

OCRテスト用ドキュメント

作成日: 2024年12月25日

1. テキスト認識テスト

日本語テキスト: これは日本語の文章です。漢字、ひらがな、カタカナが含まれています。

English Text: This is an English sentence for OCR testing purposes.

2. 表（テーブル）認識テスト

基本的な表:

項目	説明	価格
商品A	高品質な製品	¥1,200
商品B	スタンダード製品	¥800
商品C	エコノミー製品	¥500

複雑な表（数値データ）：

年度	Q1	Q2	Q3	Q4	合計
2022	1,234	2,345	3,456	4,567	11,602
2023	1,500	2,800	3,200	5,100	12,600
2024	1,800	3,100	4,200	5,500	14,600

3. リスト認識テスト

- 項目1: 最初の項目
- 項目2: 2番目の項目
- 項目3: 3番目の項目

1. 手順1: 準備をする

2. 手順2: 実行する

3. 手順3: 確認する

4. コードブロック認識テスト

```
def hello_world():
    print("Hello, World!")
    return True
```

```
for i in range(10):
    print(i)
```

5. 数式認識テスト

$$E = mc^2$$

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$$

6. エンジン比較表

エンジン	OCR	表認識	速度
MarkItDown	×	△	◎
Docling (CPU)	○	◎	○
Docling (GPU)	○	◎	◎

テストドキュメント終了