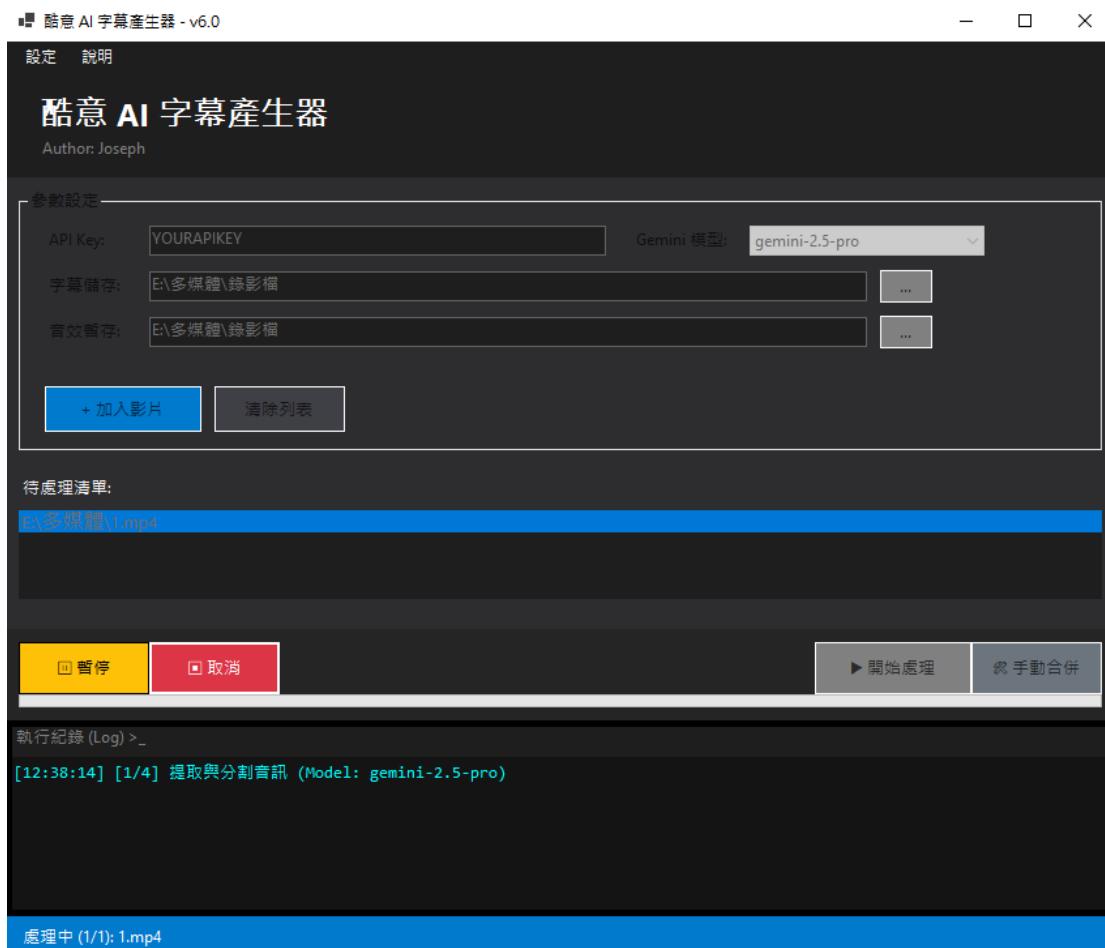


酷意 AI 字幕產生器 (CoolIdea AI Subtitle Generator) 產品使用說明

1. 產品概述

酷意 AI 字幕產生器 (**Studio v6.0**) 是一款由 CoolIdea (Author: Joseph) 開發的智慧型字幕製作工具。本軟體深度整合 Google Gemini AI 模型，旨在解決影片創作者繁瑣的字幕製作流程。透過自動化的「影音分離」、「音訊切割」與「雲端辨識」技術，使用者僅需簡單設定，即可一鍵產出高精準度的 SRT 字幕檔案。



(軟體主介面，包含參數設定、模型選擇與待處理清單)

2. 系統核心功能

- 支援最新 AI 模型**：整合 Google Gemini (如 gemini-2.5-pro)，提供優異的語意理解與聽寫能力。
- 長影片自動分割**：系統自動將長影片的音軌切割為多個片段，突破 AI Token 限制並提升處理穩定度。

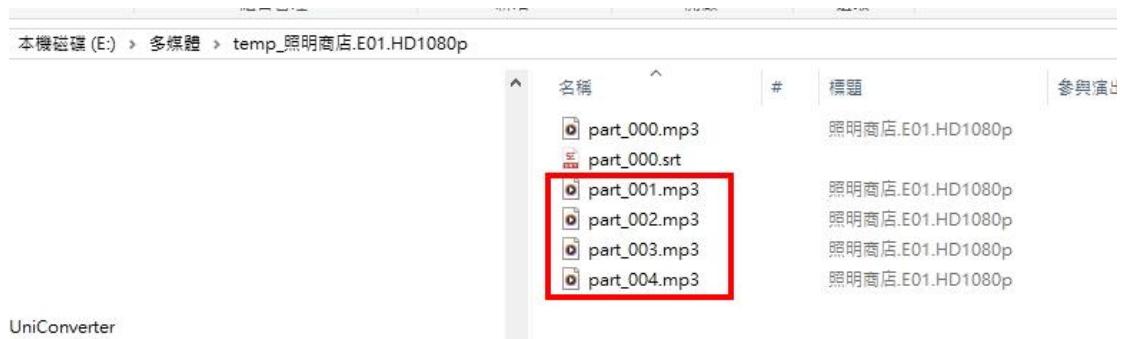
- **智慧合併技術**：分段生成的字幕會自動依據時間軸進行無縫合併，無需人工手動調整。
 - **詳細數據監控**：提供完整的執行 Log，包含 Token 使用量統計與處理進度。
-

3. 自動化作業流程

本軟體將字幕製作簡化為全自動流程，具體運作步驟如下：

步驟一：提取與分割音訊 (Extraction)

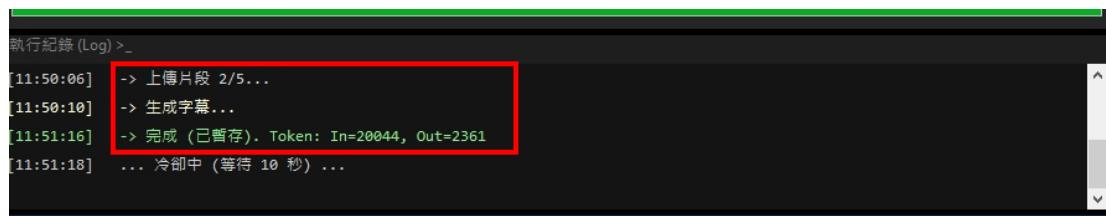
當使用者匯入影片並開始處理後，系統首先會將影片中的聲音提取出來，並依據長度自動切割成數個 MP3 音訊檔（例如 part_000.mp3, part_001.mp3...），以利於分批上傳運算。



(系統自動將原始音訊分割為多個小檔案)

步驟二：AI 雲端運算 (AI Processing)

系統依序將分割後的音訊上傳至 Gemini 進行分析。在此過程中，軟體會自動管理 API 連線，包含上傳、生成字幕以及片段之間的冷卻時間 (Cool-down)，確保流程不中斷。



(Log 視窗即時顯示上傳進度與 Token 使用狀況)

步驟三：分段字幕生成 (Generation)

AI 分析完畢後，會回傳對應的 .srt 字幕檔。您可以在暫存資料夾中看到與音訊檔對應的字幕片段（如 part_000.srt）。

名稱	#	標題	參與演出者	專
part_000.mp3		照明商店.E01.HD1080p		
part_000.srt				
part_001.mp3		照明商店.E01.HD1080p		
part_001.srt				
part_002.mp3		照明商店.E01.HD1080p		
part_003.mp3		照明商店.E01.HD1080p		
part_004.mp3		照明商店.E01.HD1080p		

(音

訊檔與生成的字幕檔成對產生)

生成的字幕內容包含精確的時間軸與對話文字，如下圖所示：

1
01:44,481 --> 01:46,316
這次的停靠站是四月二洞
2
01:47,385 --> 01:49,520
下一站是四月二洞
3
01:50,955 --> 01:55,793
車門即將關閉
4
02:06,128 --> 02:09,131
車門即將關閉
5
02:48,791 --> 02:50,126
車門即將關閉
6
02:51,260 --> 02:53,496
車門即將關閉
7
02:56,365 --> 02:57,266
啊真是的
8
03:55,270 --> 03:57,606
這次的停靠站是四月二洞
9
03:58,473 --> 04:00,675
下一站是四月二洞
10
04:01,776 --> 04:06,448
車門即將關閉
11
04:27,108 --> 04:30,011
啊你好
12
04:31,446 --> 04:35,583
不是什麼大事 只是想問你在這裡做什麼
13
04:39,454 --> 04:40,288
我在等
14
04:41,156 --> 04:42,757
可是剛剛那班是末班車了
15
04:45,560 --> 04:46,061
是的
16
04:49,497 --> 04:52,667
看起來你好像要去哪裡
17

(SRT 檔案內容預覽，包含時間碼與辨識文字)

步驟四：批次處理完成

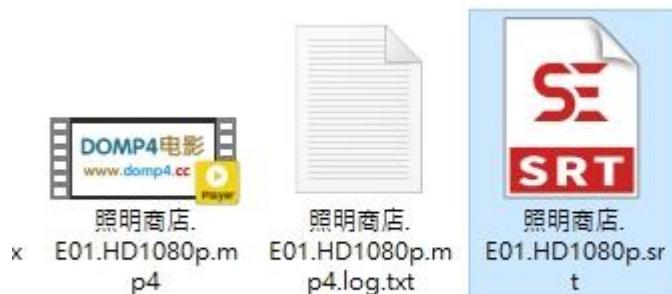
系統會持續運作，直到所有分割的音訊片段皆完成辨識。

part_000.mp3
 part_000.srt
 part_001.mp3
 part_001.srt
 part_002.mp3
 part_002.srt
 part_003.mp3
 part_003.srt
 part_004.mp3
 part_004.srt

(所有片段皆已生成對應的 SRT 檔案)

步驟五：智慧合併 (Final Merge)

當所有片段處理完畢，系統會執行「智慧合併（修正時間軸/重疊）」。此步驟會將分散的字幕檔組合成一個與原影片同名的完整 SRT 檔案，並自動校正時間序列。



```
168
00:26:09,634 --> 00:26:10,934
真是的 你相信她的話嗎

169
00:26:12,374 --> 00:26:13,674
不管了 我肚子餓了

170
00:26:14,374 --> 00:26:15,604
晚自習結束後去便利商店怎麼樣

171
00:26:16,044 --> 00:26:17,274
我要幫媽媽跑腿

172
00:27:00,484 --> 00:27:01,584
請問有什麼事嗎

173
00:27:14,694 --> 00:27:15,994
請不要亂碰東西

174
00:27:57,264 --> 00:27:58,194
下次你得自己撐傘

175
00:27:58,564 --> 00:27:59,164
啊 不要

176
00:27:59,764 --> 00:28:00,464
你有良心嗎

177
00:28:01,434 --> 00:28:02,734
喂 我會淋到雨啦 過來一點

178
00:28:03,234 --> 00:28:04,234
不是啦

179
00:42:05,382 --> 00:42:06,212
你好

180
00:42:17,351 --> 00:42:18,651
媽媽 我回來了
```

(最終產出的完整 SRT 字幕檔)

4. 資源統計與紀錄

軟體底部的執行紀錄 (Log) 視窗提供了完整的任務報告。在處理結束後，您可以看到本次任務的資源消耗統計，這對於管理 API 額度非常有幫助。

統計資訊包含：

- **Token 總用量**：本次任務消耗的總 Token 數。
- **輸入 (Input)**：上傳音訊所佔用的 Token。

- **輸出 (Output)**：AI 生成文字所佔用的 Token。

```
== 開始處理: 照明商店.E01.HD1080p.mp4 == 時間: 2025/12/12 上午 11:48:30 機型: gemini-2.5-pro[11:48:30] [1/4] 提取與分割音訊 (1/4)
[11:49:06] -> 準備好 5 個音訊片段
[11:49:06] -> 上傳片段 1/5...
[11:49:12] -> 生成字幕...
[11:49:54] -> 完成 (已暫存).Token: In=20044, Out=1754
[11:49:55] ... 冷卻中 (等待 10 秒) ...
[11:50:06] -> 上傳片段 2/5...
[11:50:10] -> 生成字幕...
[11:51:16] -> 完成 (已暫存).Token: In=20044, Out=2361
[11:51:18] ... 冷卻中 (等待 10 秒) ...
[11:51:28] -> 上傳片段 3/5...
[11:51:33] -> 生成字幕...
[11:52:40] -> 完成 (已暫存).Token: In=20044, Out=3372
[11:52:42] ... 冷卻中 (等待 10 秒) ...
[11:52:52] -> 上傳片段 4/5...
[11:52:57] -> 生成字幕...
[11:53:56] (回傳為空，視為靜音)
[11:53:56] -> 完成 (已暫存).Token: In=20044, Out=0
[11:53:58] ... 冷卻中 (等待 10 秒) ...
[11:54:08] -> 上傳片段 5/5...
[11:54:12] -> 生成字幕...
[11:54:23] -> 完成 (已暫存).Token: In=7882, Out=52
[11:54:24] [4/4] 智慧合併 (修正時間軸/重疊)...
[11:54:25] 成功儲存: E:\多媒體\照明商店.E01.HD1080p.srt
[11:54:25] 資源使用***Token 總用量**: 95597 ***輸入 Token**: 88058 ***輸出 Token**: 7539
[11:54:25] 提示: 暫存檔保留以供中斷續傳，如無需使用請手動刪除 temp 資料夾
```

(任務完成後的詳細資源使用報告)



(成果展示)