).* Keyword-based AI assistance in the generation of radiology reports: a randomized controlled trial*. European Radiology, 35(5), 1877-1885. [https://pmc.ncbi.nlm.nih.gov/articles/PMC12316847/] (https://pmc.ncbi.nlm.nih.gov/articles/PMC12316847/)
