**MVP 功能名稱：AI 影像描述器**

**核心概念：** 我們讓 AI「看」一張從手機鏡頭拍下來的照片，然後「說」出它看到了什麼。

**Input (輸入)**

**輸入有兩個部分：**

1. **一張圖片 (An Image):**
   * **來源：** 使用者透過 App，用手機相機拍下的一張照片。
   * **格式：** 一個 JPEG 或 PNG 圖片檔案。
2. **一句指令 (A Text Prompt):**
   * **來源：** 我們在程式碼中預先設定好，用來告訴 AI 要做什麼。
   * **範例：** 「請詳細描述這張圖片裡的內容。」

**簡單來說，我們的 Input 就是：「一張圖片」+「一句話指令」。**

**Output (輸出)**

**輸出只有一個部分：**

1. **一段文字 (A String of Text):**
   * **來源：** AI 模型 (例如 Google Gemini) 看完圖片和指令後，生成的回應。
   * **格式：** 一段純文字。

**簡單來說，我們的 Output 就是：「AI 的文字回答」。**

**舉一個您專案的實際例子**

假設我們來實現您專案最核心的「掃描書本封面」功能：

* **使用者操作：**
  + 打開您的 App，將相機對準《小王子》這本書的封面，然後按下「分析」按鈕。
* **系統內部的 Input/Output：**
  + **Input (輸入):**
    1. [小王子封面的圖片檔案.jpg]
    2. 「這本書的書名是什麼？」
  + **Output (輸出):**
    1. 「根據圖片，這本書的書名是《小王子》。」