

# 人臉辨識系統

# INTRODUCTION

---

- 目標:
  1. 辨識出明星照片及組內成員照片
  2. 比較哪種方式可行度較高
  3. 比較不同方式的優缺點
- 模型使用:
  1. 自訓練模型: Tensor flow-Keras 、Tensor flow-face recognition
  2. 大型語言模型(LLM): ChatGPT-4o(MoE) 、 Google Gemini(MoE, after1.5pro)
- 辨識階段:
  1. 判斷主流明星
  2. 拍攝組內成員進行訓練&辨識

# DATA SOURCE

---

1. 上網搜尋主流明星照片

- Johnny Depp
- John Cena



Johnny Depp



John Cena

2. 拍攝小組成員照片



# INTRO – TENSOR FLOW

---

Tensor flow是一個機器學習框架，用於機器學習的開源軟體庫，可以支援深度學習的各種演算法

- 處理器：可在CPU、TPU、GPU上執行
- 平台：具有跨平台能力，像是Windows、Linux、Raspberry pi...等
- 前端程式語言：可支援多種前端程式語言
- 高階API：可開發許多種高階的API，例如：Keras、TF-Learn、TF-Slim等

# INTRO – TENSOR FLOW\_KREAS

---

Keras 是一個開放原始碼，基於 Python 高階深度學習的程式庫

- 可以快速有方便運算的主要原因：

它已將訓練模型的輸入層、隱藏層、輸出層，做好架構，使用者只需要加入並且填寫正確的參數即可

# INTRO – TENSOR FLOW\_FACE RECOGNITION

---

- Face-recognition 模組的主要功能包括：  
人臉偵測、人臉特徵提取、人臉識別、人臉追蹤、人臉操作
- Face-recognition 模組的優點包括易於安裝和使用，以及強大的功能。它可以應用於多個領域，如安全認證、圖像檢索、人臉標記等。但對於大規模的人臉數據集，需要考慮算法的性能和準確性。

# INTRO – GOOGLE GEMINI

---

Gemini是一個由 Google AI 創建的大型語言模型。接受了大量的文字和程式碼訓練，能夠生成文字、翻譯語言、編寫不同類型的創意內容，並以信息豐富的方式回答問題。

# INTRO – CHATGPT-4O

---

- GPT-3:

- 發佈時間:2020/6
- LLM
- 僅支援文字處理

- GPT-3.5

- 發佈時間:2022
- LLM
- 僅支援文字處理

- GPT-4

- 發佈時間:2023/3
- MoE (Mixture of Experts)
- 支援處理文字、圖片、文件(PDF)

- GPT-4o

- 發佈時間:2024/5/13
- 原生多模態模型→ 處理速度更快
- 支援處理文字、圖片、文件(PDF)



# THE PROCESS AND RESULT

# KERAS

---

- 使用老師上課教的方式試試看是否能實現人臉辨識
- 結論：可以只辨認一個人，但無法辨認好幾位，而且辨識成功的機率非常低，無法同時餵好幾組資料（不同人）訓練模型。

GROUP 3

10

# OPEN CV

---

- 使用OPEN CV 也可以實現AI 人臉辨識。
- 使用模型：

[https://github.com/opencv/opencv/blob/4.x/data/haarcascades/haarcascade\\_frontalface\\_default.xml](https://github.com/opencv/opencv/blob/4.x/data/haarcascades/haarcascade_frontalface_default.xml)

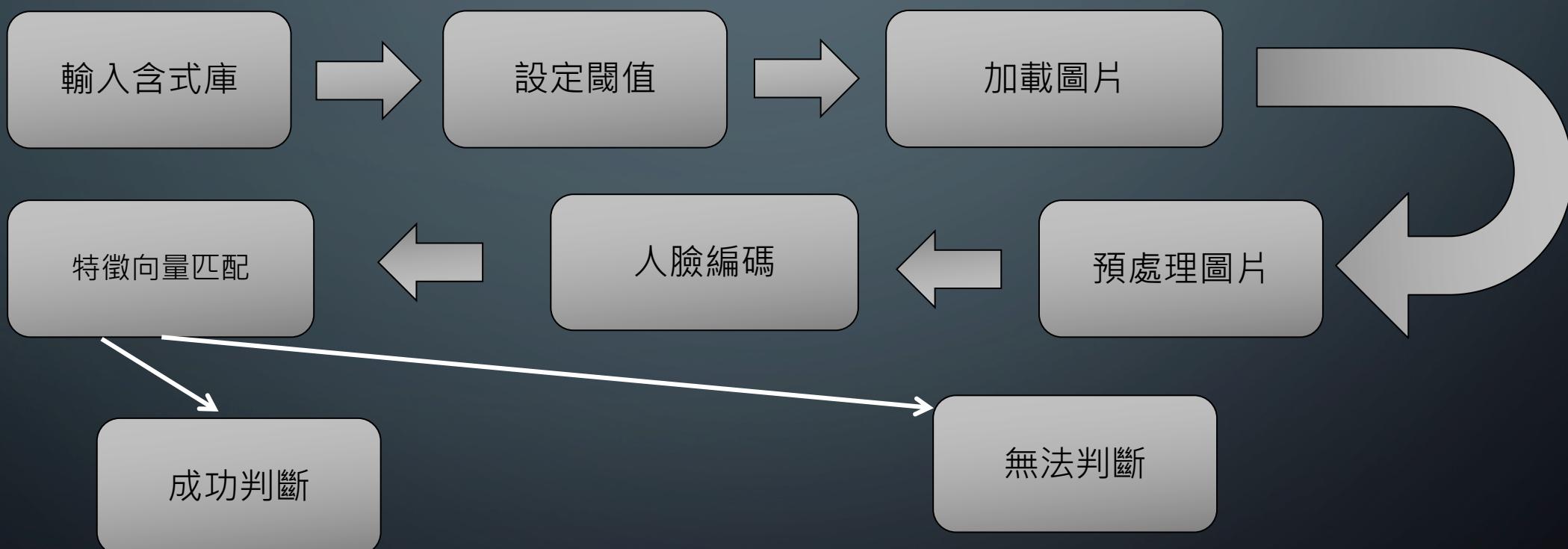
- 結論：失敗了，目前無法正確辨識人臉，餵資料訓練模型時就有些問題。

# TENSOR FLOW\_FACE RECOGNITION

---

- 使用face\_recognition實現AI 人臉辨識。
- 模型特點: 利用深度學習將人臉轉換成特徵向量
- 工作原理: 根據輸入圖片的特徵向量和已知圖片做比較，再根據距離判斷是否匹配。

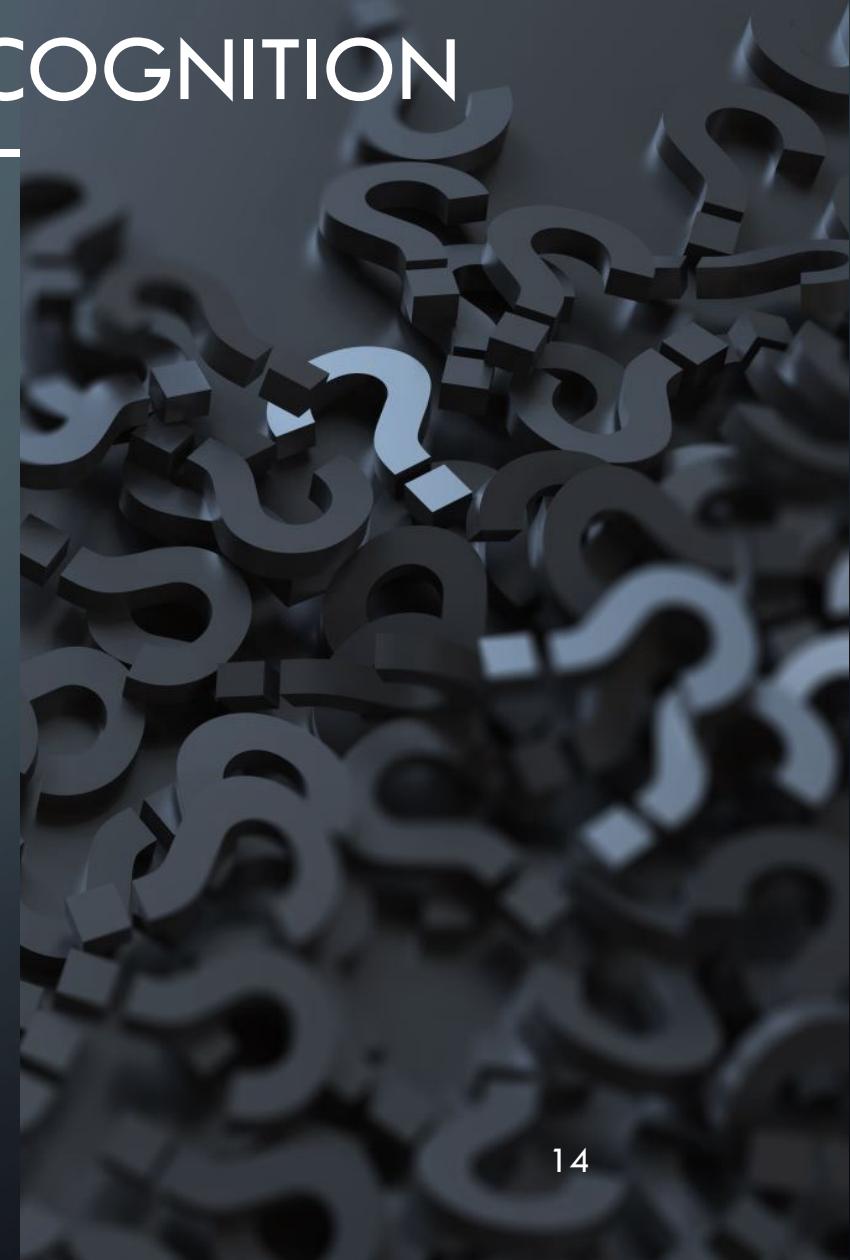
# CODE LOGIC – TENSOR FLOW \_FACE RECOGNITION



# PROBLEM – TENSOR FLOW\_FACE RECOGNITION

---

- 手機在拍照儲存時會存在EXIF元數據中，電腦在讀取時未必會正確處理元數據的方向信息，導致方向錯誤
- 長得像的人會分不出來



# SOLVE – TENSOR FLOW\_FACE RECOGNITION

## 利用PIL函式庫處理EXIF旋轉訊息

1. 使用PIL檢查圖像的EXIF訊息，特別是方向
2. 根據方向信息，對圖像進行旋轉
3. return旋轉後的圖像數據

## 加入閾值

- 設定閾值 0.4 來判斷兩個人臉編碼的歐氏距離
- 若距離<閾值：兩個人臉為同一個人
- 若距離>閾值：兩個人臉是不同的人。

```
def correct_image_orientation(image_path):
    image = Image.open(image_path)

    # 應用 EXIF 旋轉訊息
    try:
        for orientation in ExifTags.TAGS.keys():
            if ExifTags.TAGS[orientation] == 'Orientation':
                break
        exif = dict(image._getexif().items())
        # 需要逆時針轉180度
        if exif[orientation] == 3:
            image = image.rotate(180, expand=True)
        # 需要逆時針270度
        elif exif[orientation] == 6:
            image = image.rotate(270, expand=True)
        # 需要逆時針90度
        elif exif[orientation] == 8:
            image = image.rotate(90, expand=True)
    except (AttributeError, KeyError, IndexError):
        # 無 EXIF 訊息，跳過
        pass

    return np.array(image)
```

▲ 使用PIL處理圖像程式

# RESULT – FACE RECOGNITION主流明星測試

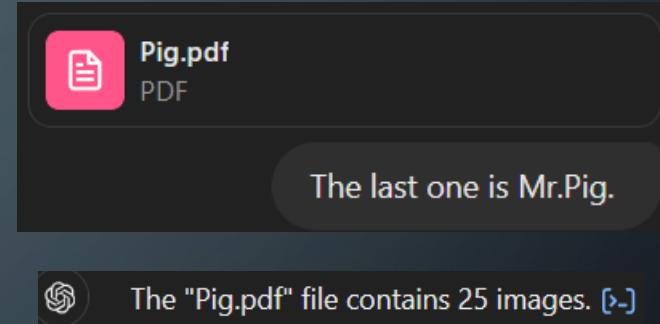
結果:

```
✓ who_is_this("https://hips.hearstapps.com/hmg-  
  
Detected: 新垣結衣
```

```
✓ who_is_this("https://cdn.hk01.com/di/media/images/dw/20240307/8424385847/385847" 0s 1us/step  
  
Downloading data from https://cdn.hk01.com/di/media/images/dw/20240307/8424385847/385847 0s 1us/step  
Detected: Unknown
```

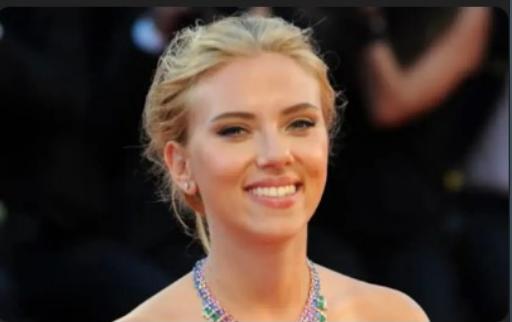
# SETUP – GPT-4O

- 使用網頁版的GPT-4o
- 輸入資料型態: 將每個人的照片製作成pdf格式的”相簿”
- 步驟:
  1. 上傳5人的相簿，要求GPT-4o記住並同時標記對應人物
  2. 識別測試:
    - a) 紿先前上傳的相簿內的照片作為測試，確保GPT-4o有成功記憶並識別
    - b) 紿予一至多張，且參雜非訓練用資料的照片
  3. 統計成功率



# RESULT – GPT-4O 主流明星測試

- 成功率: ~99%



yes, do you know who she is ?



This is a picture of Scarlett Johansson.



5



This is Johnny Depp.



yes, try this one



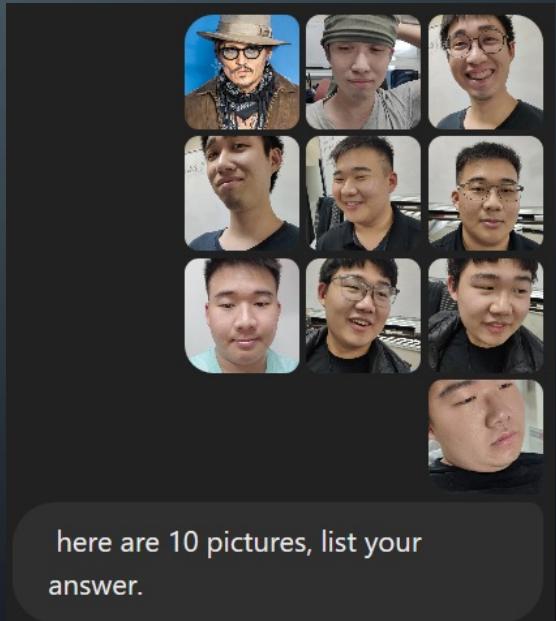
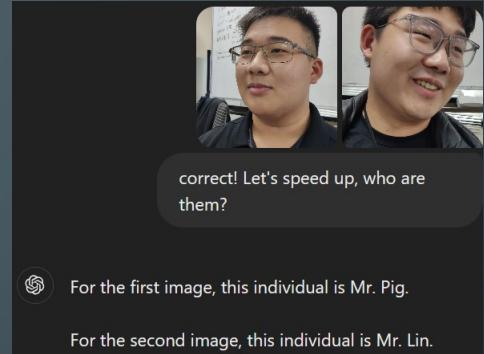
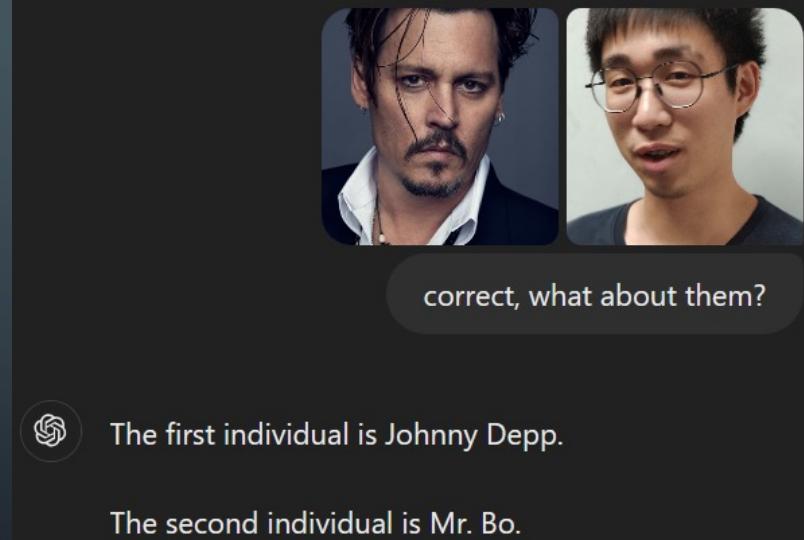
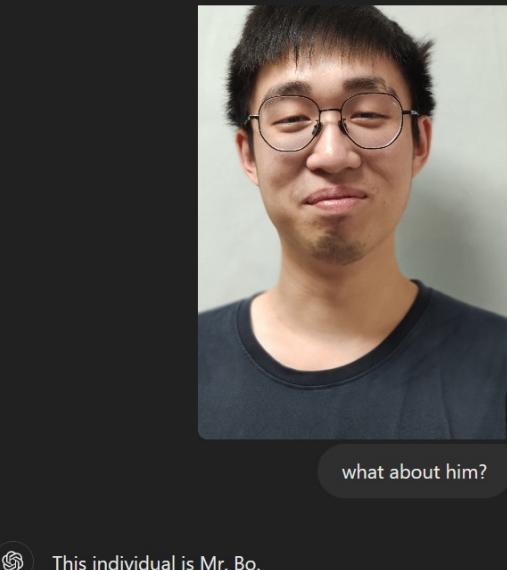
This is a picture of John Cena.

GROUP 3

18

# GPT-4O – RESULT 組員測試

- 成功率:
  - 42/100~42%



# SUMMARY

---

- 自訓練模型：
  - ✓ **Keras**：可以辨識大明星，但成功率非常低
  - ✓ **Open CV**：無法正確辨識人臉，無法正常餵資料訓練模型，可能是程式撰寫的問題
  - ✓ **Face recognition**：可以辨識大明星，可以辨識小組組員，成功率最高
  - ✓ 優點：可以客製化、不用額外花費
  - ✓ 缺點：需要自己從頭做起
- 大型語言模型：
  - ✓ **Google Gemini**：可以辨識大明星，也可以辨識小組成員（數據量不足，尚未統計成功率）
  - ✓ **GPT-4o**：可以辨識大明星，可以辨識小組組員（成功率略低於50%）
  - ✓ 優點：註冊即可直接使用免費的網頁版
  - ✓ 缺點：較無法客製化微調，需要付費才能有更完善功能