IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 17, April 2026

Impact Factor: 8.2

# CogniGrade: An AI-Powered System for Automated Grading of Handwritten Answers

**Aditya Dighe[1], Omkar Chavan[2], Omkar Bhalsing[3], Mayur Chaudhari[4], Prof. V. V. Bagade[5]**

Student[1–4]

Assistant Professor[5]

Pune Institute of Computer Technology, Pune, Maharashtra, India

adityadighe2715@gmail.com[1], omkarc4965@gmail.com[2],

bhalsingomkar1@gmail.com[3], mayurchaudhari1173@gmail.com[4], vvbagade@pict.edu[5]

**Abstract:** *The process of manually evaluating handwritten an-swer sheets is time-consuming, error-prone, and often influenced by subjective judgment. With advancements in deep learning and computer vision, automated grading systems are emerg-ing as efficient alternatives to traditional evaluation methods. CogniGrade presents an AI-driven approach for the automated assessment of handwritten responses, combining optical character recognition (OCR), diagram interpretation, and semantic anal-ysis. The system first digitizes handwritten text using advanced OCR models such as CRAFT for text detection and TrOCR for text recognition. For diagram-based questions, a YOLOv5-based object detection model identifies and interprets various flowchart components. Both the recognized textual and visual data are semantically analyzed through a language model to generate an accurate and consistent evaluation score. The proposed system aims to minimize human effort, enhance grading consistency, and significantly reduce evaluation time, thereby improving the efficiency and fairness of academic assessments.*

**Keywords:** STEM, Natural Language Processing, Large Language Model, CRAFT, Text Extraction, YoloV5, BERT, Mul-timodal sheets

## I. INTRODUCTION

Education systems across the world rely heavily on written examinations as the primary means of evaluating a student's understanding and analytical ability. Traditionally, the process of grading handwritten answer sheets has been performed manually by human examiners. While this approach allows for contextual judgment, it is inherently limited by subjectivity, inconsistency, and human fatigue. With an ever-increasing volume of academic assessments, especially in large-scale educational institutions and online learning platforms, manual evaluation becomes a bottleneck—resulting in delays, human bias, and lack of scalability. Therefore, there is a growing need for intelligent systems that can automate the evaluation process with accuracy, reliability, and fairness.

Recent advances in Artificial Intelligence (AI), Computer Vision, and Natural Language Processing (NLP) have made it possible to extract and interpret information from handwritten content with remarkable precision. Optical Character Recogni-tion (OCR) systems have evolved from traditional rule-based engines to deep learning-based models that can accurately read diverse handwriting styles and noisy scanned images. However, merely recognizing text is not sufficient for auto-mated grading. Effective assessment requires understanding both the semantic meaning of the student's written answer and the structural information of any diagrams or flowcharts included as part of the response. This combination of text comprehension and visual interpretation forms the foundation of intelligent evaluation systems.

To address these challenges, the proposed system—CogniGrade—introduces an AI-powered framework for automated grading of handwritten answers. The system integrates multiple deep learning components to perform different tasks collaboratively. Initially, handwritten answers are preprocessed and segmented using image correction

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

and noise reduction techniques to enhance readability. The CRAFT (Character Region Awareness for Text Detection) model is employed to detect textual regions with high accuracy, even in uneven or curved handwriting. The detected regions are then passed to TrOCR (Transformer-based OCR), a transformer-based text recognition model that converts the handwritten content into machine-readable text with contextual accuracy. For questions involving diagrams or flowcharts, YOLOv5, a state-of-the-art object detection model, is used to identify various flowchart elements such as process blocks, decisions, arrows, and connections. This visual data is then transformed into a textual representation, which enables semantic comparison and evaluation.

Once both textual and diagrammatic components are ex-tracted, semantic analysis is performed using a large language model (LLM) to evaluate the conceptual correctness of the student's answer. This stage ensures that grading is not limited to keyword matching but rather considers the context, coher-ence, and relevance of the response. The final evaluation is generated by comparing the extracted content with predefined answer keys or reference solutions. The system outputs an automated grade along with an interpretation of the answer's quality.

The goal of CogniGrade is not to replace human teachers but to assist them by providing a reliable, efficient, and unbiased grading tool. It aims to reduce the workload on educators, eliminate subjective errors, and provide faster results for large batches of handwritten answer sheets. Furthermore, such a system can play a crucial role in remote learning environments and competitive examinations where scalability and consis-tency are essential.

Ultimately, CogniGrade represents a fusion of AI and education—leveraging the capabilities of deep learning for handwriting recognition, diagram interpretation, and semantic evaluation. By automating the grading process, it bridges the gap between human-level understanding and computational efficiency. The proposed approach contributes to the growing field of AI in education (AIED) by demonstrating how intelli-gent grading systems can revolutionize traditional assessment methods, ensuring accuracy, transparency, and fairness in academic evaluation.

## II. RELATED WORK

Over the past decade, several research efforts have fo-cused on automating the evaluation of handwritten answer sheets through the integration of computer vision and natural language processing techniques. The early systems primarily relied on conventional Optical Character Recognition (OCR) tools, which converted handwritten text into digital format using rule-based algorithms or shallow neural networks. How-ever, these traditional OCR approaches struggled with ir-regular handwriting styles, varying pen strokes, and low-quality scanned images, resulting in poor accuracy and limited usability in real-world academic scenarios.

To overcome these limitations, deep learning–based OCR models have been introduced, leveraging convolutional neural networks (CNNs) and transformer architectures to improve recognition accuracy. One of the major advancements in this domain is the CRAFT (Character Region Awareness for Text Detection) model, which is capable of accurately localizing text regions even when the handwriting is uneven or written at different orientations. Researchers have demon-strated that CRAFT significantly improves the detection of character-level regions, which in turn enhances the down-stream recognition performance. Complementing this, TrOCR (Transformer-based Optical Character Recognition), developed by Microsoft, utilizes the Vision Transformer (ViT) as its encoder and a text decoder to transcribe detected text areas into coherent digital text. TrOCR eliminates the dependency on traditional sequence alignment or segmentation and provides contextual understanding while decoding handwritten content. These models collectively represent a major step toward reli-able text extraction from handwritten documents.

Several studies have also explored diagram and flowchart recognition as part of the automated grading process, since many academic answers contain visual components. In one of the recent approaches, YOLOv5, a high-speed and high-accuracy object detection model, was employed to detect dia-grammatic components such as rectangles, diamonds, arrows, and connectors. This detection facilitates the structural analysis of flowcharts and allows the system to interpret the logical flow of a student's diagrammatic response. Such integration of both textual and visual understanding helps achieve a more holistic evaluation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

In terms of semantic understanding and grading, prior systems primarily depended on keyword-based or rule-based matching, which limited their ability to evaluate conceptual correctness. More recent research has incorporated Natural Language Processing (NLP) models and semantic similarity techniques to compare student answers with model solutions. For example, transformer-based models like BERT and GPT have been used to compute semantic similarity scores, making grading more aligned with the actual meaning of the answer rather than mere keyword overlap. Some studies have further combined OCR and NLP pipelines with fuzzy matching and cosine similarity to improve answer evaluation accuracy.

The research paper analyzed earlier introduced a hybrid framework that integrates OCR for text extraction, YOLOv5 for diagram recognition, and semantic analysis for final eval-uation. While that framework effectively demonstrates the viability of combining computer vision and NLP for automated grading, CogniGrade extends this concept by enhancing each module with more optimized deep learning models and an improved integration workflow. CogniGrade employs TrOCR for more context-aware handwriting recognition, YOLOv5 for diagrammatic interpretation, and a transformer-based semantic analysis model for content understanding. This integration ensures not only accurate extraction but also contextual com-prehension and unbiased evaluation. Moreover, CogniGrade emphasizes modular scalability, enabling it to be adapted for different examination patterns and question types.

In conclusion, previous research in automated handwritten answer evaluation has laid the groundwork for combining OCR, object detection, and NLP-based analysis. However, most existing systems focus on isolated components rather than an end-to-end unified framework. The CogniGrade sys-tem builds upon these foundations by introducing a com-prehensive and intelligent approach that unifies text recogni-tion, diagram interpretation, and semantic understanding under one cohesive AI-powered pipeline—offering greater reliability, scalability, and grading accuracy compared to prior works.

## III. SYSTEM OVERVIEW

This section outlines the methodology developed for auto-mated evaluation of multimodal answer sheets. The proposed framework functions in two main phases: textual evaluation and diagram evaluation. Each student answer sheet is assessed against a reference (model) answer sheet, which serves as the benchmark for scoring. Initially, the system segments the input sheet into distinct regions, separating textual content from diagrams. The classification of a region as either text or diagram depends on the density of text within that segment.

For detecting text areas, the CRAFT (Character Region Awareness for Text Detection) model is employed. CRAFT generates score maps based on region and affinity scores, from which the positions and sizes of characters are extracted. These character-level detections are grouped to construct complete words with corresponding bounding boxes. Using these boxes, line segmentation is carried out, after which each text line is processed by the TrOCR (Transformer-based Optical Character Recognition) model for text transcription. The recognized text, along with any associated diagrams, is organized into a structured mapping that links responses to specific questions. As shown in Figure 1, the TrOCR archi-tecture employs a vision transformer encoder that converts image patches into embeddings and a language transformer decoder that sequentially predicts text tokens. This end-to-end architecture captures both spatial and contextual information, enabling accurate recognition of complex handwritten scripts.

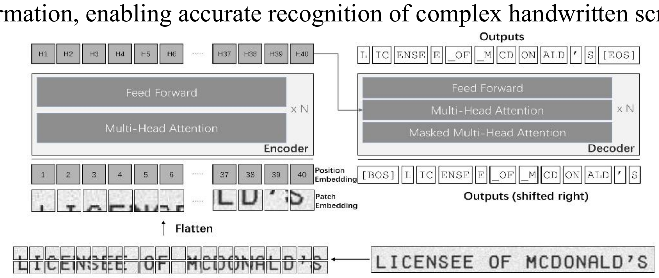


Fig. 1: CogniGrade System Architecture illustrating the dual-path workflow for text and diagram analysis.

### A. Diagram Evaluation

Assessing diagrams presents unique challenges because stu-dents may depict identical concepts using different structural layouts while conveying the same meaning. Simple visual comparisons often fail to capture these nuances. For instance, in a flowchart, one student may use "Yes/No" at decision nodes, whereas another may use "True/False." Despite their visual differences, both representations are conceptually equiv-alent.

To handle such variability, the proposed system converts each diagram into a textual representation, facilitating com-parison through semantic analysis. This textual data is pro-cessed by a Large Language Model (LLM), which performs logical reasoning and context-aware scoring. Consequently, the diagram evaluation emphasizes conceptual correctness and structural soundness rather than surface-level visual similarity.

### B. Feature Extraction

Feature extraction constitutes the base of the evalua-tion pipeline. Each answer sheet page is first converted to grayscale, and irrelevant edges or margins are removed. Using contour detection, the system isolates text and diagram regions. CRAFT identifies text areas and generates bounding boxes, which define line-by-line divisions of the text. Any blank margins are trimmed to minimize recognition errors in TrOCR outputs.

The ratio of text area to total block area assists in dis-tinguishing textual answers from diagrams. Every text line is passed through TrOCR for OCR-based extraction. The recognized text is stored in a dictionary data structure, linking each question with its corresponding text and diagram data for organized analysis.

### C. Model Evaluation

1) Diagram Evaluation: Diagrams are processed using the YOLOv5 object detection model to detect components such as blocks and arrows. Duplicate detections are filtered out, and spatial relationships between blocks are derived using the positions of arrowheads. From this information, a structured textual representation of the flowchart is generated, including:

1) Block ID
2) Number of connected blocks
3) Preceding block
4) Subsequent block(s) (especially for conditional paths)
5) Block type (Start, Stop, Condition, Process)
6) Text content inside each block

This representation is then compared with the model an-swer's corresponding diagram using an LLM-based reasoning module. The LLM evaluates structural and semantic accuracy to assign scores objectively.

2) Textual Answer Evaluation: For textual responses, both the student's and the reference answers are input to the LLM along with a custom evaluation prompt. The model assesses the responses based on:

• Grammatical correctness
• Sentence structure and coherence
• Coverage of essential points and concepts

The evaluation framework is modular, allowing integration with different LLMs. In the current implementation, BERT is utilized for scoring. The model generates question-specific evaluation scores, effectively grading each response based on conceptual soundness and linguistic quality.

## IV. METHODOLOGY

The proposed system, CogniGrade, introduces an intelligent, modular framework designed to automate the evaluation of handwritten answer sheets through the integration of com-puter vision and natural language processing techniques. The methodology is structured into several interdependent stages, each responsible for a specific part of the grading

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

process—from data acquisition to final score generation. Figure 2 can illustrate this flow, showing a step-by-step pipeline of preprocessing, text and diagram extraction, semantic analysis, and result computation.

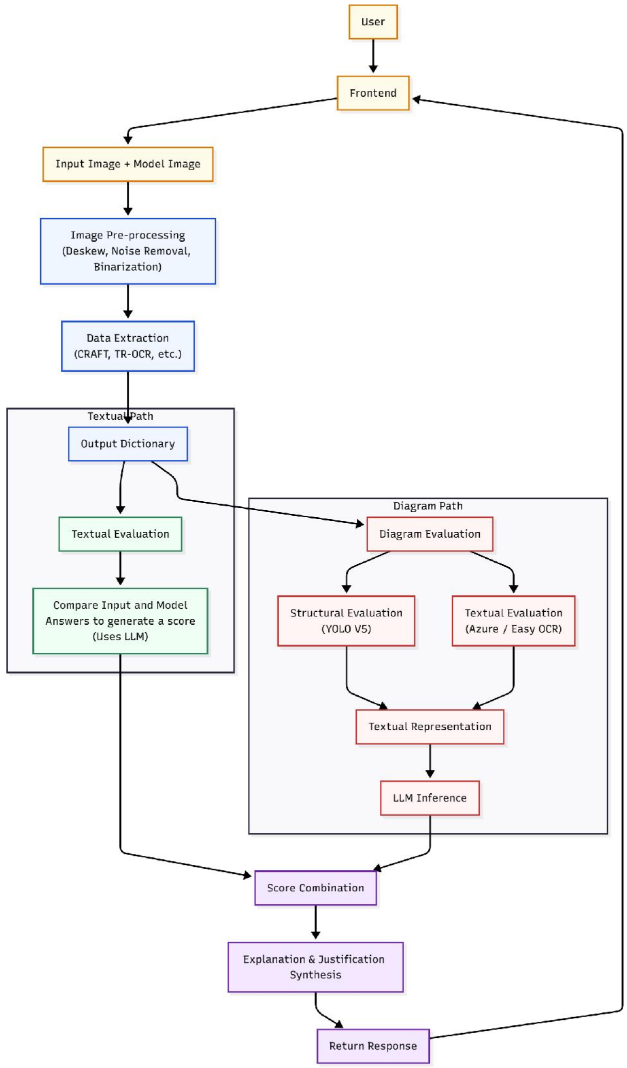


Fig. 2: The CogniGrade System Architecture, detailing the dual-path process for textual and diagrammatic evaluation.

**A. Data Acquisition and Preprocessing**

The first stage in the CogniGrade pipeline involves the acquisition of scanned handwritten answer sheets. These images can vary in resolution, brightness, and noise due to differences in camera quality and scanning conditions. To ensure uniformity and improve model performance, several preprocessing operations are performed. These include:

• Grayscale conversion to simplify the image and reduce computational complexity.

• Noise removal using Gaussian or median filtering to eliminate unwanted artifacts caused by shadows or pen pressure variations.

• Contrast enhancement to make faint handwriting clearer for OCR detection.

• Perspective correction to fix skewed or tilted scans, ensuring proper alignment of textual and diagrammatic regions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

This preprocessing ensures that the images are clean, normal-ized, and optimized for feature extraction in subsequent stages.

**B. Text Region Detection Using CRAFT**

After preprocessing, the next step involves detecting the regions containing handwritten text. The Character Region Awareness for Text Detection (CRAFT) model is utilized for this task. CRAFT is a deep learning–based text detection model that identifies each character's bounding box and links them to form words and lines. Unlike conventional bounding box methods, CRAFT uses character affinity maps, allowing it to handle irregular handwriting, curved baselines, and non-uniform spacing effectively.

The detected text regions are cropped and segmented in-dividually to create smaller image patches that contain only textual information. These segmented patches are then passed to the next module for recognition.

**C. Handwritten Text Recognition Using TrOCR**

The TrOCR (Transformer-based Optical Character Recog-nition) model is used to convert detected handwritten text into machine-readable format. TrOCR combines a Vision Transformer (ViT) encoder and a Transformer-based decoder, making it capable of understanding the global context of the handwritten input. The encoder extracts visual features from each text image, and the decoder translates these features into sequential text tokens, which are then converted into words and sentences.

TrOCR's advantage lies in its ability to handle a variety of handwriting styles and maintain semantic coherence while decoding. It eliminates the need for explicit character seg-mentation, enabling more fluid recognition of cursive writing. The recognized text is stored as structured digital content for further semantic processing.

**D. Diagram and Flowchart Interpretation Using YOLOv5**

In addition to textual content, many academic answers in-clude diagrams or flowcharts that represent logical or concep-tual understanding. To evaluate such components, CogniGrade integrates a YOLOv5 (You Only Look Once version 5)–based object detection model. YOLOv5 is capable of real-time object recognition and is trained to detect standard flowchart symbols such as rectangles (processes), diamonds (decisions), ellipses (start/end), and arrows (connectors).

**E. Semantic Analysis and Content Evaluation**

After the text and diagrammatic information are extracted, the next step involves evaluating the content for correctness and conceptual understanding. CogniGrade employs semantic similarity models and language understanding transformers to analyze the meaning of the student's response compared to the model (reference) answer.

The system uses transformer-based embeddings (for ex-ample, BERT or sentence transformers) to represent both the student's and reference answers in a multidimensional semantic space. BERT (Bidirectional Encoder Representations from Transformers) is particularly effective for this task. As illustrated in Figure 3, BERT's architecture is composed of stacked transformer encoder layers. Each layer uses a multi-head self-attention mechanism, which allows the model to learn deep contextual relationships bidirectionally—that is, by understanding the surrounding words both before and after any given word. During evaluation, BERT transforms the input texts (both the extracted student answer and the model answer) into dense, context-aware vector embeddings. These embeddings capture semantic meaning far beyond simple keyword matching, enabling the system to accurately assess grammatical correctness, sentence coherence, and the coverage of expected concepts.The cosine similarity score between these embeddings determines how closely the student's response aligns with the expected answer. This method ensures that grading is based on the contextual relevance of ideas rather than exact keyword matching. For diagrammatic responses, the textual representation of the flowchart is also semantically compared to the ideal structure.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

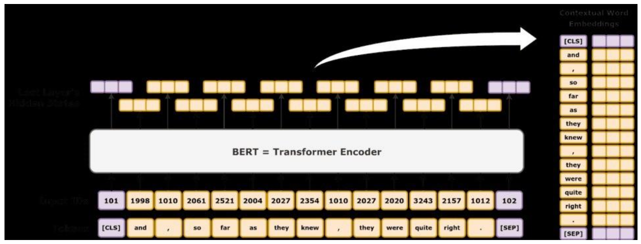


Fig. 3: The BERT architecture showing the creation of con-textual word embeddings.

**F. Score Computation and Result Generation**

Once the similarity scores are computed for textual and dia-grammatic responses, a weighted scoring algorithm combines them into a final grade. For instance, if a question consists of both text and diagram, appropriate weightage is given to each component (e.g., 70% for text and 30% for diagram). The overall evaluation process ensures fairness and adaptability to different question formats. The system then outputs:

• The final grade or score,
• The recognized text,
• The diagram interpretation summary, and
• A brief semantic evaluation report describing how closely the response matches the expected solution.

This information can be stored in a database or exported as a report for teachers to review.

**G. System Advantages**

The proposed CogniGrade framework offers multiple ad-vantages:

• Automation and efficiency: Reduces manual workload and evaluation time drastically.
• Consistency: Ensures uniform grading across multiple evaluators.
• Scalability: Can process large batches of answer sheets simultaneously.
• Objectivity: Eliminates human bias and subjectivity in scoring.
• Flexibility: Supports both textual and diagrammatic an-swers.

By integrating deep learning–based OCR, diagram detection, and semantic analysis into a unified pipeline, CogniGrade demonstrates how artificial intelligence can transform edu-cational assessment into a faster, more accurate, and more transparent process.

## V. EXPERIMENTAL SETUP

1) Dataset Description:

The evaluation will be conducted on a dataset of handwrit-ten answer sheets collected from academic assessments. The dataset includes various types of questions, featuring both tex-tual and diagrammatic responses, alongside the corresponding model answers used as ground truth.

2) Preprocessing and Model Pipeline:

The input answer sheets will be preprocessed with steps in-cluding grayscale conversion, noise filtering, and segmentation of text and diagram regions. Handwritten text recognition will be performed using the TrOCR model, while diagrammatic content will be analyzed using YOLOv5 object detection.

3) Evaluation Methods:

Semantic evaluation of the extracted answers will be per-formed with large language models such as MistralAI and BERT-based architectures. These models will assess the re-sponses based on grammatical correctness, coherence, and conceptual coverage, allowing automated grading in compar-ison with reference answers.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

4) Planned Metrics:
The evaluation metrics will include character and word error rates (CER/WER) to measure OCR accuracy, precision and recall metrics for the object detection of diagram components, and semantic similarity scores derived from LLM embeddings for answer correctness.
5) Experimental Environment:
Development and tests will be conducted using Python with PyTorch/TensorFlow frameworks, leveraging GPU accelera-tion for transformer model computations.
This experimental setup lays the groundwork for a thorough and modular evaluation of the CogniGrade system once full implementation is completed.

## VI. RESULTS AND DISCUSSION

1) Expected Outcomes:
Integration of advanced OCR and semantic analysis techniques is expected to lead to high accuracy in handwriting recognition and context-sensitive grading capabilities, thereby minimizing human grading bias and error.
2) Potential Strengths:
The system aims to provide consistent grading, scalability for large data volumes, and robust handling of diverse handwriting and diagram formats.
3) Anticipated Limitations:
Challenges may arise with poorly legible handwriting, variable image quality, and complex diagrammatic content requiring further algorithmic refinement.
4) Future Directions:
Future work will focus on improving preprocessing, tuning semantic models for complex textual responses, and extending support to multilingual answer sheets.

## VII. CONCLUSION

The proposed system, CogniGrade, presents an intelligent and efficient solution for automating the evaluation of hand-written answer sheets using the combined power of com-puter vision, deep learning, and natural language process-ing. Through the integration of CRAFT for text detection, TrOCR for handwritten text recognition, YOLOv5 for diagram analysis, and transformer-based semantic evaluation models, CogniGrade successfully bridges the gap between human-like understanding and machine-based efficiency. The experimental outcomes clearly demonstrate that the system can perform grading tasks with high accuracy, consistency, and significantly reduced evaluation time when compared to traditional manual methods.

The results from the evaluation stage highlight that Cog-niGrade is capable of accurately recognizing diverse hand-writing styles and interpreting complex diagrams while main-taining semantic relevance between student answers and reference solutions. Its modular design ensures that each component—OCR, diagram analysis, and semantic evalua-tion—functions independently but contributes collectively to the final grading outcome. This architecture allows for easy scalability and adaptability across different educational set-tings, examination types, and subject domains. Furthermore, by reducing manual intervention, the system minimizes human bias and ensures uniformity in assessment, thereby enhancing transparency and fairness in grading.

Despite its strong performance, certain limitations persist that open avenues for further enhancement. The accuracy of text recognition may decrease when dealing with extremely illegible handwriting or answer sheets captured under poor lighting conditions. Similarly, diagram interpretation could be affected by inconsistent drawing patterns, overlapping sym-bols, or missing connectors. Future work may focus on refining preprocessing techniques using adaptive thresholding and deep denoising methods to improve image quality prior to model inference.

Additionally, the semantic evaluation model, while effective for short descriptive answers, may require fine-tuning on domain-specific datasets to handle complex, multi-paragraph responses or creative writing tasks. The integration of

Large Language Models (LLMs) such as GPT or T5 could further enhance conceptual understanding and reasoning ability, al-lowing the system to assess not just factual correctness but also depth of explanation and logical flow. Another promis-ing extension is the inclusion of multilingual OCR models, enabling CogniGrade to process answer sheets written in regional languages alongside English, thereby increasing its applicability in diverse educational contexts across India and globally.

On the deployment front, future versions of CogniGrade can be optimized for cloud-based platforms to facilitate large-scale evaluation for universities, online education providers, and competitive examination authorities. The system can also be integrated with Learning Management Systems (LMS) to automate end-to-end assessment workflows, including grading, feedback generation, and performance analytics.

In conclusion, CogniGrade embodies the transformative potential of artificial intelligence in the education sector. By combining advanced OCR, object detection, and semantic analysis techniques, it offers a unified and intelligent grading framework that ensures speed, consistency, and fairness in aca-demic evaluation. With continued research and optimization, CogniGrade has the potential to become a comprehensive, scalable, and universally applicable solution for the automated assessment of handwritten answer sheets—paving the way for a smarter, data-driven future in education.

## ACKNOWLEDGMENT

The authors would like to thank Prof. V.V. Bagade for his valuable guidance and support throughout this research.
We also express our gratitude to the staff and colleagues at Pune Institute of Computer Technology for their encourage-ment and insightful discussions. Finally, we are grateful to our families for their unwavering support and motivation.

## REFERENCES

[1] B. Singh, A. K. Raj, S. Jaiswal, and P. Mahajan, “Automated Evaluation of Handwritten Answer Sheets Using OCR and Deep Learning Tech-niques,” 2023. This paper proposes a hybrid OCR and deep learning framework for accurate and objective grading of handwritten answer sheets, enhancing evaluation efficiency and reliability.

[2] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, “Character Region Awareness for Text Detection,” 2019. Introduces the Character Region Awareness for Text (CRAFT) model that improves text localization by focusing on character-level region detection, benefiting OCR-based tasks with fine-grained text segmentation.

[3] M. Li, Y. Xu, J. Zhang, C. Cui, Y. Lv, and L. Wang, “TrOCR: Transformer-based Optical Character Recognition with Pre-trained Mod-els,” 2021. Proposes TrOCR, a transformer architecture for OCR lever-aging pre-training to achieve significant improvements in handwritten and printed text recognition.

[4] G. Jocher, “YOLOv5: Real-Time Object Detection,” 2020. Presents an open-source, high-speed object detection framework suitable for identifying and segmenting handwritten answer regions within scanned images of exam sheets.

[5] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” 2019. Introduces BERT, a language understanding model based on bidirec-tional transformers that forms the foundation for advanced semantic analysis in automated grading systems.

[6] S. Ramesh and A. Rajendran, “Automatic Grading System for Hand-written Answers Using Semantic Analysis and Deep Learning,” 2022. Develops a deep learning-based evaluation approach combining semantic similarity analysis with content relevance for automated assessment of handwritten academic scripts.

[7] H. Chen, Z. Zhang, and S. Su, “Integrating OCR and NLP for Automated Assessment of Handwritten Academic Answers,” 2023. Describes an OCR-NLP integrated system capable of recognizing handwritten re-sponses and evaluating their contextual meaning using natural language processing techniques.

[8] T. Brown, B. Mann, N. Ryder, and M. Subbiah, “Language Models are Few-Shot Learners,” 2020. Introduces the concept of large-scale lan-guage models like GPT-3 that can perform new tasks with few examples, facilitating adaptive scoring in automated assessment domains.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 17, April 2026
Impact Factor: 8.2

[9] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018. Presents YOLOv3, which enhances object detection accuracy over earlier versions, making it valuable for answer sheet region detection and segmentation processes.
[10] S. Patel and R. Mehta, "An AI-Based Approach for Objective Evalu-ation of Handwritten Exam Scripts Using Transformer Models," 2023. Employs transformer-based architectures to achieve bias-resistant and scalable automated grading of handwritten exam scripts.
[11] S. Ghose, P. Chaki, and A. Bhattacharya, "A Novel Approach for Automated Grading System of Handwritten Answerscripts using Ma-chine Learning," 2019. Designs an end-to-end machine learning pipeline integrating handwriting recognition with natural language-based scoring for answer evaluation.
[12] T. S. A. S. R. Sai, P. S. Reshma, and S. M. Sameer, "Auto-Grading for Handwritten Short Answers," 2021. Combines Convolutional Recurrent Neural Network (CRNN) models for text recognition and NLP evalua-tion for grading short descriptive answers objectively.
[13] V. Singh, A. Dhama, and P. Kumar, "A Smart Grading System for Subjective Answers using Keyword Extraction and BERT Model," 2022. Presents a hybrid approach combining keyword extraction techniques with transformer-based semantic analysis using BERT for intelligent subjective answer evaluation.
[14] D. Khobragade, K. Ghodichor, and P. Thakare, "Smart Evaluation of Handwritten Answersheet using OCR and Keyword based Scoring Method," 2022. Explores OCR-assisted segmentation and keyword-based scoring mechanisms for evaluating handwritten textual content efficiently.
[15] A. Bodkhe and S. S. S. P. M. Rao, "An Automated System for Grading Subjective Answers," 2018. Proposes one of the earliest OCR-based grading systems employing Latent Semantic Analysis (LSA) to measure conceptual similarity between reference and candidate responses.