World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:19, No:02, 2025

Explainable Model for Fetal Plane Classification using VIT and LRP

Authors : Sinchana S., Nisarga Kunder, Siri Gowri H., Shravani Hiremath

Abstract : For prenatal diagnosis of anomalies and monitoring them, accurate classification of fetal anatomical planes from ultrasound images is essential because it allows clinicians to evaluate fetal development and administer treatment accordingly. However, manual classification of these planes is time-consuming and heavily dependent on the operator's expertise. In recent years, encoder-decoder-based Vision Transformers (ViT) have emerged as a better alternative to traditional convolutional neural networks (CNNs) in various natural language processing tasks like text prediction and image-based tasks like classification, achieving state-of-the-art performance. Despite these advancements, the use of ViTs within the medical domain, especially for MRI and Ultrasound images, remains relatively less due to their intricate architecture and relative difficulty in applying explainability techniques, as ViTs' attention-based mechanisms are more complex to interpret than those of CNNs. In clinical environments, model interpretability is essential; thus, explainability is crucial for delivering consistent and transparent decision support. This interpretability allows physicians to understand and verify model outputs, fostering trust in AI-driven diagnostic processes. Although CNN-based models frequently use explainability techniques like Grad-CAM, transformer-based models lack similar interpretability due to their distinct attention mechanisms and architectural structure. In this work, there is the introduction of a novel approach using Layer-wise Relevance Propagation (LRP) to provide visual explanations for Vision Transformer (ViT) predictions on fetal ultrasound images. With LRP, demonstrated the feasibility of deriving meaningful and interpretable insights from ViT models in medical imaging, paving the way for reliable and explainable AI applications in healthcare.

Keywords : fetal ultrasound, vision transformer (ViT), layer-wise relevance propagation (LRP), explainability, medical image analysis, model interpretability
Conference Title : ICCSCIT 2025 : International Conference on Computer Science, Cybersecurity and Information Technology
Conference Location : Mumbai, India
Conference Dates : February 17-18, 2025