

Memory_Exploration_case2 — step: Open the wooden cabinet top region [manipulate]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab



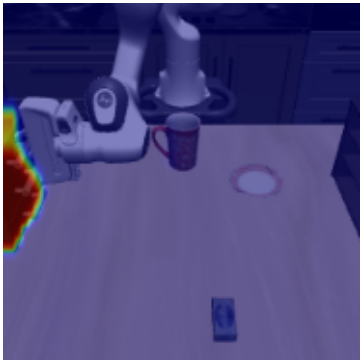
Memory_Exploration_case2 — step: Close the wooden cabinet top region [manipulate]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab

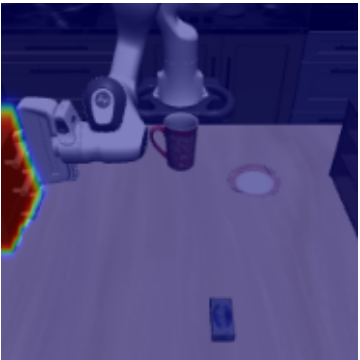
DINO-only



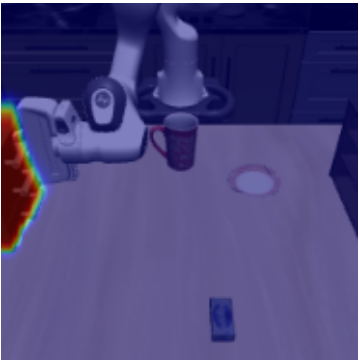
A



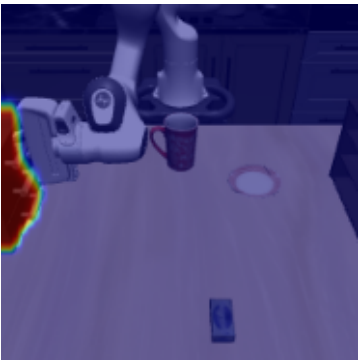
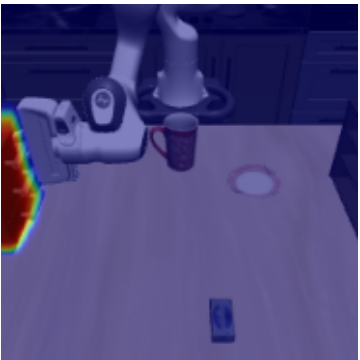
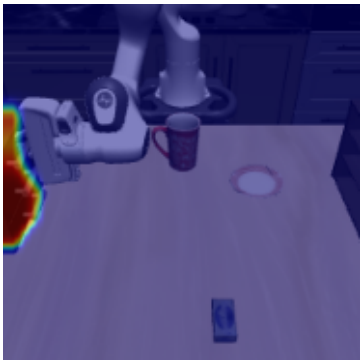
B



C

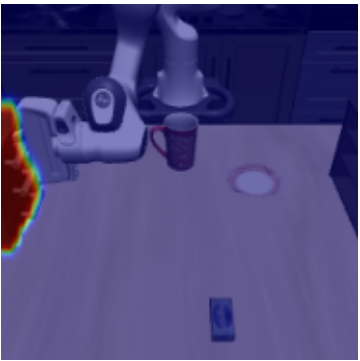
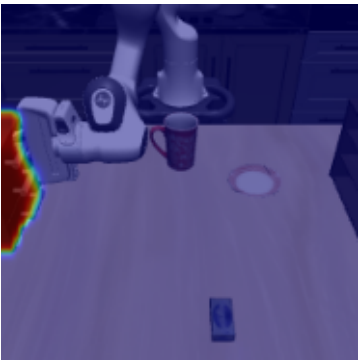
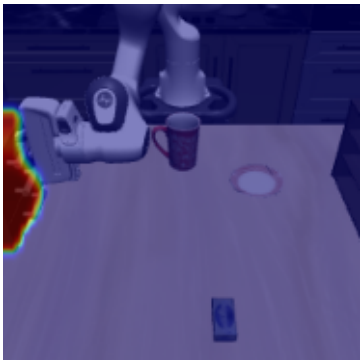


CLIP-only

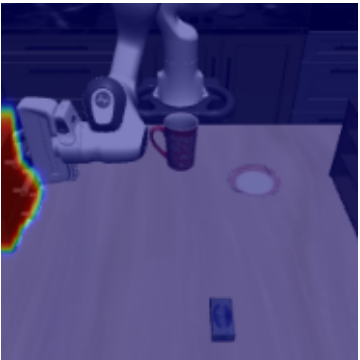
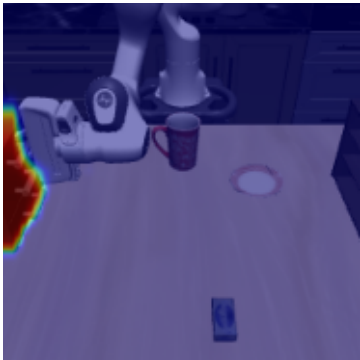


single-object GT (old) ↑ dual-object role-aware GT ↓

DINO+CLIP



role-aware



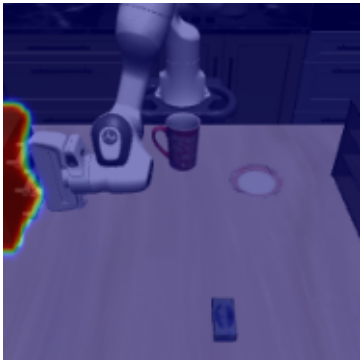
Memory_Exploration_case2 — step: Open the wooden cabinet middle region [manipulate]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab

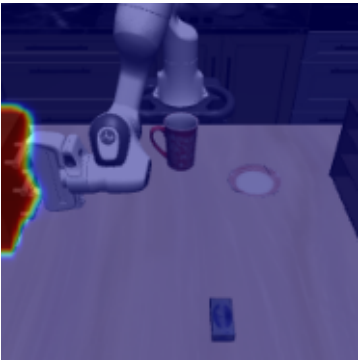
DINO-only



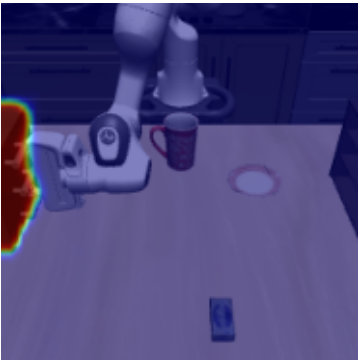
A



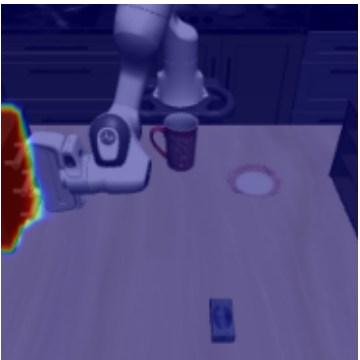
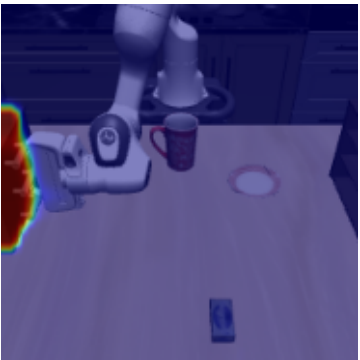
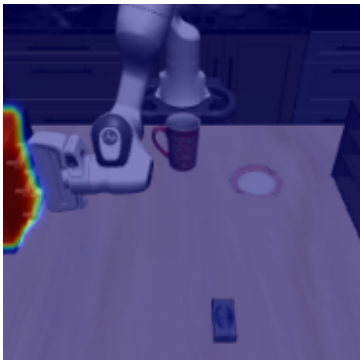
B



C

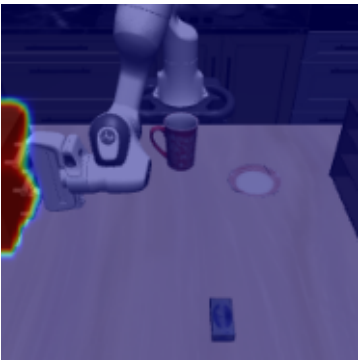
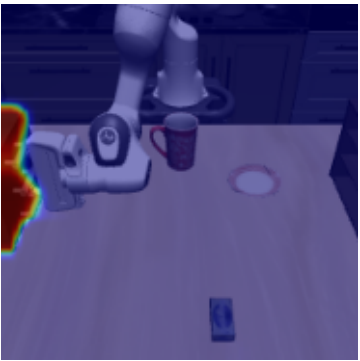


CLIP-only



single-object GT (old) ↑ dual-object role-aware GT ↓

DINO+CLIP



role-aware



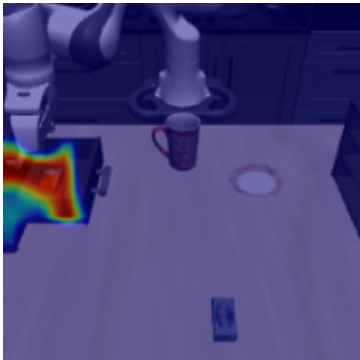
Memory_Exploration_case2 — step: Pick up the butter from the wooden cabinet middle region [grasp]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab

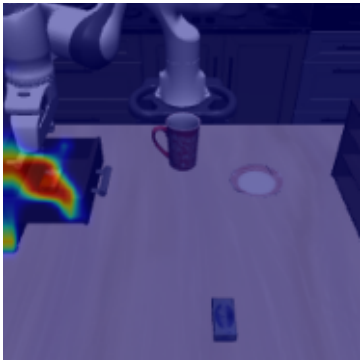
DINO-only



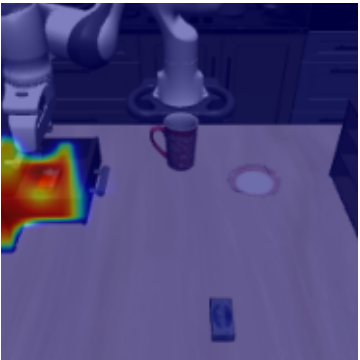
A



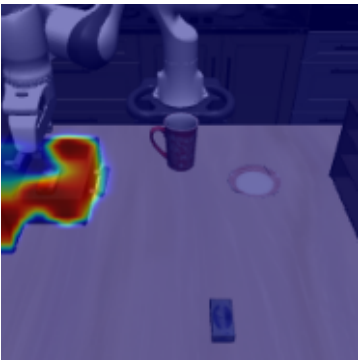
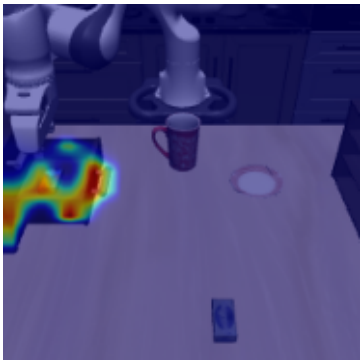
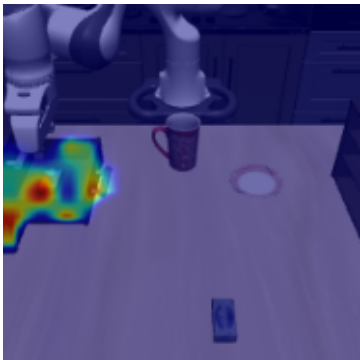
B



C

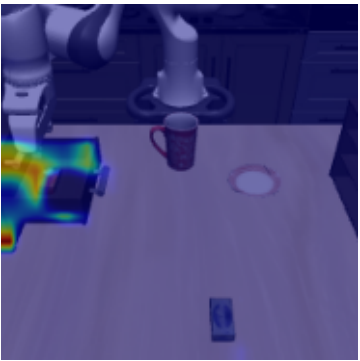
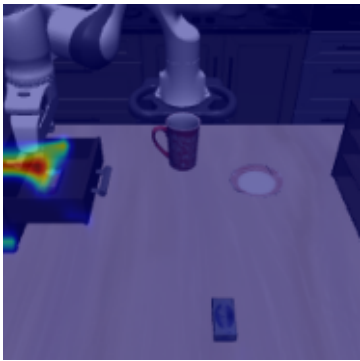
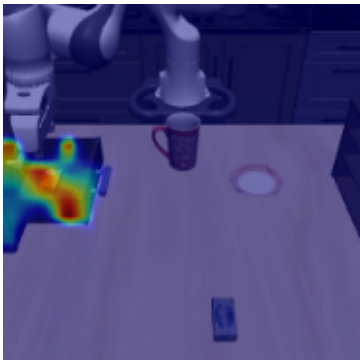


CLIP-only

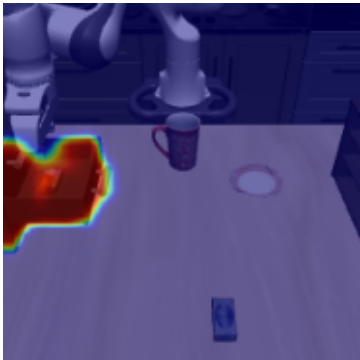
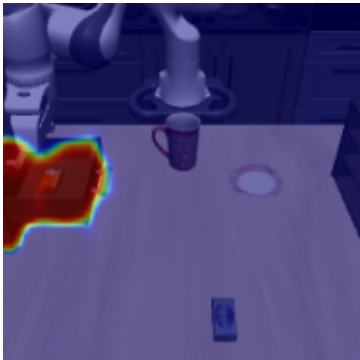


single-object GT (old) ↑ dual-object role-aware GT ↓

DINO+CLIP

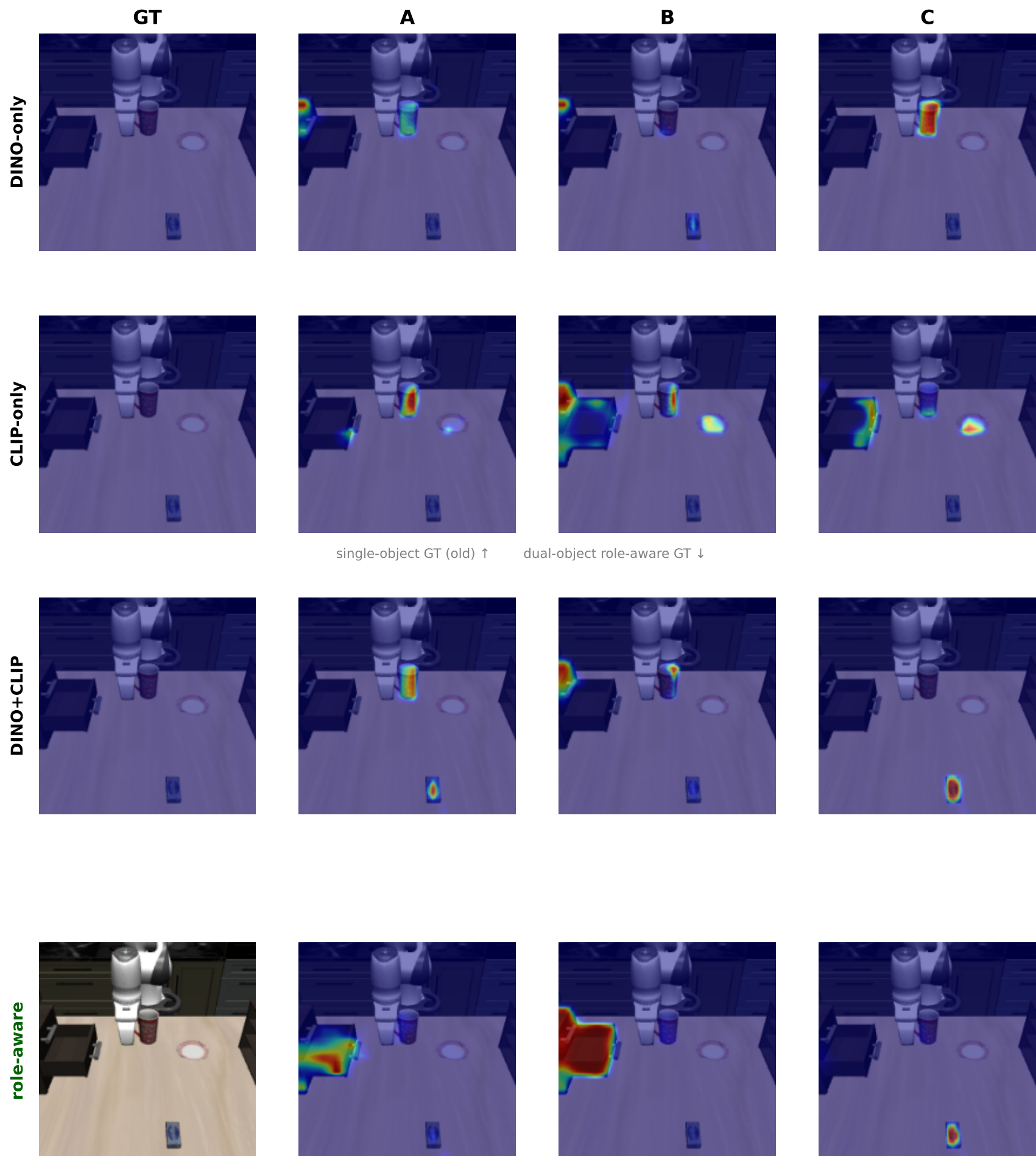


role-aware



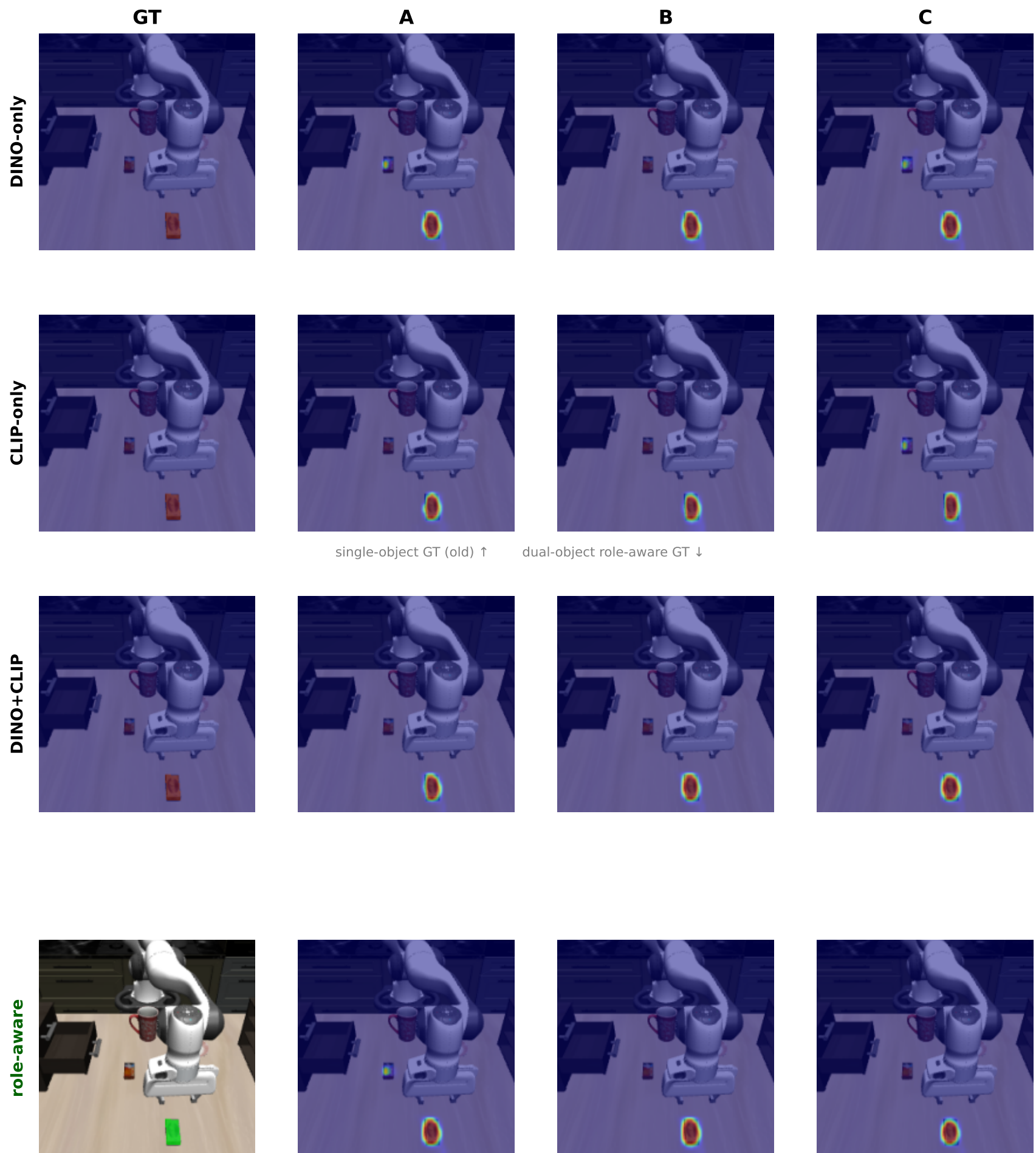
Memory_Exploration_case2 — step: Place the butter on the kitchen table [release]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab



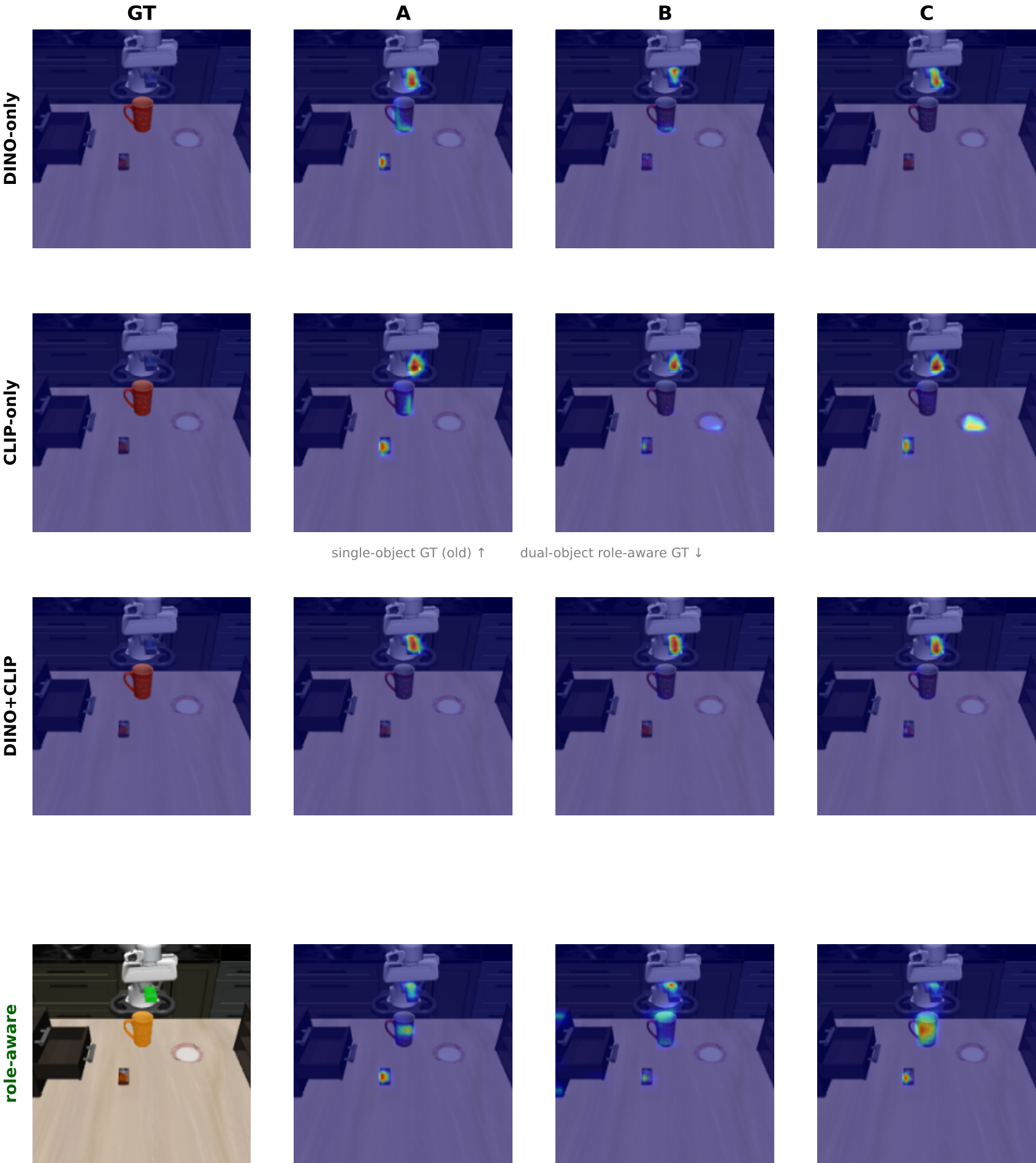
Memory_Exploration_case2 — step: Pick up the cream cheese from the kitchen table [grasp]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab



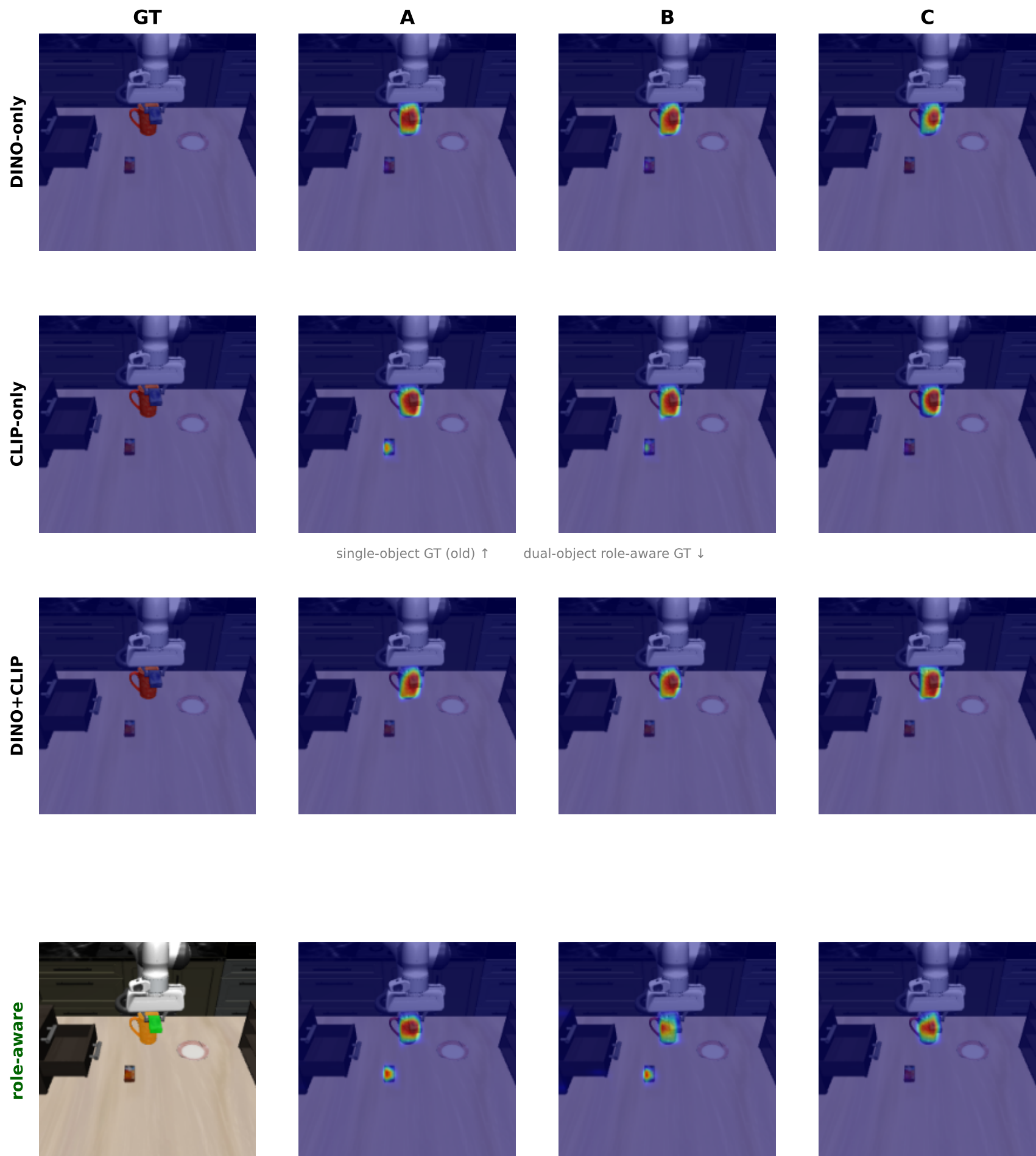
Memory_Exploration_case2 — step: Pour the cream cheese into the red coffee mug [manipulate]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab



Memory_Exploration_case2 — step: Place cream cheese on the kitchen table [release]

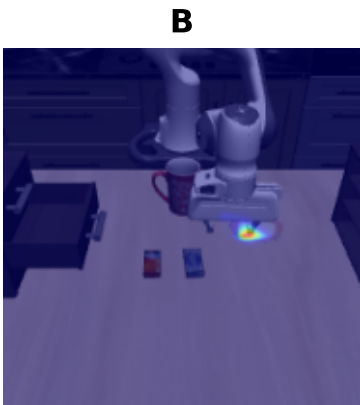
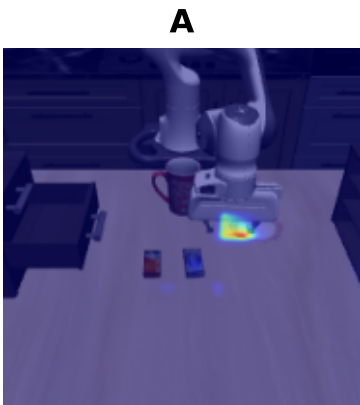
goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab



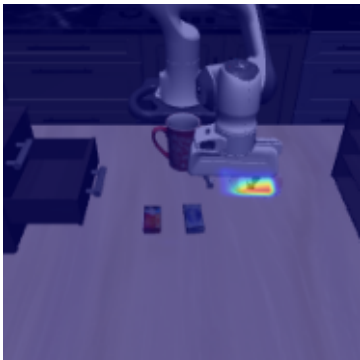
Memory_Exploration_case2 — step: Pick up the plate from the kitchen table [grasp]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab

DINO-only

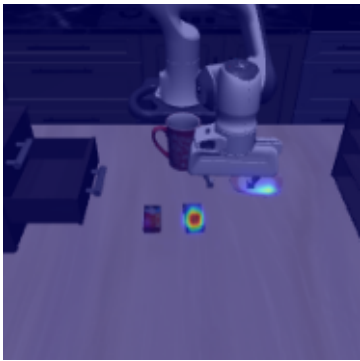
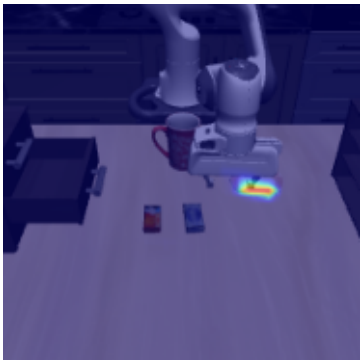
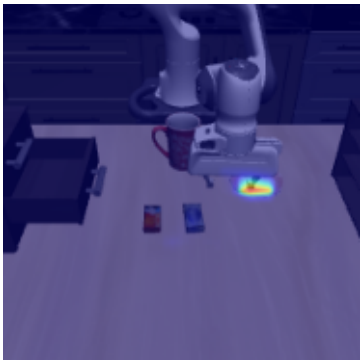


CLIP-only

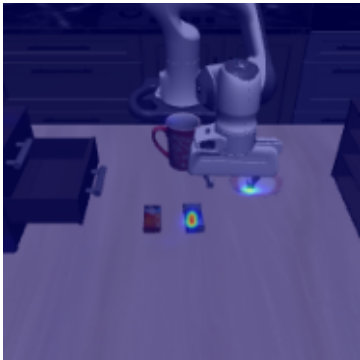


single-object GT (old) ↑ dual-object role-aware GT ↓

DINO+CLIP



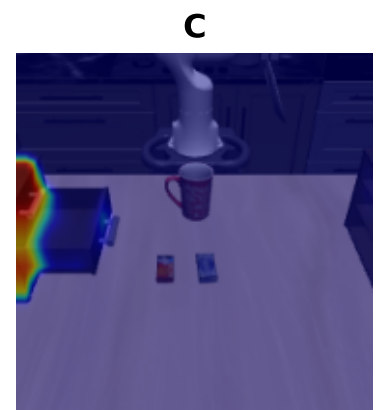
role-aware



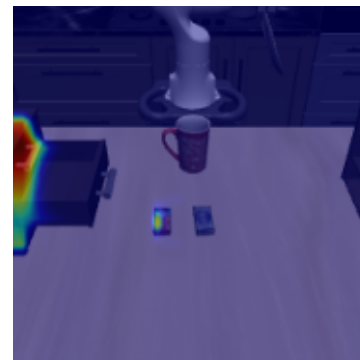
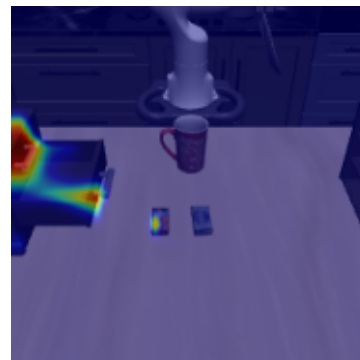
Memory_Exploration_case2 — step: Place the plate on the wooden shelf top side [release]

goal: Find and pick up the butter from the wooden cabinet, pour the cream cheese from the kitchen tab

DINO-only



CLIP-only



single-object GT (old) ↑

dual-object role-aware GT ↓

DINO+CLIP



role-aware

