Step 1: Uncertainty Quantification
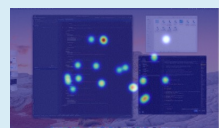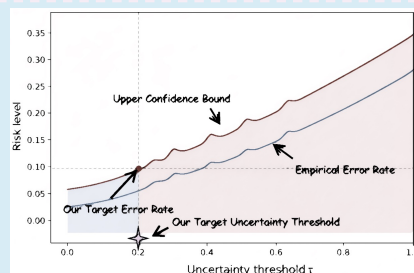
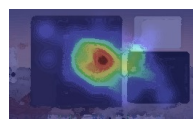(100, 820),
(342, 1576),
(1890, 433),
(756, 1294),
(1210, 98),
(1678, 1822)
...

Step 2: Selective Prediction Calibration
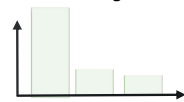
Upper Confidence Bound

Empirical Error Rate

Our Target Error Rate

Our Target Uncertainty Threshold

Risk level

Uncertainty threshold τ

Low-Uncertainty Scenario

Ranked Region Scores

✓ High Concentration
✓ Low Entropy
✓ Low Variance

Reliable Prediction!

High-Uncertainty Scenario

Ranked Region Scores

⚠ Low Concentration
⚠ High Entropy
⚠ High Variance

High Uncertainty Warning!

Decision Making

Execute:
uncertainty $\leq \tau$

✓

Abstain or cascade:
uncertainty $> \tau$

✗

User query

User's screen

SafeGround
(ours)

Local GUI
grounding
model

External
resource