




Title:

Text Identification using OpenCV, Tesseract (OCR) and Deep Neural Network



Introduction to Text Identification

► What is Text Identification?

- The process of automatically detecting and extracting text from images or scanned documents.
- Used in a wide range of applications like reading signs, invoices, handwritten notes, etc.

► Why is it Important?

- Helps convert non-digital information into machine-readable text.
- Boosts automation, especially for digitizing documents or helping visually impaired users.



Overview of the Tools

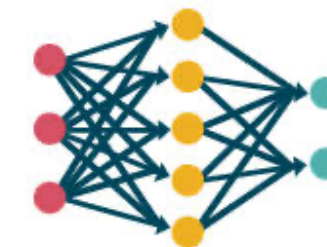


- A powerful tool for processing images and videos.
- Helps us clean, modify, and improve the quality of images before performing text extraction.



Tesseract OCR

- A leading tool for converting images of text into actual text.
- Works well with both printed and handwritten text.



DNN

- In text identification, DNNs help improve accuracy in detecting complex text, especially in challenging images.

Introduction to OCR

◆ What is OCR?

- OCR stands for Optical Character Recognition.
- Process to convert images of text (printed or handwritten) into machine readable text .
- Widely used for document digitalization , license plate recognitino, and automatic data extraction

◆ Overview of Tesseract OCR Engine

- Tesseract is open-source OCR engine developed by Google .
- It supports multiple languages and complex images text extraction .
- Highlt accurate and widely used due to its adaptability and performance .

◆ Integration of Tesseract with OpenCV

- OpenCV handles image preprocessing (resizing , thresholding , noise reduction).
- Tesserract extracts text from processed images , resulting in cleaner and more accurate recognition.
- Example workflow:
Load image with OpenCV.
Preprocess (grayscale, threshold).
Pass processed image to Tesseract for text extraction.

Preprocessing Techniques in OpenCV for OCR

◆ Grayscale Conversion

- Grayscale conversion reduces an image to shades of gray, eliminating color information.
- It uses a weighted formula, typically combining red, green, and blue values:
$$Y = 0.299R + 0.587G + 0.114B.$$
- This process simplifies image data, often used for tasks like image processing and computer vision.

◆ Noise Reduction

- Blurring: Removes noise by smoothing the image.
- Thresholding: Binarizes the image for clearer text boundaries.

◆ Edge Detection and Dilation/Erosion

- Edge Detection: Highlights text boundaries using methods like Canny Edge Detection.
- Dilation/Erosion: Enhances or reduces features like lines and gaps between characters.

Challenges in Text Identification

◆ Common Challenges:

- Varying Fonts and Sizes: Diverse fonts and text sizes affect recognition accuracy.
- Image Quality: Low-resolution images or noisy backgrounds complicate text extraction.
- Rotated or Skewed Text: Misaligned text can reduce OCR effectiveness.

◆ Solutions:

- Advanced Preprocessing: Apply rotation correction, adaptive thresholding, and image sharpening.
- Deep Learning Models: Use neural networks (e.g., EAST text detector) to better detect and extract irregular or complex text.



Real-Life Applications

01

Digitizing Documents: Convert physical documents into searchable, digital formats (like PDFs or Word documents).

02

Automatic Number Plate Recognition: Detecting and reading vehicle number plates automatically in parking lots or traffic systems.

03

Helping Visually Impaired Users: Apps that read out loud the text on a sign or document for people who cannot see.

04

Banking and Finance: Automating data entry from checks, receipts, and invoices using OCR.



Colab notebook



THANK YOU