

**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](#_Toc153803678)

[二、 安裝環境 5](#_Toc153803679)

[1.1 安裝Anaconda 5](#_Toc153803680)

[1.2 使用YAML創建虛擬環境 5](#_Toc153803681)

[三、 使用圖像轉文字功能 5](#_Toc153803682)

[範例一： 6](#_Toc153803683)

[範例二： 6](#_Toc153803684)

[四、 未來展望 7](#_Toc153803685)

# 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

一張含有 文字, 螢幕擷取畫面, 軟體, 多媒體軟體 的圖片

自動產生的描述

進到想要的資料夾，按下右鍵，並點擊Open Terminal。

一張含有 文字, 螢幕擷取畫面, 軟體, 多媒體軟體 的圖片

自動產生的描述

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

一張含有 文字, 電子產品, 螢幕擷取畫面, 軟體 的圖片

自動產生的描述

# 安裝環境

## 安裝Anaconda

進到 [Free Download | Anaconda](https://www.anaconda.com/download) 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：



# 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：

一張含有 文字, 字型, 螢幕擷取畫面 的圖片

自動產生的描述

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

一張含有 文字, 戶外, 天空, 招牌 的圖片

自動產生的描述

AI 的回答是：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



# 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。