PROJECT ABSTRACT

Handwritten Text Recognition using CRNN

CRNN model that transcribes handwritten English text lines from scanned images — evaluated with CER/WER on the IAM Handwriting Database.

01 Title

Handwritten Text Recognition using CRNN — an AI/ML software project that transcribes handwritten English text lines from scanned images using a Convolutional Recurrent Neural Network with CTC decoding.

02 Introduction / Background

Handwriting remains one of the hardest problems in optical character recognition: every writer's slant, stroke width and letter joins differ, so off-the-shelf print OCR collapses on handwritten notes. The CRNN (Convolutional Recurrent Neural Network), introduced by Shi et al., has become a standard approach for sequence-like text recognition: a CNN extracts visual features, a recurrent layer models sequence context, and CTC decoding produces the character string without character-level alignment labels.

03 Problem Statement

Digitizing handwritten notes, answer scripts, forms and historical documents requires transcription that is accurate, fast and reproducible. Manual re-typing is slow and expensive; generic print OCR fails on handwriting; and students need an end-to-end deep-learning project with a measurable, demoable outcome — exactly what this build delivers.

04 Objectives

- Build a CRNN model that transcribes handwritten English text lines from scanned images.
- Train on the IAM Handwriting Database (13,353 labeled text lines, line-level task).
- Evaluate with Character Error Rate (CER) and Word Error Rate (WER) computed by the training notebook on the validation split.
- Ship a web demo that uploads a scan and returns line-by-line transcription with confidence scores.
- Deliver the complete viva kit: report, PPT and Q&A preparation document.

05 Proposed System / Working

The pipeline deskews and binarizes the scanned page, segments it into text-line images, and normalizes each line to a fixed height. A CNN encoder converts each line into a feature sequence; a bidirectional LSTM reads it with full left-right context; CTC decoding converts frame predictions into the character string without alignment labels. The web app displays the transcription with per-line confidence, and the training notebook logs CTC loss curves plus CER/WER on the validation split.

06 Technologies / Components Used

- Python 3.10, PyTorch (CRNN training, CTC loss)
- OpenCV and Pillow (preprocessing, line segmentation)
- Flask or Streamlit web app (live transcription demo)
- IAM Handwriting Database — 13,353 handwritten text lines
- NumPy, Matplotlib (CER/WER evaluation, loss curves)

07 Key Features

- Line-level transcription from scanned or photographed pages
- CNN + BiLSTM + CTC decoding architecture
- Per-line confidence scores in the web demo
- CER/WER evaluation and error analysis in the training notebook
- Line-segmentation utility for full pages
- Augmentation (slant, stretch, noise) for writer robustness
- CPU-friendly inference — under one second per line

08 Applications / Use Cases

- Digitizing handwritten notes, answer scripts and forms
- Archival document transcription
- Accessibility tools for reading handwritten content
- Document-AI pipelines needing a handwriting stage

09 Results & Scope

The included training notebook computes Character Error Rate and Word Error Rate on the IAM validation split and logs CTC loss curves during the build; design targets are CER below 10% and WER below 30% on the validation split. No performance number is pre-claimed — the deliverable is the working model, the evaluation code and the reproducible measurements it produces. Scope covers English line-level handwriting; cursive extremes, other scripts and handwritten paragraphs without clean line segmentation are stated limitations.

10 Conclusion

This project delivers a complete, demoable deep-learning OCR system — from the IAM dataset through CRNN training and CTC decoding to a live web demo with quantified CER/WER evaluation — plus the full documentation kit expected of a final-year submission.

11 References

IAM Handwriting Database (Marti & Bunke, Univ. of Bern) — 13,353 labeled text lines; Shi, Bai & Yao, 'An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition', arXiv:1507.05717 (CRNN); Graves et al., 'Connectionist Temporal Classification' (CTC loss); PyTorch documentation.