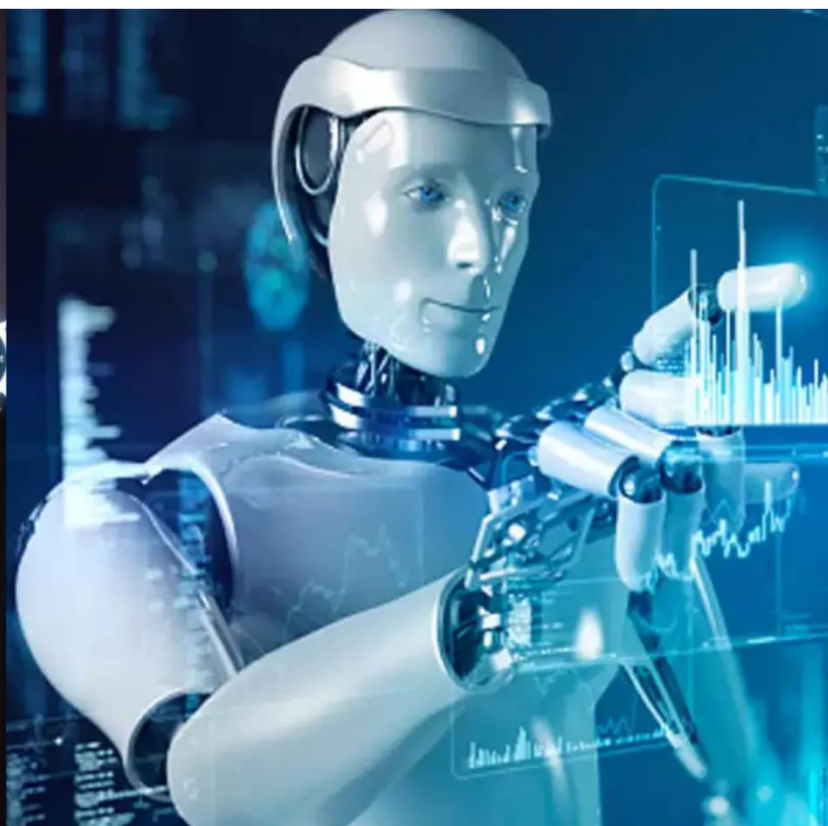




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 12, December 2025

AI-Powered Handwritten Script Evaluation

Prof. Naveen G¹, Avani A A², Nagaraj Gowda C Patil², Spoorti V Sangur², Srujan Chilukuri²

Assistant Professor, Department of Artificial Intelligence & Machine Learning, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India¹

Student, Department of Artificial Intelligence & Machine Learning, Bapuji Institute of Engineering and Technology, Davangere, Karnataka, India²

ABSTRACT: The proposed AI-driven handwritten script evaluation framework introduces a comprehensive methodology for assessing student responses by integrating advanced Optical Character Recognition (OCR) and Natural Language Processing (NLP) techniques. The system reduces subjectivity in grading by converting handwritten answers into digital text through a high-precision OCR pipeline, followed by semantic interpretation using Transformer-based language models such as BERT. Evaluation is performed by analysing contextual relevance, conceptual completeness, and the overall quality of student responses. Preprocessing operations—including image enhancement, noise removal, and text normalization—ensure reliable extraction prior to semantic comparison. The model is trained on curated datasets with instructor-assigned scores, enabling supervised learning for objective performance prediction. The system presents an efficient alternative to manual evaluation and contributes to more consistent and scalable academic assessment practices

KEYWORDS: Automated assessment, Handwritten script evaluation, Optical Character Recognition (OCR), Natural Language Processing (NLP), BERT, Semantic matching, Deep learning, Educational automation.

I. INTRODUCTION

Accurate assessment of student assignments is a central component of educational practice, yet traditional manual grading of handwritten scripts is constrained by slow turnaround times, evaluator fatigue, and inherent inconsistency. As academic institutions grow in scale, maintaining fairness, reliability, and rapid feedback becomes increasingly difficult. Recent advances in Artificial Intelligence—specifically in OCR and NLP—offer a transformative opportunity to automate the evaluation process while preserving meaningful pedagogical insights. OCR technologies enable the conversion of handwritten responses into machine-readable text, whereas Transformer-based NLP models, such as BERT, facilitate contextual understanding of student content by capturing semantic and conceptual relationships.

The system developed in this work leverages AI, Computer Vision, and NLP to evaluate descriptive responses in both handwritten and typed formats. Using robust OCR models like TrOCR, the system extracts text across diverse handwriting styles. Semantic analysis modules then compare student responses with reference answers, enabling context-aware scoring and detailed feedback generation. The proposed architecture reduces evaluator workload, enhances grading consistency, and aligns with the increasing digitization of assessment frameworks. Its adaptability makes it suitable for academic environments including schools, colleges, internal assessments, and practice evaluations.

II. LITERATURE SURVEY

- **Foundations of handwriting recognition:** Early OCR advancements were driven by deep learning models such as RNNs for sequence recognition [1] and CNNs for document processing and segmentation tasks [2], forming the base for modern handwritten script evaluation systems.
- **OCR + NLP integration:** Hybrid systems combining OCR outputs with Transformer-based models like BERT enabled semantic similarity-based evaluation, shifting grading from keyword matching to meaning-based assessment [3], [4].
- **Deep learning-based answer evaluation:** End-to-end frameworks using CNNs for segmentation and BERT/LSTM models for semantic scoring improved accuracy, consistency, and scalability in large-volume assessments [4], [6].

- **Multimodal grading systems:** Recent works introduced dual-pipeline architectures using image layout detection along with textual semantic evaluation, enhancing grading performance for scripts containing diagrams, tables, and complex layouts [5], [7].
- **Knowledge-enhanced evaluation:** Methods integrating handwritten text recognition with knowledge graphs enabled deeper conceptual understanding, while reinforcement learning approaches improved grading accuracy by simulating human evaluator feedback loops [8], [9].
- **Handling handwriting diversity and multilingual scripts:** Research also focused on improving OCR reliability through handwriting correction modules [10] and extending automated grading techniques to multilingual handwritten responses, increasing system applicability across regions and languages [11].

III. PROBLEM STATEMENT

Manual evaluation of handwritten student assignments is slow, inconsistent, and prone to human error. Existing automated methods do not fully combine accurate OCR with deep semantic understanding, making it difficult to perform fair and meaningful evaluation at scale.

IV. OBJECTIVES

- Develop a reliable OCR-to-text pipeline using models like TrOCR and integrate advanced NLP techniques such as BERT to extract, interpret, and analyze the meaning and conceptual accuracy of handwritten responses.
- Design a semantic scoring framework that evaluates answers based on contextual relevance and similarity rather than simple keyword matching, ensuring fair, consistent, and scalable assessment while significantly reducing teacher workload.
- Incorporate an optional automated feedback mechanism that provides constructive insights to help students improve their understanding and overall learning outcomes.

V. PROPOSED SOLUTION

The proposed solution employs a unified AI-based framework for automated evaluation of handwritten scripts. The system initiates with preprocessing of scanned answer sheets to enhance image clarity and remove noise. A high-performance OCR model such as TrOCR then converts handwriting into structured digital text. This extracted text undergoes normalization to correct minor inconsistencies before being processed through a BERT-based semantic analysis engine. By comparing student responses with instructor-provided reference answers, the system computes a semantic similarity score that reflects conceptual understanding and relevance. Marks and optional feedback are subsequently generated, offering a scalable, objective, and pedagogically meaningful alternative to manual evaluation.

VI. SYSTEM DESIGN

1. SYSTEM ARCHITECTURE

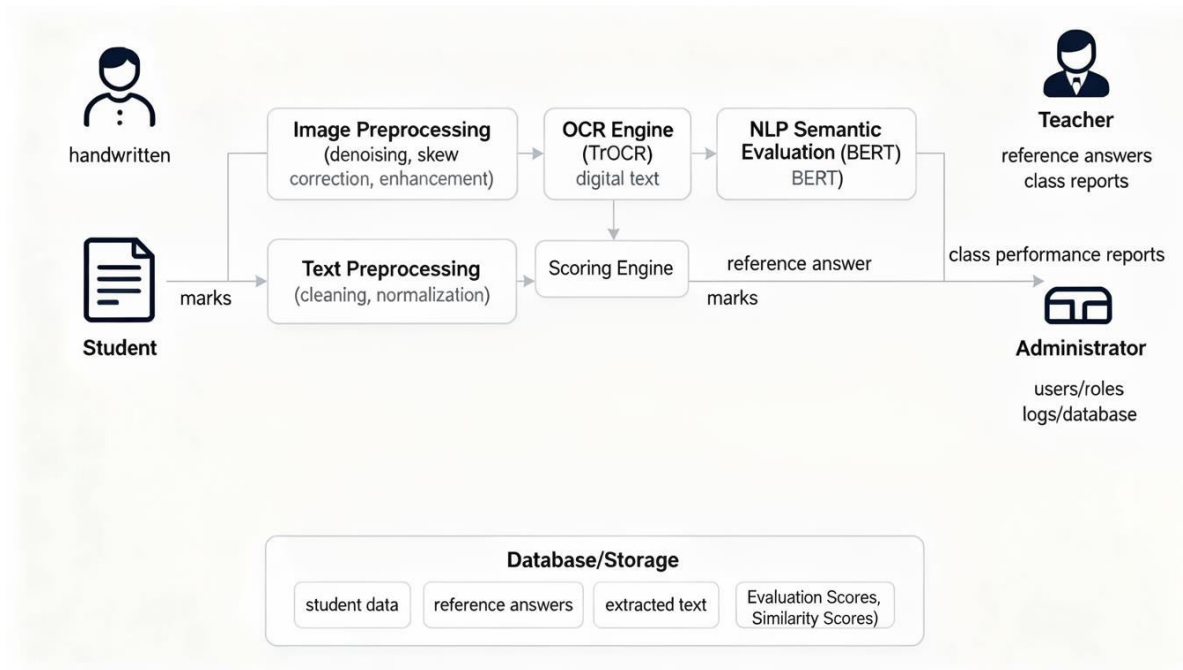


Figure1: System Architecture of Janmitra

The system architecture follows a modular, multi-stage pipeline that integrates OCR, text preprocessing, semantic analysis, and scoring into a cohesive evaluation workflow. Each module operates independently while contributing to an end-to-end assessment framework capable of handling diverse handwriting styles and content variations.

The process begins with image preprocessing, where scanned answer sheets undergo denoising, enhancement, and normalization. The enhanced image is processed by a Transformer-based OCR model such as TrOCR, which extracts digital text while accommodating handwriting variability. This text is cleaned and normalized to remove artifacts and ensure structured input for semantic processing.

A BERT-based semantic evaluation module then interprets the contextual meaning of the student's response and compares it with the reference answer. The resulting similarity measure is used by the scoring engine to assign marks and optionally generate instructional feedback. This multi-stage pipeline enables fair, fast, and reliable assessment suitable for deployment in academic institutions.

VII. RESULTS

1. Snapshots Of the Proposed System:

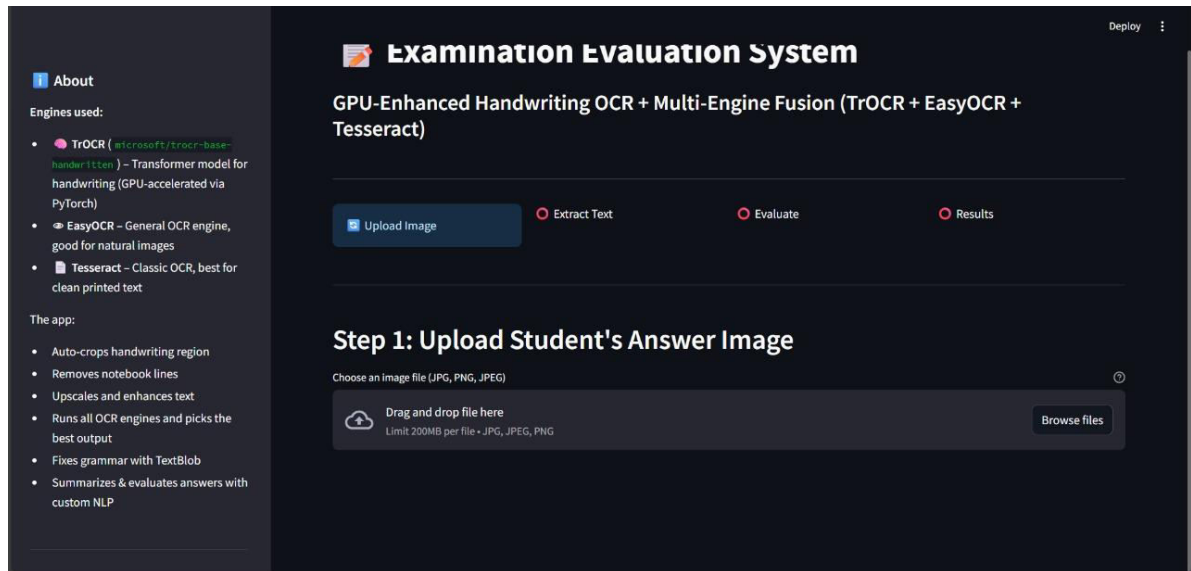


Figure 2: System Homepage – Examination Evaluation Interface

Displays the primary interface for uploading scripts, extracting text, evaluating responses, and viewing results using an integrated multi-engine OCR pipeline.

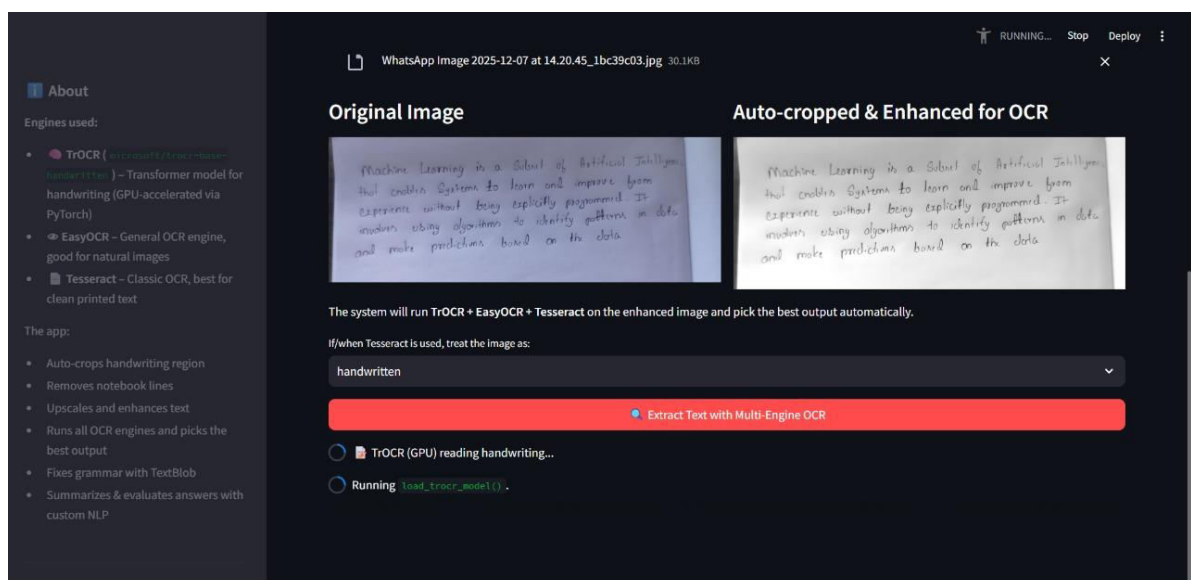


Figure 3: Image Preprocessing and Multi-Engine OCR Extraction

Shows the original answer image alongside its auto-cropped and enhanced version, demonstrating improved readability prior to OCR.

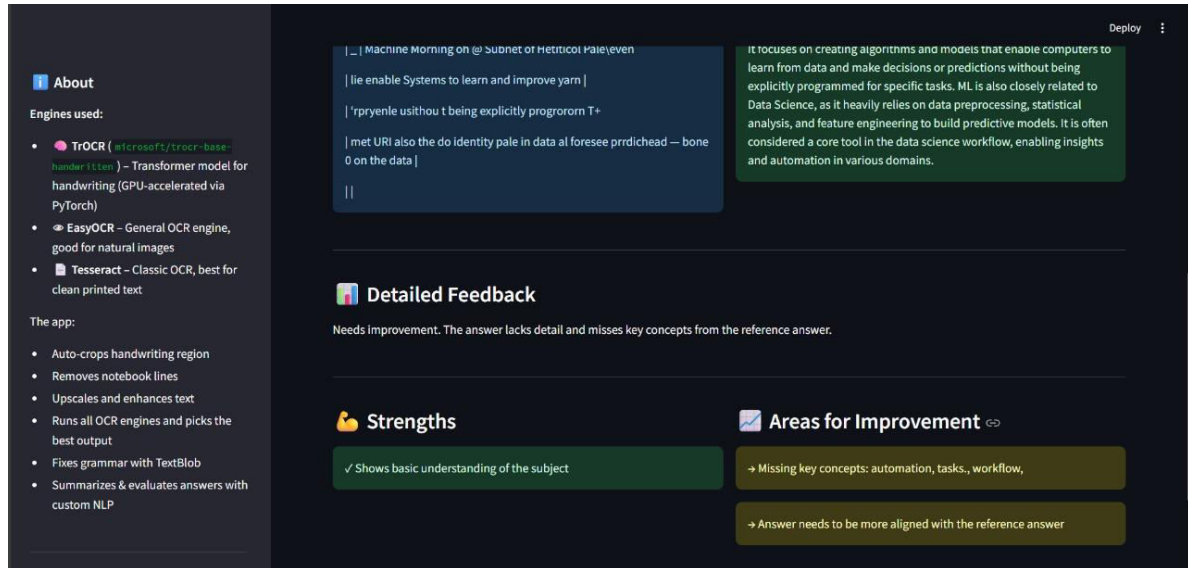


Figure 4: Detailed Feedback and Semantic Analysis Output

Presents AI-generated evaluation details, strengths, and improvement suggestions based on semantic comparison between student and reference responses.

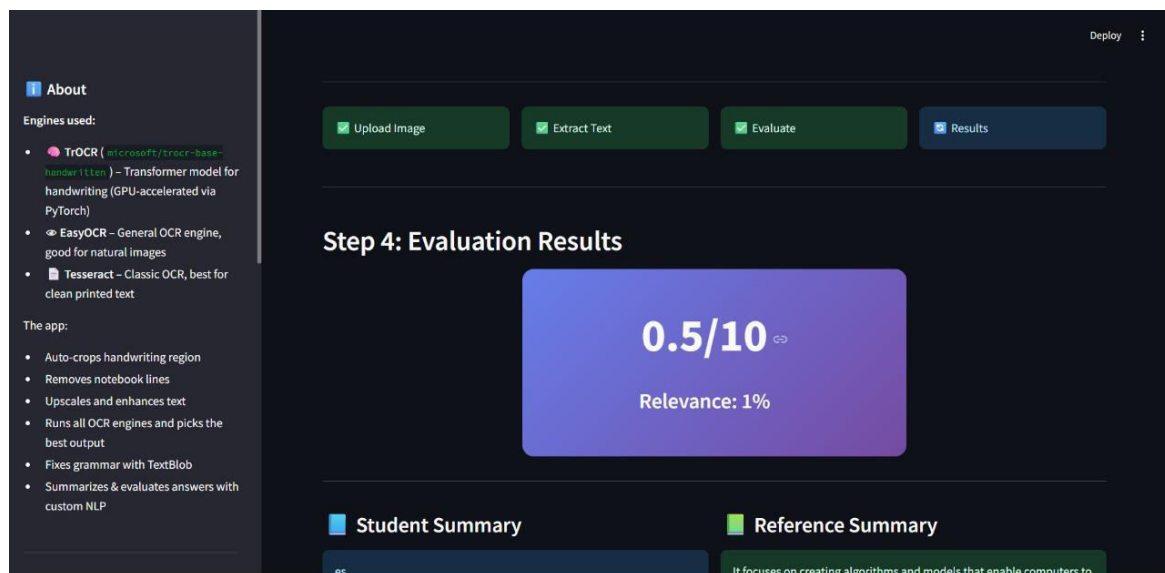


Figure 5: Final Scoring and Evaluation Summary

Displays the computed semantic similarity score, overall relevance, and summarized evaluation outcomes.

VIII. CONCLUSION

The AI-Powered Handwritten Script Evaluation System provides an effective, scalable, and objective alternative to manual grading by integrating state-of-the-art OCR and NLP technologies. Through accurate text extraction, semantic interpretation, and automated scoring, the system minimizes errors, accelerates evaluation, and ensures consistent

grading across large datasets. Its modular architecture enables seamless deployment in diverse academic settings, underscoring the transformative potential of AI in modernizing educational assessment practices.

REFERENCES

- [1] A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, and J. Schmidhuber, "A Novel Connectionist System for Unconstrained Handwriting Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009.
- [2] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [3] K. Ramachandran, S. K. Verma, and R. Mehta, "Automated Answer Evaluation Using OCR and BERT-Based Semantic Similarity," *International Journal of Intelligent Systems*, 2019.
- [4] P. Sharma, R. Singh, and A. Gupta, "End-to-End Automated Grading of Handwritten Answers Using CNN and BERT," *IEEE Access*, 2021.
- [5] M. Ahmed, T. Rahman, and S. Chowdhury, "A Multimodal Deep Learning Framework for Automated Evaluation of Handwritten Answer Sheets," *Neural Computing and Applications*, 2023.
- [6] R. Singh and S. Yadav, "Handwritten Answer Evaluation Using CNN-LSTM Hybrid Networks," *Journal of Image and Vision Computing*, 2020.
- [7] X. Zhang, J. Liu, and K. Zhou, "Semantic Answer Understanding Using Transformer Models: A T5-Based Approach," *Expert Systems with Applications*, 2022.
- [8] A. Gupta, P. Arora, and V. Nair, "Knowledge Graph-Enhanced Automated Answer Evaluation," *International Journal of Data Science and Analytics*, 2020.
- [9] Y. Wang, L. Chen, and Q. Zhou, "Reinforcement Learning for Automated Grading of Handwritten Responses," *IEEE Transactions on Learning Technologies*, 2021.
- [10] L. Chen, H. Wu, and M. Li, "Improving OCR Accuracy Using Handwriting Correction Models for Educational Assessment," *Pattern Recognition Letters*, 2021.
- [11] J. Lee, S. Park, and H. Kim, "Multilingual Handwritten Answer Evaluation Using Deep Learning," *IEEE Transactions on Artificial Intelligence*, 2022.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details