

What are image inputs and how do they work in ChatGPT?

ChatGPT now has image capabilities to understand and interpret images you add to conversations as image inputs.

How should I use image inputs in conversations?

Basic Use: Upload a photo to start. Ask about objects in images, analyze documents, or explore visual content. Add more images in later turns to deepen or shift the discussion. Return anytime with new photos.

Annotating Images: To draw attention to specific areas, consider using a photo edit markup tool on your image before uploading. This guides ChatGPT to focus on elements you deem important.

Which plans can use image inputs?

Plus and ChatGPT Enterprise.

Which models can accept image inputs?

Which platforms are image inputs available on?

All platforms including web (chat.openai.com) and mobile (iOS / Android).

Are my images used to improve your models?

Our approach to using content, including images, remains the same for each product.

Please refer to [How your data is used to improve model performance](#) to better understand how content on ChatGPT may be used to improve model performance and the choices that users have.

For [ChatGPT Enterprise](#), we do not use content to train our models.

How do I add image inputs in ChatGPT?

Ensure the [model selector](#) is set to GPT-4 then tap the + icon in the prompt area to add image inputs.

Do the image inputs support videos?

No it can not handle videos. It currently supports processing static images only.

What files types are supported?

[PNG](#) (.png), [JPEG](#) (.jpeg and .jpg), and non-animated [GIF](#) (.gif).

How many images can I upload at once?

The number of images you can add to a conversation depends on various factors, including the size of the images and the amount of text accompanying them. As a general guideline, if you encounter issues, consider reducing the image quantity or size.

What is the size limit per image?

How do the image capabilities handle ambiguous or unclear images?

If an image is ambiguous or unclear, the model will do its best to interpret it. However, the results may be less accurate.

What limitations should users be aware of when using ChatGPT with Image Inputs?

If you're using ChatGPT's new image input feature, it's important to be aware of these limitations:

1. **Medical:** The model is not suitable for interpreting specialized medical images like CT scans and shouldn't be used for medical advice.
2. **Non-English:** The model does not perform as well handling images with text of non-Latin alphabets, such as Japanese or Korean.
3. **Big text:** Enlarge text within the image to improve readability, but avoid cropping important details.
4. **Rotation:** The model may misinterpret rotated / upside-down text or images.
5. **Visual elements:** The model may struggle to understand graphs or text where colors or styles like solid, dashed, or dotted lines vary.

6. **Spatial:** The model struggles with tasks requiring precise spatial localization, such as identifying chess positions.
7. **Accuracy:** The model may generate incorrect descriptions or captions in certain scenarios.
8. **Shape:** The model struggles with panoramic and fisheye images.
9. **Metadata and resizing:** The model doesn't process original file names or metadata, and images are resized before analysis, affecting their original dimensions.
10. **Counting:** May give approximate counts for objects in images.

Related Articles

[What is ChatGPT?](#)

[ChatGPT — Release Notes](#)

[ChatGPT Android app - FAQ](#)

[Voice chat FAQ](#)

[File uploads FAQ](#)