

# 文字提示驅動之桌面機器人流程自動化

## Desktop RPA by Prompting

指導教授: 李信杰教授

專題成員: 李維軒、歐陽主敬

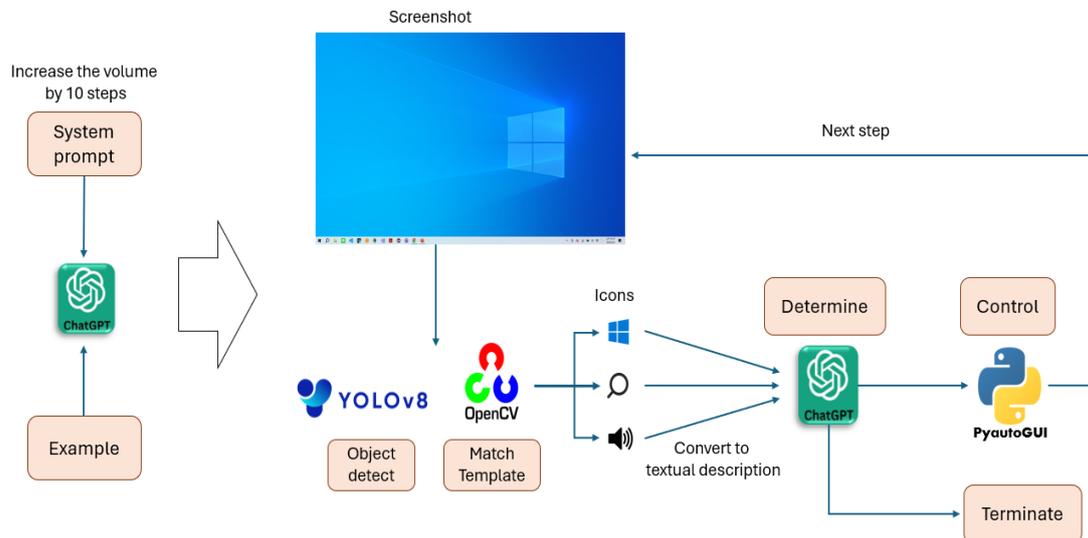
開發工具: GPT4、OpenCV、Python、YOLOv8

測試環境: Windows

### 一、簡介：

近年來，大型語言模型(LLM)的發展推動了許多與自然語言處理(NLP)相關的研究與發想。本專題結合了 GPT4、OpenCV 的 Match Template、Python 的 PyautoGUI 模組以及 YOLOv8，實作出了一個基於文字提示(Prompt)的桌面機器人流程自動化(Desktop RPA)。只需使用者用自然語言輸入「請幫我關機」、「幫我調高螢幕亮度」、「幫我打開 Edge 搜尋 PChome」等，就可以完全不用動手，單純靠我們的軟體自動去執行這些操作。

以下為系統架構圖：



## 二、 測試結果：

可以看到，當我們下了「幫我調高螢幕亮度 20 格」這個指令時，GPT 成功的依序點擊了開始按鈕->設定->系統，在看到調整螢幕亮度的 slider 後，GPT 滑動了這個 slider 成功調高螢幕亮度

```
1 GOAL: increase the screen light by 20 steps
2 1.
3 icons:
4 volume(1) start_menu (Detect the icons of current screenshot, and used as prompt)
5 ChatGPT:
6 start_menu click None (Click Start menu, no text input)
7 2.
8 icons:
9 volume(1) power start_menu setting
10 ChatGPT:
11 setting click None (Click Setting, no text input)
12 3.
13 icons:
14 searchbar_0 volume(1) start_menu device personalize system
15 ChatGPT:
16 system click None (Click System, no text input)
17 4.
18 icons:
19 searchbar_0 volume(1) adjust_slider power start_menu
20 ChatGPT:
21 adjust_slider drag 56 (Drag adjust slider right by 56 pixels)
22 5.
23 icons:
24 volume(1) adjust_slider power start_menu
25 ChatGPT:
26 -1 terminate None (Complete the task)
```

在下面這個截圖例子中，擷取到了以下 Icon

