



IMAGE TO TEXT

Image-to-text 操作手冊

開發者：邱之宇、林瑞庭

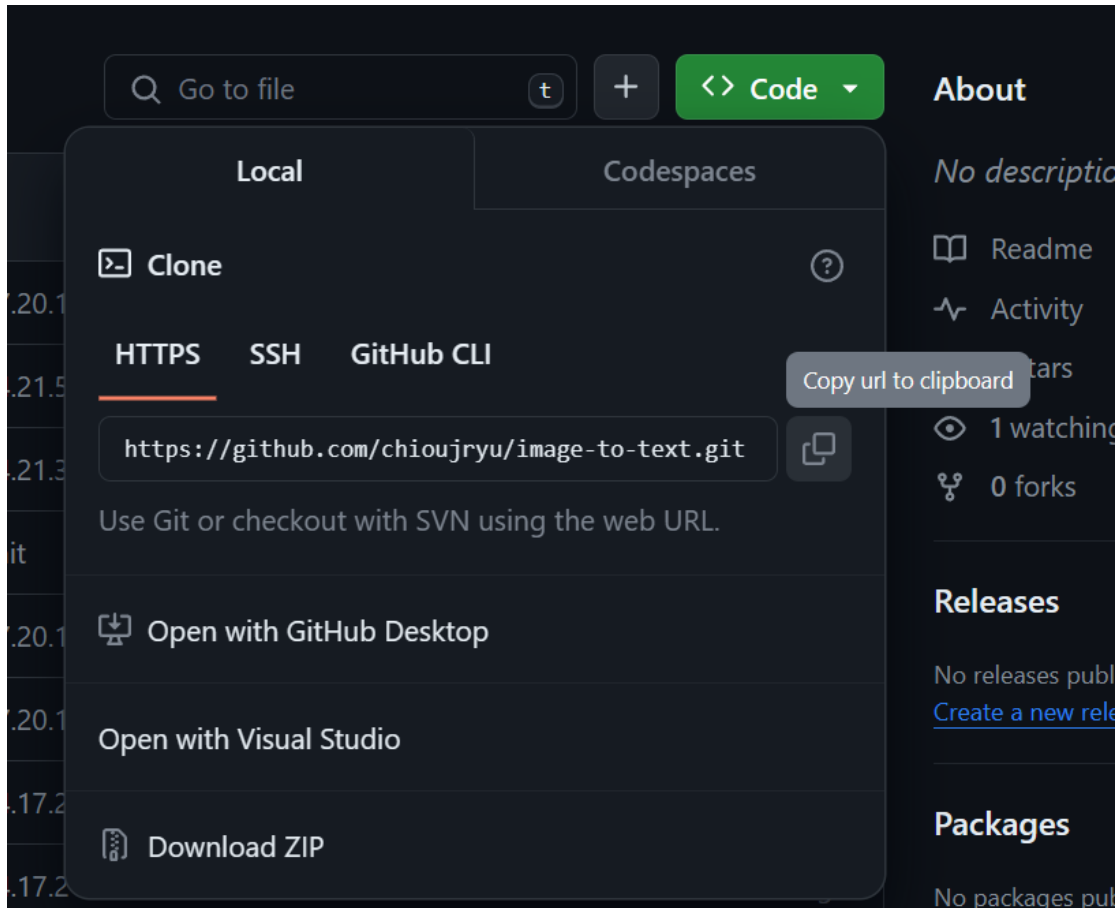
目錄

- 一、 下載程式檔案3
- 二、 安裝環境4
 - 1.1 安裝 Anaconda4
 - 1.2 使用 YAML 創建虛擬環境4
- 三、 使用圖像轉文字功能4
 - 範例一：5
 - 範例二：5
- 四、 未來展望6

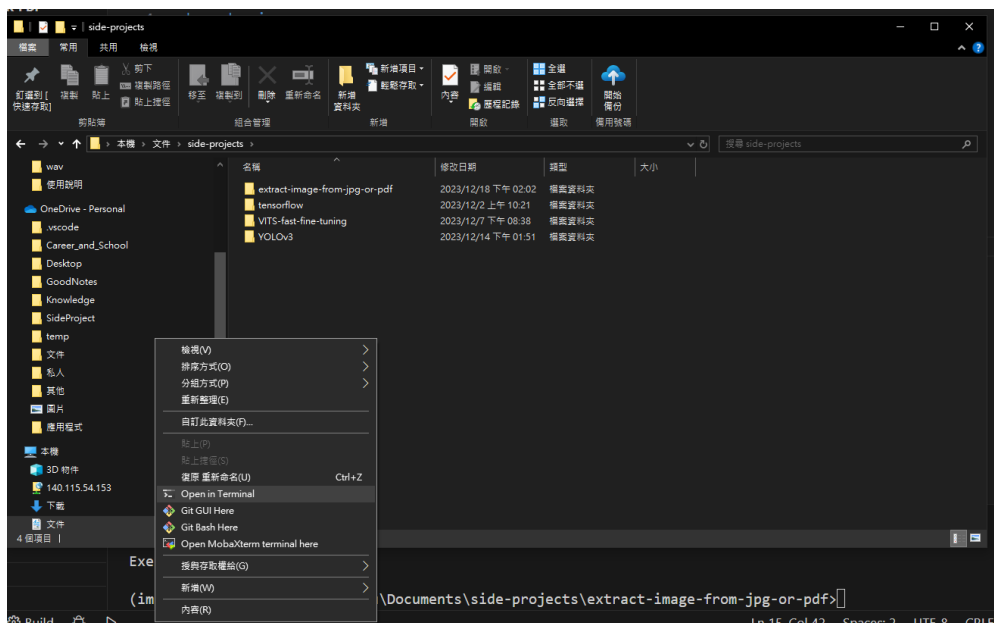
一、下載程式檔案

請利用 [chioujryu/image-to-text \(github.com\)](https://github.com/chioujryu/image-to-text) 這個網址下載程式碼。

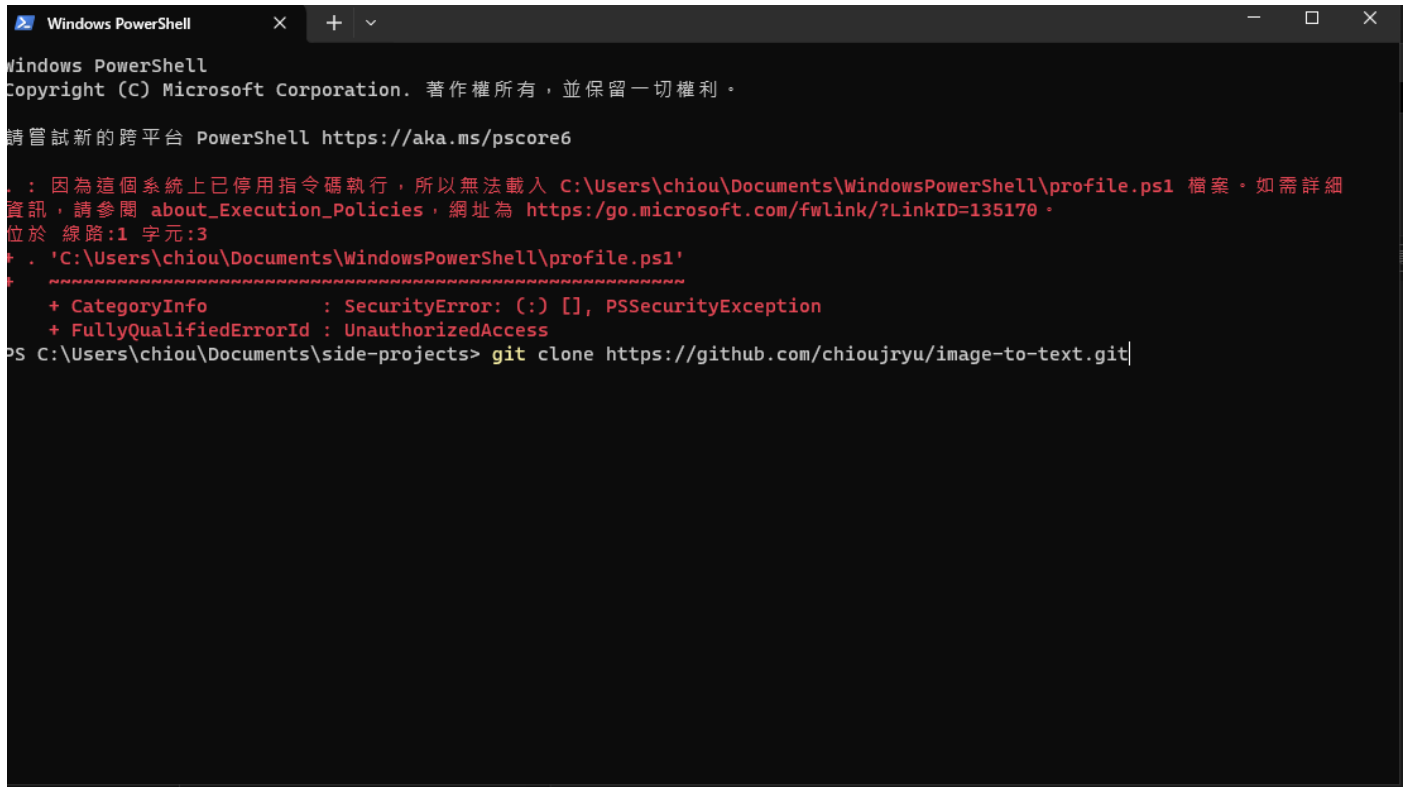
點擊 Copy url to clipboard



進到想要的資料夾，按下右鍵，並點擊 Open Terminal。



接著輸入 `git clone https://github.com/chioujryu/image-to-text.git`，然後按下 Enter



```
Windows PowerShell
Copyright (C) Microsoft Corporation. 著作權所有，並保留一切權利。

請嘗試新的跨平台 PowerShell https://aka.ms/pscore6

. : 因為這個系統上已停用指令碼執行，所以無法載入 C:\Users\chiou\Documents\WindowsPowerShell\profile.ps1 檔案。如需詳細
資訊，請參閱 about_Execution_Policies，網址為 https://go.microsoft.com/fwlink/?LinkID=135170。
位於 線路:1 字元:3
+ ~~~~~
+ CategoryInfo          : SecurityError: (:) [], PSSecurityException
+ FullyQualifiedErrorId : UnauthorizedAccess

PS C:\Users\chiou\Documents\side-projects> git clone https://github.com/chioujryu/image-to-text.git
```

二、安裝環境

1.1 安裝 Anaconda

進到 [Free Download | Anaconda](#) 安裝虛擬環境，或是可以選擇任意的虛擬環境安裝。

1.2 使用 YAML 創建虛擬環境

激活 Anaconda 後，在 Terminal 輸入：

```
conda env create -f environment.yml
```

三、使用圖像轉文字功能

打開 test.py 檔案，更改 image_url 以及 prompt，即可詢問 AI 圖片裡面的資訊。

範例一：

```
# 初始化
prompt=" Question: What is in this picture? Answer:"
image_url = "https://img.onl/sDBysN"
```

圖片來自網路上的圖片，是一個標誌，上面有 SPEED LIMIT 40 的字樣。

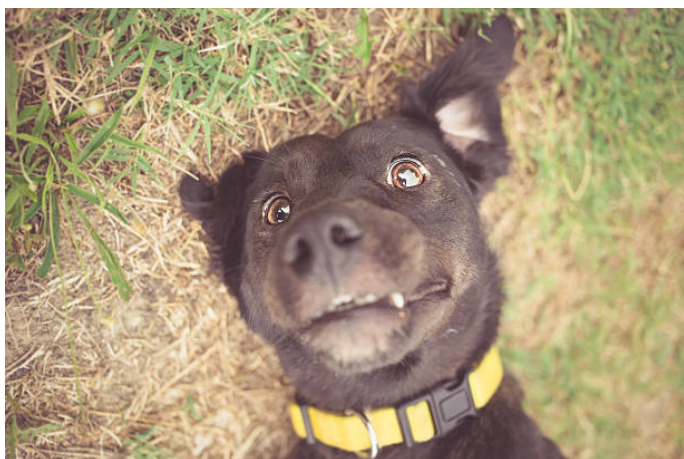


AI 的回答是：speed limit 40

```
image = <PIL.JpegImagePlugin.JpegImageFile image mode=RGB size=497x331 at 0x181C1573070>
speed limit 40
```

範例二：

給 AI 看一隻狗的照片，Prompt 則是 Question: What is in this picture? tell me more detail. Answer:



AI 的回答是：A dog

```
image = <PIL.JpegImagePlugin.JpegImageFile image mode=RGB size=612x408 at 0x1FCA66D5AB0>  
A dog
```

四、 未來展望

隨著 OpenAI 的 ChatGPT 的 Image-recognize 功能愈來愈強，當 OpenAI 的 Image-recognize API 推出之後，便可以拿來使用並取代我們現在的專案技術，OpenAI 的 Image-recognize 可以分析圖表、分析流程圖、更重要的是可以讓大語言多模態模型學習如何閱讀 Grafcet，期待有朝一日 OpenAI 能推出相關的 API。