Name: Yash Gaikwad
Roll No: 381071
PRN: 22420141

Assignment 5

Problem Statement: Text identification using OpenCV, Tesseract (OCR) and deep neural network.

Introduction:

Text identification from images is an important task in computer vision and natural language processing, with applications in document digitization, license plate recognition, automated form reading, and assistive technologies. The process typically involves two major stages: **detecting text regions** in an image and **recognizing the text content**.
In this assignment, we use **OpenCV** for image preprocessing and text region detection, **Tesseract OCR** for text extraction, and a **Deep Neural Network (DNN)** for enhancing recognition accuracy and handling challenging cases such as distorted, noisy, or complex background images.

Objective:

The objectives of this assignment are:

- To understand how computer vision and OCR techniques can be combined for text identification.
- To use **OpenCV** for preprocessing and detecting text-containing regions.
- To integrate **Tesseract OCR** for character-level text recognition.
- To build and train a **Deep Neural Network** to improve recognition performance and handle noisy data.
- To evaluate the system's accuracy in identifying text from images.

Theory:

1 Optical Character Recognition (OCR)

- OCR is the process of converting images of typed, handwritten, or printed text into machine-encoded text.
- **Tesseract OCR**, an open-source OCR engine, supports multiple languages and can be trained on custom datasets.

2 Role of OpenCV

- **Image Preprocessing**: Improves the quality of the input image before OCR.
    - Grayscale conversion.
    - Noise removal (Gaussian/Median filtering).
    - Thresholding (binarization).
    - Edge detection and contour detection for identifying text regions.

- **Text Region Detection**: Helps isolate relevant portions of an image where text is likely to be present.

3 Deep Neural Network (DNN) for Text Identification

- While OCR engines like Tesseract perform well on clean images, performance drops with noisy, low-resolution, or distorted inputs.
- A **Deep Neural Network** (CNN-based or LSTM-based) can be used for:
    - Text detection (locating text in natural images).
    - Character recognition (classifying cropped character images).
- Example DNN workflow:
    - **Input Layer**: Preprocessed image patches.
    - **Convolutional + Pooling Layers**: Extract features such as edges, curves, and patterns.
    - **Fully Connected Layers**: Combine features for decision-making.
    - **Softmax Layer**: Outputs the probability distribution over character classes.

4 Workflow of Text Identification System

1. **Image Acquisition** – Input scanned or photographed image.
2. **Preprocessing with OpenCV** – Resize, denoise, threshold, and segment the image into text regions.
3. **Text Detection** – Use contour or DNN-based region detection to locate text.
4. **Text Recognition** – Apply Tesseract OCR to extract characters/words.
5. **Post-processing** – Use language models or spell-checkers to improve accuracy.
6. **Output** – Identified text is converted to editable and searchable digital text.

Conclusion:

In this assignment, we implemented a **text identification system** using **OpenCV**, **Tesseract OCR**, and **Deep Neural Networks**. OpenCV was responsible for preprocessing and enhancing image quality, Tesseract OCR extracted textual content, and the DNN helped improve accuracy in challenging conditions. This approach demonstrates the synergy between traditional computer vision methods, OCR technology, and modern deep learning techniques. The system can be further extended for multilingual OCR, handwritten text recognition, or real-time applications such as license plate and signboard recognition.