

Image Evaluation Guidelines - Critical

Goal

We are generating images based on existing product images in order to create more visually appealing advertisements (compared to the product on a plain white background). To do this, we need to ensure that:

- All generated images accurately represent the product including its appearance, functionality, and where/how it is used.
- All images include scenes that make sense.
 - For real-world scenes, this includes ensuring that the entire scene looks plausible, with all physical interactions between objects (including gravity) and lighting/shadows/reflections appearing normal. These scenes not contain obvious "product placement" settings, such as an unnecessary pedestal or pole on which the product sits.
 - For more abstract scenes (fantasy, studio shoot, etc.) there is more room for creative freedom, including objects floating, being balanced on edge, or presented on a pedestal, but the scene should still look cohesive and not feel like parts of it are "pasted on top".
- All humans and animals appear natural and "normal". Historically, incorrect numbers of fingers or limbs, and contorted or disfigured faces have been a trait of bad AI generated images, so viewers will be quick to pick up on that and make that association. Even if they do not make that direct association, it gives the image an uncomfortable/uneasy feeling that is not suitable for advertising purposes.

Guidelines

1) Is the product perfectly represented in the enhanced image?

This means no critical parts are added to or removed from the main product, any packaging or labelling is correct, it is shown at a reasonable scale/size relative to the rest of the scene, and if it's being used it is being used correctly. It is fine for additional elements to be present in the original image, such as accessories or attachments, views from additional angles, or color swatches to be removed so long as the main product is still present and depicted in a useful way. The product should be prominent in the enhanced image, meaning it is large enough and high enough contrast that it is easily recognizable, and is not obscured by other elements within the scene, or the overlaid text/logos in the rendered ad. This should be validated by referring back to the original product image and comparing to the enhanced image to ensure the product looks identical, and then confirming in the rendered ad that the product is not covered by the text/logo elements.

Key focus: We want images where the product is the focal point of the image, and is accurately represented both visually and in how it is used (if it's being used in the image).

- **Yes** - These images show the product clearly and accurately, with no alterations to its composition or functionality. They are also shown at a plausible scale with the rest of the scene, large enough to be easy to find and not obscured by other elements of the ad. In situations where the original product image contains multiple items, it is permissible for the scene to only contain some (or even one) of them, so long as it contains the main product.
- **No** - These images show the product with parts added or removed, bent or otherwise distorted, being used incorrectly or implying non-existent functionality, or show the product at an incorrect size relative to other objects in the scene.
- **Not Sure** - These should be a rare case. These images do not contain any obvious issues, but do have some elements that are difficult to verify, and therefore cannot safely be marked as "Yes", and should be manually double-checked.

Good examples ("Yes")



This image is still considered "Yes" despite views/packaging removed from Original image:



This image is a "Yes" because it is not obscured by the text/logos

Bad examples ("No")

One headphone (key part of the product) removed



The blender has an added base



The washer has an added nozzle on the back



The product is obscured by the text/logos in the ad



2) Is the scene in the enhanced image plausible and free from defects?

This means it should not contain any unrealistic scenarios (including floating products/objects in otherwise realistic scenes), animated/cartoon elements in an otherwise photorealistic scene, physically impossible situations, impossible shadows/reflections, or similar/duplicative products (there should not be any objects that are hard to tell if they are or are not the product). The product should not be placed precariously on a pole or pedestal. Products should not be balanced or stood up implausibly on their own except when necessary to show the normal context/usage of the product.

Key focus: We want images that look like realistic, cohesive scenes. Everything should appear "normal", and every element of the scene should feel like a part of the scene rather than something pasted/layered on top of it.

- **Yes** - These images show plausible scenes that are a uniform style. All objects in the scene are stylistically similar (realistic, animated, sketched, etc.). In real-world scenes, everything is grounded and placed in physically appropriate ways, while in fantasy/studio shoot scenes floating products can be accepted. No objects appear to be passing through other objects, and there are no objects in the scene that are similar to or can be confused with the product (such as showing chocolate chip cookies in an Oreo ad). These images also have the product placed naturally, rather than situating it on an otherwise unnecessary pedestal or pole. Products that are incapable of standing on their own and do not need to do so to show their normal usage or context (such as watches), should not be depicted doing so. Products that can plausibly stand on their own or need to stand in order to show normal context or usage, such as upright vacuums or bicycles, may do so even if they typically require human intervention to stand.
- **No** - These images may contain defects that break the continuity of the scene and make it look "impossible" or "fake". They also may contain elements that compete with or can be confused with the product, or may position the product unnaturally on a pole or pedestal.
- **Not Sure** - These should be a rare case. These images do not contain any obvious issues, but do have some elements that are difficult to verify, and therefore cannot safely be marked as "Yes", and should be manually double-checked.

Good examples ("Yes")



Bad examples ("No")

Floating product in real-world scene



Floating product + table in lake



Scale issues (9oz can & cup)



Object standing on it's own that shouldn't be



Scene includes dog treats that are not the same as the depicted product



3) If there are humans or animals present in the enhanced image, are they anatomically correct and in natural looking positions/poses?

Key focus: Any humans/animals included must look "normal" and naturally fit into the scene, without any abnormalities that will lead the viewer to assume the image is poorly AI-generated. (Common triggers for this being incorrect numbers of fingers, disformed facial features, etc.)

- **Yes** - All humans and animals in the image appear anatomically correct and are realistically posed. Any humans/animals will have proper facial composition, the correct number of limbs/digits, and will be in a natural pose (including being bent where joints are supposed to be). They will have realistic features/complexion. They will also be “complete”, meaning there are no hands not attached to arms/bodies save for reasonable cropping (such as a hand protruding from the edge of the image or sticking out of a doorway). Note that this includes any images where humans or animals were present on the original image.
- **No** - These images contain humans or animals, and at least one of them has a defect that makes it look unnatural, such as extra fingers, elbows bent the wrong direction, or weird/impossible poses. These images also may include severed/incomplete humans/animals, such as a hand not connected to an arm/body, or humans/animals with unnatural features/complexion, such as a green cat.
- **No Humans/Animals Present** - These images are neither a yes or a no because they don't contain any humans/animals at all.
- **Not Sure** - These should be a rare case. These images do not contain any obvious issues, but do have some elements that are difficult to verify, and therefore cannot safely be marked as “Yes”, and should be manually double-checked.

Good examples ("Yes")



Bad examples ("No")

Cropped/partial human



6 legged dog



Disconnected human hand

