**Name :** Chandrakant Dattatrey Thakare  
**Roll No.:**382014 **PRN:**22310303

**Class :** TY CSE AI **Batch :** B1

**Subject :** DL Lab Assignments

**Assignment 05**

**Text identification using OpenCV, Tesseract (OCR) and deep neural network**

**Problem Statement:**

Implement text identification from images using **OpenCV for preprocessing**, **Tesseract OCR for text extraction**, and a **Deep Neural Network (DNN)** for improved recognition accuracy.

**Objective:**

* To understand the process of **text detection and recognition** from images.
* To apply **OpenCV** for preprocessing (grayscale, thresholding, noise removal).
* To use **Tesseract OCR** for extracting textual information.
* To design a **Deep Neural Network** that enhances OCR performance by classifying characters more accurately.
* To evaluate the model’s effectiveness on text images.

**S/W Packages and H/W apparatus used:**

* **Operating System:** Windows/Linux/MacOS
* **Kernel:** Python 3.x
* **Tools:** Jupyter Notebook, Anaconda, or Google Colab
* **Hardware:** CPU with minimum 4GB RAM; GPU optional for training DNN

**Libraries and packages used:**

* OpenCV
* Pytesseract
* TensorFlow / Keras
* Numpy
* Pandas
* Matplotlib

**Theory:**

**Definition:**  
Optical Character Recognition (OCR) is a technique to convert images of typed, handwritten, or printed text into machine-encoded text.

**Structure:**

1. **Input Image:** Raw image containing text (e.g., scanned document, signboard, license plate).
2. **Preprocessing:** Convert to grayscale, apply thresholding, and remove noise to improve text visibility.
3. **OCR Engine:** Tesseract detects characters and outputs text.
4. **Deep Neural Network:** Enhances OCR accuracy by classifying characters and correcting OCR errors.
5. **Output:** Recognized and refined text.

**Activation Functions:**  
ReLU is commonly used in hidden layers, while Softmax is used in the output layer for character classification.

**Methodology:**

**Step 1: Import Libraries**  
Load OpenCV, Pytesseract, TensorFlow/Keras, and utility libraries.

**Step 2: Load Dataset**  
Use sample text images such as scanned documents, license plates, or signboards.

**Step 3: Preprocessing with OpenCV**

* Convert image to grayscale.
* Apply thresholding (binary/OTSU).
* Use morphological operations (dilation, erosion) to clean text regions.

**Step 4: Apply OCR with Tesseract**

* Extract raw text using Pytesseract.
* Draw bounding boxes on detected text regions.

**Step 5: Build a Deep Neural Network**

* Create a Sequential model with Dense layers.
* Input layer: flattened image pixels.
* Hidden layers: ReLU activation.
* Output layer: Softmax activation for character classification.

**Step 6: Train the DNN**

* Train on labeled character dataset (A–Z, 0–9).
* Evaluate model performance.

**Step 7: Integrate OCR + DNN**

* Use OCR for initial text extraction.
* Pass uncertain/misclassified characters through the DNN for correction.

**Step 8: Evaluation**

* Compare OCR-only vs OCR+DNN outputs.
* Calculate recognition accuracy.

**Advantages:**

* **Improved Accuracy:** Preprocessing + DNN improves OCR performance.
* **Versatility:** Works on scanned documents, license plates, and natural scenes.
* **Automation:** Eliminates manual transcription.

**Limitations:**

* **Dependent on Image Quality:** Blurry or noisy images reduce performance.
* **Training Data Requirement:** DNN requires a labeled character dataset.
* **Computational Cost:** DNN training requires more resources.

**Applications:**

* Document digitization
* License plate recognition
* Passport/ID scanning
* Scene text recognition (street signs, billboards)
* Real-time translation apps

**Conclusion:**

In conclusion, this assignment demonstrated how **OpenCV preprocessing**, **Tesseract OCR**, and a **Deep Neural Network** can be combined for effective text identification. While Tesseract alone works well for clean images, the integration of preprocessing and a DNN enhances performance in noisy and complex scenarios. This system has practical applications in document scanning, license plate recognition, and real-time translation systems.