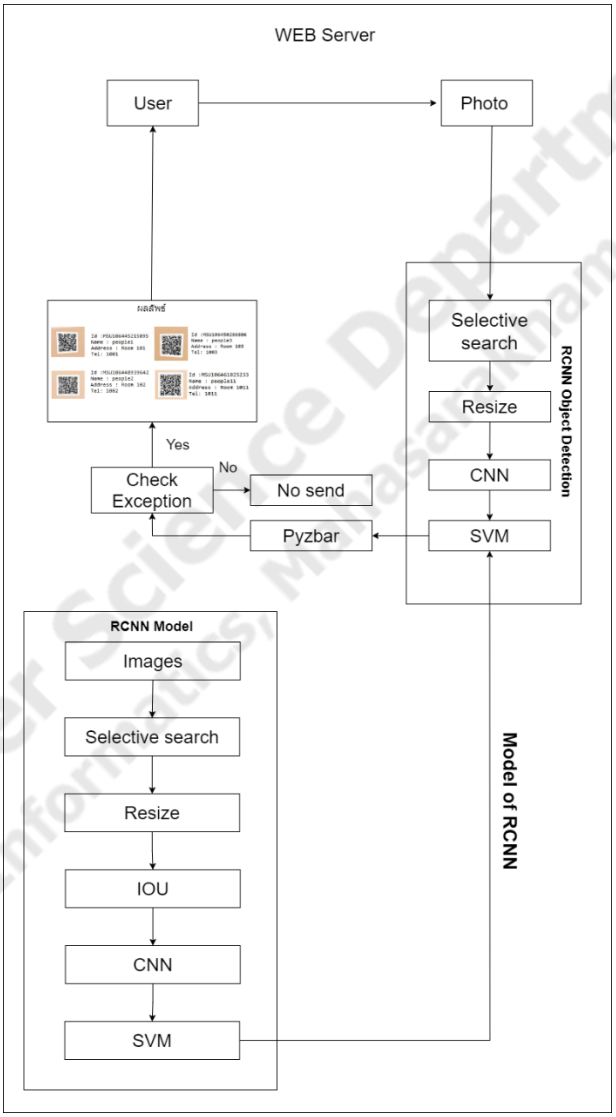
บทที่ 3
วิธีการดำเนินงานวิจัย

3.1 กรอบการดำเนินงาน

ระบบตรวจจับและรู้จำ โดยรายละเอียดแสดงดังภาพประกอบที่ 3.1



ภาพประกอบที่ 3.1 กรอบการทำงาน

กระบวนการทำงาน ผู้ใช้งานจะใช้การถ่ายภาพนิ่งหรือส่งภาพนิ่งเพื่อทำการส่งข้อมูลภาพนิ่ง ไปผ่านกระบวนการ RCNN Object Detection โดยมีขั้นตอนการทำงานคือ

1) ใช้ Selective Search กับภาพที่นำเข้า

Computer Science Department
Faculty of Informatics, Mahasarakham University

2) ปรับขนาดภาพประกอบที่ได้จากขั้นตอนที่ 1 ให้เท่ากันทั้งหมด
3) ทำการสกัดคุณลักษณะภาพด้วย CNN
4) ใช้โมเดลแยกประเภทด้วย SVM

โดยโมเดลที่คัดแยกประเภทของภาพที่ได้จากหัวข้อย่อยที่ 4 นั้นได้มาจากการการใช้ RCNN Model โดยกระบวนการดังกล่าว มีขั้นตอนการทำงานคือ

1) นำเข้ารูปภาพต้นแบบโมเดล
2) ใช้ Selective Search
3) ปรับขนาดของภาพประกอบที่ได้จาก ขั้นตอนที่ 1 ให้เท่ากันทั้งหมด
4) ใช้ IOU เพื่อทำการให้การแบ่งแยกประเภทโมเดลง่ายขึ้น
5) ทำการสกัดคุณลักษณะด้วย CNN
6) สร้างโมเดลโดนใช้ SVM

หลังจากที่ผ่านกระบวนดังกล่าวมาจะได้รูปภาพ QR Code เพื่อนำไปถอดรหัสเป็นตัวอักษร โดยใช้ Library ของ Python ชื่อว่า “Pyzbar” และจะทำการเช็คข้อมูล หากข้อมูลที่ Pyzbar ถอดรหัสออกมาเป็นตัวอักษร จะทำการตัวอักษรเหล่านั้นแนบไปกับรูปภาพ QR Code ที่ตัดแยกออกมาและส่งข้อมูลไปยังผู้ใช้งาน แต่ถ้าหากรูปภาพนั้นไม่สามารถถอดรหัสได้จะถือว่าภาพนั้นไม่ใช่รูปภาพ QR Code และจะไม่ทำการส่งค่าข้อมูลของรูปนั้นไปยังผู้ใช้งาน

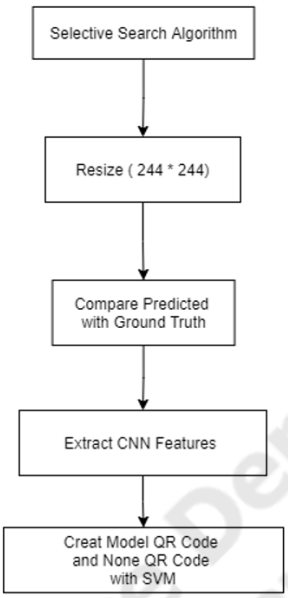
3.2 การทำงานส่วนของการสร้างโมเดล (Region Convolution Neuron Network Model)

การทำงานส่วนของการสร้างโมเดลจะใช้การเรียนรู้เชิงลึกแบบ R-CNN (Region Convolution Neuron Network) การสร้างโมเดลหรือการเรียนรู้คิวอาร์โค้ดจะต้องเตรียมภาพเพื่อเรียนรู้รูปแบบของคิวอาร์โค้ด โดยจะเตรียมข้อมูลนำเข้าเพื่อสร้างโมเดลโดยมีรูปภาพดังภาพประกอบที่ 3.2

ภาพประกอบที่ 3.2 ภาพนำเข้า RCNN Model

Computer Science Department
Faculty of Informatics, Mahasarakham University

การทำงานในส่วนของการสร้าง Model โดยใช้ RCNN มีอยู่ 5 ขั้นตอน รายละเอียดการทำงานของโมเดลแสดงดังภาพประกอบที่ 3.3



ภาพประกอบที่ 3.3 แผนผังการทำงานส่วนของการสร้าง Model โดยใช้ RCNN

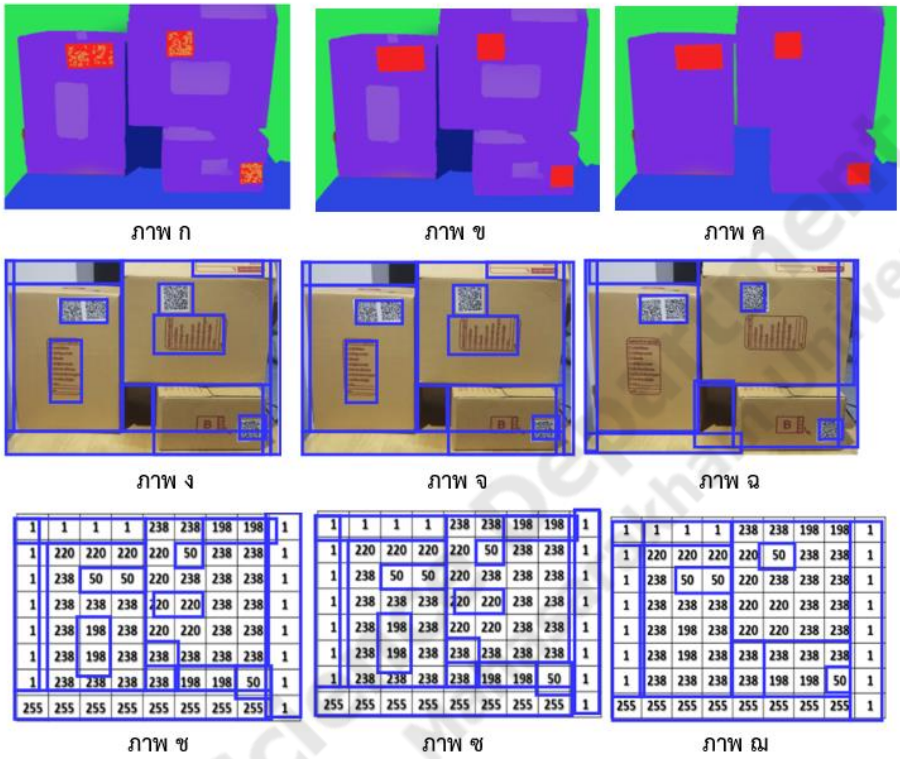
ขั้นตอนที่ 1 เสนอพื้นที่โดยใช้ Selective Search Algorithm โดยข้อมูลนำเข้า คือภาพที่ประกอบด้วยกล่องที่มี QR Code และพื้นหลังแสดงภาพประกอบที่ 3.4

ภาพประกอบที่ 3.4 ข้อมูลนำเข้า Selective search

กระบวนการทำงาน Selective search จะทำการเลือกพื้นที่จากช่วงค่าของสีของภาพแสดงดังภาพ ก ในภาพประกอบที่ 3.5 และจะทำการรวมพื้นที่ที่มีสีใกล้เคียงกันให้กลายเป็นพื้นที่สีเดียวกันดัง

Computer Science Department
Faculty of Informatics, Mahasarakham University

ภาพ ข หลังจากนั้นจะทำการยูเนี่ยนพิเซลที่มีลักษณะใกล้เคียงกันให้กลายเป็นพื้นที่สีเดียวกันดังภาพ ช เมื่อแบ่งค่าช่วงสีได้แล้วจะทำการตีกรอบเลือกพื้นที่ที่มีช่วงสีและนำไปเป็นผลลัพธ์



ภาพประกอบที่ 3.5 กระบวนการ Selective search

ภาพประกอบที่ 3.6 ผลลัพธ์พื้นที่ที่ถูกเสนอด้วย Selective Search

ผลลัพธ์ จะได้พื้นที่ที่ถูกเสนอออกมาโดยจะสร้างกรอบขอบเขต (Bounding Box) เอาไว้บริเวณที่เป็นลักษณะเด่นของภาพหรือเป็นกลุ่มของช่วงสีของแต่แต่ละช่วงสีนั้นๆ

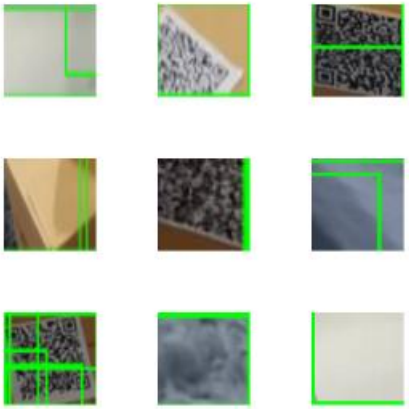
ขั้นตอนที่ 2 ปรับขนาดภาพของวัตถุ ทำการปรับขนาดภาพให้ให้เป็น 224 * 224 ให้ภาพมีขนาดเท่ากัน เนื่องจากเป็นขนาดมาตรฐานและใช้งานทั่วไป เมื่อปรับขนาดเสร็จแล้วจึงจะนำไปขั้นตอนถัดไป ข้อมูลนำเข้า เป็นภาพทั้งหมดที่ได้จาก Selective search ภาพประกอบที่ 3.7

ภาพประกอบที่ 3.7 ผลลัพธ์พื้นที่ที่ถูกเสนอด้วย Selective Search

การประมวลผล จะนำทุกภาพจากกรอบสีเขียวมาปรับให้มีขนาด 224*224 พิกเซล แสดงดังภาพประกอบที่ 3.8

ภาพประกอบที่ 3.8 การปรับขนาดรูปภาพ

ผลลัพธ์ จะได้ภาพจากกรอบสีเขียวที่มีขนาด 224*224 พิกเซล ทุกภาพแสดงดังในภาพประกอบที่ 3.9



ภาพประกอบที่ 3.9 ภาพผลลัพธ์จาก Selective search

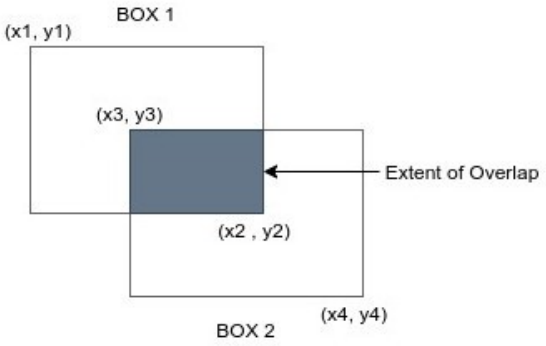
ขั้นตอนที่ 3 แบ่งประเภทของภาพเป็นกลุ่ม QR Code และ Non QR Code ด้วย IOU ข้อมูลนำเข้า จะแบ่งเป็น 2 แบบคือ 1.ภาพที่ได้จากขั้นตอนที่ 2 ดังภาพประกอบที่ 3.10 และการสร้างภาพ Ground Truth แสดงดังภาพประกอบที่ 3.10 โดยในโครงงานนี้จะตีกรอบ QR Code เป็นภาพ Ground Truth แสดงดังภาพที่ 3.10

Ground Truth

ภาพประกอบที่ 3.10 ภาพนำเข้า Ground Truth

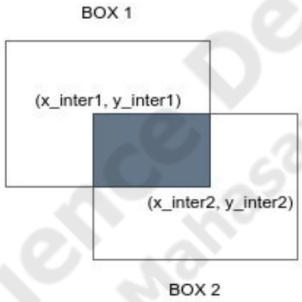
กระบวนการทำงาน นำภาพทั้งหมดที่ได้จากขั้นตอนที่ 2 มาเปรียบเทียบกับภาพ Ground Truth ทุกภาพ โดยมีการพิจารณาความเหมือนของภาพ Predict กับ ภาพ Ground Truth ดังนี้

1) คำนวณพื้นที่จุดตัดของกล่องโดยกล่องที่ 1 หรือพื้นที่ที่ถูกเสนอมาจาก Selective Search เรียกว่า พื้นที่คาดการ์ณ (Predicted) และกล่องที่ 2 คือ พื้นที่จริง (Ground Truth)



ภาพประกอบที่ 3.11 จุดตัดของสองกล่อง

2) แสดงพิกัดของสี่เหลี่ยมที่ตัดกันเป็น [x_inter1, y_inter1, x_inter2, y_inter2] เพื่อแสดงพิกัดของด้านบนซ้ายและขวาล่างของจุดตัด



ภาพประกอบที่ 3.12 การตั้งชื่อพิกัดที่ต้องคำนวณ

3) เริ่มต้นด้วยการกำหนดระบบพิกัดอ้างอิงโดยจะใช้รูปแบบคอมพิวเตอร์กราฟิกของแกนบวก โดย X จะเคลื่อนไปทางขวาและแกน Y เคลื่อนลงด้านล่าง

4) ในการคำนวณมุมบนซ้ายของจุดตัด เปรียบเทียบมุมซ้ายบนของแต่ละช่อง ในทางคณิตศาสตร์

x_inter1 สามารถพบได้โดยดูว่าช่องใดมีมุมบนซ้ายมากกว่าไปทางขวา

y_inter1 สามารถพบได้โดยดูว่าช่องใดมีมุมซ้ายบนต่ำกว่าอีกช่อง

สามารถคำนวณได้ดังนี้:

$$\text{x_inter1} = \max(x1, x3) \qquad (1)$$

$$\text{y_inter1} = \max(y1, y3) \qquad (2)$$

5) ในการคำนวณมุมขวาล่างของจุดตัด เปรียบเทียบมุมขวาล่างของแต่ละช่อง

x_inter2 สามารถพบได้โดยดูว่าช่องใดมีมุมขวาล่างอยู่ทางซ้ายมากกว่า

y_inter2 สามารถพบได้โดยดูว่าช่องใดมีมุมล่างขวาสูงกว่าอีกกล่อง

Computer Science Department
Faculty of Informatics Mahasarakham University

ในทางคณิตศาสตร์สามารถคำนวณได้ดังนี้:

$$x_inter2 = \max(x2, x4) \quad (1)$$

$$y_inter2 = \max(y2, y4) \quad (2)$$

6) กรณีที่กล่องทับซ้อนกันอย่างสมบูรณ์หรือมีค่าเดียวกันสำหรับพิกัดเฉพาะ กรณีนี้จะมีเดียวจะค่าเท่ากับ 1 สามารถคำนวนได้ดังนี้

$$\max(1,1) = 1$$

$$\min(1,1) = 1$$

7) หลังจากได้พิกัดของจุดตัด จะนำค่าทั้งหมดมาคำนวณพื้นที่ของจุดตัด สามารถคำนวณได้ดังนี้

$$width_inter = (x_inter2 - x_inter1) \quad (1)$$

$$height_inter = (y_inter2 - y_inter1) \quad (2)$$

$$area_inter = width_inter * height_inter \quad (3)$$

ในการหาพื้นที่ทั้งหมดก่อนอื่นเราจะคำนวณพื้นที่ของแต่ละกล่อง สามารถคำนวณได้ดังนี้

$$width_box1 = (x_2 - x_1) \quad (1)$$

$$height_box1 = (y_2 - y_1) \quad (2)$$

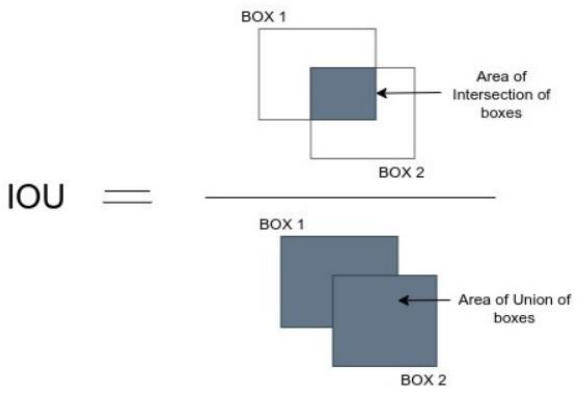
$$width_box2 = (x_2 - x_1) \quad (3)$$

$$height_box2 = (y_2 - y_1) \quad (4)$$

$$area_box1 = width_box1 * height_box1 \quad (5)$$

$$area_box2 = width_box2 * height_box2 \quad (6)$$

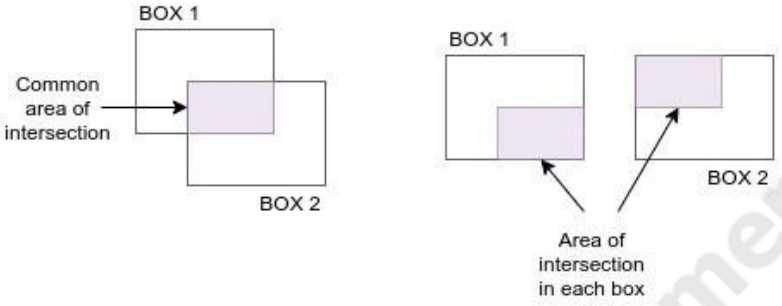
การรวมกันของกล่องคือพื้นที่ทั้งหมดที่ครอบคลุมโดยทั้งสองกล่องดังแสดงในภาพประกอบที่ 3.13



ภาพประกอบที่ 3.13 การแสดงแผนภาพของสูตรเพื่อคำนวณ IOU

Computer Science Department
Faculty of Informatics, Mahasarakham University

พื้นที่ทั้งหมดที่ครอบคลุมโดยทั้งสองกล่องจะทราบว่าส่วนของจุดตัดนั้นครอบคลุมอยู่ในทั้งสองกล่องกล่าวคือพื้นที่ของจุดตัดจะรวมอยู่ใน area_box1 และ area_box2



ภาพประกอบที่ 3.14 การแสดงพื้นที่ทางแยกที่ครอบคลุมในทั้งสองกล่อง

เนื่องจากเราต้องการคำนวณพื้นที่ส่วนกลางของจุดตัดเพียงครั้งเดียวเราจึงสามารถลบพื้นที่จุดตัดที่เราคำนวณออกจากพื้นที่ทั้งหมดของสองกล่อง สามารถคำนวณได้ดังนี้

$$\text{areaunion} = \text{area_box1} + \text{area_box2} - \text{area_intersection} \quad (1)$$

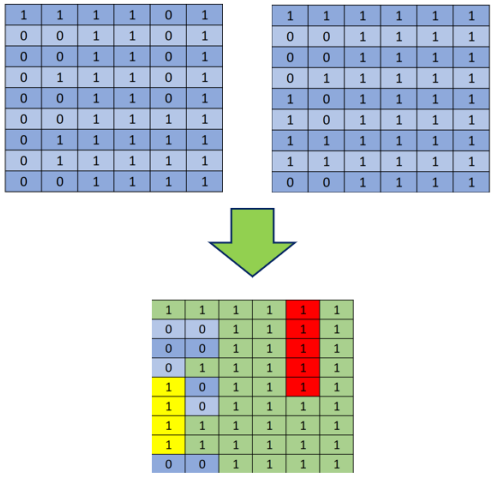
$$\text{IOU} = \frac{\text{area}}{\text{areaunion}} \quad (2)$$

(1) สูตรคำนวนพื้นที่ของการรวมกันของกล่อง

(2) สูตรคำนวณ IOU

จากสูตรทั้งหมดสามารถสรุปออกมาเป็นตัวอย่างได้ตามตัวอย่างให้เข้าใจได้ดังต่อไปนี้

ตัวอย่างที่ 1



ภาพประกอบที่ 3.15 ผลลัพธ์การทำ IOU

Computer Science Department
Faculty of Informatics, Mahasarakham University

สีเขียวคือพื้นที่ทั้งหมด

$$\text{IOU} = \frac{36}{54}$$

$$\text{IOU} = 0.66$$

ผลลัพธ์ : เมื่อคำนวณหาค่า IOU หากค่า ที่นำมาเปรียบเทียบมากกว่า 0.7 จะถือว่าเป็น QR Code แต่ถ้ามีค่า IOU น้อยกว่า 0.7 จะถือว่าเป็นภาพที่ไม่ใช่ QR Code ซึ่งผลลัพธ์ที่ได้จะแสดงดังภาพประกอบที่ 3.16 และ ภาพประกอบที่ 3.17

ภาพประกอบที่ 3.16 ภาพคิวอาร์โค้ด

ภาพประกอบที่ 3.17 ภาพที่ไม่ใช่คิวอาร์โค้ด

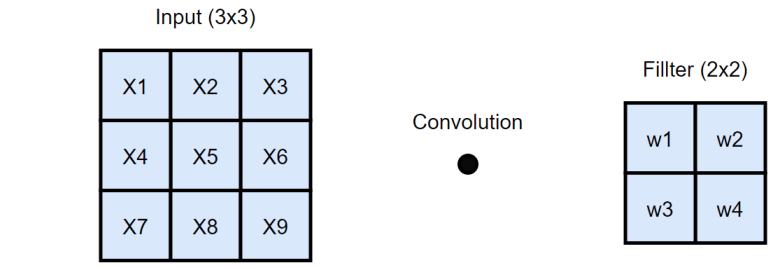
ขั้นตอนที่ 4 นำภาพในขั้นตอนที่ 3 ไปสกัดคุณลักษณะเด่นด้วย CNN

ข้อมูลนำเข้า เป็นภาพทั้งหมดที่ได้จากขั้นตอนที่ 3 ดังภาพประที่ 3.16

กระบวนการทำงาน เป็นขั้นตอนในการสกัดคุณลักษณะของภาพ โดยการทำ convolution โดยใช้ Sliding Windows (Filter) มาสแกนภาพเพื่อทำการแยกองค์ประกอบ โดยวิธีดำเนินการ เริ่มจากนำ filter สแกนไปบนภาพ และทำการคำนวณโดยการนำเอาตำแหน่งที่ตรงกันของภาพต้นฉบับและ filter มาคูณกัน แล้วจึงนำผลรวมของทุกตำแหน่งมาบวกกัน (ตำแหน่งที่ 1 + ตำแหน่งที่ 2 + ... + ตำแหน่งที่ n)

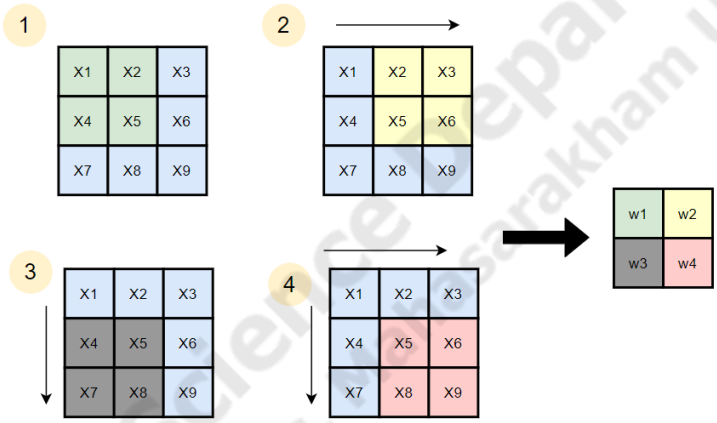
Computer Science Department
Faculty of Informatics Mahasarakham University

หลักการทำงานของ Convolution กำหนดตัวอย่าง Input ขนาด 3x3 และ Weight ขนาด



2x2

ภาพประกอบที่ 3.18 Convolution Input ขนาด 3x3 และ Weight ขนาด 2x2



ภาพประกอบที่ 3.19 ขั้นตอนการคำนวณ Convolution

ขั้นตอนการคำนวณคือ หาผลรวมของการคูณระหว่าง Input กับ weight โดยใช้ weight ชุดเดิมแล้ว สแกน ไปทั้ง Input จากซ้ายไปขวา และบนลงล่าง (1) - (4) ผลลัพธ์ (Z) ที่ได้คือ

$$z_1 = w_1x_1 + w_2x_2 + w_3x_4 + w_4x_5$$

$$z_2 = w_1x_2 + w_2x_3 + w_3x_5 + w_4x_6$$

$$z_3 = w_1w_4 + w_2x_5 + w_3x_7 + w_4x_8$$

$$z_4 = w_1x_5 + w_2x_6 + w_3x_8 + w_4x_9$$

วิธีการหาผลรวมของการคูณระหว่าง input กับ weight

Computer Science Department
Faculty of Informatics, Mahasarakham University

29	88	141	142	145	156	145	72	80	29	88	141
68	93	139	141	145	130	120	60	81	68	93	139
96	133	140	142	146	129	100	29	88	80	59	80
80	97	98	97	76	80	80	68	93	76	80	80
59	65	75	80	72	59	59	76	88	72	59	59
43	55	75	81	60	43	43	68	93	60	43	43
29	88	141	88	80	80	29	76	80	29	88	141
68	93	139	55	76	81	68	76	81	68	93	139
142	145	156	29	88	141	80	29	88	80	59	80
141	145	130	68	93	139	81	68	93	81	43	81

ภาพประกอบที่ 3.20 ค่าจากรูปภาพนำเข้าก่อนทำ Convolution

ตัวอย่างที่ 1

			142	145	156	145	72	80	29	88	141
	103		141	145	130	120	60	81	68	93	139
			142	146	129	100	29	88	80	59	80
80	97	98	97	76	80	80	68	93	76	80	80
59	65	75	80	72	59	59	76	88	72	59	59
43	55	75	81	60	43	43	68	93	60	43	43
29	88	141	88	80	80	29	76	80	29	88	141
68	93	139	55	76	81	68	76	81	68	93	139
142	145	156	29	88	141	80	29	88	80	59	80
141	145	130	68	93	139	81	68	93	81	43	81

29	88	141	142	145	156	145	72	80	29	88	141
68	93	139	141	145	130	120	60	81	68	93	139
96	133	140	142	146	129	100	29	88	80	59	80
80	97	98	97	76	80	80	68	93	76	80	80
59	65	75	80	72	59	59	76	88	72	59	59
43	55	75	81	60	43	43	68	93	60	43	43
29	88	141	88	80	80	29	76	80	29	88	141
68	93	139	55	76	81	68	76	81	68	93	139
142	145	156	29	88	141	80	29	88	80	59	80
141	145	130	68	93	139	81	68	93	81	43	81

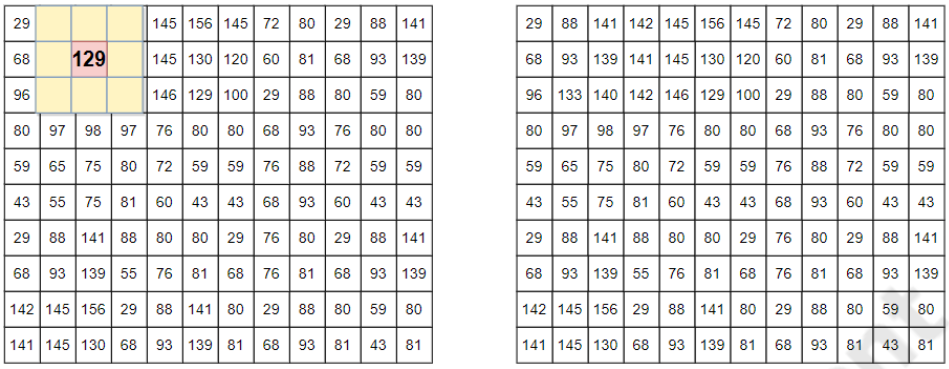
ภาพประกอบที่ 3.21 ตัวอย่างการทำ Convolution

$$z_1 = \left(29 * \frac{1}{9}\right) + \left(88 * \frac{1}{9}\right) + \left(141 * \frac{1}{9}\right) + \left(68 * \frac{1}{9}\right) + \left(93 * \frac{1}{9}\right) + \left(139 * \frac{1}{9}\right) + \left(96 * \frac{1}{9}\right) + \left(133 * \frac{1}{9}\right) + \left(140 * \frac{1}{9}\right)$$

$$= 103$$

ตัวอย่างที่ 2

Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 3.22 ตัวอย่างการทำ Convolution 2

$$z_2 = \left(88 * \frac{1}{9}\right) + \left(141 * \frac{1}{9}\right) + \left(142 * \frac{1}{9}\right) + \left(93 * \frac{1}{9}\right) + \left(139 * \frac{1}{9}\right) + \left(141 * \frac{1}{9}\right) + \left(133 * \frac{1}{9}\right) + \left(140 * \frac{1}{9}\right) + \left(142 * \frac{1}{9}\right)$$
$$= 128.77 \approx 129$$

หลักจากทำขั้นตอน convolution เสร็จแล้วในส่วนต่อไปจะเป็นการทำ RELU เพื่อให้ค่าที่ได้อยู่ในช่วงที่ต้องการ และ Max Pooling Layer ซึ่งเป็นขั้นตอนในการลดขนาดของข้อมูล แต่ยังคงรายละเอียดของ input ไว้ ส่งผลให้ความเร็วในการคำนวณเพิ่มขึ้นและหลีกเลี่ยงการเกิด overfitting และหลังจากนั้นจะนำไปทำ Flatten เพื่อเปลี่ยนข้อมูลที่ได้เป็น Vector 1 มิติ

หลักการทำงานของ RELU คือการลบค่าลบทั้งหมดออกจาก Convolution ค่าบวกทั้งหมดยังคงเหมือนเดิม แต่ค่าลบทั้งหมดจะเปลี่ยนเป็นศูนย์

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

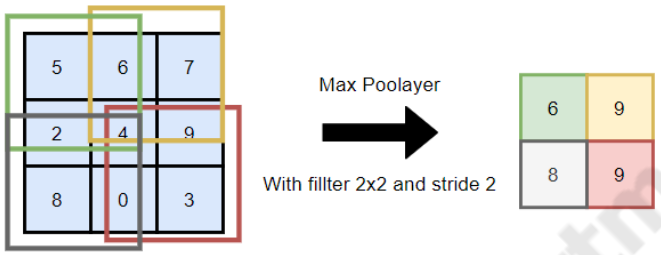
สูตรคำนวณ RELU

ตารางที่ 3.1 ตัวอย่าง RELU

x	f(x) = x	f(x)
-3	f(x) = 0	0
0	f(x) = 0	0
1	f(x) = 1	1

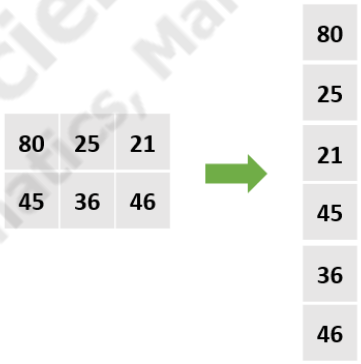
Computer Science Department
Faculty of Informatics, Mahasarakham University

หลักการทำงานของ Max Pooling Layer ทำการกำหนดขนาดของ filter และกำหนดค่า stride จากนั้นวางฟิลเตอร์ที่มุมซ้ายสุดแล้วนำค่าที่มากที่สุดที่อยู่ในบริเวณของ filter มาเป็นค่าของรูปใหม่ จากนั้นขยับ filter ตามค่า stride เช่นค่า stride = 2 ก็ให้ขยับ filter ทีละ 2 ช่อง ดังนี้



ภาพประกอบที่ 3.23 ตัวอย่างการทำ Max Pooling Layer

หลักการทำงานของ Flatten การทำ Flatten คือการทำให้ค่าข้อมูลที่เป็นภาพประกอบที่ได้จาก Maxpooling 2 มิติ กลายเป็นข้อมูล Vector แบบ 1 มิติ เพื่อนำไปใช้ในขั้นตอนต่อไปดังภาพประกอบที่ 3.23



ภาพประกอบที่ 3.24 ตัวอย่างการทำ Flatten

ผลลัพธ์ CNN จะได้คุณลักษณะเด่นของภาพแต่ละภาพออกเป็น n*1 เวกเตอร์แสดงดังภาพประกอบที่ 3.24

Computer Science Department
Faculty of Informatics, Mahasarakham University

ภาพประกอบที่ 3.25 ตัวอย่างผลลัพธ์ของ CNN

ขั้นตอนที่ **5** สร้างโมเดล QR Code และไม่ใช่ QR Code ด้วย Support Vector Machine

Support Vector Machine เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล

ข้อมูลนำเข้า เป็นเวกเตอร์ของภาพโดยแบ่งเป็นกลุ่มของ QR Code และ Non QR Code แสดงดังภาพประกอบที่ 3.26

ภาพประกอบที่ 3.26 ข้อมูลนำเข้า SVM

กระบวนการทำงาน ขั้นตอนที่ 1 เพิ่มค่า Bias ให้กับ $\mathbf{x_1}$ $\mathbf{x_2}$ และ $\mathbf{x_3}$ โดยมีค่า Bias เท่ากับ 1 ดังนั้นจะได้ค่าใหม่ดังนี้

$$\mathbf{x_1} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{x_2} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \qquad \mathbf{x_3} = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

ขั้นตอนที่ 2 นำคุณลักษณะที่ได้ไปคำนวณหาสมการเส้นตรง

$$\mathbf{\alpha_1 x_1 \cdot x_1 + a_2 x_2 \cdot x_1 + a_3 x_3 \cdot x_1 = -1(-ve\ class)}$$

$$\mathbf{\alpha_1 x_1 \cdot x_2 + a_2 x_2 \cdot x_2 + a_3 x_3 \cdot x_2 = -1(-ve\ class)}$$

$$\mathbf{\alpha_1 x_1 \cdot x_3 + a_2 x_2 \cdot x_3 + a_3 x_3 \cdot x_3 = +1(+ve\ class)}$$

Computer Science Department
Faculty of Informatics, Mahasarakham University

จากสมการจะทำการหาค่า $\boldsymbol{\alpha_1}$ $\boldsymbol{\alpha_2}$ และ $\boldsymbol{\alpha_3}$ ตามลำดับดังนี้

$$\boldsymbol{\alpha_1}\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_2}\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_3}\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}=\mathbf{-1}$$

$$\boldsymbol{\alpha_1}\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{-2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_2}\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_3}\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}=\mathbf{-1}$$

$$\boldsymbol{\alpha_1}\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_2}\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_3}\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}\cdot\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}=\mathbf{+1}$$

จากการแทนค่าในสมการด้านบนจะได้สมการใหม่ดังนี้ :

$$\mathbf{6\alpha_1+4a_2+9a_3=-1}$$

$$\mathbf{4\alpha_1+6a_2+9a_3=-1}$$

$$\mathbf{9\alpha_1+9a_2+17a_3=+1}$$

เมื่อทำการแก้สมการทั้ง 3 แล้วจะได้ $\boldsymbol{\alpha_1 = -3.25}$, $\boldsymbol{\alpha_2 = -3.25}$ และ $\boldsymbol{\alpha_3 = 3.5}$

ขั้นตอนที่ 3 หาค่าน้ำหนัก (w) จากสมการดังนี้

$$w=\sum_i \alpha_i x_i$$

ทำการแทนค่า $\boldsymbol{\alpha_i}$ และ $\mathbf{X_i}$ ลงในสมการ

$$w=\boldsymbol{\alpha_1}\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_2}\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}+\boldsymbol{\alpha_3}\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}$$

$$w=\mathbf{(-3.25)}\begin{pmatrix}\mathbf{2}\\\mathbf{1}\\\mathbf{1}\end{pmatrix}+\mathbf{(-3.25)}\begin{pmatrix}\mathbf{2}\\\mathbf{-1}\\\mathbf{1}\end{pmatrix}+\mathbf{(3.5)}\begin{pmatrix}\mathbf{4}\\\mathbf{0}\\\mathbf{1}\end{pmatrix}$$

$$w=\begin{pmatrix}\mathbf{1}\\\mathbf{0}\\\mathbf{-3}\end{pmatrix}$$

ดังนั้นค่าโมเดลที่จะนำไปใช้งานก็คือค่า w

หมายเหตุ : หากต้องการผลการทดลองที่มีความถูกต้องสูงก็จะทำการเพิ่มจำนวนข้อมูลเข้ามาเรียนรู้

Computer Science Department
Faculty of Informatics, Mahasarakham University

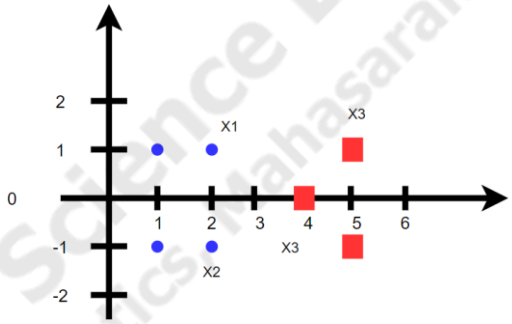
ขั้นตอนที่ 4 นำค่าโมเดล ไปใช้งานจริงจากภาพประกอบที่ $\mathbf{X_1}$ $\mathbf{X_2}$ และ $\mathbf{X_3}$ เป็นข้อมูลที่ใช้ในการเรียนรู้

ตัวอย่างการคำนวณเพื่อแยกประเภทกลุ่มของ QR Code และ Non QR Code และหากต้องการรู้ว่า $\mathbf{X_4}$ อยู่ในกลุ่มของ QR Code หรือ Non QR Code ให้นำโมเดลที่ได้จากการเรียนรู้มาคำนวณดังนี้ โดย $\mathbf{X_4} = \binom{5}{1}$ แล้วนำค่า $\mathbf{X_4}$ และ $\mathbf{w}$ ไปแทนในสมการเส้นตรง $\mathbf{y} = \mathbf{wx} + \mathbf{B}$ จะได้ค่าดังนี้

$$\mathbf{y} = \binom{\mathbf{1}}{\mathbf{0}} \cdot \binom{\mathbf{5}}{\mathbf{1}} + (\mathbf{-3})$$

$$\mathbf{y} = \mathbf{2}$$

ดังนั้น $\mathbf{X_4}$ จะอยู่ในกลุ่มของ QR Code เพราะว่า y เป็นบวก ซึ่งจากภาพประกอบที่ 3.27 $\mathbf{X_4}$ก็อยู่ในกลุ่มของ QR Code



ภาพประกอบที่ 3.27 กลุ่มของ SVM

ผลลัพธ์ เมทริกของค่าน้ำหนักที่ดีที่สุด (w) จากการเรียนรู้ของ SVM

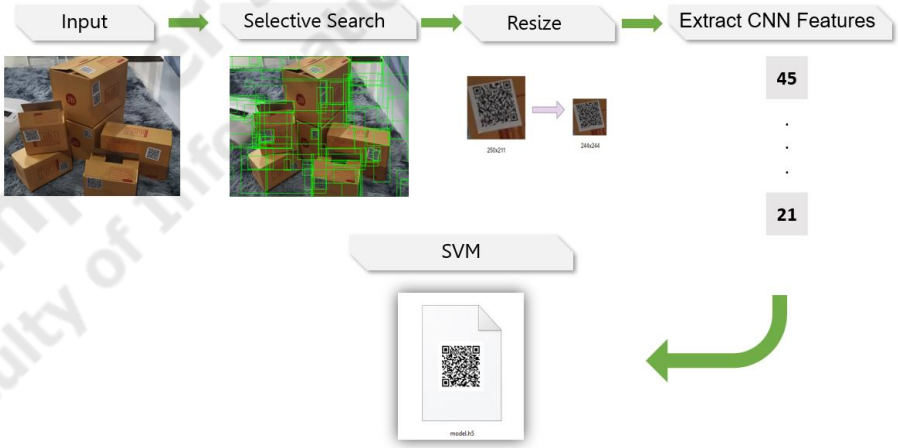
สรุปการทำงานของ RCNN จะเริ่มจากนำภาพเข้า ตัว Selective search จะเสนอพื้นที่ที่น่าสนใจมาให้เยอะพอสมควร แต่ละพื้นที่จะถูกนำไปปรับขนาดให้เท่ากัน ซึ่งจะต้องทำให้ขนาดเท่ากันเพราะเป็นข้อจำกัดของ CNN ที่ขนาดภาพจะตายตัว โดยภาพย่อยที่ตัดมาจะมีการนำมาตรวจสอบเมื่อภาพย่อยที่นำมาตรวจสอบมีค่าความมั่นใจเท่ากับหรือมากกว่า 0.7 จะนับว่าเป็นภาพเป็นที่มีคลาสเป็น QR Code ถ้าภายในภาพย่อยที่นำมาตรวจสอบ เมื่อภาพนั่นมีคลาส QR Code จะทำ bounding box บนพื้นที่นั้นดังภาพประกอบที่ 3.28

Computer Science Department
Faculty of Informatics, Mahasarakham University

ภาพประกอบที่ 3.28 ผลลัพธ์ Bounding Box ที่ได้จากการทำ RCNN

3.3 การค้น QR Code ด้วย RCNN (RCNN Object QR Code Detection)

ในขั้นตอนนี้ เป็นการค้นหา QR Code จากภาพที่ผู้ใช้ส่งมาจากสมาทโฟน โดยมีขั้นตอนแสดงดังภาพที่ 3.29



ภาพประกอบที่ 3.29 ผลลัพธ์ที่ได้จากการค้นหาวัตถุด้วย RCNN

ขั้นตอนที่ 1 Input นำภาพนำเข้าและจะเป็นภาพที่มี QR Code และไม่มี QR Code

ขั้นตอนที่ 2 Selective search จะทำค้นหาลักษณะเด่นของรูปภาพนำเข้าการตีกรอบพื้นที่

ขั้นตอนที่ 3 Resize ทำการเปลี่ยนขนาดภาพประกอบที่ได้จากขั้นตอนที่ 2 ให้เป็น 224*224 พิกเซล

ขั้นตอนที่ 4 Extract CNN Features ทำการสกัดคุณลักษณะเด่นของภาพประกอบที่ได้ทั้งหมดได้จากขั้นตอนที่ 3

ขั้นตอนที่ 5 SVMใช้ SVM ในการจัดกลุ่มของข้อมูลภาพที่ได้จากขั้นตอนที่ 4 ว่าภาพที่ได้เป็นรูปภาพประเภท QR Code หรือ Non QR Code แสดงดังภาพประกอบที่ 3.30

ภาพประกอบที่ 3.30 ผลลัพธ์ที่ได้จาก SVM

3.4 การตีความ QR Code

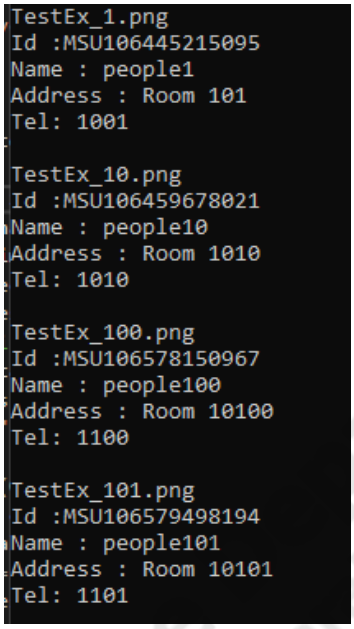
ในถอดรหัสข้อมูล QR Code จะใช้ library ของ Python ที่มีชื่อว่า “Pyzbar” เพื่อถอดรหัสข้อมูลจากภาพ QR Code ทีได้ออกมาเป็นตัวอักษรข้อมูล ข้อมูลนำเข้า เป็นภาพ QR Code แสดงดังภาพที่ 3.31

ภาพประกอบที่ 3.31 ภาพนำเข้า Pyzbar

กระบวนการทำงาน ตัว Module จะทำการแปลงรหัส QR Code ให้เป็นข้อมูล UTF-8 หรือตัวอักษร

Computer Science Department
Faculty of Informatics, Mahasarakham University

ผลลัพธ์ ข้อมูลที่ได้จาก QR Code ของโครงงานนี้จะประกอบไปด้วย รหัส , ชื่อ , ที่อยู่ และเบอร์โทร แสดงดังภาพประกอบที่ 3.32



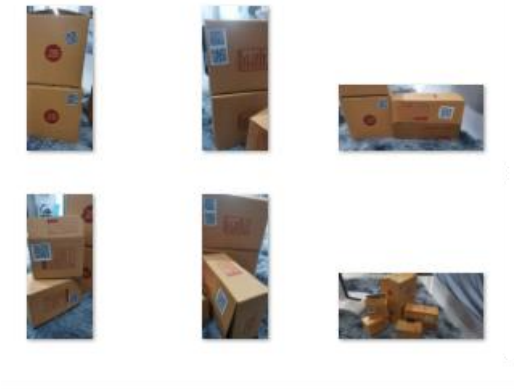
ภาพประกอบที่ 3.32 ผลลัพธ์การถอดรหัสด้วย Pyzbar

3.5 การเตรียมข้อมูล

การเตรียมข้อมูลเพื่อนำไปทำ RCNN ภาพที่นำเข้าจะถูกแบ่งเป็นชุดสอน 80% และชุดทดสอบ 20% จะเข้าสู่กระบวนการทำ Pre-Processing เพื่อเป็นการเตรียมข้อมูลก่อนเข้าสู่ Model และนำข้อมูลที่ผ่านกระบวนการ Pre-Processing แล้ว มาทำการ Resize รูปภาพให้อยู่ในขนาดรูปที่เท่ากัน เพื่อเตรียมเข้าเข้าสู่ Model โดยใช้ RCNN

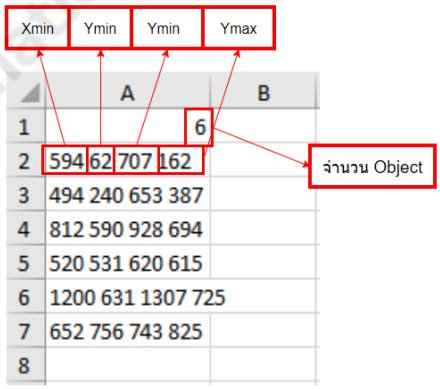
เตรียมไฟล์รูปภาพเพื่อนำไปทำ RCNN โดยจะมีขนาดข้อมูลทดสอบ 20% จากทั้งหมด และจะมีขนาดข้อมูลเรียนรู้ 80% จากข้อมูลทั้งหมดโดยมีตัวอย่างข้อมูลใน ภาพประกอบที่ 3.33 และ ภาพประกอบที่ 3.34

Computer Science Department
Faculty of Informatics, Mahasarakham University



ภาพประกอบที่ 3.33 ชุดข้อมูลการเรียนรู้

ภาพประกอบที่ 3.34 ชุดข้อมูลการทดสอบ



ภาพประกอบที่ 3.35 ชุดข้อมูลการเรียนรู้

จาก ภาพประกอบที่ 3.34 ข้อมูลที่นำมาเก็บโดยไฟล์ใช้ Csv จะเป็นไฟล์ที่ทำการเก็บจำนวนวัตถุ ของภาพ QR Code และ Ground truth โดยในไฟล์จะมีรายละเอียดดังนี้ :

บรรดทัดแรก เป็นข้อมูลจำนวนของวัตถุภายในภาพ เช่น ภาพนี้มีจำนวน QR Code ทั้งหมด 6 ภาพ

Computer Science Department
Faculty of Informatics, Mahasarakham University

บรรดทัดสอง เป็นการเก็บข้อมูลหรือข้อมูลพื้นที่จริง (Groundtruth) ของวัตถุภายในภาพ โดยค่าตัวเลขจะเป็นค่าตำแหน่งของ Xmin,Ymin,Xman,Ymax ตามลำดับเพื่อใช้ในการตีกรอบของพื้นที่จริง (Groundtruth)

3.6 การวัดประสิทธิภาพของการค้นหาคิวอาร์โค้ด

ในการวัดประสิทธิภาพการค้นหาคิวอาร์โค้ดนี้ใช้ อัตราการตรวจจับ (Detection Rate) หรืออัตราผลบวกจริง (True positive rate) เป็นการคำนวณหาอัตราความแม่นยำในการค้นหา เพื่อทดสอบอัตราการค้นคิวอาร์โค้ดโดยมีสมการคำนวนคือ

$$\mathrm{DR} = \frac{\mathrm{TP}}{TP+FP} \qquad (1)$$

DR คือ อัตราการค้นหาการค้นหา

TP คือภาพเป็นคิวอาร์โค้ดและโปรแกรมตอบเป็นคิวอาร์โค้ด

FP คือภาพที่เป็นคิวอาร์โค้ดตอบไม่เป็นคิวอาร์โค้ด

ตัวอย่างการคำนวน

ภาพประกอบที่ 3.36 ภาพประกอบการทดสอบการคำนวณอัตราการค้นหา

จากภาพประกอบที่ 3.36 จะได้ค่า TP = 8 และ = 1 นำไปคำนวณตามสมการที่ (1) จะได้

$$\mathrm{DR} = \frac{8}{8+1} = 0.88$$

หากภาพที่นำมาทดสอบมีทั้งหมด 100 ภาพจะหา DR ของทุกภาพเพื่อหาค่าเฉลี่ยของทุกภาพ

Computer Science Department Faculty of Informatics Mahasarakham University