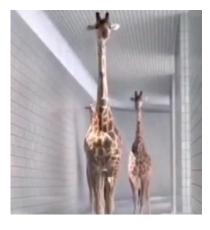|	action	location	view

HIVE-c

instruction: let the yellow cat step close to the toy on the floor
score: 0
reason: The yellow cat is completely missing in the edited image, and there is no evidence of the intended action of the yellow cat stepping closer to the toy. The scene does not reflect any attempt to implement the specified action.

instruction: move these giraffes ahead
score: 0
reason: The images are completely different, with no evidence of object relocation as instructed. The second image does not show the same scene or objects as the first image.
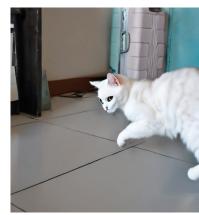
instruction: shift the view to the right until the two riders on the left come into view
score: 0
reason: No proper viewpoint transformation was detected. The images appear to show different scenes entirely, with different riders and horses, and the landscape has changed significantly.

CosXL-Edit

instruction: make the cat in from jump
score: 7
reason: The edited image successfully shows the cat in a jumping position, with its body elevated and legs extended, which is consistent with a jumping action. The cat's posture appears natural and physically plausible. However, there is a lack of motion blur that would typically accompany such a dynamic movement, and the shadows on the floor do not fully reflect the cat's new position, which slightly affects the realism. Overall, the transformation is well-executed but could benefit from more attention to motion blur and shadow interaction.

instruction: let the two sharks swim away
score: 0
reason: The two images are completely different, with no evidence of object relocation or any attempt to move the specified sharks as instructed. The scenes, lighting, and objects are entirely distinct, indicating no continuity or integration between the two images.

instruction: shift the view upward until no people can be seen
score: 0
reason: No proper viewpoint transformation was detected. The two images appear to depict entirely different scenes, with the first showing a bridge and cyclists, and the second showing a cityscape with a person standing on a road. The intended viewpoint change was to shift upward until no people can be seen, but the edited image does not reflect this transformation.

UltraEdit

instruction: have the man bend down
score: 9
reason: The man is clearly bending down in the edited image compared to the original. The action appears natural and physically plausible, with appropriate body posture. The interaction with the surroundings, such as shadows and contact points, is maintained well. There is no noticeable motion blur, but the action is subtle enough that this does not detract significantly from the realism.

instruction: make two fish a little lower
score: 0
reason: The images do not show evidence of the same fish being moved lower. Instead, the second image features different fish entirely, with no attempt to move the specified object as instructed.
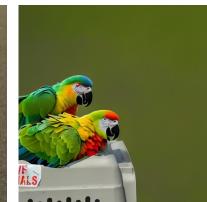
instruction: shift the view upward until no people can be seen
score: 0
reason: No proper viewpoint transformation was detected. The images appear to be entirely different scenes, with no evident change in viewpoint from the original to the edited version.

HIVE-w

instruction: let the parrots spread their wings
score: 0
reason: The images show no evidence of the parrots spreading their wings. Both images depict the parrots in a similar position with their wings folded, and there is no visible attempt to implement the specified action.
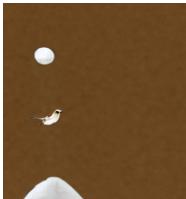
instruction: move the herds of horses on the roads forward
score: 3
reason: The herd of horses has been moved forward, but the integration is poor. The road has been replaced with grass, which is inconsistent with the original scene. The lighting and perspective are not well maintained, and the background consistency is disrupted by the change in the road surface.

instruction: shift the view to the left until only one side of the bird is visible
score: 0
reason: No proper viewpoint transformation was detected. The two images appear to show different scenes entirely, with different birds and backgrounds.