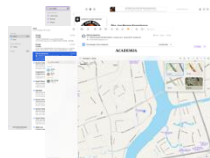


(a) Random Select



Samples with overly simple instructions, or rollout results that are all correct or all wrong, are not informative for GRPO.

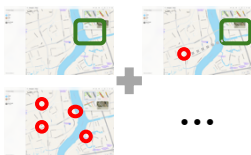


Ignoring the different grounding difficulty of each sample.

(b) GUI-D (Ours)



Retain data with training value



Difficulty score guidance

GRPO



High training efficiency

(c) Agentic RL



$\frac{|Crop \cap GT|}{|GT|}$ -style

(e.g., GUI-Eyes)



Redundancy, Time-consuming



IoU-style



Easy to miss, Time-consuming



One-step fixed size

(e.g., AdaZoom-GUI)



Inflexible

(d) GUI-C² (Ours)



Click or Crop?



Coarse to fine without thinking

Click or Crop?



Improvement-aware stage rewards that ensure each refinement is effective



High efficiency, Stable, Lightweight, Without <think>

(e) Performances

