

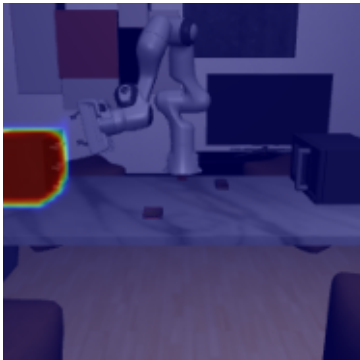
Memory_Exploration_case9 — step: Open the wooden cabinet top region. [manipulate]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put

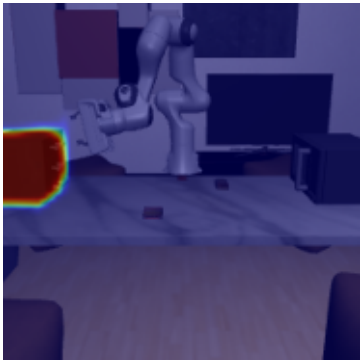
DINO-only



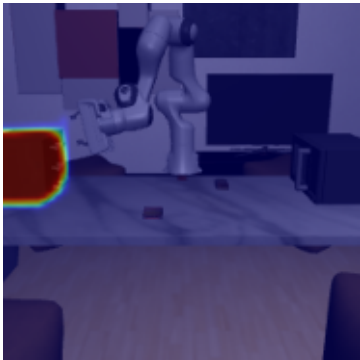
A



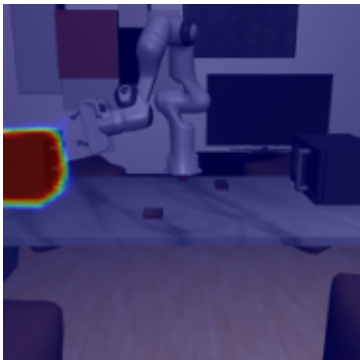
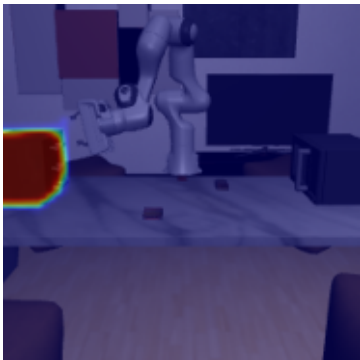
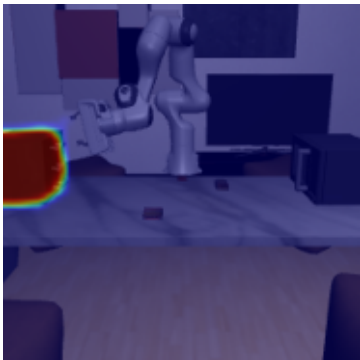
B



C

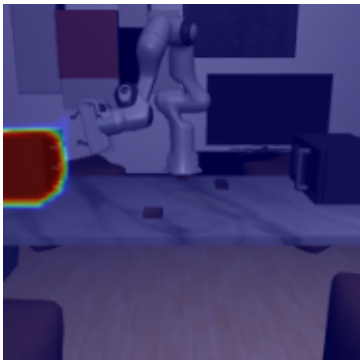
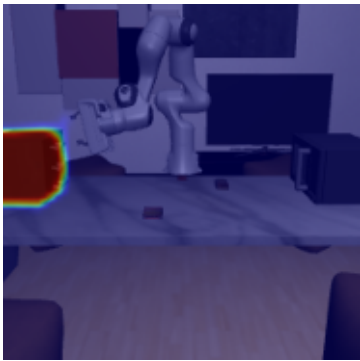
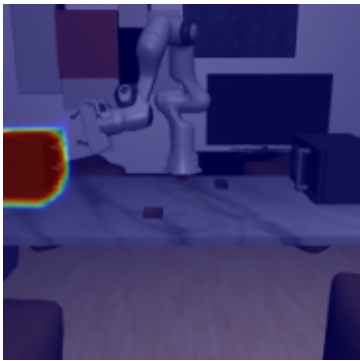


CLIP-only

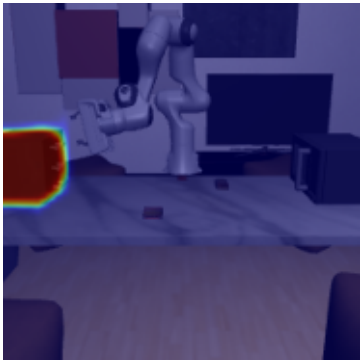
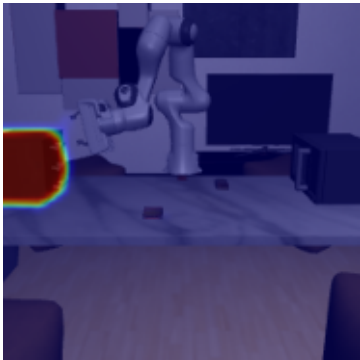
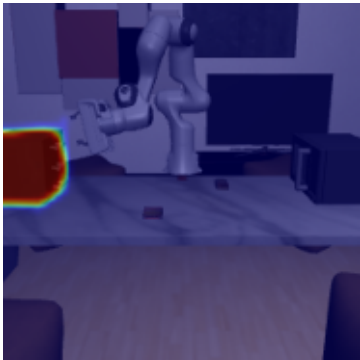


single-object GT (old) ↑ dual-object role-aware GT ↓

DINO+CLIP

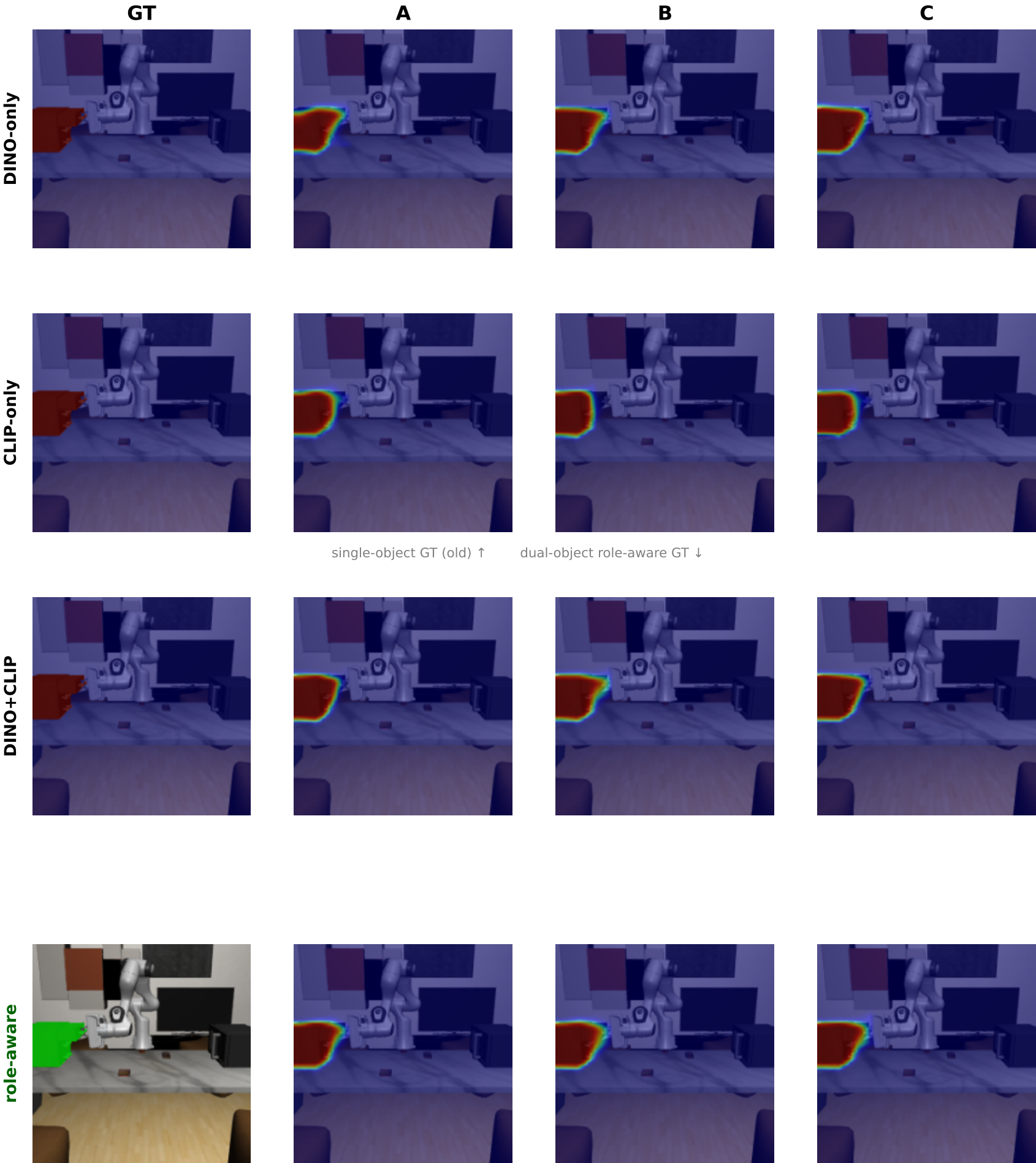


role-aware



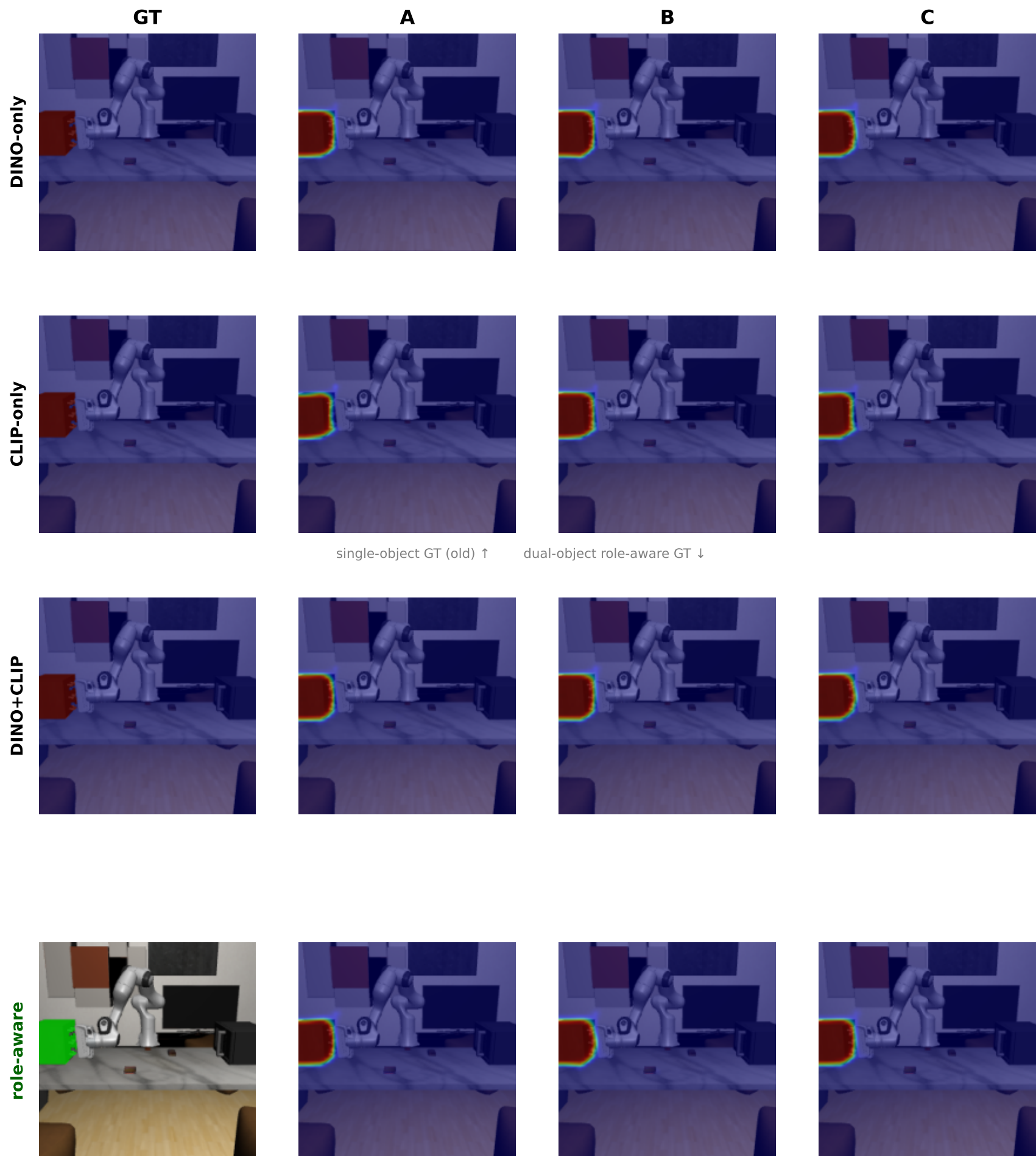
Memory_Exploration_case9 — step: Close the wooden cabinet top region. [manipulate]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



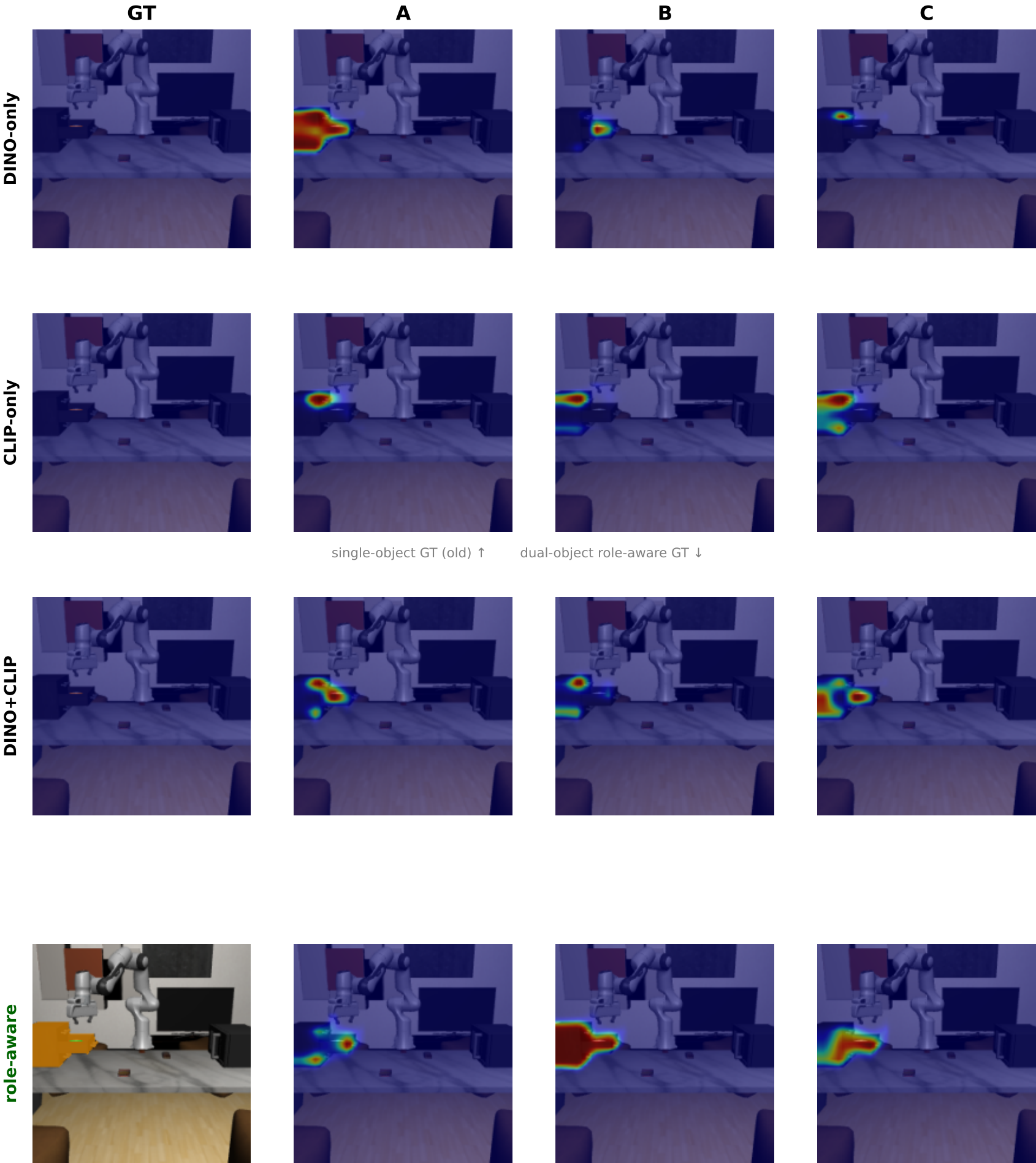
Memory_Exploration_case9 — step: Open the wooden cabinet middle region. [manipulate]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



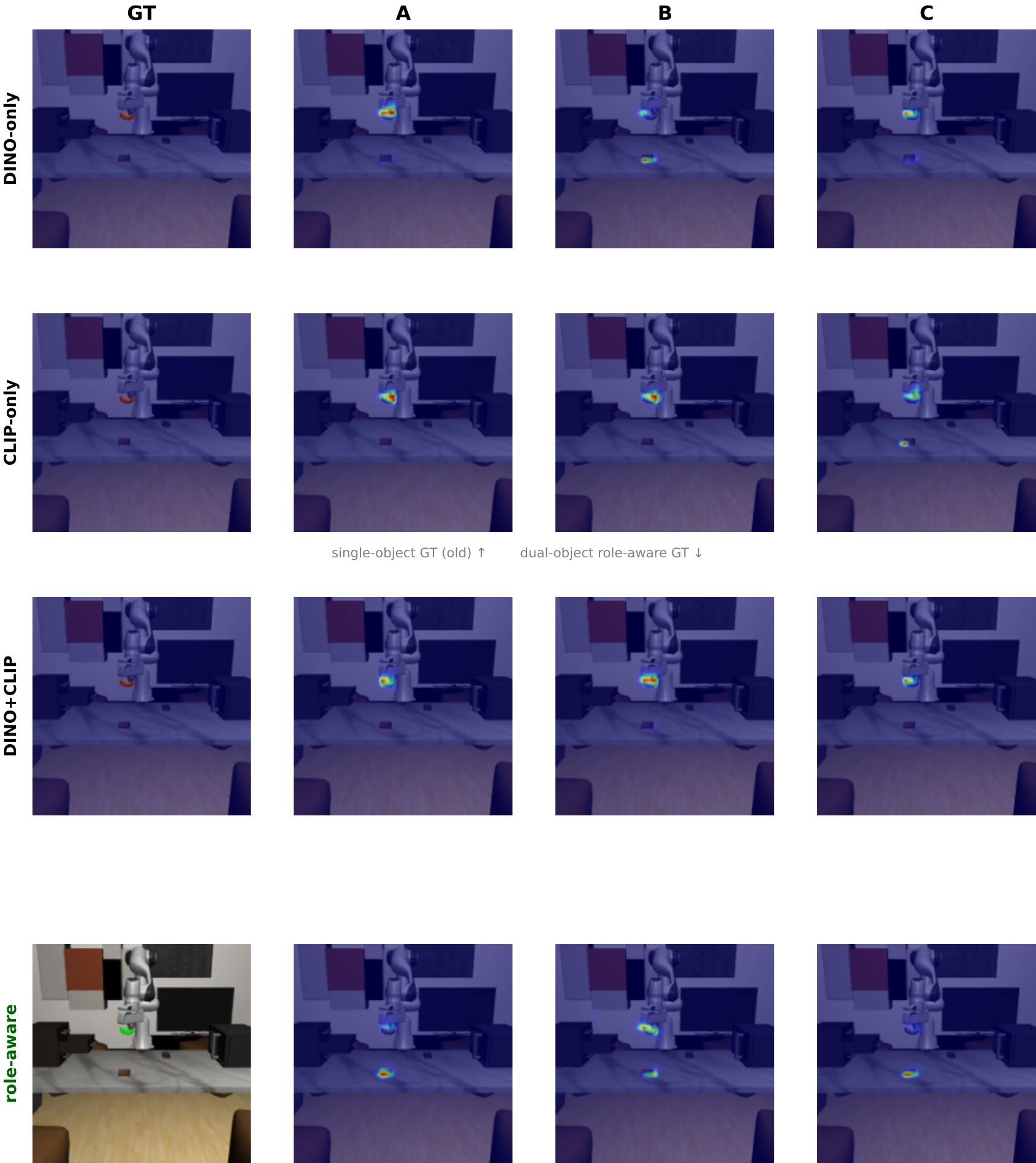
Memory_Exploration_case9 — step: Pick up the white bowl from the wooden cabinet [grasp]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



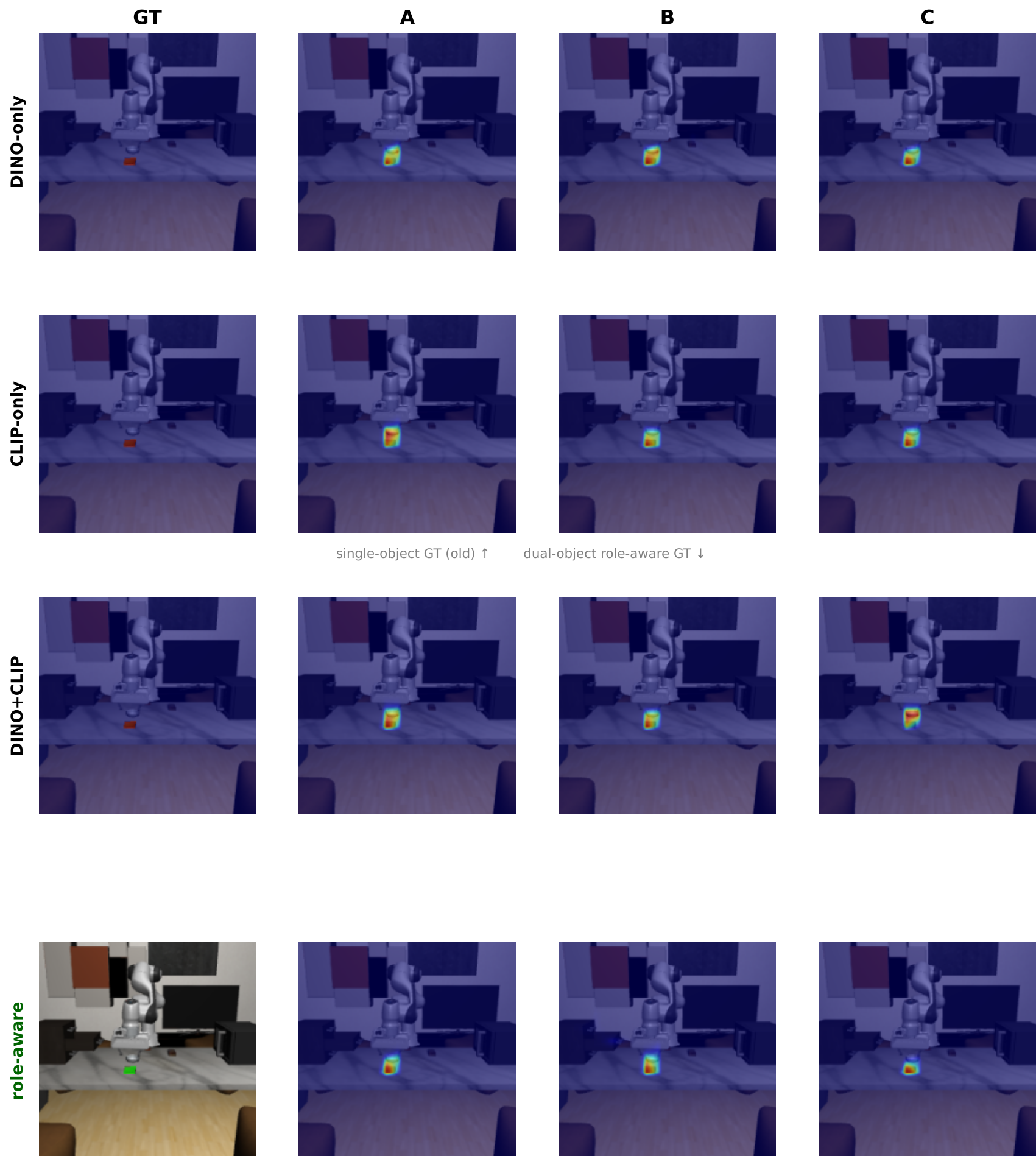
Memory_Exploration_case9 — step: Place down the white bowl on the coffee table [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



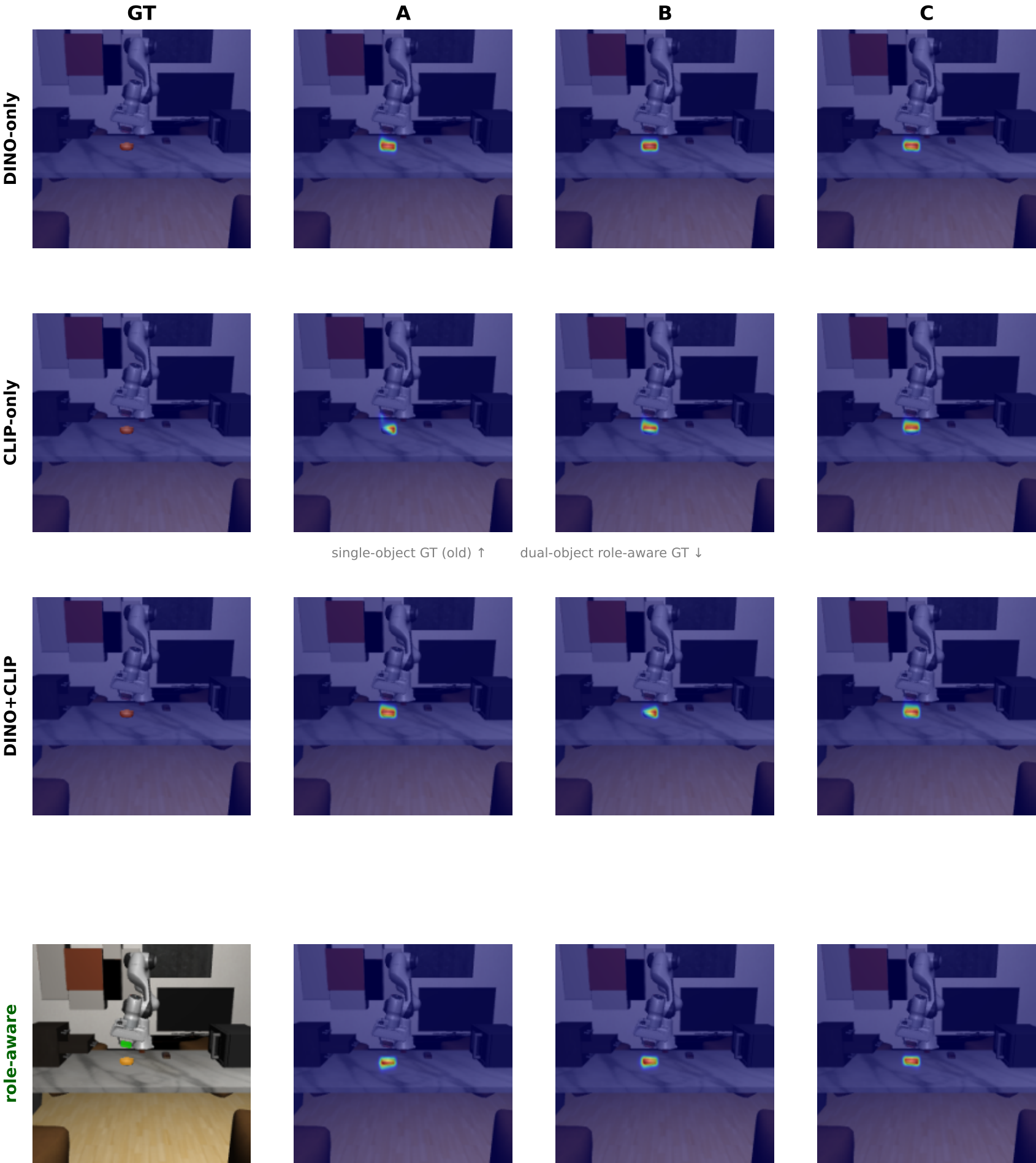
Memory_Exploration_case9 — step: Pick up the cookies from the coffee table [grasp]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



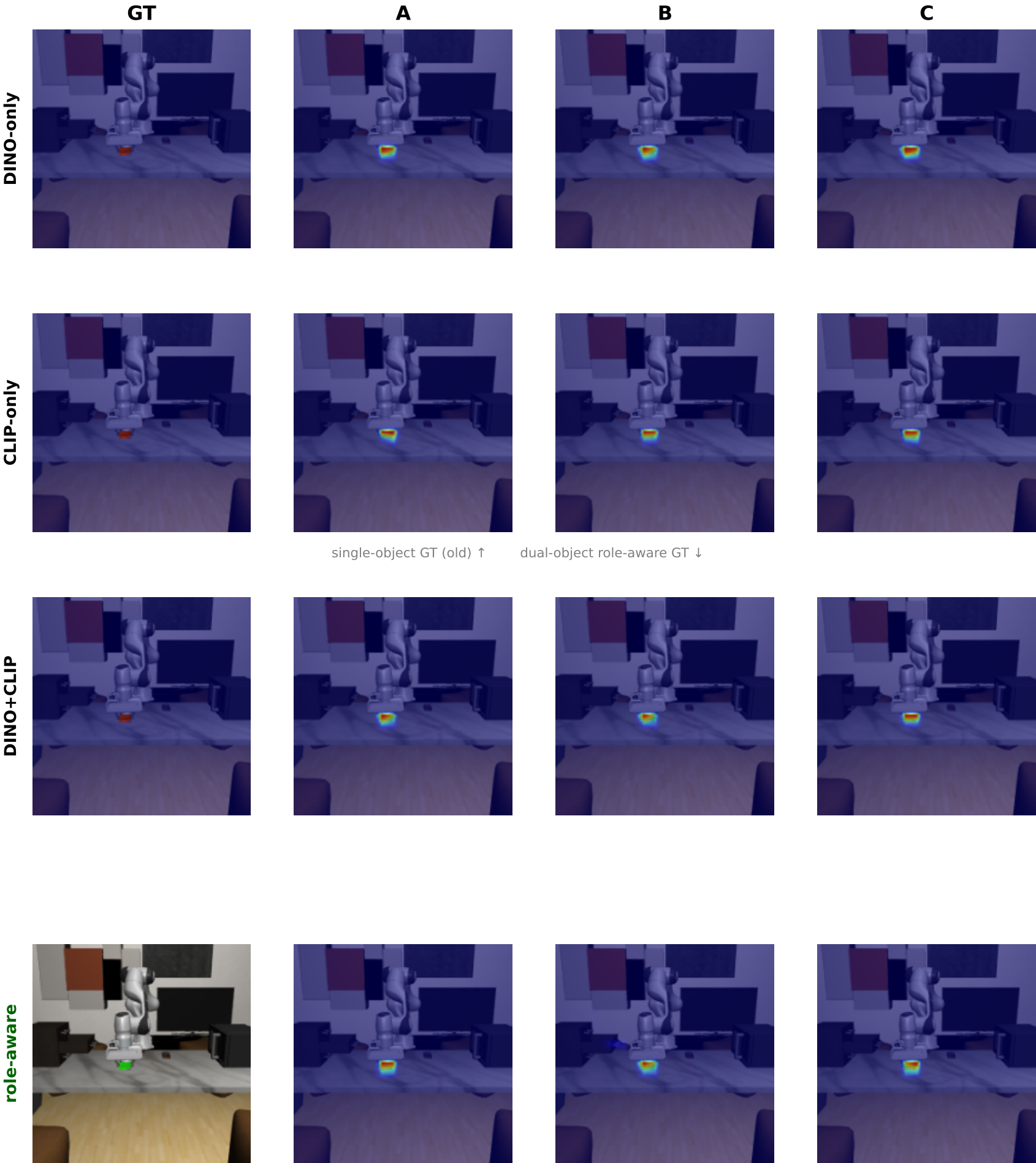
Memory_Exploration_case9 — step: Place down the cookies into the white bowl [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



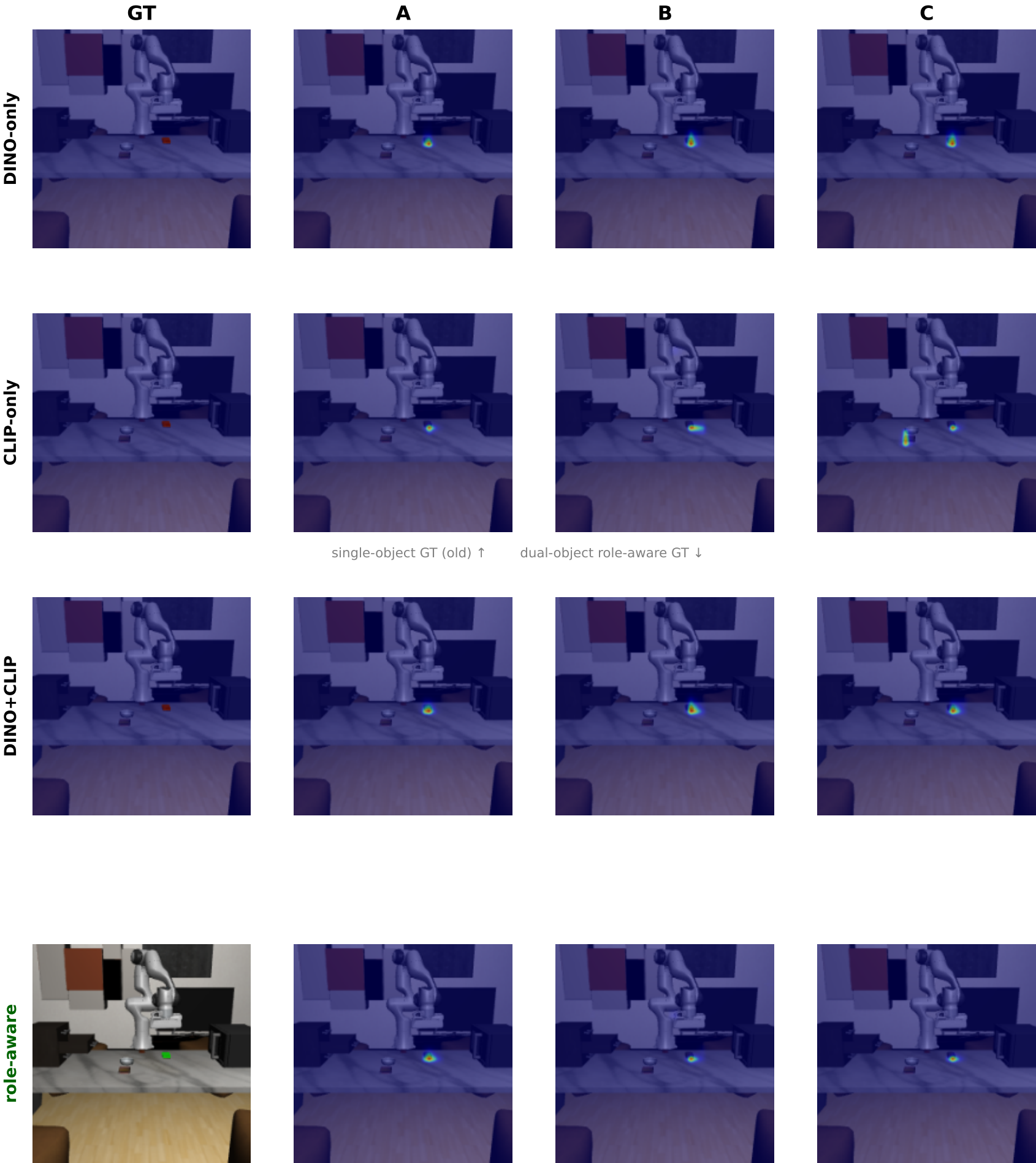
Memory_Exploration_case9 — step: Place the cookies on the coffee table [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



Memory_Exploration_case9 — step: Pick up the chocolate pudding from the coffee table [grasp]

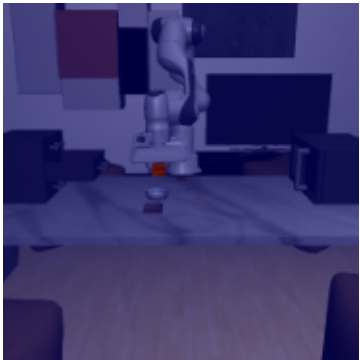
goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



Memory_Exploration_case9 — step: Place down the chocolate pudding into the white bowl [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put

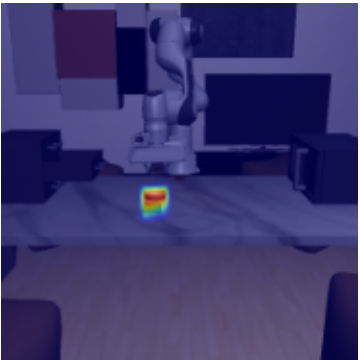
GT



A



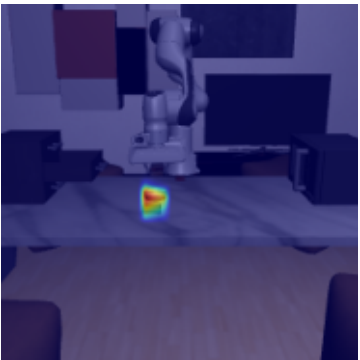
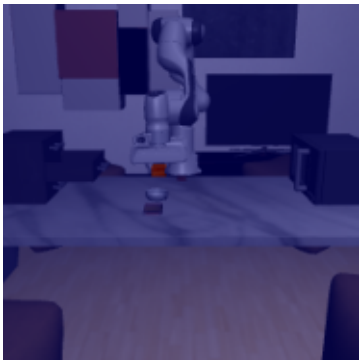
B



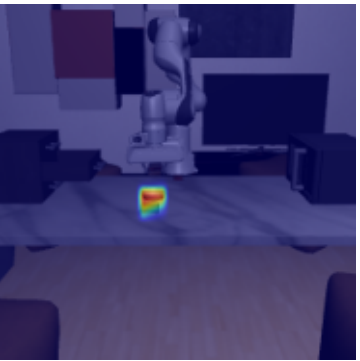
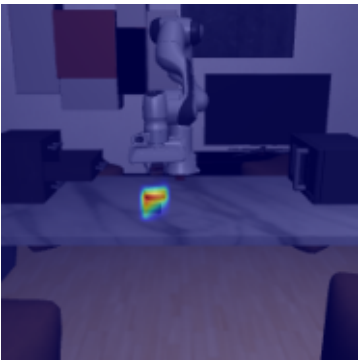
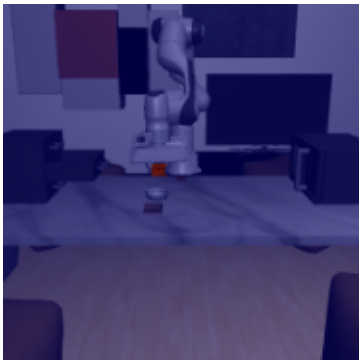
C



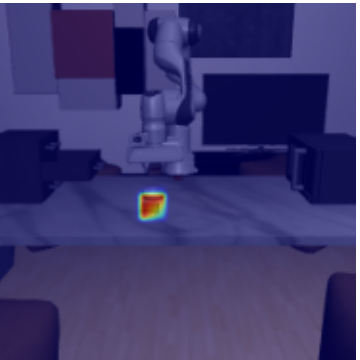
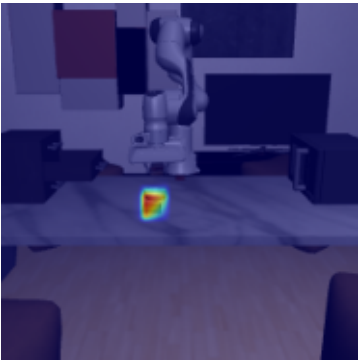
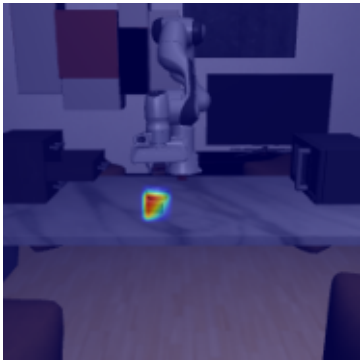
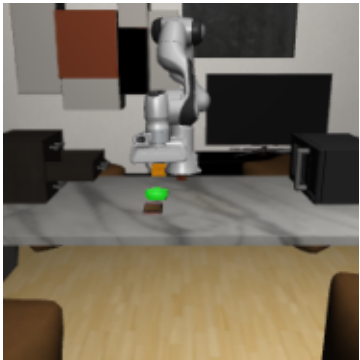
DINO-only



CLIP-only



DINO+CLIP

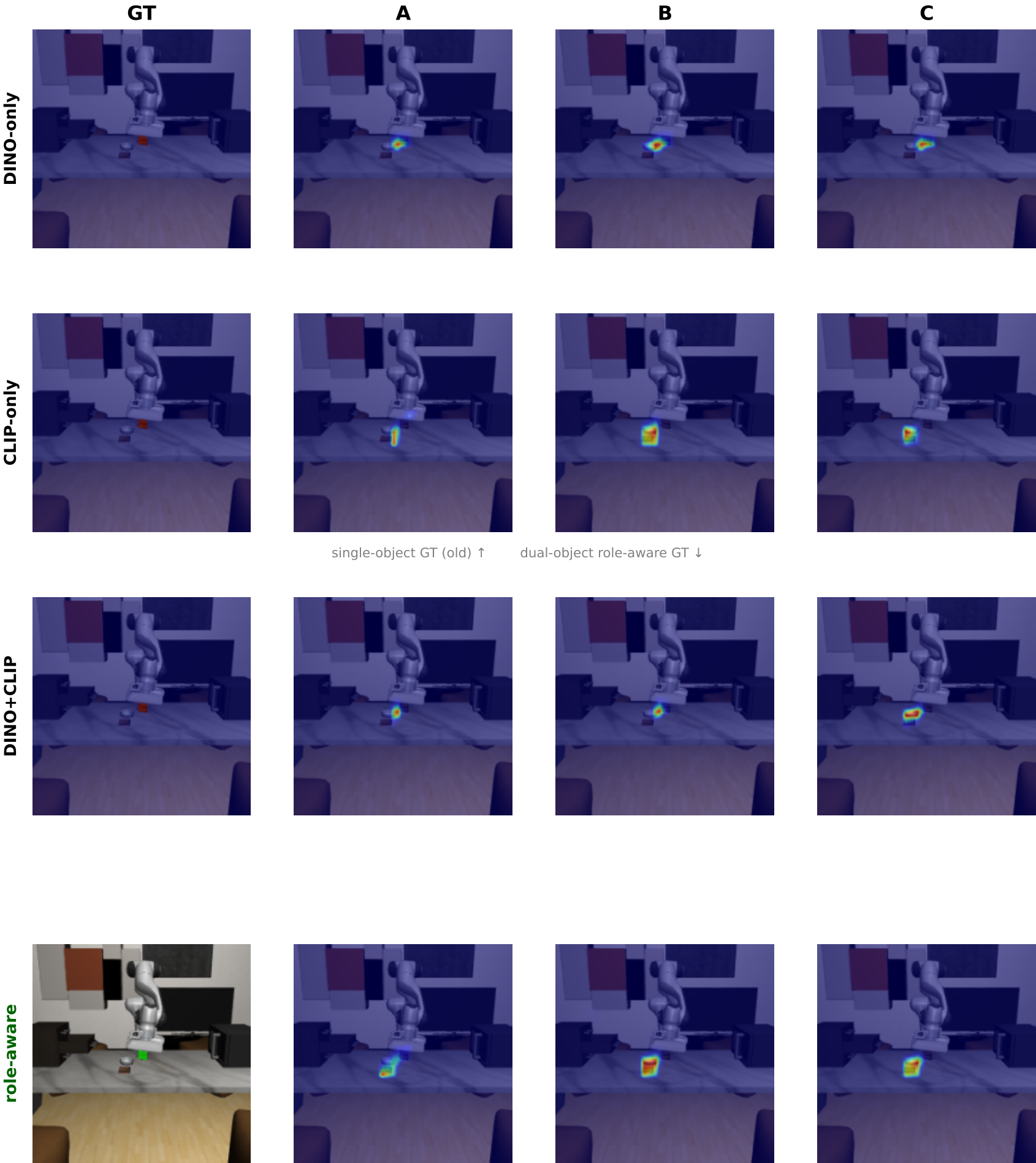


role-aware

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

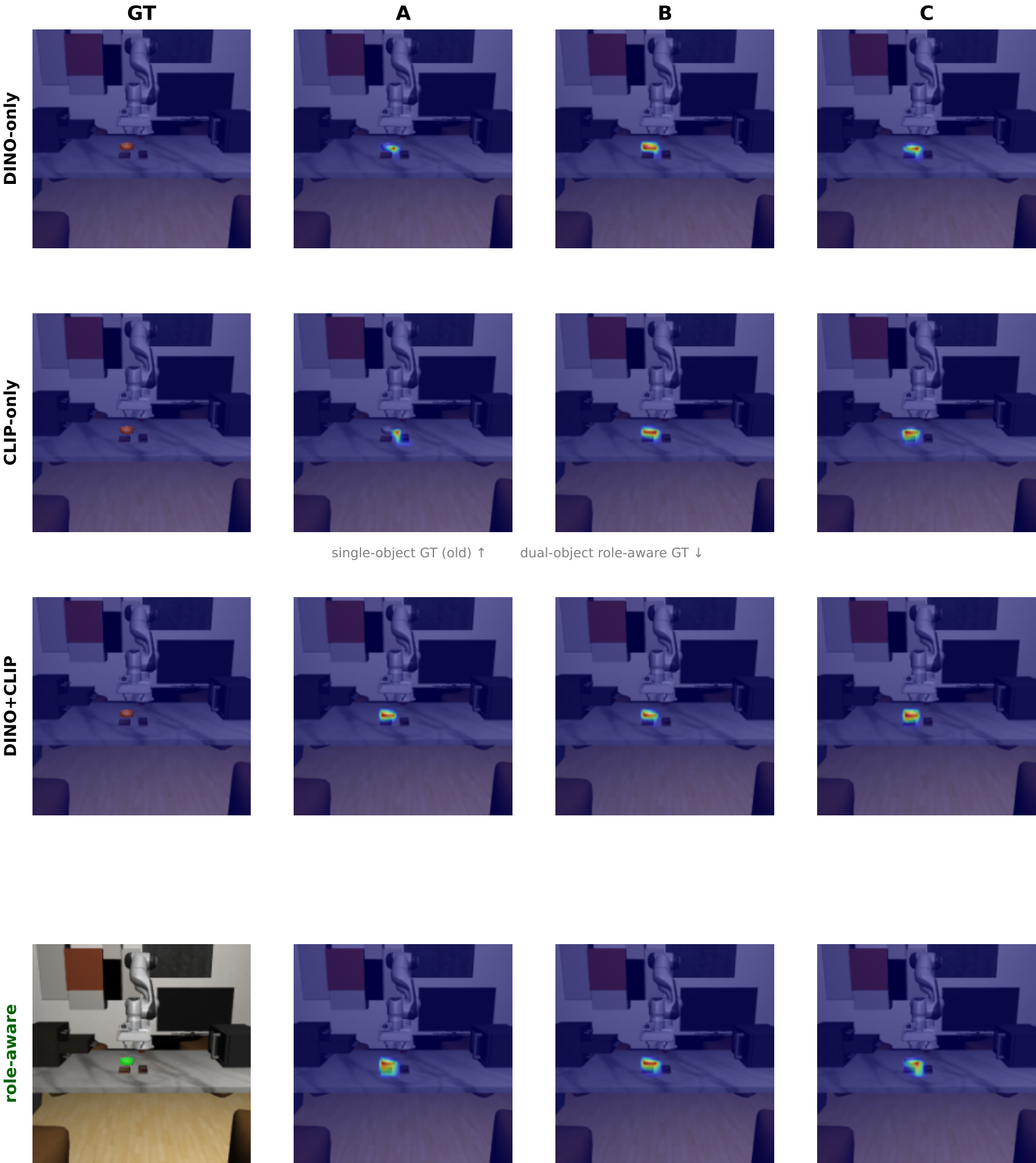
Memory_Exploration_case9 — step: Place the chocolate pudding on the coffee table [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



Memory_Exploration_case9 — step: Pick up the white bowl from the coffee table [grasp]

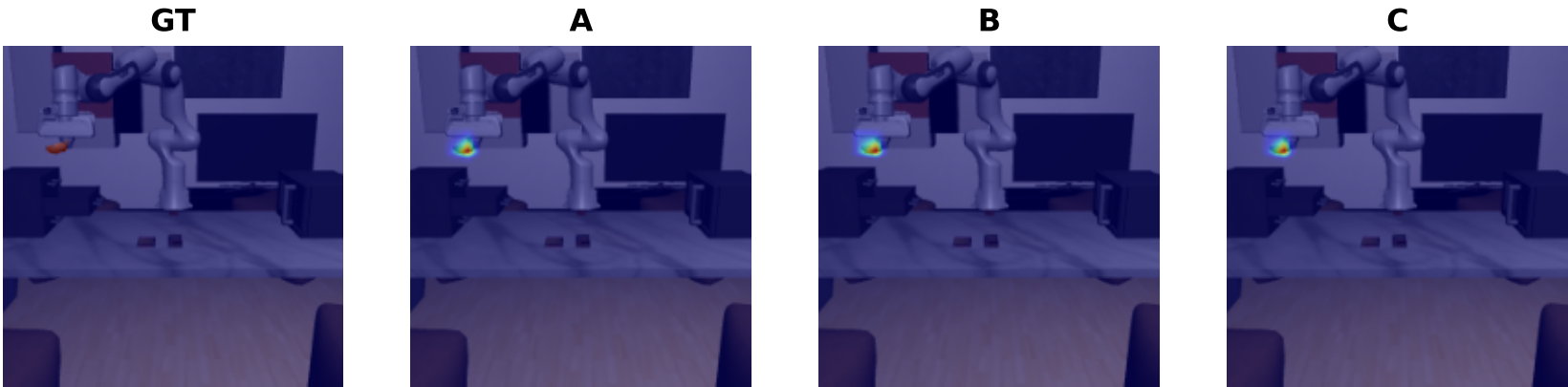
goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put



Memory_Exploration_case9 — step: Place the white bowl into the wooden cabinet at top side [release]

goal: Take the white bowl from the wooden cabinet, pour in the chocolate pudding and cookies, and put

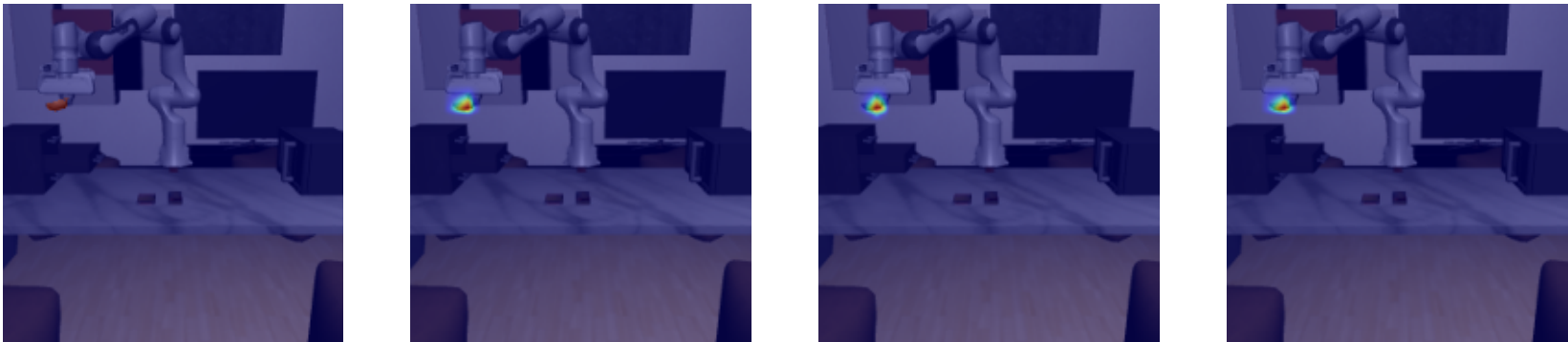
DINO-only



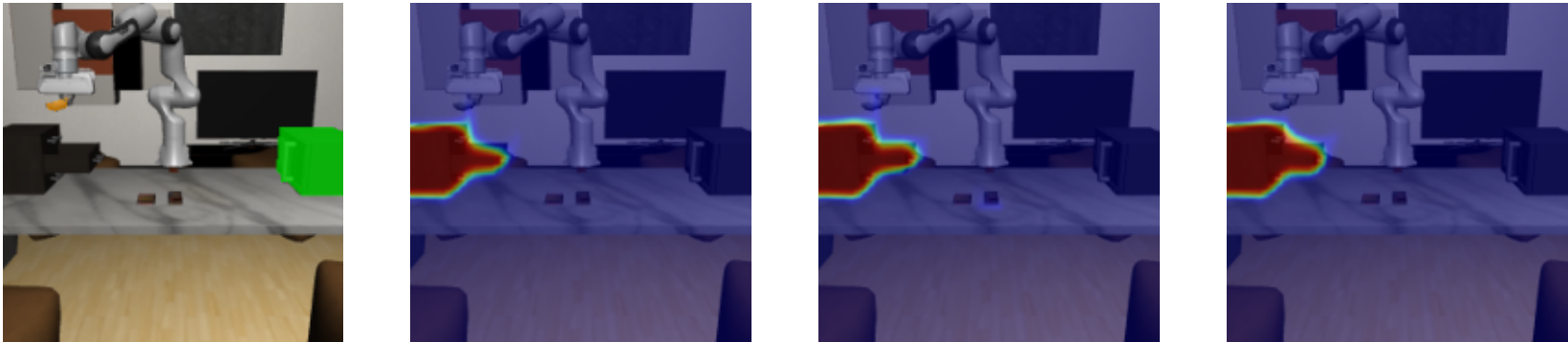
CLIP-only



DINO+CLIP



role-aware



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