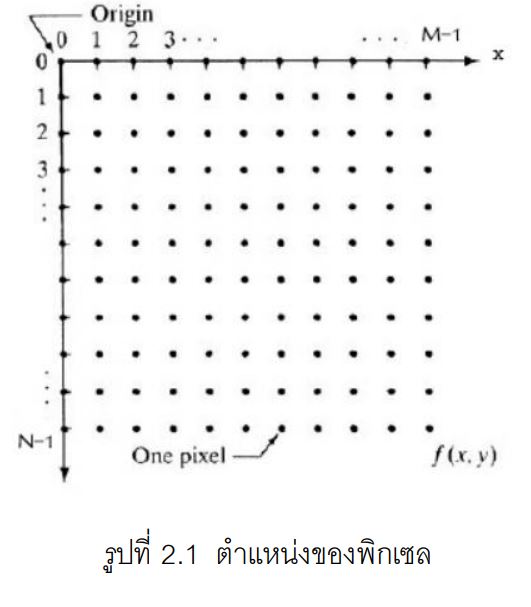
**บทที่ 2**

**หลักการและทฤษฎีที่เกี่ยวข้อง**

โครงการพัฒนาระบบตรวจพาหนะเรื่อง ตรวจจับและแจ้งเตือนอุบัติเหตุบนท้องถนนโดยโครงงานนี้จะเน้นเรื่องการประมวลผลภาพดิจิตอลสำหรับใช้ในการตรวจจับโมเดลยานพาหนะอื่น ๆ

โดยการใช้เทคนิค **Machine Learning with TensorFlow** ทฤษฎีและหลักการต่าง ๆ ที่จะกล่าวต่อไปนี้มาประยุกต์ใช้ รวมถึงทฤษฎีที่นำมาใช้กับระบบตรวจจับและแจ้งเตือนอุบัติเหตุบนท้องถนนที่จะพัฒนาขึ้นมาในโครงงานนี้และสามารถนำไปต่อยอดในอนาคต

* 1. **ภาพและความหมายของพิกเซล**
     + พิกเซล(pixel) เป็นหน่วยพื้นฐานของภาพ คือจุดภาพบนจอแสดงผล หรือ จุดภาพในรูปภาพที่รวมกันเป็นภาพขึ้น โดยภาพหนึ่งๆ จะประกอบไปด้วยจุดภาพหรือพิกเซลมากมาย และแต่ละภาพที่สร้างขึ้นจะมีความหนาแน่นของจุดภาพ หรือบางครั้งแทนว่าความละเอียด (ความคมชัด) ที่แตกต่างกันไป จึงใช้ในการบอกคุณสมบัติของภาพ จอภาพ หรือ อุปกรณ์แสดงผลภาพได้จอภาพที่มีจำนวนพิกเซลมาก จะมีความละเอียดของภาพมาก โดยมากจะระบุจำนวนพิกเซลแนวนอน x แนวตั้ง เช่น 1366 x 768 พิกเซลคำว่า "พิกเซล" (pixel) มาจากคำว่า "พิกเจอร์" (picture) ที่แปลว่า รูปภาพ และ "เอเลเมนต์" (element) ที่แปลว่า องค์ประกอบ



M = จำนวนที่มากที่สุดของพิกเซลในแกน X

Y = จำนวนที่มากที่สุดของพิกเซลในแกน Y

* 1. **ลักษณะการจัดเก็บมูลภาพแบบดิจิตอล**

โดยทั่วไปแล้ วภาพจะมีความเข้มตั้งแต่ 2 ระดับขึ้นไป แต่ที่นิยมใช้ กันมาก คือ ค่าระดับ ความเข้มของพิกเซลที่เท่ากับ 256 ระดับ ซึ่งจะท าให้ค่าของพิกเซลอยู่ในช่วง (0-255) โดยใช้ เนื้อที่ในการเก็บข้อมูลขนาด 1 ไบต์หรือ 8 บิต สำหรับข้อมูล 1 พิกเซล (256) ในกรณี ที่ต้องการภาพที่ มีความเข้มสูงอาจจะต้องการจำนวนบิตสำหรับการเก็บข้อมูลมากกว่า 8 บิต คืออาจจะเป็น 16 หรือ 24 บิตโดยจะแยกความแตกต่างของภาพแต่ละประเภทให้เห็นอย่างชัดเจนได้ดังนี้

* + - ภาพ 2 ระดับ คือ มีพิกเซลสีขาวกับสีดำเท่านั้น โดยแต่ละพิกเซลจะมีขนาดของ ข้อมูลเท่ากับ 1 บิต
    - ภาพ 16 ระดับ คือ ในแต่ละพิกเซลจะมีขนาดข้องข้อมูล 8 บิต ซึ่งสามารถแสดง ภาพได้ความเข้มถึง 256ระดับ
    - ภาพทรูคัลเลอร์ (True Color) คือในแต่ละพิกเซลจะมีขนาดของข้อมูล 24 บิตทำให้สามารถแสดงภาพออกมาได้เฉพาะภาพสีเท่านั้น ไม่สามารถแสงดเป็นภาพขาวดำได้

โดยทั่วไปวิธีการประมวลผลภาพเชิงตัวเลขที่ทำให้คอมพิวเตอร์สามารถรู้ วัตถุในภาพได้

นั้นแบ่งออกได้เป็นสองระดับด้วยกันคือ การประมวลผลภาพในระดับต่ำ (Low-Level Image)

และการประมวลผลภาพในระดับสูง (High-Level Image Processing) การประมวลผลภาพใน

ระดับต่ำจะเป็นการประมวลผลเชิงตัวเลขทั้งหมด เพื่อหาตัวแปรต่าง ๆ มาอธิบายข้อมูลภาพ

โดยมีจุดประสงค์ที่จะนำตัวแปรเหล่านั้นไปใช้ในการประมวลผลภาพระดับสูงต่อไปการประมวลผลภาพระดับต่ำจะประกอบด้วย การประมวลผลภาพก่อน (Preprocessing) เช่น

การกำจัดสัญญาณรบกวน หรือการทำให้ภาพคมชัด เป็นต้น การประมวลผลระดับสูงเป็นการนำผลลัพธ์ หรือสัญลักษณ์ที่ได้จากการประมวลผลภาพ ระดับต่ำมาตีความหรือประมวลเพื่อให้คอมพิวเตอร์สามารถรู้จักและเข้าใจภาพได้ สำหรับความ แตกต่างของการประมวลผลภาพ ทั้ง 2 ประเภทนั้นคือ การประมวลผลภาพระดับต่ำจะใช้ ค่า ความสว่างของจุดภาพ (พิกเซล) ส่วนการประมวลภาพระดับสูงนั้นข้อมูลภาพที่นำมาประมวลผล จะถูกแสดงในรูปสัญลักษณ์ ซึ่งสัญลักษณ์เหล่านี้จะแสดถึงสิ่งต่าง ๆ ที่อยู่ในภาพเช่น ขนาดของ วัตถุรูปร่าง และความสัมพันธ์กันระหว่างวัตถุภาพ

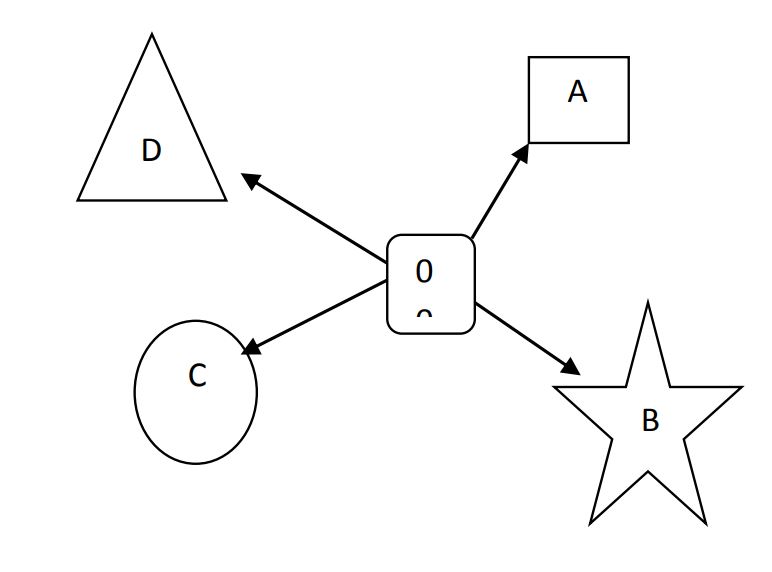
* 1. **การแยกลักษณะเฉพาะของภาพ (Image Feature Extraction)**

การแยกลักษณะเฉพาะของภาพเป็นการแยกหรือสกัดเอาข้อมูลที่สำคัญของภาพออกมา ซึ่งลักษณะเฉพาะของภาพเป็นคุณสมบัติที่สามารถหาได้โดยใช้ ขั้นตอนวิธี การประมวลผลภาพ (Image Processing) โดยที่ลักษณะเฉพาะพื้นฐานของภาพประกอบด้วย 3 ส่วนคือสี รู ปร่างและ พื้นผิว

* สี(Color) เป็นลักษณะเฉพาะของภาพที่มีบทบาทสำคัญในระบบค้นคืนภาพเช่นฮิสโต แกรมสีซึ่งเป็นลักษณะเฉพาะของสีที่ถูกนำมาใช้บ่อย ๆ เนื่องจากสีเป็นสิ่งที่สามารถมองเห็นได้ ง่าย และเป็นสิ่งแรกที่สามารถสังเกตเห็นได้จากการมองภาพนอกจากนี้สียังสามารถใช้ ในการแยกแยะ กลุ่มของภาพออกตามเนื้อหาได้เป็นอย่างดีเช่นสี ฟ้าของน้ำทะเลสี แดงของดอกไม้ เป็นต้น
* รูปร่าง (Shape) เป็นลักษณะเฉพาะของภาพที่ใช้อธิบายถึงรูปร่างและลักษณะรวมถึง ขนาดของวัตถุภายในภาพซึ่งทำให้สามารถแยกวัตถุออกจากพื้นหลังหรือแยกแยะระหว่างวัตถุที่มี รูปร่างแตกต่างกันออกจากกันได้
* พื้นผิว (Texture) เป็นลักษณะเฉพาะที่ใช้อธิบายความหยาบความละเอียดหรือความ ซับซ้อนของวัตถุภายในภาพซึ่งแต่ละภาพอาจจะประกอบด้วยวัตถุที่มีลักษณะพื้นผิวที่แตกต่าง กันออกไปการวิเคราะห์พื้นผิวจะช่วยให้สามารถแยกแยะความแตกต่างของวัตถุได้ ดียิ่งขึ้นการค้น คืนภาพที่ใช้พื้นผิวเป็นลักษณะเฉพาะของภาพส่วนใหญ่จะถูกนำไปใช้ในการค้นหาภาพจากกลุ่ม ภาพพื้นผิวเช่นชุดภาพพื้นผิวของหินชุดภาพพื้นผิวของใบไม้
  1. **การจำแนกหมวดหมู่ของพาหนะ (Classification)**

Classifications เป็นขั้นตอนในการจำแนกและตัดสินใจว่าอินพุตที่เข้ามานั้นเป็นภาพรถ ชนิดใด โดยในขั้นตอนนี้มีหลายวิธีการด้วยกัน เช่น การเปรียบเทียบอินพุตกับโครงสร้างของภาพ รถยนต์ในฐานข้อมูล การเปรียบเทียบอินพุตกับกฎเพื่อการตัดสินใจ การใช้ โครงข่ายประสาท เทียม การเรียนรู้ของเครื่องจักร การคำนวณเชิงวิวัฒนาการ ในส่วนของเนื้อหาที่เกี่ยวข้องกับการ ทำโปรเจคจะเป็นส่วนของการเรียนรู้ของเครื่องจักร การเรียนรู้ของเครื่องจักร (Machine Learning) เป็นเทคนิคหรือกระบวนการที่ใช้ สำหรับ ปรับแต่งเครื่องจักรหรือคอมพิวเตอร์ให้ มีพฤติกรรมเฉพาะตัวที่สนับสนุน การเรียนรู้ โดยจะมี เทคนิคต่าง ๆ หลายวิธี แต่ที่เทคนิคที่สนใจคือ เทคนิค Nearest Neighbor Classification Nearest Neighbor Classification เป็นการเรียนรู้ประเภท Unsupervised Learning ซึ่ง เป็นการจำแนกหรือจัดกลุ่มที่มีวิธี การไม่ซับซ้อน โดยพิจารณาจากชุดข้อมูลที่ใกล้ เคียงกัน กล่าวคือ ข้อมูลที่ใกล้ เคียงกับค่าของข้อมูลที่พิจารณามากที่สุด

ในที่ นี้ค่าความใกล้ เคียงจะ หมายถึง ระยะทาง ที่มีค่าน้อยที่สุดระหว่างชุดข้อมูลกับข้อมูลที่พิจารณา



ภาพที่ 2.4 แสดงตัวอย่างระยะทางระหว่างชุดข้อมูล

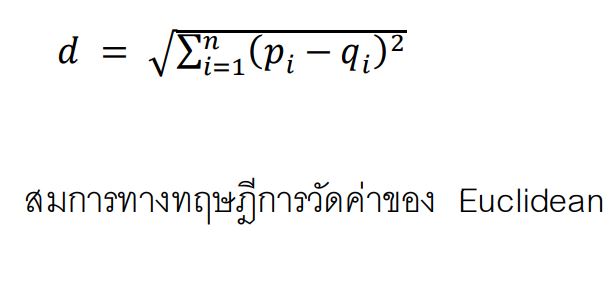
จากรูปที่2.4 จะเห็นได้ว่า ข้อมูลที่มีระยะทางใกล้ กับจุดศูนย์กลางมากที่สุด คือ A ซึ่งมี

ระยะเป็น 1 ในทฤษฎีของ Nearest Neighbor การจำแนกชุดข้อมูลที่มีระยะห่าง 1 จะเรียกว่า

1NN (One Nearest Neighbor) โดยระยะห่างของข้อมูลนั้นสามารถกำหนดได้ ว่าต้องการมาก

น้อยเพียงใดด้วยเหตุนี้จึงมีการจำแนกที่เรียกว่า k NN ซึ่ง k แทนด้วยค่าระยะทางระหว่างข้อมูลที่

ต้องการ สำหรับการหาค่าระยะทางจะใช้สมการทางทฤษฎีการวัดค่าของ Euclidean



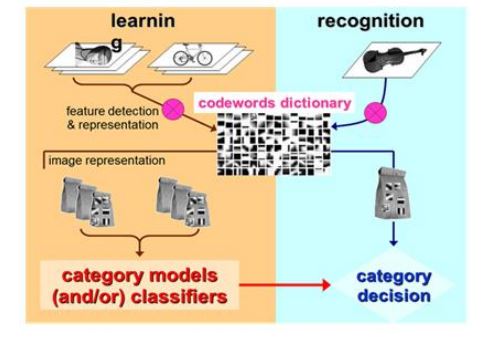
* 1. **OpenCV (Open Source Computer Vision Library)**

OpenCV หรือ Open Source Computer Vision Library ใช้ ในการประมวลผลภาพ และงานทางด้าน การมองเห็นของคอมพิวเตอร์ (Computer Vision) Library นี้ถูกพัฒนาขึ้นด้วย ภาษา C และ C++ และยังมี interface ที่ไว้ เชื่อมต่อกับ tool อื่นด้วยเช่น Python, Ruby, Matlab เป็นต้น นอกจากนี้ OpenCV เป็น library ที่สร้างขึ้นเพื่อให้ ผู้ ใช้ หรือนักพัฒนาสามารถใช้ ฟังก์ชัน ใน library

มาพัฒนาชิ้นงาน ที่มีความซับซ้อนโด ยใช้ เวลาเพียงไม่นาน OpenCV

ประกอบด้วย Data Structure และ Algorithm

* Data Structure ใช้เก็บข้อมูลต่าง ๆ อาทิ เช่น รูปภาพ เมทริกซ์ พิกัด
* Algorithm เพื่อการประมวลผลต่าง โดยเฉพาะการประมวลผลทางรูปภาพ
  1. **หลักการทำงานโครงสร้างแบบถุงคำ (bag of words)**

****

รูปที่ 2.6 ส่วนของโครงสร้างแบบถุงคำ

**ส่วนที่หนึ่ง Image representation** หลักการทำงานในรูป (ด้านซ้าย) แสดงให้ เห็นว่า

เมื่อเรารับภาพเข้ามา ภาพที่เรารับเข้ามานั้นก็จะถูกนำไปตรวจสอบในฐานข้อมูล ว่าเป็นภาพที่มี

ส่วนของภาพ คล้ายกับภาพที่เรารับเข้ามาหรือไม่ถ้ามีก็จะแสดงออกมาว่า ภาพที่เรารับเข้ามานั้น

เป็นภาพอะไร

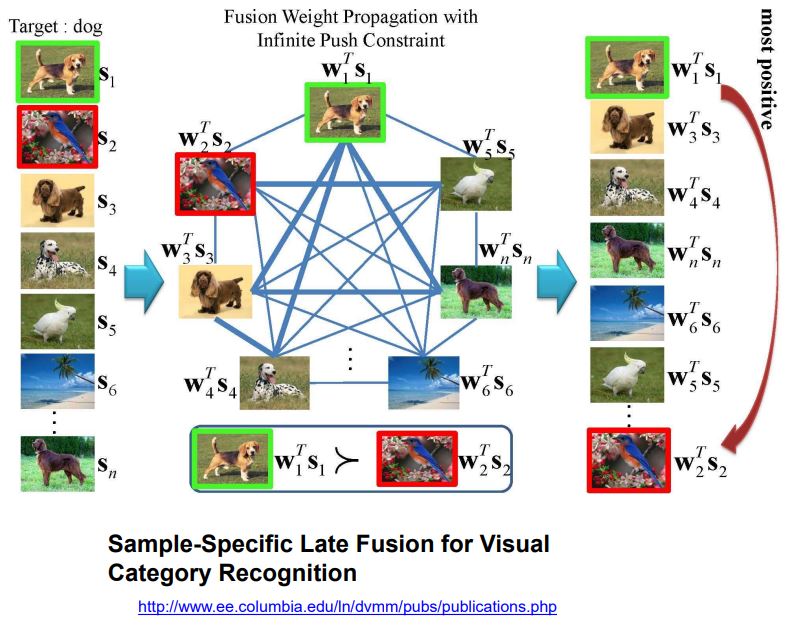
**ส่วนที่สอง ส่วนของการรู้ จำ (Recognition)** จะเป็นส่วนที่ท าการค้นหาว่า input ที่

รับเข้ามา มีอยู่ในฐานข้อมูลหรือไม่ โดยสามารถอธิบายด้วยรูปภาพได้ ดังรู ป (ด้านขวา) ส่วนของ

ระบบรู้จำจะเป็นส่วนที่นำข้อมูลมาทำการ segmentation แล้วนำไป เลือกเฉพาะส่วนที่เป็น

จุดเด่นของภาพนำไปเก็บลงในฐานข้อมูลโดยจะถูกเรียกใช้ ก็ต่อเมื่อ มีการรับ input เข้ามา

* 1. **การรู้จำ (Recognition)**



การรู้จำภาพเป็นแขนงหนึ่งของการรู้ จำแบบรูป(Pattern Recognition) ที่มี

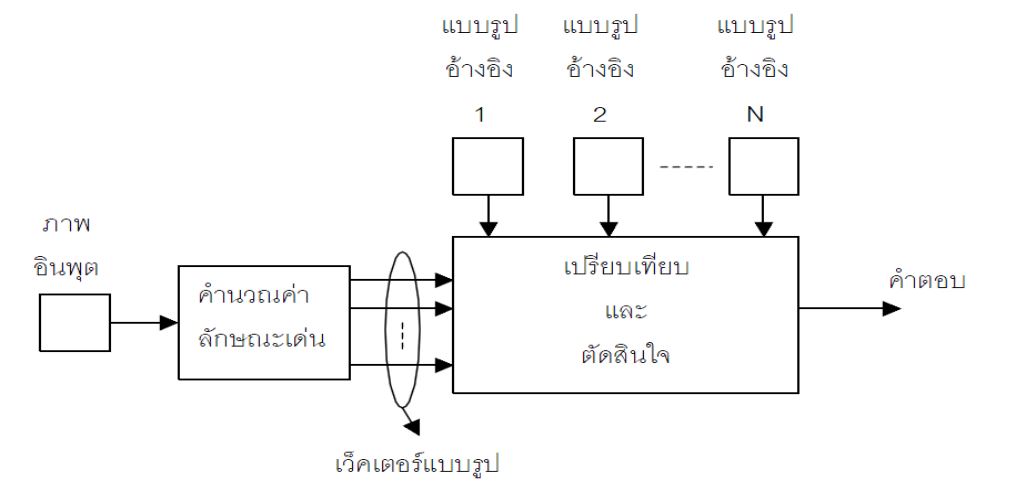
การศึกษาอย่างแพร่หลายคาดว่าแบบรูป (Pattern) สามารถแปลความได้หลายแบบในที่ นี้รูปแบบ

หมายถึงกลุ่มตัวเลขที่บรรยายลักษณะของภาพในการรู้จำภาพจะต้องรู้จำแบบรูปของแต่ละภาพ

เพื่อแยกแยะภาพที่ต่างกันออกจากกัน แบบรูปที่ดีจะต้องบ่งถึงลักษณะเด่นของภาพซึ่งอาจได้จาก

การวัดเช่นอัตราส่วนความกว้างต่อความยาว จำนวนยอดแหลมและการพิจารณาจากฮิสโตแกรม

เป็นต้น ในบทนี้จะกล่าวถึงหลักการรู้จำภาพ โครงสร้างพื้นฐานของระบบรู้จำภาพ แสดงโครงสร้างพื้นฐานของระบบรู้จำภาพ ส่วนแรกคือส่วนคำนวณค่าลักษณะเด่น (Feature Extraction)ของภาพ ผลลัพธ์ที่ได้คือ แบบรูปซึ่งอาจอยู่ในรูปของเวกเตอร์ ค่าลักษณะเด่นเหล่านี้จะนำไปเปรียบเทียบกับแบบรูปอ้างอิงต่าง ๆ จากนั้นจะทำการตัดสินใจโดยให้คำตอบเป็นแบบรูปอ้างอิงที่คล้ายกับแบบรูปอินพุตมากที่สุด ดังภาพ



จากภาพ ส่วนของการเปรียบเทียบและตัดสินใจมักถูกเรียกว่าเป็นขั้นตอนของ

การจำแนกแบบรูป (Pattern Classification)การจำแนกแบบรูปนี้มีอยู่หลายวิธี ซึ่งในบทนี้

จะแนะนำวิธีง่ายๆเพื่อให้ผู้อ่านเข้าใจหลักการของระบบรู้จำภาพดีขึ้นขึ้นตอนที่แสดงในภาพ

เป็นขั้นตอนของการรู้จำ (Recognition Phrase) ซึ่งเป้ าหมายก็คือการให้คำตอบว่าแบบรูป

ของภาพอินพุตมีความคล้ายกับแบบรูปของภาพอ้างอิงภาพใดมากที่สุดการได้มาของแบบรูป

อ้างอิงนั้นสามารถทำได้หลายวิธี แบบรูปอ้างอิงอาจอยู่ในรูปของแบบจำลองทางคณิตศาสตร์

ซึ่งจะต้องมีวิธีเฉพาะในการเปรียบเทียบการสร้างแบบจำลองทางคณิตศาสตร์สามารถทำได้จาก

ขั้นนตอนการฝึกฝน (Training Phrase)ซึ่งโดยทั่วไปแล้วจะต้องมีตัวอย่างภาพที่มีลักษณะเดียวกัน

หลายๆภาพจากนั้นจะทำการคำนวณหาค่าลักษณะเด่นของแต่ละภาพ ซึ่งผลลัพธ์ที่ได้ก็คือแบบรูปของภาพเหล่านั้นนั่นเอง

* 1. **สี่เหลี่ยมในการปิดดล้อมวัตถุ (Bounding Box)**

คือกล่องสี่เหลี่ยมสำหรับปิดล้อมรอบบริเวณนั้น ๆ หรือบริเวณที่ถูกกำหนดหมายเลข ซึ่ง

กล่องสี่เหลี่ยมนี้จะช่วยในการคำนวณจุดศูนย์กลางพื้นที่ ความยาวแกนเอก-แกนโท จุดพิกัด ฯ

ของบริเวณหรือวัตถุในภาพที่สนใจ โดยจากรูปที่ 2.12 จะแสดงแนวความคิดการสร้างกรอบภาพ

โดยรูป 2.12(a) คือภาพวัตถุสีขาว-ดำ ส่วนในรูปที่ 2.12(b) คือขอบและจุดศูนย์กลาง (Centroid)

ของวัตถุ ต่อจากนั้นจะมีการหาแกนเอกและแกนโทของวัตถุดังรู ปที่ 2.12(c) จากนั้นจะมีการหา

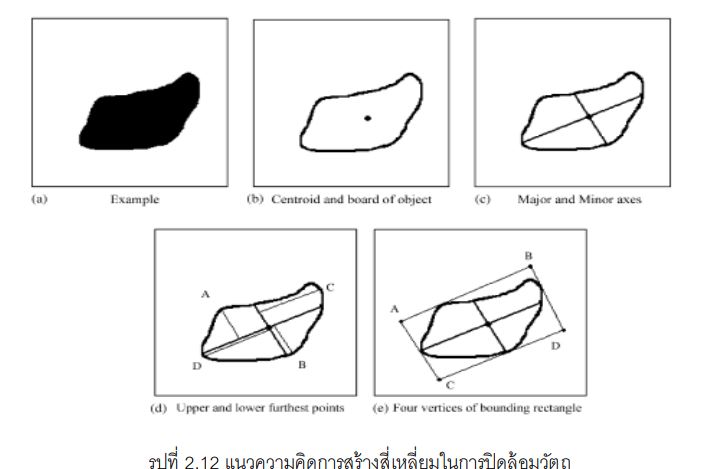
จุดที่มีระยะห่างจากจุดศูนย์กลางวัตถุถึงขอบวัตถุที่ยาวที่สุ ดโดยพิจารณาจากส่วนด้านบนของ

แกนเอกและส่วนล่างของแกนเอก (A และ B) รวมทั้งการหาจุดที่มีระยะห่างจากจุดศูนย์กลางวัตถุ

ถึงขอบวัตถุที่ยาวที่สุดโดยพิจารณาจากบริเวณส่วนซ้ายของแกนโทและส่วนขวาของแกนโท(C

และ D) สุดท้ายก็จะได้ จุดพิกัดใน การสร้างกล่องสี่เหลี่ยมในการปิดล้อมวัตถุ ( Minimum

bounding box) นั้น การหา ( xmax , xmin , ymax , ymin , xcenter , ycenter)



จากรูปทำการคำณวนพิกัดที่ล้อมรอบรูปจะได้ว่า

xmin = C , xmax = D , ymin = C , ymax = B , xcenter = C+D / 2 ycenter = C+B/2

* 1. **Machine Learning**

**(TensorFlow API )**

* + - ประเภทของ Machine Learning
      1. การเรียนรู้แบบไม่มีผู้สอน (unsupervised Learning )ไม่มีการสอนใด ๆ และให้คอมพิวเตอร์หารูปแบบของข้อมูลขาเข้าเอง
      2. การเรียนรู้แบบมีผู้สอน (supervised learning) ข้อมูลตัวอย่างและผลลัพธ์ที่"ผู้สอน"ต้องการถูกป้อนเข้าสู่คอมพิวเตอร์ เป้าหมายคือการสร้างกฎทั่วไปที่สามารถเชื่อมโยงข้อมูลขาเข้ากับขาออกได้

ในโปรเจ็คนี้จะใช้การเรียนรู้แบบมีผู้สอนเป็นหลัก