MiniGPT 自動看圖說故事

如何開始使用MiniGPT

善用 AI 以圖生文的優勢

① 自動化和高效率

AI可以自動識別圖像中的物體、場景和概念，並迅速檢索相關資料。這比人工查找更快、更有效率。

② 多領域查詢

AI可以識別各種類型的圖像，包括顏色、形狀、紋理、背景資訊等，給予多個領域的資料。

③ 節省人力資源

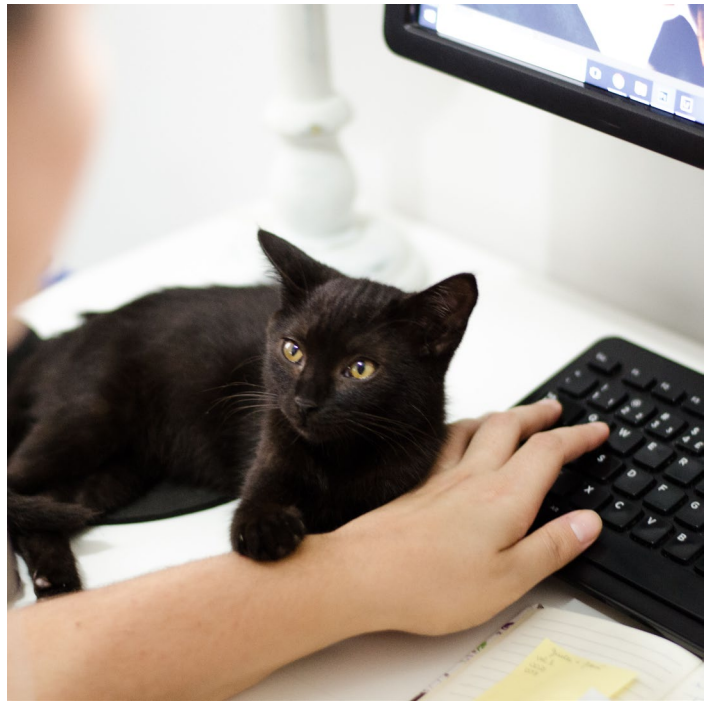
自動化圖像識別和相關資料檢索可以減少人工操作的需求。

④ 支持多語言檢索

透過 AI 圖片辨識，可跨語言檢索圖片的背景資訊。

⑤ 挖掘圖片資訊

透過 AI 進行大數據分析圖片中的相關資訊，進而取得圖片中的更多延伸資訊。



Google 搜尋 minigpt

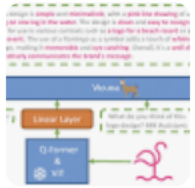
Google minigpt

登入

影片 購物 圖片 4 4 github 4 Demo 新聞 書籍 地圖 所有篩選器 工具 安全搜尋

約有 48,200 項結果 (搜尋時間：0.32 秒)

GitHub Pages
https://minigpt-4.github.io · 翻譯這個網頁



Minigpt-4

To examine this phenomenon, we present MiniGPT-4, which aligns a frozen visual encoder with a frozen LLM, Vicuna, using just one projection layer.

電器系女孩
https://tracyting.com › minigpt-4

【免費使用MiniGPT-4 】MiniGPT-4 是什麼? 用圖像與AI 對話

2023年7月30日 — 這是一款咖啡機，旨在製作咖啡和濃縮咖啡飲料。它有一個熱水箱和兩個沖泡杯，還有一個用於卡布奇諾和拿鐵的牛奶起泡器。它還有一個用於保持食物溫暖的熱板 ...

MiniGPT-4 是什麼？ · MiniGPT-4 的特色 · MiniGPT-4 的使用方法？

GitHub
https://github.com › Vision-CAIR › Min... · 翻譯這個網頁

選擇 Space 進入

MiniGPT-4:

Enhancing Vision-language Understanding with Advanced Large Language Models

Deyao Zhu* , Jun Chen*, Xiaoqian Shen, Xiang Li, Mohamed Elhoseiny

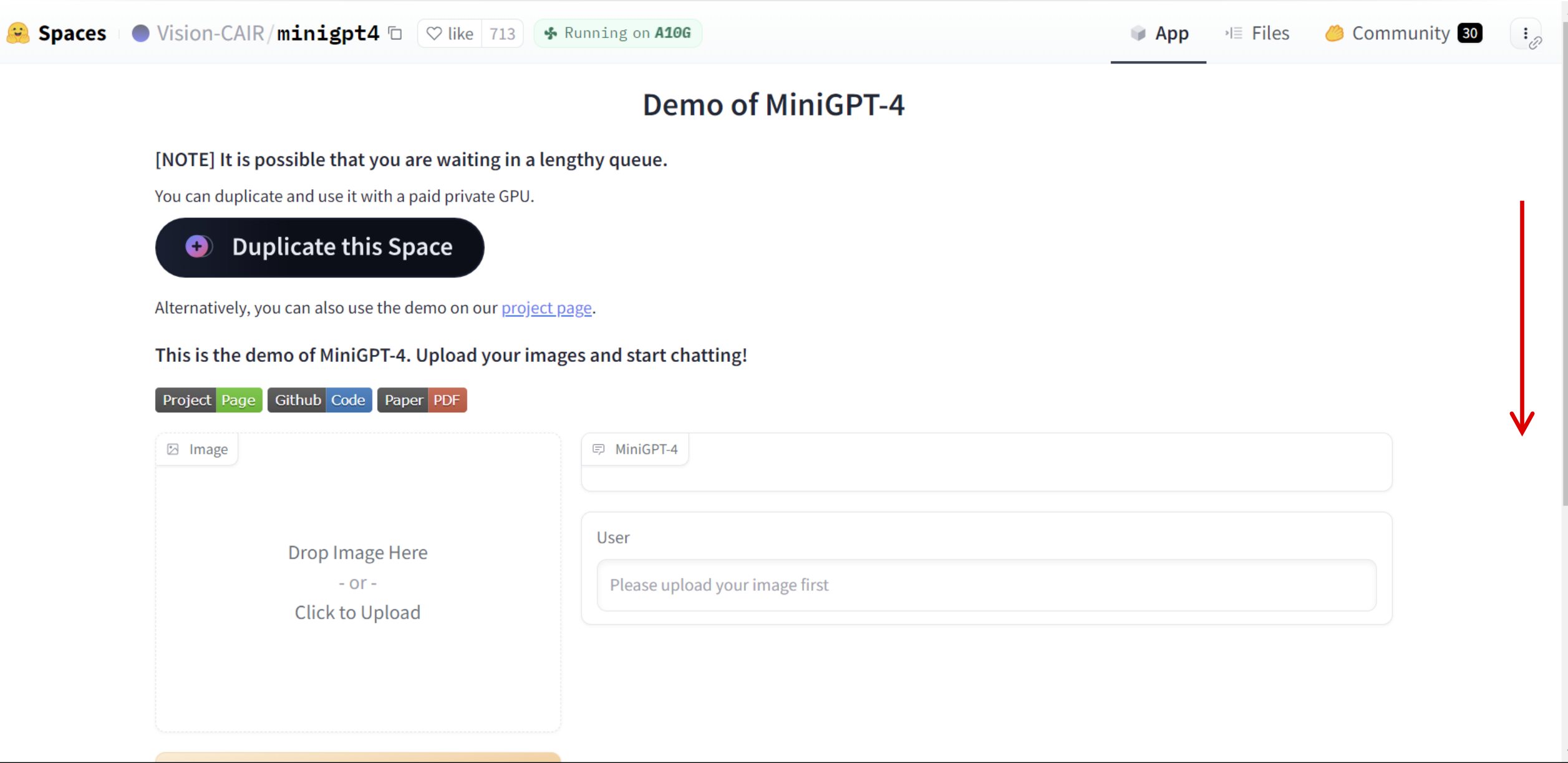
▶ King Abdullah University of Science and Technology

*Equal Contribution

Paper | Code | Space | Video | Dataset | Model

Abstract

來到已建立的 Demo 站台頁面（可以直接使用）



載入圖片的位置，可點擊選擇圖檔

Project Page | Github Code | Paper PDF

Image

Drop Image Here
- or -
Click to Upload

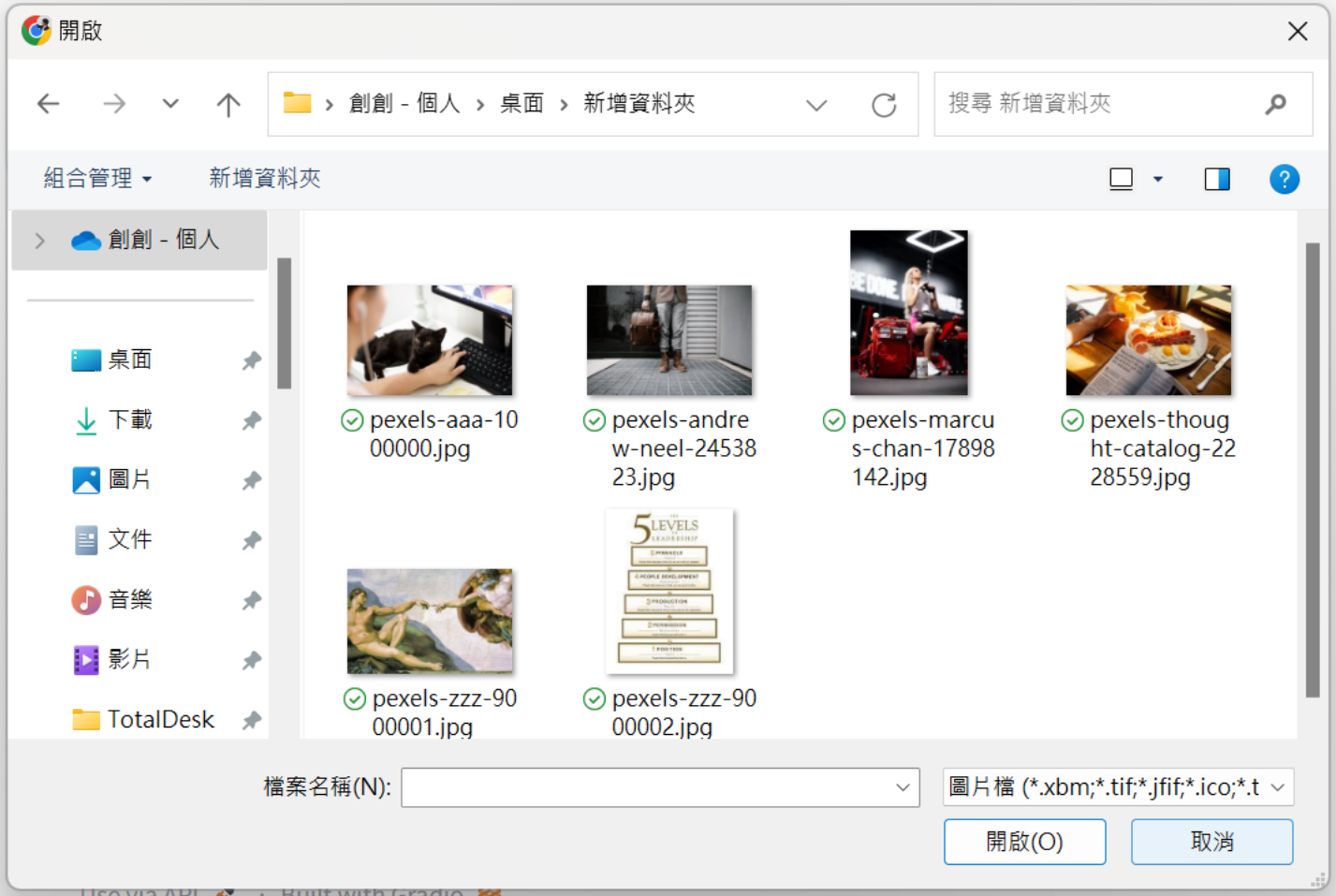
Upload & Start Chat

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4



Use via API · Built with Gradio

載入圖片的位置，可拖曳選擇圖檔

Project Page | Github Code | Paper PDF

Image

Drop Image Here
- or -
Click to Upload

Upload & Start Chat

Restart

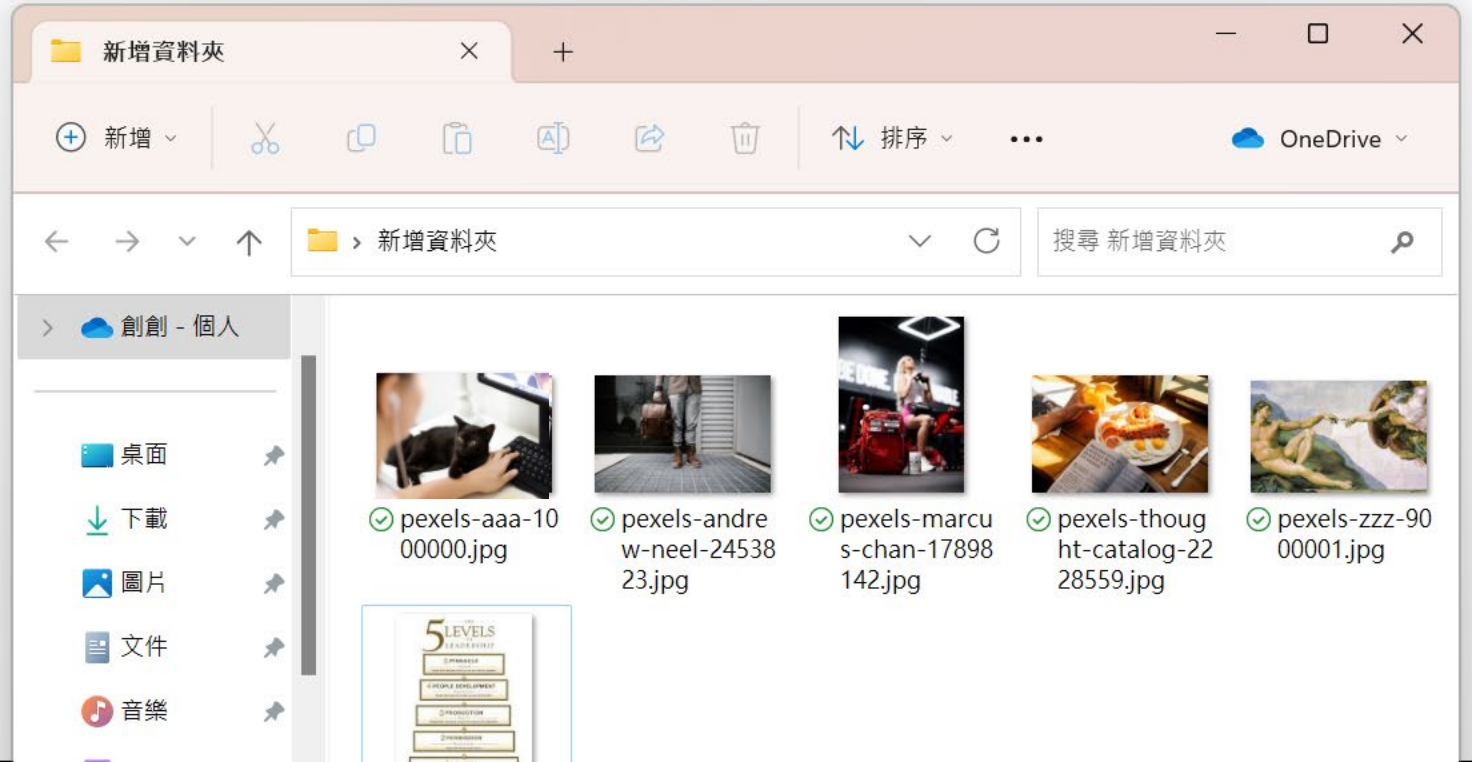
beam search numbers) 1

Temperature 1

MiniGPT-4

User

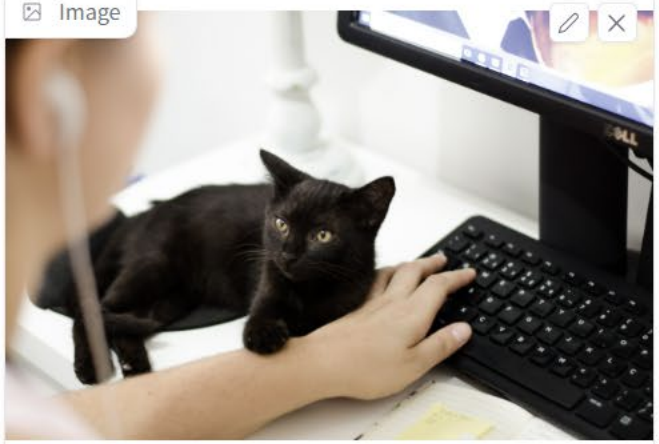
Please upload your image first



仍須點擊「Upload & Start Chat」才能開始提問 Chat

Project Page | Github Code | Paper PDF

Image



MiniGPT-4

User

Please upload your image first

Upload & Start Chat

Restart

beam search numbers) 1

Temperature 1

Use via API · Built with Gradio

上傳圖片中（約要跑十幾秒）

Project Page | Github Code | Paper PDF

queue: 1/1 | 16.4/18.4s

MiniGPT-4

queue: 1/1 | 16.4/18.4s

Upload & Start Chat

Restart

beam search numbers) 1

Temperature 1

Use via API · Built with Gradio

設定參數 Beam search numbers 和 Temperature

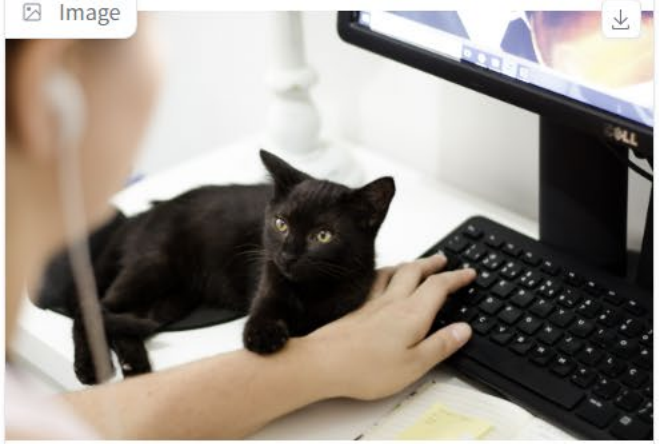
Beam search numbers 預設為 **1**，這個參數在訓練的時候影響比較大，預設為**1**時只會生成一個結果，加大這項數值會生成多個結果；然後從中找出最適合的答案，但會佔用額外的顯示卡記憶體。

Temperature 預設為 **1**，數值區間為 **0.1 ~ 2**，這項數值用來控制結果的隨機性，不會佔用額外的顯示卡記憶體。

提示語輸入框

Project Page | Github Code | Paper PDF

Image



Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

User

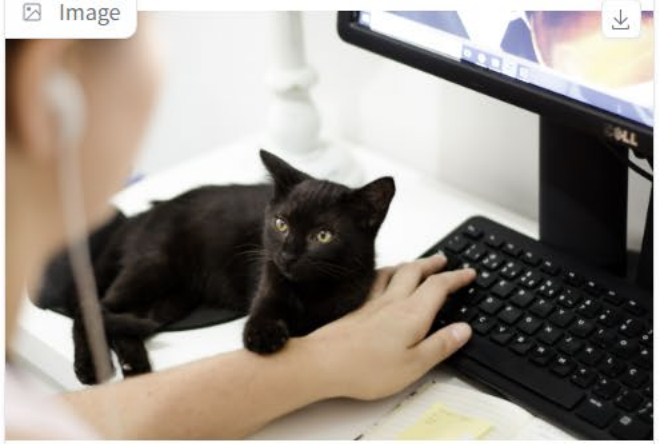
Type and press Enter

Use via API · Built with Gradio

輸入提示語

Project Page | Github Code | Paper PDF

Image



Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

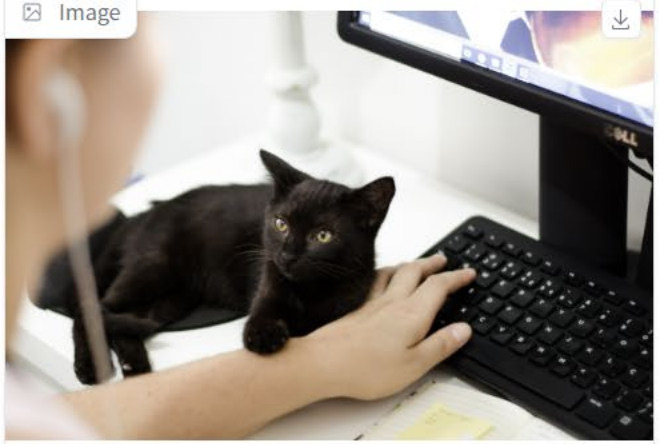
User

請說明本圖所代表的狀況，並且依此給我三個職場中的生存法則。

輸入提示語，按下 Enter 鍵送出（約要跑十幾秒）

Project Page | Github Code | Paper PDF

Image



Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

請說明本圖所代表的狀況，並且依此給我三個職場中的生存法則。

• • •

User

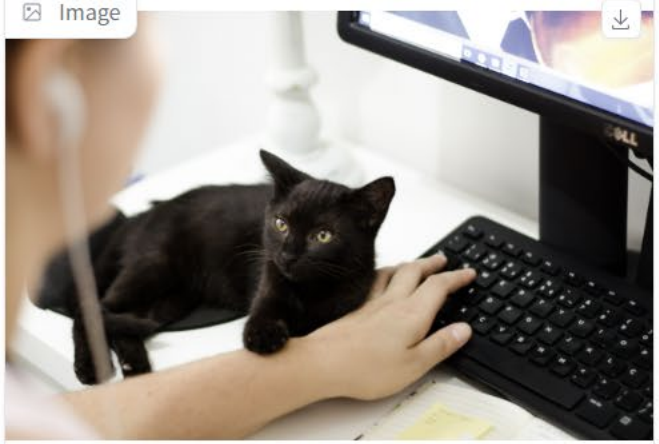
Type and press Enter

Use via API · Built with Gradio

使用中文提示語，理解力很差、中翻英也很差

Project Page | Github Code | Paper PDF

Image



Start Chatting

Restart

beam search numbers) 1

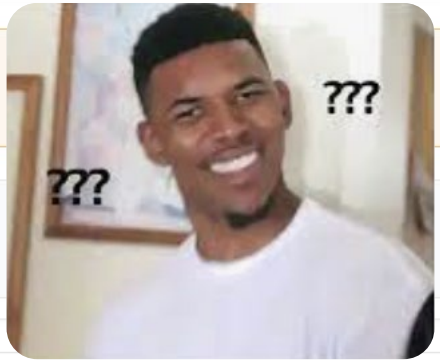
Temperature 1

MiniGPT-4

請說明本圖所代表的狀況，並且依此給我三個職場中的生存法則。

這個圖像顯示了一個人在輸入機鍵板上輕輕摸著一隻小黑貓。這個人可能是在使用電腦，或者是在輕輕摸著貓的感覺。

請翻譯"杭州在哪裡"?



“Where is Nanchang?”

杭州?? 南昌??

User

Type and press Enter

Use via API · Built with Gradio

使用英文提示語，效果好很多

建議！先翻成英文再貼入提示語輸入框：

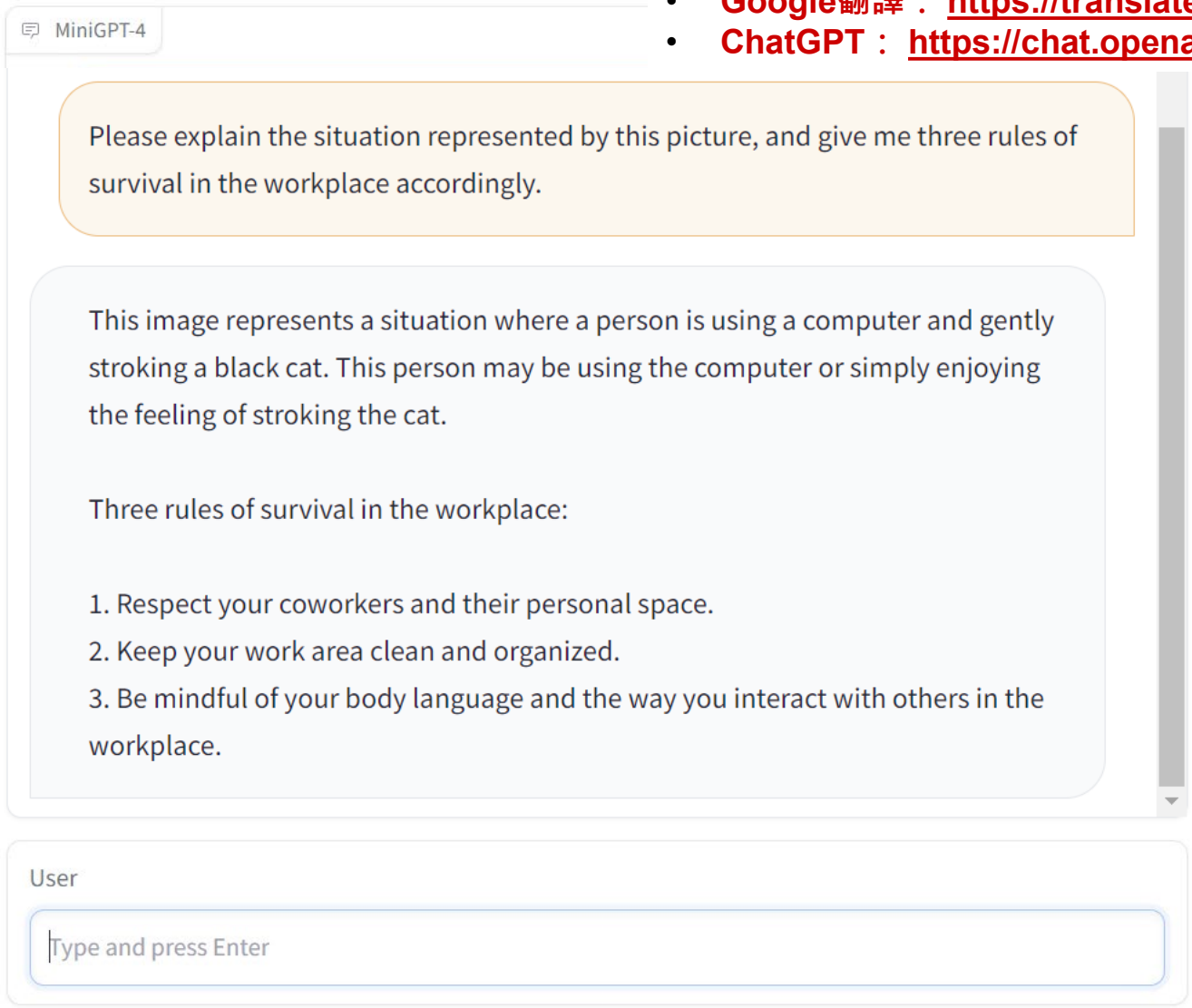
- Google翻譯： https://translate.google.com/
- ChatGPT： https://chat.openai.com/

請說明本圖所代表的狀況，並且依此給我三個職場中的生存法則。

該圖像代表了一個人正在使用計算機並輕輕撫摸一隻黑貓的情況。這個人可能正在使用電腦或者只是享受撫摸貓的感覺。

職場生存三法則：

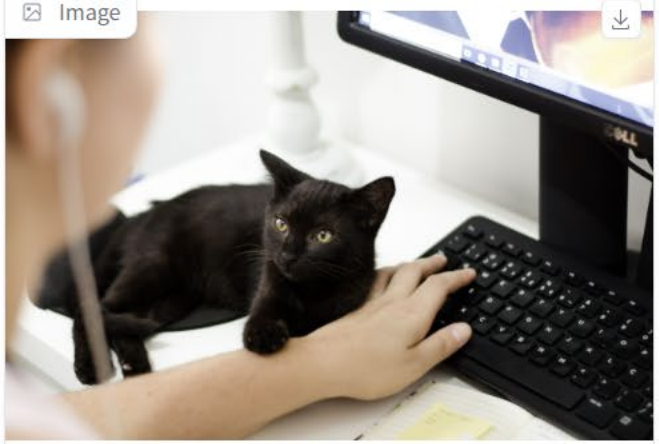
1. 尊重同事及其個人空間。
2. 保持工作區域乾淨整潔。
3. 注意您的肢體語言以及您在工作場所與他人互動的方式。



雖然中翻英很差，但是英翻中很好

Project Page | Github Code | Paper PDF

Image



Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Three rules of survival in the workplace:
1. Respect your coworkers and their personal space.
2. Keep your work area clean and organized.
3. Be mindful of your body language and the way you interact with others in the workplace.
翻譯成中文

在工作場所中的生存規則：

1. 尊重同事和他們的個人空間。
2. 保持你的工作區域清潔和整潔。
3. 意識到你的身體語言和與同事在工作場所互動方式。

User

Type and press Enter

Use via API · Built with Gradio

結論：

- **多使用「英文提示語」與 MiniGPT 溝通**
- **在 MiniGPT 的提示語輸入框中，盡量使用英文**

MiniGPT 自動看圖說故事

MiniGPT 的使用範例

根據圖片寫一篇「商品介紹」

提示語：

中文 根據圖片寫一篇商品介紹。

英文翻譯 Write a product introduction based on the picture.

參考圖片：

示範 1：根據圖片寫一篇「商品介紹」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Write a product introduction based on the picture

This image shows a man standing in front of a building, holding a brown leather briefcase. He is wearing a brown suit, a white shirt, and black shoes. The briefcase is made of leather and has a brown handle. The man is standing on a gray tile floor and there is a white wall behind him. The building behind him appears to be made of concrete and has large windows. The man is looking at something in his briefcase.

這張圖片顯示一名男子站在建築物前，手裡拿著一個棕色皮革公文包。他穿著棕色西裝、白色襯衫和黑色鞋子。公文包由皮革製成，配有棕色手柄。男人站在灰色的瓷磚地板上，身後是一堵白色的牆。他身後的建築物似乎是由混凝土製成的，並有大窗戶。男人正在看公文包裡的東西。

Use via API · Built with Gradio

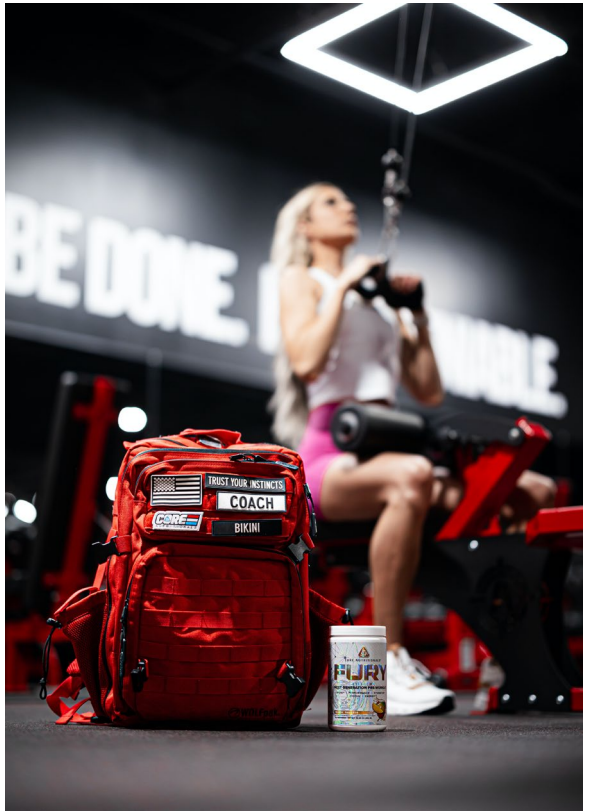
根據圖片寫一篇「Blog推薦文章」

提示語：

中文 根據圖片寫一篇Blog推薦文章。

英文翻譯 Write a blog recommendation article based on the picture.

參考圖片：



示範 2：根據圖片寫一篇「Blog推薦文章」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Write a blog recommendation article based on the picture

The image shows a woman in a gym, working out on a weight machine. She is wearing a red sports bra and shorts, and has a water bottle on the ground next to her. The gym is equipped with weight machines, dumbbells, and other fitness equipment. The walls are painted in a dark color, and there are mirrors on the walls to allow the gym goers to see themselves while working out. The overall atmosphere of the gym is professional and well-maintained.

圖片顯示一名女子在健身房裡用舉重機鍛煉。她穿著紅色運動胸罩和短褲，旁邊的地上有一個水瓶。健身房配有舉重機、啞鈴等健身器材。牆壁被漆成了深色，牆上還裝有鏡子，讓健身的人在鍛煉時可以看到自己。健身房的整體氛圍專業且維護良好。

該圖像很好地體現了健身房以及健身在生活中的重要性。圖中的女士正在專注於鍛煉，健身房設備齊全，可幫助她實現健身目標。牆壁和鏡子的深色營造出一種專業和嚴肅的感覺，這對於想要提高健身水平的人們來說非常重要。健身房的整體氛圍維持得很好，這對於想要在乾淨、安全的環境中鍛煉的人們來說非常重要。

根據圖片寫一篇「料理教學」

提示語：

中文 根據圖片寫一篇料理教學。

英文翻譯 Write a cooking tutorial based on the picture.

參考圖片：

示範 3：根據圖片寫一篇「料理教學」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Write a cooking tutorial based on the picture

Sure, here's a cooking tutorial based on the image:

How to Make a Delicious Breakfast:

Start by setting up your kitchen with all the necessary ingredients. In this case, you'll need eggs, bacon, orange slices, and a glass of orange juice.

1. Crack the eggs into a bowl and beat them with a fork.
2. Heat up the bacon in a pan until it's crispy.
3. Slice the orange slices into thin rounds.
4. Pour yourself a glass of orange juice.
5. Place a plate on top of the bacon and sprinkle some salt and pepper on it.
6. Pour the beaten eggs onto the bacon and sprinkle some cheese on top.
7. Place the orange slices on top of the eggs.
8. Serve the dish immediately and enjoy your delicious breakfast!

User

示範 3：根據圖片寫一篇「料理教學」

MiniGPT 的回應：

Sure, here’s a cooking tutorial based on the image:

How to Make a Delicious Breakfast:

Start by setting up your kitchen with all the necessary ingredients. In this case, you’ll need eggs, bacon, orange slices, and a glass of orange juice.

1. Crack the eggs into a bowl and beat them with a fork.
2. Heat up the bacon in a pan until it’s crispy.
3. Slice the orange slices into thin rounds.
4. Pour yourself a glass of orange juice.
5. Place a plate on top of the bacon and sprinkle some salt and pepper on it.
6. Pour the beaten eggs onto the bacon and sprinkle some cheese on top.
7. Place the orange slices on top of the eggs.
8. Serve the dish immediately and enjoy your delicious breakfast!

Note: You can also add other ingredients to this breakfast, such as sausage, ham, or even avocado for a more filling meal.

中文翻譯：

當然，這是一個基於圖像的烹飪教程：

如何製作美味的早餐：

首先用所有必要的原料來設置你的廚房。在這種情況下，你需要雞蛋、培根、橙片和一杯橙汁。

1. 將雞蛋打入碗中，用叉子攪打。
2. 將培根放入鍋中加熱至酥脆。
3. 將橙片切成薄圓片。
4. 給自己倒一杯橙汁。
5. 將一個盤子放在培根上，撒上一些鹽和胡椒粉。
6. 將打散的雞蛋倒在培根上，撒上一些奶酪。
7. 將橙片放在雞蛋上。
8. 立即上菜，享用美味的早餐！

注意：您還可以在早餐中添加其他成分，例如香腸、火腿，甚至牛油果，以獲得更飽腹的一餐。

根據圖片寫一篇「圖表說明」

提示語：

中文 請說明一下這張圖的由來，並解釋本圖的含意。

英文翻譯 Please explain the origin of this picture and explain the meaning of this picture.

參考圖片：

示範 4：根據圖片寫一篇「圖表說明」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Please explain the origin of this picture and explain the meaning of this picture.

The image shows the five levels of leadership, which are:

1. Position Power
2. Personal Power
3. Role Power
4. Task Power
5. Positional Power

Each level of leadership has its own unique characteristics and responsibilities. Position Power refers to the authority and influence that a leader has due to their position in an organization. Personal Power refers to the leader's ability to influence others through their personal qualities and reputation. Role Power is the power that a leader has due to their role in an organization. Task Power is the power that a leader has to accomplish specific tasks or objectives. Positional Power is the power that a leader has due to their position in an organization

User

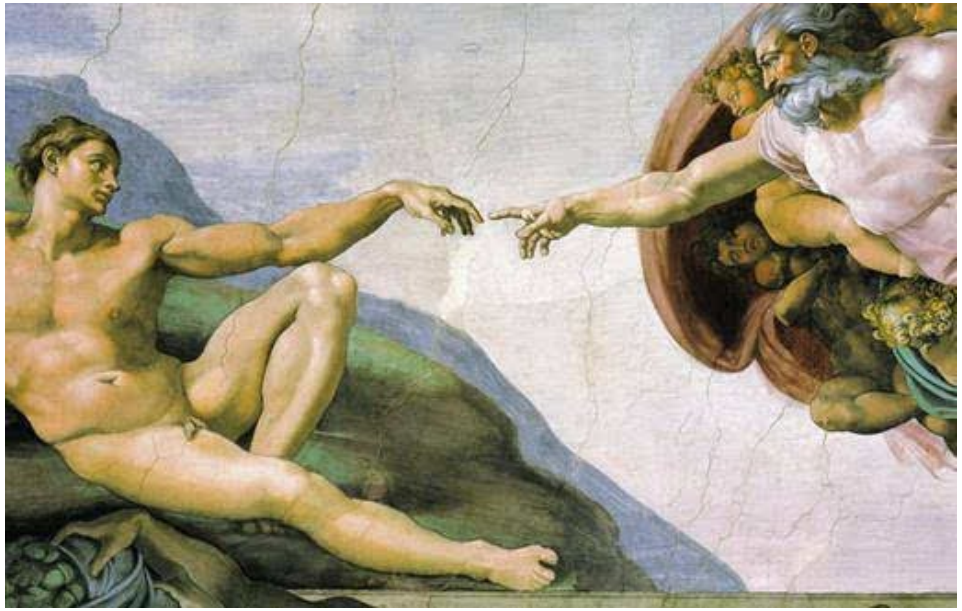
根據圖片寫一篇「有關於ＸＸＸ的故事」

提示語：

中文 用這張圖來說明一個神與人的故事。

英文翻譯 Use this picture to illustrate a story of God and man.

參考圖片：



示範 5-1：根據圖片寫一篇「有關於ＸＸＸ的故事」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Use this picture to illustrate a story of God and man

The image shows the creation of Adam, as depicted in the Sistine Chapel ceiling by Michelangelo Buonarroti. The image shows God reaching out to touch Adam's hand, with Adam looking up at God in awe. The image is a representation of the biblical story of the creation of Adam, the first man, by God. The image is a masterpiece of Renaissance art, and is considered one of the most famous works of art in the world.

該圖像顯示了亞當的創造，正如米開朗基羅·博納羅蒂在西斯廷教堂天花板上所描繪的那樣。圖像顯示上帝伸出手來觸摸亞當的手，亞當敬畏地抬頭看著上帝。該圖像再現了聖經中上帝創造第一個人亞當的故事。該圖像是文藝復興時期藝術的傑作，被認為是世界上最著名的藝術作品之一。

Use via API · Built with Gradio

根據圖片寫一篇「有關於ＸＸＸ的故事」

提示語：

先說目的

我是一名神話故事專家，專門分享神話故事的內容，

參考資料

我想要根據本圖片中的神話故事，寫一篇關於創世紀的腳本，需要條列3個創世紀的重點內容，而且世人皆知，

輸出限制

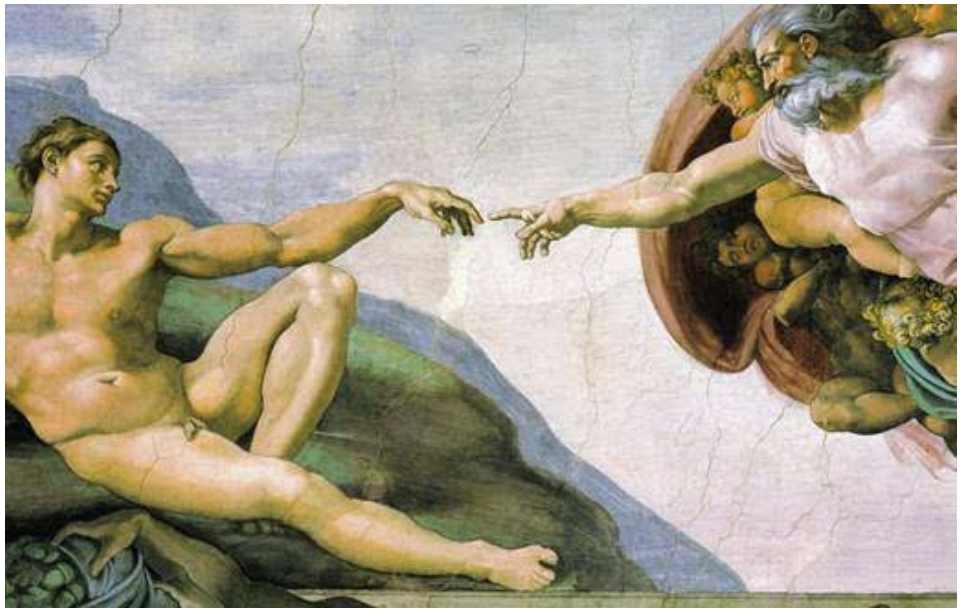
請用第一人稱給新手、一般民眾也能聽得懂的口語化方式來撰寫，只能限定於300~500字內。

I am a mythology expert, specializing in sharing the content of mythology,

I want to write a script about Genesis based on the myths in this picture, and I need to list 3 key points of Genesis, which are well known to the world,

Please use the first person to write in a colloquial way that beginners and ordinary people can understand, and it can only be limited to 300~500 words.

參考圖片：



示範 5-2：根據圖片寫一篇「有關於ＸＸＸ的故事」

Project Page | Github Code | Paper PDF

Image

Start Chatting

Restart

beam search numbers) 1

Temperature 1

MiniGPT-4

Use this picture to illustrate a story of God and man

The image shows the creation of Adam, as depicted in the Sistine Chapel ceiling by Michelangelo Buonarroti. The image shows God reaching out to touch Adam's hand, with Adam looking up at God in awe. The image is a representation of the biblical story of the creation of Adam, the first man, by God. The image is a masterpiece of Renaissance art, and is considered one of the most famous works of art in the world.

User

I am a mythology expert, specializing in sharing the content of mythology,
I want to write a script about Genesis based on the myths in this picture, and I need to list 3 key points of Genesis, which are well known to the world,
Please use the first person to write in a colloquial way that beginners and ordinary people can understand, and it can only be limited to 300~500 words.

Use via API · Built with Gradio

示範 5-2：根據圖片寫一篇「有關於ＸＸＸ的故事」

MiniGPT 的回應：

Genesis is the first book of the Bible and tells the story of the creation of the world and the first humans, Adam and Eve. The key points of Genesis are:

1. God created the world in six days and rested on the seventh. This is the foundation of the biblical creation story and sets the stage for the rest of the Bible.
2. God created Adam and Eve, the first humans, in the Garden of Eden. They were given the task of tending the garden and were warned not to eat from the tree of knowledge of good and evil.
3. Adam and Eve disobeyed God’ s command and ate from the tree, which led to their expulsion from the Garden and the introduction of sin into the world. This event is known as the Fall and has significant implications for humanity, as it sets the stage for the need for salvation through Jesus Christ.

Overall, Genesis is a foundational book of the Bible that sets the stage for the rest of the scriptures and provides important context for understanding the nature of God, the origin of humanity, and the relationship between humans and God.

中文翻譯：

創世記是聖經的第一本書，講述了創造世界和第一批人類亞當和夏娃的故事。 創世記的要點是：

1. 上帝用六天創造了世界，第七天休息了。 這是聖經創世故事的基礎，並為聖經的其餘部分奠定了基礎。
2. 上帝在伊甸園創造了第一批人類亞當和夏娃。 他們被賦予了照料花園的任務，並被警告不要吃善惡知識樹上的果子。
3. 亞當和夏娃違背了上帝的命令，吃了樹上的果子，這導致他們被逐出伊甸園，並將罪惡帶入了世界。 這一事件被稱為墮落，對人類具有重大影響，因為它為通過耶穌基督獲得救贖的需要奠定了基礎。

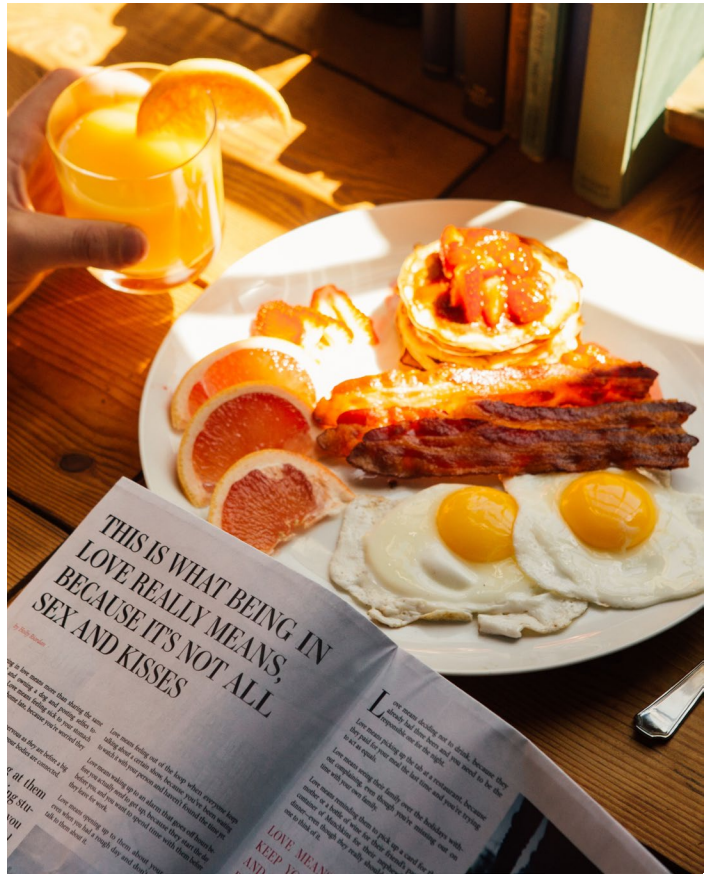
總的來說，《創世記》是聖經的基礎書，為聖經的其餘部分奠定了基礎，並為理解神的本質、人類的起源以及人與神之間的關係提供了重要的背景。

MiniGPT 自動看圖說故事

課後練習

課後練習

① 拍攝一張家電照片，上傳給 MiniGPT 進行解析，並且要求一篇「商品介紹」。

② 拍攝一張料理照片，上傳給 MiniGPT 進行解析，並且要求一篇「符合影片腳本提問公式」的「料理教學」。

（以上兩題，請提供你的提示語和MiniGPT的回應，皆以英文提供）