Springboard OCR Comparison - Readme

- **Overview**:
  This project is a comparison of two popular Optical Character Recognition (OCR) libraries: Tesseract and EasyOCR. The goal is to create a user interface that allows users to upload an image and see the OCR results from both libraries side-by-side, along with the confidence scores for the identified text.

- **Development Process:**

  1. **Install Required Libraries**
  The first step is to install the necessary libraries for the project. The code snippet provided installs the following libraries:

  - **tesseract-ocr:** The Tesseract OCR engine
  - **pytesseract:** A Python wrapper for Google's Tesseract-OCR Engine
  - **easyocr:** A deep learning-based OCR library
  - **opencv-python**: A popular computer vision library
  - **matplotlib:** A plotting library
  - **gradio:** A library for creating interactive web interfaces

  2. **Implement Tesseract OCR**
  The ocr_with_tesseract function takes an input image, converts it to RGB format, and then uses the Tesseract OCR engine to detect and extract text from the image. The function also draws rectangles around the detected words and collects the identified words and their confidence scores.

  3. **Implement EasyOCR**
  The ocr_with_easyocr function performs a similar task, but it uses the EasyOCR library instead of Tesseract. It converts the input image to RGB format, initializes the EasyOCR reader, performs the OCR, draws rectangles around the detected words, and collects the identified words and their confidence scores.

  4. **Compare OCR Results**
  The compare_ocr function takes the input image and runs both the Tesseract and EasyOCR OCR functions. It then returns the output images, the identified words, and the confidence scores for each library, which are then displayed in the user interface.

Saladi V S S Siva Hemanth Kumar , Hemanthsaladi2004@gmail.com

5. **Create the User Interface**

The final step is to create the user interface using the Gradio library. The gr.Interface function is used to define the inputs (an image), the outputs (the Tesseract and EasyOCR results), and the title and description of the application.

- ## Conclusion

This project demonstrates how to create a simple yet effective OCR comparison tool using Tesseract and EasyOCR. The user interface allows users to upload an image and see the OCR results from both libraries side-by-side, along with the confidence scores for the identified text. This can be useful for evaluating the performance of different OCR engines and choosing the one that best suits the user's needs.