Redefining MRI:
From Image to Insight

ICMRI 2026
THE 14th INTERNATIONAL CONGRESS ON MRI
& 31st ANNUAL SCIENTIFIC MEETING OF KSMRM

AUGUST 27 - 28, 2026
Grand Walkerhill Seoul, Seoul, Korea
https://icmri.ksmrm.org/2026m

Meet the New Experts

SY16-1

From Text Understanding to Visual Grounding: Current Progress and Future Directions of Large Language Models in Neuroradiology

Pae Sun Suh

Research Institute of Radiological Science and Center for Clinical Imaging Data Science, Yonsei University College of Medicine, Korea

The rapid evolution of large language models (LLMs) has generated considerable interest in their potential applications in neuroradiology. Initial investigations primarily evaluated diagnostic accuracy using medical image challenges and quiz-based datasets. Although these studies demonstrated encouraging performance, their findings remain difficult to translate into routine clinical practice.

LLMs have shown promising performance in a variety of text-based tasks, including radiology report generation, structured information extraction, and automated clinical history generation. These findings suggest that language understanding and report-level reasoning have advanced to a level at which LLMs can effectively support neuroradiology workflow. In contrast, image-based tasks have remained substantially more challenging. Previous studies consistently demonstrated that diagnostic performance is strongly influenced by accompanying textual information, while image interpretation remains susceptible to hallucinations, including fabricated imaging findings and incorrect anatomical localization.

Nevertheless, recent investigations indicate that modern vision-language models (VLMs) have substantially improved their performance in neuroradiology. Despite these improvements, an important limitation remains: current VLMs largely function as black boxes, providing diagnostic outputs without explicitly revealing which imaging features or anatomical structures contributed to their decisions. Existing studies have focused predominantly on diagnostic accuracy and, in some cases, textual explanations of reasoning, while the underlying visual reasoning process has remained largely unexplored.

To address this challenge, our ongoing work investigates visual grounding in neuroradiology. Rather than evaluating diagnostic accuracy alone, this approach examines whether VLMs can explicitly localize anatomically relevant regions during image interpretation, providing a more transparent representation of visual understanding. Although visual grounding alone is unlikely to fully explain VLM reasoning, it may represent an important first step toward systematically investigating how these models interpret neuroimaging, integrate visual evidence, and generate diagnostic conclusions, thereby providing a foundation for future studies on explainable AI in neuroradiology.

Keywords: *Large language model, Vision language model, Neuroradiology*