GradeZen: automated grading ecosystem using deep learning for educational assessments

Murugavalli Elangovan, Rajeswari Kaleeswaran, Kathirvel Mohankumar, Shreya Rangachari, Ubasini Manivannan, Rishapa Pandidurai, Maria Bellarmin Joel, Preethi Meenatchi Murugesan
Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, India

Article Info

Article history:

Received Sep 2, 2024
Revised Feb 6, 2025
Accepted Mar 15, 2025

Keywords:

Automation
EAST
Education 4.0
Grading process
Mixtral 8x7B
TrOCR
YOLOv8

ABSTRACT

This study introduces a groundbreaking software solution poised to revolutionize grading procedures in higher education through advanced artificial intelligence and machine learning techniques. Leveraging cutting-edge technologies such as YOLOv8 for real-time object detection, transformer-based optical character recognition (TrOCR), and Mixtral 8x7B transformer models for complex data analysis, the software automates the grading process. By significantly reducing the time and effort required for manual grading, it aims to streamline educational practices while ensuring consistency and scalability. The study provides a comprehensive analysis of use cases, identifies key issues in current grading methods, and elucidates the rationale driving its development. This innovative approach holds immense promise for transforming educational practices, fostering student success through efficient and artificial intelligence assisted automated assessment methodologies.

Corresponding Author:

Rajeswari Kaleeswaran
Department of Electronics and Communication Engineering, Thiagarajar College of Engineering
Madurai-15, Tamilnadu, India
Email: rajeswari@tce.edu

1. INTRODUCTION

In contemporary educational discourse, the symbiotic relationship between education and examination stands as a pivotal focus for scholarly inquiry. The importance of education lies in its transformative power, serving as the conduit through which individuals acquire knowledge, hone skills, and cultivate critical thinking abilities. Examination, as a mechanism for evaluation within the educational framework, assumes a dual role: it acts as a barometer of student comprehension and performance while also serving as a catalyst for pedagogical refinement and institutional accountability [1]. The examination process not only assesses academic achievement but also fosters a culture of diligence, perseverance, and intellectual inquiry among students. Thus, the intersection of education and examination emerges as a rich terrain for research, offering insights into best practices for assessment, the impact of examination on student motivation and learning outcomes, and the role of examinations in shaping educational policies and practices [2]. By delving into this dynamic relationship, researchers can illuminate the complexities of educational assessment, inform evidence-based instructional strategies, and contribute to the ongoing discourse on educational excellence and equity [3]. In the realm of educational assessment, manual examination grading remains a foundational practice entrenched in academic tradition. The reliance on human assessors introduces subjectivity and variability into the grading process, potentially compromising the fairness and consistency of evaluations. Moreover, manual grading is labor-intensive and time-consuming, placing significant burdens on educators and limiting their capacity for innovative pedagogical endeavors [4]. Through automated scanning and processing of answer sheets, optical mark recognition (OMR) systems ensure consistency and fairness in evaluations, minimizing the impact of human bias [5]. Furthermore, OMR grading streamlines the grading process, significantly reducing the time and labor required for assessment tasks. OMR grading presents a promising solution primarily tailored for objective-type questions. However, this specificity inherently limits its applicability, as OMR systems are ill-equipped to evaluate a brief answer type of questions. By leveraging cutting-edge technologies such as artificial intelligence and machine learning, we can revolutionize the assessment process, enhancing efficiency, accuracy, and scalability [6]–[23]. Recognizing these challenges, the development of innovative software applications integrated with artificial intelligence to automate grading workflows has emerged as a promising solution.

This research paper delves into the analysis of a ground-breaking software application poised to revolutionize how college professors and teachers assign grades. The software in question harnesses the formidable capabilities of cutting-edge technologies like YOLOv8, transformer based optical character recognition (TrOCR), and Mixtral 8x7B transformer models, marking the dawn of a new era characterized by enhanced efficiency and precision in grading procedures. By automating pivotal aspects of the grading process, this software streamlines manual assessment, liberating educators to invest their time and energy into more substantive interactions with students. This manuscript embarks on a voyage through the myriad dimensions of the software's genesis and execution, commencing with a comprehensive exploration of user requisites and the underlying logic guiding its conception. It then proceeds to illuminate the technical intricacies inherent in YOLOv8, TrOCR, and Mixtral 8x7B transformer models, elucidating how these technological marvels underpin the automation of grading workflows. The Mixtral 8x7B is one of the top most model ranking 17th place in large model systems (LMSYS) chatbot arena leader-board. The major contributions of this paper can be summarized as follows:

– By utilizing the technologies, such as TrOCR and Mixtral 8x7B transformer models the software significantly enhances efficiency and accuracy, marking a significant stride forward in educational technology.
– A robust solution tailored to the needs of educators and educational institutions.

The rest of the paper is organized as follows: section 2 deals with the proposed methodology followed by implementation in section 3. Results are discussed in section 4. Finally, section 5 concludes the work.

2. PROPOSED METHODOLOGY

The proposed solution aims to revolutionize the grading process for college instructors through the development of a comprehensive software system. This system integrates three key subsystems-image-to-text conversion, picture segmentation, and natural language processing (NLP) for grading-to provide a robust and efficient solution for assessment automation. The Figure 1 describes the functional flow of the solution proposed.

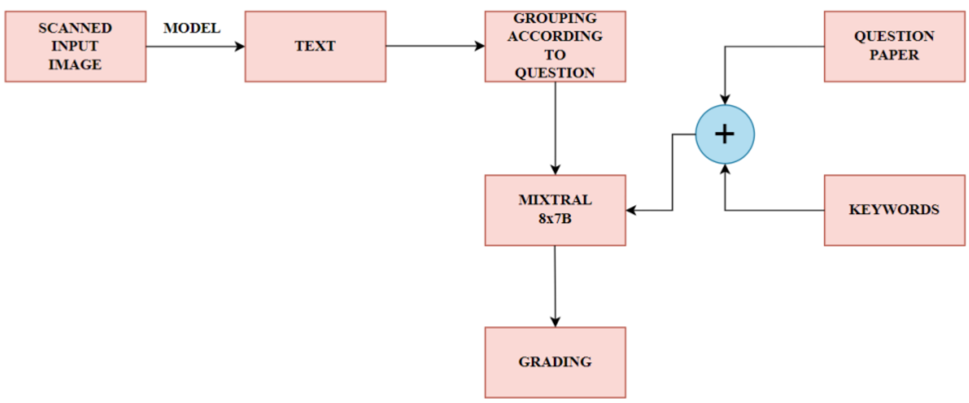

Figure 1. Functional flow of proposed system

2.1. Text detection and segmentation

The image is segmented and structured using the text detection and segmentation subsystem. This component employs YOLOv8-text detector to identify and segment individual components within the text, such as paragraphs, sentences, and individual responses. By segmenting the image, this subsystem prepares the data for further analysis and evaluation.

2.2. Text recognition

From the segmented images, the text recognition subsystem meticulously processes the images and converts them into a text format. This subsystem utilizes advanced optical character recognition (OCR) technology. This enables accurate and reliable extraction of text from scanned documents.

2.3. Natural language processing for grading

The graded responses are then evaluated using the NLP grading subsystem. This component utilizes state-of-the-art NLP algorithms to analyze and assess each response against the provided answer key. By leveraging contextual understanding and linguistic analysis, the NLP grading subsystem ensures accurate and consistent grading of student responses.

2.4. Integration and presentation

The system seamlessly integrates the outputs from each subsystem to compile and present the grades for each student. Instructors can easily access and review the graded responses, facilitating efficient feedback, and evaluation. By automating key aspects of the grading process, this integrated solution minimizes the time and effort required for manual assessment, enabling instructors to focus on more meaningful interactions with students.

3. IMPLEMENTATION

The proposed system integrates three essential subsystems-text detection and segmentation, text recognition, and NLP for grading-to offer a multifaceted approach to assessment automation. Initially, the text detection and segmentation subsystem identify and isolates text regions within scanned documents. Among the text detection model such as efficient and accurate scene text detection (EAST) [24], character region awareness for text detection (CRAFT) [25], and the YOLOv8 [26]–[29] model had been chosen for its better performance. Subsequently, the text recognition subsystem processes these segmented text regions, and convert them into a digital format. This subsystem can be developed by utilizing the models such as Paddle-OCR [30]–[32], Keras-OCR [33], [34], Tesseract-OCR [35]–[39], and TrOCR [40], [41]. The TrOCR is utilized in our paper because of its better accuracy. Finally, the NLP for grading subsystem evaluates the recognized text against predefined criteria, utilizing sophisticated linguistic analysis techniques Mixtral 8x7B [42]–[44] to provide accurate and consistent grading. Through the seamless integration of these subsystems, the system streamlines the grading process, enhances efficiency, and ensures the reliability of assessment outcomes. Figure 2 describes the use case of the system. The System involves two actors: a professor and a system Mixtral 8x7B. The professor interacts with the system through several actions: login, upload questions, upload answer sheet, view marks, and logout. The TrOCR does the image to text conversion. The system, Mixtral 8x7B, compares answers, gives marks, and answers the question.

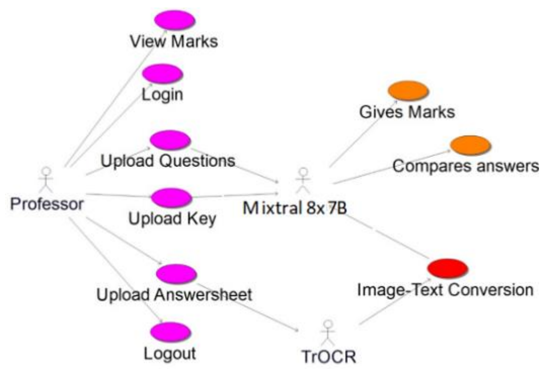

Figure 2. Use-case of the proposed system

3.1. YOLOv8

As a state-of-the-art model that builds on the successes of its predecessors while providing new features and refinements to push performance and adaptability to new heights, Ultralytics YOLOv8 is the epitome of innovation in the world of computer vision. With its foundation rooted in previous YOLO versions, YOLOv8 is meticulously engineered to embody speed, accuracy, and ease of use, positioning itself as a premier choice for a diverse array of computer vision tasks. Whether it's object detection and tracking, instance segmentation, image classification, or pose estimation, YOLOv8 showcases unparalleled proficiency, empowering users with the tools needed to tackle complex visual recognition challenges with confidence and precision. Furthermore, the integration of pretrained data further amplifies the capabilities of YOLOv8, imbuing it with a wealth of knowledge and experience garnered from extensive training on vast datasets. This pretrained data serves as a catalyst for accelerated learning and adaptation, enabling YOLOv8 to swiftly recognize and categorize objects across a myriad of contexts and environments. By leveraging pretrained data, YOLOv8 not only expedites the deployment process but also enhances its ability to generalize and extrapolate from limited training samples, ultimately leading to more robust and reliable performance in real-world scenarios. YOLOv8 is a computer vision innovation leader, combining state-of-the-art technology with pretrained data in a way never seen before, potentially changing the way humans perceive, understand, and work with visual data. Figure 3 compares the mean average precision (mAP) values of several YOLO models and displays the low latency of YOLOv8.

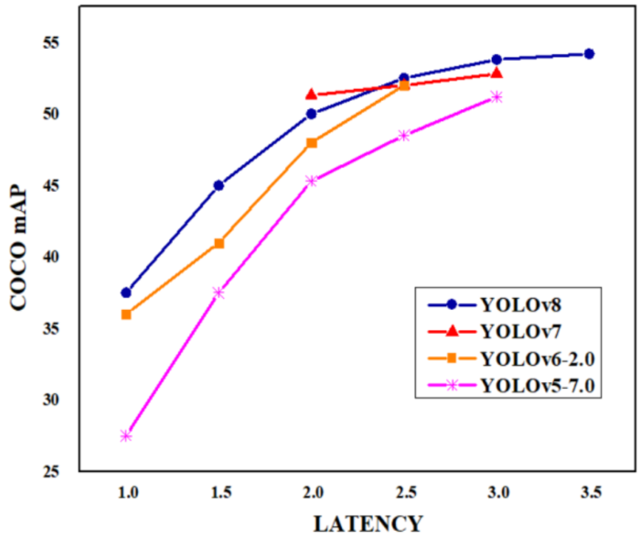

Figure 3. Comparing mAP of different YOLO models

3.2. Transformer optical character recognition

TrOCR does not rely on convolutional neural network (CNN) backbones but uses transformer architecture for both image understanding and text generation. TrOCR consists of an image transformer for visual features extraction and a text transformer for language modeling. The encoder processes image patches, while the decoder generates wordpiece sequences. Input image is resized and decomposed into patches for special tokens like [CLS] and distillation token are used for image representation and learning from pre-trained models. The TrOCR models surpassed the current state-of-the-art methods on the leaderboard of the scanned receipts OCR and key information extraction (SROIE) dataset, demonstrating superior performance in capturing visual information and language modeling without complex pre/post processing steps. Transformer-based text recognition models exhibited competitive performance compared to CNN and recurrent neural network (RNN)-based networks, reaffirming the effectiveness of transformer structures [40], [41]. This gives significant results on the identity and access management (IAM) handwriting database, highlighting the effectiveness of methods with CTC decoders and the impact of external language models on performance. Table 1 gives the performance analysis of different TrOCR model. Competitive performance was achieved, indicating the efficacy of the implemented approaches in addressing the challenges of handwritten text recognition. The SROIE dataset was evaluated using word-level precision, recall, and F1-score, while the IAM dataset was assessed based on the character error rate (CER) and scene text datasets by word accuracy. The study employed rigorous evaluation criteria to assess the accuracy and efficacy of the implemented models.

Table 1. Performance analysis of various TrOCR models

Performance metrices (%)	Small	Base	Large
Recall	96.9	96.27	96.58
Precision	95.63	96.41	96.56
F1-score	95.92	96.35	96.59

3.3. Mixtral 8x7B

Mixtral 8x7B emerges as a cutting-edge sparse mixture of expert models (SMoE) with open weights, representing a significant leap forward in model architecture. Surpassing Llama 2 70B in most benchmarks while boasting a 6x faster inference rate, Mixtral 8x7B sets a new standard for performance and efficiency. Notably, this model excels in handling 32k-token contexts and supports multiple languages, including English, French, Italian, German, and Spanish. Its prowess extends beyond NLP, as it demonstrates proficiency in coding tasks. With fine-tuning, Mixtral 8x7B has the potential to become an instruction-following model, achieving an impressive 8.3 score on the MT-Bench. Furthermore, its compatibility with existing optimization tools such as flash attention 2, bits and bytes, and parameter-efficient fine-tuning (PEFT) libraries streamlines integration and deployment processes. The architecture of Mixtral employs a sparse mixture of expert (MoEs) model, showcasing a sophisticated setup where each token undergoes processing by a specific expert. In the case of Mixtral-8x-7B, the complexity is heightened, featuring 8 experts, with 2 experts allocated for each token. Through a specialized router network, 2 of the 8 experts are selected to process each token, with their outputs merged additively. The MoE methodology, applied selectively to MoE layers rather than self-attention weight matrices, effectively reduces the total parameter count, estimated to be around 40-50 B. The efficiency of Mixtral lies in its router functionality, directing tokens so that only 7B parameters are engaged during the forward pass, significantly expediting both training and inference processes compared to traditional non-MoE models. This selective engagement of parameters underscores the efficacy of MoE-based approaches, such as Mixtral, in achieving unparalleled efficiency and performance in complex language processing tasks. Figure 4 describes the performance of the large language models (LLM) for different benchmarks. The graph infers that the Mixtral 8x7B model has better performance than other LLM models.

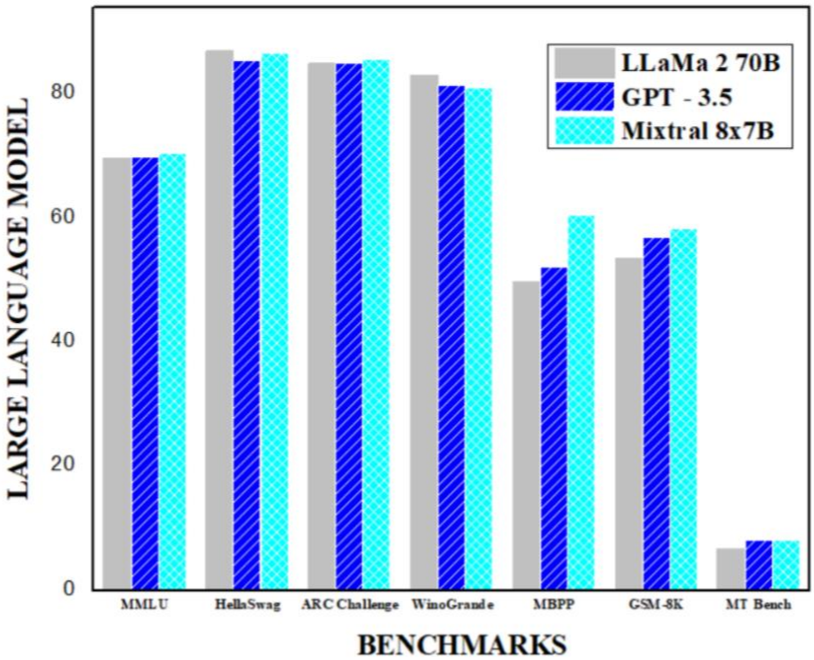

Figure 4. Performance of Mixtral 8x7B compared with other popular LLM models

4. RESULTS AND DISCUSSION

The YOLOv8 model is trained with 1280 data images. The YOLOv8 model is custom trained for the process of object detection of the handwritten sentences. The model performance is analyzed and validated with testing data and the optimum threshold is found to be 0.314. This system identifies and separates individual sentences from answer sheets, then saves them as separate images in the database. Also, the same process is carried out using EAST and CRAFT text detectors. Among the three the YOLOv8 model is chosen for text detection and segmentation subsystem. Table 2 describes the performance of YOLOv8 model. Figure 5 shows the performance analysis of YOLOv8 in a graphical format.

Table 2. Performance analysis of YOLOv8 model in comparison with other models

Models	Recall (%)	Precision (%)	F1-score (%)	Latency (ms)
EAST	74.48	90.26	81.61	80
CRAFT	84.3	89.8	86.9	116.3
YOLOv8	94	78.4	78	20

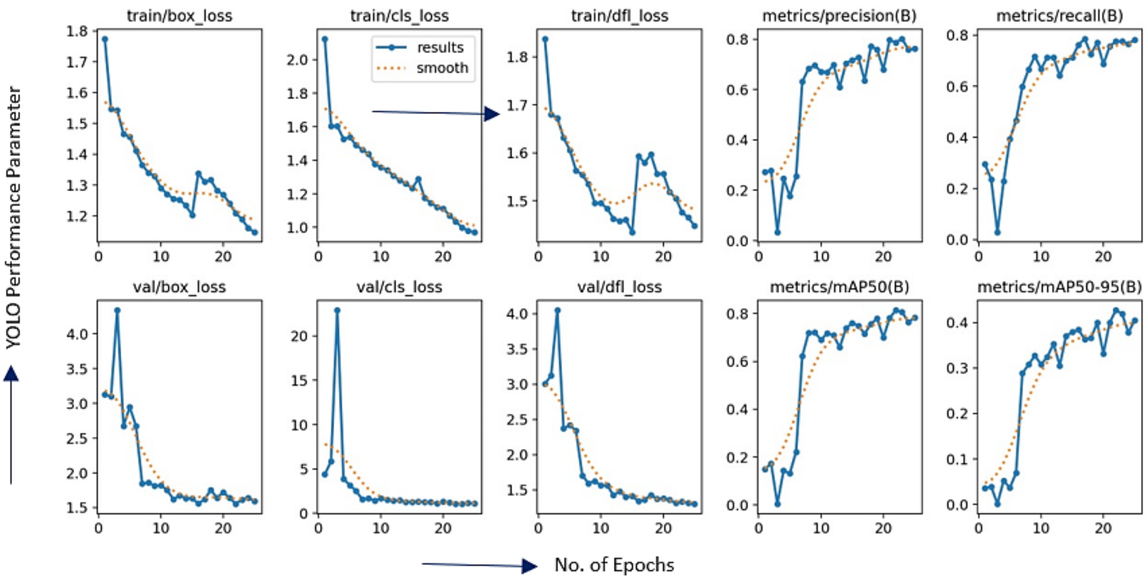

Figure 5. Performance analysis of YOLOv8 model (YOLO performance parameter for each graph is shown in the top of each graph)

The cropped figures of handwritten sentences which were stored in the database were fed as input to TrOCR model in proper order for text recognition. The output strings are combined together resulting in a successful extraction of text from the given handwritten answer sheet. This string along with a previously fed question and answer-key is formulated as a prompt. Figure 6 is the result obtained for a test data.

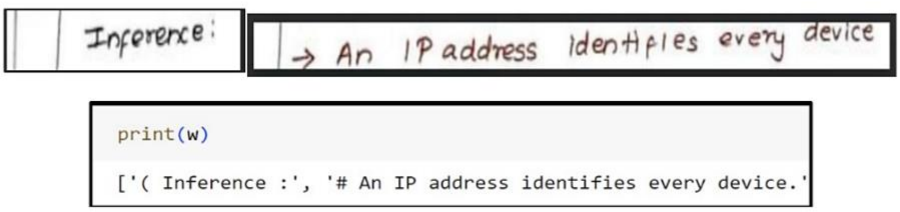

Figure 6. TrOCR output

The framed prompt is given as input to the Mixtral 8x7B model. The model evaluates the written answer and the answer in the given answer-key and allots the score for that question. Unlike other methods of automatic grading, this LLM model ignores CER that could have resulted from TrOCR inference. It also ignores missing words that could have resulted from the text detection process. It does this by understanding the underlying contextual meaning, that it finds from the prompt consisting of question, answer-key, written answer, and its own reasoning ability. So, the accuracy of the entire model is increased, even if the OCR process fails at times. The level of grading such as lenient and strict evaluation can be predefined and the model can infer the feedback of the student's performance as a separate output for better understanding of each student's needs. Figures 7 and 8 shows the sample prompt and Figure 9 is the output obtained from the system.

Give me only the marks+feedback according to the format you would give out of {Total marks} for the answer:"{answer}" to the question:"{question}" based on the answer key:"{answer key}" .type in only the marks+feedback. give me the output in this format: given corrected marks/{Total marks}+short feedback for this student like a teacher to improve them.

Figure 7. Sample prompt

Low level complexity with no Character Error Rate(CER)

qn:write about rose flower
answer key: Rose, beauty, love, affection, symbol

written answer:The rose flower is celebrated for its exquisite beauty and is often considered a symbol of love and affection.
given marks: 90/100

written answer:The sunflower stands tall and bright, its golden petals a radiant emblem of joy and vitality.
given marks: 0/100

High level complexity with High Character Error Rate(CER)

qn: write about IPAddress and computer networks
answer key: IP address, device identification, internet communication, mobile phones, IoT devices, Teredo, ISATAP, IPv6 connectivity, Windows 7, physical address, hexadecimal, hostname, nslookup command.

written answer: An IP address identifies every device. IEEE connected to the internet. This enables computers and other internet-connected devices, such as mobile phones and lot devices, to communicate over the internet. Both Teredo and I SATAP are mechanisms for IPVB connectivity on IPru-only network since Windows 7 comes with 1Pv6 enabled, I wouldn't be surprised if they were always on, just didn't appear in impumping for some strange reason. Physical address is displayed to 48 bit besca decimal value. Hostname is nothing but the name of the composer. nslookup command is used.
given marks: 70/100

written answer: Firewalls are essential components of network security, acting as barriers that monitor and control the incoming and outgoing network traffic based on predetermined security rules. They serve as the first line of defense in network security by preventing unauthorized access, filtering out unwanted traffic, and blocking potential threats. By distinguishing between secure and unsafe data, firewalls ensure that sensitive information within a network remains protected from various cyber threats
given marks: 0/100

qn: write about IPAddress and computer networks
answer key: IP address, device identification, internet communication, mobile phones, IoT devices, Teredo, ISATAP, IPv6 connectivity, Windows 7, physical address, hexadecimal, hostname, nslookup command.

written answer: An IP address identifies every device. IEEE connected to the internet. This enables computers and other internet-connected devices, such as mobile phones and lot devices, to communicate over the internet. Both Teredo and I SATAP are mechanisms for IPVB connectivity on IPru-only network since Windows 7 comes with 1Pv6 enabled, I wouldn't be surprised if they were always on, just didn't appear in impumping for some strange reason. Physical address is displayed to 48 bit besca decimal value. Hostname is nothing but the name of the composer. nslookup command is used.
given marks: 70/100. Good attempt, but some inaccuracies and irrelevant information were included. Focus on the specific question and provide only relevant details. For example, Teredo and ISATAP are not required for IPv6 connectivity on IPv4-only networks, and physical address has no relation to IP addresses or computer networks. Also, explain the use of the nslookup command.

Figure 8. Sample prompt with level of complexity

```
User: give me only the marks according to the

Mixtral: 80/100
```

Figure 9. Output from the grading system

5. CONCLUSION

This research paper presents a pioneering software application poised to redefine the assessment landscape in education. By leveraging advanced technologies such as YOLOv8, TrOCR, and Mixtral 8x7B, the proposed system offers unparalleled efficiency and accuracy in automated grading. Moreover, the system's ability to provide personalized feedback and streamline assessment procedures signifies a paradigm shift in educational practices. As technology continues to evolve, the proposed system stands as a testament to the transformative power of innovation in education, promising a future characterized by efficiency, accuracy, and equity in assessment.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Murugavalli Elangovan	✓	✓		✓	✓	✓	✓			✓		✓	✓	
Rajeswari Kaleeswaran	✓	✓		✓	✓	✓	✓			✓		✓	✓	
Kathirvel Mohankumar		✓	✓	✓			✓		✓		✓		✓	
Shreya Rangachari		✓	✓				✓		✓				✓	
Ubasini Manivannan		✓		✓			✓		✓		✓		✓	
Rishapa Pandidurai		✓			✓		✓	✓	✓				✓	
Maria Bellarmin Joel			✓			✓	✓	✓	✓					
Preethi Meenatchi Murugesan						✓		✓	✓					

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author [RK] on request.

REFERENCES

[1] R. E. Stake, “The countenance of educational evaluation,” *Teachers College Record: The Voice of Scholarship in Education*, vol. 68, no. 7, pp. 1–15, 1967, doi: 10.1177/016146816706800707.

[2] I. Szőköl, K. Szarka, and J. Hargaš, “The functions of educational evaluation,” *R&E-SOURCE: Online Journal for Research and Education*, no. 24, pp. 80–86, 2022, doi: 10.53349/resource.2022.iS24.a1112.

[3] S. Talukder, “New trends in evaluation and their significance in the field of present education system,” *Psychology and Education Journal*, vol. 58, no. 1, pp. 2255–2259, 2021, doi: 10.17762/pae.v58i1.1103.

[4] C. Geigle, C. Zhai, and D. C. Ferguson, “An exploration of automated grading of complex assignments,” in *Proceedings of 2016 the Third ACM Conference on Learning @ Scale*, 2016, pp. 351–360, doi: 10.1145/2876034.2876049.

[5] A. E. Tümer and Z. Küçükkara, “An image processing oriented optical mark recognition and evaluation system,” *International Journal of Applied Mathematics, Electronics and Computers*, vol. 6, no. 4, pp. 59–64, 2018, doi: 10.18100/ijamec.2018447788.

[6] T. Wang, X. Su, Y. Wang, and P. Ma, “Semantic similarity-based grading of student programs,” *Information and Software Technology*, vol. 49, no. 2, pp. 99–107, 2007, doi: 10.1016/j.infsof.2006.03.001.

[7] J. Shrungare, “AI in education,” *XRDS: Crossroads, The ACM Magazine for Students*, vol. 29, no. 3, pp. 63–65, 2023, doi: 10.1145/3589657.

[8] V. Tanwar, “Machine learning based automatic answer checker imitating human way of answer checking,” *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 12, pp. 101–107, 2021.

[9] D. Ramesh and S. K. Sanampudi, “An automated essay scoring systems: a systematic literature review,” *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2495–2527, 2022, doi: 10.1007/s10462-021-10068-2.

[10] S. M. Darwish and S. K. Mohamed, “Automated essay evaluation based on fusion of fuzzy ontology and latent semantic analysis,” *Advances in Intelligent Systems and Computing*, vol. 921, pp. 566–575, 2020, doi: 10.1007/978-3-030-14118-9_57.

[11] P. Doke, P. Gangane, K. S. Babu, P. Lagad, and N. Vaidya, “Online subjective answer verifying system using artificial intelligence,” *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, vol. 10, no. 11, pp. 1128–1131, 2022, doi: 10.22214/ijraset.2022.47543.

[12] H. M. Balaha and M. M. Saafan, “Automatic exam correction framework (AECF) for the MCQS, essays, and equations matching,” *IEEE Access*, vol. 9, pp. 32368–32389, 2021, doi: 10.1109/ACCESS.2021.3060940.

[13] B. M. Pushpa, K. S. Sandhya, S. R. Subedar, and M. G. Aruna, “Automatic exam correction framework for short written answers using NLP,” *International Research Journal of Modernization in Engineering Technology and Science*, vol. 4, no. 7, pp. 3942–3951, 2022.

[14] H. K. Ault and A. Fraser, “A comparison of manual vs. online grading for solid models,” in *120th ASEE Annual Conference and Exposition*, 2013, pp. 1–13, doi: 10.18260/1-2--19045.

[15] N. Süzen, A. N. Gorban, J. Levesley, and E. M. Mirkes, “Automatic short answer grading and feedback using text mining methods,” *Procedia Computer Science*, vol. 169, pp. 726–743, 2020, doi: 10.1016/j.procs.2020.02.171.

[16] P. S. Annor, E. Kayang, S. Boateng, and G. Boateng, “AutoGrad: automated grading software for mobile game assignments in SuaCode courses,” *ACM International Conference Proceeding Series*, pp. 79–85, 2021, doi: 10.1145/3507923.3507954.

[17] M. Mohler and R. Mihalcea, “Text-to-text semantic similarity for automatic short answer grading,” in *EACL 2009 - 12th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings*, 2009, pp. 567–575, doi: 10.3115/1609067.1609130.

[18] D. M. Rao, “Experiences with auto-grading in a systems course,” in *2019 IEEE Frontiers in Education Conference (FIE)*, 2019, pp. 1–8, doi: 10.1109/FIE43999.2019.9028450.

[19] H. Simanjuntak, “Proposed framework for automatic grading system of ER diagram,” in *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2015, pp. 141–146, doi: 10.1109/ICITEED.2015.7408930.

[20] S. Srikant and V. Aggarwal, “Automatic grading of computer programs: a machine learning approach,” in *2013 12th International Conference on Machine Learning and Applications*, 2013, pp. 85–92, doi: 10.1109/ICMLA.2013.22.

[21] J. Muangprathub, O. Shichim, Y. Jaroensuk, and S. Kajornkasirat, “Automatic grading of scanned multiple choice answer sheets,” *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 2, pp. 175–179, 2018, doi: 10.14419/ijet.v7i2.23.11910.

[22] S. Srikant and V. Aggarwal, “A system to grade computer programming skills using machine learning,” in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014, pp. 1887–1896, doi: 10.1145/2623330.2623377.

[23] X. Lv, “A study on the application of automatic scoring and feedback system in college english writing,” *International Journal of Emerging Technologies in Learning*, vol. 13, no. 3, pp. 188–196, 2018, doi: 10.3991/ijet.v13i03.8386.

[24] X. Zhou *et al.*, “EAST: An efficient and accurate scene text detector,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2642–2651, doi: 10.1109/CVPR.2017.283.

[25] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, “Character region awareness for text detection,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 9357–9366, doi: 10.1109/CVPR.2019.00959.

[26] M. Talib, A. H. Y. Al-Noori, and J. Suad, “YOLOv8-CAB: improved YOLOv8 for real-time object detection,” *Karbala International Journal of Modern Science*, vol. 10, no. 1, 2024, doi: 10.33640/2405-609X.3339.

[27] J. R. Terven and D. M. Cordova-Esparza, “A comprehensive review of YOLO: from YOLOv1 to YOLOv8 and beyond,” *arXiv-Computer Science*, vol. 1, pp. 1–36, 2024.

[28] N. Garg, N. Sharma, G. Jain, V. Jain, and V. Upreti, “Handwriting recognition system using YOLO and CTC,” in *2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech)*, 2023, pp. 496–502, doi: 10.1109/ICACCTech61146.2023.00087.

[29] W. Qu, S. Zhong, Y. Wu, and X. Cao, “Development of a real-time pen-holding gesture recognition system based on improved YOLOv8,” in *2023 International Conference on Image Processing, Computer Vision and Machine Learning, ICICML 2023*, 2023, pp. 1035–1039, doi: 10.1109/ICICML60161.2023.10424854.

[30] J. K. Manipatruni, R. G. Sree, R. Padakanti, S. Naroju, and B. K. Depuru, “Leveraging artificial intelligence for simplified invoice automation: paddle OCR-based text extraction from invoices,” *International Journal of Innovative Science and Research Technology*, vol. 8, no. 9, pp. 1735–1740, 2023.

[31] Y. Du *et al.*, “PP-OCR: a practical ultra lightweight OCR system,” *arXiv-Computer Science*, vol. 1, pp. 1–14, 2020.

[32] Q. Peng and L. Tu, “Paddle-OCR-based real-time online recognition system for steel plate slab spray marking characters,” *Journal of Control, Automation and Electrical Systems*, vol. 35, no. 1, pp. 221–233, 2024, doi: 10.1007/s40313-023-01062-w.

[33] R. Deepa, S. Gayathri, P. Chitra, J. J. Jasmine, R. R. Devi, and A. Thilagavathy, “An enhanced machine learning technique for text detection using Keras sequential model,” in *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1147–1151, doi: 10.1109/ICEARS56392.2023.10085174.

[34] S. Bhadra, K. Aneja, and S. Mandal, “OCR using convolution neural network in Python with Keras and TensorFlow,” *International Journal of Advanced Research in Science, Communication and Technology*, vol. 2, no. 1, pp. 285–292, 2022, doi: 10.48175/IJARSCT-2283.

[35] R. Smith, “An overview of the Tesseract OCR engine,” in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2*, 2007, pp. 629–633, doi: 10.1109/ICDAR.2007.4376991.

[36] S. Badla, “Improving the efficiency of Tesseract OCR engine,” M.Sc. Thesis, Department of Computer Science, San Jose State University, San Jose, United States, 2014, doi: 10.31979/etd.5avd-kf2g.

[37] S. Mubeen, J. Brahmani, D. P. Kalyan, A. Jagirdar, and A. P. Kumar, “Optical character recognition using Tesseract,” *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 11, pp. 672–675, 2022, doi: 10.22214/ijraset.2022.47414.

[38] G. A. Robby, A. Tandra, I. Susanto, J. Harefa, and A. Chowanda, “Implementation of optical character recognition using tesseract with the Javanese script target in Android application,” *Procedia Computer Science*, vol. 157, pp. 499–505, 2019, doi: 10.1016/j.procs.2019.09.006.

[39] J. A. Patel, “Handwritten and printed text recognition using Tesseract-OCR,” *International Journal of Creative Research Thoughts (IJCRT)*, vol. 9, no. 9, pp. 69–77, 2021.

[40] M. Li *et al.*, “TrOCR: transformer-based optical character recognition with pre-trained models,” in *The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)*, 2023, vol. 37, no. 11, pp. 13094–13102, doi: 10.1609/aaai.v37i11.26538.

[41] D. Chang and Y. Li, “DLoRA-TrOCR: mixed text mode optical character recognition based on transformer,” *arXiv-Computer Science*, vol. 1, pp. 1–8, 2024.

[42] A. Q. Jiang *et al.*, “Mixtral of experts,” *arXiv-Computer Science*, vol. 1, pp. 1–13, 2024.

[43] K. Ono and A. Morita, “Evaluating large language models: ChatGPT-4, Mistral 8x7B, and Google Gemini benchmarked against MMLU,” *Techrxiv*, vol. 1, pp. 1–7, 2024, doi: 10.36227/techrxiv.170956672.21573677/v1.

[44] Z. Chen, Y. Deng, Y. Wu, Q. Gu, and Y. Li, "Towards understanding mixture of experts in deep learning," *arXiv-Computer Science*, vol. 1, pp. 1–53, 2022.

BIOGRAPHIES OF AUTHORS

Murugavalli Elangovan received B.E. degree in electronics and communication engineering from Bharathidasan University, India in 1996, M.E. degree in optical communication from Anna University, India in 2005 and Ph.D. degree from Anna University, Chennai, India in 2018. She is now an Assistant Professor in the Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai, India. Her research interest includes wireless communication networks, millimeter wave communication, internet of things and security. She can be contacted at email: murugavalli@tce.edu.

Rajeswari Kaleeswaran received the Ph.D. degree in wireless communication from Anna University, Chennai, in 2015. She has a working experience as a lecturer for 11 years and as an Assistant Professor for 12 years. She is also a senior member of IEEE. She is currently an Associate Professor with the Department of Electronics and Communication Engineering, Thiagarajar College of Engineering, Madurai. Her research interests include 6G wireless communication and biomedical signal processing using artificial intelligence. She can be contacted at email: rajeswari@tce.edu.

Kathirvel Mohankumar is undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. His research interest includes image processing. He can be contacted at email: mkathirvel64420@gmail.com.

Shreya Rangachari is undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. Her research interest includes image processing and RF communication. She can be contacted at email: shreyarangachari@gmail.com.

Ubasini Manivannan is undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. Her research interests include image processing and VLSI. She can be contacted at email: ubasinimanivannan@gmail.com.

Rishapa Pandidurai is undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. Her research interest includes VLSI. She can be contacted at email: pandiraji2011@gmail.com.

Maria Bellarmin Joel undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. His research interest includes image processing. He can be contacted at email: mariabellarminjoel@gmail.com.

Preethi Meenatchi Murugesan undergoing Bachelor of Engineering in Electronics and Communication at Thiagarajar College of Engineering, Madurai, Tamilnadu, India. Her research interest includes VLSI. She can be contacted at email: preethimeenatchi@gmail.com.