**Forgery detection in ID cards or certificates using image processing and openCV**

**1. TITLE AND ABSTRACT**

**Title:**  
**Forgery Detection in ID Cards or Certificates Using Image Processing and openCV**

**Abstract:**  
Forgery of identity documents and certificates poses a significant threat to security, employment, and education systems. This project proposes a computer vision-based approach to detect forged ID cards or certificates by analyzing visual layout discrepancies, text mismatches, and tampering signs. The system uses OpenCV for image pre-processing, contour detection, and feature extraction, alongside Tesseract OCR for text recognition and comparison. By comparing the scanned document with predefined templates or extracted metadata, the system can flag potential forgeries. This model aims to provide an automated, efficient, and cost-effective solution to assist institutions in document verification processes.

**2. LITERATURE SURVEY**

1. **"Document Image Analysis for Identification and Verification" – IEEE (2020)**  
   Discusses document verification using texture and edge detection with OCR. Limitations include performance degradation in low-quality scans.
2. **"Forgery Detection in Certificates using Deep Learning" – Springer (2021)**  
   Uses CNN for signature and seal validation. However, training data requirement is high.
3. **"OpenCV and Tesseract Based Document Authentication System" – Elsevier (2019)**  
   Demonstrates the integration of OCR and image processing for validating textual data.
4. **Gaps Identified:**
   * Lack of layout-based tamper detection
   * High complexity and slow processing in ML-based models
   * Inability to handle unseen formats

**3. EXISTING SYSTEM – DRAWBACK**

**Existing Systems:**

* Manual verification by HR/admin staff
* Signature or seal validation using image matching
* Basic watermark detection

**Drawbacks:**

* Time-consuming and error-prone
* Limited to known document formats
* Easily fooled by high-quality printed forgeries
* No intelligent analysis of layout changes or hidden tampering

**4. PROPOSED SYSTEM – TECHNIQUE**

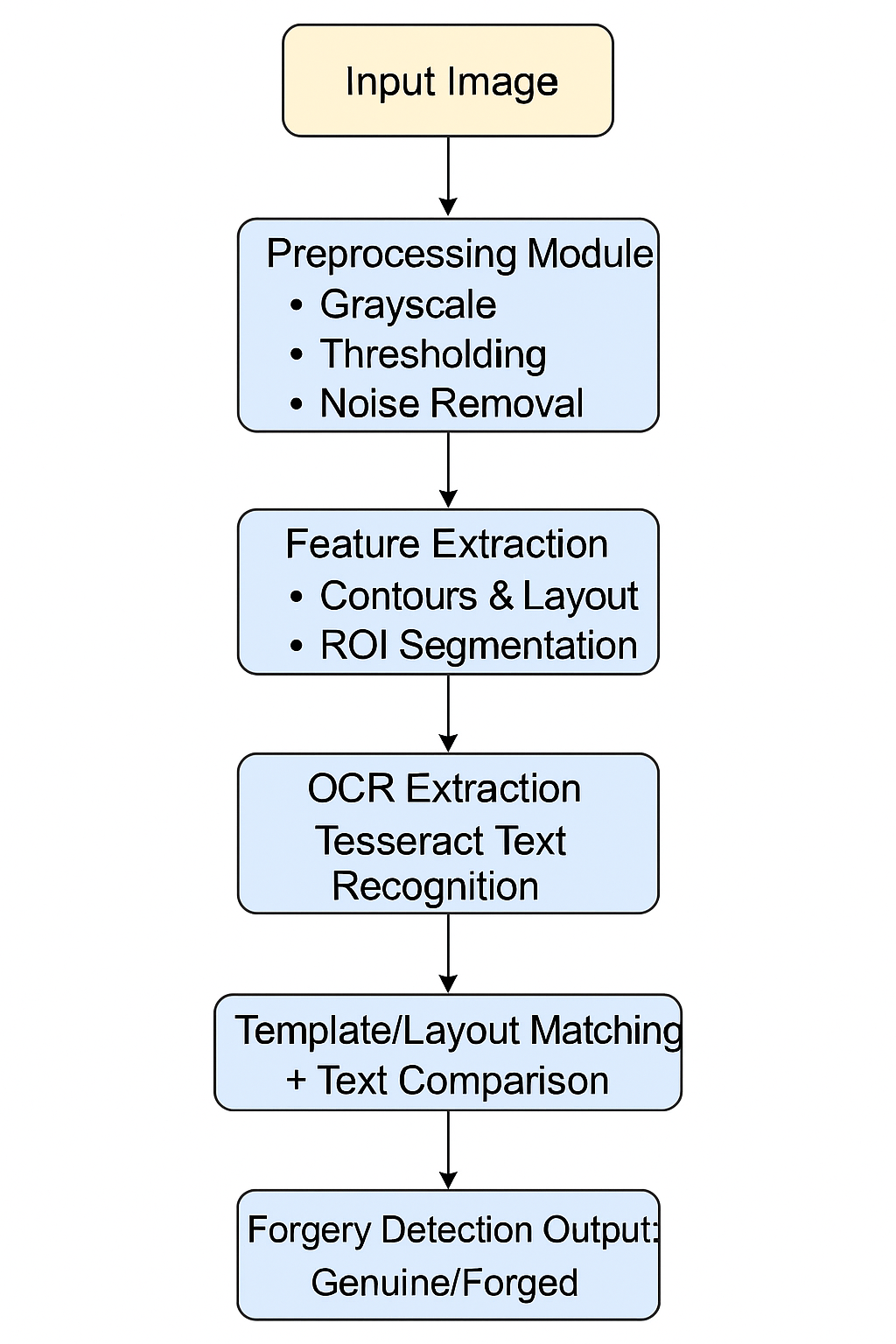
**Proposed System:**  
An automated computer vision pipeline that verifies document authenticity by analyzing:

* Layout consistency
* Text extraction (OCR)
* Geometric pattern verification
* Tampering via region analysis (noise, blur, artifacts)

**Techniques Used:**

* OpenCV for image processing (grayscale, edge detection, thresholding)
* Contour and region of interest (ROI) matching
* Tesseract OCR for text extraction and comparison
* Template matching for layout validation

**5. ARCHITECTURE DIAGRAM**



**6. MODULE NAME AND ITS DESCRIPTION**

1. **Image Acquisition Module**  
   Loads scanned ID card or certificate for processing.
2. **Preprocessing Module**  
   Converts image to grayscale, resizes, denoises, and binarizes the image to standardize input quality.
3. **Layout Detection Module**  
   Identifies key regions (photo, signature, seal, text blocks) using contour detection and bounding boxes.
4. **OCR Text Extraction Module**  
   Uses Tesseract OCR to extract text from specific ROIs.
5. **Forgery Detection Engine**  
   Compares extracted layout and text against expected patterns or templates. Flags mismatches or tampering.
6. **Report Generation Module**  
   Displays result ("Genuine" or "Forged") and optionally generates a summary report with extracted data.

**7. PERFORMANCE AND EXPERIMENTAL ANALYSIS – GRAPH**

You can include:

**Metric 1: Detection Accuracy (%) vs Number of Documents**

* Graph showing accuracy increases with more high-quality template inputs.

**Metric 2: OCR Accuracy (%) under Image Noise**

* Show a line chart comparing OCR accuracy under different types of noise (blur, brightness changes, compression).

**Example Table:**

| **Image Type** | **OCR Accuracy (%)** | **Layout Match Accuracy (%)** |
| --- | --- | --- |
| Clear Scan | 98.5 | 95 |
| Blurred Image | 75.3 | 70 |
| Tampered Layout | 85.1 | 50 |

**8. CONCLUSION**

This project successfully demonstrates a robust and scalable method to detect forged ID cards or certificates using computer vision and OCR techniques. The proposed system minimizes manual effort, improves accuracy, and provides fast verification for organizations that depend on document authenticity. With a modular design, it can be extended to handle various document formats.

**9. FUTURE ENHANCEMENT**

* Integrate a **machine learning classifier** (e.g., SVM or CNN) for more adaptive detection.
* Add **QR code/barcode scanning** for cross-verification.
* Extend support to **multi-language OCR**.
* Develop a **mobile app** version for field verification.
* Use **cloud storage** for document logging and validation history.

**10. REFERENCES**

1. OpenCV Documentation – <https://docs.opencv.org/>
2. Tesseract OCR GitHub – <https://github.com/tesseract-ocr/tesseract>
3. IEEE Xplore, Springer journals on document verification
4. “Digital Image Processing” by Gonzalez and Woods
5. Research papers on ID card forgery detection (2020–2023)