Comprehensive OCR Model Performance Report

This document provides a structured and in-depth evaluation of multiple OCR models tested for handwritten and printed text recognition. The analysis is based on extensive testing performed in two key documents:

1. Handwritten.ipynb - Focused on various OCR techniques for recognizing both handwritten and printed text.
2. OllamaOCR.pdf - Evaluated fine-tuned vision-language models, including LLaMA 3.2, Qwen-VLM, and custom approaches.

1. OCR Model Evaluations

This section presents the OCR models tested, their intended roles, and their observed performance in extracting printed and handwritten text. Models were assessed based on **accuracy, robustness, and computational efficiency**. The table below provides a structured comparison.

Model Name	OCR Type	Performance Overview	Strengths	Limitations
Tesseract OCR	Traditional OCR	Effective for printed text, weak on handwriting	High accuracy for printed text	Poor handwriting recognition
TrOCR	Transformer-based	Performs well on structured handwriting	Good for clean handwriting	Struggles with cursive and noisy input
YOLOv8 + PaddleOCR + TrOCR	Hybrid	Good text detection but OCR inconsistencies	Strong text detection	OCR model needs improvement
EasyOCR + TrOCR	Hybrid	Performs well on printed text, inconsistent handwriting	Handles multi-lingual text	Handwriting accuracy is variable
Improved Text Extractor	Custom Pipeline	Best overall accuracy but computationally expensive	Superior hybrid OCR accuracy	Slow processing
PaliGemma	Vision-Language Model	Failed to extract meaningful text	Optimized for multimodal tasks	Not designed for OCR
LLaMA 3.2 Vision	Fine-Tuned LLaMA	OCR performance unknown (CPU issue)	Strong for structured data extraction	Needs GPU for testing
Qwen2-VL-7B -	Multimodal	OCR	Potential for	Requires better

VLM	Model	performance unknown (CPU issue)	OCR tasks	hardware
Meta-LLaMA 7B	LLaMA-based	Failed due to authentication issues	Advanced NLP capabilities	Not tested on this system
Pytesseract + Preprocessing	Traditional OCR + Enhancement	Enhances printed OCR, moderate handwriting	Best for printed text	Handwriting errors persist
Ollama OCR (Fine-Tuned LLaMA 3.2)	Fine-Tuned LLaMA	Improved handwritten text extraction but still inconsistent	Custom fine-tuning showed promise	Needs more refinements

3. Key Findings & Recommendations

Based on extensive evaluations, the following conclusions and recommendations are made:
1. Tesseract OCR remains the best for printed text but is not suitable for handwriting.
2. Fine-tuned LLaMA 3.2 and hybrid approaches like EasyOCR + TrOCR show potential for handwritten OCR but need refinement.
3. The Improved Text Extractor provided the best overall results but is computationally expensive.
4. Some models like Qwen-VLM and Meta-LLaMA 7B could not be fully evaluated due to hardware limitations.

WILL WORK ON:
Fine-tune transformer-based models (like LLaMA 3.2) further for handwritten OCR.
Combine traditional and deep-learning OCR models to create a balanced hybrid solution.

OPENAI- need to try

GITHUB LINKS FOR BOTH THE .IPYNB WITH ALL THE MODELS TRIED AND RESULTS

github_link