

Mix_case8 — step: Pick up the butter from coffee table [grasp]

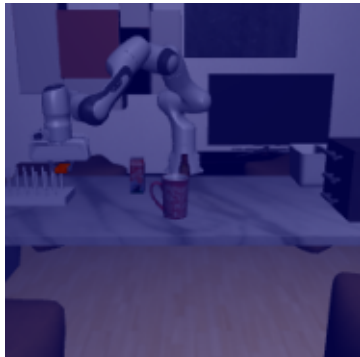
goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to



Mix_case8 — step: Place the butter onto the rack [release]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to

DINO-only



A



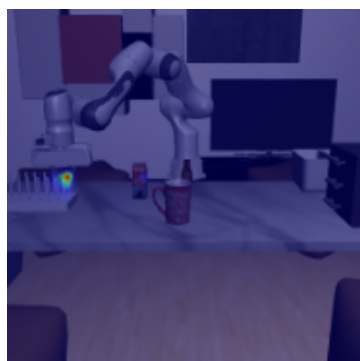
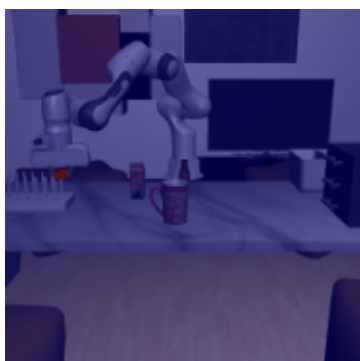
B



C



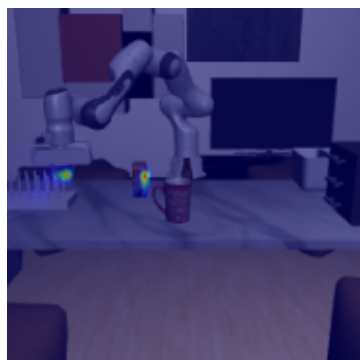
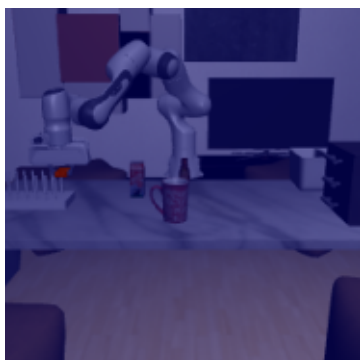
CLIP-only



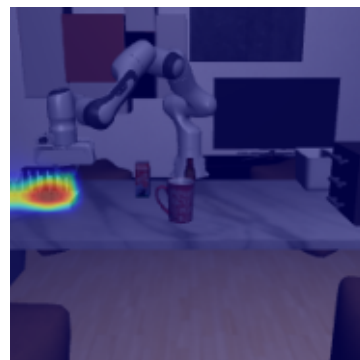
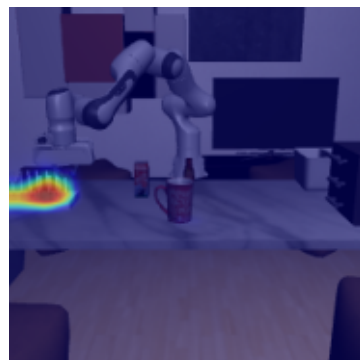
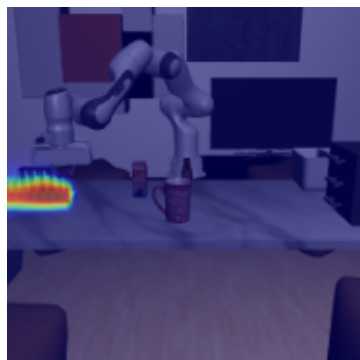
single-object GT (old) ↑

dual-object role-aware GT ↓

DINO+CLIP

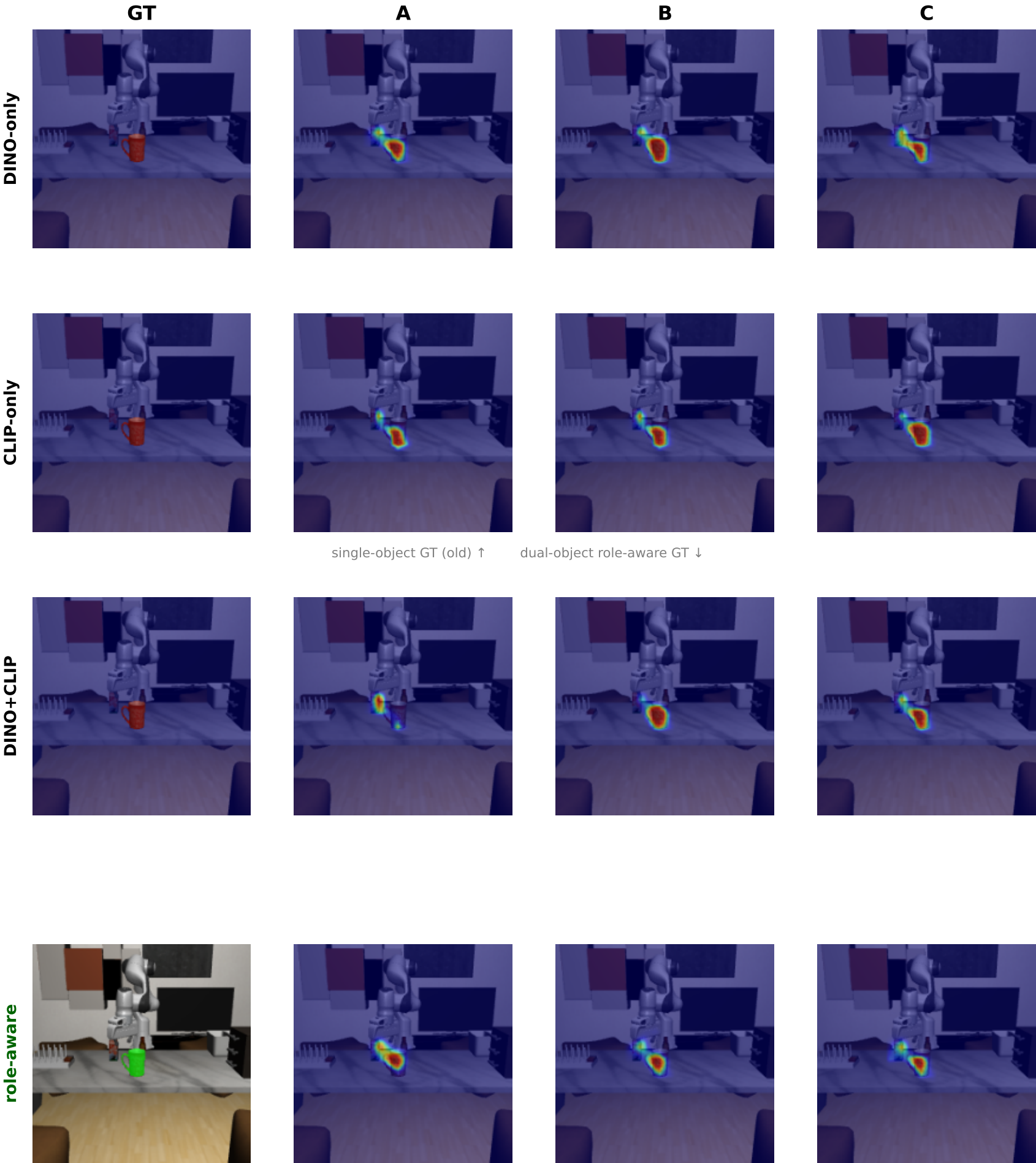


role-aware



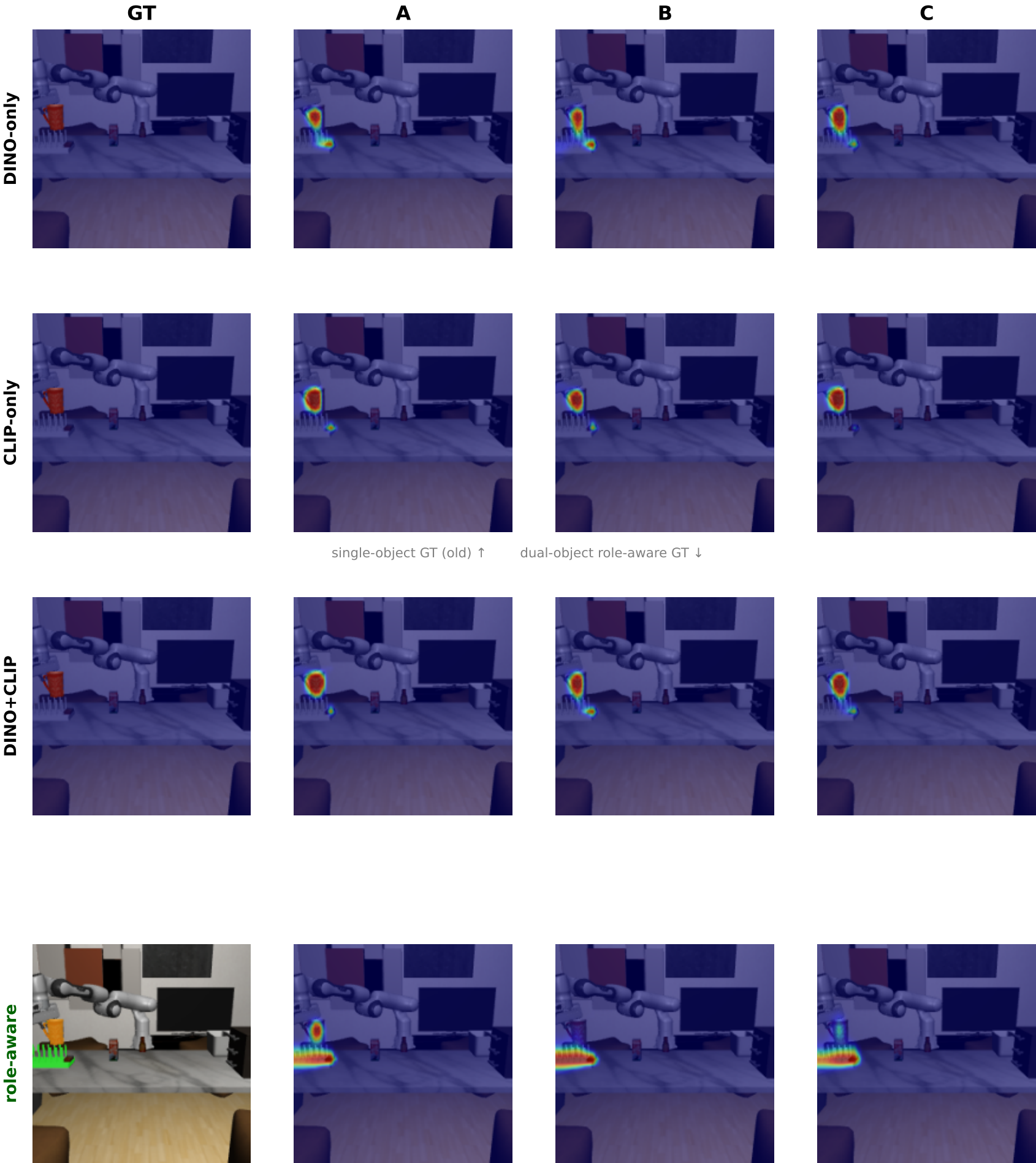
Mix_case8 — step: Pick up the red coffee mug from coffee table [grasp]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to



Mix_case8 — step: Place the red coffee mug onto the rack [release]

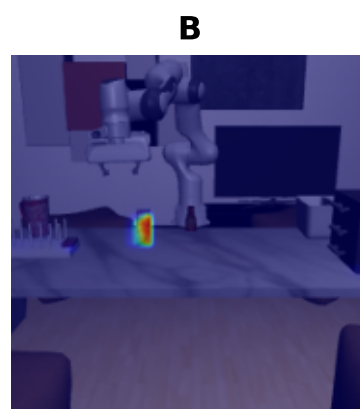
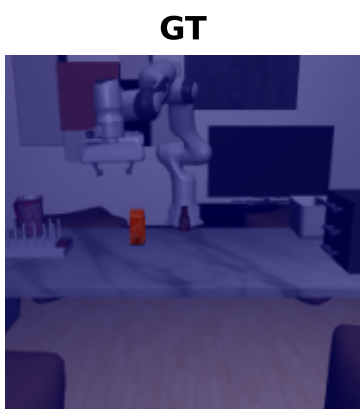
goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to



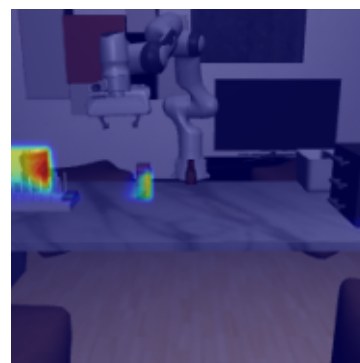
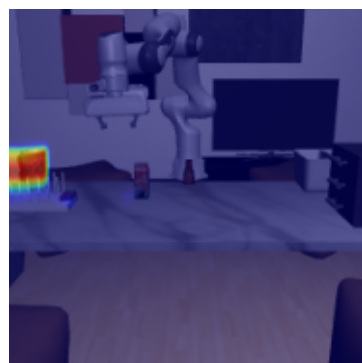
Mix_case8 — step: Pick up the milk from coffee table [grasp]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to

DINO-only



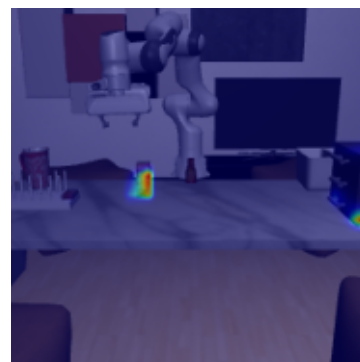
CLIP-only



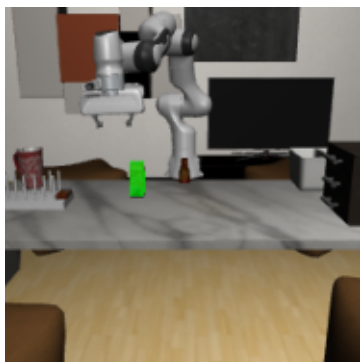
single-object GT (old) ↑

dual-object role-aware GT ↓

DINO+CLIP



role-aware



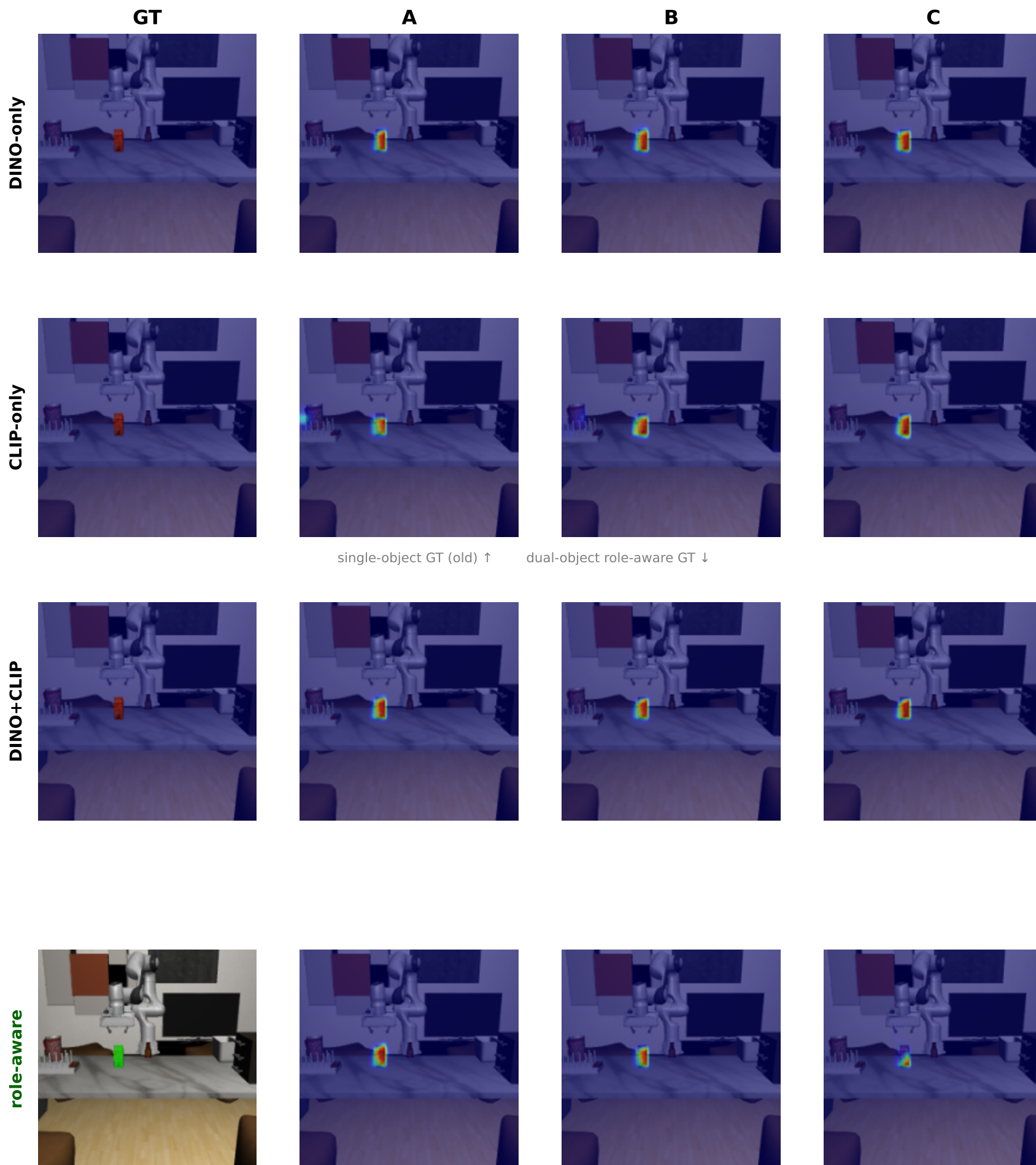
Mix_case8 — step: Pour out the milk from the milk container [manipulate]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to



Mix_case8 — step: Place down the milk onto coffee table [release]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to



Mix_case8 — step: Open the wooden cabinet top region [manipulate]

goal: After placing the butter and mug on the rack, pour in the milk, and check the wooden cabinet to

