

Prompt for chatGPT to generate GPT caption

Generate a simple description of the sound from an image description and sound labels.

Here are the rules you should follow.

1. Consider the sound labels more than the image description because the labels represent the sound.
2. Remove the name of the place, city, or country.
3. Delete the time, device names, proper name modifiers, number modifiers, and unit modifiers.
4. Do not use expressions such as "the sound of ~" or "the audio of ~" because it is obvious.
5. Generate 3 descriptions.
6. Each description will always consist of one sentence.
7. Eliminate information that can only be obtained from visual information contained in the image description, such as the color and shape of clothing and objects.

Image description : **image discription**

Sound label : **sound labels**