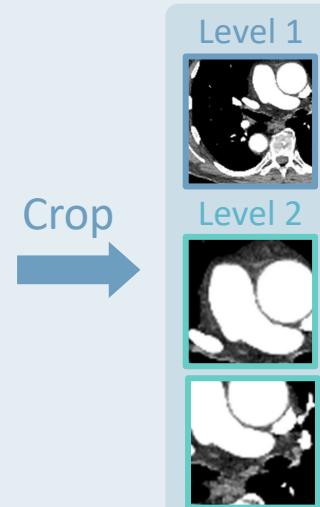
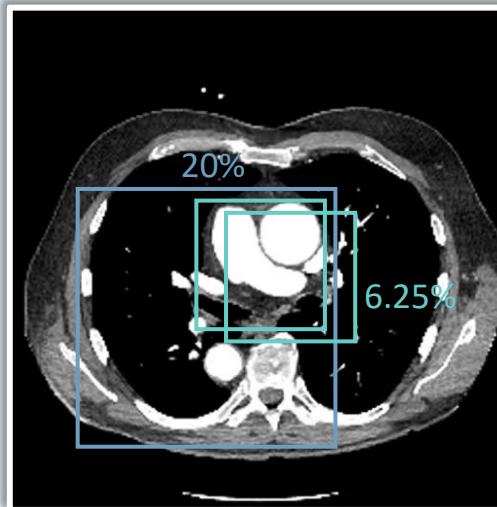
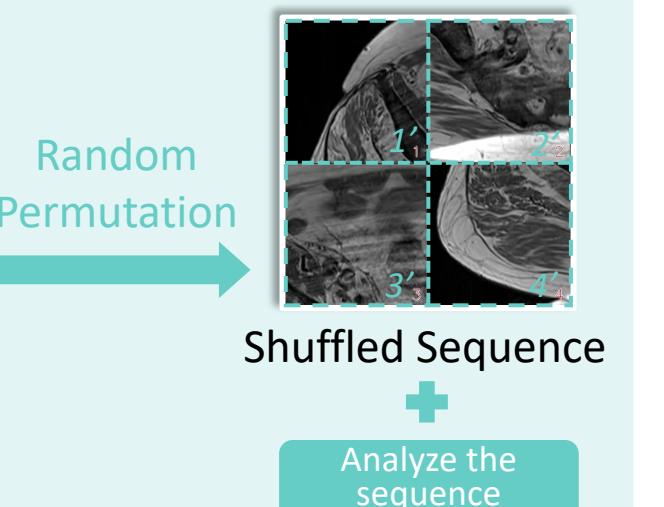
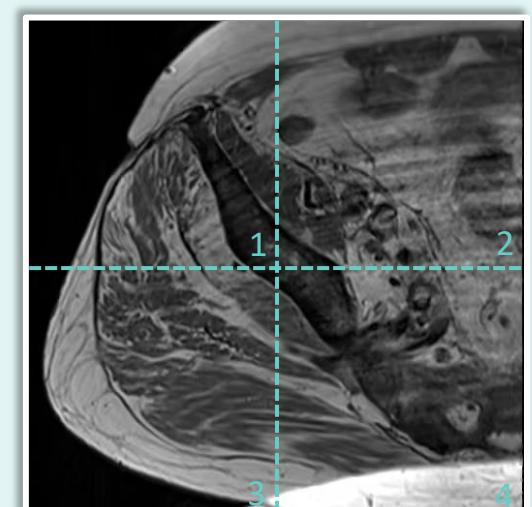


Med-Scout

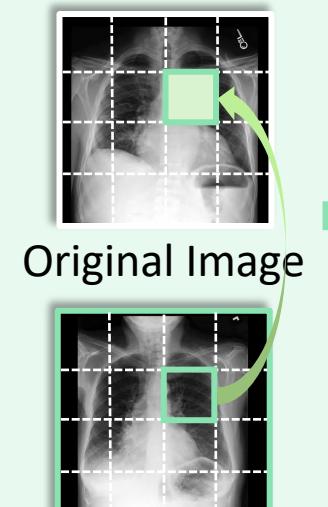
Task A: Hierarchical Scale Localization



Task B: Topological Jigsaw Reconstruction



Task C: Anomaly Consistency Detection



$\langle \text{Image Input} \rangle + \langle \text{Text Instruction} \rangle \rightarrow \langle \text{Target Answer} \rangle$

Unified VQA Stream

Image + Prompt

Ground Truth

MLLM

 Qwen3-VL

HuatuoGPT-Vision



LINGSHU

Output

Dense Geometric Reward

Task A

Level Prediction	1	1	2
Ground Truth	2	1	2
Reward: 2/3			

GT
Predicted box

Task B

Model Prediction	3	2
Ground Truth	3	1
Reward: 1/2		

Task C

GT	
Predicted center	

Reward = distance

GRPO