Interpretability Challenges in Multimodal AI

Interpretation difficulties due to multimodal model complexity, utilizing attention visualization and SHAP

Need for Interpretability

Ensuring reliability of medical AI

- Supporting clinical decision-making
- Regulatory requirements (FDA, CE)
- Patient explanation responsibility
- Model debugging and improvement

Multimodal Interpretation Challenges

Sources of complexity

- Multi-layer fusion structure
- Inter-modal interactions
- Non-linear transformations
- High-dimensional representation space

Attention Visualization

Attention weight analysis

- Cross-attention maps
- Modal-specific contributions
- Heatmap overlay
- Temporal attention patterns

XAI Techniques

Explainable AI methodology

- SHAP: Shapley values
- LIME: Local approximation
- Grad-CAM: Gradient-based
- Integrated Gradients

Modal Importance Scoring	Feature Contribution Analysis	Contrastive Case Presentation