

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AI Powered Handwritten Notes Digitization Using NLP and Deep Learning

Malini M*1, Jones Geneva D*2, Priya Dharshini S*3, Rakshani D*4

*1 Sri Shakthi Institute Of Engineering & Technology,(Anna University Affiliated),Coimbatore, Tamil Nadu, India

*2 Sri Shakthi Institute Of Engineering & Technology,(Anna University Affiliated),Coimbatore, Tamil Nadu, India

*3 Sri Shakthi Institute Of Engineering & Technology,(Anna University Affiliated),Coimbatore, Tamil Nadu, India

*4 Sri Shakthi Institute Of Engineering & Technology,(Anna University Affiliated),Coimbatore, Tamil Nadu, India

ABSTRACT

This paper introduces an AI-based system for digitizing handwritten text from notes through image-based text recognition and real-time NLP-based correction. The system takes handwritten images from a camera or device, extracts text through Easy OCR, corrects spelling through SymSpell, and corrects grammar through transformer-based NLP models. A user interface built with Streamlit supports image upload, shows output text in real-time, and supports downloading the final corrected document. The suggested system decreases the amount of manual transcription work and guarantees precision and readability in transcribed notes, thus making it ideal for students and educational institutions.

KEYWORDS : OCR, NLP, Grammar Correction, EasyOCR, SymSpell, Deep Learning, Streamlit

1. INTRODUCTION

Handwritten note-taking remains a common practice in educational and professional environments. Nevertheless, organizing and searching handwritten notes is tedious. This research solves the problem of converting handwritten notes into neat, editable text with the help of AI processes. The system employs OCR-based extraction, NLP-based correction, and unifies all modules through an interactive UI. It makes digitization easy without having to depend on manual typing or editing and is compatible with device-based uploads and real-time camera input.

2. METHODOLOGY

The system to be proposed has four primary modules:

OCR Module:

Utilizes EasyOCR to pull raw text from handwritten images.

Spell Correction:

SymSpell corrects simple spelling mistakes based on a frequency dictionary.

Grammar Correction:

Transformer-based model (prithivida/grammar_error_correcter_v1) and spaCy examine and correct contextual grammar mistakes.

Frontend Integration:

Streamlit offers an interface for input (upload/camera), output preview, and text download.

All modules are written in Python and integrated into one pipeline (app.py) to provide efficient real- time processing.

3. MODELING AND ANALYSIS

The system employs:

EasyOCR (Deep learning model for image text detection) SymSpell (Edit-distance based spell correction)

HuggingFace Transformers (Grammar correction with pretrained models)- spaCy (POS tagging and subject-verb error detection)

A red-black themed Streamlit UI supports image input and text output.

4. RESULTS AND DISCUSSION SCREENSHOTS:

	Deploy 1
Al Handwritten Notes Digitizer •	
Choose Input Type:	
Upload from Device Control from Comman	
Lapore mon carriera	
Drag and drop file here Lost 2008 core file. DO: UTC: 240.	Browse files

2.



3.



4.



5.



6.

File	Edit	View			
Proje	ct is	completed.			

System was also tested on several handwritten samples.

Example:

Input Image Text: "Project is completed" OCR Output: "psoject 8s completed"

After Spell Correction: "project is completed" After Grammar Correction: "Project is completed."

The output was very accurate and efficient. Users enjoyed the feature of the uploading/capturing and receiving downloadable clean text immediately. NLP usage guaranteed sentence meaning retention while correcting mistakes.

CONCLUSION

This paper puts forward a simple, AI-driven solution for handwritten notes digitization. It prevents the need for manual transcription, identifies and remedies spelling and grammatical errors, and produces downloadable clean text employing deep learning and NLP technologies. The platform offers real-time input choices and seamless UI integration, making it ideal for both students and instructors. Future updates include multilingual support, export to PDF, and voice-based output.

REFERENCES

- 1. Prithiviraj Damodaran, "Grammar Error Correction Model", HuggingFace, 2022.
- 2. EasyOCR: Jaided AI, GitHub, 2023.
- 3. SymSpell: Wolf Garbe, GitHub, 2022.
- 4. spaCy NLP Toolkit: Explosion AI, 2022.
- 5. Streamlit Docs streamlit.io, 2024.