Assignment No 5

Name : Samiksha More

PRN : 22420127

Roll : 381070

Problem Statement: Text Identification using OpenCV, Tesseract (OCR), and Deep Neural Networks

Objective

The objective is to develop a text identification system that can extract and recognize text from images or documents using OpenCV for preprocessing, Tesseract OCR for text extraction, and deep neural networks (DNNs) to improve recognition accuracy. The goal is to automate reading from scanned documents, license plates, street signs, or natural scene images.

Theory

Optical Character Recognition (OCR) is the process of detecting and converting text from images into machine-readable form. Traditional OCR systems often fail with noisy, distorted, or complex backgrounds.

* OpenCV: Provides tools for image preprocessing such as grayscale conversion, thresholding, edge detection, and contour detection to isolate text regions.
* Tesseract OCR: An open-source OCR engine that recognizes text characters in images and converts them to digital text.
* Deep Neural Networks (DNNs): Enhance OCR by providing robust text detection, segmentation, and recognition under challenging conditions (e.g., varying fonts, orientations, lighting).

Methodology

1. Image Acquisition: Capture or scan an image containing text.
2. Preprocessing with OpenCV:
   * Convert image to grayscale.
   * Apply noise reduction and thresholding.
   * Detect contours or bounding boxes to locate text regions.
3. Text Detection (Deep Neural Network):
   * Use deep learning-based text detectors (e.g., EAST Text Detector, CTPN) to locate text regions accurately.
   * Extract regions of interest (ROIs).
4. Text Recognition (Tesseract OCR):
   * Apply OCR to the extracted text regions.
   * Convert recognized text into machine-readable format.
5. Post-processing:
   * Correct spelling errors.
   * Format text according to application requirements.

Advantages

* Automates extraction of text from documents, signs, or images.
* Works with various image formats and fonts.
* Deep learning enhances accuracy in noisy or distorted conditions.
* Open-source tools (OpenCV, Tesseract) make implementation cost-effective.

Limitations

* Accuracy depends on image quality and preprocessing.
* Sensitive to variations in lighting, blur, and orientation.
* Requires significant computational power for DNN-based detection.
* Struggles with cursive or handwritten text compared to printed fonts.

Applications

* Document Digitization: Converting printed books, invoices, or forms into digital text.
* License Plate Recognition: Identifying vehicle registration numbers.
* Assistive Technology: Helping visually impaired people read printed text.
* Translation Systems: Detecting and translating street signs in real time.
* Data Entry Automation: Extracting structured information from forms.

Working / Algorithm

1. Input image is captured and preprocessed with OpenCV (grayscale, thresholding, noise removal).
2. Deep neural network text detector locates potential text regions.
3. Detected regions are cropped and passed to Tesseract OCR.
4. OCR converts image text to digital characters.
5. Post-processing refines results (error correction, formatting).
6. Final machine-readable text is produced for storage or further analysis.

Conclusion

The combination of OpenCV, Tesseract OCR, and Deep Neural Networks provides a robust solution for text identification. OpenCV handles preprocessing, DNNs enhance detection accuracy, and Tesseract efficiently performs text recognition. This hybrid approach makes OCR systems more reliable in real-world scenarios such as document scanning, license plate detection, and assistive technologies.