Numbers Extraction OCR

## What it Does?

-Reads image from the DATA folder

-Runs OCR using Paddle OCR (the best OCR for irregular images)

-Runs OCR using **Google Cloud Vision** and extracts numbers.

-Saves extracted numbers using both models in a csv.

## What You Need?

-**Install Dependencies**

pip install opencv-python numpy paddleocr google-cloud-vision

-**Enable Google Cloud Vision API**

**-**Create a service account and download the JSON key file.

-Set the GOOGLE\_APPLICATION\_CREDENTIALS environment variable to the file path.

## How to Run?

-Place your images in the folder specified by folder\_path.

-Run the script. The extracted numbers will be saved in extracted\_text\_numbers\_combined.csv.

-If you have no Google cloud account you can input “no” so that only paddle OCR output will be added to the CSV. Otherwise (you have a Google cloud account)you can input “yes”.