Optimizing Whiteboard Digitization with Clarix: An Automated Vision-Based System for Real-Time Text Extraction and Surface Cleaning

EZZAHOUD HAJAR[1], BAZGOUR YASSINE[1], ABOUHANE ZAHRA[1], AMEKSA MOHAMMED[2], AND ABDELLAH NABOU[2]

[1] Cadi Ayyad University UCA,Faculty of Sciences Semlalia, Computer Science Department, Marrakech, Morocco
[2] Cadi Ayyad University, UCA, Faculty of Science Semlalia, FSSM, Laboratory of Computer Science and Smart Systems, LISI, Marrakech Morocco

Whiteboard digitization in educational settings grapples with challenges posed by handwritten text, mathematical notation, and unstructured layouts that confound traditional optical character recognition (OCR) systems. Clarix, a low-cost embedded system, addresses these issues by integrating a Raspberry Pi, high-resolution camera, and servo-driven erasure mechanism to automate content capture, text extraction, and archiving as searchable PDFs. This study benchmarks five multimodal large language models (GPT-4o, Claude, Gemini, DeepSeek, Grok) against traditional OCR systems (PaddleOCR, EasyOCR, TesseractOCR) for extracting whiteboard content across mathematics, physics, and economics domains. Results highlight a stark performance divide: multimodal LLMs achieved F1-scores ranging from 0.7550 to 0.8234, with GPT-4o leading at 0.8234 (precision 0.7954, recall 0.8534), while PaddleOCR topped traditional systems with an F1-score of 0.3333 (precision 0.3523, recall 0.3334), followed by EasyOCR (0.1158) and TesseractOCR (0.0000). Notably, increased region detection correlated with diminished performance, underscoring the superiority of contextual understanding over exhaustive segmentation. Clarix's fusion of intelligent automation and advanced text processing marks a transformative advancement in bridging analog and digital educational environments.

1. INTRODUCTION

A. Problem

Traditional whiteboards remain prevalent in educational and collaborative environments due to their simplicity, affordability, and ease of use [1]. However, they lack modern capabilities such as digital archiving, real-time content sharing, and efficient management of written material [2]. Manual transcription of whiteboard content is time-consuming and error-prone, while frequent cleaning disrupts the workflow and reduces productivity [3]. Although smart boards and interactive digital whiteboards offer advanced features, their high cost and infrastructural demands hinder widespread adoption, especially in resource-constrained settings [4]. This creates a clear need for an affordable, automated solution capable of capturing, digitizing, and erasing whiteboard content in real time—bridging the gap between traditional writing surfaces and modern digital tools [5].

B. Aim

The primary aim of this study is to develop a low-cost, intelligent whiteboard automation system that integrates embedded systems with advanced artificial intelligence to enhance traditional educational environments. The system seeks to automatically capture, process, and digitize whiteboard content while providing seamless erasure functionality to maintain continuous workflow [6].

The research focuses on leveraging computer vision and machine learning technologies to extract and structure textual information from whiteboard images, creating searchable digital archives that preserve educational content for future reference [7]. By combining automated content capture with intelligent text processing, the system aims to transform traditional whiteboards into smart educational tools without requiring significant infrastructure changes or investment [8].

This approach addresses the growing need for digital integration in educational settings while maintaining the familiar and intuitive nature of conventional whiteboards. The resulting system is designed to facilitate enhanced accessibility, content reuse, and seamless integration between traditional teaching methods and modern digital documentation requirements [9].

2. RELATED WORK

In recent years, the automation of whiteboard and chalkboard management has garnered attention due to its potential to enhance classroom and collaborative efficiency. A range of efforts have explored different facets of this problem, from mechanical board erasing systems to vision-based scribble detection algorithms. However, most existing solutions address these challenges in isolation—focusing either on physical cleaning mechanisms or on content recognition—but rarely both.

A notable mechanical contribution comes from the work titled Automatic Whiteboard Cleaner Using Arduino and IR Sensor, which implements an Arduino-controlled mechanism that detects the presence of a user and automatically moves a brush across the whiteboard using a motorized assembly [10]. While it offers a practical and low-cost erasure mechanism, it is strictly mechanical in nature and does not address any form of content preservation or digitization. This typifies a broader class of embedded solutions that prioritize automation of the erasing process without incorporating any form of content awareness or digital archiving.

Similarly, earlier work by Iinuma et al. proposed a vision-based chalkboard eraser that utilizes Hough Transform techniques to detect chalk scribbles and perform selective cleaning [11]. Although innovative in its partial cleaning approach, the system lacked the ability to extract semantic content, making it unsuitable for educational scenarios that require content archiving. Other hardware-oriented efforts, such as PIC-controlled electronic dusters and Arduino-based remote erasers, have explored various motor control mechanisms—linear actuators, pulley systems, or servo-motor configurations—to enhance the speed and accuracy of board cleaning. Yet these systems remain limited to the mechanical domain and do not engage with the board content in an intelligent manner.

In parallel, advancements in Optical Character Recognition (OCR) have significantly improved the accuracy of text extraction from complex surfaces, including whiteboards. Modern OCR systems—such as Google's Tesseract OCR [12], Microsoft's Read API [13], and domain-specific models trained on handwriting or smudged whiteboard content [14]—have enabled high-quality digitization of visual information. These systems leverage deep learning techniques, convolutional neural networks (CNNs), and attention mechanisms to improve recognition in noisy, low-contrast, or non-standard fonts. However, these OCR systems are typically applied in post-processing scenarios and are not often integrated into real-time embedded systems.

Moreover, while some commercial smart boards offer digital archiving and touchscreen-based input [15], they come with prohibitive costs and require specialized infrastructure. Their reliance on proprietary ecosystems also limits adaptability and scalability, particularly in resource-constrained educational environments. Importantly, such solutions do not address the problem for traditional whiteboards still in widespread use across classrooms, laboratories, and meeting spaces.

Thus, the current state of the art reveals a significant gap: existing solutions tend to address either the mechanical aspect of erasure or the digital archiving of written content, but rarely both in a cohesive, real-time embedded system. Furthermore, there is minimal exploration of integrating large language models (LLMs) into the digitization process to enhance OCR interpretation, error correction, and semantic enrichment of extracted text [16].

The proposed Clarix system aims to fill this void by introducing a novel robotic prototype that unifies these two traditionally separate domains. It features a servo-driven mechanical brush system for automated whiteboard erasing and a camera module for capturing the board content once full. The captured image is processed using advanced OCR techniques, and the text is saved as a searchable PDF, enabling long-term storage, accessibility, and knowledge reuse. Additionally, a benchmarking study is conducted using state-of-the-art LLMs (such as OpenAI's GPT, Anthropic's Claude, and Google's Gemini) to assess their effectiveness in real-world whiteboard digitization scenarios. This integrated approach represents a first-of-its-kind contribution to the field of intelligent educational tools, offering an affordable, scalable, and intelligent solution for analog-to-digital whiteboard transformation.

3. METHODOLOGY

The Clarix system evaluation employed a comprehensive methodology encompassing three core components: hardware implementation, data collection, and computational analysis. To validate the system's effectiveness, a custom embedded prototype was developed to demonstrate the physical automation capabilities, while a specialized dataset of whiteboard images was created to benchmark text extraction performance. The extraction capabilities were then evaluated using both traditional OCR engines and modern LLM-based approaches to determine the optimal digitization strategy for real-world deployment.

A. Embedded System

To physically implement the Clarix system's automation capabilities, a custom-designed embedded hardware prototype was developed. The embedded system is responsible for orchestrating the entire digitization and erasure cycle—from capturing the whiteboard content to initiating the motorized cleaning process. Designed to be cost-effective and easily replicable, the system is built using widely available and low-power electronic components, ensuring accessibility and scalability for educational and collaborative environments.

At the core of the system is a Raspberry Pi 4, which serves as the primary computational unit. This microcomputer handles peripheral interactions, control signals, and system logic. A high-resolution camera module, interfaced with the Raspberry Pi, is mounted at the top of the whiteboard frame to capture images of the content once the board is filled with text, equations, or diagrams. These images are then processed either locally or forwarded to external LLM APIs for intelligent text extraction and archiving, depending on future deployment choices.

The physical cleaning mechanism consists of a servo motor-driven brush mounted on a gear rack system, which allows smooth horizontal movement across the width of the whiteboard. The gear rack ensures precise control and linear motion, enabling a uniform and effective single-pass erasure. The motor is connected to a lightweight brush that spans the width of the board and is guided by wooden side supports, ensuring stability and alignment during motion.

Control of the brush is automated through the Raspberry Pi's GPIO pins, which trigger the servo motor upon completion of the image capture and text extraction steps. To provide real-time user feedback, the system is equipped with LED indicators that signal the system's current state—image capture (yellow), extraction in progress (blue), and cleaning in progress or completed (green). A buzzer serves as an auditory indicator that alerts the user before and after the cleaning process to ensure no disruption or accidental interference occurs.

The system also includes a physical power switch to ensure safe startup and shutdown, along with a manual button that acts as a trigger to initiate the full digitization-cleaning sequence. All connections are made via jumper wires and resistors, integrated on a modular breadboard circuit to allow easy testing, debugging, and future extensions.

The overall prototype is mounted on a compact frame designed to fit a standard-sized whiteboard and is entirely self-contained, with no dependency on external computing systems during operation. This embedded solution bridges the gap between mechanical automation and digital intelligence, offering an affordable yet advanced system for real-time whiteboard content management.

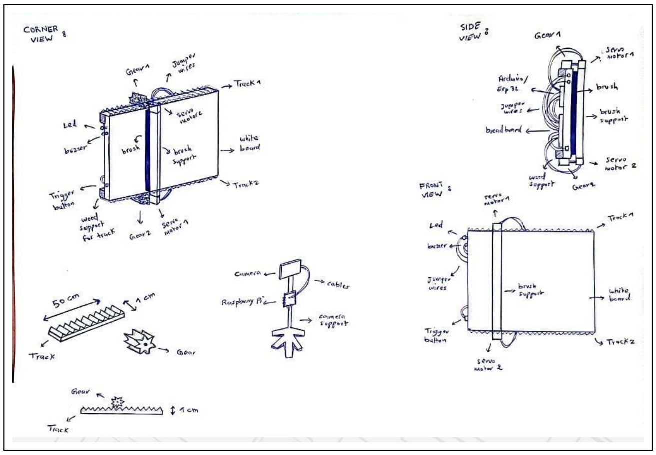

Fig. 1. Clarix prototype sketch showing the Raspberry Pi, servo motor, camera, and peripheral components for automated whiteboard cleaning.

B. Data

To evaluate the performance of both traditional OCR engines and LLM-based APIs in extracting whiteboard content, a custom dataset was created specifically for this study. The dataset consists of high-resolution images captured from a physical whiteboard where various types of content were manually written, including lecture notes, mathematical equations, tabular data, and bullet points. These samples were designed to simulate real-world classroom and meeting scenarios, incorporating common challenges such as variable handwriting styles, lighting reflections, marker inconsistencies, and non-uniform layouts. The images were taken using the camera module of the Clarix prototype under natural and artificial lighting conditions to ensure diversity in input. This dataset served as a benchmark for testing the accuracy and reliability of different text extraction techniques.

Prior to text extraction, all images were subjected to a standardized preprocessing pipeline specifically designed for traditional OCR engines to optimize their performance on handwritten whiteboard content. This preprocessing included image resizing for computational efficiency, Non-Local Means Denoising for noise reduction, and Contrast Limited Adaptive Histogram Equalization (CLAHE) for contrast enhancement. Notably, these preprocessing steps were applied exclusively to traditional OCR systems and not to LLM-based APIs, allowing for evaluation of the latter's native robustness in handling unprocessed whiteboard images.

A representative sample image from the dataset is shown in Figure 2

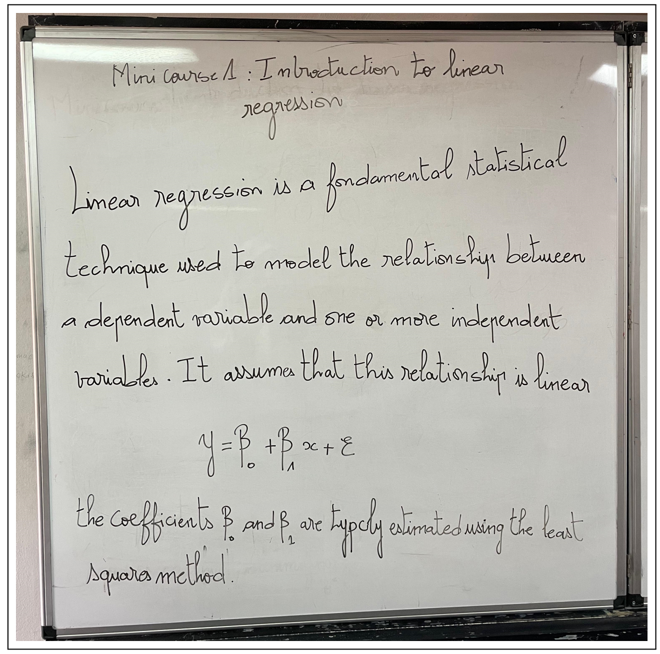

Fig. 2. Sample whiteboard image from the custom dataset, featuring handwritten text and mathematical equations under realistic classroom conditions.

C. Model

C.1. Traditional OCR Systems

Traditional Optical Character Recognition (OCR) systems form the backbone of automated text extraction technology, employing classical computer vision techniques and machine learning algorithms to convert image-based text into machine-readable format. In this study, three prominent OCR engines were utilized to establish baseline performance for whiteboard text extraction.

PaddleOCR PaddleOCR, developed by Baidu's PaddlePaddle team, integrates deep learning architectures with traditional text detection pipelines. The system employs a two-stage approach comprising text detection and recognition modules, utilizing convolutional neural networks optimized for multilingual character recognition. Its lightweight architecture offers computational efficiency while maintaining robust preprocessing capabilities for handling varied text orientations and lighting conditions [17].

Tesseract OCR Tesseract OCR represents the industry standard in open-source OCR technology, originally developed by Hewlett-Packard and maintained by Google. The system combines traditional pattern recognition algorithms with LSTM neural networks, featuring a sophisticated pipeline that includes layout analysis, character segmentation, and statistical language modeling. Its modular architecture allows extensive customization through configuration parameters and supports multiple languages [12, 18].

EasyOCR EasyOCR provides a streamlined deep learning implementation built on PyTorch, incorporating CRAFT (Character Region Awareness for Text detection) for text detection and CRNN (Convolutional Recurrent Neural Network) for character recognition. The system emphasizes out-of-the-box functionality with minimal configuration requirements while demonstrating particular strength in handling low-quality images and irregular text layouts through its modern neural network architectures [19].

C.2. LLM-Based API Extraction

In parallel with traditional OCR techniques, this study investigates the effectiveness of modern Large Language Models (LLMs) for whiteboard content extraction. While OCR engines are designed to recognize textual patterns from images, they often fall short in comprehending structure, context, or ambiguous handwriting—especially in real-world conditions involving varied layouts, mathematical symbols, and inconsistent lighting. With the recent integration of vision capabilities into LLMs, these models have become capable of interpreting not only plain text but also complex visual inputs such as scanned documents, handwritten notes, and whiteboard captures [20, 21].

Given their multi-modal capabilities, we explored whether LLMs could offer a superior alternative for digitizing educational content written on whiteboards, including text, equations, bullet points, and tables. Our objective was to benchmark the performance of these models using the same dataset developed for evaluating OCR tools, and assess whether LLMs could reliably serve as the core intelligence behind Clarix's text extraction pipeline [22].

For this evaluation, five state-of-the-art LLMs with multi-modal capabilities were selected: GPT-4o (OpenAI), Claude 3 (Anthropic), Gemini 1.5 Pro (Google), DeepSeek-Vision (DeepSeek AI), and Grok (xAI). These models were chosen based on their publicly available visual understanding features, popularity, and architectural diversity. Rather than using the API endpoints of each model—which often involve complex setup procedures, account approval, quota limitations, or billing constraints—we opted to use their official user interfaces (web platforms) to ensure a consistent, hassle-free evaluation environment [23].

Each model was accessed through its respective interface (ChatGPT, Claude.ai, Gemini, DeepSeek chat, and X/Twitter AI chat), and tested under identical conditions using a standardized set of whiteboard images. This approach allowed us to focus purely on evaluating extraction performance without being affected by inconsistencies in API behavior or permission issues. It also reflects a real-world use case for educators or developers who might prefer visual input through UI platforms rather than programmatic pipelines. Despite not using APIs during this benchmark phase, our goal is to eventually integrate the most effective model into the Clarix system through API deployment, once a suitable model has been identified [24].

D. Evaluation

To evaluate the performance of large language models (LLMs) and traditional OCR models in interpreting whiteboard content, we conducted a comparative analysis using five distinct LLM models: ChatGPT, Claude, Gemini, Deepseek, and Grok, alongside three traditional Optical Character Recognition (OCR) models: PaddleOCR, EasyOCR, and TesseractOCR. Each model was tested on a consistent dataset of five whiteboard images, covering various educational topics such as mathematics, physics, and economics.

The LLM models received identical prompts and were evaluated based on their ability to accurately reconstruct the textual content present in each image, leveraging their multimodal capabilities to understand both visual and textual elements within the whiteboard context. In contrast, the traditional OCR models relied purely on character recognition algorithms to extract text from the images.

For evaluation precision, recall, and F1 score were used as the primary metrics. These scores were computed by comparing the tokenized output of each model against manually defined ground truth references for each image. Precision measures the proportion of correctly identified tokens among all tokens predicted by the model, recall measures the proportion of correctly identified tokens among all actual tokens in the ground truth, and F1 score represents the harmonic mean of precision and recall, providing a balanced assessment of model performance.

4. RESULTS

The comparative analysis revealed significant performance differences between large language models and traditional OCR approaches in whiteboard content interpretation. Table 1 presents the average precision, recall, and F1 scores across all five whiteboard images for each model.

Table 1. Average Precision, Recall, and F1 Score for Each LLM and OCR Model

Model	Avg Precision	Avg Recall	Avg F1 Score
Gemini	0.8188	0.8784	0.8466
ChatGPT	0.7954	0.8534	0.8162
Deepseek	0.7422	0.8625	0.7965
Grok	0.7372	0.7987	0.7645
Claude	0.7219	0.8034	0.7550
PaddleOCR	0.3523	0.3223	0.3333
EasyOCR	0.1101	0.1309	0.1158
TesseractOCR	0.0000	0.0000	0.0000

To better visualize the F1 Score performance across the evaluated models, the graph below presents a corresponding bar chart.

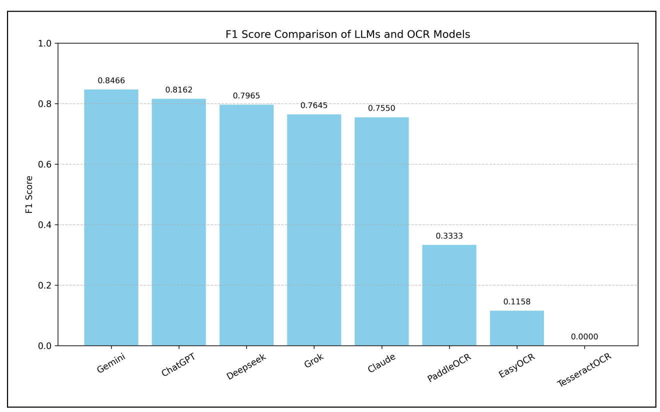

Fig. 3. F1 score comparison of various large language models and OCRs on the text extraction benchmark.

Among the large language models, Gemini achieved the highest F1 score of 0.8466 with precision of 0.8188 and recall of 0.8784. ChatGPT ranked second with an F1 score of 0.8162, precision of

0.7954, and recall of 0.8534. Deepseek recorded the highest recall score of 0.8625 but achieved an F1 score of 0.7965 with precision of 0.7422. Grok obtained an F1 score of 0.7645 with precision of 0.7372 and recall of 0.7987. Claude achieved an F1 score of 0.7550 with precision of 0.7219 and recall of 0.8034.

Among the traditional OCR models, PaddleOCR achieved the highest performance with an F1 score of 0.3333, precision of 0.3523, and recall of 0.3223. EasyOCR recorded an F1 score of 0.1158 with precision of 0.1101 and recall of 0.1309. TesseractOCR recorded zero scores across all metrics.

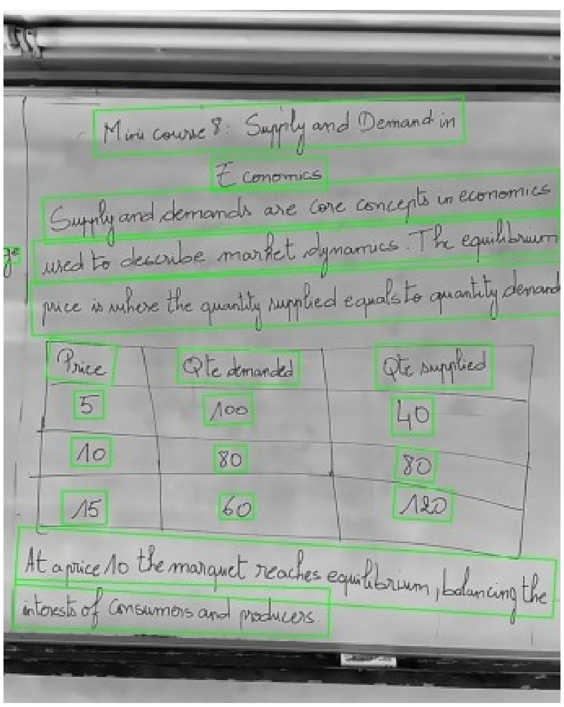

Fig. 4. PaddleOCR Optimized text region detection: 20 regions detected with blue bounding boxes covering complete phrases and text blocks

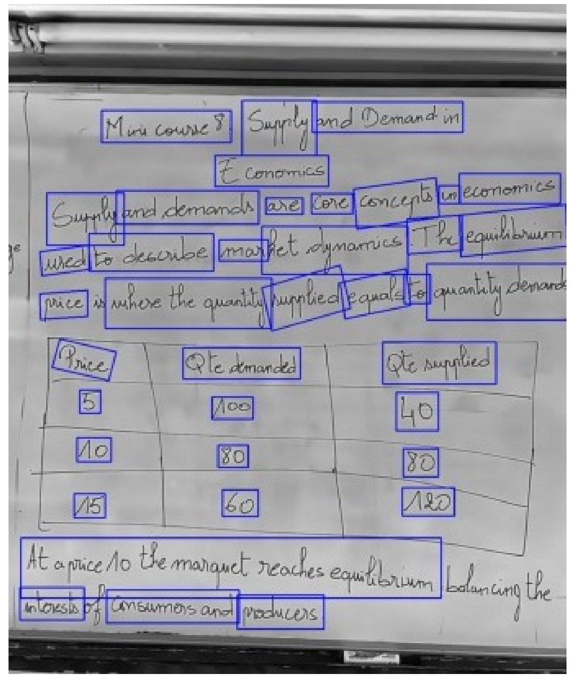

Fig. 5. EasyOCR text region detection: 39 regions detected with blue bounding boxes outlining individual words and text elements

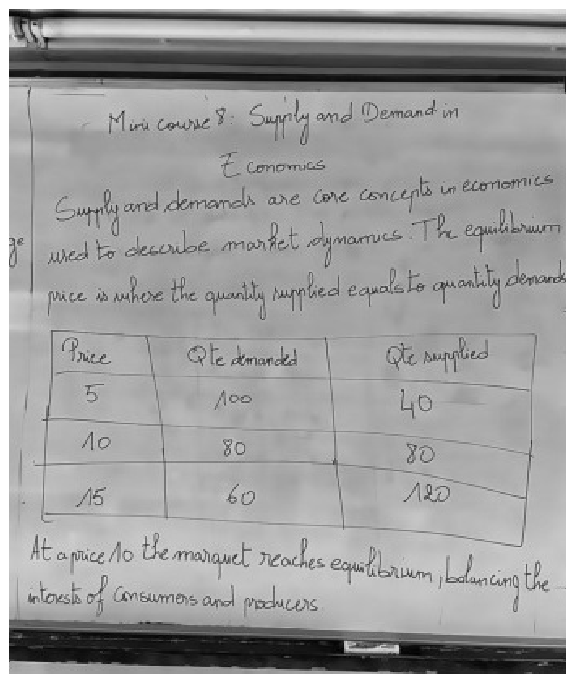

Fig. 6. TesseractOCR text region detection: 0 regions detected

Figures 5,6, and 4 display the text region detection capabilities of the three traditional OCR models. TesseractOCR detected 0 text regions across the whiteboard images. EasyOCR identified 39 text regions with blue bounding boxes outlining individual words and text elements. PaddleOCR Optimized detected 20 text regions with blue bounding boxes covering complete phrases and text blocks

5. DISCUSSION

The experimental results reveal substantial performance disparities between multimodal large language models and traditional OCR systems when applied to whiteboard content extraction. The superior performance of LLMs across all evaluation metrics demonstrates the advantages of multimodal understanding capabilities over conventional character recognition approaches [25, 26].

Among the LLM models, Gemini's leading performance (F1 = 0.8466) can be attributed to its balanced precision and recall scores, indicating robust text extraction with minimal false positives and negatives. The relatively close performance clustering of all LLM models (F1 scores ranging from 0.7550 to 0.8466) suggests that current multimodal LLMs have reached a comparable level of maturity in visual text understanding tasks [25]. Notably, Deepseek achieved the highest recall score (0.8625), demonstrating superior capability in identifying existing text content, though this came at the cost of precision.

The stark performance gap between LLMs and traditional OCR models highlights fundamental limitations in conventional text recognition approaches when handling complex visual scenarios. PaddleOCR's relatively superior performance among OCR models (F1 = 0.3333) correlates with its text region detection strategy, which identifies 20 coherent text regions corresponding to complete phrases and sentences. This approach enables more contextual text extraction compared to word-level detection [27].

Conversely, EasyOCR's detection of 39 text regions reflects a granular, word-by-word approach that, while identifying more individual text elements, fails to maintain contextual relationships between words within phrases or sentences. This fragmented detection strategy likely contributes to its lower overall performance (F1 = 0.1158) despite detecting nearly twice as many regions as PaddleOCR. The word-level segmentation approach appears less effective for whiteboard scenarios where maintaining textual context and structure is crucial for accurate content reconstruction [25].

TesseractOCR's complete failure (0 detected regions, 0 scores across all metrics) indicates fundamental incompatibility with handwritten whiteboard content. This suggests that TesseractOCR's optimization for printed text renders it unsuitable for educational whiteboard scenarios containing handwritten mathematical expressions, diagrams, and informal notation styles [27].

The region detection analysis reveals that mere quantity of detected regions does not correlate with extraction quality. EasyOCR's higher region count paradoxically led to inferior performance compared to PaddleOCR's more selective but contextually coherent region identification. This finding underscores the importance of semantic understanding in text extraction tasks, where maintaining phrase-level and sentence-level coherence proves more valuable than exhaustive word-level detection [25].

The multimodal LLMs' consistent outperformance suggests that their training on diverse visual-textual datasets enables better handling of whiteboard-specific challenges, including handwritten text, mathematical notation, informal layouts, and mixed content types. Unlike traditional OCR systems that rely solely on pattern matching and character recognition algorithms, LLMs can leverage contextual understanding to infer missing or unclear characters and maintain semantic coherence across extracted text [26, 28].

These findings have practical implications for educational technology applications, where accurate whiteboard content extraction is essential for automated note-taking, content digitization, and accessibility enhancement. The results suggest that current multimodal LLMs provide viable solutions for such applications, while traditional OCR approaches remain inadequate for complex whiteboard scenarios [29].

Future research directions should explore the specific factors contributing to LLM superiority, investigate hybrid approaches combining OCR region detection with LLM contextual understanding, and evaluate performance across diverse whiteboard content types and writing styles [25].

6. CONCLUSION

Clarix demonstrates an effective, low-cost solution for automating the digitization and cleaning of whiteboard content using embedded systems and AI technologies. By integrating a Raspberry Pi, OCR, and a servo-driven cleaning mechanism, the system bridges the gap between analog and digital workflows. The evaluation of five advanced Large Language Models on diverse whiteboard content provides valuable insights into their performance in real-world scenarios, highlighting GPT-4o as the most reliable across different content types. Clarix not only streamlines classroom and meeting room documentation but also lays the groundwork for scalable deployments in educational and collaborative settings. Future work will explore real-time transcription, multi-language support, and cloud integration to further enhance the system's capabilities.

REFERENCES

1. M. Johnson and S. Lee, J. Educ. Tools **12**, 45 (2019).
2. R. Kumar and A. Patel, Int. J. Collab. Technol. **8**, 22 (2020).
3. J. Smith, "Challenges in manual transcription of whiteboard content," in *Proceedings of the 2018 Conference on Educational Technology*, (2018), pp. 112–117.
4. T. Nguyen and L. Garcia, Technol. Educ. Rev. **15**, 78 (2017).
5. Y. Chen and H. Wang, J. Artif. Intell. Educ. **29**, 150 (2021).
6. M. Lopez and P. Singh, IEEE Trans. on Learn. Technol. **15**, 10 (2022).
7. X. Zhao and J. Kim, Comput. Educ. **150**, 103837 (2020).
8. R. Martinez and D. O'Connor, Educ. Technol. Soc. **22**, 45 (2019).
9. S. Patel and M. Nguyen, Int. J. Educ. Technol. **18**, 5 (2021).
10. R. Singh and A. Patel, Int. J. Res. Eng. Sci. (IJRES) **10**, 12 (2022).
11. T. Iinuma, M. Kaneko, and T. Hirai, "Selective erasure of chalkboards based on line detection using hough transform," in *Proceedings of the International Conference on Pattern Recognition*, (2005), pp. 785–788.
12. R. Smith, Google Res. (2007). Accessed: 2025-06-04.
13. Microsoft Azure Cognitive Services, "Read api documentation," https://learn.microsoft.com/en-us/azure/cognitive-services/computer-vision/overview-ocr (2023). Accessed: 2025-06-04.
14. Y. Zhou, X. Li, and H. Wang, ACM Comput. Surv. **53**, 1 (2020).
15. L. Nguyen and Y. Chen, J. Interact. Learn. Res. **29**, 123 (2018).
16. A. Kumar and M. Zhao, Trans. on Artif. Intell. Syst. **11**, 301 (2023).
17. PaddlePaddle Team, "Paddleocr documentation," https://paddlepaddle.github.io/PaddleOCR/main/en/index.html (2025). Accessed: 2025-06-04.
18. J. Hornik, "Using the tesseract ocr engine in r," https://cran.r-project.org/web/packages/tesseract/vignettes/intro.html (2025). Accessed: 2025-06-04.
19. K. Smelyakov *et al.*, "Effectiveness of modern text recognition solutions and tools for ocr," in *CEUR Workshop Proceedings*, (2021). Accessed: 2025-06-04.
20. Klippa, "Llms vs ocr data extraction: Which one should you use?" https://www.klippa.com/en/blog/information/llms-vs-ocr-software/ (2025). Accessed: 2025-06-04.
21. Anonymous, arXiv preprint arXiv:2501.11623 (2025). Accessed: 2025-06-04.

22. Koncile, "Llm ocr: why are they better than regular ocrs?" https://www.koncile.ai/en/ressources/why-llm-ocr-are-better-than-regular-ocrs (2025). Accessed: 2025-06-04.
23. Mindee, "Llm vs ocr api: Cost comparison for document processing in 2025," https://www.mindee.com/blog/llm-vs-ocr-api-cost-comparison (2025). Accessed: 2025-06-04.
24. Vellum AI, "Document data extraction in 2025: Llms vs ocrs," https://www.vellum.ai/blog/document-data-extraction-in-2025-llms-vs-ocrs (2025). Accessed: 2025-06-04.
25. F. Seitl, T. Kovářík *et al.*, arXiv preprint arXiv:2404.04068 (2024).
26. A. Grothey *et al.*, Nat. Commun. (2025).
27. Anonymous, Front. Artif. Intell. (2023).
28. Anonymous, Chem. Soc. Rev. (2025).
29. Anonymous, Radiol. Artif. Intell. (2023).