# Image Processing Project

**Name**: Adham Mohamed Abdel-Aziz Mahmoud

**ID**: 2022/04058

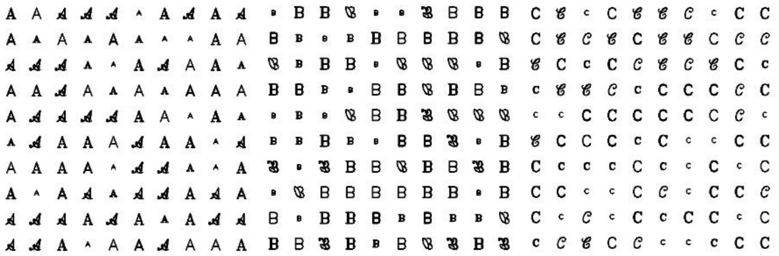
**Project**: *Text recognition in images and converting recognized text to speech*

## Project Overview:

This project aims to recognize text from images using Optical Character Recognition (OCR) and convert the recognized text into speech using a text-to-speech (TTS) system.

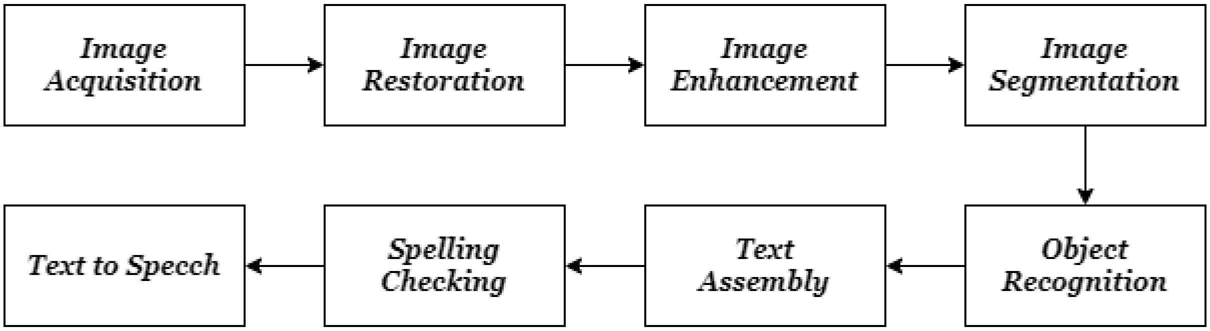
## Dataset:

- **Content**: A diverse collection of images containing characters or texts (e.g., scanned documents, street signs, handwritten notes, book pages).
- **Formats**: JPEG, PNG, TIFF, etc.
- **Diversity**:
  - **Text type**: Printed or handwritten.
  - **Languages**: Dataset should match the supported languages of the OCR tool.
  - **Image Quality**: Include both high-quality and noisy/blurred images.



## Sequence of steps:

The following diagram represent the sequence of steps used in the project



## Description of the steps:

1. **Image Acquisition**

   The process begins with obtaining the input image. This step involves either capturing an image using a camera, or loading an existing image from storage. In this project, OpenCV is used to load the image from the local file system.

2. **Image Restoration**

   Image restoration focuses on improving the quality of an image by removing unwanted noise or distortions.

   **Steps**:

   - **Dilation**: Enhances structures in the image by expanding bright areas.
   - **Erosion**: Restores the structure by shrinking the expanded areas, removing small noise.
   - **Morphological Closing**: Fills small gaps within text regions for better connectivity.
   - **Median Blurring**: Reduces salt-and-pepper noise while preserving edges.

3. **Image Enhancement**

   **Steps**:

   - **Grayscale Conversion:** Converts the image to a single intensity channel to simplify processing.
   - **Binary Thresholding:** Segments the image into text and background by converting pixel values to either black or white.
   - **Noise Removal:** Applies the noise removal process from Step 2 for further refinement.

4. **Image Segmentation**

   Segmentation is used to divide the image into meaningful regions, such as separating text from the background or identifying individual characters and lines.

   **Steps**:

   - Detect text lines using horizontal projections.
   - Segment words and characters using contour analysis or connected component labeling.

5. **Object Recognition**

   Using Machine Learning to recognize each Object

   **Steps**:

   - Train a model with the dataset we collected to classify the object according to characters classification (class for each character)

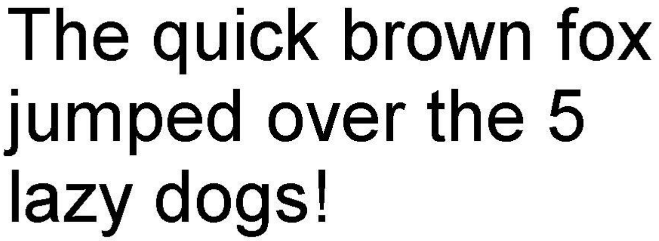
- **Feature Extraction for classification**
    - **HOG Features** are used to detect objects or characters in images by capturing gradient information (edges and textures).
    - **Horizontal and Vertical Projections:** Summing pixel intensities along rows and columns. These features help determine the overall structure and placement of characters in the image.
    - **Bounding Box**: Calculate the width and height of the bounding box surrounding the character to determine its size and aspect ratio.
    - Symmetry: Measure the symmetry of the character, which is useful for recognizing certain letters or shapes.
    - **Pixel Count**: Count the number of non-zero pixels in the image, which provides information about the area occupied by the character.
    - **Edge Density** measures the proportion of edge pixels in the image. It is calculated by applying the **Canny Edge Detection** algorithm to the image and counting how many pixels are part of edges.
    - **Corner Detection**: Detect key points in the image that are stable across different viewing angles and lighting conditions.
    - **Image Moments**: Moments are statistical properties of an image that can be used to describe the shape or distribution of pixel intensity. **Hu Moments** are invariant to translation, scaling, and rotation, making them useful for shape recognition.

6. **Text Assembly**
   If the text is segmented into parts (e.g., individual characters or words), this step assembles them into meaningful sentences or paragraphs. It ensures proper formatting and logical sequence.
   **Steps**:
    - Reformat the output into structured text.
7. **Spelling Checking**
   The extracted text often contains errors due to OCR inaccuracies. This step involves correcting spelling errors to produce a clean and accurate text.
   **Steps**:
    - Using library (pyspellchecker) for grammar and spell correction.
    - Fix common OCR misinterpretations.

8. **Text-to-Speech (TTS)**
   Once the text is extracted, it is converted into speech. This step utilizes a TTS engine to generate audio output from the recognized text. Users can save the output as an audio file or directly play it.
   **Steps**:
    - Using edge_tts engine.
    - Input the cleaned text and convert it to speech.
    - Save the speech output as an audio file.

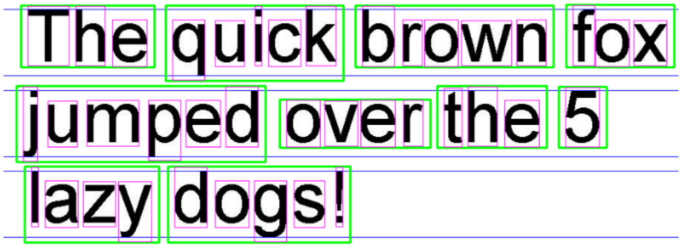
## Testing Results:

Segmentation result: Before



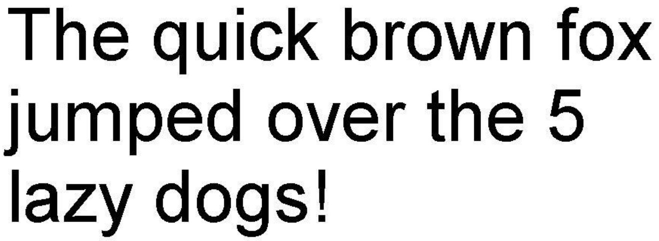
The quick brown fox
jumped over the 5
lazy dogs!

Segmentation result: After



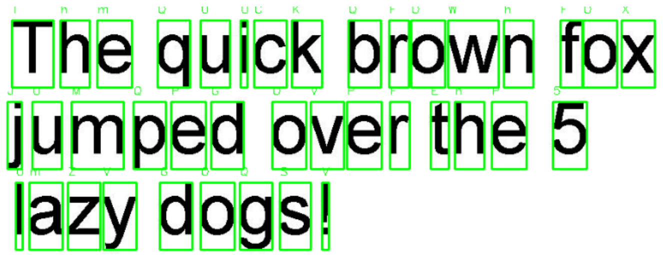
The quick brown fox
jumped over the 5
lazy dogs!

Word Recognition result: Before



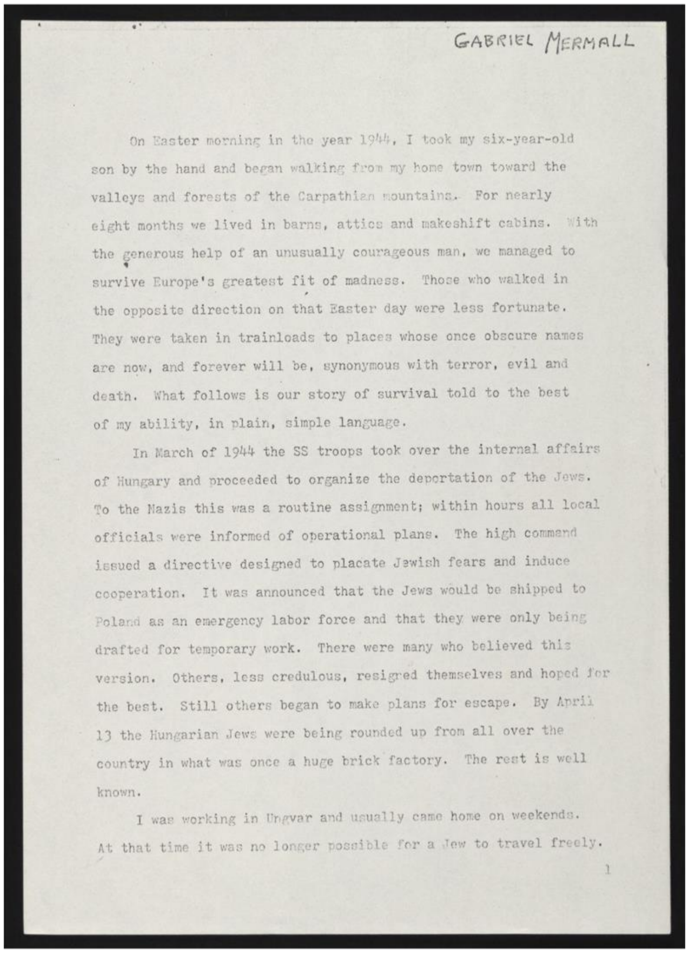
The quick brown fox
jumped over the 5
lazy dogs!

Word Recognition result: After



The quick brown fox
jumped over the 5
lazy dogs!

Word Recognition result: Before



GABRIEL MERMALL

On Easter morning in the year 1944, I took my six-year-old son by the hand and began walking from my home town toward the valleys and forests of the Carpathian mountains. For nearly eight months we lived in barns, attics and makeshift cabins. With the generous help of an unusually courageous man, we managed to survive Europe's greatest fit of madness. Those who walked in the opposite direction on that Easter day were less fortunate. They were taken in trainloads to places whose once obscure names are now, and forever will be, synonymous with terror, evil and death. What follows is our story of survival told to the best of my ability, in plain, simple language.

In March of 1944 the SS troops took over the internal affairs of Hungary and proceeded to organize the deportation of the Jews. To the Nazis this was a routine assignment; within hours all local officials were informed of operational plans. The high command issued a directive designed to placate Jewish fears and induce cooperation. It was announced that the Jews would be shipped to Poland as an emergency labor force and that they were only being drafted for temporary work. There were many who believed this version. Others, less credulous, resigred themselves and hoped for the best. Still others began to make plans for escape. By April 13 the Hungarian Jews were being rounded up from all over the country in what was once a huge brick factory. The rest is well known.

I was working in Ungvar and usually came home on weekends. At that time it was no longer possible for a Jew to travel freely.

1

Word Recognition result: After