

# Enhancing Visual Grounding for GUI Agents via Self-Evolutionary Reinforcement Learning

May 19, 2025

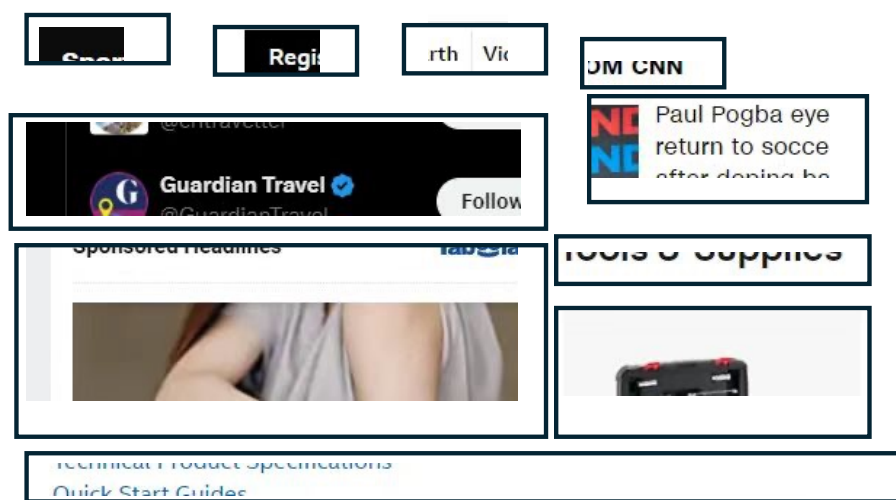
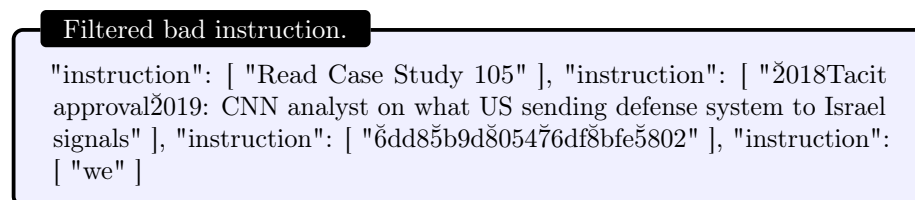


Figure 1: Filtered bad box in seed data curation stage. Most bad cases include some incomplete UI elements or some that contain messy background information.



Most web data includes a large number of bad instructions, some of which contain garbled letters, while others are ambiguous, containing only a single word or an unclear phrase .

## Average Attention Map

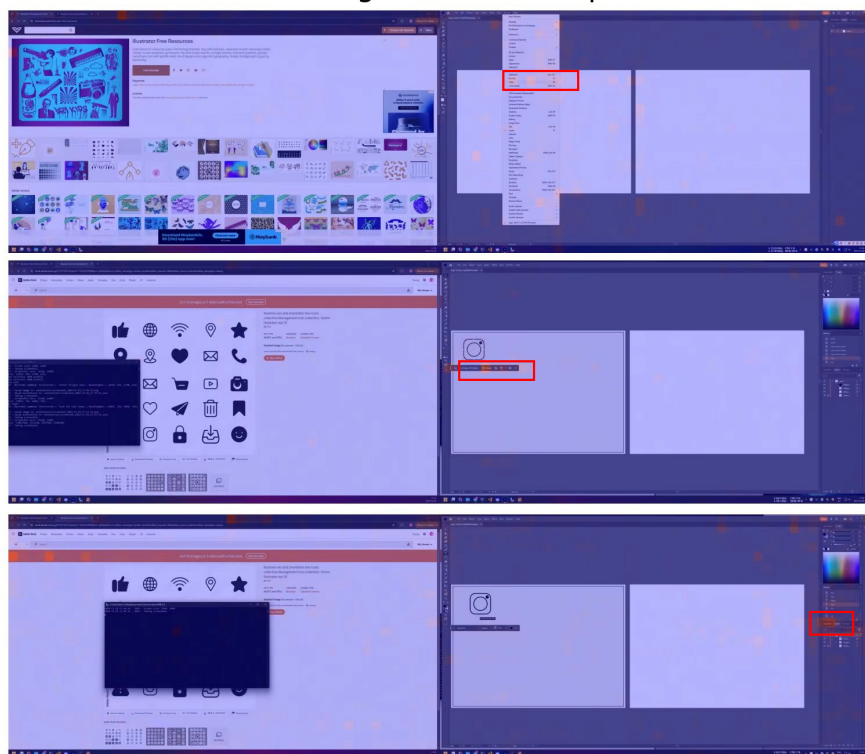


Figure 2: we present more visualization examples of attention maps.