

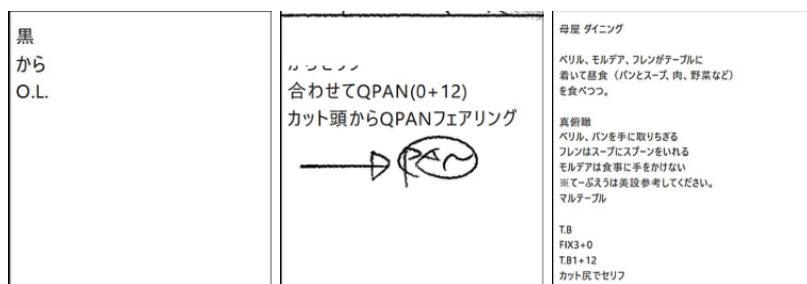
報告書

1 今週の進捗

- GPT API の料金調査
- Gemini との結果比較

2 GPT, Gemini への入力

適当に選んだ 10 ページで OCR を試した.



(a) page002 ("黒"を含む) (b) page007 (手書きの記号) (c) page018 (小さめの文字)



(d) page024 (形式ずれ) (e) page053 (手書き文字) (f) page060 (手書きベースト)



(g) page061 (手書き文字と記号) (h) page082 (イラスト) (i) page094 (文章量少なめ) (j) page105 (量多め)

2.1 GPT API の料金調査

model	料金 (input)	料金 (output)
gpt-4o-mini	0.15	0.60
gpt-4.1-mini	0.40	1.60
gpt-4.1-nano	0.10	0.40

model	料金 (表示)	料金 (計算)
gpt-4o-mini	0.13	0.130472
gpt-4.1-mini	0.01	0.013099
gpt-4.1-nano	0.01	0.003822

2.2 OCR 結果

比較項目	gpt-4o-mini	gpt-4.1-mini	gpt-4.1-nano
page002_row1	黒@から@O.L.	黒@から@O.L.	黒@から@O.L.
page007_row2	ハンセン@ 合わせて QPAN(0+12)@ カット頭から QPAN フェアリング@ →@PAC	はととつ 合わせて QPAN(0+12) カット頭から QPAN フェアリング	合せて QPAN(0+12)@ カット頭から QPAN フェアリング

3 今後の課題

- "gpt-4.1" で間違えたものを試す
- "gemma-3" モデルのプログラムコード経由での実行における動作の安定化および精度改善
- Mixture of Experts (MoE) アーキテクチャを採用した OCR モデルに関する調査

参考文献