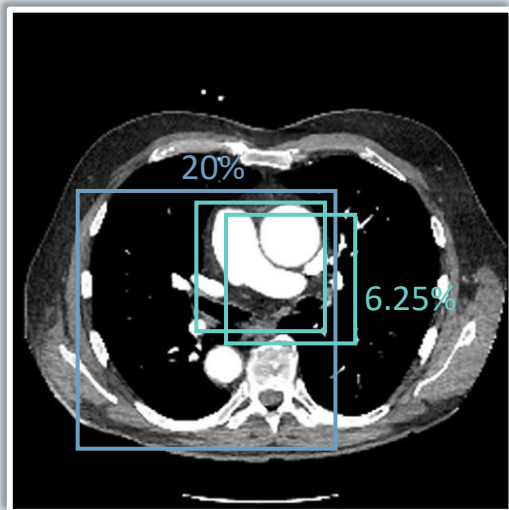


Med-Scout

Task A: Hierarchical Scale Localization



Original Image I

Crop



Patches

Resize

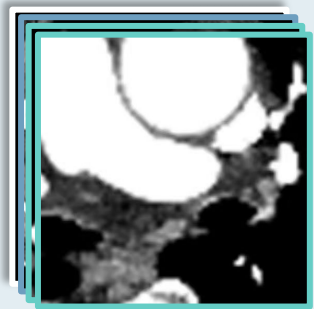
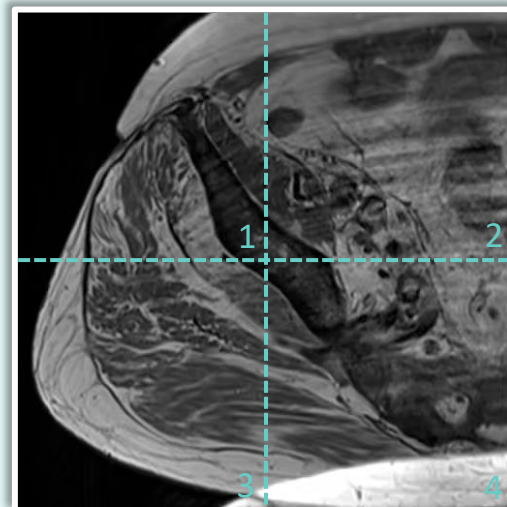


Image I & patches

Identify scale & bounding box

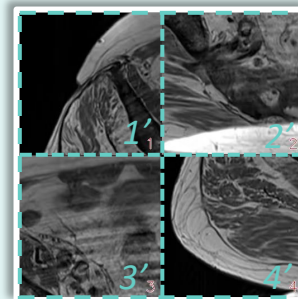
Prompt

Task B: Topological Jigsaw Reconstruction



Original Image I

Random Permutation

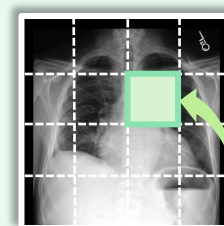


Shuffled Sequence

Analyze the sequence

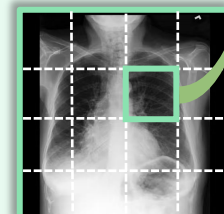
Prompt

Task C: Anomaly Consistency Detection

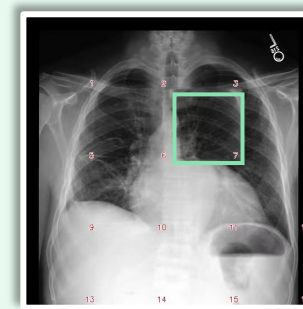


Original Image

Seamless Blending



Reference Image



Target Image

Output the inconsistent patch

Prompt

<Image Input> + <Text Instruction> → <Target Answer>

Unified VQA Stream

Image + Prompt

Ground Truth

MLLM



Qwen3-VL

HuatuoGPT-Vision



LINGSHU

...

Output

GRPO

Dense Geometric Reward

Task A

Level Prediction

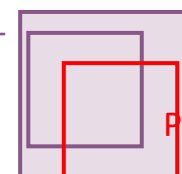
1 1 2

Ground Truth

2 1 2

Reward: 2/3

GT



Reward: IoU

Predicted box

Task B

Model Prediction

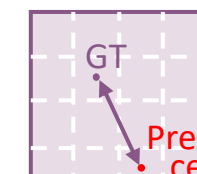
3 2
1 4

Reward: 1/2

Ground Truth

3 1
2 4

Task C



Reward = distance