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 1, February 2026
Impact Factor: 8.2

An Integrated YOLO-OCR Framework for Academic Document Verification

Ishita Kharde[1], Pranali Raikar[2], Shravani Patil[3], V. V. Sovani[4]
Department of Electronics and Telecommunication[1-4]
Pune Vidyarthi Griha's College of Engineering and Technology Pune, India
22012110@pvgcoet.ac.in , 22012111@pvgcoet.ac.in , 22012079@pvgcoet.ac.in , vvs_entc@pvgcoet.ac.in

Abstract: *This paper proposes an intelligent system for verifying the authenticity of academic documents and detecting tampering using artificial intelligence methods. The YOLOv8n deep learning model is used for detecting key areas on an image of an academic document, based on an image of the document being verified (e.g., name, roll number, marks, percentage, institutional seal).*
The text created by Optical Character Recognition (OCR) from the identified areas is checked against two different ways of determining whether there has been any tampering or forgery of the document using logical validation of the data in question. For example, whether the percentage computed matches the total amount of marks received.
If any inconsistency exists between the extracted data from the document being verified, as well as from the logical validation, the areas of the document will be marked on the webpage, which shows both visually the document and the results of the verification.
A variety of experiments with test cases show that the current system classifies academic documents into three categories (i.e., legitimate, forged and needing to be confirmed) based on the verification results. Thus, the proposed technique has reduced the time and effort required to manually verify a document and can also assist in detecting academic documents that have possibly been tampered with..

Keywords: OCR, YOLOv8n, forgery detection, document verification, machine learning, deep learning

I. INTRODUCTION

The usage of digital documents in educational and administrative spheres is on the rise, along with forgery and tampering cases. Academic certificates, marksheets, and identification documents are generally verified through manual means, which is tedious and inefficient. It is hence important to have an automated system that can effectively verify the documents and confirm their legitimacy. Contemporary advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have helped deploy intelligent systems that can process and recognize documents, and extract significant information from them. Computer vision enables recognition of vital document fields such as names, roll number, marks, percentage, and institution seal, and Optical Character Recognition (OCR) is used to recognize and extract text from images.

The proposed project offers an intelligent verification system that uses the YOLOv8n deep learning model for field detection in the documents and optical character recognition technology for text extraction. The forgery detection module uses a machine learning model, while the verification module relies on database and logical comparisons.

II. LITERATURE REVIEW

Research on document automation verification has been enjoyed as a very strong area of interest for at least the last ten years with many technological advancements including OCR, machine learning and deep learning methods advancing research and development. Large amounts of research has been conducted on the current state-of-the-art of text extraction and text forgery, however there is very little research specifically on either automated document verification or text verification from embedded platforms, and most existing research falls into either the computationally intensive category or requires significant post-processing.

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 1, February 2026 Impact Factor: 8.2

A. OCR Based text extraction methods

Traditional verification methods are totally reliant upon OCR engines such as Tesseract, however according to Smith [1], the ability of an OCR engine to read a scanned image is adversely impacted by a number of variables including the lighting conditions, resolution of the scans and the quality of the scan; thus, structures such as stamps, logos and visible watermarks may not be verified using traditional verification methods as traditional verification methods are text centric in nature and focus only on the extraction of text.

B. Document Authenticity using Machine Learning Strategies

Traditional methods of utilizing machine learning techniques to confirm the legitimacy of a given document or file typically make use of handcrafted extracted features and the results of Optical Character Recognition (OCR) as data sources to establish classification through Support Vector Machines (SVM) or Logistic Regression. Castelblanco et al. [2], for example, have shown through extensive research and experimentation that they were successful in producing exceptional results with regards to the verification of document validity when combining the three phases of the overall verification process into one collective hybrid for applying classification between documents with regard to their verification status. Other hybridization techniques were utilized for combining RGB-DCT features [3] which improved noise tolerance. Although these types of machine learning verification solutions require significant computational power and greater processing capabilities than are currently available, they generally are incapable of adapting to new formats of documents that typically arise because they rely solely on handcrafted features for extraction. Most importantly, because of the extremely complex nature of these types of solution methods using machine learning, they will not be able to operate in real-time on raspberry pi-based embedded systems as yet.

C. Detection of Tampering and Forgery Based on Deep Learning

In recent years, deep learning has become increasingly popular, with most attention focused on the use of CNN-based models used to identify forgeries in both digital and scanned copies. ELA (Error Level Analysis) based approaches using CNN-based models have shown high accuracy [4] for detecting inconsistencies at the pixel level. However, these approaches have primarily been developed for scanned images instead of camera captured ones thus negatively impacting their generalization performance in the context of embedded vision applications. Further, the implementation of GPUs in such networks substantially reduces processing performance when applied to resource-constrained embedded vision systems.

D. Document Field Detection by YOLO

Real-time object detection models based on the YOLO (You Only Look Once) algorithm are receiving a lot of attention because they enable fast processing speeds and are efficient. Several recent studies have reported that the YOLO family of algorithms can detect fields on Identification cards, Driver's licenses, and Mark sheets with high accuracy and very little delay [5]. All other studies using the YOLO algorithm have focused primarily on Localisation of fields, e.g. identifying the fields of Name, Photograph and Signature; most have not incorporated any extraction/recognition using optical character recognition (OCR), and no studies have investigated whether or not the document has been tampered with. No study has yet provided an integrated pipeline for the verification of embedded information using YOLO, OCR by field and Image Similarity evaluation.

E. Identified Research Gaps

Although many advancements have been made in OCR and other related technologies (e.g., ML, Deep Learning, and YOLO) for analyzing documents, no systems currently exist that integrate a complete and effective verification process. For example, OCR-based approaches primarily extract text from images but are susceptible to variations in image quality and lighting conditions [1], and traditional ML methods use pre-defined features that are limited when attempting to analyze different types of documents [2],[3]. Deep Learning-based forgery detection methods often require a significant amount of computational resources and cannot be implemented in an on-demand fashion [4]. YOLO-based methods perform very quickly and accurately detect fields in documents, however; most previous works

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 1, February 2026
Impact Factor: 8.2

have not developed an automated system that brings together field detection, OCR extraction of the fields, forgery detection, and validation of tampered fields through a single web-based interface [5]. Thus, it is important to build an integrated document verification system that can provide a unified process of detecting document fields, electronically extracting document fields, detecting any attempts at forgery, and validating any attempts at tampering, through a single web-interface that operates in real-time.

III. METHODOLOGY

Proposed system for Smart Document Verification: It can be noted that the proposed system for smart document verification follows a systematic approach based on deep learning, OCR, machine learning comparison, and verification logic for identifying forged and tampered documents.

A. Dataset Description

This study's dataset is made up of academic certificate images obtained from the Kaggle website. The "Certificate Tampering Detection Dataset" contains original and modified document designs. These images were utilized during the training and assessment phase for the full Document Analysis Pipeline.

Appropriate document fields such as Name, Roll Number, Marks, Percentage, and Institutional Seals were annotated with bounding boxes in order to train the YOLOv8n model for identifying significant areas on Academic Documents with object detection training. The annotated datasets were divided into training and validation datasets so that model evaluations could be made accurately.

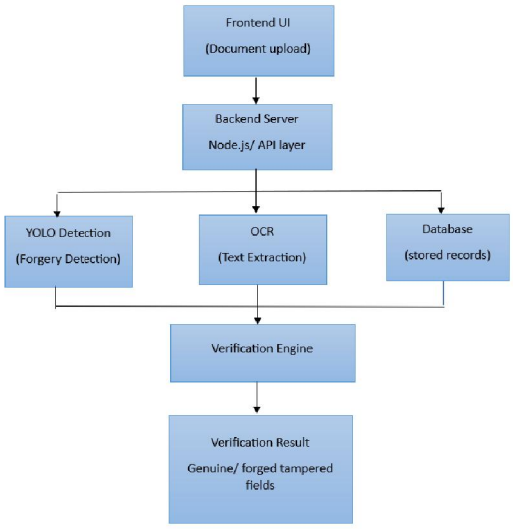

Figure 1. Overall System Architecture of the Proposed Smart Document Verifier

B. Document Upload and Input

The first part of the system is a web interface through which an image of a document is uploaded for verification. Once the image is uploaded, it is sent to the backend server using an API request. The backend server facilitates communication among the various components of the system, such as field detection, OCR, forgery detection, and database verification. The original document and the test document are provided as input to the system for comparison.

C. Image Preprocessing

The pre-processing of the uploaded image is conducted prior to any operation being performed in the document analysis section. The pre-processing operation helps to enhance the quality of the uploaded image and further improves the detection accuracy of the OCR text in the image. The pre-processing operations include changing the colour mode

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 1, February 2026 Impact Factor: 8.2

to grayscale and applying thresholding and filtering to reduce unwanted noises in the image. The image may also be resized to match specific dimensions.

D. Detection of document fields using YOLOv8n

The preprocessed document image is then fed into the YOLOv8n deep learning model, which is trained on academic documents. It detects and localizes some important ROIs in the document: Name, Roll Number, Marks, Percentage, and Institutional Seal. Using the bounding box coordinates generated by the model, these detected field images are cropped from the document image. These cropped field images are further sent for text extraction with the OCR module.

E. OCR-based text extraction

The cropped document fields are passed through Tesseract OCR in order to extract textual information for each region. Unnecessary spaces and other formatting inconsistencies are cleaned and normalized in the extracted text. Regular expression-based filtering is used to filter out field values, such as numeric marks and percentage values or textual names. Field-value pairs are then organized in structured data format for further verification.

F. Forgery Detection and Similarity Analysis

In determining whether a document has been forged, detection is achieved by comparing field information obtained from both original and test documents. Text similarity is determined by performing string matching operations on field values to assess whether modifications have been done. In addition, structural similarity will also be analyzed using the Structural Similarity Index (SSIM) algorithm [6] in determining visual differences between original and test documents. Estimation is done using a weighted approach in obtaining similarity by combining field similarity values and image similarity values.

G. Verification Engine and Result Generation

The verification engine is the heart of the system, as it is where the decisions are made. It considers all the processes of OCR, comparisons, image comparison, databases, logic, etc., while authenticating the document. If a mismatch is found, the system highlights the particular fields, which are suspected of being altered. The output of the verification process is conveyed to the web interface, which presents it to the user. The output shows whether the document is genuine or fake.

IV. RESULTS AND DISCUSSION

The proposed document verification system was assessed through three separate input scenarios. Field detection using the YOLOv8n model for uploaded document processing, using OCR for text extraction and using logic to check extracted values for further verification. The verification results were made accessible via a web-based interface, including detection visualization.

Test Case Analysis

A. Genuine Document

The system was tested with an original academic credential uploaded via the web interface and passed all fields necessary to identify a document: Name, Roll Number, Marks, Percentage, and College Seal identified using the YOLOv8n Model. The OCR module was able to extract text from all identified areas of the document; all validation checks (Logical Validation) of the extracted values indicate that value was in the proper format and within the correct range. Due to the verification process, the system classified the document as a GENUINE document demonstrating that the proposed system can verify an education document to be genuine.

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 1, February 2026
Impact Factor: 8.2

Table I: Test Case 1 Result

Parameter	Observation
Input Document	Original Certificate
Field Detection	Successful
OCR Extraction	Successful
Logical Validation	Passed
Verification Result	Authentic

B. Document Forgery

In this example, a fabricated academic certificate was uploaded with altered information on the certificate, which was detected by the YOLOv8n model within the document fields, and the OCR completed character recognition but displayed an inconsistency by the extracted value with expected values of the document for logical validation. The detection visualization highlighted the detected areas and classified them as FORGED (SUSPICIOUS). The results confirm the system can identify an altered image.

Table II: Test Case 2 Result

Parameter	Observation
Input Document	Forged Certificate
Field Detection	Successful
OCR Extraction	Successful
Logical Validation	Failed
Verification Result	Forged

C. Invalid Document

This test case is an example of a non-academic, unsupported document being uploaded into the system. The document was determined to be invalid by the system, because there were no document fields that could be detected by the system, and therefore no usable OCR extraction could occur, which resulted in verification conditions being unable to be met; therefore the document is invalid. The purpose of this test is to show that the system properly handles incorrect input data and prevents the false authentication of a user.

Table III: Test Case 3 Result

Parameter	Observation
Input Document	Invalid Document
Field Detection	Failed
OCR Extraction	Failed
Logical Validation	Not Performed
Verification Result	Invalid

Overall Result

The proposed document verification system's results show the ability to analyze academic documents among the various ways to verify those academic documents. Throughout the three test cases, the proposed system was able to detect various fields of an academic document via the YOLOv8n model, extract textual data from the document via OCR, and logically validate the extracted data.

The proposed system was able to accurately classify a genuine document as genuine, identify a fraudulently manipulated document as a forgery, and verify that an unsupported document was not valid.

The results of these experiments suggest that the proposed system can accurately and reliably differentiate between legitimate and illegitimate academic documents, and can also handle erroneous or incomplete input data. By combining object detection, text extraction, and validation logic, the proposed system has developed the capability to perform automated and accurate document verification with high levels of efficiency. Each of the proposed system's operational

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 1, February 2026 Impact Factor: 8.2

capabilities were demonstrated to be consistent across each of the three academic document tests, confirming its suitability for use as a real-time document verification application.

V. FUTURE SCOPE

The proposed system for smart document verification has the potential to be expanded to accommodate the verification of other types of documents apart from academic marksheets and certificates. Even though academic certificates and documents can be checked through online platforms such as DigiLocker, other documents such as bank receipts, transaction slips, invoices, etc., do not yet have a verification technology for these types of documents. The technology for document verification may be enhanced to check for forgery and tampering in bank receipts by using OCR to read the fields such as transaction ID, amount, date, etc.

Also, future enhancements may include integrating sophisticated deep learning-based forgery detection models, increasing the training dataset to cater for various document types, and enhancing OCR accuracy for poor image quality documents. The system may be implemented as either a cloud-based service or as a mobile application, allowing real-time document verification across various domains and sectors, including banking, administration, and e-commerce.

More security features, such as QR codes, digital signatures, and using blockchain technology to validate documents, can be added to increase the effectiveness and scalability of the system.

VI. CONCLUSION

This project created a document verification system that uses artificial intelligence to identify forgery or tampering in academic documents. This verification system uses an object detection technique (YOLOv8n), Optical Character Recognition (OCR) and logical assessment methods to fully automate academic document verification. By performing an analysis of important fields within each document and extracting text from each of those fields, this system assesses overall document consistency and helps to find documents that have undergone suspicious modification.

The results from the experiments validate that this verification system effectively classifies each academic document as genuine, forged and invalid when presented with various types of test documents. The use of detection visualization & web interface enhances the objectivity and user-friendliness of this verification system. It reduces the workload for manual verification, improves the turnaround time for document verification, and reduces the likelihood of unintentionally accepting fraudulent academic documents.

The results of the project indicate that using a combination of deep learning methodology for detecting document fields along with OCR and logical validation will provide a viable, effective solution for automating authentication of academic documentation.

REFERENCES

[1] R. Smith, “An overview of the Tesseract OCR engine,” in Proc. Int. Conf. Document Analysis and Recognition (ICDAR), 2007, pp. 629–633.

[2] Á. Castelblanco, A. Daza, and A. Pérez, “Document authenticity verification using feature fusion and machine learning,” IEEE Access, vol. 7, pp. 81296–81305, 2019.

[3] H. S. Bhateja, A. Kumar, and A. Jain, “Forgery detection in document images using RGB-DCT feature fusion,” in Proc. Int. Conf. Signal Processing and Communication (ICSC), 2018, pp. 192–197.

[4] B. Bayar and M. C. Stamm, “A deep learning approach to universal image manipulation detection using a new convolutional layer,” in Proc. ACM Workshop on Information Hiding and Multimedia Security, 2016, pp. 5–10.

[5] X. Wang, Y. Zhang, and J. Liu, “Real-time document field detection using YOLO for identity verification,” in Proc. IEEE Int. Conf. Image Processing (ICIP), 2021, pp. 1844–1848.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[7] D. Karatzas et al., “ICDAR 2015 competition on robust reading,” *Proc. Int. Conf. Document Analysis and Recognition (ICDAR)*, 2015, pp. 1156–1160.

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 1, February 2026
Impact Factor: 8.2

[8] S. Long, X. He, and C. Yao, "Scene text detection and recognition: The deep learning era," *Int. J. Computer Vision*, vol. 129, pp. 161–184, 2021.
[9] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.
[10] S. Eskenazi, P. Gomez, and J. V. Gomez, "Real-time embedded vision systems: A survey," *IEEE Access*, vol. 8, pp. 140216–140236, 2020.
[11] Raspberry Pi Foundation, "Raspberry Pi 4 Model B Datasheet," 2019.
[12] N. Wang, Y. Zhang, and X. Wang, "Document image tampering detection based on deep learning," *Multimedia Tools and Applications*, vol. 79, pp. 13147–13163, 2020.
[13] M. Li, Y. Xu, and Z. Chen, "Lightweight deep learning models for embedded vision systems," *IEEE Embedded Systems Letters*, vol. 14, no. 2, pp. 85–88, 2022.