

Gateway OCR

Document status	DRAFT
Document owner	Jun Song
Designer	Mooyoung Oh, Jun Song
Technical writers	Jun Song
QA	Jun Song

Objective

Container Transaction , Yard Container , YMS

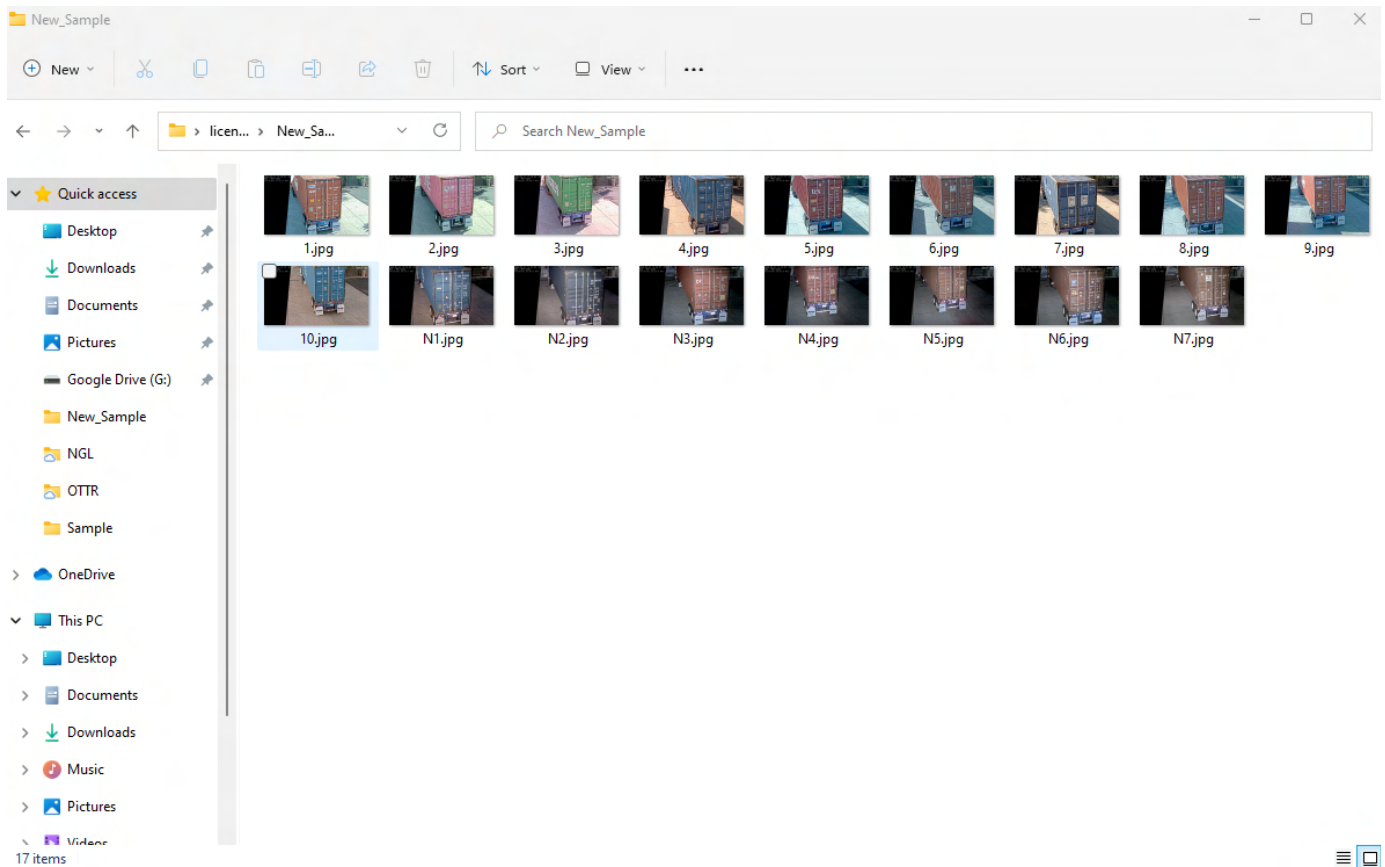
Success metrics

Goal	Metric
Sustainable development & Scalability	Readable and documented codes
No missing transactions	Increasing Detection Accuracy

WorkFlow

1. Our CCTV system - Synology
- There is motion detect system. If there is movement, The scene will be captured, and then it will be sent to the file Server by FTP. The image files will be stored as "Sample_File".





2. Detecting the container number region YOLOV5 machine learning Library

- Detect the Container number Region

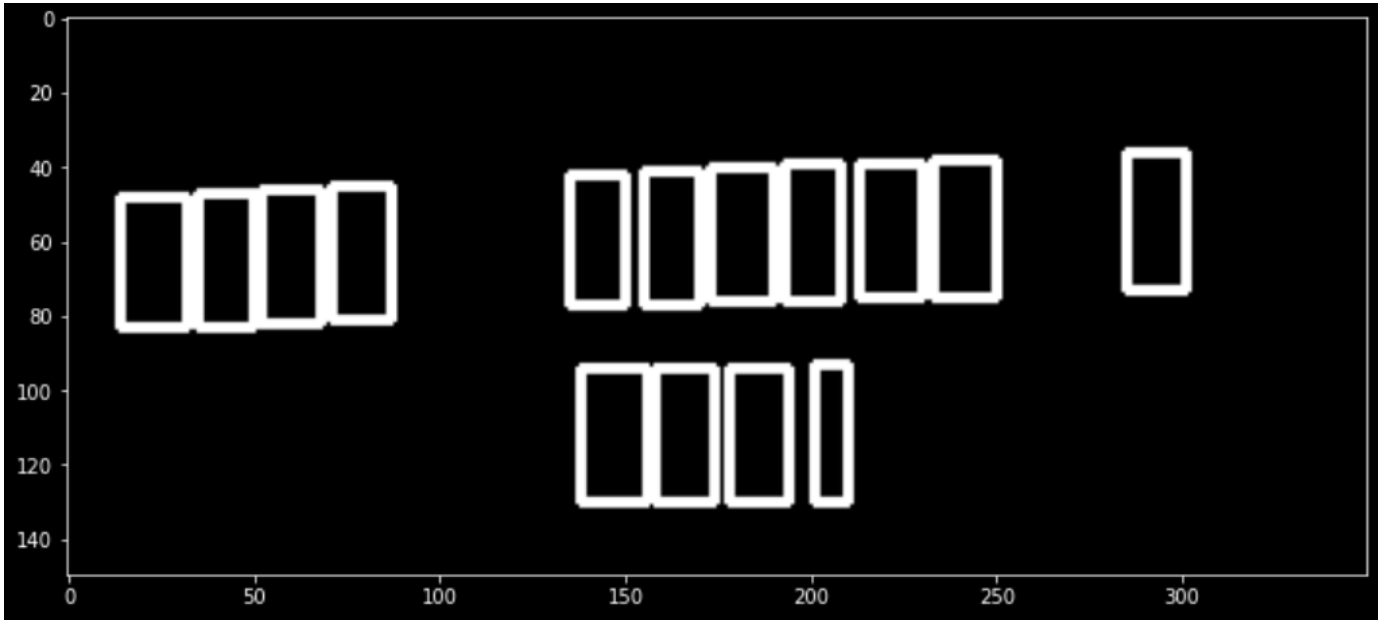


- Stack the Dataset, Images or Video file

The screenshot shows a Windows File Explorer window. On the left sidebar, the 'NAV_Noo' folder is highlighted. The main area shows a list of files under the 'Yesterday' view. The files are as follows:

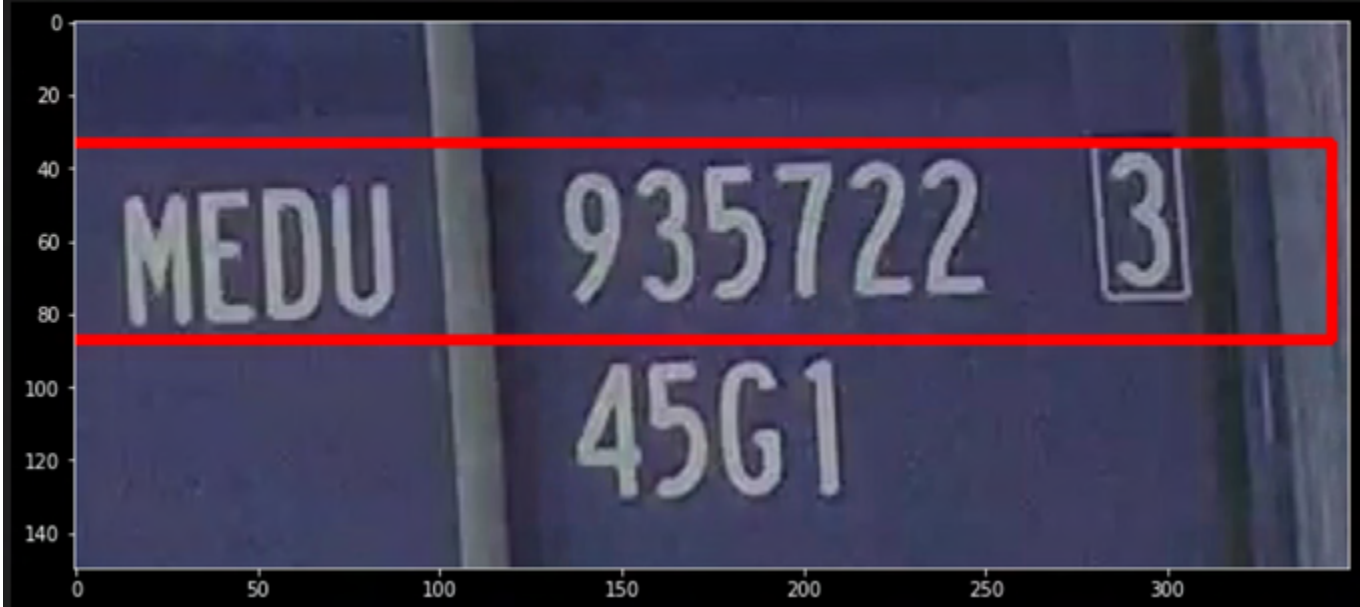
Name	Date modified
221005_034457_1_0000000.jpg	12/19/2022 3:38 PM
221005_041700_1_0000000.jpg	12/19/2022 3:38 PM
221005_003416_0_0000005.jpg	12/19/2022 3:38 PM
221005_004315_1_0000000.jpg	12/19/2022 3:38 PM
221005_005604_1_0000000.jpg	12/19/2022 3:38 PM
221004_233817_1_0000000 (2).jpg	12/19/2022 3:38 PM
221004_233817_1_0000000.jpg	12/19/2022 3:38 PM
221004_235256_1_0000005.jpg	12/19/2022 3:38 PM
221004_215806_1_0000000.jpg	12/19/2022 3:38 PM
221004_222125_1_0000000.jpg	12/19/2022 3:38 PM
221004_233817_1_0000000 (1).jpg	12/19/2022 3:38 PM
221004_213449_1_0000032 (2).jpg	12/19/2022 3:38 PM
221004_213449_1_0000032.jpg	12/19/2022 3:38 PM
221004_213812_1_0000000.ipa	12/19/2022 3:38 PM





MEDU9357223

<matplotlib.image.AxesImage at 0x1cd01db5030>



4. Uploading on YMS

- Date, Time, Container number will be uploaded on YMS.

 Develop History

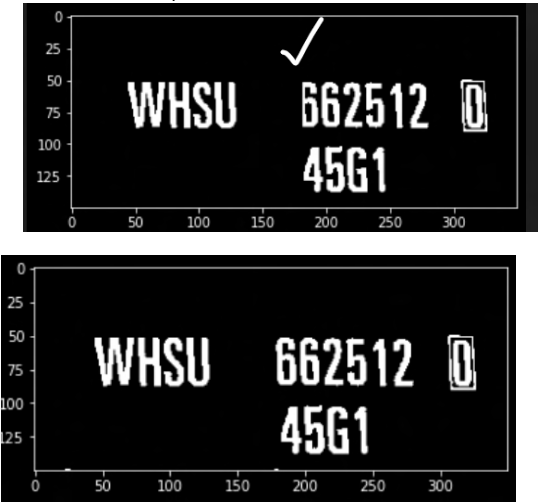
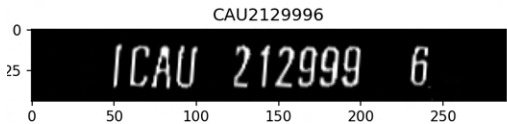
NGL_OCR_PROJECT

- https://github.com/songjeongjun320/NGL_OCR

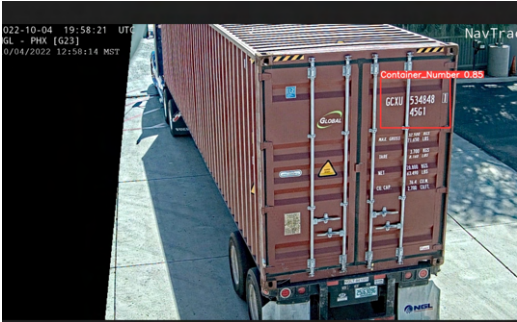
Date	User Story	Modified Strategy
------	------------	-------------------

09/02/2022	<ul style="list-style-type: none"> Establishing Strategy Mapping the Logic <ul style="list-style-type: none"> Pull out the picture on Synology automatically Extract the container number Update on YMS 	
09/05/2022	Utilize Open Source <ul style="list-style-type: none"> https://github.com/kairess/license_plate_recognition 	Selecting main Library <ul style="list-style-type: none"> Python Tesseract Opencv Numpy Matplotlib
09/08/2022	<ol style="list-style-type: none"> Modified Strategy of Extracting Container Number Reason to choose "Tesseract" <ul style="list-style-type: none"> Library is not heavy than the other Library (PaddlePaddle, YOLO) Easy to find the information <ul style="list-style-type: none"> Utilization with Opencv, Matplotlib, Numpy 	Extracting Strategy <ul style="list-style-type: none"> Read Image Cut the Image Convert Image to BlackandWhite Maximize Contrast Adaptive Thresholding Find Contours Preparing Data Select Candidates Thresholding Result
09/15/2022	Edit the Extracting Strategy <ul style="list-style-type: none"> Remove "Maximize Contrast" <ul style="list-style-type: none"> If I wanna do maximize contrast, the camera's pixel should be high, but our camera is not as much as 12million pixel. Remove "Adaptive Thresholding" <ul style="list-style-type: none"> If doing Thresholding twice, The pixels are broken. One time Thresholding is enough. 	Extracting Strategy <ul style="list-style-type: none"> Read Image Cut the Image Convert Image to BlackandWhite Find Contours Preparing Data Select Candidates Thresholding Result
9/21/2022	Add 2 strategies <ul style="list-style-type: none"> Denoise the picture – Using the fastNIMeansDenoisingColored in OpenCV <ul style="list-style-type: none"> Changing ThresHold from THRESH_BINARY to THRESH-OTSU On staging <ul style="list-style-type: none"> Removing Contour surrounding the number – Divide contours and remove the surrounded one – Fade the surrounded contour	
9/22/2022	<ul style="list-style-type: none"> Adjust Image Brightness : Make dark on sun, Make bright after sun Lines are made a little thinner so that anything other than text is erased. In particular, the last digit of the container number was surrounded by a rectangle on the outside, but it succeeded in erasing it. 	Add "Erosion" on main

10/3/2022	<ul style="list-style-type: none"> Succeeded to erase useless area. 	
10/11/2022	<ul style="list-style-type: none"> Holding the "Tensorflow" object. Issue : Last number is contoured by square. <p>– Solution : Using "Teachablemachine" from Google.</p> <p>– Educate the individual numbers which are contoured by square and detect them to exact number.</p>	<ul style="list-style-type: none"> Change the "Erosion" order Delete "Another Thresholding to find chars" <p><<Extracting Strategy>></p> <ul style="list-style-type: none"> Read Input Image Cut the Image Convert Image to BlackandWhite Find Contours Prepare Data Select Candidates Rotate Plate Images Erosion & Detect Result
10/12/2022	<ul style="list-style-type: none"> Removing "MAX_AREA_DIFF" & "MAX_WIDTH_DIFF" on Arrangement Contours Picture's size influenced the OCR accuracy. <p>– Size should be enough to detect the numbers.</p> <p>– Like First picture, It should have some space on the wall.</p> <p>– Vertical size is more important than Horizontal.</p> 	<ul style="list-style-type: none"> Change the strategy on "Cut the Image" table. <p>– Cutting the image with wide wall.</p>

10/13/2022	<ul style="list-style-type: none"> Video quality fixed as 2560 x 1440. Low (640 x 360) & High (3840 x 2160) are not efficient on OCR. Testing "Dilate the picture" on "Convert Img2BlckWht" Table 	
10/19/2022	<ul style="list-style-type: none"> Amending table "Convert to BlackandWhite" - cv2.threshold <ul style="list-style-type: none"> Expanding the range of colors recognized as letters Option should be changed by the light (depends on Shadow) . Issue : White wall & Black letters Tablet "Cut the Image" : <ul style="list-style-type: none"> cv2.fastNlMeansDenoisingColored (img_ori, None, 10,10,41,41) or (img_ori, None, 10,10,71,71) 	<ul style="list-style-type: none"> Expanding the accepted color ranges under the shadow Need to reinforce the case which "out of shadow" Images.
10/27/2022	<ul style="list-style-type: none"> Developing simulation.py, Test code <ul style="list-style-type: none"> Automation of loading and reading images. Exception Handling issues. 	
10/31/2022	<ul style="list-style-type: none"> Developed run.py, Test code <ul style="list-style-type: none"> Test success rate 20% 	
11/2/2022	<ul style="list-style-type: none"> Success of erase the bar between alphabet and numbers.  	<ul style="list-style-type: none"> Tablet "Prepare Data" <ul style="list-style-type: none"> Change erase function from "cv2.rectangle" to "cv2.minAreaRect"
11/7/2022	<ul style="list-style-type: none"> Through the Test process (run.py), Found the better Constant. <ul style="list-style-type: none"> kernel(2,2) denoisingcolored(10,10,41,41) threshold(gray, 80, 255) success rate : 24 % 45 % 	
11/10/2022	<ul style="list-style-type: none"> Test with NAV samples. Add code which if contours size is smaller than X, ignore the contour. <ul style="list-style-type: none"> To make wider range on samples Table "Visualize Possible Contours" <ul style="list-style-type: none"> Between character contours and number contours, there is space. If the space is detected, that space's every color to be black. Same process on both end side of the image. 	<ul style="list-style-type: none"> Tablet "Prepare Data" <ul style="list-style-type: none"> add (elif h < 20: continue)
11/15/2022	<ul style="list-style-type: none"> In specific condition, reverse black & white. Functionalized. 	
11/17/2022	<ul style="list-style-type: none"> Last digit's square contour erased. <ul style="list-style-type: none"> def visualize_possible_cntrs constant : erase_range  	

11/18/2022	<ul style="list-style-type: none"> • findcontours method test on "Find Contours" Tablet. - CHAIN_APPROX_TC89_L1 is selected. (53%) • Container number's header adjustment started. 	
11/23/2022	<ul style="list-style-type: none"> • Create cntr_adjustment function. - Create DB for cntr's header. ex) [APHU, BEAU, OOLU ...] - If OCR read APH or APHI, It will be adjusted as APHU. • Established Condition when the white & black reverse. - If len(result_chars) < 8, It will be reversed. 	
11/27/2022	<ul style="list-style-type: none"> • When expanding the plate size, expanded part has some issue. - Expanded part also be black to read without issue. • Enforcing cntr_adjustment fuction. - Will divide depending on the number of head read by OCR. - There are 3,4 or more head. • Gathering picture sample for night time. 	
11/30/2022	<ul style="list-style-type: none"> • Issue fix : Eroded the IMG which doesn't need to be eroded. - Before eroding the img, If the result is predicted the answer. just return the answer with erosion. • Total : 337 • Success : 254 • Failure : 83 • Success Rate : 75.37 % 	
12/02/2022	<ul style="list-style-type: none"> • Add cntr_last_digit function, through logic of how to make Container #. - https://blog.naver.com/dhrudwn12345/222180809747 - Logic is being created, this logic will be tested next week. - Total : 476 - Success : 357 - Failure : 111 - Head Issue : 20 	
12/08/2022	<ul style="list-style-type: none"> • Machine Learning environment setting - CUDA, CUDNN, Pytorch, Pycharm, Anaocnda • Creating Container_behind & Container_number dataset - Using Roboflow 	

12/19/2022	<ul style="list-style-type: none"> YOLOV5 License Check : GNU 3.0 / Commercial use permitted - https://github.com/ultralytics/yolov5/blob/master/LICENSE Trained "Vehicle_Dataset_result". - Need to implement a more precise model Installed the Camera on gate and testing. - Succeeded to connect Camera to YOLOV5 through RTSP.   <ul style="list-style-type: none"> Enforcing the Code : New Created - def cntr_head_adjust - def cntr_head_len_over4 - def cntr_head_len_under4 - def cntr_head_len_4 Test Result : 75.37 % 89.56 % 	
12/29/2022	<ul style="list-style-type: none"> YOLOV5 Live detecting the cntr # region If cntr # region is detected, It would be trigger. And OCR engine will run. When detect the result, send YMS with .json file. Created trigger.py - If cropped image is detected on file, OCR engine will read the cropped image. denoise process whole fixed. - To make code run faster, removing the denoise function. 	
1/4/2023	<ul style="list-style-type: none"> YOLOV5 - file_location issue fixed - file_location name fixed : create def f_name(): -> make file to save cropped images and video. OCR_Engine - Created def make_json(result, name): -> When the cntr # detected, it is anticipated right cntr #. -> Make .json files on "date" folder. - Created def crops(files, dir): -> pull cropped images to OCR engine. Read through Main_by_Trigger.main 	<ul style="list-style-type: none"> detect.py - def f_name(): trigger_OCR.py - def make_json(result, name): - def crops(files, dir):
1/6/2023		

- Success uploading img files on AWS S3 DB.

Amazon S3 > Buckets > ngl-yms

ngl-yms

Info

Publicly accessible

Objects

Properties

Permissions

Metrics

Management

Access Points

Objects (4)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh

Copy S3 URI

Copy URL

Download

Open

Delete





Actions

Create folder

Upload

Find objects by prefix

< 1 > ⚙

<input type="checkbox"/>	Name ▲	Type ▼	Last modified ▼	Size ▼	Storage class
<input type="checkbox"/>	 01062023/	Folder	-	-	-
<input type="checkbox"/>	 010620231/	Folder	-	-	-
<input type="checkbox"/>	 010620234/	Folder	-	-	-
<input type="checkbox"/>	 010620236/	Folder	-	-	-

- Success uploading .json files on MySQL server.

trans_id	channel_no	container_no	detected_time	img_dir	img_path	load_length	type_id	img_date	img_user_id	serial_no
1	PHX11567628	PHX11567628	2023-01-08 07:00:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
2	PHX11567628	PHX11567628	2023-01-08 08:42:34.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
3	PHX11567628	PHX11567628	2023-01-08 13:43:34.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
4	PHX11567628	PHX11567628	2023-01-08 13:41:04.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
5	PHX11567628	PHX11567628	2023-01-08 13:53:12.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
6	PHX11567628	PHX11567628	2023-01-08 14:00:27.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
7	PHX11567628	PHX11567628	2023-01-08 14:03:34.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
8	PHX11567628	PHX11567628	2023-01-08 14:12:57.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
9	PHX11567628	PHX11567628	2023-01-08 14:16:55.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
10	PHX11567628	PHX11567628	2023-01-08 14:17:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
11	PHX11567628	PHX11567628	2023-01-08 14:38:53.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
12	PHX11567628	PHX11567628	2023-01-08 14:40:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
13	PHX11567628	PHX11567628	2023-01-08 14:42:48.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
14	PHX11567628	PHX11567628	2023-01-08 14:43:38.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
15	PHX11567628	PHX11567628	2023-01-08 14:44:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
16	PHX11567628	PHX11567628	2023-01-08 14:45:17.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
17	PHX11567628	PHX11567628	2023-01-08 14:45:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
18	PHX11567628	PHX11567628	2023-01-08 14:47:32.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
19	PHX11567628	PHX11567628	2023-01-08 14:57:32.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
20	PHX11567628	PHX11567628	2023-01-08 15:03:12.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
21	PHX11567628	PHX11567628	2023-01-08 15:03:21.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
22	PHX11567628	PHX11567628	2023-01-08 15:06:38.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
23	PHX11567628	PHX11567628	2023-01-08 15:06:55.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
24	PHX11567628	PHX11567628	2023-01-08 15:14:18.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
25	PHX11567628	PHX11567628	2023-01-08 15:16:34.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
26	PHX11567628	PHX11567628	2023-01-08 15:17:36.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000
27	PHX11567628	PHX11567628	2023-01-08 15:21:15.000000	CHOUT-080000-080000	01042023\CHOUT-080000-080000	0000	0000	0000	0000	0000

- Change DB Server from MySQL to PostgreSQL.

1/9/2023

- DB Server changed from MySQL to PostgreSQL.
 - Checked the data stacked automatically on greySQL server.
- Installed "PHX-CKIN" CCTV Camera.
 - Add "Division" & "Camera location" column on .json file.

1/20/2023

- CNTR Size Detect Start
 - def cntr_size_contour
 - def cntr_size_contour_draw
 - Add parameters on def find_chars
 - Add parameters on def rotate_plate_img



- Need to revise the result to make correct Answers.

Reference

Reference	Contents
Python	Computer Language
Tesseract	Open Source Computer OCR Library https://github.com/tesseract-ocr/tesseract
Opencv	Open Source Computer Vision Library https://opencv.org/
Numpy	The fundamental package for scientific computing with Python https://numpy.org/doc/stable/

Matplotlib	Visualization Package https://matplotlib.org/
Synology	Synology CCTV Surveillance System https://www.synology.com/en-us
YOLOV5	Machine Learning Library https://github.com/ultralytics/yolov5
Reolink	Detecting Camera https://reolink.com/