

Do Do

Which of these cars is more aerodynamic? The one on the left or the right? Explain why, using specific visual details.

 OmniLMM-12B

The car on the right is more aerodynamic. This can be inferred from the shape of the car on the right, which appears to have a more streamlined design with a pointed front, compared to the car on the left, which has a more boxy and less aerodynamic shape.



Do Do

Do you think this image is funny, explaining the reason to me?

 OmniLMM-12B

Yes, I believe this image is funny because it shows a dog wearing clothes that humans typically wear, such as a blue shirt and black shorts. The dog appears to be imitating a person by taking a selfie in front of a mirror, creating a humorous contrast between the dog's behavior and its attire.



Do Do

What is this photo about? Please answer in detail.

 OmniLMM-12B

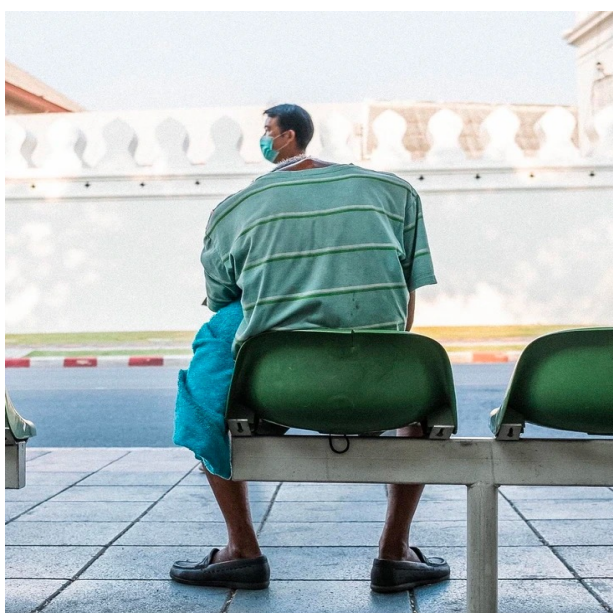
The image captures a serene winter scene at the **Forbidden City in Beijing**, China. The Forbidden City, a UNESCO World Heritage Site, stands majestically in the background, its red walls and roofs contrasting against the overcast sky. The ground is blanketed in a layer of snow, adding a touch of tranquility to the scene.

In the foreground, a group of people are seen walking towards the Forbidden City, their breath visible in the cold air. Among them, one individual is capturing the moment on their phone, perhaps to cherish the memory or share it with friends and family.

The image is taken from a low angle, looking up at the Forbidden City, emphasizing its grandeur and historical significance. The perspective also allows for a clear view of the people in the foreground, providing a sense of scale and depth to the scene.

**The image is watermarked with "CCTV 2" and "CCTV.COM"**, indicating that it was likely taken by China Central Television, a major state-owned television broadcaster in China. The watermark also suggests that the image may be part of a news report or documentary.

Overall, the image provides a glimpse into a day in the life at the Forbidden City, capturing the blend of history, culture, and everyday life in Beijing.



Do Do

How many people can you see in the picture? Please answer in detail.

 OmniLMM-12B

There are **two people in the picture**. One person is **sitting on a green bench**, and the other person is **walking in the background**.

 GPT-4V

In the picture, **there is one person visible**. They are seated on a bench facing away from the camera, wearing a mask, suggesting that the photo may have been taken during a time when mask-wearing was recommended or required, possibly due to health regulations. The person is wearing a green striped shirt and dark pants, and a blue towel is draped over one side of the bench, which might suggest that they are in a waiting area, possibly at a sports facility, pool, or similar location. There are additional empty green seats next to the person, and **no other individuals are seen in the frame**. The background is out of focus but appears to be an outdoor setting with a white structure and red and white barriers, indicating some form of boundary or restricted area.