

Steering Awareness Training Pipeline



CAA: $\text{mean}(h^+) - \text{mean}(h^-)$

50% steered + 50% replay

$r=32, \alpha=64, 4$ epochs

95.3% det, 0% FPR