```
!pip install paddlepaddle
!pip install paddleocr
```

```
Collecting paddlepaddle
  Downloading paddlepaddle-2.6.2-cp310-cp310-manylinux1_x86_64.whl.metadata (8.6 kB)
Requirement already satisfied: httpx in /usr/local/lib/python3.10/dist-packages (from paddlepaddle) (0.28.1)
Requirement already satisfied: numpy>=1.13 in /usr/local/lib/python3.10/dist-packages (from paddlepaddle) (1.26.4)
Requirement already satisfied: Pillow in /usr/local/lib/python3.10/dist-packages (from paddlepaddle) (11.0.0)
Requirement already satisfied: decorator in /usr/local/lib/python3.10/dist-packages (from paddlepaddle) (4.4.2)
Collecting astor (from paddlepaddle)
  Downloading astor-0.8.1-py2.py3-none-any.whl.metadata (4.2 kB)
Collecting opt-einsum==3.3.0 (from paddlepaddle)
  Downloading opt_einsum-3.3.0-py3-none-any.whl.metadata (6.5 kB)
Requirement already satisfied: protobuf>=3.20.2 in /usr/local/lib/python3.10/dist-packages (from paddlepaddle) (4.25.5)
Requirement already satisfied: anyio in /usr/local/lib/python3.10/dist-packages (from httpx->paddlepaddle) (3.7.1)
Requirement already satisfied: certifi in /usr/local/lib/python3.10/dist-packages (from httpx->paddlepaddle) (2024.12.14)
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.10/dist-packages (from httpx->paddlepaddle) (1.0.7)
Requirement already satisfied: idna in /usr/local/lib/python3.10/dist-packages (from httpx->paddlepaddle) (3.10)
Requirement already satisfied: h11<0.15,>=0.13 in /usr/local/lib/python3.10/dist-packages (from httpcore==1.*->httpx->paddlepaddl
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.10/dist-packages (from anyio->httpx->paddlepaddle) (1.3.1)
Requirement already satisfied: exceptiongroup in /usr/local/lib/python3.10/dist-packages (from anyio->httpx->paddlepaddle) (1.2.2
Downloading paddlepaddle-2.6.2-cp310-cp310-manylinux1_x86_64.whl (126.0 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 126.0/126.0 MB 7.4 MB/s eta 0:00:00
Downloading opt_einsum-3.3.0-py3-none-any.whl (65 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 65.5/65.5 kB 6.2 MB/s eta 0:00:00
Downloading astor-0.8.1-py2.py3-none-any.whl (27 kB)
Installing collected packages: opt-einsum, astor, paddlepaddle
  Attempting uninstall: opt-einsum
    Found existing installation: opt_einsum 3.4.0
    Uninstalling opt_einsum-3.4.0:
      Successfully uninstalled opt_einsum-3.4.0
Successfully installed astor-0.8.1 opt-einsum-3.3.0 paddlepaddle-2.6.2
Collecting paddleocr
  Downloading paddleocr-2.9.1-py3-none-any.whl.metadata (8.5 kB)
Requirement already satisfied: shapely in /usr/local/lib/python3.10/dist-packages (from paddleocr) (2.0.6)
Requirement already satisfied: scikit-image in /usr/local/lib/python3.10/dist-packages (from paddleocr) (0.25.0)
Requirement already satisfied: imgaug in /usr/local/lib/python3.10/dist-packages (from paddleocr) (0.4.0)
Collecting pyclipper (from paddleocr)
  Downloading pyclipper-1.3.0.post6-cp310-cp310-manylinux_2_12_x86_64.manylinux2010_x86_64.whl.metadata (9.0 kB)
Collecting lmdb (from paddleocr)
  Downloading lmdb-1.6.2-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (1.1 kB)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from paddleocr) (4.67.1)
Requirement already satisfied: numpy<2.0 in /usr/local/lib/python3.10/dist-packages (from paddleocr) (1.26.4)
Requirement already satisfied: rapidfuzz in /usr/local/lib/python3.10/dist-packages (from paddleocr) (3.11.0)
Requirement already satisfied: opencv-python in /usr/local/lib/python3.10/dist-packages (from paddleocr) (4.10.0.84)
Requirement already satisfied: opencv-contrib-python in /usr/local/lib/python3.10/dist-packages (from paddleocr) (4.10.0.84)
Requirement already satisfied: cython in /usr/local/lib/python3.10/dist-packages (from paddleocr) (3.0.11)
Requirement already satisfied: Pillow in /usr/local/lib/python3.10/dist-packages (from paddleocr) (11.0.0)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.10/dist-packages (from paddleocr) (6.0.2)
Collecting python-docx (from paddleocr)
  Downloading python_docx-1.1.2-py3-none-any.whl.metadata (2.0 kB)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.10/dist-packages (from paddleocr) (4.12.3)
Requirement already satisfied: fonttools>=4.24.0 in /usr/local/lib/python3.10/dist-packages (from paddleocr) (4.55.3)
Collecting fire>=0.3.0 (from paddleocr)
  Downloading fire-0.7.0.tar.gz (87 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 87.2/87.2 kB 3.3 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from paddleocr) (2.32.3)
Collecting albumentations==1.4.10 (from paddleocr)
  Downloading albumentations-1.4.10-py3-none-any.whl.metadata (38 kB)
```

```
from paddleocr import PaddleOCR, draw_ocr
import cv2
from PIL import Image

# Initialize the OCR model
ocr_model = PaddleOCR(use_angle_cls=True, lang='en')  # Set 'lang' to 'en' for English language

# Function to recognize handwritten text from an image
def recognize_handwritten_text(image_path):
    # Read the image
    image = cv2.imread(image_path)

    # Perform OCR on the image
    result = ocr_model.ocr(image, cls=True)

    # Extract recognized text
    recognized_text = ''
    for line in result[0]:
        recognized_text += line[1][0] + '\n'

    return recognized_text
```

```
# Example usage: recognize text from a single image
image_path = "/content/8255.jpg"  # Replace with your image path
recognized_text = recognize_handwritten_text(image_path)
print("Recognized Text:\n", recognized_text)
```

```
[2025/01/06 08:37:33] ppocr DEBUG: Namespace(help='==SUPPRESS==', use_gpu=False, use_xpu=False, use_npu=False, use_mlu=False, ir_opt
[2025/01/06 08:37:34] ppocr DEBUG: dt_boxes num : 9, elapsed : 0.27484846115112305
[2025/01/06 08:37:35] ppocr DEBUG: cls num  : 9, elapsed : 0.05196046829223633
[2025/01/06 08:37:35] ppocr DEBUG: rec_res num  : 9, elapsed : 0.6228659152984619
Recognized Text:
 21339
3174g Buhler is aPo9t of
31750 Builders Club was
quickly recoynised thet
it cas better than ocder
Cements, althouyh more
expensive.31752 Bcilders
Sterted to bceild the
bunker in 1942.
```

```
!pip install jiwer
```

```
Collecting jiwer
  Downloading jiwer-3.0.5-py3-none-any.whl.metadata (2.7 kB)
Requirement already satisfied: click<9.0.0,>=8.1.3 in /usr/local/lib/python3.10/dist-packages (from jiwer) (8.1.7)
Collecting rapidfuzz<4,>=3 (from jiwer)
  Downloading rapidfuzz-3.11.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (11 kB)
Downloading jiwer-3.0.5-py3-none-any.whl (21 kB)
Downloading rapidfuzz-3.11.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.1 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 3.1/3.1 MB 36.8 MB/s eta 0:00:00
Installing collected packages: rapidfuzz, jiwer
Successfully installed jiwer-3.0.5 rapidfuzz-3.11.0
```

```
#Paddle OCR
from jiwer import cer, wer

# Ground truth and recognized text samples for evaluation
ground_truth_texts = [
   "31749 Buhler is a part of Buhler USD 313 school district. 31750 Builders Club was first membership club which could only be joined t
]

# Texts recognized by the HTR model
recognized_texts = [

"3174g Buhler is aPo9t of 31750 Builders Club was quickly recoynised thet it cas better than ocder Cements, althouyh more expensive.3175
]

# Calculate CER and WER for each sample
for gt_text, rec_text in zip(ground_truth_texts, recognized_texts):
    current_cer = cer(gt_text, rec_text)
    current_wer = wer(gt_text, rec_text)
    print(f"Ground Truth: {gt_text}")
    print(f"Recognized Text: {rec_text}")
    print(f"CER: {current_cer:.2f}, WER: {current_wer:.2f}")
    print("="*50)

# Calculate average CER and WER across all samples
average_cer = sum(cer(gt, rec) for gt, rec in zip(ground_truth_texts, recognized_texts)) / len(ground_truth_texts)
average_wer = sum(wer(gt, rec) for gt, rec in zip(ground_truth_texts, recognized_texts)) / len(ground_truth_texts)

print(f"Average CER: {average_cer:.2f}")
print(f"Average WER: {average_wer:.2f}")
```

```
Ground Truth: 31749 Buhler is a part of Buhler USD 313 school district. 31750 Builders Club was first membership club which could or
Recognized Text: 3174g Buhler is aPo9t of 31750 Builders Club was quickly recoynised thet it cas better than ocder Cements, althouyh
CER: 0.44, WER: 0.65
==================================================
Average CER: 0.44
Average WER: 0.65
```

```
# ollama OCR
from jiwer import cer, wer

# Ground truth and recognized text samples for evaluation
ground_truth_texts = [
   "31749 Buhler is a part of Buhler USD 313 school district. 31750 Builders Club was first membership club which could only be joined t
```

```
]

# Texts recognized by the HTR model
recognized_texts = [

"31749 Buhler is a part of Buhler USD 313 School district. 31750 Builders Club was first membership club which could only be joined thro
]

# Calculate CER and WER for each sample
for gt_text, rec_text in zip(ground_truth_texts, recognized_texts):
    current_cer = cer(gt_text, rec_text)
    current_wer = wer(gt_text, rec_text)
    print(f"Ground Truth: {gt_text}")
    print(f"Recognized Text: {rec_text}")
    print(f"CER: {current_cer:.2f}, WER: {current_wer:.2f}")
    print("="*50)

# Calculate average CER and WER across all samples
average_cer = sum(cer(gt, rec) for gt, rec in zip(ground_truth_texts, recognized_texts)) / len(ground_truth_texts)
average_wer = sum(wer(gt, rec) for gt, rec in zip(ground_truth_texts, recognized_texts)) / len(ground_truth_texts)

print(f"Average CER: {average_cer:.2f}")
print(f"Average WER: {average_wer:.2f}")
```

```
Ground Truth: 31749 Buhler is a part of Buhler USD 313 school district. 31750 Builders Club was first membership club which could or
Recognized Text: 31749 Buhler is a part of Buhler USD 313 School district. 31750 Builders Club was first membership club which could
CER: 0.01, WER: 0.06
==================================================
Average CER: 0.01
Average WER: 0.06
```