# AI-Powered Handwritten Exam Evaluation System Using OCR and Text Similarity

DR. P SUMATHI, DR. N RATHINA KUMAR, ARUN NITHISH N S*, DIVYADHARSHINI S*, DIVYAPRABHA K G*, HARINI V*
Assistant Professor, Student*,
Department of Artificial Intelligence and Data Science, SNS College of Engineering, Coimbatore, India.

**ABSTRACT**

In an era where digital transformation is reshaping traditional educational practices, the need for automation in examination assessment has become increasingly vital. This research introduces an innovative AI-powered system designed to automate the evaluation of handwritten exam papers by integrating Optical Character Recognition (OCR) and Natural Language Processing (NLP) techniques. The system uses advanced OCR models like TrOCR and EasyOCR to extract content from scanned answer scripts and then employs text similarity techniques, such as cosine similarity and transformer-based models, to compare the extracted answers with a reference key. The final result displays the extracted response, similarity percentage, and the marks assigned. This approach enhances accuracy, reduces evaluator workload, and promotes transparency in assessment.

**Keywords** – OCR, NLP, TrOCR, EasyOCR, Cosine Similarity, Text Evaluation, AI in Education, Handwritten Answer Evaluation.

## I. INTRODUCTION

Handwritten examinations remain a fundamental part of academic evaluation, especially in regions where digital exams are not yet widely adopted. Manual evaluation of these answer sheets is not only time-consuming but also prone to inconsistencies due to human bias and fatigue. With the growth of AI, particularly in computer vision and natural language understanding, there is a unique opportunity to digitize and automate this traditional process.
This paper introduces a system that extracts handwritten responses using OCR and evaluates them through NLP-based similarity algorithms, ensuring both efficiency and fairness.

## II. RELATED SYSTEM

Current systems like Gradescope, Evalbee, and ZipGrade provide partial solutions for digital evaluation. Gradescope excels in rubric-based grading but requires institutional support. Evalbee is simple but lacks flexibility for subjective assessment. ZipGrade is effective for MCQs but not suitable for descriptive answers. CamScanner and OCR tools assist with text extraction, but evaluation is still manual. These limitations underscore the necessity for a unified system capable of end-to-end automated subjective assessment.

## III. PROPOSED SYSTEM

The proposed system comprises several integrated modules:**Image Upload**: Allows users to upload handwritten answer scripts in image format.

1. **OCR Module**: Uses TrOCR or EasyOCR to convert handwritten content into text.
2. **Preprocessing**: Cleans and normalizes extracted text for comparison.
3. **Answer Matching**: Compares the student's answer with a reference key using cosine similarity and optionally NLP transformers like BERT.
4. **Scoring Engine**: Allocates marks based on a predefined threshold mapping for similarity scores.

5. **Result Display**: Outputs extracted content, similarity percentage, and marks obtained in a user-friendly interface.

## IV. METHODOLOGY

The methodology follows a structured pipeline:

**Data Collection**: Images of handwritten answer scripts are collected and labeled with their respective correct answers.

**OCR Implementation**: TrOCR and EasyOCR are implemented for extracting text from handwritten images. TrOCR, trained on IAM datasets, shows higher accuracy in cursive writing.

**Text Preprocessing**: The text is cleaned—removal of punctuation, lowercasing, and stopword filtering—to normalize it for similarity comparison.

**Similarity Computation**: Cosine similarity is calculated between the answer key and the student's response. For deeper analysis, transformer-based models are considered.

**Scoring Strategy**: Marks are awarded based on similarity thresholds (e.g., 90–100% → full marks, 70–89% → partial credit).

**Output Interface**: The result includes detected text, similarity score, and calculated marks.

## V. SYSTEM ARCHITECTURE

The system architecture consists of the following layers:

- **Frontend Interface** – Built using HTML/CSS and Flask to allow file uploads and display results.
- **Backend API** – Processes the OCR and similarity analysis.
- **OCR Engine** – Integrates TrOCR/EasyOCR for handwriting-to-text conversion.
- **Similarity Engine** – Implements cosine similarity and NLP models for comparison.
- **Scoring & Output Module** – Finalizes marks and sends data to the frontend.

## VI. CONCLUSION

This project showcases the feasibility of applying OCR and NLP technologies to the evaluation of handwritten answer sheets. By automating the process, the system not only saves time but also enhances accuracy and ensures fairness inassessment.

The proposed solution addresses current gaps in the education sector by offering an AI-based evaluation mechanism suitable for large-scale use. Future developments could include adaptive scoring based on keyword importance, integration with student databases, and multilingual support.

## VII. REFERENCES

1. Jain, L. C., & Blum, M. (2001).

*Handwriting Recognition: Soft Computing and Probabilistic Approaches*. Springer.

2. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed.). Pearson.

3. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

5. Tunstall, L., von Werra, L., & Wolf, T. (2022). *Natural Language Processing with Transformers*. O'Reilly Media.

6. Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign.

7. Cosine Similarity – Journal of Data Science, 2021

8. Devlin et al. (2019), BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding