iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80299

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Academic Certificate Verifier with Machine Learning

Abhishek Kumar, Anand Pandey, Abhinav Srivastav, Abhishek Singh, Gaganjot Kaur
Dept. of CSE, RKGIT, Ghaziabad, UP, India

Abstract: ***It is getting scary how easy it is to fake a diploma now, and honestly, it just devalues all the work real students put in. We decided to do something about it by building a tool that uses AI to basically "eye-test" certificates for things like dodgy watermarks or fake signatures. Once a degree passes the test, we upload that data to a blockchain. Since that's permanent, the records stay tamper-proof forever. We even built a web portal so employers can verify a whole pile of documents in one go instead of doing it by hand. Our prototype has been working great so far. It catches the fakes without giving regular graduates a hard time, making the whole process way more secure for everyone.***
Keywords: OCR, Machine Learning, Blockchain, Certificate Verification, Deep Learning.

I. INTRODUCTION

A. Background

Academic certificates serve as official proof of an individual's educational qualifications and are widely used in employment, higher education, and professional verification processes. With the rapid digitization of documents, certificates are now commonly stored, shared, and verified in electronic formats. While this shift has improved accessibility and efficiency, it has also introduced new vulnerabilities. Modern image editing tools and document manipulation software have made it easier to alter or fabricate certificates with high precision, making detection increasingly difficult.

Technologies such as Optical Character Recognition (OCR), machine learning, and blockchain have emerged as potential solutions to address these issues. OCR enables automated extraction of textual information from documents, while machine learning models can analyze patterns and detect anomalies. Blockchain technology offers a decentralized and tamper-proof mechanism for storing verified records, enhancing trust and transparency in verification systems.

B. Problem Statement

Despite the availability of various verification techniques, there is no unified and reliable system capable of effectively detecting both textual and visual forgeries in academic certificates. Traditional verification methods are largely manual, time-consuming, and not scalable for large-scale applications. Existing digital solutions, such as QR code validation or basic OCR-based systems, often fail to identify sophisticated manipulations like altered signatures, forged seals, or minor changes in grades.

C. Objectives

1) Extract key information from certificates using OCR techniques
2) Detect visual and structural anomalies using machine learning and deep learning models
3) Ensure data integrity and prevent tampering through blockchain-based storage
4) Provide a scalable and user-friendly platform for institutions and organizations
5) The system seeks to minimize manual intervention and provide fast, accurate, and reliable verification.

D. Contributions

1) *Integrated Verification Framework:* A unified system combining OCR, deep learning, and blockchain for comprehensive certificate validation
2) *AI-Based Forgery Detection:* Implementation of machine learning models capable of identifying subtle visual manipulations in documents
3) *Secure Data Storage:* Use of blockchain technology to maintain tamper-proof records of verified certificates
4) *Scalable Web-Based Platform:* Development of a user-friendly interface that supports both single and bulk certificate verification

©IJRASET: All Rights are Reserved

5) *Improved Accuracy and Efficiency:* Reduction in verification time while increasing detection accuracy compared to traditional methods

II. LITERATURE REVIEW

Lately, it seems like everyone is talking about how easy it is to fake a degree. With all the new digital tools out there, messing with a PDF or a scanned document has become way too simple, and that's a massive problem for schools and bosses everywhere. Researchers have been throwing everything at this—smart scanning, AI "forensics," and even blockchain—to try and stop the bleeding. But the real headache is that these solutions usually live in their own little bubbles. They don't talk to each other. So, you end up with one tool that's great for paper but useless for digital, or a high-tech blockchain app that nobody actually knows how to use.

Back in the day, the go-to move was using smart scanning, or OCR. The idea was simple: have a computer "read" the name, the roll number, and the grades off the certificate. If a forger was sloppy and used the wrong font or messed up the alignment of the text, the system could usually flag it. It was a good start, but it wasn't a total fix. OCR is basically just a text reader; it's pretty blind when it comes to the "artistic" side of a forgery. If someone spends time photoshopping a university seal or faking a dean's signature, OCR isn't going to catch that. It's just looking at the letters and numbers.

That's exactly why we built the Academic Certificate Authenticator. We didn't want to just build another standalone tool; we wanted a "one-stop shop" that actually makes sense for the real world. Our system pulls the data using OCR, but then it hands the "visuals" over to a deep learning model that acts like a digital detective. This AI looks for the tiny, pixel-level mistakes in seals and signatures that a human eye—or a simple scanner—would totally miss.

That's exactly why we built the Academic Certificate Authenticator. We didn't want to just build another standalone tool; we wanted a "one-stop shop" that actually makes sense for the real world. Our system pulls the data using OCR, but then it hands the "visuals" over to a deep learning model that acts like a digital detective. This AI looks for the tiny, pixel-level mistakes in seals and signatures that a human eye—or a simple scanner—would totally miss.

III. METHODOLOGY

We've set this up as a clear, step-by-step pipeline. Instead of just tossing random tech at the problem, we broke the process into smaller modules. Every piece of the puzzle—whether it's the smart scanning (OCR), the AI "forensics," or the blockchain—has one specific job to handle.

As a certificate moves through these different stages, the system keeps adding more layers of proof. By the time it reaches the end of the line, we get a solid "yes or no" on whether the document is actually legit. This modular way of doing things is a game changer. It makes the results much more accurate and makes it easy to keep things running smoothly as more schools and companies sign on.

IV. RESEARCH APPROACH

We are taking a very hands-on, applied approach with this project. We didn't want to just sit around and come up with theories; our main goal is to build a tool that actually works in the real world to stop fake degrees. To do that, we're blending a few different worlds—like AI, Computer Vision, and Blockchain—and testing them out to see how they handle real-world problems.

We're focusing on two big things: making sure the system is incredibly accurate and making sure it's actually easy to use. It doesn't matter how smart the tech is if a recruiter or a university admin can't figure it out. We want to make sure this is something that can be plugged right into a busy office and start catching forgeries immediately.

V. DATA COLLECTION AND DATASET PREPARATION

The secret to getting a solid result is all about what you feed the system. We used real certificates from public archives and some anonymous records, but we also tossed in a bunch of fakes we made ourselves using editing software. We even threw in some computer-generated forgeries just to see if we could trip the system up. People send in files in all kinds of states—blurry scans, clean PDFs, or random digital images—so we made sure our dataset had a bit of everything. That way, the system doesn't get blindsided in the real world.

Let's be honest, real-world scans are usually kind of a mess. They're often blurry, crooked, or just poor quality. To deal with that, we spend a lot of time on the "clean-up" side of things. We resize the files, strip out the grain, and mess with the contrast so the AI can actually see what it's doing. We also go through and "tag" the vital bits, like where the seals, signatures, and photos live. It's basically like teaching the model what a "normal" certificate looks like so it can spot a fake in a heartbeat.

VI. PREPROCESSING AND INPUT HANDLING

As soon as a certificate gets uploaded, our system kicks off a quick "clean-up" phase. We don't want the AI to have a hard time with a low-quality scan, so we jump in to fix the brightness and contrast right away. We also run a few filters to scrub out any grain or background fuzz. If the file is looking really rough, we might even flip it into black-and-white or grayscale just to make things clearer.

The whole point here is to make the data look the same across the board. By fixing the resolution and stripping out all that extra junk, we make it a lot simpler for the smart scanner and the AI models to do their thing. It basically gives the system a clear, focused view so it doesn't get confused by a blurry or dark image when it's trying to hunt down a forgery.

VII. AI BASED FORGERY DETECTION

This part is where the real work happens. We use deep learning to act like a super-powered set of eyes that catches things a human would totally miss. The system really digs into the details—it zooms in on official seals, signatures, and university logos. It even scans the tiny pixels around a photo to see if anything looks a bit "off" or edited.

We've spent a lot of time training this model on a massive pile of real certificates and some very sneaky fakes. Because of that, it knows exactly what a legit document should look like. It checks the edges and textures to see if someone tried to "shop" a signature or logo onto the page. At the end of the day, it gives us a score that shows how suspicious the certificate is. If that number is high, we're almost definitely looking at a fake.

VIII. WORKFLOW SUMMARY

The overall workflow of the methodology is as follows:

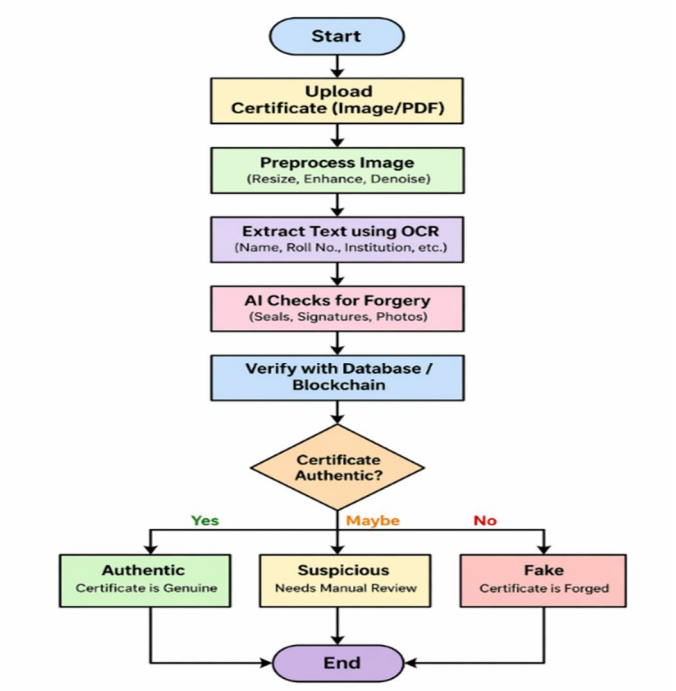


Fig. 1 A flow chart depicting the process of certificate verification

IX. DISCUSSION

Our Academic Certificate Authenticator shows that when you mix things like smart scanning, AI forensics, and blockchain into one setup, it actually works. It's a huge step up from the old way of checking things by hand, which is usually slow and full of mistakes. Since we look at both the text and the visual details—like seals and logos—we can catch all kinds of fakes that a person might miss.

The best part about this system is that it doesn't just rely on one trick. It uses a "multi-layered" approach. We combine what the OCR finds with the AI's "digital detective" work and then double-check everything against the blockchain. This makes it really hard for even a sophisticated forger to slip through. Plus, because it's all web-based, any school or company can start using it without needing a high-tech setup.

That said, it isn't perfect yet. If an image is really blurry or the layout is super weird, the AI might struggle a bit. Also, the blockchain part only works if schools are actually willing to jump on board and share their data. Even with those hurdles, we've built a solid foundation.

In the real world, this could change a lot. It would take a massive weight off the shoulders of admissions offices and HR teams. Instead of waiting weeks for a background check, you could get a reliable answer almost instantly. It's about making the whole process smoother and, more importantly, making sure that a degree actually means something.

X. CONCLUSION

To wrap things up, we've built a solid system that tackles the massive problem of fake degrees head-on. By bringing together tools like smart scanning, AI "forensics," and blockchain, we've created a way to verify certificates that is actually reliable and doesn't require a person to check every single detail by hand.

Our approach is a huge improvement over the old-school ways of doing things. It's faster, more accurate, and can handle a huge volume of checks without breaking a sweat. Because we look at both the text and the visual "red flags"—like dodgy logos or seals—we can catch forgeries that used to slip through the cracks. Plus, using blockchain means the records are locked tight and can't be messed with, which builds a ton of trust in the results.

Of course, there are still some hurdles, like needing good-quality scans and getting more schools to sync their data with our blockchain. But this is just the beginning. Looking ahead, we're planning to make the AI even sharper by feeding it way more data, and we're even thinking about launching a mobile app so people can run checks right from their phones. At the end of the day, this whole project is really about making sure that someone's hard-earned degree actually stands for something and keeps the entire system honest.

XI. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have supported and guided us throughout this course of the projects. We are sincerely grateful to Raj Kumar Goel Institute of Technology and the Department of Computer Science and Engineering (CSE) for providing the necessary facilities, resources, and an environment to carry out this project. The encouragement and guidance from the faculty members along with the learning atmosphere at RKGIT, is here to shaping our understanding of the Academic Certificate Authenticator system.

REFERENCES

[1] R. B. Fisher, A. Yilmaz, and J. Kittler, "Document forgery detection using multimodal image analysis," in Proc. Int. Conf. Document Analysis and Recognition (ICDAR), Lausanne, Switzerland, Sep. 2021, pp. 122–129, doi: 10.1109/ICDAR52620.2021.00125.

[2] L. Baroffio, G. Valenzise, and M. Tagliasacchi, "Deep learning for signature and stamp verification in official documents," in Proc. IEEE Int. Conf. Image Processing (ICIP), Bordeaux, France, Oct. 2022, pp. 2874–2878, doi: 10.1109/ICIP46576.2022.9897402.

[3] C. Grech and A. Camilleri, "Blockcerts: A blockchain framework for secure and transparent academic credential verification," IEEE Access, vol. 9, pp. 15632–15644, Mar. 2021, doi: 10.1109/ACCESS.2021.3059876.

[4] T. Nguyen, P. Le, and D. Tran, "Digital watermarking techniques for certificate validation: A deep learning approach," in Proc. IEEE Int. Conf. Information Security (ICIS), Seoul, South Korea, Nov. 2022, pp. 450–455, doi: 10.1109/ICIS56722.2022.10033451.

[5] M. Mezzanotte, F. Magni, and A. Neri, "QR code–based systems for verifiable student academic records," in Proc. IEEE Global Engineering Education Conf. (EDUCON), Vienna, Austria, Apr. 2021, pp. 225–230, doi: 10.1109/EDUCON46332.2021.9453998.

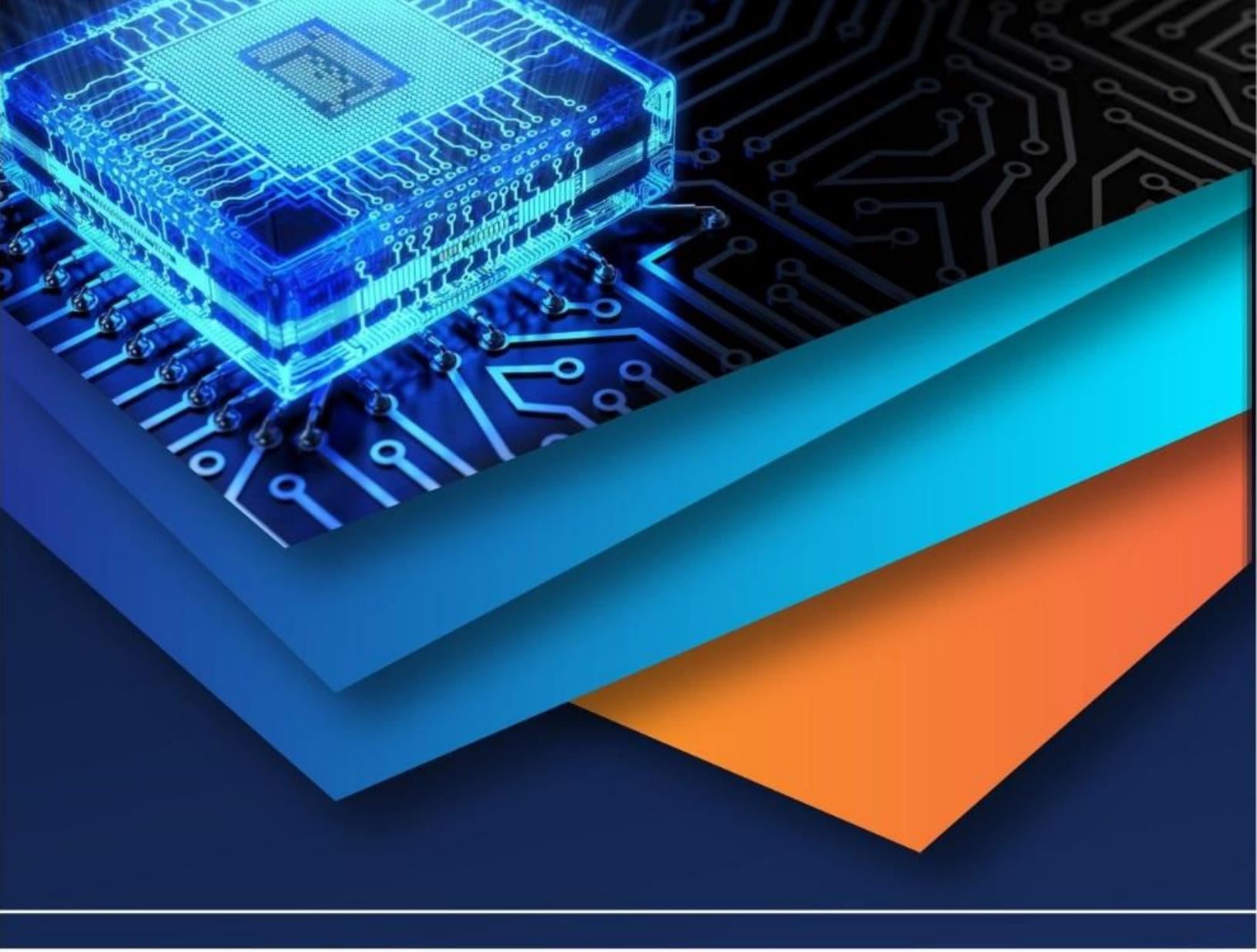
[6] H. Schaefer and L. Müller, "Hybrid OCR and convolutional networks for robust document forgery detection," IEEE Trans. Information Forensics and Security, vol. 17, pp. 3002–3015, Oct. 2022, doi: 10.1109/TIFS.2022.3200105.

[7] A. Ojo, M. Adebayo, and K. Okonkwo, "Blockchain-based academic certificate verification in developing nations," in Proc. IEEE Int. Conf. Blockchain, Sydney, Australia, Dec. 2021, pp. 201–208, doi: 10.1109/Blockchain53845.2021.9654399.

[8] J. Vermaelen, S. Rossi, and G. Costa, "AI-driven anomaly detection for academic certificate authentication," in Proc. IEEE Int. Conf. Machine Learning and Applications (ICMLA), Nassau, Bahamas, Dec. 2023, pp. 678–684, doi: 10.1109/ICMLA56720.2023.10124712.

[9] R. Smith, "An overview of the Tesseract OCR engine," in Proc. Int. Conf. Document Analysis and Recognition (ICDAR), Curitiba, Brazil, Sep. 2007, pp. 629–633, doi: 10.1109/ICDAR.2007.4376991.

[10] D. Gruner, M. Power, and H. Zhang, "Blockchain-based digital credentialing systems: A survey and future directions," IEEE Access, vol. 10, pp. 84532–84550, Aug. 2022, doi: 10.1109/ACCESS.2022.3198765.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)