

Handwritten to Text Converter using CNN

Prof. Komal Rathod*, Sanika Jori**, Pallavi Chavan***, Vaishnavi Chikane****, Divya Deshmukh*****
Information Technology ,
Zeal College of Engineering And Research Pune,
Pune -India

ABSTRACT

Handwritten text recognition has become an important research area in the field of computer vision and artificial intelligence due to the increasing need for digitizing handwritten documents. Traditional Optical Character Recognition (OCR) techniques often face challenges in accurately recognizing handwritten text because of variations in writing styles, character shapes, and document quality. To address these limitations, this paper proposes a *Handwritten-to-Text Conversion System using Convolutional Neural Networks (CNNs)* for efficient and accurate recognition of handwritten characters and words.

The proposed system employs image preprocessing techniques such as noise removal, grayscale conversion, normalization, and segmentation to enhance the quality of input handwritten images. A CNN-based deep learning model is utilized to automatically extract relevant features and classify handwritten characters with high precision. The recognized characters are then assembled to generate meaningful digital text output. The system is implemented using Python-based deep learning frameworks and integrated with a user-friendly interface for uploading and processing handwritten documents.

The results indicate that CNN-based handwritten text recognition provides a reliable, scalable, and cost-effective solution for automated document digitization and contributes to the advancement of intelligent handwriting recognition systems.

Keywords:

Handwritten Text Recognition, Convolutional Neural Network, Deep Learning, Optical Character Recognition, Image Processing, Feature Extraction, Document Digitization.

I. INTRODUCTION

The rapid advancement of digital technologies has significantly increased the demand for efficient methods of storing, processing, and retrieving information. Despite the widespread use of digital devices, handwritten documents continue to play an important role in various domains such as education, healthcare, banking, administration, and historical record preservation. Large volumes of handwritten notes, forms, applications, medical prescriptions, and archival documents are still generated daily. However, managing and processing such documents manually is time-consuming, labor-intensive, and prone to human errors.

Handwritten Text Recognition (HTR) is a challenging task in the field of Artificial Intelligence (AI), Machine Learning (ML), and Computer Vision. Unlike printed text, handwritten text exhibits significant variations in writing styles, character sizes, orientations, spacing, and stroke patterns. These variations make accurate recognition difficult for conventional Optical Character Recognition (OCR) systems, which are primarily designed for machine-printed text. Consequently, there is a growing need for intelligent systems capable of accurately recognizing and converting handwritten content into machine-readable digital text.

Recent developments in Deep Learning have revolutionized image recognition and pattern analysis tasks. Among various deep learning techniques, Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in extracting complex visual features from images and accurately classifying handwritten characters. CNNs automatically learn hierarchical feature representations, eliminating the need for

manual feature engineering and significantly improving recognition accuracy. Due to these advantages, CNN-based approaches have become one of the most effective solutions for handwritten text recognition.

This paper presents a *Handwritten-to-Text Conversion System using Convolutional Neural Networks (CNNs)* that aims to automatically convert handwritten text into editable and searchable digital text. The proposed system incorporates image preprocessing techniques such as grayscale conversion, noise reduction, normalization, and segmentation to enhance input image quality. A CNN-based recognition model is employed to extract meaningful features from handwritten characters and classify them accurately. The recognized characters are then reconstructed into words and sentences to generate meaningful textual output.

The primary objective of this work is to reduce manual transcription efforts, improve digitization efficiency, and provide an intelligent solution for handwritten document processing. The proposed system can be utilized in various real-world applications, including document digitization, educational content management, medical record processing, archival preservation, and administrative data management. By leveraging deep learning techniques, the system offers improved recognition accuracy, scalability, and adaptability to different handwriting styles.

II. RELATED WORK

Handwritten Text Recognition (HTR) has emerged as a significant research area in the domains of computer vision, pattern recognition, and artificial intelligence due to the increasing need for digitizing handwritten documents. Various

approaches have been proposed to improve the accuracy and efficiency of handwritten text conversion systems. These approaches can be broadly categorized into traditional OCR methods, machine learning techniques, and deep learning-based recognition systems.

Traditional OCR-Based Techniques:

Early handwritten text recognition systems primarily relied on Optical Character Recognition (OCR) techniques combined with image processing methods. These systems employed template matching, character segmentation, and handcrafted feature extraction to recognize individual characters. While traditional OCR systems achieved satisfactory results for printed text, their performance was limited when applied to handwritten documents because of variations in writing styles, character sizes, orientations, and spacing. Consequently, recognition accuracy often decreased when processing unconstrained handwritten text.

Feature Extraction and Pattern Recognition Methods:

To improve recognition performance, researchers introduced feature-based approaches that extract structural and statistical characteristics from handwritten characters. Commonly used features include edges, strokes, loops, intersections, contours, and geometric properties. Techniques such as Histogram of Oriented Gradients (HOG), zoning, projection profiles, and contour analysis have been widely utilized for feature extraction. These features are subsequently classified using pattern recognition algorithms. Although these methods improved recognition accuracy, their effectiveness largely depended on manual feature engineering and domain expertise.

Machine Learning-Based Recognition Approaches :

Several machine learning algorithms have been applied to handwritten character recognition, including Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Decision Trees, and Random Forest classifiers. These approaches classify handwritten characters based on extracted features and have demonstrated reasonable performance on structured datasets. However, their accuracy is highly dependent on the quality of extracted features and they often struggle to generalize across diverse handwriting styles and large-scale datasets.

Deep Learning Approaches Using Convolutional Neural Networks :

Recent advancements in deep learning have significantly transformed handwritten text recognition systems. Convolutional Neural Networks (CNNs) have become one of the most widely adopted architectures due to their ability to automatically learn hierarchical feature representations directly from image data. CNN-based systems eliminate the need for manual feature extraction and have achieved superior recognition accuracy compared to traditional machine learning methods. Their capability to capture spatial information and complex visual patterns makes them particularly effective for handwritten character recognition tasks.

Hybrid CNN-LSTM Architectures :

To address challenges associated with sequential handwriting recognition, researchers have proposed hybrid architectures that combine Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks. In these systems, CNNs are responsible for extracting visual features, while LSTMs model contextual and sequential relationships among characters. Such architectures have demonstrated improved performance in recognizing cursive handwriting and complete handwritten words. However, they require substantial computational resources and large annotated datasets for effective training.

Therefore, there is a need for an efficient and scalable handwritten text recognition system capable of accurately converting handwritten content into machine-readable text while maintaining computational efficiency. To address these challenges, the proposed work employs a Convolutional Neural Network (CNN)-based architecture integrated with image preprocessing, character segmentation, feature extraction, and text reconstruction techniques. The proposed system aims to enhance recognition accuracy, reduce manual transcription efforts, and provide an effective solution for intelligent document digitization.

III. PROPOSED METHOD

The proposed system presents an intelligent handwritten text recognition framework that integrates Convolutional Neural Networks (CNN) with Natural Language Processing (NLP) techniques to improve the accuracy, readability, and contextual correctness of handwritten text conversion. The complete methodology consists of the following stages:

1. HANDWRITTEN IMAGE ACQUISITION:

The system accepts handwritten document images captured through scanners, mobile cameras, or uploaded image files. This stage serves as the input layer of the framework and provides handwritten samples for further processing.

2. IMAGE PREPROCESSING:

The acquired image undergoes several preprocessing operations to improve image quality and reduce distortions. The preprocessing stage includes:

- Grayscale Conversion
- Noise Removal
- Image Normalization
- Thresholding

These operations enhance character visibility and improve recognition performance.

3. TEXT SEGMENTATION:

The preprocessed image is segmented into smaller textual units to facilitate character recognition. The segmentation process includes:

- Line Segmentation
- Word Segmentation
- Character Segmentation

This stage isolates handwritten characters and prepares them for feature extraction.

4. CNN-BASED FEATURE EXTRACTION :

The segmented characters are passed through a Convolutional Neural Network (CNN) model. The CNN automatically learns and extracts important visual features such as:

- Edges
- Curves
- Loops
- Stroke Patterns
- Character Structures

These features help distinguish handwritten characters with high accuracy.

5. CHARACTER RECOGNITION AND TEXT GENERATION :

The extracted feature maps are processed through convolutional, pooling, and fully connected layers. The CNN classifies individual characters and combines them to generate:

- Recognized Characters.
- Words
- Sentences
- Initial Machine-Readable Text

This stage produces the preliminary text output from the handwritten document.

6. SYSTEM ARCHITECTURE :

The proposed handwritten image-to-text conversion system follows a layered architecture consisting of image acquisition, preprocessing, feature extraction, text recognition, natural language processing, and output generation modules. The architecture integrates deep learning and natural language processing techniques to improve recognition accuracy and generate meaningful textual output.

- Image Acquisition Layer
- Image Preprocessing Layer
- Text Segmentation Layer
- CNN-Based Feature Extraction Layer
- Character Recognition Module
- NLP Processing and Correction Layer
- Language Model Verification Layer
- Output Generation Layer

This modular design ensures scalability and smooth system performance.

A. Image Acquisition Layer: The image acquisition layer is responsible for obtaining handwritten document images from different sources such as scanners, mobile cameras, or uploaded image files. This layer acts as the entry point of the system and provides input data for further processing.

B. Image Processing Layer: The preprocessing layer improves image quality before recognition. Operations such as grayscale conversion, noise removal, thresholding,

normalization, and contrast enhancement are performed to eliminate distortions and improve character visibility. Effective preprocessing significantly improves feature extraction and recognition performance.

C. Test Segmentation Layer: The segmentation layer divides the document into smaller units including lines, words, and characters. Proper segmentation helps isolate handwritten characters and enables the recognition model to process each character accurately.

D. CNN-Based Feature Extraction Layer: This layer employs a Convolutional Neural Network to automatically learn discriminative features from handwritten characters. Convolution and pooling operations extract important visual characteristics such as edges, curves, loops, and stroke patterns. These features are transformed into feature maps that represent character information.

E. Character Recognition Module: The extracted features are supplied to the classification component of the CNN architecture. The model identifies individual characters and combines them to generate words and sentences. The output of this stage is raw machine-readable text obtained directly from handwritten input.

F. NLP Processing and Correction Layer: The generated text may contain recognition errors due to handwriting variations and ambiguities. Therefore, Natural Language Processing techniques are applied to refine the extracted text. This layer performs tokenization, spelling correction, grammatical analysis, and contextual validation to improve textual quality and readability.

G. Language Model Verification Layer: The language model verification layer evaluates sentence structure and semantic consistency. Context-aware prediction techniques identify incorrect words and reconstruct meaningful sentences. This stage improves overall textual coherence and minimizes recognition errors.

H. Output Generation Layer: The final layer generates corrected and structured digital text. The generated output can be stored, edited, searched, or exported for further applications. The integration of CNN and NLP techniques enables the system to produce highly accurate and contextually meaningful text from handwritten documents.

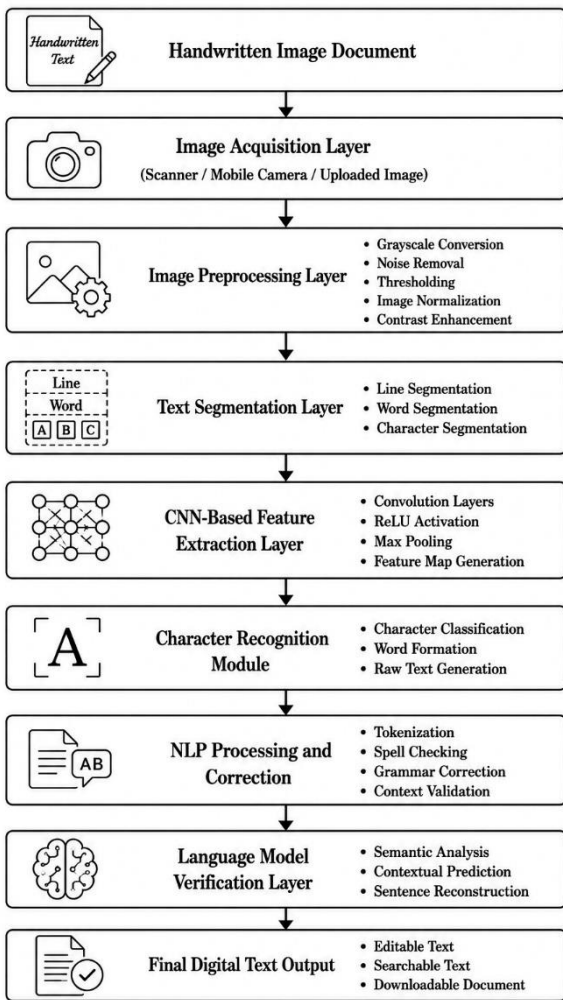


Figure 1 : System Architecture

IV. EXPERIMENTAL SETTINGS

This section describes the experimental setup, implementation environment, training configuration, and performance evaluation criteria used for the proposed handwritten text recognition system.

1. Experimental Environment:

The proposed system was developed using Python as the primary programming language due to its extensive support for machine learning, deep learning, and natural language processing libraries. The CNN model was implemented using TensorFlow and Keras frameworks, while OpenCV was utilized for image preprocessing and segmentation tasks. Natural Language Processing operations including tokenization, spell correction, and contextual analysis were implemented using NLTK and related NLP libraries.

2. Dataset Preparation:

The experimental dataset consists of handwritten character and word images collected from publicly available handwriting datasets and user-generated handwritten samples. The dataset

contains variations in writing styles, character sizes, stroke thicknesses, and image quality to ensure robustness.

The dataset was divided into:

- Training Dataset – 80%
- Validation Dataset – 10%
- Testing Dataset – 10%

This distribution enables effective model learning and unbiased performance evaluation.

3. Image Preprocessing:

Prior to model training, all images were subjected to preprocessing operations including:

- Grayscale Conversion
- Noise Reduction
- Image Normalization
- Thresholding
- Contrast Enhancement
- Image Resizing

These preprocessing steps improve character visibility and facilitate effective feature extraction by the CNN model.

4. CNN Training Configuration:

The CNN architecture consists of multiple convolutional layers, max-pooling layers, activation functions, and fully connected layers. The model was trained using the Adam optimizer with categorical cross-entropy loss function.

Training Parameters:

- Optimizer: Adam
- Learning Rate: 0.001
- Batch Size: 32
- Number of Epochs: 30–50
- Activation Function: ReLU
- Output Function: Softmax

Data augmentation techniques such as rotation, scaling, shifting, and flipping were applied to improve generalization and reduce overfitting.

5. NLP-Based Post-Processing:

After character recognition, the extracted text was processed using NLP techniques. The NLP module performs tokenization, spelling correction, contextual validation, and sentence reconstruction. Language-based correction reduces recognition errors caused by ambiguous handwriting and improves semantic consistency in the generated text.

6. Evaluation Metrics:

The performance of the proposed system was evaluated using standard recognition metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- Character Error Rate (CER)
- Word Error Rate (WER)

Accuracy measures the overall correctness of recognized characters. Precision and recall evaluate recognition quality, while F1-score provides a balanced assessment of both metrics. Character Error Rate and Word Error Rate quantify the effectiveness of text extraction and contextual correction processes.

7. Performance Evaluation:

The experimental analysis focuses on comparing CNN-only recognition results with the proposed CNN-NLP framework. The integration of NLP-based correction is expected to reduce recognition errors, improve sentence-level accuracy, and enhance overall document readability.

The proposed system demonstrates improved robustness against handwriting variations and generates more meaningful textual output than conventional handwriting recognition approaches.

V. IMPLEMENTATION AND USER INTERFACE

The implementation of the Handwritten to Text Converter using Convolutional Neural Network (CNN) focuses on converting handwritten text images into machine-readable digital text. The system is developed using Python and Deep Learning techniques. It combines image processing methods with CNN-based character recognition to achieve accurate handwriting recognition. The complete implementation process consists of several stages including image acquisition, preprocessing, segmentation, feature extraction, character recognition, text generation, and output storage. The proposed system automates the manual process of transcription and provides a reliable solution for document digitization.

A. Home Dashboard

The home dashboard acts as the central control panel of the application. It provides easy access to all major functionalities of the system. Users can upload handwritten images, view conversion history, access settings, and download previously converted documents. The dashboard is designed to provide a simple and user-friendly experience.

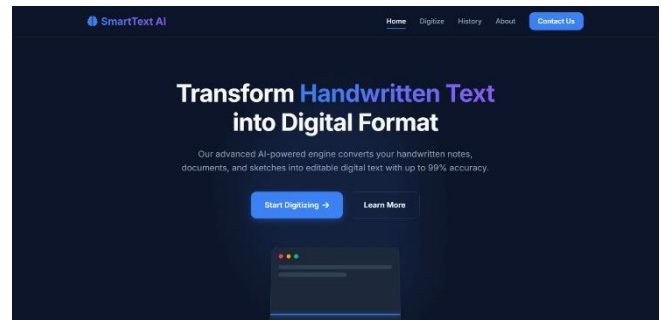


Figure 4: User Interface

B. Image Upload Module

The image upload module allows users to submit handwritten documents for conversion. Users can upload images from their device or capture images using a camera. The system supports multiple image formats such as JPG, JPEG, and PNG. Once the image is selected, a preview is displayed before processing begins.

The following is the general workflow when an image is uploaded and analyzed:

1. User uploads an image
2. Image is pre-processed (resized, normalized etc.)
3. Image is passed into the trained model
4. Detection is performed

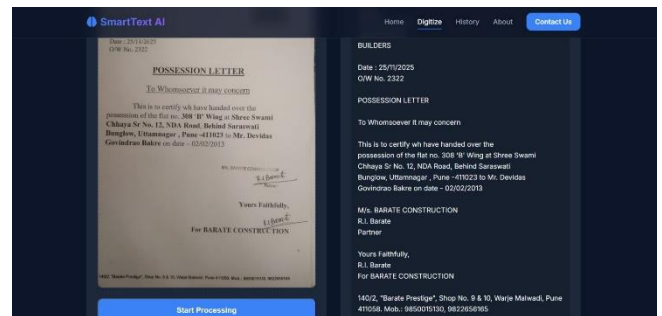


Figure 5: Uploading Images

C. Image Preprocessing Module

After uploading the image, the preprocessing module enhances the image quality. The uploaded image is converted into grayscale format, resized to a standard size, and cleaned using noise removal techniques. Contrast enhancement and normalization are also applied to improve character visibility. These steps help increase the accuracy of handwriting recognition.

number of detected microplastic particles, the average confidence score of the detections and the risk level that has been calculated based on the analysis. The output displayed is presented in a structured table as seen below:

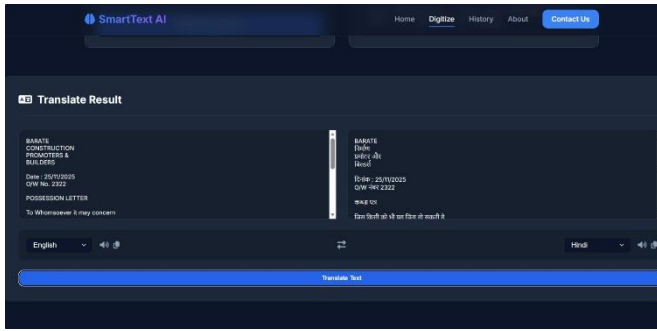


Figure 6: Detection Results

1) D. Responsive User Interface

The entire system is designed with a responsive user interface that works efficiently on desktops, laptops, tablets, and smartphones. The interface provides simple navigation, fast processing, and an improved user experience. The combination of CNN implementation and an easy-to-use interface makes the system suitable for educational, professional, and document digitization applications.

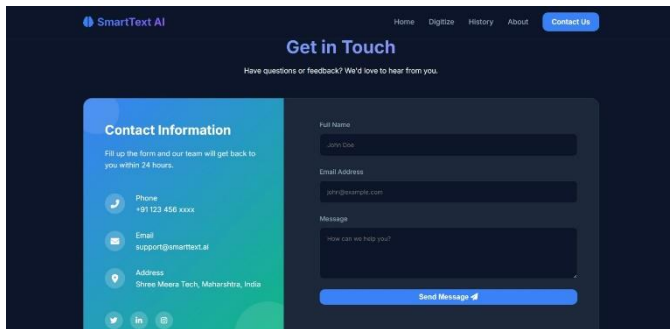


Figure 7: Feedback Overview

E. System Workflow Interface

The System Workflow Interface provides a clear and organized view of the complete handwriting-to-text conversion process.

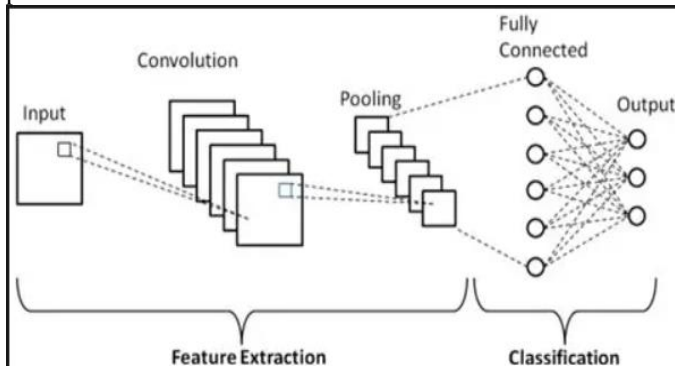


Figure 8: System Workflow

VI. CONCLUSION

The Handwritten to Text Converter using Convolutional Neural Network (CNN) is an effective solution for converting handwritten content into machine-readable digital text. The proposed system successfully combines image processing techniques with deep learning algorithms to achieve accurate handwriting recognition. The implementation of preprocessing, segmentation, feature extraction, classification, and post-processing modules enables efficient conversion of handwritten documents into editable digital text.

The CNN model demonstrates strong performance in recognizing various handwriting styles, character shapes, and writing patterns. The automatic feature extraction capability of CNN eliminates the need for manual feature engineering and significantly improves recognition accuracy. Furthermore, preprocessing techniques such as grayscale conversion, noise removal, image enhancement, and normalization contribute to better recognition results.

The developed system reduces the time and effort required for manual transcription while increasing accuracy and productivity. The user-friendly interface makes the application accessible to students, teachers, researchers, business professionals, and organizations involved in document management. The ability to store, edit, search, and share converted text further enhances the usefulness of the system.

The proposed solution can be applied in numerous domains including education, banking, healthcare, government record management, historical document preservation, and office automation. By converting handwritten information into digital form, the system supports efficient data storage, retrieval, and analysis.

In the future, the system can be enhanced by incorporating multilingual handwriting recognition, transformer-based architectures, real-time handwriting processing, cloud integration, and mobile application support. These improvements can further increase recognition accuracy and expand the practical applications of the system.

REFERENCES

- [1] IEE "Handwritten Text Recognition Using Deep Learning Techniques," IEEE Access, 2024.
- [2] Springer, "Handwriting Recognition and Conversion Using Neural Networks," Springer Nature, 2024.
- [3] R. Malhotra and M. T. Addis, "Handwritten Amharic Word Recognition Using Additive Attention Mechanism," IEEE Transactions on Multimedia, 2024.
- [4] J. Memon et al., "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review," Journal of Artificial Intelligence Research, 2023.

- [5] M. Rajalakshmi, P. Saranya, and P. Shanmugavadivu, "Recognition of Handwritten Documents Using Convolutional Neural Networks," *Pattern Recognition Journal*, 2023.
- [6] Hemanth G. R., Jayasree M., Keerthi Venii S., "Handwritten Text Recognition: An Automated System for Transcription," *E3S Web of Conferences*, 2023.
- [7] Deepthi S. A. et al., "Handwritten to Text Document Converter," *Lecture Notes in Electrical Engineering*, Springer, 2022.
- [8] Yin Y., Zhang W., Hong S., Yang J., Xiong J., and Gui G., "Deep Learning-Aided OCR Techniques for Chinese Characters," *IEEE Access*, 2019.
- [9] España-Boquera S., Castro-Bleda M. J., Gorbe-Moya J., and Zamora-Martinez F., "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.
- [10] Goodfellow I., Bengio Y., and Courville A., *Deep Learning*, MIT Press, 2016.