InstructPix2pix





instruction: make the hind leg of the brown cat step on the air conditioning score: 0

reason: The images are completely different, with no evidence of the specified action being implemented. The second image does not show the brown cat stepping on the air conditioning unit, nor does it maintain any visual coherence with the first image.







instruction: make the person feed the bear score: 8

reason: The edited image successfully depicts the person feeding the bear, with the bear holding a bowl and the person extending a hand with food. The action appears natural and physically plausible, with the bear's posture and the person's gesture consistent with feeding. The visual coherence with the scene is maintained, and the interaction between the person and the bear is believable. However, there is a slight lack of motion blur or deformation that might be expected with such an action, and the shadows and reflections could be more pronounced to enhance realism.

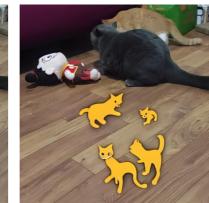




instruction: have the dog step into the room score: 3

reason: The edited image shows the dog in a different position, suggesting an attempt to depict the dog stepping into the room. However, the dog's body appears distorted and unnatural, particularly around the head and legs, which affects the realism of the movement. The shadows and contact points with the floor are not consistent with the new position, and there is no motion blur to indicate movement. The overall transformation lacks visual coherence with the scene.

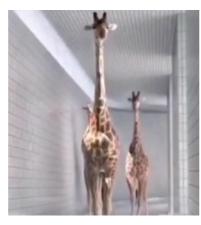


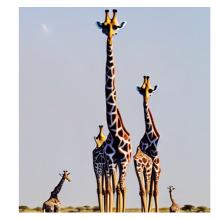


instruction: let the yellow cat step close to the toy on the floor

score: 0

reason: The edited image does not show any realistic movement of the yellow cat towards the toy. Instead, it features cartoonish yellow cat figures added to the scene, which do not represent the intended action of the real yellow cat stepping closer to the toy. There is no evidence of the described action being implemented in a realistic manner.





instruction: move these giraffes ahead score: 0

reason: The images are completely different, with no evidence of object relocation as instructed. The first image shows giraffes in an indoor setting, while the second image shows giraffes in an outdoor setting, indicating no attempt to move the specified object within the same scene.





instruction: remove the objects in the sky score: 8

reason: The objects in the sky have been successfully removed, and the background consistency is maintained well. However, there is a slight visible artifact where the objects were removed, which slightly affects the integration quality.





instruction: move the e-bike on the left to the middle score: 0

reason: The images are identical, with no evidence of the e-bike being moved from the left to the middle.





instruction: shift the view forward score: 0

reason: No proper viewpoint transformation was detected as the two images appear to depict entirely different scenes.





instruction: shift the view to the left score: 0

reason: No proper viewpoint transformation was detected as the images depict entirely different scenes.





instruction: shift the view upward until the road no longer be seen

score: 0

reason: No proper viewpoint transformation was detected. The two images appear to be different scenes entirely, as the road is still visible in the edited version, contrary to the instruction.

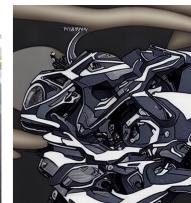




instruction: make the man pushing the bicycle closer to another bicycle a little score: 0

reason: The images do not show any evidence of the man pushing the bicycle being moved closer to another bicycle. Instead, a new person on a bicycle has been added to the scene, which does not fulfill the instruction given.





instruction: shift the view backward score: 0

reason: No proper viewpoint transformation was detected. The images appear to show different scenes entirely, with the first image depicting a road under an overpass and the second image showing a stylized motorcycle design.